

Universidade Federal de Minas Gerais  
Programa de Pós-Graduação  
em Estatística

# Modelos robustos para degradação linear e tempo de falha

**Rívert Paulo Braga Oliveira**

*lamabarista@gmail.com*

**Orientadora: Profa. Dra. Rosângela Helena Loschi**

*loschi@est.ufmg.br*

**Co-orientadora: Profa. Dra. Marta Afonso Freitas**

*marta@dep.ufmg.br*

**Belo Horizonte  
30 de Novembro de 2015**



Rívert Paulo Braga Oliveira

# Modelos robustos para degradação linear e tempo de falha

Tese apresentada ao Departamento de Pós-Graduação em Estatística da Universidade Federal de Minas Gerais, para a obtenção do Título de Doutor em Estatística

Orientadora: Rosângela Helena Loschi

Co-orientadora: Marta Afonso Freitas

**Belo Horizonte**  
**30 de novembro de 2015**

## Agradecimentos

Agradeço a todos que direta ou indiretamente me apoiaram neste trabalho.

A minha orientadora, Dr<sup>a</sup>. Rosângela Helena Loschi pela orientação, paciência e dedicação à elaboração da tese, pelo esmero com a excelência do trabalho sempre me ensinando a ser um melhor pesquisador.

A minha co-orientadora, Dr<sup>a</sup>. Marta Afonso Freitas pela orientação e iniciação na minha linha de pesquisa favorita.

A ambas por acreditarem na minha capacidade. Terão sempre minha eterna gratidão, admiração, respeito e carinho.

Aos meus pais, José Antônio e Valdelice; meus sogros, Deusdedit e Maria Auxiliadora; meus irmãos, Rodrigo e Rony; minhas cunhadas Marinalva e Yara; meus cunhados Luís e Will; pela força, dedicação, apoio e carinho.

A minha esposa Carolina, pelo companheirismo, força, incentivo, dedicação, carinho e compreensão.

Aos amigos do Objetivo, do CEFET-MG, da Estatística-UFMG, do mestrado do PPGEF-UFMG, do doutorado do DEST-UFMG, da minha infância e demais amigos que fiz ao longo dos anos.

Aos professores do programa de Pós Graduação em Estatística, pelo conhecimento transmitido.

A Rogéria, Rose, Maísa, Wilton, pelo carinho e ótimo trabalho que desempenham para o departamento, e aos demais funcionários da UFMG.

A UFMG pelo suporte e estrutura.

Ao Centro Universitário Metodista Izabela Hendrix que me permitiu descobrir a beleza e o prazer da atividade de docência, e também pelo apoio da equipe de professores e coordenação dos cursos.

A CAPES, CNPq e FAPEMIG, pelo suporte financeiro.

## *Resumo*

Em estatística a confiabilidade é um ramo que visa descrever a distribuição do tempo de falha de objetos de interesse.

Para determinadas situações, nas quais a ocorrência de falhas é pequena ou quase nula, a estimação das quantidades que descrevem os tempos de falha fica comprometida. Dessa forma foram desenvolvidos os modelos gerais de degradação, que possuem como dado experimental não a falha, mas sim alguma característica mensurável a ela atrelada. Quando tal característica é monitorada tornam-se possíveis melhorias significativas nas estimativas das quantidades de interesse.

Neste texto introduzimos classes flexíveis de modelos de tempo de falha e degradação capazes de acomodar comportamentos assimétricos e caudas pesadas.

Para os modelos de degradação assumimos que a taxa e a taxa recíproca de degradação seguem distribuições em ambas as classes de distribuição, as de famílias de mistura de escala da distribuição normal assimétrica e log normal assimétrica. No que concerne aos modelos de análise convencional de tempo de falha também consideramos as mesmas famílias. Para ambas as modelagens introduzimos algoritmos com base na técnica de aumento de dados para amostrar das distribuições *a posteriori* dos parâmetros.

As novas metodologias são aplicadas a bancos de dados simulados com características diversas, considerando presença ou ausência de assimetria, caudas pesadas e caudas leves. A modelagem de situações práticas conhecidas da literatura também é explorada, a saber dados de degradação de rodas de trem e lasers, e dados de tempo de vida em pacientes com câncer de pulmão. Os resultados da nova modelagem são comparados com as abordagens vigentes da modelagem geral de degradação e de tempos de falha convencional, e devido às características de flexibilidade se mostram promissores inclusive em casos de má especificação.

O trabalho conta ainda com uma rica discussão sobre modelos gerais de degradação e modelos convencionais de tempo de falha em termos do paradigma bayesiano. Algumas provas formais e proposições também fornecem resultados teóricos interessantes ainda não explorados na literatura.

Por fim, a modelagem desenvolvida deixa vários problemas abertos e resultados para fomentar pesquisas futuras.

**Palavras-chave:** Aumento de dados, Degradação, Confiabilidade, Normal e Log Normal Assimétrica

## *Abstract*

In statistics the reliability is a branch that seeks to describe the time to failure distribution of objects of interest.

For certain situations where failures are not frequent or virtually nonexistent, the estimation of quantities describing the failure times is compromised. The general degradation models have been developed to overcome this problem by measuring a feature linked to failure, not the failure itself. When this feature is monitored it becomes possible to improve the estimates of the time to failure quantities of interest.

In this paper we introduce flexible classes of failure time models and degradation that are able to accommodate asymmetric behavior and heavy tails.

For the degradation path models this goal is achieved by assuming the degradation and reciprocal degradation rates have distributions in both classes of distributions, the scale mixture of skew-normal and scale mixture of log skew normal families of distributions. We also consider the same families to perform conventional time to failure analysis. For both models we build up algorithms based on the data augmentation technique in order to sample parameters from the posterior distribution.

The new methodologies are applied to simulated databases with different characteristics, considering the presence or absence asymmetry, heavy tails and light tails. The modeling of known literature practical situations is also explored, namely the train wheels and lasers degradation data, and the lung cancer lifetime data. The results of the new modeling are compared with current approaches to model degradation and conventional failure times. Due to its flexibility characteristics, the proposed methods show promise even in cases of misspecification.

The work also has a rich discussion on general degradation path models, and conventional failure time models, in terms of the Bayesian paradigm. Some formal proofs

and propositions also provide interesting theoretical results not yet explored in the literature.

Finally, the developed methodology leaves several open issues and results to encourage further research.

**Keywords:** Data augmentation, Degradation, Reliability, Scale Mixture, Skew and Log Skew Normal



# Lista de Figuras

1.1	Alguns perfis de desgaste de rodas de trem . . . . .	16
2.1	Densidades da distribuição skew normal padrão para diferentes valores de $\alpha$ . . . . .	23
2.2	Densidades da distribuição log skew normal padrão para diferentes valores de $\alpha$ . . . . .	25
3.1	Estimativa proposta por Hamada (2005) <i>versus</i> estimativa via distribuição preditiva . . . . .	46
4.1	Funções densidade de probabilidade da distribuição skew normal com $\mu_\beta = 5$ e $\omega_\beta^2 = 4$ para efeito aleatório e tempo de falha, com limiar de falha para a degradação $D_f = 20$ , assumindo-se diferentes valores de $\alpha_\beta$ - modelo $y_{ij} = \beta_i/t_{ij}$ . . . . .	58
4.2	Funções densidade de probabilidade da distribuição skew-t com $\mu_\beta = 40$ , $\omega_\beta^2 = 4$ e $\alpha_\beta = \pm 3$ para efeito aleatório e tempo de falha, com limiar de falha para a degradação $D_f = 20$ , assumindo-se diferentes valores de $\nu_\beta$ - modelo $y_{ij} = \beta_i/t_{ij}$ . . . . .	60

4.3	Funções densidade de probabilidade da distribuição log skew normal com $\mu_\beta = 5$ e $\omega_\beta^2 = 4$ para efeito aleatório e tempo de falha, com limiar de falha para a degradação $D_f = 20$ , assumindo-se diferentes valores de $\alpha_\beta$ - modelo $y_{ij} = \beta_i/t_{ij}$ . . . . .	62
4.4	Funções densidade de probabilidade da distribuição log skew-t com $\mu_\beta = 4$ , $\omega_\beta^2 = 4$ e $\alpha_\beta = \pm 100$ para efeito aleatório e tempo de falha, com limiar de falha para a degradação $D_f = 20$ , assumindo-se diferentes valores de $\nu_\beta$ - modelo $y_{ij} = \beta_i/t_{ij}$ . . . . .	63
4.5	Funções densidade de probabilidade da distribuição log skew normal com $\mu_\beta = 5$ e $\omega_\beta^2 = 4$ para efeito aleatório e tempo de falha, com limiar de falha para a degradação $D_f = 20$ , assumindo-se diferentes valores de $\alpha_\beta$ - modelo $y_{ij} = \beta_i t_{ij}$ . . . . .	66
4.6	Funções densidade de probabilidade da distribuição log skew-t com $\mu_\beta = 3$ , $\omega_\beta^2 = 0.04$ e $\alpha_\beta = \pm 3$ para efeito aleatório e tempo de falha, com limiar de falha para a degradação $D_f = 10$ , assumindo-se diferentes valores de $\nu_\beta$ - modelo $y_{ij} = \beta_i t_{ij}$ . . . . .	67
4.7	Funções densidade de probabilidade da distribuição log skew normal com $\mu_\beta = 5$ e $\omega_\beta^2 = 2$ , com limiar de falha para a degradação $D_f = 20$ , assumindo-se diferentes valores de $\alpha_\beta$ - modelo $y_{ij} = \beta_i t_{ij}$ . . . . .	70
4.8	Funções densidade de probabilidade da distribuição log skew-t com $\mu_\beta = 3$ , $\omega_\beta^2 = 0.04$ e $\alpha_\beta = \pm 3$ para efeito aleatório e tempo de falha, com limiar de falha para a degradação $D_f = 10$ , assumindo-se diferentes valores de $\nu_\beta$ - modelo $y_{ij} = \beta_i t_{ij}$ . . . . .	71
4.9	Perfis de degradação dos dados artificiais . . . . .	89
4.10	Comparação da distribuição <i>a posteriori</i> dos parâmetros com os valores reais no ajuste SN: dados artificiais com efeito aleatório Skew Normal . . . . .	91

