

UNIVERSIDADE DE SÃO PAULO
Escola de Engenharia de São Carlos
Departamento de Engenharia Mecânica

PROCEDIMENTOS PARA A MODELAGEM
ESTATÍSTICA DOS DADOS DE ENTRADA
PARA A SIMULAÇÃO DE SISTEMAS DE
EVENTOS DISCRETOS



Ana Paula Kiguti Muniz



Trabalho apresentado à Escola de Engenharia de São Carlos da Universidade de São Paulo, como parte dos requisitos para a obtenção do título de mestrado.

ORIENTADOR: Prof. Dr. Ricardo Yassushi Inamasu

São Carlos
2000

Serviço de Pós-Graduação EESC/USP
EXEMPLAR REVISADO
Data de entrada no Serviço: 28.16.2000
Ass.: *Elson*

Class.	TESE-EESC
Cutt.	3781
Tombo	0130/00

31100008672

S/S 1084559

PROCEDIMENTOS PARA A MODELAGEM
ESTATÍSTICA DOS DADOS DE ENTRADA
PARA A SIMULAÇÃO DE SISTEMAS DE
EVENTOS DISCRETOS

Ana Paula Kiguti Muniz

Ficha catalográfica preparada pela Seção de Tratamento
da Informação do Serviço de Biblioteca – EESC/USP

M966p Muniz, Ana Paula Kiguti
Procedimentos para a modelagem estatística dos
dados de entrada para a simulação de sistemas de
eventos discretos / Ana Paula Kiguti Muniz. -- São
Carlos, 2000.

Dissertação (Mestrado) -- Escola de Engenharia de
São Carlos-Universidade de São Paulo, 2000.

Área: Engenharia Mecânica.

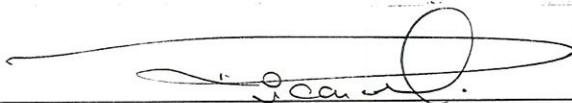
Orientador: Prof. Dr. Ricardo Yassushi Inamasu.

1. Modelagem de dados. 2. Simulação. 3. Sistemas de
manufatura. I. Título.

FOLHA DE APROVAÇÃO

Candidata: Licenciada ANA PAULA KIGUTI MUNIZ

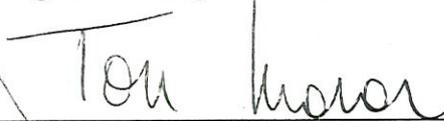
Dissertação defendida e aprovada em 28.04.2000
pela Comissão Julgadora:



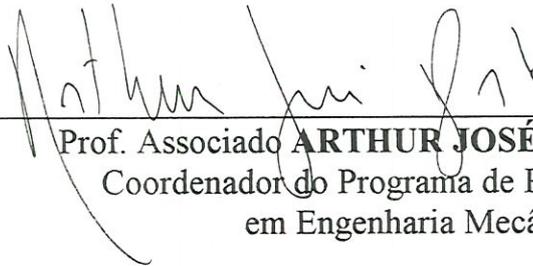
Prof. Doutor **RICARDO YASSUSHI INAMASU (Orientador)**
(Escola de Engenharia de São Carlos - Universidade de São Paulo)



Prof. Associado **ARTHUR JOSÉ VIEIRA PORTO**
(Escola de Engenharia de São Carlos - Universidade de São Paulo)



Prof. Associado **WASHINGTON LUÍS MARAR**
(Instituto de Ciências Matemáticas de São Carlos - Universidade de São Paulo)



Prof. Associado **ARTHUR JOSÉ VIEIRA PORTO**
Coordenador do Programa de Pós-Graduação
em Engenharia Mecânica



JOSÉ CARLOS A. CINTRA
Presidente da Comissão de Pós-Graduação da EESC

*Qualquer um que pára de
aprender é velho. Quem
continua a aprender é e
permanece jovem.
(Henry Ford)*

DEDICATÓRIA

*Aos meus pais pela minha primeira formação,
e ao meu marido Sérgio por ser hoje, meu
amigo e companheiro e a razão
maior do meu viver.*

AGRADECIMENTOS

Ao meu orientador *Prof. Dr. Ricardo Yassuchi Inamassu*, pela paciência e orientação fornecida durante a elaboração deste trabalho.

Ao *Prof. Dr. Arthur José Vieira Porto* pela oportunidade de realizar esta pesquisa.

À *Prof^ª Dr^ª. Maria Creusa Breta Salles (ICMC)*, *Prof. Dr. Luiz Gustavo Marcassa (IFSC)*, *Prof. Dr. Vanderley Bagnhato (IFSC)* e *Prof. Dr. Yashiro Yamamoto (IFSC)* pelo constante incentivo.

Aos amigos do Laboratório de Simulação de Eventos Discretos, *Jandira, Edilson, Orides, José Luis (famoso Zé do Pagode), Valério, Osvaldo Oshiro, Osvaldo Asato, Politano, Mamoru, Roberto, Tanaka, Renato, Gaga* pela amizade e colaboração no desenvolvimento deste trabalho. Em especial, eu agradeço a *Sayuri* pelo estímulo e preciosas sugestões, e a *Anna Cristina e ao Lobão* pelo auxílio na correção deste trabalho.

As secretárias do Depto de Mecânica *Bete e Ana Paula* por toda simpatia e apoio técnico.

Aos amigos do IFSC - Instituto de Física de São Carlos, *Daniel, Kilvia, Gugs, Marilia, Clebão (e Regina), Lino, Philippe (e Marilde)* pela amizade e apoio nos momentos difíceis.

À todos os amigos da USP, pelos bons momentos de companheirismo. Em especial a *Lú, a Lika, a Fabi, a Karen, a Érica e a Liliana (UFSCAR)*.

Aos meus avós de coração *Eliza e Mitsuo Suzuki* e família pelo carinho e amizade incondicional.

Aos meus padrinhos *Carlos e Isa* e família pelo apoio aos meus primeiros passos na vida de estudante longe do lar materno.

À minha família: minha mãe *Maria das Graças*, ao meu pai *Yoshio Kiguti* por todos os esforços para que eu chegasse até aqui. E ao meu irmãos *Douglas e Tays*, ao *Kleber* e ao meu sobrinho *Yuri* pela alegria do dia a dia.

Ao meu marido *Sérginho* por todo o apoio e compreensão pelos momentos de ausência dedicados a este trabalho.

À Deus e a todos os bons anjos que iluminaram mais esta etapa de minha vida.

À *CAPES* pela bolsa de estudo concedida.

RESUMO

MUNIZ, A.P.K. (2000). *Procedimentos para a modelagem estatística dos dados de entrada para a simulação de sistemas de eventos discretos*. São Carlos, 2000. 132 p. Dissertação (Mestrado) – Escola de Engenharia de São Carlos, Universidade de São Paulo.

Neste trabalho é estudado os procedimentos para a realização do modelamento estatístico de dados aplicados à simulação de sistemas de eventos discretos; sistematizando e integrando estas atividades à realização de um projeto de simulação.

Os procedimentos de modelagem estatística de dados são abordados em dois tópicos: modelagem de dados com coleta direta – que permite a análise de uma amostra de dados de uma variável do sistema a ser estudado pela simulação; e modelagem de dados com coleta indireta – que aborda o problema da ausência de uma amostra de dados na implementação de uma variável em um modelo de simulação.

O desenvolvimento desta pesquisa buscou permitir que os usuários de simulação utilizem os procedimentos estatísticos na modelagem de dados de entrada para os modelos de simulação de forma mais sistemática e confiável.

Palavras-chave: modelagem de dados, simulação, sistemas de manufatura

ABSTRACT

MUNIZ, A.P.K. (2000). *Procedures for statistical modelling of data of entrance for the discrete event system simulation*. São Carlos, 2000. 132 p. Dissertação (Mestrado) – Escola de Engenharia de São Carlos, Universidade de São Paulo.

In this work it is studied the procedures to implement the statistical modeling of the data used in the simulation of discrete events systems. The activities are systemized and integrated at a simulation project.

The procedures discussed here consider two different cases: first, when the direct data collection is possible; and also, when the access of data is not available, due to physical limitations or just because it is not economically feasible.

The aim of this work is to aid users to apply statistical analysis, on input data of a simulation model systematically and in a reliable way.

It is also presented results about a simulation model of a semi automatic manufacture line to which was applied several types of distributions on the input, in order to analyze the influence of these distributions to the final results of the simulation.

Key word: modelling of data, simulation, manufacturing system

SUMÁRIO

RESUMO	i
ABSTRACT	ii
SUMÁRIO	iii
LISTA DE TABELAS	v
LISTA DE FIGURAS	vi
LISTA DE SÍMBOLOS	vii
1.INTRODUÇÃO	1
1.1. JUSTIFICATIVA DO TRABALHO	1
1.2. OBJETIVOS DO TRABALHO	2
1.3. ESTRUTURA DO TRABALHO	2
2.REVISÃO BIBLIOGRÁFICA – SIMULAÇÃO	4
2.1.CONCEITOS FUNDAMENTAIS DE SIMULAÇÃO	6
2.1.1.SISTEMA E MODELO	7
2.2. SIMULAÇÃO NA FABRICAÇÃO	14
2.3.CONSIDERAÇÕES FINAIS DO CAPÍTULO	18
3.MODELAGEM ESTATÍSTICA DOS DADOS DE ENTRADA, PARA A SIMULAÇÃO DE MODELOS DE SISTEMAS DE EVENTOS DISCRETOS	20
3.1. MODELAGEM ESTATÍSTICA DOS DADOS DE ENTRADA COM A REALIZAÇÃO DE UMA COLETA DE DADOS DIRETA	23
3.1.1. IDENTIFICAÇÃO DAS VARIÁVEIS DE ENTRADA PARA ATINGIR OS OBJETIVOS DE ANÁLISE DO MODELO	24
3.1.2. O PROCESSO DE COLETA DE DADOS	25
3.1.3. VERIFICAÇÃO DA SUPOSIÇÃO BÁSICA: DE DADOS INDEPENDENTES E IDENTICAMENTE DISTRIBUÍDOS (<i>iid</i>)	26
3.1.4. CONSTRUÇÃO DE UM HISTOGRAMA	29
3.1.5. SELEÇÃO DA DISTRIBUIÇÃO DE PROBABILIDADE REPRESENTATIVA PARA OS DADOS COLETADOS	31
3.1.6. IDENTIFICAR OS PARÂMETROS DA DISTRIBUIÇÃO DE PROBABILIDADE	37
3.1.7. VERIFICAÇÃO DA REPRESENTAÇÃO DA DISTRIBUIÇÃO DE PROBABILIDADE SELECIONADA	38
3.1.8.VALIDAÇÃO DOS DADOS	39
3.1.9. DOCUMENTAÇÃO DOS DADOS DO MODELO	41
3.2. MODELAGEM ESTATÍSTICA DOS DADOS DE ENTRADA COM REALIZAÇÃO DE UMA COLETA DE DADOS INDIRETA, PARA MODELOS UNIVARIÁVEIS	42
3.2.1. ESPECIFICANDO UMA DISTRIBUIÇÃO DE PROBABILIDADE NA AUSÊNCIA DE UMA AMOSTRA DE DADOS	44
3.3.CONSIDERAÇÕES FINAIS DO CAPÍTULO	50

4. CASO PARTICULAR DE UM MODELO DE SIMULAÇÃO DE UMA LINHA DE PRODUÇÃO	52
4.1. PROCEDIMENTOS DE ANÁLISE DOS RESULTADOS GERADOS PELO MODELO DE SIMULAÇÃO	57
4.2. CONSIDERAÇÕES FINAIS DO CAPÍTULO	63
5. RESULTADOS E DISCUSSÕES	64
5.1. CONSIDERAÇÕES FINAIS DO CAPÍTULO	77
6. CONCLUSÃO	79
7. REFERÊNCIAS BIBLIOGRÁFICAS	83
APÊNDICE 1. DISTRIBUIÇÕES DE PROBABILIDADE	89
I. DISTRIBUIÇÕES CONTÍNUAS	89
I.1. DISTRIBUIÇÃO UNIFORME CONTÍNUA – $U(a,b)$	90
I.3. DISTRIBUIÇÃO GAMA (α, β)	93
I.4. DISTRIBUIÇÃO WEIBULL (α, β)	95
I.5. DISTRIBUIÇÃO NORMAL – $N(\mu, \sigma^2)$	97
I.6. DISTRIBUIÇÃO TRIANGULAR – $TRIANG(a,b,c)$	99
II. DISTRIBUIÇÃO DISCRETA	100
II.1. DISTRIBUIÇÃO DE BERNOULLI	100
II.2. DISTRIBUIÇÃO BINOMIAL	101
II.3. DISTRIBUIÇÃO GEOMÉTRICA	103
II.4. DISTRIBUIÇÃO NEGATIVA BINOMIAL	105
II.5. DISTRIBUIÇÃO POISSON	106
APÊNDICE 2. PROCEDIMENTOS DE ANÁLISE ESTATÍSTICA PARA SIMULAÇÃO TERMINATIVA	108
1. INTERVALO DE CONFIANÇA PARA UM NÚMERO FIXO DE REPLICAÇÕES	110
2. INTERVALO DE CONFIANÇA COM PRECISÃO ESPECIFICADA PELO MODELADOR	112
ANEXO 1. LISTAGEM DO PROGRAMA IMPLEMENTADO NO MODELO	115

LISTA DE TABELAS

TABELA 3.1 – COEFICIENTE DE CORRELAÇÃO.....	28
TABELA 4.1 – DEFINIÇÃO DOS <i>CONTROL-POINTS</i> DAS ESTAÇÕES DE TRABALHO, DAS ESTEIRA E FILAS DO SISTEMA	54
TABELA 4.2 – LISTAGEM DAS DISTRIBUIÇÕES FORNECIDA PELO SOFTWARE EXPERTFIT - PARTE I.....	58
LISTAGEM DAS DISTRIBUIÇÕES FORNECIDAS PELO EXPERTFIT.....	58
TABELA 4.2 – LISTAGEM DAS DISTRIBUIÇÕES FORNECIDA PELO SOFTWARE EXPERTFIT - PARTE II.....	59
TABELA 4.2 – LISTAGEM DAS DISTRIBUIÇÕES FORNECIDA PELO SOFTWARE EXPERTFIT - PARTE III	60
TABELA 4.3 – SEMENTE DO CONJUNTO DE NÚMEROS ALEATÓRIOS APLICADO.....	62
EM CADA REPLICAÇÃO	62
TABELA 5.1 – DADOS DE SAÍDA DO MODELO DE SIMULAÇÃO	69
TABELA 5.2 – RESULTADOS DA DISTRIBUIÇÃO EXPONENCIAL(0,1).....	70
TABELA 5.3 – RESULTADOS DA DISTRIBUIÇÃO EXPONENCIAL (0,09846).....	70
TABELA 5.4 - RESULTADOS DA DISTRIBUIÇÃO GAMA(0,09940;0,99053).....	71
TABELA 5.5 – RESULTADOS DA DISTRIBUIÇÃO WEIBULL(0,09819;0,99358)	71
TABELA 5.6 – RESULTADOS DA DISTRIBUIÇÃO LOGNORMAL (-2,90154;1,29292).....	72
TABELA 5.7 - RESULTADOS DA DISTRIBUIÇÃO WEIBULL (0,5; 1,0).....	72
TABELA 5.8 - RESULTADOS DA CONSTANTE 10 PEÇAS/MINUTO	73
TABELA 5.9 – RESULTADOS DO MODELO DE SIMULAÇÃO	74
TABELA 5.10 – COMPARAÇÃO ENTRE AS DISTRIBUIÇÕES SELECIONADAS COM A DISTRIBUIÇÃO HIPOTÉTICA IDEAL.....	75
TABELA 5.11 – INTERVALO DE CONFIANÇA PARA A DIFERENÇA ENTRE O NÚMERO MÉDIO DE PEÇAS PRODUZIDAS PELAS DISTRIBUIÇÕES	76

LISTA DE FIGURAS

FIGURA 2.1 - ELEMENTOS DE UM MODELO IMPLEMENTADO	7
FIGURA 2.2 - CLASSIFICAÇÃO DOS SISTEMAS DE ACORDO COM O COMPORTAMENTO DE SUAS VARIÁVEIS	9
FIGURA 2.3 - GRÁFICO DO COMPORTAMENTO DE UM SISTEMA DISCRETO	10
FIGURA 2.4 - GRÁFICO DO COMPORTAMENTO DE UM SISTEMA CONTÍNUO	11
FIGURA 2.5 - PROCEDIMENTOS PARA ANÁLISE DE UM SISTEMA	13
FIGURA 3.1 - ETAPAS DA MODELAGEM DOS DADOS DE ENTRADA COM A REALIZAÇÃO DE UMA COLETA DIRETA	23
FIGURA 3.2 - DIAGRAMA DE DISPERSÃO	27
FIGURA 3.3 - HISTOGRAMA	30
FIGURA 3.4 - <i>PP PLOT</i> QUANDO UMA DISTRIBUIÇÃO DE PROBABILIDADE É UMA BOA REPRESENTAÇÃO	34
FIGURA 3.5 - <i>P-P PLOT</i> QUANDO UMA DISTRIBUIÇÃO DE PROBABILIDADE NÃO É UMA BOA REPRESENTAÇÃO	34
FIGURA 3.6 - <i>Q-Q PLOT</i> QUANDO UMA DISTRIBUIÇÃO DE PROBABILIDADE É UMA BOA REPRESENTAÇÃO PARA O CONJUNTO DE DADOS	35
FIGURA 3.7 - <i>Q-Q PLOT</i> QUANDO UMA DISTRIBUIÇÃO DE PROBABILIDADE NÃO É UMA BOA REPRESENTAÇÃO PARA O CONJUNTO DE DADOS	36
FIGURA 3.8 - <i>BOX PLOT</i> DO CONJUNTO DE DADOS E DA DISTRIBUIÇÃO DE PROBABILIDADE AJUSTADA	37
FIGURA 4.1 - LAYOUT DA LINHA DE PRODUÇÃO	53
FIGURA 5.1 - HISTOGRAMA COM O AJUSTE DAS DISTRIBUIÇÕES SELECIONADAS PARA AVALIAÇÃO	68

LISTA DE SÍMBOLOS

- θ - estimativa, valor numérico
- $\hat{\theta}$ - estimador, função analítica que se baseia nos dados da amostra
- $\hat{\theta}_r$ - estimador da replicação r
- E - evento que limita a corrida de simulação
- t_e - tempo que limita o tamanho da corrida de simulação
- R - número de replicações do modelo de simulação
- r, i, s - índice
- μ - média da amostra
- H - semi-amplitude do intervalo
- α - probabilidade de erro
- ϵ - erro absoluto

CAPÍTULO 1

INTRODUÇÃO

Atualmente, a concorrência industrial está cada vez mais acirrada. Pois, se de um lado, o avanço tecnológico e a maior velocidade das informações, permitem maior rapidez no desenvolvimento; por outro lado também, a competição entre as empresas aumenta constantemente. Além disso, o mercado consumidor que antes tinha uma atitude passiva, hoje, devido a grande variedade de produtos, qualidade, preço e atendimento; passa a cada vez mais exigir qualidade e baixo custo. Segundo AGUIAR (1994), as dificuldades das empresas tem aumentado muito, pois se encontram pressionadas tanto por fatores externos, como o novo papel do consumidor. Como por fatores internos, traduzidos pelas inadequações dos sistemas administrativos e dos processos produtivos, ligados as vezes ao próprio desenvolvimento tecnológico.

Dentro deste contexto, onde rápidas mudanças ocorrem constantemente, as empresas se vêem obrigadas a continuamente inovar os seus meios de produção e de administração. Além de adotar novas filosofias, visando o aumento do nível de competitividade e suas chances de sobrevivência. Uma ferramenta que auxilia a atingir estes objetivos é a simulação, que permite inovar com maior rapidez e a um custo menor.

1.1. JUSTIFICATIVA DO TRABALHO

A simulação, cada vez mais, tem despertado o interesse tanto do meio empresarial, cujos principais objetivos são a diminuição dos custos, aumento da produtividade e conseqüentemente aumento dos lucros. Como do meio acadêmico, que procura compreender o funcionamento das novas ferramentas e processos, sempre com a

finalidade de auxiliar a formação de novos profissionais, e atualizar os conhecimentos daqueles que já atuam na área.

Realizando um levantamento bibliográfico sobre os motivos que podem levar um projeto de simulação a um insucesso, verificou-se que um projeto de simulação pode falhar devido a diversos fatores, dentre os quais temos a modelagem incorreta dos dados de entrada no modelo de simulação (ANNINO & RUSSELL (1981), CLARK (1994), CONWAY (1988), GROSS & JUTTIJUDATA (1997), ISHIHARA (1999), LAW et al (1994), MARTIN & BRADFORD (1993), WILLIAMS (1994)). Que se não invalidam o modelo como um todo, reduzem sua eficiência.

1.2. OBJETIVOS DO TRABALHO

Este trabalho tem por objetivo sistematizar o processo de coleta e modelamento estatístico dos dados de entrada. Com a finalidade de auxiliar os usuários na implementação de um modelo de simulação, de forma confiável. Para ressaltar a importância deste processo, foi desenvolvido um modelo de simulação de um sistema de manufatura hipotético, usando o software *AUTOMOD*.

1.3. ESTRUTURA DO TRABALHO

O trabalho apresenta a seguinte estrutura: levantamento bibliográfico; pesquisa e sistematização dos procedimentos para modelagem estatística dos dados de entrada; descrição do modelo de simulação proposto; execução do modelo de simulação; apresentação, discussão dos resultados e conclusão.

Inicialmente, no capítulo 2, realiza-se um levantamento bibliográfico sobre os tipos de sistemas, modelos e principalmente sobre os procedimentos necessários para se realizar um projeto de simulação de forma mais confiável.

Com base nesta revisão foi possível identificar as falhas mais frequentes que podem ocorrer em um projeto de simulação, entre as quais se encontra uma incorreta modelagem dos dados de entrada.

No capítulo 3 é apresentado uma sistematização para o processo de modelagem estatística dos dados de entrada. Esta sistematização, foi subdividida em dois casos:

- ♦ quando os dados podem ser coletados diretamente no sistema a ser modelado, e
- ♦ quando os dados são obtidos de fontes secundárias, ou seja, quando é realizado uma coleta de dados indireta.

A sistematização apresentada neste capítulo é aplicada no modelo de simulação proposto no capítulo seguinte.

O capítulo 4 é uma descrição do modelo de simulação usado como teste. Apresenta o *layout* do modelo, as variáveis implementadas no sistema, as condições iniciais consideradas para a execução do modelo e os critérios de prioridade do sistema.

O capítulo 5 mostra os resultados gerados pelo modelo de simulação. É uma discussão sobre as soluções obtidas, de acordo com cada um dos teste realizado na comparação dos resultados.

E finalmente, no capítulo 6, apresentam-se as conclusões finais deste trabalho.

CAPÍTULO 2

REVISÃO BIBLIOGRÁFICA – SIMULAÇÃO

A simulação é definida de variadas formas de modo a satisfazer diversas particularidades e escopos. TUMAY & HARREL(1997), definem a simulação como uma ferramenta que permite analisar o comportamento de um sistema real, com base no comportamento de um modelo, cujas relações de causa e efeito são equivalentes ou similares ao sistema modelado. Nesta mesma linha de abordagem, SHANNON et al.(1990) caracterizam simulação como uma metodologia experimental, que consiste em projetar e realizar experimentos com um modelo para compreender o comportamento do sistema real, avaliando desta forma novas estratégias de operação. Enquanto GOGG(1993) apresenta a simulação como uma arte e uma ciência que cria a representação de um processo ou sistema para os propósitos de experimento e análise.

De uma maneira geral, simulação pode ser compreendida como uma técnica e/ou ferramenta que permite obter conclusões sobre o comportamento de um sistema real, através da análise do comportamento de um modelo, com características essenciais muito próximas ao sistema real a ser analisado. A análise permite ao usuário obter um conhecimento antecipado dos efeitos que futuras ações e/ou modificações podem vir a gerar sobre o sistema estudado.

Esta ferramenta fornece uma série de vantagens, tais como:

- ◆ Permite obter conclusões sobre o comportamento de um sistema real, com base na análise do comportamento de um modelo, sem a necessidade de construir ou alterar o sistema real;

- ♦ Alternativas propostas para o planejamentos de sistemas (ou regras alternativas de operação de um sistema) podem ser comparadas via simulação para eleger a que melhor se adapta ao requisito especificado (LAW & KELTON,1991);
- ♦ Permite visualizar o sistema de forma global obtendo uma maior compreensão do sistema (SHANNON et al.,1990);
- ♦ Permite que o projeto do produto e/ou projeto do processo de produção sejam estudados a um nível bastante detalhado, pois na execução da simulação o ritmo do tempo pode ser alterado – acelerado ou desacelerado (AKBAY (1996) apud LOBÃO & PORTO (1997));
- ♦ Estudos sobre a utilização da capacidade instalada, níveis de inventário, lógica de controle, refinamento de projeto, integração, seqüenciamento, gargalos do sistema, melhor arranjo físico e melhor índice de produtividade dos funcionários;
- ♦ Serve como auxílio na tomada de decisões administrativas, funcionais e outras;
- ♦ Treinamento de operadores, testes de inicialização de equipamentos, etc.

E possui algumas dificuldades que devem ser consideradas:

- ♦ A implementação de um modelo requer treinamento especial;
- ♦ Um mesmo sistema pode ser representado por diversos modelos diferentes, que quando não bem analisados pode dificultar a tomada de decisão;

Devido a sua grande flexibilidade, a simulação é aplicada em diversas áreas de pesquisa. Segundo SHANNON et al.(1990), quase todo tipo de sistema foi (ou pode ser) simulado, e a larga série de aplicações de modelagem quase desafia a classificação. Algumas aplicações representativas são:

Fabricação. Sistemas de manuseio de material, linhas de montagem, automatização da produção, recursos de estocagem automatizada, sistemas de controle de inventário, estudos de segurança e manutenção, *layout* em planta baixa, projeto de máquina.

Negócios. Análise de mercadoria e estoque, policiamento de preços, estratégias de mercado, estudo de aquisições, análise de fluxo de caixa, planejamento, alternativas de transporte, planejamento do potencial humano.

Governo. Armas militares e seu uso, táticas militares, planejamento de população, uso da terra, serviços de polícia, justiça criminal, projeto de estradas, controle do tráfego, serviços de saneamento.

Ecologia e meio ambiente. Poluição e purificação de água, controle de devastação, poluição do ar, controle de pragas, previsão do tempo, análises de terremotos e tempestades, extração e exploração de minerais, sistemas de energia solar, produção das colheitas.

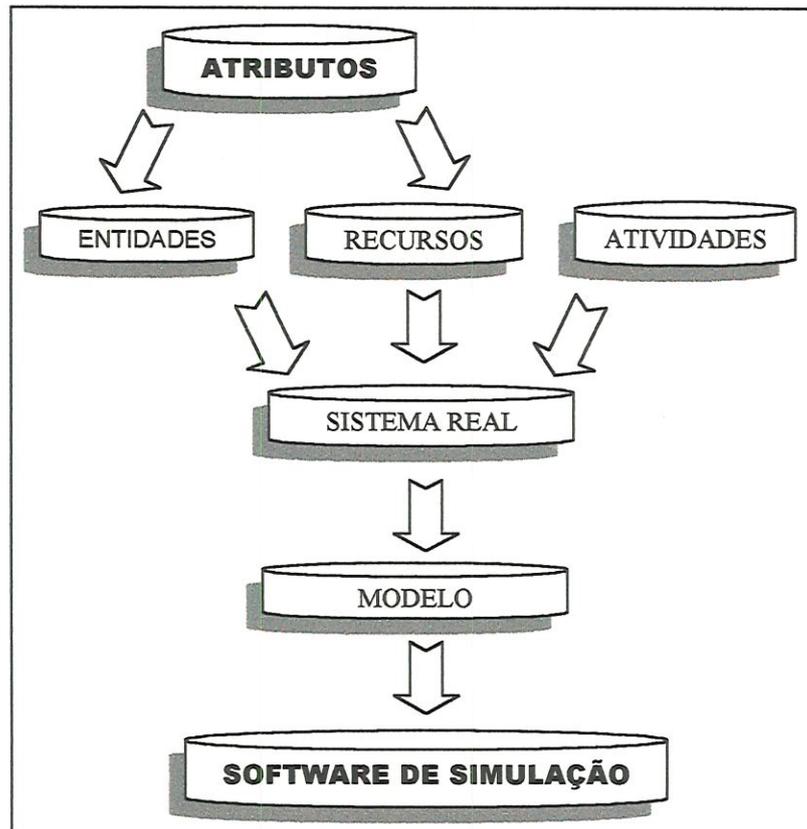
Sociedade e comportamento. Análise alimentação/população, política educacional, estrutura organizacional, análise do sistema social, administração de universidade.

Este trabalho tem como enfoque a aplicação da simulação na área de fabricação.

2.1. CONCEITOS FUNDAMENTAIS DE SIMULAÇÃO

Neste capítulo, serão apresentados alguns conceitos fundamentais para melhor compreensão do estudo de simulação, tais como: sistema, evento, entidade, recurso, e modelo.

A FIGURA 2.1 apresenta as relações existentes entre os elementos de um modelo implementado.



**FIGURA 2.1 - Elementos de um modelo implementado
(KELTON & VICENT,1992)**

2.1.1. SISTEMA E MODELO

Segundo SCHMIDT & TAYLOR apud LAW & KELTON (1991), um sistema é uma coleção de elementos, que se interagem com a finalidade de atingir algum objetivo. Em geral, os sistemas são influenciados por fatores externos, ou seja, fatores que ocorrem fora dos limites do sistema. Este lado externo é denominado de ambiente de

sistema (BANKS et al.,1996). O limite entre o ambiente de sistema e o sistema é definido pelos objetivos do estudo de simulação. Uma simulação só pode ser realizada se estes limites estiverem claramente determinados.

Geralmente, um sistema é composto pelos seguintes elementos: entidades ou atividades e recursos. CARSON (1992) define uma entidade como uma representação de algum elemento de interesse do sistema real que será modelado, classificando-a em:

- ♦ Dinâmicas – que possuem capacidade de locomoção dentro do sistema;
- ♦ Estáticas – que não se locomovem e em geral, servem outras entidades.

Recursos são entidades que fornecem serviços para outras entidades, e freqüentemente são representados pelo *status*: ocupado ou desocupado. As características pertencentes as entidades ou recursos são nomeadas de atributos.

