

**UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL**  
**ESCOLA DE ENGENHARIA**  
**PROGRAMA DE PÓS-GRADUAÇÃO EM ENGENHARIA DE PRODUÇÃO**

**ESTIMATIVAS DE CONFIABILIDADE DE PRODUTO A PARTIR DAS  
CONTAGENS MENSAS DE VENDAS E FALHAS AO LONGO DO  
PERÍODO DE GARANTIA.**

**Marcos Rebello Dillenburg**

**Porto Alegre, 2005**

**UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL**  
**ESCOLA DE ENGENHARIA**  
**PROGRAMA DE PÓS-GRADUAÇÃO EM ENGENHARIA DE PRODUÇÃO**

**ESTIMATIVAS DE CONFIABILIDADE DE PRODUTO A PARTIR DAS  
CONTAGENS MENSAS DE VENDAS E FALHAS AO LONGO DO  
PERÍODO DE GARANTIA.**

**Marcos Rebello Dillenburg**

**Dissertação apresentada ao Programa de Pós-Graduação em  
Engenharia de Produção da Universidade Federal do Rio  
Grande do Sul como requisito parcial à obtenção do título de  
Mestre em Engenharia.**

**Área de concentração: Engenharia de Qualidade**

**Porto Alegre, 2005**

**ESTIMATIVAS DE CONFIABILIDADE DE PRODUTO A PARTIR DAS  
CONTAGENS MENSAS DE VENDAS E FALHAS AO LONGO DO  
PERÍODO DE GARANTIA.**

**Marcos Rebello Dillenburg**

Esta dissertação foi julgada adequada para obtenção do título de MESTRE EM ENGENHARIA e aprovada em sua forma final pelo orientador e pela banca examinadora do curso de Pós-Graduação em Engenharia de Produção.

Orientador: \_\_\_\_\_

Prof. Dr. José Luís Duarte Ribeiro

Banca Examinadora:

- Prof. Flavio Sanson Fogliatto, Ph.D.
- Prof<sup>a</sup>. Dr<sup>a</sup>. Liane Werner
- Prof. Dr. Marcelo Maia Rocha

Coordenador: \_\_\_\_\_

Prof. Dr. José Luís Duarte Ribeiro

## **AGRADECIMENTOS**

À minha esposa e filhos, pela paciência ao longo das jornadas noturnas à frente do computador.

À Novus Produtos Eletrônicos Ltda, pelo apoio durante estes dois anos.

Ao meu orientador, professor José Luis Duarte Ribeiro, pelas contribuições, revisões e embasamento oferecidos ao longo do curso e durante a realização deste trabalho.

Aos professores e colegas do Programa de Pós-Graduação em Engenharia de Produção, pelos ensinamentos, apoio e bons momentos dentro e fora da sala de aula.

## SUMÁRIO

<b>1</b>	<b>INTRODUÇÃO.....</b>	<b>13</b>
1.1	COMENTÁRIOS INICIAIS.....	13
1.2	TEMAS E OBJETIVOS .....	14
1.3	JUSTIFICATIVA .....	14
1.4	MÉTODO .....	15
1.5	LIMITAÇÕES .....	16
1.6	ESTRUTURA.....	17
<b>2</b>	<b>REFERENCIAL TEÓRICO.....</b>	<b>18</b>
2.1	SISTEMAS REPARÁVEIS E NÃO REPARÁVEIS .....	18
2.2	FONTES DE DADOS PARA ESTUDOS DE CONFIABILIDADE .....	19
2.2.1	Tabelas de dados de vida de partes ou componentes do produto .....	19
2.2.2	Testes de vida.....	20
2.2.3	Análise de dados de falha no campo.....	20
2.3	CARACTERÍSTICAS DOS DADOS DE VIDA.....	21
2.4	FUNÇÕES BÁSICAS DA CONFIABILIDADE.....	22
2.5	MODELOS PARAMÉTRICOS DA CONFIABILIDADE .....	24
2.5.1	Distribuição Lognormal .....	25
2.5.2	Distribuição de Weibull .....	26
2.5.3	Estimativa de parâmetros Weibull por mínimos quadrados .....	27
2.5.4	Estimativa de parâmetros Weibull pelo método dos momentos.....	28
2.5.5	Estimativa de parâmetros Weibull pelo método dos momentos modificado.....	29
2.5.6	Estimativa de parâmetros Weibull pelo método da máxima verossimilhança... 30	
2.6	MODELOS NÃO-PARAMÉTRICOS DA CONFIABILIDADE .....	32
2.6.1	Estimador de Kaplan-Meier.....	32
2.6.2	Estimador de Nelson-Aalen .....	33
2.6.3	Análise de dados completos de vendas e falhas.....	34
2.6.4	Análise de dados marginais de vendas e falhas .....	36
2.7	ESTATÍSTICA BAYESIANA EM ESTUDOS DE CONFIABILIDADE .....	38
2.7.1	Estimativas de confiabilidade utilizando o modelo Bayesiano Gama .....	41

2.8	TESTE DE HIPÓTESE <i>WILCOXON SIGNED RANK TEST</i> .....	42
2.8.1	Procedimento .....	42
<b>3</b>	<b>ANÁLISE DE DADOS AGRUPADOS DE VENDAS E FALHAS .....</b>	<b>45</b>
3.1	REQUISITOS DA METODOLOGIA .....	45
3.1.1	Adequação a diferentes produtos, partes ou componentes .....	45
3.1.2	Facilidade de aquisição e registro dos dados de falha .....	46
3.1.3	Usuários novatos em técnicas de estimativa da confiabilidade .....	47
3.1.4	Características dos dados de falha .....	47
3.2	SIMULAÇÃO DE DADOS DE FALHAS .....	48
3.2.1	Simulador de dados completos de falha.....	48
3.2.2	Simulador de dados incompletos de falha .....	50
3.3	IMPLEMENTAÇÃO DA ANÁLISE DE DADOS AGRUPADOS COMPLETOS .....	51
3.3.1	Condições para aplicabilidade .....	51
3.3.2	Entrada dos dados de falha .....	52
3.3.3	Estimativa não paramétrica das probabilidades de falha .....	53
3.3.4	Suavização da estimativa pelo ajuste de modelo paramétrico .....	55
3.3.5	Verificação dos resultados .....	57
3.3.6	Validação da implementação .....	58
3.4	IMPLEMENTAÇÃO DA ANÁLISE DE DADOS AGRUPADOS INCOMPLETOS .....	61
3.4.1	Condições para aplicabilidade .....	62
3.4.2	Entrada dos dados de falha .....	62
3.4.3	Estimativa não paramétrica das probabilidades de falha .....	63
3.4.4	Suavização da estimativa pelo ajuste de modelo paramétrico .....	64
3.4.5	Verificação dos resultados .....	65
3.4.6	Validação da implementação .....	65
<b>4</b>	<b>ESTUDO DE CASOS .....</b>	<b>68</b>
4.1	METODOLOGIA DE AVALIAÇÃO .....	68
4.2	BAIXO NÚMERO DE FALHAS .....	68
4.2.1	Condições estabelecidas para o teste .....	69
4.2.2	Resultados das análises .....	69
4.3	VENDAS VARIÁVEIS .....	72
4.3.1	Condições estabelecidas para o teste .....	72
4.3.2	Vendas aleatoriamente distribuídas .....	72

4.3.3	Vendas crescentes .....	74
4.3.4	Vendas decrescentes .....	78
4.3.5	Meses sem vendas .....	78
4.4	IMPERFEIÇÕES NA DISTRIBUIÇÃO DE ORIGEM .....	79
4.4.1	Condições estabelecidas para o teste .....	79
4.4.2	Distribuição Weibull com ruído de 50% .....	80
4.4.3	Distribuição Lognormal com ruído de 50%.....	81
4.5	FATORES COMBINADOS.....	81
4.5.1	Condições estabelecidas para o teste .....	82
4.5.2	Resultados das análises .....	82
4.6	ANÁLISE DE RESULTADOS E RECOMENDAÇÕES PARA APLICAÇÃO.....	84
4.6.1	Estimativas não-paramétricas das probabilidades de falha.....	84
4.6.2	Estimativas paramétricas das probabilidades de falha.....	84
4.6.3	Recomendações para aplicação.....	85
<b>5</b>	<b>CONSIDERAÇÕES FINAIS .....</b>	<b>87</b>
5.1	CONCLUSÕES .....	87
5.2	TRABALHOS FUTUROS .....	89

## LISTA DE FIGURAS

Figura 1 – Gráficos hipotéticos das quatro funções básicas da confiabilidade .....	23
Figura 2 – Planilha de simulação de dados completos de falhas.....	50
Figura 3 – Planilha de entrada de dados completos de falha.....	53
Figura 4 – Visão geral da planilha de análise de dados completos .....	54
Figura 5 – Dados para validação gerados a partir de distribuição Weibull.....	59
Figura 6 – Resultado da validação com dados de distribuição Weibull.....	60
Figura 7 – Dados para validação gerados a partir de distribuição Lognormal .....	60
Figura 8 – Resultado da validação com dados de distribuição Lognormal.....	61
Figura 9 – Planilha de entrada de dados incompletos de falha.....	63
Figura 10 – Visão geral da planilha de análise de dados incompletos .....	64
Figura 11 – Resultado da validação com dados de distribuição Weibull.....	66
Figura 12 – Resultado da validação com dados de distribuição Lognormal.....	67
Figura 13 – Dados de avaliação com baixo número de falhas mensais .....	69
Figura 14 – Resultado da análise de dados completos com poucas falhas.....	70
Figura 15 – Resultado da análise de dados incompletos com poucas falhas.....	70
Figura 16 – Dados de avaliação com vendas variáveis de forma aleatória.....	73
Figura 17 – Resultado da análise de dados completos com vendas variáveis.....	73
Figura 18 – Resultado da análise de dados incompletos com vendas variáveis.....	74
Figura 19 – Dados de avaliação com vendas variáveis de forma crescente.....	75
Figura 20 – Resultado da análise de dados completos com vendas crescentes.....	76
Figura 21 – Resultado da análise de dados incompletos com vendas crescentes.....	77
Figura 22 – Dados de falha Weibull com ruído 50% e vendas aleatórias.....	83
Figura 23 – Análise de dados completos Weibull com ruído 50% e vendas aleatórias .....	83
Figura 24 – Análise de dados incompletos Weibull com ruído 50% e vendas aleatórias.....	84

## LISTA DE TABELAS

Tabela 1 – Estrutura geral de dados mensais completos de vendas e falhas .....	35
Tabela 2 – Estrutura geral de dados marginais de vendas e falhas.....	37
Tabela 3 – Exemplo de teste <i>Wilcoxon Signed Rank Test</i> .....	44
Tabela 4 – Exemplo de resultado do simulador de dados incompletos de falha.....	51
Tabela 5 – Valores iniciais e limites para os parâmetros das distribuições de suavização .....	57
Tabela 6 – Avaliação do efeito da variação no número de falhas nas estimativas.....	71
Tabela 7 – Avaliação do efeito da variação aleatória das vendas nas estimativas .....	74
Tabela 8 – Avaliação do efeito de vendas crescentes nas estimativas .....	77
Tabela 9 – Estimativas paramétricas de falhas com distribuição Weibull e ruído 50%.....	80
Tabela 10 – Estimativas paramétricas de falhas com distribuição Lognormal e ruído 50%....	81
Tabela 11 – Estimativas paramétricas de falhas Weibull, ruído 50% e vendas aleatórias .....	82

## LISTA DE ABREVIATURAS

- EM:** *Expectation Maximization* – Maximização da Esperança (Algoritmo EM)
- KM:** Kaplan-Meier (Estimador da função de confiabilidade)
- MLE:** *Maximum Likelihood Estimates* – Estimativa de máxima verossimilhança
- MTBF:** *Mean Time Between Failures* – Tempo médio entre falhas
- MTTF:** *Mean Time to Failure* – Tempo médio até a falha
- P&D:** Pesquisa e Desenvolvimento
- ROCF:** *Rate of Occurrence of Failures* – Taxa de ocorrência de falhas
- VBA:** *Visual Basic for Applications* (linguagem de programação)

## **RESUMO**

Esta dissertação aborda a estimativa das probabilidades de falha de um produto ao longo do período de garantia. As fontes de dados para esta estimativa são a quantidade de produtos vendidos e o número de falhas detectadas em cada mês. Duas metodologias não-paramétricas para esta análise são apresentadas e validadas. A metodologia de análise de dados completos requer o conhecimento da data de venda e de falha de cada produto. A metodologia de análise de dados incompletos requer apenas os totais de vendas e falhas em cada mês ao longo do período de garantia. Para os dois casos, é ainda implementada a suavização das probabilidades de falha estimadas, utilizando distribuições paramétricas Weibull ou Lognormal. As duas técnicas são implementadas em planilha eletrônica e aplicadas na análise de dados simulados. O desempenho de cada metodologia é avaliado com dados de diferentes características, resultando em recomendações para escolha e aplicação da metodologia mais adequada em cada caso.

## **ABSTRACT**

The subject of this work is the assessment of failure probabilities for a product during its warranty period. Data for this assessment are the sales and failure counts for each month of this period. Two nonparametric methods for this analysis are presented and validated. The Complete Data Analysis requires knowledge of the date of sale and failure for each product. The Incomplete Data Analysis requires only totals for sales and failures on each month. For both methods, the estimated failure probabilities are smoothed using parametric Weibull or Lognormal distributions. Both methods are implemented in electronic spreadsheet, and applied to the analysis of simulated data. Based on the performance of each method in the analysis of a number of different data, recommendations for selection and usage of the most suitable method are established.

## CAPÍTULO 1

### 1 INTRODUÇÃO

#### 1.1 COMENTÁRIOS INICIAIS

Ao longo da última década, muito foi feito nas empresas para aprimorar a gestão da qualidade e estimular a melhoria contínua dos processos, serviços e produtos. Na fabricação de produtos de alta tecnologia, a confiabilidade é uma métrica importante da qualidade final, mas sua previsão, monitoração e controle são muitas vezes esquecidos ou tratados de forma superficial, principalmente nas pequenas empresas.

A previsão da confiabilidade na fase inicial do projeto a partir de modelos de falha padronizados e a realização de testes acelerados de vida nos protótipos são algumas das ferramentas disponíveis para obtenção de uma estimativa inicial da confiabilidade e correção precoce de falhas que só seriam detectadas no uso real do produto, após seu lançamento. Os pequenos fabricantes de equipamentos de alta tecnologia, pressionados pela necessidade de desenvolvimento acelerado de seus produtos e limitados em recursos financeiros e humanos, geralmente desconhecem ou não aplicam estas ferramentas nos estágios iniciais do desenvolvimento, e muitas vezes o resultado é o lançamento de produtos com problemas crônicos de confiabilidade (MEEKER & HAMADA, 1995).

Uma vez no mercado, o produto é operado pelos clientes e eventualmente apresenta falhas que resultam no seu retorno ao fabricante, principalmente durante o período de garantia. É exatamente neste momento que existe uma excelente oportunidade de coleta de informações referentes à sua confiabilidade, que adequadamente registradas e processadas, permitem estimar parâmetros de confiabilidade do produto, sem gerar custos excessivos para o fabricante. A monitoração destes parâmetros permite detectar precocemente problemas, prever custos futuros, dimensionar estoque e equipe de manutenção e prover a área de Pesquisa e Desenvolvimento (P&D) com informações úteis ao aprimoramento deste e de futuros produtos (WU & MEEKER, 2002; MEEKER & HAMADA, 1995).

A adoção de uma metodologia de acompanhamento da confiabilidade baseada em informações de manutenção em garantia não deve, no entanto, ser a única iniciativa no

sentido de estimar e melhorar a confiabilidade. A detecção e correção de problemas de confiabilidade antes do lançamento do produto, preferencialmente nos estágios iniciais de seu desenvolvimento, sempre será menos onerosa, apesar de exigir maior investimento inicial em equipamentos e treinamento de pessoal (MARCORIN & ABACKERLI, 2002).

A análise de confiabilidade a partir de informações de garantia é um ponto de partida conveniente para a implantação gradual de um programa de confiabilidade, pois pode ser implantado rapidamente e com baixo investimento, contribuindo para difundir a cultura e os conceitos de confiabilidade na empresa.

## 1.2 TEMAS E OBJETIVOS

Os temas deste trabalho são a confiabilidade de produtos de alta durabilidade e os modelos de falha no período de garantia, estruturados a partir de informações de vendas e falhas por período.

O objetivo é propor uma metodologia de acompanhamento da confiabilidade aplicável por empresas pequenas, utilizando como fonte de dados os registros de vendas e de serviços de manutenção realizados dentro do período de garantia. A metodologia proposta é adequada para aplicação em produtos que, em caso de falha, são encaminhados pelo usuário para manutenção no fabricante ou seu agente autorizado. Além de estabelecer uma sistemática para a análise matemática dos dados, são estabelecidos critérios para a coleta e apresentação destes.

Este trabalho tem ainda como objetivo difundir e facilitar a aplicação de técnicas de monitoração de confiabilidade, estabelecendo uma metodologia simples e implementável em planilhas de cálculo, sendo assim acessível a empresas de qualquer porte.

## 1.3 JUSTIFICATIVA

A maior parte das empresas enfrenta ou enfrentará problemas com a confiabilidade de seus produtos, freqüentemente causados por um ou mais dos seguintes motivos: um modo de falha não previsto, condições operacionais piores do que as previstas, mudanças inesperadas na qualidade da matéria-prima ou alterações de projeto mal validadas. A adequada monitoração das falhas em garantia pode identificar estes problemas antecipadamente, permitindo a redução dos custos tangíveis e intangíveis da baixa confiabilidade (WU & MEEKER, 2002).

Ao medir a confiabilidade dos produtos em campo, o fabricante pode obter importantes informações para a melhoria do projeto e processo de fabricação, prever recursos para a assistência técnica, dimensionar prazos de garantia e detectar precocemente falhas relacionadas à confiabilidade do produto.

Para viabilizar a realização de estudos de confiabilidade nos produtos fabricados por empresas de pequeno porte, é importante que estes possam ser realizados a baixo custo e por pessoal não especializado na teoria de confiabilidade. Tipicamente as informações de falhas em garantia e volumes de venda já são registradas nestas empresas para fins de controle de custos da área de assistência técnica, controle e planejamento de produção e previsão de faturamento. Com procedimentos simples de refinamento destes dados é possível sua aplicação também no estudo de confiabilidade, gerando um mínimo de custos adicionais (WU & MEEKER, 2002; MARCORIN & ABACKERLI, 2002).

As informações de falhas em campo são tipicamente imprecisas. Não há informações exatas das datas em que o produto começou a ser utilizado e em que apresentou a falha e nem de sua taxa de utilização, o que dificulta estimar o tempo de operação até a falha, informação básica para os estudos usuais de confiabilidade. As incertezas existentes nos dados motivam a análise da confiabilidade a partir de informações mais simples de ser coletadas e analisadas, como as vendas e falhas mês a mês, sem a necessidade de rastrear a data de venda de cada produto que falhou (RAI & SINGH, 2003; KARIM, YAMAMOTO & SUZUKI, 2001).

Para análise dos dados coletados podem ser utilizados programas gratuitos de análise estatística, mas esses têm utilização complexa e pouco difundida. Programas específicos para análise de confiabilidade são caros e incorporam recursos desnecessários ao objetivo proposto. A implementação dos algoritmos de análise em planilha eletrônica não gera custos adicionais se a empresa já utiliza este tipo de programa, contribuindo para a redução de custo do estudo, sem prejuízo ao resultado final.

#### 1.4 MÉTODO

O desenvolvimento desta dissertação dividiu-se em cinco etapas. Na etapa de revisão da literatura foram analisados trabalhos anteriores referentes ao tema, com o objetivo de verificar alternativas para sua abordagem bem como aperfeiçoar os conhecimentos teóricos necessários à sua realização. A partir das informações obtidas na etapa inicial foram definidas as fontes de dados e as metodologias de análise mais alinhadas aos objetivos estabelecidos. Na etapa

seguinte foi realizada a implementação do método em planilha eletrônica, acompanhada de validação da implementação pela análise de dados simulados para os quais os resultados esperados eram conhecidos.

Na etapa de avaliação de desempenho do método de análise, optou-se pela utilização de dados simulados ao invés de dados reais. Esta abordagem permitiu analisar a robustez do método a dados oriundos de diferentes distribuições de falhas e melhor estabelecer suas limitações em futuras aplicações reais. Na etapa final foram elaboradas recomendações para aplicação do método, incluindo formas de coleta das informações de vendas e falhas e alternativas de análise para produto completo ou em partes (componentes).

## 1.5 LIMITAÇÕES

Os objetivos estabelecidos para este trabalho impõem algumas limitações à sua aplicabilidade. Uma primeira limitação se refere ao tipo de dado utilizado na análise. As descrições das falhas ocorridas são muitas vezes imprecisas, e não há tipicamente informações sobre a intensidade de uso dos produtos que falharam.

Para muitos tipos de produtos a ocorrência de falhas está mais relacionada à intensidade de uso (tempo e condições de operação) do que ao tempo decorrido entre a venda e a falha. Neste trabalho o tempo decorrido entre venda e falha é adotado como estimativa do tempo de uso do produto, o que pode criar distorção nos resultados caso um número significativo dos produtos que falham sejam operados a uma intensidade atípica ou fora de especificação. Não são estabelecidos mecanismos para minimizar este efeito.

Uma segunda limitação se origina na utilização apenas de dados de falha em garantia. Esta limitação é imposta pela metodologia proposta, que assume que todas as falhas ocorridas são levadas ao conhecimento do fabricante e incluídas no estudo, o que tipicamente só ocorre no período de garantia. Se o produto em análise apresenta uma taxa de falhas inicial alta (mortalidade infantil elevada) ou modos de falha que se manifestam apenas após um tempo de uso superior ao prazo de garantia, a estimativa de confiabilidade pode não refletir a realidade. Nenhum mecanismo para contornar esta limitação é proposto neste trabalho.

A terceira limitação se refere à validação do método apenas pela utilização de dados simulados. A utilização de dados reais originados por falhas de produtos diferentes e

fabricados por empresas diferentes permitiria uma melhor avaliação de seu desempenho em condições não-ideais.

## 1.6 ESTRUTURA

O primeiro capítulo desta dissertação apresenta os objetivos e justificativas de seu desenvolvimento, as limitações em sua aplicabilidade e a metodologia adotada em sua execução.

O segundo capítulo é a síntese da revisão bibliográfica realizada. Apresenta os tópicos básicos de confiabilidade necessários à compreensão e execução do trabalho, bem como as técnicas e resultados obtidos em trabalhos recentes na área de interesse.

O terceiro capítulo apresenta a metodologia de análise de confiabilidade proposta nesta dissertação e resultados de análises de dados simulados em planilha eletrônica.

O quarto capítulo estende a análise para diferentes tipos de dados simulados, apresentando situações limite de sua aplicação e comparando os resultados obtidos com os esperados. São apresentadas também recomendações para a aplicação em diferentes situações.

O quinto capítulo apresenta as conclusões obtidas e propõe sugestões de trabalhos futuros que podem aprimorar ou dar continuidade ao tema apresentado.

## CAPÍTULO 2

### 2 REFERENCIAL TEÓRICO

O objetivo deste capítulo é apresentar conceitos básicos necessários à compreensão do texto e revisar trabalhos já realizados que contribuem ao desenvolvimento do tema desta dissertação. Serve também como resumo das principais técnicas aplicadas na análise de dados de falha, citando as referências onde informações mais detalhadas podem ser obtidas.

Muitos dos tópicos abordados neste capítulo não são diretamente aplicados no desenvolvimento deste trabalho, mas sua análise e compreensão são necessárias para o bom entendimento do tema e para o desenvolvimento de estudos futuros.

#### 2.1 SISTEMAS REPARÁVEIS E NÃO REPARÁVEIS

Sistemas reparáveis são aqueles que após a ocorrência de uma falha podem ser colocados em condições satisfatórias de operação por alguma ação, incluindo substituição de peças ou ajustes. Na análise deste tipo de sistema são usualmente calculados a taxa de ocorrência de falha (ROCF – *Rate of Occurrence of Failure*) e o tempo médio entre falhas (MTBF – *Mean Time Between Failures*). A análise de sistemas reparáveis é aplicada por usuários de sistemas complexos, visando o planejamento da manutenção destes sistemas (NIST/SEMATECH, 2004).

Sistemas não reparáveis não podem ser novamente colocados em operação após a ocorrência de uma falha, sendo excluídos da população. Na análise deste tipo de sistema são usualmente calculados o tempo médio até a falha – MTTF (*Mean Time to Failure*) e a probabilidade de falha em diferentes tempos de uso. Muitas vezes sistemas reparáveis são analisados como não reparáveis, considerando-se apenas a primeira falha de cada sistema. A análise de sistemas não reparáveis é tipicamente aplicada pelo fabricante do produto ou sistema, visando controle e melhoria da confiabilidade, bem como dimensionamento de prazos e custos de garantia (NIST/SEMATECH, 2004).

A diferença básica entre as técnicas de análise para produtos reparáveis e não reparáveis está no cálculo da população de produtos em risco de falha. No caso de produtos não

reparáveis, unidades que falham são excluídas da população em risco para falhas posteriores. Em produtos reparáveis, as unidades que falham não são excluídas da população em risco para falhas posteriores, mas o estado do produto reparado precisa ser considerado no estudo.

Na análise de falhas em garantia de produtos reparáveis com alta durabilidade, o número de falhas esperado é baixo quando comparado com a população de produtos em risco. Como consequência, não é relevante excluir da população em risco as unidades que falharam e nem considerar a possibilidade de ocorrência de novas falhas no mesmo produto, o que permite a aplicação de técnicas para produtos não reparáveis, de implementação mais simples. Sistemas não reparáveis, ou tratados como tal, são o foco deste trabalho.

## 2.2 FONTES DE DADOS PARA ESTUDOS DE CONFIABILIDADE

De acordo com Marcorin & Abackerli (2002), a escolha das fontes de informação que serão utilizadas para realizar estudos de confiabilidade depende de diversos fatores: objetivo do estudo, tipo de produto em análise, estágio do produto em seu ciclo de vida, tempo disponível para testes e recursos disponíveis para o estudo. Para cada situação existem uma ou mais possíveis fontes de dados.

