

Universidade Federal De Minas Gerais
Instituto de Ciências Exatas
Departamento de Estatística
Programa de Pós-Graduação em Estatística

**MÉTODOS NUMÉRICOS DE ESTIMAÇÃO PARA
MODELOS DE DEGRADAÇÃO COM EFEITOS
ALEATÓRIOS NÃO NORMAIS**

Mestrando: Clódio Pereira de Almeida
clodioalm@gmail.com

Orientadora: Profa Dra. Marta Afonso de Freitas
marta@dep.ufmg.br

Co-orientador: Prof. Dr. Enrico Antônio Colosimo
enricoc@est.ufmg.br

Belo Horizonte
Junho 2011

Clódio Pereira de Almeida

**MÉTODOS NUMÉRICOS DE ESTIMAÇÃO PARA
MODELOS DE DEGRADAÇÃO COM EFEITOS
ALEATÓRIOS NÃO NORMAIS**

Dissertação apresentada ao Programa de Pós-Graduação do Departamento de Estatística do Instituto de Ciências Exatas (ICEX) da Universidade Federal de Minas Gerais, para a obtenção do Título de Mestre em Estatística, Concentração em Análise de Sobrevivência e Confiabilidade.

Orientadora: Marta Afonso Freitas

Co-orientador: Enrico Colosimo

**Belo Horizonte
Fevereiro
2011**

Agradecimentos

Diversas pessoas foram responsáveis pela materialização desta dissertação, entre as quais especialmente,

agradeço à minha orientadora, Dr^a. Marta Afonso Freitas pelo apoio incansável, paciência, incentivo, e pela capacidade de esclarecer dúvidas e iluminar caminhos, tornando mais fácil a caminhada;

ao meu co-orientador, Dr. Enrico Antônio Colosimo, pela perspicácia e entusiasmo, objetividade das análises, sugestões e conselhos;

à professora aposentada e tia querida, Ana Lúcia Alves Pereira, pelo “empurrão” inicial e entusiasmo contagiante;

aos professores e funcionários do Programa de Pós-Graduação do departamento de Estatística do ICEX;

aos colegas do programa de pós-graduação em Engenharia de Produção, especialmente Rívert Oliveira e Maria Luíza Toledo, tutores e incentivadores sempre presentes;

à CAPES e CNPq pelo financiamento da pesquisa;

e, sobretudo, à minha querida esposa Marília, pela compreensão e companheirismo, pelo amor e carinho e pelo suporte material e espiritual constantes durante a elaboração deste trabalho.

Resumo

Modelos de degradação são cada vez mais usados para se avaliar a confiabilidade de diversos tipos de produtos, especialmente aqueles com pequena probabilidade de falhar durante o tempo dos testes de vida tradicionais ou testes de vida acelerados. Os pacotes computacionais existentes para se ajustar modelos de degradação, assumem que os efeitos aleatórios nestes modelos têm distribuição normal. Estimativas para os parâmetros do modelo quando é violada a suposição de normalidade mostraram-se viciadas. Neste trabalho, implementamos solução para análise de dados de degradação modelados por perfis lineares simples e efeito aleatório com distribuição não normal, que consiste numa aproximação numérica da verossimilhança com a transformação sugerida por Nelson *et al.* (2006). A partir de simulações e estudo de casos reais, comparamos as estimativas para a distribuição do tempo de falha obtidas pelo método implementado com aquelas obtidas com a utilização dos softwares estatísticos padrões (por exemplo, os pacotes LME e NLME do *Software R 2.11*) e pelos métodos aproximado e análise de falha tradicional (FTA).

Palavras-chave: Modelos de Degradação, Máxima Verossimilhança, Probability Integral Transformation (PIT)

Lista de Figuras

1.1	Croquis do Trem Unidade Elétrico	11
1.2	Codificação identificativa das rodas dos TUE	11
1.3	Gráfico dos dados de degradação das 14 rodas MA11	12
1.4	Gráfico dos dados de degradação das unidades de LASER	14
1.5	Crescimento de trincas por fadiga - Bogdanoff e Kozin (1985)	17
4.1	Função densidade de probabilidade - Weibull	52
4.2	Perfis de degradação simulados Caso 1: modelo linear com efeito aleatório Weibull	54
4.3	Medidas de desempenho estimativas do MTTF Caso 1: modelo linear com efeito aleatório Weibull	56
4.4	Medidas de desempenho estimativas do quantil 0,01 Caso 1: modelo linear com efeito aleatório Weibull	57
4.5	Medidas de desempenho estimativas da mediana Caso 1: modelo linear com efeito aleatório Weibull	58
4.6	Função densidade de probabilidade - Lognormal	59
4.7	Perfis de degradação simulados Caso 2: modelo linear com efeito aleatório Lognormal	61

4.8	Medidas de desempenho estimativas do MTTF Caso 2: modelo linear com efeito aleatório Lognormal	62
4.9	Medidas de desempenho estimativas do quantil 0,01 Caso 2: modelo linear com efeito aleatório Lognormal	63
4.10	Medidas de desempenho estimativas da mediana Caso 2: modelo linear com efeito aleatório Lognormal	64
5.1	Detalhe dos perfis de degradação das rodas MA11	72
5.2	Gráfico de $g^*(a_i)$ (integrando de 5.2) para cada uma da 14 rodas (exceto 7 e 13)	74
5.3	Gráfico de $g^*(a_i)$ (integrando de 5.4) para cada uma da 14 rodas (exceto 7 e 13)	76
5.4	Gráfico de probabilidade para $\hat{\beta}_i$	82

Lista de Tabelas

4.1	Limites de degradação (D_f) por percentual de censuras Caso 1: modelo linear com efeito aleatório Weibull	53
4.2	Figuras de mérito reais Caso 1: modelo linear com efeito aleatório Weibull	53
4.3	Limites de degradação (D_f) por percentual de censuras Caso 2: modelo com efeito aleatório Lognormal	60
4.4	Figuras de mérito reais para perfis simulados Caso 2: modelo com efeito aleatório Lognormal	60
5.1	Medidas de degração (mm.) das rodas MA11	69
5.2	Rodas - estimativas iniciais para β_i (Weibull)	70
5.3	Rodas - estimativas iniciais para β_i (Lognormal)	78
5.4	Estimativas pontuais e intervalos de confiança	81
5.5	Medidas de degração (porcentagem) das unidades de laser	85
5.6	Laser - estimativas iniciais para β_i	86
5.7	Resumo de algumas figuras de mérito para unidades de laser	88
A.1	Resultado das simulações Caso 1: modelo linear com efeitos aleatórios com distribuição de Weibull	101
A.2	Resultado das simulações Caso 2: modelo linear com efeitos aleatórios com distribuição Lognormal	102

Sumário

1	Introdução	10
1.1	Situações práticas motivadoras	10
1.1.1	Degradação de rodas de trem unidade elétrico (TUE)	10
1.1.2	Lasers	13
1.2	Revisão de literatura	14
1.3	Modelo geral de degradação	21
1.3.1	Avaliação de $F_t(t)$	22
1.3.2	O problema da estimação dos parâmetros	24
1.4	Objetivo do trabalho	26
1.5	Estrutura do texto	26
2	Métodos padrão de análise de dados de degradação	27
2.1	O método de análise aproximado	27
2.2	O modelo geral de degradação - estimação por MV	29
2.2.1	O modelo baseado na distribuição normal multivariada	34
3	Métodos computacionais para estimativas MV	36
3.1	O método “PIT” (probability integral transformation)	37
3.2	Método de reformulação da verossimilhança	39
3.3	Maximização da verossimilhança - solução numérica baseada no PIT	40

3.3.1	Efeitos Aleatórios com distribuição de Weibull	41
3.3.2	Efeitos Aleatórios com distribuição Lognormal	46
4	Os resultados de Toledo (2007) e Freitas <i>et al.</i> (2009) revisitados	50
4.1	Simulando efeitos aleatórios com distribuição de Weibull	51
4.2	Simulando efeitos aleatórios com distribuição Lognormal	59
4.3	Conclusões	65
5	Situações motivadoras revisitadas	68
5.1	Rodas de trem	68
5.1.1	Modelo de degradação linear - Weibull	70
5.1.2	Modelo de degradação linear - lognormal	78
5.1.3	Conclusões	80
5.2	Unidades de laser	84
5.2.1	Conclusões	88
6	Conclusões e trabalhos futuros	90
	Referências Bibliográficas	94
A	Tabelas de resultados das simulações	100
B	Rotinas implementadas para o <i>Software R-11</i>	103
B.1	Simulação de modelo com efeitos aleatórios Weibull	104
B.2	Simulação de modelo com efeitos aleatórios Lognormal	110
B.3	Dados rodas de trens - distribuição Weibull	116
B.4	Bootstrap paramétrico e não paramétrico para rodas trens	119

Capítulo 1

Introdução

1.1 Situações práticas motivadoras

1.1.1 Degradação de rodas de trem unidade elétrico (TUE)

Defeitos em rodas são responsáveis por boa parte dos acidentes com veículos ferroviários, com potencial de gerar custos extraordinários de bilhões de dólares para as empresas ao redor do mundo. Especialmente no transporte de passageiros, as consequências de acidentes são muito mais sérias, por exporem a risco vidas humanas. Para ajudar a evitar eventos catastróficos, bem como reduzir custos operacionais, as ferrovias monitoram o desempenho das rodas de seus veículos, com o objetivo de antecipar qualquer defeito, permitindo que as mesmas sejam substituídas antes que uma falha ocorra.

As áreas responsáveis pela manutenção mecânica nas ferrovias normalmente mantêm banco de dados com informações detalhadas das intervenções preventivas, corretivas ou preditivas em seus carros, vagões e locomotivas. Iremos analisar um pequeno subconjunto de um banco de dados apresentado por Freitas *et al.* (2009), e que faz parte de um estudo abrangente conduzido por uma empresa ferroviária brasileira. O banco de dados

completo inclui, entre outras informações, medidas do diâmetro das rodas, obtidas em 13 inspeções realizadas a cada 50.000 *km* rodados: $t_1 = 0 \text{ km}$, $t_2 = 50.000 \text{ km}$, $t_3 = 100.000 \text{ km}$, ..., $t_{13} = 600.000 \text{ km}$. Estas medidas foram registradas para 14 TUEs, cada um deles composto de 2 carros motores (MA e MB) e dois carros reboques (RA e RB), como ilustrado na figura 1.1.

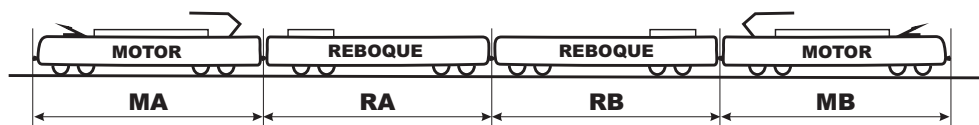


Figura 1.1: Croquis do Trem Unidade Elétrico

O estudo focará os carros MA, que são os responsáveis por tracionar os outros três num determinado sentido de tráfego. O processo de degradação das rodas é mais acelerado no carro motor, particularmente nas rodas do primeiro eixo deste carro.

A identificação de uma roda num carro em particular é codificada pelo prefixo do carro (MA, RA, RB ou MB), pelo número do eixo (são quatro eixos por carro) e pelo lado da roda no eixo: esquerda (1) ou direita (2), conforme croquis mostrado na figura 1.2.

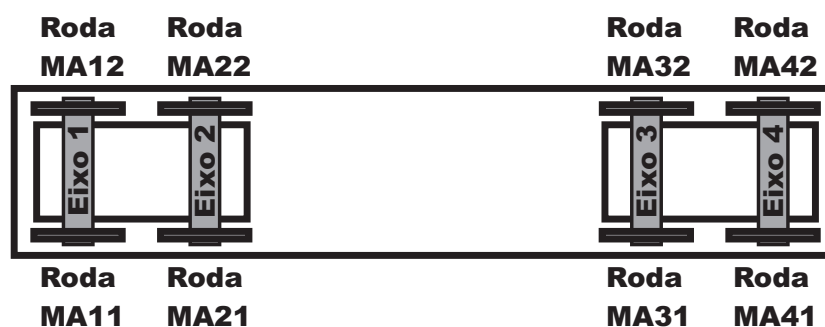


Figura 1.2: Codificação identificativa das rodas dos TUE

Neste trabalho usaremos os dados referentes às medidas do diâmetro das rodas do lado esquerdo do eixo número 1 de cada um dos quatorze carros MA (rodas MA11). O

diâmetro de uma roda nova é 966 mm. À medida que os trens são operados as rodas vão se desgastando, e quando seu diâmetro atinge 889 mm, ela é substituída por uma nova. Na modelagem serão usadas as medidas de desgaste observadas por ocasião de cada inspeção, obtidas por:

$$Y_{ij} = 966 \text{ mm} - \varphi_{ij} \quad \text{com } i = 1, \dots, 14 \text{ e } j = 1, \dots, 13;$$

em que φ_{ij} é o diâmetro medido na i -ésima roda na j -ésima inspeção e Y_{ij} o desgaste correspondente. A figura 1.3 apresenta o gráfico de degradação das 14 rodas estudadas.

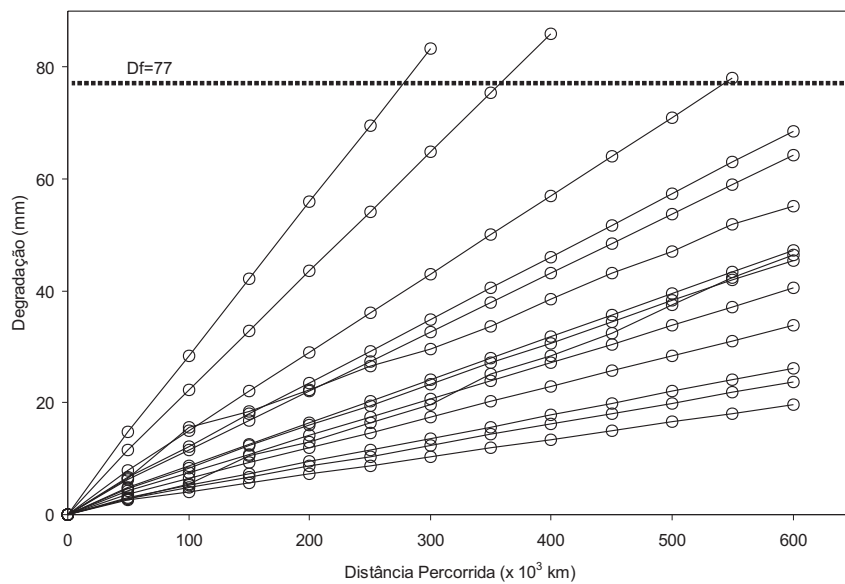


Figura 1.3: Gráfico dos dados de degradação das 14 rodas MA11

Define-se então que ocorre uma “falha” numa roda, quando a degradação atingir o nível limiar crítico $D_f = 77$ mm ($966 \text{ mm} - 889 \text{ mm}$). Três dentre as quatorze rodas estudadas atingiram o nível limiar crítico durante o período do estudo.

Freitas *et al.* (2009) utilizaram estas medidas de degradação para estimar a distri-

buição do tempo de vida $F_T(t)$ das rodas ferroviárias, e a partir desta distribuição, outras características importantes como o tempo médio até a falha (MTTF) e alguns quantis principais. Iremos analisar os dados de degradação das mesmas rodas, utilizando a solução sugerida neste trabalho, comparando os resultados com os apresentados por Freitas *et al.* (2009)

1.1.2 Lasers

Alguns dispositivos para amplificação da luz por emissão estimulada de radiação, conhecidos pela sigla em inglês LASER (*Light Amplification by Stimulated Emission of Radiation*) apresentam degradação ao longo do tempo, que leva à redução da luminosidade emitida. No entanto, é possível manter esta luminosidade praticamente constante, desde que se aumente a corrente de operação. Quando esta corrente atinge um valor muito alto, considera-se que houve uma falha no dispositivo.

Meeker e Escobar (1998) apresentaram um estudo com dados de degradação de 15 unidades de LASER do tipo GaAs (composto com os elementos Gálio e Arsênio), acompanhadas durante 4.000 horas de operação, onde se mediu o acréscimo percentual de corrente em relação à corrente nominal do início do estudo. Estas medidas foram obtidas em intervalos igualmente espaçados, a cada 250 horas. Neste experimento as medidas de degradação são os percentuais de aumento da corrente para cada unidade, e definiu-se que um aumento de 10% na corrente de trabalho caracteriza uma falha. Os autores utilizaram um método que calcula pseudotempos de falhas, para estimar a distribuição do tempo de vida das unidades, construindo intervalos de confiança aproximados para as estimativas.

Hamada (2005) revisou os dados das unidades de LASER e utilizou um método de inferência bayesiana para analisá-los, apresentando entre outros resultados, a distribuição *a posteriori* do quantil 0,10 da distribuição do tempo até a falha. Comparou estes resultados com os obtidos por análises bayesianas utilizando os mesmos valores *a priori*

para os parâmetros, em outros dois métodos. No primeiro ele usou os tempos até a falha (censurados por intervalo em 3 unidades e censurados à direita em 12 unidades); e no segundo pseudotempos de falhas.

A Figura 1.4 mostra os perfis de degradação das 15 unidades.

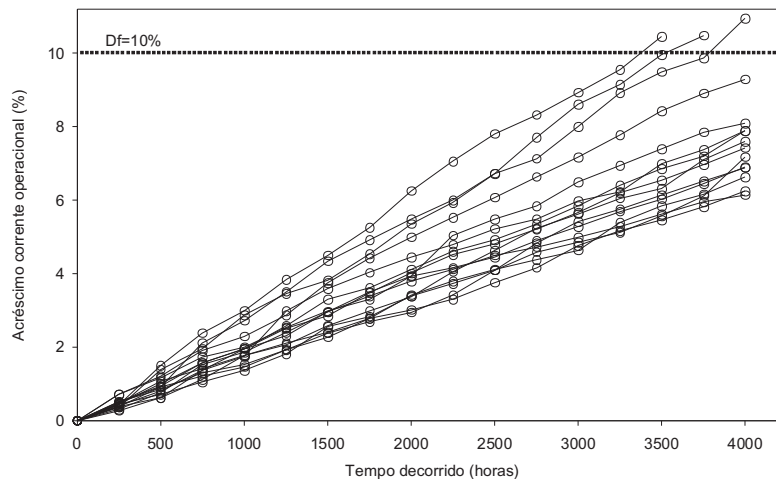


Figura 1.4: Gráfico dos dados de degradação das unidades de LASER

Iremos aplicar a solução implementada no presente trabalho aos dados das unidades de LASER, comparando os resultados com os apresentados pelos autores citados.

1.2 Revisão de literatura

Em um mundo cada vez mais globalizado e competitivo, as margens de lucro das empresas estão em constante queda, tornando-se imperativo para qualquer negócio a manutenção e ampliação da base de clientes. Por outro lado, o aumento exponencial do acesso à informação e a abertura de novos mercados fornecedores, aumenta o nível de exigência dos clientes e reduz sua fidelização a marcas e a fabricantes. Por isso, a busca por melhoria constante da qualidade e confiabilidade de seus produtos é, cada vez mais,

uma questão de sobrevivência para as empresas.

Confiabilidade pode ser definida, como em Lewis (1996), como “a probabilidade de que um sistema irá desempenhar sua função esperada por um período específico de tempo, sob dado conjunto de condições”, e é caracterizada pela distribuição do tempo até a falha ou, de maneira equivalente, pelo tempo de vida T (neste texto os termos “tempo de vida” e “tempo até a falha” serão usados indiscriminadamente). O conhecimento desta função de distribuição nos permite estabelecer algumas quantidades importantes denominadas **figuras de mérito** entre as quais:

- MTTF (Mean Time to Failure): tempo médio até a falha;
- p -ésimo quantil da distribuição (t_p): tempo no qual se espera que $p \times 100\%$ dos itens venham a falhar ($\mathbb{P}[T \leq t_p] = p$);
- Tempo mediano ($t_{0,50}$): é o quantil 0,50;
- Função de confiabilidade - $R(t)$: esta função fornece a probabilidade do item funcionar por um tempo superior a t , ou seja $R(t) = \mathbb{P}[T \geq t]$.

Durante muito tempo, a distribuição do tempo até a falha era obtida por abordagens estatísticas tradicionais, como os testes de vida. Nestes testes, amostras dos itens em estudo eram submetidas a condições de trabalho idênticas às reais, observando-se o tempo até a ocorrência da falha. Com base nestas observações estimava-se as **figuras de mérito**. Entretanto, em ensaios deste tipo, é comum a presença de censuras, ou seja, ao final do ensaio um certo número de unidades pode ainda estar em funcionamento. Mesmo incorporando as observações censuradas na análise, se o número de censuras for muito alto, perde-se precisão na estimativa das quantidades de interesse. Esta dificuldade, aliada à pressão cada vez maior pela redução no tempo de duração dos ensaios e consequentemente de seus custos, levou ao surgimento dos ensaios de vida acelerados e de técnicas

estatísticas para modelagem e análise dos resultados de tais testes. Diversos artigos, livros e publicações trataram destas abordagens, entre os quais Nelson (1982), Nelson (1990), Meeker e Escobar (1998), Elsayed (1996), Freitas e Colosimo (1997), e Colosimo e Giolo (2006), entre outros. Entretanto produtos cada vez mais confiáveis provocam situações em que pouquíssimas ou nenhuma falha ocorre durante o período de testes, mesmo nos testes de vida acelerados.

Se o objeto em estudo possuir características cuja degradação ao longo do tempo estão relacionadas com o tempo de vida deste produto, então informações sobre a evolução desta degradação podem fornecer dados importantes sobre a confiabilidade do item estudado, mesmo nos casos em que não ocorram falhas durante o período do estudo. Surgem assim os modelos de degradação, nos quais observa-se a evolução de características físicas indicadoras do desempenho do item estudado ao longo do experimento. Exemplos destas características são: trincas, corrosão, oxidação, redução dimensional, desgaste, etc. Estas características são chamadas “mecanismos de degradação”. Um nível crítico para a medida de degradação de interesse é pré-definido, e a partir deste nível configura-se a falha. Esta medida é conhecida na literatura como “limiar de falha” ou “*failure threshold*”. Durante o período do experimento, obtém-se medidas de degradação em instantes de tempo pré-fixados e define-se que a falha ocorre, para uma dada unidade experimental, quando a quantidade de degradação para a mesma exceder o nível crítico. Estas medidas de degradação podem fornecer informações úteis para se estimar a confiabilidade mesmo se o limiar de degradação não for ultrapassado por nenhuma unidade amostral.

Existem referências importantes que usaram dados de degradação para estimar a confiabilidade. Gertsbackh e Kordonskiy (1969) estudaram a degradação de um ponto de vista da engenharia, chamando a atenção para a importância de se analisar medidas de degradação e seus respectivos perfis amostrais para se estimar a confiabilidade, e apresentaram a distribuição Bernstein, que descreve a distribuição dos tempos até a falha por

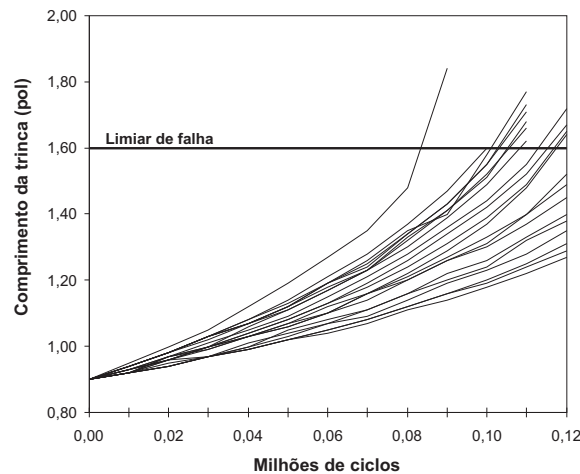


Figura 1.5: Crescimento de trincas por fadiga - Bogdanoff e Kozin (1985)

um modelo linear simples com intercepto e inclinação aleatórios. Nelson (1981) estudou a situação especial em que a medida de degradação é destrutiva (uma só medida por item). Bogdanoff e Kozin (1985) usaram uma abordagem probabilística para modelar comprimento de trincas por fadiga em metal, e apresentaram diversos gráficos de amostras de perfis para descrever diferentes tipos de degradação, como o exemplo da Figura 1.5. Nelson (1990) revisou a literatura sobre degradação e aplicações em sobrevivência, e descreveu ideias básicas para modelos de testes acelerados de degradação; usando um exemplo específico com dados de voltagem de ruptura de isolantes submetidos a diferentes temperaturas e tempos de exposição, mostrou como analisar dados de degradação. As duas situações práticas motivadoras apresentadas na seção 1.1 são exemplos de dados de degradação.

Na literatura existem basicamente duas abordagens para a modelagem de dados de degradação. A primeira considera a degradação como um processo aleatório que depende apenas do tempo. Doksum (1991) usou um modelo de processo de Wiener para analisar dados de degradação. Tang e Chang (1995) modelaram dados de degradação acelerada não destrutiva de unidades fornecedoras de energia como uma coleção de processos es-

tocásticos. Whitmore e Shenkelberg (1997) consideraram que o processo de degradação em seu modelo é um processo de difusão de Wiener com transformação na escala do tempo e ilustraram seus modelos e métodos de inferência com um estudo de caso envolvendo cabos de aquecimento auto-reguláveis.

Na segunda abordagem, a degradação é modelada por modelos estatísticos mais gerais, que são função do tempo e de algumas variáveis aleatórias, eventualmente multidimensionais. Estes modelos são chamados modelos gerais de perfis de degradação.

Nos modelos gerais de perfis de degradação, a análise é, em geral, implementada em duas etapas:

1. A primeira etapa consiste na construção de um modelo misto (linear ou não) que explique o perfil de degradação ao longo do tempo e que permita a estimação dos parâmetros desconhecidos deste modelo. Esta etapa é, portanto, uma análise de dados longitudinais utilizando modelo não linear (ou linear) misto, isto é, que incorpora em sua forma funcional tanto efeitos aleatórios quanto fixos (Pinheiro e Bates, 2000). Esta primeira etapa da análise dos dados de degradação (ajuste do modelo para dados longitudinais), pode por sua vez ser vista como uma modelagem de dois estágios, na qual inicialmente (primeiro estágio) parâmetros não observáveis (os efeitos aleatórios) são amostrados de uma distribuição que por si só, possui parâmetros (fixos) desconhecidos. Em seguida (segundo estágio), os dados observados no ensaio são então considerados como sendo amostras (realizações) de distribuições Normais independentes, cujas médias são funções lineares (ou não lineares) destes efeitos aleatórios e possivelmente de outros parâmetros populacionais (efeitos fixos) como em Verbeke e Molenberghs (2000). Uma vez que uma família paramétrica é escolhida para os efeitos aleatórios é preciso estimar os parâmetros desta distribuição e os demais parâmetros fixos que fazem parte da forma funcional do perfil de degradação.

2. A segunda etapa da análise consiste na estimação da distribuição do tempo de falha $F_T(t)$ e outras características de interesse, a partir do modelo ajustado para os perfis de degradação e utilizando as estimativas dos parâmetros do modelo obtidas na primeira etapa. Quando o modelo de degradação não é suficientemente simples para que se obtenha uma forma fechada para a distribuição do tempo de falha $F_T(t)$ nesta segunda etapa, pode-se conseguir uma estimativa desta distribuição através de simulação de Monte Carlo.