De um modo geral, as entidades estão anexadas aos recursos pela lista de ordens que esta libera. As listas de ordens são especificações que uma entidade deve cumprir, tais como filas, tempos de espera ou outras condições. As operações realizadas pelas entidades são denominadas de atividades.

Geralmente os sistemas são classificados de acordo com as mudança de estado de suas variáveis em: estáticos, dinâmicos, discretos, contínuos ou uma combinação de sistema discreto e contínuo. O estado de um sistema pode ser compreendido como um conjunto de variáveis que definem o comportamento do sistema em um determinado período de tempo (BANKS et al.,1996). Cada mudança que ocorre no estado do sistema é um evento.

Segundo KHOSHNEVIS (1994), as variáveis podem ser definidas como medidas que caracterizam o sistema e que dependem de parâmetros ou de outras variáveis. Os parâmetros também são medidas controláveis de interesse de um sistema, que ao contrário das variáveis, são independente de outras medidas que configuram as condições de entrada do sistema. A FIGURA 2.2 mostra a classificação dos sistemas de acordo com o comportamento de suas variáveis.

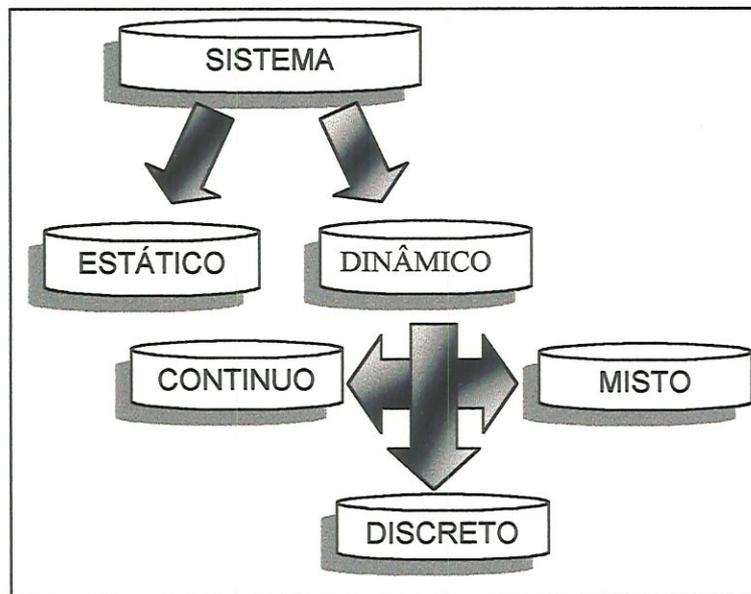


FIGURA 2.2 - Classificação dos Sistemas de acordo com o comportamento de suas variáveis

(KHOSHNEVIS,1994)

Um sistema estático é aquele cujas variáveis não se alteram com o decorrer do tempo. Em caso contrário, o sistema é dinâmico. Os sistemas dinâmicos podem ser: discretos, contínuo ou misto.

BANKS et al. (1996), definem os sistemas discretos como aqueles cujo estado das variáveis mudam em pontos discretos no tempo. Ou seja, a mudança de comportamento nestes sistemas ocorre aos saltos. Deve-se ressaltar que a maioria dos sistemas de manufatura e de serviço são sistemas de eventos discretos – SED'S (TUMAY & HARREL,1997). A FIGURA 2.3 mostra o comportamento de um sistema discreto.

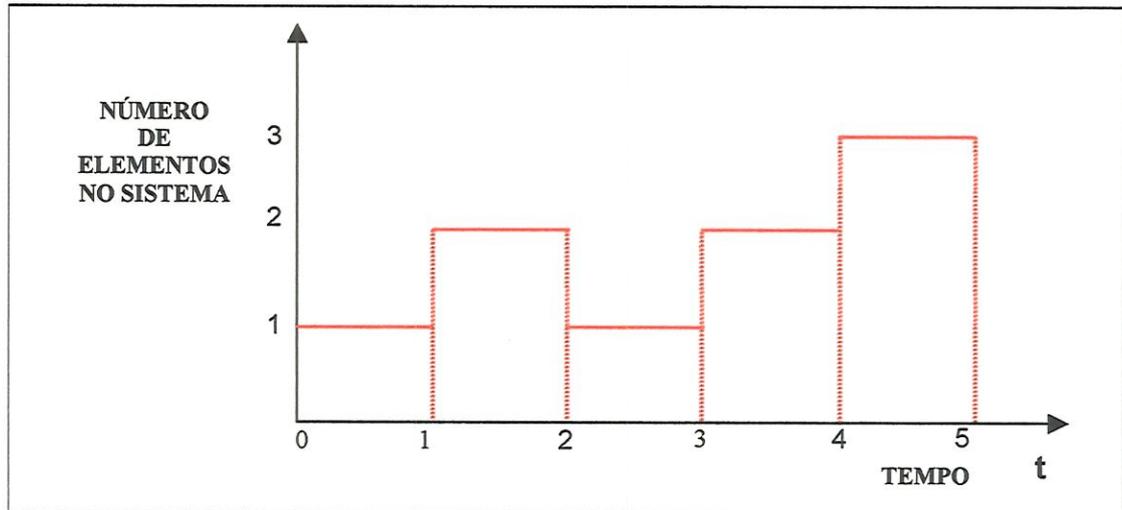


FIGURA 2.3 - Gráfico do comportamento de um sistema discreto

Já nos sistemas contínuos, o estado das variáveis se alteram sem interrupção ou quebra dentro de um considerado intervalo de tempo. A FIGURA 2.4, apresenta o comportamento de um sistema contínuo.

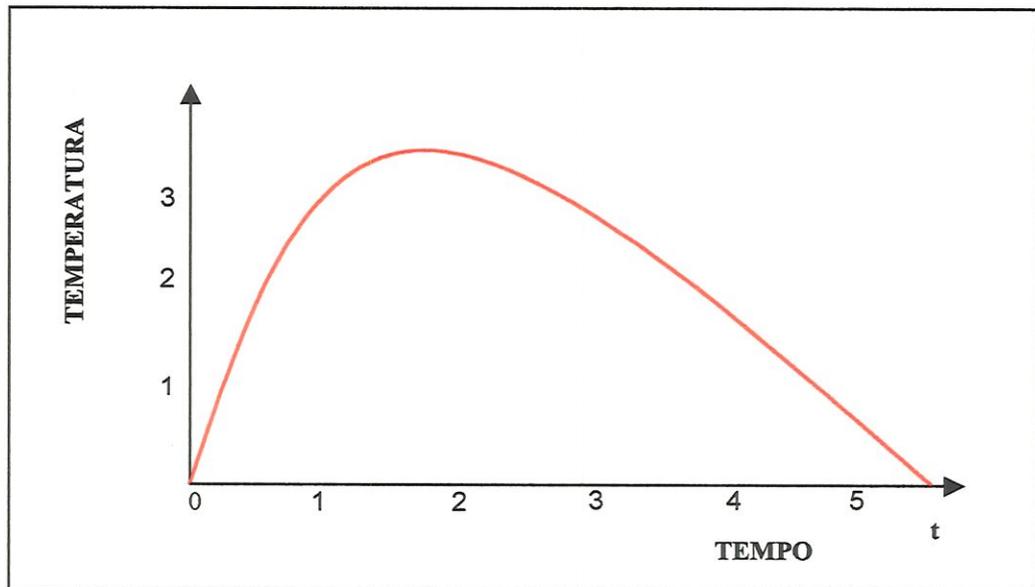


FIGURA.2.4 - Gráfico do comportamento de um sistema contínuo

De acordo com LAW& KELTON (1991), poucos sistemas na prática podem ser completamente discretos ou contínuos, mas desde que um tipo seja predominante sobre outro, o sistema será classificado de acordo com a predominância. Os sistemas podem ser analisados de duas formas, através de experimentos realizados no próprio sistema real ou através de modelos.

Um modelo é uma representação de um sistema real, utilizado para analisar o comportamento do mesmo (BANKS et al.,1996).

Segundo SHANNON et al. (1990), “desde que o homem começou a tentar entender e manipular o ambiente, os modelos tem sido usados para representar e expressar idéias e objetos.” Esta prática, atualmente é o principal meio utilizado para tentar encontrar a solução de problemas complexos.

Os modelos possuem algumas classificações muito semelhantes à classificação dos sistemas, podendo ser caracterizados de acordo com os seguintes casos:

1. Em relação a importância que a passagem do tempo possui na análise do sistema estudado, os modelos são divididos em dinâmicos e estáticos. De

acordo com LAW & KELTON(1991), modelos de simulação dinâmicos são aqueles com capacidade para representar um sistema com suas respectivas alterações sobre o decorrer do tempo. Como exemplo, pode ser citado o comportamento de um sistema bancário em um determinado período de tempo. Ao contrário dos modelos dinâmicos, os modelos estáticos representam um sistema em um instante particular de tempo. Para estes modelos, as vezes chamados de Método de Monte Carlo, o decorrer do tempo é insignificante (BANKS et al.,1996).

2. Em relação a presença de variáveis aleatórias, os modelos são classificados em: estocásticos e determinísticos. Segundo TUMAY & HARREL (1997), modelos estocásticos são aqueles que possuem uma ou mais variáveis aleatórias como dados de entrada. Enquanto que os modelos determinísticos são aqueles que não possuem variáveis aleatórias. E permitem através da especificação dos dados de entrada, determinar antecipadamente a solução que será obtida. Como exemplo, podemos citar um sistema de equações diferenciais que descrevem uma reação química. Uma diferença marcante entre os modelos determinísticos e estocásticos, se encontra no modo de analisar os resultados. Enquanto nos modelos determinísticos apenas uma única corrida de simulação é suficiente para obter um resultado exato; nos modelos estocásticos várias simulações devem ser corridas para se obter a estimativa mais precisa.
3. Em relação ao comportamento do estado do sistema modelado, os modelos são classificados de forma análoga à classificação dos sistemas em: discretos e contínuos. Segundo LAW & KELTON(1991), é comum ocorrer a relação entre sistemas discretos e modelos discretos, ou sistemas contínuos com modelos contínuos. No entanto, modelos discretos não necessariamente devem modelar sistemas discretos, de maneira análoga modelos contínuos não devem ser usados somente para os sistemas contínuos. A decisão por

qual tipo de modelo se deve optar cabe aos objetivos de estudo estabelecidos. Dependendo da precisão exigida para o modelo, é mais fácil trabalhar com um sistema contínuo em um modelo de simulação de sistema discretos.

4. Em relação ao tipo de procedimento utilizado para analisar um sistema, os modelos podem ser classificados em: modelo físico em escala e modelo matemático. A FIGURA 2.5 apresenta os tipos de análise.

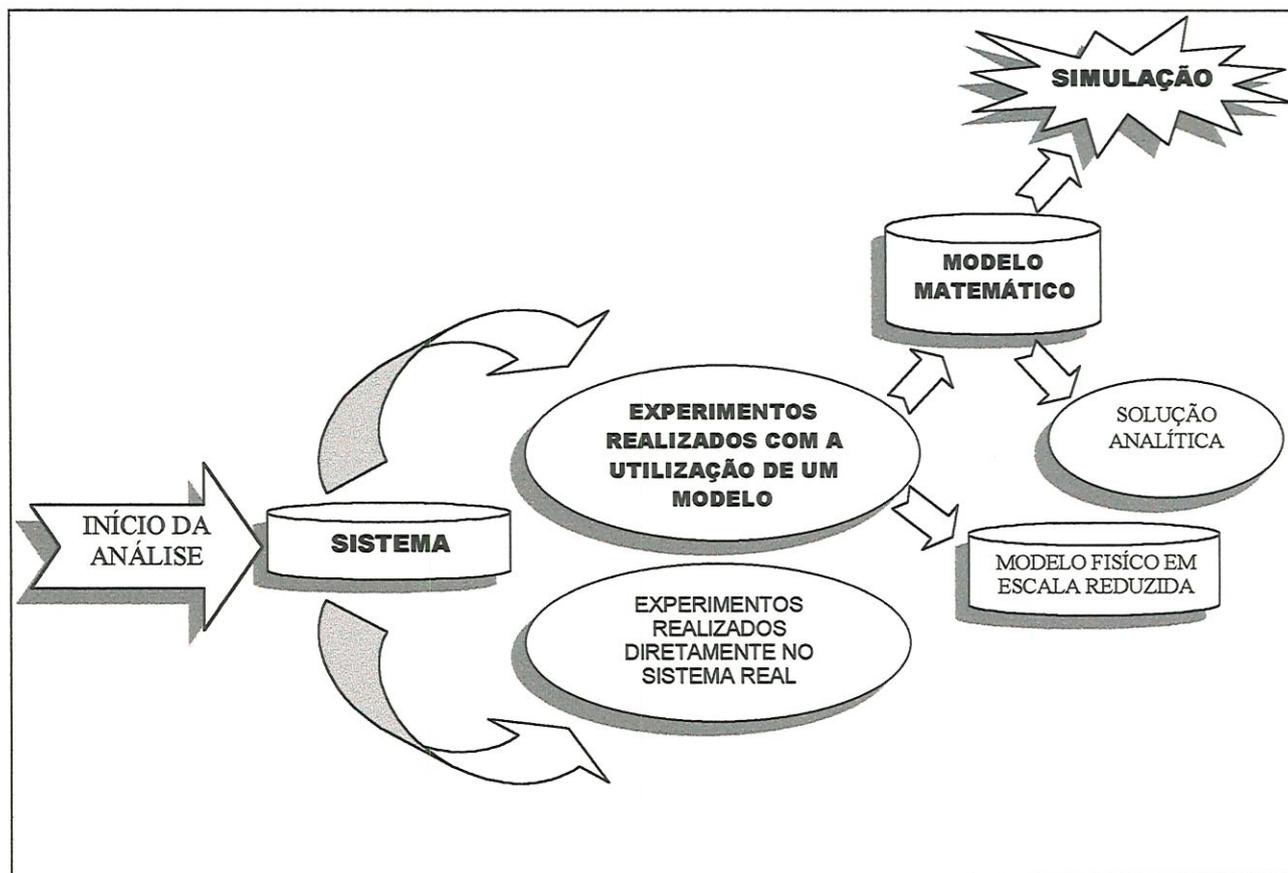


FIGURA 2.5 – Procedimentos para análise de um sistema
(LAW & KELTON,1991)

Modelos matemáticos são aqueles que representam as propriedades e características do sistema real em forma simbólica e/ou matemática, como por exemplo: uma equação do tipo $s = v \cdot t$ onde t é o tempo gasto e v é a

velocidade (SHANNON et al.,1990). Podem ser solucionados de duas formas (LAW & KELTON,1991):

- ♦ Soluções analíticas: só podem ser aplicadas quando o sistema é suficientemente simples. No entanto, em geral, a maioria dos sistemas são muito complexos para serem solucionados desta forma, sendo então necessário recorrer à simulação.
- ♦ Simulação: aplicada para estudar o comportamento de sistemas complexos. Consiste em analisar as saídas de um modelo, aplicando nas entradas as suas várias possibilidades.

Os modelos físicos em escala são aqueles que se utilizam da reprodução de objetos em proporções exatas com uso de material equivalente nos sistemas reais. De acordo com LOBÃO & PORTO(1996), apesar do avanço da tecnologia os modelos físicos em escala apresentam desvantagens como a inflexibilidade de readaptação para outros sistemas, elevado custo e complexidade de seu processo de montagem. Sendo portanto, geralmente aconselhável a aplicação dos modelos matemáticos, e em particular a utilização da simulação, quando necessita-se maior flexibilidade.

2.2. SIMULAÇÃO NA FABRICAÇÃO

A manufatura, uma das primeiras áreas a se beneficiar com a simulação vem aplicando de forma crescente esta ferramenta na implementação e inovação de seus sistemas (CLARK,1994). Este fato, que pode ser verificado através de um número crescente de trabalhos que abordam a manufatura em conjunto com a simulação, nos simpósios, congressos e outros encontros científicos.

Entre os diversos trabalhos existentes, encontra-se o modelo de um *job-shop* desenvolvido por MARIA (1997). Este modelo fornece uma visão geral, sobre como realizar e analisar um modelo de simulação. Para modelos mais complexos tem-se o trabalho de GUPTA et al. (1993) apud KENNE & GHARBI (1999). Eles desenvolveram um modelo detalhado de simulação, para estudar o planejamento de operações e os problemas de programação, em sistemas de manufatura. Nesta mesma abordagem, CONWAY et al (1988) analisam o efeito do WIP (*work in process*) sobre uma linha de produção. Considerando o problema de paradas e/ou quebras aleatórias de máquina, tem-se o modelo de simulação de uma linha de transferência, desenvolvido por MARTIN & BRADFORD (1993). CLARK (1994) também analisa um modelo com parada de máquinas. O efeito das falhas na manipulação de materiais em um sistema JIT, quando os tempos de processamento são estocásticos, foram avaliados por GUPTA & AL-TURKI (1998).

De um modo geral, a história da evolução da simulação, aplicada aos sistemas de manufatura, ocorre simultaneamente ao desenvolvimento da tecnologia e pode ser dividida em 4 fases (LOBÃO & PORTO, 1996):

Fase 1: Nesta primeira fase, a simulação era realizada baseando-se nos conhecimentos disponíveis da época, que consistiam em conhecimentos científicos, matemáticos, estatísticos e habilidades manuais. Os modelos eram de dois tipos: modelos matemáticos e os modelos físicos em escala. Com o avanço da tecnologia, os modelos físicos em escala receberam mais recursos tais como microprocessadores e microsensores eletrônicos. Paralelamente os modelos matemáticos foram aprimorados e atualmente existem diversos softwares que permitem simular modelos mais complexos tais como: *MATLAB* e *MATRIX*.

Fase 2: Inicialmente os sistemas de simulação através de computadores eram desenvolvidos em linguagens de programação, tais como : *FORTRAN*, *PASCAL*, *C*, entre outros. O problema nesta fase é a exigência de elevados conhecimentos de programação, e a grande

quantidade de tempo consumida para a elaboração dos modelos. Ainda nesta fase, surgem então as linguagens de programação voltadas a simulação, tais como *GPSS* e *SIMAN*, com o objetivo de facilitar o processo de construção dos modelos. O software *ARENA*, aplicativo de edição gráfica, foi desenvolvido com base na linguagem *SIMAN* (*ARENA SIMAN*,1994).

Fase 3: É marcada pelo desenvolvimento de programas simuladores específicos para determinados sistemas. Os simuladores, embora não apresentem a mesma flexibilidade que as linguagens de simulação, permitem maior facilidade na modelagem, dispensando o usuário da necessidade de profundos conhecimentos de programação. Com o aumento do poder de processamento dos computadores pessoais, surgiram vários sistemas simuladores comerciais tais como: *ARENA*, *PROMODEL* e *AUTOMOD*.

Fase 4: É o próximo passo da evolução da simulação, que será marcada pelos sistemas interativos inteligentes, que empregam técnicas de realidade virtual, inteligência artificial e sistemas especialistas.

Atualmente, o uso da simulação na fabricação tem aumentando de forma crescente. Isto ocorre devido ao aumento da concorrência industrial e as novas exigências do mercado consumidor; que em conjunto com as diversas vantagens que a simulação oferece, tem estimulado as empresas a investir na simulação, com a finalidade de melhorar seus sistemas produtivos, de forma a aumentar a qualidade de seus produtos, a um baixo custo, satisfazendo desta forma as necessidades de seus clientes.

De um modo geral, uma simulação pode ser compreendida como um projeto complexo, que pode ser sistematizado nas seguintes etapas (*MUNIZ & CARVALHO*,1998):

Fase 1: Elaboração

- ◆ Definição do problema e dos objetivos a serem atingidos;
- ◆ Coleta de dados e informações para elaboração do modelo e
- ◆ Refinamento dos dados obtidos.

Fase 2: Desenvolvimento

- ◆ Construção do modelo;
- ◆ Validação e verificação do modelo;
- ◆ Execução da simulação e
- ◆ Analisar e Ajustar o Experimento.

Fase 3: Análise

- ◆ Análise da Tomada de Decisão e
- ◆ Documentação final

É importante ressaltar que, as etapas apresentadas nesta sistematização variam de projeto para projeto.

Embora o planejamento sistematizado seja de grande auxílio na elaboração de um modelo de simulação, deve-se ressaltar que diversos erros podem ocorrer de modo a comprometer o sucesso de um projeto. Isto ocorre devido a existência de diversos fatores que influenciam na elaboração do modelo, que vão desde o tipo de software selecionado, habilidade do modelador e/ou equipe responsável, procedimentos de análise dos resultados, entre outros. LAW & KELTON(1991), BANKS & GIBSON(1997a), ANNINO & RUSSELL(1981) citam alguns erros comuns que devem ser evitados, para aumentar as chances de sucesso em um projeto de simulação:

- ◆ Utilização de fórmulas estatística tradicionais quando os dados não são *iid* (independente e identicamente distribuídas);
- ◆ Falta de uma concepção clara da definição do modelo;
- ◆ Uso arbitrário de uma distribuição de entrada na simulação;

- ◆ Nível inapropriado de detalhamento do modelo;
- ◆ Uso incorreto de uma animação;
- ◆ Análise de saída dos dados usando fórmulas estatísticas que assumem independência;
- ◆ Realização de uma única replicação de um modelo particular e tratar esta saída como uma resposta verdadeira;
- ◆ Execução de simulação em sistema que pode ser solucionado analiticamente, e
- ◆ Restrição no tempo de execução da simulação.

O presente trabalho possui foco na qualidade da distribuição dos dados de entrada em simulação de sistemas de eventos discretos. O processo que determina uma distribuição, para representar uma variável que será implementada em um modelo é denominado de modelagem estatística dos dados de entrada.

2.3. CONSIDERAÇÕES FINAIS DO CAPÍTULO

A simulação é uma técnica que muitas empresas têm empregado para aperfeiçoar seus processos de produção. E devido ao crescente avanço da tecnologia, vem sendo continuamente aperfeiçoada.

Atualmente existe uma grande quantidade de sofisticados *softwares* de simulação no mercado, com recursos gráficos e de programação que fornecem maior facilidade na elaboração de modelos, fornecendo um número maior de informações sobre o sistema estudado com mais rapidez.

Apesar deste crescente avanço, os princípios básicos para a elaboração de um modelo de simulação continuam sendo os mesmos. As dificuldades atuais para realização de um modelo de simulação se concentram na seleção de um *software* adequado para o problema em estudo, de modo a utilizar seus recursos com confiabilidade.

Embora o planejamento sistematizado de um projeto de simulação seja de grande auxílio, não existem garantias de que erros não serão cometidos na elaboração do modelo, comprometendo o sucesso do projeto. Sendo que, muitas vezes, recursos humanos qualificados contribuem de forma decisiva para realização adequada de um modelo de simulação.

CAPÍTULO 3

MODELAGEM ESTATÍSTICA DOS DADOS DE ENTRADA, PARA A SIMULAÇÃO DE MODELOS DE SISTEMAS DE EVENTOS DISCRETOS

O processo de modelagem dos dados de entrada consiste basicamente em encontrar uma distribuição de probabilidade, para representar o comportamento da variável a ser implementada em um modelo. Uma variável pode ser independente ou dependente. Uma variável independente é aquela que não possui influência sobre o comportamento de outra variável e vice-versa (KELTON & VICENT, 1992). Em caso contrário, a variável é dependente. Em geral, nos modelos de simulação, as variáveis são fontes de aleatoriedade do sistema. Ou seja, são variáveis de natureza estocástica, que registram as variações que ocorrem no sistema durante um intervalo de tempo. Estas variações podem ser causadas por fontes externas e/ou internas tais como: escassez de matéria prima, quebra de máquina, falhas de operação, variação no tempo manual das operações, e outros (BUZACOTT, 1972).

Nos modelos de simulação de sistemas de manufatura, é comum a existência de algumas das seguintes variáveis aleatórias (LAW & KELTON, 1994):

- ◆ Tempo entre chegadas de peças ou matérias primas;
- ◆ Tempo de processamento ou de montagem de peças;
- ◆ Tempo entre falhas de uma máquina;
- ◆ Tempos de reparo de máquinas e
- ◆ Tempos de *setup*.

Modelos de simulação que possuem apenas uma variável $X = (X_1, X_2, X_3, \dots, X_n)$ de interesse, onde X representa uma seqüência de variáveis aleatórias X_i , com $i = 1, 2, 3, \dots$ *iid* (independente e identicamente distribuídas), são ditos modelos de entrada univariável (NELSON et al., 1995). Uma variável X_i é *iid* quando esta seqüência de variáveis é mutuamente independente e se, todas as variáveis X_i , com $i = 1, 2, 3, \dots$ têm a mesma distribuição.

A importância de como os dados foram coletados e como serão introduzidos em um modelo, ou seja, o tipo de distribuição de probabilidade selecionada para representá-los, se concentra no fato de que esta distribuição possui grande influência sobre todos os resultados da simulação. Visto que, dados de entrada não confiáveis irão gerar resultados não confiáveis, afetando desta forma a qualidade dos resultados do modelo de simulação. E conseqüentemente nas decisões realizadas no sistema com base neste estudo. INGELS apud LOBÃO & PORTO (1997), ressalta este fato pelo seguinte axioma: “entra lixo, sai lixo”.

Um problema muito comum que pode ocorrer na modelagem de dados é a ausência de dados para realização deste processo. Este fato geralmente obriga o modelador a implementar o modelo, com estimativas sobre o comportamento da variável. Estas estimativas são, geralmente, baseadas em informações obtidas de fontes secundárias - ou seja, com baixa acuracidade.

Diante deste problema, este trabalho abordará o processo de modelagem dos dados de entrada em dois tópicos :

- ♦ Modelagem estatística com a realização de uma coleta direta – consiste na modelagem dos dados de entrada, quando estes foram coletados diretamente no sistema real a ser analisado pela simulação. Ressalta-se que para realizar uma coleta de dados direta, o sistema deve estar implementado. Um sistema implementado pode ser interpretado como um sistema que se encontra em pleno funcionamento de suas atividades e/ou operações.
- ♦ Modelagem estatística com a realização de uma coleta indireta – consiste na modelagem dos dados através de informações obtidas a partir de relatórios,

experiências de profissionais ou outras fontes indiretas. A principal característica desta modelagem, se concentra na impossibilidade do modelador realizar uma coleta de dados direta.

Deve-se ressaltar que, embora o processo de modelagem dos dados de entrada não seja o único fator a influenciar nos resultados de um modelo de simulação, este processo tem cada vez mais despertado o interesse de profissionais da área. Isto ocorre porque uma maior qualidade dos dados de entrada permite que as decisões voltadas para o sistema em estudo tenham menores chances de erro. Este fato pode ser comprovado através de experiências positivas como a Honda, que ao utilizar os dados de seus sistemas, conseguiu diminuir as falhas existentes na produção, atingindo desta forma melhorias em seus processos produtivos (ISHIHARA, 1999).

Esta crescente preocupação, voltada para a modelagem de dados, também pode ser observada pelo aparecimento de softwares estatísticos com a finalidade de auxiliar os profissionais da simulação na modelagem dos dados. Visto que, em geral, os modelos de simulação possuem uma grande quantidade de variáveis a serem analisadas, a modelagem de dados manualmente fica sendo, portanto, um processo muito trabalhoso. Com a utilização destes software é possível determinar com maior rapidez e acuracidade uma distribuição de probabilidade para representar o comportamento de uma variável. Além disto, eles fornecem uma maior facilidade para a elaboração de gráficos estatísticos.

Deve-se observar entretanto que, apesar das diversas vantagens oferecidas pelos softwares estatísticos, isto não elimina a necessidade do modelador entender o ambiente que produziu os dados e ter conhecimentos em estatística. Ou seja, erros podem ocorrer no processo de modelagem, devido a falhas do profissional de simulação.

Em vista deste problema, este trabalho objetiva apresentar os procedimentos de modelagem de dados de entrada para a simulação de modelos de manufatura, com a finalidade de auxiliar os profissionais da área. Tendo como principal enfoque, o processo de modelagem diante do problema da ausência de dados.

3.1. MODELAGEM ESTATÍSTICA DOS DADOS DE ENTRADA COM A REALIZAÇÃO DE UMA COLETA DE DADOS DIRETA

A modelagem de dados obtidos a partir de uma coleta direta, permite determinar com maior precisão uma distribuição de probabilidade para representar o comportamento de uma variável.

De um modo geral, este tipo de modelagem pode ser desenvolvido nas seguintes etapas ilustradas no diagrama da FIGURA 3.1.

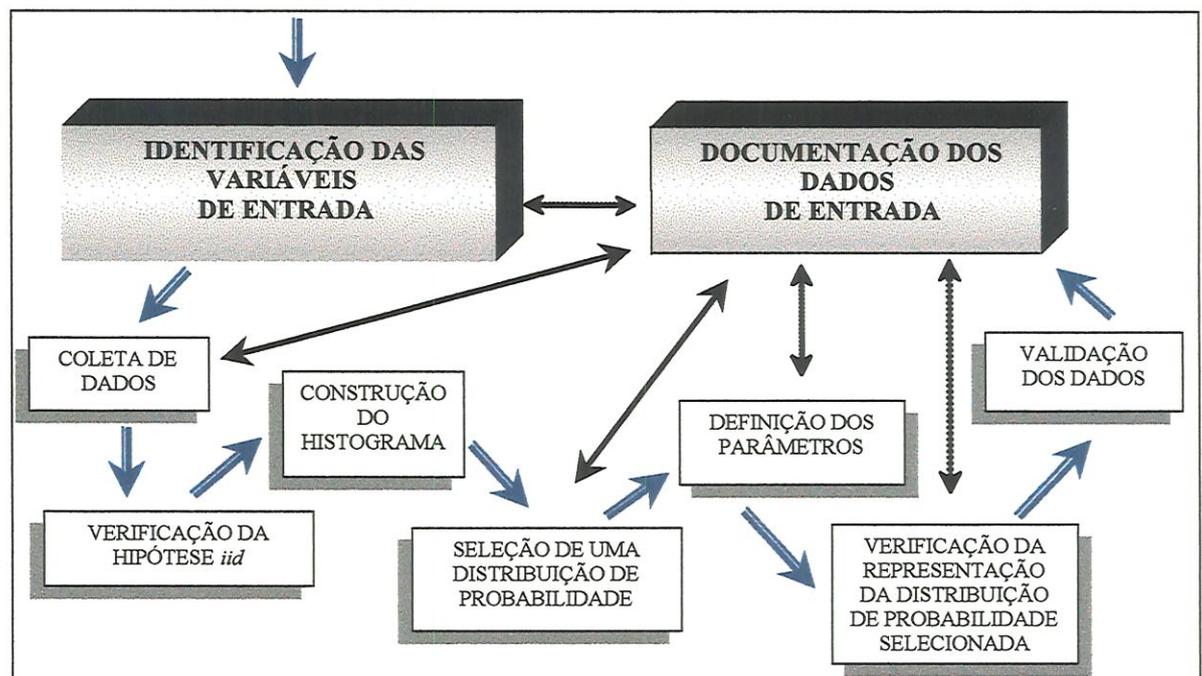


FIGURA 3.1 – Etapas da modelagem dos dados de entrada com a realização de uma coleta direta

A seguir serão analisados e comentados cada uma das etapas propostas para a modelagem dos dados com a realização de uma coleta direta.