4.11	Comparação da distribuição <i>a posteriori</i> dos parâmetros com os valores reais no ajuste ST: dados artificiais com efeito aleatório Skew T . . . . .	92
4.12	Comparação da distribuição <i>a posteriori</i> dos parâmetros com os valores reais no ajuste LST: dados artificiais com efeito aleatório Log Skew T . . . . .	93
4.13	Comparação da distribuição <i>a posteriori</i> dos parâmetros com os valores reais no ajuste W: dados artificiais com efeito aleatório Weibull . . . . .	94
4.14	Função densidade de probabilidade <i>a posteriori</i> dos tempos de falha para uma unidade futura, pelos métodos de Hamada (2005) (esquerda) e proposta pela preditiva a posteriori (direita), para todos os modelos ajustados aos dados B-SN (primeira linha), B-ST (segunda linha), B-LST (terceira linha) e B-W (quarta linha) . . . . .	96
4.15	Função distribuição acumulada <i>a posteriori</i> dos tempos de falha para uma unidade futura, pelos métodos de Hamada (2005) (esquerda) e proposta pela preditiva a posteriori (direita), para todos os modelos ajustados aos dados B-SN (primeira linha), B-ST (segunda linha), B-LST (terceira linha) e B-W (quarta linha) . . . . .	97
4.16	Função densidade de probabilidade dos tempos de falha para uma unidade futura pelo método proposto via preditiva a posteriori do modelo ajustado T, dados B-ST . . . . .	98
4.17	$CPO_i$ por degradação $y_{ij}$ (esquerda) e perfil $y_i$ (direita), nos cenários B-SN (primeira linha), B-ST (segunda linha), B-LST (terceira linha) e B-W (quarta linha). . . . .	115
4.18	$CPO_i$ por tempo de falha (esquerda) e pseudo tempo de falha (direita), nos cenários B-SN (primeira linha), B-ST (segunda linha), B-LST (terceira linha) e B-W (quarta linha). . . . .	116

4.19	$KL_i$ por degradação $y_{ij}$ (esquerda) e por perfil $\mathbf{y}_i$ (direita), nos cenários B-SN (primeira linha), B-ST (segunda linha), B-LST (terceira linha) e B-W (quarta linha). . . . .	119
4.20	$p_i$ por degradação $y_{ij}$ (esquerda) e por perfil $\mathbf{y}_i$ (direita), nos cenários B-SN (primeira linha), B-ST (segunda linha), B-LST (terceira linha) e B-W (quarta linha). . . . .	120
4.21	$KL_i$ por tempo de falha $t_i$ (esquerda) e por pseudotempo de falha $\hat{t}_i$ (direita), nos cenários B-SN (primeira linha), B-ST (segunda linha), B-LST (terceira linha) e B-W (quarta linha). . . . .	121
4.22	$p_i$ por tempo de falha $t_i$ (esquerda) e por pseudotempo de falha $\hat{t}_i$ (direita), nos cenários B-SN (primeira linha), B-ST (segunda linha), B-LST (terceira linha) e B-W (quarta linha). . . . .	122
4.23	Distribuição <i>a posteriori</i> $\beta_i \mathbf{y}$ sob todos os modelos, cenário B-SN. As linhas pretas horizontais indicam os valores reais $\beta_i$ . . . . .	125
4.24	Distribuição <i>a posteriori</i> $\beta_i \mathbf{y}$ sob todos os modelos, cenário B-ST. As linhas pretas horizontais indicam os valores reais $\beta_i$ . . . . .	126
4.25	Distribuição <i>a posteriori</i> $\beta_i \mathbf{y}$ sob todos os modelos, cenário B-LST. As linhas pretas horizontais indicam os valores reais $\beta_i$ . . . . .	127
4.26	Distribuição <i>a posteriori</i> $\beta_i \mathbf{y}$ sob todos os modelos, cenário B-W. As linhas pretas horizontais indicam os valores reais $\beta_i$ . . . . .	128
4.27	Croqui - vista principal inferior de um vagão motor. . . . .	131
4.28	Perfis de desgaste . . . . .	132
4.29	Função distribuição acumulada (esquerda) e densidade de probabilidade (direita) <i>a posteriori</i> dos tempos de falha para uma unidade futura, pelos métodos de Hamada (2005) (acima) e proposta pela preditiva a posteriori (embaixo), para todos os modelos ajustados aos dados da roda de trem MA31 . . . . .	134

4.30	Função distribuição acumulada (esquerda) e densidade de probabilidade (direita) <i>a posteriori</i> dos tempos de falha para uma unidade futura, pelos métodos de Hamada (2005) (acima) e proposta pela preditiva a posteriori (embaixo), para todos os modelos ajustados aos dados da roda de trem MA12 . . . . .	135
4.31	Função distribuição acumulada (esquerda) e densidade de probabilidade (direita) <i>a posteriori</i> dos tempos de falha para uma unidade futura, pelos métodos de Hamada (2005) (acima) e proposta pela preditiva a posteriori (embaixo), para todos os modelos ajustados aos dados da roda de trem MA11 . . . . .	136
4.32	Perfil de degradação das unidades emissoras laser. . . . .	145
4.33	Função distribuição acumulada (esquerda) e densidade de probabilidade (direita) <i>a posteriori</i> dos tempos de falha para uma unidade futura, pelos métodos de Hamada (2005) (acima) e proposta pela preditiva a posteriori (embaixo), para todos os modelos ajustados aos dados de laser . . . . .	146
5.1	Distribuições <i>a posteriori</i> dos parâmetros sob os modelos SN centrado (esquerda) e SN não centrado (direita), dados BT-SN1. . . . .	165
5.2	Distribuições <i>a posteriori</i> dos parâmetros sob os modelos SN centrado (esquerda) e SN não centrado (direita), dados BT-SN2. . . . .	167
5.3	Comparação da distribuição <i>a posteriori</i> dos parâmetros com os valores reais no ajuste LST: dados simulados log skew-t . . . . .	168
5.4	Confiabilidade empírica via Kaplan-Meyer (abscissa) versus modelos ajustados pela média <i>a posteriori</i> (ordenada) para os tempos de falha. N, LN e T (topo). Tempos LST. . . . .	170

5.5	Confiabilidade empírica via Kaplan-Meyer (abscissa) versus modelos ajustados pela distribuição preditiva <i>a posteriori</i> (ordenada) para os tempos de falha. N, LN e T (topo). Tempos LST. . . . .	171
5.6	Confiabilidade Real, empírica via Kaplan-Meyer, e distribuição preditiva <i>a posteriori</i> sob os modelos LSN, ST, LST e W para os tempos de falha de uma unidade futura da população . . . . .	172
5.7	Sobrevivência empírica via Kaplan-Meyer (abscissa) versus modelos ajustados pela média <i>a posteriori</i> (ordenada) para os tempos de vida. Dados de câncer de pulmão. . . . .	175
5.8	Sobrevivência empírica via Kaplan-Meyer (abscissa) versus modelos ajustados pela distribuição preditiva <i>a posteriori</i> (ordenada) para os tempos de vida. Dados de câncer de pulmão. . . . .	176
5.9	Sobrevivência empírica via Kaplan-Meyer, e distribuição preditiva <i>a posteriori</i> sob os modelos LSN, LST e W, para os tempos de falha de uma unidade futura da população . . . . .	177

# Lista de Tabelas

3.1	Obtenção da amostra de <i>QIPOST</i> da distribuição preditiva <i>a posteriori</i> de T via Hamada (2005) . . . . .	42
3.2	Amostra de $F_{T_{n+1} \beta_{n+1},\sigma_{\varepsilon}^2}(\bullet)$ pelo método de Robinson e Crowder (2000)	48
4.1	Força da evidência sobre $FB(\mathcal{M}_1, \mathcal{M}_2)$ . . . . .	82
4.2	Resumos <i>a posteriori</i> nos ajustes de mesma distribuição das degradações geradas . . . . .	94
4.3	Percentis da distribuição <i>a posteriori</i> dos tempos de falha usando a Distribuição Preditiva <i>a posteriori</i> e o método de Hamada (2005) em todos os modelos ajustados para o cenário B-SN . . . . .	100
4.4	Percentis da distribuição <i>a posteriori</i> dos tempos de falha usando a Distribuição Preditiva <i>a posteriori</i> e o método de Hamada (2005) em todos os modelos ajustados para o cenário B-ST . . . . .	101
4.5	Percentis da distribuição <i>a posteriori</i> dos tempos de falha usando a Distribuição Preditiva <i>a posteriori</i> e o método de Hamada (2005) em todos os modelos ajustados para o cenário B-LST . . . . .	102
4.6	Percentis da distribuição <i>a posteriori</i> dos tempos de falha usando a Distribuição Preditiva <i>a posteriori</i> e o método de Hamada (2005) em todos os modelos ajustados para o cenário B-W . . . . .	103

4.7	Melhores (topo) e piores (pé) modelos segundo os critérios LPML, DIC, WAIC, FB, baseados na degradação (Deg), tempos reais (T) e pseudo tempos de falha (PTF); e teste KS em relação à fda empírica (KS-FDA EMP), fda real (KS-FDA REAL), fda real com gradeamento do eixo do tempo (KSG) . . . . .	107
4.8	Modelos ordenados do melhor (topo) para o pior (pé) segundo LPML, DIC, WAIC, Fator de Bayes (FB), Teste de Kolmogorov-Smirnov (KS), baseados na degradação (Deg) e pseudo tempos de falha (PTF), Rodas de Trem . . . . .	139
4.9	Percentis da distribuição <i>a posteriori</i> dos tempos de falha para a Roda MA31 . . . . .	141
4.10	Percentis da distribuição <i>a posteriori</i> dos tempos de falha para a Roda MA12 . . . . .	142
4.11	Percentis da distribuição <i>a posteriori</i> dos tempos de falha para a Roda MA11 . . . . .	144
4.12	Modelos ordenados do melhor (topo) para o pior (pé) segundo LPML, DIC, WAIC, Fator de Bayes (FB), Teste de Kolmogorov-Smirnov (KS), baseados na degradação (Deg) e pseudo tempos de falha (PTF), Dados de Laser . . . . .	147
4.13	Percentis da distribuição <i>a posteriori</i> dos tempos de falha para os emissores de laser . . . . .	148
5.1	Percentis de ordem 5%, 10%, 25%, 50%, 75%, 90% para os tempos de falha reais simulados do modelo LST, as estimativas segundo o estimador de Kaplan-Meyer, distribuição preditiva <i>a posteriori</i> e média <i>a posteriori</i>	173



