7

Projeto de Experimentos com Modelos de Simulação

Dentre os inúmeros problemas com que se deparam os tomadores de decisão, encontra-se o de compreender como o comportamento dos sistemas, medido por uma ou mais variáveis de interesse, é influenciado por uma ou mais variáveis controláveis ou mesmo por outras não controláveis. Neste capítulo apresenta-se uma técnica estatística que pode ser empregada para este propósito, o projeto experimental.

Tópicos

7.1. Introdução

O processo ou técnica estatística empregada com o propósito de determinar como o comportamento de um sistema pode ser influenciado pelos possíveis valores de uma ou mais variáveis é o da experimentação. Considerando o jargão da estatística, as variáveis que medem o desempenho do sistema são aquelas conhecidas como respostas ou variáveis de resposta enquanto que aquelas que o experimentador está manipulando são conhecidas por fatores.

Segundo Montgomery (1997), um experimento é um teste ou uma série de testes, nos quais alterações controladas são realizadas sobre os fatores envolvidos num sistema ou processo, de tal forma que se possa observar e identificar as razões das mudanças ocorridas sobre as respostas. A Figura 7.1 procura ilustrar esta definição.

Na figura, o sistema ou processo é o objeto de estudo. Como visto no capítulo 2, pode ser visto como uma combinação de recursos (pessoas, máquinas, métodos, etc.) os quais atuam e transformam entradas em algum tipo de resposta observável. Dentre as variáveis envolvidas no processo (fatores), algumas são controláveis e outras não. No entanto, quando se faz uso de modelos de sistemas ou processos, os fatores não controláveis podem ser fixados para fins de teste.

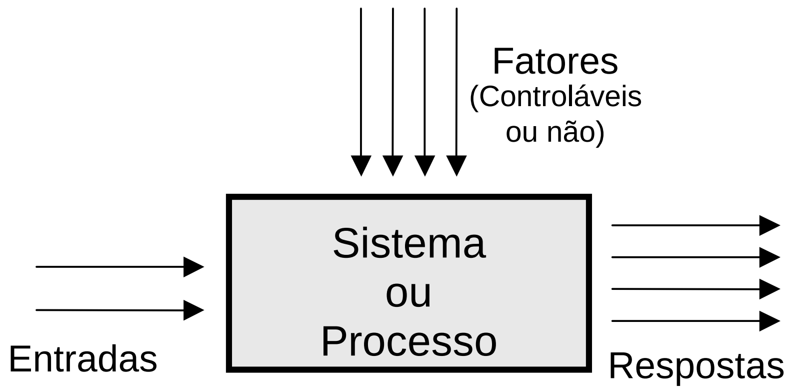

Figura 7.1: Sistema ou processo, com respostas dependentes dos fatores.

Ainda segundo Montgomery (1997), no processo experimental, os objetivos perseguidos podem ser:

1. Determinar quais variáveis têm mais influência sobre as respostas;
2. Determinar que valores associar a estas variáveis, de forma que as respostas permaneçam próximas a seus valores nominais;
3. Determinar que valores associar a estas variáveis, de forma que a variabilidade das respostas seja mínima;
4. Determinar que valores associar a estas variáveis, de forma que a influência dos fatores não controláveis seja minimizada.

Um ponto importante a ser discutido ao longo deste capítulo diz respeito à estratégia a ser adotada para a realização dos experimentos. O fato de alguns sistemas envolverem um grande número de fatores a serem testados, pode resultar num excessivo número de experimentos. A idéia associada as diferentes estratégias que podem ser empregadas, é uma só: obter o máximo de informações com o mínimo de esforço.

Para exemplificar esta questão, considere um assíduo praticante de *windsurf* (ou prancha a vela), chegando á praia para velejar numa manha de domingo. Para melhorar seu desempenho, este atleta necessita tomar algumas decisões antes de se aventurar ao mar com seu equipamento. Alguns dos fatores, que ele imagina, podem fazer com que seu desempenho venha a ser razoável, glorioso ou um total fracasso são os seguintes:

1. O tipo de prancha utilizada: média (M) ou grande(G);
2. O tipo de vela utilizada: pequena (P) ou grande (G);
3. O tipo de quilha: pequena (P) ou grande(G);
4. As condições de vento no local: fraco/moderado (F) ou moderado/forte (M);
5. As condições do mar: calmo (C) ou agitado (A);
6. Sua noite/madrugada anterior: dormindo (D) ou festa (F).

Fica claro, que inúmeros outros fatores podem influenciar o desempenho do velejador, mas, para os propósitos deste exercício, estes seis são suficientes. Imaginando ainda, que para uma boa performance, a noite anterior à velejada tenha sido de um bom sono e descanso, o atleta passa a desconsiderar o fator seis. Como os cinco fatores restantes podem influenciar seu desempenho? A melhor forma de compreender tais influências é testar suas diversas combinações. É nesse ponto que as diferentes estratégias de experimentação podem e devem ser consideradas. No item 7.3 discutem-se as estratégias mais empregadas. Para assegurar sua perfeita compreensão apresentam-se no subtítulo 7.2 os principais elementos da terminologia empregada.

7.2. Terminologia

Esta seção explica os termos que são usados em análise e projeto experimental. Isto é feito com eventuais referencias ao exemplo do velejador de Windsurf, acima apresentado.

Os termos seguintes são freqüentemente usados na análise e projeto de experimentos:

- **Variável de Resposta:** Os resultados de um experimento são observados por meio dos valores (respostas) apresentados por algumas variáveis previamente eleitas, chamadas variáveis de resposta. Uma variável de resposta permite observar o comportamento de um sistema. No exemplo do velejador, algumas variáveis de resposta poderiam ser: a velocidade alcançada em determinado trecho de um percurso e a facilidade com que ele (velejador) encontra para a realização de manobras. Em determinados seguimentos (sistemas computacionais, por exemplo), o termo métrica é também empregado para denotar quem serão as variáveis responsáveis por demonstrar o desempenho alcançado por um sistema. Cabe citar também a utilização do termo variáveis de desempenho, no lugar de variáveis de resposta. No entanto, como as técnicas de projeto experimental são aplicáveis para qualquer tipo observação que se possa fazer de sistemas, não apenas do desempenho, o termo *resposta* é preferível em lugar de desempenho.
- **Fatores:** São as variáveis controladas pelo experimentador e que podem afetar a(s) variável (eis) de resposta. Cada fator apresenta, pelo menos duas alternativas. No exemplo envolvendo a decisão sobre velejar com o equipamento de windsurf, existem cinco fatores envolvidos no estudo.
- **Níveis:** Os valores que um fator pode assumir são chamados de níveis. Em outras palavras, cada nível do fator constitui uma alternativa para aquele fator. No exemplo do windsurf, cada um dos fatores apresentados possui dois níveis. Estes poderiam, se necessário serem ampliados. Por exemplo, as velas poderiam ser: pequena (P), média (M) ou Grande (G).
- **Replicações:** A repetição de todos ou de alguns experimentos é chamada de replicação. Por exemplo, se cada experimento em um estudo é repetido três vezes, é dito que o estudo teve três replicações.
- **Projeto:** Um projeto experimental consiste na especificação de uma estratégia de ação que resulta na determinação do número de experimentos (combinações entre fatores e seus níveis) a serem realizados, do número de replicações de cada experimento e da forma que os mesmos serão efetivados. Por exemplo, no estudo do windsurf, pode-se executar experimentos considerando todas as combinações possíveis para os níveis dos cinco

fatores. Para tanto, 2 x 2 x 2 x 2 x 2, ou $2^5 =$ 32 experimentos seriam necessários. Se cada um fosse repetido cinco vezes, uma total de 160 observações seria obtida.

7.3. Estratégias de Projetos Experimentais

Dentre as várias estratégias que podem ser adotadas para a execução de um projeto experimental, pode-se citar:

- Bom senso;
- Um fator por vez;
- Projeto fatorial completo;
- Projeto fatorial fracionário;
- Projeto fatorial com replicações.

7.3.1. Bom Senso

A primeira estratégia de experimentação é bastante empregada, principalmente, quando o experimentador possui um bom conhecimento do sistema, bem como sobre as diversas variáveis envolvidas no processo. Desta maneira, os procedimentos costumam ser iniciados compondo-se um quadro preliminar (valores atribuídos aos diversos fatores envolvidos) sobre o qual já se possui uma boa estimativa dos prováveis resultados.

Para o exemplo, o velejador considera o fato de existirem cinco fatores, cada um deles com duas diferentes alternativas. Com um cálculo rápido ele verifica a necessidade de realizar 32 experimentos para que possa combinar os cinco fatores com suas duas alternativas. No entanto, usando sua experiência prévia, o velejador parte para uma combinação de fatores que ele considera ideal. Como ele não pode controlar os fatores 4 e 5, sua atenção recai sobre os três primeiros fatores. Neste caso, o número necessário de experimentos cai de 32 para 8. Assim, considerando as condições do mar e do vento, o velejador decide empregar uma prancha de tamanho médio, uma vela pequena e uma quilha pequena. Uma vez no mar, ele verifica sua performance considerando, principalmente, sua velocidade em relação a outros velejadores e suas próprias condições para a realização de manobras básicas. Com este tipo de sentimento em relação a sua performance, ele realiza mais um ou dois experimentos e conclui sobre qual a melhor combinação de equipamentos (fatores 1, 2 e 3) lhe dará as melhores condições de velejo, diante do cenário que lhe é imposto (fatores não controláveis 4 e 5).

Esta estratégia de experimentação é muito empregada no dia a dia de engenheiros e pesquisadores. Em geral costuma trazer bons resultados, principalmente, quando associada a uma experiência prática, técnica e teórica razoável sobre o sistema e as condições de teste que o envolve. Obviamente, existem contra indicações ao emprego desta estratégia. Considere, por exemplo, a situação de um velejador com pouca experiência. Embora este tenha as mesmas oito combinações iniciais disponíveis, é possível que a partir de uma escolha inicial, ele tenha que realizar vários outros testes antes de decidir qual a melhor combinação. Além disso, é possível que um resultado inicial razoável o iniba de realizar outros experimentos. Tais contra indicações, não implicam que a estratégia do bom senso deixe de ser empregada. Bom senso sempre foi e será bem vindo em qualquer processo decisório, principalmente, quando empregado em conjunto com técnicas mais completas e que não dependam tão fortemente da experiência prática do experimentador.

7.3.2. Um Fator por Vez

Outra estratégia de experimentação muito utilizada é a de variar um fator por vez, mantendo os demais fixos. Esta estratégia também é conhecida como projeto experimental simples. A idéia é que o experimentador inicie com uma configuração típica para todos os fatores e observe o resultado. No experimento seguinte, um determinado fator escolhido assume outro valor (nível), enquanto os demais permanecem nos mesmos níveis anteriores. Quando se esgotarem todas as possibilidades de variação deste primeiro fator, este se torna fixo e um segundo passa a sofrer variações controladas. O processo segue até que todos os testes tenham sido realizados, considerando-se todos os fatores e seus respectivos níveis.

Neste tipo de estratégia as análises são, geralmente, efetuadas a partir de tabelas ou gráficos construídos com os resultados dos experimentos. Na figura 7.2 pode-se observar os resultados do desempenho (D) a partir dos experimentos realizados pelo velejador de windsurf, na tentativa de encontrar a melhor combinação entre os três fatores considerados relevantes e controláveis. A análise dos gráficos leva o decisor a considerar o emprego de uma prancha média e de uma vela grande, para qualquer tipo de quilha disponível, uma vez que, segundo sua medição, este fator não proporcionou diferenças observáveis sobre seu desempenho. Neste caso, as conclusões do velejador consideram o cenário definido pelos fatores não controláveis (condições de vento no local e condições do mar).

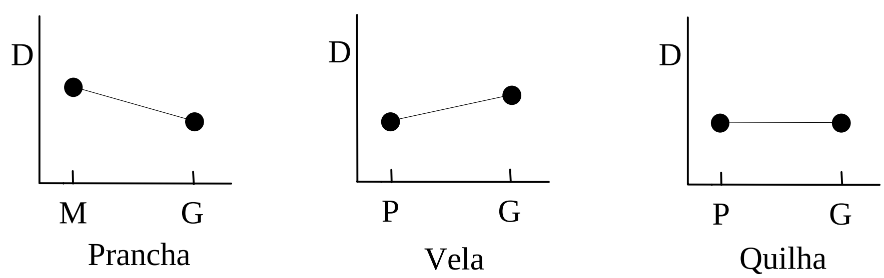

Figura 7.2: Resultados de um projeto experimental simples para o exemplo do windsurf.

Embora muito popular por ser intuitiva esta estratégia desconsidera um elemento fundamental, quando se trata de experimentação, a interação entre os fatores. Por interação entende-se a mútua influência entre dois ou mais fatores. Assim, dados dois fatores A e B diz-se que eles interagem, se o efeito de um depende do nível de outro. Por exemplo, a tabela 7.1 mostra o desempenho de um sistema com dois fatores. Quando o fator A é trocado do nível A_1 para o nível A_2, o desempenho é incrementado em 2, independente do nível do fator B. Neste caso não há interação. A tabela 7.2 mostra outra possibilidade. Neste caso, quando o fator A é trocado do nível A_1 para o nível A_2, o desempenho aumentou em 2 ou em 3 dependendo se B está no nível B_1 ou no nível B_2, respectivamente. Neste caso, diz-se que os dois fatores interagem.

	A_1	A_2
B_1	3	5
B_2	6	8

Tabela 7.1: Fatores sem interação

	A_1	A_2
B_1	3	5
B_2	6	9

Tabela 7.2: Fatores com interação

Uma representação gráfica deste exemplo é dada nas figuras 7.3 e 7.4. Na primeira as linhas estão em paralelo, indicando que não há interação. No segundo caso as linhas são divergentes, indicando que há interação.

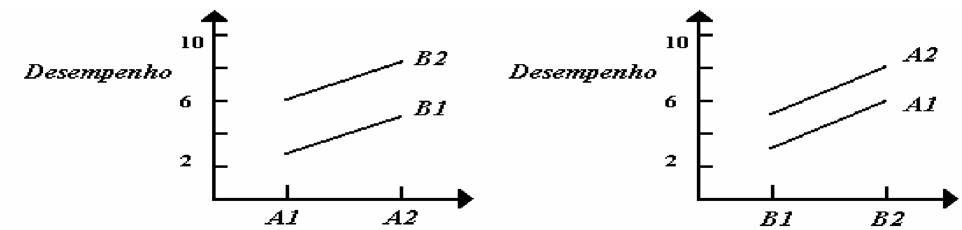

Figura 7.3: Representação gráfica de fatores sem interação

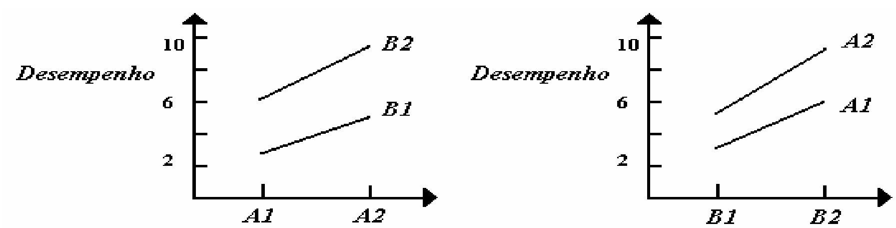

Figura 7.4: Representação gráfica de fatores com interação

Uma vez que a interação entre fatores é muito comum, o emprego desta técnica não é recomendado, pois não faz o melhor uso do esforço realizado. Em outras palavras, pode, na maioria das situações não ser estatisticamente eficiente. Se os fatores têm interação, este tipo de projeto pode levar a conclusões erradas. Por exemplo, se o efeito da vela depende do tamanho do tipo de quilha empregada, a combinação ótima não poderá ser determinada até que todas possibilidades sejam testadas. Este projeto pode, portanto, não ser recomendado.