3.1.1. IDENTIFICAÇÃO DAS VARIÁVEIS DE ENTRADA PARA ATINGIR OS OBJETIVOS DE ANÁLISE DO MODELO

Inicialmente, para se realizar a modelagem dos dados, deve-se identificar quais variáveis serão implementadas no modelo de simulação. De acordo com SHANNON et al.(1990), é fundamental definir de forma clara os objetivos e metas a serem alcançados pela simulação. Limitando, desta forma, as respostas que devem ser atingidas pelo estudo e conseqüentemente identificando as variáveis que serão implementadas no modelo. Os dados coletados devem ser direcionados para :

- ◆ Identificar as relações de causa e efeito existentes no modelo;
- ◆ Identificar os fatores de impacto. Deve-se preocupar somente com dados que fornecem alguma alteração no sistema. Dados que fornecem tempo de atividade com pequena variação sobre o comportamento do sistema, podem ser implementados como tempo constante, e
- ◆ Distinguir atividades que dependem de um período de tempo para executar uma operação, daquelas que só podem ser efetuadas se certas condições do sistema forem satisfeitas.

LAW & KELTON(1991) ressaltam que um modelo deve ser iniciado com poucos detalhes, sendo que estes devem ser acrescentados no decorrer da simulação se necessários. Para determinar os dados de entrada deve-se abordar as seguintes questões (USER'S GUIDE DA PROMODEL – "PROMODEL" (1994):

- ◆ Identificar todas as entidades do sistema e seus respectivos atributos;
- ◆ Determinar onde, quando e em que quantidade as entidades entram no sistema;
- ◆ Determinar a rota para cada entidade, assim como as prioridades que serão relevantes em caso de múltipla alternativa de rota para uma entidade e/ou para o caso de múltipla entidade e uma rota;

- ♦ Identificar as causas que disparam uma entidade de um local para outro, o tempo necessário para percorrer tal percurso e o tipo e a quantidade de recurso necessário;
- ♦ Identificar todos os recursos do sistema e suas respectivas quantidades, e
- ♦ Determinar como os recursos se movem de um local para o outro e as tomadas de decisões que devem ser realizadas.

Dependendo dos objetivos do modelo de simulação, as considerações acima são alteradas.

3.1.2. O PROCESSO DE COLETA DE DADOS

O processo de coletar os dados diretamente no sistema, consiste em registrar o comportamento da variável de interesse que será implementada no modelo. Ou seja, para realizar a coleta de dados “*in locu*”, é necessário que o sistema a ser modelado esteja implementado e seja acessível ao modelador e/ou equipe responsável pela simulação. Um sistema implementado é um sistema que se encontra em pleno funcionamento de suas operações e/ou atividades. Em geral, este processo requer mais de um responsável para ser executado.

Os dados de um sistema podem ser coletados manualmente através de observações pessoais do modelador e/ou equipe responsável pela simulação, e/ou por equipamentos de coleta automática. A coleta de dados deve ser realizada de forma homogênea em sucessivos períodos de tempo, ou seja, dias, meses ou anos dependendo dos objetivos do estudo. Coletar os dados de forma homogênea pode ser interpretado como o registro do comportamento dos dados em intervalos de tempo iguais, como por exemplo das 11:00 as 14:00 hs (BANKS et al. ,1996).

Segundo MUSSELMAN da *Prisker Corporation* (LAW,1993), a coleta de dados deve ser realizada por aqueles que possuem maiores conhecimentos em estatística e sobre os processos envolvidos no sistema. Além disso, deve-se ter conhecimento de como os dados foram coletados e como estes serão aplicados no modelo (COWLEY da

General Motors Associates (LAW,1993)). COSTA NETO (1977) adverte que, nesta fase não basta apenas descrever convenientemente os dados da amostra e ter conhecimentos sobre todas as técnicas estatísticas, é necessário garantir que a amostra ou amostras foram obtidas por processos adequados. Se erros forem cometidos no processo de coleta, o trabalho final ficará comprometido.

Uma das maiores dificuldades para coletar os dados é determinar quando a coleta de novos dados não trará maior precisão para a amostra de dados. De acordo com HARREL et al (1992), uma boa interação com o cliente ou outros beneficiários dos estudo é essencial para determinar este limiar.

Os dados quando coletados de um sistema podem ser implementados diretamente em um modelo de simulação. O inconveniente desta proposta é que a simulação se encontra limitada a reproduzir somente o que ocorre para os dados coletados. Desta forma, valores que se encontram entre um dado coletado e seu consecutivo dado registrado, não possuem uma representação. Além disto, freqüentemente existe o problema da insuficiência de dados para executar as diversas replicações. Por tal motivo, geralmente utiliza-se uma distribuição de probabilidade para representá-los.

3.1.3. VERIFICAÇÃO DA SUPOSIÇÃO BÁSICA: DE DADOS INDEPENDENTES E IDENTICAMENTE DISTRIBUÍDOS (*iid*)

É necessário garantir que os dados coletados, ou seja, a variável de interesse é *iid* (LEEMIS, 1996). Visto que sem esta condição, não é possível tentar encontrar uma distribuição de probabilidade para se ajustar ao conjunto de dados. Uma variável X_i com $i=1,2,3,\dots,n$ é *iid*, se todos os valores desta seqüência de variáveis aleatórias possuírem a mesma distribuição de probabilidade e forem independentes entre si.

Para se verificar a independência destes dados, pode-se utilizar os seguintes métodos:

- a) Diagrama de dispersão : Consiste em apresentar os dados de entrada $X_{[1]}$, $X_{[2]}$, ..., $X_{[n]}$, em pares $(X_{[i]}, X_{[i+1]})$ para $i = 1, 2, \dots, n - 1$ em um sistema

de coordenadas. Se $X_{[i]}$ são independentes os pontos serão dispersos aleatoriamente no gráfico (LAW & KELTON ,1991).

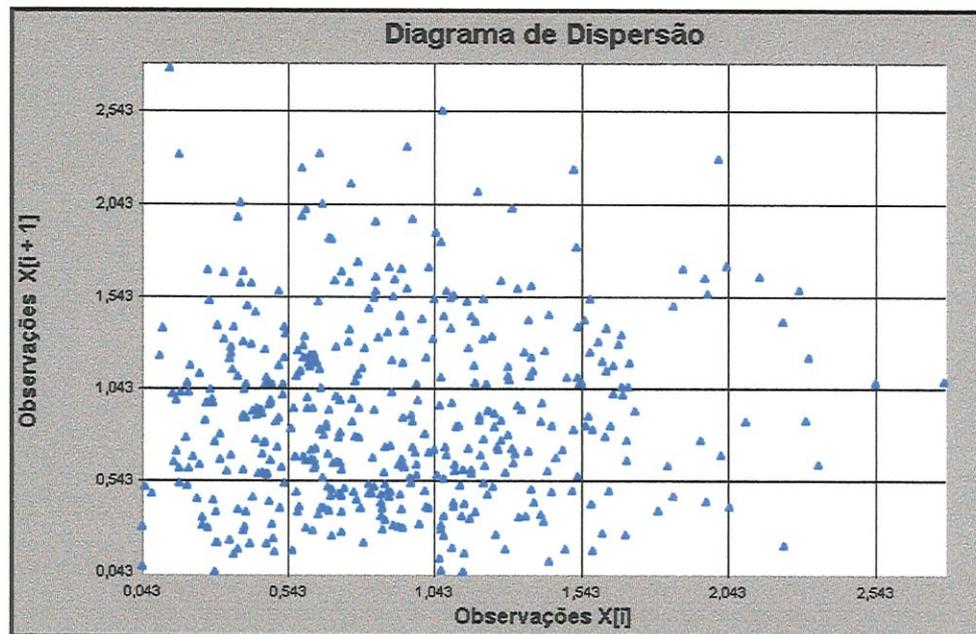


FIGURA 3.2 – Diagrama de Dispersão

b) Gráfico de correlação: Mede o grau de relação entre as variáveis através do coeficiente de correlação $\rho \in [-1,1]$. O valor de ρ pode ser calculado pela seguinte equação:

$$\rho = \frac{\text{cov}(x, y)}{S_x S_y} \quad (3.1)$$

Onde $\text{cov}(x, y)$ indica a covariância entre as variáveis x e y , e S_x e S_y são respectivamente o desvio padrão de x e y . Se as variáveis forem independentes, então $\rho = 0$, ou seja, não existe correlação. Entretanto, deve-se ressaltar que a recíproca não é válida. Além disto, quanto mais próximos ρ de $+1$ ou de -1 maior a dependência entre as variáveis. A seguir apresenta-se na TABELA 1 uma relação para os valores de ρ e a dependência das variáveis.

Tabela 3.1 – Coeficiente de Correlação
(FRANCISCO,1993)

Valores de ρ	Grau de dependência das Variáveis
$\rho = 0$	Não existe correlação
$\rho = 1$	Correlação positiva, os valores de uma variável aumentam
$\rho = -1$	Correlação negativa, os valores de uma variável diminuem
$-0.3 < \rho < 0.3$	A correlação é muito fraca e pouco pode-se concluir sobre as relações
$-0.6 < \rho < -0.3$ ou $0.3 < \rho < 0.6$	Existe uma correlação relativamente baixa entre as variáveis
Para ρ com valor diferente dos casos anteriores	A medida que ρ cresce de 0.6 a 1 ou decresce de -0.6 a -1 a correlação se torna mais forte entre as variáveis

Deve-se ressaltar que, as seqüências de variáveis aleatórias são caracterizadas em duas simplificações que abrangem a maioria dos problemas: variáveis independentes e identicamente distribuídas; e o caso de dependência de Markov.

Quando os dados não se encaixam em nenhuma destas duas caracterizações, a seqüência de variáveis são dependentes, o modelo é dito ser um modelo multivariável ou um modelo de tempos de série.

O presente trabalho tem como escopo variáveis independentes e identicamente distribuídas.

3.1.4. CONSTRUÇÃO DE UM HISTOGRAMA

Um histograma é útil para apresentar a freqüência dos dados de uma amostra e/ou a freqüência relativa (SARGENT,1996). Para construir o histograma deve-se agrupar os dados em k intervalos adjacentes. Uma das grandes dificuldades nesta etapa é escolher o número de k intervalos devido a inexistência de uma regra geral. Segundo LAW & KELTON (1991), a melhor opção para escolher o valor de k é a aplicação da *regra de Sturges*:

$$k = 1 + 3.322 \log_{10} n \quad (3.2)$$

onde n é o tamanho da amostra. Ao determinar o valor de k deve-se verificar se este valor cobre todo o intervalo existente entre o menor e o maior valor da amostra. Caso isto não ocorra, aconselha-se a tentar $k+1$ intervalos, e novamente fazer a verificação.

A amplitude Δb de cada intervalo pode ser calculada pela seguinte equação:

$$\Delta b = \frac{R}{k} \quad (3.3)$$

Onde $R = X_{max} - X_{min}$, é a diferença entre o maior valor da amostra definido por X_{max} , e o menor valor definido por X_{min} . Na prática, é aconselhável testar diversos valores de Δb para encontrar um histograma que apresente uma distribuição sem grandes irregularidades.

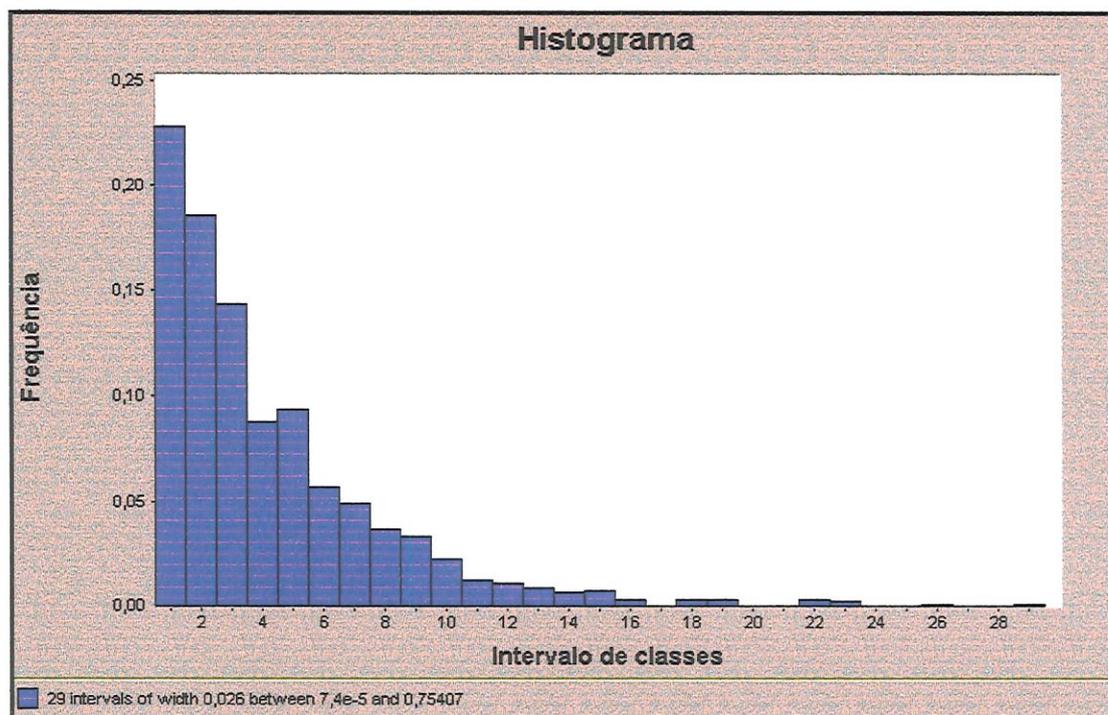


FIGURA 3.3 - Histograma

3.1.5. SELEÇÃO DA DISTRIBUIÇÃO DE PROBABILIDADE REPRESENTATIVA PARA OS DADOS COLETADOS

Com base no formato do histograma, deve-se atribuir uma distribuição de probabilidade teórica para representar a amostra dos dados. Em geral, as distribuições utilizadas são as mais conhecidas. Nesta etapa, a decisão sobre qual distribuição pode representar os dados é baseada somente no formato do histograma, sendo irrelevante os parâmetros da distribuição. Logo, é necessário ter conhecimentos sobre os possíveis formatos das distribuições para atribuir uma hipótese viável. A seguir, apresenta-se algumas importantes distribuições de probabilidade:

Beta: é aplicada em modelos onde são conhecidos somente os limites de mínimo, máximo, a moda e o valor mais provável. Sendo uma de suas principais aplicações a representação de proporções ou frações. Como exemplo, pode-se citar: proporção de elementos defeituosos em lotes de produtos. Esta distribuição possui dois parâmetros, e fornece uma ampla variedade de formas.

Erlang: é usada para modelar fenômenos aleatórios que podem ser descritos como uma série de processos independentes com distribuições exponenciais. Sendo aplicada também na modelagem de tempos de reparos ou tempos decorridos entre falhas. Esta distribuição que é uma extensão da distribuição exponencial, e é um caso especial da distribuição Gama.

Exponencial: é amplamente aplicada para modelar os tempos entre eventos independentes. Possui grande variabilidade, não sendo portanto recomendada para caracterizar os tempos de serviços. A principal característica desta distribuição é sua falta de memória. Isto significa que a probabilidade de chegada ou de uma partida em algum intervalo de

tempo futuro não é afetada pela história passada do processo. Ou seja, a cadeia de acontecimentos que ocorreram antes do sistema ter atingido o estado atual, não tem influência nenhuma sobre o evento futuro.

Gama: é uma generalização da distribuição Erlang, sendo aplicada em situações semelhantes as da distribuição Erlang.

Lognormal: é aplicada na modelagem de processo que podem ser analisados como o produto de vários outros processo, como por exemplo, a taxa de retorno de um investimento, quando a variável de interesse é um valor composto, como o produto dos lucros para vários períodos.

Normal: é a mais importante distribuição de probabilidade contínua no campo da estatística. Seu gráfico tem o formato de um sino, descreve fenômenos simétricos em torno da média. Sua importância se concentra no teorema do limite central que estabelece que a soma ou a média resultante de um grande número de valores aleatórios e independentes é aproximadamente normal, independente da distribuição dos valores individuais. Com base neste teorema, pode-se por exemplo, agregar os tempos de inúmeros sub-processos independentes, somando-os e substituindo-os por um único valor (soma ou média) cujo resultado tende a normalidade na medida que cresce o número de sub-processos.

Poisson: é aplicada para modelar o números de eventos independentes que ocorrem em um intervalo de tempo contínuo. Esta distribuição é particularmente apropriada na modelagem de chegadas independentes. Este processo de chegadas é denominado de processo de Poisson.

Triangular: é aplicada quando somente os limites de mínimo, máximo e moda (ou valor mais provável) são conhecidos.

Uniforme contínua: é aplicada quando as únicas informações sobre o comportamento da variável são os limites de máximo e mínimo. Esta distribuição permite gerar valores entre 0 e 1. Função indispensável para a aplicação dos inúmeros métodos numéricos destinados a geração de variáveis aleatórias.

Weibull: Esta distribuição que possui uma grande variedade de formatos, é aplicada principalmente para modelar falhas de sistemas ou componentes.

Estas distribuições são apresentadas detalhadamente no apêndice A1.

Existem alguns procedimentos que auxiliam o modelador a selecionar a hipótese de uma distribuição, tais como: *P-P Plot*, *Q-Q Plot*, e o *Box Plot*

- ♦ *P-P Plot*: compara graficamente a função de distribuição do modelo com a função de distribuição da amostra, onde $X[i]$ valores da amostra se encontram em ordem crescente. Se a distribuição selecionada for uma boa representação, o gráfico apresentará uma linearidade da função densidade com a reta que intercepta $x = 0$ com inclinação 1. A FIGURA 3.4 apresenta o *P-P Plot* quando um modelo de distribuição de probabilidade é uma boa representação.

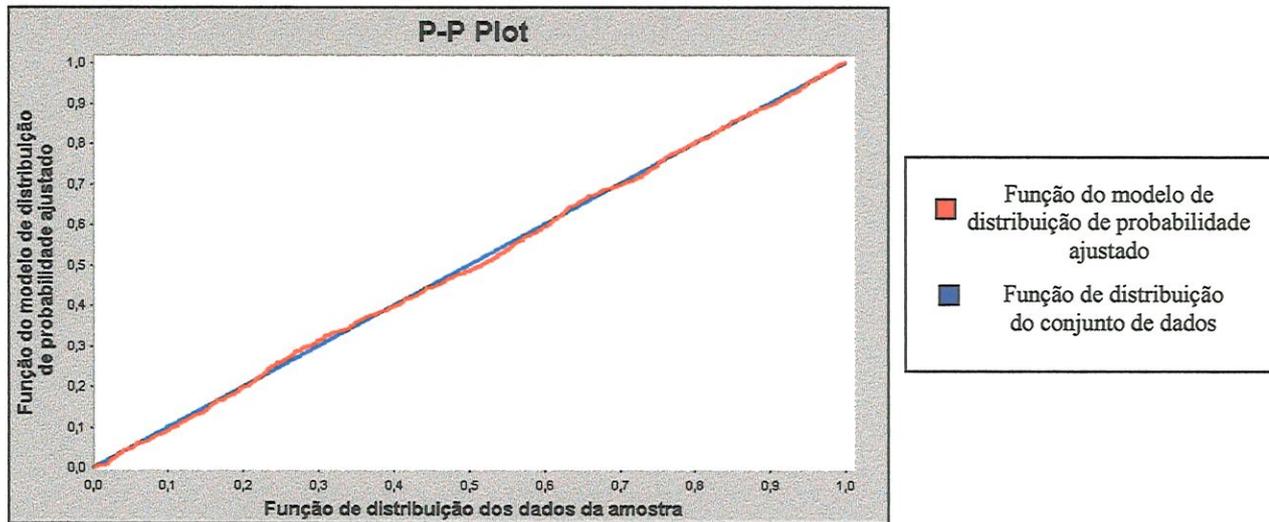


FIGURA 3.4 – PP PLOT quando uma distribuição de probabilidade é uma boa representação

A FIGURA 3.5 apresenta o *P-P Plot* quando um modelo de distribuição de probabilidade não é uma boa representação.

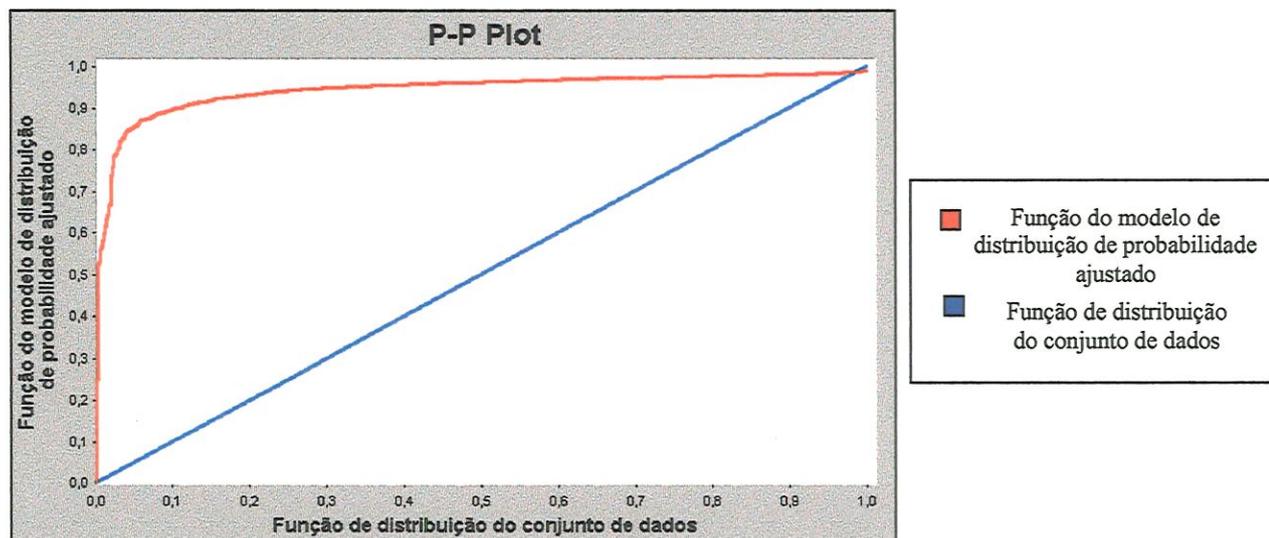


FIGURA 3.5 – P-P PLOT quando uma distribuição de probabilidade não é uma boa representação

O *P-P Plot* amplia as diferenças existentes entre as médias da função distribuição selecionada e da amostra (LAW & KELTON, 1991).

- ♦ *Q-Q plot*: É um gráfico do 100º percentil da função de distribuição selecionada com o 100º percentil da amostra, limitado no alcance de 0 e 1. De forma análoga ao *P-P Plot*, se a distribuição selecionada for uma boa representação, o gráfico mostrará uma linearidade da função densidade com a reta que intercepta $x = 0$ com inclinação 1. A FIGURA 3.6 apresenta o *Q-Q Plot* quando um modelo de distribuição de probabilidade é uma boa representação.

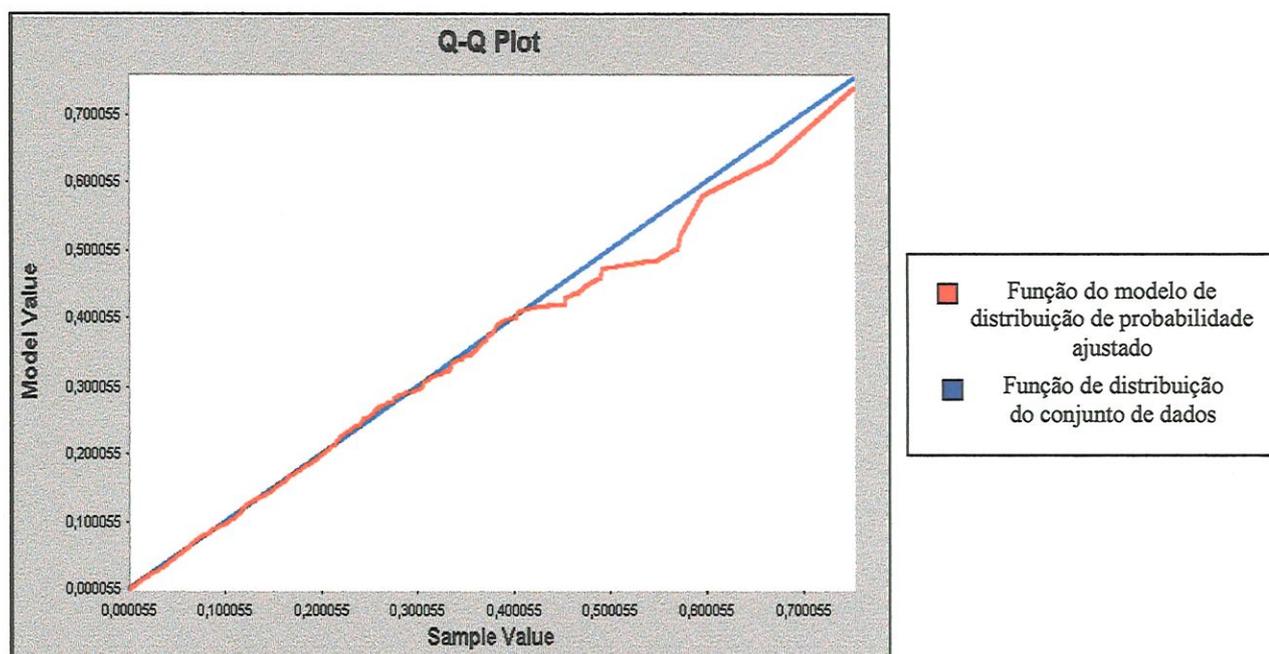


FIGURA 3.6 – Q-Q PLOT quando uma distribuição de probabilidade é uma boa representação para o conjunto de dados

A FIGURA 3.7 apresenta o *Q-Q Plot* quando um modelo de distribuição de probabilidade não é uma boa representação.

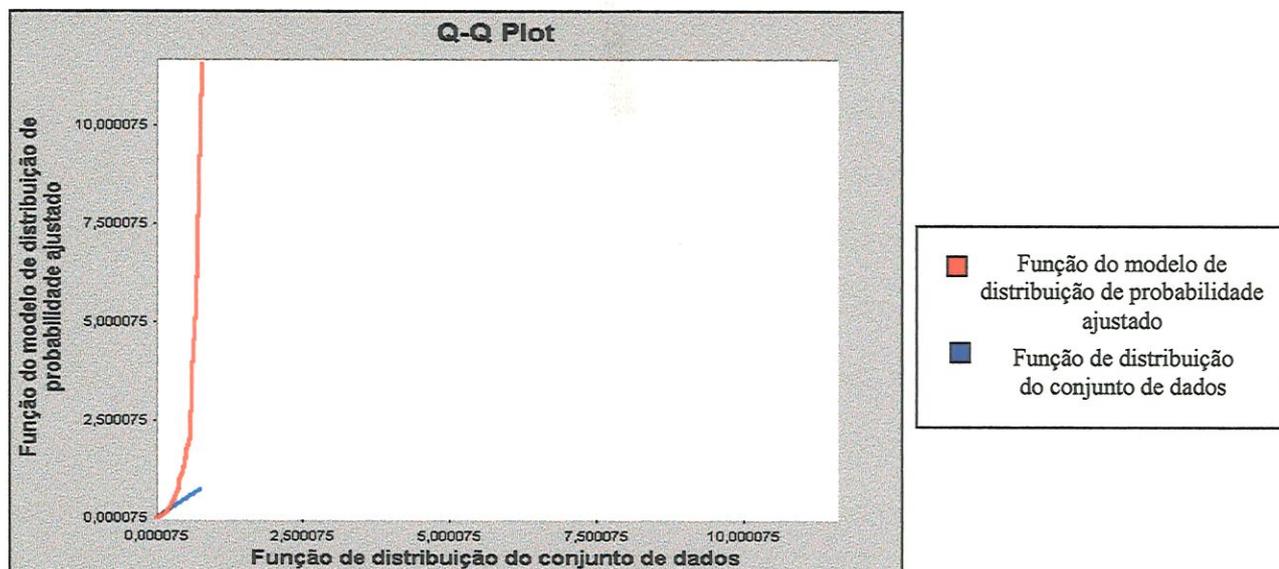


FIGURA 3.7 – Q-Q PLOT quando uma distribuição de probabilidade não é uma boa representação para o conjunto de dados

Uma limitação deste método é que este só pode ser aplicado a dados reais. O *Q-Q Plot* amplia as diferenças existentes entre as caudas da função de distribuição selecionada e da amostra (BANKS et al., 1996).

- ♦ *Box plot*: De acordo com VOELKER & ORTON (1993), este método gráfico permite verificar se a distribuição é simétrica, ou se esta possui uma cauda a direita ou a esquerda. O *Box Plot* é criado com os seguintes valores: limite inferior da amostra, 1º quartil, 1º octil, mediana, 3º quartil, 7º octil e o limite superior da amostra. Calcula-se os pontos médios entre : o 1º quartil e 3º quartil, que será denominado de “*média 1*” para melhor compreensão; o 1º octil e 7º octil, denominado de “*média 2*”; e entre os limites da amostra denominado de “*média 3*”. Se os valores entre a mediana e as “*médias 1, 2 e 3*” nesta respectiva ordem forem aproximadamente iguais, a distribuição apresenta simetria, logo é a normal. Caso estes valores

apresentem uma discrepância crescente entre os valores, a distribuição possui uma cauda a direita e em caso contrário, apresenta uma cauda a esquerda. A FIGURA 3.8 mostra o *Box Plot* de um conjunto de dados com cauda a direita.



FIGURA 3.8 – BOX PLOT do conjunto de dados e da distribuição de probabilidade ajustada

Apesar deste método auxiliar a descartar algumas hipóteses de distribuições de probabilidade, selecionando outras; deve-se ressaltar que somente esta técnica não é o suficiente para garantir que a distribuição selecionada seja uma boa representação. Segundo LAW & KELTON (1991) este método só deve ser aplicado em funções contínuas.

