

6

ANÁLISE DOS RESULTADOS DA SIMULAÇÃO

Esse capítulo é dedicado ao estudo e a compreensão dos métodos e procedimentos relacionados com a análise de resultados de simulações de sistemas. Como se verá, os tratamentos empregados implicam no conhecimento preliminar de técnicas estatísticas básicas e de alguns tópicos relacionados com teoria das probabilidades. Na medida do necessário, serão realizadas pequenas incursões e revisões destes elementos. Ao longo do capítulo serão tratadas as diferentes abordagens e técnicas que devem ser empregadas, considerando o tipo de sistema sob análise. Estes são qualificados em dois grandes grupos: sistemas terminais e sistemas não-terminais. Estes são os principais pontos abordados neste capítulo.

Tópicos

- 6.1 Introdução**
- 6.2 Experimentação e Análise de Resultados**
- 6.3 Confiança Estatística**
- 6.4 Classificando os Sistemas em Terminais ou Não-Terminais**
- 6.5 Análise de Sistemas Terminais**
- 6.6 Usando o *Arena Output Analyzer* para tratar Sistemas Terminais**
- 6.7 Análise de Sistemas Não-Terminais**
- 6.8 Usando o *Arena Output Analyzer* para tratar Sistemas Não-Terminais**
- Sumário**
- Exercícios**

6.1 Introdução

O processo de análise dos resultados de simulações realizadas com um modelo computacional, lida com os dados obtidos a partir dos experimentos. Seu principal objetivo é permitir a realização de inferências e previsões sobre o comportamento e o desempenho do sistema real sob análise.

A principal razão para uma maior atenção aos processos de análise dos resultados das simulações, baseia-se no fato de, em geral, os modelos apresentarem um comportamento estocástico à semelhança dos sistemas que estão imitando. Desta maneira, adota-se, com relação aos resultados dos modelos, o mesmo tratamento estatístico que seria empregado sobre sistemas reais.

Para se possa concluir sobre o desempenho de um sistema, por meio de resultados obtidos a partir de um modelo, é necessário observar o comportamento das variáveis de resposta, na medida da realização de experimentos simulados. Uma vez realizados os experimentos, estima-se o comportamento do sistema real por um processo de inferência a partir do conjunto de resultados obtidos.

Como em qualquer processo de estimação, a preocupação em minimizar erros deve ser constante. Portanto, a análise e o tratamento dos dados resultantes dos experimentos realizados servirá, também, como etapa condutora de procedimentos técnicos que permitam reduzir os erros resultantes do processo de inferência.

Os procedimentos necessários à análise dependem do tipo de sistema que está sendo tratado. Com este propósito, os sistemas estarão sendo classificados como *terminais* ou *não-terminais*. Nos tópicos que seguem, aborda-se as técnicas empregadas e as diferenças de suas aplicações, considerando os dois tipos de sistemas.

6.2 Experimentação e Análise de Resultados

O procedimento de análise dos resultados da simulação inicia com a seleção de variáveis relacionadas ao tipo de desempenho que se deseja auferir no sistema. Estas variáveis podem ser, por exemplo, elementos contadores de ocorrências ou ainda serem obtidas de alguma expressão incluindo médias, variâncias, etc.. Alguns exemplos destas variáveis são: tempo no sistema, tempo nas filas, número de entidades nas filas, utilização de recursos, vazão (*throughput*), etc..

Toda a simulação de sistemas que contém variáveis aleatórias gera resultados aleatórios. Para que se possa fazer qualquer tipo de inferência sobre estes resultados ou, mais especificamente, sobre aqueles que medem o desempenho do sistema modelado, é necessário realizar análises apropriadas. Algumas questões que devem ser inicialmente tratadas dizem respeito aos seguintes aspectos:

- Qual é a duração apropriada de uma rodada de simulação?
- Como interpretar corretamente seus resultados?
- Como analisar corretamente as diferenças obtidas em cada uma das replicações?

Estas e outras indagações são objeto dos tópicos que serão abordados a seguir. Várias técnicas são utilizadas ao longo deste processo, desde a avaliação visual de dados dispostos graficamente, até técnicas estatísticas um pouco mais sofisticadas, como médias móveis, correlação e a definição de intervalos de confiança.

6.3 Confiança Estatística

A determinação de intervalos de confiança para as variáveis de interesse que medem o desempenho do sistema é um componente fundamental no processo de análise de resultados. Um

o intervalo de confiança compreende um intervalo numérico que possui uma probabilidade igual a $(1 - \alpha)$ de incluir o verdadeiro valor da variável ou medida de desempenho sob análise, onde, $(1 - \alpha)$ é denominado o nível de confiança do intervalo. Observe que $(1 - \alpha)$ pode assumir valores como 99% ou 90%. Desta forma, α , será o erro admitido ao se concluir sobre a presença do verdadeiro valor da variável no intervalo calculado.

Exemplo 6.1 Para os dados da tabela 6.1 deseja-se construir intervalos de confiança de 95% ($\alpha=0,05$) e de 99% ($\alpha=0,01$). Assumindo que os valores para a variável aleatória X (Tempo Médio na Fila) são normalmente distribuídos, o semi-intervalo h permitirá o cálculo do intervalo de confiança a $1 - \alpha$, para o verdadeiro valor da média μ , que estará centrado em X .

Número da Replicação	Tempo Médio na Fila
1	63,2
2	69,7
3	67,3
4	64,8
5	72,0

Tabela 6.1: Dados para o exemplo 6.1.

O semi-intervalo h é calculado por:

$$h = t_{n-1, 1-\alpha/2} \frac{S}{\sqrt{n}} \quad (6.1)$$

onde o valor de $t_{n-1, 1-\alpha/2}$ é o valor tabulado de t (ver anexo 3 - tabela distribuição t).

Uma vez que α reflete o nível de confiança para todo intervalo, no sentido de obter-se um intervalo bilateral usa-se o valor de $\alpha/2$ para computar o valor de h . Usando-se as equações 5.1 e 5.2 calcula-se o valor de \bar{X} , S^2 e S . Adicionalmente, busca-se os valores tabelados de t com 95% e 99% de confiança. Os valores calculados e tabelados são os seguintes:

$$\bar{X} = 67,4; \quad S^2 = 12,74 \quad S = 3,57$$

Assim, para 95% de confiança ($\alpha = 5\%$), $1 - \alpha/2 = 0,975$ e $t_{4, 0,975} = 2,78$. Já para 99% de confiança ($\alpha = 1\%$), $1 - \alpha/2 = 0,995$ e $t_{4, 0,995} = 4,60$.

Desta forma, os resultados dos valores de h são os seguintes:

$$\begin{aligned} h &= 4,44 \text{ (95\% de confiança)} \\ h &= 7,34 \text{ (99\% de confiança)} \end{aligned}$$

Como o intervalo é simétrico em torno da média, os limites serão os seguintes:

$$[\bar{X} - h, \bar{X} + h]$$

Portanto, o intervalo com 95% de confiança é [62,6; 71,84] e o intervalo com 99% de confiança é [60,06; 74,74].

Como se pode verificar, o tamanho do intervalo depende do nível de confiança desejado. O nível mais elevado (99%) requer um intervalo maior em comparação com o nível de confiança mais baixo. Verifica-se que o número de replicações n e desvio-padrão s , também são utilizados na obtenção de t e, por consequência, do intervalo de confiança. Os n valores dos dados da tabela foram obtidos das n replicações realizadas. Desta forma, três fatores influenciam a largura do intervalo de confiança:

1. O número de replicações n ;
2. O nível de confiança $(1 - \alpha)$ predefinido pelo analista;
3. Variação (S^2) associada à medida de desempenho sob análise.

O relacionamento entre estes três fatores ocorre da seguinte maneira:

1. Na medida em que se aumenta o número de replicações, reduz-se a largura do intervalo de confiança;
2. Na medida em que o nível de confiança aumenta, aumenta também a largura do intervalo de confiança. O intervalo com 99% de confiança será maior do que o aquele com 95% de confiança, uma vez mantidos os demais elementos;
3. Na medida que a variação (S^2) aumenta, cresce a largura do intervalo de confiança.

A variância da medida de desempenho de um sistema, e por consequência de seu modelo, é dependente dos parâmetros deste sistema. Não sendo possível modificar os elementos do modelo que alterem sua natureza estocástica, a liberdade de atuação do analista fica restrita a possibilidade de lidar com alterações no número de replicações e no nível de confiança desejado aos resultados. Desta maneira, com base nas afirmações acima têm-se três situações que definem as relações entre o número de replicações e o nível de confiança:

1. Se o nível de confiança é fixo, um grande número de replicações resultará num menor intervalo de confiança.
2. Se o tamanho do intervalo de confiança é fixo, um grande número de replicações resultará em um maior grau de confiança.
3. Se o número de replicações é fixo, um alto nível de confiança resultará um grande intervalo de confiança.

Pelas afirmações acima, nota-se que a definição do número de replicações que deverão ser realizadas em um estudo onde se emprega a técnica da simulação é um fator relevante na determinação do intervalo de confiança. Como se verá mais adiante neste capítulo, a definição sobre o quanto e como deve um sistema ser simulado depende, fundamentalmente, do tipo de sistema que se está tratando.

6.4 Classificando os Sistemas em Terminais ou Não-Terminais

Tendo em vista as análises pretendidas, dividem-se os sistemas em duas classes: sistemas terminais e sistemas não-terminais. Os classificados como sistemas terminais, apresentam condições iniciais fixas (condições estas que o sistema volta a assumir ao início de cada rodada de simulação) e um evento que determina um fim natural para o processo de simulação (isto é, pode-se programar o fim de uma rodada de simulação). Como exemplos de sistemas terminais pode-se citar os serviços, tais como, restaurantes, bancos, lojas comerciais, etc.

As condições iniciais da maioria dos sistemas terminais apontam para um sistema *vazio e disponível*, isto é, todos os recursos do sistema estão disponíveis e não existem entidades ou *clientes* esperando em filas. Para os sistemas comerciais, as condições de término geralmente estão associadas a um evento que aponta para o fim de um dia de trabalho com 8 a 10 horas de funcionamento. Já os sistemas considerados não-terminais não possuem condições iniciais fixas, nem um evento que determina o fim do processo de simulação. Como exemplos mais comuns deste tipo de sistemas pode-se citar os serviços de 24 horas, os hospitais, os sistemas de comunicação, etc.

Nem todos os sistemas são facilmente classificados. Um caso típico é o de uma indústria funcionando em turnos de oito horas por dia durante cinco dias por semana. Precipitadamente pode-se imaginar este sistema como terminal no qual diariamente se iniciariam os processos de manufatura que se encerrariam ao fim do dia. Pode-se, talvez, até imaginar um ciclo um pouco mais longo que teria início toda segunda-feira e término as sextas. No entanto, considerando que ao final de um turno de oito horas, todo o trabalho a ser realizado possa ser retomado no dia seguinte do ponto em que foi suspenso no dia anterior, incluindo-se aí os intervalos de fins de semana, este sistema pode ser considerado como não-terminal.

Um contra exemplo a este tipo de consideração é quando da ocorrência de uma troca do lote de produtos a serem fabricados. Geralmente, a ocorrência deste fato provoca uma série de conseqüências sobre o sistema ocasionando uma quebra das rotinas e, portanto, do estado *em regime* do sistema. Sistemas de manufatura que funcionam desta forma são considerados terminais e seu ciclo inicia com o *setup* dos equipamentos para a produção do lote especificado pela programação e finda com o alcance das metas de produção. Este é um regime típico de células de fabricação com alta flexibilidade.

Outro tipo de sistema que pode gerar dúvidas quanto a sua classificação são aqueles envolvendo projetos de longa duração. Projetos envolvendo a construção de prédios, de navios, de equipamentos industriais de grande porte (um forno ou caldeira, por exemplo), etc., são alguns exemplos. Também se pode incluir aí a programação de eventos envolvendo projetos de manutenção preventiva, especialmente a de equipamentos de grande porte. Ocorre que estes costumam ser de longa duração o que, muitas vezes nos leva a confundi-los com sistemas não-terminais. O fato é que estes sistemas possuem condições iniciais determináveis e um evento término. Se, por acaso, o projeto for repetido, as mesmas condições iniciais e evento término podem ser aplicados, tratando-se, portanto, de sistemas terminais de longo prazo.

Uma vez que o sistema sob análise tenha sido classificado, o passo seguinte no processo de análise dos resultados da simulação é o procedimento a ser adotado na análise. Estes procedimentos seguem metodologias distintas de acordo com a classificação adotada.

É importante citar que para os projetos classificados como não-terminais, o interesse das análises recai sobre o comportamento do sistema quando “em regime”, isto é, quando as variáveis de estado não mais sofrem a ação ou dependência do período transiente inicial. O que se deseja é a avaliação do sistema sobre um longo período simulado, uma vez que não existe nenhuma condição de término da simulação. Ao contrário, nos sistemas considerados terminais se conhecem perfeitamente as condições iniciais e o chamado evento término. Desta forma, pode-se simular o sistema quantas vezes se desejar. A cada vez repetem-se as mesmas condições iniciais e o mesmo evento termina a rodada de simulação. Os procedimentos a serem adotados na realização das análises sobre os dois tipos de sistemas serão abordados nos tópicos que seguem.

6.5 Análise de Sistemas Terminais

O objetivo de simular sistemas terminais é compreender seu comportamento ao longo de um período predeterminado e com duração fixa. Uma vez que as condições iniciais e o período simulado são fixos, o único fator controlável é o número de replicações. Desta forma, os procedimentos de análise para sistemas terminais consistem em realizar um determinado número de replicações, verificar a variância da medida de desempenho selecionada e determinar se o intervalo de confiança resultante encontra-se dentro de limites aceitáveis. O fluxograma apresentado na figura 6.1 sintetiza os passos a serem realizados neste procedimento.

Por exemplo, se a variável de desempenho for o número médio de entidades na fila F , o primeiro passo é conduzir uma execução piloto de n replicações. O segundo passo é computar o intervalo de confiança para a variável de desempenho (número médio de entidades na fila F). Para tanto, deve-se utilizar os valores do número médio de entidades em F , obtidos em cada uma das n replicações (geralmente armazenados em um arquivo de dados). O terceiro passo é testar se o intervalo de confiança obtido para a variável é aceitável. Se for maior que o desejado, determina-se o número necessário de replicações adicionais, de tal forma que o novo intervalo de confiança encontre-se dentro dos limites preestabelecidos. Finalmente, após a execução das replicações adicionais recalcula-se o novo intervalo de confiança usando todos os dados obtidos. Os dois últimos passos são repetidos até que os limites do intervalo sejam satisfatórios.

6.5.1 Tamanho da Amostra (número de replicações) para a Determinação da Média

Suponha que se deseja estimar o desempenho de um sistema considerando o valor médio de uma variável de controle qualquer, com um nível de precisão de $\pm r\%$ (sobre esta variável) e um nível de confiança de $100(1-\alpha)\%$. O tamanho da amostra piloto pode ser definido de acordo com um procedimento estatístico, como o que se apresenta a seguir ou de uma forma intuitiva,

dependendo da experiência do analista e de seu conhecimento sobre o comportamento geral do sistema sob análise.

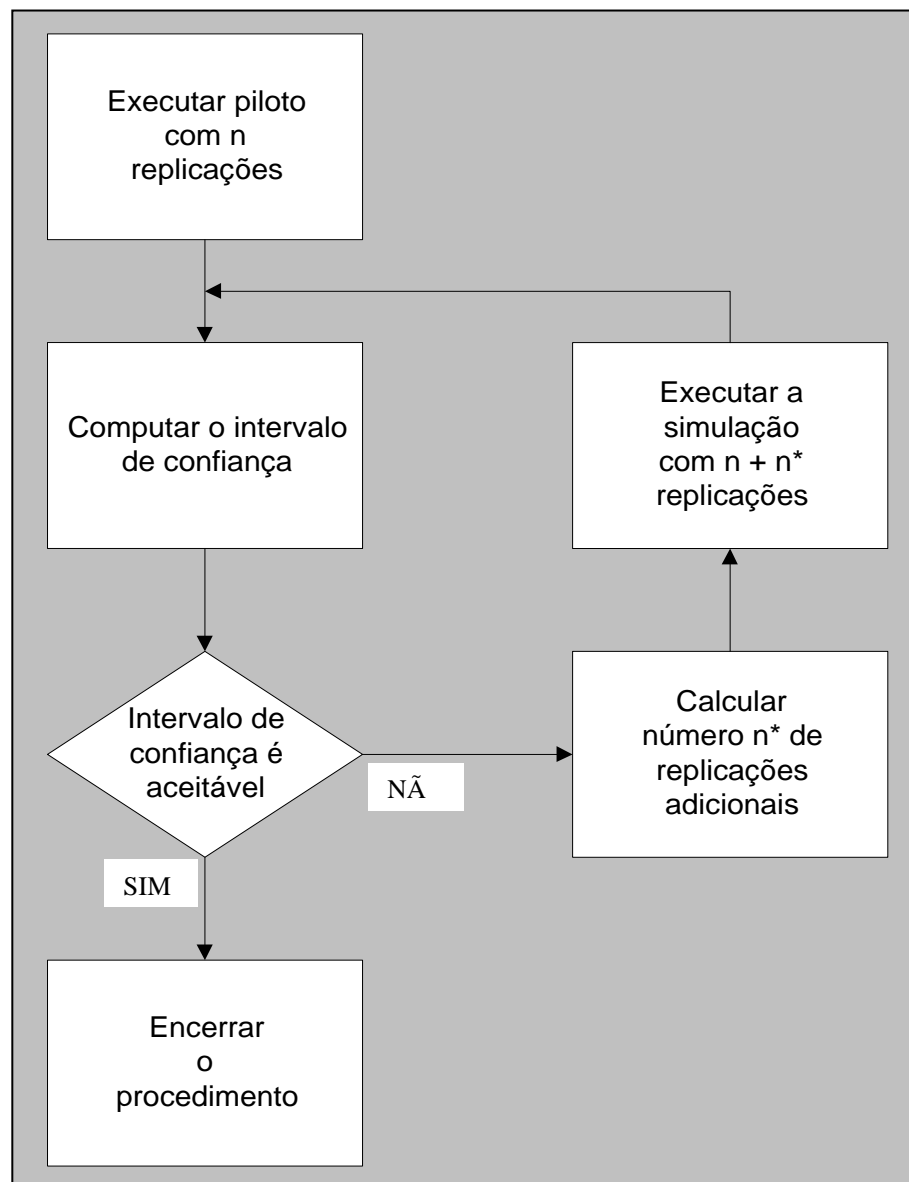


Figura 6.1: Fluxograma do procedimento de análise de sistemas terminais

Sabe-se que para uma amostra de tamanho n , o intervalo de confiança $100(1-\alpha)\%$ da média populacional da variável de interesse é dado por:

$$\bar{x} \pm z \frac{s}{\sqrt{n}} \quad (6.2)$$

A precisão desejada para $r\%$ implica que o intervalo de confiança deve ser:

$$[\bar{x} (1-r/100), \bar{x} (1+r/100)]. \text{ Logo,}$$

$$\begin{aligned}\bar{x} \pm z \frac{s}{\sqrt{n}} &= \bar{x} \left(1 \pm \frac{r}{100}\right) \\ z \frac{s}{\sqrt{n}} &= \bar{x} \frac{r}{100} \\ n &= \left(\frac{100zs}{r\bar{x}}\right)^2\end{aligned}\tag{6.3}$$

Na última fórmula, z é o valor da variável normal padronizada para o nível de confiança desejado.