### 2.2.1 Tabelas de dados de vida de partes ou componentes do produto

Muitos componentes largamente utilizados na indústria têm características de confiabilidade conhecidas e divulgadas. É o caso de muitos componentes eletrônicos, rolamentos, vedações, dentre outros, que por seu volume de aplicação têm os parâmetros médios de distribuição de falhas conhecidos (MARCORIN & ABACKERLI, 2002). Por se tratar de médias, as condições de operação nem sempre coincidem com as condições da aplicação real, o que pode tornar impreciso o estudo.

Especificamente em eletrônica, a técnica mais comum para predição de confiabilidade se baseia em dados empíricos apresentados pelo Departamento de Defesa norte-americano no manual MIL-HDBK-217 (U.S. DEPARTMENT OF DEFENSE, 1991). As previsões são estabelecidas a partir da contagem e tipo de componentes utilizados no produto, além de considerações referentes ao seu ambiente e condições de operação. Os modelos de predição lá descritos são largamente aceitos na indústria, mas não são precisos em muitas situações, além de estarem desatualizados em relação aos novos tipos de encapsulamentos de componentes eletrônicos (KLEYNER & BENDER, 2003).

O acesso a este tipo de informação é rápido e de baixo custo, sendo o método aplicável desde o início do desenvolvimento do produto, resultando em boas estimativas iniciais para a confiabilidade.

### **2.2.2 Testes de vida**

Em testes de laboratório é possível simular condições específicas de utilização do produto, o que permite avaliar com precisão sua confiabilidade na condição estabelecida. Estes testes verificam o produto em poucos modos de falha previstos como mais críticos, produzindo resultados que podem não coincidir com os encontrados na utilização real do produto. Os testes exigem investimentos significativos em equipamentos, e sua duração pode ser muito longa se mecanismos de aceleração, medição de degradação ou censura não forem aplicáveis.

Os testes acelerados de vida são aplicados a componentes, materiais ou processos para determinar sua vida útil. O objetivo é identificar e quantificar as falhas e modos de falha que impedem a continuidade de uso do produto. A aplicação de testes acelerados requer que os mecanismos potenciais de falha e o modelo de aceleração sejam previamente conhecidos (MEEKER & HAMADA, 1995).

Testes de melhoria da confiabilidade são utilizados para determinar os tipos e níveis de estresse que podem causar falhas em um produto sabidamente sem defeitos. Estes testes são importantes para determinar a robustez do projeto do produto, detectando e corrigindo os pontos fracos antes do início de sua produção e comercialização.

Estes testes podem ser aplicados desde a fase de protótipo, principalmente em produtos simples ou sujeitos a poucos modos de falha predominantes. O investimento e o tempo de teste podem ser elevados, mas os resultados são bastante precisos.

### **2.2.3 Análise de dados de falha no campo**

Uma vez iniciada a utilização do produto, todo dado de falha é uma fonte potencial de informações para estudos de confiabilidade. No período de garantia, em que é mais provável que o consumidor acione o fabricante ou revendedor em caso de falha, são obtidas as informações mais completas, coletadas pela infra-estrutura de assistência técnica já existente na empresa, reduzindo o custo do processo de coleta de dados.

Ao coletar e analisar dados de falha em campo, fabricantes podem: prever falhas futuras, definir a necessidade de melhorias no produto ou processo, monitorar o efeito de mudanças no processo e condições de uso na confiabilidade, comparar taxas de falha entre produtos similares ou concorrentes. Os dados de garantia ainda descrevem custos, causas de falhas e variações do tipo de reclamação em função do período e local de fabricação (KARIM, YAMAMOTO & SUZUKI, 2001).

Muitos dos modos de falha dependentes do tempo de uso são pouco prováveis de ocorrer dentro do período de garantia. Mesmo dados incompletos de falhas fora deste período devem ser incorporados ao estudo para evitar estimativas excessivamente otimistas da confiabilidade em longo prazo (OH & BAI, 2001).

Apesar de ricos, por se originarem da utilização real do produto, os dados obtidos são tipicamente imprecisos em diversos aspectos: faltam informações referentes à taxa de uso do produto, tempo entre fabricação e início do uso e tempo entre a falha e o envio para manutenção. Também a descrição das falhas pelos clientes é muitas vezes imprecisa. A principal desvantagem desta fonte de dados é sua característica reativa, visto que os dados só se tornam disponíveis após o lançamento do produto no mercado, quando a correção de problemas de confiabilidade é mais complicada e onerosa. Sua utilização é indicada como importante complemento em um programa de melhoria da confiabilidade.

### 2.3 CARACTERÍSTICAS DOS DADOS DE VIDA

Na análise de dados de vida é desejável a utilização da maior quantidade de dados disponível, mas muitas vezes estes dados são incompletos ou incertos. De forma geral, os dados de vida podem ser classificados em: completos (toda informação é disponível) ou censurados (falta parte da informação) (RELIASOFT, 2004).

Quando os tempos até a falha de todas as unidades testadas são conhecidos, tem-se um conjunto de dados completos. Se os dados de tempo até a falha disponíveis são limitados a um tempo máximo, há censura do tipo I. Se os dados de tempo até a falha são limitados a um número máximo de falhas, há censura do tipo II (RAI & SINGH, 2003).

A censura é dita à direita se o tempo da falha não é conhecido, mas é sabido que o equipamento sobreviveu até um tempo determinado sem apresentar falha. Quando se sabe que um equipamento já havia falhado em um determinado instante de tempo, mas o tempo exato

da falha não é conhecido, tem-se uma censura à esquerda. Na censura por intervalo, um conjunto de equipamentos em teste é monitorado a intervalos regulares e, quando uma falha é encontrada, sabe-se que a mesma ocorreu dentro do intervalo, mas não seu momento exato. Se em um mesmo experimento de análise de vida existe mais de um ponto de censura, esta é dita censura múltipla (ESCOBAR & MEEKER, 1999).

Um caso típico de censura múltipla ocorre na análise de múltiplos grupos de produtos com entrada em operação escalonada. Neste tipo de análise, a idade dos produtos, as probabilidades de falha e o número de unidades em risco de falha variam de grupo para grupo, dificultando a análise (ESCOBAR & MEEKER, 1999). Um exemplo deste caso é a análise de confiabilidade a partir das informações de vendas e falhas registradas em intervalos fixos (KARIM, YAMAMOTO & SUZUKI, 2001).

A análise de dados de vida pode ser realizada tanto para dados completos quanto para dados censurados, utilizando as técnicas adequadas para cada caso.

## 2.4 FUNÇÕES BÁSICAS DA CONFIABILIDADE

As quatro funções básicas em estudos de confiabilidade, adaptadas de Fritsch (1996) e Ribeiro (2003) são:

- Função densidade de probabilidade de falha –  $f(t)$
- Função acumulada de falha –  $F(t)$ : Probabilidade de falha até a idade  $t$ .
- Função confiabilidade –  $R(t)$ : Probabilidade de sucesso além da idade  $t$ .
- Taxa de risco –  $h(t)$ : Taxa de falha dos sobreviventes na idade  $t$ .

Estas funções se relacionam conforme as equações a seguir.

$$F(t) = \int_0^t f(x)dx \quad (2.1)$$

$$R(t) = 1 - F(t) \quad (2.2)$$

$$f(t) = \frac{dF(t)}{dt} = \frac{-dR(t)}{dt} \quad (2.3)$$

$$h(t) = \frac{f(t)}{R(t)} \quad (2.4)$$

Supondo que  $n_0$  componentes idênticos são testados e que ao final do tempo  $t$ ,  $n_f(t)$  componentes falharam e  $n_s(t)$  componentes sobreviveram, as funções básicas da confiabilidade podem ser estimadas como:

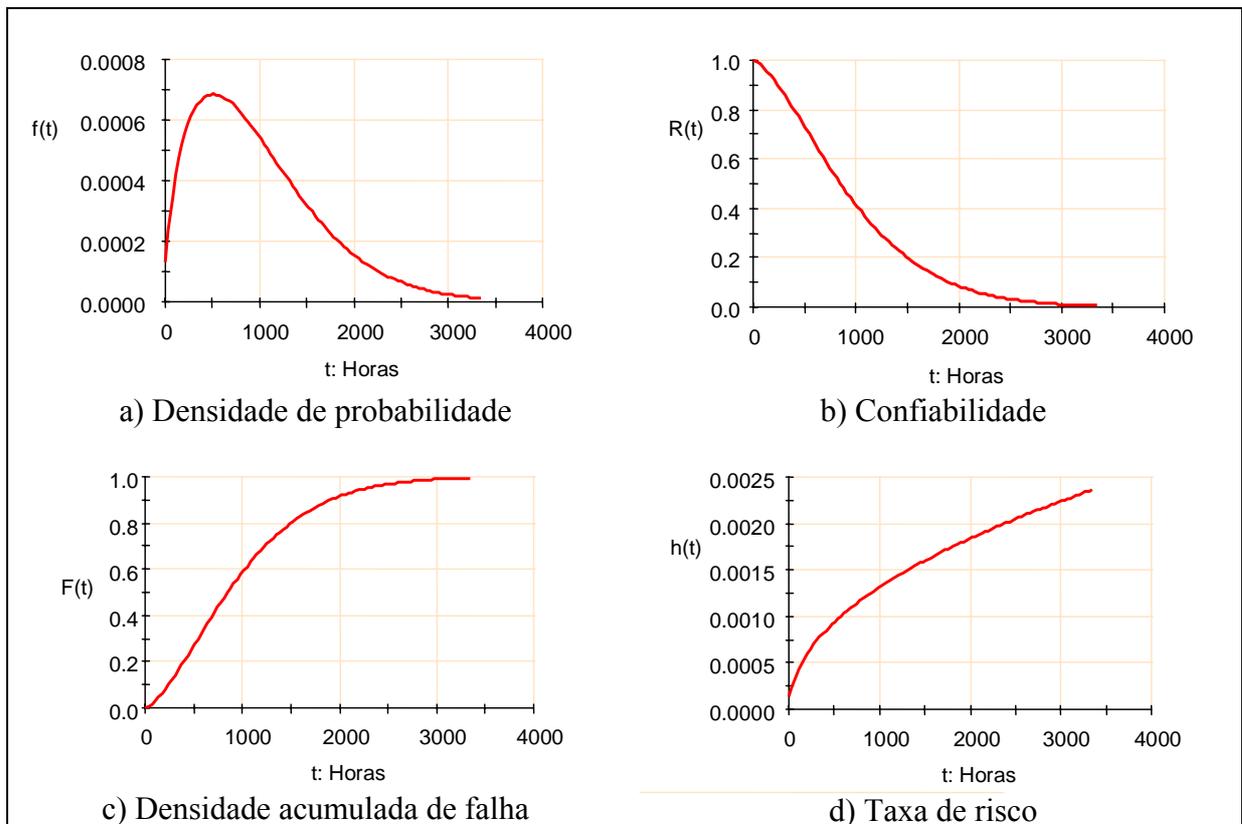
$$\hat{F}(t) = \frac{n_f(t)}{n_0} \quad (2.5)$$

$$\hat{R}(t) = \frac{n_s(t)}{n_0} \quad (2.6)$$

$$\hat{f}(t) = \frac{n_f(\Delta t)}{n_0 \Delta t} \quad (2.7)$$

$$\hat{h}(t) = \frac{n_f(\Delta t)}{n_s(\Delta t) \Delta t} \quad (2.8)$$

A Figura 1 a seguir apresenta representações gráficas hipotéticas para as quatro funções básicas da confiabilidade apresentadas:



Fonte: Elaborada pelo autor

Figura 1 – Gráficos hipotéticos das quatro funções básicas da confiabilidade

## 2.5 MODELOS PARAMÉTRICOS DA CONFIABILIDADE

Modelos matemáticos pré-definidos para a função confiabilidade podem ser ajustados aos dados de tempo até a falha, permitindo sua descrição concisa, uma vez que passam a ser descritos por um pequeno conjunto de parâmetros que caracterizam a distribuição adotada. A adoção de um modelo paramétrico permite previsões de confiabilidade pela extrapolação no tempo ao longo da distribuição, além de resultar em uma estimativa suave da distribuição de confiabilidade (MEEQUER, 2003).

A seleção de um modelo para os dados em análise se baseia em um ou mais critérios como: há um argumento físico ou estatístico que associa os dados de falha a um modelo específico, um modelo específico já foi utilizado com sucesso em um estudo de dados similares, a análise de aderência indica a adequação do modelo aos dados em análise (NIST/SEMATECH, 2004).

Argumentos físicos ou estatísticos para a adoção de um modelo para a confiabilidade são tipicamente baseados em modos de falha específicos (um determinado tipo de falha), e não refletem necessariamente as falhas ao nível de componente ou sistema. É necessário determinar o modelo adequado para cada modo de falha para permitir estudos paramétricos ao nível de componentes e sistemas, em uma análise de baixo para cima. Para combinar os diferentes modos de falha de um componente, pode ser utilizado o modelo *Competing Risk* e para combinar os modelos de falha dos diferentes componentes de um sistema são utilizados modelos para sistemas série, paralelo ou complexos (NIST/SEMATECH, 2004).

Em estudos de vida de componentes e sistemas eletrônicos, as distribuições mais utilizadas são: Exponencial, Weibull e Lognormal. A distribuição Exponencial se caracteriza por apresentar taxa de falha constante ao longo de toda a vida, sendo muitas vezes aplicada a componentes eletrônicos. A distribuição Lognormal é uma distribuição bastante genérica em estudos de confiabilidade, muitas vezes aplicada em estudos de número de ciclos até a fadiga. A distribuição de Weibull é a distribuição mais importante em estudos de confiabilidade, e dependendo dos valores de seus parâmetros pode representar taxas de falha decrescentes, constantes ou crescentes. Tem aplicação em ensaios de resistência à fadiga e em equipamentos ou sistemas eletrônicos (DODSON, 1994; RAC, 2001; RIBEIRO, 2003).

Uma vez escolhida a distribuição para os tempos até a falha, a determinação dos parâmetros que melhor descrevem os dados é feita por diferentes técnicas analíticas, como:

análise de regressão (mínimos quadrados), método da máxima verossimilhança, método dos momentos.

As diferentes técnicas têm em comum a complexidade matemática e, muitas vezes, a impossibilidade de automatizar completamente o cálculo, pela necessidade de julgamento dos resultados intermediários para a escolha da solução final mais adequada. A seguir são apresentadas resumidamente as distribuições Lognormal e Weibull, e técnicas de estimativa de parâmetros para a distribuição de Weibull.

### 2.5.1 Distribuição Lognormal

Uma variável aleatória  $X$  tem uma distribuição Lognormal se a variável  $Y=LN(X)$  tem uma distribuição normal, com  $LN$  denotando o logaritmo natural. A distribuição Lognormal é resultante da multiplicação de um conjunto de variáveis aleatórias. Iterações multiplicativas entre variáveis são encontradas em muitos processos naturais, sendo observadas em falhas de semicondutores, em processos de fadiga de materiais, crescimento geométrico de bactérias, fungos e corrosão (RIBEIRO, 2003; RAC, 2001).

A distribuição Lognormal para uma variável  $X$  pode ser especificada por seus parâmetros de média e variância, mas é mais usual sua representação utilizando os parâmetros de média ( $\mu$ ) e variância ( $\sigma^2$ ) de  $LN(X)$ . Sua função densidade de probabilidade expressa a três parâmetros, adaptada de RAC (2001) é:

$$f(t) = \frac{1}{(t-L)\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{\ln(t-L)-\mu}{\sigma}\right)^2} \quad (2.9)$$

onde  $\sigma$ ,  $\mu$  e  $L$  são os parâmetros da distribuição Lognormal.

No caso da distribuição Lognormal não há distinção clara entre parâmetro de forma e escala. O parâmetro de localização  $L$  não é usual, e só assume valor diferente de zero se há alguma evidência da impossibilidade de ocorrência de falhas nos tempos iniciais. Não há solução fechada para as funções de densidade acumulada de falha, confiabilidade e taxa de risco para a distribuição Lognormal. O valor de  $MTTF$  e variância  $var$  podem ser calculados como (RIBEIRO, 2003):

$$MTTF = e^{\left(\mu + \frac{\sigma^2}{2}\right)} \quad (2.10)$$

$$var = MTTF^2 \left( e^{\sigma^2} - 1 \right) \quad (2.11)$$

### 2.5.2 Distribuição de Weibull

A partir da teoria dos valores extremos é possível demonstrar que a distribuição de Weibull é o modelo para a distribuição dos mínimos de  $n$  valores provenientes de diversas distribuições iniciais limitadas à esquerda (RIBEIRO, 2003). Sistemas em que a falha de qualquer componente ou subsistema resulta na falha total (sistemas série) são fortes candidatos ao modelo Weibull. A capacidade de modelar tanto taxas de risco crescentes como decrescentes contribui para a ampla adoção dessa distribuição em estudos de confiabilidade de sistemas.

As equações (2.12), (2.13), (2.14) e (2.15) adaptadas de Ribeiro (2003) e Dodson (1994) representam, respectivamente, a função densidade de probabilidade de falha, a função confiabilidade, a função acumulada de falha e a função de risco para uma distribuição Weibull a três parâmetros.

$$f(t) = \frac{\gamma}{\theta} \left( \frac{t-L}{\theta} \right)^{\gamma-1} e^{-\left( \frac{t-L}{\theta} \right)^{\gamma}} \quad (2.12)$$

$$R(t) = e^{-\left( \frac{t-L}{\theta} \right)^{\gamma}} \quad (2.13)$$

$$F(t) = 1 - e^{-\left( \frac{t-L}{\theta} \right)^{\gamma}} \quad (2.14)$$

$$h(t) = \frac{\gamma}{\theta} \left( \frac{t-L}{\theta} \right)^{\gamma-1} \quad (2.15)$$

onde  $\gamma$  é o parâmetro de forma,  $\theta$  é o parâmetro de escala e  $L$  é o parâmetro de localização da distribuição.

O parâmetro de localização representa o deslocamento da distribuição ao longo do tempo, e permite maior flexibilidade na representação de processos que não apresentam falhas nos tempos iniciais. A distribuição de Weibull a dois parâmetros considera  $L = 0$ .

O valor do tempo médio até a falha (MTTF) no modelo Weibull pode ser calculado como:

$$MTTF = L + \theta \Gamma \left( 1 + \frac{1}{\gamma} \right) \quad (2.16)$$

onde  $\Gamma(\cdot)$  representa a função Gama.

### 2.5.3 Estimativa de parâmetros Weibull por mínimos quadrados

A estimativa de parâmetros pelo método dos mínimos quadrados é simples para funções que podem ser linearizadas, como é o caso da maior parte das distribuições aplicadas em estudos de confiabilidade. Para estas distribuições os cálculos são relativamente simples e diretos, tendo solução fechada que produz resultados sem iteração ou consulta a tabelas. Além disto, pode ser calculado o coeficiente de correlação que é uma medida adequada da aderência do modelo aos dados do experimento. O método dos mínimos quadrados é usualmente aplicado na análise de dados completos de tempos até a falha (RELIASOFT, 2004).

Se  $t_1, t_2, \dots, t_n$  formam um conjunto ordenado de  $n$  tempos até a falha, as estimativas dos parâmetros de forma ( $\gamma$ ) e escala ( $\theta$ ) de uma distribuição de Weibull utilizando regressão é obtida pelas equações a seguir, adaptadas de Fawzan (2000) e Reliasoft (2004):

$$\hat{\gamma} = \frac{n \sum_{i=1}^n x_i y_i - \sum_{i=1}^n y_i \sum_{i=1}^n x_i}{n \sum_{i=1}^n x_i^2 - \left( \sum_{i=1}^n x_i \right)^2} \quad (2.17)$$

$$\hat{\theta} = e^{(\bar{x} - \bar{y}/\hat{\gamma})} \quad (2.18)$$

onde:

$$x_i = \ln(t_i) \quad (2.19)$$

$$y_i = \ln \left( \ln \left( \frac{1}{1 - \frac{i}{n+1}} \right) \right) \quad (2.20)$$

$$\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i \quad (2.21)$$

$$\bar{y} = \frac{1}{n} \sum_{i=1}^n y_i \quad (2.22)$$

A fração no denominador da equação (2.20) é a estimativa para a função acumulada de falha baseada no cálculo *Mean Rank*. É usual também a adoção da aproximação de Benard para o *Median Rank* ou do estimador de Kaplan-Meier (RELIASOFT, 2004).

O coeficiente de correlação  $\rho$  é uma medida de quanto o modelo de regressão linear se adapta aos dados. Quanto mais próximo for de  $\pm 1$ , melhor o modelo de regressão. O coeficiente de correlação da população pode ser estimado pelo coeficiente de correlação da amostra, dado pela equação (2.23).

$$\hat{\rho} = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^n (x_i - \bar{x})^2 \sum_{i=1}^n (y_i - \bar{y})^2}} \quad (2.23)$$

#### 2.5.4 Estimativa de parâmetros Weibull pelo método dos momentos

Se  $t_1, t_2, \dots, t_n$  são amostras ordenadas de uma variável aleatória com função densidade de probabilidade  $f(t)$ , seu  $k$ -ésimo momento é dado por:

$$M_k = \int_{-\infty}^{+\infty} t^k f(t) dt \quad (2.24)$$

e sua estimativa amostral é dada por:

$$\hat{M}_k = \frac{1}{n} \sum_{i=1}^n t_i^k \quad (2.25)$$

Se as integrais teóricas dos momentos do modelo têm solução, os  $k$  parâmetros da distribuição podem ser obtidos a partir da igualdade entre as equações (2.24) e (2.25) para os  $k$  primeiros momentos.

Para a distribuição de Weibull, o  $k$ -ésimo momento é calculado como:

$$M_k = \theta^k \Gamma\left(\frac{k}{\gamma} + 1\right) \quad (2.26)$$

onde  $\Gamma(\cdot)$  representa a função Gama. A partir dos dois primeiros momentos da distribuição de Weibull pode ser calculado o Coeficiente de Variação CV, que depende somente de  $\gamma$  (FAWZAN, 2000):

$$CV = \frac{\sqrt{\Gamma\left(1 + \frac{2}{\hat{\gamma}}\right) - \Gamma^2\left(1 + \frac{1}{\hat{\gamma}}\right)}}{\Gamma\left(1 + \frac{1}{\hat{\gamma}}\right)} \quad (2.27)$$

Adotando o coeficiente de variação dos dados amostrais como estimativa para o resultado da equação (2.27), esta pode ser resolvida iterativamente, resultando na estimativa para o parâmetro  $\gamma$ . A estimativa para o parâmetro de escala  $\theta$  pode ser determinada a partir da equação (2.28):

$$\hat{\theta} = \left[ \frac{\bar{t}}{\Gamma\left(\frac{1}{\hat{\gamma}} + 1\right)} \right]^{\hat{\gamma}} \quad (2.28)$$

onde  $\bar{t}$  é a média dos dados de tempo até a falha da amostra.

### 2.5.5 Estimativa de parâmetros Weibull pelo método dos momentos modificado

Um melhor resultado na estimativa é esperado com a adoção de modificações no método dos momentos. Quando comparado com o método da máxima verossimilhança, o método dos momentos modificado apresenta melhores resultados em dados com forte assimetria e não é tendencioso em respeito à média ou variância (DODSON, 1994). As equações necessárias à estimativa dos dois parâmetros da distribuição Weibull são:

$$\frac{s^2}{(\bar{t} - t_1)^2} = \frac{\Gamma_2 - \Gamma_1^2}{\left[ \left( 1 - n^{-1/\gamma} \right) \Gamma_1 \right]^2} \quad (2.29)$$

$$\hat{\theta} = \frac{s}{\sqrt{\Gamma_2 - \Gamma_1^2}} \quad (2.30)$$

onde:

$$\Gamma_i = \Gamma \left( 1 + \frac{i}{\gamma} \right) \quad (2.31)$$

$s$  é o desvio padrão dos dados da amostra

$\bar{t}$  é a média dos dados de falha da amostra

$t_1$  é o primeiro dado de falha

### 2.5.6 Estimativa de parâmetros Weibull pelo método da máxima verossimilhança

Conforme Ribeiro (2003), a estimativa pelo método da máxima verossimilhança se baseia na função de verossimilhança da amostra, que expressa a probabilidade de se obter um conjunto de dados a partir de um determinado modelo de distribuição. Esta função contém os parâmetros desconhecidos do modelo. Os valores destes parâmetros que maximizam a verossimilhança da amostra são conhecidos como estimativas de máxima verossimilhança (MLE – *Maximum Likelihood Estimates*).

O método de estimativa de máxima verossimilhança é um procedimento analítico de maximização que se aplica a todas as formas de dados censurados ou completos e que apresenta propriedades desejáveis na análise de grandes amostras:

- À medida que o tamanho da amostra cresce, a estimativa se torna não tendenciosa.
- Intervalos de confiança e testes de hipóteses podem ser estabelecidos para os parâmetros, uma vez que a estimativa tende a uma distribuição normal.

A utilização do método tem, no entanto, alguns obstáculos importantes:

- Para amostras pequenas (menos de 10 falhas), a estimativa se torna altamente tendenciosa e métodos de correção devem ser aplicados.
- A solução das equações pode se tornar complexa, exigindo a utilização de programas estatísticos ou métodos iterativos em planilhas de cálculo.

A estimativa de máxima verossimilhança dos dois parâmetros de uma distribuição de Weibull a partir de dados completos ou censurados é obtida da solução das equações (2.32) e (2.33).

$$\frac{\sum_{i=1}^n t_i^{\hat{\gamma}} \ln(t_i)}{\sum_{i=1}^n t_i^{\hat{\gamma}}} - \frac{1}{\hat{\gamma}} - \sum_{i=1}^r \frac{\ln(t_i)}{r} = 0 \quad (2.32)$$

$$\hat{\theta} = \left[ \frac{\sum_{i=1}^n t_i^{\hat{\gamma}}}{r} \right]^{1/\hat{\gamma}} \quad (2.33)$$

onde:

$t_i$  são os tempos de falha ordenados, iniciando pelas observações não censuradas.

$r$  é o número de falhas.

$n$  é o número total de observações, censuradas e não censuradas.

A solução iterativa da equação (2.32) determina o parâmetro de forma  $\gamma$ , que aplicado à equação (2.33) resulta no parâmetro de escala  $\theta$ .