5.2	Percentis de ordem 5%, 10%, 25%, 50%, 75%, 90% para os pacientes de câncer de pulmão inoperável, segundo o estimador de Kaplan-Meier, distribuição preditiva <i>a posteriori</i> e média <i>a posteriori</i> . . . . .	178
C.1	Medidas de qualidade de ajuste de modelos baseadas nas degradações $y_{ij}$ e nos perfis $y_i$ , dados artificiais . . . . .	218
C.2	Medidas de qualidade de ajuste de modelos baseadas nos tempos de falha (PTF) reais, dados artificiais . . . . .	219
C.3	Medidas de qualidade de ajuste de modelos baseadas nos pseudotempos de falha $\hat{t}_i$ , dados artificiais . . . . .	220
C.4	Fator de Bayes baseado nas degradações $y_{ij}$ e nos perfis $y_i$ , dados artificiais	221
C.5	Fator de Bayes baseado nos tempos de falha $t_i$ , dados artificiais . . . . .	222
C.6	Fator de Bayes baseado nos pseudo tempos de falha $\hat{t}_i$ , dados artificiais . . . . .	223
C.7	Teste de Kolmogorov-Smirnov baseado nos tempos $t_i$ e pseudotempos de falha $\hat{t}_i$ : comparação com a distribuição empírica no cenário B-SN . . . . .	224
C.8	Teste de Kolmogorov-Smirnov baseado nos tempos $t_i$ e pseudotempos de falha $\hat{t}_i$ : comparação com a distribuição empírica no cenário B-ST . . . . .	224
C.9	Teste de Kolmogorov-Smirnov baseado nos tempos $t_i$ e pseudotempos de falha $\hat{t}_i$ : comparação com a distribuição empírica no cenário B-LST . . . . .	225
C.10	Teste de Kolmogorov-Smirnov baseado nos tempos $t_i$ e pseudotempos de falha $\hat{t}_i$ : comparação com a distribuição empírica no cenário B-W . . . . .	225
C.11	Teste de K-S comparando as distribuições estimadas e a real baseado nos tempos de falha e pseudotempos de falha, todos os cenários . . . . .	226
C.12	Distância KS com particionamento dos espaço amostral dos tempos de falha para comparação dos modelos ajustados com a distribuição real, todos os cenários . . . . .	226

C.13	Qualidade de ajuste baseado na degradação (Deg) e pseudotempos de falha (PTF), Rodas de Trem . . . . .	227
C.14	Fator de Bayes baseado na degradação (Deg) e pseudo tempos de falha (PTF), Roda MA31 . . . . .	228
C.15	Fator de Bayes baseado na degradação (Deg) e pseudo tempos de falha (PTF), Roda MA12 . . . . .	228
C.16	Fator de Bayes baseado na degradação (Deg) e pseudo tempos de falha (PTF), Roda MA11 . . . . .	229
C.17	Teste de Kolmogorov-Smirnov baseado nos pseudotempos de falha $\hat{t}_i$ para as Rodas MA31, MA12 e MA11 . . . . .	229
C.18	Qualidade de ajuste baseado na degradação (Deg) e pseudo tempos de falha (PTF), Dados de Laser . . . . .	230
C.19	Fator de Bayes baseado na degradação (Deg) e pseudo tempos de falha (PTF), Dados de Laser . . . . .	230
C.20	Teste de Kolmogorov-Smirnov baseado nos pseudo tempos de falha $\hat{t}_i$ para os Dados de Laser . . . . .	231

# Lista de Quadros

4.1	Cenários dos dados artificiais de degradação . . . . .	86
-----	--	----

# Sumário

<b>1</b>	<b>Introdução</b>	<b>1</b>
1.1	Literatura . . . . .	1
1.2	O modelo geral de degradação e a estimação dos parâmetros . . . . .	7
<b>2</b>	<b>Algumas famílias de distribuições assimétricas</b>	<b>21</b>
2.1	Algumas distribuições de probabilidade . . . . .	22
2.1.1	A família de distribuições skew normal univariada . . . . .	22
2.1.2	A família de distribuições log skew normal univariada . . . . .	24
2.1.3	Famílias de distribuições de mistura na escala da skew normal univariada . . . . .	25
2.1.4	Famílias de distribuições de mistura na escala da log skew normal univariada . . . . .	27
<b>3</b>	<b>Modelos gerais de degradação: inferência sobre a distribuição dos tempos de falha</b>	<b>30</b>
3.1	Inferência clássica . . . . .	31
3.1.1	O método aproximado . . . . .	31
3.1.2	O método analítico . . . . .	32
3.1.3	O método de dois estágios . . . . .	34
3.1.4	O método numérico . . . . .	37

3.2	Inferência bayesiana . . . . .	40
3.2.1	Método proposto por Hamada (2005) . . . . .	41
3.2.2	Método de Robinson e Crowder (2000) . . . . .	47
3.2.3	Distribuição preditiva <i>a posteriori</i> de $T_{n+1}$ . . . . .	49
3.2.4	Nova unidade na verossimilhança via aumento de dados . . . . .	51
3.2.5	Interpretação nas abordagens bayesianas . . . . .	52
<b>4</b>	<b>Modelos de degradação robustos</b>	<b>55</b>
4.1	Método analítico . . . . .	56
4.1.1	Modelo linear com efeito aleatório SMSN e SMLSN . . . . .	56
4.1.2	Outro modelo robusto de degradação linear . . . . .	64
4.2	Inferência sobre os parâmetros do modelo linear de degradação . . . . .	69
4.3	Crêterios e medidas: seleçãõ de modelos . . . . .	79
4.3.1	Ordenada preditiva condicional (CPO) . . . . .	79
4.3.2	Deviance information criterion (DIC) . . . . .	80
4.3.3	Watanabe-Akaike information criterion (WAIC) . . . . .	81
4.3.4	Fator de bayes . . . . .	81
4.3.5	Teste de Kolmogorov-Smirnov (KS) . . . . .	82
4.3.6	Diagnóstico de influênciã via divergênciã de Kullback-Leibler (KL) . . . . .	83
4.3.7	Seleçãõ de modelos em modelos de degradaçãõ . . . . .	84
4.4	Análise de dados de degradaçãõ artificiais . . . . .	85
4.4.1	Geraçãõ dos perfis artificiais e especificações para inferênciã bayesiana . . . . .	85
4.4.2	Resultados . . . . .	89
4.4.3	Conclusões . . . . .	118
4.5	Dados reais: modelos de degradaçãõ . . . . .	129
4.5.1	Estudo de caso 1: dados de desgaste de rodas de trem . . . . .	130



# Capítulo 1

## Introdução

### 1.1 Literatura

Grande parte da literatura em confiabilidade faz o uso de dados de tempo de falha (ou tempo de vida) em estudos nos quais o objetivo é caracterizar a confiabilidade de produtos. Esses dados são oriundos de ensaios de vida nos quais a medida observada é o tempo de falha. Neste texto, os termos “tempo de vida”, “tempo de falha” e “tempo até a falha” serão utilizados indiscriminadamente. É sabido, entretanto, que para produtos que possuem elevada confiabilidade, tais como componentes eletrônicos modernos ou mecânicos de alta durabilidade, é comum encontrar-se diante de uma situação na qual, ao final do experimento, poucas ou até mesmo nenhuma falha é observada. O resultado experimental finda num alto índice de censuras. Em casos como estes, em geral, os custos de realização dos ensaios são elevados e os mesmos acabam não fornecendo informações suficientes para a inferência estatística. Em algumas situações é possível implementar testes acelerados (ver Nelson, 1990), mas, mesmo nestes casos, o índice de censuras resultante pode ser elevado. Dados com esta natureza (alto índice de censuras) fornecem pouca informação a respeito da proporção de produtos que conseguem, por exemplo, operar além do período de garantia estipulado.

Por esta razão, a literatura específica na área de confiabilidade tem dado maior ênfase à utilização de dados de degradação como uma alternativa aos dados de tempo até a falha. Alguns exemplos importantes são os trabalhos de (Lu e Meeker, 1993; Tseng *et al.*, 1995; Lu *et al.*, 1996; Chiao e Hamada, 1996, 2000; Robinson e Crowder, 2000; de Oliveira e Colosimo, 2004; Hamada, 2005; Hong *et al.*, 2015). Nos testes de degradação, a variável observada não é o tempo de falha, mas uma medida de degradação de alguma característica de qualidade do produto de interesse tomada ao longo do tempo. Esta medida deve estar diretamente relacionada à falha. Além disto, a própria definição da falha deve estar associada a algum nível crítico pré-especificado da medida de degradação. Este nível crítico é denominado limiar de falha ou “*failure threshold*”. A justificativa para a utilização desse tipo de abordagem reside no fato de que muitas falhas são o resultado de um mecanismo de degradação em atuação para o qual existem características que se degradam com o tempo. Do ponto de vista da engenharia, alguns mecanismos de degradação comuns incluem fadiga, trincas, corrosão e oxidação. Para conduzir um teste (ensaio) de degradação, é necessário pré-especificar um limiar crítico de degradação ou limiar de falha, obter medidas da degradação em tempos pré-fixados e distintos e definir que a falha ocorre quando a quantidade de degradação para uma unidade sob teste excede àquele nível crítico. A principal vantagem da utilização de dados de degradação sobre os dados de tempos de falha, oriundos de testes de vida, é que a análise pode ser feita de maneira satisfatória, ainda que nenhuma falha tenha ocorrido, isto é, mesmo que nenhum dos perfis das unidades sob observação tenha atingido o limiar considerado para definir uma “falha”.

A análise de dados de degradação, assim como dos dados de tempo de falha, tem como objetivo estimar a distribuição do tempo até a falha ou características da mesma, tais como o  $p$ -ésimo quantil ( $t_p$ ), o tempo médio até a falha, ou a função de confiabilidade  $R(t)$ , por exemplo. Existem referências importantes abordando a utilização de dados de degradação para estimar a confiabilidade. Gertsbackh e Kordonskiy (1969), discutem



o problema de degradação sob o ponto de vista da engenharia. Os autores apresentam a distribuição de Bernstein, a qual descreve a distribuição do tempo até a falha para um modelo linear simples com intercepto e inclinação ambos aleatórios. Nelson (1981) discute uma situação especial na qual a medida de degradação é destrutiva, em que somente uma medida pode ser feita em cada item. Uma revisão detalhada sobre este tema pode ser encontrada em (Nelson, 1990, Capítulo 11), onde também se encontram exemplos importantes de aplicações e a descrição de idéias básicas de como se analisar certo tipo de dados de degradação.