3.1.6. IDENTIFICAR OS PARÂMETROS DA DISTRIBUIÇÃO DE PROBABILIDADE

Após especificar algumas hipóteses sobre as distribuições de probabilidade que podem representar o conjunto de dados, deve-se determinar seus respectivos parâmetros. Os parâmetros podem ser estimados de duas maneiras: por ponto e por intervalo. A estimativa de parâmetro por ponto, consiste em determinar uma única estimativa para o parâmetro procurado. Enquanto que a estimativa por intervalo consiste em determinar um intervalo, no qual o parâmetro deve estar contido. De acordo com LAW & KELTON (1991), uma estimativa é um valor numérico de uma função analítica que se baseia nos dados da amostra. A esta função analítica denomina-se de estimador.

Existem diversas maneiras de se determinar os estimadores, tal como o método do MLE. Este método consiste em determinar, através da função de densidade da amostra em relação ao parâmetro θ desconhecido, um ponto máximo que fornecerá a estimativa de máxima verossimilhança do parâmetro θ . Em geral, as estimativas podem ser facilmente encontradas desde que a função não tenha mais de um parâmetro. Em caso contrário, os cálculos se tornam mais complicados, sendo as vezes necessário a aplicação de métodos iterativos e com auxílio de computadores.

3.1.7. VERIFICAÇÃO DA REPRESENTAÇÃO DA DISTRIBUIÇÃO DE PROBABILIDADE SELECIONADA

Conhecendo-se as hipóteses sobre a distribuição de probabilidade que podem representar o conjunto de dados coletados e seus respectivos parâmetros, deve-se verificar se a distribuição selecionada fornece uma boa representação para a amostra. Segundo LOBÃO & PORTO (1997), existem diversos procedimentos que permitem auxiliar o modelador a analisar se uma boa representação foi obtida, entre as quais pode-se citar :

- ♦ *testes de ajuste (ou teste de aderência)* : consiste em verificar analiticamente se a distribuição hipotética selecionada para a amostra fornece uma boa representação. Existem os seguintes testes: *teste do Qui-quadrado (χ^2)* Smirnov (COSTA NETO,1977), *teste de Kolmogorov-*

Smirnov (COSTA NETO,1977), *teste de Anderson-Darling* (LAW & KELTON, 1991), *teste de Turing* (SCHRUBEN, 1980).

- ♦ *comparação das freqüências dos dados de entrada com os da distribuição de probabilidade pelo histograma* : verifica visualmente, através do histograma de freqüência a semelhança existente entre a amostra e a distribuição hipotética.
- ♦ *comparação do formato da distribuição de probabilidade selecionada com o histograma dos dados*: verifica visualmente, através da função densidade da amostra e da distribuição hipotética a semelhança existente entre ambas distribuições.

Em geral, o método utilizado para avaliar a representação da distribuição selecionada, são os testes de ajuste. Visto que estes testes são de fácil aplicação e permitem avaliar analiticamente a representação fornecida pela distribuição hipotética. Deve observar que os testes de ajuste só podem ser aplicados em dados *iid* (LAW & KELTON,1991).

Quando as distribuições de probabilidade hipotéticas selecionadas não fornecem uma boa representação para a amostra de dados, deve-se construir uma distribuição empírica para implementar os dados no modelo de simulação. Uma distribuição empírica permite gerar os eventos de um modelo, dentro de um conjunto de dados limitados pelo maior e menor valor da amostra. O inconveniente desta proposta é que esta representação se restringe a este intervalo de dados, não possuindo uma representação para o comportamento da variável fora destes limites. Além disto, os dados também podem se encontrar em pequena quantidade, o que tornará a distribuição empírica pouco representativa.

3.1.8. VALIDAÇÃO DOS DADOS

A entrada de dados de um modelo tem um grande impacto sobre os dados de saída que serão gerados pela simulação e, conseqüentemente, sobre a validação do

modelo analisado. Uma etapa de fundamental importância em um projeto de simulação, que visa aumentar a acuracidade dos dados de entrada em um modelo, e respectivamente sua validação, é o processo de validação dos dados.

Validar os dados consiste em verificar a consistência da distribuição de probabilidade representativa ao conjunto de dados, implementada em um modelo.

Segundo LOBÃO & PORTO (1997), existem diversas formas de validar os dados, entre as quais podemos citar:

- ♦ Confrontar dados de diferentes fontes e verificar a semelhança existente entre estes;
- ♦ Verificar, realizando a primeira corrida de simulação com os dados coletados, se os mesmos produzem resultados que não sejam compatíveis com a realidade;
- ♦ Teste de Turing;
- ♦ Análise de sensibilidade.

Entre os diversos procedimentos existente, em geral, o processo de validação dos dados de entrada é realizado pela análise de sensibilidade. Esta análise permite ao modelador identificar a variável, parâmetro, distribuição ou subsistema, que possui maior ou menor impacto nos resultados do modelo. Determina ainda as variáveis que necessitam de uma maior precisão, direcionando assim novos esforços para a coleta de novos dados. Sua lógica consiste em verificar se grandes variações ocorrem nos dados de saída, através da “perturbação” e/ou modificações dos valores de entrada do modelo (IMAN,1981). Se isto não for observado, a distribuição de probabilidade representativa aos dados de entrada e o modelo de simulação são ditos válido. Em caso contrário, uma outra representação do valor de entrada deve ser encontrada.

Embora a análise de sensibilidade seja importante para o modelador obter uma melhor compreensão sobre o comportamento do modelo de simulação. A falta de variação nos dados de saída não implica que tenha sido definida uma boa representação para o conjunto de dados de entrada. Visto que, o modelo pode apresentar nenhuma

variação na variável de saída em estudo. E apresentar por exemplo, uma grande variação em alguma outra variável do sistema, tal como no aumento no tempo de fila de entidades no sistema. Ou seja, o aparecimento de gargalos. Além disto, de acordo com KELTON & VICENT (1992), uma vez que uma distribuição de probabilidade tenha sido definida como representativa aos dados de entrada, que não apresente variação nos resultados do modelo. Uma outra forma mais simples para representar este conjunto, que também não apresente variação, pode ser obtido.

3.1.9. DOCUMENTAÇÃO DOS DADOS DO MODELO

Todas informações referentes aos dados que foram incluídos e excluídos do modelo, devem ser registrados na documentação do modelo (USER'S GUIDE DA PROMODEL – "PROMODEL" , 1994)). LAW & KELTON (1991) advertem que estas informações servem como referências tanto para modelos que podem ser construídos posteriormente, como para análise da simulação.

Uma das maiores dificuldades para se realizar a modelagem de dados com coleta direta, é justamente coletar os dados do sistema em estudo. Em geral, isto ocorre porque os sistemas são muito complexos, sendo portanto necessário um elevado tempo, recursos e uma equipe de profissionais para se realizar a coleta de dados.

Embora a modelagem por coleta direta, em geral, forneça uma maior qualidade nos resultados de uma simulação. Essa, muitas vezes, não pode ser realizada devido a restrição de prazo e/ou recurso. Além, é claro, das próprias limitações físicas, como por exemplo, o fato da fábrica ou processo (que esta sendo modelado) ainda estar na fase de projeto. Em geral, diante destes problemas a modelagem de dados é realizada via coleta indireta dos dados. Na próxima seção discutiremos como proceder em tais situações.

3.2. MODELAGEM ESTATÍSTICA DOS DADOS DE ENTRADA COM REALIZAÇÃO DE UMA COLETA DE DADOS INDIRETA, PARA MODELOS UNIVARIÁVEIS

A *modelagem estatística dos dados de entrada com a realização de uma coleta indireta*, é caracterizada pela impossibilidade de coletar os dados da variável a ser implementada no modelo. O problema da ausência de dados, muito comum entre os profissionais da simulação, dificulta a especificação de uma distribuição de probabilidade que represente o comportamento da variável. Em geral, a impossibilidade de coletar os dados de um sistema, ocorre nos seguintes casos:

- ♦ O sistema real a ser modelado se encontra em fase de projeto;
- ♦ Tempo existente para realizar a simulação é muito restrito, não permitindo portanto a realização da coleta de dados;
- ♦ Sistema real a ser modelado é inacessível para se realizar a coleta de dados e
- ♦ Sistema muito complexo.

A distribuição de probabilidade, devido a inexistência de uma amostra de dados, é definida com base em estimativas hipotéticas sobre o comportamento da variável. Estas estimativas, que são obtidas de fontes secundárias, são informações, tais como: média, tempo mínimo de uma operação, tempo máximo de uma operação, tempo mais provável de uma operação (*moda*), variância e outras. Segundo BANKS et al (1996), HARREL et al (1992), LAW & KELTON(1991) e LOBÃO & PORTO (1997), estas estimativas podem ser obtidas das seguintes fontes:

- ◆ Entrevistas com profissionais da área, sendo recomendado neste caso, cautela ao considerar dados provenientes de pessoas intimamente envolvidas com o processo;
- ◆ Especificações de equipamentos;
- ◆ Dados de sistemas similares;
- ◆ Experiência do operador; e outros.

BANKS & GIBSON(1998) aconselham cautela ao aplicar estimativas na modelagem de dados, pois em geral estas não fornecem grande acuracidade sobre o verdadeiro comportamento da variável. Para melhor exemplificar este fato, considere o problema a seguir: seja KX um novo tipo de máquina a ser implementada em um modelo de simulação, e considere que o tempo de reparo seja desconhecido. Devido a impossibilidade de coletar estes dados, as seguintes informações foram obtidas com o fornecedor A da máquina: tempo mínimo necessário para reparar a máquina é de 30 minutos e o tempo máximo de 2 horas. Deve-se observar que o fornecedor A sempre apresentará estimativas otimistas sobre o comportamento da máquina, procurando desta forma ressaltar as qualidades do equipamento. Considere também que o Sr. M., que trabalha diretamente com a máquina, forneça as seguintes informações: tempo mínimo de reparo é igual a 60 minutos e o tempo máximo de 4 horas. Surge uma dúvida, qual informação é a mais próxima do verdadeiro tempo de reparo da máquina? Deve-se observar que em geral, pessoas são fontes de informações pouco confiáveis. A mente humana frequentemente consegue se lembrar muito bem sobre um fato do presente, mas se esquece facilmente de outro fato que ocorreu no passado.

Deste modo, diante de problemas deste tipo deve-se obter novas informações, de forma a procurar alcançar o comportamento da variável mais próximo possível de seu comportamento real. Em geral, a variável é implementada no modelo, com base nestas estimativas, e na experiência e bom senso do profissional de simulação.

De um modo geral, a modelagem de dados com coleta indireta, pode ser elaborada nas seguintes etapas:

- i) Identificar as variáveis que serão implementadas no modelo de simulação, e determinar a variável de interesse para atingir os objetivos do estudo de simulação
- ii) Coletar informações e/ou estimativas sobre o comportamento da variável que será implementada no modelo;
- iii) Especificar uma distribuição de probabilidade com base nas estimativas e/ou informações obtidas;
- iv) Validação dos dados e
- v) Documentação dos dados.

Como pode-se observar, comparando os dois tipos de modelagem tem-se que as etapas (i), (iv) e (v) são idênticas as etapas (3.1.1), (3.1.8) e (3.1.9) do processo de modelagem dos dados com a realização de uma coleta direta. A etapa (ii) é similar a etapa (3.1.2), com a diferença de que é realizado uma coleta de informações sobre a variável. Enquanto que a etapa (iii), que substitui as etapas (3.1.3), (3.1.4), (3.1.5), (3.1.6) e (3.1.7) fornece a principal diferença entre os dois tipos de modelagem. Diante desta fato, será analisado somente a etapa (iii) da modelagem de dados com a realização de uma coleta indireta.

3.2.1. ESPECIFICANDO UMA DISTRIBUIÇÃO DE PROBABILIDADE NA AUSÊNCIA DE UMA AMOSTRA DE DADOS

Em geral, diante da ausência de dados, as distribuições de probabilidade frequentemente aplicadas são: triangular, uniforme e beta. Visto que estas distribuições permitem gerar uma função, dentro de um intervalo $[a,b]$, onde $a < b$; através do conjunto de números pseudo-aleatórios existente na maioria dos softwares de simulação. Os limites “a” e “b” de um intervalo que podem ser interpretados como o menor e o maior tempo de uma operação, são estimativas obtidas de fontes secundárias. Para os sistemas de manufatura, estas estimativas frequentemente representam o tempo padrão da operação. Um tempo padrão é o tempo gasto para se executar uma operação no ciclo

de manufatura (AGOSTINHO,1985). Para casos onde somente as estimativas de mínimo e máximo são determinadas, a distribuição aplicada é a uniforme. Entretanto, como esta distribuição fornece pouca informação sobre o comportamento da variável, sendo portanto aconselhável obter mais uma estimativa da variável em estudo, tal como: tempo mais provável de operação (moda, caracterizado por “c”) ou a média, e aplicar a distribuição triangular.

Uma vez que as estimativas de mínimo, máximo, média e *moda* são definidas tem-se a caracterização da distribuição Beta, cujos parâmetros α_1 e α_2 com $\alpha_2 > \alpha_1$, podem ser calculados pelas seguintes equações:

$$\alpha_1 = \frac{(\mu - a) * (2 * c - a - b)}{(c - \mu) * (b - a)} \quad e \quad \alpha_2 = \frac{(b - \mu) * \alpha_1}{\mu - a} \quad (3.4)$$

Deve-se observar que para $\alpha_1 = \alpha_2 = 1$, a distribuição beta é um caso especial, pois representa a distribuição uniforme.

Ressalta-se que entre estas três distribuições, a distribuição beta é a mais utilizada para representar o comportamento de uma variável, pois possui uma maior variedade de formatos. Isto pode ser confirmado pela sua aceitação nos softwares estatísticos, como o *Expertfit* e o *Bestfit*.

Uma outra estratégia para representar o comportamento de uma variável, quando não se possui uma amostra de dados, é recorrer aos conhecimentos estatísticos teóricos, tais como, os processos de Poisson e/ou teoria de filas.

O processo de Poisson é um processo de contagem $\{N(t), \text{ para todo } t \geq 0 \}$, que registra o número de eventos que ocorrem durante um intervalo de tempo $[0,t]$, com uma taxa de chegada constante $\lambda > 0$. É de fundamental importância para os modelos de simulação de sistemas de manufatura e de serviços, pois estes sistemas são tipicamente direcionados para a entrada de elementos tal como: chegada de clientes, ordens de produção, materiais, informações e outros. Este processo satisfaz as seguintes condições (BANKS et al ,1996):

- ♦ a entrada de elementos no sistema no intervalo de tempo t e $(t + s)$ depende somente do tamanho do intervalo s , e não do instante t ;
- ♦ os elementos que chegam no sistema são independentes, ou seja, a ocorrência da entrada de uma grande quantidade de elementos em um intervalo de tempo t , não tem influência sobre a quantidade de elementos que entrarão no sistema no próximo intervalo;
- ♦ a possibilidade de dois ou mais elementos chegarem num mesmo intervalo de tempo é muito pequena, quando comparado à de um único elemento.

Os intervalos de tempo entre chegadas sucessivas de entidades em um sistema, obedecem a distribuição exponencial com média $1/\lambda$, onde $\lambda > 0$ é a taxa de entrada constante por unidade de tempo. É importante ressaltar que a distribuição exponencial não tem memória, ou seja, satisfaz a propriedade de esquecimento de Markov. Segundo (CLARKE & DISNEY, 1979), esta propriedade estabelece que a probabilidade de chegada ou de uma partida, em algum intervalo de tempo futuro, não é afetada pela história passada do processo. Ou seja, a cadeia de acontecimentos que ocorreram antes do sistema ter atingido o estado atual, não tem influência nenhuma sobre um evento futuro.

Deve-se observar que os tempos entre chegadas é independente do tipo de sistema em estudo. Logo, a chegada aleatória de elementos em qualquer tipo de sistema, pode ser representada pelo Processo de Poisson.

Por tal motivo, é muito comum a aplicação da distribuição exponencial para representar o comportamento de uma variável de chegada, quando não existem uma amostra para se realizar a modelagem de dados de entrada. Visto que o processo de Poisson, se fundamenta em condições matemáticas que se adaptam a maioria dos sistemas oferecendo uma boa aproximação a muitas situações, além de ser de fácil aplicação. Já que existe a necessidade de estimar apenas o valor de λ para utilizá-lo.

Diversos trabalhos que se beneficiaram do processo de Poisson podem ser encontrados na literatura. Como exemplo, pode-se citar o modelo de um *job-shop*

desenvolvido por LAW & KELTON(1991), cuja entrada de materiais obedece ao processo de Poisson. E o estudo de NELSON(1959) apud MARTIN & BRADFORD (1993), que ao avaliar um *job-shop* com 15 máquinas, determinou que na maioria dos casos, o tempo entre chegadas e o tempo de operação são um processo de Poisson. Para os sistemas de produção em série, tem-se o trabalho desenvolvido por HUNT(1956) apud CONWAY (1988) que analisa o efeito da variação do tempo de processamento em um sistema com 2 e 3 estágios, com *buffers* finitos e infinitos.

O processo de Poisson também é fundamental para alguns modelos de filas. Estes modelos, que são estudados e avaliados pela teoria de filas, são constituídos basicamente por elementos de entrada e de serviço de atendimento. A teoria de filas analisa um sistema com base no comportamento de entrada de elementos e o tempo de atendimento destes, através de estimativas, tais como: tamanho de fila, número de elementos em espera, proporção de utilização de servidores e outras. Os modelos de fila do tipo M/M/1, M/Erk/1, M/G/1, M/D/1 , cujo intervalo de entrada de elementos no sistema é um processo de Poisson, e possuem os tempos de atendimento representados respectivamente pelas seguintes distribuições de probabilidade: Exponencial, Erlang, Generalizada e Determinística. Com base nesta teoria, diversos trabalhos foram realizados, entre os quais pode-se citar o modelo M/M/1 de um sistema de manufatura desenvolvido por MARIA (1997). E o modelo M/Erk/1 também de uma manufatura analisado por HILLIER & BOLING (1966,1967) apud CONWAY (1988).

Apesar das distribuições citadas acima serem mais comuns para representar o tempo de atendimento, pode-se encontrar outras distribuições de probabilidade, tais como a distribuição normal e a distribuição Weibull. Isto foi analisado por BARTEN (1969) apud CONWAY (1988) em um modelo de manufatura que avalia o efeito da capacidade de *buffers* idênticos em linhas balanceadas, com tempo de processamento Normal. Com relação a distribuição Weibull tem-se o trabalho de CARNALL & WILD (1976) apud CONWAY (1988), que analisaram a variabilidade de tempos de processo com a aplicação desta distribuição.

Deve-se entretanto observar que, as análises obtidas nos estudos de simulação apresentados, dependem fortemente da condição de que o sistema modelado obedece as condições de um processo de Poisson.

Ressalta-se que, embora os tempos entre chegadas e os tempo de processamento sejam, em geral, variáveis aleatórias estocásticas. É possível encontrar casos onde estas variáveis são representadas por um tempo constante, como por exemplo: a entrada de peças em uma linha de montagem que se movimenta a uma velocidade constante. E o tempo do processo de soldagem automatizado, em uma linha do mesmo tipo, BANKS et al.(1996). Estes tempos são denominados de tempo-padrão. Segundo AGOSTINHO (1985), “nas operações de usinagem, os tempos de máquina podem ser normalmente determinados por equações conhecidas de usinagem dos metais. Onde deduções análogas podem ser feitas para as operações de fresamento, furação, alargamento, mandrilamento e outras. Os tempos mais difíceis de serem determinados são os tempos manuais, tal como os tempos de *setup*, visto que os tempos necessários para o operador perfazer o seu trabalho não podem ser matematicamente calculados.”

Outro evento importante, que deve ser considerado na modelagem de dados dos sistemas de manufatura, é a ocorrência de quebras ou paradas de máquinas. Em geral, a ocorrência de quebra e/ou parada de uma máquina está associado diretamente ao seu tempo de uso. Ou seja, quanto maior o tempo de utilização do equipamento maior a possibilidade de ocorrer uma quebra. Este tipo de problema, denominado de “problema de substituição”, engloba equipamentos cujo custo de manutenção aumenta com o decorrer do tempo; enquanto seu valor de venda no mercado diminui. Um outro tipo de quebra de máquina é aquele que ocorre de forma aleatória. Ou seja, a máquina funciona eficientemente até que falha repentinamente. Este tipo de problema, denominado de problema de reposição é fator de preocupação da manutenção preventiva. Esta manutenção procura substituir o equipamento antes que ele falhe. Tentando, desta forma, evitar o prejuízo que ocorre com sua ausência (FUNDAÇÃO CARLOS A. VANZOLINI, 1997). As quebras e/ou paradas de máquinas, geralmente ocorrem de acordo com um ciclo, composto por dois intervalos de tempo: *uptime* – é o intervalo de tempo no qual a máquina está em pleno funcionamento. Este intervalo é composto pelo

tempo de processamento mais o tempo no qual a máquina fica esperando uma peça para ser processada. E *downtime* – é o intervalo de tempo que se inicia no instante de quebra de uma máquina, e vai até o instante em que esta retorna a funcionar. Este intervalo de tempo é composto pelo tempo de espera do técnico para realizar o conserto mais o tempo de reparo.

Para modelar a ocorrência de quebras aleatórias de equipamentos, é necessário determinar a distribuição de probabilidade representativa para os tempos médio de reparo (MTTR - *Mean Time To Repair*) e o tempo médio entre as quebras do equipamento (MTBF – *Mean Time Between Failure*). Entre as diversas distribuições de probabilidade existentes, as mais aplicadas são: Weibull, Geométrica, Exponencial e Gama. Isto pode ser observado no trabalho de KAY apud BRADFORD & MARTIN (1993), que determina a distribuição Weibull como a mais representativa para os ciclos de falha de máquina. De modo análogo, JAIN & HANCOCK apud BRADFORD & MARTIN (1993) também desenvolveram um trabalho, no qual utilizam a distribuição de Poisson para representar os tempos de entre falhas e tempo de reparo. Já BUZACOTT (1967,1968) apud CONWAY (1988) e YAMASHINA & OKAMURA(1983) apud CONWAY (1988), avaliaram modelos de manufatura com a utilização da distribuição geométrica na representação dos tempos de reparo de equipamentos. Deve-se observar que para os softwares estatísticos *Expertfit* e *Bestfit*, os tempos entre falhas e tempo de reparo são sempre representadas pela distribuição Gama.

Após se determinar uma distribuição de probabilidade para representar o comportamento dos dados, deve-se verificar a acuracidade dos mesmos; através do processo de validação dos dados. HARREL et al. (1992) ressalta a importância deste processo, principalmente para os casos onde uma distribuição hipotética está sendo implementada no modelo.

3.3. CONSIDERAÇÕES FINAIS DO CAPÍTULO

Neste capítulo, procurou-se sistematizar os procedimentos mais usuais que são aplicados no modelamento dos dados de entrada. Tendo como finalidade, auxiliar os profissionais de simulação a implementá-los num modelo, da forma mais confiável possível. Ressalta-se que não é nosso objetivo aqui propor uma nova metodologia, visto que muito deste assunto podem ser encontrados na literatura da área. Mas sim, apresentá-los de uma forma organizada e coerente.

Os procedimentos apresentados foram abordados em dois tópicos principais: modelagem estatística dos dados de entrada com a realização de uma coleta direta, e modelagem estatística dos dados de entrada com a realização de uma coleta indireta.

Para o processo de modelagem dos dados com a realização de uma coleta direta, as etapas apresentadas se concentram basicamente em procurar determinar uma distribuição de probabilidade padrão, para o conjunto de dados em estudo. Este processo, embora seja trabalhoso, apresenta uma maior precisão na representação dos dados no modelo.

Já o processo de modelagem dos dados de entrada com coleta indireta, devido a ausência de uma amostra de dados, é um dos maiores desafios para os profissionais de simulação. Visto que resulta em uma grande dificuldade na definição de uma distribuição de probabilidade para representar a variável no modelo. Embora este tipo de modelagem não forneça a melhor precisão nos resultados do modelo. Ressalta-se que este tipo de modelagem é bastante comum, visto que, em geral os sistemas a serem analisados pela simulação são de grande porte, o que dificulta bastante a realização da coleta de dados direta.

Diante deste contexto, é possível observar uma maior aplicação de algumas distribuições de probabilidade em relação a outras. No entanto, deve-se ressaltar que as particularidades de cada variável de um modelo são totalmente dependentes das particularidades do sistema em estudo. Não sendo portanto possível determinar uma distribuição de probabilidade, que forneça uma aproximação ao comportamento de uma



variável, de modo a abranger todas as particularidades de todos os sistemas de manufatura existentes.

CAPÍTULO 4

CASO PARTICULAR DE UM MODELO DE SIMULAÇÃO DE UMA LINHA DE PRODUÇÃO

O modelo de simulação em estudo é uma linha de produção semi-automática composta por 3 estações de trabalho. A FIGURA 4.1 apresenta o layout do sistema. A primeira estação, denominada grupo de máquinas 1, é composto por 3 máquinas idênticas, cujo tempo de processamento de cada máquina obedece uma distribuição Gama (2; 0,25) minutos. A estação 2, denominada grupo de máquinas 2, é composta por 2 máquinas idênticas, possui o tempo de processamento de acordo com uma distribuição Gama (2; 0,05) minutos. E a estação 3, uma estação de inspeção, possui 2 operários, que trabalham conforme uma distribuição triangular (0,3; 1,2). O modelo possui ainda 10 AGVS, 7 esteiras e 3 *buffers* com capacidade de 10 peças.

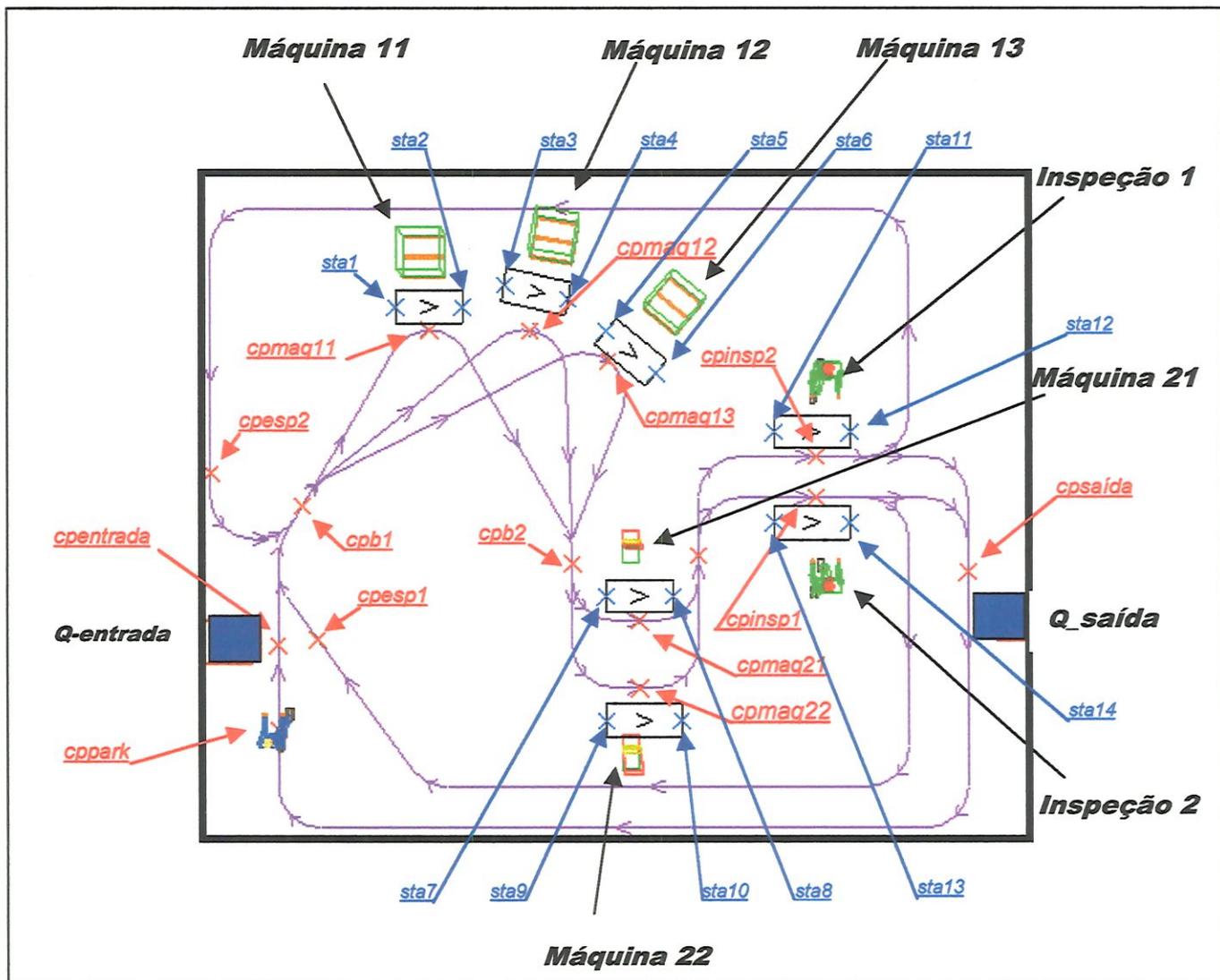


FIGURA 4.1 – Layout da Linha de Produção

A tabela 4.1 apresenta as estações de trabalho do sistema.

TABELA 4.1 – Definição dos *control-points* das estações de trabalho, das esteira e filas do sistema

CONTROL-POINT DAS ESTAÇÕES DE TRABALHO DO SISTEMA			
Estação de entrada das cargas do sistema <i>cpentrada</i>			
Buffers do sistema	<i>cpb1</i>	<i>cpb2</i>	<i>cpb3</i>
Grupo de máquinas 1	<i>cpmaq11</i>	<i>cpmaq12</i>	<i>cpmaq13</i>
Grupo de máquinas 2	<i>cpmaq21</i>	<i>cpmaq22</i>	
Estações de inspeção	<i>cpinsp 1</i>	<i>cpinsp2</i>	
Estações de espera para iniciar reprocesso da carga	<i>cpesp1</i>	<i>cpesp2</i>	
Estação de parada do agv	<i>cppark</i>		
Estação de saída das cargas do sistema <i>cpsaída</i>			
ESTAÇÕES DAS ESTEIRAS DO SISTEMA			
Grupo de máquinas 1			
Máquina 11	<i>sta1</i>	<i>sta2</i>	
Máquina 12	<i>sta3</i>	<i>sta4</i>	
Máquina 13	<i>sta5</i>	<i>sta6</i>	
Grupo de máquinas 2			
Máquina 21	<i>sta7</i>	<i>sta8</i>	
Máquina 22	<i>sta9</i>	<i>sta10</i>	
Grupo de inspeção 1			
Inspeção1	<i>sta11</i>	<i>sta12</i>	
Inspeção2	<i>sta13</i>	<i>sta14</i>	
FILAS			
Entrada	<i>Q-entrada</i>		
Saída	<i>Q-saída</i>		

O estudo de simulação neste sistema tem como objetivo verificar e analisar a variação que pode ocorrer na sua capacidade de produção, quando a distribuição do conjunto de dados de entrada é alterada. Para isto, foi considerado que o sistema modelado é um projeto, ou seja, não se encontra implementado. E conseqüentemente, não existe possibilidades de coletar dados para as variáveis do modelo. Deste modo as distribuições de probabilidade implementadas no modelo são hipóteses obtidas de fontes secundárias, sobre o comportamento de cada variável. Ou seja, não possuem uma grande acuracidade.