Para amostras pequenas ou com grande censura, a estimativa do parâmetro de forma é tendenciosa, e precisa ser corrigida. Bain-Engelhardt *apud* Fritsch (1996) propõe a adoção de um fator de correção  $G_n$ :

$$\hat{\gamma} = G_n \hat{\gamma}_{MLE} \quad (2.34)$$

que pode ser aproximado pela equação (2.35) proposta por Ribeiro *apud* Fritsch (1996):

$$G_n = 1,0 - \frac{1,346}{n} - \frac{0,8334}{n^2} \quad (2.35)$$

Diferentes alternativas para corrigir as estimativas de máxima verossimilhança para os parâmetros de uma distribuição Weibull são discutidas em Cacciari, Mazzanti e Montanari (1996). Aproximações para os intervalos de confiança podem ser encontradas em Dodson (1994).

## 2.6 MODELOS NÃO-PARAMÉTRICOS DA CONFIABILIDADE

Freqüentemente em análises de confiabilidade se deseja estimar parâmetros de confiabilidade sem restringi-los a uma distribuição paramétrica padronizada. Estas estimativas são chamadas não-paramétricas, e usualmente estimam a função confiabilidade ou a taxa de risco do produto em análise (FORSTER, 2004). A estimativa não-paramétrica pode ser aplicada no estudo de um modo de falha, componente ou sistema, mas tem validade apenas para o intervalo de tempo analisado. Extrapolações para previsões em tempos diferentes requerem informações complementares.

A habilidade em analisar dados sem assumir uma distribuição matemática para a confiabilidade evita a ocorrência de erros grosseiros ao se assumir erradamente uma distribuição. Por outro lado, os intervalos de confiança associados a uma análise não-paramétrica são usualmente muito mais largos que aqueles associados a uma análise paramétrica (RELIASOFT, 2004).

Na ausência de dados censurados, a estimativa óbvia da função de confiabilidade, baseada nos tempos até a falha  $t_1, \dots, t_n$  é (FORSTER, 2004):

$$\hat{R}(t) = \frac{\#(t_i > t)}{n} \quad (2.36)$$

onde:

$n$  é o número total de unidades em análise

$\#(t_i > t)$  é o número de unidades em risco (ainda operando) em  $t$

Na presença de dados censurados, a equação (2.36) não é mais adequada, pois considera a população ( $n$ ) constante ao longo de todo o teste. Na presença de censura, pode ser aplicado o estimador de Kaplan-Meier ou de Nelson-Aalen.

### 2.6.1 Estimador de Kaplan-Meier

Também conhecido como estimador produto limite, é utilizado no cálculo não-paramétrico da confiabilidade de dados contendo múltiplas falhas e censuras. O estimador de Kaplan-Meier é calculado conforme a seguir, adaptado de Forster (2004) e Reliasoft (2004).

A taxa de risco  $h(t)$  na idade  $t$  depende do número de falhas e da população em risco nesta idade, e é calculada conforme a equação (2.37) a seguir.

$$h(t_i) = \frac{r_i}{NR_i} \quad (2.37)$$

onde  $NR_i$  é o número de unidades em risco imediatamente antes de  $t_i$ , e  $r_i$  é o número de falhas em  $t_i$ . A população em risco deve ser calculada em cada instante em que a taxa de risco é calculada.

O estimador de Kaplan-Meier é utilizado para o cálculo da confiabilidade do produto na idade  $t$ , a partir dos valores das taxas de risco calculadas até o instante  $t$ , conforme a equação a seguir:

$$\hat{R}(t) = \prod_{t_i \leq t} (1 - h(t_i)) \quad (2.38)$$

onde  $h(t_i)$  são as taxas de falha calculadas nos instantes  $t_i$  até a idade  $t$ , usando (2.37).

A variância do estimador de Kaplan-Meier é dada pela fórmula de Greenwood:

$$\hat{Var}(\hat{R}(t)) = (\hat{R}(t))^2 \sum_{t_i \leq t} \frac{r_i}{NR_i(NR_i - r_i)} \quad (2.39)$$

Considerando que o estimador é assintoticamente normal, pode ser estabelecido seu intervalo de confiança pontual de 95% como:

$$\left[ \hat{R}(t) - 1,96\sqrt{\hat{Var}(\hat{R}(t))}, \hat{R}(t) + 1,96\sqrt{\hat{Var}(\hat{R}(t))} \right] \quad (2.40)$$

A taxa de risco cumulativa –  $H(t)$  – pode ser calculada diretamente a partir do estimador de Kaplan-Meier como:

$$\hat{H}(t) = -\log \hat{R}(t) \quad (2.41)$$

## 2.6.2 Estimador de Nelson-Aalen

Apesar de a taxa de risco cumulativa poder ser calculada diretamente a partir da estimativa de Kaplan-Meier para a função de confiabilidade, o estimador de Nelson-Aalen é mais utilizado quando o principal interesse é a função taxa de risco cumulativa, dada por (WOLFE, 2004):

$$\hat{H}(t) = \sum_{t_i \leq t} h(t_i) \quad (2.42)$$

A variância do estimador de Nelson-Aalen pode ser calculada como:

$$\hat{V}ar(\hat{H}(t)) = \sum_{t_i \leq t} \frac{r_i}{NR_i^2} \quad (2.43)$$

e o intervalo de confiança pontual de 95% como:

$$\left[ \hat{H}(t) - 1,96\sqrt{\hat{V}ar(\hat{R}(t))}, \hat{H}(t) + 1,96\sqrt{\hat{V}ar(\hat{R}(t))} \right] \quad (2.44)$$

A equação (2.45) apresentada por Jauw & Vassiliou (2000) e baseada no estimador de Nelson-Aalen, leva em conta o prazo de garantia do produto, resultando no cálculo da taxa média de risco dentro do período de garantia.

$$\hat{H}(t) = \sum_{j=t-(U-1)}^t \frac{r'_j}{\sum_{i=j-(U-1)}^j NR'_i} \quad (2.45)$$

onde  $t$  é o período de tempo atual (mês, semana),  $U$  é o período de garantia,  $r'_j$  é o número de falhas em garantia reportadas no período  $j$  e  $NR'_i$  é o número de unidades vendidas no período  $i$ .

### 2.6.3 Análise de dados completos de vendas e falhas

Nos momentos em que produtos em garantia apresentam falhas, seu fabricante tipicamente coleta informações referentes à data de venda, data da falha e tipo de defeito apresentado (KALBFLEISH, LAWLESS & ROBINSON, 1991). O fabricante também dispõe de informações referentes ao número de produtos vendidos ao longo do tempo. O conjunto de dados composto pelas datas e quantidades de produtos vendidos e pelas datas e quantidades de falhas entre estes produtos, pode ser agrupado em intervalos, tipicamente mensais, conforme ilustrado na Tabela 1. A probabilidade de falha na idade  $t$  pode ser estimada a partir destes dados, conforme adaptado de Kalbfleish, Lawless & Robinson (1991); Karim, Yamamoto & Suzuki (2001) e George (2004).

Tabela 1 – Estrutura geral de dados mensais completos de vendas e falhas

Mês da Venda	Vendas	Mês da Falha				
$X$	$N_x$	1	2	3	...	S
1	$N_1$	$r_{10}$	$r_{11}$	$r_{12}$	...	$r_{1,S-1}$
2	$N_2$		$r_{20}$	$r_{21}$	...	$r_{2,S-2}$
3	$N_3$			$r_{30}$	...	$r_{3,S-3}$
...	...				...	...
S	$N_S$					$r_{S0}$

Fonte: adaptado de Karim, Yamamoto & Suzuki (2001) e Yang & Zaghati (2002).

Para cada mês  $x$  ao longo de um período  $S$ , são registrados o número de produtos vendidos ( $N_x$ ) e o número de produtos vendidos no mês  $x$  e que apresentaram falha após  $t$  meses de uso, representados por  $r_{xt}$ .

O somatório de cada diagonal composta pelos valores  $r_{xt}$  com o mesmo valor de  $t$  resulta no número de produtos que falharam nesta idade, calculado usando a equação (2.46).

$$r_{\bullet t} = \sum_{x=1}^{S-t} r_{xt} \quad t = 0, \dots, S-1 \quad (2.46)$$

onde  $S$  é o número total de períodos em análise,  $r_{xt}$  são as contagens de falha e  $r_{\bullet t}$  é o número de produtos que falharam na idade  $t$ .

A determinação do número de unidades em risco de falha na idade  $t$  depende do número de unidades em operação com pelo menos esta idade e ainda dentro do intervalo de análise  $[1, S]$ . Na Tabela 1 pode ser observado que os  $N_S$  produtos vendidos no mês  $S$  são elegíveis para falhas apenas na idade  $t=0$ , enquanto que os  $N_1$  produtos vendidos no mês 1 são elegíveis para falhas em todas as idades analisadas. A equação (2.47) é utilizada no cálculo da população em risco de falha na idade  $t$ .

$$\begin{cases} NR_t = \sum_{x=1}^{S-t} N_x & t = 0 \\ NR_t = \sum_{x=1}^{S-t} \left[ N_x - \sum_{i=0}^{t-1} r_{xi} \right] & t = 1, \dots, S-1 \end{cases} \quad (2.47)$$

onde  $N_x$  é o número de unidades vendidas no mês  $x$ ,  $r_{xi}$  são as contagens de falhas,  $S$  é o número de períodos de análise e  $NR_t$  é o número de unidades em risco na idade  $t$ . O termo  $\sum r_{xi}$  em (2.47) exclui da população em risco as unidades que já falharam.

A partir das equações (2.46), (2.47) e (2.38) é possível calcular a estimativa para a Função Confiabilidade  $\hat{R}(t)$  na idade  $t$ :

$$\hat{R}(t) = \prod_{t_i \leq t} \left( 1 - \frac{r_{\bullet t_i}}{NR_{t_i}} \right) \quad (2.48)$$

A estimativa para a probabilidade de falha  $\hat{f}(t)$  pode ser calculada a partir das equações (2.3) e (2.48), resultando em:

$$\hat{f}(t) = -\frac{d}{dt} \hat{R}(t) \quad (2.49)$$

Considerando que as idades  $t$  são discretas e a diferença entre duas idades contíguas é sempre unitária ( $dt = 1$ ), a estimativa para a probabilidade de falha na idade  $t$  pode ser reescrita como:

$$\hat{q}_t = \hat{f}(t) = \hat{R}(t-1) - \hat{R}(t) \quad (2.50)$$

onde  $\hat{q}_t$  é a probabilidade de falha na idade  $t$ .

A partir dos valores de  $\hat{q}_t$  é possível estimar as falhas mês-a-mês para diferentes volumes de vendas, prever as falhas até a idade  $S-1$ , prever o percentual de falhas dentro do período  $S$  e estabelecer referências de *performance* para análise de melhorias no produto.

Em George (2004) é apresentado exemplo de planilha de cálculo para estimativa de confiabilidade a partir de dados completos de vendas e falhas, utilizando o estimador de Kaplan-Meier.

#### 2.6.4 Análise de dados marginais de vendas e falhas

Em muitas empresas a única informação disponível para estimativas de confiabilidade é o total de produtos vendidos a cada período (tipicamente mês) e o total de falhas registradas neste mesmo período. Esta informação – denominada contagem marginal – é incompleta, pois não relaciona as falhas à idade ou tempo de uso dos produtos. As técnicas estatísticas para

análise deste tipo de dado são descritas a seguir, sintetizando o trabalho de Karim, Yamamoto & Suzuki (2001). O principal benefício desta análise é a simplicidade em se coletar e registrar os dados de falha, pois não há necessidade de estabelecer a idade (ou tempo de uso) de cada produto que falhou.

Os dados marginais de falha podem ser resumidos como:  $N_x$ , número de produtos vendidos no mês  $x$  ( $x = 1, \dots, S$ );  $r_{xt}$ , número de produtos vendidos no mês  $x$  e que falharam após  $t$  meses de uso ( $t = 0, \dots, S-x$ );  $r_j$ , número de falhas ocorridas no mês  $j$ ,

$$r_j = \sum_{x=1}^{\min(S,j)} r_{x,j-x} \quad j = 1, \dots, S \quad (2.51)$$

Este conjunto de dados é muito semelhante ao descrito no item 2.6.3, mas na análise de dados marginais os valores  $r_{xt}$  não são conhecidos. As únicas informações utilizadas na estimativa de confiabilidade são as contagens de vendas  $N_x$  e a contagem de falhas  $r_j$ , coletadas para cada mês de interesse. Os dados marginais de vendas e falhas são organizados conforme a Tabela 2.

Tabela 2 – Estrutura geral de dados marginais de vendas e falhas

$x$	$N_x$	$r_x$
1	$N_1$	$r_1$
2	$N_2$	$r_2$
3	$N_3$	$r_3$
...	...	...
$S$	$N_S$	$r_S$

Fonte: adaptado de Karim, Yamamoto & Suzuki (2001)

No modelo estatístico adotado por Karim, Yamamoto & Suzuki (2001), os valores de  $r_{xt}$  para cada mês de venda são considerados como originários de uma distribuição de Poisson com média  $N_x q_t$ , onde  $q_t$  é o número médio de falhas por produto na idade  $t$ . Neste modelo, os valores de  $r_j$  também seguem uma distribuição de Poisson quando o número de unidades vendidas é alto e a probabilidade de falha é baixa. Para produtos bastante duráveis, o valor de  $q_t$  é baixo e pode ser interpretado como uma aproximação da probabilidade de falha na idade  $t$  (KALBFLEISH, LAWLESS & ROBINSON, 1991).

O objetivo é estimar  $q_t$  para os diferentes tempos de uso do produto. A solução é obtida pela aplicação do Algoritmo EM (*Expectation Maximization*), resultando nas equações (2.52) e (2.53), que iterativamente estimam a probabilidade de falha na idade  $t$  ( $\hat{q}_t$ ) a partir das observações de vendas ( $N_x$ ) e falhas ( $r_j$ ) no mês  $x$ .

$$E_{q_t}^{(k)}[r_{x,j-x} | r_j] = \frac{r_j N_x q_{j-x}^{(k)}}{\sum_{i=1}^{\min(S,j)} N_i q_{j-i}^{(k)}} \quad x = 1, \dots, \min(S, j) \quad (2.52)$$

$$\hat{q}_t^{(k+1)} = \frac{\sum_{x=1}^{\min(S,T-t)} E_{\hat{q}_t}^{(k)}[r_{xt} | r_{x+t}]}{\sum_{x=1}^{\min(S,T-t)} N_x} \quad t = 0, \dots, T-1 \quad (2.53)$$

A equação (2.52) corresponde à etapa *E* do algoritmo e corresponde à esperança dos valores  $r_{xt}$  dados os valores  $r_j$  observados. A equação (2.53) corresponde à etapa *M* e calcula o valor de  $\hat{q}_t$  que maximiza a esperança calculada. As duas equações são resolvidas iterativamente, sendo  $k$  o contador de iterações. O algoritmo tem convergência garantida e é concluído quando não há mais alteração significativa no valor de  $\hat{q}_t$  para novas iterações.

O modelo estatístico adotado na obtenção das equações (2.52) e (2.53) é adequado à análise de produtos reparáveis, mas pode ser aplicado na análise de produtos não reparáveis com baixa probabilidade de falha (produtos de alta durabilidade).

## 2.7 ESTATÍSTICA BAYESIANA EM ESTUDOS DE CONFIABILIDADE

A estatística Bayesiana tem aplicações crescentes em estudos de confiabilidade. A vantagem de sua utilização é a possibilidade de combinar informações de confiabilidade já conhecidas (predições, testes, avaliação de engenharia) com informações recentes (dados de campo ou novos testes) para obter uma estimativa de confiabilidade baseada em toda a informação disponível, tanto subjetiva como objetiva. Seu uso permite que estimativas de confiabilidade sejam continuamente atualizadas à medida que mais dados de testes e utilização em campo são obtidos (PONGO, FOLLEDO & BUENO NETO, 1996).

A estimativa inicial da confiabilidade é etapa essencial para o sucesso da aplicação do método. Informações subjetivas da equipe de desenvolvimento e resultados de testes em protótipos irão determinar o modelo matemático para a confiabilidade do produto.

Informações posteriores irão refinar os parâmetros deste modelo. A escolha arbitrária ou descuidada do modelo inicial para a confiabilidade pode tornar o resultado da análise Bayesiana impreciso ou enganoso (U.S. DEPARTMENT OF DEFENSE, 1998; NIST/SEMATECH, 2004).

O teorema de Bayes estudado na teoria probabilística básica é dado pela equação:

$$P(A|B) = P(A) \frac{P(B|A)}{P(B)} \quad (2.54)$$

A lei da probabilidade total é utilizada para o cálculo de  $P(B)$  no denominador de (2.54):

$$P(B) = \sum_{i=1}^n P(B|A_i)P(A_i) \quad (2.55)$$

O exemplo a seguir, adaptado de (U.S. DEPARTMENT OF DEFENSE, 1998), não tem aplicação prática, mas contextualiza na teoria da confiabilidade os vários termos das equações (2.54) e (2.55).

- $A$  Uma hipótese. Exemplo: “A confiabilidade desta peça após um ano de uso é 0,9”.
- $A_i$  Conjunto exaustivo de hipóteses mutuamente excludentes, incluindo a hipótese  $A$ . É o conjunto de todas as hipóteses para a confiabilidade da peça.
- $B$  Uma evidência (resultado de teste ou campo) que contribui para a credibilidade da hipótese. Exemplo: “A peça falhou com um ano de uso”.
- $P(A)$  Probabilidade inicial. É a probabilidade estabelecida antes de a evidência  $B$  estar disponível. Exemplo: “Há 50% de chance de a confiabilidade desta peça ser próxima a 0,9 após um ano de uso.  $P(A)=50\%$ ”.
- $P(B|A)$  Probabilidade de ocorrência da evidência  $B$  assumindo como verdadeira a hipótese  $A$ . Exemplo: “A probabilidade da falha com um ano de uso, considerando real a confiabilidade de 0,9 é obviamente 10%”
- $P(B)$  Probabilidade de ocorrência da evidência  $B$  considerando-se todo o conjunto de hipóteses  $A_i$ .
- $P(A|B)$  Confiabilidade resultante de  $A$ , ocorrida a evidência  $B$ .

A noção básica relacionada à abordagem Bayesiana é o uso da probabilidade subjetiva, isto é, o grau de convicção em um evento ou proposição. Para uma variável aleatória  $X$  (por exemplo, tempos até a falha) com função densidade de probabilidade  $f(x)$ , supõe-se que esta densidade depende de um parâmetro (ou conjunto de parâmetros)  $\lambda$ . Na teoria estatística

clássica o parâmetro é considerado uma constante fixa. Na teoria Bayesiana se introduz a convicção pessoal nos possíveis valores de  $\lambda$ , considerando-a como uma variável aleatória com densidade  $g(\lambda)$ . Ao se adotar uma distribuição para os valores de  $\lambda$  não se está dizendo que este não é uma constante, mas apenas que, pelo fato de não sabermos seu valor exato, mas termos uma razoável idéia de quais são seus possíveis valores com certezas diferentes, modelamos este conhecimento por uma distribuição de probabilidade (PONGO, FOLLEDO & BUENO NETO, 1996).

A equação básica de Bayes pode ser estendida a distribuições contínuas de probabilidade:

$$g(\lambda | x) = f(x | \lambda) \frac{g(\lambda)}{\int_0^{\infty} f(x | \lambda) g(\lambda) d\lambda} \quad (2.56)$$

onde  $g(\lambda)$  é o modelo inicial (*a priori*) da distribuição de valores do parâmetro de confiabilidade,  $f(x|\lambda)$  é a função densidade de probabilidade da informação observada, dado o parâmetro  $\lambda$  e  $g(\lambda|x)$  é o modelo final (*a posteriori*) da distribuição de valores do parâmetro de confiabilidade, dadas as observações  $x$ .

O denominador da equação (2.56) é uma constante em termos de  $\lambda$ , resultando na proporcionalidade:

$$g(\lambda | x) \propto f(x | \lambda) g(\lambda) \quad (2.57)$$

Se  $g(\lambda)$  e  $g(\lambda|x)$  são da mesma família de distribuições, então  $g(\lambda)$  e  $f(x|\lambda)$  são ditas distribuições conjugadas. Nestes casos, a aplicação do método é muito simplificada, pois a atualização do modelo se resume a uma simples troca nos valores dos parâmetros de  $g(\lambda)$ , e a metodologia pode ser repetida indefinidamente, uma vez que não há mudança no tipo de distribuição de  $g(\lambda)$ .

Formam pares conjugados as distribuições Gama–Exponencial, Normal–Normal e Binomial–Beta. Infelizmente a distribuição de Weibull a dois parâmetros, largamente adotada em estudos de confiabilidade, não possui uma distribuição conjugada contínua que simplifique a análise. Propostas e referências adicionais da aplicação de Bayes com a distribuição de Weibull são apresentadas em Kaminskiy & Kristov (2004) e Bacha & Celeux (1996).

O par de distribuições conjugadas Gama–Exponencial é muito utilizado em estudos Bayesianos de confiabilidade, sendo o modelo Gama associado ao parâmetro taxa de falhas. A taxa de falhas, por sua vez, é o parâmetro do modelo Exponencial da função de confiabilidade do produto. Esta análise é tipicamente aplicada em estudos de confiabilidade na região plana da “Curva da Banheira”, onde a taxa de falha é constante.

### 2.7.1 Estimativas de confiabilidade utilizando o modelo Bayesiano Gama

A distribuição Gama é conjugada da distribuição Exponencial e amplamente aplicada para expressar a função densidade de probabilidade para a taxa de falha de sistemas operando com taxa de risco constante (NIST/SEMATECH, 2004). Neste modelo, os tempos até a falha do produto em análise são modelados segundo uma distribuição Exponencial, que tem como parâmetro a taxa de falha  $\lambda$ . A primeira estimativa subjetiva de  $\lambda$  define sua faixa de valores prováveis na forma de uma distribuição Gama. Os parâmetros desta distribuição serão corrigidos quando novas informações de falha forem obtidas. A partir da distribuição de valores de  $\lambda$ , pode ser determinado seu valor médio e intervalo de confiança.

Os valores da taxa de falha  $\lambda$  seguem uma distribuição Gama:

$$g(\lambda) = \frac{\lambda^{\gamma-1}}{\theta^\gamma \Gamma(\gamma)} e^{-\frac{\lambda}{\theta}} \quad (2.58)$$

onde os parâmetros  $\gamma$  e  $\theta$  determinados de informações prévias (por exemplo, estimativas de confiabilidade em diferentes tempos de uso feitas por um conjunto de especialistas no produto). Os tempos até a falha seguem uma distribuição exponencial:

$$f(t) = \lambda e^{-\lambda t} \quad (2.59)$$

Supondo que um teste de vida com duração  $T$  resultou em  $r$  falhas, pode ser demonstrado a partir da equação (2.56) que a distribuição a posteriori  $g(\lambda|t)$  é também uma distribuição Gama com parâmetros:

$$\gamma' = \gamma + r \quad \theta' = \frac{\theta}{1 + T\theta} \quad (2.60)$$

Os novos parâmetros da distribuição a posteriori, calculados segundo (2.60), acumulam o conhecimento prévio e o novo conhecimento adquirido a partir dos dados do teste realizado.

Pela aplicação da função inversa desta nova distribuição, podem ser calculados o novo MTBF e seus limites de confiança. À medida que novas informações de vida são obtidas, o mesmo processo é repetido, e a distribuição de possíveis valores para o parâmetro  $\lambda$  é refinada.

## 2.8 TESTE DE HIPÓTESE *WILCOXON SIGNED RANK TEST*

O *Wilcoxon Signed Rank Test* é um teste não-paramétrico aplicado na verificação da diferença entre medianas de dados pareados, sendo o equivalente não-paramétrico ao teste-t para dados pareados.

O teste se baseia na magnitude da diferença entre pares de observações. Para aplicação de um teste-t é necessário que os valores das diferenças sigam uma distribuição normal. O teste de Wilcoxon não tem esta restrição, pois se baseia no *rank* de classificação das diferenças ao invés dos valores das diferenças. No entanto ainda se faz necessário supor que a distribuição das diferenças é simétrica (CRICHTON, 2000).

As hipóteses em teste são:

$H_0$ : A mediana da diferença de cada par é zero ( $\tilde{d} = 0$ ).

$H_1$ : A mediana da diferença de cada par é diferente de zero ( $\tilde{d} \neq 0$ ).

O nível de significância  $\alpha$  representa a probabilidade de a hipótese  $H_0$  ser rejeitada mesmo quando verdadeira, sendo usual o valor 5% em testes de hipótese.

### 2.8.1 Procedimento

As seguintes etapas, adaptadas de IFA (2004), são necessárias para aplicação do teste:

- i. Ordenar as diferenças de forma crescente, desconsiderando o sinal (valor absoluto). Omitir diferenças nulas e contar o número de pares com diferenças não nulas ( $N$ ).
- ii. Atribuir uma seqüência numérica crescente (*rank*) às diferenças ordenadas. Pares com o mesmo valor absoluto de diferença recebem o mesmo valor de *rank*, correspondente ao valor médio dos *ranks* caso os mesmos recebessem valores seqüenciais. Utilizar a Tabela 3 como exemplo.
- iii. Aplicar o sinal de cada diferença ao valor de *rank* correspondente.
- iv. Somar todos os valores de *rank* positivos ( $W_+$ ) e negativos ( $W_-$ ). O valor  $W$  é o menor valor absoluto entre  $W_+$  e  $W_-$ .

Para valores grandes de  $N$  ( $N \geq 10$ ), a distribuição das diferenças se aproxima da normal e sua variável reduzida  $Z_0$  pode ser calculada como:

$$Z_0 = \frac{W - 0,5 - N(N + 1)/4}{\sqrt{N(N + 1)(2N + 1)/24}} \quad (2.61)$$

O nível de significância bi-caudal (*ValorP*) da estatística  $Z_0$  calculada em (2.61) é determinado a partir da distribuição normal acumulada conforme a seguir:

$$\text{ValorP} = 2 \times (1 - D(Z_0)) \quad (2.62)$$

onde  $D(Z_0)$  é a função normal acumulada da variável padronizada  $Z$ .

O *ValorP* representa a probabilidade de obtenção dos resultados analisados sendo verdadeira a hipótese  $H_0$ . A hipótese  $H_0$  é rejeitada se  $\text{ValorP} < \alpha$ , onde  $\alpha$  é o nível de significância do teste, usualmente 0,05.

Para valores menores de  $N$  ( $N < 10$ ), a aproximação normal não é mais válida e a distribuição exata de todas as probabilidades precisa ser determinada. O algoritmo para cálculo do *ValorP* utilizando a distribuição exata de todas as  $2^N$  possibilidades de resultados em um teste com  $N$  diferenças não nulas é apresentado em IFA (2004). O algoritmo proposto se torna excessivamente lento para  $N > 10$ , devendo então ser adotada a aproximação normal.