Exemplo 6.2 Com base num teste preliminar a média amostral de uma variável de controle é de 20 segundos com desvio padrão de 5 segundos. Quantas replicações são necessárias para que a variável de controle tenha uma precisão de 1 segundo com 95% de confiança?

Precisão requerida é de 1 em 20 = 5%

Os parâmetros para a aplicação da fórmula são:

$$\bar{x} = 20 \quad s = 5 \quad z = 1,960 \quad r = 5$$

$$n_{5\%} = [(100 \times 1,960 \times 5) / (5 \times 20)]^2 = (9,8)^2 = 96,04$$

Arredondando-se para cima o valor calculado obtêm-se o número de replicações necessárias: 97 replicações. Com este número, deve-se alcançar o intervalo de confiança com a amplitude requerida, de tal forma que a verdadeira média populacional encontre-se dentro de seus limites.

Exemplo 6.3 Determinar os limites de um intervalo de confiança, considerando $\alpha = 5\%$, e uma precisão em torno de 5% para o valor da média amostral. Os parâmetros fornecidos são:

$$\bar{X} = 115,3 \quad s = 5,7 \quad z = 1,960 \quad r = 5$$

Aplicando-se a formulação (6.3) anteriormente desenvolvida, obtêm-se:

$$n_{5\%} = [(100 \times 1,960 \times 5,7) / (5 \times 115,3)]^2 = (1,94)^2 = 3,75 \cong 4$$

Nos exemplos acima descritos vê-se a grande importância da variância amostral. No exemplo 6.2, a variância é muito grande e, como consequência, o número necessário de replicações também é. Já no exemplo 6.3, com uma variância menor em relação ao valor médio, o número de replicações decresceu. É interessante citar que os parâmetros utilizados no problema foram obtidos a partir de uma amostra preliminar (piloto) com 10 elementos (replicações). A

partir desta, determinou-se a necessidade de apenas quatro replicações para o alcance do intervalo desejado. De fato, todos estes cálculos não garantem, com 100% de certeza, o valor ideal de n . A metodologia utilizada apenas aponta para um valor aproximado. O processo é iterativo. Tomam-se os valores de média e desvio-padrão de uma amostra piloto e, a partir destes, procura-se chegar ao tamanho mínimo da amostra ou, neste caso, do número de replicações.

Alguns autores propõem que este processo iterativo seja estendido até o cálculo do intervalo. Ou seja, uma vez determinado o tamanho mínimo da amostra n , calcula-se o intervalo de confiança. Se este apresentar limites além dos desejados (r), faz-se uma nova estimativa para n . A nova estimativa n^* é obtida aplicando-se a seguinte fórmula:

$$n^* = [n(h/h^*)^2] \quad (6.4)$$

onde:

n^* = a nova estimativa para n

h = semi-intervalo de confiança obtido

h^* = semi-intervalo de confiança desejado

Exemplo 6.4 Considerando os valores obtidos no exemplo 6.3, após a simulação com $n = 4$ replicações, o intervalo de confiança para a média foi: [108,9; 121,34]. O valor obtido para o semi-intervalo ($h = 6,04$) é maior do que o valor desejado (5% da média = 5,77). Logo, a primeira estimativa para n , fornecida por (6.3) resultou em um intervalo maior do que o desejado. Aplicando-se a fórmula (6.4) para reestimar o valor de n^* , obtém-se:

$$n^* = [n(h/h^*)^2] = [4 (6,04 / 5,77)^2] = 4,38 \cong 5$$

O novo intervalo de confiança para a média, com $n = 5$ replicações é [109,58; 120,62]. Logo, simulando-se o mesmo problema com 5 replicações obtém-se o intervalo desejado para a variável de interesse. O valor obtido para h foi de 5,52 para uma média = 115,1.

6.6 Usando o *Arena Output Analyzer* para tratar Sistemas Terminais

No item 6.5 foi visto que para analisar o comportamento dos sistemas classificados como terminais é preciso eleger uma ou mais medidas de desempenho e obter, para cada uma delas uma amostra representativa que permita estimar, com a necessária confiança, seu valor verdadeiro.

Constatou-se que os procedimentos necessários à determinação do tamanho da amostra implicam, de fato, na realização de um determinado número de replicações, a partir do qual pode-se verificar a variância da medida de desempenho selecionada e determinar se o intervalo de confiança resultante encontra-se dentro de limites aceitáveis. Conforme sintetizado no fluxograma da figura 6.1, esta busca é um processo iterativo em que, a cada passo, procura-se chegar a valores mais acurados relativos ao valor real da variável de desempenho adotada.

Neste tópico, apresenta-se e faz-se uso de outra ferramenta computacional presente no ambiente *Arena*, o *Output Analyzer*. Assim como o *Input Analyzer*, o propósito desta ferramenta

também é auxiliar o analista nas tarefas que anteriormente abordadas no item 6.5 e cujo ponto principal é o cômputo do intervalo de confiança para a variável de interesse.

O *Arena Output Analyzer*, também é uma ferramenta independente. Uma vez instalada, pode ser acessada a partir do conjunto de ferramentas do Arena, via menu Programas do Windows (figura 6.2).

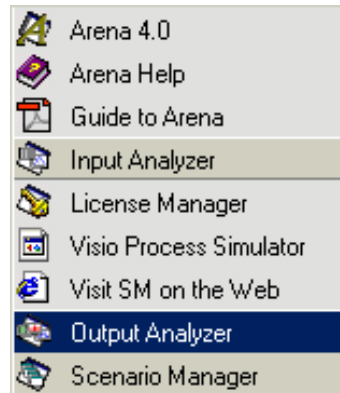


Figura 6.2: Ativando o *Output Analyzer* a partir do *Programas/Arena*.

No caso do *Output Analyzer*, a construção do intervalo de confiança requer, necessariamente, a existência de um conjunto de dados previamente armazenados em um arquivo. Embora existam algumas opções para a importação destes dados, neste texto trata-se especificamente da opção mais comum, isto é, aquela em que os dados para análise tenham sido gerados a partir da execução de um modelo no Arena.

Para um melhor entendimento dos passos que devem ser realizados, novamente considere-se o exemplo do modelo “*Transações Bancárias.doe*”, visto no capítulo 5. Trata-se do modelo de um sistema que pode ser, tipicamente, classificado como terminal e, portanto, adequado aos propósitos desta análise. As seguintes variáveis de desempenho foram designadas:

- Tempo médio no sistema de um cliente que faz uso dos caixas;
- Número médio de clientes na fila dos caixas.

O objetivo de se escolher estas duas variáveis ficará claro quando se iniciar as análises dos resultados. Uma delas, o número médio de clientes na fila possui alta variabilidade enquanto que a outra, tempo médio no sistema, é mais bem comportada. Na figura 6.3 é possível visualizar o setor do banco e a fila que se deseja analisar. Uma das variáveis de interesse (Fila Caixas) aparece em destaque.

Seguindo a metodologia proposta, inicialmente se estabelece um número de replicações para a determinação da amostra piloto. Em geral este número gira em torno de 15 replicações.

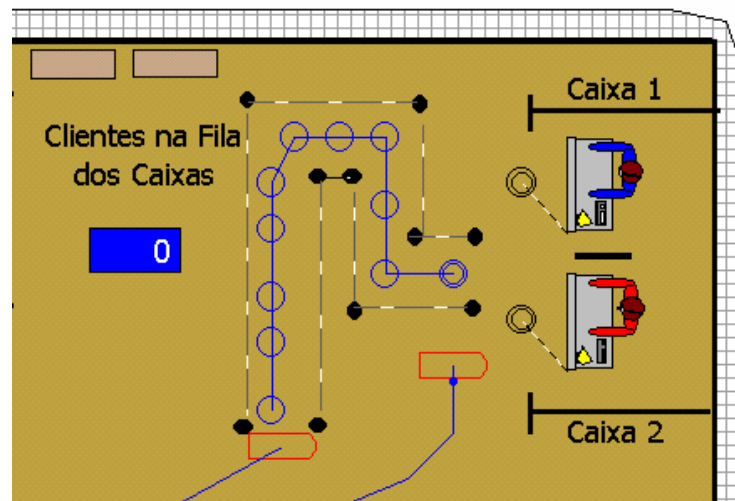


Figura 6.3: Fila dos Caixas no modelo Transações Bancárias

Você deve estar se perguntando: será este mais um daqueles números mágicos da estatística? Na verdade, pode-se estabelecer o valor que se deseja. Se este número for pequeno demais e a variável de resposta apresentar muita variabilidade, com toda certeza não será possível alcançar o intervalo de confiança desejado (pois a amostra será pequena) e, as estatísticas obtidas, não ajudarão a determinar, com maior precisão, o número ideal de replicações no processo iterativo. Por outro lado, se a amostra piloto for muito grande, pode-se estar desperdiçando recursos ao obter amostras maiores do que o necessário.

O bom senso aponta em estabelecer este valor com base nas seguintes variáveis: o esforço (tempo) necessário à realização de uma replicação; a variabilidade da variável de interesse e a precisão do intervalo de confiança desejado.

Na prática, como se está experimentando com um modelo computacional, o tempo de execução do programa pode ser um bom indicador. Se o tempo necessário à realização de uma replicação for pequeno, de tal maneira que sejam necessários apenas uns poucos minutos para se obter grandes amostras (50 ou mais elementos), então não há motivo algum para preocupações. Iniciar com uma amostra piloto com 30 ou mais elementos é plenamente justificado.

Se o tempo necessário à realização de uma replicação for grande (mais do que alguns minutos por replicação) então se deve julgar com base nas demais variáveis. Realize uma simulação com apenas uma replicação (é o *default* do Arena) e observe se existem grandes diferenças entre os valores mínimo e máximo da variável de interesse. Embora este fato, isoladamente não seja conclusivo, é um bom indicativo de sua variabilidade. Se a diferença for pequena, uma amostra piloto contendo de 10 a 15 elementos deve ser suficiente. Se, no entanto, esta diferença for grande e seu nível de exigência (precisão do intervalo de confiança) também for grande (menos do que 5% da média), então caro leitor prepare-se para longas simulações. Um café e um bom livro ou revista ajudarão a passar o tempo.

Definir uma rodada de simulação com várias replicações é uma tarefa simples no Arena. A figura 3.11 no capítulo 3 indica onde determinar o número de replicações (*Number of*

Replications), a partir da janela de diálogo *Run/Setup* localizada no menu principal do Arena. Para este exemplo, se estabelece o valor 10 para o número de replicações que constituirão a amostra piloto.

Antes de se iniciar a simulação, é necessário trabalhar o modelo no sentido de capturar (armazenar) os dados referentes as variáveis de interesse.

6.6.1 Armazenando os Dados da Simulação

Quando se executa um modelo de simulação, é possível visualizar, não somente todo o dinamismo das entidades e recursos, como também acompanhar, por meio de gráficos ou outras formas de animação, as variáveis sobre as quais se tem maior interesse. Vê-se na figura 6.3, por exemplo, a variável Fila Caixa. Durante uma simulação pode-se observar o valor desta variável uma vez que o mesmo será animado. A expressão NQ (Fila Caixas), que guarda seu valor, é exibida na tela (na figura o valor 0 aparece, uma vez que o modelo não se encontra em tempo de execução).

Assim como se pode observar os valores das variáveis de interesse durante uma simulação com animação, é possível compreender, pelo menos parcialmente, o comportamento destas quando, ao final da simulação, seus valores são apresentados nos relatórios do Arena. Ali são apresentados os valores médios, mínimos, máximos, etc. Embora, na maioria das vezes estas sejam informações suficientes para que se entenda o comportamento do sistema ou da variável, em algumas circunstâncias é preciso mais do que isso.

Por exemplo, quando se deseja analisar com mais detalhes o comportamento do estado da fila dos caixas do banco, faz-se necessário a construção de um gráfico ou histograma. Outro exemplo é caso que se está tratando: o cálculo do intervalo de confiança para esta variável. Tanto num caso como no outro, será necessário informar ao programa que se necessita guardar os dados decorrentes das replicações para uma futura análise.

O primeiro passo neste sentido é declarar ao programa que os dados que serão empregados no cálculo do intervalo de confiança devem ser armazenados em arquivos. Esta declaração é feita por meio do módulo *Statistics*.

Este módulo permite determinar que o programa calcule diferentes tipos de estatísticas e/ou armazene dados para posterior análise pelo *Output Analyser*. Na sua grande maioria, estas estatísticas já são automaticamente solicitadas quando o modelo está sendo construído, e apresentadas nos relatórios ao final das simulações.

Statistic - Advanced Process					
	Name	Type	Expression	Report Label	Output File
1	Statistic 1	Output	TAVG(Tempo Trans Caixas)	Tempo nos Caixas	TTCx.dat
2	Statistic 2	Output	DAVG(Fila Caixas.Queue)	Clientes Fila Caixas	NMFCx.dat

Figura 6.4: O módulo *Statistics* do painel *Advanced Process*.

Conforme visto anteriormente, o tratamento de sistemas terminais exige a geração de um conjunto de dados a partir da simulação do sistema. Esta amostra com N elementos é obtida a partir das N replicações simuladas.

Ao final destas replicações, o Arena apresenta, no relatório padrão, os resultados das estatísticas para cada uma delas. Como se deseja construir um intervalo de confiança para uma ou mais variáveis de interesse, é preciso “dizer” ao programa que este construa um arquivo com dados desta(s) variável(eis) para posterior análise. É a opção *Output* (veja a coluna sob o título *Type* da figura 6.4) que permite realizar esta tarefa.

Ao utilizar a opção *Output*, informa-se qual a variável e o nome do arquivo destino. O Arena se encarrega do resto. No caso, dois arquivos denominados “NMFCx.DAT” e “TTCx.DAT”, estão sendo criados. Como são 15 replicações, cada arquivo conterá 15 valores. Detalhes da montagem do *Outputs* aparecem na figura 6.5.

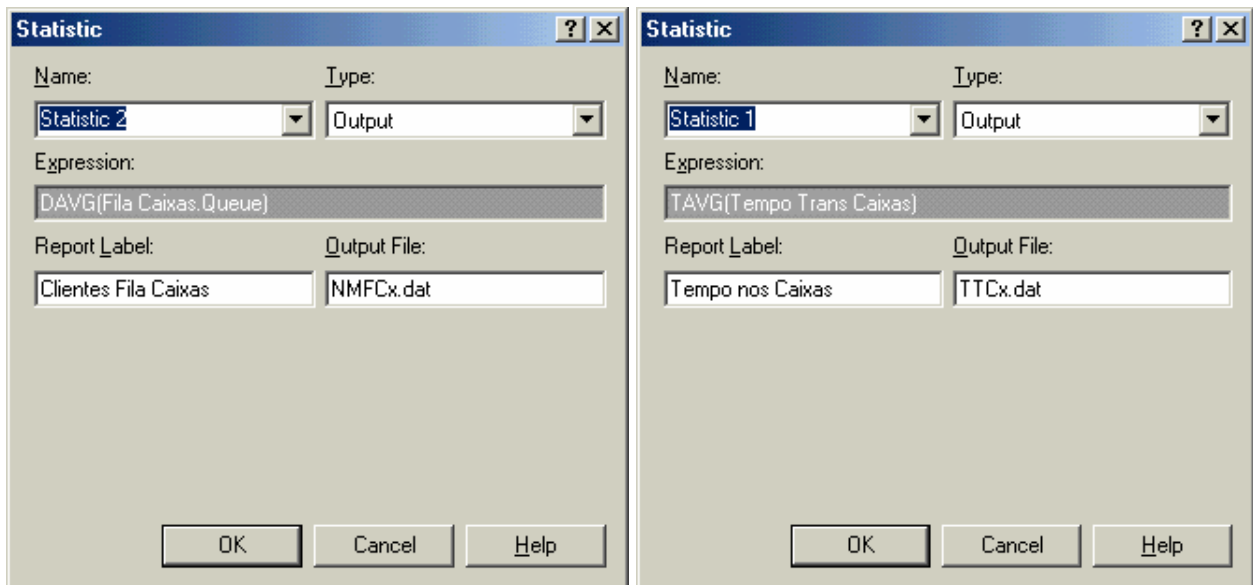


Figura 6.5: Solicitando ao Arena a criação de arquivos de dados.

A extensão DAT é o padrão do Arena para o *Output Analyzer*. Examinando a figura 6.5, o arquivo “NMFCx.DAT” conterá dados referentes a uma estatística cujo nome de referência é: *Fila Caixas.Queue*. O dado que se procura guardar refere-se ao valor médio dos tempos de fila obtidos (*Average Value*). A variável do Arena que guarda este valor é *DAVG(Fila Caixas.Queue)*. Já o arquivo “TTCx.DAT” conterá dados referentes a estatística *Tempo nos Caixas*. Os dados também referem-se aos valores médios obtidos (*Average Value*) cuja expressão, neste caso, é *TAVG(Tempo nos Caixas)*. Cabe aqui uma observação. Até a versão 3.x do Arena, as expressões eram geradas automaticamente pelo software. Infelizmente, a partir da versão 4.0, o usuário terá que conhecer ou buscar tais expressões nos manuais.

Cada um dos valores armazenados refere-se ao valor médio do tamanho da fila dos caixas, obtido em cada uma das replicações (o mesmo valor médio mostrado no *Summary Report* ao final da simulação) e ao tempo médio que os clientes permaneceram no banco (tempo na fila + tempo

de atendimento nos caixas + tempos de deslocamento). A primeira é uma variável dependente do tempo. A estatística final resulta de uma média ponderada que considera os diversos estados assumidos pela variável tamanho da fila e os tempos de permanência em cada um destes estados. A segunda é uma estatística mais simples, obtida pela média simples dos valores observados. A figura 6.6 ilustra a origem e o destino dos dados, tomando como exemplo a estatística tamanho da fila.