O modelo foi implementado no software *AUTOMOD*. O sistema trabalha com um turno diário de 8 horas. A simulação foi corrida no período de 2 meses, ou seja, 480 horas. O modelo opera da seguinte forma, as cargas são geradas infinitamente em *Q_entrada*, chegam no sistema por *cpentrada*, se este estiver desocupado. Caso *cpentrada* esteja ocupado, a carga fica aguardando sua liberação. Em *cpentrada*, a lista de ordens da simulação avança para o programa *PB1*, que verifica se o buffer *cpb1* possui capacidade para receber novas cargas. Se $cpb1 \geq 10$, a carga aguarda a liberação deste *control point*. Caso $cpb1 < 10$, a carga é deslocada até este *control point* iniciando o programa *PV1*. Este programa, verifica a existência de máquinas livres no grupo de máquinas 1. Se todas as máquinas estiverem ocupadas, a carga fica novamente aguardando a liberação de uma máquina livre para se locomover. Quando existir alguma máquina livre, a carga é deslocada para o respectivo *control point* desocupado. A peça é movida para a esteira, onde é processada de acordo com a distribuição Gama (2; 0,25) minutos. Ao término deste processo, a lista de ordens da simulação vai para o programa *PB2*, que de forma idêntica a *PB1*, verifica a capacidade do buffer *cpb2* de receber novas cargas. Se *cpb2* estiver livre, as cargas são deslocadas para este *control point* e a lista de ordens da simulação é direcionada para o programa *PV2*. Este programa, idêntico a *PV1*, verifica a existência de máquinas livres no grupo de máquinas 2. Caso todas as máquinas estejam ocupadas, a carga fica esperando a liberação de pelo menos uma das máquinas deste grupo. Existindo alguma máquina livre no grupo 2, a carga é deslocada até o respectivo *control point* desocupado. A peça é movida para a esteira, onde é processada conforme uma distribuição Gama (2; 0,05) minutos. Novamente, a

lista de operações da simulação é direcionada para o programa seguinte, *PB3*. Que de forma análoga aos outros programas dos *buffers*, verifica a capacidade de *cpb3* de receber novas cargas. Identicamente aos outros processos, a carga é deslocada para *cpb3* somente quando este *control point* estiver desocupado. De *PB3*, a simulação segue para o programa *PV3*, que desloca a carga para o operário em *cpinsp1* ou em *cpinsp2*, se este estiver livre. No processo de inspeção, 90% das cargas são aceitas e tem seu processos finalizados. E portanto saem do sistema em estudo, em *cpsaída*, enquanto 10% são rejeitadas e são reprocessadas. As cargas rejeitadas por *cpinsp1* são enviadas para *cpesp1* e as cargas rejeitadas por *cpinsp2* são enviadas para *cpesp2*. Em *cpinsp1* e *cpinsp2*, a simulação segue para *PB1*, iniciando novamente o processamento da carga.

Ressalta-se a existência de 3 prioridades de operações no decorrer do processamentos das cargas no modelo. A primeira delas ocorre no programa *PENTRADA*. A carga em *cpentrada* só é enviada para *cpb1* se *cpesp1* e *cpesp2* estiverem desocupados. Esta prioridade foi definida, devido ao fato de que as peças para serem reprocessadas são em geral, mais rápidas para atingir a finalização do seu processo do que as novas peças que entram no sistema. No entanto, com a finalidade de facilitar nossos estudos, as peças são reprocessadas com o mesmo tempo de processamento de uma nova peça que entra no sistema. A segunda prioridade que ocorre em *PESP1* e *PESP2*, determina que uma carga que se encontra em *cpesp1* só é disparada para *cpb1*, se *cpesp2* estiver desocupado. De forma idêntica, uma carga em *cpesp2* só é enviada para *cpb1*, se *cpesp1* estiver desocupado. Assim, quando uma carga ao se deslocar para *cpesp1* e encontrar *cpesp2* ocupado, esta carga ficará aguardando a liberação de *cpesp2* para prosseguir para *cpb1*. O mesmo ocorre para uma carga que se desloca para *cpesp2* e encontra *cpesp1* ocupado. Esta prioridade foi criada para ordenar o deslocamento das cargas para *cpesp1* e outra para *cpesp2*. Garante-se o desta forma que a primeira carga que alcança o *control point* de espera, irá iniciar o reprocessamento primeiro. A última prioridade, também definida em *PESP1* e *PESP2*, e é executada em *PINSP*, consiste em garantir que uma carga reprocessada não seja reprocessada novamente, enviando-a para *cpsaída*.

O tempo de carregamento e descarregamento da carga para cada máquina está incluído no tempo de processamento. Em *PSAIDA*, foi criado um programa para armazenar o tempo de espera de cada carga, diante de cada recurso, assim como o tempo de utilização dos recursos. Este programa ainda, visando facilitar a análise dos resultados, calcula o tempo médio de fila do sistema para cada um dos recursos e o tempo médio de utilização. Além de fornecer o número de cargas que são processadas no período do estudo de simulação.

4.1. PROCEDIMENTOS DE ANÁLISE DOS RESULTADOS GERADOS PELO MODELO DE SIMULAÇÃO

O modelo de simulação analisado é do tipo terminativo, visto que, existe um tempo limite para o tamanho da corrida de simulação. Para este estudo, foi considerado que o estado inicial do sistema é vazio, a simulação é corrida sob condições iniciais idênticas e o sistema a ser modelado se encontra em fase de projeto.

Devido ao fato do sistema em estudo não estar implementado, para avaliar os efeitos que a distribuição dos dados de entrada possui sobre os resultados de um modelo de simulação, inicialmente foi considerado uma distribuição hipotética para representar estes dados no modelo. A distribuição de probabilidade considerada, que foi selecionada com base na teoria do processo de Poisson é a distribuição exponencial com parâmetro $\lambda = 0,1$ (10 cargas por minuto). Este parâmetro foi definido à partir do estudo de um modelo teórico de simulação.

Desta forma, admitindo a distribuição Exponencial(0,1) como a distribuição ideal para representar os dados de entrada no modelo, utilizou-se o software estatístico *EXPERTFIT* para gerar uma amostra de dados com chances de representar a variável de entrada no modelo.

Para determinar esta listagem de distribuições foi gerado um total de 30 amostras com 900 dados. A listagem de distribuições selecionada foi aquela que forneceu o parâmetro λ mais próximo de 0,1 da distribuição exponencial.

Esta listagem de distribuições fornecida pelo *EXPERTFIT* é apresentada na TABELA 4.2. As distribuições se encontram no formato do software *AUTOMOD*.

TABELA 4.2 – Listagem das distribuições fornecida pelo software Expertfit - PARTE I

LISTAGEM DAS DISTRIBUIÇÕES FORNECIDAS PELO EXPERTFIT

1 – Erlang(1;0,09846)

2 – Exponential (0;09846)

3 – Gamma (0,99053;0,09940)

4 - Weibull 0,99358;0,09819

5 - Erlang(E) 0,000075 + gamma 1,000000; 0,098381

6 - Exponential(E) 0,000075 + exponential 0,098381

7 - Weibull(E) 0,000075 + weibull 0,986767; 0,097823

8 - Gamma(E) 0,000075 + gamma 0,974578; 0,100947

9 - Pearson Type 6

First Do:

set <temp1> to (gamma 1,059178; 1,)

set <temp2> to (gamma 11,716026; 1,)

set <temp3> to <temp1>/<temp2>

Then Use:

<temp3>

Where:

<temp1>, <temp2>, and <temp3> are temporary real variables

10 - Pearson Type 6(E)

First Do:

set <temp1> to (gamma 1,040795; 1,)

set <temp2> to (gamma 11,529870; 1,)

set <temp3> to <temp1>/<temp2>

Then Use: 0,000075 + <temp3>

Where:

<temp1>, <temp2>, and <temp3> are temporary real variables

11 - Log-Logistic

First Do:

set <temp> to EXP(LOG(1/(uniform 0,5;0,5)-1,0)/1,432809)

Then Use:

*0,061921 * (<temp>)*

Where:

<temp> is a temporary real variable

TABELA 4.2 – Listagem das distribuições fornecida pelo software Expertfit - PARTE II

12 - Log-Logistic(E)

First Do:

set <temp> to EXP(LOG(1/(uniform 0,5;0,5)-1)/1,416992)

Then Use:

*0,000075 + 0,061761 * (<temp>)*

Where: <temp> is a temporary real variable

13 - Log-Laplace(E)

Use:

0,000075 + EXP(RVLaplace(-2,730123;0,978679;<stream>))

Note: Use of this distribution is explained in the UAutoMod.C file, which is available using the Simulation Software Files item in the Files menu.

14 - Log-Laplace

Use:

EXP(RVLaplace(-2,728979;0,967813;<stream>))

Note: Use of this distribution is explained in the UAutoMod.C file, which is available using the Simulation Software Files item in the Files menu.

15 - Lognormal -2,901543; 1,292922

16 - Random Walk

Use:

1,/RVInvGauss(0,;77,615752;11,686179;<stream>)

Note: Use of this distribution is explained in the UAutoMod.C file, which is available using the Simulation Software Files item in the Files menu

17 - Lognormal(E) 0,000075 + lognormal -2,912989; 1,351381

18 - Random Walk(E)

Use:

0,000075 + 1,/RVInvGauss(0,;5784,026030;10,182504;<stream>)

Note: Use of this distribution is explained in the UAutoMod.C file, which is available using the Simulation Software Files item in the Files menu.

19 - Inverted Weibull 1/(weibull 0,601487; 36,797182) 0;02718

20 - Pearson Type 5 1/(gamma 0,445927; 174,054848)

TABELA 4.2 – Listagem das distribuições fornecida pelo software Expertfit - PARTE III**21 - Wald***Use:* $RVInvGauss(0,000000;1,000000;0,013208;<stream>)$ *Note:*

Use of this distribution is explained in the UAutoMod.C file, which is available using the Simulation Software Files item in the Files menu.

22 - Inverted Weibull(E) $0,000075 + 1./(weibull\ 0,406634; 38,942529)$ **23 - Inverse Gaussian***Use:* $RVInvGauss(0,000000;0,098455;0,014824;<stream>)$ *Note:*

Use of this distribution is explained in the UAutoMod.C file, which is available using the Simulation Software Files item in the Files menu.

24 – Rayleigh $weibull\ 2,000000; 0,140189$ **25 - Rayleigh(E)** $0,000075 + weibull\ 2,000000; 0,140137$ **26 - Pareto(E)***First Do:* $set\ <temp>\ to\ EXP(-9,508227-LOG(uniform\ 0,5;0,5)/0,151362)$ *Then Use:* $set\ <temp>\ to$ *Where:* $<temp>\ is\ a\ temporary\ real\ variable$ **27 - Pearson Type 5(E)** $0,000075 + 1./(gamma\ 0,135612; 42651,186930)$ **28 - Wald(E)***Use:* $RVInvGauss(0,000075;1,000000;0,000173;<stream>)$

Note: Use of this distribution is explained in the UAutoMod.C file, which is available using the Simulation Software Files item in the Files menu.

29 - Inverse Gaussian(E)*Use:* $RVInvGauss(0,000075;0,098381;0,000173;<stream>)$

Note: Use of this distribution is explained in the UAutoMod.C file, which is available using the Simulation Software Files item in the Files menu.

Com base na listagem de distribuições fornecida pela TABELA 4.2, pode-se observar que a distribuição que mais se aproxima da Exponencial (0,1) é a Exponencial (0,09846). Esta distribuição que aparece em segundo lugar na listagem de distribuições, tem o mesmo formato que a distribuição Erlang (1;0,09846).

Para avaliar a influência do “input” dos dados de entrada no modelo de simulação selecionou-se as seguintes distribuições da TABELA4.2 :

2 – Exponencial (0,09846)

3 – Gama (0,99053;0,09940)

4 – Weibull (0,99358;0,09819)

15 – Lognormal (-2,901543; 1,292922)

Além das distribuições citadas acima, foi também avaliada a distribuição Weibull (0,5;1) e a constante igual a 10 peças/minuto que não se encontram na listagem de distribuições. Pretendeu-se com isto ressaltar as diferenças que podem ocorrer ao selecionar uma distribuição que não fornece uma boa representação. Deve-se ressaltar que o software de simulação *AUTOMOD* possui as distribuições de probabilidade Constante, Uniforme, Normal, Triangular, Exponencial, Lognormal, Gama e Weibull para serem implementadas diretamente no modelo. Para implementar qualquer outra distribuição neste software é necessário criar um programa lógico para o modelo.

Com base nas distribuições selecionadas, executou-se a simulação num total de 10 replicações para cada uma das distribuições. O método aplicado para estimar os resultados, foi o método das replicações independentes. Ou seja, foi utilizado um conjunto de “stream” diferente em cada replicação, garantindo desta forma a independência entre as replicações.

A TABELA 4.3 apresenta a semente aplicada na entrada dos dados e em cada recurso do sistema, para cada replicação da simulação.

TABELA 4.3 – Semente do conjunto de números aleatórios aplicado em cada replicação

REPLICACÃO	ENTRADA	M11	M12	M13	M21	M22	INSP1	INSP2
1	100	231	345	412	590	678	989	888
2	231	345	412	590	678	989	888	100
3	345	412	590	678	989	888	100	231
4	412	590	678	989	888	100	231	345
5	590	678	989	888	100	231	345	412
6	678	989	888	100	231	345	412	590
7	989	888	100	231	345	412	590	678
8	888	100	231	345	412	590	678	989
9	770	220	170	150	999	745	832	000
10	000	832	745	999	150	170	220	770

Onde:

Entrada - Semente usada na entrada dos dados

M11, M12, M13 - Semente usada no grupo de máquinas 1 para as máquinas 11, 12 e 13

M21, M22 - Semente usada no grupo de máquinas 2, para as máquinas 21 e 22

INSP1, INSP2 - Semente usada na estação de inspeção, para as inspeções 11 e 12

Deve-se observar que as sementes apresentadas na TABELA 4.3 são modificadas de uma replicação para outra. No entanto, o conjunto de números pseudo-aleatórios é o mesmo para todas as distribuições analisadas em cada replicação. Desta forma, foi

possível aplicar a proposta da amostra correlacionada, para comparar os resultados gerados do modelo de simulação

4.2. CONSIDERAÇÕES FINAIS DO CAPÍTULO

A finalidade deste estudo de simulação foi testar o quanto pode influenciar a escolha de uma distribuição de probabilidade em um modelo. Para isto, ressalta-se que o tipo do sistema a ser analisado pela simulação é irrelevante. Isto ocorre devido a impossibilidade de padronizar as distribuições de probabilidade para as variáveis de um sistema, com base no tipo do sistema. Visto que cada sistema é um caso particular, cujo comportamento da variável depende de todos os fatores que ocorrem no sistema, tal como: as ordens de produção e outros.

O modelo desenvolvido para análise foi suposto em fase de projeto, e assim sendo não permite a realização de uma coleta direta dos dados. Diante do problema da ausência de dado foi utilizado a teoria do Processo de Poisson, para determinar uma distribuição hipotética para os dados de entrada. A distribuição hipotética admitida como ideal foi a Exponencial (0,1). Com base nesta hipótese e com a utilização do software estatístico *EXPERTFIT*, foi gerado uma listagem de distribuições de probabilidade. Dentre as quais selecionou-se um *rank* de distribuições para ser avaliado.

Ressalta-se que sendo o modelo analisado hipotético, o comportamento das variáveis implementadas pode se encontrar muito distantes do verdadeiro comportamento da variável, quando o sistema for implementado.

CAPÍTULO 5

RESULTADOS E DISCUSSÕES

Para avaliar se as distribuições Exponencial (0,09846), Gama (0,99053;0,09940), Weibull (0,99358;0,09819), Lognormal (-2,901543; 1,292922) e Weibull (0,5;1) são representativas ao conjunto de dados de entrada, foram aplicados os seguintes testes de ajuste pelo software *EXPERTFIT*.

1) Teste de Anderson-Darling:

Exponencial (0,09846)						
Tamanho da amostra	900					
Estatística de teste	0,28078					
	Valores Críticos para o Nível de Significância					
	<i>0,25</i>	<i>0,1</i>	<i>0,05</i>	<i>0,025</i>	<i>0,01</i>	<i>0,005</i>
	0,736	1,061	1,32	1,59	1,958	2,243
Rejeitado?	Não					
Gama (0,09940; 0,99053)						
Tamanho da amostra	900					
Estatística de teste	0,28078					
	Valores Críticos para o Nível de Significância					
	<i>0,25</i>	<i>0,1</i>	<i>0,05</i>	<i>0,025</i>	<i>0,01</i>	<i>0,005</i>
	1,248	1,933	2,492	3,07	3,857	4,5
Rejeitado?	Não					
Weibull(0,09819;0,99358)						
Tamanho da amostra	900					
Estatística de teste	0,29043					
	Valores Críticos para o Nível de Significância					
	<i>0,25</i>	<i>0,1</i>	<i>0,05</i>	<i>0,025</i>	<i>0,01</i>	<i>0,005</i>
	0,471	0,633	0,752	0,871	1,031	
Rejeitado?	Não					

Lognormal(-2,90154; 1,29292)						
Tamanho da amostra	900					
Estatística de teste	0,28078					
	Valores Críticos para o Nível de Significância					
	<i>0,25</i>	<i>0,1</i>	<i>0,05</i>	<i>0,025</i>	<i>0,01</i>	<i>0,005</i>
	0,47	0,63	0,751	0,872	1,034	1,158
Rejeitado?	Sim					
Weibull(0,5;1)						
Tamanho da amostra	900					
Estatística de teste	424,93499					
	Valores Críticos para o Nível de Significância					
	<i>0,25</i>	<i>0,1</i>	<i>0,05</i>	<i>0,025</i>	<i>0,01</i>	<i>0,005</i>
	1,248	1,933	2,492	3,07	3,857	4,5
Rejeitado?	Sim					

2) Teste de Kolmogorov- Smirnov:

Exponencial (0,09846)						
Tamanho da amostra	900					
Estatística de teste	0,534					
	Valores Críticos para o Nível de Significância					
	<i>0,15</i>	<i>0,1</i>	<i>0,05</i>	<i>0,025</i>	<i>0,01</i>	
	0,924	0,988	1,091	1,186	1,303	
Rejeitado?	Não					
Gama (0,09940;0,99053)						
Tamanho da amostra	900					
Estatística de teste	0,58803					
	Valores Críticos para o Nível de Significância					
	<i>0,15</i>	<i>0,1</i>	<i>0,05</i>	<i>0,025</i>	<i>0,01</i>	
	1,133	1,219	1,352	1,474	1,621	
Rejeitado?	Não					
Weibull(0,09819;0,99358)						
Tamanho da amostra	900					
Estatística de teste	0,5966					
	Valores Críticos para o Nível de Significância					
	<i>0,1</i>	<i>0,05</i>	<i>0,025</i>	<i>0,01</i>		
infinity	0,803	0,874	0,939	1,007		
Rejeitado?	Não					

Lognormal(-2,90154;1,29292)						
Tamanho da amostra	900					
Estatística de teste	2,54297					
Valores Críticos para o Nível de Significância						
	0,15	0,1	0,05	0,025	0,01	
	0,775	0,818	0,894	0,994	1,034	
Rejeitado?	Sim					
Weibull(0,5;1)						
Tamanho da amostra	900					
Estatística de teste	16,12373					
Valores Críticos para o Nível de Significância						
	0,1	0,05	0,025	0,01		
	1,133	1,219	1,352	1,474	1,621	
Rejeitado?	Sim					

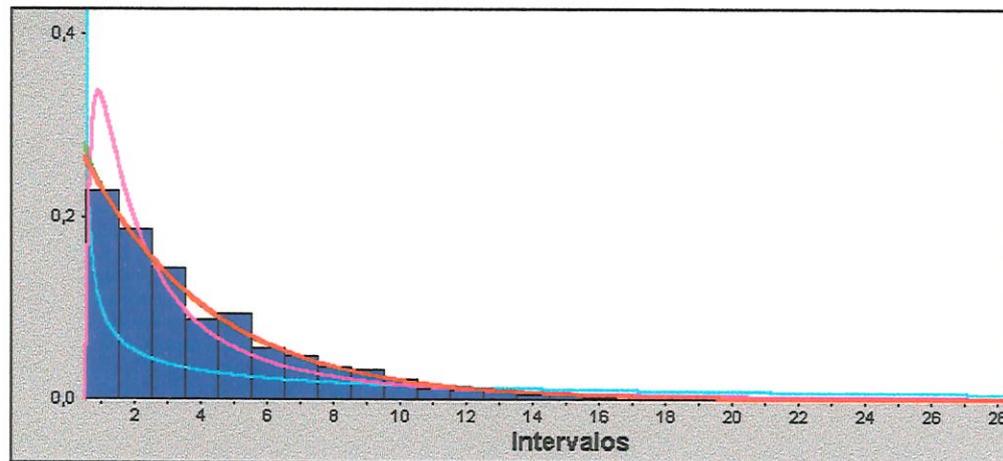
3) Teste do Qui-Quadrado:

Exponencial (0,09846)						
Número de intervalos	40					
Média esperada	22,5					
Estatística de teste	30,75556					
Graus de Liberdade	Nível de significancia	Valores Críticos para o Nível de Significancia				
		0,25	0,15	0,1	0,05	0,01
38	0,792	43,462	47,007	49,513	53,384	61,162
39	0,824	44,539	48,126	50,66	54,572	62,428
	Rejeitado?	Não				
Gamma (0,09940;0,99053)						
Número de intervalos	40					
Média esperada	22,5					
Estatística de teste	29,77778					
Graus de Liberdade	Nível de significancia	Valores Críticos para o Nível de Significancia				
		0,25	0,15	0,1	0,05	0,01
37	0,795	42,383	45,886	48,363	52,192	59,893
39	0,856	44,539	48,126	50,66	54,572	62,428
	Rejeitado?	Não				
Weibull(0,09819;0,99358)						
Número de intervalos	40					
Média esperada	22,5					
Estatística de teste	30,4					
Graus de Liberdade	Nível de significancia	Valores Críticos para o Nível de Significancia				
		0,25	0,15	0,1	0,05	0,01
37	0,77	42,383	45,886	48,363	52,192	59,893
39	0,836	44,539	48,126	50,66	54,572	62,428
	Rejeitado?	Não				

Lognormal (-2,90154;1,29292)						
Número de intervalos	40					
Média esperada	22,5					
Estatística de teste	136,53333					
Graus de Liberdade	Nível de significancia	Valores Críticos para o Nível de Significancia				
		0,25	0,15	0,1	0,05	0,01
37	0	42,383	45,886	48,363	52,192	59,893
39	0	44,539	48,126	50,66	54,572	62,428
	Rejeitado?	Sim				
Weibull(0,5;1)						
Número de intervalos	40					
Média esperada	22,5					
Estatística de teste	1.466,84					
Graus de Liberdade	Nível de significancia	Valores Críticos para o Nível de Significancia				
		0,25	0,15	0,1	0,05	0,01
39	0,836	44,539	48,126	50,66	54,572	62,428
	Rejeitado?	Sim				

Os testes de ajuste consistem em comparar a estatística de teste com o valor crítico, para determinar se uma boa aderência da distribuição de probabilidade foi encontrada. Uma estatística de teste é uma variável na qual é armazenado a hipótese H_0 a ser testada pelo teste de ajuste. Seu valor é calculado de acordo com o tipo do teste de ajuste aplicado e a distribuição de probabilidade e/ou o conjunto de dados da amostra. O valor crítico é definido pelo nível de significância α e pela quantidade de graus de liberdade da amostra de dados. O teste de ajuste rejeita uma hipótese H_0 quando o valor crítico tabelado é maior que a estatística de teste, ou seja, um bom ajuste não foi obtido. Em caso contrário, a hipótese H_0 não é rejeitada e portanto um bom ajuste foi obtido.

De acordo com os testes de ajuste aplicados nas distribuições selecionadas, pode-se observar que a distribuição Lognormal (-2,90154; 1,29292) e Weibull (0,5;1) são rejeitadas, ou seja, não fornecem uma boa representação para o conjunto de dados. Este fato pode ser comprovado pela FIGURA 6.1 apresenta o ajuste destas distribuições para o conjunto de dados gerados.



- | | |
|---|---------------------------------|
| ■ | 1. Exponencial (0.09846) |
| ■ | 2. Gamma (0.09940, 0.99053) |
| ■ | 3. Weibull (0.09819, 0.99358) |
| ■ | 4. Lognormal (2.90154, 1.29292) |
| ■ | 5. Weibull (0.5, 1) |

FIGURA 5.1 – Histograma com o ajuste das distribuições selecionadas para avaliação

Ainda analisando visualmente o histograma da FIGURA 5.1, pode-se observar que as distribuições Gama (0,09940; 0,99053) e Weibull (0,09819;0,99358) apresentam um comportamento muito próximo à Exponencial(0,09846). Enquanto que a distribuição Lognormal (2,90154; 1,29292) e Weibull (0,5;1) se encontram bem distantes do comportamento da variável.

Para avaliar a influência destas distribuições no modelo, considere os resultados gerados pela simulação em cada replicação. Estes resultados são apresentados nas TABELAS 5.2 à 5.8. Onde:

TABELA 5.1 – Dados de saída do modelo de simulação

<i>Número de cargas no sistema</i> - indica o número de peças produzidas
<i>T_m das cargas no sistema</i> - indica o tempo médio em que as cargas ficam no sistema
<i>T_m de espera no processo 1</i> - indica o tempo médio de fila para as cargas entrarem no sistema
<i>T_m de espera no processo 2</i> - indica o tempo médio de fila para as cargas entrarem no buffer 1
<i>T_m de espera no processo 3</i> - indica o tempo médio de fila para as cargas entrarem no grupo de máquinas 1
<i>T_m de espera no processo 4</i> - indica o tempo médio de fila para as cargas entrarem no buffer 2
<i>T_m de espera no processo 5</i> - indica o tempo médio de fila para as cargas entrarem no grupo de máquinas 2
<i>T_m de espera no processo 6</i> - indica o tempo médio de fila para as cargas entrarem no buffer 3
<i>T_m de espera no processo 7</i> - indica o tempo médio de fila para as cargas entrarem na inspeção
<i>T_m de espera no sistema</i> - indica o tempo médio das cargas no sistema
<i>T_m de utilização do recurso 1</i> - indica o tempo médio de utilização da máquinas do grupo 1
<i>T_m de utilização do recurso 2</i> - indica o tempo médio de utilização da máquinas do grupo 2
<i>T_m de utilização do recurso 3</i> - indica o tempo médio de permanência nas estações de inspeção
<i>T_m de utilização dos recursos no sistema</i> - indica o tempo médio de utilização da máquinas do grupo 1

*EXPONENCIAL (0,1)***TABELA 5.2 – Resultados da distribuição Exponencial(0,1)**

EXP(0.1)	1	2	3	4	5	6	7	8	9	10
Numero de cargas no sistema	674	669	670	668	689	675	677	671	685	674
T_m das cargas no sistema	10,691	10,732	10,699	10,785	10,451	10,627	10,635	10,689	10,512	10,651
T_m de espera no processo 1	0,149	0,149	0,149	0,149	0,149	0,149	0,149	0,1487	0,1487	0,149
T_m de espera no processo 2	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,0000	0,0000	0,000
T_m de espera no processo 3	0,199	0,219	0,178	0,188	0,198	0,205	0,188	0,2002	0,1847	0,203
T_m de espera no processo 4	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,0000	0,0000	0,000
T_m de espera no processo 5	0,220	0,228	0,227	0,223	0,235	0,235	0,235	0,2255	0,2307	0,219
T_m de espera no processo 6	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,0000	0,0000	0,000
T_m de espera no processo 7	0,601	0,589	0,610	0,596	0,599	0,596	0,603	0,5975	0,5880	0,597
T_m de espera no sistema	1,169	1,185	1,163	1,156	1,181	1,185	1,175	1,172	1,152	1,168
T_m de utilização do recurso 1	0,398	0,400	0,400	0,402	0,401	0,400	0,400	0,3985	0,3995	0,402
T_m de utilização do recurso 2	0,179	0,179	0,179	0,179	0,179	0,179	0,180	0,1792	0,1796	0,178
T_m de utilização do recurso 3	0,737	0,739	0,747	0,741	0,738	0,741	0,742	0,7429	0,7424	0,742
T_m de utilização dos recursos no sistema	1,314	1,318	1,326	1,322	1,317	1,321	1,322	1,321	1,322	1,321

*2. EXPONENCIAL (0,09846)***TABELA 5.3 – Resultados da distribuição Exponencial (0,09846)**

EXP(0.09846)	1	2	3	4	5	6	7	8	9	10
Numero de cargas no sistema	674	669	670	668	689	675	677	671	685	674
T_m das cargas no sistema	10,691	10,732	10,699	10,785	10,451	10,627	10,635	10,689	10,512	10,651
T_m de espera no processo 1	0,149	0,149	0,149	0,149	0,149	0,149	0,149	0,149	0,149	0,149
T_m de espera no processo 2	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000
T_m de espera no processo 3	0,199	0,219	0,178	0,188	0,198	0,205	0,188	0,200	0,185	0,203
T_m de espera no processo 4	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000
T_m de espera no processo 5	0,220	0,228	0,227	0,223	0,235	0,235	0,235	0,226	0,231	0,219
T_m de espera no processo 6	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000
T_m de espera no processo 7	0,601	0,589	0,610	0,596	0,599	0,596	0,603	0,598	0,588	0,597
T_m de espera no sistema	1,169	1,185	1,163	1,156	1,181	1,185	1,175	1,172	1,152	1,168
T_m de utilização do recurso 1	0,398	0,400	0,400	0,402	0,401	0,400	0,400	0,398	0,400	0,402
T_m de utilização do recurso 2	0,179	0,179	0,179	0,179	0,179	0,179	0,180	0,179	0,180	0,178
T_m de utilização do recurso 3	0,737	0,739	0,747	0,741	0,738	0,741	0,742	0,743	0,742	0,742
T_m de utilização dos recursos no sistema	1,314	1,318	1,326	1,322	1,317	1,321	1,322	1,321	1,322	1,321