A Tabela 3 apresenta um exemplo de aplicação do teste ao conjunto de dados apresentado em Lowry (2005).

Tabela 3 – Exemplo de teste *Wilcoxon Signed Rank Test*

$X_A$	$X_B$	$X_A - X_B$	$ X_A - X_B $	Rank	Signed Rank
78	78	0	0	—	—
24	24	0	0	—	—
64	62	2	2	1	1
45	48	-3	3	2	-2
64	68	-4	4	3,5	-3,5
52	56	-4	4	3,5	-3,5
30	25	5	5	5	5
50	44	6	6	6	6
64	56	8	8	7	7
50	40	10	10	8,5	8,5
78	68	10	10	8,5	8,5
22	36	-14	14	10	-10
84	68	16	16	11	11
40	20	20	20	12	12
90	58	32	32	13	13
72	32	40	40	14	14
				<b>W<sub>+</sub></b>	86
				<b>W<sub>-</sub></b>	-19
				<b>W</b>	19
				<b>Z<sub>0</sub></b>	2,07
				<b>ValorP</b>	0,038
				<b><math>\alpha</math></b>	0,05
				<b>Hipótese aceita:</b>	$H_1$

Fonte: Lowry (2005)

## CAPÍTULO 3

### 3 ANÁLISE DE DADOS AGRUPADOS DE VENDAS E FALHAS

Neste capítulo são apresentadas duas técnicas de análise de dados agrupados de vendas e falhas. A primeira técnica adaptada dos trabalhos de Kalbfleish, Lawless & Robinson (1991); Karim, Yamamoto & Suzuki (2001) e George (2004), aplica-se quando os dados de falhas são ditos completos, ou seja, se conhece a idade de cada produto que falhou. Os dados de falha são agrupados em intervalos de um mês e apresentados em forma tabular. A segunda técnica, proposta por Karim, Yamamoto & Suzuki (2001), aplica-se quando os dados de falha são ditos incompletos. Este tipo de dado, também chamado contagem marginal de falhas, não incorpora a idade de cada produto que falhou. As únicas informações disponíveis são o número de produtos vendidos e o número de produtos que falharam, agrupados em intervalos de um mês. O objetivo das duas técnicas é estabelecer uma estimativa não paramétrica da probabilidade de falha do produto nas diferentes idades.

As duas técnicas de análise são implementadas em planilha eletrônica *Excel* e validadas quanto a sua capacidade de reproduzir os dados de falha originais a partir dos parâmetros de confiabilidade estimados.

#### 3.1 REQUISITOS DA METODOLOGIA

Na seleção das metodologias de análise mais adequadas aos objetivos estabelecidos para este trabalho, quatro fatores principais foram considerados e são analisados a seguir.

##### 3.1.1 Adequação a diferentes produtos, partes ou componentes

As técnicas aplicadas neste trabalho se destinam à análise de confiabilidade de produtos com alta durabilidade. Pela baixa ocorrência de falhas no período de garantia, estes produtos são tratados como não reparáveis mesmo quando este não é o caso. Esta é tipicamente a análise de interesse para o fabricante do produto, sendo a análise de produtos reparáveis mais importante para usuários de um conjunto de equipamentos que precisam ser mantidos em operação.

Para garantir a aplicabilidade em diferentes tipos de produtos, partes ou componentes, não é assumido nenhum modelo paramétrico para modelagem dos tempos até a falha. A estimativa não-paramétrica da confiabilidade nas diferentes idades do produto é informação suficiente para os objetivos deste trabalho.

O ajuste de um modelo paramétrico é contemplado com o propósito de suavizar a estimativa não-paramétrica, e é estimado pelo ajuste de uma distribuição paramétrica às probabilidades de falha obtidas pela análise não-paramétrica.

### **3.1.2 Facilidade de aquisição e registro dos dados de falha**

Quando um produto falha e retorna ao seu fabricante ou assistência técnica autorizada, sua data de venda é verificada para estabelecer se o produto é elegível para a manutenção em garantia. A partir das datas de venda e de recebimento para assistência técnica, é possível estimar o tempo até a falha (ou idade do produto) como sendo a diferença entre estas datas (WU & MEEKER, 2002).

Informações apenas sobre os produtos que falham não são suficientes para estudos de confiabilidade. É necessária também informação quantitativa dos que não falharam para que possam ser estabelecidas proporções (taxas ou probabilidades) de falhas nas diferentes idades do produto. Esta informação é obtida a partir do número de produtos vendidos e ainda dentro do período de garantia.

Duas formas de aquisição e registro de dados de falha em garantia foram adotadas neste trabalho:

- Dados completos: A data de venda e falha de cada produto é conhecida.
- Dados incompletos: Apenas a data de falha é conhecida.

A análise de dados incompletos é necessária nos casos em que é difícil determinar a data da venda do produto que falhou, o que muitas vezes ocorre em produtos com grande volume de vendas e sem identificação de número de série ou lote de fabricação. A análise de dados completos pode ser usada nos casos em que a idade do produto pode ser estimada.

### **3.1.3 Usuários novatos em técnicas de estimativa da confiabilidade**

Uma vez que as técnicas descritas neste trabalho foram implementadas em uma planilha eletrônica, esta possivelmente será utilizada por pessoal não qualificado na teoria de confiabilidade. Para estimular a adoção da estimativa e monitoração da confiabilidade mesmo por empresas sem pessoal qualificado ou dedicado a esta tarefa, deve ser possível a obtenção de resultados úteis a partir de informações mínimas.

Muitas das técnicas de modelagem paramétrica requerem informações prévias ou análise de resultados intermediários de testes de aderência que podem inviabilizar sua adoção por estes usuários. Sem descartar a modelagem paramétrica, a técnica implementada produz resultados úteis a partir da modelagem não-paramétrica da probabilidade de falha nas diferentes idades do produto, resultado este que pode ser obtido instantaneamente após a entrada das informações de falhas e vendas na planilha eletrônica.

A adoção da metodologia Bayesiana foi descartada por apresentar maior benefício quando existe um conhecimento prévio sobre a confiabilidade do produto, o que não é o caso nas empresas que este estudo pretende atingir. Os resultados da aplicação deste trabalho na análise de um produto podem servir de informação prévia da confiabilidade para o estudo de outro produto semelhante, sendo, portanto, a análise Bayesiana uma possível extensão futura deste trabalho.

### **3.1.4 Características dos dados de falha**

A informação de tempo de uso baseada na diferença entre a data de venda e de entrada para assistência técnica apresenta incertezas: não é conhecido o retardo entre a venda e início de uso, não é conhecido o retardo entre a falha e o envio para manutenção e não é conhecida a intensidade de uso de cada produto que falhou. Se ao invés da data de venda é adotada a data de fabricação, há ainda a incerteza referente ao retardo entre a fabricação e venda (tempo de prateleira). Estas incertezas podem totalizar dias ou semanas no tempo de uso do produto ao longo do prazo de garantia. Para manter a simplicidade na aquisição dos dados, este trabalho desconsidera as incertezas referentes ao tempo e intensidade de uso do produto.

Estas incertezas na determinação do real tempo de uso do produto são acompanhadas ainda das incertezas referentes à caracterização da falha ocorrida. A correta determinação do problema ocorrido é importante para realimentação à equipe de desenvolvimento do produto.

Quando o foco do estudo é a melhoria de uma parte ou componente do produto, a análise de confiabilidade pode ser aplicada em separado aos modos de falha referentes a esta parte.

Para simplificar o registro e análise dos dados de vendas e falha, estes são agrupados em períodos mensais, mas as técnicas descritas aplicam-se para outros períodos de agrupamento. O agrupamento em períodos mensais é também adotado nos trabalhos de Karim, Yamamoto & Suzuki (2001); Wu & Meeker (2002); Yang & Zaghati (2002); Kalbfleish, Lawless & Robinson (1991) e Rai & Singh (2003).

### 3.2 SIMULAÇÃO DE DADOS DE FALHAS

A seguir são apresentados dois métodos para gerar dados de falhas a partir de volumes mensais de vendas e de valores discretos de probabilidades de falha nas diferentes idades. Estes métodos são aplicados na simulação de dados e também na previsão de falhas a partir dos modelos não-paramétricos estabelecidos.

Quando aplicados para simulação de dados de falhas, as probabilidades são geradas a partir de distribuições de Weibull ou Lognormal incluídas no simulador. Quando aplicados na previsão de falhas, as probabilidades de falha são as estimadas pelas metodologias de análise não-paramétrica.

#### 3.2.1 Simulador de dados completos de falha

O objetivo deste simulador é gerar um conjunto de dados de falha agrupados em intervalos mensais conforme apresentado na Tabela 1. O número de falhas ocorridas na idade  $t$  (em meses) para os produtos vendidos no mês  $x$  é representado como  $r_{xt}$ , e pode ser calculado utilizando a equação (3.1):

$$r_{xt} = N_x q_t \quad x = 1, \dots, S; t = 0, \dots, S - 1; t < x \quad (3.1)$$

onde  $N_x$  é o número de produtos vendidos no mês  $x$  e  $q_t$  é a probabilidade de falha na idade  $t$ . O resultado  $r_{xt}$  deve ser arredondado para um número inteiro de falhas.

A seguinte convenção referente à idade dos produtos é adotada em todo este trabalho:

- A idade da falha é igual à diferença em meses entre a falha e a venda.
- No cálculo da probabilidade de falha na idade  $t$  a partir de uma distribuição contínua, o valor de tempo utilizado é  $t + 0,5$ , que é o centro do intervalo de idades  $[t, t+1)$ .

Na implementação em planilha deste simulador, os valores de  $q_t$  podem ser gerados automaticamente a partir de funções densidade de probabilidade do tipo Weibull ou Lognormal. Um ruído aleatório pode ser incorporado a cada valor de  $q_t$  originado da distribuição, permitindo a simulação de distribuições imperfeitas. A amplitude deste ruído é estabelecida em percentual do valor de  $q_t$  para cada instante de tempo, sendo calculado conforme a seguir:

$$q_t = q_t + rnd\left(-q_t \frac{r_{\%}}{100}; q_t \frac{r_{\%}}{100}\right) \quad (3.2)$$

onde  $rnd(a ; b)$  é um gerador de números aleatórios entre  $a$  e  $b$  e  $r_{\%}$  é a amplitude definida para o ruído, em percentual de  $q_t$ .

As vendas mensais podem ser geradas aleatoriamente ou em forma de rampa linear entre dois limites estabelecidos, para até 36 meses. Rotinas em VBA (*Visual Basic for Applications*) associadas a botões são utilizadas para simulação destes dados. A Figura 2 apresenta parte da planilha implementada para simulação de dados de falha. Nesta figura as contagens de falha foram geradas a partir do modelo Weibull.

	A	B	C	D	E	F	G	H	I	J
1	<b>Weibull</b>		<b>Lognormal</b>		<b>Vendas</b>		Ruído na distribuição:		0%	
2	$\gamma$	1,7	$\sigma$	1	Inicial:	500	% Falhas no período:		2,7%	
3	$\theta$	50	$\mu$	3,4	Final:	2000				
4	L	0	L	0	Meses:	6				
5	MTTF	44,6	MTTF	49,4						
6	Aplicar		Aplicar		Aplicar		<input checked="" type="radio"/> Aleatória			
7							<input type="radio"/> Rampa			
8										
9	Idade - t			0	1	2	3	4	5	
10	Probabilidade de falha - $q_t$			0,001353	0,002913	0,004150	0,005228	0,006197	0,007084	
11	Falhas na idade - $r_t$			11	19	18	15	13	8	
12	Mês Falha			1	2	3	4	5	6	
13	Falhas no mês			2	4	9	14	23	32	
14	Mês Venda	Vendas	Falhas grupo							
15	1	1181	31	2	3	5	6	7	8	
16	2	1035	19		1	3	4	5	6	
17	3	725	10			1	2	3	4	
18	4	1557	13				2	5	6	
19	5	1894	9					3	6	
20	6	1295	2						2	

Fonte: Elaborada pelo autor

Figura 2 – Planilha de simulação de dados completos de falhas

O percentual de falhas no período indicado na Figura 2 é a soma acumulada dos valores de  $q_t$  até a última idade simulada, e corresponde ao percentual esperado de falhas até esta idade.

### 3.2.2 Simulador de dados incompletos de falha

O objetivo deste simulador é calcular o número de falhas no mês  $x$  a partir do total acumulado de vendas até este mês e das probabilidades de falha nas idades até  $x-1$ . Este simulador foi implementado na forma de uma função em VBA, e pode ser utilizado em qualquer planilha que requeira a simulação do total de falhas mensais. Para o cálculo do total de falhas no mês  $x$ , é utilizada a seguinte equação:

$$r_x = \sum_{i=1}^x N_i q_{x-i} \quad x = 1, \dots, S \quad (3.3)$$

onde  $r_x$  é o total de falhas no mês  $x$ ,  $q_{x-i}$  é a probabilidade de falha na idade  $x-i$ ,  $N_i$  é o número de produtos vendidos no mês  $i$  e  $S$  é o total de períodos analisados. O resultado  $r_x$  deve ser arredondado para um número inteiro de falhas.

A Tabela 4 é um exemplo de utilização do simulador. Os dados na coluna Falhas Simuladas foram gerados pela função VBA que implementa a Equação (3.3). Os dados de entrada para a função são as colunas ‘Mês Venda’, ‘Vendas’, ‘ $q_t$ ’ e o número do mês para o

qual é gerada a contagem de falhas. A contagem de falhas resultante neste exemplo é igual ao total de falhas a cada mês gerado pelo simulador de dados completos apresentado na Figura 2, uma vez que os dados de entrada utilizados (Vendas e  $q_t$ ) foram os mesmos.

Tabela 4 – Exemplo de resultado do simulador de dados incompletos de falha

<b>Mês Venda</b>	<b>Vendas</b>	<b>Falhas Simuladas</b>	<b>Idade</b>	<b><math>q_t</math></b>
1	1181	2	0	0,001353
2	1035	4	1	0,002913
3	725	9	2	0,004150
4	1557	14	3	0,005228
5	1894	23	4	0,006197
6	1295	32	5	0,007084

Fonte: Elaborada pelo autor

### 3.3 IMPLEMENTAÇÃO DA ANÁLISE DE DADOS AGRUPADOS COMPLETOS

A técnica descrita no item 2.6.3 se aplica na análise de contagens de vendas e falhas, estando os dados de falhas agrupados de acordo com a data de venda e data de falha dos produtos. Esta análise tem como principal vantagem a capacidade de produzir boas estimativas das probabilidades de falha mesmo quando o número de falhas é pequeno. Sua principal desvantagem é a necessidade de se estabelecer a idade de cada produto que falhou. A idade pode ser estimada como sendo a diferença entre a data de venda e de falha. A data da venda pode ser determinada por diferentes métodos: a partir da nota fiscal de compra apresentada pelo proprietário do produto; por consulta do número de série nos registros de vendas ou por marcação aplicada ao produto no momento da fabricação ou venda.

A implementação em planilha de cálculo pode ser toda realizada sem a utilização de *Scripts* ou funções VBA. Duas planilhas foram implementadas, uma para entrada dos dados de falha e outra para a análise de confiabilidade.

#### 3.3.1 Condições para aplicabilidade

Em duas situações específicas a adoção desta técnica é recomendada:

- Quando a empresa já tem implementada uma sistemática que permita rastrear a data de venda de cada produto que falha, e conta com um sistema adequado de bancos de dados para registro e posterior seleção das informações para análise. Se esta condição é atendida, fica atenuada a principal desvantagem do método.

- Quando o número de falhas dentro do período de análise é pequeno, tanto pela alta confiabilidade do produto quanto pelo pequeno número de unidades vendidas. Nesta condição, os resultados são mais precisos que os obtidos pela técnica de análise de dados marginais de vendas e falhas.

Se nenhuma das situações anteriores se aplica, a opção é determinada pela possibilidade ou não de implantar um sistema de registro de vendas e falhas que viabilize a obtenção de dados conforme apresentados à Tabela 1. É importante ainda considerar algumas restrições:

- Toda a falha ocorrida deve se originar dos produtos vendidos incluídos na análise. Esta ressalva é importante para o caso de a técnica ser aplicada para um produto que já estava sendo vendido antes do primeiro mês analisado. Neste caso, as falhas de produtos oriundos destas vendas anteriores devem ser eliminadas da análise, mesmo que ainda estejam ocorrendo dentro do prazo de garantia.
- Todas as falhas ocorridas devem ser do conhecimento do fabricante e incluídas na análise. Esta condição é tipicamente atendida se o período de análise  $S$  é no máximo igual ao período de garantia do produto.

### **3.3.2 Entrada dos dados de falha**

Os dados de falha são agrupados em contagens mensais de falhas de acordo com o mês de venda e falha de cada produto. A Figura 3 apresenta parte da planilha de entrada de dados. Esta planilha está preparada para receber dados de até 36 meses de vendas e falhas. O botão “Simular Falhas” é utilizado para importar dados de vendas e falhas do simulador de dados completos descrito no item 3.2.1.

	A	B	C	D	E	F	G	H	I
1									
2									
3	Simular Falhas								
4				<b>Mês da Falha</b>					
5	<b>Mês da Venda</b>			jun/04	jul/04	ago/04	set/04	out/04	nov/04
6	<b>Rótulo</b>	<b>Mês</b>	<b>Vendas</b>	<b>1</b>	<b>2</b>	<b>3</b>	<b>4</b>	<b>5</b>	<b>6</b>
7	jun/04	1	1181	2	3	5	6	7	8
8	jul/04	2	1035		1	3	4	5	6
9	ago/04	3	725			1	2	3	4
10	set/04	4	1557				2	5	6
11	out/04	5	1894					3	6
12	nov/04	6	1295						2

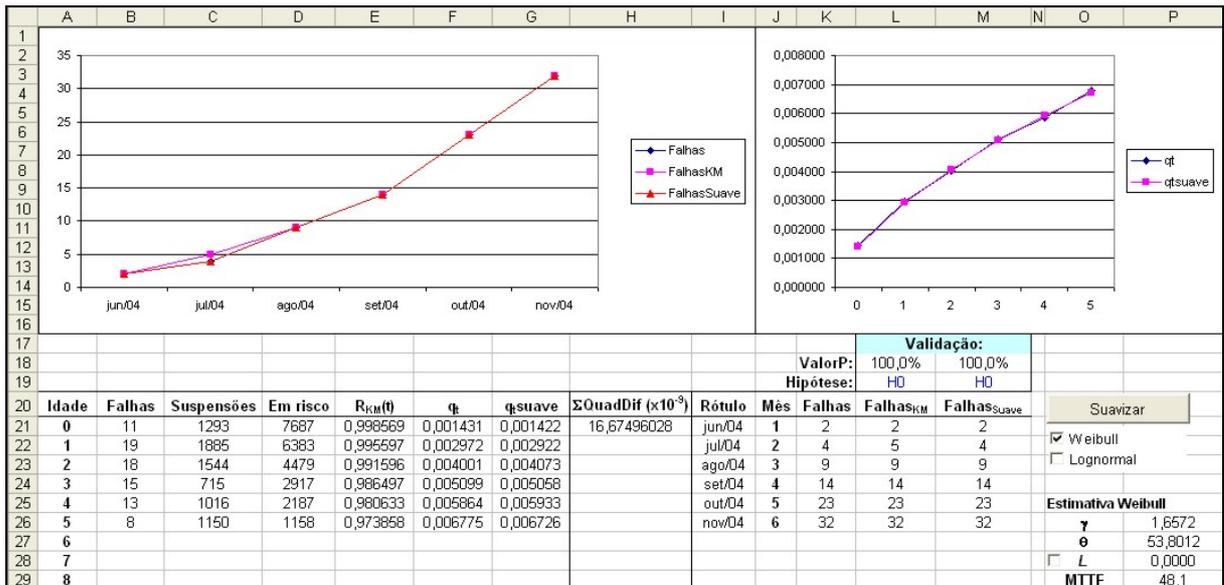
Fonte: Elaborada pelo autor

Figura 3 – Planilha de entrada de dados completos de falha

Cada linha representa um grupo de unidades vendidas (todas as unidades vendidas em um dado mês), e cada célula de falha é preenchida com o número de unidades do grupo que falharam em cada mês a partir do próprio mês da venda. Pelo fato de um produto não apresentar falha em garantia antes de sua venda, o conjunto das contagens de falha forma uma matriz triangular superior, onde a soma de cada diagonal representa o total de falhas ocorridas entre produtos com uma mesma idade.

### 3.3.3 Estimativa não paramétrica das probabilidades de falha

Uma segunda planilha é utilizada na análise dos dados. Os cálculos nesta planilha são divididos em duas etapas: estimativa não paramétrica da probabilidade de falha nas diferentes idades e suavização da estimativa por ajuste de modelo paramétrico. A Figura 4 apresenta uma visão geral desta planilha.



Fonte: Elaborada pelo autor

Figura 4 – Visão geral da planilha de análise de dados completos

A coluna “Falhas” contém o número de falhas ocorridas em cada idade do produto. É calculada a partir dos dados completos apresentados à Figura 3, pela aplicação da equação (2.46), com valor  $S = 6$ . Cada célula nesta coluna corresponde à soma de uma diagonal do conjunto completo de falhas.

A coluna “Suspensões” contém o número de unidades que serão excluídas da análise na idade seguinte por estarem impossibilitadas de atingir esta idade por censura no tempo. Para melhor compreender as suspensões, observar a Figura 3. Nesta, só são incluídas na análise as falhas ocorridas até o mês seis. Para as unidades vendidas no próprio mês seis, apenas as falhas ocorridas na idade zero são consideradas, devendo todas as unidades vendidas neste mês serem desconsideradas na análise de falhas para idades posteriores a zero. O mesmo ocorre para todos os grupos de unidades vendidas, sendo elegíveis para falhas em todas as idades analisadas apenas as unidades vendidas no primeiro mês da análise. A censura nesta análise é dita múltipla, pois apresenta valores de tempo diferentes para cada grupo de unidades vendidas.

A coluna “Em risco” contém o número de unidades em risco de falha em cada idade. Na idade zero, o total de unidades vendidas em todos os meses compõe as unidades em risco, mas nas idades seguintes devem ser excluídas as unidades que já falharam (coluna “Falhas”) e unidades não vendidas (coluna “Suspensões”). O cálculo das unidades em risco a partir do total de unidades vendidas e do abatimento de falhas e suspensões produz o mesmo resultado

que a equação (2.47), e foi adotado nesta planilha para permitir a implementação na planilha eletrônica sem a necessidade de implementação de uma função VBA.

A partir da população em risco e do número de ocorrências de falha em cada idade, é aplicada a equação (2.48) para o cálculo da estimativa de confiabilidade apresentada na coluna “ $R_{KM}(t)$ ”. A probabilidade de falha  $\hat{q}_t$  é calculada a partir da confiabilidade pela aplicação da equação (2.50) e resulta na coluna “ $q_t$ ”.

Os valores na coluna “ $q_t$ ” formam a estimativa não paramétrica da probabilidade de falha na idade  $t$ , objetivo desta análise. A partir desta estimativa, podem ser realizadas simulações de falhas para diferentes quantidades de unidades vendidas ou prever as falhas subsequentes censuradas na análise. Utilizando a equação (3.1) é possível estimar as ocorrências mensais de falhas entre os produtos vendidos em um dado mês, e utilizando a equação (3.3) é possível estimar o total mensal de falhas entre todos os produtos vendidos.

A análise comparativa dos valores de  $\hat{q}_t$  entre dois produtos ou de um mesmo produto antes e após alterações em seu projeto podem ser informações úteis para a área de desenvolvimento. O somatório de todos os valores  $\hat{q}_t$  entre a idade zero e a idade máxima de um produto em garantia é a probabilidade acumulada de falha em garantia, que indica a fração dos produtos vendidos que apresentam falhas neste período.

### **3.3.4 Suavização da estimativa pelo ajuste de modelo paramétrico**

A estimativa não paramétrica dos valores de  $\hat{q}_t$  é resultado suficiente para os objetivos propostos neste trabalho. No entanto, benefícios adicionais podem ser obtidos pelo ajuste de uma distribuição paramétrica aos dados de falha. Um primeiro benefício é a suavização da estimativa não paramétrica. Não é de se esperar mudanças abruptas na probabilidade de falha em curtos períodos de tempo, mas muitas vezes as imprecisões em sua estimativa resultam em valores que variam excessivamente, aumentando e diminuindo repetidamente em diferentes idades, situação pouco provável de estar associada a alguma característica do produto.

Um segundo benefício é a associação de um modelo matemático à probabilidade de falha, o que em teoria permite extrapolar as estimativas para tempos fora do intervalo de idades pesquisado e calcular o tempo médio entre falhas. Na prática este benefício dificilmente poderá ser aplicado neste trabalho, principalmente pela limitação do tempo de

análise, restrito ao prazo de garantia do produto. Os modos de falha que irão predominar no prazo de garantia são aqueles associados à mortalidade infantil de componentes e a falhas de montagem, sendo pouco provável que alguma falha associada à degradação de componentes do produto seja capturada e incluída no modelo (WILKINS, 2002).

O modelo paramétrico é obtido pelo ajuste de uma distribuição de Weibull ou Lognormal às probabilidades de falha calculadas, utilizando a técnica dos mínimos quadrados. A minimização da soma é calculada iterativamente pela variação dos parâmetros da distribuição utilizando a ferramenta *Solver* do *Excel*. A técnica aqui adotada para a estimativa de parâmetros não deve ser confundida com a técnica descrita no item 2.5.3, que utiliza mínimos quadrados para estimar os parâmetros da distribuição Weibull a partir de um conjunto de tempos até a falha.

A determinação dos parâmetros do modelo diretamente a partir dos tempos até a falha foi descartada neste trabalho pelos seguintes motivos: i) o objetivo principal é a suavização da estimativa não paramétrica das probabilidades de falha, devendo ser estes os dados de entrada para o ajuste, ii) a disponibilidade no *Excel* da ferramenta de solução numérica *Solver*, iii) a facilidade de se incluir novas distribuições para ajuste, o que requer apenas a implementação no *Excel* do cálculo de sua função densidade de probabilidade.