Quando, comprovadamente, não existem interações entre fatores, esta estratégia pode ser uma opção prática.

Dados K fatores, com o *i-ésimo* fator tendo n_i níveis, um projeto simples de análise do tipo “um fator por vez” requer somente n experimentos, onde:

$$n = 1 + \sum_{i=1}^{k} (n_i - 1) \quad (7.1)$$

Exemplo 1.

Considere um projeto fatorial com dois fatores A e B, cada um deles com dois níveis: A_1, A_2 e B_1, B_2, respectivamente. Considerando a fórmula 7.1, quantos experimentos são necessários para analisar a influência destes fatores sobre uma determinada variável de resposta, considerando a não existência de interações entre os dois fatores?

Dado que o número de experimentos n é dado por $n = 1 + \sum_{i=1}^{k} (n_i - 1)$ onde, k =2 e n_i = 2, obtém-se: n = 1 + (2-1) + (2-1) = 3 experimentos.

Os seguintes experimentos poderiam ser determinados: A_1B_1, A_2B_1 e A_1 ou A_2 com B_2.

Exemplo 2.

Considere um projeto fatorial com três fatores A, B e C, cada um deles com dois níveis: A_1, A_2, e B_1, B_2 e C_1 e C_2, respectivamente. Considerando a mesma fórmula 7.1, quantos experimentos são necessários para analisar a influência destes fatores sobre uma determinada variável de resposta, considerando que neste caso também não existam interações entre os três fatores?

Dado que o número de experimentos n é dado por $n = 1 + \sum_{i=1}^{k} (n_i - 1)$ onde, k =3 e n_i = 2, obtém-se: n = 1 + (2-1) + (2-1) + (2-1) = 4 experimentos.

Um exemplo destes quatro experimentos poderia ser: $A_1B_1C_1$, $A_2B_1C_1$, $A_2B_2C_1$ e $A_2B_2C_2$.

7.3.3. Projeto Fatorial Completo

Para lidar com a questão de avaliar as interações entre os fatores é necessário realizar o que os estatísticos chamam de projeto fatorial completo. Neste tipo de estratégia, empregam-se todas as combinações possíveis entre os fatores e seus diferentes níveis. Os fatores (em seus diversos níveis) são variados juntos, ao invés de um por vez, como na estratégia anterior. Um estudo de desempenho com k fatores, com o i-ésimo fator tendo n_i níveis, requer n experimentos, onde:

$$n = \prod_{i=1}^{k} n_i \quad (7.2)$$

Exemplo 3.

Considere um projeto fatorial com três fatores A, B e C, cada um deles com dois níveis: A_1, A_2, e B_1, B_2 e C_1 e C_2, respectivamente. Considerando a fórmula 7.2, quantos experimentos são necessários para analisar a influência destes fatores sobre uma determinada variável de resposta, considerando que, neste caso, podem existir interações entre os três fatores?

O número de experimentos n é dado por $n = \prod_{i=1}^{k} n_i$, onde, k =3 e n_i = 2, obtém-se:

n = 2 x 2 x 2 = 2^3 = 8 experimentos. Os experimentos seriam as combinações:

A1	B1	C1
A1	B1	C2
A1	B2	C1
A1	B2	C2
A2	B1	C1
A2	B1	C2
A2	B2	C1
A2	B2	C2

O número de fatores e seus correspondentes níveis permitem a concepção de diferentes tipos de projetos fatoriais. Alguns destes (os mais empregados) serão objeto de analises nas sessões subseqüentes deste capítulo.

Voltando a idéia de representar graficamente o processo experimental, considere novamente o exemplo do velejador. Para simplificar, admita apenas dois fatores, tipo de vela e tipo de prancha, cada um deles com seus dois níveis já conhecidos. A figura 7.5 é uma representação gráfica do projeto experimental completo resultante, isto é, 2^2 = 4 experimentos.

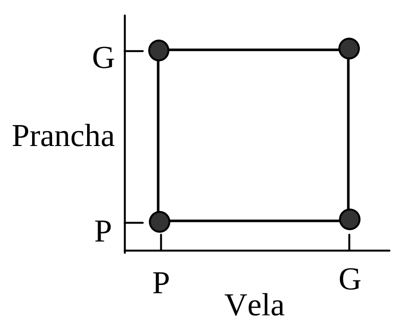

Figura 7.5: Representação geométrica de um projeto fatorial completo considerando dois fatores (Prancha e Vela) cada um com dois possíveis níveis (P e G).

É possível estender esta mesma idéia para mais fatores e níveis. Incluindo-se um terceiro fator a ser analisado pelo velejador, o tipo de quilha, também com dois níveis, implica em um aumento do número de experimentos de quatro para oito. A figura 7.6 mostra a representação geométrica dos oito experimentos.

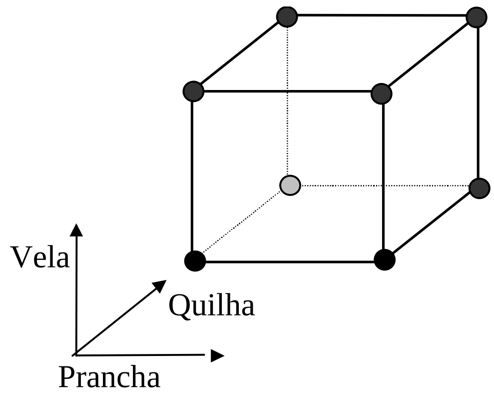

Figura 7.6: Representação geométrica de um projeto fatorial completo considerando três fatores (prancha, vela e quilha) cada um com dois possíveis níveis (P e G).

A vantagem de um projeto fatorial completo é que todas as combinações possíveis são examinadas. Pode-se encontrar o efeito de todos fatores incluindo fatores secundários e suas interações. A principal desvantagem é o custo deste tipo de estudo. Este pode ser muito alto, especialmente quando se considera a possibilidade de cada experimento ser repetido várias vezes. Existem três caminhos para reduzir o número de experimentos e, portanto, os custos associados:

- Reduzir o número de níveis por fator;
- Reduzir o número de fatores;
- Usar projetos fatoriais fracionários.

A primeira alternativa é especialmente recomendada. Em alguns casos, é possível tentar somente dois níveis por fator e determinar a importância relativa de cada fator. Um projeto fatorial completo no qual cada um dos k fatores é usado em dois níveis requer 2^k experimentos. Este é um projeto muito popular e é chamado de projeto 2^k. Depois da lista de fatores estar substancialmente reduzida, pode-se tentar mais níveis por fator. A terceira alternativa, projeto fatorial fracionário, é descrita na próxima seção.

7.3.4. Projeto fatorial fracionário

Em determinadas situações, o número de experimentos requeridos para um projeto fatorial completo muito grande. Por exemplo, um experimento com 10 fatores, cada um deles com apenas dois níveis implicará em $2^{10} = 1024$ experimentos. De forma semelhante, um experimento com quatro fatores onde cada um deles tenha 3, 4, 5, e 6 níveis, respectivamente, terá 3 x 4 x 5 x 6 = 360 experimentos. Estes números podem impossibilitar economicamente a execução do projeto, quando se considera o tempo e os recursos necessários a sua execução.

Felizmente, a matemática e a estatística permitem que se possa realizar boas estimativas sobre as variáveis de resposta, realizando apenas parte ou fração do total de experimentos de um projeto fatorial completo. Este tipo de estratégia é conhecido como projeto fatorial fracionário.

Para exemplificar esta situação, observe a representação geométrica da figura 7.7. Nesta estão representados os mesmos oito experimentos (2^3) do projeto fatorial completo relativo ao problema do velejador. No projeto visto na figura, quatro são os fatores analisados.

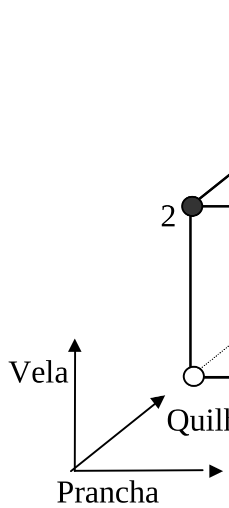

Figura 7.7: Projeto Fatorial Fracionário com quatro fatores.

Além dos três fatores anteriores (prancha, vela e quilha), um quarto fator, Experiência do Velejador, também com dois níveis (Novato e Sênior), foi incluído. Considerando um projeto fatorial completo, quatro fatores, com dois níveis cada um deles, resultam em $2^4 = 16$ experimentos. Na figura 7.7, os dois cubos apresentam nos seus cantos bolas com preenchimento (pretas) e sem preenchimento (brancas). As que estão preenchidas, simbolizam as combinações de níveis e fatores que terão o correspondente experimento realizado (os números indicam o experimento mostrado na tabela 7.3). O quarto fator, experiência, assumirá o nível Novato quando combinado com os valores de prancha, vela e quilha, indicados no primeiro cubo e assumirá o nível Sênior quando combinado com os demais valores de prancha, vela e quilha, indicados no segundo cubo. Observe então, que do total necessário de 16 experimentos, apenas oito serão realizados. Simbolicamente, representa-se este tipo de estratégia por 2^{4-1} experimentos.

Exemplo 4.

A tabela 7.3 mostra de forma mais clara as diferentes combinações entre os quatro fatores considerando apenas 8 experimentos.

Experimento	**Prancha**	**Vela**	**Quilha**	**Experiência**
1	G	P	P	Novato
2	P	G	P	Novato
3	P	P	G	Novato
4	G	G	G	Novato
5	P	P	P	Sênior
6	G	G	P	Sênior
7	P	G	G	Sênior
8	G	P	G	Sênior

Tabela 7.3: As diferentes combinações do projeto fatorial fracionário 2^{4-1}.

Embora o nível de informações obtido numa estratégia como a apresentada na tabela 7.3 não seja tão completa, quanto àquela que se obteria no caso da adoção de um projeto fatorial completo, é possível conseguir boas informações sobre os efeitos causados pelos principais fatores, bem como, alguma informação sobre suas possíveis interações. Este tipo de projeto será discutido posteriormente com um maior nível de detalhamento.

7.3.5. Projeto Fatorial com Replicações

Quando se está realizando experimentos associados a determinado projeto, é comum que ao repetir determinado experimento, o analista observe diferenças entre um ensaio e outro, embora todos os cuidados em procurar manter as mesmas condições, sob todos os aspectos.

Na adoção de modelos computacionais para simular o ambiente onde se realizam os experimentos, embora o controle seja muito mais facilitado (se está tratando de um modelo lógico/matemático), sabe-se que, pelo fato destes modelos incluírem variáveis aleatórias, a cada realização de um experimento (replicação de uma simulação) a possibilidade de se obter diferentes resultados para as variáveis de resposta é grande.

As diferenças observadas nas respostas de cada ensaio são chamadas pelos estatísticos de *erros experimentais*. Tais erros podem ser quantificados pelo repetição das medidas realizadas para cada combinação dos fatores e seus níveis. Desta maneira, na estratégia denominada projeto fatorial com replicações, *r* replicações (ou repetições) de cada ensaio são realizadas. Novamente é importante lembrar que se o número de experimentos definidos pela combinação de fatores e níveis for grande, neste tipo de estratégia este número estará sendo multiplicado por *r* repetições. Assim como as demais estratégias, esta também estará sendo explorada com mais profundidade nas sessões subseqüentes.

7.4. Projetos Fatoriais 2^k

No item 7.3 vimos as diferentes estratégias de experimentação que um analista pode empregar para estudar os efeitos dos diferentes fatores sobre uma variável de resposta. Considerando a classe das estratégias que envolvem projetos fatoriais, uma delas, em especial, merece destaque: a dos projetos fatoriais tipo 2^k. Este tipo de estratégia é empregado para determinar o efeito de k fatores, cada um dos quais com duas alternativas ou níveis. Esta classe de estratégia é importante por vários motivos:

1. Sua simplicidade de implementação;
2. Sua facilidade de compreensão por parte dos analistas menos experientes;
3. A facilidade para entender os efeitos de cada fator sobre as variáveis de resposta;
4. A possibilidade de ordenação dos fatores por sua ordem de importância.

Quando se estudam efeitos de fatores sobre respostas, é muito comum que se observe uma certa continuidade unidirecional no comportamento destas variáveis, na medida em que se alteram os valores dos fatores (aumento ou redução). Por exemplo, imagine um experimento com dois fatores A e B, cada um com cinco níveis e uma variável de resposta R. Ao se realizar observações, verifica-se que R cresce continuamente na medida em que se incrementam os valores dos dois fatores, de seus níveis mais baixos para os mais altos. Se isto de fato ocorre, realizar 25 experimentos para verificar estes efeitos pode significar um desperdício de recursos. Em tais casos pode-se iniciar experimentando os níveis mínimos e máximos (dois níveis portanto) de cada fator. Esta abordagem ajudará a decidir se a diferença de desempenho é significante o bastante para justificar um exame mais detalhado. Para o caso do exemplo, isto significaria uma redução de 25 para 4 experimentos.

Para uma melhor compreensão sobre os conceitos envolvidos em projetos tipo 2^k, é interessante que se inicie pelo seu caso mais simples, isto é, com somente dois fatores (k=2). Uma vez que se tenha compreendido os principais conceitos para este caso, faz-se, adiante, uma generalização para um número maior de fatores.

7.4.1. Projeto Fatorial 2^2

Um projeto experimental 2^2 é um caso especial de um projeto fatorial 2^k com k=2. Neste caso, há dois fatores com dois níveis. Trata-se do caso mais simples de um projeto fatorial. Para estudar toda a sua mecânica, emprega-se um exemplo ilustrativo.

Exemplo 5.

Considere o problema de estudar o impacto do tamanho da memória principal e o tamanho da memória *cache* sobre a performance de um servidor de arquivos de uma rede local de computadores (LAN). Cada um destes dois fatores possui dois níveis, conforme indicado na tabela 7.4. A performance do servidor de arquivos foi medida considerando como variável de resposta o seu *throughput* (neste caso, número de operações de transferência de arquivos por segundo).

Memória *Cache* (kbytes)	Memória Principal 128 Mbytes	Memória Principal 256 Mbytes
256	500	1100
512	900	1900

Tabela 7.4: Performance do servidor de Arquivos (Operações/seg.)

Considere a obtenção de uma equação que permita o cálculo dos efeitos de cada um dos fatores sobre a variável de resposta. Conforme um desdobramento semelhante desenvolvido por Raj Jain (JAIN, 1991), pode-se considerar e definir as variáveis X_A (memória principal) e X_B (memória *cache*), como segue:

X_A = -1 se 128 Mbytes de memória principal.
X_A = 1 se 256 Mbytes de memória principal.
X_B = -1 se 256 kbytes de memória cache.
X_B = 1 se 512 kbytes de memória cache.

A performance *R* em (operações/seg.) pode agora ser avaliada em X_A e X_B usando-se um modelo de regressão não linear na forma:

$$R = q_0 + q_A.x_A + q_B.x_B + q_{AB}.x_A.x_B \qquad (7.3)$$

Substituindo as quatro observações (r_1, r_2, r_3 e r_4) no modelo, obtém-se as seguintes equações:

$$\begin{aligned} 500 &= q_0 - q_A - q_B + q_{AB} \\ 1100 &= q_0 + q_A - q_B - q_{AB} \\ 900 &= q_0 - q_A + q_B - q_{AB} \\ 1900 &= q_0 + q_A + q_B + q_{AB} \end{aligned}$$

Estas quatro equações podem ser resolvidas unicamente para as quatro incógnitas e a equação de regressão fica:

$$R = 1100 + 400x_A + 300x_B + 100x_Ax_B$$

O resultado é interpretado como segue. O desempenho médio é de 1100 operações/seg; o efeito da memória principal é 400 operações /seg; o efeito da memória cache é de 300 operações /seg. e a interação entre memória principal e a cache respondem por 100 operações /seg.