Na literatura, vários trabalhos que utilizaram esta abordagem podem ser citados. Lu e Meeker (1993) usaram um modelo não linear de efeitos mistos e desenvolveram um método de dois estágios para obter estimativas pontuais e intervalos confiança para os percentis da distribuição do tempo de falha, e aplicaram o método de dois estágios aos dados de crescimento de trincas por fadiga de Bogdanoff e Kozin (1985). Tseng *et al.* (1995) apresentaram um estudo de caso onde utilizaram dados de degradação em um experimento planejado, seguindo um modelo fatorial fracionado para melhorar a confiabilidade de lâmpadas fluorescentes. Yacout *et al.* (1996) usaram dados de degradação para estimar a distribuição do tempo de falha de bombas combustível nuclear, com base em um modelo de falha com limiar fixo e a abordagem de estimação em dois estágios proposta por Lu e Meeker (1993). Lu *et al.* (1997) propuseram um modelo com coeficientes de regressão aleatórios e função de desvio padrão para analisar dados de degradação de semicondutores. Su *et al.* (1999) consideraram um modelo de degradação com coeficientes aleatórios usando máxima verossimilhança para estimação de parâmetros, ilustrado por um aplicativo com dados de semicondutor. Wu e Shao (1999) estabeleceram as propriedades assintóticas dos estimadores de mínimos quadrados ponderados sob o modelo NLME (*Non Linear Mixed Effects*) e obtiveram, a partir destas propriedades, estimativas pontuais e intervalos de confiança aproximados para percentis da distribuição do tempo de falha, aplicando seu método em conjuntos de dados de resistores metálicos de Zhuang

(1994) e do crescimento do comprimento de trincas por fadiga em metais apresentado por Lu e Meeker (1993). Wu e Tsai (2000) apresentaram um método de estimação em que os perfis de degradação eram agrupados de acordo com características especiais comuns, e cada grupo recebia peso diferente (ponderação difusa) para modificar o procedimento de dois estágios proposto em Lu e Meeker (1993); compararam seu método e o de Lu e Meeker (1993), utilizando os dados de resistores metálicos de Wu e Shao (1999) obtendo intervalos de confiança mais estreitos para as estimativas da distribuição do tempo de falha. Crk (2000) apresentou nova metodologia para tratar análise de falhas decorrentes de um ou mais mecanismos de degradação: para cada mecanismo é estabelecido um modelo cujos parâmetros podem ser aleatórios, correlacionados ou dependentes do estresse, e uma regressão multivariada múltipla com os parâmetros de cada mecanismo permite determinar o modelo final. Jiang e Zhang (2002) propuseram um modelo dinâmico para dados de degradação, analisando o crescimento aleatório de trincas por fadiga, em que a variável resposta é modelada como um processo de Markov, admitindo que medidas sucessivas em uma mesma unidade são correlacionadas. Oliveira e Colosimo (2004) compararam os métodos analítico, aproximado e da verossimilhança com efeitos aleatórios com distribuição normal (“numérico”) para obter estimação dos tempos de falha em modelos de degradação usando dados simulados de desgaste da banda de rodagem de pneus de automóveis; concluindo que no caso em que o perfil de degradação é linear ou suficientemente simples, os métodos aproximado e analítico fornecem resultados similares, mas piores que os resultados do método da verossimilhança normal, que fornece intervalos de confiança mais estreitos; no entanto para modelos não lineares ou com mais de um efeito aleatório recomendam somente este último método. Finalmente, Freitas *et al.* (2009) compararam através de um estudo por simulação, as estimativas de quantis da distribuição do tempo até a falha obtidas pelos métodos da verossimilhança com efeitos aleatórios normais, aproximado e por uma análise FTA (*Failure Time Analysis*) convencional; si-

mularam perfis de degradação com forma funcional próxima a uma reta com intercepto zero e coeficientes associados à inclinação (efeito aleatório) com distribuições Lognormal, de Weibull e Normal (Gaussiana); e concluíram a partir das simulações que o método da verossimilhança normal é altamente afetado pela violação da suposição de normalidade dos efeitos aleatórios, em particular a magnitude do vício das estimativas dos quantis é significativamente aumentada quando a distribuição utilizada não é a distribuição normal. Alternativas para superar este problema serão discutidas ao longo deste trabalho.

1.3 Modelo geral de degradação

Em testes de degradação, para uma amostra aleatória de unidades experimentais, obtemos medidas da característica física escolhida à medida que a mesma se degrada ao longo do tempo, para cada unidade individualmente. A degradação de cada unidade é modelada usando-se a mesma forma funcional e as diferenças entre as unidades incorporadas ao modelo através dos efeitos aleatórios. O modelo é:

$$Y_{ij} = D_{ij} = D(t_{ij}; \alpha; \beta_i) + \varepsilon_{ij} , \quad (1.1)$$

em que Y_{ij} é a variável aleatória que representa a medida de degradação para a i -ésima unidade amostral no j -ésimo instante de medição, com $i = 1, 2, \dots, n$ e $j = 1, 2, \dots, m_i$; $D(t_{ij}; \alpha; \beta_i)$ é o perfil real de degradação da unidade i no tempo t_{ij} ; $\alpha = (\alpha_1; \alpha_2; \dots; \alpha_p)'$ é o vetor de efeitos fixos (características populacionais comuns a todas as unidades); $\beta_i = (\beta_{i1}, \beta_{i2}, \dots, \beta_{ik})'$ o vetor de efeitos aleatórios referentes à i -ésima unidade; ε_{ij} o erro aleatório associado à i -ésima unidade no instante t_{ij} . Além disto, assumimos que ε_{ij} são independentes e identicamente distribuídos (i.i.d.) segundo uma distribuição normal com média 0 e variância σ_ε^2 (desconhecida).

Os vetores aleatórios $\beta_i = (\beta_{i1}, \beta_{i2}, \dots, \beta_{ik})'$, $i = 1, 2, \dots, n$; possuem distribuição

multivariada $\Lambda(\beta|\theta)$, em que o vetor de parâmetros $\theta = (\theta_1, \dots, \theta_q)'$ (fixo) deve ser estimado a partir dos dados de degradação. Assumimos que ε_{ij} e β_i são independentes. Eventualmente Y e t poderão estar em escalas transformadas, se necessário.

A proporção de falhas no instante t é equivalente à proporção de perfis de degradação que excedem o nível crítico D_f até o instante t . Assim sendo, podemos definir a distribuição do tempo até a falha T para modelo (1.1) como:

$$F_T(t) = F_T(t; \alpha; \Lambda(\beta|\theta); D_f; D) = \mathbb{P}(T \leq t) = \mathbb{P}[D(t; \alpha; \beta) \geq D_f],$$

quando as medidas degradação são crescentes no tempo, ou

$$F_T(t) = F_T(t; \alpha; \Lambda(\beta|\theta); D_f; D) = \mathbb{P}(T \leq t) = \mathbb{P}[D(t; \alpha; \beta) \leq D_f],$$

quando as medidas degradação são decrescentes no tempo.

Sob este modelo de degradação, é necessário obter-se estimativas de α (o vetor de efeitos fixos) e de θ , o vetor de parâmetros da distribuição dos efeitos aleatórios $\Lambda(\beta|\theta)$, para se estimar a distribuição do tempo até a falha $F_T(t)$ e as figuras de mérito de interesse.

1.3.1 Avaliação de $F_t(t)$

Quando modelos simples podem ser utilizados para descrever os perfis de degradação, uma vez conhecida a distribuição dos efeitos aleatórios $\Lambda(\beta|\theta)$, podemos expressar $F_T(t)$ de forma fechada, permitindo fácil obtenção das figuras de mérito.

Por exemplo, suponha que o perfil de degradação real de uma unidade em particular seja dado por

$$D(t) = \alpha + \beta t ,$$

em que α é fixo e representa a quantidade inicial de degradação comum a todas as unidades

experimentais no início do teste ($D(0) = \alpha$) e, por conseguinte, corresponde a um efeito fixo; β é a taxa de degradação que varia de unidade para unidade (corresponde a um efeito aleatório) de acordo, por exemplo, com uma distribuição lognormal com parâmetros de escala $\exp(\mu_\beta)$ e parâmetro de forma σ_β . Para o nível crítico D_f temos:

$$D_f = \alpha + \beta T \quad \text{e} \quad T = g(\beta; \alpha; D_f) = \frac{D_f - \alpha}{\beta},$$

em que g é a transformação da variável aleatória β na variável aleatória T , tempo até a falha. A função de distribuição acumulada de T , $F_T(t)$ é dada por:

$$\begin{aligned} F_T(t) &= \mathbb{P}(T \leq t) = \mathbb{P}\left(\frac{D_f - \alpha}{\beta} \leq t\right) \\ &= \mathbb{P}\left(\beta \geq \frac{D_f - \alpha}{t}\right) = 1 - \Phi_{\text{nor}}\left(\frac{\log(D_f - \alpha) - \log(t) - \mu_\beta}{\sigma_\beta}\right) \\ &= \Phi_{\text{nor}}\left(\frac{\log(t) - [\log(D_f - \alpha) - \mu_\beta]}{\sigma_\beta}\right), \quad t > 0, \end{aligned}$$

em que $\Phi_{\text{nor}}(\cdot)$ é a função de distribuição acumulada de uma variável aleatória com distribuição normal padrão. Neste caso, T tem também uma distribuição lognormal com parâmetro de escala $t_{0,50} = \exp(\mu_T)$ em que $\mu_T = [\log(D_f - \alpha) - \mu_\beta]$ e parâmetro de forma $\sigma_T = \sigma_\beta$. Outras distribuições podem ser usadas e os mesmos procedimentos para obter-se a distribuição de T também podem ser aplicados. Resultados com outras distribuições como a Weibull, Normal (Gaussiana) e Bernstein podem ser encontrados em Lu e Meeker (1993).

Entretanto, especialmente nos casos em que $D(t_{ij}; \alpha; \beta_i)$ não é linear e o modelo tem mais de um efeito aleatório, a obtenção de $F_T(t)$ pode ser complicada. Nestes casos é, geralmente, possível a obtenção da distribuição do tempo de vida numericamente, para valores fixos de α , $\Lambda(\beta|\theta)$, D_f e D , através de simulação de Monte Carlo. Mas mesmo após especificarmos uma forma funcional para $D(t_{ij}; \alpha; \beta_i)$, definirmos o limiar

de falha D_f , e escolhermos uma distribuição para $\Lambda(\beta|\theta)$, teremos a tarefa de estimar $\alpha = (\alpha_1; \alpha_2; \dots; \alpha_p)'$ e $\theta = (\theta_1, \dots, \theta_q)'$, tanto nos casos em que é possível obter $F_T(t)$ analiticamente (como no exemplo apresentado) como nos casos em que é preciso recorrer à simulação por Monte Carlo.

1.3.2 O problema da estimação dos parâmetros

A literatura está repleta de artigos cujos autores propõem métodos para obter estimativas dos parâmetros desconhecidos dos modelos de degradação e, a partir delas, estimar a distribuição do tempo de falha. Lu e Meeker (1993) propuseram o método de dois estágios para modelos nos quais o vetor de efeitos aleatórios β , ou alguma transformação do mesmo, segue uma distribuição Normal Multivariada com média μ_β e matriz de variância-covariância Σ_β . Posteriormente Pinheiro e Bates (1995) desenvolveram rotinas computacionais para a estimação dos parâmetros por máxima verossimilhança e máxima verossimilhança restrita ou residual, ainda utilizando a suposição de que os efeitos aleatórios seguem uma distribuição Normal Multivariada. Mesmo abordagens alternativas publicadas na década de 1990 assumem erros aleatórios com distribuição normal, como Breslow e Clayton (1993) que utilizaram um método de estimação baseado em quase-verossimilhança penalizada, McGilchrist (1994) que comparou estimadores baseados no BLUP (*Best Linear Unbiased Prediction*) com os obtidos por máxima verossimilhança e máxima verossimilhança residual e Wolfinger e Lin (1997) que apresentaram dois métodos de cálculo aproximado da verossimilhança baseados em expansões de série de Taylor. Entretanto a suposição de normalidade dos efeitos aleatórios nem sempre é verdadeira em situações práticas. Dada a dificuldade de verificação de sua validade, muitos trabalhos na literatura têm tentado quantificar o efeito de desvios desta suposição nas estimativas dos parâmetros fixos tanto para esta classe de modelos (Modelos Lineares e Não Lineares de Efeitos Mistos) quanto para outras mais abrangentes, como a dos Modelos Lineares Gene-

realizados de Efeitos Mistos - GLMM (Generalized Linear Mixed Effects Models). Verbeke e Lesaffre (1997), Agresti *et al.* (2004), Litière *et al.* (2008), e Freitas *et al.* (2009) são alguns exemplos.

Em função da grande dificuldade de estimação dos parâmetros do modelo em situações nas quais a suposição de normalidade não se aplica, e dos efeitos desta má especificação, uma outra frente de pesquisa foi aberta. Nesta frente o foco é o desenvolvimento de métodos computacionais alternativos que permitam a obtenção de estimadores de máxima verossimilhança para modelos de efeitos mistos, utilizados na modelagem de dados de degradação, com efeitos aleatórios não-normais. Pinheiro *et al.* (2001) propuseram um modelo linear misto com efeitos aleatórios de uma distribuição T multivariada. Para modelos gerais com efeitos aleatórios não-normais, alguns métodos de estimação que aparecem na literatura incluem a verossimilhança hierárquica de Lee e Nelder (1996) e a maximização por partes de Song *et al.* (2005). Entretanto, a implementação destes métodos para utilização intensiva em análise de dados, mostrou-se uma tarefa desafiadora. A quadratura Gaussiana é uma ferramenta que tem sido utilizada com sucesso na estimação de parâmetros em modelos com efeitos aleatórios normais. Para modelos mistos com efeitos aleatórios não-normais, Nelson *et al.* (2006) propuseram um método computacional simples que utiliza a PIT (“Probability Integral Transformation”) para obter estimativas de máxima verossimilhança. Liu e Yu (2008) por sua vez propuseram um novo método de estimação, que reformula a verossimilhança condicional de efeitos aleatórios não-normais, através da divisão e multiplicação por uma função de densidade Normal. Os métodos apresentados por estes dois últimos trabalhos serão apresentados no próximo capítulo. Ressalta-se que há trabalhos em que a modelagem de dados de degradação foi feita sob o enfoque bayesiano, como por exemplo em Hamada (2005) e Freitas *et al.* (2010). Entretanto, esta abordagem foge do escopo desta dissertação.

1.4 Objetivo do trabalho

O objetivo deste trabalho é apresentar a implementação de uma solução para análise de dados de degradação quando os perfis podem ser modelados por funções simples, mas com efeitos aleatórios não normais. Após analisarmos os métodos “PIT” proposto por Nelson *et al.* (2006) e o método da reformulação da verossimilhança proposto por Liu e Yu (2008) optamos por utilizar o primeiro, com soluções de integração numérica e otimização diferentes das mostradas no artigo original. Aferimos a precisão das estimativas de quantis da distribuição do tempo até a falha assim obtidas, através de simulações como as apresentadas por Toledo (2007) e Freitas *et al.* (2009), comparando os resultados com os destes autores. Finalmente aplicamos a solução aos bancos de dados apresentados na Seção 1.1.

1.5 Estrutura do texto

Apresentamos no Capítulo 2 um resumo de dois métodos de análise de dados de degradação apresentados por Meeker e Escobar (1998) e utilizados nas simulações de Toledo (2007) e Freitas *et al.* (2009): o método aproximado e o baseado na verossimilhança com efeitos aleatórios com distribuição normal (“numérico”). No Capítulo 3 apresentamos com mais detalhes os novos métodos propostos por Nelson *et al.* (2006) e Liu e Yu (2008) e encerramos com a proposta de implementação de uma solução de maximização da verossimilhança para modelos lineares simples, utilizando a transformação sugerida por Nelson *et al.* (2006). No Capítulo 4 repetimos os estudos de simulações realizados por Freitas *et al.* (2009) e Toledo (2007), utilizando a solução sugerida neste trabalho. Finalmente esta solução é aplicada às situações práticas motivacionais, como mostrado no Capítulo 5. No Capítulo 6 estão as conclusões do trabalho e propostas de futuras pesquisas.

Capítulo 2

Métodos padrão de análise de dados de degradação

Neste capítulo apresentamos dois métodos de modelagem e análise de dados de degradação apresentados por Meeker e Escobar (1998) e que foram utilizados por Oliveira e Colosimo (2004) e Toledo (2007) em estudos por simulação e por Freitas *et al.* (2009) também em simulações e em análise de dados reais.

O primeiro é na verdade uma abordagem alternativa, aproximada (“Método de Análise Aproximado”) sugerida por Meeker e Escobar (1998). Para problemas mais simples, sobretudo aqueles cuja forma funcional do perfil é linear nos parâmetros, este método tem apelo prático pois os cálculos são relativamente fáceis de serem implementados.

O segundo utiliza o modelo geral de degradação apresentado na Seção 1.3 e a estimação dos parâmetros é feita pelo método de máxima verossimilhança.

2.1 O método de análise aproximado

Este método foi apresentado por Meeker e Escobar (1998) como uma alternativa à modelagem e análise baseada em estimação por máxima verossimilhança. Entretanto,

conforme os próprios autores ressaltam, este é um procedimento aproximado apenas, e que possui uma série de deficiências conforme será discutido ao final desta Seção.

No método aproximado analisamos cada unidade amostral individualmente, para prevermos o tempo em que essa unidade atingirá o nível crítico de degradação (“pseudo” tempo de falha). A partir dos “pseudo” tempos de falha de todas as unidades sob estudo, estimamos a função de distribuição do tempo até a falha $F_T(t)$. Formalmente o método se constitui das seguintes etapas:

- Para a unidade i , usamos o modelo de perfil definido na equação (1.1): $Y_{ij} = D_{ij} + \varepsilon_{ij}$ e os dados amostrais do perfil $(t_{i1}, y_{i1}), \dots, (t_{im_i}, y_{im_i})$ para se encontrar as estimativas de máxima verossimilhança condicional de $\alpha_i = (\alpha_1; \alpha_2; \dots; \alpha_p)'$ e $\beta_i = (\beta_{i1}, \beta_{i2}, \dots, \beta_{ik})'$, digamos $\hat{\alpha}_i$ e $\hat{\beta}_i$. Estas estimativas podem ser obtidas por mínimos quadrados. Note que ao fixarmos i , as estimativas estão sendo obtidas para uma unidade experimental específica. Assim sendo, o vetor $\beta_i = (\beta_{i1}, \beta_{i2}, \dots, \beta_{ik})'$ é tratado como fixo e o vetor $\alpha_i = (\alpha_1; \alpha_2; \dots; \alpha_p)'$ de efeitos fixos, que no modelo (1.1) é comum a todas as unidades, será estimado aqui para cada uma separadamente.
- A partir destas estimativas, solucionamos a equação $D(t, \hat{\alpha}_i, \hat{\beta}_i) = D_f$ para t , obtendo-se a solução \hat{t}_i (denominado “pseudo tempo de falha”).
- Repetimos o procedimento para cada perfil amostral para obter-se $\hat{t}_1, \dots, \hat{t}_n$.
- Aplicamos as técnicas tradicionais de análise de tempos de falha (isto é, ajuste de distribuições tais como Weibull, Lognormal e obtenção de parâmetros via máxima verossimilhança) para os “pseudo” tempos $\hat{t}_1, \dots, \hat{t}_n$; obtendo-se daí a estimativa para $F_T(t)$ (Freitas e Colosimo (1997)).

O método aproximado é simples e intuitivamente atrativo. No entanto, ele só é adequado nos casos em que o perfil de degradação $D(t)$ é relativamente simples; existem

dados de degradação suficientes para estimativas precisas de α_i e β_i ; a variância dos erros (ε_{ij}) é pequena, e finalmente a extrapolação necessária para prever o tempo de falha não é muito grande.

Deficiências do método aproximado foram apontadas por alguns autores tais como Meeker e Escobar (1998) e Oliveira e Colosimo (2004): (1) ele ignora a incerteza na predição dos “pseudo” tempos de falha \hat{t}_i e não considera os erros envolvidos nas observações dos valores de degradação, (2) a distribuição dos “pseudo” tempos de falha geralmente não corresponde àquela que será indicada pelo modelo de degradação e (3) em alguns casos, o volume de dados de degradação coletado pode ser insuficiente para se estimar todos os parâmetros do modelo. Os autores também enfatizam que nestes cenários pode ser necessário ajustar-se diferentes modelos para diferentes unidades para prever os “pseudo” tempos de falha.

2.2 O modelo geral de degradação com estimação de parâmetros pelo método de máxima verossimilhança

Esta abordagem é recomendada principalmente em situações práticas descritas por modelos não lineares, que incluem mais de um efeito aleatório em sua forma funcional.

Retomemos o modelo geral de degradação dado pela equação (1.1):

$$Y_{ij} = D_{ij} = D(t_{ij}; \alpha; \beta) + \varepsilon_{ij} .$$

Neste modelo:

- Y_{ij} é a variável aleatória que representa a medida de degradação para a i -ésima unidade amostral no j -ésimo instante de medição ($i = 1, \dots, n; j = 1, \dots, m_i$);

- $D(t_{ij}; \underset{\sim}{\alpha}; \underset{\sim}{\beta}_i)$ é o perfil real de degradação da unidade i no tempo t_{ij} ;
- $\underset{\sim}{\alpha} = (\alpha_1; \alpha_2; \dots; \alpha_p)'$ é o vetor de efeitos fixos (características populacionais comuns a todas as unidades);
- $\underset{\sim}{\beta}_i = (\beta_{i1}, \beta_{i2}, \dots, \beta_{ik})'$ é o vetor de efeitos aleatórios referentes à i -ésima unidade (isto é, diferem de unidade para unidade);
- ε_{ij} o erro aleatório associado à i -ésima unidade no tempo t_{ij} .

Para o modelo (1.1), as suposições gerais são as seguintes:

1. Os erros aleatórios ε_{ij} são independentes e identicamente distribuídos (i.i.d.) segundo uma distribuição normal com média $\mu = 0$ e variância σ_ε^2 (desconhecida). Ressaltamos entretanto que como as medidas Y_{ij} são tomadas de maneira seriada na mesma unidade, existe a probabilidade de haver autocorrelação entre os valores de ε_{ij} ($i = 1, \dots, n; j = 1, \dots, m_i$), especialmente no caso de medidas tomadas com curto espaçamento de tempo entre si. Entretanto, em muitas situações práticas envolvendo inferência a partir de dados de degradação de unidades de uma população ou processo, esta correlação é fraca. Além disto, ela é dominada pela variabilidade de unidade para unidade nos valores de β_1, \dots, β_k e pode ser ignorada. Em situações nas quais esta autocorrelação não pode ser ignorada, ela deve ser modelada (por exemplo, nos modelos de séries temporais) e incorporadas ao modelo - Meeker e Escobar (1998).
2. Os vetores $\underset{\sim}{\beta}_i = (\beta_{i1}, \beta_{i2}, \dots, \beta_{ik})'$ $i = 1, \dots, n$ são i.i.d. segundo uma distribuição multivariada $\Lambda(\underset{\sim}{\beta} | \underset{\sim}{\theta})$ em que o vetor de parâmetros fixos $\underset{\sim}{\theta} = (\theta_1, \dots, \theta_q)'$ deve ser estimado a partir dos dados de degradação.
3. ε_{ij} e $\underset{\sim}{\beta}_i$ são independentes.

Sejam agora:

- $\underset{\sim}{Y}_i = (Y_{i1}, Y_{i2}, \dots, Y_{im_i})'$ o vetor aleatório representando as medidas de degradação da i -ésima unidade e $\underset{\sim}{y}_i = (y_{i1}, y_{i2}, \dots, y_{im_i})'$ o vetor com os dados de degradação efetivamente observados.
- $\underset{\sim}{Y} = (Y_1, Y_2, \dots, Y_n)'$ o vetor aleatório representando o conjunto completo das medidas de degradação das n unidades sob estudo e $\underset{\sim}{y} = (y_1, y_2, \dots, y_n)'$ o vetor completo com todas as medidas de degradação observadas.
- $\underset{\sim}{\beta}_i = (\beta_{i1}, \beta_{i2}, \dots, \beta_{ik})'$ o vetor representando os efeitos aleatórios da i -ésima unidade e $\underset{\sim}{B} = (\beta_1, \beta_2, \dots, \beta_n)'$ o vetor com todos os efeitos aleatórios de todas as n unidades amostrais.

Se denotarmos a função de densidade de $\underset{\sim}{Y}_i$ como $f(\underset{\sim}{y}_i | \underset{\sim}{\alpha}, \underset{\sim}{\beta}_i, \underset{\sim}{\theta}, \underset{\sim}{\sigma}_\varepsilon^2)$ e a função de densidade de $\underset{\sim}{\beta}_i$ como $f(\underset{\sim}{\beta}_i | \underset{\sim}{\theta})$, usando a suposição de independência de entre os $\underset{\sim}{\beta}_i$ ($i = 1, \dots, n$) e entre $\underset{\sim}{\beta}_i$ e ε_{ij} ($i = 1, \dots, n; j = 1, \dots, m_i$) podemos escrever as funções de densidade de $\underset{\sim}{Y}$ e $\underset{\sim}{B}$ como:

$$f(\underset{\sim}{y} | \underset{\sim}{\alpha}, \underset{\sim}{\beta}_i, \underset{\sim}{\theta}, \underset{\sim}{\sigma}_\varepsilon^2) = \prod_{i=1}^n f(\underset{\sim}{y}_i | \underset{\sim}{\alpha}, \underset{\sim}{\beta}_i, \underset{\sim}{\theta}, \underset{\sim}{\sigma}_\varepsilon^2), \quad (2.1)$$

$$f(\underset{\sim}{B} | \underset{\sim}{\theta}) = \prod_{i=1}^n f(\underset{\sim}{\beta}_i | \underset{\sim}{\theta}). \quad (2.2)$$

Embora os valores de β_1, \dots, β_k para as unidades individuais possam ser de interesse em algumas aplicações (por exemplo, para prever a degradação futura de uma unidade i em particular, baseado em um número pequeno de medidas de degradação já observadas), o desenvolvimento subsequente neste texto será concentrado no uso dos dados

de degradação para inferências à respeito do processo ou população, ou a respeito de unidades futuras.

Neste caso os parâmetros (fixos) do modelo a serem estimados são α , θ e σ_ε^2 e para tanto escrevemos a expressão geral da verossimilhança observada por:

$$\begin{aligned}
 f(\underset{\sim}{y}|\underset{\sim}{\alpha}, \underset{\sim}{\theta}, \underset{\sim}{\sigma_\varepsilon^2}) &= \int_{\Xi_B} f(\underset{\sim}{y}, \underset{\sim}{B}|\underset{\sim}{\alpha}, \underset{\sim}{\theta}, \underset{\sim}{\sigma_\varepsilon^2}) d\underset{\sim}{B} \\
 &= \int_{\Xi_{\beta_1}} \cdots \int_{\Xi_{\beta_n}} f(\underset{\sim}{y}, \underset{\sim}{B}|\underset{\sim}{\alpha}, \underset{\sim}{\theta}, \underset{\sim}{\sigma_\varepsilon^2}) d\underset{\sim}{\beta_1} \cdots d\underset{\sim}{\beta_n} \\
 &= \int_{\Xi_{\beta_1}} \cdots \int_{\Xi_{\beta_n}} \left\{ f(\underset{\sim}{y}|\underset{\sim}{\alpha}, \underset{\sim}{B}, \underset{\sim}{\theta}, \underset{\sim}{\sigma_\varepsilon^2}) f(\underset{\sim}{B}|\underset{\sim}{\theta}) \right\} d\underset{\sim}{\beta_1} \cdots d\underset{\sim}{\beta_n} \\
 &= \int_{\Xi_{\beta_1}} \cdots \int_{\Xi_{\beta_n}} \left\{ \left[\prod_{i=1}^n f(\underset{\sim}{y}_i|\underset{\sim}{\alpha}, \underset{\sim}{\beta}_i, \underset{\sim}{\theta}, \underset{\sim}{\sigma_\varepsilon^2}) \right] \left[\prod_{i=1}^n f(\underset{\sim}{\beta}_i|\underset{\sim}{\theta}) \right] \right\} d\underset{\sim}{\beta_1} \cdots d\underset{\sim}{\beta_n} \\
 &= \int_{\Xi_{\beta_1}} \cdots \int_{\Xi_{\beta_n}} \left\{ \prod_{i=1}^n f(\underset{\sim}{y}_i|\underset{\sim}{\alpha}, \underset{\sim}{\beta}_i, \underset{\sim}{\theta}, \underset{\sim}{\sigma_\varepsilon^2}) f(\underset{\sim}{\beta}_i|\underset{\sim}{\theta}) \right\} d\underset{\sim}{\beta_1} \cdots d\underset{\sim}{\beta_n} \\
 &= \prod_{i=1}^n \left\{ \int_{\Xi_{\beta_i}} f(\underset{\sim}{y}_i|\underset{\sim}{\alpha}, \underset{\sim}{\beta}_i, \underset{\sim}{\theta}, \underset{\sim}{\sigma_\varepsilon^2}) f(\underset{\sim}{\beta}_i|\underset{\sim}{\theta}) d\underset{\sim}{\beta}_i \right\}, \tag{2.3}
 \end{aligned}$$

em que Ξ_B e Ξ_{β_i} representam respectivamente o espaço de variação de $\underset{\sim}{B}$ e de $\underset{\sim}{\beta}_i$.