Na Figura 4, o botão “Suavizar” é utilizado para iniciar a execução de uma função em VBA que executa a suavização dos dados na coluna “ $q_t$ ”, utilizando o modelo paramétrico Weibull ou Lognormal, conforme seleção nas caixas de opção correspondentes. O resultado é apresentado na coluna “ $q_t$ suave” e os parâmetros de ajuste da distribuição são apresentados na coluna “Estimativa Weibull” ou “Estimativa Lognormal”, conforme o caso. As etapas executadas na suavização são:

- Estimativa do parâmetro de localização;
- Configuração dos parâmetros iniciais do modelo;
- Configuração das restrições aos valores dos parâmetros;
- Execução do *Solver*.

O parâmetro de localização para a distribuição é estabelecido pelo operador ou automaticamente a partir da contagem do número de probabilidades de falha na coluna “ $q_t$ ” que são menores que um limiar estabelecido para a probabilidade nula. Esta contagem se

inicia na idade zero e é encerrada na primeira ocorrência de probabilidade acima do limiar. O limiar adotado neste trabalho é 0,000001.

Pela natureza iterativa da estimativa, valores iniciais para os parâmetros da distribuição precisam ser estabelecidos. A escolha destes parâmetros tem influência no número de iterações e na convergência da solução. Experimentalmente foram determinados os valores iniciais apresentados na Tabela 5.

A ferramenta *Solver* permite que restrições sejam estabelecidas para os parâmetros variáveis do ajuste. Para reduzir o número de iterações e evitar a convergência com parâmetros não usuais, os valores mínimo e máximo para cada parâmetro da distribuição são estabelecidos conforme apresentado na Tabela 5.

Tabela 5 – Valores iniciais e limites para os parâmetros das distribuições de suavização

<b>Modelo Weibull</b>		<b>Modelo Lognormal</b>	
$\gamma_0$	4	$\sigma_0$	1
$\theta_0$	50	$\mu_0$	1
$\gamma_{min}$	1	$\sigma_{min}$	0,1
$\gamma_{máx}$	4	$\sigma_{máx}$	10
$\theta_{min}$	1	$\mu_{min}$	1
$\theta_{máx}$	500	$\mu_{máx}$	100

Fonte: Elaborada pelo autor

Na execução do *Solver*, o objetivo é minimizar a soma dos erros quadráticos entre as colunas “ $q_t$ ” e “ $q_{t\text{suave}}$ ”, apresentada na coluna “ $\Sigma\text{QuadDif}$ ”, variando para isto os parâmetros da distribuição nas células P26 e P27 da Figura 4. O gráfico à direita nesta figura apresenta comparação entre os valores de “ $q_t$ ” e “ $q_{t\text{suave}}$ ”. Se as duas caixas de opção “Weibull” e “Lognormal” estiverem marcadas, o *Solver* é automaticamente executado duas vezes, uma para o ajuste de cada distribuição, e aquela que produzir o menor erro quadrático será adotada na suavização.

### 3.3.5 Verificação dos resultados

O resultado da estimativa é verificado pela comparação entre o número de falhas mensais reais e estimadas, utilizando o teste *Wilcoxon Signed Rank Test*. Para o cálculo do número de falhas a partir das estimativas “ $q_t$ ” e “ $q_{t\text{suave}}$ ” é utilizada a técnica descrita no item

3.2.2. O resultado deste cálculo é apresentado nas colunas “Falhas<sub>KM</sub>”, “Falhas<sub>Suave</sub>” e no gráfico à esquerda na Figura 4.

O teste de hipótese *Wilcoxon Signed Rank Test*, descrito no item 2.8, é aplicado aos pares de dados das colunas Falhas|Falhas<sub>KM</sub> e Falhas|Falhas<sub>Suave</sub>, para verificar a validade da hipótese de igualdade entre as falhas simuladas e reais. As hipóteses do teste são:

$H_0$ : A contagem de falhas simulada é igual à real.

$H_1$ : A contagem de falha simulada é diferente da real.

O nível de significância adotado no teste é  $\alpha = 5\%$ . Se a hipótese  $H_0$  for aceita, as probabilidades de falha calculadas correspondem às probabilidades reais, e o modelo correspondente é capaz de reproduzir as falhas observadas.

No exemplo à Figura 4, a hipótese  $H_0$  foi aceita para os dois modelos com  $ValorP = 100\%$ , indicando que tanto a estimativa não paramétrica quanto a paramétrica são adequadas aos dados analisados.

### 3.3.6 Validação da implementação

Para fins de validação da implementação em planilha eletrônica foram simulados dados completos de falha a partir de distribuições conhecidas. Aplicada a metodologia, espera-se reproduzir os parâmetros da distribuição de origem nos parâmetros da distribuição de suavização das probabilidades de falha. Na validação inicial da implementação é importante criar um conjunto de dados de entrada que permita uma estimativa precisa. Os volumes de vendas adotados foram altos e constantes e os parâmetros de distribuição escolhidos resultam em um número elevado de falhas no período analisado. Nestas condições são produzidos dados em quantidade suficiente para uma estimativa quase exata das probabilidades de falha.

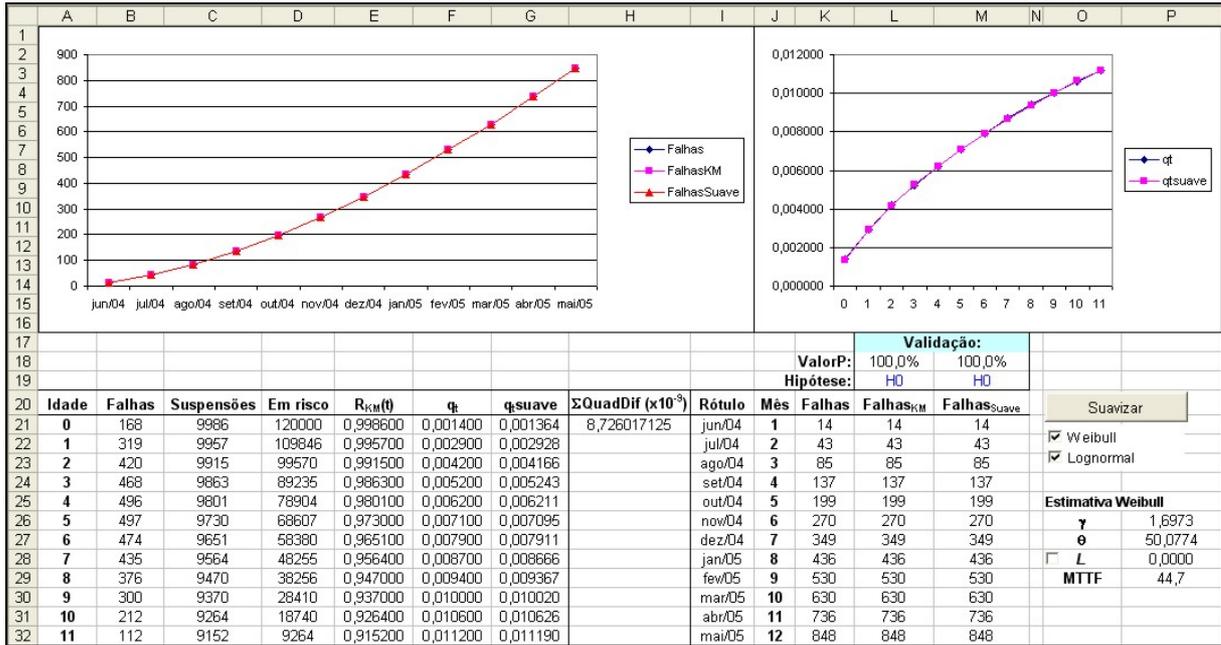
O primeiro teste utilizou uma distribuição Weibull para simulação dos dados de falha. A Figura 5 apresenta os parâmetros desta distribuição e os valores de vendas e falhas mensais gerados.

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O
1	<b>Weibull</b>		<b>Lognormal</b>		<b>Vendas</b>		Ruído na distribuição:		0%						
2	$\gamma$	1,7	$\sigma$	1	Mínimo:	10000	% Falhas no período:		8,5%						
3	$\theta$	50	$\mu$	3,4	Máximo:	10000									
4	L	0	L	0	Meses:	12									
5	MTTF	44,6	MTTF	49,4											
6	Aplicar		Aplicar		Aplicar										
7															
8															
9			Idade - t	0	1	2	3	4	5	6	7	8	9	10	11
10	Probabilidade de falha - $q_t$			0,001353	0,002913	0,004150	0,005228	0,006197	0,007084	0,007901	0,008659	0,009363	0,010019	0,010628	0,011194
11	Falhas na idade - $r_t$			168	319	420	468	496	497	474	435	376	300	212	112
12	Mês Falha			1	2	3	4	5	6	7	8	9	10	11	12
13	Falhas no mês			14	43	85	137	199	270	349	436	530	630	736	848
14	Mês Venda	Vendas	Falhas grupo												
15	1	10000	848	14	29	42	52	62	71	79	87	94	100	106	112
16	2	10000	736		14	29	42	52	62	71	79	87	94	100	106
17	3	10000	630			14	29	42	52	62	71	79	87	94	100
18	4	10000	530				14	29	42	52	62	71	79	87	94
19	5	10000	436					14	29	42	52	62	71	79	87
20	6	10000	349						14	29	42	52	62	71	79
21	7	10000	270							14	29	42	52	62	71
22	8	10000	199								14	29	42	52	62
23	9	10000	137									14	29	42	52
24	10	10000	85										14	29	42
25	11	10000	43											14	29
26	12	10000	14												14

Fonte: Elaborada pelo autor

Figura 5 – Dados para validação gerados a partir de distribuição Weibull

O resultado da análise destes dados é apresentado na Figura 6. Nos gráficos pode ser verificada a perfeita sobreposição entre as falhas reais e estimadas, bem como o excelente ajuste da distribuição paramétrica Weibull às estimativas não-paramétricas das probabilidades de falha. Os parâmetros estimados da distribuição Weibull são praticamente os mesmos da distribuição de origem. Os testes de hipótese utilizados para validação resultam no aceite da hipótese  $H_0$  com  $ValorP=100\%$  para as 2 estimativas. Estes resultados demonstram a capacidade da metodologia proposta em estimar corretamente as probabilidades de falha a partir de um conjunto de dados idealmente gerado a partir de uma distribuição de Weibull.



Fonte: Elaborada pelo autor

Figura 6 – Resultado da validação com dados de distribuição Weibull

O segundo teste utilizou uma distribuição Lognormal para simulação dos dados de falha. A Figura 7 apresenta os parâmetros desta distribuição e os valores de vendas e falhas mensais gerados.

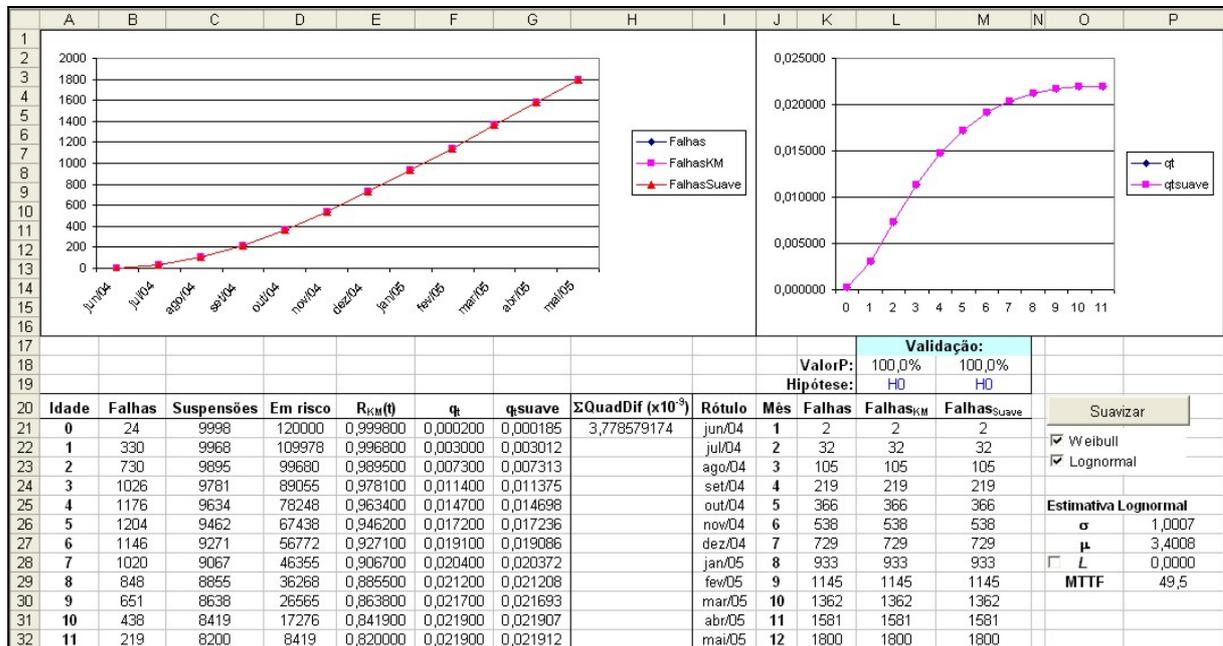
	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O
1	<b>Weibull</b>		<b>Lognormal</b>		<b>Vendas</b>		Ruído na distribuição:		0%						
2	$\gamma$	1,7	$\sigma$	1	Mínimo:	10000	% Falhas no período:		18,0%						
3	$\theta$	50	$\mu$	3,4	Máximo:	10000									
4	L	0	L	0	Meses:	12									
5	MTTF	44,6	MTTF	49,4											
6	Aplicar		Aplicar		Aplicar										
9	Idade - t		0	1	2	3	4	5	6	7	8	9	10	11	
10	Probabilidade de falha - $q_t$		0,000184	0,003003	0,007302	0,011367	0,014695	0,017238	0,019093	0,020382	0,021222	0,021710	0,021925	0,021931	
11	Falhas na idade - $r_z$		24	330	730	1026	1176	1204	1146	1020	848	651	438	219	
12	Mês Falha		1	2	3	4	5	6	7	8	9	10	11	12	
13	Falhas no mês		2	32	105	219	366	538	729	933	1145	1362	1581	1800	
14	Mês Venda	Vendas	Falhas grupo												
15	1	10000	1800	2	30	73	114	147	172	191	204	212	217	219	219
16	2	10000	1581		2	30	73	114	147	172	191	204	212	217	219
17	3	10000	1362			2	30	73	114	147	172	191	204	212	217
18	4	10000	1145				2	30	73	114	147	172	191	204	212
19	5	10000	933					2	30	73	114	147	172	191	204
20	6	10000	729						2	30	73	114	147	172	191
21	7	10000	538							2	30	73	114	147	172
22	8	10000	366								2	30	73	114	147
23	9	10000	219									2	30	73	114
24	10	10000	105										2	30	73
25	11	10000	32											2	30
26	12	10000	2												2

Fonte: Elaborada pelo autor

Figura 7 – Dados para validação gerados a partir de distribuição Lognormal

O resultado da análise destes dados é apresentado na Figura 8. Assim como no caso de dados originados da distribuição Weibull, os resultados apresentados demonstram que a

metodologia é capaz de estimar corretamente as probabilidades de falha de dados idealmente gerados a partir de uma distribuição Lognormal.



Fonte: Elaborada pelo autor

Figura 8 – Resultado da validação com dados de distribuição Lognormal

Nas duas simulações realizadas a decisão do modelo de distribuição que melhor se ajusta aos dados foi feita automaticamente baseada na minimização do erro quadrático. Nos dois casos a distribuição escolhida foi a mesma que gerou os dados de falha.

As validações realizadas até aqui servem apenas para provar a correta implementação da metodologia de análise. No capítulo 4 são realizadas validações adicionais para estimar a robustez da metodologia quando aplicada na análise de dados menos favoráveis.

### 3.4 IMPLEMENTAÇÃO DA ANÁLISE DE DADOS AGRUPADOS INCOMPLETOS

Dados incompletos de falha não incluem a idade em que os produtos foram vendidos; apenas o total de unidades que falharam a cada mês é conhecida. Esta informação, complementada pelo número de unidades vendidas a cada mês, permite estimativa das probabilidades de falha nas diferentes idades do produto utilizando a técnica descrita no item 2.6.4.

A principal vantagem desta análise é a simplicidade na coleta das informações. A partir de relatórios mensais de número de unidades vendidas e número de unidades recebidas para manutenção em garantia, podem ser estabelecidas estimativas para as probabilidades de falha

nas diferentes idades do produto. O ônus desta simplificação é uma maior imprecisão na estimativa, principalmente quando o número de falhas é pequeno. Também a complexidade matemática é maior que no caso da análise de dados completos, mas esta complexidade se torna pouco significativa com o uso de funções em VBA no *Excel*.

### **3.4.1 Condições para aplicabilidade**

A técnica de análise de dados incompletos é recomendada quando:

- É difícil determinar a data de venda dos produtos recebidos para manutenção, e não há outros usos para este dado que motive sua determinação e registro.
- O volume de vendas mensal é alto e o produto tem baixa probabilidade de falha.

Sempre que for viável determinar a idade de cada produto que falhou, é mais indicado aplicar a técnica de análise de dados completos descrita no item 3.3.

Na coleta dos dados para análise, devem ser observadas as seguintes restrições:

- Toda a falha ocorrida deve se originar dos produtos vendidos incluídos na análise. Esta ressalva é importante para o caso da técnica ser aplicada para um produto que já estava sendo vendido antes do primeiro mês analisado. Neste caso, as falhas de produtos oriundos destas vendas anteriores devem ser eliminadas da análise, mesmo que ainda estejam ocorrendo dentro do prazo de garantia.
- Todas as falhas ocorridas devem ser do conhecimento do fabricante e incluídas na análise. Esta condição é tipicamente atendida se o período de análise  $S$  é no máximo igual ao período de garantia do produto.

### **3.4.2 Entrada dos dados de falha**

As únicas informações necessárias para a análise são as contagens de vendas e falhas em cada mês do calendário. A entrada destes dados é feita em uma planilha preparada para receber dados para até 36 meses. Um botão permite a inclusão de dados de vendas e falhas a partir do simulador de dados completos descrito no item 3.2.1. Apenas as contagens de vendas e o total de falhas para cada mês são incluídos. A Figura 9 apresenta parte da planilha de entrada de dados.

	A	B	C	D	E	F
1						
2	<b>Rótulo</b>	<b>Mês</b>	<b>Vendas</b>	<b>Falhas</b>	Simular Falhas	
3	jun/04	1	1181	2		
4	jul/04	2	1035	4		
5	ago/04	3	725	9		
6	set/04	4	1557	14		
7	out/04	5	1894	23		
8	now/04	6	1295	32		

Fonte: Elaborada pelo autor

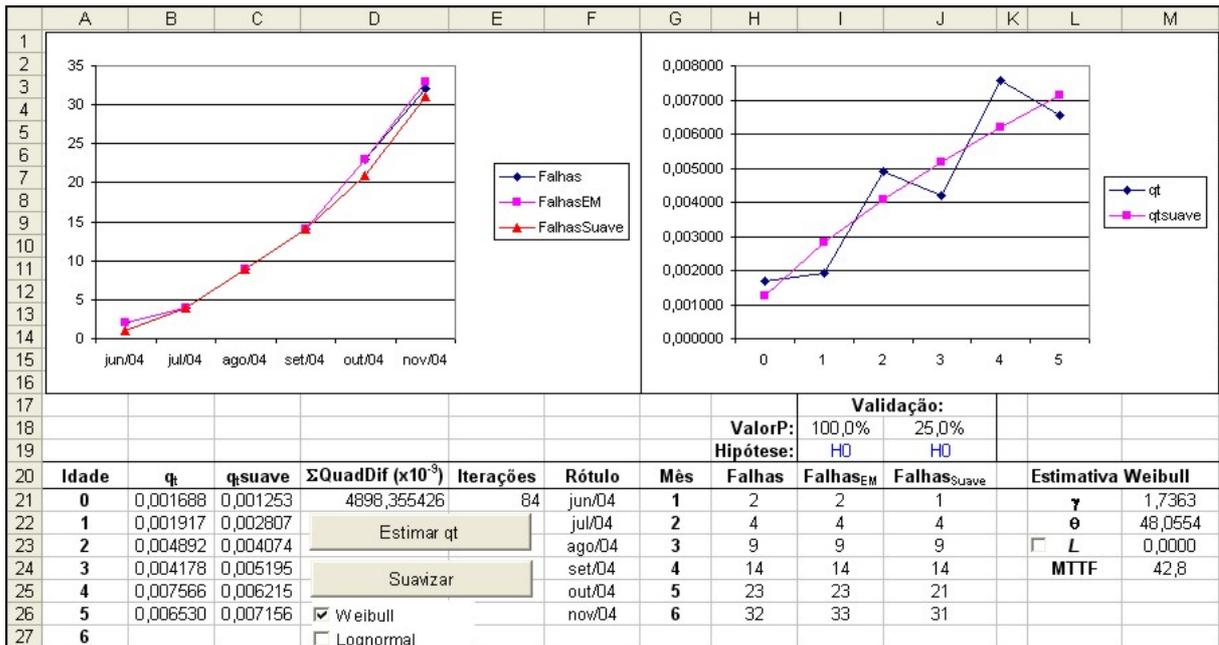
Figura 9 – Planilha de entrada de dados incompletos de falha

### 3.4.3 Estimativa não paramétrica das probabilidades de falha

Para estimativa das probabilidades de falha nas diferentes idades é utilizada a técnica de análise de dados marginais de vendas e falhas descrita no item 2.6.4. Por ser um processo iterativo, é necessária a implementação das equações (2.52) e (2.53) utilizando VBA. As iterações são repetidas até que o valor médio dos valores  $\hat{q}_t$  em uma iteração não varie significativamente quando comparado com a iteração anterior, conforme a seguir:

$$\bar{\hat{q}}_t^{k+1} - \bar{\hat{q}}_t^k < 0,000001 \quad (3.4)$$

Os dados de entrada da análise são as colunas “Vendas” e “Falhas” da planilha apresentada à Figura 9 e o resultado é o conjunto de valores  $\hat{q}_t$ , que são as estimativas para as probabilidades de falha na idade  $t$ . A Figura 10 apresenta uma visão geral da planilha que implementa a análise.



Fonte: Elaborada pelo autor

Figura 10 – Visão geral da planilha de análise de dados incompletos

A coluna “ $q_t$ ” apresenta as probabilidades de falha estimadas para cada idade do produto, calculadas após pressionar o botão “Estimar  $q_t$ ”. O número de iterações necessárias à convergência é apresentado na coluna “Iterações”. Os valores de  $\hat{q}_t$  já são suficientes para a realização de simulações de falhas para diferentes volumes de vendas ou prever as falhas subsequentes censuradas na análise. A área de desenvolvimento do produto pode comparar os valores  $\hat{q}_t$  de dois produtos semelhantes ou de um mesmo produto após implementadas melhorias, para avaliar as soluções de projeto implementadas.

### 3.4.4 Suavização da estimativa pelo ajuste de modelo paramétrico

As mesmas justificativas e técnicas já descritas no item 3.3.4 se aplicam neste caso. Na Figura 10, o botão “Suavizar” inicia a execução da função VBA que estima os parâmetros de uma distribuição Weibull ou Lognormal que melhor ajusta os valores da coluna “ $q_{t\text{suave}}$ ” aos valores da coluna “ $q_t$ ”. Os parâmetros estimados para a distribuição são apresentados na coluna “Estimativa Weibull” ou “Estimativa Lognormal”, conforme o caso.

No exemplo apresentado à Figura 10 foram analisados os mesmos dados da Figura 2 e já analisados na Figura 4. Os parâmetros da distribuição Weibull obtido nas duas análises são muito próximos aos parâmetros utilizados na geração dos dados.

### 3.4.5 Verificação dos resultados

A verificação dos resultados das estimativas não paramétricas e paramétricas é feita pela comparação entre o número de falhas reais e previstas pelas probabilidades estimadas. Esta verificação é apresentada nas colunas “Falhas”, “Falhas<sub>EM</sub>” e “Falhas<sub>Suave</sub>”, e no gráfico à esquerda da Figura 10.

O teste de hipótese *Wilcoxon Signed Rank Test*, descrito em 2.8, é aplicado aos pares de dados das colunas Falhas|Falhas<sub>EM</sub> e Falhas|Falhas<sub>Suave</sub>, para verificar a validade da hipótese de igualdade entre as falhas simuladas e reais. As hipóteses do teste são:

$H_0$ : A contagem de falhas simulada é igual à real.

$H_1$ : A contagem de falha simulada é diferente da real.

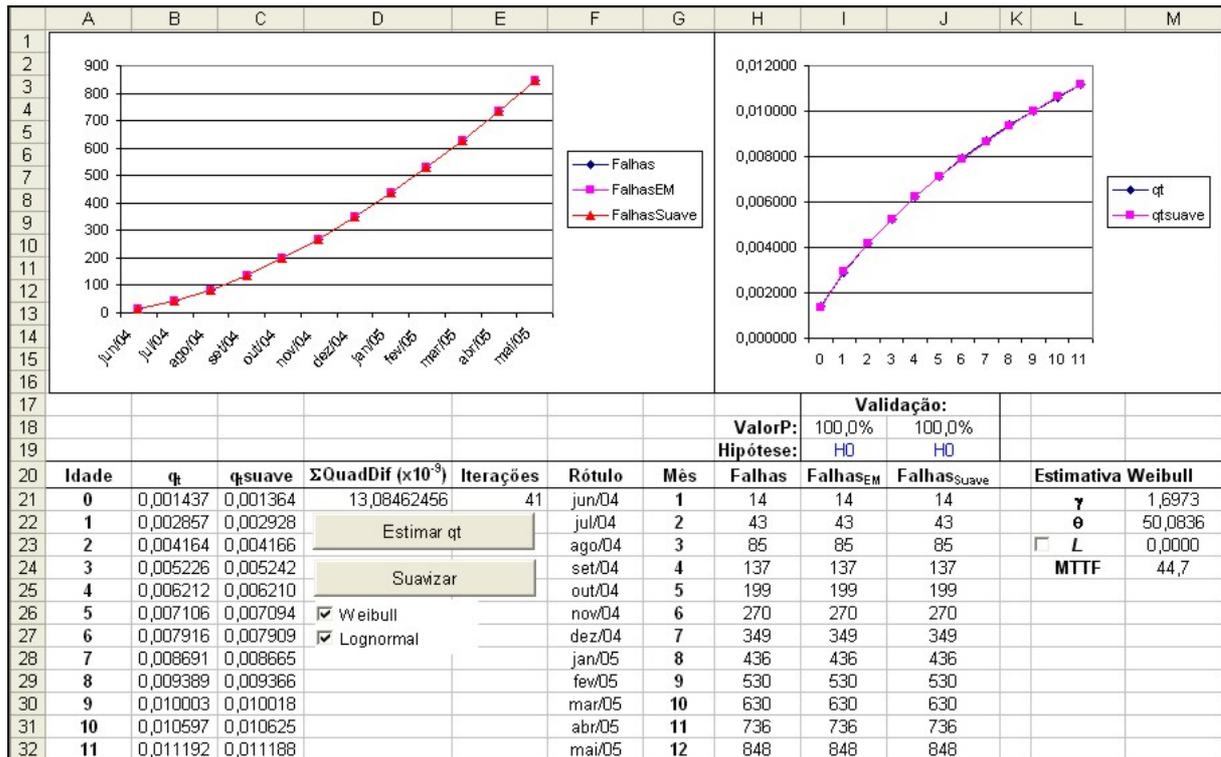
O nível de significância adotado no teste é  $\alpha = 5\%$ . Se a hipótese  $H_0$  for aceita, as probabilidades de falha calculadas correspondem às probabilidades reais, e o modelo correspondente é válido.