7.4.1.1. Computação dos Efeitos dos Fatores em Projetos 2^2

Em geral qualquer projeto 2^2 pode ser analisado usando-se o método empregado no exemplo anterior. No caso geral, suponha que r_1, r_2, r_3, e r_4 representem as quatro respostas observadas. Para o caso de projetos 2^2, a correspondência entre os níveis dos fatores e as respostas r_1, r_2, r_3, e r_4 pode ser mais bem compreendida observando-se a tabela 7.5.

Experimento	X_A	X_B	*R*
1	-1	-1	$r_1 = 500$
2	1	-1	$r_2 = 1100$
3	-1	1	$r_3 = 900$
4	1	1	$r_4 = 1900$

Tabela 7.5: Correspondência entre níveis, fatores e respostas de um projeto 2^2

O modelo de regressão para um projeto 2^2 é o mesmo verificado anteriormente na equação 7.3.. Desta forma, substituindo-se as quatro observações no modelo e os valores dos coeficientes, têm-se:

$$\begin{aligned} r_1 &= q_0 - q_A - q_B + q_{AB} \\ r_2 &= q_0 + q_A - q_B - q_{AB} \\ r_3 &= q_0 - q_A + q_B - q_{AB} \\ r_4 &= q_0 + q_A + q_B + q_{AB} \end{aligned}$$

Resolvendo estas equações para q_i's, obtêm-se:

$$q_0 = \tfrac{1}{4}(r_1 + r_2 + r_3 + r_4)$$
$$q_A = \tfrac{1}{4}(-r_1 + r_2 - r_3 + r_4)$$
$$q_B = \tfrac{1}{4}(-r_1 - r_2 + r_3 + r_4)$$
$$q_{AB} = \tfrac{1}{4}(r_1 - r_2 - r_3 + r_4)$$

As expressões para q_A, q_B, e q_{AB} são claramente combinações lineares das respostas, com a soma dos coeficientes igual zero. Estas expressões são chamadas pelos estatísticos de *contrastes*.

Uma interpretação geométrica para os efeitos dos fatores (efeitos principais) sobre a variável throughput (*R*) permite uma melhor compreensão do conceito de *contrastes*. Considere os gráfico 7.1, 7.2 e 7.3. Os experimentos e seus resultados são representados num sistema de eixos cartesianos, onde cada fator é alocado a um eixo. No caso, como se trata de apenas dois fatores, o espaço definido é um plano. Os quatro experimentos (respostas) aparecem nos cantos de um quadrado. Atribuem-se sinais algébricos (+/-) aos valores empregados (níveis) de cada um dos fatores, conforme também pode ser visto na tabela 7.4. O que se pode então observar são os *contrastes*, ou melhor, as diferenças entre valores situados em arestas opostas e perpendiculares ao eixo do fator em questão. Tais diferenças ilustram os efeitos de cada um dos fatores (e respectivos níveis) sobre a variável de resposta.

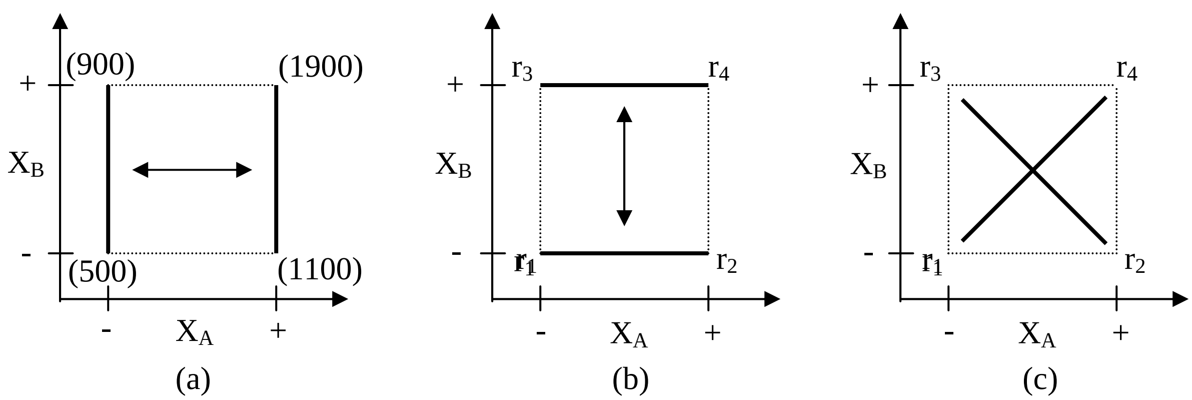

Figura 7.8: Efeitos dos fatores num planejamento 2^2.

Os efeitos principais são vistos nos gráficos das figuras 7.8 (a) e 7.8 (b). Para um melhor entendimento do conceito, na figura 7.8 (a) os símbolos r_i foram substituídos pelos valores alcançados nos ensaios. A figura 7.8 (c) é a interpretação geométrica do efeito da interação entre os fatores. Seu contraste é observado entre as duas diagonais.

Para ilustrar o efeito das interações entre os fatores, observe a figura 7.9, que considera os valores do exemplo 5.

Verifica-se, claramente, o não paralelismo entre os dois segmentos de reta. Tal situação indica, claramente, a existência de interação entre os fatores, conforme ilustrado na figura 7.4, do item 7.3.2. De fato, sabe-se, de ensaios e da prática, que o efeito da memória principal de um computador sobre a performance do mesmo, costuma ser potencializado na medida do aumento de sua memória *cache*.

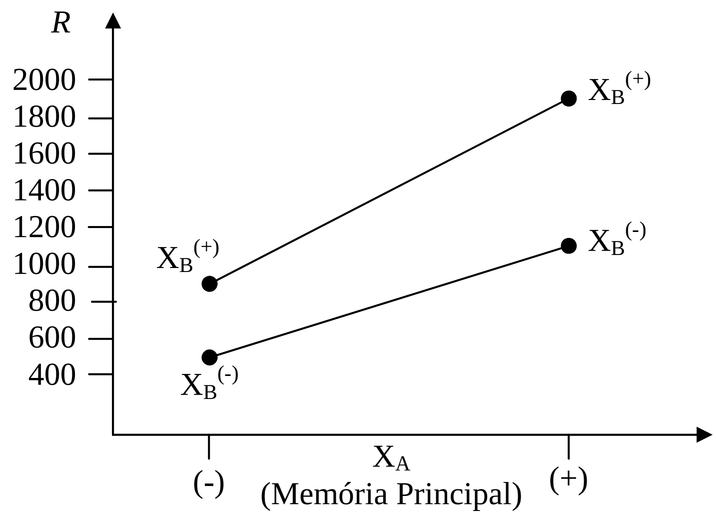

Figura 7.9: Projeto fatorial 2^2 com interação entre os fatores.

7.4.1.2. Método da Tabela de Sinais para o Cálculo dos Efeitos de Projetos 2^2

No exemplo desenvolvido no item anterior, os efeitos foram calculados deduzindo-se uma equação de regressão não-linear (equação 7.3). Posteriormente, verificou-se que ao substituir-se os coeficientes da equação por representações simbólicas de seus dois níveis (valores –1 e +1), é possível obter-se um conjunto de equações que são combinações lineares das respostas.

Com base naquela dedução, nesta seção apresenta-se um método bastante simples para calcular os efeitos de fatores sem maiores dificuldades. O método é conhecido como Método da tabela de sinais. Para descreve-lo utiliza-se o mesmo exemplo de desempenho do servidor de rede que se vem tratando, numa adaptação da descrição feita por Neto, Scarmínio e Bruns (Neto, 1995).

Considere inicialmente as colunas X_A e X_B da tabela 7.5, de correspondência entre os dois fatores e seus dois níveis e as respectivas respostas do projeto Servidor de Arquivos. Monte uma nova tabela (7.6) acrescentando a estas duas colunas outras duas: uma com valores iguais a um positivos (coluna I) e outra cujos valores correspondem ao produto das colunas X_A e X_B (coluna X_{AB}). A tabela 7.6 apresenta o resultado desta operação.

I	X_A	X_B	X_{AB}
1	-1	-1	+1
1	+1	-1	-1
1	-1	+1	-1
1	+1	+1	+1

Tabela 7.6: Primeira etapa da tabela dos sinais para o projeto 2^2

Conforme foi visto na seção anterior, a tabela 7.6 corresponde aos coeficientes de um conjunto de equações que são combinações lineares das respostas. Sendo assim, é possível, calcular o valor dos efeitos de qualquer dos fatores realizando-se operações algébricas apropriadas. Desta forma, pode-se calcular os efeitos de cada fator, a menos de uma divisão, pelo produto escalar do seu vetor de coeficientes (cada coluna da tabela 7.6) pelo vetor de respostas. Assim, por exemplo, para se calcular o efeito de X_A (memória principal), faz-se o produto escalar:

$$X_A = (1/4) * \left\{ \begin{bmatrix} -1 & 1 & -1 & 1 \end{bmatrix} * \begin{bmatrix} 500 \\ 1100 \\ 900 \\ 1900 \end{bmatrix} \right\} = 400$$

O valor obtido a partir do produto escalar é dividido pelo número experimentos ou de combinações necessárias ao projeto, isto é, $2^k = 4$ experimentos.

Considerando a representação de matrizes e vetores em negrito, a representação matricial da operação acima realizada seria:

$$X_A = (1/4) * (\boldsymbol{X_A})^t.\boldsymbol{R}$$

onde $(\boldsymbol{X_A})^{\mathbf{t}}$ é o vetor linha obtido pela transposição do vetor coluna $\boldsymbol{X_A}$ e $\boldsymbol{R}$ é o vetor resposta.

Para o caso geral, isto é, para se calcular os efeitos de todos os fatores e da média, considera-se a própria matriz dos coeficientes ($\boldsymbol{X}$), multiplicando-se sua transposta $(\boldsymbol{X})^{\mathbf{t}}$ pelo vetor dos resultados $\boldsymbol{R}$. Cada valor obtido é então dividido por 2^k. Desta maneira obtém-se:

$$X^t R = \begin{bmatrix} 1 & 1 & 1 & 1 \\ -1 & 1 & -1 & 1 \\ -1 & -1 & 1 & 1 \\ 1 & -1 & -1 & 1 \end{bmatrix} * \begin{bmatrix} 500 \\ 1100 \\ 900 \\ 1900 \end{bmatrix} = \begin{bmatrix} 4400 \\ 1600 \\ 1200 \\ 400 \end{bmatrix}$$

Para o cálculo final dos efeitos e da média, divide-se cada um dos elementos do vetor por $2^k = 4$. Assim, obtém-se o vetor resposta com os seguintes valores:

$$\begin{bmatrix} Média \\ X_A \\ X_B \\ X_{AB} \end{bmatrix} = \begin{bmatrix} 1100 \\ 400 \\ 300 \\ 100 \end{bmatrix}$$

O resultado é interpretado da mesma forma como foi feito na seção 7.4.1, isto é:

- Média = 1100 operações/seg;
- Efeito da memória principal é 400 operações/seg;
- Efeito da memória cache é de 300 operações/seg;
- Efeito da interação entre memória principal e a cache 100 operações/seg.

O procedimento acima desenvolvido pode ser facilmente adaptado para o formato de tabelas. Uma vez assim formatado, o emprego de planilhas eletrônicas facilita seu emprego e difusão. Na tabela 7.7 (uma extensão da tabela 7.6) pode-se observar os mesmos resultados obtidos com as operações matriciais do exemplo anterior. Fórmulas internas executam as mesmas operações realizadas. Os vetores resultantes das multiplicações encontram-se nas duas últimas linhas da tabela.

I	$\mathbf{X_A}$	$\mathbf{X_B}$	$\mathbf{X_{AB}}$	***R***
1	-1	-1	1	500
1	1	-1	-1	1100
1	-1	1	-1	900
1	1	1	1	1900
4400	1600	1200	400	←Total
1100	400	300	100	←Total/4

Tabela 7.7: Tabela de sinais para cálculo de efeitos do projeto 2^2

7.4.1.3. Distribuição da Variação para Projetos 2^2

Uma vez que se tenha calculado o valor dos efeitos provocados pelos fatores, de forma individual ou de suas interações, é importante determinar a contribuição de cada um deles sobre as variações nas variáveis de resposta. A importância de um fator é medida pela proporção do total da variação na variável de resposta que é por ele explicado. Assim, se dois fatores apresentam, respectivamente, 80% e 5% de variação da resposta, o segundo fator pode ser considerado menos importante que o primeiro. Se a contribuição de determinado fator considerada pequena (menos do que 5%, por exemplo), tal fator poderá ser descartado do estudo.

A fórmula 7.4, abaixo definida, permite calcular o se chama de Variação Total de *R*. Esta variação também é conhecida como Soma dos Quadrados Totais (SQT). Tal valor compreende uma soma das contribuições de cada um dos fatores e de suas interações.

$$\text{Variação total de } R = \text{SQT} = \sum_{i=1}^{2^2} \left(r_i - \bar{r}\right)^2 \tag{7.4}$$

onde $\bar{r}$ compreende a média das respostas dos quatro experimentos realizados.

A Soma dos Quadrados Totais ou SQT pode ser definida como segue:

$$\text{SQT} = \text{SQA} + \text{SQB} + \text{SQAB} \tag{7.5}$$

As três parcelas do lado direito da equação representam a porção da variação total explicados pelos efeitos de A, B e da interação AB, respectivamente. De acordo com as fórmulas 7.4 e 7.5, para um projeto 2^2, a variação pode então ser dividida em três partes:

$$\text{SQT} = 2^2 q_A^2 + 2^2 q_B^2 + 2^2 q_{AB}^2 \tag{7.6}$$

onde $2^2 q_A^2$ é a porção de SQT que é explicado pelo fator A. Similarmente, SQB é $2^2 q_B^2$ e SQAB (devido para interação AB) é $2^2 q_{AB}^2$. Portanto, a contribuição individual de um fator ou de uma interação de fatores pode ser expressa com uma fração, por exemplo:

$$\text{Fração da variação explicada pelo fator } A = \frac{SQA}{SQT}$$

Quando expressa como uma porcentagem, esta fração provê um modo simples para medir a importância de um fator. Os fatores que explicam uma alta porcentagem da variação são considerados os mais importantes.

Exemplo 6.

Para o exemplo do servidor de rede, calcula-se as frações, ou percentuais de contribuição de variação do throughput, de cada um dos fatores e de suas interações, usando as fórmulas 7.4 e 7.6.

Empregando-se a fórmula 7.4.

$\bar{r}$ = ¼ (500 + 1100 + 900 + 1900) = 4400/4 = 1100

$$\text{Variação Total de } R = \sum_{i=1}^{2^2} \left(r_i - \bar{r}\right)^2 = (500\text{-}1100)^2 + (1100\text{-}1100)^2 + (900–1100)^2 + (1900–1100)^2$$

$$\text{Variação Total de } R = 600^2 + 0^2 + 200^2 + 800^2 = 1.040.000$$

Chega-se ao mesmo resultado empregando-se a fórmula 7.6:

$$\text{SQT} = \text{SQA} + \text{SQB} + \text{SQAB} = 2^2 q_A^2 + 2^2 q_B^2 + 2^2 q_{AB}^2 = 4 \text{ x } 400^2 + 4 \text{ x } 300^2 + 4 \text{ x } 100^2$$

$$\text{SQT} = \text{SQA} + \text{SQB} + \text{SQAB} = 1.040.000$$

Para o cálculo das frações, faz-se:

$$\text{Fração da variação explicada pelo fator } A = \frac{SQA}{SQT} = 4 \text{ x } 400^2 / 1.040.000 = 0{,}6154$$

Fração da variação explicada pelo fator $B = \frac{SQB}{SQT}$ = 4 x 300^2 / 1.040.000 = 0,3461

Fração da variação explicada pela interação $AB = \frac{SQAB}{SQT}$ = 4 x 100^2 / 1.040.000 = 0,0384

Portanto, a contribuição sobre a variação da variável de resposta throughput, do servidor de arquivos da rede local, pode ser atribuída ao fator A (memória principal) com 61,54 % da variação, ao fator B (memória cache) com 34,61 % da variação e a interação entre os dois fatores, com 3,84 % da variação.