Duas abordagens básicas para modelagem e análise de dados de degradação podem ser encontradas na literatura. Numa delas assume-se que a degradação é um processo aleatório no tempo. Dentro desta linha, Doksum (1991) utiliza um processo de Wiener para analisar dados de degradação. Tang e Chang (1995) modelam dados oriundos de ensaios de degradação acelerados não destrutivos como uma coleção de processos estocásticos. Whitmore e Schenkelberg (1997) modelam o processo de degradação como um processo de difusão de Wiener. Pettit e Young (1999) relacionam a degradação, modelada por um processo Wiener, com um limiar de falha, em que os resultados das estimativas sobre a distribuição do tempo de falha foram superiores a modelos que não consideram a degradação. Bagdonavicius e Nikulin (2001) e Bagdonavicius *et al.* (2010) modelam a degradação por um processo gama e incluem possíveis covariáveis dependentes do tempo. Lawless e Crowder (2004) inserem covariáveis e efeitos aleatórios num processo gama com aplicação para degradação e falha. Peng e Tseng (2013) consideram que a variabilidade entre unidades, com estrutura de dependência temporal, de um modelo linear de processo estocástico para a degradação pode ser descrita pela distribuição skew-normal, o que confere maior flexibilidade ao modelo e reduz problemas relacionados à má especificação da distribuição dos efeitos aleatórios na presença de assimetria.

Uma outra abordagem comum para tratamento de dados de degradação é considerar

modelos estatísticos em que a degradação é descrita como uma função do tempo e, possivelmente, algumas outras variáveis aleatórias (por exemplo, fabricantes e materiais diversos). Tratam-se dos modelos gerais de perfis de degradação (“*general degradation path models*”). A análise de dados de degradação por esta abordagem é implementada em duas etapas. A primeira consiste na construção de um modelo misto, que pode ser linear ou não linear, que descreva o perfil de degradação ao longo do tempo, seguida da estimação dos parâmetros desconhecidos do mesmo. A segunda etapa consiste em utilizar as estimativas dos parâmetros para estimar a distribuição do tempo até a falha, bem como outras características de interesse a ela associadas.

Em geral, a análise é implementada em duas etapas. É importante ressaltar que dados de degradação possuem características de dados longitudinais (Fitzmaurice *et al.*, 2004). Portanto, a primeira etapa da análise é na realidade uma análise de dados longitudinais com resposta contínua, implementada através de um modelo não linear ou linear misto. A utilização de um modelo misto, isto é, que incorpora em sua forma funcional tanto efeitos aleatórios quanto fixos, pode, por sua vez, ser vista como uma modelagem em dois estágios, na qual primeiramente os parâmetros não observáveis (os efeitos aleatórios) são amostrados de uma distribuição que por si só, possui parâmetros (fixos) desconhecidos. Em seguida (2º. Estágio) os dados observados no ensaio são 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). Uma vez que uma família paramétrica para os efeitos aleatórios é escolhida, é preciso estimar os parâmetros fixos do modelo, quais sejam os parâmetros que indexam a distribuição dos efeitos aleatórios e outros parâmetros fixos que fazem parte da forma funcional do perfil de degradação.

De posse das estimativas dos parâmetros do modelo, a segunda etapa da análise dos dados de degradação consiste em estimar a distribuição do tempo até a falha  $F_T(t)$ . Se o modelo de degradação não é suficientemente simples para que se obtenha uma forma

fechada da distribuição do tempo de falha, uma estimativa pode ser obtida através de simulação de Monte Carlo ou algum outro procedimento computacional. Um bom exemplo é o artigo recente de Hong *et al.* (2015), em que os autores constroem um modelo de degradação com covariáveis dinâmicas no tempo e o comportamento destas covariáveis é modelada através de splines de forma restrita. A distribuição dos tempos de falha é obtida via simulação assumindo uma distribuição normal para os efeitos aleatórios.

Métodos de estimação baseados em inferência clássica ou bayesiana podem ser empregados para obter estimativas dos parâmetros desconhecidos nos modelos de degradação e, a partir delas, estimar a distribuição do tempo de falha. Um bom exemplo de estimação no contexto bayesiano é o artigo de Robinson e Crowder (2000).

Lu e Meeker (1993) desenvolvem métodos estatísticos utilizando dados de degradação para estimar a distribuição do tempo até a falha para uma classe ampla de modelos de degradação. Os autores consideram um modelo não linear misto (*NLME - Nonlinear Mixed-Effects Model*) e propõem um método de estimação em dois estágios para obter estimativas pontuais e intervalos de confiança para percentis da distribuição do tempo até a falha. O procedimento foi ilustrado com dados de evolução de trincas em ensaios de fadiga. Yacout *et al.* (1996) utilizam o procedimento proposto por Lu e Meeker (1993) em dados de degradação de elementos de combustível. Lu *et al.* (1996) estendem o modelo de degradação de Lu e Meeker (1993) e sugerem métodos de estimação baseados na função de verossimilhança. Lu *et al.* (1997) propõem um modelo com coeficientes de regressão aleatórios e uma função linear no tempo para o logaritmo do desvio padrão (o que significa assumir que o desvio padrão dos erros do modelo não é constante) para analisar dados de semicondutores com perfis lineares de degradação. Su *et al.* (1999) consideram um modelo de degradação com coeficientes e tamanho de amostra aleatórios e utilizam máxima verossimilhança para estimar os parâmetros. Os métodos foram aplicados para analisar um conjunto de dados de semicondutores. Wu e Shao (1999)

estudam as propriedades assintóticas dos estimadores de mínimos quadrados (ponderados) do modelo NLME. Os autores utilizam estas propriedades para obter estimativas pontuais e intervalos de confiança aproximados para percentis da distribuição do tempo de falha. Os métodos propostos são ilustrados com dados de comprimento de trincas em metais submetidos a ensaios de fadiga. Uma boa referência em modelos gerais de perfis de degradação é Meeker e Escobar (1998). Wu e Shao (2000) apresentam um método de estimação ponderado com base em agrupamento difuso (*fuzzy clustering*) para modificar o procedimento em dois estágios proposto por Lu e Meeker (1993). O método proposto e o de dois estágios de Lu e Meeker (1993) são ambos utilizados para analisar o banco de dados apresentado no trabalho de Wu e Shao (1999).

Outros trabalhos importantes no estudo de degradação são os trabalhos de Crk (2000), Jiang e Zhang (2002) e de Oliveira e Colosimo (2004). Neste último, os autores compararam os três métodos propostos por Meeker e Escobar (1998) (analítico, aproximado e numérico) analisando o conjunto de dados de desgaste de bandas de rodagem de pneus automotivos.

Os métodos analítico e aproximado aplicam-se melhor a situações nas quais o modelo de degradação é simples (por exemplo uma reta) com poucos parâmetros aleatórios. Já o método numérico é mais abrangente e permite utilizar modelos mais complexos (não lineares por exemplo) com parâmetros fixos e aleatórios. A grande dificuldade da abordagem por inferência clássica reside justamente na primeira etapa da análise dos dados, ou seja, na estimação dos parâmetros do modelo. Esta questão será melhor apresentada a seguir na Seção 1.2 e, acompanhada por trabalhos como o de Robinson e Crowder (2000) e Hamada (2005), servirá como um dos motes para opção pela inferência bayesiana.

## 1.2 O modelo geral de degradação e a estimação dos parâmetros

Em um ensaio de degradação, uma amostra de  $n$  unidades é colocada sob teste. Para cada uma delas observa-se, ao longo do tempo, uma característica relacionada ao mecanismo de falha de interesse. Em outras palavras, observa-se a “degradação” desta unidade através das medidas desta característica tomadas ao longo do tempo. A abordagem geral consiste em modelar os perfis de degradação das unidades individuais utilizando a mesma forma funcional e as diferenças entre as unidades amostrais são explicadas através da incorporação de efeitos aleatórios no modelo. Assim, para a unidade  $i$ , a verdadeira degradação no tempo  $t$  é dada por

$$Y_{it} = D(t; \boldsymbol{\alpha}; \boldsymbol{\beta}_i), \quad (1.1)$$

em que  $D(t; \boldsymbol{\alpha}; \boldsymbol{\beta}_i)$  assume uma forma funcional que depende de parâmetros e cuja natureza será explicitada mais adiante. Para uma dada unidade  $i$ , a falha ocorre quando a degradação atinge o limiar pré-especificado  $D_f$ . Assim, o tempo de falha da unidade  $i$  é o tempo  $T_i$  tal que

$$Y_{iT} = D_f = D(T_i; \boldsymbol{\alpha}; \boldsymbol{\beta}_i). \quad (1.2)$$

Entretanto, os dados observados nos ensaios de degradação para uma dada unidade  $i$  podem ser vistos como amostras do seu verdadeiro perfil, que é dado por (1.1), obtidas em  $m_i$  tempos  $t_{ij}$  pré-especificados, em geral  $j = 1, 2, \dots, m_i$ . Como estas medidas estão sujeitas a erros, a verdadeira degradação no tempo  $t_{ij}$ , isto é,  $D(t_{ij})$ , é observada com erro  $\varepsilon_{ij}$ .

Portanto, o modelo geral de degradação é dado por:

$$Y_{ij} = D(t_{ij}; \boldsymbol{\alpha}; \boldsymbol{\beta}_i) + \varepsilon_{ij}, \quad (1.3)$$

em que  $Y_{ij}$  é a variável aleatória que representa a  $j$ -ésima medida para a unidade  $i$ ,  $D(t_{ij}; \boldsymbol{\alpha}; \boldsymbol{\beta}_i)$  é o real perfil de degradação da unidade  $i$  em um tempo pré-especificado

$t_{ij}$ ,  $i = 1, 2, \dots, n$  e  $j = 1, 2, \dots, m_i$ , em que  $\boldsymbol{\alpha} = (\alpha_1, \alpha_2, \dots, \alpha_r)'$  denota o vetor de efeitos fixos que descreve as características da população a qual se mantém constante para todas as unidades.  $\boldsymbol{\beta}_i = (\beta_{i1}, \beta_{i2}, \dots, \beta_{ip})'$  é um vetor de efeitos aleatórios associado à unidade  $i$ , representando características individuais de cada unidade (por exemplo, variações em propriedades da matéria prima, no processo de produção, nas dimensões do componente, etc.), e  $\varepsilon_{ij}$  é o erro aleatório não observável (da medida) associado à unidade  $i$  no tempo  $t_{ij}$ .