Usando a suposição do modelo de que $\varepsilon_{ij} \stackrel{\text{i.i.d.}}{\sim} N(0, \sigma_\varepsilon^2)$ concluímos que

$$\underset{\sim}{Y}_i|\underset{\sim}{\alpha}, \underset{\sim}{\beta}_i, \underset{\sim}{\theta}, \underset{\sim}{\sigma_\varepsilon^2} \stackrel{\text{indep.}}{\sim} N_{m_i}(\mu_{Y_i}, \Sigma_{Y_i}),$$

isto é, $\underset{\sim}{Y}_i$ tem distribuição normal multivariada de ordem m_i , com vetor de médias dado

por:

$$\mu_{\tilde{Y}_i} = \mathbb{E}(\tilde{Y}_i) = \left[D(t_{i1}, \alpha, \tilde{\beta}_i); \dots; D(t_{im_i}, \alpha, \tilde{\beta}_i) \right]',$$

e matriz de variância e co-variância igual a

$$\Sigma_{\tilde{Y}_i} = \sigma_\varepsilon^2 I_{m_i}, \quad \text{em que } I_{m_i} \text{ é a matriz identidade de ordem } m_i.$$

Em outras palavras, condicional em $\tilde{\beta}_i$, as variáveis aleatórias Y_{ij} que representam as observações na i -ésima unidade ($j = 1, \dots, m_i$) são independentes e distribuídas segundo uma distribuição normal.

Mais especificamente:

$$Y_{ij} | \alpha, \tilde{\beta}_i, \theta, \sigma_\varepsilon^2 \stackrel{\text{indep.}}{\sim} N \left[D(t_{ij}, \alpha, \tilde{\beta}_i); \sigma_\varepsilon^2 \right]; \quad j = 1, \dots, m_i.$$

E portanto $f(y_i | \alpha, \tilde{\beta}_i, \theta, \sigma_\varepsilon^2) = \prod_{j=1}^{m_i} N[D(t_{ij}, \alpha, \tilde{\beta}_i); \sigma_\varepsilon^2]$. Assim podemos reescrever a função de verossimilhança em (2.3) como:

$$\begin{aligned} f(\tilde{y} | \alpha, \tilde{\theta}, \sigma_\varepsilon^2) &= \prod_{i=1}^n \left\{ \int_{\Xi_{\tilde{\beta}_i}} f(\tilde{y}_i | \alpha, \tilde{\beta}_i, \tilde{\theta}, \sigma_\varepsilon^2) f(\tilde{\beta}_i | \tilde{\theta}) d\tilde{\beta}_i \right\} \\ &= \prod_{i=1}^n \left\{ \int_{\Xi_{\tilde{\beta}_i}} \left[\prod_{j=1}^{m_i} N(D(t_{ij}, \alpha, \tilde{\beta}_i); \sigma_\varepsilon^2) \right] f(\tilde{\beta}_i | \tilde{\theta}) d\tilde{\beta}_i \right\}. \end{aligned} \quad (2.4)$$

A estimação de α , $\tilde{\theta}$ e σ_ε^2 é obtida a partir da maximização da função de verossimilhança (2.4). No entanto a solução analítica das integrais quase nunca é trivial, forçando a utilização de aproximações numéricas. Mesmo estas aproximações numéricas somente são viáveis para modelos simples, com vetores de parâmetros de pequena dimensão.

2.2.1 O modelo baseado na distribuição normal multivariada

Diversos trabalhos surgidos na literatura a partir da década passada consideram que os efeitos aleatórios seguem uma distribuição normal multivariada de ordem p . Neste caso então $\tilde{\theta} = (\mu_{\tilde{\beta}}, \Sigma_{\tilde{\beta}})$ e $\tilde{\beta}_i \stackrel{\text{i.i.d.}}{\sim} N_p(\mu_{\tilde{\beta}}, \Sigma_{\tilde{\beta}})$. Em outras palavras $f(\tilde{\beta}_i|\tilde{\theta})$ em (2.4) é função de densidade de uma normal p -variada com vetor de médias $\mu_{\tilde{\beta}}$ e matriz de variância-covariância $\Sigma_{\tilde{\beta}}$.

Pinheiro e Bates (2000) usaram os resultados desenvolvidos por Lindstrom e Bates (1990) para obter a estimativa aproximada dos parâmetros $\tilde{\alpha}$, $\tilde{\theta} = (\mu_{\tilde{\beta}}, \Sigma_{\tilde{\beta}})$ e σ_ε . As funções LME e NLME escritas na linguagem *S-plus* foram desenvolvidas para obter este resultado. Outra opção é a PROC NLMIXED disponível no *Software SAS*.

Conforme discutido na Seção 1.3.1, após a estimação de $\tilde{\alpha}$, $\mu_{\tilde{\beta}}$, $\Sigma_{\tilde{\beta}}$ e σ_ε^2 , a função de distribuição do tempo de vida, $F_T(t)$, pode ser obtida analiticamente (somente em casos muito simples) ou numericamente por integração direta. Uma vez mais o custo computacional para avaliar a integral multidimensional irá crescer exponencialmente com a dimensão da integral.

Um procedimento alternativo (também já mencionado em 1.3.1) é avaliar $F_T(t)$ numericamente usando simulação de Monte Carlo. A simulação pode ser executada usando as estimativas dos parâmetros $\tilde{\alpha}$, $\mu_{\tilde{\beta}}$, $\Sigma_{\tilde{\beta}}$ e σ_ε^2 que foram fornecidas pelas funções LME ou NLME. Geram-se NB possíveis perfis de degradação $D(t)$ e para cada um deles o “tempo de falha” t_j^* ($j = 1, \dots, NB$) (tempo de transposição ou de forma equivalente o tempo em que o perfil de degradação cruza a linha $y = D_f$) é obtido. Para calcular os valores de $\hat{F}_T(t)$ usa-se então a expressão

$$\hat{F}_T(t) = \frac{\sum_{j=1}^N \mathbb{I}_{(t_j^* \leq t)}}{NB} \quad t > 0, \quad (2.5)$$

em que \mathbb{I} é a função indicadora, t é um instante de tempo fixo e NB deve ser um número bem grande (usualmente $\text{NB} \geq 10^5$).

Para simular os NB perfis de $D(t)$ é necessário gerar NB possíveis realizações do vetor $\Theta' = (\alpha; \beta) = (\alpha_1, \dots, \alpha_p; \beta_1, \dots, \beta_k)$ de uma distribuição normal multivariada com média $\hat{\mu}_\Theta$ e matriz de variância e covariância $\hat{\Sigma}_\Theta$. Note que como α é um vetor de efeitos

fixos, $\mu_\Theta = (\alpha_1, \dots, \alpha_p; \mu_{\beta_1}, \dots, \mu_{\beta_k})'$ e $\Sigma_\Theta = \begin{bmatrix} A & B \\ B^t & \Sigma_\beta \end{bmatrix}$; em que $A = Cov(\alpha)$, $B = Cov(\alpha, \beta)$, e $\Sigma_\beta = Cov(\beta)$. A e B são matrizes com elementos nulos.

O último passo consiste em se aplicar (2.5). Um algoritmo mostrando todas as etapas das sequências de estimações para o método numérico foi apresentado por Yacout *et al.* (1996). Intervalos de confiança podem ser obtidos usando métodos de reamostragem, como o Bootstrap apresentado em Efron (1985).

Capítulo 3

Métodos computacionais para obtenção de estimativas de máxima verossimilhança em modelos de degradação com efeitos aleatórios não normais

Neste capítulo apresentamos métodos computacionais que foram concebidos para possibilitar inferências em modelos de degradação em que a distribuição dos efeitos aleatórios não é normal.

Detalhamos os métodos de Nelson *et al.* (2006) (Seção 3.1) e de Liu e Yu (2008) (Seção 3.2). Finalmente na Seção 3.3 apresentamos a solução implementada nesta dissertação.

3.1 O método “PIT” (probability integral transformation)

Nelson *et al.* (2006) apresentaram um método computacional simples para obter estimativas de máxima verossimilhança em modelos não lineares de efeitos mistos, quando os efeitos aleatórios possuem distribuição não normal. Retornando à expressão da verossimilhança (2.3), e considerando primeiramente um modelo em que os efeitos aleatórios são unidimensionais (isto é, $\beta_i = \beta_i$ escalar, $i = 1, \dots, n$), vemos que a contribuição do i -ésimo perfil para a verossimilhança marginal é:

$$\int_{\Xi_{\beta_i}} \left\{ \prod_{j=1}^{m_i} f(y_{ij} | \alpha, \beta_i, \theta, \sigma_\varepsilon^2) f(\beta_i | \theta) \right\} d\beta_i .$$

No caso em que os efeitos aleatórios possuem distribuição $f(\beta_i | \theta)$ não normal, os autores propuseram a transformação descrita a seguir:

Suponha que a_i seja um efeito aleatório com distribuição normal padrão, isto é $a_i \sim N(0; 1)$. Um resultado fundamental da teoria de probabilidade, afirma que se X é uma variável aleatória contínua com função de distribuição F e densidade f , então a variável aleatória $Y = F(X)$ é uniformemente distribuída em $(0, 1)$. Como consequência, se Y é uma variável aleatória contínua uniformemente distribuída em $(0, 1)$ então a variável aleatória $F^{-1}(Y)$ tem função de distribuição acumulada F ou densidade f , como mostrado por Hoel *et al.* (1971). Assim, tomando $u_i = \Phi(a_i) \sim \text{uniforme}(0, 1)$, em que $\Phi(\cdot)$ é a função de distribuição acumulada (CDF) da distribuição normal padrão e aplicando novamente a mesma transformação, $F_\theta(\beta_i | \theta)$, a função de distribuição acumulada de $\beta_i | \theta$ também tem distribuição uniforme em $(0, 1)$. Segue daí que $F_\theta^{-1}(u_i)$, a CDF inversa de $\beta_i | \theta$, tem densidade $f(\beta_i | \theta)$. Portanto $\beta_i = F_\theta^{-1}(\Phi(a_i))$ tem a distribuição não normal de interesse.

Desta forma, em termos do efeito aleatório normal a_i e usando a teoria de probabilidade para transformações, os autores reescrevem a contribuição da i -ésima unidade (perfil) para a verossimilhança como sendo:

$$\begin{aligned} f(\underset{\sim}{y}_i|\alpha, \theta, \sigma_\varepsilon^2) &= \int_{\Xi_{\beta_i}} \left\{ \prod_{j=1}^{m_i} f(y_{ij}|\alpha, \beta_i, \theta, \sigma_\varepsilon^2) f(\beta_i|\theta) \right\} d\beta_i \\ &= \int_{\Xi_{a_i}} \left\{ \prod_{j=1}^{m_i} f(y_{ij}|\alpha, F_\theta^{-1}(\Phi(a_i)), \theta, \sigma_\varepsilon^2) \phi(a_i) \right\} da_i, \end{aligned} \quad (3.1)$$

em que $\Phi(\cdot)$ e ϕ são a função de distribuição acumulada e a função de densidade da normal padrão, respectivamente.

Métodos numéricos de integração podem agora ser usados para aproximar a solução da função de densidade (3.1) e outros modelos NLME quando não existe solução de forma fechada. Os autores utilizam a Quadratura Gaussiana que aproxima (3.1) pela média ponderada do integrando avaliado em um número, Q , de abcissas predeterminadas (pontos de quadratura) $d_q (q = 1, \dots, Q)$ sobre os efeitos aleatórios a_i , como mostrado em Davidian e Gallant (1992), Liu e Pierce (1994), Pinheiro e Bates (1995) e Lesaffre e Spiessens (2001). Assim a aproximação sugerida é:

$$\begin{aligned} f(\underset{\sim}{y}_i|\alpha, \theta, \sigma_\varepsilon^2) &= \int_{\Xi_{a_i}} \left\{ \prod_{j=1}^{m_i} f(y_{ij}|\alpha, F_\theta^{-1}(\Phi(a_i)), \theta, \sigma_\varepsilon^2) \phi(a_i) \right\} da_i \\ &\approx \sum_{q=1}^Q \prod_{j=1}^{m_i} f(y_{ij}|\alpha, F_\theta^{-1}(\Phi(d_q)), \theta, \sigma_\varepsilon^2) \phi(d_q) w_q, \end{aligned} \quad (3.2)$$

em que $d_q = \sqrt{2}z_q$, $w_q = \sqrt{2}\eta_q e^{-z_q^2}$ e os pesos padrão de Gauss-Hermite η_q e as abcissas z_q podem ser obtidos de tabelas, como em Abramowitz e Stegun (1972).

Portanto a função de verossimilhança a ser maximizada é:

$$f_{\sim}(y|\alpha, \theta, \sigma_{\varepsilon}^2) = \prod_{i=1}^n \sum_{q=1}^Q \prod_{j=1}^{m_i} f(y_{ij}|\alpha, F_{\theta}^{-1}(\Phi(d_q)), \theta, \sigma_{\varepsilon}^2) \phi(d_q) w_q. \quad (3.3)$$

Os autores apresentam dois exemplos de implementação de seu método utilizando a rotina PROC NLMIXED do *software SAS*. Afirmam também ser possível estender a transformação para os modelos com dois ou mais efeitos aleatórios independentes não normais, ou seja, os casos em que $\beta_i = (\beta_{i,1}, \beta_{i,2}, \dots)$.

3.2 Método de reformulação da verossimilhança

Liu e Yu (2008) partiram do artigo de Nelson *et al.* (2006) para apresentar um novo método para tratar o mesmo problema. Os autores argumentaram que o método “PIT” só consegue lidar com efeitos aleatórios para os quais a inversa da função distribuição acumulada (CDF) tem forma fechada, o que limita a sua aplicação. O método pode ser aplicado também a modelos muito limitados de efeitos aleatórios correlacionados não normais (exemplo: modelo Gaussiano de cópula). Entretanto o método “PIT” não é aplicável quando $F^{-1}(\cdot)$ não pode ser obtida facilmente. Por exemplo, para grande parte dos casos de efeitos aleatórios não normais bivariados, a função de distribuição acumulada (CDF) e a inversa da CDF em geral não têm forma fechada e tampouco estão disponíveis como funções em softwares estatísticos como o *SAS* e *R-11*.

Liu e Yu (2008) argumentam ainda que embora Nelson *et al.* (2006) tenham afirmado que uma versão multivariada do “PIT” (Genest e Rivest (2001)) possa ser utilizada com efeitos aleatórios correlacionados, o tratamento matemático altamente complicado pode inviabilizar a sua implementação.

Assim sendo, os autores propuseram uma nova abordagem, reformulando a verossi-

milhança condicionada a efeitos aleatórios não normais em uma verossimilhança que seja condicionada a efeitos aleatórios normais. A proposta consistiu em multiplicar e dividir o integrando de (2.3) pela densidade da normal padrão $\phi(\cdot)$, obtendo:

$$f(\underset{\sim}{y}_i | \underset{\sim}{\alpha}, \underset{\sim}{\theta}, \underset{\sim}{\sigma}_\varepsilon^2) = \int_{\Xi_{\underset{\sim}{\beta}_i}} \left\{ \prod_{j=1}^{m_i} f(y_{ij} | \underset{\sim}{\alpha}, \underset{\sim}{\beta}_i, \underset{\sim}{\theta}, \underset{\sim}{\sigma}_\varepsilon^2) \frac{f(\underset{\sim}{\beta}_i | \underset{\sim}{\theta})}{\underset{\sim}{\phi}(\underset{\sim}{\beta}_i)} \right\} \underset{\sim}{\phi}(\underset{\sim}{\beta}_i) d\underset{\sim}{\beta}_i. \quad (3.4)$$

Os autores então usam a técnica da quadratura Gaussiana no pacote NLMIXED do *software SAS* para aproximar a integração de

$$\left\{ \prod_{j=1}^{m_i} f(y_{ij} | \underset{\sim}{\alpha}, \underset{\sim}{\beta}_i, \underset{\sim}{\theta}, \underset{\sim}{\sigma}_\varepsilon^2) \frac{f(\underset{\sim}{\beta}_i | \underset{\sim}{\theta})}{\underset{\sim}{\phi}(\underset{\sim}{\beta}_i)} \right\}. \quad (3.5)$$

Segundo os autores este método somente requer que a função de densidade (ao invés da inversa da CDF) dos efeitos aleatórios não normais possua uma forma fechada e que esteja disponível no pacote escolhido. Além disso, argumentam que o custo computacional é menor quando comparado com Nelson *et al.* (2006).

3.3 Maximização da verossimilhança - solução numérica baseada no PIT

Nesta seção apresentamos uma proposta de solução numérica para estimação do tempo médio de vida e demais figuras de mérito de interesse para os modelos de degradação simples como os descritos nas situações práticas motivadoras da Seção 1.1, baseada no método apresentado por Nelson *et al.* (2006).

Se observarmos os gráficos dos dados de degradação das rodas dos TUE (Figura 1.3) ou

dos Lasers (Figura 1.4), vemos que os perfis são linhas aproximadamente retas passando pela origem, sugerindo a utilização de um modelo linear simples (uma reta) sem intercepto. Estes modelos serão desenvolvidos nas Seções 3.3.1 (modelo com efeitos aleatórios Weibull) e 3.3.2 (modelo com efeitos aleatórios com distribuição Lognormal).

3.3.1 Efeitos Aleatórios com distribuição de Weibull

Uma das distribuições utilizadas para os efeitos aleatórios tanto por Freitas *et al.* (2009) para os dados de rodas de trens, quanto por Hamada (2005) para os dados dos lasers, é a distribuição de Weibull. Neste caso ambos os autores utilizaram o seguinte modelo linear simples:

$$Y_{ij} = \frac{1}{\beta_i} t_{ij} + \varepsilon_{ij}, \quad i = 1, \dots, n \text{ e } j = 1, \dots, m_i, \quad (3.6)$$

em que

- Y_{ij} é a j -ésima medida de degradação (desgaste das rodas dos trens ou o aumento percentual na corrente de operação no caso dos lasers), tomada no tempo t_{ij} na i -ésima unidade;
- $\beta_i (i = 1, \dots, n)$ é o recíproco da taxa de degradação da i -ésima unidade, correspondendo ao efeito aleatório. Desta forma, β_i indica o quanto e como o perfil médio de degradação específico da unidade i ($\mathbb{E}(Y_{ij}|\beta_i)$) se difere do perfil médio populacional ($\mathbb{E}(Y_{ij})$);
- ε_{ij} é o erro aleatório associado à j -ésima medida tomada na i -ésima unidade (no tempo t_{ij}). Os ε_{ij} 's ($i = 1, \dots, n$ e $j = 1, \dots, m_i$) estão associados à variação em cada tempo t_{ij} das respostas Y_{ij} acima ou abaixo do perfil médio específico da unidade i .

Supomos ainda que:

- Os ε_{ij} 's ($i = 1, \dots, n$ e $j = 1, \dots, m_i$) são i.i.d. $N(0, \sigma_\varepsilon^2)$;
- Os β_i 's e ε_{ij} 's são independentes;
- Os β_i 's são i.i.d. segundo uma distribuição Weibull (α, δ) (α é o parâmetro de forma e δ o parâmetro de escala) com função densidade dada por:

$$f(\beta|\alpha, \delta) = \frac{\alpha}{\delta} \left(\frac{\beta}{\delta}\right)^{\alpha-1} \exp\left\{-\left(\frac{\beta}{\delta}\right)^\alpha\right\} \quad (\alpha > 0, \delta > 0, \beta > 0). \quad (3.7)$$

A forma funcional (3.6) assumida para o perfil de degradação permite uma solução analítica para $F_T(t)$ e conseqüentemente para as figuras de mérito de interesse tais como os quantis, MTTF, etc.

De fato, se o verdadeiro perfil de degradação é dado por $Y = \frac{1}{\beta} t$, o nível crítico de degradação D_f será atingido no tempo de falha T , em outras palavras:

$$D_f = \frac{1}{\beta} T \implies T = D_f \beta \implies \beta^{-1}(T) = \frac{T}{D_f}.$$

Portanto a função de densidade de T será dada por:

$$\begin{aligned} f_T(t) &= f_\beta(\beta^{-1}) \frac{\partial \beta^{-1}(t)}{\partial t} \\ &= \frac{\alpha}{\delta} \left(\frac{t}{D_f \delta}\right)^{\alpha-1} \exp\left\{-\left(\frac{t}{D_f \delta}\right)^\alpha\right\} \frac{1}{D_f} \\ &= \frac{\alpha}{(D_f \delta)} \left(\frac{t}{D_f \delta}\right)^{\alpha-1} \exp\left\{-\left(\frac{t}{D_f \delta}\right)^\alpha\right\}. \end{aligned}$$

Ou seja, T tem distribuição de Weibull com parâmetro de forma igual a α e parâmetro de escala igual a $D_f \delta$.

Se o modelo adotado é o (3.6), então sendo $\theta = (\alpha, \delta)'$:

$$Y_i | \beta_i, \theta, \sigma_\varepsilon^2 \stackrel{\text{indep.}}{\sim} N(\mu_i, \Sigma_{Y_i}) \text{ em que } \mu_i = \left(\frac{1}{\beta_i} t_{i1}, \dots, \frac{1}{\beta_i} t_{im_i} \right)^t \text{ e } \Sigma_{Y_i} = \sigma_\varepsilon^2 I_{m_i} .$$

Assim a verossimilhança (2.4) toma a forma:

$$\begin{aligned} f(\underset{\sim}{y} | \underset{\sim}{\theta}, \sigma_\varepsilon^2) &= \prod_{i=1}^n \left\{ \int_{\Xi_{\beta_i}} \left[\prod_{j=1}^{m_i} \frac{\exp \left[-\frac{(y_{ij} - t_{ij}/\beta_i)^2}{2\sigma_\varepsilon^2} \right]}{\sqrt{2\pi}\sigma_\varepsilon} \right] f(\beta_i | \underset{\sim}{\theta}) d\beta_i \right\} \\ &= \prod_{i=1}^n \left\{ \int_{\Xi_{\beta_i}} \left[\prod_{j=1}^{m_i} \frac{\exp \left(-\frac{(y_{ij} - t_{ij}/\beta_i)^2}{2\sigma_\varepsilon^2} \right)}{\sqrt{2\pi}\sigma_\varepsilon} \right] \frac{\alpha}{\delta} \left(\frac{\beta_i}{\delta} \right)^{\alpha-1} \exp \left[-\left(\frac{\beta_i}{\delta} \right)^\alpha \right] d\beta_i \right\} . \end{aligned} \quad (3.8)$$

Para encontrarmos as estimativas de máxima verossimilhança $\hat{\alpha}_{MV}$, $\hat{\delta}_{MV}$ e $\hat{\sigma}_{\varepsilon MV}^2$ de α , δ e σ_ε^2 respectivamente, precisamos maximizar (3.8) ou de maneira equivalente, maximizar

$$\log f(\underset{\sim}{y} | \underset{\sim}{\theta}, \sigma_\varepsilon^2) = \sum_{i=1}^n \log \left\{ \int_{\Xi_{\beta_i}} \left[\prod_{j=1}^{m_i} \frac{\exp \left(-\frac{(y_{ij} - t_{ij}/\beta_i)^2}{2\sigma_\varepsilon^2} \right)}{\sqrt{2\pi}\sigma_\varepsilon} \right] \frac{\alpha}{\delta} \left(\frac{\beta_i}{\delta} \right)^{\alpha-1} \exp \left[-\left(\frac{\beta_i}{\delta} \right)^\alpha \right] d\beta_i \right\} \quad (3.9)$$

Para isto seguimos os seguintes passos:

Passo 1 Fornecemos valores iniciais para α , δ e σ_ε^2 , digamos $\hat{\alpha}^{(0)}$, $\hat{\delta}^{(0)}$ e $\hat{\sigma}_{\varepsilon}^{2(0)}$;

Passo 2 Substituímos estes valores em (3.8);

Passo 3 Resolvemos as integrais em $\beta_i (i = 1, \dots, n)$;

Passo 4 Encontramos $\log f(\underset{\sim}{y} | \hat{\alpha}^{(0)}, \hat{\delta}^{(0)}, \hat{\sigma}_{\varepsilon}^{2(0)})$;

Passo 5 Atualizamos os valores de $\hat{\alpha}$, $\hat{\delta}$ e $\hat{\sigma}_{\varepsilon}^2$;

Passo 6 Repetimos os passos 2 até 4 até encontrar os valores de α , δ e σ_ε^2 que maximizem

(3.8), ou seja, até encontrar $\hat{\alpha}_{MV}$, $\hat{\delta}_{MV}$ e $\hat{\sigma}_{\varepsilon MV}^2$ que são as estimativas dos parâmetros pelo método de máxima verossimilhança.

Estes passos são detalhados a seguir:

Passo 1: Determinação dos valores iniciais de α , δ e σ_ε^2 :

Os valores $\hat{\alpha}^{(0)}$, $\hat{\delta}^{(0)}$ e $\hat{\sigma}_\varepsilon^{2(0)}$ foram obtidos da seguinte forma:

1. Ajustamos para cada perfil individualmente o modelo (3.6), obtendo assim as estimativas $\hat{\beta}_1, \dots, \hat{\beta}_n$ de β_1, \dots, β_n (por mínimos quadrados).
2. Utilizando $\hat{\beta}_1, \dots, \hat{\beta}_n$ como se fossem observações de uma variável aleatória Weibull (α, δ) , encontramos as estimativas de máxima verossimilhança para α e δ . Estas estimativas serão os valores iniciais $\hat{\alpha}^{(0)}$ e $\hat{\delta}^{(0)}$.
3. Valor inicial para σ_ε^2 : De posse das estimativas $\hat{\beta}_1, \dots, \hat{\beta}_n$ obtidas em 1, calculamos os valores ajustados \hat{Y}_{ij} , dados por

$$\hat{Y}_{ij} = \frac{1}{\hat{\beta}_i} t_{ij} \quad i = 1, \dots, n \text{ e } j = 1, \dots, m_i,$$

e conseqüentemente $\hat{\varepsilon}_{ij} = Y_{ij} - \hat{Y}_{ij}$, $i = 1, \dots, n$ e $j = 1, \dots, m_i$. O valor inicial $\hat{\sigma}_\varepsilon^{2(0)}$, foi então calculado como:

$$\hat{\sigma}_\varepsilon^{2(0)} = \frac{\sum_{i=1}^n \sum_{j=1}^{m_i} (\hat{\varepsilon}_{ij} - \bar{\varepsilon})^2}{N - 1},$$

em que $N = \sum_{i=1}^n m_i$ e $\bar{\varepsilon} = \frac{\sum_{i=1}^n \sum_{j=1}^{m_i} \hat{\varepsilon}_{ij}}{N}$.

Passo 2: Substituindo $\hat{\alpha}^{(0)}$, $\hat{\delta}^{(0)}$ e $\hat{\sigma}_\varepsilon^{2(0)}$ em (3.9) obtemos $\log f(y| \hat{\theta}^{(0)}, \hat{\sigma}_\varepsilon^{2(0)})$

$$\sum_{i=1}^n \log \left\{ \int_{\Xi_{\beta_i}} \left[\prod_{j=1}^{m_i} \frac{\exp \left(-\frac{(y_{ij} - t_{ij}/\beta_i)^2}{2\hat{\sigma}_\varepsilon^{2(0)}} \right)}{\sqrt{2\pi\hat{\sigma}_\varepsilon^{2(0)}}} \right] \frac{\hat{\alpha}^{(0)}}{\hat{\delta}^{(0)}} \left(\frac{\beta_i}{\hat{\delta}^{(0)}} \right)^{\hat{\alpha}^{(0)}-1} \exp \left[-\left(\frac{\beta_i}{\hat{\delta}^{(0)}} \right)^{\hat{\alpha}^{(0)}} \right] d\beta_i \right\}. \quad (3.10)$$

Passo 3: Solução das integrais em (3.10):

Um dos complicadores para resolução das integrais de (3.10) é a determinação de limites de integração apropriados para utilização dos métodos de integração numérica visto que $\Xi_{\beta_i} = (0, \infty)$.

Para contornar esta dificuldade, utilizamos a transformação sugerida por Nelson *et al.* (2006) na equação (3.1). Assim (3.10) pode ser reescrita como:

$$\begin{aligned} \log f(Y|\theta, \sigma_\varepsilon^2) &= \sum_{i=1}^n \log \left\{ \int_{\Xi_{a_i}} \left[\prod_{j=1}^{m_i} \frac{\exp \left(-\frac{[y_{ij} - t_{ij}/F^{-1}(\Phi(a_i))]^2}{2\hat{\sigma}_\varepsilon^{2(0)}} \right)}{\sqrt{2\pi\hat{\sigma}_\varepsilon^{(0)}}} \right] \phi(a_i) da_i \right\} \\ &= \sum_{i=1}^n \log \left\{ \int_{\Xi_{a_i}} \left[\prod_{j=1}^{m_i} \frac{\exp \left(-\frac{[y_{ij} - t_{ij}/\hat{\delta}^{(0)}(-\log(1 - \Phi(a_i))^{1/\hat{\alpha}^{(0)}})]^2}{2\hat{\sigma}_\varepsilon^{2(0)}} \right)}{\sqrt{2\pi\hat{\sigma}_\varepsilon^{(0)}}} \right] \phi(a_i) da_i \right\}, \end{aligned} \quad (3.11)$$

em que $\Phi(a_i)$ e $\phi(a_i)$ são respectivamente a distribuição acumulada e a densidade de uma normal padrão; $F^{-1}(\Phi(a_i)) = \delta(-\log(1 - \Phi(a_i))^{1/\alpha})$ é a forma da inversa da função de distribuição acumulada da Weibull com parâmetros α (forma) e δ (escala), avaliada em $\Phi(a_i)$; e $\Xi_{a_i} = (-\infty, \infty)$.