7.4.2. Projetos Fatoriais 2^k

Um projeto experimental 2^k é usado para determinar o efeito de k fatores, onde cada um deles possui duas alternativas ou níveis. Na seção 7.4.1, apresentou-se com detalhes o caso de dois fatores com dois níveis, isto é, k = 2. Nesta seção, generaliza-se a análise para mais de dois fatores.

Uma vez compreendidos os conceitos vistos para os projetos 2^2, pode-se facilmente estendê-los para projetos nos quais k assume valores maiores ou iguais a três. O método de tabela de sinais desenvolvido para o caso de k = 2 pode também ser empregado para o caso de mais níveis.

Considerando um estudo com três fatores, A, B e C, cada um deles com dois níveis, um projeto fatorial completo terá um total de 2^3 = 8 experimentos, que correspondem as oito linhas de uma tabela padronizada de sinais, como a tabela 7.8 abaixo.

Experimento	Média	A	B	C	AB	AC	BC	ABC	R
1	1	-1	-1	-1	1	1	1	-1	r_1
2	1	1	-1	-1	-1	-1	1	1	r_2
3	1	-1	1	-1	-1	1	-1	1	r_3
4	1	1	1	-1	1	-1	-1	-1	r_4
5	1	-1	-1	1	1	-1	-1	1	r_5
6	1	1	-1	1	-1	1	-1	-1	r_6
7	1	-1	1	1	-1	-1	1	-1	r_7
8	1	1	1	1	1	1	1	1	r_8

Tabela 7.8: Tabela padronizada de sinais para projetos 2^3

Esta tabela conterá também oito colunas de valores unitários com sinais positivos ou negativos, correspondentes aos dois níveis de cada fator. A primeira delas é dedicada ao cálculo da média, com todos os valores positivos. As três colunas seguintes são voltadas ao cálculo dos efeitos principais, isto é, dos fatores A, B e C. Cada uma destas três colunas inicia por um valor negativo (-) e depois os sinais se alternam. Um a um na primeira coluna, - + - + ..., dois a dois na segunda, - - + + ... e, finalmente, quatro valores negativos e quatro positivos na terceira destas colunas. Se houvesse um quarto fator, sua coluna teria oito valores negativos e oito positivos, uma vez que o projeto teria 16 experimentos. As quatro colunas seguintes correspondem aos valores necessários aos cálculos dos efeitos das interações entre os três fatores principais. As três primeiras, destas quatro colunas, estão associadas às interações de dois fatores, como por exemplo AB ou AC. A última destas colunas refere-se a interação entre os três fatores, isto é, ABC. Completando a tabela, aparece a coluna dedicada aos valores $r_{i's}$ da variável de resposta *R*, obtidos nos ensaios.

No exemplo a seguir, apresenta-se um caso de k = 3 níveis, cada um deles com dois fatores. Emprega-se o modelo da tabela de sinais 7.8 para o cálculo dos efeitos principais e das interações. A partir destes efeitos, calcula-se as contribuições dos fatores sobre a variação na variável de resposta.

Exemplo 7.

Considere uma ampliação do problema da performance de um servidor de arquivos de uma rede local de computadores (LAN). Além de estudar o impacto do tamanho da memória principal (Fator A) e o tamanho da memória *cache* (Fator B), verificou-se a necessidade de ampliar a pesquisa, incluindo-se agora um terceiro fator (C): disco rígido ou HD (*Hard Disc*).

Duas alternativas são propostas para este fator: Disco A (disco atualmente utilizado no servidor) e Disco B (nova alternativa). Para diferenciá-los, considera-se três características básicas destes equipamentos: *Seektime*, Latência Rotacional e Taxa de Transferência. A primeira característica se refere ao tempo médio necessário para posicionar o cabeçote de leitura sobre o local onde se encontram os dados solicitados (quanto menor, melhor). A segunda característica indica um período de tempo decorrido entre o fim do posicionamento do cabeçote de leitura e o efetivo início da transferência dos dados (quanto menor, melhor). Finalmente, a última característica trata da capacidade do disco em transferir dados para a memória principal (quanto maior, melhor). Os valores associados aos dois discos são descritos na tabela 7.9 abaixo.

	Características dos Discos		
Discos	***SeekTime***	**Latência Rotacional**	**Taxa Transferência**
A	12	6,2	12
B	9	4,5	20

Tabela 7.9: Características das alternativas do fator Disco

A performance do servidor de arquivos foi medida considerando-se a mesma variável de resposta, isto é, o *throughput* (número de operações de transferência de arquivos por segundo). Como o projeto é do tipo 2^3, faz-se necessário um total de oito experimentos. A tabela 7.10 apresenta o resultado dos ensaios, considerando os três fatores e seus níveis.

	Memória Principal 128 MB		**Memória Principal 256 MB**	
Cache	**Disco A**	**Disco B**	**Disco A**	**Disco B**
256 KB	500	630	1100	1420
512 KB	900	1180	1900	2460

Tabela 7.10: Resultados dos oito experimentos do projeto 2^3.

A partir dos resultados obtidos nos ensaios com o servidor é possível preencher a coluna *R* de resultados. A tabela 7.11 apresenta a tabela de sinais completa.

Experimento	***Média***	***A***	***B***	***C***	***AB***	***AC***	***BC***	***ABC***	***R***
1	1	-1	-1	-1	1	1	1	-1	500
2	1	1	-1	-1	-1	-1	1	1	1100
3	1	-1	1	-1	-1	1	-1	1	900
4	1	1	1	-1	1	-1	-1	-1	1900
5	1	-1	-1	1	1	-1	-1	1	630
6	1	1	-1	1	-1	1	-1	-1	1420
7	1	-1	1	1	-1	-1	1	-1	1180
8	1	1	1	1	1	1	1	1	2460

Tabela 7.11: Tabela padronizada de sinais para projetos 2^3

Uma vez preenchida esta coluna, pode-se então realizar a operação de multiplicação da matriz transposta dos sinais pelo vetor de resposta ***R***. Desta maneira, obtém-se:

$$X^t R = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ -1 & 1 & -1 & 1 & -1 & -1 & 1 & -1 \\ -1 & -1 & 1 & 1 & -1 & -1 & 1 & 1 \\ -1 & -1 & -1 & -1 & 1 & 1 & 1 & 1 \\ 1 & -1 & -1 & 1 & 1 & -1 & -1 & 1 \\ 1 & -1 & 1 & -1 & -1 & 1 & -1 & 1 \\ 1 & 1 & -1 & -1 & -1 & -1 & 1 & 1 \\ -1 & 1 & 1 & -1 & 1 & -1 & -1 & 1 \end{bmatrix} * \begin{bmatrix} 500 \\ 1100 \\ 900 \\ 1900 \\ 630 \\ 1420 \\ 1180 \\ 2460 \end{bmatrix} = \begin{bmatrix} 10090 \\ 3670 \\ 2790 \\ 1290 \\ 890 \\ 470 \\ 390 \\ 90 \end{bmatrix}$$

Cada valor obtido deverá ser dividido por 2^k. Assim, para o cálculo final dos efeitos e da média, divide-se cada um dos elementos do vetor por $2^k = 8$. O vetor resposta terá os seguintes valores:

$$\begin{bmatrix} Média \\ A \\ B \\ C \\ AB \\ AC \\ BC \\ ABC \end{bmatrix} = \begin{bmatrix} 1261{,}25 \\ 458{,}75 \\ 348{,}75 \\ 161{,}25 \\ 111{,}25 \\ 58{,}75 \\ 48{,}75 \\ 11{,}25 \end{bmatrix}$$

Calculados os valores dos efeitos provocados pelos fatores individuais e suas interações, determina-se a contribuição de cada um deles sobre as variações no throughput do servidor de arquivos.

Relembrando, a fórmula 7.5 permite calcular o se chama de Variação Total de *R*, por meio da chamada Soma dos Quadrados Totais (SQT). Tal valor compreende uma soma das contribuições de cada um dos fatores e de suas interações.

Para o caso de projetos 2^3, Soma dos Quadrados Totais é definida como segue:

$$SQT = SQA + SQB + SQC + SQAB + SQAC + SQBC + SQABC \qquad (7.7)$$

De acordo com as fórmulas 7.6 e 7.7, para um projeto 2^3, a variação pode então ser dividida em sete parcelas:

$$SQT = 2^3 q_A^2 + 2^3 q_B^2 + 2^3 q_C^2 + 2^3 q_{AB}^2 + 2^3 q_{Ac}^2 + 2^3 q_{BC}^2 + 2^3 q_{ABC}^2 \qquad (7.8)$$

Empregando-se a fórmula acima se obtém os percentuais das contribuições de cada fator sobre a variação do throughput.

$$SQT = 8 \text{ x } \{(458{,}75)^2 + (348{,}75)^2 + (161{,}25)^2 + (111{,}25)^2 + (58{,}75)^2 + (48{,}75)^2 + (11{,}25)^2\}$$

$$SQT = 3.011.287{,}5$$

A partir do cálculo do total da variação, as contribuições individuais podem ser obtidas.

Fração da variação explicada pelo fator $A = \frac{SQA}{SQT}$ = 8 x $458{,}75^2$ /3.011.287,50 = 0,5591

Percentual da variação explicada pelo fator *A* = 0,5591 x 100% = 55,91%.

Realizando-se cálculos semelhantes para os demais fatores chega-se aos seguintes resultados:

O fator A (memória principal) contribui com 55,91% da variação do throughput do servidor de arquivos da rede local. O fator B (memória cache) contribui com 32,31 % da variação e o fator C (disco) é responsável por 6,91% da variação. As interações entre dois fatores contribuem com 3,29% da variação (interação AB), com 0,92% da variação (interação AC) e com 0,63% da variação (interação BC). Finalmente, conclui-se que apenas 0,03% da variação é atribuída a interação dos três fatores (interação ABC).

7.5. Projetos Fatoriais com Replicações – $2^k.r$

Sempre que se lida com projetos experimentais, é preciso estar atento ao fato de que a repetição de determinado ensaio pode levar a diferentes respostas. Estas diferenças podem ser em maior ou menor grau, dependendo das condições do teste (ambiente, instrumentos, etc.) e dos tipos de variáveis envolvidas. Desconsiderar tal realidade pode trazer sérias conseqüências, uma vez que as análises podem levar a conclusões errôneas sobre as relações de influência dos fatores sobre variáveis de resposta.

Quando se realizam projetos fatoriais sem a repetição dos experimentos, não é possível avaliar se as possíveis diferenças nos resultados permitiriam que se chegasse a mesma conclusão. Quando se repetem os ensaios, sob as mesmas combinações fatores/níveis, é possível calcular o chamado erro experimental e avaliar o nível qualitativo das análises e das conclusões. Se cada um dos 2^k experimentos do projeto é repetido *r* vezes, necessita-se de 2^k x r experimentos. Este tipo de projeto é chamado projeto fatorial 2^k.r. Nas subseção que segue será explorado o caso especial de projetos envolvendo dois fatores com dois níveis cada um deles. Na seqüência estende-se o estudo para o caso de k fatores com dois níveis.

7.5.1. Projetos Fatoriais 2^2.*r*

Nos projetos fatoriais do tipo 2^2.*r*, são realizados *r* repetições de cada uma das quatro possíveis combinações entre os dois fatores e seus níveis. Desta forma, para o caso de se repetirem três vezes cada um dos ensaios, um total de 4 x 3 = 12 experimentos são realizados.

Uma vez que mais de um ensaio é realizado para cada combinação de fatores e níveis, torna-se possível o cálculo do erro experimental que cometido, uma vez que para fins do cálculo dos efeitos, toma-se o valor médio das respostas. O modelo de desempenho dado pela fórmula 7.3 vista na seção 7.4.1 (Projeto Fatorial 2^2) pode ser estendido (fórmula 7.9, abaixo), incluindo-se naquela equação uma parcela referente ao erro.

$$R = q_0 + q_{A}.x_A + q_{B}.x_B + q_{AB}.x_A.x_B + \boldsymbol{e} \qquad (7.9)$$

Neste modelo ***e*** é o erro experimental, X_A e X_B são os dois fatores e os q´s são os efeitos. Para a determinação dos valores dos efeitos, pode-se adotar o método da tabela dos sinais, como visto anteriormente.

7.5.2. Tabela dos Sinais para o Computo dos Efeitos em Projetos 2^2.*r*

O emprego da tabela de sinais para calcular os efeitos dos fatores e suas interações sobre as variáveis de resposta exige, inicialmente, apenas o acréscimo de colunas referentes aos valores das respostas, para cada uma das repetições. Uma vez completadas as colunas de respostas, acrescenta-se uma nova coluna, a qual conterá os valores das médias das várias replicações. Estes valores médios são então empregados no cálculo dos efeitos, de forma semelhante ao que já foi realizado anteriormente. Para ilustrar esta mecânica de cálculo, repete-se aqui o *Exemplo 5*, agora com cada um dos experimentos repetidos três vezes.

Exemplo 8.

Considere o mesmo problema de estudar o impacto do tamanho da memória principal (Fator A) e o tamanho da memória *cache* (Fator B) sobre a performance de um servidor de arquivos de uma rede local de computadores (LAN). Cada um destes dois fatores possui dois níveis, conforme indicado na tabela 7.12. A performance do servidor de arquivos foi medida considerando como variável de resposta o seu *throughput* (neste caso, número de operações de transferência de arquivos por segundo).

Memória *Cache* (kbytes)	Memória Principal 128 Mbytes	Memória Principal 256 Mbytes
256	500	1100
512	900	1900

Tabela 7.12: Performance do servidor de Arquivos (Operações/seg.)

Observe os dados da tabela 7.13. Verifica-se que para cada combinação de níveis dos dois fatores, houve três repetições do ensaio. Estes experimentos resultam em 12 observações.

Média	A	B	AB	Replicação 1	Replicação 2	Replicação 3	$R_{\text{médio}}$
1	-1	-1	+1	500	550	540	530
1	+1	-1	-1	1100	1140	1090	1110
1	-1	+1	-1	900	850	920	890
1	+1	+1	+1	1900	1950	1910	1920

Tabela 7.13: Os 12 resultados de um projeto experimental $2^2.3$

Na tabela 7.13, os ensaios repetidos encontram-se sob os títulos Replicação 1, 2 e 3. Na última coluna da tabela, sob o título $R_{\text{médio}}$, encontra-se o vetor de respostas. Estas são agora representadas pelas médias calculadas a partir dos três resultados das replicações. Como se pode claramente verificar, existem diferenças entre os três resultados de cada combinação de fatores. Desta forma, a representação do conjunto dos resultados por seu valor médio implica a admissão de um erro, o chamado erro experimental. Aqui não se está discutindo a origem das diferenças encontradas em cada resultado das replicações. Os motivos podem ser diversos, tais como um erro de medição, uma falha do instrumento de medida, erros de transcrição ou mesmo, diferenças devido à própria natureza randômica da(s) variável(is) envolvida(s) no(s) processo(s).

É possível, a partir destas repetições de experimentos, calcular o erro experimental. A tabela 7.14 abaixo é uma extensão da tabela anterior. A ela são anexadas mais três colunas (e_{i1}, e_{i2}, e_{i1}) que retratam a diferença entre os valores de cada replicação e a própria média.