A forma funcional de  $D(t_{ij}; \boldsymbol{\alpha}; \boldsymbol{\beta}_i)$  é, em geral, obtida a partir da análise empírica dos perfis de degradação dos processos sob estudo, mas sempre que possível deve ser baseada em algum fenômeno físico-químico associado à ele (Lu e Meeker, 1993). Assume-se, em geral, que os erros  $\varepsilon_{ij}$ 's, para  $i = 1, 2, \dots, n; j = 1, 2, \dots, m_i$ , são independentes e identicamente distribuídos (*iid*), seguindo uma distribuição normal com média zero e variância desconhecida  $\sigma_\varepsilon^2$ .

Entretanto, há trabalhos nos quais  $\sigma_\varepsilon^2$  varia com o tempo ou o nível de degradação (Lu *et al.*, 1997; Wakefield *et al.*, 1994). Se o processo de medição induz uma correlação serial entre os erros, então uma forma de covariância estruturada pode ser adotada. Neste trabalho será adotado o modelo geral (1.3) com erros *iid*  $N(0, \sigma_\varepsilon^2)$ .

Em modelos de degradação é costumaz assumir que os efeitos aleatórios  $\boldsymbol{\beta}_i$ 's, para  $i = 1, \dots, n$ , são independentes, com distribuição  $\Lambda_{\boldsymbol{\beta}|\boldsymbol{\theta}}(\boldsymbol{\beta}|\boldsymbol{\theta})$ , a qual depende de um vetor de parâmetros (fixos) desconhecidos  $\boldsymbol{\theta} = (\theta_1, \theta_2, \dots, \theta_q)'$ . O vetor de parâmetros  $\boldsymbol{\theta}$  também precisa ser estimado a partir dos dados de degradação. Além disso,  $\{\boldsymbol{\beta}_i\}$  é independente de  $\{\varepsilon_{ij}\}$  por suposição, para todo  $i$  e  $j$ . Assume-se também que  $Y$  e  $t$  estão em escalas transformadas, caso necessário. Por exemplo,  $Y$  e  $t$  podem estar em escala logarítmica.

A proporção de falhas no tempo  $t$  é equivalente à proporção de perfis de degradação que excedem o limiar crítico  $D_f$ , o qual é pré-definido, até o tempo  $t$ . Portanto, é possível definir a distribuição do tempo de falha  $T$  para o modelo (1.1) como sendo

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

quando as medidas de degradação são não decrescentes com o tempo, ou como sendo

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

quando as medidas de degradação são não crescentes com o tempo.

Repare que, a rigor,  $D(t;\alpha;\beta)$  não precisa ser não decrescente para que seja possível obter  $P[D(t;\alpha;\beta) \geq D_f|\alpha,\beta,\theta]$ , nem  $D(t;\alpha;\beta)$  carece ser não crescente para se obter  $P[D(t;\alpha;\beta) \leq D_f|\alpha,\beta,\theta]$ . Por exemplo, a acumulação de matéria orgânica numa turbina de uma usina hidrelétrica nem sempre é estritamente não decrescente, mas a turbina pára de funcionar caso essa matéria acumulada exceda um determinado nível crítico.

Para que se possa estimar, por exemplo, os quantis da distribuição do tempo de falha  $T$  com base neste modelo de degradação, é necessário estimar  $\alpha$ , vetor de efeitos fixos, e  $\theta = (\theta_1, \theta_2, \dots, \theta_q)'$  o vetor de parâmetros, também fixos, da distribuição dos efeitos aleatórios  $\Lambda_{\beta|\theta}(\beta|\theta)$ .

Para casos cujos perfis de degradação assumem formas funcionais simples, uma vez que  $\Lambda_{\beta|\theta}(\beta|\theta)$  é conhecida, a distribuição  $F_{T|\alpha,\beta,\theta}(t|\alpha,\beta,\theta)$  pode ser expressa, em grande parte das vezes, em forma fechada. Entretanto, em muitos casos, isto não é possível. Quando a forma funcional de  $D(t_{ij}; \alpha; \beta_i)$  é não linear e o modelo tem mais de um efeito aleatório (ou seja, quando o vetor  $\beta_i$  tem dimensão  $k > 1$ ), a obtenção em forma analítica de  $F_{T|\alpha,\beta,\theta}(t|\alpha,\beta,\theta)$  torna-se complicada. Em casos como este essa avaliação é feita de forma numérica. De maneira mais geral, é possível obter numericamente a distribuição de  $T$  para quaisquer  $\alpha$ ,  $\Lambda_{\beta|\theta}(\beta|\theta)$ ,  $D_f$  e  $D$  especificados, utilizando simulação de Monte Carlo (p. 330, Meeker e Escobar, 1998). Todavia, este procedimento só pode ser implementado se o vetor de parâmetros fixos  $\alpha$  e o vetor de parâmetros  $\theta$  (fixo) da distribuição dos efeitos aleatórios  $\Lambda_{\beta|\theta}(\beta|\theta)$  puderem ser estimados. Portanto, mesmo para uma dada distribuição  $\Lambda_{\beta|\theta}(\beta|\theta)$  o problema

continua sendo a estimação de  $\boldsymbol{\alpha}$  e  $\boldsymbol{\theta} = (\theta_1, \dots, \theta_q)'$ .

Lu e Meeker (1993) trabalham no problema de estimação de parâmetros e propõem o método de dois estágios para o caso no qual o vetor de efeitos aleatórios  $\boldsymbol{\beta}$ , ou alguma reparametrização apropriada do mesmo, segue uma distribuição normal multivariada (NM) com vetor de médias  $\boldsymbol{\mu}_\beta$  e matriz de variância-covariância  $\boldsymbol{\Sigma}_\beta$ . Em outras palavras, neste caso,  $\boldsymbol{\theta} = (\boldsymbol{\mu}_\beta, \text{vech}[\boldsymbol{\Sigma}_\beta]')$ , em que  $\text{vech}(A)$  de uma matriz  $n \times n$  simétrica  $A$  denota o vetor coluna  $n(n+1)/2 \times 1$  obtido pela vetorização da parcela triangular inferior da matriz  $A$ . Como a estimação por máxima verossimilhança dos parâmetros da distribuição dos efeitos aleatórios  $\boldsymbol{\mu}_\beta$  e  $\text{vech}[\boldsymbol{\Sigma}_\beta]'$  é, frequentemente, algebricamente complicada ou computacionalmente intensiva quando estes aparecem de forma não linear no modelo postulado para o perfil, Lu e Meeker (1993) propõem este método de dois estágios como uma alternativa aos métodos computacionalmente intensivos para estimá-los. Estudos por simulação mostram que os resultados obtidos pelo método de dois estágios são comparáveis aos dos métodos computacionalmente intensivos.

Posteriormente, a fim de tratar especificamente do problema inferencial para dados longitudinais, Pinheiro e Bates (1995) utilizam o método desenvolvido por Lindstrom e Bates (1990) para dados com medidas repetidas para obter uma estimativa de máxima verossimilhança dos parâmetros  $\boldsymbol{\mu}_\beta$ ,  $\boldsymbol{\Sigma}_\beta$  e  $\sigma_\varepsilon^2$  em modelos de efeitos mistos com suposição de normalidade para os efeitos aleatórios. As funções LME (*linear mixed-effects models*) e NLME escritas na linguagem S-PLUS, foram desenvolvidas para este objetivo (Pinheiro e Bates, 2000). Em outras palavras, estas funções são desenvolvidas para o caso específico de modelos cujos efeitos aleatórios seguem uma distribuição normal multivariada.

Neste texto revisaremos o método analítico, aproximado e o numérico. O método analítico é na verdade utilizado como uma forma de se obter a distribuição do tempo de falha a partir do conhecimento da distribuição dos efeitos aleatórios do modelo. Utilizando a forma funcional proposta para o verdadeiro perfil de degradação e o limiar



crítico que define a falha, a relação entre o tempo de falha e os efeitos aleatórios é obtida. Assim, supondo que se conhece a distribuição dos efeitos aleatórios chega-se analiticamente à distribuição do tempo até a falha  $F_T(t)$  através de transformação de variáveis. Os parâmetros da distribuição de  $F_T(t)$  podem ser escritos como funções do limiar crítico  $D_f$  e dos parâmetros da distribuição dos efeitos aleatórios. Entretanto, este procedimento torna-se razoavelmente complicado para formas funcionais mais complexas e nem sempre é possível chegar a uma forma fechada para a distribuição do tempo de falha. Além disso, mesmo que uma forma fechada seja encontrada e se obtenha a relação existente entre seus parâmetros e os da distribuição dos efeitos aleatórios, ainda é preciso estimar tais parâmetros a partir dos dados de degradação.

O método aproximado consiste em ajustar um modelo de degradação para o perfil de cada unidade separadamente e, a partir dele, estimar os respectivos tempos até a falha isto é, o tempo no qual cada perfil irá atingir o limiar  $D_f$ . Estas estimativas são denominadas “pseudo tempos de falha”. Em seguida uma análise de tempo de falha tradicional utilizando distribuições tais como Lognormal, weibull etc é implementada com base nestas quantidades. O que significa que os pseudo tempos de falha são tratados como se fossem os resultados observados do ensaio. As estimativas dos parâmetros da distribuição escolhida são obtidas por máxima verossimilhança. Este método é simples e tem apelo prático. Entretanto, ele pode subestimar a variabilidade das estimativas das quantidades de interesse como percentis por exemplo, visto que não se leva em conta na análise dos pseudo tempo de falhas, os erros nas estimativas dos mesmos. Tanto o método analítico como o aproximado são mais indicados para situações nas quais a forma funcional do modelo é razoavelmente simples (uma reta por exemplo e com um parâmetro aleatório).