No exemplo à Figura 10, a hipótese  $H_0$  foi aceita para os dois modelos com  $ValorP = 100\%$  para a estimativa não paramétrica “ $q_t$ ” e  $ValorP = 25\%$  para a estimativa paramétrica Weibull “ $q_t$ suave”, indicando que ambas são adequadas aos dados analisados.

### 3.4.6 Validação da implementação

O objetivo da validação é verificar a correta implementação da técnica em VBA e nas planilhas *Excel*. Foram utilizados dois conjuntos de dados de falhas, um originado a partir de uma distribuição Weibull e outro de uma distribuição Lognormal.

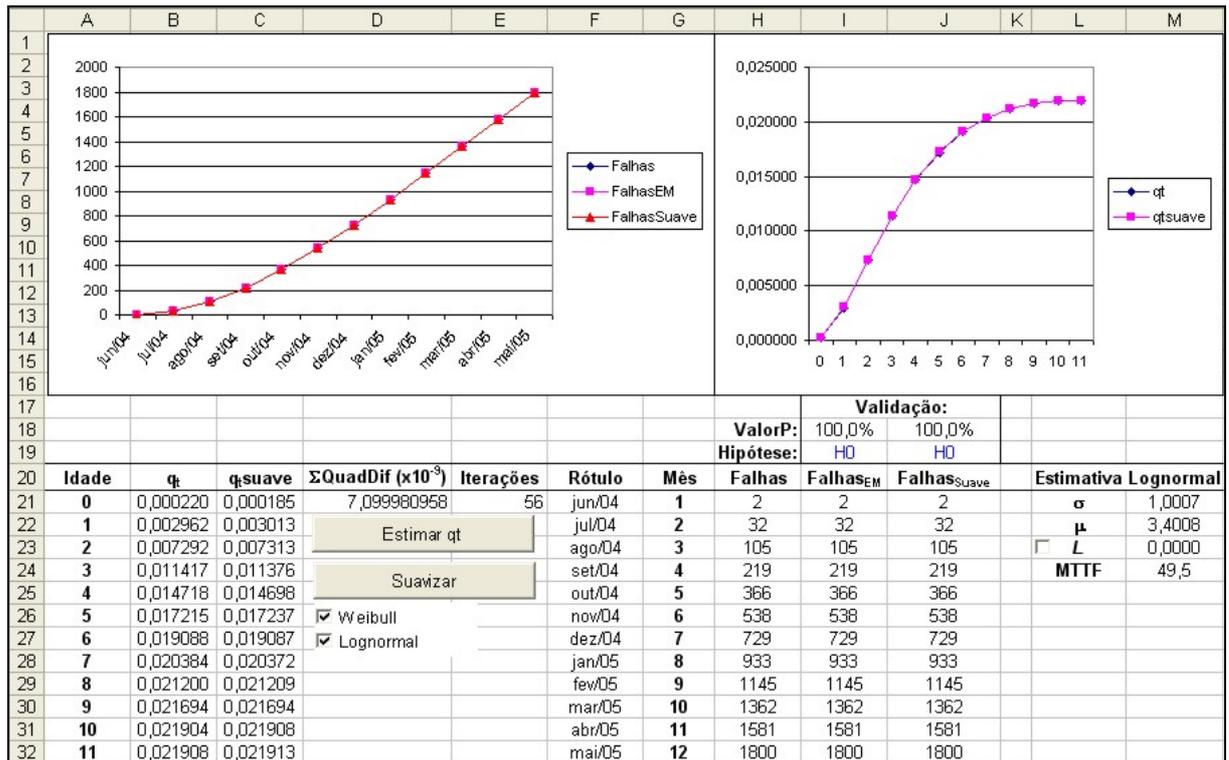
O primeiro teste utilizou os dados de vendas e falhas totais no mês da Figura 5, gerados a partir de uma distribuição Weibull. O resultado da análise é apresentado na Figura 11. Pode ser observada a perfeita reprodução do número de falhas a partir das estimativas paramétrica e não-paramétrica das probabilidades de falha. Os parâmetros obtidos para a distribuição Weibull são praticamente coincidentes com os da distribuição de origem.



Fonte: Elaborada pelo autor

Figura 11 – Resultado da validação com dados de distribuição Weibull

O segundo teste utilizou os dados de vendas e falhas totais no mês da Figura 7, gerados a partir de uma distribuição Lognormal. O resultado da análise é apresentado na Figura 12. Novamente pode ser observada a perfeita reprodução do número de falhas a partir das estimativas das probabilidades de falha. Os parâmetros obtidos para a distribuição Lognormal são praticamente coincidentes com os da distribuição de origem.



Fonte: Elaborada pelo autor

Figura 12 – Resultado da validação com dados de distribuição Lognormal

Estes resultados confirmam a correta implementação da metodologia de análise de dados incompletos.

Nos resultados apresentados fica evidente que, na presença de um número elevado de falhas originadas de distribuições bem definidas, as técnicas de análise de dados completos e incompletos produzem resultados praticamente idênticos, não ficando evidenciada nenhuma diferença de desempenho. As condições destes testes foram ideais e no capítulo 4 os testes são estendidos para verificar as condições que evidenciam o melhor ou pior desempenho de cada uma das técnicas.

## CAPÍTULO 4

### 4 ESTUDO DE CASOS

Neste capítulo são apresentados resultados de diferentes aplicações das duas metodologias de análise propostas. Os dados utilizados são simulados e os testes têm como principal objetivo avaliar o desempenho na análise de dados com diferentes características, estabelecendo diretrizes para a sua boa aplicação.

#### 4.1 METODOLOGIA DE AVALIAÇÃO

As metodologias de análise de dados agrupados completos e incompletos foram validadas em sua capacidade de estimar corretamente as taxas de falha a partir de dados idealmente gerados. Em aplicações reais, no entanto, os dados de falha apresentam características diversas, sendo o objetivo desta avaliação estabelecer a metodologia mais adequada para cada condição. Os fatores a serem investigados são:

- Baixo número de falhas mensais
- Vendas variáveis mês-a-mês
- Imperfeições na distribuição que simula os dados de falha

O efeito de cada um destes fatores é inicialmente analisado isoladamente, mantendo os demais na condição ideal, e a seguir a combinação destes fatores é analisada, simulando dados mais próximos aos esperados em aplicações reais. Os dados de vendas e falhas são analisados em um período de doze meses.

O resultado das análises é avaliado por teste de hipótese que compara as contagens de falhas ocorridas com as calculadas a partir das estimativas da confiabilidade. Os parâmetros estimados para a distribuição de tempos até a falha são também comparados com os da distribuição de origem.

#### 4.2 BAIXO NÚMERO DE FALHAS

O objetivo deste teste é verificar o efeito isolado da redução do número de falhas no resultado da análise. Quando o número de falhas é baixo, é esperado um pior desempenho na

estimativa, pela perda de sensibilidade decorrente do fato de o número de falhas ser sempre inteiro.

#### 4.2.1 Condições estabelecidas para o teste

Na realização deste teste os dados de falha são gerados sempre a partir de uma mesma distribuição de tempos até a falha. Mudanças no número de falhas são simuladas pela mudança do número de unidades vendidas.

Para geração das falhas foi utilizada uma distribuição Weibull com parâmetros  $\gamma = 1,7$ ,  $\theta = 50$  e  $L = 0$ , que resulta em  $MTTF = 44,6$  meses e uma probabilidade de falha de 8,5% no período analisado. As vendas mensais simuladas foram de 10000 e 200 unidades, sendo o mesmo valor repetido para todos os meses.

#### 4.2.2 Resultados das análises

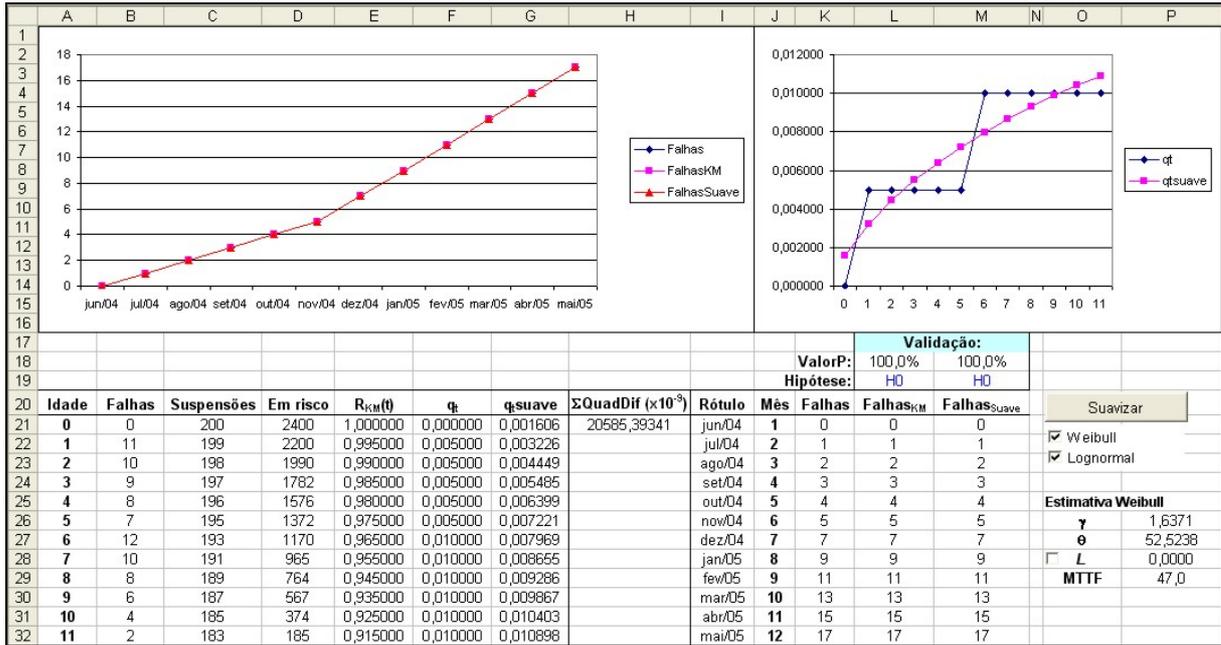
Os resultados para vendas de 10000 unidades/mês já foram apresentados nas figuras 6 e 11. Os dados simulados de falhas para vendas de 200 unidades/mês são apresentados na Figura 13.

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O
1	<b>Weibull</b>		<b>Lognormal</b>		<b>Vendas</b>		Ruído na distribuição:		0%						
2	$\gamma$	1,7	$\sigma$	1	Inicial:	200	% Falhas no período:		8,5%						
3	$\theta$	50	$\mu$	3,4	Final:	200									
4	L	0	L	0	Meses:	12									
5	MTTF	44,6	MTTF	49,4											
6	Aplicar		Aplicar		Aplicar		<input checked="" type="radio"/> Aleatória <input type="radio"/> Rampa								
9	Idade - t			0	1	2	3	4	5	6	7	8	9	10	11
10	Probabilidade de falha - $q_z$			0,001353	0,002913	0,004150	0,005228	0,006197	0,007084	0,007901	0,008659	0,009363	0,010019	0,010628	0,011194
11	Falhas na idade - $r_z$			0	11	10	9	8	7	12	10	8	6	4	2
12	Mês Falha			1	2	3	4	5	6	7	8	9	10	11	12
13	Falhas no mês			0	1	2	3	4	5	7	9	11	13	15	17
14	Mês Venda	Vendas	Falhas grupo												
15	1	200	17	0	1	1	1	1	1	2	2	2	2	2	2
16	2	200	15		0	1	1	1	1	1	2	2	2	2	2
17	3	200	13			0	1	1	1	1	1	2	2	2	2
18	4	200	11				0	1	1	1	1	1	2	2	2
19	5	200	9					0	1	1	1	1	1	2	2
20	6	200	7							0	1	1	1	1	2
21	7	200	5								0	1	1	1	1
22	8	200	4									0	1	1	1
23	9	200	3										0	1	1
24	10	200	2											0	1
25	11	200	1												0
26	12	200	0												

Fonte: Elaborada pelo autor

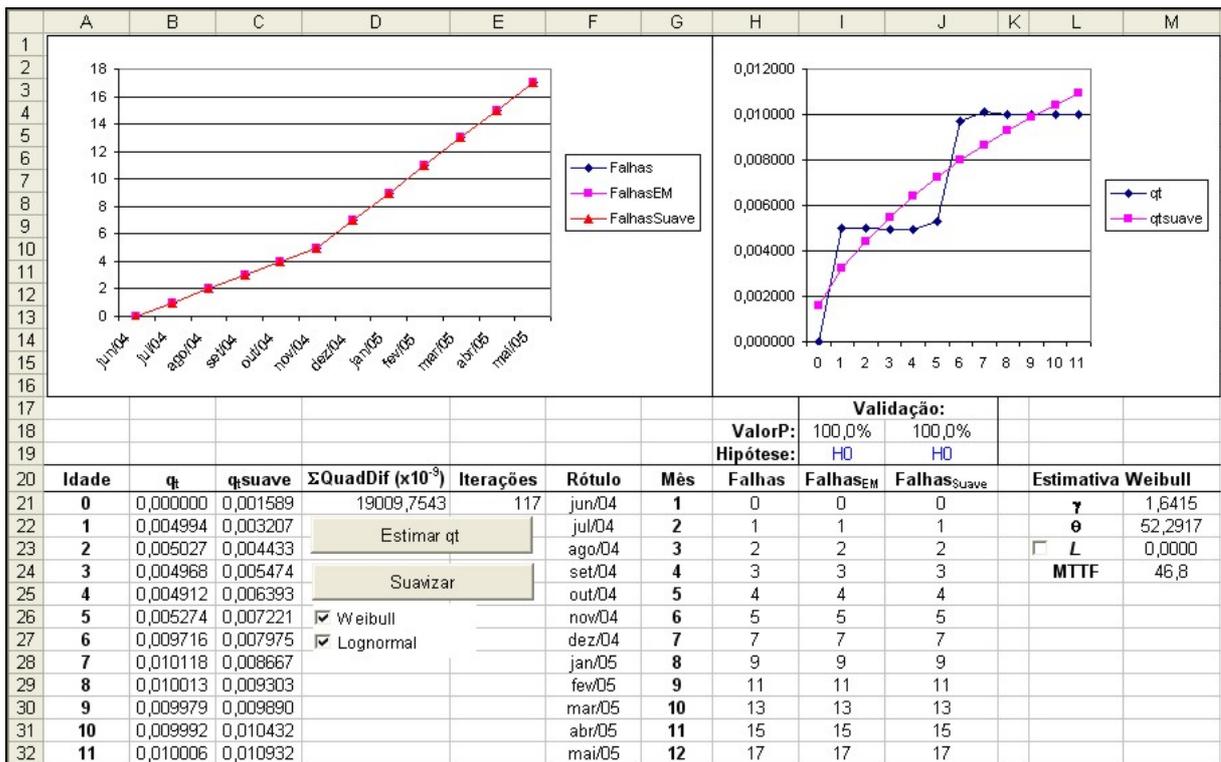
Figura 13 – Dados de avaliação com baixo número de falhas mensais

As figuras a seguir apresentam os resultados das duas metodologias de análise para os dados da Figura 13.



Fonte: Elaborada pelo autor

Figura 14 – Resultado da análise de dados completos com poucas falhas



Fonte: Elaborada pelo autor

Figura 15 – Resultado da análise de dados incompletos com poucas falhas

A Tabela 6 resume os resultados das duas técnicas para os dois conjuntos de dados simulados. Os testes de hipótese indicam que as estimativas paramétricas e não paramétricas obtidas a partir da análise de dados completos e incompletos são capazes de bem reproduzir as

falhas observadas, mesmo quando estas se apresentam em pequenas quantidades. Não há diferença significativa de desempenho entre a análise de dados completos e incompletos.

Mesmo na análise de dados com baixo número de falhas, as distribuições paramétricas ajustadas são muito próximas à distribuição de origem. A distribuição de Weibull foi selecionada como a de melhor ajuste, e o maior erro na estimativa de seus parâmetros foi de 5,1%, resultado considerado bom face ao reduzido número de falhas.

Tabela 6 – Avaliação do efeito da variação no número de falhas nas estimativas

	Poucas Falhas		Muitas Falhas	
	Dados Completos	Dados Incompletos	Dados Completos	Dados Incompletos
<b>Estimativa não paramétrica</b>	Aceita ValorP=100%	Aceita ValorP=100%	Aceita ValorP=100%	Aceita ValorP=100%
<b>Estimativa paramétrica</b>	Aceita ValorP=100%	Aceita ValorP=100%	Aceita ValorP=100%	Aceita ValorP=100%
<b><math>\gamma</math></b>	1,6371	1,6415	1,6973	1,6973
<b><math>\theta</math></b>	52,5238	52,2917	50,0774	50,0836
<b>MTTF</b>	47,0	46,8	44,7	44,7
<b>Erro % em <math>\gamma</math></b>	-3,7%	-3,4%	-0,2%	-0,2%
<b>Erro % em <math>\theta</math></b>	+5,1%	+4,6%	+0,2%	+0,2%
<b>Erro % em MTTF</b>	+5,4%	+4,9%	+0,2%	+0,2%

Fonte: Elaborada pelo autor

No gráfico à direita nas Figuras 14 e 15 pode ser observado que a estimativa não paramétrica da probabilidade de falha ( $q_t$ ) apresenta basicamente três valores distintos. Este resultado é decorrente da existência de apenas três valores distintos para a contagem de falhas em cada grupo de produtos vendidos em um mesmo mês, como pode ser observado à Figura 13. Não é de se esperar que a probabilidade de falha de um produto varie aos saltos em suas diferentes idades, sendo neste caso a estimativa paramétrica ( $q_{tsuave}$ ) mais próxima do comportamento real esperado.

Os resultados obtidos nas análises de dados com grande quantidade de falhas foram muito precisos. Os parâmetros da distribuição de Weibull foram estimados com erro máximo de 0,2%.

### 4.3 VENDAS VARIÁVEIS

Em muitos tipos de produtos é esperada variação significativa no número de produtos vendidos em cada mês. Algumas das causas para esta variação são:

- Estágio do produto em seu ciclo de vida.
- Efeitos sazonais, como férias, estações do ano.
- Venda eventual de grandes lotes.
- Falta do produto.

Quatro condições de vendas variáveis são analisadas: vendas aleatoriamente variáveis, vendas crescentes, vendas decrescentes e meses sem vendas.

#### 4.3.1 Condições estabelecidas para o teste

Na realização deste teste os dados de falha são gerados a partir de uma distribuição de tempos até a falha tipo Weibull com parâmetros  $\gamma = 1,7$ ,  $\theta = 50$  e  $L = 0$ , que resulta em  $MTTF = 44,6$  meses e uma probabilidade de falha de 8,5% no período analisado.

#### 4.3.2 Vendas aleatoriamente distribuídas

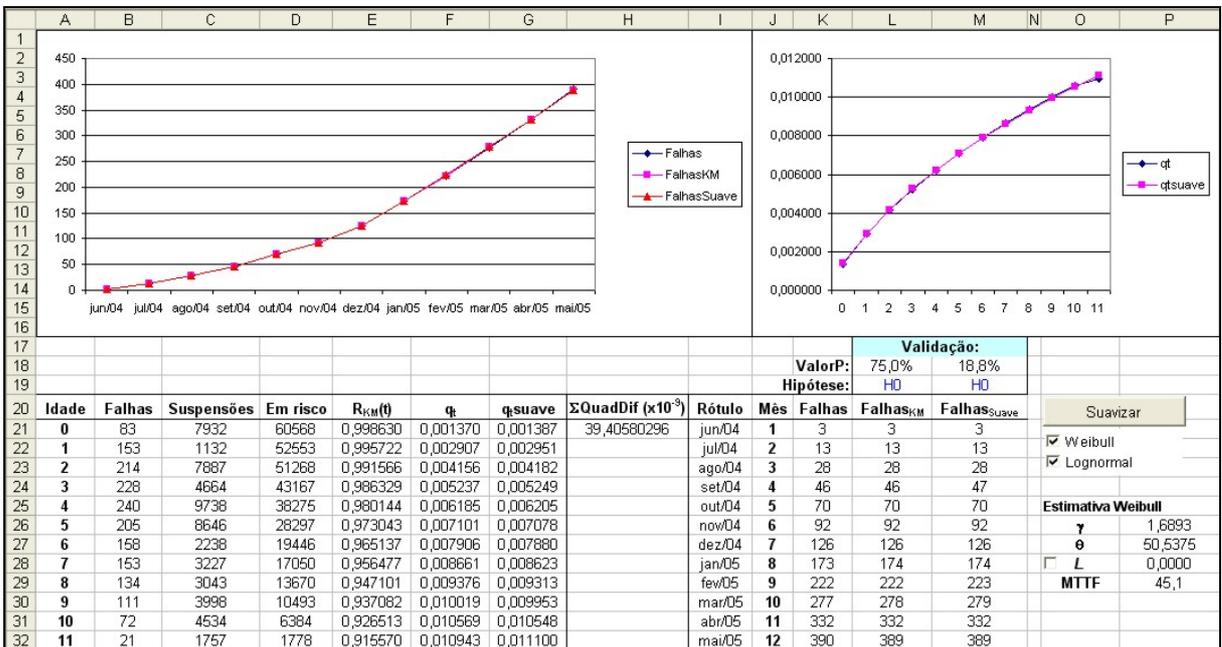
Os valores de vendas em cada mês foram gerados aleatoriamente entre os valores de 1000 e 10000 unidades mensais, conforme a Figura 16. Repetidas análises de nove outros conjuntos de dados como este foram realizadas para verificar se diferentes combinações aleatórias de vendas produzem resultados semelhantes. Os resultados obtidos foram sempre semelhantes, exceto quando os valores de vendas nos meses iniciais eram repetidamente próximos ao limite inferior estabelecido. Nesta condição, as estimativas baseadas na análise de dados incompletos apresentavam pior desempenho. Esta ocorrência motivou a análise deste efeito em separado pela simulação de vendas crescentes e decrescentes.

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O
1	<b>Weibull</b>		<b>Lognormal</b>		<b>Vendas</b>		Ruído na distribuição: 0%								
2	$\gamma$	1,7	$\sigma$	1	Inicial:	1000	% Falhas no período: 8,5%								
3	$\theta$	50	$\mu$	3,4	Final:	10000									
4	L	0	L	0	Meses:	12									
5	MTTF	44,6	MTTF	49,4											
6	Aplicar		Aplicar		Aplicar		<input checked="" type="radio"/> Aleatória								
7												<input type="radio"/> Rampa			
8															
9	Idade - t														
10	Probabilidade de falha - $q_t$														
11	Falhas na idade - $r_t$														
12	Mês Falha														
13	Falhas no mês														
14	Mês Venda	Vendas	Falhas grupo												
15	1	1920	163	3	6	8	10	12	14	15	17	18	19	20	21
16	2	4894	360		7	14	20	26	30	35	39	42	46	49	52
17	3	4266	268			6	12	18	22	26	30	34	37	40	43
18	4	3212	169				4	9	13	17	20	23	25	28	30
19	5	3375	148					5	10	14	18	21	24	27	29
20	6	2318	80						3	7	10	12	14	16	18
21	7	8885	239							12	26	37	46	55	63
22	8	9935	197								13	29	41	52	62
23	9	4729	65									6	14	20	25
24	10	7954	67										11	23	33
25	11	1137	5											2	3
26	12	7943	11												11

Fonte: Elaborada pelo autor

Figura 16 – Dados de avaliação com vendas variáveis de forma aleatória

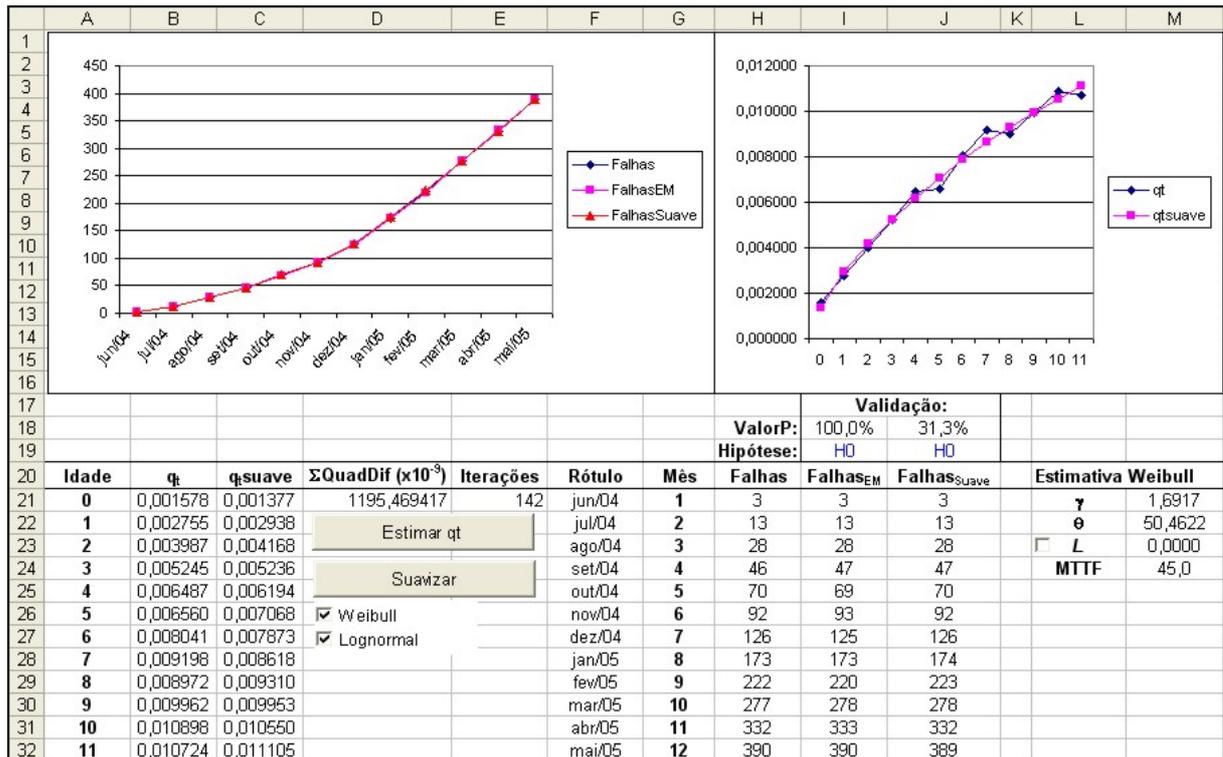
Os resultados da análise de dados completos apresentada à Figura 17 mostram que tanto a estimativa não-paramétrica quanto a paramétrica foram excelentes.