Média	A	B	AB	Rep. 1	Rep. 2	Rep. 3	$R_{\text{médio}}$	e_{i1}	e_{i2}	e_{i3}
1	-1	-1	+1	500	550	540	530	-30	20	10
1	+1	-1	-1	1100	1140	1090	1110	-10	30	-20
1	-1	+1	-1	900	850	920	890	10	-40	30
1	+1	+1	+1	1900	1950	1910	1920	-20	30	-10

Tabela 7.14: Computação dos erros em um projeto experimental $2^2.3$

Uma vez calculados os erros absolutos, é possível determinar a contribuição ou o efeito do erro experimental que consta da fórmula 7.9. Para o cálculo deste efeito calcula-se a parcela de contribuição SQE referente ao erro:

$$SQE = \sum_{i=1}^{2^k} \sum_{j=1}^{r} e_{ij}^2$$

Para o caso do exemplo, conforme os dados da tabela 7.14, o valor de SQE será dado pela soma dos quadrados das 12 diferenças. Estes valores são tomados ao quadrado para evitar que se anulem. Desta forma obtém-se:

$$SQE = (-30)^2 + 20^2 + 10^2 + ... + 30^2 + (-10)^2 = 5000$$

Para completar o exemplo, computa-se os efeitos dos demais fatores empregando-se a matriz dos coeficientes (***X***), multiplicando-se sua transposta $(\boldsymbol{X})^{\mathbf{t}}$ pelo vetor dos resultados ***R***, agora representado pelos valores médios.

$$X^{t}R = \begin{bmatrix} 1 & 1 & 1 & 1 \\ -1 & 1 & -1 & 1 \\ -1 & -1 & 1 & 1 \\ 1 & -1 & -1 & 1 \end{bmatrix} * \begin{bmatrix} 530 \\ 1110 \\ 890 \\ 1920 \end{bmatrix} = \begin{bmatrix} 4450 \\ 1610 \\ 1170 \\ 450 \end{bmatrix}$$

Para o cálculo final dos efeitos e da média, divide-se cada um dos elementos do vetor por $2^k = 4$. Assim, obtém-se o vetor resposta com os seguintes valores:

$$\begin{bmatrix} Média \\ X_A \\ X_B \\ X_{AB} \end{bmatrix} = \begin{bmatrix} 1112{,}5 \\ 402{,}5 \\ 292{,}5 \\ 112{,}5 \end{bmatrix}$$

7.5.3. Alocação das Variações

Assim como no projeto 2^2, o percentual de variação atribuído a cada uma dos fatores, ajuda a decidir o quanto um determinado fator é importante, na medida do impacto que causa sobre a resposta. Uma vez que se calculou a parcela relativa ao erro experimental, é possível determinar sua contribuição sobre as variações sofridas pelo throughput.

A fórmula da Soma dos Quadrados Totais fica:

$$SQT = SQA + SQB + SQAB + SQE$$

Para o projeto $2^2.r$, a variação pode ser dividida em quatro partes:

$$SQT = 2^2 rq_A^2 + 2^2 rq_B^2 + 2^2 rq_{AB}^2 + SQE$$

Complementando o exemplo 8, a alocação das variações devido aos fatores principais, sua interação e ao erro experimental pode ser calculada como segue:

$SQA = 12 \text{ x } (402{,}5)^2 = 1.944.075$

$SQB = 12 \times (292,5)^2 = 1.026.675$
$SQAB = 12 \times 112,5^2 = 151.875$
$SQE = 5000$
$SQT = 1.944.075 + 1.026.675 + 151.875 + 5.000 = 3.127.625$

Assim,

Variação devida ao fator A = 1.944.075/3.127.625 = 62,16 %.
Variação devida ao fator B = 1.026.675/3.127.625 = 32,83 %.
Variação devida à interação entre A e B = 151.875/3.127.625 = 4,85 %.

Os restantes 0,16 % não são explicáveis e podem ser atribuídos aos erros.

No caso do exemplo acima, verifica-se que muito pouco da variação sofrida pela variável de resposta pode ser atribuída a erros experimentais. No entanto, nem sempre este valor é tão pequeno como o que foi observado. Em algumas situações, quando um número razoável de candidatos a fatores encontra-se disponível, é possível que o analista venha a escolher fatores que tenham pouca ou nenhuma influência sobre a variável de resposta. Neste caso, haverá uma concentração de responsabilidade sobre os fatores mais influentes. Se todos os fatores escolhidos tiverem pouca influência sobre a variável de resposta, o analista observará uma grande concentração de responsabilidade atribuída ao erro amostral.

Na próxima seção serão revisados os conceitos vistos neste capítulo por meio de dois estudos de caso. O primeiro refere-se a um estudo envolvendo um sistema de manufatura e o segundo um sistema computacional. Ambos os sistemas foram modelados no Arena. Desta maneira, todos os ensaios e replicações referentes à estratégia experimental adotada serão realizados por meio de simulações dos referidos modelos demonstrando-se, mais uma vez, a ampla abrangência e aplicabilidade desta técnica.

7.6. Projetos Experimentais Empregando Modelos de Simulação

Nesta seção será desenvolvido, passo-a-passo, dois estudos relacionados a desempenho de sistemas. O primeiro deles trata de um pequeno ambiente de manufatura. No entanto, mesmo aqueles não familiarizados com sistemas de produção, não terão dificuldades de entender o estudo. Da mesma maneira, não será necessário ser um especialista em tecnologias de informação para compreender o segundo estudo de caso, o qual envolve um sistema computacional cuja arquitetura e topologia é bastante familiar aos mais iniciados nesta área do conhecimento. Nesta oportunidade se estará enfatizando, mais uma vez, a importância e a facilidade de se empregar modelos de simulação aliados a projetos experimentais, em estudos voltados a análise de desempenho de sistemas.

7.6.1. Primeiro Estudo de Caso – Sistema de Manufatura

Descrição do Sistema

Uma célula de manufatura é constituída de duas estações em série (ver Figura 7.10). Na primeira, produtos semi-acabados são processados em uma máquina (centro de usinagem). Logo após, são enviados a uma estação de inspeção. Os tempos de deslocamento entre as estações são desprezados. Os produtos chegam automaticamente ao sistema, com intervalos de 1 minuto. O processo na estação de usinagem segue uma distribuição Uniforme (0,65; 0,70) minutos e os tempos de inspeção também seguem uma distribuição Uniforme (0,75; 0,80) minutos. Após a inspeção, as peças seguem para expedição. Verificou-se que no processo de inspeção 10 % das peças falham e retornam para serem reprocessadas no centro de usinagem.

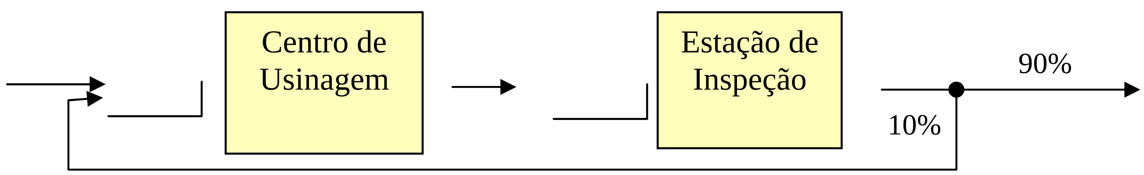

Figura 7.10: Descrição esquemática do sistema.

O centro de usinagem está necessitando de manutenção e como não tem sido possível esta parada, tem apresentado falhas. O tempo decorrido entre estas falhas foi detectado como sendo uma variável aleatória que pode ser descrita por uma distribuição Exponencial com média de 6 horas. Já o tempo de reparo da máquina é uma variável aleatória que segue uma distribuição Uniforme (8, 12) minutos.

Este sistema foi modelado e simulado por 160 horas (aproximadamente 20 dias de produção). Foi considerado um período de *warm-up* de 40 horas.

Estratégia de Experimentação

A Engenharia Industrial tem interesse especial por uma variável de resposta que esta sendo monitorada no modelo: o tempo total para completar o processo de usinagem/inspeção (*leadtime*). Depois de várias análises, os engenheiros decidiram que cinco variáveis de controle (fatores) devem ser consideradas nas análises. Um projeto experimental deve ser realizado para verificar as relações entre estes fatores e a variável de resposta. Os fatores, listados na tabela abaixo, terão dois níveis: o atual (-) e um proposto (+).

Fator	Descrição	Nível atual (-1)	Nível proposto (+1)
A	Tempo entre falhas da maquina	EXPO (360) min.	EXPO (720) min.
B	Tempo reparo da máquina	UNIF (8, 12) min.	UNIF (4, 6) min.
C	Probabilidade de falha na inspeção	10%	5%
D	Tempo de processamento	UNIF (0,65; 0,70)	UNIF (0,45; 0,50)
E	Tempo de inspeção	UNIF (0,75; 0,80)	UNIF (0,675; 0,720)

Tabela 7.15: Fatores que podem ser analisados.

Na tabela 7.15 verifica-se que os tempos das atividades associadas aos dois primeiros fatores expressam uma expectativa da gerência em observar a influência de um novo processo de manutenção sobre o Leadtime. Uma vez implantado este novo processo, espera-se que dobre o período médio de tempo entre paradas da máquina (fator 1). Da mesma forma, a expectativa é que os tempos de reparo da mesma (fator 2) caiam pela metade, em média. Com a melhoria da manutenção, espera-se, também, que as falhas no processo (fator 3) reduzam a taxa dos produtos que necessitam retrabalho. Assim, pretende-se uma redução de 10% para 5% neste número. Com relação aos dois últimos fatores, os valores associados aos níveis propostos espelham, também, esperanças de melhorias no processo de produção/inspeção em relação aos tempos dos níveis atuais.

Uma analise do número de fatores revela que para um procedimento do tipo Projeto Fatorial Completo, tem que ser realizados $2^5 = 32$ experimentos. Para se resguardarem, os analistas exigem pelo menos três replicações de cada um deles (quase todos os fatores são variáveis aleatórias). Desta forma, chega-se a um total de 96 simulações. Este número não é considerado grande, pois o modelo de simulação é bastante simples. As três replicações de cada experimento demoram cerca de 15 segundos em uma máquina com 132 MB de memória RAM, com um processador de 500 MHZ e rodando sob o sistema operacional Windows. Desta forma, fica decidida a realização de um projeto fatorial completo. No entanto, os analistas decidem que apenas os efeitos principais e os efeitos das interações entre pares de fatores serão averiguados.

EXP	Média	A	B	C	D	E	AB	AC	AD	. . .	DE
1	1	1	1	1	1	1	1	1	1	. . .	1
2	1	1	1	1	1	-1	1	1	1	. . .	-1
3	1	1	1	1	-1	1	1	1	-1	. . .	-1
4	1	1	1	1	-1	-1	1	1	-1	. . .	1
5	1	1	1	-1	1	1	1	-1	1	. . .	1
6	1	1	1	-1	1	-1	1	-1	1	. . .	-1
7	1	1	1	-1	-1	1	1	-1	-1	. . .	-1
8	1	1	1	-1	-1	-1	1	-1	-1	. . .	1
9	1	1	-1	1	1	1	-1	1	1	. . .	1
10	1	1	-1	1	1	-1	-1	1	1	. . .	-1
11	1	1	-1	1	-1	1	-1	1	-1	. . .	-1
12	1	1	-1	1	-1	-1	-1	1	-1	. . .	1
13	1	1	-1	-1	1	1	-1	-1	1	. . .	1
14	1	1	-1	-1	1	-1	-1	-1	1	. . .	-1
15	1	1	-1	-1	-1	1	-1	-1	-1	. . .	-1
16	1	1	-1	-1	-1	-1	-1	-1	-1	. . .	1
17	1	-1	1	1	1	1	-1	-1	-1	. . .	1
18	1	-1	1	1	1	-1	-1	-1	-1	. . .	-1
19	1	-1	1	1	-1	1	-1	-1	1	. . .	-1
20	1	-1	1	1	-1	-1	-1	-1	1	. . .	1
21	1	-1	1	-1	1	1	-1	1	-1	. . .	1
22	1	-1	1	-1	1	-1	-1	1	-1	. . .	-1
23	1	-1	1	-1	-1	1	-1	1	1	. . .	-1
24	1	-1	1	-1	-1	-1	-1	1	1	. . .	1
25	1	-1	-1	1	1	1	1	-1	-1	. . .	1
26	1	-1	-1	1	1	-1	1	-1	-1	. . .	-1
27	1	-1	-1	1	-1	1	1	-1	1	. . .	-1
28	1	-1	-1	1	-1	-1	1	-1	1	. . .	1
29	1	-1	-1	-1	1	1	1	1	-1	. . .	1
30	1	-1	-1	-1	1	-1	1	1	-1	. . .	-1
31	1	-1	-1	-1	-1	1	1	1	1	. . .	-1
32	1	-1	-1	-1	-1	-1	1	1	1	. . .	1

Tabela 7.16: Planilha do Excel mostrando os 32 experimentos realizados

Desta maneira, para se efetivarem a análise, foi construído um modelo do sistema visto na Figura 7.10, no ambiente Arena (ver arquivo “Estudo Caso Manufatura.doe”). Para a realização dos cálculos dos efeitos dos fatores, foi construída uma planilha no Excel (ver arquivo “Estudo Caso Manufatura.xls”).

Considerando-se o cálculo dos efeitos principais e dos efeitos das combinações dois a dois dos fatores principais, a tabela 7.16 mostra parcialmente a planilha eletrônica construída no Excel com os coeficientes da tabela de sinais para projetos fatoriais tipo 2^5.r.

Observando-se a tabela 7.16, verifica-se que serão 15 os efeitos computados (A, B, C, ...DE). As primeiras colunas sob os títulos A, B, C, D e E, apresentam os sinais referentes aos efeitos primários e as demais os sinais decorrentes das interações entre dois fatores. Neste estudo, se está ignorando os efeitos dos fatores tomados três a três (ABC, ABD, ABE, ACD, ACE, ADE, BCD, BCE, BDE e CDE), quatro a quatro (ABCD, ABCE, ABDE, ACDE e BCDE) e cinco a cinco (ABCDE).

A Tabela 7.17 apresenta os resultados dos experimentos. Para cada ensaio são realizadas três replicações. O valor final da variável de resposta de cada replicação encontra-se sob os títulos Y1, Y2 eY3. A média aritmética destes três valores pode ser vista na coluna Ymédio. As três colunas seguintes apresentam as diferenças entre o valor médio e a resposta de cada replicação.

EXP	Y1	Y2	Y3	Ymédio	e1	e2	e3
1	1,3487	1,3699	1,3679	1,3622	-0,0135	0,0077	0,0057
2	1,5070	1,4591	1,5584	1,5082	-0,0012	-0,0491	0,0502
3	1,5426	1,6200	1,5399	1,5675	-0,0249	0,0525	-0,0276
4	1,7106	1,6968	1,7277	1,7117	-0,0011	-0,0149	0,0160
5	1,5433	1,5123	1,5040	1,5199	0,0234	-0,0076	-0,0159
6	1,6783	1,7425	1,7539	1,7249	-0,0466	0,0176	0,0290
7	1,7232	1,7022	1,7409	1,7221	0,0011	-0,0199	0,0188
8	2,0358	1,9585	2,0114	2,0019	0,0339	-0,0434	0,0095
9	1,7299	1,4422	1,5642	1,5788	0,1511	-0,1366	-0,0146
. . .	. . .	. . .	. . .	. . .	. . .	. . .	. . .
30	2,4396	2,7676	2,2326	2,4799	-0,0403	0,2877	-0,2473
31	2,0152	2,1106	1,9601	2,0286	-0,0134	0,0820	-0,0685
32	2,7628	2,6133	2,5548	2,6436	0,1192	-0,0303	-0,0888

Tabela 7.17: Resultados das três replicações de cada combinação dos níveis e fatores

Uma vez montada a matriz dos sinais (Tabela 7.16), calculados os valores da variável de resposta (Ymédio) e dos erros de cada um dos experimentos (Tabela 7.17), é possível se obter a equação da Soma Total dos Quadrados, considerando-se os efeitos dos fatores principais, das combinações e do erro. A equação resultante possui a seguinte formação:

SQT = SQA + SQB + SQC + SQD + SQE + SQAB + SQAC + SQAD + SQ AE + SQBC + . . . + SQDE + SQErros

Os valores das contribuições de cada um dos fatores e de suas combinações sobre as variações da variável de resposta são calculados automaticamente pela planilha. O resultado final pode ser observado na figura 7.11 abaixo.