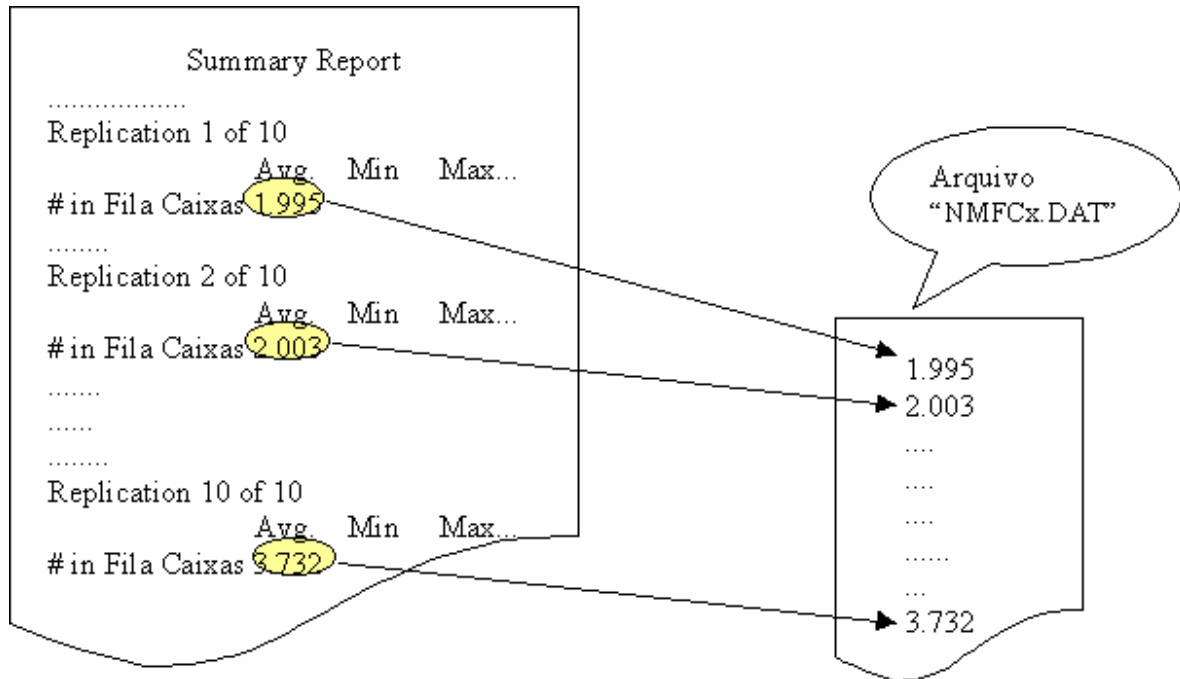


Figura 6.6: Origem das informações geradas e armazenadas no arquivo destino

Do lado esquerdo da figura 6.6 tem-se uma representação, resumida, do *Summary Report*. Para cada uma das réplicas o valor médio (*Average Value*) da Fila Caixas é apresentado. É a partir destes valores que se monta o arquivo com os dados para o posterior cálculo do intervalo de confiança.

Uma vez terminada a simulação do modelo com os dados salvos em arquivos, pode-se prosseguir com o processo de análise invocando o *Arena Output Analyzer*.

6.6.2 Usando o *Arena Output Analyzer*

Assim como no *Arena Input Analyzer*, aqui também se iniciam os procedimentos com a criação de um novo documento. Uma vez criado o documento, arquivos contendo dados para análise devem ser a ele anexados.

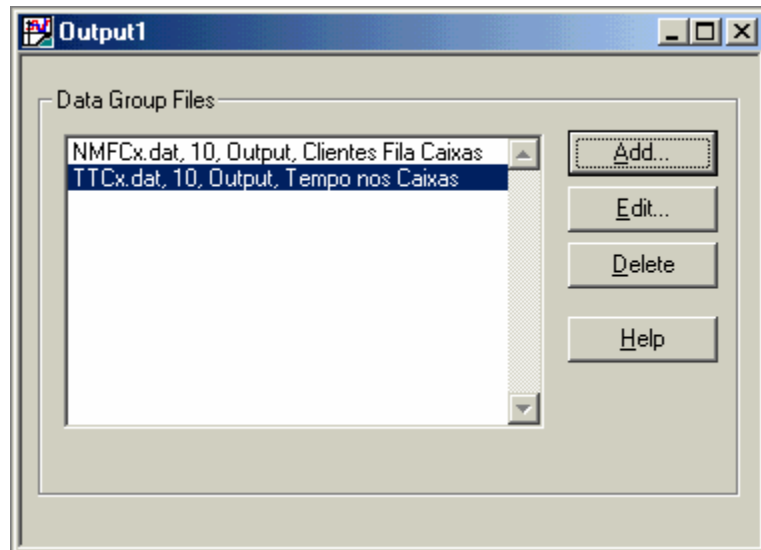


Figura 6.7: Anexando os arquivos NMFCx.DAT e TTCx.DAT ao *Output Analyzer*

A figura 6.7 ilustra o procedimento. Ao clicar o botão *Add*, a janela do *Windows* permite buscar os arquivos de interesse. Na figura vê-se os dois arquivos “NMFCx.DAT” e “TTCx.DAT” em destaque. Uma vez anexados os arquivos, pode-se iniciar as análises.

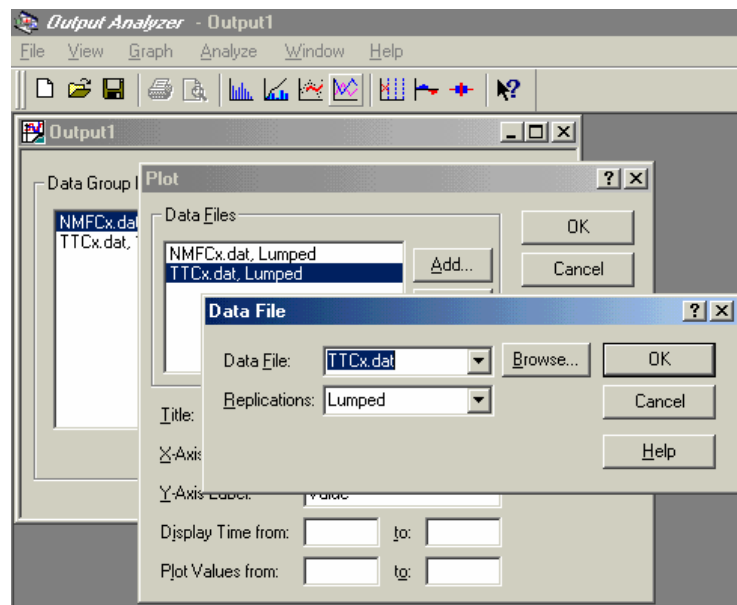


Figura 6.8: Solicitação de gráficos de linhas (*Plot*)

As análises podem iniciar, por exemplo, pela simples visualização gráfica dos dados. O *Output Analyzer* oferece algumas opções: gráfico de barras, histogramas e gráficos de linhas. Todas elas podem ser acessadas via o menu *Graph* ou via os ícones presentes na barra de ferramentas principal. Ao clicar uma das opções, uma janela de diálogo aparece. Seu preenchimento permite informar qual o arquivo de dados que se deseja visualizar (pode-se ter vários arquivos disponíveis ao mesmo tempo), bem como a edição do gráfico a ser visualizado. A

figura 6.8 apresenta a janela de diálogo para a visualização de gráficos de linhas dos arquivos “NFCx.DAT” e “TTCx.DAT”.

Observe na figura 6.8, que o diálogo inclui a determinação de qual ou de quais replicações se deseja apresentar no gráfico. Neste caso, a opção *Lumped* foi escolhida. Esta opção faz com que o programa trate os resultados de todas as replicações em conjunto. Na figura 6.9 é exibido o gráfico desenhado.

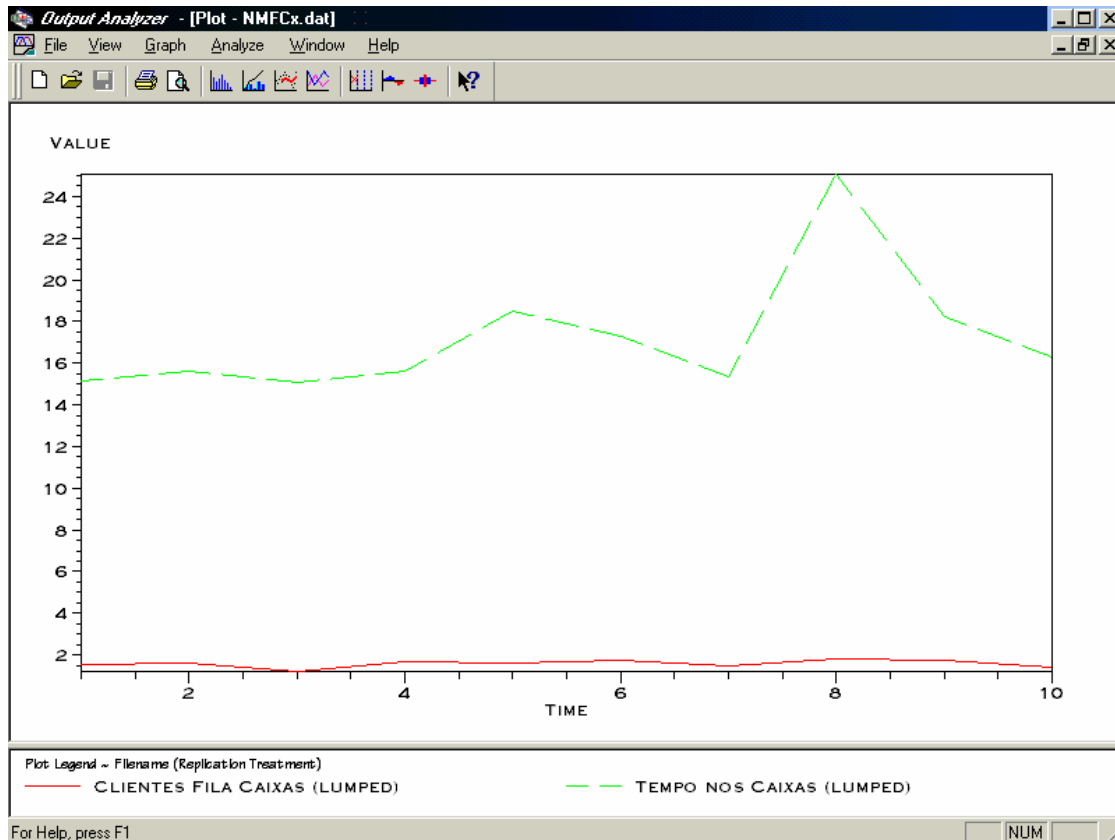


Figura 6.9: Gráficos do tamanho médio da fila dos caixas e do tempo médio no sistema

Uma vez que se tem uma idéia mais clara do comportamento do sistema para as várias replicações realizadas, pode-se passar ao cálculo do intervalo de confiança para as duas variáveis.

O cálculo do intervalo de confiança é tão simples de realizar quanto qualquer outra tarefa. Primeiramente busca-se no menu *Analyse*, a opção *Confidence Interval on Mean*. A partir daí, a opção *Classical*. Outra maneira é apontar e clicar o ícone correspondente. A figura 6.10 apresenta a busca desta função.

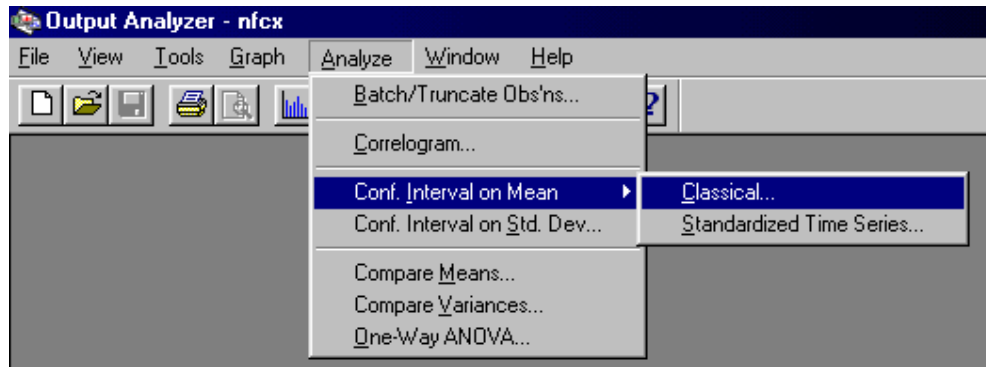


Figura 6.10: Busca da função para o cálculo do intervalo de confiança

Uma vez ativada a função, o diálogo exibido na figura 6.11 aparece. Indicam-se os arquivos contendo os dados, da mesma maneira que se fez para anteriormente para a montagem dos gráficos.

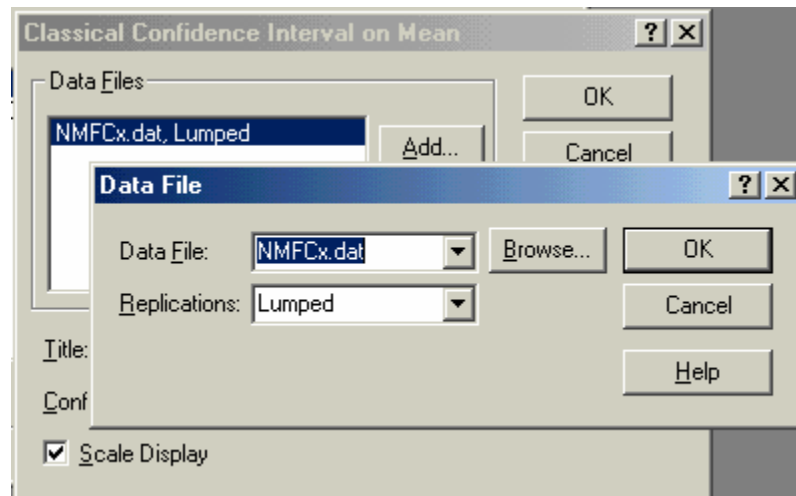


Figura 6.11: Diálogo para o cálculo do intervalo de confiança

A figura 6.11 mostra que o conjunto de dados referentes ao tamanho da fila dos caixas é que será tratado. Observe-se, novamente, a opção *Lumped*, adotada para as replicações. Lembre-se que se deseja que os dados sejam tratados como pertencentes a uma única amostra. Esta opção faz com que o programa entenda que os valores provenientes de cada uma das 10 replicações sejam tratados em conjunto. Outra observação importante é o nível de confiança adotado. No caso, se está utilizando a opção padrão 95% ($Confidence\ Level = 0.95$).

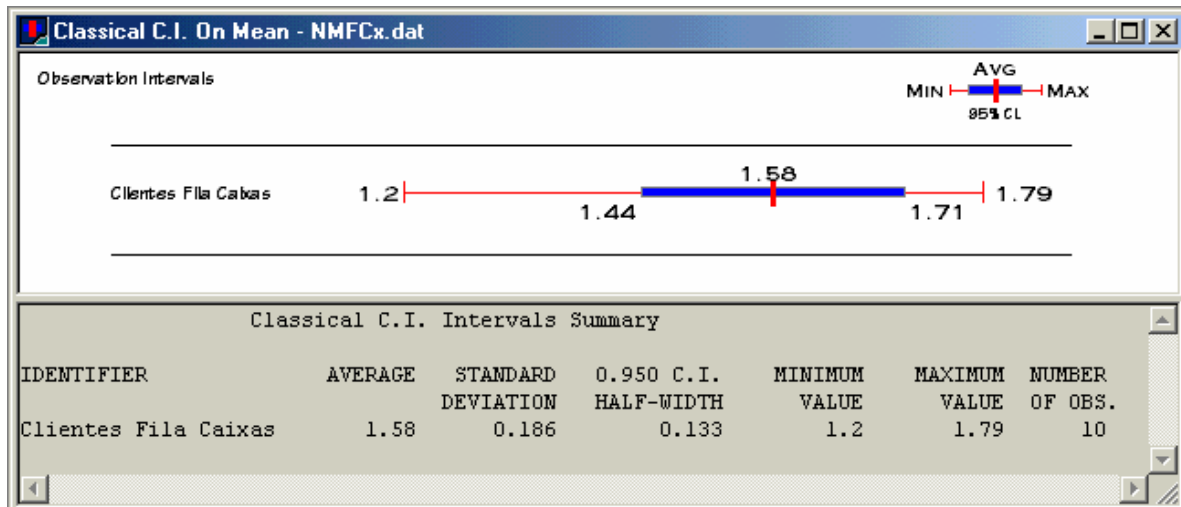


Figura 6.12: Intervalo de confiança para a variável de desempenho – Fila dos Caixas.

A figura 6.12 exibe a forma com que o Arena *Output Analyser* apresenta seus resultados para os intervalos de confiança calculados. No canto superior direito observa-se a legenda. Na parte superior da janela, os resultados na sua forma gráfica. Na parte inferior na forma de texto.

A amostra relativa a primeira variável, clientes na fila dos caixas, contém 10 observações (conforme pode ser visto na coluna mais à direita) e apresenta uma média de 1,58 clientes com um desvio padrão de 0,186. O valor do semi-intervalo de confiança, dado por h (*Half-width*) é de 0,133. Já com relação à segunda amostra de dados (Figura 6.13), que contém os valores dos tempos médios no sistema dos clientes que visitam os caixas, conclui-se a partir dos valores de média e desvio padrão (16,7 e 3,03, respectivamente) que esta é menos estável. Pode-se confirmar isto, comparando-se a variabilidade das duas amostras pelo cálculo de seus coeficientes de variação. Este é dado pela razão entre o desvio-padrão e a média amostral. Obtém-se $0,186/1,58 = 0,12$ para a primeira amostra e $3,10/16,1 = 0,18$ para a segunda.

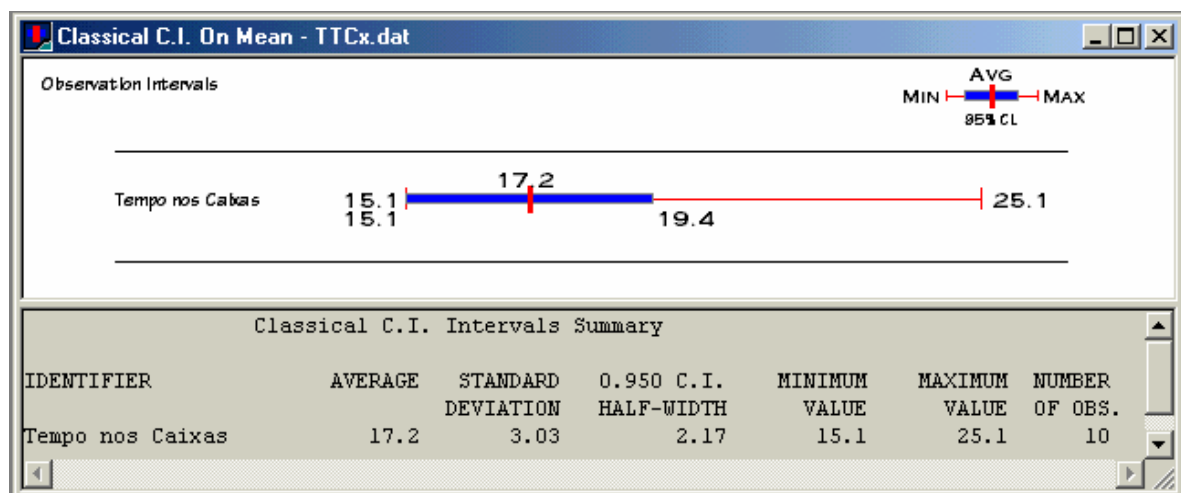


Figura 6.13: Intervalo de confiança para a variável de desempenho – Tempo nos Caixas.