Como o integrando está multiplicado por $\phi(a_i)$, que é a densidade da normal padrão, somente teremos massa significativa no intervalo $(-3.5 \text{ a } 3.5)$. Podemos portanto considerar $\Xi_{a_i} = (-3,5; 3,5)$.

Na rotina implementada para o *software R.11* e reproduzida no anexo B.1 deste trabalho, utilizamos a soma de Riemann no intervalo $(-3.5 \text{ a } 3.5)$ e $\Delta x = 0,005$ como aproximação da integral em (3.11).

A utilização de quadratura gaussiana sugerida por Nelson *et al.* (2006) não funcionou aqui devido ao perfil bastante irregular da função de verossimilhança nos casos estudados.

Passos 4, 5 e 6: Otimização

Para otimização usamos a rotina ConstrOptim implementada no *software R.11*, que usa um algoritmo de otimização baseado no método de barreira adaptativa. No nosso caso usamos a restrição de que os parâmetros a serem estimados são positivos. Para maiores detalhes veja Lange (2010).

De posse das estimativas de máxima verossimilhança para α , δ e σ_ε^2 podemos calcular analiticamente as figuras de mérito de interesse, já que como vimos, conhecemos a distribuição do tempo até a falha neste modelo.

3.3.2 Efeitos Aleatórios com distribuição Lognormal

Outra distribuição utilizada por Freitas *et al.* (2009) para os efeitos aleatórios no estudo de degradação das rodas de trem, foi a Lognormal, desta vez utilizando o modelo (3.12):

$$Y_{ij} = \beta_i t_{ij} + \varepsilon_{ij}, \quad i = 1, \dots, n \text{ e } j = 1, \dots, m_i . \tag{3.12}$$

Se $\beta \sim \text{Lognormal}(\mu, \sigma)$ em que $\sigma > 0$ é o parâmetro de forma e e^μ é o parâmetro de escala (Meeker e Escobar, 1998), sua função densidade de probabilidade é dada por:

$$f(\beta | \mu, \sigma) = \frac{1}{\beta \sqrt{2\pi} \sigma} \exp \left\{ -\frac{[\log(\beta) - \mu]^2}{2\sigma^2} \right\}, \quad 0 \leq x < \infty .$$

Novamente, se D_f é o nível crítico de degradação e T o tempo para atingir este nível (tempo de falha) podemos escrever:

$$D_f = \beta T \implies T = \frac{D_f}{\beta} \implies \beta^{-1}(T) = \frac{D_f}{T} .$$

E a função de densidade de T será dada por:

$$\begin{aligned} f_T(t) &= f_\beta(\beta^{-1}) \frac{\partial \beta^{-1}(t)}{\partial t} \\ &= \frac{t}{D_f \sqrt{2\pi\sigma}} \exp \left\{ -\frac{[\log(D_f/t) - \mu]^2}{2\sigma^2} \right\} \frac{D_f}{t^2} \\ &= \frac{1}{t\sqrt{2\pi\sigma}} \exp \left\{ -\frac{[\log(t) - (\log(D_f) - \mu)]^2}{2\sigma^2} \right\}. \end{aligned}$$

Portanto agora T tem distribuição Lognormal como os efeitos aleatórios, mas com parâmetro de escala igual a $\exp[\log(D_f) - \mu]$ e parâmetro de forma igual a σ .

A verossimilhança (2.4) toma a forma:

$$f(\underline{y}|\underline{\theta}, \sigma_\varepsilon^2) = \prod_{i=1}^n \left\{ \int_{\Xi_{\beta_i}} \left[\prod_{j=1}^{m_i} \frac{\exp \left(-\frac{(y_{ij} - \beta_i t_{ij})^2}{2\sigma_\varepsilon^2} \right)}{\sqrt{2\pi\sigma_\varepsilon}} \right] \frac{1}{\beta_i \sqrt{2\pi\sigma}} \exp \left\{ -\frac{[\log(\beta_i) - \mu]^2}{2\sigma^2} \right\} d\beta_i \right\}. \quad (3.13)$$

Neste modelo os parâmetros a serem estimados são μ , σ e σ_ε^2 . Portanto é preciso maximizar (3.13) ou, de maneira equivalente:

$$\log f(\underline{y}|\underline{\theta}, \sigma_\varepsilon^2) = \sum_{i=1}^n \log \left\{ \int_{\Xi_{\beta_i}} \left[\prod_{j=1}^{m_i} \frac{\exp \left(-\frac{(y_{ij} - \beta_i t_{ij})^2}{2\sigma_\varepsilon^2} \right)}{\sqrt{2\pi\sigma_\varepsilon}} \right] \frac{1}{\beta_i \sqrt{2\pi\sigma}} \exp \left\{ -\frac{[\log(\beta_i) - \mu]^2}{2\sigma^2} \right\} d\beta_i \right\}. \quad (3.14)$$

Os passos para se encontrar as estimativas de máxima verossimilhança para os parâmetros são similares aos descritos na Seção 3.3.1 e são apresentados a seguir de forma sucinta:

Passo 1: Determinação dos valores iniciais de μ , σ e σ_ε^2 :

Ajustamos (3.12) para cada perfil individualmente e obtivemos $\hat{\beta}_1, \dots, \hat{\beta}_n$. Considerando-

se estas estimativas como observações de uma variável aleatória Lognormal (μ, σ) , encontramos os estimadores de máxima verossimilhança $\hat{\mu}^{(0)}$ e $\hat{\sigma}^{(0)}$. Em seguida, a partir de $\hat{\beta}_1, \dots, \hat{\beta}_n$ calculamos \hat{Y}_{ij} e $\hat{\varepsilon}_{ij}$ que permitiram a obtenção de $\hat{\sigma}_\varepsilon^{2(0)}$ conforme já descrito.

Passo 2: Substituindo $\hat{\mu}^{(0)}$, $\hat{\sigma}^{(0)}$ e $\hat{\sigma}_\varepsilon^{2(0)}$ em (3.14) obtemos $\log f(y|\theta, \sigma_\varepsilon^2)$:

$$\sum_{i=1}^n \log \left\{ \int_{\Xi_{\beta_i}} \left[\prod_{j=1}^{m_i} \frac{\exp \left(-\frac{(y_{ij} - \beta_i t_{ij})^2}{2\sigma_\varepsilon^{2(0)}} \right)}{\sqrt{2\pi}\sigma_\varepsilon^{(0)}} \right] \frac{1}{\beta_i \sqrt{2\pi}\sigma^{(0)}} \exp \left\{ -\frac{[\log(\beta_i) - \mu^{(0)}]^2}{2\sigma^{2(0)}} \right\} d\beta_i \right\}. \quad (3.15)$$

Passo 3: Solução das integrais em (3.15):

Utilizando a mesma transformação sugerida por Nelson *et al.* (2006), (3.15) é reescrita neste caso como:

$$\begin{aligned} \log f(y|\theta, \sigma_\varepsilon^2) &= \sum_{i=1}^n \left\{ \int_{\Xi_{a_i}} \left[\prod_{j=1}^{m_i} \frac{\exp \left(-\frac{[y_{ij} - F^{-1}(\Phi(a_i))t_{ij}]^2}{2\hat{\sigma}_\varepsilon^{2(0)}} \right)}{\sqrt{2\pi}\hat{\sigma}_\varepsilon^{(0)}} \right] \phi(a_i) da_i \right\} \\ &= \sum_{i=1}^n \left\{ \int_{\Xi_{a_i}} \left[\prod_{j=1}^{m_i} \frac{\exp \left(-\frac{[y_{ij} - Qlnorm(\Phi(a_i), \hat{\mu}^{(0)}, \hat{\sigma}^{(0)})t_{ij}]^2}{2\hat{\sigma}_\varepsilon^{2(0)}} \right)}{\sqrt{2\pi}\hat{\sigma}_\varepsilon^{(0)}} \right] \phi(a_i) da_i \right\}, \end{aligned} \quad (3.16)$$

em que $\Phi(a_i)$ e $\phi(a_i)$ são respectivamente a distribuição acumulada e a densidade de a_i (que tem distribuição normal padrão); $F^{-1}(\Phi(a_i)) = Qlnorm(\Phi(a_i), \hat{\mu}^{(0)}, \hat{\sigma}^{(0)})$ é a inversa da função de distribuição acumulada da Lognormal, chamada função quantílica por alguns autores, com parâmetros $\hat{\mu}^{(0)}$ e $\hat{\sigma}^{(0)}$, aplicada em $\Phi(a_i)$; e Ξ_{a_i} é o suporte do integrando. Novamente somente teremos massa significativa no intervalo $(-3.5 a 3.5)$, e consideramos este intervalo como o suporte de a_i .

Passos 4, 5 e 6: Otimização

As estimativas de máxima verossimilhança para μ , σ e σ_ε^2 são obtidas como descrito na Seção 3.3.1 e a partir delas calculamos analiticamente as figuras de mérito de interesse, já que neste modelo o tempo até a falha também tem distribuição conhecida. O script implementado para este modelo se encontra no anexo B.2.

Capítulo 4

Os resultados de Toledo (2007) e Freitas *et al.* (2009) revisitados

Toledo (2007) e Freitas *et al.* (2009) apresentaram simulações comparando os métodos aproximado e de máxima verossimilhança normal (denominado de “Método Numérico” pelos autores) propostos em Meeker e Escobar (1998) e a análise de tempo de vida (*FTA - Failure Time Analysis*) tradicional para se estimar a distribuição do tempo até a falha em modelos de degradação lineares sem intercepto. Utilizaram o modelo (3.12) com efeitos aleatórios com distribuição Lognormal e o modelo (3.6) com a distribuição de Weibull. Mostraram que apesar de não haver alteração significativa no desvio padrão das estimativas das figuras de mérito obtidas pelos três métodos, os métodos aproximado e de máxima verossimilhança normal apresentavam estimativas com desvio padrão ligeiramente menor que o daquelas obtidas pelo método de tempo de vida, especialmente para percentuais de censura maiores. No entanto, o vício dos estimadores obtidos pelo método de máxima verossimilhança eram muito grandes, quando comparados com o resultado dos outros dois métodos. Isso se devia ao fato de que as estimativas dos parâmetros da verossimilhança definida na equação (2.4) foram obtidas pelo pacote LME no *Software R* – 11 que, con-

forme relatado na seção 2.2, utiliza pressuposto de normalidade para os efeitos aleatórios do modelo.

Nosso objetivo é repetir exatamente as simulações daqueles trabalhos, e utilizar a solução descrita na Seção 3.3 para obter estimativas das mesmas figuras de mérito, comparando os resultados obtidos com aqueles dos trabalhos citados.

4.1 Simulando efeitos aleatórios com distribuição de Weibull

O ponto de partida para as simulações é a geração de perfis de degradação obedecendo a critérios pré-determinados. De posse destes perfis, se aplicam métodos de inferência para determinação da distribuição do tempo de vida e outras figuras de mérito.

O primeiro modelo utilizado nas simulações foi

$$Y_{ij} = \frac{1}{\beta_i} t_{ij} + \varepsilon_{ij}, \quad i = 1, \dots, 50 \text{ e } j = 1, \dots, 20 ,$$

com $\beta \sim Weibull(\alpha, \delta)$, sendo α o parâmetro de forma e δ o de escala. Os valores dos parâmetros da distribuição de β_i utilizados por Freitas *et al.* (2009) e Toledo (2007), com base em Oliveira e Colosimo (2004), foram $\alpha = 6,230596$ e $\delta = 10.767,69$. A Figura 4.1 ilustra a função de densidade de probabilidade de uma distribuição de Weibull com estes parâmetros.

Cada amostra simulada tem 50 perfis de degradação e cada perfil é composto de 20 medidas igualmente espaçadas, tomadas de 2000 em 2000 (km), no intervalo $[0, 38.000]$. Em outras palavras, $t_{i1} = t_1 = 0$, $t_{i2} = t_2 = 2.000 \text{ km}$, \dots , $t_{i20} = t_{20} = 38.000 \text{ km}$ para todo $i = 1, \dots, 50$. Como vimos na Seção 3.3.1, sob este modelo o tempo de vida tem distribuição de Weibull com parâmetros de forma e escala respectivamente α e

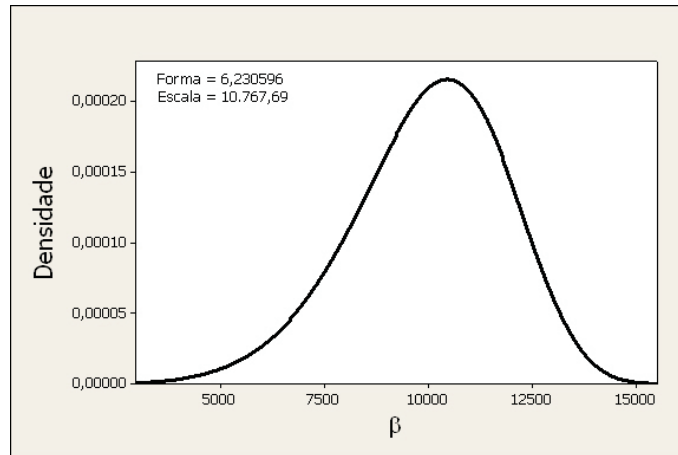


Figura 4.1: Função densidade de probabilidade - Weibull

$D_f\delta$. Podemos então obter diretamente os valores reais da média e do p -ésimo quantil da distribuição do tempo de falha pelas expressões:

$$\mathbb{E}(T) = MTTF = D_f\delta\Gamma\left(1 + \frac{1}{\alpha}\right), \quad (4.1)$$

e

$$t_p = D_f\delta[-\ln(1 - p)]^{1/\alpha}. \quad (4.2)$$

A partir de (4.2) podemos então calcular os valores de D_f que induzam diferentes níveis de censura:

$$D_f = \frac{t_p}{10.767,69(-\ln(1 - p))^{1/6,230596}}. \quad (4.3)$$

O objetivo dos autores era aferir se variações nos percentuais de censura, que sabidamente afetam a qualidade das estimativas na análise tradicional de tempo até a falha, influenciariam também as estimativas obtidas pelos outros métodos.

Seja $t_f = 38.000$ (tempo limite ou final de acompanhamento) e $q = 1 - p$ a fração de

censuras desejada em t_f . Queremos achar o D_f tal que:

$$\mathbb{P}(T \leq t_f) = p \quad \text{ou seja,} \quad \mathbb{P}(T > t_p) = q = 1 - p .$$

Substituindo em (4.3) t_p por $t_f = 38.000$, e $q = 1 - p = 0,10; 0,30; 0,60; 0,80$ teremos os valores de D_f que irão gerar $(1 - p) \times 100\%$ de censuras. Estes valores estão na Tabela 4.1.

Tabela 4.1: Limites de degradação (D_f) por percentual de censuras
Caso 1: modelo linear com efeito aleatório Weibull

Percentual de censura	D_f
10%	3,09
30%	3,43
60%	3,93
80%	4,49

Levando estes valores de D_f em (4.1) e (4.2) podemos obter os verdadeiros valores do MTTF e de quantis de interesse para o modelo simulado, para cada um dos percentuais de censura desejados. A Tabela 4.2 apresenta estes valores.

Tabela 4.2: Figuras de mérito reais
Caso 1: modelo linear com efeito aleatório Weibull

Percentual de censura	MTTF	$t_{0,01}$	$t_{0,05}$	$t_{0,10}$	Mediana	$t_{0,80}$
10%	30.931	15.901	20.656	23.186	31.371	35.913
30%	34.335	17.651	22.929	25.737	34.823	39.865
60%	39.340	20.224	26.271	29.489	39.900	45.676
80%	44.945	23.106	30.015	33.691	45.585	52.184

Para repetirmos as simulações, inicialmente geramos aleatoriamente os valores para

β_i ($i = 1, \dots, 50$), de uma Weibull(6,230596; 10.767,69). Em seguida geramos, de uma distribuição normal com média 0 e desvio padrão 0,01 (mesmos valores utilizados na simulação original), 1000 (50×20) valores para os ε_{ij} . Podemos agora calcular os 50 perfis da amostra usando $Y_{ij} = (1/\beta_i)t_j + \varepsilon_{ij}$. Este procedimento foi repetido 10.000 vezes para cada um dos quatro níveis de censura. Como o valor de D_f determina a ocorrência de falha, as medidas subsequentes àquelas que ultrapassavam os valores de D_f em cada unidade amostral simulada eram abandonadas.

A Figura 4.2 mostra um exemplo de um conjunto de 50 perfis simulados com $D_f = 3,09$. Neste exemplo 6 unidades (aproximadamente 10%) não apresentaram falha (não ultrapassaram o valor limite de degradação D_f) no período simulado.

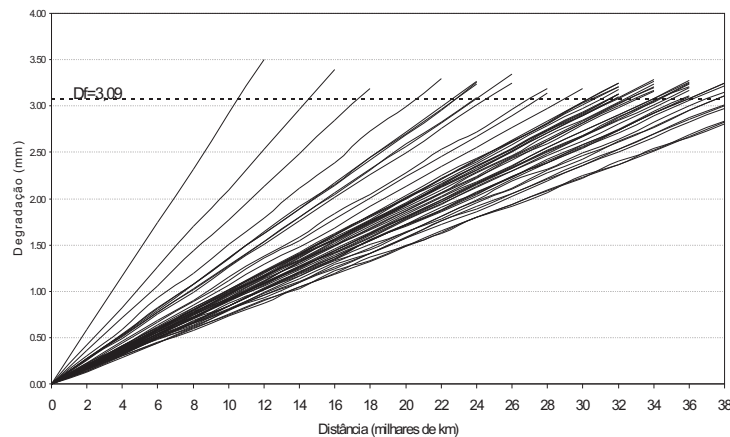


Figura 4.2: Perfis de degradação simulados
Caso 1: modelo linear com efeito aleatório Weibull

Para cada conjunto de perfis simulados obtivemos as estimativas de máxima verossimilhança para α , δ e σ_ε^2 conforme descrito em 3.3.1. A partir destas estimativas calculamos, com (4.1) e (4.2), os 10.000 valores estimados para o MTTF e para os quantis 0,01; 0,05; 0,10; 0,50 (mediana) e 0,80. Utilizamos as mesmas medidas de desempenho mostradas por Toledo (2007) e Freitas *et al.* (2009), comparando os valores médios estimados com os reais (Tabela 4.2), para cada uma das figuras de mérito.

Estas medidas são:

- O desvio padrão, calculado diretamente com as 10.000 estimativas.
- O vício relativo obtido dividindo-se o valor absoluto da diferença entre o valor médio estimado e o valor real, pelo valor real ($\times 100$ para obtermos a percentagem).
- O Erro Quadrático Médio, calculado pela soma da variância e o quadrado do vício.

Todos os resultados são apresentados na tabela A.1, no anexo A, e para ilustração mostramos nas Figuras 4.3, 4.4 e 4.5 o resultado das três medidas de três figuras de mérito: MTTF, quantil 0,01 e Mediana.

MTTF

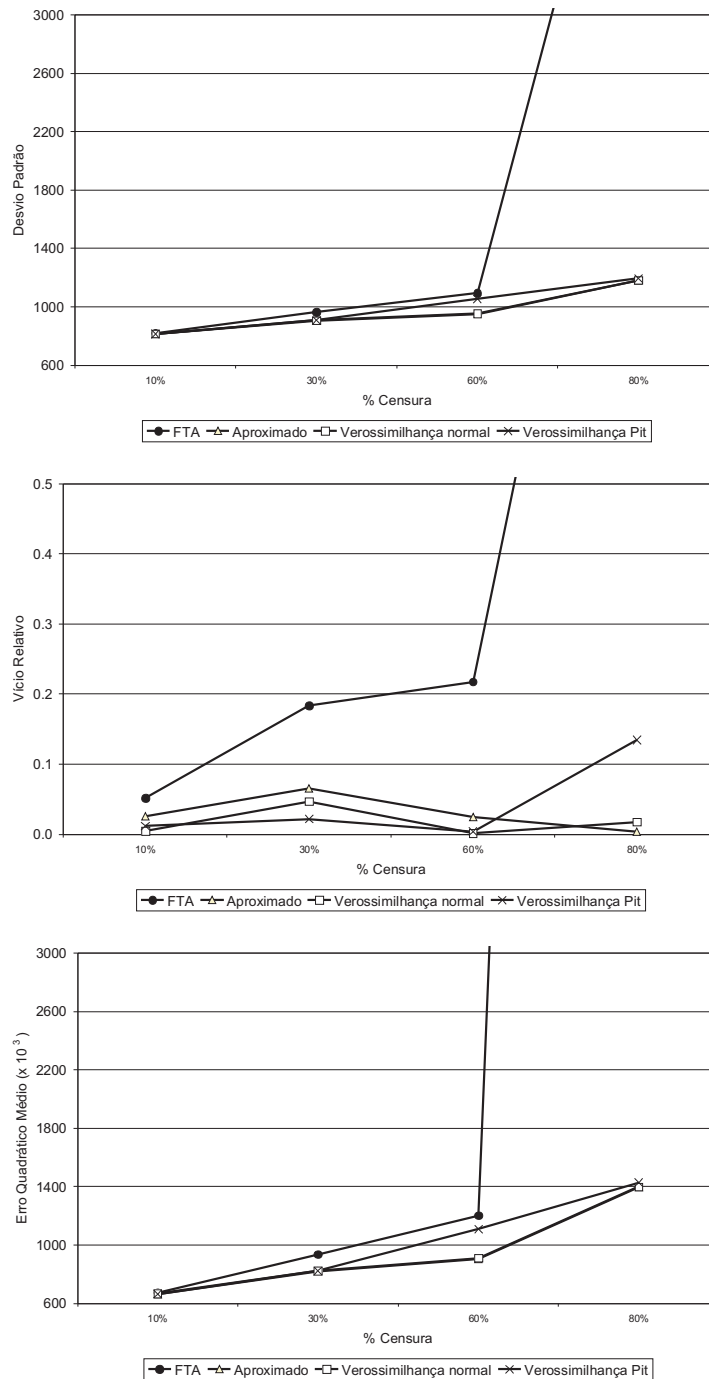


Figura 4.3: Medidas de desempenho estimativas do MTTF
 Caso 1: modelo linear com efeito aleatório Weibull

Quantil 0,01

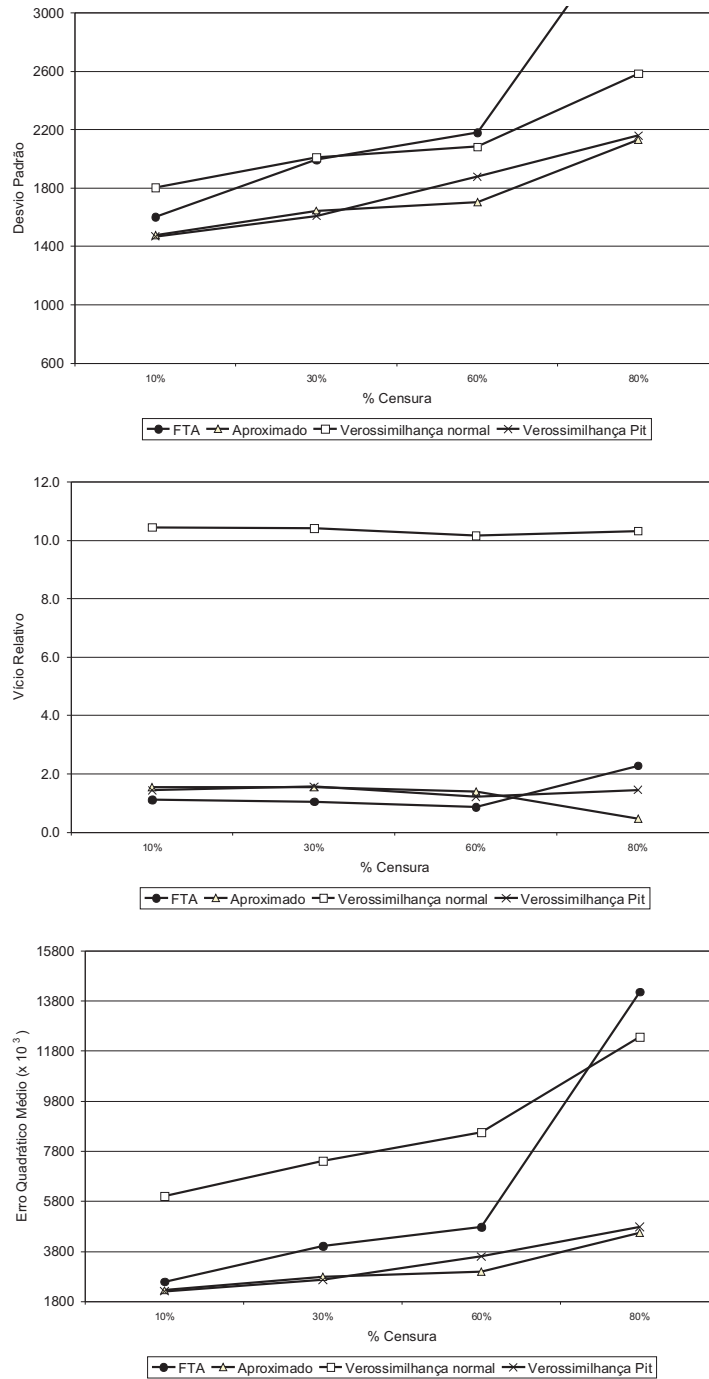


Figura 4.4: Medidas de desempenho estimativas do quantil 0,01
Caso 1: modelo linear com efeito aleatório Weibull

Mediana

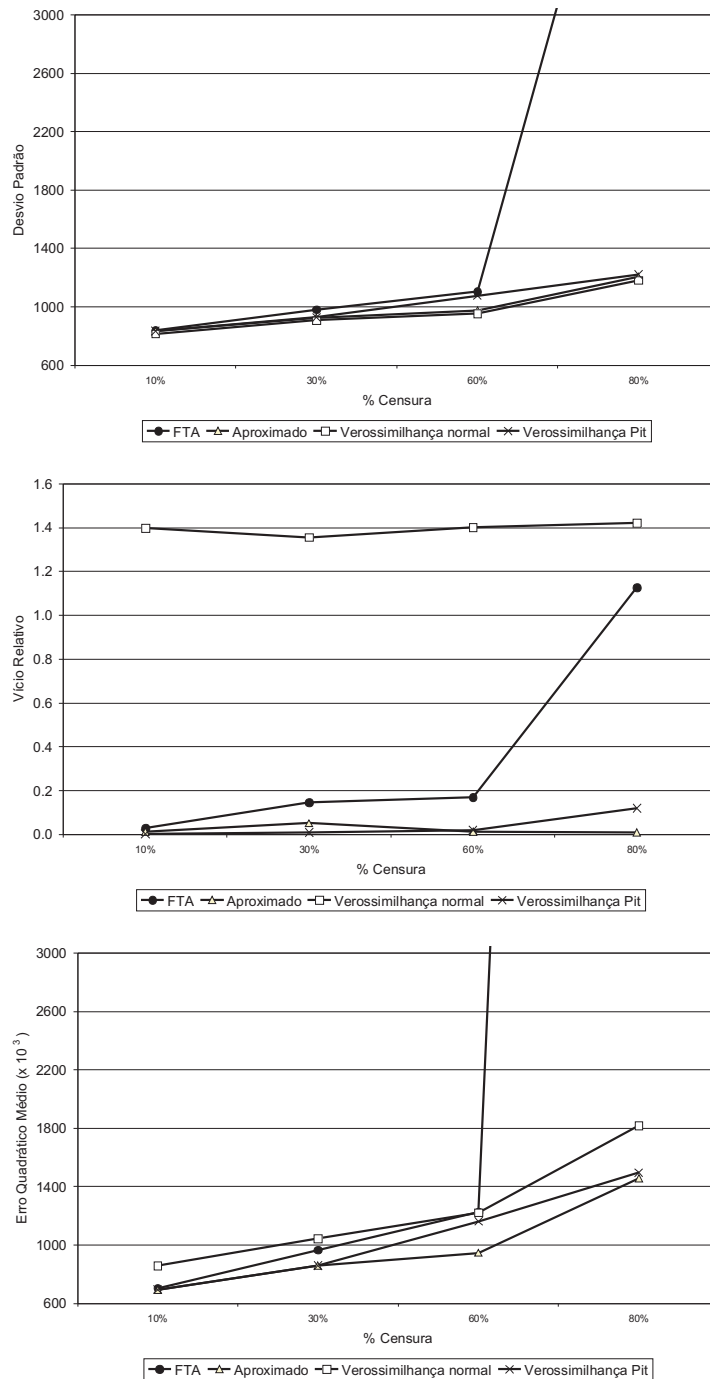


Figura 4.5: Medidas de desempenho estimativas da mediana
 Caso 1: modelo linear com efeito aleatório Weibull

4.2 Simulando efeitos aleatórios com distribuição Lognormal

O outro modelo cuja simulação reproduzimos foi

$$Y_{ij} = \beta_i t_{ij} + \varepsilon_{ij}, \quad i = 1, \dots, n \text{ e } j = 1, \dots, m_i,$$

com $\beta \sim \text{Lognormal}(\mu, \sigma)$. Novamente usamos os valores dos parâmetros dos trabalhos de referência, que foram $\mu = -9,12265$ e $\sigma = 0,235929$. A Figura 4.6 ilustra a função densidade de probabilidade de uma distribuição Lognormal com estes parâmetros.