3. *GAMA(0,09940;0,99053)*

TABELA 5.4 - Resultados da distribuição Gama(0,09940;0,99053)

GAMMA(0.09940,0.99053)	1	2	3	4	5	6	7	8	9	10
Numero de cargas no sistema	677	680	678	664	669	673	677	677	668	680
T_m das cargas no sistema	10,630	10,563	10,597	10,855	10,801	10,736	10,634	10,547	10,756	10,592
T_m de espera no processo 1	0,149	0,149	0,149	0,149	0,149	0,149	0,149	0,149	0,149	0,149
T_m de espera no processo 2	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000
T_m de espera no processo 3	0,184	0,191	0,188	0,217	0,217	0,204	0,188	0,193	0,221	0,226
T_m de espera no processo 4	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000
T_m de espera no processo 5	0,234	0,227	0,222	0,224	0,227	0,234	0,235	0,240	0,230	0,234
T_m de espera no processo 6	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000
T_m de espera no processo 7	0,594	0,595	0,600	0,602	0,592	0,591	0,603	0,600	0,596	0,595
T_m de espera no sistema	1,161	1,162	1,159	1,192	1,184	1,177	1,175	1,182	1,196	1,205
T_m de utilização do recurso 1	0,399	0,399	0,399	0,398	0,398	0,400	0,400	0,402	0,400	0,399
T_m de utilização do recurso 2	0,179	0,179	0,179	0,180	0,178	0,180	0,180	0,179	0,178	0,180
T_m de utilização do recurso 3	0,7350	0,7391	0,7496	0,745	0,738	0,748	0,742	0,741	0,735	0,746
T_m de utilização dos recursos no sistema	1,313	1,318	1,327	1,324	1,315	1,329	1,322	1,322	1,313	1,325

4. *WEIBULL(0,09819;0,99358)*

TABELA 5.5 – Resultados da distribuição Weibull(0,09819;0,99358)

WEIBULL(0.09819,0.99358)	1	2	3	4	5	6	7	8	9	10
Numero de cargas no sistema	677	669	670	668	689	675	677	671	685	674
T_m das cargas no sistema	10,630	10,732	10,699	10,785	10,451	10,627	10,635	10,689	10,512	10,651
T_m de espera no processo 1	0,149	0,149	0,149	0,149	0,149	0,149	0,149	0,149	0,149	0,149
T_m de espera no processo 2	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000
T_m de espera no processo 3	0,184	0,219	0,178	0,188	0,198	0,205	0,188	0,200	0,185	0,203
T_m de espera no processo 4	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000
T_m de espera no processo 5	0,234	0,228	0,227	0,223	0,235	0,235	0,235	0,226	0,231	0,219
T_m de espera no processo 6	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000
T_m de espera no processo 7	0,594	0,589	0,610	0,596	0,599	0,596	0,603	0,598	0,588	0,597
T_m de espera no sistema	1,161	1,185	1,163	1,156	1,181	1,185	1,175	1,172	1,152	1,168
T_m de utilização do recurso 1	0,399	0,400	0,400	0,402	0,401	0,400	0,400	0,398	0,400	0,402
T_m de utilização do recurso 2	0,179	0,179	0,179	0,179	0,179	0,179	0,180	0,179	0,180	0,178
T_m de utilização do recurso 3	0,735	0,739	0,747	0,741	0,738	0,741	0,742	0,743	0,742	0,742
T_m de utilização dos recursos no sistema	1,313	1,318	1,326	1,322	1,317	1,321	1,322	1,321	1,322	1,321

5. LOGNORMAL (-2,90154;1,29292)**TABELA 5.6 – Resultados da distribuição Lognormal (-2,90154;1,29292)**

LOGNORMAL (-2,90154,1,29292)	1	2	3	4	5	6	7	8	9	10
Numero de cargas no sistema	673	669	671	668	689	675	677	671	674	674
T_m das cargas no sistema	10,692	10,732	10,701	10,785	10,451	10,627	10,635	10,689	10,634	10,651
T_m de espera no processo 1	0,149	0,149	0,149	0,149	0,149	0,149	0,149	0,149	0,149	0,149
T_m de espera no processo 2	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000
T_m de espera no processo 3	0,199	0,219	0,178	0,188	0,198	0,205	0,188	0,200	0,193	0,203
T_m de espera no processo 4	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000
T_m de espera no processo 5	0,220	0,228	0,227	0,223	0,235	0,235	0,235	0,226	0,213	0,219
T_m de espera no processo 6	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000
T_m de espera no processo 7	0,601	0,589	0,610	0,596	0,599	0,596	0,603	0,598	0,595	0,597
T_m de espera no sistema	1,169	1,185	1,163	1,156	1,181	1,185	1,175	1,172	1,150	1,168
T_m de utilização do recurso 1	0,398	0,400	0,400	0,402	0,401	0,400	0,400	0,398	0,396	0,402
T_m de utilização do recurso 2	0,179	0,179	0,179	0,179	0,179	0,179	0,180	0,179	0,178	0,178
T_m de utilização do recurso 3	0,737	0,739	0,747	0,741	0,738	0,741	0,742	0,743	0,734	0,742
T_m de utilização dos recursos no sistema	1,314	1,318	1,326	1,322	1,317	1,321	1,322	1,321	1,309	1,321

6. WEIBULL (0,5; 1,0)**TABELA 5.7 - Resultados da distribuição Weibull (0,5; 1,0)**

WEIBULL (0.5, 1.0)	1	2	3	4	5	6	7	8	9	10
Numero de cargas no sistema	433	409	421	424	339	325	374	378	333	452
T_m das cargas no sistema	7,465	7,463	7,168	7,435	7,212	7,090	7,505	7,761	7,543	8,033
T_m de espera no processo 1	0,146	0,146	0,146	0,146	0,146	0,146	0,146	0,146	0,146	0,146
T_m de espera no processo 2	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000
T_m de espera no processo 3	0,218	0,215	0,222	0,221	0,217	0,192	0,224	0,220	0,218	0,230
T_m de espera no processo 4	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000
T_m de espera no processo 5	0,226	0,243	0,234	0,233	0,224	0,224	0,220	0,234	0,231	0,226
T_m de espera no processo 6	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000
T_m de espera no processo 7	0,583	0,572	0,567	0,584	0,568	0,576	0,578	0,567	0,580	0,571
T_m de espera no sistema	1,173	1,176	1,170	1,184	1,156	1,138	1,169	1,168	1,175	1,174
T_m de utilização do recurso 1	0,367	0,368	0,369	0,366	0,364	0,37	0,37	0,37	0,37	0,37
T_m de utilização do recurso 2	0,162	0,161	0,160	0,161	0,161	0,16	0,16	0,16	0,16	0,16
T_m de utilização do recurso 3	0,749	0,747	0,737	0,740	0,752	0,74	0,74	0,74	0,74	0,75
T_m de utilização dos recursos no sistema	1,277	1,276	1,265	1,266	1,277	1,270	1,275	1,268	1,278	1,276

7.CONSTANTE: 10 PEÇAS/MINUTO**TABELA 5.8 - Resultados da constante 10 peças/minuto**

CONSTANTE 10 PEÇAS/MINUTO	
Numero de cargas no sistema	80
Tempo médio das cargas no sistema	5,536597
Tempo médio de espera no processo 1	0
Tempo médio de espera no processo 2	0
Tempo médio de espera no processo 3	0
Tempo médio de espera no processo 4	0
Tempo médio de espera no processo 5	0
Tempo médio de espera no processo 6	0
Tempo médio de espera no processo 7	0
Tempo médio de espera no sistema	0
Tempo médio de utilização do recurso 1	0,497947
Tempo médio de utilização do recurso 2	0,101231
Tempo médio de utilização do recurso 3	1,109075
Tempo médio de utilização dos recursos no sistema	1,708253

Avaliando os resultados gerados pela simulação é possível observar que as distribuições Exponencial (0,09846) e Exponencial (0,1) possuem soluções idênticas para todas as variáveis do processo, e em todas as replicações. Logo a listagem de distribuições geradas para a distribuição hipotética ideal (Exponencial (0,1)) é válida. Deve-se observar também que a distribuição Gama (0,99053;0,09940), Weibull (0,99358;0,09819) e Lognormal (-2,90154, 1,29292) apresentam resultados muito próximo entre si. As estimativas para cada replicação, podem ser obtidas pela equação (A.2). Estas estimativas são por ponto, e portanto não fornecem uma boa precisão. Sendo portanto aconselhável construir um intervalo de confiança. A solução do modelo, ou seja, o resultado final para o qual o intervalo de confiança será construído pode ser calculado pela equação (A.3). A TABELA 5.9 apresenta os resultados gerados pelo modelo de simulação em cada replicação

TABELA 5.9 – Resultados do modelo de simulação

NUMERO DE PEÇAS PRODUZIDAS NO SISTEMA EM CADA REPLICAÇÃO					
	EXP 1	EXP 2	EXP 3	EXP 4	EXP 5
REPLICAÇÃO 1	674,00	677	677	433	673
REPLICAÇÃO 2	669,00	680	669	409	669
REPLICAÇÃO 3	670,00	678	670	421	671
REPLICAÇÃO 4	668,00	664	668	424	668
REPLICAÇÃO 5	689,00	669	689	339	689
REPLICAÇÃO 6	675,00	673	675	325	675
REPLICAÇÃO 7	677,00	677	677	374	677
REPLICAÇÃO 8	671,00	677	671	378	671
REPLICAÇÃO 9	685,00	668	685	333	674
REPLICAÇÃO 10	674,00	680	674	452	674
MÉDIA	675,20	674,3	675,5	388,80	674,1
VARIANCIA	47,51	30,67777778	47,61111111	2070,178	34,988889

Onde:

EXP1: Exponencial (0,09846)

EXP2: Gama(0,09940;0,99053)

EXP3: Weibull(0,09819;0,99358)

EXP4: Weibull(0,5;1)

EXP5: Lognormal(-2,90154;1,29292)

De acordo com os resultados apresentados na TABELA 5.9, para o número médio de peças produzidas com a aplicação das distribuições selecionadas pode-se observar que as distribuições Lognormal(-2,90154;1,29292) e Weibull (0,5;1) que foram rejeitadas pelos testes de ajuste, possuem um erro proporcional em relação a média da distribuição ideal de respectivamente 0,16% e 42,42%. Enquanto que a distribuição Gama(0,09940;0,99053) e Weibull(0,09819;0,99358), que não foram rejeitadas pelos testes de ajuste, apresentam um erro proporcional de respectivamente 0,1% e 0,04%. Desta forma, ao comparar o erro proporcional entre a distribuição Lognormal e a distribuição Gama, observa-se que a Lognormal embora tenha sido rejeitada apresenta uma estimativa para a média bem próxima da solução fornecida pela distribuição hipotética ideal. Deve-se ressaltar que esta comparação fornece apenas o erro percentual

entre as estimativas, não fornecendo uma garantia de que uma distribuição é mais representativa do que outra.

Para melhor avaliar os resultados obtidos realizou-se a comparação entre as diferenças do número médio de peças produzidas para cada uma das distribuições selecionadas em relação a distribuição Exponencial (0,09846). A TABELA 5.10 apresenta as diferenças encontradas em cada replicação.

TABELA 5.10 – Comparação entre as distribuições selecionadas com a distribuição Hipotética ideal

COMPARAÇÃO DAS DISTRIBUIÇÕES SELECIONADAS COM A DISTRIBUIÇÃO HIPOTÉTICA IDEAL				
	EXP1-EXP2	EXP1-EXP3	EXP1-EXP4	EXP1-EXP5
<i>REPLICAÇÃO 1</i>	-3,00	-3,00	241,00	1,00
<i>REPLICAÇÃO 2</i>	-11,00	0,00	260,00	0,00
<i>REPLICAÇÃO 3</i>	-8,00	0,00	249,00	-1,00
<i>REPLICAÇÃO 4</i>	4,00	0,00	244,00	0,00
<i>REPLICAÇÃO 5</i>	20,00	0,00	350,00	0,00
<i>REPLICAÇÃO 6</i>	2,00	0,00	350,00	0,00
<i>REPLICAÇÃO 7</i>	0,00	0,00	303,00	0,00
<i>REPLICAÇÃO 8</i>	-6,00	0,00	293,00	0,00
<i>REPLICAÇÃO 9</i>	17,00	0,00	352,00	11,00
<i>REPLICAÇÃO 10</i>	-6,00	0,00	222,00	0,00
<i>MÉDIA</i>	0,90	-0,30	286,40	1,10
<i>VARIANCIA</i>	13,71	0,10	1674,61	11,81

Comparando os resultados da TABELA 5.10 pela variação do número médio de peças produzidas, pode-se observar que a distribuição Weibull (0,09819;0,99358) (3^o colocada no *rank* de distribuições) fornece uma menor variação em relação a média da distribuição hipotética ideal (Exponencial (0,09846)). Sendo portanto a melhor representação para o comportamento dos dados de entrada, entre as distribuições analisadas. Isto contradiz o fato da distribuição Gama(0,09940;0,99053) (2^o colocada no *rank* de distribuições) ser a melhor representação ao conjunto de dados gerados. Além disto, pode-se observar nesta comparação que a Gama(0,09940;0,99053) apresenta uma variação maior do que a distribuição Lognormal, que foi rejeitada pelo *EXPERTFIT*. Em relação a distribuição Weibull (0,5;1) também rejeitada pelos testes de ajuste, verifica-se

que existe uma grande variação em relação a média da hipótese ideal, o que confirma o fato de não ser uma boa representação para o comportamento dos dados.

Com a finalidade de reavaliar os resultados obtidos, construiu-se o intervalo de confiança no nível de significância de 95%, para a diferença do número médio de peças produzidas entre as distribuições. Estes intervalos são apresentados na TABELA 5.11. Onde “a” e “b” são respectivamente os limites inferiores e superiores, ou seja, os pontos extremos do intervalo.

TABELA 5.11 – Intervalo de Confiança para a diferença entre o número médio de peças produzidas pelas distribuições

<i>INTERVALO DE CONFIANÇA</i>			
	Limite inferior	Média	Limite superior
<i>EXP1-EXP2</i>	-1,75	0,90	3,55
<i>EXP1-EXP3</i>	-0,526	-0,3	-0,074
<i>EXP1-EXP4</i>	257,154	286,400	315,646
<i>EXP1-EXP5</i>	-1,356	1,100	3,556

Pode-se verificar pelos resultados apresentados na TABELA 5.11 que o intervalo EXP1-EXP2 e EXP1-EXP5 não se encontram nem a direita e nem a esquerda do valor zero. Não fornecendo portanto nenhuma garantia sobre a melhor representação existente entre as distribuições analisadas. Para o intervalo EXP1-EXP4, que se encontra a direita do valor zero; temos que $EXP1-EXP4 > 0$. Logo $EXP1 > EXP4$, ou seja, a distribuição exponencial (0,09846) fornece uma maior quantidade de peças produzidas em relação à Weibull(0,5;1) quando implementada no modelo. Isto pode ser comprovado pelos resultados obtidos pela TABELA 6.9 De forma análoga, o intervalo EXP1-EXP3 se encontra à esquerda de zero, logo $EXP1-EXP3 < 0$. Ou seja, $EXP1 < EXP3$ e portanto o número de peças produzidas pela distribuição exponencial fornece uma menor quantidade de peça produzidas. Ressalta-se entretanto que o número de peças produzida em um sistema só admite valores inteiros positivos.

5.1. CONSIDERAÇÕES FINAIS DO CAPÍTULO

Foram realizados os seguintes testes para avaliar a influência dos dados de entrada no modelo de simulação:

- ◆ Histograma com o ajuste das distribuições de probabilidade para comparação visual;
- ◆ Testes de ajuste: Teste de Anderson-Darling, Teste de Kolmogorov- Smirnov e Teste do Qui-Quadrado;
- ◆ Comparação do erro proporcional entre o número médio de peças produzidas para cada uma das distribuições selecionadas com o número médio de peças produzidas pela distribuição hipotética ideal;
- ◆ Comparação da variação da diferença entre o número médio de peças produzidas para cada uma das distribuições selecionadas em relação a distribuição hipotética ideal;
- ◆ Comparação do intervalo de confiança no nível de 95% para a diferença do número médio de peças produzidas

Analisando os resultados gerados pela simulação através dos testes apresentados tem-se os seguintes resultados:

- ◆ Na análise do histograma pode-se observar que as distribuições Lognormal (2,90154; 1,29292) e Weibull (0,5;1) não se ajustam a amostra de dados;
- ◆ Em todos os testes de ajuste realizados as distribuições Lognormal (-2,90154; 1,29292) e Weibull (0,5;1) também são rejeitadas;
- ◆ Erro proporcional do número médio de peças produzidas em relação a distribuição hipotética ideal é de 0,1% e 0,04% para as distribuições Gama(0,09940;0,99053) e Weibull(0,09819;0,99358). Deve-se observar que estas distribuições não foram rejeitadas nos testes (a) e (b). Para as distribuições

Lognormal(-2,90154;1,29292) e Weibull (0,5;1) o erro proporcional foi respectivamente de 0,16% e 42,42%;

- ◆ Comparando a variância do número médio de peças produzidas tem-se os seguintes resultados:

Gama(0,09940;0,99053)	13,71
Weibull(0,09819;0,99358)	0,10
Lognormal(-2,90154;1,29292)	11,81
Weibull (0,5;1)	1674,61

Como pode ser observado a distribuição Weibull(0,09819;0,99358) fornece a menor variância do número médio de peças produzidas em relação a distribuição hipotética ideal.

- ◆ Teste do intervalo de confiança comprova os resultados gerados pelo modelo de simulação para as distribuições Weibull (0,09819;0,99358) e Weibull (0,5;1).

Analisando os resultados gerados pela simulação através dos testes apresentados pode-se concluir que a distribuição Weibull (0,5;1) e a Constante 10 peças/minuto não representam o comportamento da variável. Visto que foram rejeitadas em todos os testes realizados.

Já a distribuição Lognormal (-2,90154;1,29292) é uma hipótese de incerteza visto que é rejeitada nos testes (a) e (b), e no entanto fornece a variância do número médio de peças produzidas menor do que a distribuição Gama (0,09940;0,99053). Deve-se observar que a distribuição Gama (0,09940;0,99053) e a distribuição Weibull (0,09819;0,99358) são respectivamente a segunda e a terceira colocadas no *rank* de distribuições fornecidas pelo *Expertfit*. Ou seja, são as hipóteses mais representativas para o comportamento da variável. Isto pode ser verificado pela aceitação de ambas nas avaliações realizadas neste trabalho.

CAPÍTULO 6

CONCLUSÃO

Neste trabalho, procurou-se sistematizar os procedimentos mais usuais aplicados no processo de modelagem dos dados de entrada, em um modelo de simulação. Os procedimentos apresentados foram abordados em dois tópicos principais:

- ♦ modelagem estatística dos dados de entrada com a realização de uma coleta direta e
- ♦ modelagem estatística dos dados de entrada com a realização de uma coleta indireta.

Para o processo de modelagem dos dados com a realização de uma coleta direta, as etapas apresentadas se concentram basicamente em procurar determinar uma distribuição de probabilidade padrão, para o conjunto de dados em estudo. Este processo, embora seja trabalhoso, fornece maior confiabilidade sobre a correta representação matemática da variável a ser implementada no modelo.

Já o processo de modelagem de dados com coleta indireta apresenta um maior grau de incerteza na representação matemática da variável a ser implementada no modelo, devido a ausência de uma amostra de dados da variável em estudo. Em decorrência deste problema há maiores possibilidades de se introduzir erros durante a especificação de uma distribuição de probabilidade representativa à variável. Deve-se observar que embora isto ocorra, este tipo de modelagem é mais comum. Visto que, em geral os sistemas a serem analisados pela simulação são complexos, o que dificulta bastante a realização da coleta de dados direta. Devido a este fato, o presente trabalho teve como enfoque principal abordar a modelagem de dados com coleta indireta.

Em geral, no processo de modelagem com coleta indireta, é comum a aplicação de conhecimentos teóricos de estatística para justificar a utilização de determinadas distribuições de probabilidade, tais como: o Processo de Poisson e a Teoria de Filas. Com base nestas teorias pode-se observar uma maior utilização de algumas distribuições de probabilidade. Sendo, portanto, estas distribuições as primeiras hipóteses a serem testadas. Deste modo, de acordo com o levantamento realizado, tem-se as seguintes distribuições:

- ♦ Tempo entre chegadas : representado pela distribuição Exponencial quando as chegadas são aleatórias (quando a entrada de entidades no sistema é programada, utiliza-se a distribuição Constante);
- ♦ Tempo de processamento : de forma análoga a variável “tempo entre chegadas”, pode ser representada pela distribuição Exponencial e a distribuição Constante. Utiliza-se a distribuição Uniforme, Triangular ou Beta quando existem poucas informações sobre esta variável. A distribuição Erlang é aplicada, quando a operação a ser executada é subdividida em etapas, que devem ser realizadas para a operação ser finalizada. As distribuições: Gama, Normal e Weibull também são utilizadas neste caso;
- ♦ Tempo entre falhas : Pode ser representada pela distribuição Weibull. Nos softwares estatísticos *Expertfit* e *Bestfit*, a única distribuição de probabilidade representativa é a Gama.
- ♦ Tempo de reparo: Pode ser representada pela distribuição Beta. A distribuição Gama é a única representação aceita pelos softwares estatísticos *Bestfit* e *Expertfit*.

Para avaliar a influência dos dados de entrada nos resultados de um modelo de simulação, considere as seguintes avaliações realizadas no modelo em estudo:

- ◆ A distribuição Weibull (0,5;1) e Lognormal (-2,90154;1,29292) são rejeitadas na análise visual do histograma e em todos os testes de ajustes. Deve-se observar que a distribuição Lognormal é a décimo quinto do *rank* de distribuições e a distribuição Weibull (0,5;1) não se encontra nesta listagem;
- ◆ Embora a distribuição Lognormal (-2,90154;1,29292) tenha sido rejeitada nos testes mencionados no item (i) anterior. Esta distribuição fornece um diferença proporcional de 0,16% do número médio de peças produzidas em relação a distribuição hipotética ideal. Enquanto a distribuição Gama (0,09940;0,99053) e Weibull (0,09819;0,99358) que não foram rejeitadas, fornecem um diferença de 0,1% e 0,04%. Ou seja, o número médio de peças produzidas pela distribuição hipotética Lognormal (-2,90154;1,29292) e pela distribuição ideal Exponencial (0,1) são bem próximos. Logo, embora a distribuição Lognormal (-2,90154;1,29292) tenha sido rejeitada, esta distribuição fornece resultados compatíveis com o sistema em estudo;
- ◆ Já o mesmo não ocorre com a distribuição Weibull (0,5;1) que foi rejeitada nos testes do item (i) e fornece uma diferença de 42,42%. Enquanto a distribuição constante 10 peças/minuto fornece uma diferença de 88,15%;
- ◆ Com base nos resultados do modelo pode-se observar que existe mais de uma possível distribuição de probabilidade representativa para a variável. As distribuições são: Gama (0,09940;0,99053), Weibull (0,09819;0,99358) e Lognormal (-2,90154;1,29292);
- ◆ O modelo de simulação executou 10 replicações. Ressalta-se que quanto maior o número de replicações maior será a precisão dos resultados. Ou seja, o número de replicações executadas poderá influenciar os resultados do modelo.

Embora a modelagem de dados não seja o único fator a influenciar nos resultados de um modelo de simulação, é possível notar uma crescente preocupação voltada para este processo. Isto pode ser verificado através de experiências como o caso da Honda (ISHIHARA, 1999), que ao utilizar dados de seus sistemas produtivos atingiu melhorias em seus processos de produção.

De um modo geral, pode-se concluir que tanto a modelagem de dados com a realização de uma coleta indireta, quanto a modelagem com a realização de uma coleta direta são processos complexos e de difícil realização. Sendo portanto possível, mesmo com os devidos cuidados introduzir erros a partir de dados que são aparentemente corretos.

CAPÍTULO 7

REFERÊNCIAS BIBLIOGRÁFICAS

AGOSTINHO, O.L. (1985) . *Estudo da Flexibilidade dos Sistemas Produtivos*. São Carlos. 243p. Tese (Doutorado) – Escola de Engenharia de São Carlos, Universidade de São Paulo.

AGUIAR, A.F.S.; ROZENFELD,H.; RENTES, A.F.; BREMER, C.F.; ALLIPRANDINI, D.H. (1994). Integração da manufatura: o caminho para a modernização. *Revista Máquinas e Metais*, p.98-113, setembro.

ANNINO,J.S.;RUSSELL,E.C.(1981). The seven most frequent causes of simulation analysis failure – and how to avoid them. *Interfaces*,v.11,n.3,p.59-63,June.

ARENA SIMAN (1994). Reference guide. Systems Modeling Corporation, Sewickley, PA, USA.

BUZACOTT,J. (1972). The effect of station Breakdowns and random processing times on the capacity of flow lines with in-process storage. *AIIE Transactions*, v.4, n.4,p.308-312,Apr.

BANKS,J.; GIBSON,R.R.(1998). Simulation input Data. *IIE Solutions*,p.27-36

BANKS,J.;CARSON II,J.S.;NELSON,B.L.(1996). *Discrete-event system simulation*. Prentice Hall.

BANKS,J.;GIBSON,R. (1997a). Simulation modeling some programming required. *IIE Solutions*,p.26-31,Feb.

BUZACOTT,J.A.;SHANTHIKUMAR,J.G.(1993). *Stochastic models of manufacturing systems*. Prentice Hall, Englewood Cliffs, New Jersey

CARSON, J. S.(1992). Modeling. In: PROCEEDINGS OF THE 1992 WINTER SIMULATION CONFERENCE.p.83-87.

CLARK,A.B.;DISNEY,R.L.(1979).*Probabilidade e Processos Estocásticos*. Livros Técnicos e Científicos Editora.

CLARK, G.M.(1994). Introduction to manufacturing applications. In: PROCEEDINGS OF THE 1994 WINTER SIMULATION CONFERENCE. p.15-21.

CONWAY, R; et al (1988). The role of work-in-process inventory in serial production lines. *Operation Research*. v.36,n.2,Mar.

COSTA NETO, P.L. (1977).*Estatística*. Editora Edgard Blucher Ltda

E.MARIAN SCOTT,E.M.(1996). Uncertainty and sensitivity studies of models of environmental systems. IN: PROCEEDING OF THE 1996 WINTER SIMULATION CONFERENCE. p.255-259.

EXPERTFIT USER'S GUIDE - 1995

FRANCISCO, W.F. (1993). *Estatística Básica*. EDITORA UNIMEP

FUNDAÇÃO CARLOS ALBERTO VANZOLINI (1997). *Gestão de Operações*. Edt. Edgard Blucher Ltda.

GOGG,T.; MOTT,J.R.A.(1993). Introduction to simulation. In: PROCEEDING OF THE 1993 WINTER SIMULATION CONFERENCE. p.09-17.

GROSS,D.; JUTTIJUDATA,M. (1997) Sensitivity of output performance measures to input distributions in queueing simulation modeling.IN: PROCEEDING OF THE 1997 WINTER SIMULATION CONFERENCE. p.296-302.

GUPTA,S.M., AL-TURKI,Y.A.Y. (1998). The effect of sudden material handling system breakdown on the performance of a JIT system. *Int. J. Prod. Res.*. v.36,n.7,p.1935-1960, June.

HARREL, C. J.; BATEMAN, R.E.; GOGG, T.J.; MOTT, J.R.A. (1992). *System improvement using simulation*. McGraw-Hill,Inc.

HINES,W.W.; MONTGOMERY,D.C.(1990). *Probability and Statistics in Engineering and management science*3rd ed. Wiley, New York

IMAN,R.L., et al. (1981). Na approach to sensitivity analysis of computer models: Part I – Introduction, Input Variable Selection and Preliminary Variable Assessment. *Journal of quality Technology*. v.13,n.3,July.

ISHIHARA,T.(1999). Aumento da confiabilidade com base nos dados de manutenção. *Revista Máquinas e Metais*.p.134-143,Aug.

KELTON,W.D; VICENT, S.G.(1992). Distribution Selection and Validation. In: PROCEEDING OF THE 1992 WINTER SIMULATION CONFERENCE. p.300-304.

KELTON, W.D.S(1997). Statistical analysis of simulation output. IN: PROCEEDING OF THE 1997 WINTER SIMULATION CONFERENCE. p.23-30.

KHOSHNEVIS,B.(1994).*Discrete Systems Simulation*. McGraw-Hill,Inc.

LAW, A . M. (1993). A Forum on Crucial Issues in Simulation Modeling. *Industrial Engineering*,p.32-36, May.

LAW,A.M.;McCOMAS,M.G. (1989). Pitfalls to avoid in the simulation of manufacturing systems. *Industrial Engineering*,p.28-30 e 69, May.

LAW,A.M.;McCOMAS,M.G.; VICENT,S.G. (1994). The crucial role of input modeling in successful simulation studies. *Industrial Engineering*,p.55-59, July.

LAW,A.M.; KELTON, W.D.(1991) *Simulation modeling and analysis*. McGraw-Hill, Inc.

LEEMIS, L.M. (1996). Discrete-event simulation input process modeling. In: PROCEEDINGS OF THE 1996 WINTER SIMULATION CONFERENCE.p.39-46.

LOBÃO,E.C.;PORTO, A.J.V. (1996). Evolução das técnicas de simulação em acordo com a tecnologia. CD ROOM. IN: XVI ENEGEP - CONGRESSO NACIONAL DE ENGENHARIA DE PRODUÇÃO.

LOBÃO,E.C.;PORTO, A.J.V. (1997).Proposta para sistematização de estudos de simulação. CD ROOM. IN: XVII ENEGEP - CONGRESSO NACIONAL DE ENGENHARIA DE PRODUÇÃO.

MARIA A.(1997), Introduction to modeling and simulation. In: PROCEEDINGS OF THE 1997 WINTER SIMULATION CONFERENCE.p.7-13.

MARTIN, K.F.; BRADFORD,T.C.(1993). Modelling the breakdown behavior of transfer line machines for use in computer simulation. *International Journal of Modelling & Simulation*. V.13,n.1,p.4-8.

NELSON,L.B. et al.(1995) Input modeling when simple model fail. In: PROCEEDINGS OF THE 1995 WINTER SIMULATION CONFERENCE.p.93-100.