Fonte: Elaborada pelo autor

Figura 17 – Resultado da análise de dados completos com vendas variáveis

Os resultados da análise de dados incompletos apresentados à Figura 18 indicam também ótimos resultados, mas pode ser observado no gráfico que a estimativa não-paramétrica ( $q_t$ ) apresenta pequenas variações não observadas na análise de dados completos.



Fonte: Elaborada pelo autor

Figura 18 – Resultado da análise de dados incompletos com vendas variáveis

A Tabela 7 resume os resultados das estimativas pelas duas técnicas implementadas. Os testes de hipótese indicam que as estimativas paramétricas e não paramétricas obtidas a partir da análise de dados completos e incompletos são capazes de bem reproduzir as falhas observadas. Não há diferença significativa de desempenho entre a análise de dados completos e incompletos, e os parâmetros estimados para a distribuição de Weibull foram quase exatos.

Tabela 7 – Avaliação do efeito da variação aleatória das vendas nas estimativas

	Dados Completos	Dados Incompletos
<b>Estimativa não paramétrica</b>	Aceita, ValorP=75%	Aceita, ValorP=100%
<b>Estimativa paramétrica</b>	Aceita, ValorP=18,8%	Aceita, ValorP=31,3%
$\gamma$	1,6893	1,6917
$\theta$	50,5375	50,4622
<b>MTTF</b>	45,1	45,0
<b>Erro % em <math>\gamma</math></b>	-0,6%	-0,5%
<b>Erro % em <math>\theta</math></b>	+1,1%	+0,9%
<b>Erro % em MTTF</b>	+1,1%	+0,9%

Fonte: Elaborada pelo autor

### 4.3.3 Vendas crescentes

Na estimativa de confiabilidade de um novo produto é de se esperar que as vendas mensais sejam crescentes, e o desempenho das metodologias de análise precisa ser avaliado

nesta condição. Para esta análise os valores de vendas foram distribuídos em crescimento linear ao longo dos 12 meses analisados.

O valor inicial e final de vendas foi variado e constatou-se que quando o volume de vendas inicial é suficientemente alto para permitir que ocorram falhas entre os primeiros produtos vendidos, não há piora nas estimativas de confiabilidade. Se a ocorrência de falhas entre os primeiros produtos vendidos é nula ou pequena, as estimativas de confiabilidade são negativamente afetadas. Isto ocorre porque são estas primeiras unidades vendidas que formam a população em risco de falhar nas idades finais da análise. Se não forem detectadas falhas nestas idades finais, as estimativas resultam imprecisas.

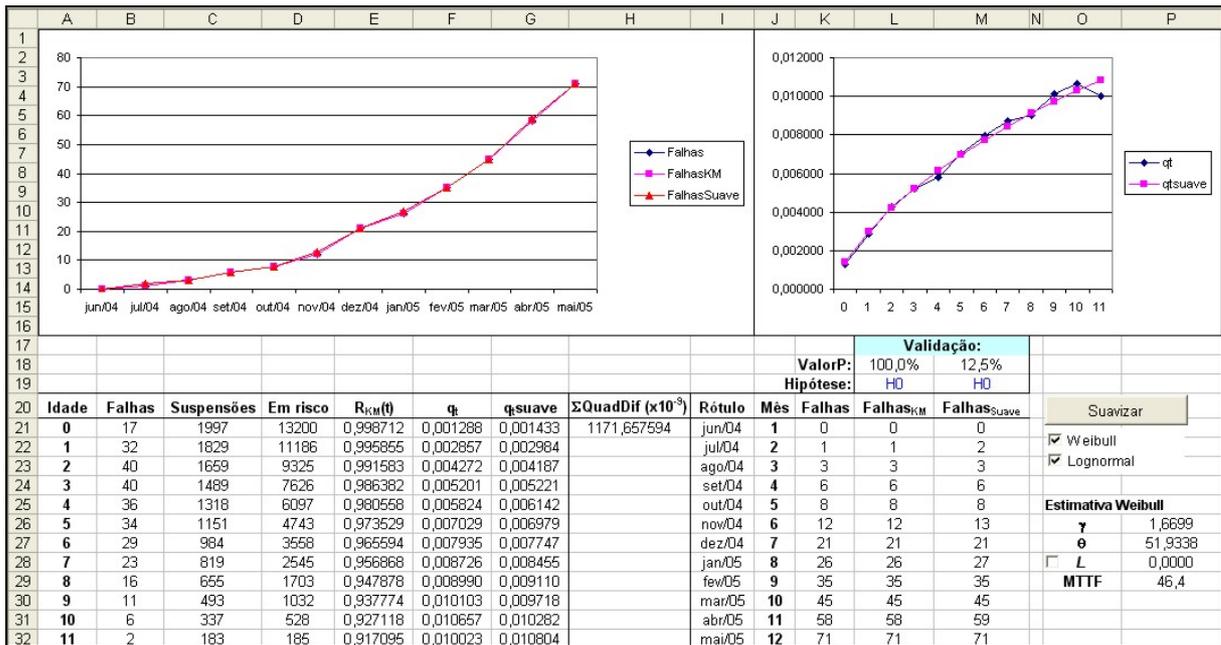
Para expor as deficiências na análise deste tipo de dado, as vendas foram variadas de forma linearmente crescente entre 200 e 2000 unidades ao longo dos doze meses da análise. A escolha destes valores permite que ocorram algumas falhas entre as primeiras 200 unidades vendidas e não é excessivamente baixo a ponto de comprometer a análise pelo baixo número de falhas. Os dados de falha simulados nesta condição são apresentados à Figura 19.

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O
1	<b>Weibull</b>		<b>Lognormal</b>		<b>Vendas</b>		Ruído na distribuição:		0%						
2	$\gamma$	1,7	$\sigma$	1	Inicial:	200	% Falhas no período:		8,5%						
3	$\theta$	50	$\mu$	3,4	Final:	2000									
4	L	0	L	0	Meses:	12									
5	MTTF	44,6	MTTF	49,4											
6	Aplicar		Aplicar		Aplicar		<input type="radio"/> Aleatória								
7							<input checked="" type="radio"/> Rampa								
8															
9			<b>Idade - t</b>	<b>0</b>	<b>1</b>	<b>2</b>	<b>3</b>	<b>4</b>	<b>5</b>	<b>6</b>	<b>7</b>	<b>8</b>	<b>9</b>	<b>10</b>	<b>11</b>
10	<b>Probabilidade de falha - <math>q_t</math></b>			0,001353	0,002913	0,004150	0,005228	0,006197	0,007084	0,007901	0,008659	0,009363	0,010019	0,010628	0,011194
11	<b>Falhas na idade - <math>r_t</math></b>			17	32	40	40	36	34	29	23	16	11	6	2
12	<b>Mês Falha</b>			<b>1</b>	<b>2</b>	<b>3</b>	<b>4</b>	<b>5</b>	<b>6</b>	<b>7</b>	<b>8</b>	<b>9</b>	<b>10</b>	<b>11</b>	<b>12</b>
13	<b>Falhas no mês</b>			0	1	3	6	8	12	21	26	35	45	58	71
14	<b>Mês Venda</b>	<b>Vendas</b>	<b>Falhas grupo</b>												
15	<b>1</b>	200	17	0	1	1	1	1	1	2	2	2	2	2	2
16	<b>2</b>	364	27		0	1	2	2	2	3	3	3	3	4	4
17	<b>3</b>	527	34			1	2	2	3	3	4	4	5	5	5
18	<b>4</b>	691	36				1	2	3	4	4	5	5	6	6
19	<b>5</b>	855	36					1	2	4	4	5	6	7	7
20	<b>6</b>	1018	34						1	3	4	5	6	7	8
21	<b>7</b>	1182	31							2	3	5	6	7	8
22	<b>8</b>	1345	27								2	4	6	7	8
23	<b>9</b>	1509	20									2	4	6	8
24	<b>10</b>	1673	14										2	5	7
25	<b>11</b>	1836	7											2	5
26	<b>12</b>	2000	3												3

Fonte: Elaborada pelo autor

Figura 19 – Dados de avaliação com vendas variáveis de forma crescente

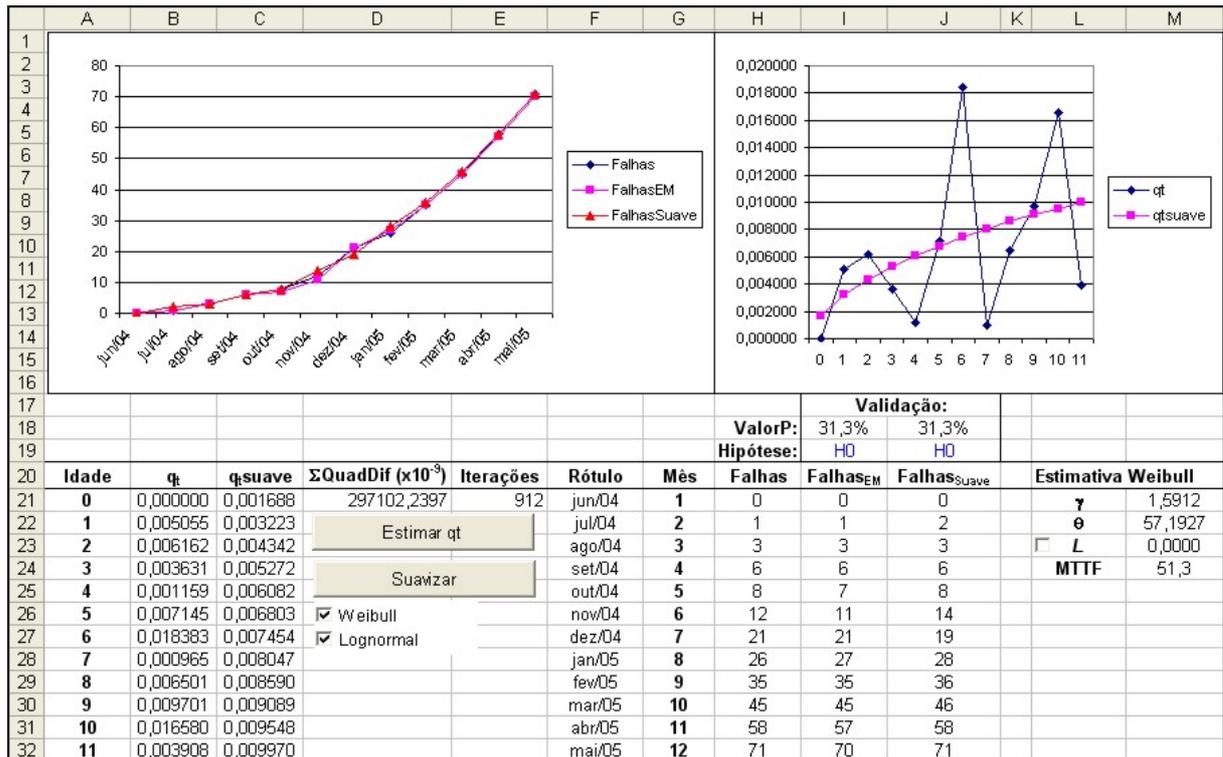
Na Figura 20 é apresentado o resultado da análise de dados completos para os dados com vendas crescentes. Os testes de hipóteses e a inspeção visual indicam o bom ajuste das estimativas.



Fonte: Elaborada pelo autor

Figura 20 – Resultado da análise de dados completos com vendas crescentes

A Figura 21 apresenta o resultado da estimativa baseada nos dados incompletos de falha. Os testes de hipótese indicam que tanto a estimativa paramétrica quanto a não-paramétrica são adequadas. A representação gráfica de  $q_t$  já indica visualmente que esta estimativa apresenta problemas, principalmente considerando-se que os dados de falha se originam de uma distribuição pura, em que oscilações nas probabilidades de falha não existem. A estimativa paramétrica resultou em uma distribuição Weibull próxima à distribuição de origem, mas novamente a análise visual dos gráficos de  $q_t$  e  $q_{tsuave}$  indica que este ajuste é frágil. Esta condição pode ser verificada matematicamente pela comparação entre os valores da soma quadrática dos erros ( $\Sigma QuadDif$ ) na Figura 20 e Figura 21, muito mais elevada na análise de dados incompletos.



Fonte: Elaborada pelo autor

Figura 21 – Resultado da análise de dados incompletos com vendas crescentes

A Tabela 8 resume os resultados das estimativas pelas duas técnicas implementadas. Os testes de hipótese indicam que as estimativas paramétricas e não paramétricas obtidas a partir da análise de dados completos e incompletos são capazes de bem reproduzir as falhas observadas. O desempenho da análise de dados completos foi bom, estando os parâmetros da distribuição de Weibull estimada próximos aos valores da distribuição de origem. A análise de dados incompletos apresenta desempenho pior, estando os parâmetros estimados para a distribuição de Weibull desviados em até 14% dos valores reais.

Tabela 8 – Avaliação do efeito de vendas crescentes nas estimativas

	Dados Completos	Dados Incompletos
<b>Estimativa não paramétrica</b>	Aceita, ValorP=100%	Aceita, ValorP=31,3%
<b>Estimativa paramétrica</b>	Aceita, ValorP=12,5%	Aceita, ValorP=31,3%
$\gamma$	1,6699	1,5912
$\theta$	51,9338	57,1927
<b>MTTF</b>	46,4	51,3
<b>Erro % em <math>\gamma</math></b>	-1,8%	-6,4%
<b>Erro % em <math>\theta</math></b>	+3,9%	+14,0%
<b>Erro % em MTTF</b>	+4,0%	+15,0%

Fonte: Elaborada pelo autor

#### **4.3.4 Vendas decrescentes**

O desempenho das metodologias de análise foi avaliado também na condição de vendas decrescentes, nas mesmas condições estabelecidas para a análise de vendas crescentes. Os resultados obtidos com vendas decrescentes são praticamente idênticos aos obtidos com vendas variáveis aleatoriamente, indicando que não ocorre a perda de desempenho observada no caso de vendas crescentes.

A ocorrência de vendas em maior quantidade nos meses iniciais da análise garante a manutenção de um número elevado para a população em risco de falha em todas as idades, resultando em um bom desempenho para a análise. As poucas unidades vendidas nos meses finais não podem contribuir com falhas nas idades finais pela censura no tempo, mas estas idades finais já têm dados suficientes gerados pelas falhas das primeiras unidades vendidas.

Os resultados detalhados da análise de vendas decrescentes não são aqui apresentados e sua análise foi realizada para melhor compreensão da causa de perda de desempenho na condição de vendas crescentes.

#### **4.3.5 Meses sem vendas**

O efeito da inexistência de unidades vendidas em um ou mais meses foi investigada, pois esta situação pode ocorrer em casos de falta do produto para venda, períodos de férias coletivas ou vendas fortemente sazonais.

Para este teste foram utilizados os dados da Figura 5, tendo sido anulados os valores de vendas para diferentes meses. Os testes indicam que a falta de vendas em qualquer mês que não seja o inicial não afeta significativamente o resultado da análise. Isto ocorre porque o grupo de unidades vendidas no primeiro mês está em risco de falha em todas as idades analisadas, sendo as falhas deste grupo informação suficiente para a estimativa da confiabilidade. Neste teste, a grande quantidade de produtos vendidos no mês inicial garante a ocorrência de um número de falhas suficiente para que a confiabilidade possa ser estimada apenas a partir das falhas deste grupo de produtos. Em situações reais não é de se esperar que os produtos vendidos em um único mês produzam falhas suficientes para uma boa estimativa da confiabilidade.

O método de análise de dados incompletos resulta em uma operação de divisão por zero se o número de produtos vendidos no primeiro mês é nulo, não sendo obtido nenhum

resultado. O método de análise de dados completos, na mesma condição, consegue estimar a probabilidade de falha para as idades anteriores à última (onze meses neste teste). Na análise de um novo produto, o primeiro mês da análise deve ser o primeiro mês com vendas não nulas.

Os resultados desta análise não são apresentados pois em pouco diferem dos resultados obtidos na análise dos dados da Figura 5.

#### 4.4 IMPERFEIÇÕES NA DISTRIBUIÇÃO DE ORIGEM

Nas avaliações realizadas até aqui os dados de falha foram gerados a partir de uma distribuição de Weibull. Na análise de produtos reais as falhas não apresentam ajuste perfeito a uma distribuição, e é provável que diferentes modos de falha sigam diferentes distribuições. Mesmo no caso em que as falhas obedecem a uma distribuição padronizada, não é garantido que a falha seja notificada ao fabricante no mês em que ela ocorreu. Isto não altera significativamente o total de falhas ocorridas no período de análise, mas desloca falhas de um mês para outro, sendo esta a condição simulada pela adição de uma parcela aleatória nas probabilidades de falha da distribuição.

As estimativas não paramétricas calculadas pelos dois métodos de análise implementados não são afetadas pela distribuição de origem dos dados, mas a suavização destas estimativas é feita por uma distribuição paramétrica que deve coincidir com a distribuição dos tempos até a falha. O objetivo desta avaliação é verificar se a distribuição de origem é corretamente identificada quando esta apresenta um ruído aleatório sobreposto aos seus valores de probabilidade de falha.

##### 4.4.1 Condições estabelecidas para o teste

Os dados de origem são gerados a partir de duas distribuições: Weibull com  $\gamma = 1,7$ ,  $\theta = 50$  e  $L = 0$  (MTTF de 44,6 meses) e Lognormal com  $\sigma = 1,1$ ,  $\mu = 4$  e  $L = 0$  (MTTF de 100 meses). As duas distribuições escolhidas apresentam probabilidade de falha no período analisado de 8,5%, para que o número de falhas total no período seja semelhante nos dois casos. Nas simulações de falhas com ruído aleatório, são adotadas apenas as simulações em que a probabilidade de falha resultante no período é mantida em  $8,5\% \pm 0,1$ , ou seja, nas quais o ruído acrescentado tem média próxima a zero.

O número de unidades vendidas em cada mês é mantido alto e constante, para analisar unicamente o efeito da imperfeição da distribuição nas estimativas de falha. Cinco conjuntos de dados de falha originados de cada distribuição, mas com diferentes conjuntos de parcelas aleatórias, são analisados pelas duas metodologias propostas. A suavização das estimativas não-paramétricas das probabilidades de falha é através de uma distribuição Weibull ou Lognormal, sendo adotada a que resulta em melhor ajuste pelo método dos mínimos quadrados.

#### 4.4.2 Distribuição Weibull com ruído de 50%

Os resultados de cinco análises realizadas com dados originados da distribuição Weibull acrescida de um ruído com amplitude de 50% são resumidos na Tabela 9. Em três das cinco análises realizadas, a distribuição de Weibull foi automaticamente selecionada como a de melhor ajuste às probabilidades de falha estimadas. Nestes casos, os parâmetros da distribuição foram estimados com um erro máximo de 15,5% na estimativa dos parâmetros e de 15,4% no MTTF.

Tabela 9 – Estimativas paramétricas de falhas com distribuição Weibull e ruído 50%

Conjunto	Dados	Melhor ajuste	$\gamma$ ou $\sigma$	$\theta$ ou $\mu$	MTTF	Hipótese	ValorP
1	Completos	Weibull	1,8582	44,1601	39,2	H0	100,0%
	Incompletos	Weibull	1,8606	44,0854	39,1	H0	5,5%
2	Completos	Weibull	1,5468	57,7719	52,0	H0	100,0%
	Incompletos	Weibull	1,5547	57,2370	51,5	H0	37,5%
3	Completos	Lognormal	1,1934	4,1248	126,1	H0	100,0%
	Incompletos	Lognormal	1,1938	4,1246	126,1	H0	75,0%
4	Completos	Lognormal	1,3359	4,3250	184,5	H0	100,0%
	Incompletos	Lognormal	1,3241	4,3081	178,5	H0	15,6%
5	Completos	Weibull	1,8382	45,1757	40,1	H0	100,0%
	Incompletos	Weibull	1,7994	46,4140	41,3	H0	100,0%

Fonte: Elaborada pelo autor

Nas análises dos conjuntos 3 e 4, a distribuição Lognormal foi a de melhor ajuste aos dados pelo critério dos mínimos quadrados. Nestes casos, ao se forçar o ajuste de uma distribuição de Weibull, o ajuste obtido foi semelhante aos outros três casos, indicando que os dados 3 e 4 não se ajustam mal ao modelo Weibull, mas são melhor representados por uma distribuição Lognormal. Todos os modelos foram aceitos pelos testes de hipóteses,

evidenciando que estes testes não são capazes de identificar o ajuste do modelo inesperado com o nível de ruído adotado.

#### 4.4.3 Distribuição Lognormal com ruído de 50%

Os resultados de cinco análises realizadas com dados originados da distribuição Lognormal acrescida de um ruído com amplitude de 50% são resumidos na Tabela 10. Em apenas duas das cinco análises o modelo Lognormal foi corretamente identificado. O erro máximo na estimativa de parâmetros nestes casos foi de 11,3%, e no MTTF de 33,9%.

Tabela 10 – Estimativas paramétricas de falhas com distribuição Lognormal e ruído 50%

Conjunto	Dados	Melhor ajuste	$\gamma$ ou $\sigma$	$\theta$ ou $\mu$	MTTF	Hipótese	ValorP
1	Completos	Lognormal	1,2246	4,1475	133,9	H0	100,0%
	Incompletos	Lognormal	1,2241	4,1470	133,8	H0	50,0%
2	Completos	Weibull	1,9413	41,2922	36,6	H0	100,0%
	Incompletos	Weibull	1,9393	41,3444	36,7	H0	50,0%
3	Completos	Lognormal	1,1671	4,0841	117,4	H0	100,0%
	Incompletos	Lognormal	1,1680	4,0849	117,6	H1	3,1%
4	Completos	Weibull	1,8505	43,8992	39,0	H0	100,0%
	Incompletos	Weibull	1,8510	43,8885	39,0	H0	37,5%
5	Completos	Weibull	2,0211	39,8277	35,3	H0	100,0%
	Incompletos	Weibull	2,0224	39,7971	35,3	H0	15,6%

Fonte: Elaborada pelo autor

Nas análises dos conjuntos dois, quatro e cinco, o modelo Weibull resultou em um melhor ajuste às probabilidades de falha estimadas. Ao se forçar o ajuste Lognormal nestes casos, as estimativas obtidas foram semelhantes às obtidas para os conjuntos um e três.

O teste de hipóteses rejeitou apenas o modelo paramétrico estimado pela análise de dados incompletos para os dados do conjunto três. Todas as demais estimativas paramétricas e não-paramétricas foram aceitas pelo teste. Fica evidenciado que o teste de hipóteses aplicado não é capaz de identificar o ajuste do modelo inesperado.

#### 4.5 FATORES COMBINADOS

Nesta avaliação alguns dos fatores analisados individualmente são combinados para simular um conjunto de dados mais próximo aos esperados em aplicações reais: baixo número de falhas, vendas variáveis aleatoriamente e ruídos na distribuição de origem dos dados.

#### 4.5.1 Condições estabelecidas para o teste

Os dados para esta avaliação são gerados a partir de uma distribuição Weibull com parâmetros  $\gamma = 1,9$ ,  $\theta = 93$  e  $L = 0$  (MTTF de 82,5 meses), que resulta em um total de falhas no período de 2%. Com esta redução no percentual de falhas e pela escolha de volumes de vendas entre 500 e 2000 unidades/mês, o número de falhas é mantido baixo. A esta distribuição é acrescentado um ruído aleatório de amplitude 50%, com média zero.

Para cada análise um novo conjunto aleatório de vendas e de ruído na distribuição é gerado. Cinco dados de falha distintos são analisados pelas duas metodologias propostas.

#### 4.5.2 Resultados das análises

A Tabela 11 resume o resultado das análises realizadas nos cinco dados de falha gerados.

Tabela 11 – Estimativas paramétricas de falhas Weibull, ruído 50% e vendas aleatórias

Conjunto	Dados	$\gamma$	$\theta$	MTTF	Hipótese	ValorP
1	Completo	2,0107	84,9537	75,3	H0	12,9%
	Incompleto	2,0406	82,9131	73,5	H0	12,9%
2	Completo	1,7194	116,6324	104,0	H0	62,5%
	Incompleto	1,9336	87,3477	77,5	H0	82,4%
3	Completo	2,1193	77,4732	68,6	H1	1,4%
	Incompleto	2,1539	73,9942	65,5	H0	10,9%
4	Completo	1,9438	91,9174	81,5	H0	100,0%
	Incompleto	1,9786	87,9894	78,0	H0	100,0%
5	Completo	1,8501	99,3482	88,2	H0	68,8%
	Incompleto	1,7616	110,7832	98,6	H0	81,3%

Fonte: Elaborada pelo autor

Para os conjuntos dois, três e cinco, a distribuição Lognormal foi apontada como de melhor ajuste. Para estes casos a suavização foi novamente executada, limitada ao ajuste de uma distribuição Weibull. Estes três conjuntos foram os que apresentaram o maior erro nos parâmetros estimados para a distribuição Weibull, evidenciando que o ruído adicionado à distribuição de origem a distorceu excessivamente.