Os números apresentados pelos intervalos de confiança exibidos nas figuras 6.12 e 6.13 permitem ainda que o analista decida se os resultados satisfazem ou não suas expectativas. Buscando outro daqueles “números mágicos”, oriundos não da estatística, mas de quem com ela lida no dia a dia, pode-se dizer que é comum que se busque intervalos de confiança para os quais o valor de h seja, aproximadamente, menor ou igual a 10% da média amostral. Obviamente que este valor dependerá muito do processo que se está lidando, mas como uma primeira meta, este é um bom número.

Buscando uma maior precisão em nos resultados, considere o objetivo de alcançar valores para $h \leq 10\%$ da média amostral. Neste caso, com relação a este exercício, se esta estabelecendo como metas os seguintes valores:

Para a variável Clientes na Fila, $h \leq 0,158$

Para a variável Tempo no Sistema, $h \leq 1,72$

A amostra da variável – Clientes na Fila, apresentou como resultado da simulação piloto um valor para $h = 0,133$. Logo, neste caso, os resultados já satisfazem a meta almejada. Isto é, uma amostra de 10 elementos foi suficiente para garantir a obtenção de um intervalo de confiança aceitável. Já para a variável – Tempo no Sistema, o valor de h é maior do que a meta almejada ($2,17 > 1,72$). Assim é preciso determinar qual o tamanho ideal da amostra. Para tanto, aplica-se a fórmula 6.4.

$$n^* = [n(h/h^*)^2]$$

onde:

n^* = a nova estimativa para n

n = número de replicações já realizadas = 15

h = semi-intervalo de confiança já obtido = 2,17

h^* = semi-intervalo de confiança desejado = 1,72

$$n^* = [n(h/h^*)^2] = [10.(2,17 / 1,72)^2] = 10.(1,26)^2 \cong 16$$

Portanto, a estimativa para que se alcance a meta desejada para h , é que se tenha uma amostra de 16 observações. Para tanto, repete-se o processo executar a simulação, agora com o número de replicações alterado de 10 para 16. A figura 6.14 mostra as respostas da simulação.

As respostas mostram que a meta foi alcançada. O valor obtido de $h = 0,161$ é menor que o desejado (10% de $16,8 = 0,168$). Portanto, mais nenhum aumento no número de dados se faz necessário e o procedimento termina aqui.

Uma pergunta que pode ser colocada é: já que com 16 replicações obteve-se um valor para h menor do que a meta desejada, não seria possível alcançar a meta com menos simulações? É provável. Mas lembre-se, que a metodologia proposta implica na aplicação de um processo de aproximações sucessivas. Nem sempre a aplicação da fórmula 6.4 utilizada nos leva à meta almejada já na primeira tentativa. Muitas vezes acontece o contrário, e tem-se que repetir o procedimento mais de uma vez.

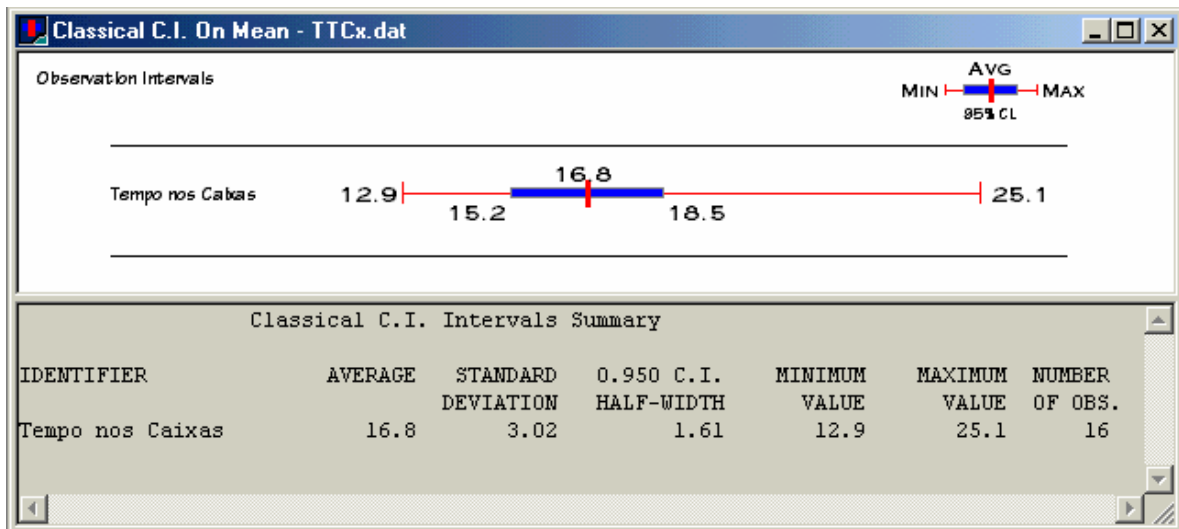


Figura 6.14: Resultado do Intervalo de Confiança com 16 replicações.

Estes são os procedimentos básicos que se deve executar para a realização de análises estatísticas adequadas aos sistemas considerados terminais. No item que segue, o objetivo é mostrar a os procedimentos adequados à condução de experimentos para os sistemas classificados como não-terminais.

6.7 Análise de Sistemas Não-Terminais

A maioria dos sistemas modelados é rotulado como não-terminal. Como já foi dito, diferentemente dos sistemas terminais, não se tem um estado inicial predefinido tem tampouco um evento caracterizando o encerramento do período de simulação. Além disso, dois problemas básicos devem ser contornados. O primeiro deles trata do descarte das observações que pertencem ao período transiente, isto é, aquelas observações que são fortemente influenciadas pelo estado inicial do sistema. O segundo problema diz respeito ao período de simulação. Uma vez que não mais se tem um evento característico de término da simulação, para que se possa avaliar a variância da medida de interesse, considerando os descartes da fase transiente, as técnicas de avaliação exigem longas simulações. Quão longas? Este é o segundo problema a ser tratado. Nos próximos dois itens abordam-se tais problemas.

6.7.1 Remoção da Fase Transiente

No que se refere a sistemas não-terminais, o interesse das análises recai sobre o desempenho do sistema quando o mesmo se encontra em regime. Desta forma, os resultados iniciais das simulações, referentes ao estado transiente do sistema, não devem ser incluídos na computação das estatísticas finais. O problema, neste caso, é determinar o final desta fase e é conhecido pelo nome de *remoção da fase transiente*.

A maior dificuldade na remoção da fase transiente é a impossibilidade de determinar o seu término com muita clareza. Todos os métodos associados à remoção dos dados deste período são, por consequência, heurísticos. Alguns destes métodos são apresentados a seguir:

1. Longa simulação;
2. Inicialização apropriada;
3. Truncagem;
4. Observação visual.

Longa Simulação

O método mais simples consiste na realização de uma longa simulação, isto é, tão longa que assegure que as condições iniciais de instabilidade não afetem os resultados. A vantagem deste método consiste em sua simplicidade. No entanto, ele apresenta pelo menos duas desvantagens. A primeira desvantagem é o desperdício de recursos. Se os recursos são caros, as simulações não devem ser mais longas do que o absolutamente necessário. Mesmo que estes não sejam escassos e a que geração de novas observações não consuma muitos recursos, é difícil assegurar que a rodada simulada é suficientemente longa. Esta é a segunda desvantagem. Por tais razões este método, por si só não é recomendado.

Inicialização Apropriada

Este método requer a inicialização das variáveis do sistema em um estado próximo das condições de estabilidade. Por exemplo, as condições iniciais das filas do sistema não mais serão vazias e os recursos não mais estarão inicialmente disponíveis. O número de entidades nas filas poderá ser fixado considerando-se simulações pilotos ou análises anteriores do sistema. Embora o método contribua para uma substancial redução do período de simulação, a determinação das condições iniciais nem sempre é uma tarefa fácil de ser realizada.

Truncagem

Este método pressupõe que a variabilidade durante o período transiente é maior do que quando em regime. Geralmente tal hipótese é aceitável. A variabilidade é testada de acordo com o seu intervalo de variação, isto é, os valores mínimo e máximo observados. Desenhando-se os valores em um gráfico verifica-se que, após o período transiente, estes começam a se estabilizar. No método da truncagem, descarta-se então os valores iniciais que claramente encontrem-se fora dos limites da estabilização.

Desta forma, dada uma amostra de n observações $\{x_1, x_2, \dots, x_n\}$, simplificada o método consiste em ignorar as primeiras l observações e então verificar os valores de mínimo e de máximo das demais $n - l$ observações. Este passo é repetido para $l = 1, 2, \dots, n - 1$, até que a observação $(l + 1)$ não seja nem mínimo nem máximo das demais observações. Considere o exemplo a seguir.

Exemplo 6.5 Veja a seguinte seqüência de observações: 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 10, 9, 10, 11, 10, 9, 10, 11, 10, 9, ...

Ignorando a primeira observação ($l = 1$), os limites das demais observações são [2; 11]. Uma vez que a segunda observação é igual ao mínimo, a fase transiente é maior do que 1. Ignorando as primeiras duas observações ($l = 2$), os limites das demais observações são agora [3; 11]. Desta vez a terceira observação é igual ao mínimo. A truncagem continua com $l = 3$ e assim por diante.

Finalmente, com $l = 9$, os limites das demais observações são [9; 11] e a décima observação (10) não é nem o mínimo nem o máximo. Desta forma o período transiente é igual a 9 e as primeiras nove observações são então descartadas. O gráfico 6.1 exibe os valores amostrados. Embora de fácil implementação, nem sempre os resultados obtidos são corretos.

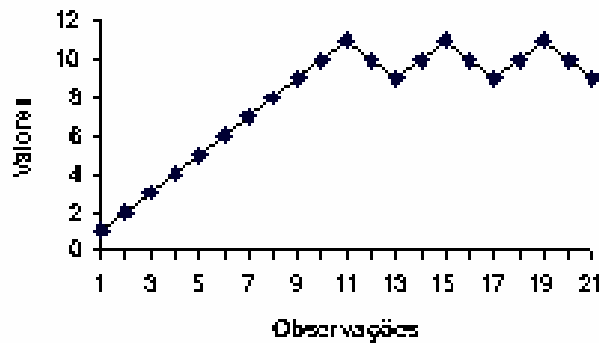


Gráfico 6.1: Valores das observações

Observação Visual

O método mais simples e prático para a determinação do ponto de término do período transiente é o da observação visual. A partir da construção de um gráfico, o qual aponta o comportamento da variável de resposta ao longo do tempo, procura-se observar, de forma aproximada, em que momento as respostas passam a ter uma conduta mais estabilizada. Em alguns sistemas a flutuação das respostas é bastante acentuada, mesmo quando em regime, dificultando à observação. Neste caso, traçar gráficos com médias móveis da variável é aconselhável.

A média móvel é construída calculando-se a média aritmética das k mais recentes observações em cada ponto do conjunto de dados. O valor de k (tamanho da média móvel) é selecionado pelo analista, e poderá variar de acordo com o comportamento da variável sob investigação. Na medida em que se aumenta o valor de k , suaviza-se o gráfico tornando mais clara a observação do ponto de truncagem. Para $k=1$, os gráficos da média móvel e dos dados brutos coincidem. Os gráficos 6.2, 6.3 e 6.4 apresentam médias móveis para $k=1$, $k=10$ e $k=50$, respectivamente.

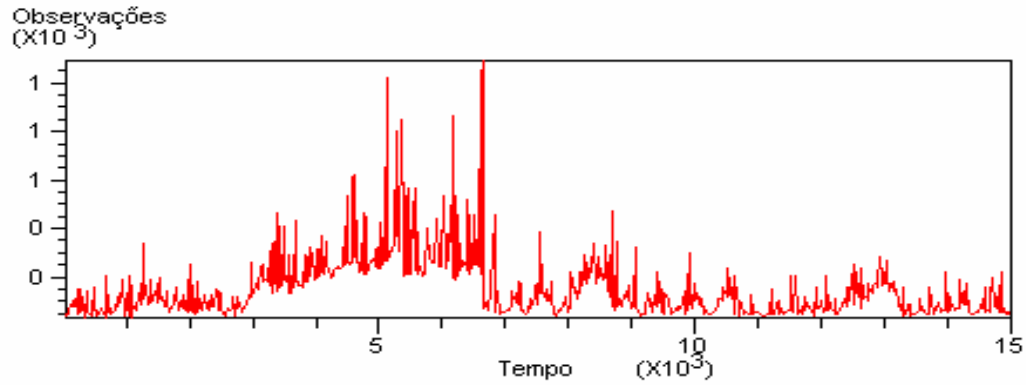


Gráfico 6.2: Média móvel com $k=1$

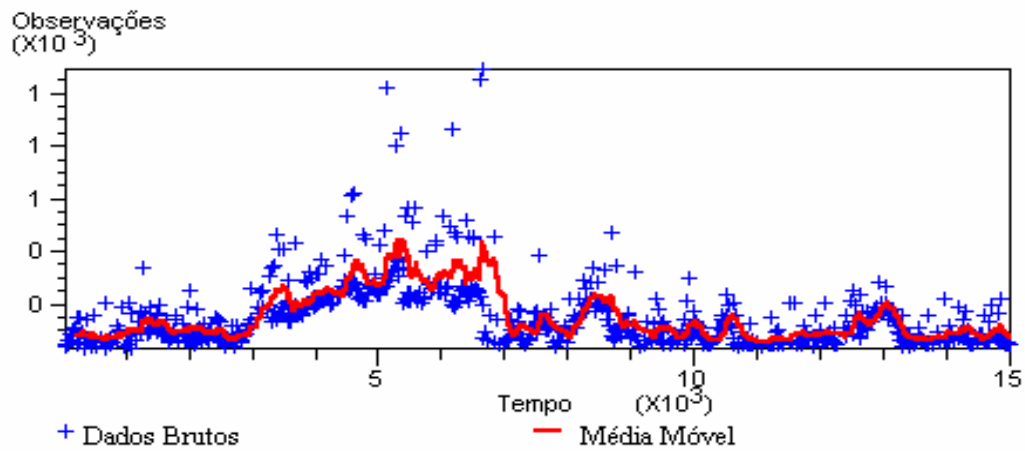


Gráfico 6.3: Média móvel com $k=10$

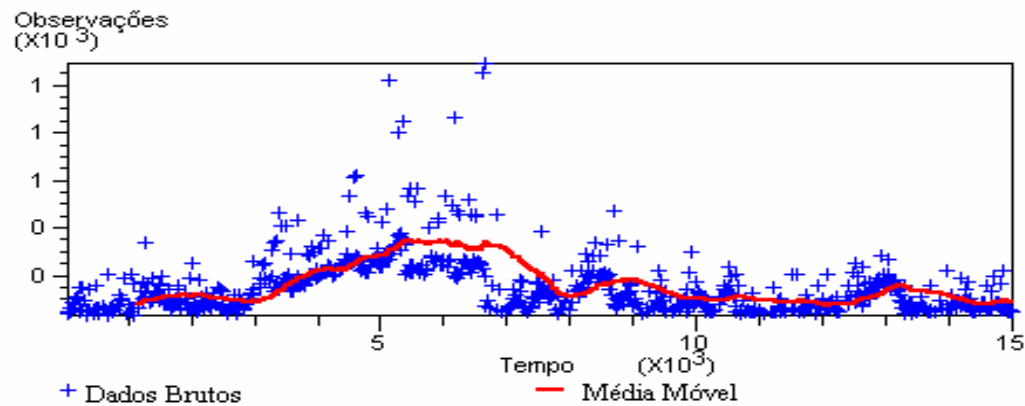


Gráfico 6.4: Média móvel com $k=50$

6.7.2 Determinação do Tamanho do Período de Simulação

Uma vez estabelecida a extensão do período transiente, o próximo problema a ser tratado é o da determinação do período de simulação, propriamente dito. Nos sistemas terminais,

anteriormente discutidos, estimava-se a variância da média da medida de desempenho sob observação, gerando uma amostra cujas observações eram constituídas pelos valores médios da variável de interesse, obtidos a partir das diversas replicações realizadas. Nestas replicações, o estado inicial do sistema (definido pelas variáveis de estado) permanece sempre o mesmo (geralmente vazio e disponível), alterando-se apenas as sementes dos geradores de números aleatórios. Desta forma, garante-se que os valores gerados por cada uma das replicações sejam independentes, exigência para que se aplique as formulas 6.2 e 6.3 apresentadas no item 6.5.1 A mesma abordagem pode ser aplicada aos sistemas não-terminais.

Nos sistemas não-terminais, embora seja possível a aplicação deste mesmo método, será visto que, em função da necessidade de descarte das observações iniciais, da natureza destes sistemas e de outras considerações, as replicações serão bem mais longas do que aquelas aplicadas em sistemas terminais.

Além do método das replicações, dois outros, que se aplicam a estes sistemas, serão apresentados: o método da regeneração e o método do loteamento.

Replicações Independentes

A semelhança do que foi visto para sistema terminais, o método das replicações independentes consiste em conduzir m replicações de tamanho $n + n_0$ observações, onde n_0 é o tamanho do período transiente. As primeiras n_0 observações de cada replicação devem ser descartadas. Os seguintes passos devem ser seguidos:

1. Calcular a média de cada replicação:

$$\bar{x}_i = \frac{1}{n} \sum_{j=n_0+1}^{n_0+n} x_{ij}, \quad i = 1, 2, \dots, m \quad (6.5)$$

2. Calcular a média geral de todas as replicações (a média das médias):

$$\hat{x} = \frac{1}{m} \sum_{i=1}^m \bar{x}_i \quad (6.6)$$

3. Calcular a variância das médias das replicações:

$$Var(\bar{x}) = \frac{1}{m-1} \sum_{i=1}^m (\bar{x}_i - \hat{x})^2 \quad (6.7)$$

O intervalo de confiança para a média da variável de desempenho é então obtido por:

$$\left[\hat{x} \pm z_{1-\alpha/2} Var(\bar{x}) \right] \quad (6.8)$$

Neste método e nos seguintes, apresentados nos itens 6.6.2.2 e 6.6.2.3, caso as amostras sejam pequenas ($m < 30$), é recomendado a substituição de $Z_{1-\alpha/2}$ por $t_{(1-\alpha/2; m-1)}$.