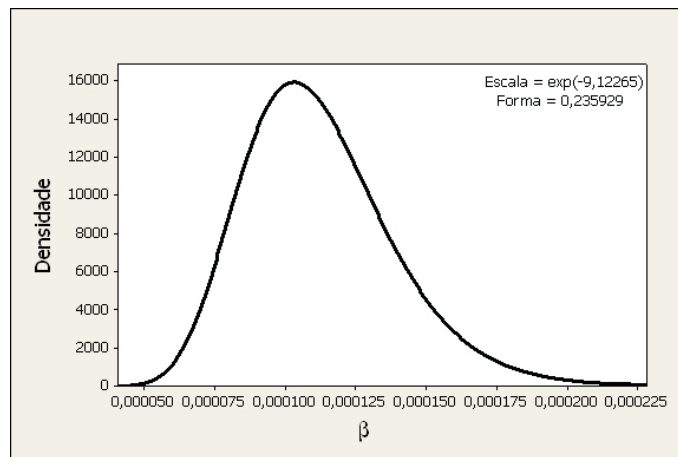


Figura 4.6: Função densidade de probabilidade - Lognormal

A estrutura dos dados é a mesma, isto é, 50 perfis de degradação cada um deles com 20 medidas de 2000 em 2000 (km), no intervalo $[0, 38.000]$. Como vimos na Seção 3.3.2, sob este modelo o tempo de vida tem distribuição Lognormal com parâmetro de escala igual a $\exp[\log(D_f) - \mu]$ e parâmetro de forma igual a σ . Podemos então obter diretamente os valores reais da média e do p -ésimo quantil da distribuição do tempo de falha pelas expressões:

$$\mathbb{E}(T) = MTTF = \exp \left[\ln(D_f) - \mu + \frac{\sigma^2}{2} \right], \quad (4.4)$$

$$\text{e } t_p = \exp[z_p\sigma + \log(D_f) - \mu]. \quad (4.5)$$

O cálculo dos valores de D_f que induzam os níveis de censura desejados pode ser feito a partir de (4.5):

$$D_f = \exp\{\log(t_p) - 0,235929z_p - 9,12265\}.$$

Assim como no caso Weibull, fazemos $t_f = 38.000$ e $q = 1 - p = 0,10; 0,30; 0,60; \text{ e } 0,80$; obtendo D_f para cada percentual de censura (veja Tabela 4.3).

Tabela 4.3: Limites de degradação (D_f) por percentual de censuras
Caso 2: modelo com efeito aleatório Lognormal

Percentual de censura	D_f
10%	3,07
30%	3,67
60%	4,40
80%	5,06

Os valores de D_f fornecem, a partir de (4.4) e (4.5), os verdadeiros valores do MTTF e de quantis de interesse para o modelo simulado. A tabela 4.4 apresenta o resultado.

Tabela 4.4: Figuras de mérito reais para perfis simulados
Caso 2: modelo com efeito aleatório Lognormal

Percentual de censura	MTTF	$t_{0,01}$	$t_{0,05}$	$t_{0,10}$	Mediana	$t_{0,80}$
10%	30.931	15.901	20.656	23.186	31.371	35.913
30%	34.335	17.651	22.929	25.737	34.823	39.865
60%	39.340	20.224	26.271	29.489	39.900	45.676
80%	44.945	23.106	30.015	33.691	45.585	52.184

Repetimos os mesmos procedimentos de simulação, agora gerando inicialmente valores para β_i ($i = 1, \dots, 50$), de uma Lognormal($-9,12265$; $0,235929$). Novamente trabalhamos com 10.000 conjuntos de perfis para cada um dos quatro níveis de censura. A título de ilustração, a Figura 4.7 apresenta um conjunto de perfis simulados com $D_f = 5,06$. Neste exemplo 10 unidades apresentaram falha (ultrapassaram o valor limite de degradação D_f) no período simulado, ou seja, 40 em 50 não “falharam” (80%).

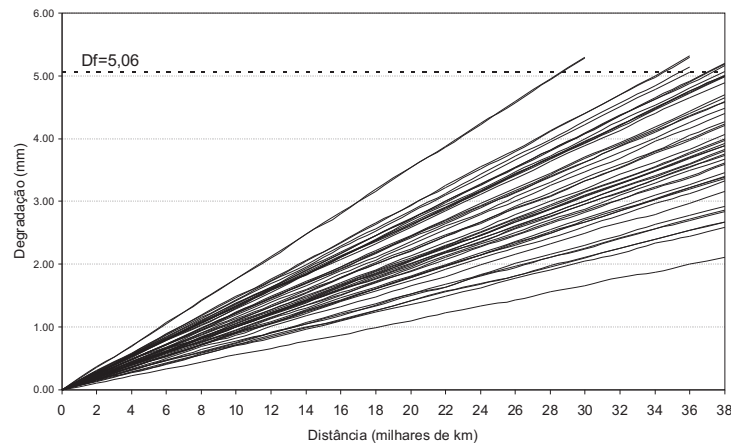


Figura 4.7: Perfis de degradação simulados
Caso 2: modelo linear com efeito aleatório Lognormal

Após a geração dos perfis, obtivemos as estimativas de máxima verossimilhança para μ , σ e σ_ε^2 conforme descrito em 3.3.2 e em seguida calculamos, com (4.4) e (4.5), as estimativas para o MTTF e quantis escolhidos. Mais uma vez utilizamos as mesmas medidas de desempenho. Ilustramos nas figuras 4.8, 4.9 e 4.10 o resultado do desvio padrão, vício relativo e erro quadrático médio para três figuras de mérito: MTTF, quantil 0,01 e Mediana. Todos os resultados podem ser consultados no anexo A, na Tabela A.2.

MTTF

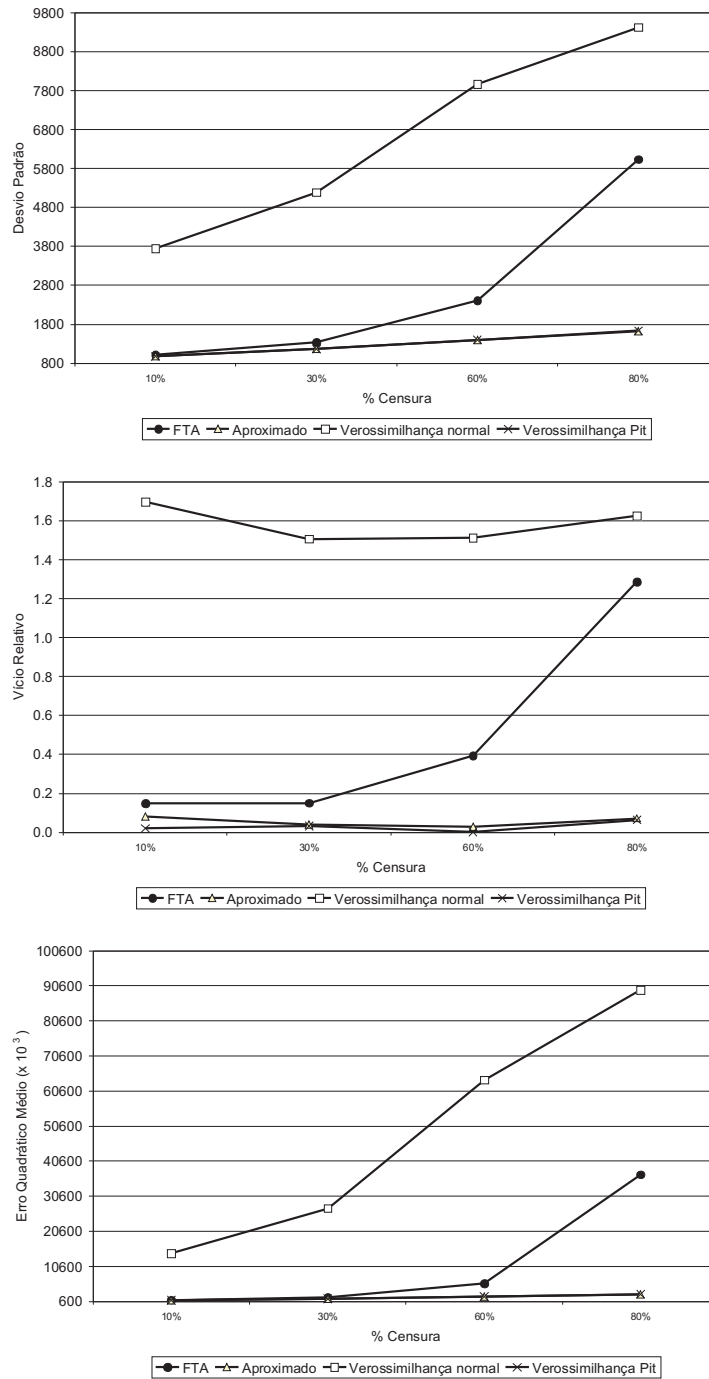


Figura 4.8: Medidas de desempenho estimativas do MTTF
 Caso 2: modelo linear com efeito aleatório Lognormal

Quantil 0,01

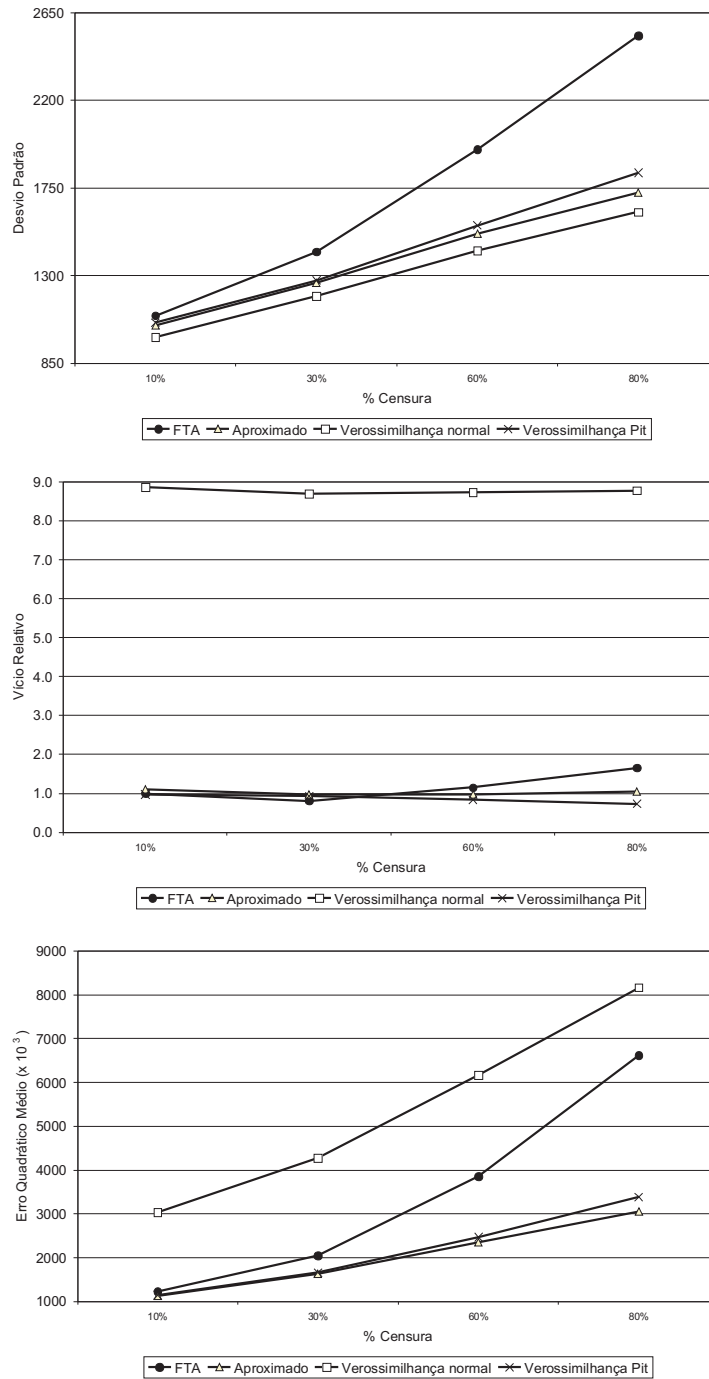


Figura 4.9: Medidas de desempenho estimativas do quantil 0,01
 Caso 2: modelo linear com efeito aleatório Lognormal

Mediana

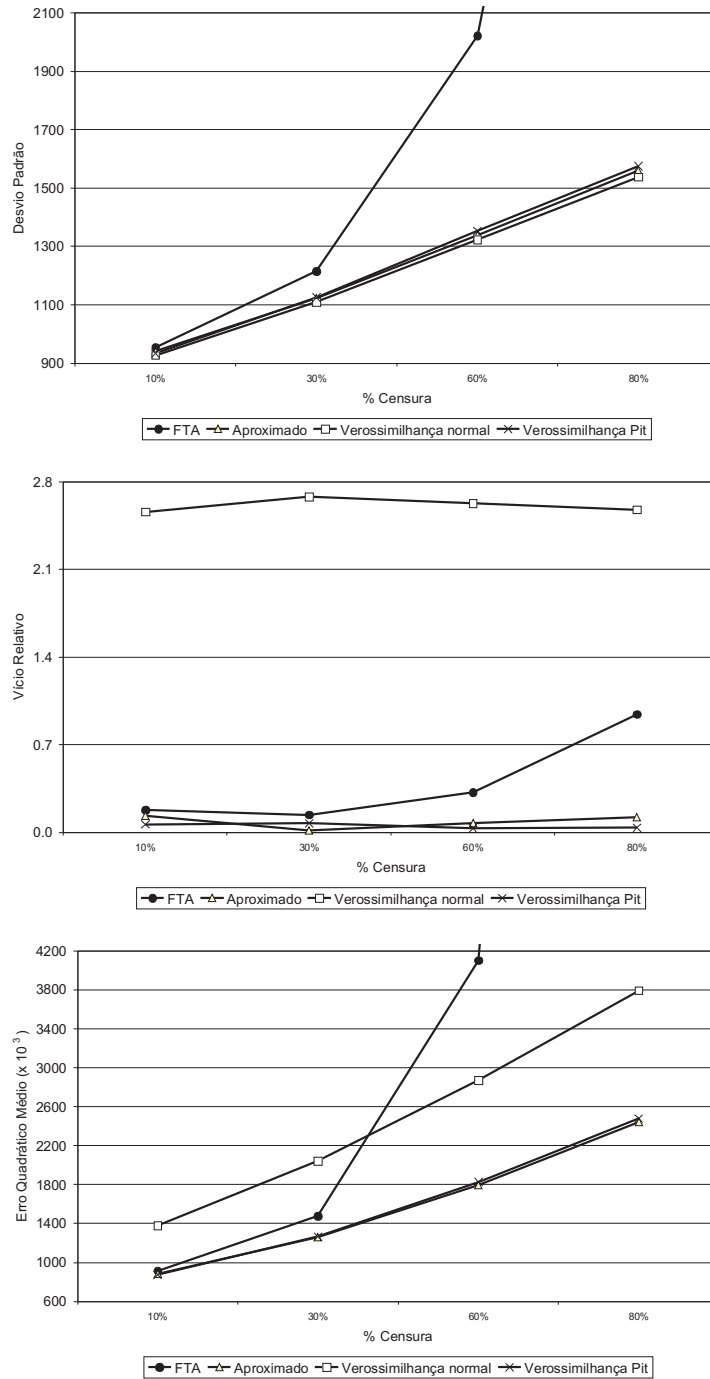


Figura 4.10: Medidas de desempenho estimativas da mediana
 Caso 2: modelo linear com efeito aleatório Lognormal

4.3 Conclusões

Analisando cada método individualmente (FTA, Aproximado, Verossimilhança com efeitos aleatórios com distribuição normal e verossimilhança com PIT e efeitos aleatórios Weibull e Lognormal) podemos notar que:

1. **Precisão das estimativas (avaliada pelo desvio padrão):** para todos os métodos considerados, a precisão das estimativas de todas as figuras de mérito decresce com o aumento do percentual de censuras. Este resultado já era esperado no caso da análise de tempo de falha (FTA) tradicional, visto que é de conhecimento comum que a quantidade de observações censuradas tem impacto na precisão das estimativas. Esta é justamente uma das motivações para o uso de dados de degradação quando esta possibilidade for viável. No caso do método aproximado de análise, um maior número de perfis que não atinge o limiar crítico implica em um maior número de pseudotempos de falha “extrapolados”, que serão posteriormente utilizados como se fossem de fato observações dos tempos de falha (portanto, “dados” sem censura) em uma análise de falha tradicional. Além disto quanto maior o percentual de censuras, maior o valor de D_f e conseqüentemente maior o grau de extrapolação necessária para obter-se os pseudotempos de falha. Portanto, é natural que decresça a precisão das estimativas com o aumento do percentual de censuras também no método aproximado. Nos métodos de máxima verossimilhança as inferências são baseadas nas observações (medidas de degradação). No entanto, quanto maior o percentual de censura, menor o número de perfis que atingem o limiar de falha (D_f) durante o período do estudo. Assim sendo, com um menor número de falhas a informação disponível para inferências a respeito do tempo de vida é menor, gerando estimativas de menor precisão. Isto explica o aumento do desvio padrão das estimativas obtidas pelos métodos baseados na máxima verossimilhança à medida que se aumenta o

percentual de censuras.

2. **Vício relativo das estimativas:** nesta medida, não há um padrão tão claro de comportamento em função do percentual de censuras, seja qual for a figura de mérito considerada. Entretanto percebe-se que, para qualquer um dos métodos, o vício relativo das estimativas de quantis mais baixos, como por exemplo $t_{0,01}$ e $t_{0,05}$ é bem maior que o vício relativo para quantis mais altos, como a mediana.

Comparando agora os métodos de análise entre si, os principais pontos são os seguintes:

1. Não existem diferenças consideráveis entre o desvio padrão das estimativas obtidas pelo método aproximado e pelos métodos da verossimilhança para todas as figuras de mérito, sendo que estas estimativas são sempre melhores que as obtidas pela análise de tempo de falha (FTA) tradicional. A exceção são as estimativas para o MTTF fornecidas pelo método da verossimilhança normal, que apresentou o maior desvio padrão entre todos os métodos, para todos os percentuais de censura.
2. Na medida do vício relativo é que o método da verossimilhança PIT tanto para a distribuição de Weibull quanto para a Lognormal apresenta ótimos resultados: de fato o vício destas estimativas é o menor entre todos os métodos para praticamente todas as figuras de mérito, independentemente do percentual de censura. Já sabíamos, como apontado por Freitas *et al.* (2009) e Toledo (2007), que o método da verossimilhança normal produzia estimativas com vício elevado devido a violação do pressuposto de normalidade dos efeitos aleatórios.
3. O Erro Quadrático Médio das estimativas obtidas pelo método da verossimilhança PIT para as duas distribuições utilizadas foi menor que o das estimativas obtidas pelo método da verossimilhança normal, para todas as figuras de mérito em quase todos os percentuais de censura. Mesmo que eventualmente em alguns casos o

desvio padrão destas estimativas seja maior do que o das estimativas obtidas pela verossimilhança normal, o vício consideravelmente menor compensa esta diferença. Os resultados do Erro Quadrático Médio foram praticamente os mesmos dos obtidos pelo método aproximado, ainda que ligeiramente maiores.

O resultado das simulações demonstraram portanto que a utilização da solução numérica de maximização da verossimilhança baseada no PIT apresentada na Seção 3.3, é eficiente na análise de modelos de degradação com perfis lineares simples, fornecendo boas estimativas para a distribuição do tempo de vida.

Capítulo 5

Situações motivadoras revisitadas

5.1 Rodas de trem

Nesta seção retornamos ao problema descrito na Seção 1.1.1, analisando os dados de degradação de rodas de trem, apresentados na Figura 1.3. Iremos analisar as 14 rodas *MA11* que possuem o mesmo diâmetro inicial (degração em $t_0 = 0$) com o objetivo de obter a distribuição do tempo de vida, que neste caso é o tempo gasto até que a roda atinja o limiar crítico de degradação, definido como $D_f = 77 \text{ mm}$. Utilizamos a solução de maximização da verossimilhança numericamente com a transformação PIT, e vamos comparar os resultados deste método com os obtidos por Freitas *et al.* (2009). Intervalos de confiança para as estimativas foram obtidos por bootstrap paramétrico e não paramétrico, (Meekeer e Escobar, 1998) e também serão apresentados. As rotinas desenvolvidas para o *Software R*, tanto para estimativa dos parâmetros do modelo quanto para as simulações bootstrap estão no anexo B.

Na Tabela 5.1 estão as medidas de degração das 14 rodas analisadas. Apenas as rodas 11, 12 e 14 atingiram o limiar de degradação no período de observação, configurando um nível de censura de 78% (11/14).

Tabela 5.1: Medidas de degradação (mm.) das rodas MA11

Unidade	Distância percorrida (mil quilômetros)													
	0	50	100	150	200	250	300	350	400	450	500	550	600	
Amostrai	0	2,7000	4,0659	5,6609	7,2631	8,6850	10,4017	12,0305	13,4421	14,9470	16,6277	18,1337	19,6648	
Roda 1	0	2,8998	4,9251	6,6984	8,6252	10,4220	12,4066	14,3119	16,1846	17,9799	19,8565	21,8192	23,7242	
Roda 2	0	3,7028	6,4751	9,2832	11,8857	14,6862	17,3731	20,2155	22,8683	25,6566	28,3640	31,0904	33,8826	
Roda 3	0	4,6583	8,3716	12,2653	15,9242	19,5283	23,2358	27,0731	30,6517	34,3705	38,2276	41,8718	45,4724	
Roda 4	0	6,6897	12,2041	17,9402	23,5024	29,1228	34,8227	40,4531	46,0059	51,6525	57,3747	62,9516	68,5743	
Roda 5	0	4,8609	8,6231	12,5503	16,3594	20,2660	24,0793	27,9391	31,7847	35,6754	39,5580	43,3766	47,2591	
Roda 6	0	2,8000	5,5000	10,6000	12,9000	16,4000	19,6000	25,1000	28,4000	32,5000	37,6000	42,3000	46,5000	
Roda 7	0	3,0743	5,2000	7,3039	9,5605	11,4660	13,6586	15,6574	17,7745	19,9214	22,0137	24,1333	26,2075	
Roda 8	0	6,3538	11,4653	16,7972	22,0818	27,3303	32,5504	37,9657	43,0866	48,4552	53,6830	59,0720	64,3083	
Roda 9	0	4,2360	7,5336	10,7680	14,1162	17,4094	20,7214	23,8505	27,1674	30,4357	33,7930	37,0533	40,4565	
Roda 10	0	11,5675	22,2514	32,8311	43,5843	54,1357	64,7981	75,3103	85,9533					
Roda 11	0	14,7473	28,4222	42,2152	55,8604	69,5656	83,2889							
Roda 12	0	6,5000	15,7000	18,4000	22,2000	26,5000	29,6000	33,6000	38,5000	43,2000	47,1000	51,8000	55,1000	
Roda 13	0	7,9946	14,9741	22,0686	28,9677	35,9954	42,9094	49,9912	56,9346	64,0608	70,9726	77,9534		
Roda 14	0													

5.1.1 Modelo de degradação linear simples com efeito aleatório Weibull

Inicialmente utilizamos o modelo apresentado em (3.6):

$$Y_{ij} = \frac{1}{\beta_i} t_{ij} + \varepsilon_{ij}, \quad i = 1, \dots, 14 \text{ e } j = 1, \dots, m_i,$$

em que, conforme visto na tabela 5.1, $m_{11} = 9$, $m_{12} = 7$ e $m_{14} = 12$. Para as demais rodas, $m_i = 13$. Além disso $t_{i1} = t_1 = 0$, $t_{i,2} = t_2 = 50.000$, \dots , $t_{i,13} = t_{13} = 600.000 \forall i$. Neste modelo supomos que $\beta \sim Weibull(\alpha, \delta)$, com função de densidade dada por (3.7).

Conforme descrito na Seção 3.3.1, para utilizarmos a solução proposta, o primeiro passo é conseguir os valores iniciais $(\hat{\alpha}^{(0)}, \hat{\delta}(0) \text{ e } \hat{\sigma}_\varepsilon^{2(0)})$ para os parâmetros do modelo.

Inicialmente efetuamos o ajuste de modelos lineares sem intercepto para cada roda, obtendo, por mínimos quadrados, estimativas para a inclinação de cada perfil. Os recíprocos destas inclinações são as estimativas iniciais para β_i , conforme mostrado na Tabela 5.2.

Tabela 5.2: Rodas - estimativas iniciais para β_i (Weibull)

Roda número	Estimativa $\hat{\beta}_i$
1	29.7974
2	24.8349
3	17.4985
4	13.0441
5	8.6929
6	12.5883
7	13.5898
8	22.5055
9	9.2729
10	14.7148
11	4.6297
12	3.5848
13	10.3967
14	7.0149

A inclinação do perfil de degradação neste modelo é inversamente proporcional a β , e portanto os perfis de degradação das unidades 11, 12, e 14, que tiveram falhas antes do fim do período do estudo e que são os de maior inclinação correspondem aos menores valores de $\hat{\beta}$. Por outro lado, as rodas 1, 2, e 8, que tiveram a menor degradação no período e portanto têm perfis com menor inclinação, correspondem aos maiores valores de $\hat{\beta}$. Veja detalhes na Figura 5.1.