Por outro lado, o método numérico é bastante abrangente e permite utilizar modelos mais complexos (não lineares por exemplo) com vários parâmetros fixos e aleatórios. A estimação dos parâmetros é feita pelo método de máxima verossimilhança. Todavia, as

rotinas disponíveis para a estimação dos parâmetros por tal método utilizam a suposição de que os efeitos aleatórios seguem uma distribuição normal multivariada com vetor de médias e matriz de variância-covariância desconhecidos (Pinheiro e Bates, 1995, 2000) (lembrando que estas rotinas lidam apenas com a parte do problema referente aos dados longitudinais da degradação e não com a obtenção da distribuição dos tempos de falha). No caso em que os efeitos aleatórios não têm distribuição normal, a obtenção das estimativas não é uma tarefa trivial. Meeker e Escobar (1998) apresentam o método numérico supondo a normalidade para os efeitos aleatórios e utilizam os algoritmos de (Pinheiro e Bates, 1995, 2000) para obtenção das estimativas dos parâmetros nos exemplos apresentados. Entretanto, para muitas situações práticas tais como as que serão tratadas neste trabalho, esta suposição não é razoável de ser assumida. Tal dificuldade computacional impõe uma restrição prática ao método numérico.

Dada a dificuldade de verificação da validade dessa suposição de normalidade, muitos trabalhos na literatura têm tentado estudar 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 tais como a dos modelos lineares generalizados de efeitos mistos - GLMM (*generalized linear mixed effects models*). Verbeke e Lesaffre (1997), Agresti *et al.* (2004), Litière *et al.* (2008), Alonso e Molenberghs (2008) e Freitas *et al.* (2009) são alguns exemplos. Em particular, com o objetivo de estudar o impacto de desvios da suposição de normalidade nas estimativas de percentis, Freitas *et al.* (2009) comparam, por simulação, os resultados obtidos usando os métodos numérico e aproximado com aqueles oriundos de uma análise tradicional de tempo até a falha, isto é, utilizando apenas o tempo no qual o perfil atinge o limiar  $D_f$  e as observações censuradas. Os autores utilizam um modelo de degradação linear simples no qual o efeito aleatório associado à cada perfil de degradação era oriundo das distribuições lognormal, normal e weibull, e também constam que o desempenho do método numérico é altamente afetado pela violação da suposição de

normalidade dos efeitos aleatórios, chegando este a ter o pior desempenho dentre os três métodos comparados. Toledo (2007) constata que o efeito maior desta má especificação é no vício relativo das estimativas obtidas pelo método numérico.

Em função da grande dificuldade na estimação dos parâmetros do modelo longitudinal em situações nas quais a suposição de normalidade não é razoável, e dos efeitos desta má especificação, uma outra frente de pesquisa foi aberta e cujo foco é o desenvolvimento de métodos computacionais que permitam obter estimadores de máxima verossimilhança para modelos de efeitos mistos, também utilizados na modelagem de dados de degradação, com efeitos aleatórios não-normais. Pinheiro *et al.* (2001) propõem 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 (Lee e Nelder, 1996) e a maximização por partes (Song *et al.*, 2005). Entretanto, a implementação destes métodos para utilização intensiva em análise de dados mostra-se uma tarefa desafiadora porque carece procedimentos específicos sempre que se varia a quantidade e distribuição dos efeitos aleatórios, e formas assumidas para modelar os dados longitudinais. A quadratura gaussiana é uma ferramenta que tem sido utilizada com sucesso na estimação de parâmetros em modelos com efeitos aleatórios normalmente distribuídos. Para modelos mistos com efeitos aleatórios que não são normalmente distribuídos, Nelson *et al.* (2006) propõem um método computacional que utiliza o método PIT (*Probability Integral Transformation*) para obter estimativas de máxima verossimilhança. Uma restrição desta abordagem é exigir que a inversa da função distribuição acumulada (fda) dos efeitos aleatórios tenha forma fechada, o que limita a sua aplicação. Liu e Yu (2008) propõem um novo método de estimação que reformula a função de verossimilhança condicional de efeitos aleatórios não-normais. Os autores reformulam a verossimilhança, dividindo e multiplicando a verossimilhança condicional por uma função densidade de probabilidade (fdp) normal. Os mesmos mostram que os resultados encontrados são

similares aos obtidos pelo método PIT. Além disso, há vantagens adicionais tais como a redução do esforço computacional e a possibilidade de considerar distribuições para os efeitos aleatórios cuja fda não tenha forma fechada.

Outra abordagem para o tratamento de dados de degradação é a utilização de métodos bayesianos. Nesta abordagem o tratamento dos efeitos aleatórios e fixos ocorre de maneira natural visto que são tratados como parâmetros. Em fármaco-cinética, métodos bayesianos são utilizados em (Wakefield *et al.*, 1994; Gelman *et al.*, 1996; Wakefield, 1996), entretanto só recentemente esta abordagem começou a ser utilizada em dados de degradação para fins de estimação da distribuição de tempos de falha. Robinson e Crowder (2000) fazem uma revisão dos métodos clássicos de inferência para modelos de degradação e apresentam uma abordagem bayesiana completa para o problema. A abordagem inclui a estimação da distribuição do tempo de falha de unidades futuras e para as unidades sob teste. No artigo os autores aplicam o enfoque bayesiano no modelo de degradação não linear utilizado por Lu e Meeker (1993) originário da Lei de Paris-Edorgan (e.g., Sobczyk e Spencer, 1992, Capítulo 1). O banco de dados utilizado foi primeiramente analisado por Hudak *et al.* (1978), e trata de medidas da evolução do tamanho da trinca em eixos. Os autores eliciam distribuições *a priori* não informativas no sentido de Jeffreys para os parâmetros da distribuição e hiperparâmetros da distribuição dos efeitos aleatórios, posteriormente, utilizam as distribuições marginais *a posteriori* dos mesmos para obter a distribuição dos tempos de falha das unidades sob teste e de uma unidade futura através da distribuição preditiva *a posteriori* dos tempos de falha. Hamada (2005) utiliza inferência bayesiana em dados de degradação associados a falhas em lasers. O autor considera um modelo linear de degradação com estrutura hierárquica, no qual uma distribuição *a priori* weibull é escolhida para os efeitos aleatórios e distribuições *a priori* gamma são utilizadas para os dois hiperparâmetros. Amostras da distribuição *a posteriori* dos parâmetros do modelo são obtidas através de métodos MCMC (Gamerman e Lopes, 2006, por exemplo). As distribuições

*a posteriori* de quantis da distribuição do tempo de falha  $T$ , bem como da confiabilidade *a posteriori*, também são obtidas. Para isso, condicional nos parâmetros, as expressões analíticas para o cálculo as quantidades  $p$ -ésimo quantil e a confiabilidade em um tempo  $t$ ,  $R(t)$ , da distribuição weibull, são aplicadas às amostras da distribuição *a posteriori* dos parâmetros do modelo, o que resulta numa amostra *a posteriori* de cada uma destas quantidades. Contudo, algumas inconsistências teóricas são observadas no referido artigo, uma vez que não é possível obter a prova analítica referente ao procedimento utilizado para a obtenção dos percentis *a posteriori* dos tempos de falha.

Diante desta constatação um dos objetivos da presente pesquisa é obter a distribuição preditiva *a posteriori* para o tempo de vida de uma nova unidade, baseada no modelo de degradação, e comparar os sumários *a posteriori* desta distribuição preditiva com aqueles propostos por Hamada (2005). Esta comparação será conduzida através de dados simulados e também dos dados dos emissores de laser e rodas de trem. Esta metodologia é mais coerente com o paradigma bayesiano para estimação da confiabilidade, e simples o bastante para aplicações imediatas por interessados que não dominem a compreensão da estatística a nível teórico. Alguns resultados preliminares que atendem parte desses objetivos podem ser encontrados em Oliveira (2011).

Outro trabalho interessante é o de Freitas *et al.* (2010), em que os autores analisam um banco de dados de desgaste de rodas de trem e comparam os resultados obtidos pelas abordagens bayesiana e clássica. No artigo apenas o desgaste da roda de uma única posição foi analisada, entretanto o banco de dados conta com oito posições de roda, com perfis de degradação de comportamentos variados (maiores detalhes na Seção 4.5.1). Estes comportamentos variados servem como motivação prática deste texto. Observando na Figura 1.1 a evolução da degradação das rodas das posições MA12, MA31 e MA41 notamos que alguns perfis se destacam dos demais, indicando que os dados poderiam ser melhor modelados por distribuições de caudas pesadas ou misturas de distribuições.

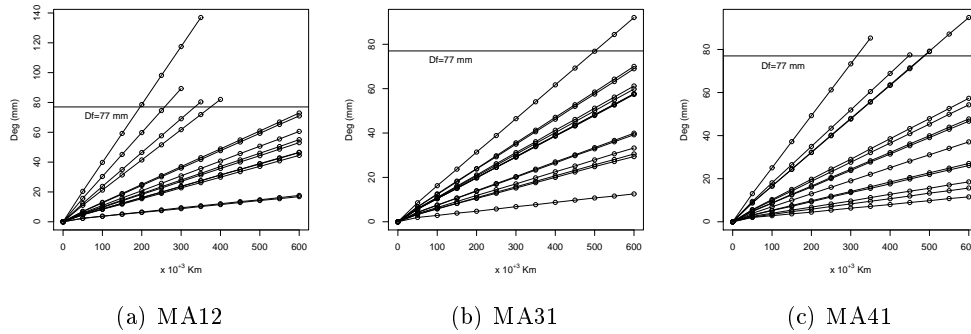


Figura 1.1: Alguns perfis de desgaste de rodas de trem

Partindo desta motivação prática, neste trabalho também tem-se como outro objetivo propor modelos de degradação mais flexíveis e capazes de acomodar simultaneamente características de assimetria e caudas pesadas. Para este fim, no caso de modelos de degradação, será assumido que os efeitos aleatórios têm seu comportamento aleatório descrito pelas famílias de mistura na escala de distribuições skew normais (SMSN, introduzidas por Branco e Dey, 2001) e pelas famílias log-SMSN introduzidas por Marchenko e Genton (2010). Em ambos os casos a relação entre degradação e o tempo de falha será fornecida analiticamente para duas formas funcionais lineares da degradação (optamos por lidar com modelos lineares porque abrangem grande variedade de situações práticas e também pela possibilidade de aplicar tanto a proposta de Hamada (2005) quanto Oliveira (2011)). Uma nova família de distribuições surgirá como consequência de um destes resultados, a qual é nomeada família de mistura na escala de distribuições skew normais recíprocas (SMRSN). As distribuições condicionais completas *a posteriori* dos objetos aleatórios de todos os modelos serão fornecidas de modo a possibilitar a implementação de procedimentos de amostragem via métodos MCMC. Vale ressaltar que tais distribuições condicionais completas *a posteriori* são unificadas para uma série de modelos lineares de degradação possíveis dentro das famílias SMSN e log-SMSN. Com o intuito de disseminar a utilização da nova metodologia os modelos