Para o caso dos sistemas não-terminais, alguns problemas se contrapõem ao uso das replicações para a geração da amostra e conseqüente cálculo do intervalo de confiança. O primeiro problema diz respeito ao critério de parada. Lembre-se que nestes sistemas não existem eventos típicos que permitam definir um momento de término da simulação. Ao contrário, longas simulações são exigidas para representar o comportamento dos sistemas. Portanto, no caso do emprego do método das replicações, existe a necessidade de se realizar longas ($n + n_0$) e inúmeras (m) replicações o que poderá inviabilizar os estudos. Outro problema é que cada replicação apresentará um conjunto de observações iniciais (n_0) as quais pertencerão ao estado transiente do sistema e que, portanto, devem ser descartadas. Se a fase transiente for significativa (muito longa), a quantidade de dados gerados e descartados nas diversas replicações será muito grande implicando em grande desperdício ($m \times n_0$ observações).

Veja que o intervalo de confiança é inversamente proporcional a $\sqrt{m \cdot n}$. Desta maneira, a obtenção de um intervalo de confiança mais estreito poderá ser conseguido tanto pelo aumento de m como de n . No entanto, para minimizar o desperdício, uma redução de m e um aumento de n é sempre preferível. Um número razoável de replicações, para a maioria dos problemas é 10. Não mais do que 20 devem ser realizadas uma vez que acima deste valor, as melhorias apresentadas pela estatística t são mínimas. Por estas razões, o uso desta técnica torna-se viável somente para aqueles sistemas cujo período transiente é bastante curto.

Regeneração

Para melhor compreender o conceito do método da regeneração, considere o exemplo de uma fila diante de um terminal de atendimento eletrônico de um banco que opera entre 8:00 horas e 20:00 horas. Uma vez que, inicialmente, a fila do terminal se encontre vazia ao início das operações diárias, é possível que o comportamento da variável *Tamanho da Fila* ao longo do período de operação do terminal, possa ser parecida como o gráfico 6.5.

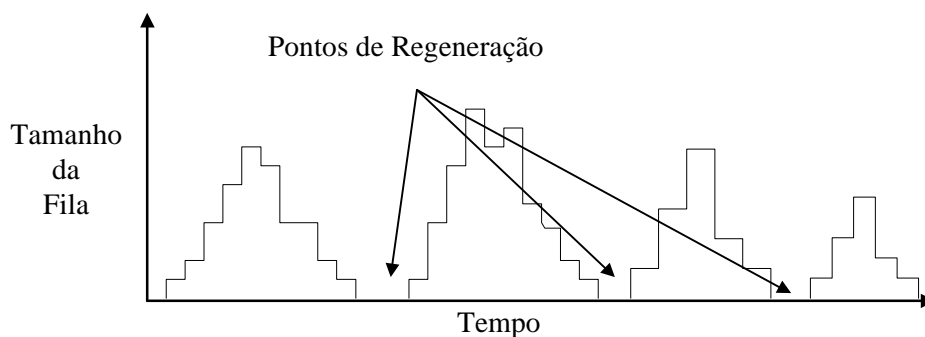


Gráfico 6.5: Comportamento do tamanho da fila ao longo do tempo.

Observa-se no gráfico que, ao longo da operação deste sistema, a variável sob investigação retorna, por diversas vezes, ao seu estado inicial. A trajetória da variável após a

volta a este estado, não mais depende de seu comportamento anterior. Sabe-se que o tempo de espera na fila do terminal por um cliente depende da demanda deste recurso por parte de clientes anteriores. Entretanto, o tempo no sistema de um cliente que chega para ser servido e encontra o sistema sem filas, não mais depende do ocorrido com os clientes que anteriormente tenham passado pelo sistema. O que ocorre na prática é que, após estes períodos em que o sistema encontra-se novamente vazio, inicia-se uma nova fase no sistema. Em outras palavras o sistema renasce de forma totalmente independente de sua trajetória ou comportamento anterior. Este fenômeno é conhecido por *regeneração do sistema*. Os momentos em que o sistema inicia uma nova fase são chamados de *pontos de regeneração*. O tempo decorrido entre dois pontos sucessivos de regeneração é chamado ciclo de regeneração. No gráfico 6.5 vê-se que a trajetória da variável passa por três destes ciclos, os quais são independentes. Desta forma, a média do tamanho da fila no segundo ciclo é independente da média do primeiro ciclo. O argumento da independência é válido também para outras variáveis.

Um sistema regenerativo, isto é, com ciclos de regeneração, pode ser analisado usando-se o método da regeneração. Nem todos os sistemas são regenerativos. Por exemplo, um sistema com duas filas estaria regenerado, quando ambas as filas estivessem vazias. Na medida do aumento do número de filas os pontos de regeneração tornam-se raros e os ciclos cada vez mais longos. Considera-se que sistemas com uma *memória* regenerativa muito longa tornam-se não-regenerativos. Computar a variância destes sistemas usando os ciclos de regeneração é um mais complexo que fazer uso dos métodos anteriores (loteamento e replicações independentes).

Loteamento

O método do loteamento consiste da realização de uma longa simulação, seguido do descarte das observações iniciais e da divisão das observações restantes numa série de lotes ou de subamostras. Dada uma longa simulação de $N + n_0$ observações, n_0 é o número de observações que pertencem ao período transiente e são descartadas. As N observações remanescentes são divididas em $m = [N/n]$ lotes de n observações cada uma.

Para a definição do intervalo de confiança os seguintes passos devem ser seguidos:

1. Calcular a média de cada replicação:

$$\bar{x}_i = \frac{1}{n} \sum_{j=1}^n x_{ij}, \quad i = 1, 2, \dots, m \quad (6.9)$$

2. Calcular a média geral de todas as replicações (a média das médias):

$$\hat{x} = \frac{1}{m} \sum_{i=1}^m \bar{x}_i \quad (6.10)$$

3. Calcular a variância das médias das replicações:

$$Var(\bar{x}) = \frac{1}{m-1} \sum_{i=1}^m (\bar{x}_i - \hat{x})^2 \quad (6.11)$$

O intervalo de confiança para a média da variável de desempenho é então obtido por:

$$\left[\hat{x} \pm z_{1-\alpha/2} \sqrt{Var(\bar{x})} \right] \quad (6.12)$$

Observa-se que os cálculos realizados seguem os mesmos procedimentos do método das replicações independentes. Todavia, no método do loteamento o desperdício de dados observados é bem menor, uma vez que apenas um único descarte de n_0 observações é realizado. Novamente o tamanho do intervalo de confiança é inversamente proporcional a $\sqrt{m.n}$. Sendo assim, tanto o aumento de m como de n permitem a redução do intervalo. O tamanho n do lote deverá ser suficientemente longo (isto é, com um número razoável de elementos) de tal forma que a correlação entre as médias de cada lote seja muito pequena. Isto permite a aplicação da formulação acima a qual pressupõe a independência entre as médias de cada lote. Se os lotes forem suficientemente longos, as médias de dois lotes adjacentes serão aproximadamente independentes, mesmo que as observações do final do lote j estejam correlacionadas àquelas do início do lote $j + 1$. A idéia é que o tamanho do lote seja maior do que o mínimo intervalo (k) entre duas observações correlacionadas. Desta forma, para garantir esta independência entre os valores médios dos lotes é necessário definir este intervalo mínimo. Uma forma de encontrar o tamanho correto para n é calcular a covariância entre as médias de lotes sucessivos. O valor da covariância é dado por:

$$Cov(\bar{x}_i, \bar{x}_{i+1}) = \frac{1}{m-2} \sum_{i=1}^{m-1} (\bar{x}_i - \hat{x})(\bar{x}_{i+1} - \hat{x}) \quad (6.13)$$

Em geral, quanto mais congestionado for o sistema, maior deverá ser o tamanho do lote devido à alta correlação entre as os tempos de processos de entidades sucessivas que disputem um recurso comum. Por exemplo, o tempo que um cliente c_i permanece num sistema de auto-atendimento em um banco (tempo de fila + tempo de serviço) recebe influência direta daqueles clientes que já se encontrem na fila (clientes c_{i+1} , c_{i+2} ,...) do recurso (caixa automático) que ele está buscando. O cálculo da covariância é realizado entre duas variáveis \bar{x}_i e \bar{x}_{i+1} , que são os valores médios de dois lotes adjacentes. Quanto maior o lote menor a correlação entre \bar{x}_i e \bar{x}_{i+1} . Uma alternativa para a definição de n envolve um procedimento iterativo no qual define-se um tamanho de lote, calcula-se as médias de cada um deles e a correlação entre estes valores. Uma alternativa é iniciar com um pequeno valor e dobrá-lo a cada interação.

A tabela 6.2 demonstra a relação entre o tamanho de cada lote, a covariância e a variância das médias, obtidos a partir da modelagem e simulação de um sistema de rede de computadores. Após uma simulação piloto e o devido descarte das observações da fase transiente, a definição do tamanho do lote foi realizada por meio de um procedimento iterativo.

Tamanho do Lote	Covariância entre \bar{x}_i e \bar{x}_{i+1}	Variância
-----------------	---	-----------

1	-0,18792	1,79989
2	0,02643	0,81173
4	0,11024	0,42003
8	0,08979	0,26437
16	0,04001	0,17650
32	0,01108	0,10833
64	0,00010	0,06066
128	-0,00378	0,02992
256	0,00027	0,01133
512	0,00069	0,00503
1024	0,00078	0,00202

Tabela 6.2: Relação entre Covariância e Tamanho do Lote

Inicialmente, cada observação correspondeu a um lote ($n=1$). O tamanho do lote foi então sucessivamente dobrado. Observa-se na tabela que a variância da média de cada lote vai reduzindo-se na medida do aumento de n . Da mesma forma, a correlação entre as médias de dois lotes sucessivos vai, paulatinamente, abreviando-se. A covariância tornou-se menor do que 1% do lote original quando o tamanho de n é igual ou maior do que 64. Portanto, o tamanho mínimo de lote para este sistema deverá ser 64 observações.

Como se vê, a análise de sistemas não-terminais pressupõe duas etapas principais: a remoção das observações da fase transiente e a definição do período de simulação do sistema quando em regime, permitindo ao analista a realização das observações necessárias à estimação dos parâmetros das variáveis de controles. A figura 6.15 apresenta um resumo geral da metodologia a ser adotada no tratamento de sistemas não-terminais.

No item a seguir aplica-se o procedimento apresentado e resumido na figura 6.15. Um exemplo, fazendo uso de um sistema simples será adotado, de forma que a condução de cada uma das etapas seja mais bem compreendida. Todos os passos serão realizados com o auxílio do *Arena Output Analyzer*.

6.8 Usando o Arena Output Analyzer para tratar Sistemas Não-terminais

No exemplo tratado, será feito uso do método da observação visual para realizar o descarte das observações iniciais.

Exemplo 6.7: Para à análise de um sistema não-terminal, considera-se o sistema hipotético representado na figura 6.16.

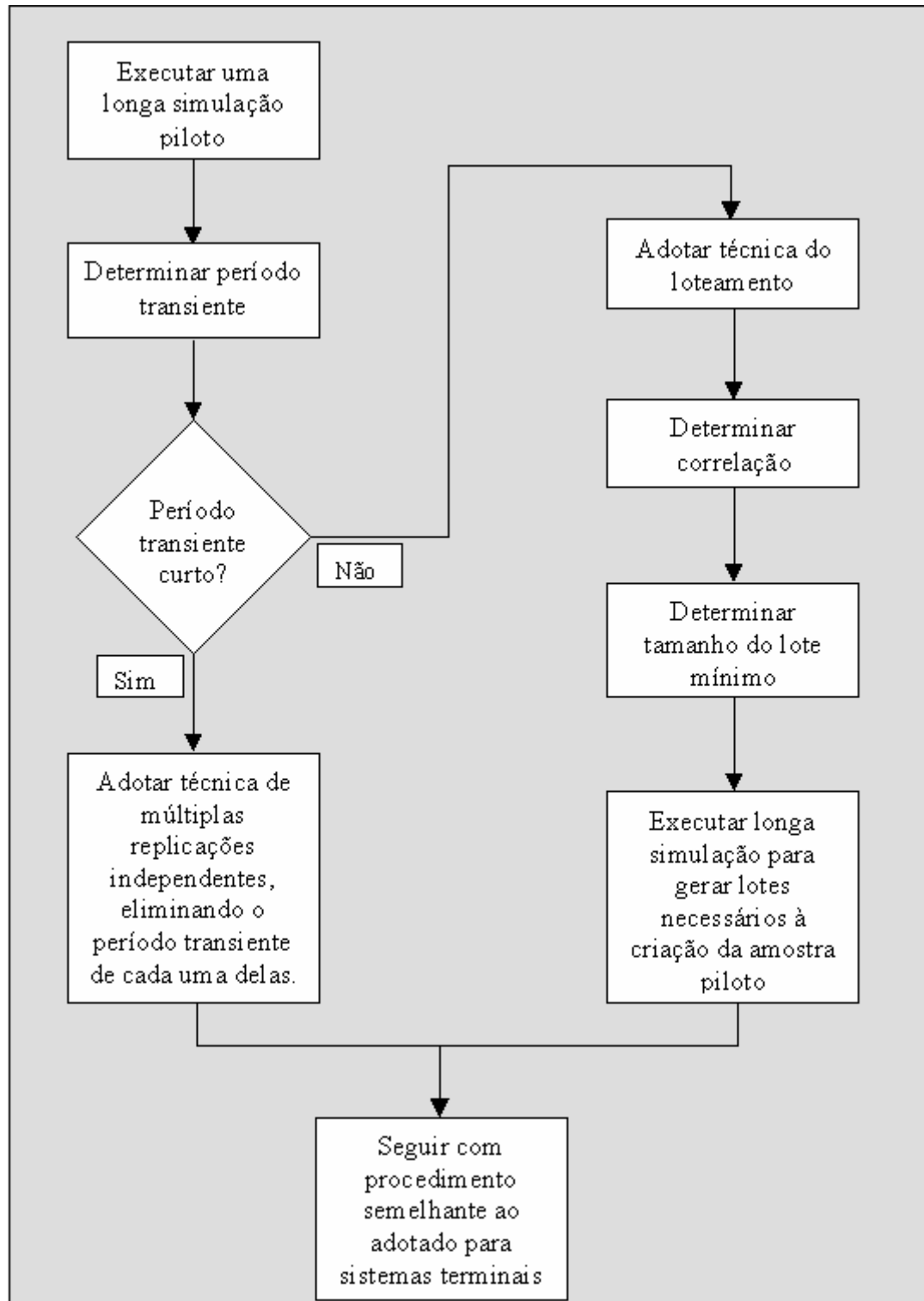


Figura 6.15: Fluxo do processo metodológico para análise de sistemas não-terminais.

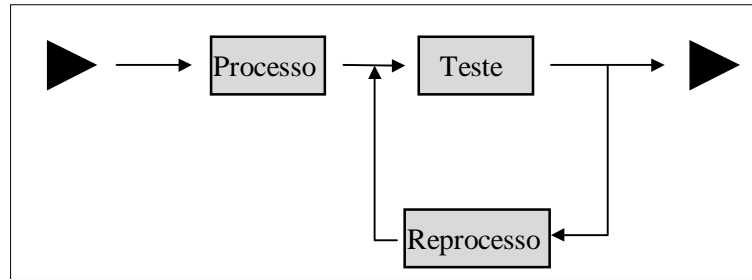


Figura 6.16: Representação esquemática do sistema de processo e inspeção

Neste sistema, entidades (peças) chegam para realizar um processo em uma estação de trabalho. Nesta estação, o recurso possui capacidade de servir até duas entidades ao mesmo tempo. Após realizarem este processo, as peças seguem para um setor de inspeção onde são testadas. Os tempos de serviços em cada uma das estações seguem as distribuições apresentadas na tabela 6.3.

Serviço	Tempo (min.)
Processo	Normal (34, 9)
Teste de Inspeção	Uniforme (10, 18)
Reprocessamento	Triangular (25, 40, 45)
Tempo entre Chegadas	Exponencial (20)

Tabela 6.3: Tempos dos Serviços e Tempo entre Chegadas no sistema.

O tempo de deslocamento entre as estações é de 5 minutos. Os resultados dos testes na área de inspeção definem o caminho que deve ser tomado pelas entidades. Se forem aprovadas nos testes (80% dos casos) as peças deixam o sistema. Se forem reprovadas devem seguir para um segundo servidor onde são reprocessadas. Uma vez servidas pela estação de reprocessamento, retornam à estação de testes. Este sistema de processamento e inspeção é totalmente automatizado funcionando cinco dias por semana em dois turnos de 8 horas (das 8:00 às 24:00 horas). Ao final de cada dia, todos os processos que se encontrarem no sistema e que por ventura não puderem ser atendidos, ou por encerramento do turno diário (às 24:00 horas) ou por encerramento da semana, são retomados no reinício dos trabalhos (às 8:00 do dia seguinte ou às 8:00 da segunda-feira).

Existe interesse por parte da gerência do sistema em uma análise do *tempo médio de passagem* (TMP – *Lead Time*) das entidades por este seguimento do processo produtivo. Desta forma, o sistema foi modelado e deve ser simulado. A variável de desempenho será o TMP das entidades pelo sistema. As análises devem indicar os procedimentos experimentais de tal forma que os resultados obtidos sejam estatisticamente garantidos. Esta garantia é definida com o estabelecimento de uma confiança desejada para os resultados da variável de desempenho. Desta forma, se quer determinar: qual o período transiente e qual o período a ser simulado, de tal maneira que se possa garantir que, em 95% dos casos, a estimativa obtida para o TMP encontre-se em um intervalo cujos limites não sejam superiores a 10% da média obtida para aquela variável.

O primeiro procedimento será a simulação do sistema por um período relativamente longo. Esta replicação *piloto* será usada para que se tenha uma visão mais clara do comportamento da variável de interesse. Servirá também para uma primeira observação do período transiente. O tempo de replicação escolhido foi de 20.000 min. Este tempo reflete um período de aproximadamente um mês de funcionamento deste sistema. O modelo deste sistema foi denominado “Exemplo65.doe”, e encontra-se no CD que acompanha este livro.