Em seguida, considerando os valores da Tabela 5.2 como sendo uma amostra de uma variável aleatória Weibull (α, δ) , encontramos com base nos mesmos as estimativas de máxima verossimilhança de α e δ , as quais, conforme já mencionado, serão utilizadas como valores iniciais dos parâmetros. Assim:

$$\hat{\alpha}^{(0)} = 1,9779 \quad \text{e} \quad \hat{\delta}^{(0)} = 15,54574 .$$

No próximo passo, levamos as estimativas $\hat{\beta}_i$ em $\hat{Y}_{ij} = t_j/\hat{\beta}_i$ e calculamos $\hat{\varepsilon}_{ij}$, cuja variância fornece a estimativa inicial para σ_ε^2 :

$$\hat{\sigma}_\varepsilon^{2(0)} = 0,7726 .$$

Podemos agora escrever a verossimilhança apresentada em (3.8):

$$f(\mathbf{y}|\hat{\theta}^{(0)}, \hat{\sigma}_\varepsilon^{2(0)}) = \prod_{i=1}^n \left\{ \int_{\Xi_{\beta_i}} \left[\prod_{j=1}^{m_i} \frac{\exp\left(-\frac{(y_{ij} - t_{ij}/\beta_i)^2}{2\hat{\sigma}_\varepsilon^{2(0)}}\right)}{\sqrt{2\pi\hat{\sigma}_\varepsilon^{(0)}}} \right] \frac{\hat{\alpha}^{(0)}}{\hat{\delta}^{(0)}} \left(\frac{\beta_i}{\hat{\delta}^{(0)}}\right)^{\hat{\alpha}^{(0)}-1} \exp\left[-\left(\frac{\beta_i}{\hat{\delta}^{(0)}}\right)^{\hat{\alpha}^{(0)}}\right] d\beta_i \right\}, \quad (5.1)$$

ou seja, a contribuição da i -ésima roda para a verossimilhança já com os valores iniciais $\hat{\alpha}^{(0)}$, $\hat{\delta}^{(0)}$ e $\hat{\sigma}_\varepsilon^{2(0)}$ substituídos é dada por:

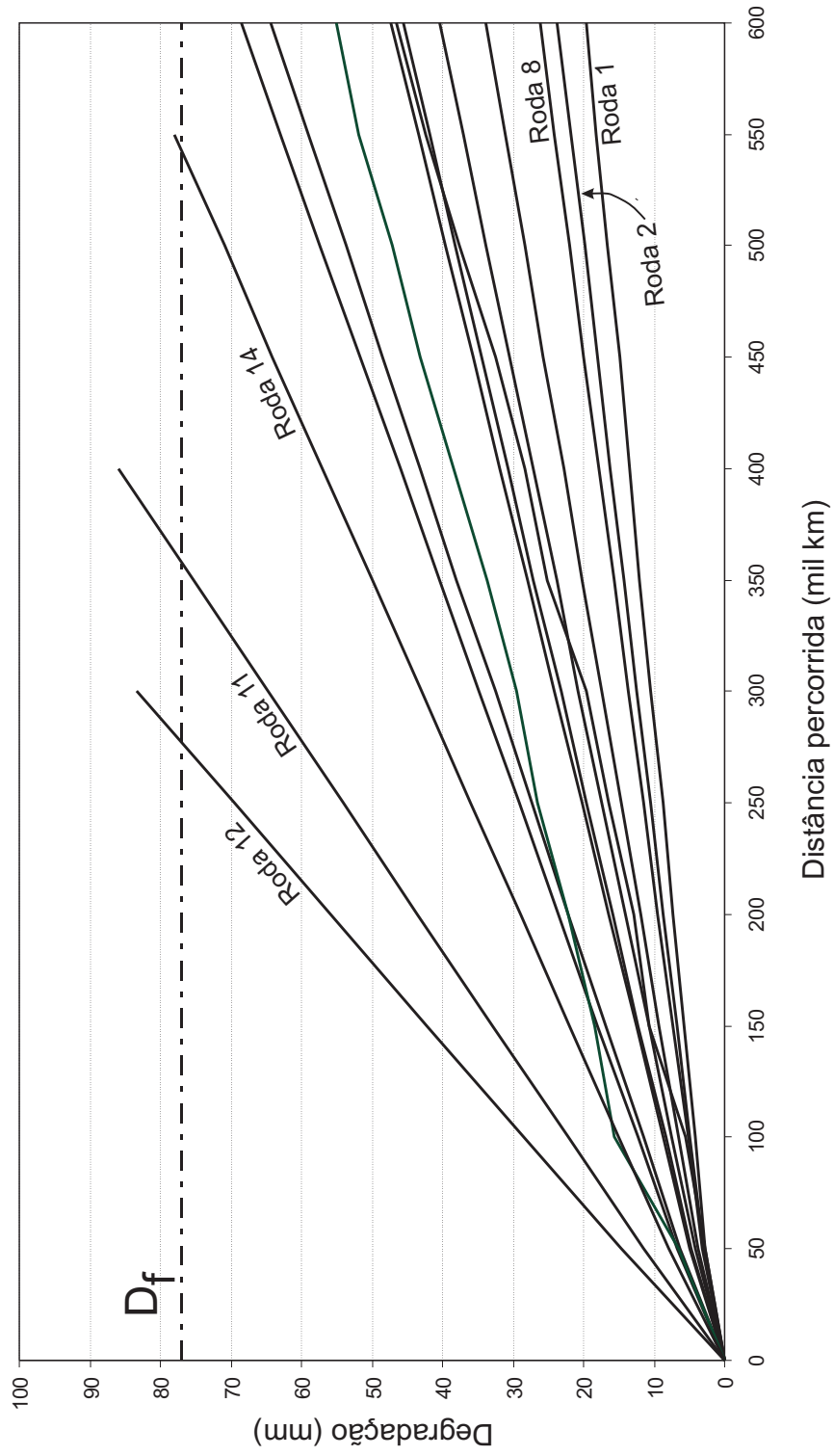


Figura 5.1: Detalhe dos perfis de degradação das rodas MA11

$$\begin{aligned}
 & f(y_i|\hat{\theta}_{\sim}^{(0)}, \hat{\sigma}_{\varepsilon}^{2(0)}) = \\
 & = \int_{\Xi_{\beta_i}} \left\{ \left[\prod_{j=1}^{m_i} \frac{\exp\left(-\frac{(y_{ij} - t_{ij}/\beta_i)^2}{2 \times 0,7726}\right)}{\sqrt{2\pi} \sqrt{0,7726}} \right] \frac{1,9779}{15,54574} \left(\frac{\beta_i}{15,54574}\right)^{1,9779-1} \exp\left\{-\left(\frac{\beta_i}{15,54574}\right)^{1,9779}\right\} d\beta_i \right\} \\
 & = \int_{\Xi_{\beta_i}} \left\{ \left[\prod_{j=1}^{m_i} \frac{\exp\left(-\frac{(y_{ij} - t_{ij}/\beta_i)^2}{1,5451}\right)}{2,2033} \right] 0,0087\beta_i^{0,9779} \exp\left(-0,0044\beta_i^{1,9779}\right) d\beta_i \right\} = \int_{\beta_i} g(\beta_i) d\beta_i. \quad (5.2)
 \end{aligned}$$

A título de ilustração, a Figura 5.2 mostra o gráfico da função $g(\beta_i)$, ou seja, o integrando da expressão (5.2) para cada uma das 14 rodas. As rodas 7 e 13 por terem menores valores para $\int_{\beta_i} g(\beta_i) d\beta_i$ não aparecem no gráfico na escala utilizada.

Os estimadores de máxima verossimilhança são aqueles que maximizam (5.1) ou de maneira equivalente, $\log f(y_i|\hat{\theta}_{\sim}^{(0)}, \hat{\sigma}_{\varepsilon}^{2(0)})$:

$$\log f(y_i|\hat{\theta}_{\sim}^{(0)}, \hat{\sigma}_{\varepsilon}^{2(0)}) = \sum_{i=1}^{14} \log \left\{ \int_{\Xi_{\beta_i}} \left[\prod_{j=1}^{m_i} \frac{\exp\left(-\frac{(y_{ij} - t_{ij}/\beta_i)^2}{1,5451}\right)}{2,2033} \right] 0,0087\beta_i^{0,9779} \exp\left(-0,0044\beta_i^{1,9779}\right) d\beta_i \right\} \quad (5.3)$$

ou seja, tornam máxima a soma do logaritmo das áreas sob as curvas representadas na Figura 5.2.

Para aproximarmos o valor das integrais acima usamos, como citado na Seção 3.3.1, a soma de Riemann: se $f(x) \geq 0$ podemos aproximar $\int_a^b f(x) dx$ por $\sum_{i=1}^n f(x_i^*) \Delta x$ para um número adequado n de pontos x_i^* igualmente espaçados por Δx no intervalo (a, b) . No caso aqui estudado, para obtermos uma aproximação da ordem de 10^{-3} na logverossimilhança temos que usar $\Delta x = 0,01$ no intervalo $(0, 35)$. Como consequência, $n = 3500$. A escolha deste intervalo foi feita à luz do gráfico apresentado na Figura 5.2, o que no entanto pode não ser tarefa simples para análise de modelos com maior número de unidades

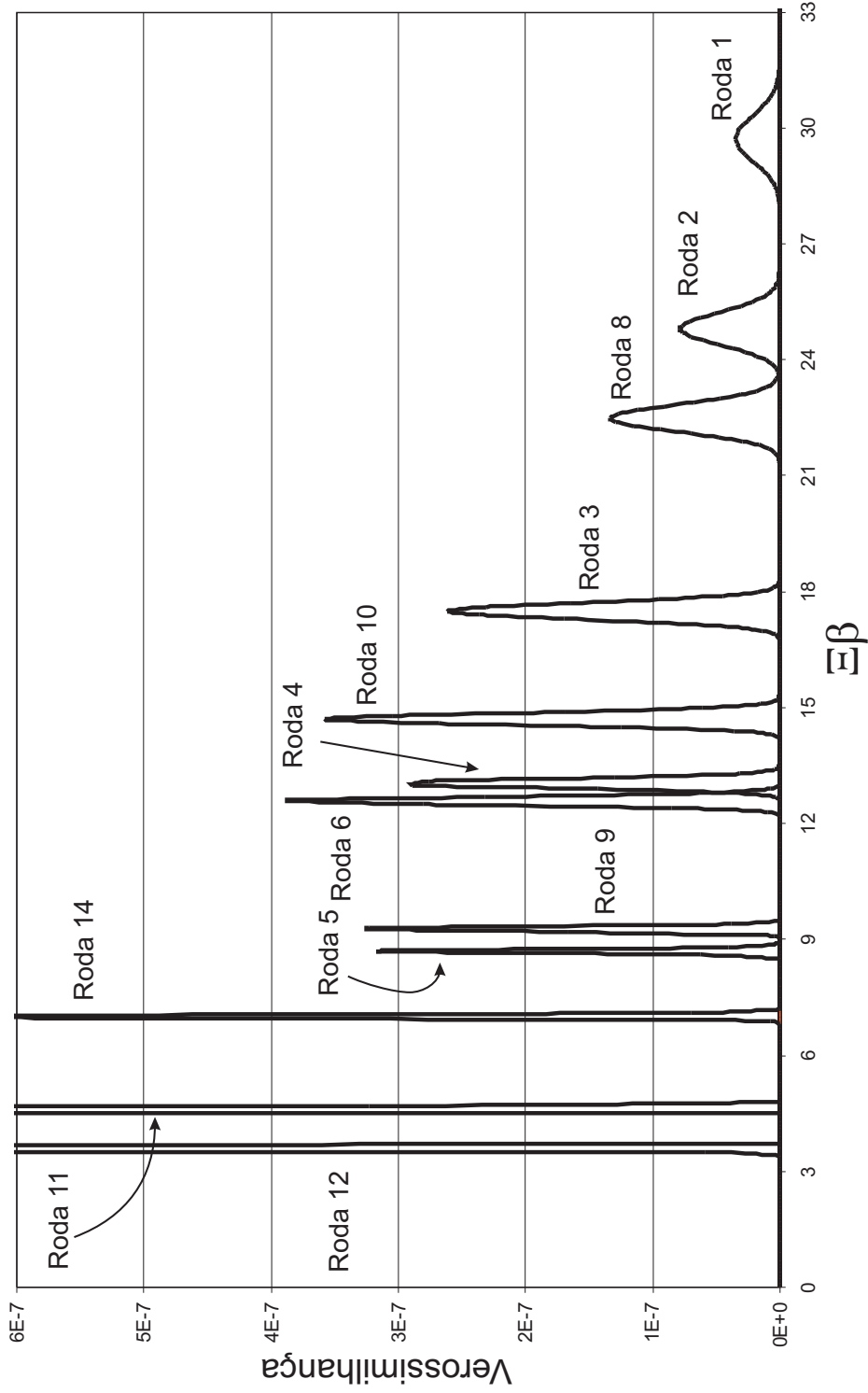


Figura 5.2: Gráfico de $g^*(a_i)$ (integrando de 5.2) para cada uma da 14 rodas (exceto 7 e 13)

experimentais ou com maior quantidade de medidas por unidade, o que configura mais uma vantagem de utilizarmos a mudança de variáveis descrita a seguir.

Implementando a mudança de variáveis sugerida por Nelson *et al.* (2006), reduzimos o número de intervalos na soma de Riemann sem reduzirmos a precisão da aproximação. Usando a logverossimilhança apresentada em (3.11), com os mesmos valores de partida, (5.3) pode ser reescrita como:

$$\begin{aligned} \log f(y_i | \hat{\theta}^{(0)}, \hat{\sigma}_\varepsilon^{2(0)}) &= \\ &= \sum_{i=1}^{14} \log \left\{ \int_{\Xi_{a_i}} \left[\prod_{j=1}^{m_i} \frac{\exp \left(-\frac{[y_{ij} - t_{ij}/15,54574(-\log(1 - \Phi(a_i))^{1/1,9779})^2]}{1,5451} \right)}{2,2033} \right] \phi(a_i) da_i \right\} \\ &= \sum_{i=1}^{14} \log \left\{ \int_{\Xi_{a_i}} g^*(a_i) da_i \right\}. \end{aligned} \tag{5.4}$$

A figura 5.3 apresenta o gráfico dos integrandos da expressão 5.4 ($g^*(a_1), \dots, g^*(a_{14})$), ou seja, dos integrandos para cada uma das rodas ($i = 1, 2, \dots, 14$). Com esta mudança, ao invés de integrarmos em β , cujo suporte ($\Xi\beta$) era $(0 \text{ a } 35)$, integramos em a , com suporte (Ξa) no intervalo $(-3 \text{ a } 3)$. Obtemos a mesma precisão na soma de Riemann com o mesmo $\Delta x = 0,01$, mas com redução do número de pontos para $n = 600$. O método de quadratura Gaussiana sugerido por Nelson *et al.* (2006) para aproximar o valor da integral não funcionou adequadamente para os dados de degradação aqui analisados, devido à falta de suavidade da função de verossimilhança. Mesmo se tomarmos o número máximo de abscissas predeterminadas (pontos de quadratura) o resultado obtido não produz uma aproximação razoável.

A redução de custo computacional na implementação da rotina de otimização utilizada para maximização da soma das áreas da função de verossimilhança é considerável quando utilizamos (5.4) ao invés de (5.3). A rotina de otimização (função `constrOptim` do *Software R-11*) usando um computador com processador AMD Athlon 64 × 2 Dual com 2,40 GHZ

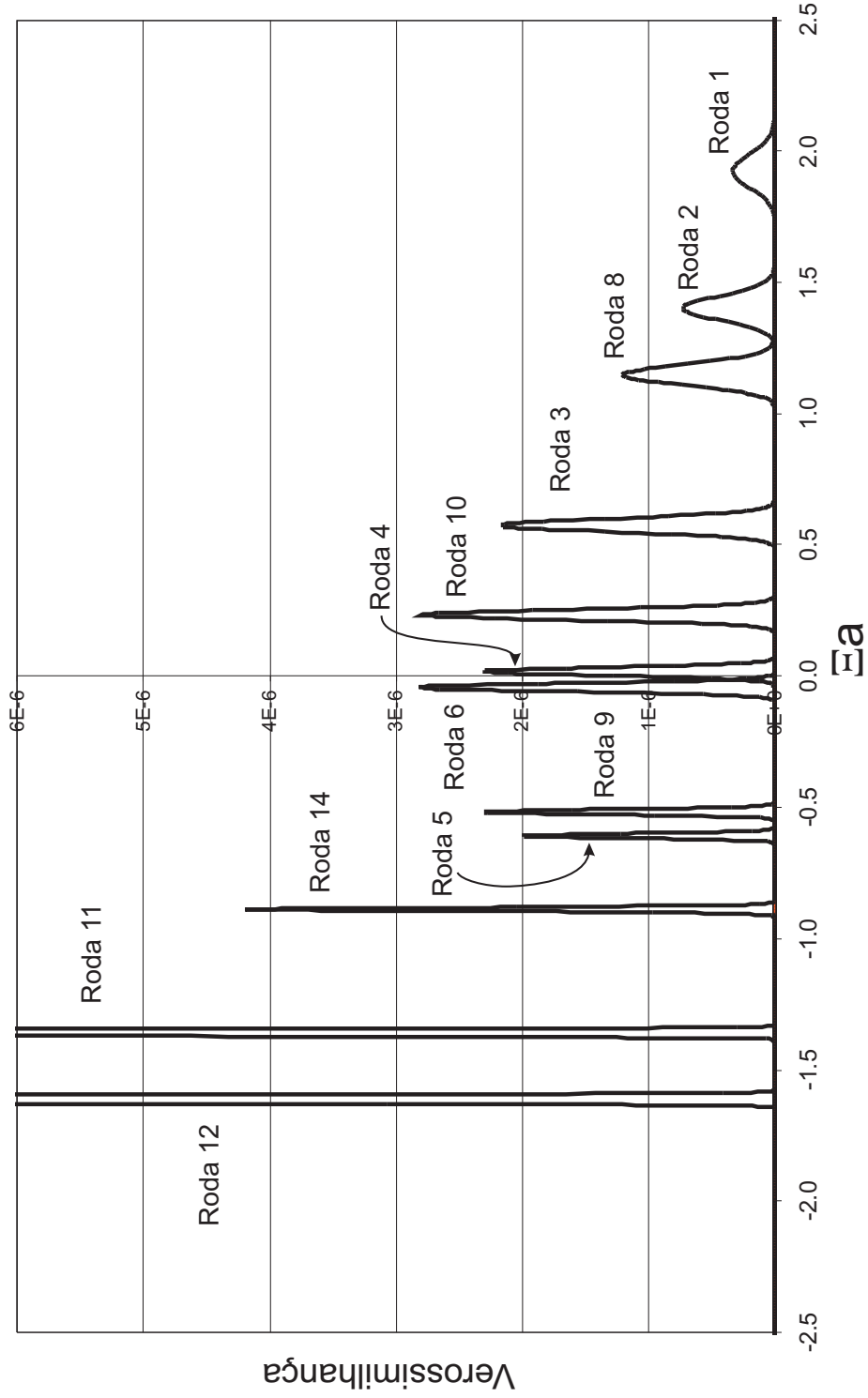


Figura 5.3: Gráfico de $g^*(a_i)$ (integrando de 5.4) para cada uma da 14 rodas (exceto 7 e 13)

e 3,25 GB de RAM gastou 146,67 segundos com a verossimilhança original, e apenas 25,33 segundos com a versão transformada.

As estimativas de máxima verossimilhança encontradas para cada parâmetro foram:

$$\hat{\alpha} = 1,9765 \quad \hat{\delta} = 15,56133 \quad \hat{\sigma}_\varepsilon^2 = 0,8836 .$$

De posse destes valores, calculamos com auxílio das expressões (4.1) e (4.2), as estimativas para as mesmas figuras de mérito apresentadas por Freitas *et al.* (2009): o MTTF, os quantis 0,10 e 0,50 e a confiabilidade após 300.000 km. Além disso, construímos intervalos de confiança bootstrap para estas estimativas, usando dois métodos diferentes. No bootstrap não paramétrico, amostramos com reposição dos 14 perfis originais, 10.000 conjuntos de 14 perfis. Para cada um dos conjuntos, aplicamos a solução de maximização da verossimilhança Weibull com transformação “PIT” e geramos “estimativas bootstrap” das quatro grandezas. Os percentis 2,5% e 97,5% destas estimativas forneceram os limites dos intervalos de 95% confiança. No bootstrap paramétrico, usando as estimativas de máxima verossimilhança acima para os parâmetros α , δ e σ_ε^2 , e o modelo assumido, geramos também 10.000 amostras de 14 perfis. Cada um dos perfis simulados era censurado considerando-se $D_f = 77 \text{ mm}$. Novamente aplicamos a solução para cada conjunto simulado e calculamos os intervalos de confiança como anteriormente. Os scripts utilizados encontram-se no anexo B.4. Os resultados são apresentados na tabela 5.4, onde são também reproduzidas as estimativas pontuais e intervalos de confiança obtidos pelos métodos aproximado, verossimilhança normal (“numérico”) e análise de falha tradicional nas simulações descritas por Freitas *et al.* (2009).

5.1.2 Modelo de degradação linear simples com efeito aleatório lognormal

Utilizamos agora o modelo (3.12):

$$Y_{ij} = \beta_i t_{ij} + \varepsilon_{ij}, \quad i = 1, \dots, 14 \text{ e } j = 1, \dots, m_i,$$

com $\beta \sim \text{Lognormal}(\mu, \sigma)$.

Repetimos o mesmo passo inicial, usando diretamente as estimativas retornadas pelo ajuste de modelos lineares para cada perfil, por mínimos quadrados. Estas estimativas estão na Tabela 5.3.

Tabela 5.3: Rodas - estimativas iniciais para β_i (Lognormal)

Roda número	Estimativa $\hat{\beta}_i$
1	0,03356
2	0,04027
3	0,05715
4	0,07666
5	0,11504
6	0,07944
7	0,07358
8	0,04443
9	0,10784
10	0,06796
11	0,21600
12	0,27895
13	0,09618
14	0,14255

Note que neste modelo, inversamente ao que ocorria no modelo de Weibull, a inclinação dos perfis é diretamente proporcional a β e portanto os perfis de unidades que falharam durante o período do estudo correspondem aos maiores valores de $\hat{\beta}$, e os que tiveram

menor degradação, aos menores.

Considerando agora que os dados da Tabela 5.3 correspondem a observações de uma variável aleatória lognormal, obtivemos as seguintes estimativas de máxima verossimilhança para μ e σ :

$$\hat{\mu}^{(0)} = -2,462153 \quad \text{e} \quad \hat{\sigma}^{(0)} = 0,585486 .$$

Os valores de $\hat{\beta}_i$ (Tabela 5.3) substituídos em $\hat{Y}_{ij} = \hat{\beta}_i t_{ij}$ permitem o cálculo de estimativas iniciais de ε_{ij} ($\hat{\varepsilon}_{ij}^{(0)}$) e de σ_ε^2 ($\hat{\sigma}_\varepsilon^{2(0)}$), como descrito na Seção 5.1.1:

$$\hat{\sigma}_\varepsilon^{2(0)} = 0,7726 .$$

Utilizamos estas três estimativas para escrevermos a verossimilhança mostrada em (3.13), ou seja:

$$f(y|\hat{\theta}_{\sim}, \hat{\sigma}_\varepsilon^{2(0)}) = \prod_{i=1}^n \left\{ \int_{\Xi_{\beta_i}} \left[\prod_{j=1}^{m_i} \frac{\exp\left(-\frac{(y_{ij} - \beta_i t_{ij})^2}{2\hat{\sigma}_\varepsilon^{2(0)}}\right)}{\sqrt{2\pi}\hat{\sigma}_\varepsilon^{(0)}} \right] \frac{1}{\beta_i \sqrt{2\pi}\hat{\sigma}} \exp\left\{-\frac{[\log(\beta_i) - \hat{\mu}^{(0)}]^2}{2\hat{\sigma}^2(0)}\right\} d\beta_i \right\} . \quad (5.5)$$

Portanto a contribuição de cada unidade experimental (roda) para a verossimilhança, já com os valores numéricos iniciais dos parâmetros é dada por:

$$\begin{aligned} f(y|\hat{\theta}_{\sim}, \hat{\sigma}_\varepsilon^{2(0)}) &= \\ &= \int_{\Xi_{\beta_i}} \left\{ \left[\prod_{j=1}^{m_i} \frac{\exp\left(-\frac{(y_{ij} - \beta_i t_{ij})^2}{2 \times 0,7726}\right)}{\sqrt{2\pi}\sqrt{0,7726}} \right] \frac{1}{\beta_i \sqrt{2\pi}\sqrt{0,342793}} \exp\left\{-\frac{[\log(\beta_i) + 2.462153]^2}{2 \times 0,342793}\right\} d\beta_i \right\} = \\ &= \int_{\Xi_{\beta_i}} \left\{ \left[\prod_{j=1}^{m_i} \frac{\exp\left(-\frac{(y_{ij} - \beta_i t_{ij})^2}{1,5451}\right)}{2,2033} \right] 1,4676\beta_i^{-1} \exp\left\{-\frac{[\log(\beta_i) + 2.462153]^2}{0,6856}\right\} d\beta_i \right\} . \quad (5.6) \end{aligned}$$

Os estimadores de máxima verossimilhança serão obtidos pela maximização de (5.6), ou seja, de maneira equivalente, pela maximização de $\log f(y|\hat{\theta}^{(0)}, \hat{\sigma}_\varepsilon^{2(0)})$:

$$\sum_{i=1}^{14} \log \left\{ \int_{\Xi_{\beta_i}} \left[\prod_{j=1}^{m_i} \frac{\exp \left(-\frac{(y_{ij} - \beta_i t_{ij})^2}{1,5451} \right)}{2,2033} \right] 1,4676 \beta_i^{-1} \exp \left\{ -\frac{[\log(\beta_i) + 2.462153]^2}{0,6856} \right\} d\beta_i \right\}, \quad (5.7)$$

que usando a transformação apresentada em (3.16) é reescrita em (5.8):

$$\sum_{i=1}^{14} \log \left\{ \int_{\Xi_{a_i}} \left[\prod_{j=1}^{m_i} \frac{\exp \left(-\frac{(y_{ij} - Q\text{lnorm}(\Phi(a_i); -2,4622; 0,3428)t_{ij})^2}{1,5451} \right)}{2,2033} \right] \phi(a_i) da_i \right\}. \quad (5.8)$$

As estimativas de máxima verossimilhança para μ , σ e σ_ε^2 encontradas são:

$$\hat{\mu} = -2,462008 \quad \hat{\sigma} = 0,585563 \quad \hat{\sigma}_\varepsilon^2 = 0,882825 .$$

Substituindo estas estimativas nas equações (4.4) e (4.5) obtemos as estimativas pontuais para as grandezas de interesse, cujos intervalos de confiança foram obtidos pelos dois métodos de bootstrap referidos na Seção 5.1.1. Estes resultados também são apresentados na tabela 5.4.

5.1.3 Conclusões

No trabalho de Freitas *et al.* (2009) os autores já tinham concluído que havia indicações de que, para os dados de desgaste das rodas, a suposição de normalidade para os efeitos aleatórios não é adequada. Os autores basearam esta conclusão em gráficos de probabilidade Weibull e lognormal construídos para os “pseudotempos de falha” (estimados a partir do método aproximado de análise). Ambos indicavam que estas distribuições eram adequadas para a modelagem dos tempos até a “falha” das rodas e, conseqüentemente,

Tabela 5.4: Estimativas pontuais e intervalos de confiança

Método	Distrib.	Estimativa ($\times 10^3$ km)			Log	
		MTTF	$t_{0,10}$	Mediana		R(300)
Verossimilhança Transformada PIT	Weibull	1062,14 [789; 1362] (1) [779; 1371] (2)	383,78 [210; 694] [236; 661]	995,42 [714; 1321] [726; 1326]	0,937 [0,836; 0,994] [0,856; 0,990]	-289,197
	Lognormal	1071,97 [764; 1477] (1) [778; 1398] (2)	426,40 [297; 668] [290; 711]	903,08 [669; 1233] [660; 1219]	0,970 [0,897; 1,000] [0,891; 1,000]	-289,317
Aproximado	Weibull	1060,88 [804; 1400]	383,35 [208; 707]	994,25 [727; 1359]	0,937 [0,771; 0,984]	
	Lognormal	1071,87 [769; 1494]	426,28 [282; 645]	902,64 [664; 1227]	0,970 [0,884; 0,997]	
Max. Verosim. Normal ("Numérico")	Normal	1914,15 [850; 9913]	400,00 [311; 534]	701,33 [540; 980]	0,984 [0,918; 1,000]	
	Weibull	971,77 [422; 2239]	471,60 [228; 766]	937,13 [459; 1912]	0,952 [0,689; 0,994]	
Tempo de Falha Tradicional	Lognormal	1390,50 [359; 5387]	411,59 [237; 716]	1059,29 [450; 2576]	0,956 [0,768; 0,996]	

(1) I.C. 95% - Bootstrap paramétrico

(2) I.C. 95% - Bootstrap não paramétrico

devido à forma funcional assumida para o modelo de degradação (linear simples) estas mesmas distribuições poderiam ser utilizadas para os efeitos aleatórios dos modelos considerados. Os gráficos construídos para os “pseudotempos de falha” equivalem aos gráficos das figuras 5.4 (a) e (b) construídos para as estimativas $\hat{\beta}_i$ de cada perfil. Pensando nestes valores como se fossem uma amostra de uma variável aleatória, vemos que tanto a distribuição de Weibull como a lognormal poderiam descrever a distribuição da mesma.