são implementados no software WinBugs 1.4 (Spiegelhalter *et al.*, 2000), além disso adiciona-se a distribuição *a priori* de Jeffreys independente para os graus de liberdade apresentada por Fonseca *et al.* (2008). Através de vários bancos de dados artificiais as medidas de qualidade de ajuste são estudadas tanto com foco na degradação quanto nos tempos (pseudo tempos) de falha. Por fim utiliza-se a modelagem para uma série de situações práticas para dados de desgaste de rodas de trem e de emissores de laser. Toda a modelagem apresentada é facilmente extensível para formas funcionais de degradação, lineares ou não, se abordada conforme proposta em Oliveira (2011). Nesta tese, a referida proposta ainda ganha uma série de justificativas teóricas mais rigorosas e a de inclusão de um objeto futuro na verossimilhança via aumento de dados, o que generaliza a obtenção da distribuição preditiva *a posteriori* dos tempos de falha para qualquer forma funcional de degradação, mesmo que não se conheça a degradação inicial. É importante ressaltar a inovação do trabalho. A utilização de distribuições assimétricas para dados de degradação foi utilizada no artigo recente de Peng e Tseng (2013), em que o autor modela a degradação por um processo estocástico de movimento browniano combinado com incremento linear de coeficiente aleatório skew normal. No entanto, até o momento não temos conhecimento de qualquer outro trabalho que considere modelos de degradação assumindo distribuições de caudas pesadas ou assimétricas de caudas pesadas.

Modelos robustos (no sentido de lidar com observações atípicas através de distribuições de caudas pesadas) e mais flexíveis para modelar tempos de falha convencionais também serão propostos. O caso de análise convencional de tempos de falha, onde o que de fato é observado é uma falha ou uma censura, surgiu naturalmente no trabalho por apresentar procedimentos inferenciais muito próximos aos dos dados de degradação, com diferencial na necessidade do trato com dados censurados (mais especificamente censura à direita). Propõe-se então a modelagem de tempos de falha na presença de censura pelas famílias de distribuições SMSN e log-SMSN, a qual é realizada com su-

porte na técnica de aumento de dados. As distribuições condicionais completas dos objetos aleatórios são apresentadas e a amostragem é realizada por implementação no software WinBugs 1.4 (Spiegelhalter *et al.*, 2000). Toda a modelagem é avaliada através de um banco de dados artificial e outro real. A distribuição SN centrada também é explorada e aplicada num banco de dados artificial, onde é realizado um contra-ponto com sua versão não centrada. Em testes de vida o uso de distribuições de caudas pesadas já foi abordado por Barros *et al.* (2008) e Cancho *et al.* (2010), em que os autores assumem tempos de falha distribuídos segundo a distribuição generalizada Birnbaum-Saunders-Student-t proposta por Díaz-García e Leiva-Sánchez (2005), e no recente artigo de Vallejos e Steel (2015), onde os autores fazem uso da família de mistura na forma da distribuição lognormal (SMLN). Não se tem notícia da utilização de distribuições assimétricas, nem assimétricas de cauda pesada, para análise de tempo de falha convencional. Portanto, a aplicação das famílias SMSN e log-SMSN também são modelagens inovadoras neste contexto.

A fim de facilitar a leitura, introduzimos a notação que será utilizada ao longo deste texto e as simplificações que se fizeram necessárias. A variável aleatória  $Y_{ij}$  representa a degradação a ser observada para a unidade  $i$  no  $j$ -ésimo tempo e o valor observado dessa variável aleatória (“os dados”) é denotado por  $y_{ij}$ . Para a maior parte deste texto essa convenção será mantida. Entretanto, se ocorrerem desvios desta notação, estará claro pelo contexto a referência à variável aleatória ou à sua realização (seu valor observado).

Além disso, a notação  $f_{\mathbf{Y}|\boldsymbol{\theta}(\mathbf{y}|\boldsymbol{\theta})}$  (ou  $f(\mathbf{y}|\boldsymbol{\theta})$ ) será utilizada para representar a função densidade de probabilidade (fdp) do vetor aleatório  $\mathbf{Y}$ , indexado pelo vetor de parâmetros  $\boldsymbol{\theta}$ , avaliada no vetor de valores observados (“os dados”)  $\mathbf{y}$ . De maneira equivalente,  $F_{\mathbf{Y}|\boldsymbol{\theta}(\mathbf{y}|\boldsymbol{\theta})}$  (ou  $F(\mathbf{y}|\boldsymbol{\theta})$ ) é a função de distribuição acumulada (fda) de  $\mathbf{Y}$  associada a  $f_{\mathbf{Y}|\boldsymbol{\theta}(\mathbf{y}|\boldsymbol{\theta})}$ .

Ao longo deste trabalho também denotamos por  $\Phi_n(\cdot|\mu, \sigma^2)$  e  $\phi_n(\cdot|\mu, \sigma^2)$  (respectivamente,  $\Phi_n(\cdot|\sigma^2)$  e  $\phi_n(\cdot|\sigma^2)$ ,  $\Phi_n(\cdot)$  e  $\phi_n(\cdot)$ ) a fda e a fdp de uma variável aleatória



(va)  $X \sim N(\mu, \sigma^2)$  (respectivamente,  $X \sim N(0, \sigma^2)$ ,  $X \sim N(0, 1)$ ). Nos casos em que a dimensionalidade “ $n$ ” da distribuição normal for omitida assumamos que  $n = 1$ .

Por fim, o texto está organizado da seguinte forma. No Capítulo 2 apresentamos algumas distribuições de probabilidade com propriedades interessantes para modelar características de assimetria e (ou) caudas pesadas, as quais avaliaremos se são capazes de aprimorar a modelagem de degradação para uma série de situações práticas. No Capítulo 3, Seção 3.1 apresentamos alguns dos principais modelos e métodos inferenciais disponíveis na inferência clássica para a modelagem geral de degradação, além de comentarmos brevemente algumas de suas virtudes e limitações. Na Seção 3.2 apresentamos alguns dos modelos gerais de degradação utilizados para inferir sobre tempo de falha utilizando a abordagem pela inferência bayesiana, em que também enriquecemos estes métodos com provas mais técnicas não apresentadas pelos autores e com uma nova aplicação para a técnica de aumento de dados, além de criticar e fornecer uma discussão detalhada sobre estes métodos a luz do paradigma bayesiano. No Capítulo 4 apresentamos uma modelagem completamente nova para dados de degradação, a qual é baseada nas famílias de distribuição de mistura na escala da distribuição skew normal e log skew normal (SMSN e SMLSN). Na Seção 4.1 apresentamos o método analítico aplicado a estas famílias de distribuições para fins de obtenção da distribuição dos tempos de falha em duas formas funcionais lineares para a degradação. Algumas proposições são construídas e a partir de uma delas uma nova distribuição de probabilidade emerge desta aplicação prática. A Seção 4.2 fornece as proposições necessárias para permitir a construção dos procedimentos de amostragem via métodos Markov chain Monte Carlo necessários para a realização da inferência bayesiana apresentada na Seção 3.2. Alguns métodos para avaliação de qualidade de ajuste e robustez são brevemente introduzidos na Seção 4.3. Quatro cenários de dados simulados são analisados na Seção 4.4, em que apresentamos o potencial da modelagem, suas qualidades e fraquezas. Na Seção 4.5 essas características são colocadas à prova em três bancos de dados desgaste de rodas

de trem, e num de degradação de emissores de laser, cada banco de dados com uma característica específica e que visa demonstrar a flexibilidade dos modelos propostos e vigentes. No Capítulo 5 fornecemos o ferramental teórico aplicação das distribuições nas famílias SMSN e SMLSN a dados de tempos de falha na presença de censura. A Seção 5.1 aplica a técnica de aumento de dados para a construção da função de verossimilhança. Na Seção 5.2 apresentamos a modelagem robusta de tempos de falha pelos modelos SMSN e SMLSN assim como os meios para amostrar da distribuição *a posteriori* dos parâmetros destes modelos. A distribuição skew normal centrada também é explorada. Na Seção 5.3, três cenários de dados simulados são utilizados para averiguar o potencial dos modelos propostos. A Seção 5.4 finaliza o capítulo com uma aplicação a dados reais.

## Capítulo 2

# Algumas famílias de distribuições assimétricas

Nos próximos capítulos serão apresentados a análise convencional para tempo até a falha (Capítulo 5) e modelos gerais de degradação assumindo-se um modelo linear sem intercepto para caminhos de degradação (Capítulo 4). O elemento novo trazido para modelagem de tempo de falha e degradação será a aplicação de algumas das distribuições presentes na família de distribuições com mistura na escala da skew-normal e log-skew-normal. Estas distribuições podem acomodar assimetria e caudas pesadas para a distribuição dos objetos aleatórios, o que justifica sua aplicabilidade para a análise de tempo de falha e de dados de degradação.

Algumas propriedades destas distribuições serão amplamente utilizadas e, por esta razão, este capítulo faz uma breve revisão destas distribuições. Maiores detalhes das contribuições referentes à utilização destas famílias de distribuições no contexto de confiabilidade podem ser encontradas na Seção 1.

A Seção 2.1 apresenta as distribuições de probabilidade nas famílias com mistura na escala da skew-normal e log-skew-normal. As distribuições mais conhecidas da literatura são apresentadas como casos particulares.

## 2.1 Algumas distribuições de probabilidade

Nesta seção apresentam-se brevemente algumas distribuições de probabilidade que serão consideradas ao longo deste trabalho. Tais distribuições acomodam comportamentos simétricos, assimétricos, e algumas delas também acomodam caudas pesadas.