Os dados necessários às análises foram obtidos pela incorporação de um módulo *Statistics*. Neste módulo, foi solicitado que o tempo de passagem de cada uma das peças processadas pelo sistema fosse armazenado em um arquivo chamado “TMP.DAT”. A figura 6.17 ilustra o módulo editado.

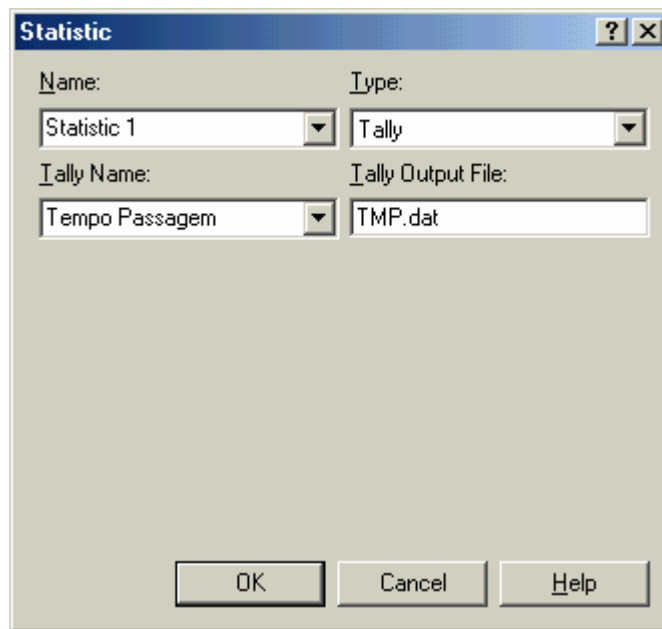


Figura 6.17: O Módulo *Statistics* e a edição da estatística *Tally*.

Os resultados apresentados pela variável TMP para esta primeira replicação são apresentados no gráfico 6.6. Observa-se pelo gráfico, que a variável de desempenho apresenta uma certa variabilidade. Como todo sistema envolvendo filas ou redes de filas (como no caso) e frequências de chegadas obedecendo a uma distribuição exponencial, é de se esperar um comportamento cíclico para o tempo de passagem pelo sistema. Observa-se que no início da simulação, os valores de TMP são pequenos. Isto ocorre porque as primeiras entidades encontram o sistema vazio e disponível. No entanto, a partir dos pontos próximos do tempo 1000 minutos, o valor de TMP começa a crescer. Ocorre então um congestionamento que só ira se acomodar, com o sistema entrando em regime próximo do tempo 5000 minutos.

Embora neste caso o gráfico seja bem sugestivo quanto ao comportamento do sistema, nem sempre a determinação sobre quando o sistema entra em regime, com base visual, é tarefa das mais simples. No entanto, acredite, outros métodos demandam um esforço ainda maior e, muitas vezes, não se obtém um benefício equivalente. É possível que uma estimativa mais otimista seja em torno de 2000 minutos. Outros (mais pessimistas ou que preferem mais

segurança) poderiam sugerir que o período logo após o primeiro grande congestionamento (em torno de 5000 minutos) seria o melhor momento.

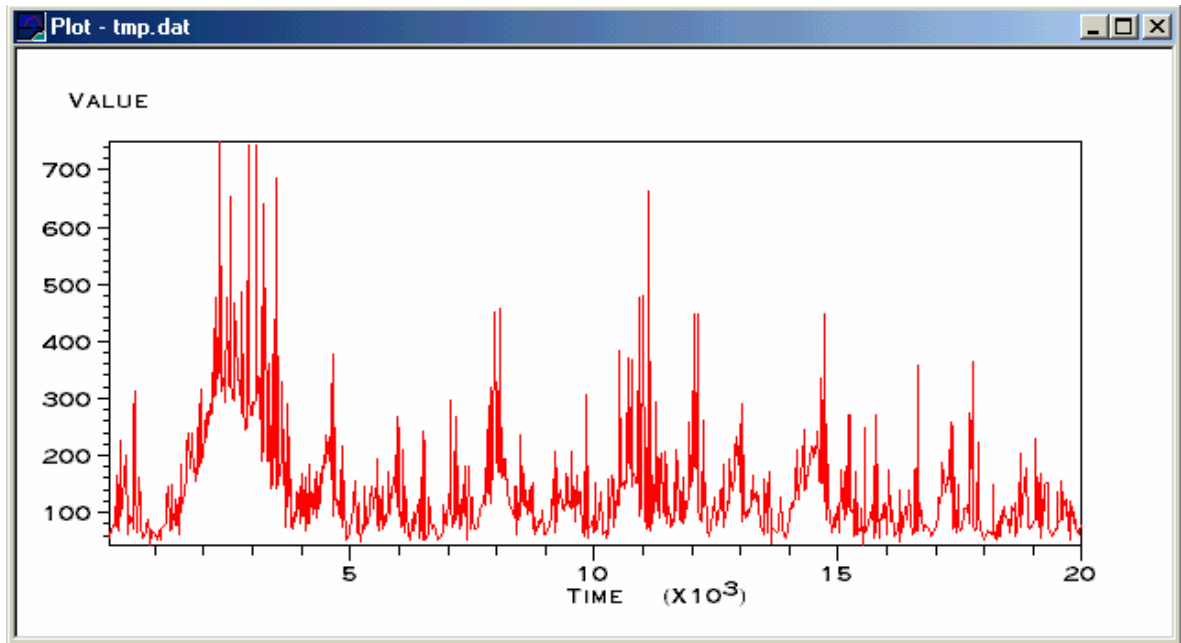


Gráfico 6.6: Tempo médio de passagem das entidades pelo sistema

Para melhor visualizar o ponto em que o sistema entra em regime, pode-se utilizar técnicas de “suavização” de curvas. Uma delas é a chamada média móvel. Para empregá-la usando o *Output Analyzer*, basta chamar a função junto ao *Menu/Graph/Moving Average*. Uma vez que a janela de diálogo esteja aberta, basta indicar o arquivo contendo os dados, e o parâmetro da média móvel desejada. No caso do exemplo das figuras 6.18 e do gráfico correspondente 6.7, a média móvel solicitada foi de 10 pontos.

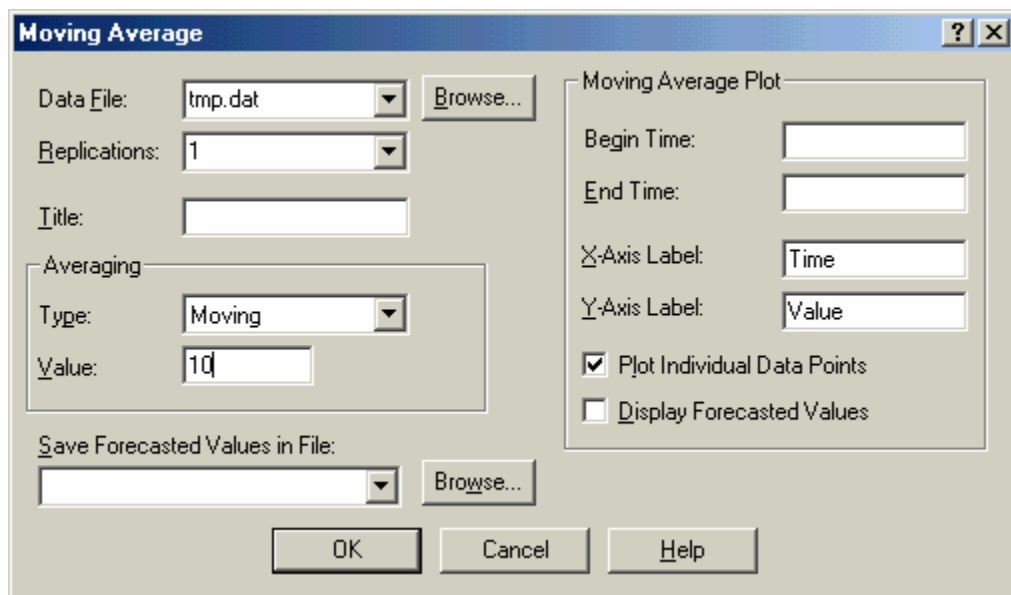


Figura 6.18: Solicitação da média móvel

Desta forma, quando da realização da simulação completa, as observações geradas neste período (de zero a 5.000 minutos) deverão ser descartadas.

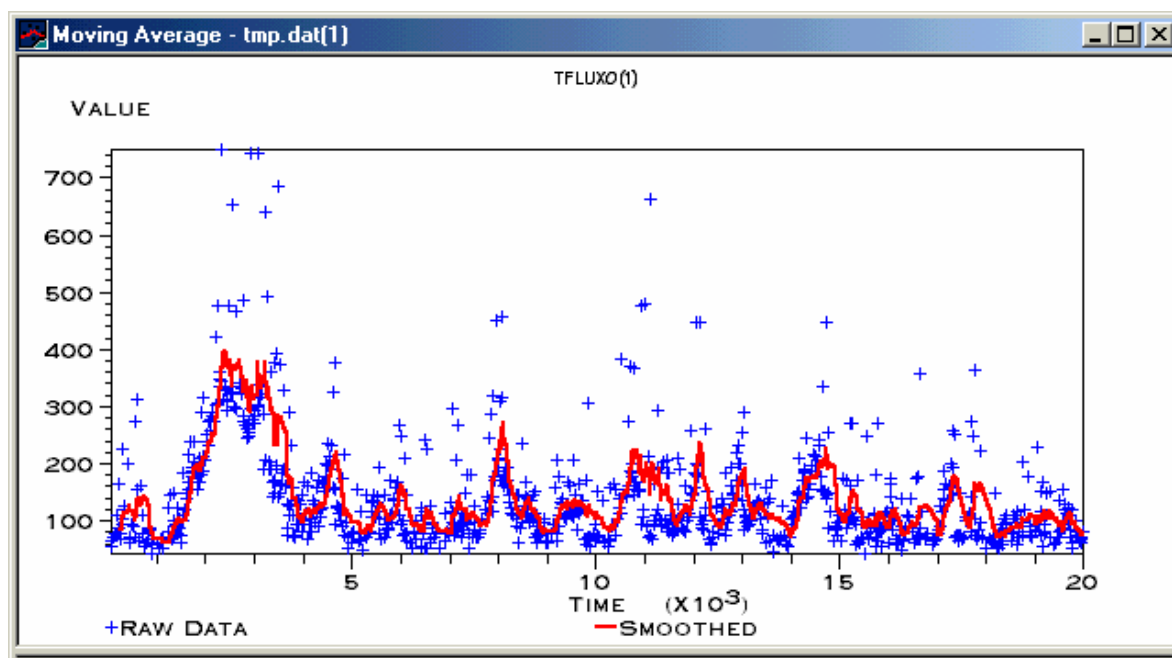


Gráfico 6.7: Média móvel de parâmetro 10.

Uma vez definido o período transiente, se passa à segunda fase dos procedimentos de análise apresentados no fluxograma da figura 6.15, isto é, determinar o tipo de técnica que será adotada: múltiplas replicações independentes ou loteamento? Como o objetivo do texto é, antes de tudo, didático, aplicam-se as duas técnicas, comparando-se seus resultados e comentando-se suas diferenças.

Aplicação do Método do Loteamento

Recordando a aplicação do método do loteamento, após o descarte das n_0 observações pertencentes ao período transiente, as N observações remanescentes são divididas em $m = [N/n]$ lotes de n observações cada uma. Como foi visto, o tamanho do intervalo de confiança é inversamente proporcional a $\sqrt{m \cdot n}$. Desta forma, o tamanho de cada lote deverá ser suficiente para que a correlação entre as médias de cada lote seja pequena. Seguindo-se os procedimentos sugeridos, deve-se calcular a autocovariância (isto é, a covariância entre elementos pertencentes a mesma amostra) entre os diversos lotes, aumentando-se paulatinamente os valores de n .

Para que se possa determinar a covariância deve-se, inicialmente, realizar o descarte das observações pertencentes a fase considerada transiente. Duas formas podem ser adotadas. Na primeira, programa-se o modelo de forma que este acumule estatísticas somente a partir do tempo 5.000 minutos. Para tanto, basta definir no menu *Run/Setup/Replication Parameters* um período de *Warm-up*. A figura 6.19 ilustra este procedimento realizado no Arena.

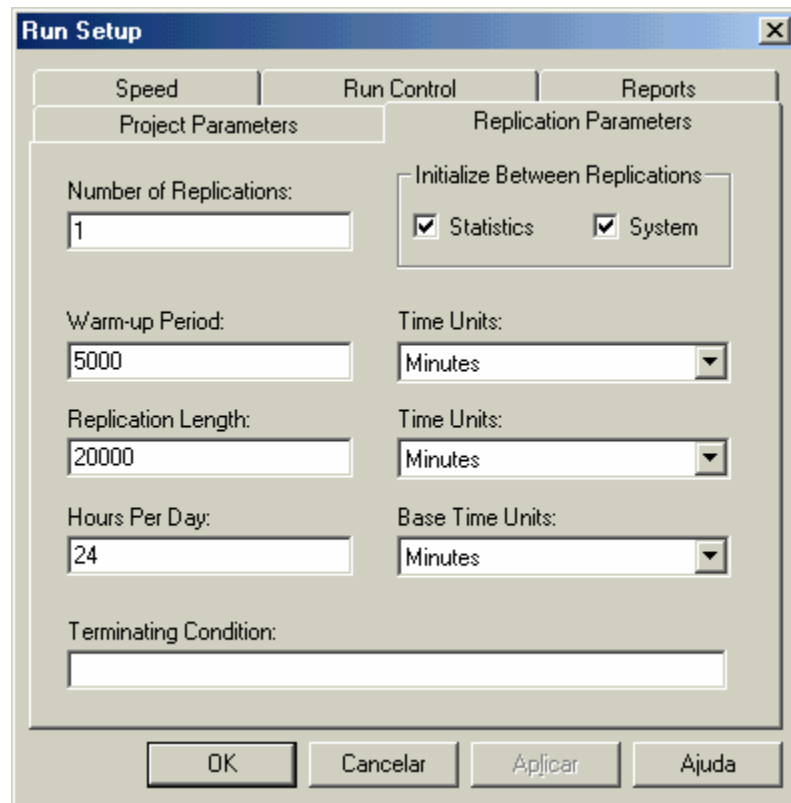


Figura 6.19: *Warm-up* de 5000 minutos definido em *Run/Setup/Replication Parameters*.

O efeito que o *Warm-up* produz é anular as estatísticas no tempo de simulação $t = 5.000$ minutos, manter o estado do sistema inalterado naquele instante e recomençar a coletar estatísticas. Desta forma, ao final da simulação de 20.000 minutos, as estatísticas se referem apenas aos últimos 15.000 minutos de simulação nos quais, acredita-se o sistema esteja em regime.

A outra maneira de proceder é coletar estatísticas desde o primeiro momento da simulação, guardá-las num arquivo de dados e descartar observações diretamente deste arquivo quando o mesmo estiver sendo tratado no *Output Analyzer*. É este o procedimento que se adotará, uma vez que já se tem o arquivo “TMP.DAT” contendo todos os dados. Neste caso, o descarte exige que se empregue a opção *Menu/Analyze/Batch/Truncate Observations*, conforme mostra a figura 6.20.

Na figura 6.20 pode-se ver as opções do comando que devem ser preenchidas. No primeiro campo o nome do arquivo a ser tratado. Em *Truncation*, observa-se *Type = Time*. Isto implica que se adotou um corte por tempo, isto é, todas as observações contidas no arquivo que tenham sido geradas antes do tempo 5.000 minutos serão descartadas. O campo *Batches* será usado logo adiante quando se proceder ao loteamento. Finalmente, se completa o campo indicando o nome de um novo arquivo já sem as observações iniciais. Você poderá salvar sobre o arquivo anterior, mas a sugestão é que se crie um novo arquivo, de forma a não se perder as informações completas presentes em “tmp.dat”. Escolheu-se o mesmo nome, trocando-se apenas

a extensão *dat* para *tru*, em alusão a truncado. Ao término do procedimento o *Output Analyzer* emite uma mensagem informando a criação do arquivo.

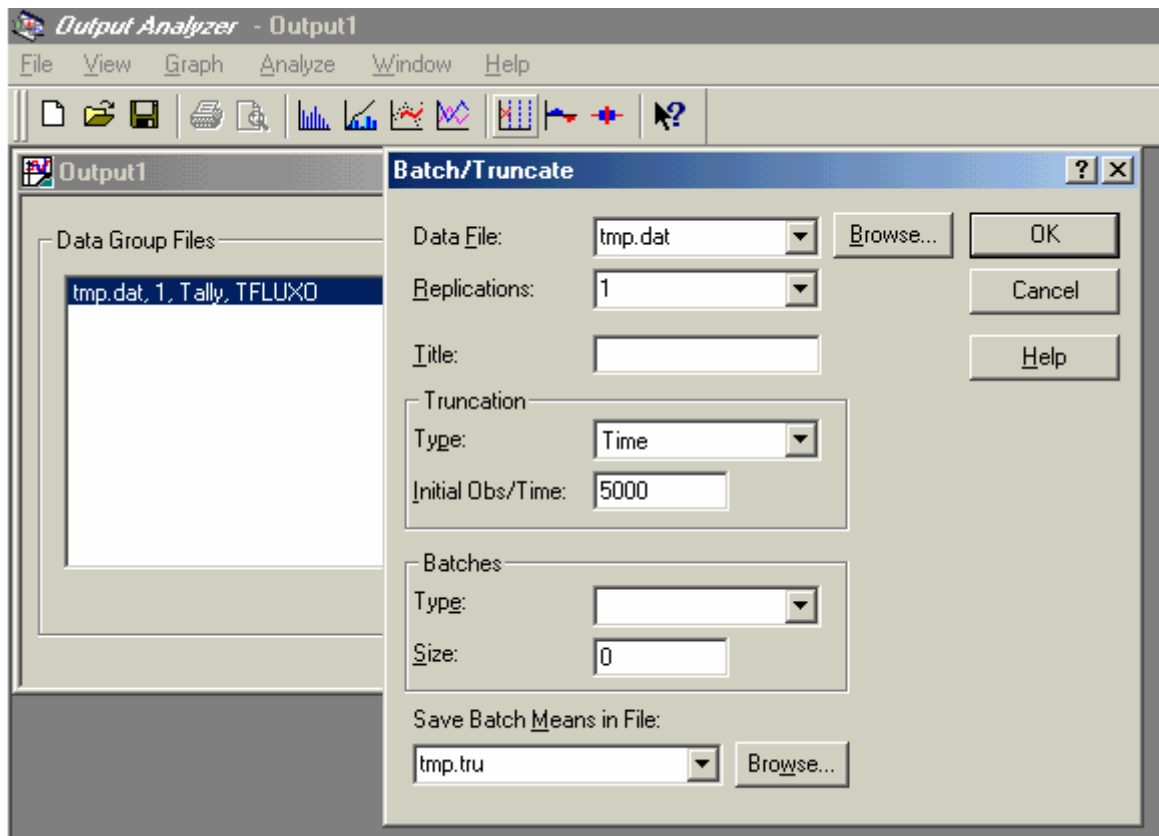


Figura 6.20: Comando *Menu/Analyze/Batch/Truncate Observations*

Uma vez que os dados estejam prontos, se passa a realização do próximo passo, isto é: determinar a correlação entre as observações. Este também se torna um procedimento simplificado, pelo uso do *Output Analyzer*. Invoque a opção *Menu/Analyze/Correlogram* para visualizar o comando e seus principais campos. A figura 6.21 mostra a interface.