NORMAN, C.V.B. et al (1993). Simulation Practices in Manufacturing IN: PROCEEDING OF THE 1993 WINTER SIMULATION CONFERENCE. p.15-23.

SALLES, M. C. (1997) *Notas de Aula – Simulação Estocástica 1997*. USP – ICMSC.

SALLES,M.C.B.(1998) *Seminário apresentado ao INPE – Instituto nacional de pesquisas espaciais*.INPE

SARGENT (1996). Verifying and Validating simulation models. IN: PROCEEDING OF THE 1996 WINTER SIMULATION CONFERENCE. p.55-64.

SCHRUBEN, LEE W. (1980). Establishing the credibility of simulation. *Simulation* p.101-105.

SHANNON, R. E.; C. PEGDEN, D.; SADOWSKI, R. P.(1990). *Introduction to simulation using siman*. McGraw-Hill, Inc.

SPIEGEL, M.R. (1984). *Estatística. Coleção Schaum*. McGRAW – HILL

TUMAY, K.;HARREL,C.(1997). Simulation made easy. *IIE Solutions*,p.39-41,July.

UNIVERSIDADE DE SÃO PAULO. Escola de Engenharia de São Carlos. Serviço de Biblioteca (1996). *Diretrizes para elaboração de dissertações e teses na EESC-USP*. São Carlos.

USER’S GUIDE DA PROMODEL - “PROMODEL”(1994)

KENNE, J.P.; GHARBI, A. (1999). Experimental design in production and maintenance control problem of a single machine, single product manufacturing system. *Int. J. Prod. Res.* v.37, n.3, p.621-637, Dec.

VOELKER, D. H.; ORTON, P. Z. (1993). *Statistics. Cliffs Quick Review*. Cliffs Notes Incorporated – Lincoln, Nebraska

WILLIAMS, E.J. (1994). Downtime Data – Its Collection, Analysis, and Importance. In: *PROCEEDING OF THE 1994 WINTER SIMULATION CONFERENCE*. p.1040-1043.

APÊNDICE 1

DISTRIBUIÇÕES DE PROBABILIDADE

As distribuições de probabilidade podem ser divididas em: contínuas e discretas. Segundo BANKS et al (1995), distribuições contínuas são aquelas que descrevem fenômenos aleatórios cujas variáveis assumem qualquer valor real em algum intervalo. E distribuições discretas são aquelas que descrevem fenômenos aleatórios cujas variáveis assumem somente valores inteiros. A seguir, descreveremos algumas das distribuições de probabilidade mais utilizadas.

I. DISTRIBUIÇÕES CONTÍNUAS

Os parâmetros para a função densidade das distribuições contínuas são 3 tipos:

1. parâmetro de localização (γ) especifica o ponto de alcance dos valores da distribuição, geralmente γ é o ponto médio ou o menor ponto do alcance da distribuição. Algumas vezes, este parâmetro é denominado de parâmetro de mudança pois somente altera o alcance da distribuição para a direita ou para a esquerda.
2. parâmetro de escala (β) determina a medida da escala (ou a unidade) dos valores no alcance da distribuição. Este parâmetro permite comprimir ou expandir a figura sem alterar sua forma.
3. parâmetro de formato (α) define através do formato da distribuição, uma família de distribuição de probabilidade a qual a distribuição de interesse pode ser associada. A modificação desta parâmetro altera as

propriedades da distribuição. Algumas distribuições não tem este parâmetro, enquanto outras possuem mais de um.

A seguir apresentaremos as distribuições contínuas mais relevantes.

1.1. DISTRIBUIÇÃO UNIFORME CONTÍNUA- $U(a,b)$

Algumas vezes conhecida como distribuição retangular, a distribuição uniforme é geralmente aplicada em casos onde o modelador não possui os dados sobre o comportamento do sistema. Tendo em geral o conhecimento das estimativas de máximo e mínimo valores mais prováveis respectivamente “a” e “b” dos dados desejados, como por exemplo, tempo de operação de uma máquina. Esta distribuição é usada para gerar números aleatórios e funções aleatórias entre os pontos. A distribuição $U(0,1)$ é fundamental para gerar valores de outras distribuições.

PARÂMETROS: Esta distribuição possui dois parâmetros: a parâmetro de localização, e $(b-a)$ o parâmetro de escala. Onde a e b são números reais, tal que $a < b$.

FUNÇÃO DENSIDADE DE PROBABILIDADE:

$$f(x) = \begin{cases} \frac{1}{(b-a)} & \text{se } a \leq x \leq b \\ 0 & \text{Outros valores} \end{cases}$$

FUNÇÃO DE DISTRIBUIÇÃO:

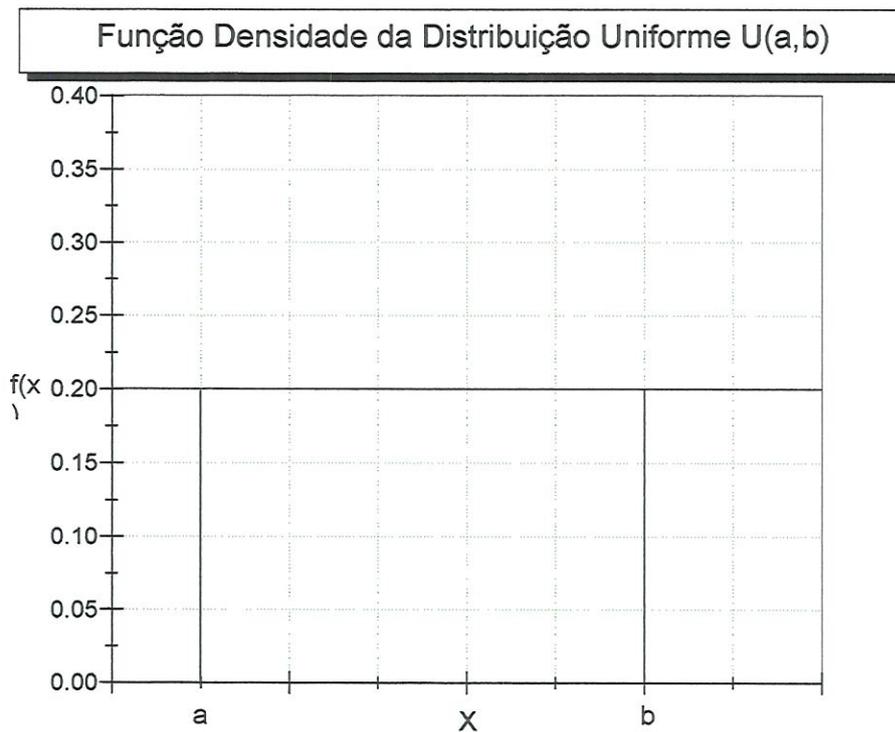
$$F(x) = \begin{cases} 0 & \text{se } x < a \\ \frac{(x-a)}{(b-a)} & \text{se } a \leq x \leq b \\ 1 & \text{se } b < x \end{cases}$$

ALCANCE: $[a,b]$

$$\text{MÉDIA: } \left(\frac{a+b}{2} \right)$$

$$\text{VARIÂNCIA: } \frac{(b-a)^2}{12}$$

MODA: não existe



COMENTÁRIOS: A distribuição uniforme é um caso especial da distribuição beta, quando $\alpha_1 = \alpha_2 = 1$.

I.2. DISTRIBUIÇÃO EXPONENCIAL

A distribuição exponencial é amplamente aplicada na análise de tempos entre eventos independentes tais com tempo entre chegadas, tempo de vida de componentes eletrônicos, tempos entre a chegada de ordens de fabricação em uma empresa, tempo de vida de mecanismos com taxa de falha constante, e outros.

PARÂMETROS: $\beta > 0$

FUNÇÃO DENSIDADE DE PROBABILIDADE:

$$f(x) = \begin{cases} \frac{1}{\beta} e^{-x/\beta} & \text{se } x \geq 0 \\ 0 & \text{outros valores} \end{cases}$$

FUNÇÃO DE DISTRIBUIÇÃO:

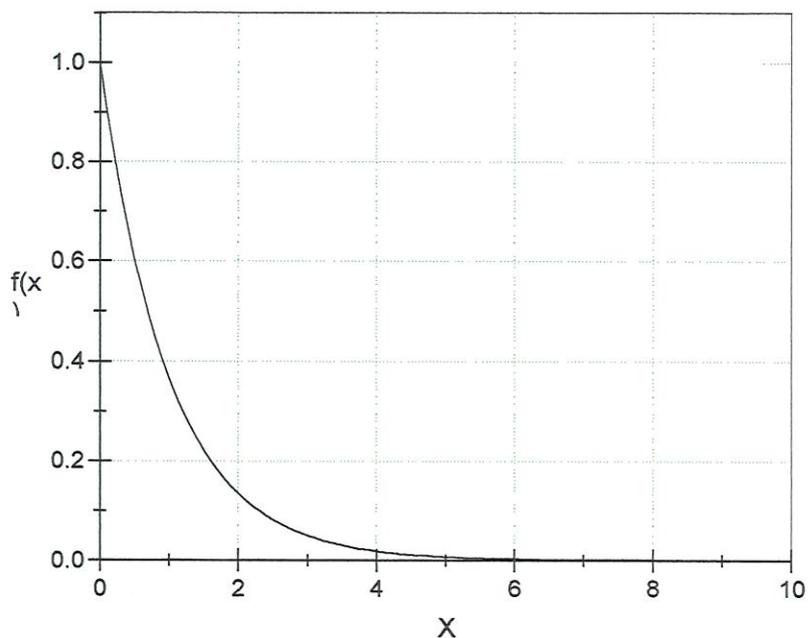
$$F(x) = \begin{cases} 1 - e^{-x/\beta} & \text{se } x \geq 0 \\ 0 & \text{outros valores} \end{cases}$$

ALCANCE: $[0, \infty)$

MÉDIA: β

VARIÂNCIA: β^2

MODA: zero

Função Densidade da Distribuição Exponencial (0,1)**COMENTÁRIOS:**

A distribuição $\exp(\beta)$ é um caso especial da distribuição Gama e da distribuição Weibull, para $\alpha=1$.

Se $X_1, X_2, X_3, \dots, X_m$ são variáveis aleatórias independentes representadas pela distribuição $\exp(\beta)$. Então $X_1 + X_2 + X_3 + \dots + X_m \cong \text{gamma}(m, \beta)$, também é denominada de distribuição m-Erlang, quando m é um inteiro.

1.3. DISTRIBUIÇÃO GAMA (α, β)

A distribuição Gama é aplicada em sistemas que possuem uma quantidade de tempo para completar alguma tarefa; como por exemplo, tempo de reparo de uma máquina, tempo de atendimento de um cliente.

PARÂMETROS: Possui o parâmetro de formato α e o parâmetro de escala β , onde $\alpha > 0$ e $\beta > 0$.

FUNÇÃO DENSIDADE DE PROBABILIDADE:

$$f(X) = \begin{cases} \frac{\beta^{-\alpha} x^{\alpha-1} e^{-x/\beta}}{\Gamma(\alpha)} & \text{se } x > 0 \\ 0 & \text{outros valores} \end{cases}$$

Onde $\Gamma(\alpha)$ é a função gamma, definido pela integral $\Gamma(z) = \int_0^{\infty} t^{z-1} e^{-t} dt$ para algum $z > 0$.

Algumas propriedades da função gamma:

$$\Gamma(z+1) = z\Gamma(z) \quad \text{para algum } z > 0$$

$$\Gamma(k+1) = k! \sqrt{\pi} \quad \text{para algum } k \text{ inteiro não negativo}$$

$$(k+1/2)! = .1.3.5\dots(2k-1)/2^k \quad \text{para algum } k \text{ inteiro positivo}$$

$$\Gamma(1/2) = \sqrt{\pi}$$

FUNÇÃO DE DISTRIBUIÇÃO: Se α não é inteiro, então não existe a definição de uma função. Se α é um inteiro positivo então:

$$F(x) = \begin{cases} 1 - e^{-x/\beta} \sum_{j=0}^{\alpha-1} \frac{(x/\beta)^j}{j!} & \text{se } x > 0 \\ 0 & \text{outros valores} \end{cases}$$

ALCANCE: $[0, \infty)$

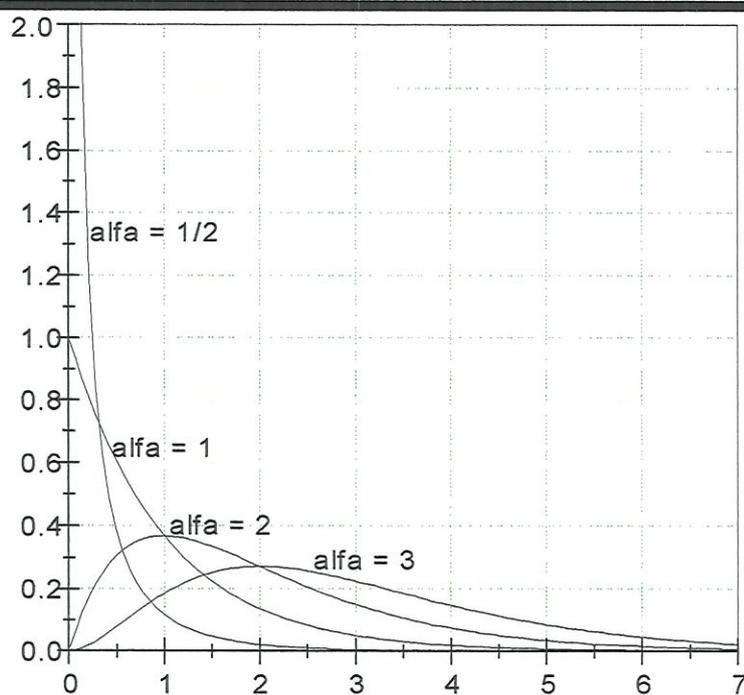
MÉDIA: $\alpha \beta$

VARIÂNCIA: $\alpha\beta^2$

MODA:

$$\begin{cases} \beta(\alpha - 1) & \text{se } \alpha \geq 1 \\ 0 & \text{outros valores} \end{cases}$$

Função Densidade da Distribuição Gamma (alfa,1)



1.4. DISTRIBUIÇÃO WEIBULL (α, β)

É amplamente aplicada em modelos que envolvem a questão de confiabilidade de equipamentos, tal como tempo de vida de equipamentos eletrônicos. Algumas vezes, esta distribuição é aplicada em sistemas com sistemas que possuem uma quantidade de tempo para completar alguma tarefa.

PARÂMETROS: Possui o parâmetro de formato α e o parâmetro de escala β , onde $\alpha > 0$ e $\beta > 0$.

FUNÇÃO DENSIDADE DE PROBABILIDADE:

$$f(X) = \begin{cases} \alpha\beta^{-\alpha} x^{\alpha-1} e^{-(x/\beta)^\alpha} & \text{se } x > 0 \\ 0 & \text{outros valores} \end{cases}$$

FUNÇÃO DE DISTRIBUIÇÃO:

$$F(x) = \begin{cases} 1 - e^{-(x/\beta)^\alpha} & \text{se } x > 0 \\ 0 & \text{outros valores} \end{cases}$$

ALCANCE: $[0, \infty)$

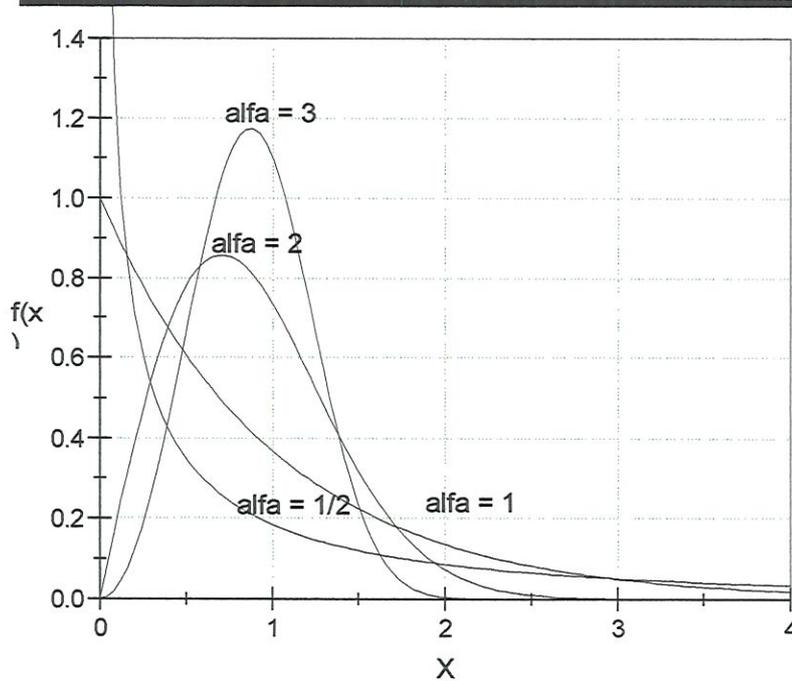
$$\text{MÉDIA: } \frac{\beta}{\alpha} \Gamma\left(\frac{1}{\alpha}\right)$$

$$\text{VARIÂNCIA: } \frac{\beta^2}{\alpha} \left\{ 2\Gamma\left(\frac{2}{\alpha}\right) - \frac{1}{\alpha} \left[\Gamma\left(\frac{1}{\alpha}\right) \right]^2 \right\}$$

MODA:

$$\begin{cases} \frac{\beta(\alpha-1)^{1/\alpha}}{\alpha} & \text{se } \alpha \geq 1 \\ 0 & \text{se } \alpha < 1 \end{cases}$$

Função Densidade da Distribuição Weibull (alfa, 1)



1.5. DISTRIBUIÇÃO NORMAL – $N(\mu, \sigma^2)$

Descreve fenômenos com uma variação simétrica em relação a média. A importância desta distribuição se encontra no teorema do limite central. Em geral, esta distribuição é aplicada em situações onde um processo pode ser representado como a soma de outros processos independentes, como por exemplo, tempo de montagem de um produto que pode ser representado como a soma de todos os tempos necessários para cada operação de montagem. A distribuição $N(0,1)$ é frequentemente denominada de distribuição normal padrão.

PARÂMETROS: Possui o parâmetro de localização $\mu \in (-\infty, \infty)$ e o parâmetro de escala $\sigma > 0$.

$$\begin{cases} \frac{1}{\sqrt{2\pi\sigma^2}} e^{-(x-\mu)^2 / 2\sigma^2} & \text{para } x \in \mathbb{R} \end{cases}$$

FUNÇÃO DENSIDADE: $f(x) =$

FUNÇÃO DE DISTRIBUIÇÃO: não possui uma fórmula

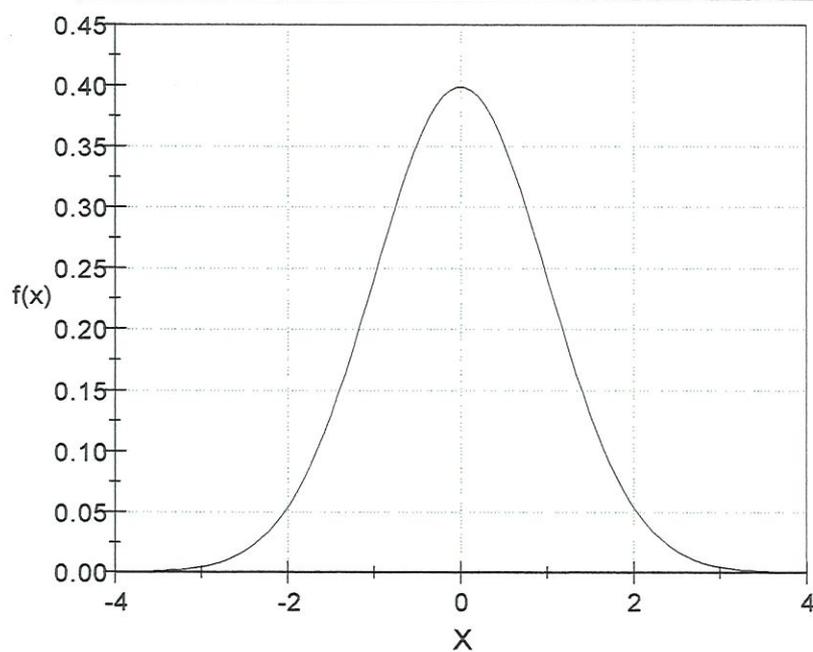
ALCANCE: $(-\infty, \infty)$

MÉDIA: μ

VARIÂNCIA: σ^2

MODA: μ

Função Densidade da Distribuição Normal (0,1)



1.6. DISTRIBUIÇÃO TRIANGULAR – $TRIANG(a,b,c)$

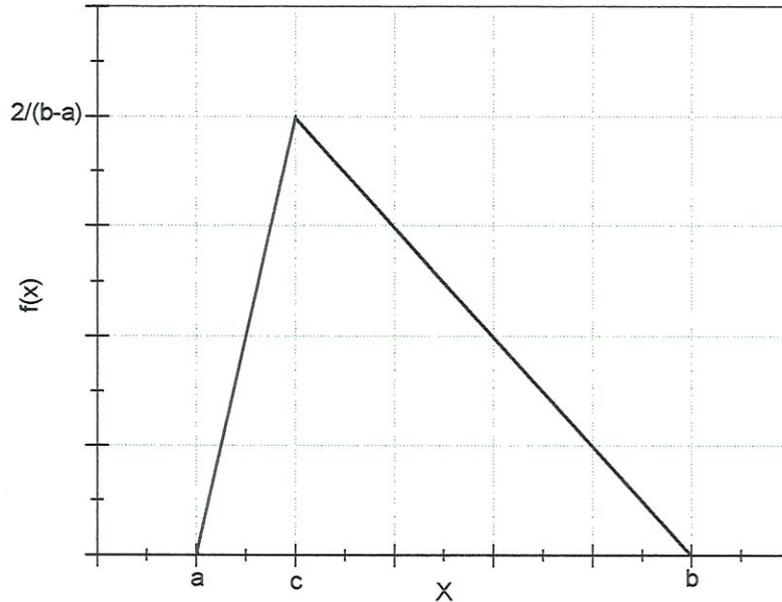
É aplicada em modelos onde são conhecidos apenas o maior valor, o mais provável e o menor valor de uma distribuição.

$$\text{MÉDIA: } \frac{a+b+c}{3}$$

$$\text{VARIÂNCIA: } \frac{a^2 + b^2 + c^2 - ab - ac - bc}{18}$$

$$\text{MODA: } c$$

Função Densidade da Distribuição Triangular (a,b,c)



II. DISTRIBUIÇÃO DISCRETA

De acordo com LAW & KELTON(1991), MORETIN(1979) D BANKS et al (1996), descreveremos 5 distribuições discretas.

II.1. DISTRIBUIÇÃO DE BERNOULLI

Aplicada em situações onde existe uma única tentativa de um experimento aleatório. Onde só é possível obter um sucesso ou fracasso com solução.

$$\text{FUNÇÃO DE PROBABILIDADE: } p(x) = \begin{cases} 1-p & \text{se } x = 0 \\ p & \text{se } x = 1 \\ 0 & \text{outros valores} \end{cases}$$

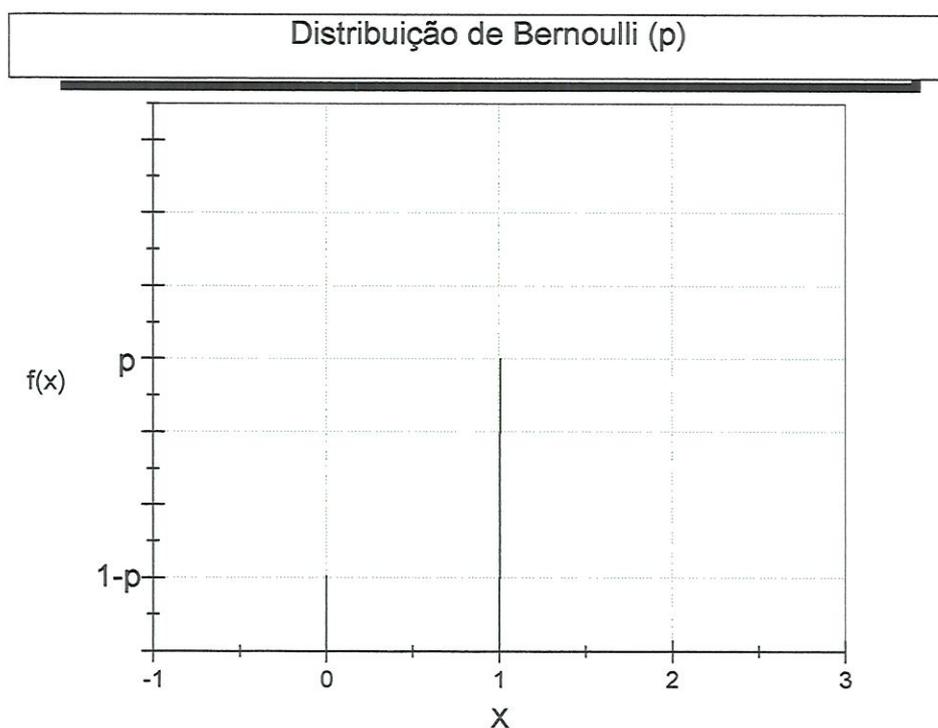
$$\text{DISTRIBUIÇÃO: } F(x) = \begin{cases} 0 & \text{se } x < 0 \\ 1-p & \text{se } 0 \leq x < 1 \\ 1 & \text{se } 1 \leq x \end{cases}$$

MÉDIA: p

VARIÂNCIA: $p(1-p)$



$$\text{MODA: } \begin{cases} 0 & \text{se } p < 1/2 \\ 0 \text{ e } 1 & \text{se } p = 1/2 \\ 1 & \text{se } p > 1/2 \end{cases}$$



II.2. DISTRIBUIÇÃO BINOMIAL

Aplicada em modelos que consideram n tentativas independentes de um mesmo experimento aleatório. Onde cada tentativa admite dois resultados: fracasso com probabilidade q e sucesso com probabilidade p , $p + q = 1$. As probabilidades de sucesso são as mesmas para cada tentativa.

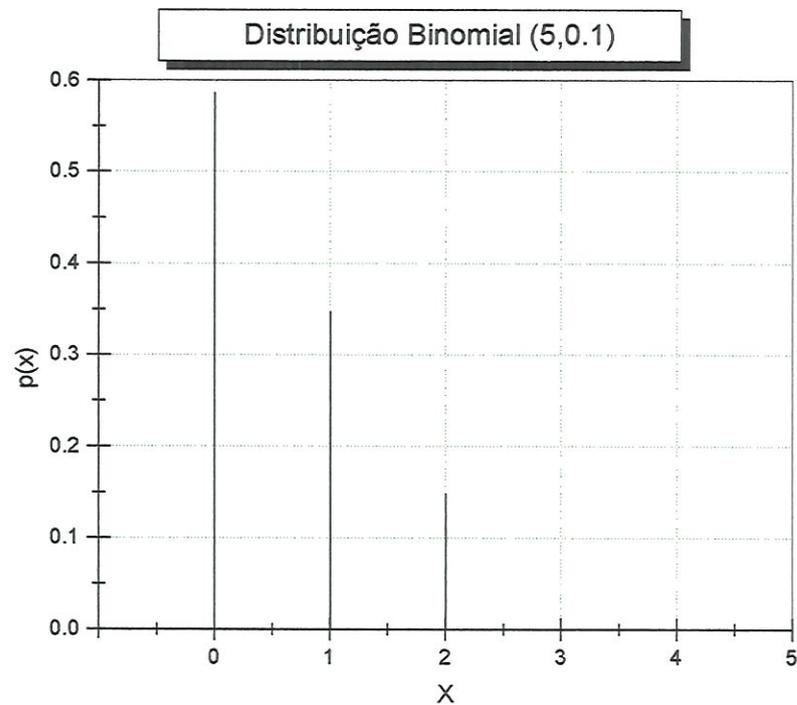
$$\text{FUNÇÃO DE PROBABILIDADE: } p(x) = \begin{cases} \binom{t}{x} p^x (1-p)^{t-x} & \text{se } x \in \{0, 1, \dots, t\} \\ 0 & \text{outros valores} \end{cases}$$

$$\text{DISTRIBUIÇÃO: } F(x) = \begin{cases} 0 & \text{se } x < 0 \\ \sum_{i=0}^{\lfloor x \rfloor} \binom{t}{i} p^i (1-p)^{t-i} & \text{se } 0 \leq x \leq t \\ 1 & \text{se } x > t \end{cases}$$

$$\text{MÉDIA: } tp$$

$$\text{VARIÂNCIA: } tp(1-p)$$

$$\text{MODA: } \begin{cases} p(t+1) - 1 \text{ e } p(t+1) & \text{se } p(t+1) \text{ é um inteiro} \\ p(t+1) & \text{outros valores} \end{cases}$$



II.3. DISTRIBUIÇÃO GEOMÉTRICA

Aplicada em tentativas sucessivas e independentes de um mesmo experimento, onde cada tentativa possui sucesso com probabilidade p e fracasso com probabilidade q .

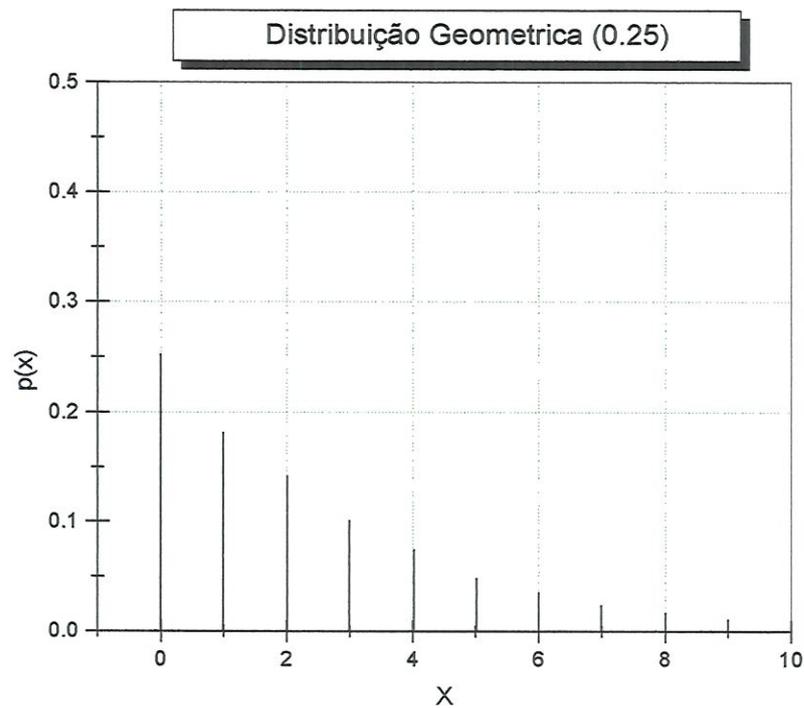
$$\text{FUNÇÃO DE PROBABILIDADE: } p(x) = \begin{cases} p(1-p)^x & \text{se } x \in \{0,1,\dots\} \\ 0 & \text{outros valores} \end{cases}$$

DISTRIBUIÇÃO: $F(x) \begin{cases} 1 - (1-p)^{|x|+1} & \text{se } x \geq 0 \\ 0 & \text{outros valores} \end{cases}$

MÉDIA: $\frac{1-p}{p}$

VARIÂNCIA: $\frac{1-p}{p^2}$

MODA: 0



II.4. DISTRIBUIÇÃO NEGATIVA BINOMIAL

É aplicada em experimentos aleatórios que são repetidos independentemente até que um evento A ocorra pela r -ésima vez.