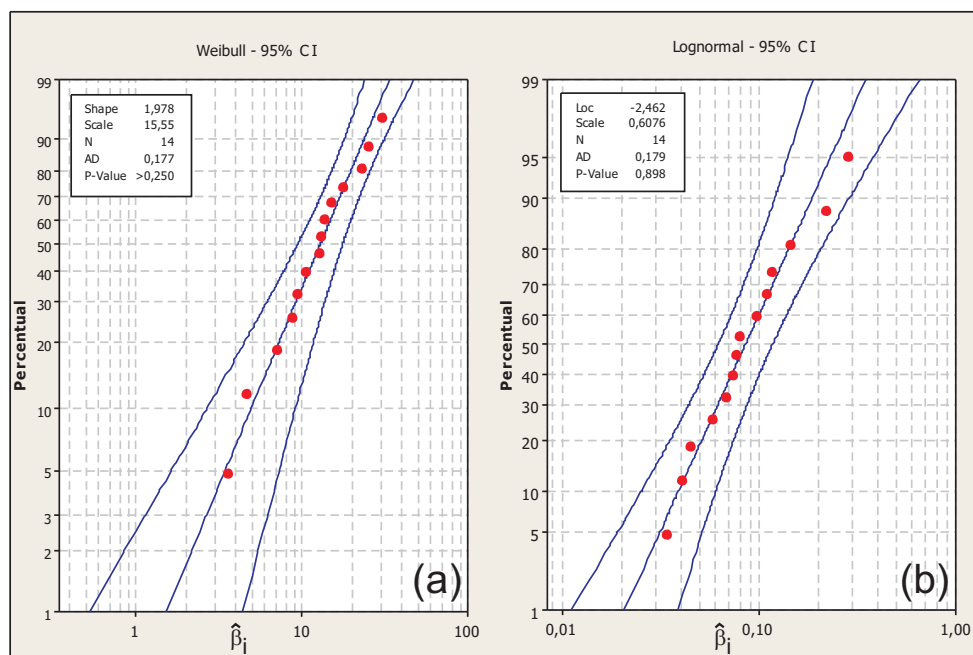


Figura 5.4: Gráfico de probabilidade para $\hat{\beta}_i$

Os principais pontos a serem destacados são:

1. Os resultados obtidos pela maximização da verossimilhança transformada PIT supondo que o tempo de vida tem distribuição de Weibull foram bem semelhantes aos obtidos considerando distribuição lognormal. Como em ambos os modelos o número de parâmetros a serem estimados é o mesmo (3) poderíamos optar por um deles com base na log-verossimilhança. Mas este valor também é praticamente o

mesmo, independentemente da distribuição escolhida. Por outro lado, a largura dos intervalos de confiança fornecidos pelo modelo log-normal, à exceção daqueles referentes à distância média percorrida pela roda (MTTF), são menores que seus equivalentes para o modelo Weibull, indicando uma precisão ligeiramente melhor das estimativas fornecidas por aquele modelo em relação às fornecidas por este último. Esta é a mesma conclusão a que Freitas *et al.* (2009) haviam chegado em seu artigo, com base nos intervalos de confiança fornecidos pelo método aproximado.

2. Não existe uma tendência clara que nos leve a preferir um dos dois métodos utilizados para construção dos intervalos de confiança para as figuras de mérito após a obtenção das estimativas pontuais pela maximização da verossimilhança transformada PIT. O bootstrap paramétrico forneceu intervalos mais estreitos para algumas figuras e o não paramétrico para outras.
3. Os resultados obtidos pela solução implementada neste trabalho são muito próximos daqueles fornecidos pelo método aproximado, para as duas distribuições utilizadas; mas o comprimento dos intervalos de confiança são menores em praticamente todos os casos, sugerindo maior precisão dessas estimativas.
4. As estimativas pontuais obtidas sob a suposição de normalidade dos efeitos aleatórios (método da máxima verossimilhança normal - “numérico”) são muito distintas das obtidas pelos dois primeiros métodos apresentados, o que parece confirmar a inadequação deste método para o caso estudado, conforme já mostrado por Freitas *et al.* (2009). Especialmente a estimativa pontual da distância média percorrida (MTTF), assim como o intervalo de confiança correspondente, demonstra a influência do vício introduzido pela violação do pressuposto de normalidade. Curiosamente, como também já ressaltado por Freitas *et al.* (2009), os intervalos de confiança das outras estimativas fornecidas pela verossimilhança com distribuição normal têm o menor

comprimento entre todos os apresentados.

5. Também os resultados fornecidos pela análise tradicional de tempo de falha não são adequados ao caso estudado, devido principalmente ao elevado percentual de censuras (quase 80%). A baixa precisão deste método, para este nível de censuras, foi demonstrada pelo resultado das simulações.

5.2 Unidades de laser

Nesta seção retomamos o exemplo apresentado por Meeker e Escobar (1998), revisitado por Hamada (2005), e introduzido na Seção 1.1.2, para analisarmos os dados de degradação mostrados na Figura 1.4. Como já descrito, o estudo acompanha 15 unidades de laser com medições do acréscimo percentual de corrente a cada 250 horas até o final do estudo (4.000 horas). A degradação inicial é considerada zero para todas as unidades (corrente inicial de operação igual para todas as unidades) e o limiar crítico de degradação é definido como $D_f = 10\%$, ou seja, quando o acréscimo da corrente de trabalho necessária para manter a luminosidade constante for superior a 10%, caracteriza-se uma falha. Mais uma vez vamos utilizar a solução de maximização da verossimilhança numericamente com a transformação PIT apresentada na Seção 3.3 e comparar os resultados com os apresentados nos trabalhos de referência.

A Tabela 5.5 apresenta as medidas de degradação para as 15 unidades amostrais. Como podemos ver, apenas as unidades 1, 6 e 10 falharam durante o período do estudo, o que representa um nível de “censura” de 80% (12/15).

O modelo proposto por Hamada (2005) foi o mesmo de 3.6:

$$Y_{ij} = \frac{1}{\beta_i} t_{ij} + \varepsilon_{ij}, \quad i = 1, \dots, 15 \text{ e } j = 1, \dots, m_i,$$

Tabela 5.5: Medidas de degração (porcentagem) das unidades de laser

Unidade Amostrai	Tempo decorrido (horas)																
	0	250	500	750	1000	1250	1500	1750	2000	2250	2500	2750	3000	3250	3500	3750	4000
1	0	0,47	0,93	2,11	2,72	3,51	4,34	4,91	5,48	5,99	6,72	7,13	8,00	8,92	9,49	9,87	10,94
2	0	0,71	1,22	1,90	2,30	2,87	3,75	4,42	4,99	5,51	6,07	6,64	7,16	7,78	8,42	8,91	9,28
3	0	0,71	1,17	1,73	1,99	2,53	2,97	3,30	3,94	4,16	4,45	4,89	5,27	5,69	6,02	6,45	6,88
4	0	0,36	0,62	1,36	1,95	2,30	2,95	3,39	3,79	4,11	4,50	4,72	4,98	5,28	5,61	5,95	6,14
5	0	0,27	0,61	1,11	1,77	2,06	2,58	2,99	3,38	4,05	4,63	5,24	5,62	6,04	6,32	7,10	7,59
6	0	0,36	1,39	1,95	2,86	3,46	3,81	4,53	5,35	5,92	6,71	7,70	8,61	9,15	9,95	10,49	
7	0	0,36	0,92	1,21	1,46	1,93	2,39	2,68	2,94	3,42	4,09	4,58	4,84	5,11	5,57	6,11	7,17
8	0	0,46	1,07	1,42	1,77	2,11	2,40	2,78	3,02	3,29	3,75	4,16	4,76	5,16	5,46	5,81	6,24
9	0	0,51	0,93	1,57	1,96	2,59	3,29	3,61	4,11	4,60	4,91	5,34	5,84	6,40	6,84	7,20	7,88
10	0	0,41	1,49	2,38	3,00	3,84	4,50	5,25	6,26	7,05	7,80	8,32	8,93	9,55	10,45		
11	0	0,44	1,00	1,57	1,96	2,51	2,84	3,47	4,01	4,51	4,80	5,20	5,66	6,20	6,54	6,96	7,42
12	0	0,39	0,80	1,35	1,74	2,98	3,59	4,03	4,44	4,79	5,22	5,48	4,96	6,23	6,99	7,37	7,88
13	0	0,30	0,74	1,52	1,85	2,39	2,95	3,51	3,92	5,03	5,47	5,84	6,50	6,94	7,39	7,85	8,09
14	0	0,44	0,70	1,05	1,35	1,80	2,55	2,83	3,39	3,72	4,09	4,83	5,41	5,76	6,14	6,51	6,88
15	0	0,51	0,83	1,29	1,52	1,91	2,27	2,78	3,42	3,78	4,11	4,38	4,63	5,38	5,84	6,16	6,62

em que $m_6 = 16$ e $m_{10} = 15$. Para todas as outras unidades amostrais, $m_i = 17$. Além disso $t_{i1} = t_1 = 0$, $t_{i2} = t_2 = 250$, $t_{i3} = t_3 = 500$, \dots , $t_{i,17} = t_{17} = 4.000 \forall i$. Seguindo a mesma distribuição utilizada pelos autores dos trabalhos de referência, $\beta \sim Weibull(\alpha, \delta)$, com função de densidade dada por (3.7).

Novamente o primeiro passo é o ajuste de modelos lineares sem intercepto para cada unidade. Os valores iniciais estimados para os β_i ($i = 1, \dots, 15$) obtidos pelo recíproco da inclinação de cada perfil ajustada por mínimos quadrados, são mostrados na Tabela 5.6.

Tabela 5.6: Laser - estimativas iniciais para β_i

Unidade número	Estimativa $\hat{\beta}_i$
1	370,6974
2	417,3440
3	562,1394
4	598,2559
5	543,3376
6	360,8645
7	614,1350
8	641,5467
9	506,5799
10	331,1882
11	526,7593
12	494,8298
13	478,1293
14	581,9450
15	612,1327

As unidades 10, 6 e 1 apresentam os menores valores de $\hat{\beta}_i$, correspondendo às maiores inclinações. De posse destas estimativas iniciais, e supondo que estas são uma amostra aleatória de uma variável aleatória Weibull (α, δ) obtemos os estimadores de máxima verossimilhança para os parâmetros:

$$\hat{\alpha}^{(0)} = 6,6002 \quad \hat{\delta}^{(0)} = 548,275 .$$

Posteriormente, substituindo as estimativas $\hat{\beta}_1, \dots, \hat{\beta}_{15}$ na forma funcional do modelo, obtemos $\hat{\varepsilon}_{ij} = Y_{ij} - \hat{Y}_{ij}$ e a estimativa inicial de σ_ε^2 :

$$\hat{\sigma}_\varepsilon^{2(0)} = 0,037734 .$$

A partir destas estimativas iniciais, a log-verossimilhança já transformada a ser maximizada é dada pela expressão (5.9):

$$\begin{aligned} & \log f(y|\hat{\alpha}^{(0)}, \hat{\delta}^{(0)}, \hat{\sigma}_\varepsilon^{2(0)}) = \\ & = \sum_{i=1}^{15} \log \left\{ \int_{\Xi_{a_i}} \left[\prod_{j=1}^{m_i} \frac{\exp \left(-\frac{[y_{ij} - t_{ij}/548,275(-\log(1 - \Phi(a_i)))^{1/6,6002}]^2}{0,07547}}{0,4869} \right)}{\phi(a_i)} da_i \right] \right\} . \end{aligned} \quad (5.9)$$

Assim as estimativas de máxima verossimilhança para os parâmetros do modelo são:

$$\hat{\alpha} = 6,612 \quad \hat{\delta} = 548,2711 \quad \hat{\sigma}_\varepsilon^2 = 0,039963 .$$

A partir destas estimativas calculamos as figuras de mérito que são apresentadas na tabela 5.7, juntamente com intervalos de 95% de confiança obtidos a partir de amostras bootstrap (paramétrico e não paramétrico). Estas figuras são a confiabilidade em 4.500 horas, o quantil 0,10 e a mediana do tempo de vida. Para efeito comparativo, na mesma tabela estão também os resultados de Meeker e Escobar (1998), que mostraram graficamente estimativas de máxima verossimilhança para $F(t)$ e intervalos de confiança pontuais de 95% usando pseudo-tempos de falhas, assim como os resultados da análise bayesiana

Tabela 5.7: Resumo de algumas figuras de mérito para unidades de laser

Figura de Mérito	Verossimilhança Transformada PIT	Hamada (2005)	Meeker e Escobar (1998)
R(4500)	0.7627	0.7107	
IC 95% Boot Paramétrico	[0.5776; 0.9307]		
IC 95% Boot Não Paramétrico	[0.5563; 0.9436]		
Intervalo de Credibilidade 95%		[0.5236; 0.8477]	
t_{0,1}	3901	3650	3810
IC 95% Boot Paramétrico	[3252; 4663]		
IC 95% Boot Não Paramétrico	[3183; 4741]		
Intervalo de Credibilidade 95%		[2852; 4206]	
IC 95% por verossimilhança			[2900; 4300]
Mediana	5187		5030
IC 95% Boot Paramétrico	[4698; 5640]		
IC 95% Boot Não Paramétrico	[4651; 5650]		
IC 95% por verossimilhança			[4450; 5610]

de Hamada (2005). Este último partiu de uma distribuição *a priori* Weibull para β_i e distribuições *a priori* Gamma para os hiperparâmetros e obteve por *MCMC* a distribuição *a posteriori* do quantil 0,1 da distribuição do tempo até a falha e da confiabilidade em 4.500 horas de funcionamento.

5.2.1 Conclusões

A análise dos resultados apresentados na Tabela 5.7 sugerem que:

1. Os intervalos de confiança fornecidos pelo bootstrap paramétrico são mais estreitos para as três figuras apresentadas.
2. A estimativa pontual da mediana (5187 horas) obtida pela verossimilhança Weibull com transformação PIT, deve ter maior precisão que a apresentada por Meeker e Escobar (1998) (5030 horas) como indicam os intervalos de confiança mais estreitos.

3. As estimativas pontuais para o quantil 0,10 fornecidas pela verossimilhança PIT e por Meeker e Escobar (1998) são bem próximas, assim como a largura dos intervalos de confiança, indicando precisão similar. No entanto ambas são maiores que a apresentada por Hamada (2005), cujo intervalo de credibilidade é cerca de 4% mais estreito que os primeiros.

4. O intervalo de confiança para a estimativa da confiabilidade em 4500 horas fornecida pela verossimilhança PIT é ligeiramente maior que o intervalo de credibilidade para a mesma medida fornecida por Hamada (2005), podendo indicar menor precisão para aquela estimativa.

Capítulo 6

Conclusões e trabalhos futuros

Modelos de dados de degradação são cada vez mais usados na estimação da confiabilidade de itens, produtos, peças e equipamentos, através de análises que permitam inferências sobre a distribuição do seu tempo de vida. A modelagem tradicional supõem efeitos aleatórios com distribuição normal atuando na degradação e a função de distribuição do tempo de vida é estimada com base neste pressuposto. No entanto, exemplos práticos e simulações mostraram que a utilização inadequada da suposição de normalidade levava à baixa precisão das estimativas obtidas.

Neste trabalho apresentamos duas abordagens para modelos de dados de degradação com efeitos aleatórios não normais publicadas recentemente: Nelson *et al.* (2006) e Liu e Yu (2008). À luz de duas situações práticas motivadoras, dados de degradação de rodas de trens e de unidades de laser, propusemos a implementação de solução numérica baseada na maximização da verossimilhança com transformação sugerida por Nelson *et al.* (2006). Para testar a efetividade da solução, implementamos estudos de simulação como os realizados por Toledo (2007) e Freitas *et al.* (2009) e comparamos os nossos resultados (maximização da verossimilhança com efeitos aleatórios não normais - Weibull e Log-normal - com transformação PIT) com os obtidos por três métodos utilizados por estes

autores (aproximado, verossimilhança com efeitos normais e análise de falha tradicional).

Pudemos concluir que:

1. De forma similar ao que acontece com estimativas fornecidas pelos três métodos originais, o desvio padrão das estimativas obtidas por nossa solução aumenta com o aumento do percentual de censura. Como as inferências são feitas com base em medidas de degradação, à medida que aumenta o percentual de censuras há uma redução do número de perfis simulados que apresentam falha no período do estudo, reduzindo a “quantidade” de informação disponível para inferências sobre a distribuição do tempo de vida.
2. O desvio padrão das estimativas obtidas na simulação utilizando nossa implementação é praticamente igual às obtidas pelos métodos aproximado e verossimilhança normal (às vezes ligeiramente menor outras vezes ligeiramente maior). A exceção são as estimativas da distância média (MTTF) obtidas pelo método da verossimilhança normal para o modelo com efeitos com distribuição lognormal, que são muito maiores que os outros dois métodos. Já os resultados do FTA tradicional apresentam desvios muito maiores, especialmente para os maiores percentuais de censura.
3. O vício das estimativas obtidas pela verossimilhança Weibull e Lognormal com transformação PIT é menor que o obtido pela verossimilhança normal em praticamente todas as 48 situações simuladas (6 figuras de mérito x 4 níveis de censura x 2 distribuições). Isto comprova o grave prejuízo à precisão das estimativas quando se supõe que os efeitos aleatórios têm distribuição normal. Este vício é também menor que o apresentado por boa parte (68%) das estimativas obtidas pelo método aproximado.

Há portanto indicações da boa adequação da solução implementada na estimação do tempo de vida em perfis de degradação simples, linear sem intercepto, como os usados nas simulações.

Com relação aos casos práticos apresentados podemos concluir:

1. Desgaste das rodas do trem:

- (a) As distribuições Weibull e Lognormal utilizadas para a distribuição do tempo de vida parecem ser adequadas, com ligeira vantagem para esta última, como já havia sido apontado por Freitas et al (2009).
- (b) As estimativas fornecidas pela solução implementada neste trabalho apresentaram boa precisão e corroboram com a afirmação da inadequação da suposição de efeitos aleatórios normais.
- (c) A grande semelhança dos resultados obtidos pela solução de maximização de verossimilhança com transformação PIT, tanto no modelo Weibull quanto lognormal, com aqueles obtidos pelo método aproximado, nos leva a concluir que para este nível de censuras, número de unidades amostrais e número de medições por unidade, parece ser perfeitamente aceitável a utilização do método aproximado.

2. Unidades de laser:

- (a) A estimativa da mediana do tempo de vida das unidades fornecida pela maximização da verossimilhança Weibull com transformação PIT aparentemente tem maior precisão que a obtida pelo método aproximado utilizado por Meeker e Escobar (1998).
- (b) As estimativas da confiabilidade em 4500 horas e do quantil 10% da distribuição do tempo de vida foram superiores às obtidas pela análise bayesiana de Hamada (2005), com intervalos de confiança ligeiramente mais largos (9% e 4%) que os intervalos de credibilidade deste autor.

De modo geral, nos casos de modelos de degradação simples (lineares sem intercepto) como os simulados e dos exemplos práticos apresentados, a solução de maximização da

verossimilhança com efeitos aleatórios Weibull e Lognormal com a transformação PIT apresentou resultados satisfatórios e parece ser uma boa alternativa para os métodos que analisam dados de degradação com suposição de normalidade dos efeitos aleatórios.

Como sugestão para trabalhos futuros fica a extensão da solução para modelos mais complexos, como os lineares com mais de um efeito aleatório e os não lineares. Também seria interessante o estudo de alternativas computacionais para aproximação numérica da integral em (3.1), já que a quadratura gaussiana com abcissas fixas não foi adequada e o custo computacional da nossa solução aumenta proporcionalmente com o número de unidades em estudo ou com o número de medidas de degradação por unidade.

Referências Bibliográficas

- Abramowitz, M. e Stegun, I. (1972). *Handbook of Mathematical Functions*. Dover: New York.
- Agresti, A., Caffo, B., e Strickland-Ohman, P. (2004). Examples in which misspecification of random effects distribution reduces efficiency, and possible remedies. *Computational Statistics & Data Analysis*, 47:639–653.
- Bogdanoff, J. L. e Kozin, F. (1985). *Probabilistic Models of Cumulative Damage*. John Wiley: New York.
- Breslow, N. e Clayton, D. (1993). Approximate inference in generalized linear mixed models. *Journal of the American Statistical Association*, 88(421):9–25.
- Colosimo, E. A. e Giolo, S. R. (2006). *Análise de Sobrevida Aplicada*. Edgard Blücher.
- Crk, V. (2000). Reliability assessment from degradation data. *Proceedings of the Annual Reliability and Maintainability Symposium*, pages 155–161.
- Davidian, M. e Gallant, A. R. (1992). Smooth nonparametric maximum likelihood estimation for population pharmacokinetics, with an application to quinidine. *Journal of Pharmacokinetics and Biopharmaceutics*, 80:475–488.
- Doksum, K. A. (1991). Degradation rate models for failure time and survival data. *CWI Quarterly*, 4:195–203.

-
- Efron, B. (1985). Bootstrap confidence intervals for a class of parametric problems. *Biometrika*, 72(1):45–58.
- Elsayed, E. A. (1996). *Reliability Engineering*. Addison Wesley.
- Freitas, M. A. e Colosimo, E. A. (1997). *Confiabilidade: Análise de Tempo de Falha e Testes de Vida Acelerados*. Fundação Christiano Ottoni, Escola de Engenharia da UFMG.
- Freitas, M. A., Colosimo, E. A., Santos, T. R., e Pires, M. C. (2010). Reliability assessment using degradations models: bayesian and classic approaches. *Pesquisa Operacional*, 30(1):195–219.
- Freitas, M. A., Toledo, M. L. G., Colosimo, E. A., e Pires, M. C. (2009). Using degradation data do assess reliability: a case study on train wheel degradation. *Quality and Reliability Engineering International*, 25:607–629.
- Genest, C. e Rivest, L. P. (2001). On the multivariate probability integral transformation. *Statistics and Probability Letters*, 53:391–399.
- Gertsbackh, I. B. e Kordonskiy, K. B. (1969). *Models of Failure*. Springer: New York.
- Hamada, M. (2005). Using degradation data to assess reliability. *Quality Engineering*, 17:615–620.
- Hoel, P. G., Port, S. C., e Stone, C. J. (1971). *Introduction to Probability Teory*. Houghton Mifflin Company.
- Jiang, M. e Zhang, Y. (2002). Dynamic modeling of degradation data. *Proceedings of the Annual Reliability and Maintainability Symposium*, pages 607–611.
- Lange, K. (2010). *Numerical Analysis for Statisticians*. Springer.

-
- Lee, E. e Nelder, J. (1996). Hierarchical generalized linear models (with discussion). *Journal of the Royal Statistical Society, Series B*, 58:619–678.
- Lesaffre, E. e Spiessens, B. (2001). On the effect of the number of quadrature points in a logistic random-effects model: an example. *Applied statistics*, 50:325–335.
- Lewis, E. E. (1996). *Introduction to Reliability Engineering*. John Wiley & Sons: New York, 2 edition.
- Lindstrom, M. J. e Bates, D. M. (1990). Nonlinear mixed-effects models for repeated measures data. *Biometrics*, 46:673–687.
- Litière, S., Alonso, S., e Molenberghs, G. (2008). The impact of a misspecified random-effects distribution on the estimation and the performance of inferential procedures in generalized linear mixed models. *Statistics in Medicine*, 27:3125–3144.
- Liu, L. e Yu, Z. (2008). A likelihood reformulation method in non-normal random effects models. *Statistics in Medicine*, 27:3105–3124.
- Liu, Q. e Pierce, D. A. (1994). A note on gauss-hermite quadrature. *Biometrika*, 81:624–629.
- Lu, C. e Meeker, W. (1993). Using degradation measurements to estimate a time-to-failure distribution. . *Technometrics*, 35:161–174.
- Lu, J., Park, J., e Yang, Q. (1997). Statistical inference of a time-to-failure distribution derived from linear degradation data. *Technometrics*, 39(4):391–400.
- McGilchrist, C. (1994). Estimation in generalized mixed models. *Journal of the Royal Statistical Society*, 56(1):61–69.
- Meeker, W. e Escobar, L. A. (1998). *Statistical Methods for Reliability Data*. Wiley Series in Probability and Statistics: New York, 1 edition.

-
- Nelson, K. P., Lipsitz, S. R., Fitzmaurice, G. M., Ibrahim, J., Parzen, M., e Strawderman, R. (2006). Use of the probability integral transformation to fit nonlinear mixed-effects models with nonnormal random effects. *Journal of Computational and Graphical Statistics*, 15(1):39–57.
- Nelson, W. (1981). Analysis of performance degradation data from accelerated tests. *IEEE Transactions on Reliability*.
- Nelson, W. (1982). *Applied Life Data Analysis*. John Wiley & Sons: New York.
- Nelson, W. (1990). *Accelerated Testing: Statistical Models, Test Plans and Data Analyses*. John Wiley: New York.
- Oliveira, V. R. B. e Colosimo, E. A. (2004). Comparison of methods to estimate the time-to-failure distribution in degradation tests. *Quality and Reliability Engineering International*, 20:363–373.
- Pinheiro, J., Liu, C., e Wu, Y. (2001). Efficient algorithms for robust estimation in linear mixed-effects models using the multivariate t distribution. *Journal of Computational and Graphical Statistics*, 10:249–276.
- Pinheiro, J. C. e Bates, D. C. (1995). Approximations to the log-likelihood function in the nonlinear mixed-effects model. *Journal of Computational and Graphical Statistics*, 1(4):12–35.
- Pinheiro, J. C. e Bates, D. M. (2000). *Mixed-Effects Models in S and S-PLUS*. Statistics and Computing Series, Springer-Verlag, New York.
- Song, P., Fan, Y., e Kalbfleisch, J. (2005). Maximization by parts in likelihood inference (with discussion). *Journal of the American Statistical Association*, 100:1145–1167.

-
- Su, C., Lu, J. C., Chen, D., e Hughes-Oliver, J. M. (1999). A random coefficient degradation model with random sample size. *Lifetime Data Anal*, 5:173–183.
- Tang, L. C. e Chang, D. S. (1995). Reliability prediction using nondestructive accelerated degradation data: Case study on power supplies. *IEEE Transactions on Reliability*, 44(4):562–566.
- Toledo, M. L. G. (2007). Ensaio de degradação: estudo comparativo de métodos de análise de dados. Dissertação de Mestrado, Departamento de Estatística - ICEX - UFMG.
- Tseng, S., Hamada, M., e Chiao, C. (1995). Using degradation data from a factorial experiment to improve fluorescent lamp reliability. *Journal of Quality Technology*, 27:363–369.
- Verbeke, G. e Lesaffre, E. (1997). The effect of misspecifying the random-effects distribution in linear mixed effects distribution in linear mixed models for longitudinal data. *Computational Statistics & Data Analysis*, 23:541–556.
- Verbeke, G. e Molenberghs, G. (2000). *Linear Mixed Models for Longitudinal Data*. Springer-Verlag: New York, 2 edition.
- Whitmore, G. A. e Shenkelberg, F. (1997). Modelling accelerated degradation data using wiener diffusion with a time scale transformation. *Lifetime Data Analysis*, 3:27–45.
- Wolfinger, R. e Lin, X. (1997). Two taylor-series approximation methods for non linear mixed models. *Computational Statistics and Data Analysis*, 25:465–490.
- Wu, S. J. e Shao, J. (1999). Reliability analysis using the least squares method in nonlinear mixed-effect degradation models. *Statist Sinica*, 9:855–877.

-
- Wu, S. J. e Tsai, T. R. (2000). Estimation of time-to-failure distribution derived from a degradation model using fuzzy clustering. *Quality and Reliability Engineering International*, 16:261–267.
- Yacout, A. M., Salvatores, S., e Orechwa, Y. (1996). Degradation analysis estimates of the time-to-failure distribution of irradiated fuel elements. *Nuclear Technology*, 113:177–189.
- Zhuang, D. (1994). *Degradation failure model and its statistical analysis*. PhD thesis, East China Normal University.