SQA =	8,87%
SQB =	33,52%
SQC =	16,37%
SQD =	11,24%
SQE =	20,61%
SQAB =	1,48%
SQAC =	0,01%
SQAD =	0,00%
SQAE =	0,36%
SQBC =	0,03%
SQBD =	0,00%
SQBE =	0,52%
SQCD =	0,01%
SQCE =	0,94%
SQDE =	0,00%
SQErros =	6,05%
SQT =	100,00%

Figura 7.11: Resultados dos efeitos dos fatores.

Os resultados apresentados na Figura 7.11 mostram que cerca de 91% da variação dos resultados da variável de resposta são de responsabilidade dos fatores, individualmente, e muito pouco é devido às combinações entre esses fatores. O principal responsável pela variação é o fator B (tempo de reparo da máquina), com cerca de 34% do total. A contribuição deste fator, somada com a contribuição do fator A (tempo entre falhas da máquina), de aproximadamente 9%, e com a contribuição da interação entre estes dois fatores (SQAB) de quase 2%, perfaz um total de 44% das contribuições de todos os fatores sobre a variação do tempo total para completar o processo de usinagem/inspeção (*leadtime*). Portanto, tem razão a engenharia industrial, em verificar a influência do novo processo de manutenção sobre o Leadtime.

Quanto aos demais fatores, se observa que o fator E (tempo de inspeção) é responsável por aproximadamente 21% da variação do Leadtime. Explica-se tal fato observando-se a seqüência de atividades. O processo de inspeção é a atividade com maior duração e, neste caso, o ponto de estrangulamento (gargalo) na célula de manufatura. Como os gargalos são grandes responsáveis pelos atrasos no Leadtime, melhorias realizadas sobre estes processos costumam ter grande impacto nas variáveis de resposta. Outro fator cuja responsabilidade merece comentários é o fator C (probabilidade de falha na inspeção). A influência deste sobre o tempo total de processo na célula é razoável (16,37%). Como, ao falhar na inspeção, uma peça é enviada de volta ao processamento no centro de usinagem, esta falha provoca um maior afluxo de entidades (peças) nos dois processos, com um conseqüente aumento das filas. Uma redução na probabilidade de falha de 10% para 5% implica que, na segunda alternativa, metade das falhas não mais ocorrerá, isso acarreta uma sensível melhora do leadtime. Finalmente, cabe comentar a contribuição do tempo de processo sobre o leadtime. Mesmo que a redução de um nível para o outro deste fator seja grande (cerca de 30%) e que este tempo tenha influência direta sobre o tempo total de produção, sua contribuição sobre a variação não é proporcional. Isto se explica pelo fato do centro de usinagem não ser um ponto de estrangulamento no fluxo produtivo. Não existe um processo de contenção neste ponto, pois a relação entre freqüência de chegadas de peças no sistema (uma peça/minuto) e a freqüência de término de serviços na máquina (uma peça/0,7 minuto no pior caso) é amplamente favorável a maquina. Desta maneira, por não se tratar de um gargalo, melhorias neste processo têm poucas conseqüências sobre o tempo total ou leadtime.

Neste estudo de caso o leitor pode ter uma boa perspectiva da importância de se aliar o emprego de modelos de simulação a técnicas estatísticas. Resultados de modelos de simulação desprovidos de análises com base estatística têm pouca validade científica e podem facilmente induzir ao erro.

7.6.2. Segundo Estudo de Caso – Sistema Cliente/Servidor

Descrição do Sistema

Em uma rede local (LAN) estão conectados computadores (estações de trabalho) de operadores de um serviço de atendimento a clientes e um computador (servidor) no qual são processadas operações sobre um banco de dados, o qual contém uma base de dados armazenada em um único disco. A LAN tem capacidade nominal de transmissão de 10Mbps (Figura 7.12).

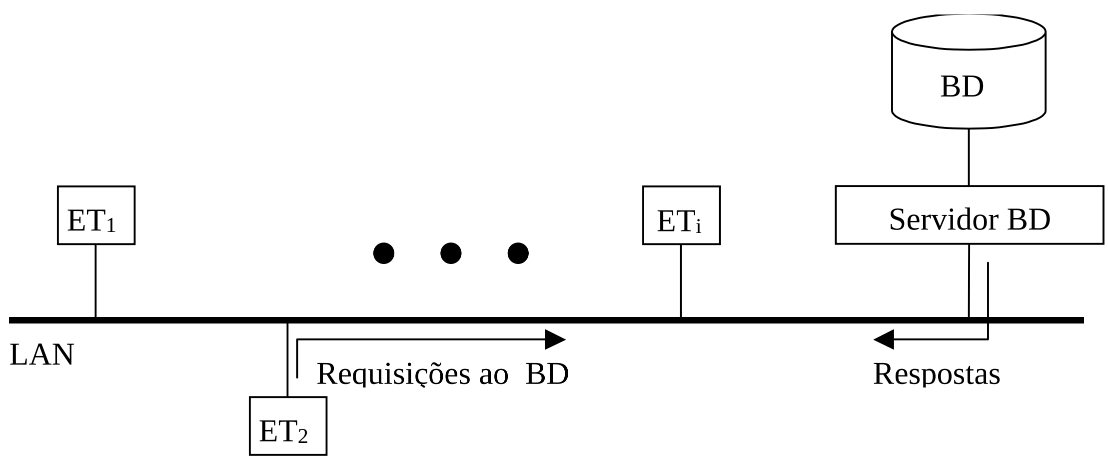

Figura 7.12: Sistema Cliente-Servidor

Neste sistema, as estações de trabalho (ET´s) são operadas por funcionários de uma empresa que presta informações por telefone. Como esta empresa é muito solicitada, considera-se que os operadores estão sempre ocupados durante o tempo de simulação. Ao receber uma pergunta via telefone, o operador submete uma requisição ao servidor da base de dados (Servidor BD) via rede local. Esta requisição tem sua sintaxe e semântica verificada no processador local antes de ser submetido à rede de transmissão. Se OK, a requisição é então enviada ao servidor.

O servidor recebe requisições provenientes de todas as estações e as submete ao sistema gerenciador da base de dados (SGBD). O SGBD é um conjunto de programas executado no servidor. Após a consulta, o servidor envia a resposta para a estação que originou o comando. A resposta trafega pela rede e chega ao operador. O operador examina a resposta obtida e responde a pergunta do cliente da empresa via telefone. Após um tempo que chamaremos de Tempo de Raciocínio, o qual é determinado por uma distribuição Exponencial com média de 10 seg., o operador (devido a um novo questionamento do cliente na linha ou devido a sua própria necessidade de informações) envia um novo comando ao servidor. Este processo se repete em todas as estações.

O servidor BD possui uma CPU capaz de executar 25 MIPS (milhões de instruções por segundo). Por sua vez, cada uma das ET possui uma CPU de 1 MIPS. Uma requisição à base de dados possui um tamanho, em bytes, que segue uma distribuição Triangular com parâmetros (60, 110, 140). Já as respostas geradas pelo SGBD seguem uma Triangular com parâmetros (800, 960, 1200) bytes. O pré-processamento local dos comandos nas ET´s implica na execução de um número de instruções de máquina fornecidos por uma distribuição Uniforme (10.000, 20.000).

O disco do servidor BD tem como principal característica sua capacidade de transferir dados para a memória principal. Sua taxa nominal de transferência é de 5 MB/segundo. Cada acesso ao disco lê ou escreve 4 Kbytes de dados.

Uma vez no servidor, cada requisição gera um número de instruções que devem ser processadas. Este número é dado por uma Uniforme (50.000, 150.000). Para simplificar a modelagem, determina-se que o número de operações de I/O depende do número de ciclos necessários. A cada ciclo a CPU processa de 5.000 a 15.000 instruções, de acordo com uma distribuição Uniforme. Este valor corresponderá à fatia de tempo (*Time Slice*) que a CPU do servidor dedicará a cada processo. Assim, a cada ciclo serão realizados um processo na CPU e uma operação de I/O no disco.

As mensagens contendo as requisições ao servidor de Banco de Dados precisam obter acesso e trafegar pelo meio de transmissão. O acesso ao meio depende de um protocolo de acesso. É este que verifica se o meio (rede) encontra-se livre para que a mensagem poça ser enviada. No caso deste sistema, supõe-se que a rede seja do tipo Ethernet e que o protocolo seja o CSMA/CD. Novamente, para simplificar a modelagem, não será preciso modelar o protocolo de acesso ao meio (esta modelagem é relativamente complexa). Considera-se apenas um pequeno atraso (DELAY) médio, dado por uma Exponencial (0.05) segundos.

A partir dos dados acima definidos, este sistema foi modelado no Arena e permite responder questões relativas a duas variáveis de resposta:

1. O tempo decorrido entre a emissão de uma requisição e o recebimento da respectiva resposta;
2. O throughput da rede (LAN).

Estratégia de Experimentação

Uma vez que a demanda pelos serviços prestados pelo sistema vem apresentando taxas de crescimento positivas, os responsáveis por sua gerência desejam realizar um estudo visando determinar que componentes deste sistema tem ligação com sua performance. Além disso, é importante estabelecer, também, o grau de responsabilidade de cada um destes componentes. A gerência pretende, a partir deste conhecimento, realizar um planejamento de capacidade visando manter a qualidade dos serviços. A principal variável a ser considerada será o tempo de resposta provido pelo SGBD.

Argüidos pela gerência, os técnicos do setor responderam que os principais elementos relacionados com a performance deste sistema são:

1. Velocidade da $CPU_{operador}$
2. Velocidade $CPU_{Servidor}$
3. Disco do Servidor
4. Capacidade da Rede
5. Tamanho das Requisições
6. Tamanho das Respostas
7. Intensidade das Requisições

Chamado para opinar, um analista com conhecimentos em simulação de sistemas e em projeto de experimentos sugere que seja realizado um projeto fatorial tipo 2^k, visando melhor conhecer as relações entre os elementos citados e a variável de resposta. Os valores sugeridos para os dois níveis dos *k* fatores são apresentados na tabela 7.18.

	Fator	**Nível mais Baixo (-1)**	**Nível mais Alto (+1)**
1	Velocidade da $CPU_{Operador}$	1 MIPS	5 MIPS
2	Velocidade $CPU_{Servidor}$	25 MIPS	100 MIPS
3	Capacidade da Rede	10 Mbps	100 Mbps
4	Intensidade das Requisições	20 telefonistas	50 telefonistas
5	Tamanho das Requisições	Tria (60, 110, 140) bytes	Tria (600, 1100, 1400) bytes
6	Tamanho das Respostas	Tria (200, 300, 400) bytes	Tria (1300, 1400, 1500) bytes
7	Disco do Servidor	Tx. transf. - 5 MB/seg.	Tx. transf. - 20 MB/seg.

Tabela 7.18: Valores associados as níveis de cada fator.

Na tabela 7.18 pode-se observar os valores associados a cada um dos dois níveis de cada fator. Os quatro primeiros fatores estão associados a parâmetros do próprio sistema, enquanto que os três últimos estão relacionados à carga de trabalho suportada pelo sistema.

A velocidade das CPU´s, tanto das estações de trabalho ($CPU_{Operador}$) quanto do servidor ($CPU_{Servidor}$) é determinada pela sua capacidade de processamento, dada em MIPS. As diferenças associadas aos dois níveis do terceiro fator, disco do servidor, são denotadas pela sua capacidade de transferir dados para a memória principal (MB/segundo). De maneira semelhante, a diferença entre os dois níveis do quarto fator, a rede, também está associada a sua capacidade de transferir dados entre os postos de trabalho e o servidor de banco de dados.

Os três últimos fatores estão associados à carga submetida ao sistema. Esta carga se divide em dois elementos: intensidade e demanda sobre recursos. O primeiro elemento (fator cinco) é determinado pela freqüência com que as requisições são enviadas ao servidor e depende do número de operadores no sistema. A demanda da carga depende dos tamanhos (em bytes) das requisições (fator seis) e das respostas (fator sete) trocadas entre os clientes do sistema (estações de trabalho dos operadores) e o servidor de banco de dados. Neste caso, esta demanda ocorre sobre o recurso rede.

Uma vez que vários parâmetros do sistema e alguns dos fatores envolvem o uso de variáveis aleatórias, é conveniente que se adote a estratégia de repetição dos experimentos, visando minimizar o erro experimental. Assim, de acordo com as definições vistas na seção 7.5, a tendência natural é realizar um projeto fatorial com replicações do tipo $2^k.r$. Considerando que um total de sete fatores foram listados e que se deseja repetir pelo menos três vezes cada um dos ensaios, serão necessários: 2^7 x 3 = 128 x 3 = 384 experimentos.

Devido ao elevado número de experimentos, outras estratégias podem ser adotadas. A primeira, e mais simples, é realizar uma avaliação mais criteriosa dos fatores e reduzir o seu número. Desta forma foi solicitado aos técnicos da área de informática que apresentassem uma lista contendo os quatro fatores mais relevantes, dentre aqueles inicialmente listados. Com apenas quatro fatores, o número de ensaios cai de 384 para apenas 48 experimentos, considerando as replicações. A adoção desta estratégia permite obter mais rapidamente os resultados e, com base nestes, vir a descartar fatores que venham a demonstrar pouca importância, substituindo-os por outros. Com o objetivo de mostrar que se pode chegar a resultados diferenciados, dependendo da escolha dos fatores, imagine que dois conjuntos de fatores tenham sido apontados pelos técnicos. Para o primeiro ensaio, os fatores 1 ($CPU_{Operador}$), 2 ($CPU_{Servidor}$), 3 (Rede) e 5 (TamRequisição) foram incluídos. Já para o segundo ensaio, a escolha recaiu sobre os fatores 2 ($CPU_{Servidor}$), 4 (Intensidade ou freqüência das requisições), 6 (tamanho das Respostas) e 7 (Capacidade do Disco).

Ensaios

Para a realização dos ensaios será utilizado um modelo do sistema desenvolvido no ambiente Arena (CSPExpAuto.doe). Especialmente para este estudo, foram integrados ao modelo do Arena rotinas desenvolvidas em VBA (*Visual Basic for Applications*). O ambiente de modelagem do Arena permite esta integração com VBA desde a versão 3.0. Por meio destas rotinas, o usuário pode realizar, de forma automatizada, os experimentos desejados, designando valores convenientes aos parâmetros (fatores) os quais pretende que sejam analisados.

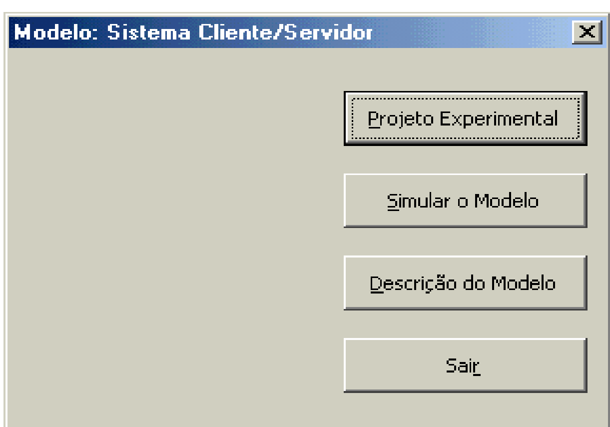

Figura 7.13: Tela inicial do modelo com interface para Projeto Experimental

Assim, através de uma interface amigável, o analista pode solicitar os experimentos, e verificar os resultados em uma planilha eletrônica gerada pelas rotinas em VBA, que realiza os cálculos necessários à determinação das influências e responsabilidades de cada fator, sobre a variação dos resultados observados na variável de resposta (tempo de resposta).