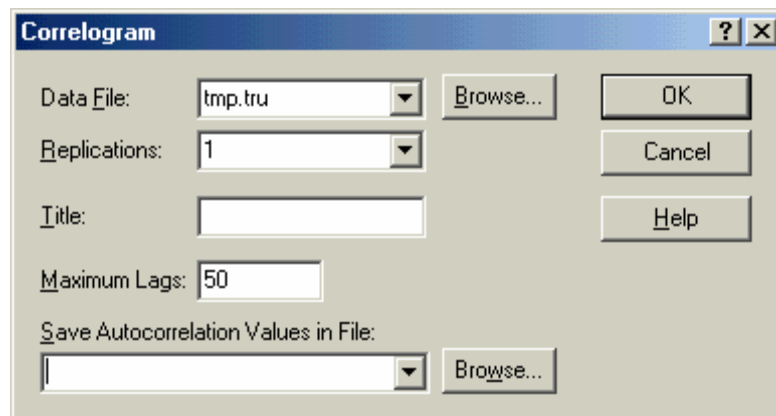


Figura 6.21: Comando *Correlogram*

Observe que no campo *Data File*, é o arquivo truncado que está presente. O campo *Maximum Lags*, é onde se determina a composição dos lotes de observações para se testar a correlação. O padrão do *Output Analyzer* é 50. O gráfico 6.7 mostra o resultado.

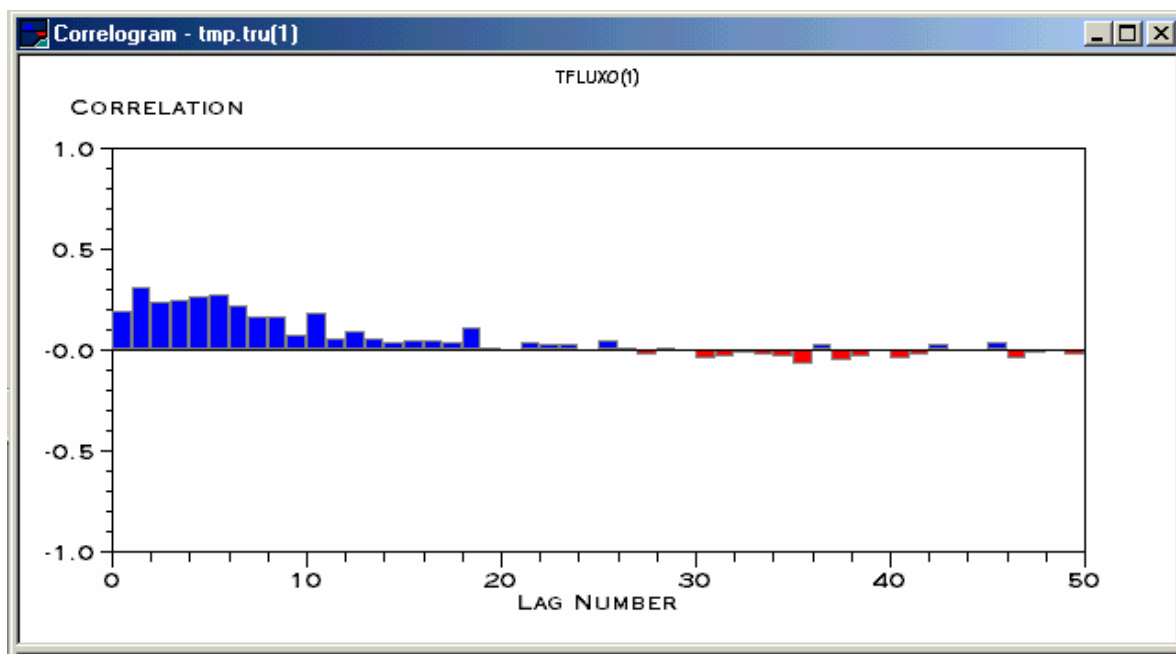


Gráfico 6.8: Correlação para a variável TMP.

Os resultados são apresentados em dois formatos: gráfico e texto. O texto encontra-se logo abaixo do gráfico. A tabela 6.4 apresenta os resultados no modo texto.

Tamanho do Lote	Covariância	Correlação
1	888,42	0,19359
2	1422,8	0,31004
3	1099,8	0,23965
4	1133,4	0,24698
5	1193,5	0,26007
.....
27	38,586	0,0084082
28	-128,55	-0,028012
29	35,254	0,0076822
30	-4,9892	-0,001087
31	-221,83	-0,048339
32	-184,01	-0,040097
.....

Tabela 6.4: Relacionamento entre tamanho de lote, covariância e correlação.

Seguindo-se os procedimentos sugeridos, foram medidas as covariâncias e correlações entre os valores médios dos lotes. Na medida do aumento do tamanho do lote n , verifica-se uma redução na covariância e no índice de correlação. Nas proximidades do lote com 30 elementos,

observa-se uma correlação no entorno de zero. Em outras palavras, pode-se pressupor que a influência do *tempo de passagem* pelo sistema de uma determinada entidade e_i sobre o *tempo de passagem* da entidade e_{i+30} será praticamente nula.

Nos procedimentos práticos, como o que agora se está realizando, os analistas costumam multiplicar n (tamanho do lote mínimo) por um coeficiente de segurança, como forma de “garantir” a necessária independência entre as médias destes lotes. Um número comum na literatura para este índice de segurança é 10 [ver Pegden (1990), Banks (1984) e Law (1991)]. Desta maneira, cada lote simulado deverá conter 300 observações (10 x 30). Em termos de tempo simulado, 300 observações são obtidas, em média, a cada 6.000 min. (300 x 20 min/observação), uma vez que, segundo os dados do problema, o tempo entre chegadas segue uma distribuição exponencial com média de 20 minutos. Considerando, por exemplo, uma simulação piloto de 15 amostras, o tempo total de simulação é dado pela fórmula empírica abaixo:

$$\text{Tempo de Simulação} = \text{Tempo Descarte} + \text{N}^\circ \text{ de Amostras} \times \text{Tempo Simulação cada Amostra}$$

Para o problema acima:

$$\text{Tempo de Simulação} = 5.000 \text{ min.} + 15 \times 6.000 \text{ min.} = 95.000 \text{ min.}$$

Este valor pode ser arredondado para cima, garantindo o número necessário de observações uma vez que, em função da variabilidade envolvida nos tempos entre chegadas, não se pode garantir que a cada exatos 6.000 minutos sejam realizadas 300 observações. Assim, assume-se 100.000 minutos como o tempo total de simulação.

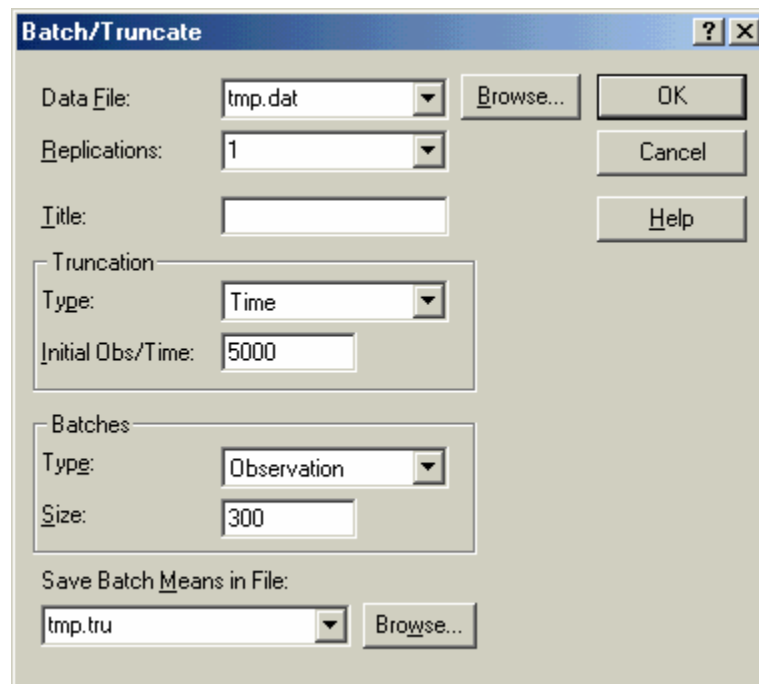


Figura 6.22: Montagem do arquivo de lotes.

O passo seguinte é simular novamente o sistema, agora com um tempo total de 100.000 minutos. Novamente os dados serão armazenados no arquivo “TMP.DAT”.

Uma vez executada a simulação, invoca-se novamente a opção *Menu/Analyze/Batch/Truncate Observations*, conforme exibe a figura 6.22. Observe que desta vez o campo *Batches* também é preenchido, informando que se deseja montar lotes com 300 observações cada um. Uma vez executado o comando, o *Output Analyzer* cria um arquivo (“TMP.TRU”) contendo um conjunto de valores, os quais estão subdivididos em lotes de 300 observações cada um. A figura 6.23 abaixo mostra como os lotes foram montados.

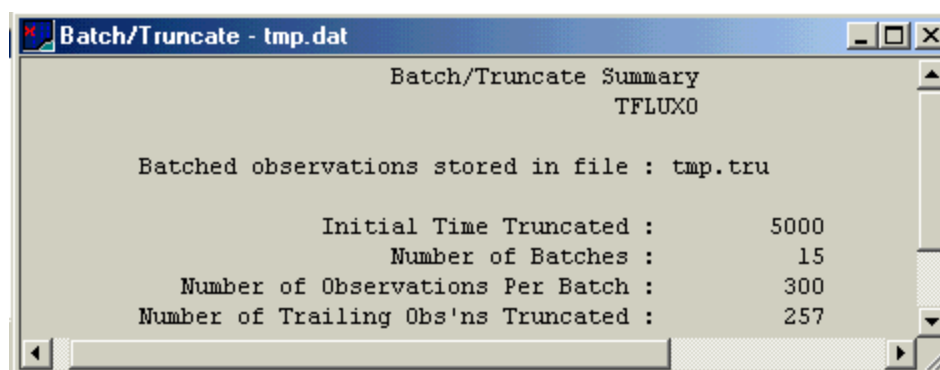
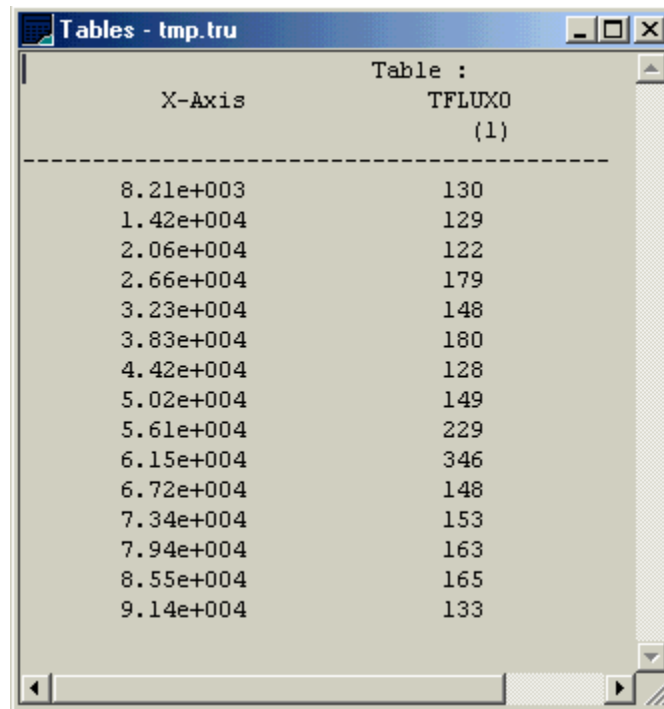


Figura 6.23: Informações sobre a montagem dos lotes.

O *Output Analyzer* informa que com o número de observações disponíveis ((15 x 300) + 257 = 4757), foi possível a criação de 15 lotes. Além disso, 257 observações restantes sobraram, pois, não podem constituir um novo lote. O analista, neste caso, pode, se o desejar, refazer a montagem dos lotes aumentando seu tamanho, de forma a aproveitar as observações que sobraram. Se aumentar o tamanho do lote de 300 para 317 observações, restarão apenas 2 observações sem uso. Outra opção é reduzir um pouco o tamanho dos lotes, aumentando a quantidade formada. Se reduzir de 300 para 297 observações/lote, poderá contar com 16 lotes. Tais procedimentos ficam a critério da analista. Novamente, o bom senso é que deve comandar. Se o custo de simular é pequeno, você poderá tranquilamente descartar as 257 observações restantes. No nosso caso, a simulação de 100.000 minutos nos custou apenas 8 segundos.

Após a simulação por 100.000 minutos e o descarte das observações ocorridas nos 5.000 minutos iniciais, forma-se lotes de 300 observações, cada um. O momento da formação dos lotes e seus respectivos valores de TMP podem ser verificados. A obtenção destes valores é feita pela solicitação ao *Output Analyzer* do conteúdo do arquivo de dados TMP.TRU”, pela opção *Menu/Graph/Table....* A figura 6.24 apresenta a tabela.



X-Axis	TFLUX0 (1)
8.21e+003	130
1.42e+004	129
2.06e+004	122
2.66e+004	179
3.23e+004	148
3.83e+004	180
4.42e+004	128
5.02e+004	149
5.61e+004	229
6.15e+004	346
6.72e+004	148
7.34e+004	153
7.94e+004	163
8.55e+004	165
9.14e+004	133

Figura 6.24: Tabela contendo os valores do arquivo TMP.TRU

É possível notar que os intervalos não são de exatamente 6.000 min., uma vez que as chegadas seguem uma distribuição exponencial. O primeiro lote, por exemplo, foi formado no tempo 8210 minutos de simulação. Verifica-se também, que o último lote foi formado por volta de um tempo simulado de 91.400 minutos. Portanto, as observações (257 delas) obtidas no intervalo de 91.400 minutos até 100.000 min., não serão aproveitadas. A última coluna da tabela apresenta os valores médios do tempo de passagem para cada um dos lotes de 300 observações.

Finalmente, chega-se à última etapa do processo. Os procedimentos finais do método do loteamento aplicado a sistemas considerados não-terminais são semelhantes ao método das replicações independentes usado para sistemas terminais. Foi obtido um arquivo contendo dados na forma de lotes, de forma semelhante aquele gerado para guardar os resultados das múltiplas replicações que se realizam para os sistemas terminais. Sendo assim, basta solicitar ao sistema que calcule o intervalo de confiança do arquivo indicado, considerando-o como um conjunto de replicações a serem tratadas de forma agrupada (opção *Lumped*).

Considerando as precauções estatísticas tomadas anteriormente, estes valores devem ser independentes. A amostra apresenta os seguintes valores para média e desvio-padrão, conforme os resultados apresentados na figura 6.25:

Média Amostral:	167
Desvio-padrão Amostral:	56,6

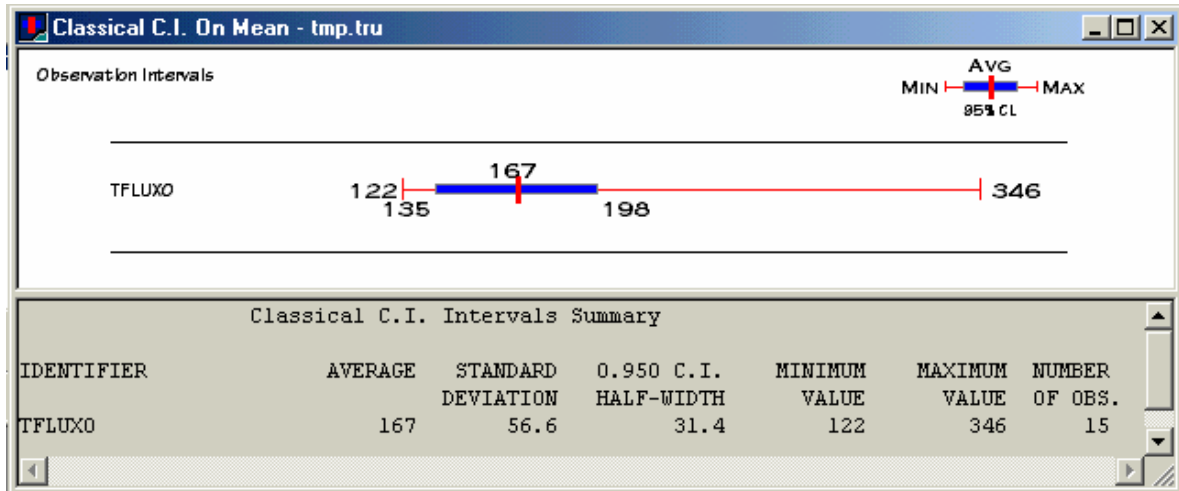


Figura 6.25: Intervalo de confiança da variável TMP

Com estes elementos, aplica-se a formulação desenvolvida no item 6.3 para a determinação do semi-intervalo h . O valor de h é dado por:

$$h = t_{n-1, 1-\alpha/2} \frac{S}{\sqrt{n}}$$

Assim tem-se: $h = 2,145 \times 56,6/3,87 = 31,37$.

O mesmo valor calculado acima é também observado na figura 6.25. Uma vez que se deseja a obtenção de um intervalo de confiança cujos limites estejam no entorno de 10% da média amostral (isto é, $167 \times 0,10 = 16,7$) e que o valor de h (31,37) é maior do que este limite, conclui-se que as 15 amostras geradas não são suficientes para garantir que a média do tempo de passagem das entidades pelo sistema encontre-se no intervalo $[150,3; 183,7]$, em 95% das observações. Desta forma, retoma-se o procedimento antes, porém, recalculando o número de amostras necessárias, dado pela equação 6.4.

$$n^* = [n(h/h^*)^2]$$

onde:

n^* = a nova estimativa para n

n = número de replicações já realizadas = 15

h = semi-intervalo de confiança já obtido = 31,4

h^* = semi-intervalo de confiança desejado = 16,7

$$n^* = [n(h/h^*)^2] = [15.(31,4 / 16,7)^2] = 15.(1,88)^2 \cong 54$$

Portanto, repete-se o processo considerando a formação de 54 lotes. O tempo de simulação necessário é dado pela expressão abaixo:

$$\text{Tempo de Simulação} = 5.000 \text{ min.} + 54 \times 6.000 \text{ min.} = 329.000 \text{ min} \cong 330.000 \text{ min.}$$

Arredondando para cima o valor obtido, se estabelece um tempo total de simulação igual a 330.000 minutos. Realizando a simulação e repetindo os procedimentos anteriores, os novos resultados aparecem na figura 6.26.