Anexo A

Tabelas de resultados das simulações

Tabela A.1: Resultado das simulações
 Caso 1: modelo linear com efeitos aleatórios com distribuição de Weibull

Figura de Mérito	% Censura	DESVIO PADRÃO				VÍCIO RELATIVO				MSE (milhares)			
		FTA	Aprox	Veros Normal	Veros PIT	FTA	Aprox	Veros Normal	Veros PIT	FTA	Aprox	Veros Normal	Veros PIT
MTTF	10%	821.7	814.0	816.1	815.2	0.052	0.026	0.005	0.012	675	663	666	665
	30%	966.2	905.1	906.8	906.7	0.184	0.066	0.047	0.022	938	820	823	822
	60%	1093.5	951.6	954.5	1053.3	0.218	0.025	0.001	0.004	1203	906	911	1109
	80%	5225.1	1181.6	1182.9	1194.3	1.346	0.003	0.018	0.135	27667	1396	1399	1430
$t_{0,01}$	10%	1600.6	1479.6	1802.0	1468.8	1.114	1.550	10.445	1.439	2593	2250	6005	2210
	30%	1992.9	1645.4	2007.2	1608.1	1.058	1.537	10.409	1.572	4006	2781	7405	2663
	60%	2179.3	1702.9	2082.1	1878.9	0.870	1.408	10.160	1.222	4780	2981	8557	3591
	80%	3725.4	2131.6	2584.3	2158.5	2.286	0.464	10.322	1.453	14157	4555	12367	4772
$t_{0,05}$	10%	1421.2	1336.1	1444.4	1328.1	0.608	0.887	3.984	0.815	2035	1819	2764	1792
	30%	1710.0	1485.6	1609.6	1455.1	0.577	0.894	3.978	0.903	2942	2249	3423	2160
	60%	1837.6	1543.0	1674.9	1701.5	0.446	0.802	3.824	0.674	3391	2425	3815	2926
	80%	2625.3	1926.8	2073.2	1951.2	1.230	0.828	3.904	0.869	7029	3774	5671	3875
$t_{0,10}$	10%	1282.6	1222.1	1267.0	1216.1	0.419	0.623	1.645	0.567	1655	1514	1751	1496
	30%	1505.3	1359.0	1410.8	1334.0	0.416	0.639	1.651	0.635	2277	1874	2171	1806
	60%	1597.8	1414.9	1471.4	1559.7	0.320	0.562	1.532	0.461	2562	2029	2369	2451
	80%	2225.2	1764.2	1820.4	1786.1	1.000	0.574	1.583	0.639	5065	3150	3598	3236
Mediana	10%	839.7	832.6	816.3	833.8	0.029	0.011	1.398	0.002	705	693	859	695
	30%	981.9	925.7	907.0	926.9	0.145	0.051	1.357	0.009	967	857	1046	859
	60%	1104.6	973.4	954.6	1077.5	0.169	0.011	1.401	0.020	1225	948	1224	1161
	80%	5045.1	1208.3	1183.0	1222.0	1.126	0.010	1.420	0.119	25717	1460	1818	1496
$t_{0,80}$	10%	895.6	814.7	832.7	817.6	0.098	0.224	0.398	0.220	803	670	714	675
	30%	1281.6	904.5	923.1	914.5	0.094	0.174	0.342	0.235	1644	823	871	845
	60%	1592.6	946.0	965.1	1059.1	0.186	0.199	0.356	0.197	2544	903	958	1130
	80%	8383.5	1179.9	1210.2	1199.0	1.545	0.235	0.403	0.076	70933	1407	1509	1439

Tabela A.2: Resultado das simulações
 Caso 2: modelo linear com efeitos aleatórios com distribuição Lognormal

Figura de Mérito	% Censura	DESVIO PADRÃO				VÍCIO RELATIVO				MSE (milhares)			
		FTA	Aprox	Veros Normal	Veros PIT	FTA	Aprox	Veros Normal	Veros PIT	FTA	Aprox	Veros Normal	Veros PIT
MTTF	10%	1010.1	981.4	3750.4	976.2	0.149	0.080	1.697	0.020	1022	964	14306	953
	30%	1334.5	1170.5	5186.0	1170.0	0.151	0.039	1.507	0.032	1784	1370	27166	1369
	60%	2408.4	1397.0	7966.7	1404.6	0.393	0.028	1.512	0.001	5827	1952	63861	1973
	80%	6035.1	1627.9	9423.1	1632.5	1.285	0.073	1.626	0.064	36797	2651	89394	2666
t0,01	10%	1094.7	1043.5	985.2	1058.3	0.996	1.110	8.859	0.970	1225	1121	3041	1145
	30%	1422.6	1263.4	1194.9	1275.1	0.809	0.974	8.691	0.930	2048	1632	4276	1658
	60%	1946.6	1516.1	1427.5	1558.4	1.153	0.978	8.729	0.838	3861	2350	6168	2467
	80%	2533.1	1724.7	1626.7	1828.9	1.658	1.051	8.773	0.735	6613	3054	8163	3384
t0,05	10%	1008.3	974.6	970.5	983.4	0.719	0.792	3.428	0.671	1035	973	1370	984
	30%	1269.1	1178.4	1174.6	1189.2	0.559	0.660	3.272	0.646	1627	1411	1937	1436
	60%	1608.9	1411.6	1401.4	1454.1	0.799	0.680	3.314	0.568	2636	2027	2785	2138
	80%	1900.8	1613.1	1602.5	1705.9	1.183	0.747	3.359	0.473	3751	2657	3683	2932
t0,10	10%	956.7	933.9	956.2	938.7	0.584	0.634	1.177	0.524	930	890	974	893
	30%	1179.3	1127.6	1155.8	1137.3	0.445	0.503	1.029	0.506	1403	1287	1401	1309
	60%	1435.7	1348.9	1378.5	1389.1	0.647	0.532	1.072	0.436	2098	1845	2002	1946
	80%	1763.5	1547.1	1580.4	1628.6	1.018	0.596	1.119	0.345	3231	2435	2644	2666
Mediana	10%	955.8	940.7	927.2	934.3	0.178	0.134	2.559	0.065	916	886	1378	873
	30%	1215.0	1123.7	1109.6	1124.8	0.142	0.013	2.680	0.073	1479	1263	2043	1266
	60%	2021.6	1339.7	1323.0	1352.7	0.318	0.073	2.627	0.034	4103	1796	2871	1830
	80%	4783.7	1560.6	1538.3	1574.6	0.944	0.122	2.577	0.037	23075	2439	3794	2480
t0,80	10%	1422.8	1335.4	1378.8	1339.5	0.028	0.143	0.112	0.184	2025	1786	1903	1798
	30%	2020.2	1591.4	1647.6	1589.8	0.028	0.258	0.208	0.160	4081	2544	2722	2532
	60%	3934.2	1908.3	1982.5	1914.0	0.279	0.176	0.143	0.175	15496	3649	3935	3671
	80%	9561.5	2211.4	2290.9	2225.3	1.351	0.139	0.090	0.231	92006	4897	5251	4969

Anexo B

Rotinas implementadas para o

Software R-11

B.1 Simulação de modelo com efeitos aleatórios Wei-

bull

```
rm(list=ls())
require(nlme)
# #
# Parâmetros iniciais - os mesmos utilizados nas simulações de
# Oliveira e Colosimo (2004) e Toledo, M.L.G. (2007)
# #
shape< -6.230596
scale< -10767.69
Df< -3.09 # Valor do limiar de degradação que induz 10% de censuras
# (também simulado com 3.43 (30%), 3.93 (60%) e 4.49 (80%))
MTTF_real< -Df*scale*gamma(1+(1/shape))
Perc1_real< -Df*scale*(-log(1-0.01))^(1/shape)
Perc5_real< -Df*scale*(-log(1-0.05))^(1/shape)
Perc10_real< -Df*scale*(-log(1-0.10))^(1/shape)
Perc50_real< -Df*scale*(-log(1-0.50))^(1/shape)
Perc80_real< -Df*scale*(-log(1-0.80))^(1/shape)
# #
# Loglik -j Função que calcula a verossimilhança transformada pelo
# método PIT no intervalo -3.5 a 3.5, com leituras a cada 0.005
# #
# Variável que recebe os valores de f(x)
joint <- array(list(NULL), c(1400,1))
acum=matrix(c(rep(0,50)),ncol=50)
```

```

# função de verossimilhança transformada
Loglik <- function(data, alpha, delta, sigma){
int <- function(yx, a, d, s){
y <- yx[,1]
t <- yx[,2]
for (p in 1:1400){
joint[p] <-
exp( sum(dnorm(y, mean=t/qweibull(pnorm((p/200-3.5), log=TRUE),a,d,log=TRUE),
sd=s, log=TRUE)) + dnorm((p/200-3.5), log=TRUE) ) }
Joint <- Vectorize(joint)
return(Joint) }
lista <- lapply(data, int, a=alpha, d=delta, s=sigma)
for (i in 1:50){
acum[,i] <- log(sum(unlist(as.vector(lista[i]))) * 0.005)}
loglik <- sum(acum)
return(loglik) }
# #
# Gerando perfis - 50 unidades com 20 medidas a cada 2000km (de 0 a 38000)
# #
Geração <- function(B) {
PerfisSim <- matrix(0,1000,B)
for (k in 1:B){
beta <- rweibull(50,shape,scale)
epsilon <- rnorm(1000,0,0.01)
perfis <- matrix(c(rep(1:50,each=20),rep(seq(0,38000,2000),50),
rep(0,1000)),ncol=3, byrow=F)

```

```

for(i in 1:1000) {
  if (perfis[i,2]==0){
    perfis[i,3]< -0}
  else{
    if (perfis[i,2]/beta[perfis[[i,1]]] + epsilon[i]>0)
      {perfis[i,3]< -0}
    else{
      perfis[i,3]< -perfis[i,2]/beta[perfis[[i,1]]] + epsilon[i]}}
  colnames(perfis)=c("Roda", "Distância", "Desgaste")
  perfis <- as.data.frame(perfis)
  ## Eliminando as medidas após a ultrapassagem de Df
  min <- vector(mode = "numeric", length = 50)
  for (i in 1:50) {
    min[i] <- 20
    for (j in 1:20) {
      if(perfis$Desgaste[perfis$Roda==i][j]>=Df) {
        if (j <= min[i]) {
          min[i] <- j } } } }
    for (i in 1:50){
      if (min[i]!= 20)
        {for (j in (min[i]+1):20)
          perfisDesgaste[perfisRoda==i][j]< - NA } }
      PerfisSim[,k] <- perfis[,3] }
  ## Salvando os perfis simulados
  write.table(PerfisSim, file = "perfisSim.csv", sep = ",",
  col.names = FALSE,row.names=T)

```

```

return(PerfisSim)
}
# Gerando B perfis simulados
Geração(B)
# #
# Estimando os parâmetros iniciais, otimizando a verossimilhança
# e exibindo os resultados
# #
recover <- read.table(file="perfisSim.csv",sep="," ,head=F)
recover <- recover[,-1]
Boot <- function(B){
Resultado <- matrix(0,B,8)
for (k in 1:B){
perfis <- matrix(c(rep(1:50,each=20),rep(seq(0,38000,2000),50),
rep(0,2000)),ncol=4, byrow=F)
perfis[,3] <- recover[,k]
if(sum (is.na(perfis[,3]))!=0)
{perfis <- perfis[-(which(is.na(perfis[,3]))),]}
colnames(perfis)=c("Roda","Distância","Desgaste","Baprox")
## Cálculo de  $b_i \wedge \alpha$  através de ajuste modelo linear para cada unidade
mod <- lmList(Desgaste Distância-1|Roda,perfis)
FitAprox <- array((1/coef(mod))[,1])
## Estimadores de máxima verossimilhança para alfa e delta
VerWeib <- function(y,alphaap)
((sum(y^alphaap*log(y)))/(sum(y^alphaap))-
(1/alphaap)-(sum(log(y))/length(FitAprox)))^2

```

```

alpha=optimize(VerWeib, c(0,100), tol = 0.00001, y=FitAprox)
hatalphaap=alpha$minimum
hatdeltaap=(1/length(FitAprox))*
sum(FitAprox^hatalphaap)^(1/hatalphaap)
## Estimador para sigma^2 erro
Fitaprox <- c(rep(0,length(perfis[,4])))
Erroaprox <- c(rep(0,length(perfis[,4])))
perfis <- cbind(perfis,Fitaprox,Erroaprox)
for(i in 1:length(perfis[,4])){
j=perfis[i,1]
perfis[i,4] <- FitAprox[j] }
perfis <- as.data.frame(perfis)
perfis$Fitaprox <- perfis$Distância/perfis$Baprox
perfis$Erroaprox <- perfis$Desgaste-perfis$Fitaprox
hatsigma2ap=var(perfis$Erroaprox)
Estimadores=matrix(c(hatalphaap,hatdeltaap,hatsigma2ap), ncol=3)
colnames(Estimadores) <- c("alpha","delta","sigma2")
Estimadores
# #
# Verossimilhança completa e otimização
# #
## Formatando os dados
Aux=matrix(c(rep(0,1020)),ncol=51)
dimnames(Aux)=list(seq(1:20),c("X",paste("T",seq(1:50),sep=)))
Aux[,1] <- seq(0,38000,2000)
for (j in 2:51) {

```

```

for (i in 1:length(which(perfis[,1]==j-1))) {
  Aux[i,j] <- - perfis[which(perfis[,1]==j-1)[i],3] } }
for (i in 2:20) {
  for (j in 2:51) {
    if (Aux[i,j]==0)
      {Aux[i,j]<- NA} else {Aux[i,j]<- -Aux[i,j]} } }
Aux <- as.data.frame(Aux)
Dados <- na.omit(reshape(Aux,idvar="id",varying=list(2:51),
v.names="y",direction="long")) names(Dados) <- c("x","roda","y","id")
Dados <- Dados[,c("y","x","roda")]
# Criando lista
Y <- vector("list",length=50)
for(j in 1:50){Y[[j]] <- Dados[Dados$roda==j,1:2]}
# #
# Calculando os estimadores de máxima verossimilhança
# #
otimo <- function(par) Loglik(data=Y,alpha=par[1], delta=par[2],sigma=par[3])
estim <- try(constrOptim(c(Estimadores[1],Estimadores[2],
sqrt(Estimadores[3])),NULL, f=otimo,ui=diag(3),
ci=0,control=list(fnscale=-1)),silent=TRUE)
# # Armazenando o resultado para cada conjunto de perfis
Resultado[k,1] <- round(Estimadores[1],4)
Resultado[k,2] <- round(Estimadores[2],4)
Resultado[k,3] <- round(Estimadores[3],9)
if(class(try(round(Loglik(data = Y, alpha=Estimadores[1],
delta=Estimadores[2],sigma=Estimadores[3]),4), silent=TRUE)))!= "try-error"){

```

```

Resultado[k,4] <- round(Loglik(data = Y, alpha=Estimadores[1],
delta=Estimadores[2],sigma=sqrt(Estimadores[3])),4)}
if(class(estim)!= "try-error"){
Resultado[k,5] <- round(estim$par[1],4)
Resultado[k,6] <- round(estim$par[2],4)
Resultado[k,7] <- round(estim$par[3],9)
Resultado[k,8] <- round(estim$value,4)}
} return(Resultado) }
Boot(B)

```

B.2 Simulação de modelo com efeitos aleatórios Log-normal

```

rm(list=ls())
require(nlme)
# #
# Parâmetros iniciais - os mesmos utilizados nas simulações de
# Oliveira e Colosimo (2004) e Toledo, M.L.G. (2007)
# #
media <- -9.12265 # Parâmetro de locação da Lognormal
desvio <- 0.235929 # Parâmetro de escala da Lognormal
Df <- 3.07 # Valor do limiar de degradação que induz 10% de censuras
# (também simulado com 3.67 (30%), 4.40 (60%) e 5.06 (80%))
## Cálculo das figuras de mérito reais para os parâmetros usados
Z1<- qnorm(0.01)
Z5<- qnorm(0.05)

```

```

Z10 <- qnorm(0.1)
Z50 <- qnorm(0.5)
Z80 <- qnorm(0.8)
MTTF_real <- -exp(log(Df)-media+desvio ^ 2/2)
Perc1_real <- -exp(Z1*desvio+log(Df)-media)
Perc5_real <- -exp(Z5*desvio+log(Df)-media)
Perc10_real <- -exp(Z10*desvio+log(Df)-media)
Perc50_real <- -exp(Z50*desvio+log(Df)-media)
Perc80_real <- -exp(Z80*desvio+log(Df)-media)

# #
# Loglik: Função que calcula a verossimilhança transformada pelo
# método PIT no intervalo -3.0 a 3.0, com leituras a cada 0.005
# #
# variável que recebe os valores de f(x)
joint <- array(list(NULL), c(1200,1))
# variável recebe a contribuição de cada unidade
acum <- matrix(c(rep(0,50)),ncol=50)
# função de verossimilhança transformada
Loglik <- function(data, mu, sigma, sigmae){
int <- function(yx, m, s, se){
y <- yx[,1]
t <- yx[,2]
for (p in 1:1200){
joint[p] <-
exp( sum(dnorm(y, mean=t*qlnorm(pnorm((p/200-3),log=TRUE),m,s,log=TRUE),
sd=se, log=TRUE)) + dnorm((p/200-3), log=TRUE) ) }

```

```

Joint <- Vectorize(joint)
return(Joint) }

lista <- lapply(data, int, m=mu, s=sigma, se=sigmae)
for (i in 1:50){
acum[,i] <- log(sum(unlist(as.vector(lista[i]))) * 0.005)}
loglik <- sum(acum)
return(loglik) }

# #
# Gerando perfis - 50 unidades com 20 medidas cada a 2000 km (De 0 a 38000)
# #

Geração <- function(B) {
PerfisSim <- matrix(0,1000,B)
for (k in 1:B){
beta <- rlnorm(50,media,desvio)
epsilon <- rnorm(1000,0,0.01)
perfis <- matrix(c(rep(1:50,each=20),rep(seq(0,38000,2000),50),
rep(0,1000)),ncol=3, byrow=F)
for(i in 1:1000) {
if (perfis[i,2]==0){
perfis[i,3]< -0}
else{
if (perfis[i,2]*beta[perfis[[i,1]]]+epsilon[i]>0) {
perfis[i,3]< -0}
else{
perfis[i,3]< -perfis[i,2]*beta[perfis[[i,1]]]+epsilon[i]}}
colnames(perfis)=c("Roda", "Distância", "Desgaste")

```

```

perfis <- as.data.frame(perfis)
## Eliminando as medidas após ultrapassagem de Df
min <- vector(mode = "numeric", length = 50)
for (i in 1:50){
  min[i] <- 20
  for (j in 1:20){
    if(perfis$Desgaste[perfis$Roda==i][j]>=Df){
      if (j <= min[i]){
        min[i] <- j}}}}
  for (i in 1:50){
    if (min[i]!= 20){
      for (j in (min[i]+1):20)
        perfis$Desgaste[perfis$Roda==i][j]<- NA }}
    PerfisSim[k] <- perfis[,3]}
## Salvando os perfis simulados
write.table(PerfisSim, file = "perfisSim.csv", sep = ",",
col.names = FALSE,row.names=T)
return(PerfisSim)}
# Gerando B perfis simulados
Geração(B)
# #
# Estimando os parâmetros iniciais, otimizando a verossimilhança
# e exibindo resultados
# #
recover <- read.table(file="perfisSim.csv",sep="," ,head=F)
recover <- recover[,-1]

```

```

Boot <- function(B){
Resultado <- matrix(0,B,8)
for (k in 1:B){
perfis <- matrix(c(rep(1:50,each=20),rep(seq(0,38000,2000),50),
rep(0,2000)),ncol=4, byrow=F)
perfis[,3] <- recover[,k]
if(sum (is.na(perfis[,3]))!=0){
perfis <- perfis[-(which(is.na(perfis[,3]))),]
colnames(perfis)=c("Roda", "Distância", "Desgaste", "Baprox")
## Cálculo de  $b_i \wedge ap$  através de ajuste modelo linear para cada unidade
mod <- lmList(Desgaste~ Distância-1|Roda,perfis)
FitAprox <- array(coef(mod)[,1])
## Estimadores de máxima verossimilhança para mu e sigma
hatmediaap <- sum(log(FitAprox))/length(FitAprox)
hatdesvioap <- (sum((log(FitAprox)-hatmediaap)^2)/length(FitAprox))
## Estimador para  $\sigma^2$  erro
Fitaprox <- c(rep(0,length(perfis[,4])))
Erroaprox <- c(rep(0,length(perfis[,4])))
perfis <- cbind(perfis,Fitaprox,Erroaprox)
for(i in 1:length(perfis[,4])){
j=perfis[i,1]
perfis[i,4] <- FitAprox[j] }
perfis <- as.data.frame(perfis)
perfis$Fitaprox <- perfis$Distância*perfis$Baprox
perfis$Erroaprox <- perfis$Desgaste-perfis$Fitaprox
hatsigmaeap=var(perfis$Erroaprox)

```

```

Estimadores=matrix(c(hatmediaap,sqrt(hatdesvioap),sqrt(hatsigmaeap)), ncol=3)
colnames(Estimadores)=c("Scale", "Location", "Sigma Erro")

Estimadores

# #
# Verossimilhança transformada e otimização
# #
## Formatando os dados como lista
Aux=matrix(c(rep(0,1020)),ncol=51)
dimnames(Aux)=list(seq(1:20),c("X",paste("T",seq(1:50),sep=)))
Aux[,1] <- seq(0,38000,2000)
for (j in 2:51){
for (i in 1:length(which(perfis[,1]==j-1))){
Aux[i,j] <- perfis[which(perfis[,1]==j-1)[i],3] }
for (i in 2:20){ for (j in 2:51){ if (Aux[i,j]==0){
Aux[i,j]<- NA} else {Aux[i,j]<-Aux[i,j]}}
Aux <- as.data.frame(Aux)
Dados <- na.omit(reshape(Aux,idvar="id",varying=list(2:51),
v.names="y",direction="long"))
names(Dados) <- c("x", "roda", "y", "id")
Dados <- Dados[,c("y", "x", "roda")]
# Criando lista
Y <- vector("list",length=50)
for(j in 1:50){ Y[[j]] <- Dados[Dados$roda==j,1:2]}
# #
# Calculando os estimadores de máxima verossimilhança
# #

```

```

otimo <- function(par) Loglik(data=Y,mu=par[1], sigma=par[2],sigmae=par[3])
estim <- try(constrOptim(c(Estimadores[1],Estimadores[2],
Estimadores[3]),NULL, f=otimo,ui=matrix(c(-1,0,0,0,1,0,0,0,1),ncol=3),
ci=0,control=list(fnscale=-1)),silent=TRUE)

## Armazenando o resultado para cada conjunto de perfis
Resultado[k,1] <- round(Estimadores[1],4)
Resultado[k,2] <- round(Estimadores[2],7)
Resultado[k,3] <- round(Estimadores[3],9)
if(class(try(round(Loglik(data = Y, mu=Estimadores[1],
sigma=Estimadores[2],sigmae=Estimadores[3]),4), silent=TRUE))!= "try-error"){
Resultado[k,4] <- round(Loglik(data = Y, mu=Estimadores[1],
sigma=Estimadores[2],sigmae=Estimadores[3]),4)}
if(class(estim)!= "try-error"){
Resultado[k,5] <- round(estim$par[1],4)
Resultado[k,6] <- round(estim$par[2],7)
Resultado[k,7] <- round(estim$par[3],9)
Resultado[k,8] <- round(estim$value,4)}
} return(Resultado) }

```

B.3 Dados rodas de trens - distribuição Weibull

```

rm(list=ls())
require(nlme)
# Lendo dados
perfis <- read.table(file="rodas.csv",sep=";",head=T)
# # Calculando os valores iniciais para os parâmetros, que

```

```

# # servirão de valores de partida para a maximização da verossimilhança
# #
# Cálculo de bi
mod <- lmList(Desgaste~Distância-1|Roda,perfis)
FitAprox <- array((1/coef(mod))[,1]) # bis aproximados
# Estimadores de máxima verossimilhança para alfa e delta
VerWeib <- function(y,alphaap)
((sum(y^ alphaap*log(y)))/(sum(y^ alphaap))
- (1/alphaap)-(sum(log(y))/dim(FitAprox)))^2
alpha=optimize(VerWeib, c(0,100), tol = 0.00001, y=FitAprox)
hatalphaap=alpha$minimum
hatdeltaap=(1/dim(FitAprox) *sum(FitAprox^hatalphaap))^(1/hatalphaap)
# Estimador para sigma^2
Baprox <- c(rep(0,length(perfis[,3])))
Fitaprox <- c(rep(0,length(perfis[,3])))
Erroaprox <- c(rep(0,length(perfis[,3])))
perfis <- cbind(perfis,Baprox,Fitaprox,Erroaprox)
for(i in 1:length(perfis[,3]))
{ j=perfis[i,1] perfis[i,4] <- FitAprox[j] }
perfis$Fitaprox=perfis$Distância/perfis$Baprox
perfis$Erroaprox = perfis$Desgaste-perfis$Fitaprox
hatsigma2ap=var(perfis$Erroaprox)
Estimadores=matrix(c(hatalphaap,hatdeltaap,hatsigma2ap),ncol=3)
colnames(Estimadores)=c("alpha", "delta", "sigma2")
Estimadores
# #

```

```

# # Formatando a verossimilhança e preparando a otimização
# #
# Formatando os dados
Aux=matrix(c(rep(0,195)),ncol=15)
dimnames(Aux)=list(seq(1:13), c("X",paste("T",seq(1:14),sep=)))
Aux[,1] <- seq(0,600,50)
for (j in 2:15) {
for (i in 1:length(which(perfis[,1]==j-1))) {
Aux[i,j] <- perfis[which(perfis[,1]==j-1)[i],3] } }
for (i in 2:13) {
for (j in 2:15) {
if (Aux[i,j]==0) {Aux[i,j]<- NA else Aux[i,j]<-Aux[i,j]}
} }
Aux <- as.data.frame(Aux)
Dados <- na.omit(reshape(Aux,idvar="id",varying=list(2:15), v.names="y",direction="long"))
names(Dados) <- c("x","roda","y","id")
Dados <- Dados[,c("y","x","roda")]
# Criando lista
Y <- vector("list",length=14) for(j in 1:14){
Y[[j]] <- Dados[Dados$roda==j,1:2]}
# Função que calcula a log-verossimilhança
joint <- array(list(NULL), c(600,1))
acum=matrix(c(rep(0,14)),ncol=14)
Loglik <- function(data, alpha, delta, sigma){
int <- function(yx, a, d, s){
y <- yx[,1]
t <- yx[,2]
for (p in 1:600){
joint[p] <- exp( sum(dnorm(y, mean=t/qweibull(pnorm((p/100-3),log=TRUE),a,d,log=TRUE),

```

```

sd=s, log=TRUE)) + dnorm((p/100-3), log=TRUE) )
} Joint <- Vectorize(joint)
return(Joint) }
lista <- lapply(data, int, a=alpha, d=delta, s=sigma)
for (i in 1:14){
acum[,i] <- log(sum(unlist(as.vector(lista[i]))) * 0.01) }
loglik <- acum
return(loglik) }
# Calculando a log-verossimilhança com valores iniciais
Loglik(data = Y, alpha=hatalphaap, delta=hatdeltaap, sigma=sqrt(hatsigma2ap))
# Rotina que calcula estimadores de máxima verossimilhança
otimo <- function(par) Loglik(data=Y, alpha=par[1], delta=par[2], sigma=par[3])
estim <- constrOptim(c(Estimadores[1], Estimadores[2], sqrt(Estimadores[3])),
NULL, f=otimo, ui=diag(3), ci=0, control=list(fnscale=-1))
estim$par # Exibição dos resultados
# Calculando a log-verossimilhança com estimadores de MV
Loglik(data = Y, alpha=estim$par[1], delta=estim$par[2], sigma=estim$par[3])

```

B.4 Bootstrap paramétrico e não paramétrico para rodas trens

```

# As duas rotinas geram um arquivo chamado perfisSim.csv
# no qual são armazenados as amostras geradas. A partir deles usa-se
# o mesmo script do anexo B3 para se obter as estimativas bootstrap
# #
# Bootstrapping paramétrico
# #
Bootp <- function(B)
PerfisSim <- matrix(0,182,B)
for (k in 1:B)
betaboop <- rweibull(14,estim$par[1],estim$par[2])

```

```

epsilonboop <- rnorm(182,0,estim$par[3])
perfisboop <- matrix(c(rep(1:14,each=13),rep(seq(0,600,50),14),
rep(0,182)),ncol=3, byrow=F)
for(i in 1:182) if (perfisboop[i,2]==0)
perfisboop[i,3]<-0
Else
if (perfisboop[i,2]/betaboop[perfisboop[[i,1]]]+epsilonboop[i]>0)
perfisboop[i,3]<-0
Else
perfisboop[i,3]<-perfisboop[i,2]/betaboop[perfisboop[[i,1]]]+ epsilonboop[i]
colnames(perfisboop)=c("Roda","Distância","Desgaste")
perfisboop <- as.data.frame(perfisboop)
min <- vector(mode = "numeric", length = 14)
for (i in 1:14)
min[i] <- 13
for (j in 1:13)
if(perfisboop$Desgaste[perfisboop$Roda==i][j]<=77)
if (j <= min[i])
min[i] <-j
for (i in 1:14)
if (min[i]!= 13)
for (j in (min[i]+1):13)
Perfisboop$Desgaste[perfisboop$Roda==i][j]<- NA
PerfisSim[,k] <- perfisboop[,3]
write.table(PerfisSim, file = "perfisSim.csv", sep = ",",
col.names = FALSE,row.names=T)
return(PerfisSim)
Bootp(10000)
# #
# Bootstrapping não paramétrico
# #

```

```
recover <- matrix(0,182,10000)
x <- seq(1:14)
seq <- matrix(0,14,10000)
for (j in 1:10000)
seq[,j] <- sample(x,size=14,replace=TRUE)
write.table(t(seq),file = "perfisnaopar.csv", sep = ",",col.names = FALSE)
perfiscomp <- read.table(file="Rodascomp.csv",sep=";",head=T)
for (a in 1:10000)
for (c in 1:14)
for (d in 1:13)
recover[(13*(c-1))+d,a] <-
perfiscomp$Desgaste[perfiscomp$Roda==seq[c,a]][d]
write.table(perfiscomp, file = "perfisSim.csv", sep = ",",
col.names = FALSE,row.names=T)
```