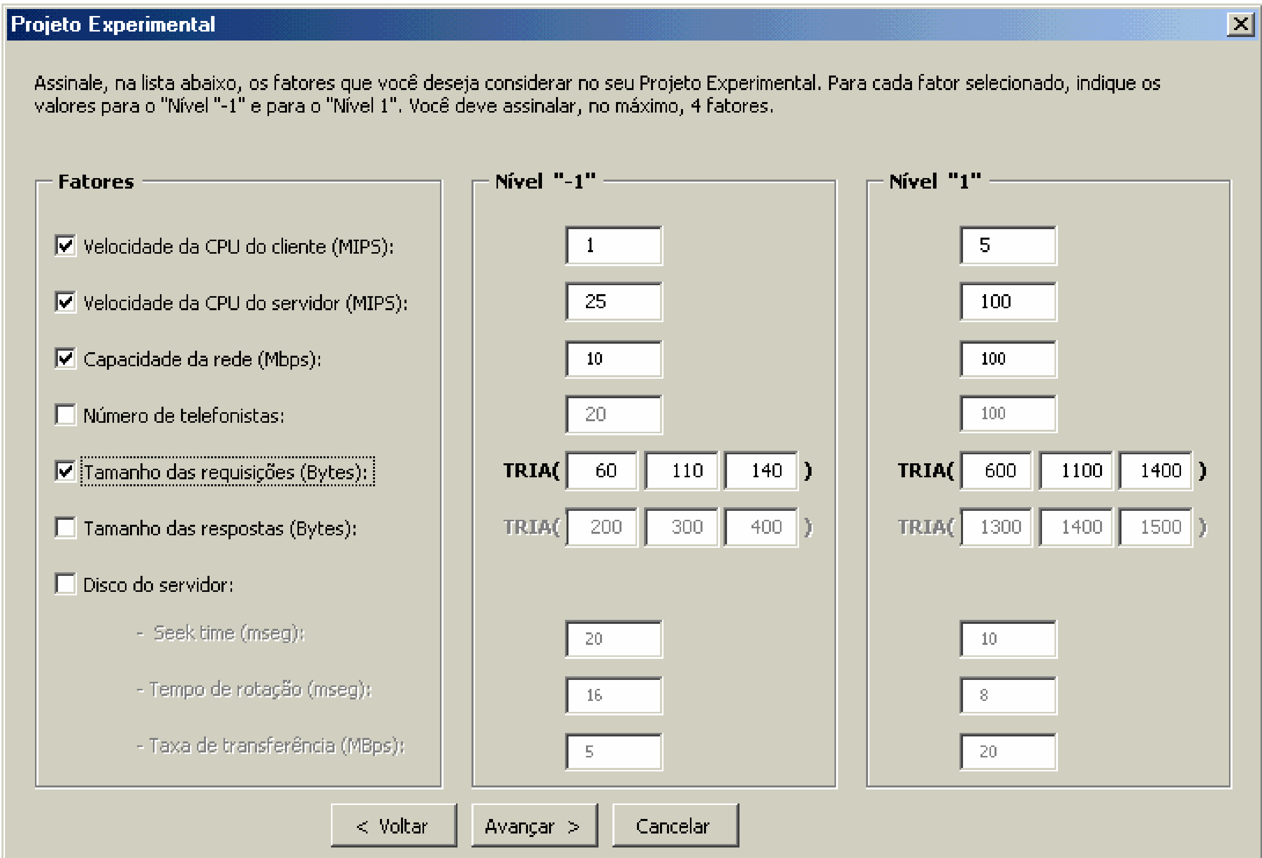

Figura 7.14: Tela da interface que permite ao analista escolher quatro entre os sete fatores disponíveis

Ao iniciar a execução do modelo, as rotinas em VBA forçam a apresentação da interface voltada à realização do projeto experimental com o modelo de Sistema Cliente/Servidor. A primeira tela (Figura 7.13), permite ao usuário iniciar os procedimentos.

Dando início ao projeto experimental, a tela da Figura 7.14 se apresenta ao analista. Através dela, é possível escolher os fatores desejados e seus respectivos limites. Na figura 7.14 observa-se que os quatro fatores referentes ao primeiro ensaio foram assinalados.

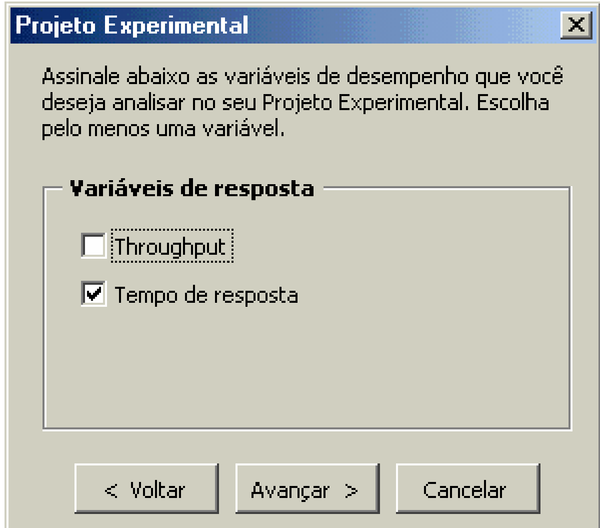

Figura 7.15: Escolha das variáveis de resposta.

Na tela da Figura 7.15, o analista define as variáveis de resposta sobre as quais deseja que sejam verificados os efeitos dos fatores.

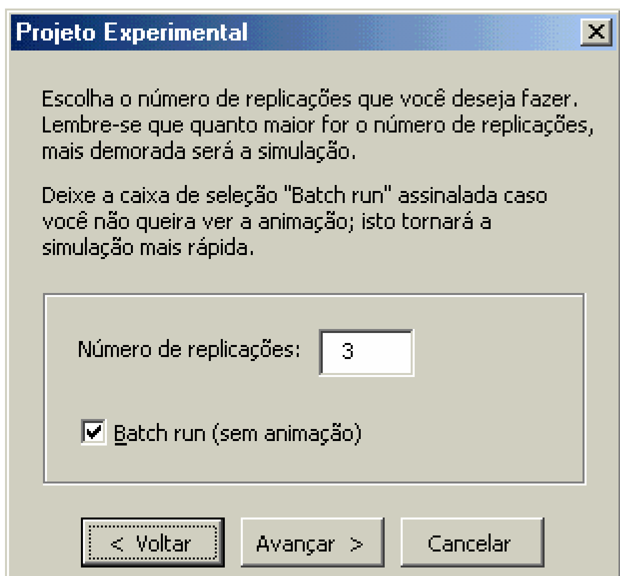

Figura 7.16: Definição do número de replicações

Finalmente, apresenta-se a tela da Figura 7.16. Com a definição do número de replicações que devem ser realizadas para cada combinação de fatores e níveis, encerra-se a fase de definições dos parâmetros para o projeto de experimentos. A tela seguinte, mostrada na figura 7.17, traz um resumo de todas as informações fornecidas pelo usuário.

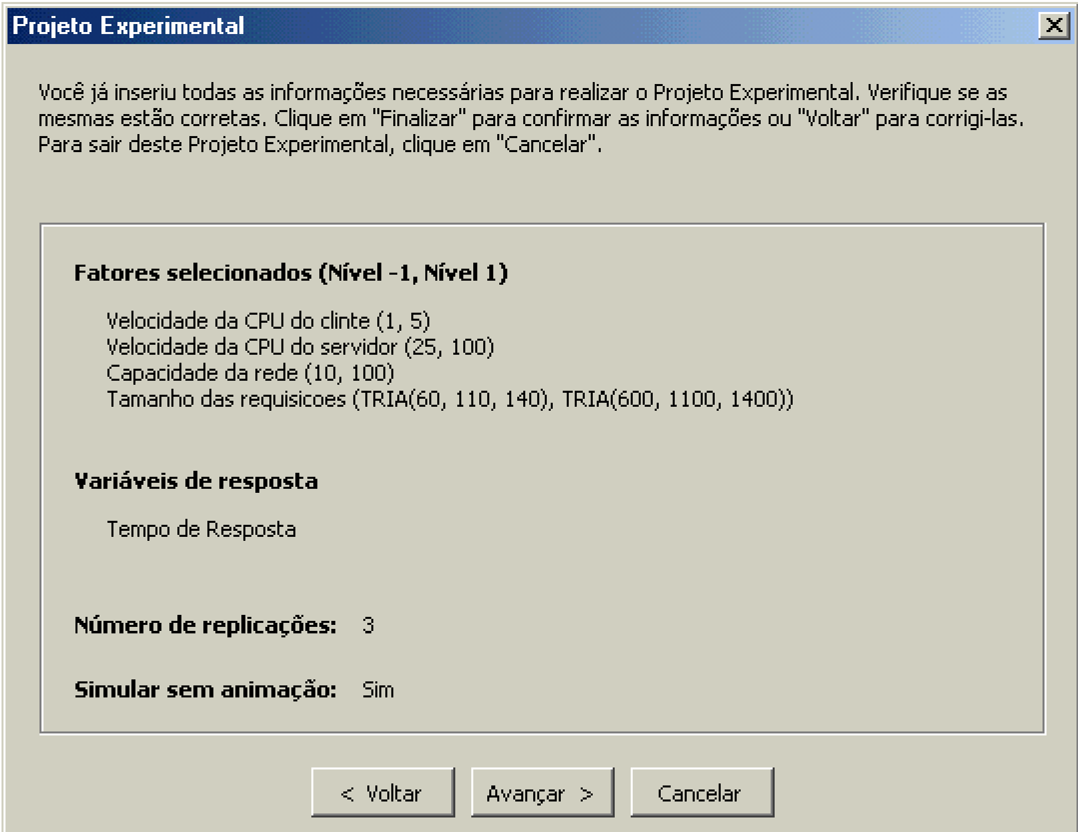

Figura 7.17: Resumo das informações fornecidas para o projeto experimental.

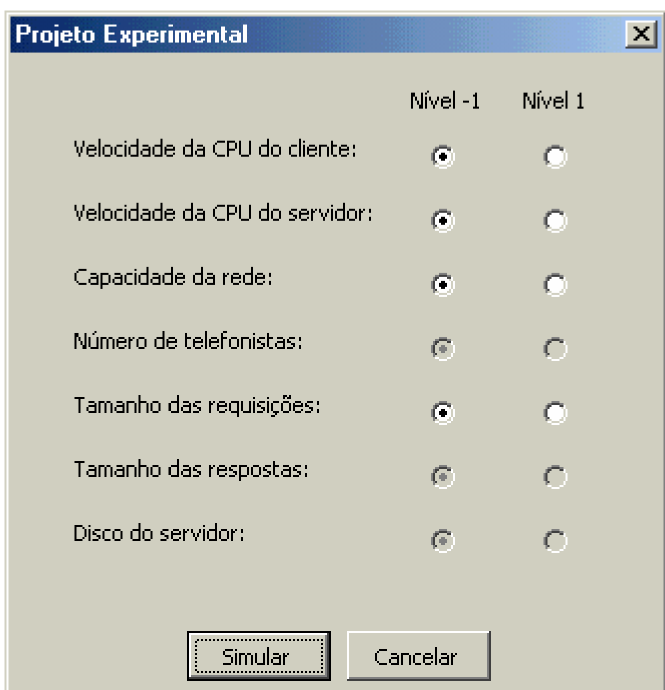

Figura 7.18: Visualização da combinação do primeiro experimento.

Uma vez que o usuário aceite as informações resumidas na Figura 7.17 (ele poderá retornar as telas anteriores e modificar os dados), o passo seguinte é o início das simulações. Antes de cada uma das 16 simulações, a combinação entre os níveis dos quatro fatores referentes àquele experimento, aparece ao usuário telas semelhantes a apresentada na Figura 7.18. Nesta figura a primeira combinação está definida (todos os fatores no nível (-1)). Na medida em que cada experimento é realizado, telas semelhantes se apresentam para que o usuário defina a próxima combinação entre os níveis dos fatores. Embora este passo possa ser automatizado, o objetivo aqui é forçar o analista a raciocinar sobre cada um dos experimentos realizados.

Na medida em que se iniciam as simulações, o programa cria e exibe planilhas eletrônicas com os resultados das variáveis de resposta e de outras variáveis sobre as quais o analista possa ter interesse. Estes resultados serão apresentados logo a seguir.

Simulados os 16 experimentos relativos ao primeiro ensaio, o analista pode executar novamente o programa para realizar os experimentos definidos pelas combinações dos fatores escolhidos para o segundo ensaio. A tela mostrada na Figura 7.19, semelhante à tela da Figura 7.14, mostra o inicio do novo procedimento. Observe que agora estão os novos fatores escolhidos.

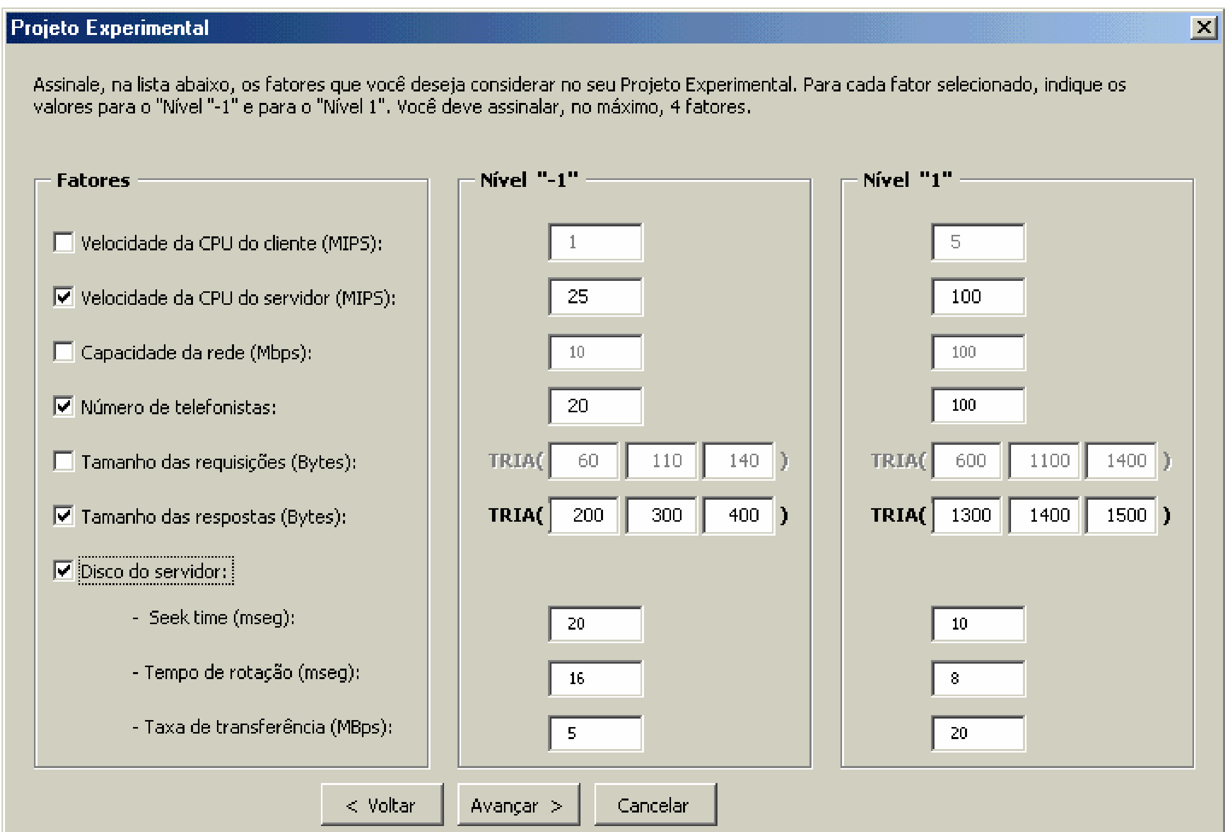

Figura 7.19: Nova configuração para o projeto experimental do ensaio dois.