0

FUNÇÃO DE PROBABILIDADE:
$$p(x) = \begin{cases} \left(\frac{S+X-1}{X}\right) p^s (1-p)^x & \text{se } x \in \{0,1,\dots\} \\ 0 & \text{outros valores} \end{cases}$$

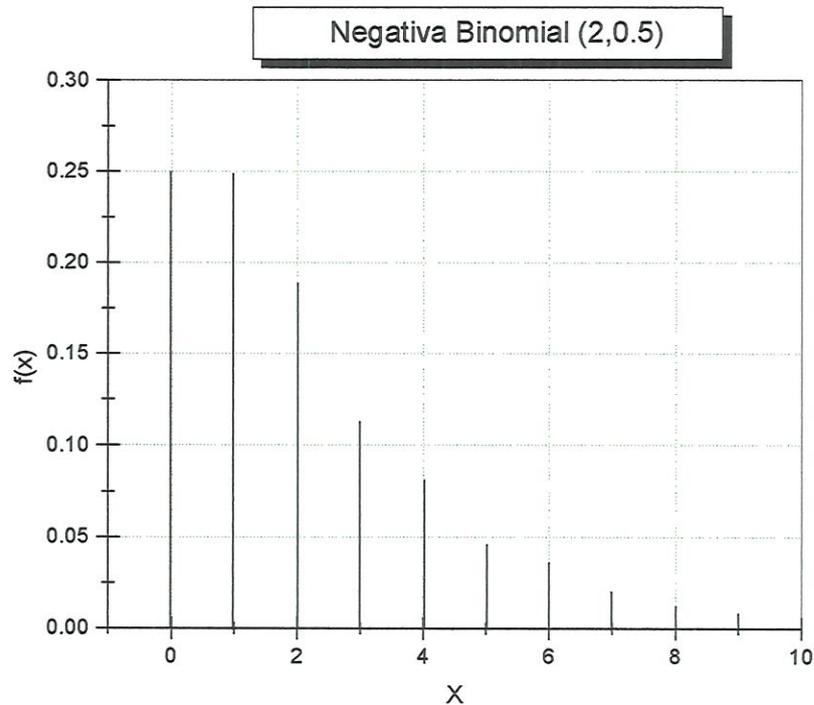
DISTRIBUIÇÃO:
$$F(x) = \begin{cases} \sum_{i=0}^{|x|} \binom{s+i-1}{i} p^s (1-p) & \text{se } x \geq 0 \\ 0 & \text{outros valores} \end{cases}$$

MÉDIA:
$$\frac{s(1-p)}{p}$$

VARIÂNCIA:
$$\frac{s(1-p)}{p^2}$$

MODA: Seja $y = [s(1-p)-1]/p$; então

$$moda = \begin{cases} y \text{ e } y+1 & \text{se } y \text{ é inteiro} \\ \lfloor y \rfloor + 1 & \text{outros valores} \end{cases}$$



II.5. DISTRIBUIÇÃO POISSON

É aplicada quando ocorre números de eventos em um intervalo de tempo, sendo que estes eventos ocorrem com uma taxa constante.

FUNÇÃO DE PROBABILIDADE:

$$p(x) = \begin{cases} \frac{e^{-\lambda} \lambda^x}{x!} & \text{se } x \in \{0, 1, \dots\} \\ 0 & \text{outros valores} \end{cases}$$

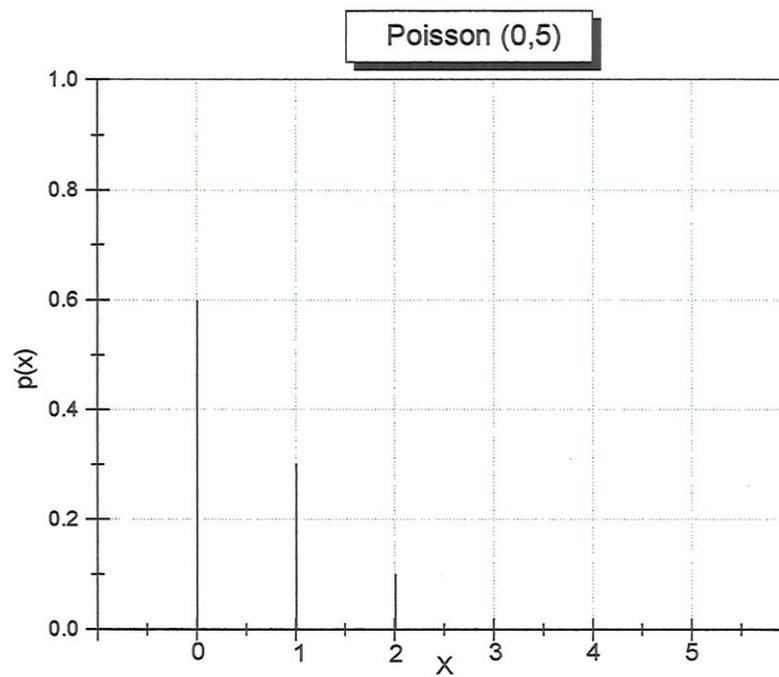
DISTRIBUIÇÃO:

$$F(x) = \begin{cases} 0 & \text{se } x < 0 \\ e^{-\lambda} \sum_{i=0}^{\lfloor x \rfloor} \frac{\lambda^i}{i!} & \text{se } 0 \leq x \end{cases}$$

MÉDIA: λ

VARIÂNCIA: λ

MODA: $\begin{cases} \lambda - 1 \text{ e } \lambda \text{ se } \lambda \text{ é um inteiro} \\ \lfloor \lambda \rfloor \text{ outros valores} \end{cases}$



APÊNDICE 2

PROCEDIMENTOS DE ANÁLISE ESTATÍSTICA PARA SIMULAÇÃO TERMINATIVA

Para analisar os dados de saída de uma simulação terminativa, deve-se considerar inicialmente as seguintes questões (*user's guide* da PROMODEL - "PROMODEL", 1994):

- ♦ estado inicial do modelo – indica sob que condições iniciais o modelo começa a ser executado;
- ♦ evento ou tempo terminativo – delimita o tamanho da corrida da simulação;
- ♦ número de replicações – é o número de corridas executadas de um modelo.

O estado inicial do modelo no instante t_0 , em geral, é assumido como vazio, ou seja, seus recursos encontram-se desocupados e prontos para serem ativados. O evento E que limita a simulação, é uma ocorrência que indica o término do processo de simulação como por exemplo a quantidade de peças a serem produzidas por uma fábrica. O tempo terminativo, com a mesma função do evento E, é o intervalo de tempo de interesse para o estudo, tal como a taxa de produção no período de um mês. Na simulação terminativa, tanto o estado inicial, quanto evento E (ou tempo terminativo) são fornecidos pelas características e/ou objetivos do modelo.

Para calcular os resultados do modelo de simulação terminativa, apresenta-se o método das replicações independentes. Este método, aplicado pela maioria dos profissionais de simulação consiste em "rodar" o modelo R vezes, onde $R \geq 10$. Considere $y_{r1}, y_{r2}, y_{r3}, y_{r4}, \dots, y_{rm}$, os resultados de uma simulação no período $[0, t_e]$ e seja

$$\theta = E\left(\frac{1}{n} \sum_{i=1}^n y_{ri}'s\right) \quad (A.1)$$

a estimativa procurada pelo estudo.

Deve-se observar que esta simulação, limitada pelo intervalo de tempo t_e , poderia também ser limitada por um evento. Os y_{ri} 's resultados do modelo, onde r é fixo, são autocorrelacionados, ou seja, $y_{r1}, y_{r2}, \dots, y_{rn}$ não são independentes. No entanto, para cada replicação, quando diferentes *stream* são aplicados nos dados de entrada, os y_{ri} 's são *iid*, ou seja, $y_{ri} \neq y_{si}$ quando $s \neq r$. Seja

$$\bar{\theta}_r = \frac{1}{n} \sum_{i=1}^n y_{ri} \quad r = 1, \dots, R \quad (A.2)$$

o estimador da média da amostra dos resultados. Embora existam outros estimadores, o estimador da média da amostra é o mais eficiente para a média populacional. Um estimador é dito mais eficiente, quando a variância da distribuição amostral de uma medida é menor que a variância da distribuição amostral de outra medida (SPIEGEL, 1985).

Ao determinar θ , deve-se encontrar um ponto que se aproxima do parâmetro populacional correspondente. A estimativa de um parâmetro por um único ponto, denominada de estimativa por ponto, é uma medida com precisão insuficiente. Uma vez que a estimativa encontrada geralmente não coincide com o verdadeiro valor do parâmetro populacional correspondente. Em outras palavras, a estimativa da amostra encontrada possui uma probabilidade muito pequena de ser o exato valor do parâmetro populacional correspondente. Desta forma, provavelmente existe um erro de estimativa, quando estima-se por ponto o parâmetro da população (COSTA NETO, 1977).

Diante deste problema, é aconselhável construir um intervalo de confiança para o parâmetro θ procurado. Ou seja, deve-se determinar um limite de aceitação para a variação de θ . Para isto, é necessário especificar dois pontos extremos simétricos a μ (média da população), de modo que a probabilidade deste intervalo de confiança conter θ seja igual a $1-\alpha$. Seja H a semi-amplitude do intervalo, onde $1-\alpha$ é o nível de confiança, e α é a probabilidade de erro. Ou seja, é a probabilidade de introduzir erro ao

afirmar que o intervalo de confiança contém o parâmetro procurado. Os pontos extremos a μ são definidos como $\mu-H$ e $\mu+H$.

A seguir será apresentado os procedimentos para construir um intervalo de confiança para a estimativa θ procurada, a partir de um determinado número R de replicações. Estes procedimentos são caracterizados em dois tipos:

- ◆ Intervalo de confiança para um número fixo de replicações e
- ◆ Intervalo de confiança com precisão especificada pelo modelador.

Para cada um dos procedimentos, foi considerado que a simulação tenha sido executada sob condições iniciais idênticas em cada replicação, utilizando um “*stream*” diferente.

1. INTERVALO DE CONFIANÇA PARA UM NÚMERO FIXO DE REPLICAÇÕES

Este método consiste em construir um intervalo de confiança para uma estimativa de $\underline{\theta}$, de acordo com um número R de replicações. Onde $R \geq 10$ é especificado inicialmente. O parâmetro θ para qual o intervalo será construído, pode ser obtido pela equação A.3:

$$\underline{\theta} = \frac{1}{R} \sum_{r=1}^R \theta_r \quad (A.3)$$

Onde $\underline{\theta}_r$ para $r = 1, \dots, R$ são as médias obtidas em cada replicação, são obtidas pela equação 4.2. Note que $\underline{\theta}_r$ são *iid*. Quando $R \geq 30$ replicações, o intervalo de confiança para θ , pode ser calculado pela seguinte equação:

$$\underline{\theta} \pm Z_{\alpha/2} \sqrt{\frac{S^2(R)}{R}} \quad (A.4)$$

onde α é o nível de significância, R é o número de replicações (tamanho da amostra de observações) e S^2 é a variância. Onde

$$Z_{\alpha/2} \sqrt{\frac{S^2(R)}{R}}$$

é a semi-amplitude do intervalo de confiança, e

$$\underline{\theta} \pm Z_{\alpha/2} \sqrt{\frac{S^2(R)}{R}}$$

define os pontos extremos do intervalo. Note que $Z_{\alpha/2}$ é a distribuição da variável normal padronizada Z , e possui seus valores tabelados na maioria dos livros de cálculo de probabilidade. Ressaltando que a probabilidade do intervalo de confiança conter θ é igual a $1-\alpha$. A equação A.5 apresenta esta idéia:

$$P\left(\underline{\theta} - Z_{\alpha/2} \sqrt{\frac{S^2(R)}{R}} \leq \theta \leq \underline{\theta} + Z_{\alpha/2} \sqrt{\frac{S^2(R)}{R}}\right) = 1 - \alpha \quad (A.5)$$

Para $R < 30$, a amostra dos resultados é denominada de *pequenas amostras*, não fornece boa aproximação para um intervalo construído com a distribuição da variável normal padronizada Z . Sendo portanto necessário aplicar a aproximação da distribuição *t-student*, na construção do um intervalo. A equação 4.6 fornece o intervalo de confiança para θ com a distribuição *t-student*.

$$\underline{\theta} \pm t_{R-1, 1-\alpha/2} \sqrt{\frac{S^2(R)}{R}} \quad (A.6)$$

Onde de modo análogo, a probabilidade do intervalo de confiança conter θ é $1-\alpha$. A equação A.7 mostra esta idéia:

$$P\left(\underline{\theta} - t_{R-1, 1-\alpha/2} \sqrt{\frac{S^2(R)}{R}} \leq \theta \leq \underline{\theta} + t_{R-1, 1-\alpha/2} \sqrt{\frac{S^2(R)}{R}}\right) = 1 - \alpha \quad (A.7)$$

Note que para dados de saída contínuos, onde $\{\underline{\theta}_r(t), 0 \leq t \leq t_e\}$ para $r = 1, \dots, R$ replicações independentes, a estimativa θ pode ser calculada pela seguinte equação A.8, onde:

$$\underline{\theta}_r = \frac{1}{t_e} \int_0^{t_e} Y_r(t) dt \quad (A.8)$$

O intervalo de confiança pode ser construído através da aplicação das equação A.6 quando $R < 30$ e a equação A.4 para $R \geq 30$.

A variância S^2 da amostra dos resultados, que também é um estimador não viciado, pode ser obtida pela seguinte fórmula:

$$S^2(R) = \frac{\sum_{r=1}^R (\underline{\theta}_r - \theta)^2}{R-1} \quad (A.9)$$

Uma das desvantagens deste método é que não se possui controle sobre a precisão de θ . O procedimento seguinte resolve esta falha.

2. INTERVALO DE CONFIANÇA COM PRECISÃO ESPECIFICADA PELO MODELADOR

Consiste em calcular o número R de replicações que devem ser realizadas, para que o intervalo de confiança de θ possua uma determinada precisão, ou seja, um erro absoluto ϵ . Este erro absoluto ϵ , que deve ser especificado inicialmente pelo modelador, é a diferença máxima de aceitação entre a média da população e a média da amostra. Em outras palavras, este método determina o número R de replicações, de modo que a diferença entre a média da amostra e a média da população seja menor ou igual a precisão de erro especificada ϵ .

Seja $1-\alpha$ a probabilidade do intervalo de confiança contém θ . Desta forma, considerando a diferença entre a média da amostra e a média da população igual a ϵ , de modo que θ esteja situado neste intervalo, tem-se:

$$-t_{R-1,1-\alpha/2}\sqrt{\frac{S^2(R)}{R}} \leq \theta - \mu \leq +t_{R-1,1-\alpha/2}\sqrt{\frac{S^2(R)}{R}} \quad (A.10)$$

Onde θ é a média da amostra dos resultados, μ é a média da população

$t_{R-1,1-\alpha/2}\sqrt{\frac{S^2R}{R}}$ e é a semi-amplitude do intervalo, de acordo com a distribuição t-student. Desta forma,

$$P(|\theta - \mu| = \epsilon) = 1 - \alpha$$

Assim, como ϵ é o valor máximo de erro admitido entre as médias da amostra e da população, tem-se a seguinte desigualdade:

$$|\theta - \mu| \leq \epsilon$$

Portanto o intervalo de confiança possui pelo menos a probabilidade de $1-\alpha$ para conter a diferença esta desigualdade.

$$P(|\theta - \mu| \leq \epsilon) \geq 1 - \alpha$$

Como a desigualdade das médias está situada na semi-amplitude do intervalo, e tendo esta desigualdade o tamanho máximo definido por ϵ . Logo o valor máximo ϵ não deve ser maior que a semi-amplitude do intervalo, ou seja, o máximo valor admitido para ϵ é o tamanho da semi-amplitude, pois em caso contrário a desigualdade não estará situada no intervalo determinado. Desta forma tem-se a seguinte equação:

$$\epsilon \leq +t_{R-1, 1-\alpha/2} \sqrt{\frac{S^2(R)}{R}} \quad (A.11)$$

Para calcular a quantidade R de replicações que devem ser executadas para atingir a precisão de erro ϵ , deve-se inicialmente executar R_0 replicações, onde $R_0 \geq 2$. As R_0 replicações são denominadas de replicações piloto. LOBÃO & PORTO (1997) advertem que quanto maior o número de replicações maior será a precisão dos cálculos. A quantidade R de replicações pode ser obtido pela equação A.11 substituindo R para R_0 :

$$R \geq \left(\frac{t_{R_0-1, \alpha/2} S_0}{\epsilon} \right)^2 \quad (A.12)$$

onde S_0^2 é a variância das R_0 replicações iniciais. Como $t_{R-1, \alpha/2} \geq Z_{\alpha/2}$, então R também pode ser estimado pela seguinte equação:

$$R \geq \left(\frac{z_{\alpha/2} S_0}{\epsilon} \right)^2 \quad (A.13)$$

Após calcular R , deve-se calcular $R-R_0$ replicações adicionais, e estimar o valor θ para este novo tamanho da amostra de resultados. Para $R \geq 30$ a equação A.13 fornece uma boa aproximação. O intervalo de confiança será obtido pela seguinte equação A.4 quando $R \geq 30$, e pela equação A.6 quando $R < 30$.

ANEXO I**LISTAGEM DO PROGRAMA IMPLEMENTADO
NO MODELO****Variáveis:**

V_timetotal – somatória de “Dif”

V_numpecas – contador do número de peças processadas pelo sistema à partir do tempo de 28.000 minutos

K – indica se a peça já foi reprocessada

te - armazena o tempo corrente da simulação no momento em que a peça entra no *control point*

espera - indica o tempo em que a peça ficou esperando a liberação do *control point*. $espera = Tempo - te$

Tespera – somatória de “espera”

N – conta o número de peças que ficaram esperando a liberação do *control point*

tem - armazena o tempo corrente da simulação no momento em que a peça começa a utilizar o recurso

TME – tempo médio de espera no grupo de máquina 1, 2 e na estação de inspeção

TTME – somatória dos “TME”

Tutil - somatória de “util”

NCM – conta o número de cargas que passaram pelas máquinas dos grupos 1,2 individualmente.

TMUM – tempo médio de utilização do recurso

TTMUM – somatório dos TMUM

Dif - armazena o tempo que a carga permaneceu no sistema

Tempo - armazena o tempo corrente da simulação no momento em que a peça abandona o *control point*, ou o recurso.

Dif – tempo médio das cargas processadas pelo sistema à partir do tempo de 28.000 minutos

util – armazena o tempo em que a carga utiliza o recurso

ind - índice

NCPB – contador do número de cargas do buffer

Recursos:

Man – operários da estação de inspeção

maq1 – máquinas do grupo 1

maq2 – máquinas do grupo 2

Filas:

Q_entrada – fila da entrada do sistema

Q_saida – fila da saída do sistema

Control-Point Das Estações De Trabalho Do Sistema

Estação de entrada das cargas do sistema cpenrada

Buffers do sistema cpb1 cpb2 cpb3

Grupo de máquinas 1 cpmaq11 cpmaq12 cpmaq13

Grupo de máquinas 2 cpmaq21 cpmaq22
Estações de inspeção cpinsp 1 cpinsp2
Estações de espera para iniciar reprocesso da carga cpep1
cpep2 **Estação de parada do agv** cppark
Estação de saída das cargas do sistema cpsaída

Estações das Esteiras do Sistema

Grupo de máquinas 1

Máquina 11 sta1 sta2

Máquina 12 sta3 sta4

Máquina 13 sta5 sta6

Grupo de máquinas 2

Máquina 21 sta7 sta8

Máquina 22 sta9 sta10

Grupo de inspeção 1

Inspeção1 sta11 sta12

Inspeção2 sta13 sta14

Processos:

Pentrada

Psaida

Pesp1

Pesp2

Pmaq1

Pmaq2

Pinsp

PB1

PB2

PB3

PV1

PV2

PV3

Listagem dos Processos:

Pentrada

```
begin
  move into Q_entrada
  set k to 0
  if agv:cpentrada current=1 then
    begin
      set te(1) to ac
      increment N(1) by 1
      wait until agv:cpentrada current =0
      set Tempo to ac
      set espera(1) to Tempo - te(1)
      set Tespera(1) to Tespera(1) + espera(1)
      move into agv:cpentrada
    end
  else
    move into agv:cpentrada
    set A_time to ac
    send to PB1
```

end

Psaida

```
begin
  travel to agv:cpsaida
  move into Q_saida
  set Tempo to ac
  if Tempo > 28000.0 then
    begin
      print this," CARGA PROCESSADA " to message
      set Dif to ac - A_time
      print this load,"Tempo da carga no sistema",Dif to message
      set V_timetotal to V_timetotal + Dif
      print this load,"Tempo total das cargas no
sistema",V_timetotal to message
      increment V_numpeças by 1
      print this load,"Numero de cargas no sistema
",V_numpeças to message
      set Dif to V_timetotal/V_numpeças
      print this load,"Tempo medio das cargas no sistema",Dif to
message
      set ind to 1
    end
  end
end
```

```
print this," TEMPO DE ESPERA DA PECA PARA
PROCESSAMENTO" to message
while ind<=7 do
begin
print this load,"Tempo de espera para
processamento ", Tespera(ind) to message
print this load,"Numero de cargas em espera ",
N(ind) to message
if N(ind) < 0 then set TME(ind) to
Tespera(ind)/N(ind)
else set TME(ind) to 0
print this load," Tempo medio de
espera",TME(ind),"no processo", ind to message
increment ind by 1
end

set ind to 1
set TTME to 0
while ind<=7 do
begin
set TTME to TTME+ TME(ind)
print this load,"Somatoria dos tempos medios de
espera",TTME to message
increment ind by 1
end
print this load,"Tempo medio total de espera no
sistema",TTME to message

set ind to 1
print this," TEMPO DE UTILIZACAO "to message
```

```
while ind<=3 do
begin
    print this load,"Tempo de utilizacao", Tutil(ind),"no
grupo de recursos", ind to message
    print this load,"Numero de cargas",NCM(ind),"no
grupo de recursos ", ind to message

    if NCM(ind) <> 0 then set TMUM(ind) to
Tutil(ind)/NCM(ind)
    else set TMUM(ind) to 0

    print this load," Tempo medio de utilizacao dos
recursos",TMUM(ind),"no grupo", ind to message
    increment ind by 1
end
set ind to 1
set TTMUM to 0
while ind<=3 do
begin
    set TTMUM to TTMUM+ TMUM(ind)
    print this load,"Somatoria dos tempos medio de
utilização",TTMUM to message
    increment ind by 1
end
print this load,"Tempo medio total de utilizacao dos
recursos no sistema",TTMUM to message
end
send to die
end
```

Pesp1

```
begin
  set k to 1
  if agv:cpep2 current=1 then wait until agv:cpep2 current=0
  travel to agv:cpep1
  send to PB1
end
```

Pesp2

```
begin
  set k to 1
  if agv:cpep1 current=1 then wait until agv:cpep1 current=0
  travel to agv:cpep2
  send to PB1
end
```

Pmaq1

```
begin

  if agv:cpmaq11 current=0 then
    begin
      travel to agv:cpmaq11
      decrement NCPB(1) by 1
    end
  end
```

```
        move into conv:sta1
        travel to conv:sta2
        set tem(1) to ac
        use maq1(1) for random stream stream_m11 gamma 2,0.25 min
        set Tempo to ac
        move into agv:cpmaq11
        increment NCM(1) by 1
        set util(1) to Tempo - tem(1)
        set Tutil(1) to Tutil(1) + util(1)
        send to PB2

    end
    else
    begin
        if agv:cpmaq12 current=0 then
        begin
            travel to agv:cpmaq12
            decrement NCPB(1) by 1
            move into conv:sta3
            travel to conv:sta4
            set tem(1) to ac
            use maq1(2) for random stream stream_m12 gamma 2,0.25

min
            set Tempo to ac
            move into agv:cpmaq12
            increment NCM(1) by 1
            set util(1) to Tempo - tem(1)
            set Tutil(1) to Tutil(1) + util(1)
            send to PB2
```

```
end
else
begin
  if agv:cpmaq13 current=0 then
  begin
    travel to agv:cpmaq13
    decrement NCPB(1) by 1
    move into conv:sta5
    travel to conv:sta6
    set tem(1) to ac
    use maq1(3) for random stream stream_m13
    gamma 2,0.25 min
    set Tempo to ac
    move into agv:cpmaq13
    increment NCM(1) by 1
    set util(1) to Tempo - tem(1)
    set Tutil(1) to Tutil(1) + util(1)
    send to PB2
  end
end
end
end
```

Pmaq2

```
begin
```

```
if agv:cpmaq21 current=0 then
begin
    decrement NCPB(2) by 1
    travel to agv:cpmaq21
    move into conv:sta7
    travel to conv:sta8
    set tem(2) to ac
    use maq2(1) for random stream stream_m21 gamma 2,0.05 min
    set Tempo to ac
    move into agv:cpmaq21
    increment NCM(2) by 1
    set util(2) to Tempo - tem(2)
    set Tutil(2) to Tutil(2) + util(2)
    send to PB3

end
else
begin
    if agv:cpmaq22 current=0 then
    begin
        decrement NCPB(2) by 1
        travel to agv:cpmaq22
        move into conv:sta9
        travel to conv:sta10
        set tem(2) to ac
        use maq2(2) for random stream stream_m22 gamma 2,0.05
min
        set Tempo to ac
        move into agv:cpmaq22
```

```
        increment NCM(2) by 1
        set util(2) to Tempo - tem(2)
        set Tutil(2) to Tutil(2) + util(2)
        send to PB3
    end
end
end
```

Pinsp

```
begin

    if agv:cpinsp1 current=0 then
        begin
            travel to agv:cpinsp1
            decrement NCPB(3) by 1
            move into conv:sta11
            travel to conv:sta12
            set tem(3) to ac
            use Man(1) for triangular 0.30,1,2 min
            set Tempo to ac
            move into agv:cpinsp1
            increment NCM(3) by 1
            set util(3) to Tempo - tem(3)
            set Tutil(3) to Tutil(3) + util(3)
            if k=1 then send to Psaida
            else send to oneof (0.1:Pesp1,0.9:Psaida)
        end
    end
end
```

```
end
else
begin
  if agv:cpinsp2 current=0 then
  begin
    travel to agv:cpinsp2
    decrement NCPB(3) by 1
    move into conv:sta13
    travel to conv:sta14
    set tem(3) to ac
    use Man(2) for triangular 0.30,1,2 min
    set Tempo to ac
    move into agv:cpinsp2
    increment NCM(3) by 1
    set util(3) to Tempo - tem(3)
    set Tutil(3) to Tutil(3) + util(3)
    if k=1 then send to Psaida
    else send to oneof (0.1:Pesp2,0.9:Psaida)
  end
end
end
```

PB1

```
begin

  if NCPB(1)>10 then
  begin
```

```
        set te(2) to ac
        increment N(2) by 1
        wait until NCPB(1)<11
        set Tempo to ac
        set espera(2) to Tempo - te(2)
        set Tespera(2) to Tespera(2) + espera(2)
        travel to agv:cpb1
        increment NCPB(1) by 1
        send to PV1
    end

    travel to agv:cpb1
    increment NCPB(1) by 1
    send to PV1
end
```

PB2

```
begin

    if NCPB(2)<11 then
        begin
            travel to agv:cpb2
            increment NCPB(2) by 1
            send to PV2
        end
    else
        set te(4) to ac
```

```
        increment N(4) by 1
        wait until NCPB(2)<11
        set Tempo to ac
        set espera(4) to Tempo - te(4)
        set Tespera(4) to Tespera(4) + espera(4)
        travel to agv:cpb2
        increment NCPB(2) by 1
    send to PV2

end
```

PB3

```
begin
    if NCPB(3)<11 then
        begin
            travel to agv:cpb3
            increment NCPB(3) by 1
            send to PV3
        end
    else
        set te(6) to ac
        increment N(6) by 1
        wait until NCPB(3)<11
        set Tempo to ac
        set espera(6) to Tempo - te(6)
        set Tespera(6) to Tespera(6) + espera(6)
        travel to agv:cpb3
    end
end
```

```
        increment NCPB(3) by 1
    send to PV3
end
```

PV1

```
begin

    if agv:cpmaq11 current=1 and agv:cpmaq12 current =1 and agv:cpmaq13
current =1 then
        begin
            set te(3) to ac
            increment N(3) by 1
            wait until ((agv:cpmaq11 current=0) or (agv:cpmaq12
current=0) or (agv:cpmaq13 current =0))
            set Tempo to ac
            set espera(3) to Tempo - te(3)
            set Tespera(3) to Tespera(3) + espera(3)
            send to Pmaq1

        end
    else
        begin
            send to Pmaq1
        end
    end
end
```

PV2

```
begin
  if agv:cpmaq21 current=1 and agv:cpmaq22 current =1 then
    begin
      set te(5) to ac
      increment N(5) by 1
      wait until agv:cpmaq21 current=0 or agv:cpmaq22 current=0
      set Tempo to ac
      set espera(5) to Tempo - te(5)
      set Tespera(5) to Tespera(5) + espera(5)
      send to Pmaq2
    end
  else
    send to Pmaq2
  end
```

PV3

```
begin
  if agv:cpinsp1 current=1 and agv:cpinsp2 current =1 then
    begin
      set te(7) to ac
      increment N(7) by 1
      wait until (agv:cpinsp1 current=0) or (agv:cpinsp2 current=0)
      set Tempo to ac
      set espera(7) to Tempo - te(7)
      set Tespera(7) to Tespera(7) + espera(7)
    end
  end
```

```
        send to PInsp  
    end  
    else  
        send to PInsp  
end
```