O ajuste obtido para o conjunto de dados três foi o mais desviado dos parâmetros da distribuição de origem, com o parâmetro de forma  $\gamma$  desviado em 13,4% e o parâmetro de

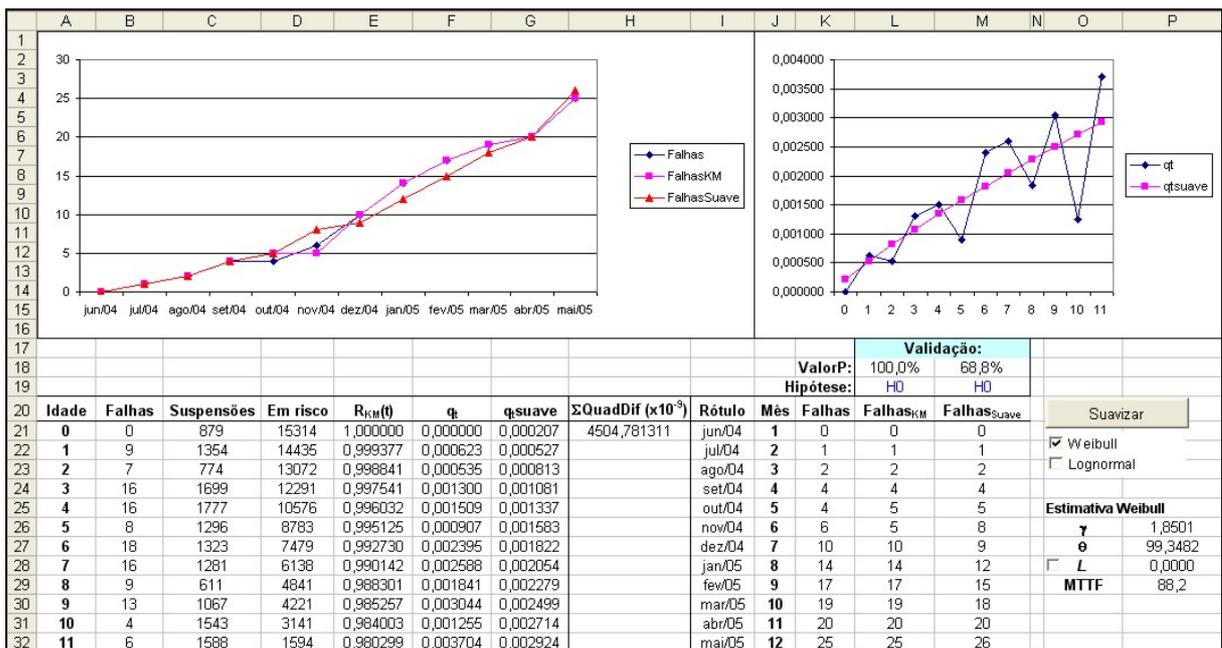
escala  $\theta$  em -20,4%. O ajuste obtido para o conjunto quatro foi o melhor, com desvio de 4,1% para  $\gamma$  e de -5,4% para  $\theta$ . A análise de dados completos aplicada ao conjunto três foi a única em que o modelo paramétrico foi rejeitado. Para todos os conjuntos analisados o modelo não-paramétrico foi aceito pelo teste de hipóteses.

As planilhas com os dados e resultados da análise número cinco são apresentadas nas Figuras 22, 23e 24, para melhor ilustrar a análise realizada.

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O
1	<b>Weibull</b>		<b>Lognormal</b>		<b>Vendas</b>		Ruído na distribuição:		50%						
2	$\gamma$	1,9	$\sigma$	1,4	Inicial:	500	% Falhas no período:		2,0%						
3	$\theta$	93	$\mu$	5,35	Final:	2000									
4	L	0	L	0	Meses:	12									
5	MTTF	82,5	MTTF	561,2											
6	Aplicar		Aplicar		Aplicar		<input checked="" type="radio"/> Aleatória <input type="radio"/> Rampa								
9			Idade - t	0	1	2	3	4	5	6	7	8	9	10	11
10	Probabilidade de falha - $q_t$			0,000191	0,000485	0,000412	0,001230	0,001494	0,000931	0,002410	0,002641	0,001752	0,002980	0,001414	0,003655
11	Falhas na idade - $r_t$			0	9	7	16	16	8	18	16	9	13	4	6
12	Mês Falha			1	2	3	4	5	6	7	8	9	10	11	12
13	Falhas no mês			0	1	2	4	4	6	10	14	17	19	20	25
14	Mês Venda	Vendas	Falhas grupo												
15	1	1620	32	0	1	1	2	2	2	4	4	3	5	2	6
16	2	1568	25		0	1	1	2	2	1	4	4	3	5	2
17	3	1083	16			0	1	0	1	2	1	3	3	2	3
18	4	618	7				0	0	0	1	1	1	1	2	1
19	5	1294	13					0	1	1	2	2	1	3	3
20	6	1333	10						0	1	1	2	2	1	3
21	7	1303	7							0	1	1	2	2	1
22	8	1784	7								0	1	1	2	3
23	9	1703	4									0	1	1	2
24	10	774	0										0	0	0
25	11	1355	1											0	1
26	12	879	0												0

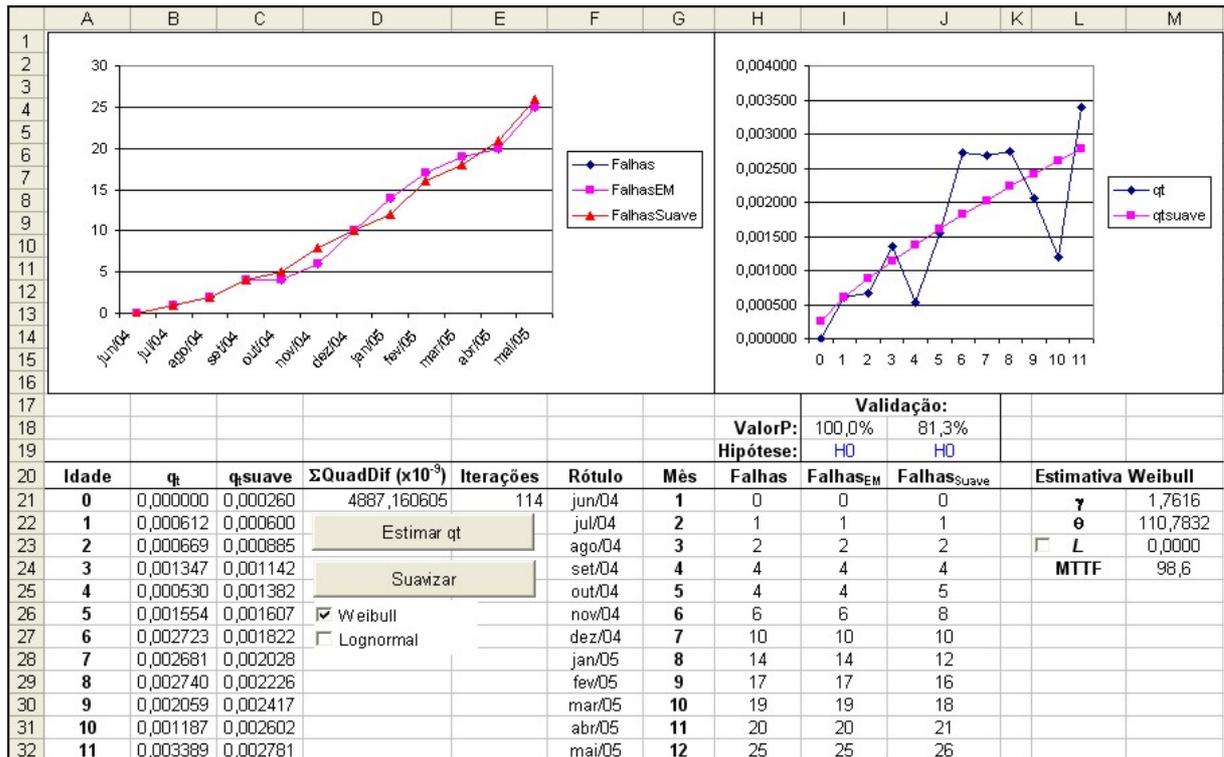
Fonte: Elaborada pelo autor

Figura 22 – Dados de falha Weibull com ruído 50% e vendas aleatórias



Fonte: Elaborada pelo autor

Figura 23 – Análise de dados completos Weibull com ruído 50% e vendas aleatórias



Fonte: Elaborada pelo autor

Figura 24 – Análise de dados incompletos Weibull com ruído 50% e vendas aleatórias

#### 4.6 ANÁLISE DE RESULTADOS E RECOMENDAÇÕES PARA APLICAÇÃO

A partir das análises realizadas com dados de falha de diferentes características, podem ser estabelecidas as limitações das metodologias de análise propostas neste trabalho e estabelecidas recomendações para sua melhor utilização.

##### 4.6.1 Estimativas não-paramétricas das probabilidades de falha

Em todos os testes realizados, as estimativas não paramétricas (valores de  $q_t$ ) obtidas foram aprovadas no teste de hipótese que verifica a igualdade entre as falhas realizadas e previstas a partir destas estimativas. A diferença de desempenho entre as análises de dados completos e incompletos só ficou evidenciada no caso de falhas crescentes (Figura 20 e Figura 21), onde o método baseado nas informações completas de falha resultou em uma melhor estimativa não paramétrica. Em todos os outros casos os desempenhos foram próximos, e a simplicidade de coleta dos dados incompletos estimula sua adoção nestes casos.

##### 4.6.2 Estimativas paramétricas das probabilidades de falha

A estimativa paramétrica foi adotada para suavização da estimativa não-paramétrica, utilizando distribuições de probabilidade padronizadas do tipo Weibull e Lognormal. As

planilhas implementadas buscam o ajuste mais adequado utilizando o método dos mínimos quadrados, e sugere a distribuição que melhor suaviza as probabilidades de falha nas diferentes idades. A opção automática pela melhor distribuição só se mostrou eficaz quando os dados de falha são originados a partir de uma distribuição exata. Na presença de distorções originadas pelo acréscimo de uma parcela aleatória às probabilidades de falha da distribuição de origem a identificação correta do modelo não foi mais possível.

Quando o modelo paramétrico a ser ajustado é previamente estabelecido com o mesmo modelo que deu origem aos dados, os parâmetros estimados são próximos aos originais. O maior desvio encontrado entre parâmetros estimados e originais foi de 20,4%, no teste do conjunto de dados três à Tabela 11.

O uso do modelo paramétrico para extrapolação de falhas além das idades analisadas, bem como para o cálculo do MTTF, não é recomendado quando a análise de confiabilidade é restrita a um período pequeno em relação à vida útil do produto. Neste período inicial de utilização irão predominar as falhas causadas pela mortalidade infantil de componentes e por deficiências no processo produtivo, enquanto que falhas resultantes da degradação de componentes do produto não serão incluídas na análise.

#### **4.6.3 Recomendações para aplicação**

Diretrizes básicas para a opção entre a análise de dados completos e incompletos foram apresentadas nos itens 3.3.1 e 3.4.1. Os resultados obtidos ao longo deste capítulo indicam novas recomendações, tanto para a seleção quanto para a aplicação da análise de dados agrupados de falha.

Na análise de dados de vendas e falhas de um novo produto, a análise de dados completos é a mais indicada, por seu melhor desempenho quando os volumes de vendas são crescentes. O menor volume de vendas e o maior envolvimento do pessoal de pós-venda e engenharia do produto nesta fase, propiciam melhores condições para o rastreamento da data de venda e falha de cada produto, bem como a determinação precisa da falha detectada e sua causa.

Para análise de produtos já em comercialização, onde o rápido crescimento inicial de vendas já cessou, é recomendada a adoção da análise de dados incompletos. O desempenho

desta análise é semelhante à análise de dados completos, mas a coleta dos dados de falha é simplificada. O início da análise, no entanto, requer cuidado.

Ao se iniciar o registro de vendas e falhas, só podem ser consideradas na análise as falhas ocorridas nos produtos vendidos a partir do primeiro mês analisado. Falhas ocorridas em produtos vendidos em meses anteriores não devem ser contadas, pois os produtos que as originaram não estão incluídos na contagem de vendas. A implementação deste procedimento requer a determinação da data de venda dos produtos que falham, pelo menos durante o período equivalente ao prazo de garantia. Quando a análise alcança todo o período de garantia, todas as falhas estarão ocorrendo entre os produtos vendidos dentro deste período, tornando-se a partir de então desnecessária a determinação da data de venda dos produtos.

A restrição da análise ao prazo de garantia maximiza a probabilidade de todas as falhas ocorridas serem reportadas ao fabricante, premissa básica para aplicação das metodologias apresentadas. A adoção do prazo de garantia como prazo de análise, resulta ainda em informações precisas da probabilidade de falha dentro deste período e a possibilidade de previsão de falhas futuras em diferentes cenários de vendas, o que contribui para o melhor dimensionamento de pessoal, insumos e custos da assistência técnica.

A partir das informações de vendas e falhas de um único produto podem ser realizados estudos de sua confiabilidade como um todo ou de suas partes e componentes. Se o número de ocorrências de cada modo de falha detectado no produto é registrado, as metodologias de estimativa podem ser aplicadas independentemente e simultaneamente para cada modo de falha e ainda para o produto como um todo. Este estudo, especialmente quando aplicado a um novo produto, permite a obtenção de informações adicionais para a equipe de desenvolvimento, contribuindo para sua contínua melhoria.

## CAPÍTULO 5

### 5 CONSIDERAÇÕES FINAIS

#### 5.1 CONCLUSÕES

Esta dissertação apresenta e discute duas metodologias para a análise de confiabilidade de produtos a partir de informações de número de falhas e número de unidades vendidas em um período. As metodologias para análise de dados completos e incompletos foram propostas pelo autor da dissertação, desenvolvidas a partir dos trabalhos de Kalbfleish, Lawless & Robinson (1991); Karim, Yamamoto & Suzuki (2001) e George (2004).

A metodologia de análise de dados completos é aplicada quando é possível determinar tanto o mês da venda quanto o mês da falha de cada produto. A análise de dados incompletos é aplicada quando as únicas informações disponíveis são o número de unidades vendidas e o número de falhas em cada mês. Nos dois casos, o objetivo é a obtenção de estimativas não-paramétricas das probabilidades de ocorrência de falhas nas diferentes idades do produto analisado.

Duas condições importantes devem ser observadas na entrada de dados para estas análises: todas as falhas devem ser conhecidas e incluídas na análise, e todas as falhas devem se originar dos produtos vendidos incluídos na análise. A primeira condição é tipicamente atendida ao se adotar para análise o período de garantia. Para produtos duráveis e de custo elevado, é de se esperar que todas as falhas ocorridas neste período sejam reportadas ao fabricante. A segunda condição é atendida naturalmente quando a análise é aplicada a um novo produto desde o início das vendas. Para aplicação a um produto que já é comercializado, as falhas originadas das vendas anteriores ao início da análise não devem ser consideradas.

A opção entre as metodologias é definida pela possibilidade ou não de determinar a data da venda de cada produto que falha. Se esta não pode ser determinada, a análise de dados incompletos deve ser adotada. Na maioria das situações analisadas, o desempenho observado utilizando as duas metodologias foi equivalente, e a simplicidade de aquisição das informações necessárias à análise de dados incompletos torna esta a escolha mais adequada.

Na análise de confiabilidade de um novo produto, a condição de vendas crescentes e a importância de se acompanhar as ocorrências de falhas de forma mais detalhada, indicam a metodologia de análise de dados completos como a mais adequada.

As análises não-paramétricas apresentadas resultam em informações suficientes para os objetivos deste trabalho. A partir das probabilidades de falha estimadas é possível prever falhas ao longo do calendário futuro para diferentes cenários de vendas, calcular o percentual de falhas em garantia, dimensionar recursos para assistência técnica e realimentar a área de desenvolvimento do produto com informações referentes a sua confiabilidade.

A partir dos valores de probabilidades de falha estimadas, um modelo paramétrico pode ser ajustado com o objetivo de suavizar estas estimativas. Neste trabalho são estudados os casos de ajuste de modelos baseados na distribuição de Weibull e Lognormal. A estimativa paramétrica permite a determinação do MTTF e a extrapolação das probabilidades de falha no tempo, mas estes resultados são de precisão duvidosa, pois a análise de dados limitada ao período de garantia pode não expor modos de falha que se manifestem em período posterior ao analisado.

O desempenho das metodologias apresentadas é avaliado pela análise de dados que simulam algumas condições esperadas em aplicações reais: baixo número de falhas, vendas variáveis mês-a-mês, imperfeição na distribuição de origem das falhas e combinação destas duas últimas condições. Nesta validação fica evidenciado que o desempenho das duas metodologias é semelhante, exceto para o caso de vendas crescentes, onde a análise de dados completos apresenta resultado melhor. O ajuste de um modelo paramétrico só é adequado se o tipo de distribuição correto é definido pelo usuário, pois a sistemática implementada para escolha da distribuição falha quando os dados não seguem a distribuição exata.

Para os casos em que a análise conjunta das falhas do produto não produz resultado útil, as metodologias podem ser aplicadas individualmente para cada modo de falha detectado. Nesta condição, as probabilidades dos diferentes tipos de falhas podem ser avaliadas. Esta informação mais detalhada permite melhor acompanhamento da confiabilidade do produto por sua equipe de desenvolvimento.

## 5.2 TRABALHOS FUTUROS

A partir das limitações citadas no capítulo inicial, podem ser estabelecidas sugestões para novos trabalhos. Uma primeira sugestão é a incorporação de dados de falha posteriores ao período de garantia. Para isto, é necessário adotar ou estimar a probabilidade de uma falha ocorrida fora do período de garantia ser reportada ao fabricante. Pela inclusão desta probabilidade na análise, um maior volume de informações será incorporado ao estudo, e modos de falha não revelados no período de garantia podem ser detectados. Oh & Bai (2001) apresenta um método para incorporação de falhas posteriores ao período de garantia em estudos paramétricos baseados no modelo de Weibull.

Uma segunda sugestão é a realização de um estudo comparativo de desempenho das metodologias apresentadas utilizando dados de falha reais, para os quais existam estimativas confiáveis de confiabilidade obtidas por outros métodos.

Uma terceira sugestão, baseada no trabalho de Wu & Meeker (2002), é a estimativa de limites de controle para as probabilidades de falha calculadas. O objetivo é criar um mecanismo estatisticamente sensível de detecção do aumento de falhas, que permita ações de correção precoces, reduzindo os custos e desgastes associados à falhas recorrentes no produto.

A partir do trabalho de Werner (1996), a simulação de falhas ao longo do calendário pode ser aprimorada pela incorporação de estimativas para o tempo de uso diário e do intervalo de tempo entre manufatura e venda do produto.

A última sugestão é a adoção de conceitos da estatística Bayesiana para incorporar ao estudo informações subjetivas originadas da área de desenvolvimento do produto. Na análise de confiabilidade de um novo produto, esta extensão permite a obtenção de estimativas de confiabilidade mesmo antes de se completar o período de garantia das primeiras unidades vendidas, além de contribuir para uma estimativa inicial de um modelo paramétrico.

## REFERÊNCIAS

- BACHA, M.; CELEUX, G.. **Bayesian estimation of a Weibull distribution in a highly censored and small sample setting**. Unité de recherche INRIA Rhône-Alpes, França, 1996.
- CACCIARI, M.; MAZZANTI, G.; MONTANARI, G.C. Comparison of Maximum Likelihood Unbiasing Methods for the Estimation of Weibull Parameters. **IEEE Transactions on Dielectrics and Electrical Insulation**, v.3, n.1, p.18-27, 1996.
- CRICHTON, N. Wilcoxon Signed Rank Test. **Journal of Clinical Nursing**, v.9, p.584, 2000.
- DODSON, Bryan. **Weibull Analysis**. Milwaukee: ASQC Quality Press, 1994.
- ESCOBAR, Luis A.; MEEKER, William Q. Statistical Prediction Based on Censored Life Data. **Technometrics**, v.41, n.2, p.113-124, maio 1999.
- FAWZAN, Mohammad A. **Methods for Estimating the Parameters of the Weibull Distribution**. King Abdulaziz City for Science and Technology, p.1-11. Arábia Saudita, maio 2000.
- FORSTER, Jon. **Survival Analysis**. University of Southampton, Southampton, 2004.  
Disponível em: [www.maths.soton.ac.uk/staff/jjforster/teaching/math6021/chapter2.pdf](http://www.maths.soton.ac.uk/staff/jjforster/teaching/math6021/chapter2.pdf).  
Acesso em: maio 2004.
- FRITSCH, Celso. **Modelos de Confiabilidade para Testes de Sobrevida**. Dissertação de Mestrado, Programa de Pós-Graduação em Engenharia de Produção, UFRGS. Porto Alegre, 1996.
- GEORGE, Larry. **Kaplan-Meier Reliability Estimator Spreadsheet**. Disponível em: <http://home.comcast.net/~pstlarry/>. Acesso em: jun. 2004.
- IFA. **The Wilcoxon Matched-Pairs Signed-Ranks Test**. Amsterdam: Institute of Phonetic Sciences, 2004. Disponível em: [http://www.fon.hum.uva.nl/Service/Statistics/Signed\\_Rank\\_Test.html](http://www.fon.hum.uva.nl/Service/Statistics/Signed_Rank_Test.html). Acesso em: 28/01/2005.

- JAUW, James; VASSILIOU, Pantelis. Field Data is Reliability Information: Implementing an Automated Data Acquisition and Analysis System. **IEEE Proceedings Annual Reliability and Maintainability Symposium**, p.86-93, 2000.
- KALBFLEISH, J.D.; LAWLESS, J. F.; ROBINSON, J. A. Methods for the Analysis and Prediction of Warranty Claims. **Technometrics**, v.33, n.3, p.273-285, ago. 1991.
- KAMINSKIY, M. P.; KRISTOV, V. V.. An approach to evaluating the prior distribution of Weibull parameters. **Fourth International Conference on Mathematical Methods in Reliability**, USA, 2004.
- KARIM, Rezaul; YAMAMOTO, Wataru; SUZUKI, Kazuyuki. Statistical Analysis of Marginal Count Failure Data. **Lifetime Data Analysis**, v.7, p.173-186, 2001.
- KLEYNER, Andre; BENDER, Mark. Enhanced Reliability Prediction Method Based on Merging Military Standards Approach with Manufacturer's Warranty Data. **IEEE Proceedings Annual Reliability and Maintainability Symposium**, p.202-206, 2003.
- LOWRY, R. **Concepts and Applications of Inferential Statistics**. Vassar College, 2005. Disponível em: <http://faculty.vassar.edu/lowry/webtext.html>. Acesso em: 23/01/2005.
- MARCORIN, José Adilson; ABACKERLI, Alvaro J. Estudo de Fontes Potenciais de Dados de Falhas para a Previsão de Confiabilidade. **Anais do XXII Encontro Nacional de Engenharia de Produção**. Curitiba, PR, out. 2002.
- MEEKER, William Q. **Location-Scale-Based Parametric Distributions**. Statistics 533 Class notes. Department of Statistics, Iowa State University. Disponível em: [www.public.iastate.edu/~wqmeeker/stat533stuff/psnups/chapter04\\_psnup.pdf](http://www.public.iastate.edu/~wqmeeker/stat533stuff/psnups/chapter04_psnup.pdf). Acesso em: jun 2003.
- MEEKER, William Q; HAMADA, Michael. Statistical Tools for the Rapid Development & Evaluation of High Reliability Products. **IEEE Transactions on Reliability**, v.44, n.2, p.187-198, 1995.
- NIST/SEMATECH. **e-Handbook of Statistical Methods**. National Institute of Standards and Technology, Gaithersburg, 2004. Disponível em: [www.itl.nist.gov/div898/handbook/](http://www.itl.nist.gov/div898/handbook/). Acesso em: jun. 2004.

OH, Y.S.; BAI, D.S. Field Data Analysis with Additional After-Warranty Failure Data.

**Reliability Engineering and System Safety**, v.72, p.1-8, 2001.

PONGO, Ruth M.R.; FOLLEDO, Manuel; BUENO NETO, Pedro R. Uma Metodologia

Inovativa para Estudos de Confiabilidade de Produto. **Anais do XVI Encontro**

**Nacional de Engenharia de Produção**. Piracicaba, SP, out. 1996.

RAC. **Common Reliability Distributions**. Rome, NY: Reliability Analysis Center, 2001.

Disponível em: <http://rac.alionscience.com/pdf/CommonReliabilityDistributions.pdf>.

Acesso em: 29/12/2004.

RAI, Bharatendra; SINGH, Nanua. Hazard Rate Estimation from Incomplete and Unclean

Warranty Data. **Reliability Engineering and System Safety**, v.81, p.79-92, 2003.

RELIASOFT. **Life Data Analysis Reference**. Reliasoft Corporation. Disponível em:

[www.weibull.com/lifedatawebcontents.htm](http://www.weibull.com/lifedatawebcontents.htm). Acesso em: jun. 2004.

RIBEIRO, J.L.D. **Confiabilidade de Componentes e Sistemas**. UFRGS/PPGEP, Porto

Alegre, 2003. Paginação irregular.

U.S. DEPARTMENT OF DEFENSE. Reliability Prediction of Electronic Equipment.

**Military Handbook MIL-HDBK-217F**. 205 p., dez. 1991.

U.S. DEPARTMENT OF DEFENSE. Electronic Reliability Design Handbook. **Military**

**Handbook MIL-HDBK-338B**. p.37-44, out. 1998.

WERNER, Liane. **Modelagem dos Tempos de Falhas ao Longo do Calendário**.

Dissertação de Mestrado, Programa de Pós-Graduação em Engenharia de Produção, UFRGS. Porto Alegre, 1996.

WILKINS, D. J. The Bathtub Curve and Product Failure Behavior – Part One. **Reliability**

**HotWire**, v. 21, 10 p., 2002. Reliasoft Corporation. Disponível em:

<http://www.weibull.com/hotwire/issue21/hottopics21.htm>. Acesso em: 13/11/2004.

WOLFE, R. **Kaplan-Meier and Nelson-Aalen with right censored and left truncated**

**data**. University of Michigan, Michigan, 2004. Disponível em: [www-personal.umich.edu/~bobwolfe/675/4Survival/survival.pdf](http://www-personal.umich.edu/~bobwolfe/675/4Survival/survival.pdf). Acesso em: 12/07/2004.

WU, Huaiqing; MEEKER, William.Q. Early Detection of Reliability Problems Using Information from Warranty Databases. **Technometrics**, v.44, p.120-133, maio 2002.

YANG, Guangbin.; ZAGHATI, Ziad. Two-Dimensional Reliability Modeling From Warranty Data. **IEEE Proceedings Annual Reliability and Maintainability Symposium**, p.272-278, 2002.