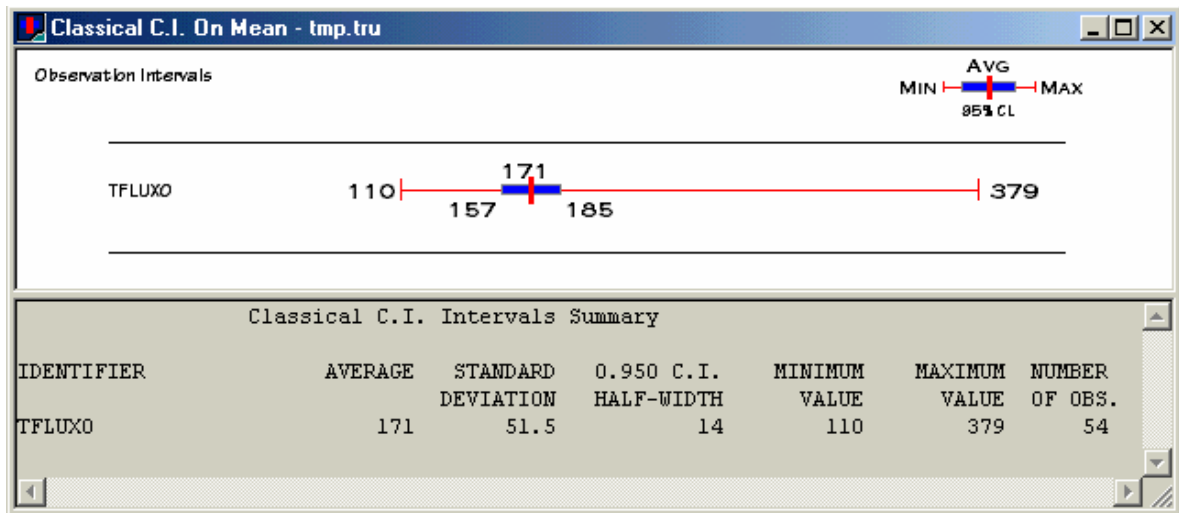


Figura 6.26: Resultados da simulação da segunda interação

Pela figura pode-se verificar que os resultados mostram que os objetivos de precisão para a estimativa da média dos tempos de passagem pelo sistema foram alcançados. O valor de h obtido para o intervalo de confiança é 14, que é menor que 10% da média amostral (17,1). Desta maneira, pode-se afirmar, com 95% de confiança que a variável TMP encontra-se no intervalo média \pm 10%, como se desejava.

Aplicação do Método das Múltiplas Replicações Independentes

A aplicação da técnica de múltiplas replicações independentes é feita à semelhança do que foi realizado para os sistemas terminais. Aqui, porém, deve-se descartar o período transiente de cada uma das replicações. Uma vez que se deseja comparar os dois métodos, o número de replicações será igual ao número de lotes do último exemplo. Assim, os 54 lotes equivalem a 54 replicações independentes. Uma vez que o número de replicações já está determinado, não será realizado o procedimento de efetuar primeiro uma amostra piloto, etc.

As estatísticas de cada replicação serão acumuladas por 6.000 minutos, a mesma duração de cada um dos lotes anteriores. No entanto, devido à necessidade de descarte das observações da fase transiente, cada uma das replicações terá, de fato, 11.000 minutos de simulação. Deste total, descartam-se as estatísticas acumuladas até o tempo de 5.000 minutos. Para tanto, basta definir no menu *Run/Setup/Replication Parameters* um período de *Warm-up*. A figura 6.27 ilustra este procedimento realizado no Arena.

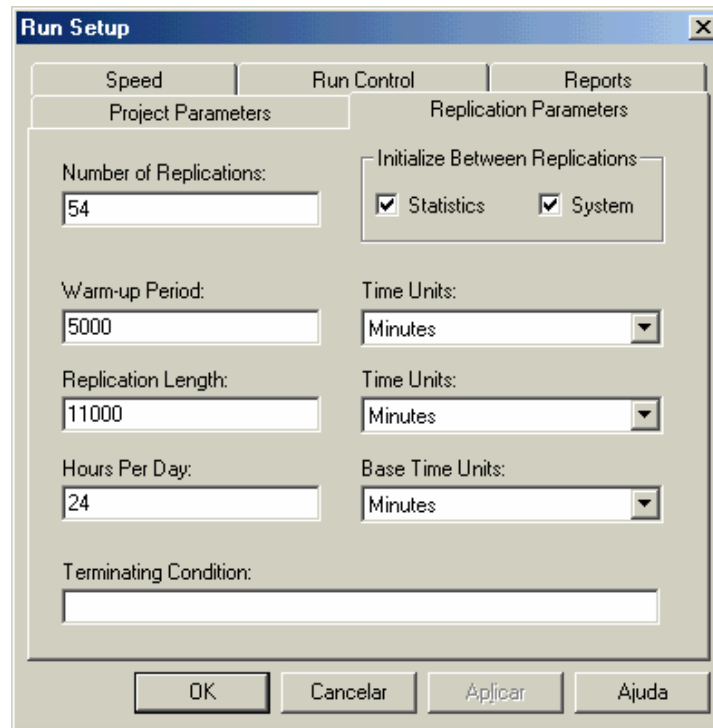


Figura 6.27: Definição do número de replicações e do *Warm-up*.

Para que o Arena colete as estatísticas dos valores médios dos tempos de passagem das entidades, em cada uma das 54 replicações, deve-se alterar o módulo *Statistics*. Lembre-se que agora o interesse é criar um arquivo contendo apenas os valores médios finais de cada replicação.

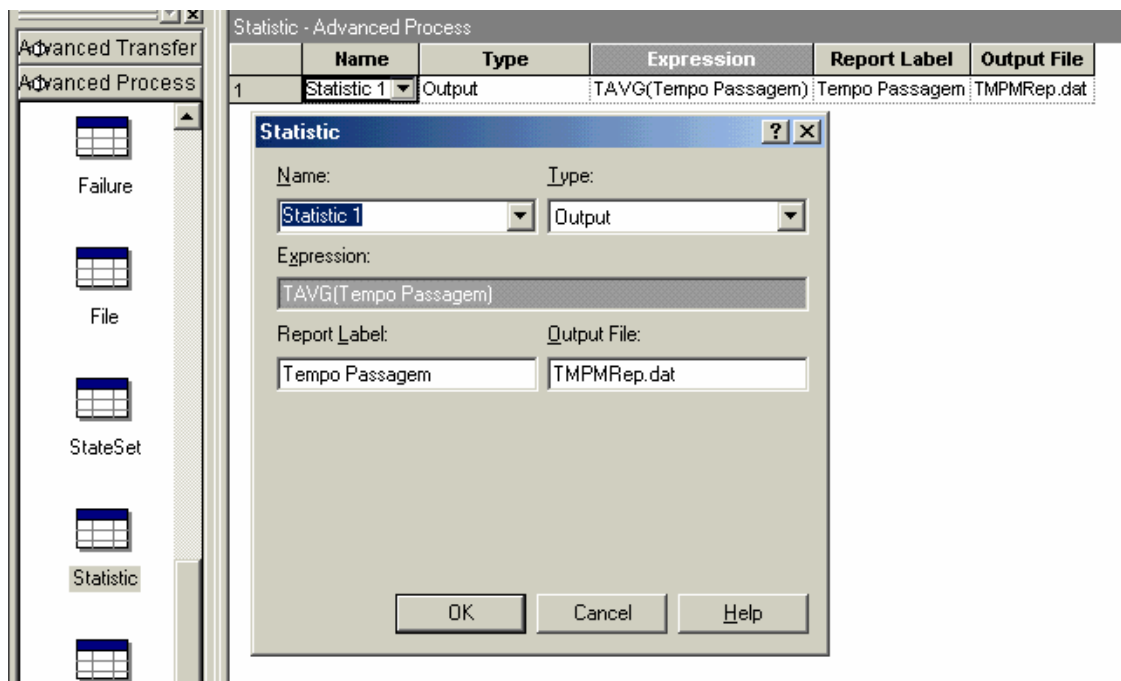


Figura 6.28: Solicitação de criação do arquivo de resultados tipo *Output* no módulo *Statistics*.

Desta forma, ativa-se no módulo o campo *Type* a estatística *Output*, como foi feito para os sistemas terminais. A figura 6.28 abaixo apresenta o comando. Observa-se que foi alterado o nome do arquivo de dados que será tratado no *Output Analyzer*, de “TMP.DAT”, para “TMPMR.DAT”, numa alusão ao processo de múltiplas replicações. Uma vez terminada a simulação, analisam-se os resultados. Uma vez ativado o *Output Analyzer*, solicita-se que este carregue o arquivo “TMPMR.DAT”. A figura 6.29 mostra o valor da estatística ao longo das 54 replicações.

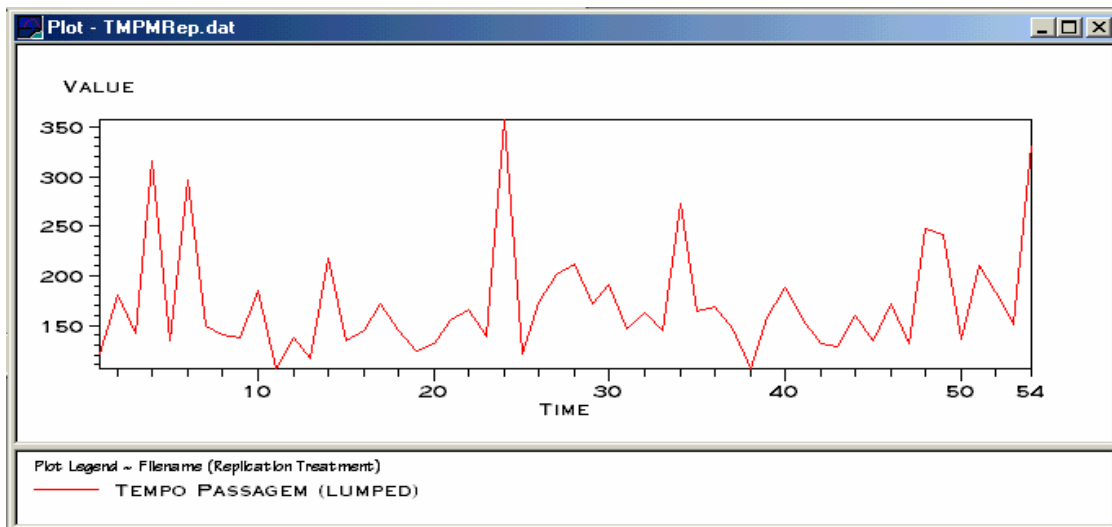


Figura 6.29: Resultado das 54 replicações.

A solicitação do cálculo do intervalo de confiança apresenta os resultados conforme apresentados na figura 6.30 abaixo.

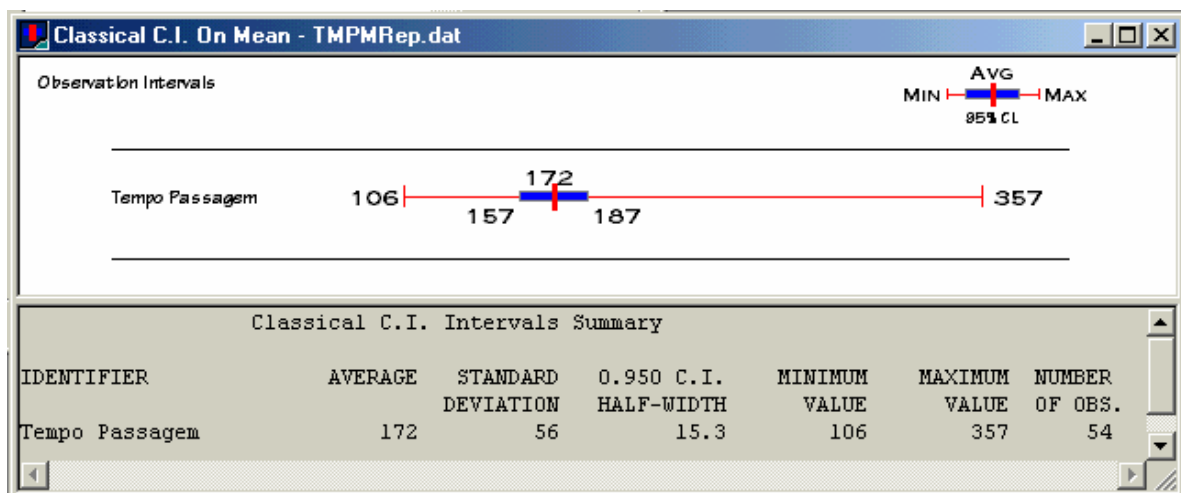


Figura 6.30: Intervalo de confiança para o experimento com as 54 replicações com descarte.

Analisando-se os resultados percebe-se que as exigências de se alcançar uma precisão para o intervalo de confiança em torno dos 10% do valor médio da variável de desempenho foi

alcançado. Para uma média final da amostra contendo 54 pontos igual a 172 minutos, foi obtido um valor de $h = 15,3$.

Além deste resultado, conclui-se que os dois métodos empregados chegam a resultados semelhantes. A tabela 6.5 apresenta uma breve comparação entre os resultados da técnica do loteamento com a das múltiplas replicações.

Técnica	Média	D. Padrão	h	V. Min.	V. Máx.	Lotes/Repl
Loteamento	171	51,5	14	110	379	54
M. Replicações.	172	56,0	15,3	106	357	54

Tabela 6.5: Comparação entre os resultados das duas técnicas.

A menos dos custos associados à obtenção dos intervalos de confiança, vê-se que os dois métodos são, de fato, equivalentes. No primeiro experimento, foram realizados 330.000 minutos de simulação com 325.000 minutos úteis e descarte de apenas 5.000 minutos. No último experimento, foram 594.000 minutos de simulação, com apenas 324.000 minutos úteis (descarte de 54 períodos de 5.000 minutos = 270.000 minutos). Se tal fator não for decisivo, isto é, se a complexidade do modelo não implicar em demoras no seu processamento (simulação), a recomendação é: use o método das múltiplas replicações com descarte.

Sumário

Neste capítulo tratou-se das principais técnicas para a abordagem e análise de resultados obtidos com a modelagem e simulação de sistemas discretos. Verificou-se a grande correspondência existente entre as técnicas aplicadas ao tratamento dos resultados de simulações de sistemas discretos e aquelas envolvendo a experimentação estatística. Foi visto que qualquer conclusão obtida com base em resultados simulados necessita uma prévia e adequada análise estatística. O propósito dos experimentos simulados é a obtenção de estimativas aceitáveis para as medidas de desempenho desejadas. Já o propósito das análises, é assegurar que tais medidas tenham a necessária e desejada precisão.

Exercícios

Sistema Terminais

1. O Dep. de Transito (DETRAN) de uma pequena cidade do interior possui dois tipos principais de clientes que buscam os seguintes serviços:
 - Ajuda para novas licenças (com novas placas) com Tempos entre Chegadas (TEC) distribuídos de acordo com uma EXPO (6.8) e Tempos de Serviço (TS) distribuídos de acordo com uma TRIA (8.7, 13.7, 15.2);

- Ajuda para renovações de licenças com TEC dado por uma EXPO (8.7) e TS providos por uma TRIA (16.7, 20.5, 29.2);.

O departamento possui duas filas, uma para cada tipo de cliente. Cinco funcionários prestam os serviços: Pedro, João, Maria, Clara e Mônica. Os dois primeiros atendem o setor de novas placas. Os dois últimos o setor de renovação. Maria funciona como uma espécie de “coringa” podendo atender os dois tipos de clientes. Maria atenderá o cliente que estiver esperando a mais tempo. Todos os funcionários estão disponíveis durante 8 horas por dia.

Realize 30 replicações, calcule o intervalo de confiança de 95% para o tempo esperado no sistema dos dois tipos de clientes.

2. Modifique o exercício anterior considerando a possibilidade de o funcionário João também ser polivalente. Compare os dois sistemas e verifique a existência de diferenças entre os dois sistemas. Compare usando as mesmas estatísticas (tempos no sistema).

Sistema Não-Terminais

3. Desenvolver um modelo de um sistema de recepção/expedição de materiais realizado por um conjunto de empilhadeiras. O sistema contém 5 docas de recepção e 3 docas de expedição. Os caminhões chegam as docas com cargas em pallets. Os tempos (todos em minutos) entre as chegadas de caminhões, em cada doca, seguem distribuições UNIF (30, 60). Cada caminhão chega com uma carga variável de pallets distribuída de acordo com uma UNIF (10, 25), que necessita ser transferida para as docas de expedição (no modelo os valores serão arredondados para o inteiro mais próximo). Assumir igual probabilidade de qualquer pallet ter que ir para qualquer das 3 docas de saída. Os tempos de descarregamento dos caminhões são assumidos como zero. Um equipamento especial encarrega-se da tarefa.

Uma vez que chegam nos caminhões, os pallets estão prontos para a transferência. Os pallets são transportados por um conjunto de 5 empilhadeiras. A distância entre as docas de chegada e as de expedição é de 25 metros. A velocidade das empilhadeiras é de 25 metros por minuto. O tempo necessário para carregar um pallet na empilhadeira é dado por uma UNIF (0.3, 0.4). Para descarregar é de UNIF (0.2, 0.4).

O desempenho do sistema é medido pelo tempo que os pallets despendem no sistema. Existem três propostas para comparar este sistema. Todas estão descritas abaixo. Implemente-as e escolha a melhor. Argumente sobre a escolha e justifique suas conclusões com base em uma análise estatística corretamente implementada. Além disso, para as opções (a) e (c), existem propostas alternativas de reduzir/aumentar o número de empilhadeiras. Teste estas alternativas com 4 e 6 empilhadeiras, respectivamente. Novamente, justifique suas respostas com uma análise estatística adequada.

- a) Assuma que a prioridade deva ser a transferência dos pallets que estejam a espera por transporte a mais tempo.
- b) Por causa do potencial congestionamento, dedique cada empilhadeira a cada uma das docas de chegada, modificando o modelo anterior.

- c) Como forma de garantir espaço para o descarregamento dos caminhões que chegam, modifique o modelo do item (a), de tal forma que a prioridade seja dada à doca de chegada com o maior número de pallets.