Resultados do Primeiro Ensaio

A tabela 7.19 apresenta os quatro fatores escolhidos para o primeiro ensaio. Esta tabela é apresentada na planilha eletrônica construída pelo programa.

Fatores selecionados	Nível -1	Nível 1
Velocidade da CPU do cliente	1	5
Velocidade da CPU do servidor	25	100
Capacidade da rede	10	100
Tamanho das requisicoes	TRIA(60, 110, 140)	TRIA(600, 1100, 1400)

Tabela 7.19: Fatores selecionados para o primeiro ensaio

Após a simulação de cada experimento, o programa acrescenta linhas à tabela de resultados, reproduzida aqui na Tabela 7.20. Observando-se a linha de títulos (cabeçalho) verifica-se que os valores sob os títulos A, B, C e D de cada linha, indicam, em cada experimento, que nível foi empregado para cada um dos fatores. Observa-se, também, que as seis colunas seguintes, sob os títulos AB, AC,..., CD, representam as interações entre pares de fatores. As colunas seguintes, Y1, Y2 e Y3, mostram os resultados da variável de resposta, em cada uma das três replicações. Sob o título Ymédio, encontra-se a média aritmética destes três resultados. Finalmente, as três últimas colunas apresentam as diferenças (erros experimentais) dos resultados de cada replicação em relação ao valor médio.

Exp	I	A	B	C	D	AB	AC	AD	BC	BD	CD	Y1	Y2	Y3	Ymédio	e1	e2	e3
1	1	-1	-1	-1	-1	1	1	1	1	1	1	3814,62	3725,15	3750,62	3763,46	51,1602	-38,314	-12,846
2	1	-1	-1	-1	1	1	1	-1	1	-1	-1	3814,62	3725,15	3750,62	3763,46	51,1602	-38,314	-12,846
3	1	-1	-1	1	-1	1	-1	1	-1	1	-1	3814,62	3725,15	3750,62	3763,46	51,1602	-38,314	-12,846
4	1	-1	-1	1	1	1	-1	-1	-1	-1	1	3814,62	3725,15	3750,62	3763,46	51,1602	-38,314	-12,846
5	1	-1	1	-1	-1	-1	1	1	-1	-1	1	3814,62	3725,15	3750,62	3763,46	51,1602	-38,314	-12,846
6	1	-1	1	-1	1	-1	1	-1	-1	1	-1	3814,62	3725,15	3750,62	3763,46	51,1602	-38,314	-12,846
7	1	-1	1	1	-1	-1	-1	1	1	-1	-1	3814,62	3725,15	3750,62	3763,46	51,1602	-38,314	-12,846
8	1	-1	1	1	1	-1	-1	-1	1	1	1	3814,62	3725,15	3750,62	3763,46	51,1602	-38,314	-12,846
9	1	1	-1	-1	-1	-1	-1	-1	1	1	1	3781,76	3696,66	3793,44	3757,28	24,476	-60,627	36,1514
10	1	1	-1	-1	1	-1	-1	1	1	-1	-1	3781,76	3696,66	3793,44	3757,28	24,476	-60,627	36,1514
11	1	1	-1	1	-1	-1	1	-1	-1	1	-1	3781,76	3696,66	3793,44	3757,28	24,476	-60,627	36,1514
12	1	1	-1	1	1	-1	1	1	-1	-1	1	3781,76	3696,66	3793,44	3757,28	24,476	-60,627	36,1514
13	1	1	1	-1	-1	1	-1	-1	-1	-1	1	3781,76	3696,66	3793,44	3757,28	24,476	-60,627	36,1514
14	1	1	1	-1	1	1	-1	1	-1	1	-1	3781,76	3696,66	3793,44	3757,28	24,476	-60,627	36,1514
15	1	1	1	1	-1	1	1	-1	1	-1	-1	3781,76	3696,66	3793,44	3757,28	24,476	-60,627	36,1514
16	1	1	1	1	1	1	1	1	1	1	1	3781,76	3696,66	3793,44	3757,28	24,476	-60,627	36,1514

Tabela 7.20: Resultados das simulações primeiro ensaio.

De posse dos dados com os resultados das simulações e suas replicações, fórmulas embutidas sob as células da planilha, calculam os efeitos de cada um dos fatores e de suas interações sobre a variável de interesse, no caso, o tempo de resposta do servidor de base de dados do sistema Cliente/Servidor. A tabela com os valores dos efeitos é aqui reproduzida na tabela 7.21.

Os resultados dos cálculos dos efeitos mostrados na tabela 7.21 podem causar surpresa se não analisados no detalhe. Os principais elementos a serem observados encontram-se na primeira e na última coluna. A primeira, apresenta os fatores (A, B, C, D, AB, ..., CD e Erro) e a última a responsabilidade de cada um deles (provida em valores percentuais) sobre a variação da variável de resposta. Observe que, praticamente, todo o peso da responsabilidade, recai sobre o erro. A interpretação, neste caso, é que nenhum dos fatores escolhidos é responsável direto por qualquer variação sobre o Tempo de Resposta. Propositalmente, escolheu-se fatores que se sabia teriam pouca influencia. Neste caso, o modelo de análise foi sensível e mostrou-se adequado.

É possível pressentir que estes seriam resultados esperados. Verificando-se a coluna Ymédio da tabela 7.20, nota-se que as respostas são praticamente iguais ao longo dos 16 experimentos. Assim, era mesmo de se esperar que o modelo indicasse que estes fatores não influenciam a variável escolhida. Como se verá, quando se examinar os resultados do segundo ensaio, as respostas serão outras.

	Soma	Soma/16	SQ	Variação	%
I	60166	3760,37			
A	-49,439	-3,09	458,296	0,00579	0,57928
B	-0,0005	-3E-05	4E-08	5E-13	5E-11
C	-0,0005	-3E-05	5,5E-08	7E-13	7E-11
D	0,00035	2,2E-05	2,3E-08	2,9E-13	2,9E-11
AB	1,4E-05	8,7E-07	3,6E-11	4,6E-16	4,6E-14
AC	-4E-06	-2E-07	2,7E-12	3,4E-17	3,4E-15
AD	7,2E-06	4,5E-07	9,8E-12	1,2E-16	1,2E-14
BC	1,4E-12	8,5E-14	3,5E-25	4,4E-30	4,4E-28
BD	1,8E-12	1,1E-13	6,2E-25	7,8E-30	7,8E-28
CD	-0,0003	-2E-05	1,6E-08	2E-13	2E-11
Erro			78656,3	0,99421	99,4207
Total			79114,6	1	100

Tabela 7.21: Cálculos dos efeitos do primeiro ensaio.

Resultados do Segundo Ensaio

A tabela 7.22 apresenta os quatro fatores escolhidos para o segundo ensaio. Esta tabela é apresentada na planilha eletrônica construída pelo programa para este segundo conjunto de experimentos.

Fatores selecionados	Nível -1	Nível 1
A - Velocidade da CPU do servidor	25	100
B - Número de telefonistas	20	100
C - Tamanho das respostas	TRIA(200, 300, 400)	TRIA(1300, 1400, 1500)
D - Disco do servidor (ST, TR, TT)	20, 16, 5	10, 8, 20

Tabela 7.22: Fatores selecionados para o segundo ensaio.

A tabela 7.23 abaixo reproduz os resultados dos 16 novos experimentos realizados. Assim como no primeiro ensaio, cada um deles também foi replicado três vezes. Diferentemente dos números apresentados sob a coluna Ymédio da Tabela 7.20, aqui observa-se um razoável variabilidade nos resultados, na medida em que se alteram as combinações entre os níveis dos fatores.

Exp	I	A	B	C	D	AB	AC	AD	BC	BD	CD	Y1	Y2	Y3	Ymédio	e1	e2	e3
1	1	-1	-1	-1	-1	1	1	1	1	1	1	3814,62	3725,15	3750,62	3763,46	51,1602	-38,314	-12,846
2	1	-1	-1	-1	1	1	1	-1	1	-1	-1	1892,79	1844,39	1940,61	1892,6	0,19413	-48,211	48,0173
3	1	-1	-1	1	-1	1	-1	1	-1	1	-1	3814,62	3725,15	3750,62	3763,46	51,1602	-38,314	-12,846
4	1	-1	-1	1	1	1	-1	-1	-1	-1	1	1892,79	1844,39	1940,61	1892,6	0,19414	-48,211	48,0173
5	1	-1	1	-1	-1	-1	1	1	-1	-1	1	16723,3	16970,1	16955,4	16882,9	-159,61	87,138	72,4742
6	1	-1	1	-1	1	-1	1	-1	-1	1	-1	9278,81	8851,31	9371,67	9167,26	111,546	-315,96	204,41
7	1	-1	1	1	-1	-1	-1	1	1	-1	-1	16723,3	16970,1	16955,4	16882,9	-159,61	87,138	72,4742
8	1	-1	1	1	1	-1	-1	-1	1	1	1	9278,81	8851,31	9371,67	9167,26	111,546	-315,96	204,41
9	1	1	-1	-1	-1	-1	-1	-1	1	1	1	3814,62	3725,15	3750,62	3763,46	51,1602	-38,314	-12,846
10	1	1	-1	-1	1	-1	-1	1	1	-1	-1	1892,79	1844,39	1940,61	1892,6	0,19414	-48,211	48,0173
11	1	1	-1	1	-1	-1	1	-1	-1	1	-1	3814,62	3725,15	3750,62	3763,46	51,1602	-38,314	-12,846
12	1	1	-1	1	1	-1	1	1	-1	-1	1	1892,79	1844,39	1940,61	1892,6	0,19414	-48,211	48,0173
13	1	1	1	-1	-1	1	-1	-1	-1	-1	1	16723,3	16970,1	16955,4	16882,9	-159,61	87,1379	72,4743
14	1	1	1	-1	1	1	-1	1	-1	1	-1	9278,81	8851,31	9371,67	9167,26	111,546	-315,96	204,41
15	1	1	1	1	-1	1	1	-1	1	-1	-1	16723,3	16970,1	16955,4	16882,9	-159,61	87,1379	72,4743
16	1	1	1	1	1	1	1	1	1	1	1	9278,81	8851,31	9371,67	9167,26	111,546	-315,96	204,41

Tabela 7.23: Resultados do segundo ensaio.

Uma análise do cálculo da responsabilidade pela variação dos resultados do tempo de resposta mostra que esta está dividida principalmente, entre três dos quatro fatores escolhidos (Tabela 7.24 abaixo).

	Soma	Soma/16	SQ	Variação	%
I	126825	7926,57			
A	-0,001	-6E-05	1,9E-07	1,2E-16	1,2E-14
B	81576,6	5098,54	1,2E+09	0,76702	76,7024
C	0,00201	0,00013	7,6E-07	4,7E-16	4,7E-14
D	-38346	-2396,6	2,8E+08	0,16948	16,9481
AB	-0,0007	-4E-05	8,3E-08	5,1E-17	5,1E-15
AC	-5E-12	-3E-13	5,6E-24	3,4E-33	3,4E-31
AD	0,00039	2,4E-05	2,8E-08	1,7E-17	1,7E-15
BC	0,00143	8,9E-05	3,8E-07	2,4E-16	2,4E-14
BD	-23379	-1461,2	1E+08	0,063	6,29997
CD	-0,0008	-5E-05	1,1E-07	6,7E-17	6,7E-15
Erro			805027	0,00049	0,04949
Total			1,6E+09	1	100

Tabela 7.24: Cálculos dos efeitos do segundo ensaio.

O fator B (número de telefonistas) é o principal responsável, com 76,7% do total. Como era de se esperar, na medida do aumento do número de operadores ligados ao sistema, intensificasse o número de requisições a base de dados. Assim, a limitação daquele recurso, aumenta as filas e, como conseqüência, o tempo de resposta. Além de confirmar as suspeitas do analista, o modelo fornece quantifica esta responsabilidade. Estas afirmações podem ser reafirmadas quando se observa que o segundo fator em responsabilidade sobre a variação, é o disco do servidor, com cerca de 17% da responsabilidade. Este recurso do sistema é claramente um gargalo, como se pode verificar na linha com o título Taxa de Utilização Disco Servidor da Tabela 7.25, que apresenta outras estatísticas do sistema, também geradas pelo programa.

Fatores	1	2	3	4	5	6	7	8
Velocidade da CPU do servidor	-1	-1	-1	-1	-1	-1	-1	-1
Número de telefonistas	-1	-1	-1	-1	1	1	1	1
Tamanho das respostas	-1	-1	1	1	-1	-1	1	1
Disco do servidor (ST, TR, TT)	-1	1	-1	1	-1	1	-1	1
Outras Estatísticas								
Throughput	2,09562	4,179267	7,625074	15,39532	1,930157	4,010961	6,426581	14,22129
Tempo de resposta (médio)	3763,465	1892,597	3763,465	1892,597	16882,94	9167,263	16882,94	9167,263
Tempo de resposta (mínimo)	1589,995	765,9152	1589,995	765,9155	6462,75	3156,672	6462,75	3156,672
Tempo de resposta (máximo)	6401,191	3332,659	6401,191	3332,659	31876,79	16024,83	31876,79	16024,83
Num. requisições respondidas	180,3333	366,3333	180,3333	366,3333	147	334	147	334
Taxa utilização CPU servidor (média)	2,12E-05	4,22E-05	2,12E-05	4,22E-05	2,2E-05	4,3E-05	2,2E-05	4,3E-05
Num. fila CPU servidor (médio)	0	0	0	0	1,54E-09	1,33E-09	1,54E-09	1,33E-09
Num. fila CPU servidor (máximo)	0	0	0	0	0,333333	0,333333	0,333333	0,333333
Num. fila CPU servidor (final)	0	0	0	0	0	0	0	0
Taxa utilização disco servidor (média)	0,999973	0,999981	0,999973	0,999981	0,999996	0,999995	0,999996	0,999995
Num. fila disco servidor (médio)	18,94913	18,8941	18,94912	18,89409	98,93158	98,87872	98,93158	98,87871
Num. fila disco servidor (máximo)	19	19	19	19	99	99	99	99
Num. fila disco servidor (final)	19	19	19	19	99	98,66667	99	98,66667

Tabela 7.25: Outras estatísticas geradas pelo programa.

Os dois ensaios mostrados neste estudo de caso deixam claro que o emprego de técnicas estatísticas, como projetos experimentais, associado a modelos de simulação, formam uma poderosa combinação de ferramentas que permitem o alcance de resultados promissores sobre o desempenho de sistemas complexos. Os modelos resultantes destes projetos autorizam, não só confirmar ou rejeitar hipóteses sobre que fatores são responsáveis sobre o melhor ou pior desempenho dos sistemas, bem como quantificar tais responsabilidades.

Sumário

Ao longo deste texto sempre se procurou deixar claro que seu principal objetivo não é o de ensinar ao leitor o emprego desta ou daquela linguagem de simulação. Seu propósito é levar ao usuário da simulação o conhecimento de algumas técnicas estatísticas, fundamentais e necessárias, para que este possa empregar corretamente a modelagem e simulação de sistemas. Este texto pretende ser uma boa referencia sobre como e quando empregar estas técnicas na solução de problemas. É pressuposto básico o conhecimento de alguma linguagem de simulação. Como visto neste capítulo e em capítulos anteriores, sempre que possível se procura exemplificar as diversas situações utilizando-se uma ferramenta de propósito geral para a modelagem e simulação de sistemas, no caso, o ambiente Arena. Sua escolha se deve basicamente à sua popularidade, tanto no âmbito acadêmico quanto profissional, de áreas ligadas à indústria e aos serviços.