### 2.1.1 A família de distribuições skew normal univariada

Azzalini (1985) prova que uma fdp assimétrica pode ser construída a partir de uma fdp  $f_0$  simétrica, no entorno de zero, e de uma fda  $F$ , com fdp  $f$  também simétrica no entorno de zero, da seguinte forma

$$f(x) = 2f_0(x)F(\alpha x), \quad x \in \mathfrak{R}, \quad (2.1)$$

em que  $\alpha \in \mathfrak{R}$  é o parâmetro que controla o grau de assimetria da fdp. A família de distribuições definida em (2.1) contém como caso particular a distribuição basal  $f_0$ , bastando para isto fazer  $\alpha = 0$ .

Assumindo  $f_0$  e  $F$  respectivamente como sendo fdp e fda da distribuição normal padronizada, Azzalini (1985) introduz a família de distribuições skew normal padronizada cuja fdp é:

$$f(z) = 2\phi(z)\Phi(\alpha z), \quad z \in \mathfrak{R}, \quad (2.2)$$

a qual será denotada por  $Z \sim \text{SN}(\alpha)$ , para  $\alpha \in \mathfrak{R}$ , onde  $\phi$  e  $\Phi$  denotam respectivamente a fdp e a fda da distribuição normal padrão.

É possível adicionar parâmetros de locação e escala à distribuição skew normal considerando transformação  $X = \omega Z + \mu$ , em que  $\mu \in \mathfrak{R}$  e  $\omega^2 > 0$ . Assim, a fdp da família skew normal com locação e escala é:

$$f(x) = \frac{2}{\omega} \phi\left(\frac{x-\mu}{\omega}\right) \Phi\left(\alpha \left(\frac{x-\mu}{\omega}\right)\right), \quad z \in \mathfrak{R}, \quad (2.3)$$

em que  $\mu \in \mathfrak{R}$  é o parâmetro de locação,  $\omega^2 > 0$  é parâmetro de escala e  $\alpha \in \mathfrak{R}$  é o parâmetro de assimetria. Uma va com distribuição definida na equação (2.3) é denotada por  $X \sim \text{SN}(\mu, \omega^2, \alpha)$ . Denotamos por  $\Phi_{\text{SN}}(\cdot|\alpha)$  a fdp de uma va  $X \sim \text{SN}(\alpha)$ .

As Figuras 2.1(a) e 2.1(b) mostram o comportamento da distribuição skew normal padrão para alguns valores de  $\alpha$ .

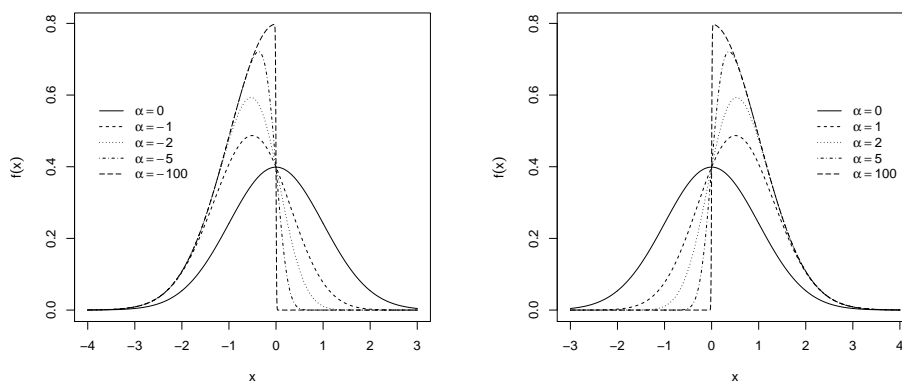
(a) Valores negativos de  $\alpha$ (b) Valores positivos de  $\alpha$ 

Figura 2.1: Densidades da distribuição skew normal padrão para diferentes valores de  $\alpha$

Algumas propriedades importantes da família skew normal são mencionadas a seguir:

1. Se  $X \sim \text{SN}(\mu, \omega^2, 0)$ , então  $X \sim \text{N}(\mu, \omega^2)$ , pois  $\Phi(0) = \frac{1}{2}$  (ver equação (2.3)).
2. Se  $\alpha \rightarrow \infty$ , então  $f(x|\mu, \omega^2, \alpha) \rightarrow \frac{2}{\omega} \phi\left(\frac{x-\mu}{\omega}\right) I_{\{x>0\}}$ , pois  $\phi(\alpha x) \rightarrow 1$ , se  $x > 0$ , e  $\phi(\alpha x) \rightarrow 0$ , se  $x < 0$ , que é a fdp de uma va com distribuição half-normal.
3. A função na equação (2.3) é log-côncava e fortemente unimodal (detalhes em Klaassen (1985)).
4.  $1 - \Phi_{\text{SN}}(-x|\alpha) = \Phi_{\text{SN}}(x|\alpha)$ .
5. Se  $Z \sim \text{SN}(\alpha)$  então  $\Phi_{\text{SN}}(x) = [\Phi(x)]^2 = \Phi_2(\mathbf{x})$ .
6. Se  $X \sim \text{SN}(0, 1, \alpha)$ , então  $X^2 \sim \chi^2$ .
7. Se  $X \sim \text{SN}(\mu, \omega^2, \alpha)$ , então  $E(X) = \mu + \omega \sqrt{\frac{2}{\pi}} \frac{\alpha}{\sqrt{1+\alpha^2}}$  e  $\text{Var}(X) = \omega^2 \left(1 - \frac{2\alpha^2}{\pi(1+\alpha^2)}\right)$ .

8. Se  $Z \sim \text{SN}(\alpha)$ , então  $X = \mu + \omega Z \sim \text{SN}(\mu, \omega^2, \alpha)$ .

9. Se  $X \sim \text{SN}(\mu, \omega^2, \alpha)$ , então  $x_p = \Phi_{\text{SN}}^{-1}(p)\omega + \mu$  é o percentil de ordem  $p$  de  $X$ .

Por questões inferenciais pode ser mais fácil trabalhar com outra forma de construção da distribuição skew normal. Seguindo Henze (1986), se uma va  $Z \sim \text{SN}(\alpha)$ , então  $Z$  pode ser representada estocasticamente por:

$$Z \stackrel{D}{=} \delta |U_0| + (1 - \delta^2)^{1/2} U_1, \quad (2.4)$$

em que  $U_0 \perp U_1$  e  $U_0 \stackrel{D}{=} U_1 \sim \text{N}(0,1)$ , com  $\delta = \frac{\alpha}{\sqrt{1+\alpha^2}}$  e  $\alpha \in \Re \Rightarrow |\delta| < 1$ .

Conseqüentemente, se  $X \sim \text{SN}(\mu, \omega^2, \alpha)$  então tem-se que  $X \stackrel{D}{=} \omega Z + \mu$ , em que  $Z \sim \text{SN}(\alpha)$ . Assim, de (2.4) segue que

$$X \stackrel{D}{=} \mu + \omega \left( \delta |U_0| + (1 - \delta^2)^{1/2} U_1 \right). \quad (2.5)$$

Observe que  $X|U_0 = u_0 \sim \text{N}(\mu + \omega\delta |u_0|, \omega^2 (1 - \delta^2))$ , em que  $U \sim \text{N}(0,1)$ . Equivalentemente pode-se escrever  $X|U = u \sim \text{N}(\mu + \omega\delta u, \omega^2 (1 - \delta^2))$ , em que  $U \sim \text{HN}(0,1)$ , em que  $\text{HN}(0,1)$  é a notação para uma distribuição half-normal padrão.

### 2.1.2 A família de distribuições log skew normal univariada

Similar ao que ocorre no caso normal, considerando uma transformação apropriada de uma variável skew normal, Azzalini *et al.* (2003) constrói uma família de distribuições que coloca massa nula em valores negativos da va.

Seja  $X \sim \text{SN}(\mu, \omega^2, \alpha)$  e considere a transformação  $Y = e^X$ . Então  $Y$  tem distribuição log skew normal, denotada por  $Y \sim \text{LSN}(\mu, \omega^2, \alpha)$  e sua fdp é dada por

$$f(y) = \frac{2}{\omega y} \phi\left(\frac{\ln y - \mu}{\omega}\right) \Phi\left(\alpha \left(\frac{\ln y - \mu}{\omega}\right)\right), \quad y > 0, \quad (2.6)$$

em que  $\mu \in \Re$  é o parâmetro de locação,  $\omega^2$  é o parâmetro de escala e  $\alpha \in \Re$  é parâmetro de assimetria.

A família de distribuições log skew normal, definida na equação (2.6), tem a distribuição lognormal como caso particular quando  $\alpha = 0$ . Podemos notar das Figuras 2.2(a) e 2.2(b) o comportamento da distribuição log skew normal padrão para diferentes

valores de  $\alpha$ .

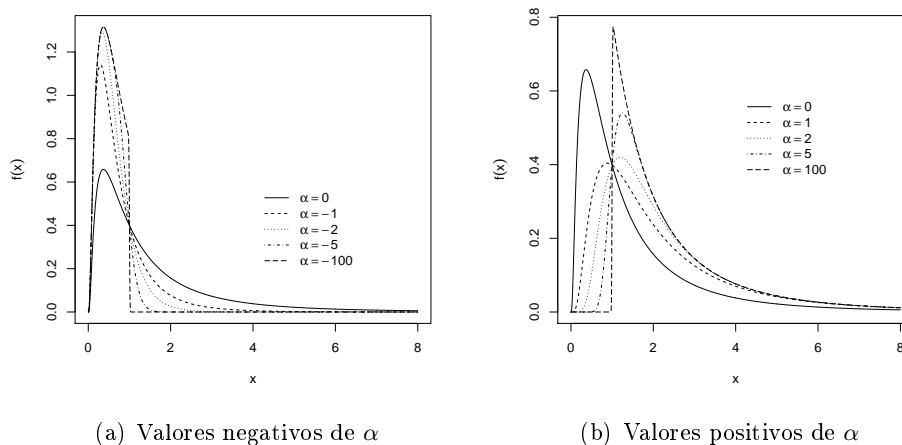


Figura 2.2: Densidades da distribuição log skew normal padrão para diferentes valores de  $\alpha$

Algumas propriedades da distribuição log skew normal são fornecidas a seguir. Se  $Y \sim \text{LSN}(\mu, \omega^2, \alpha)$ , então

1.  $E(Y) = 2e^{\left(\mu + \frac{\omega^2}{2}\right)} \Phi\left(\frac{\alpha\omega}{\sqrt{1+\alpha^2}}\right)$ .
2.  $V(Y) = 2e^{\left(\mu + \frac{\omega^2}{2}\right)} \left[ e^{\omega^2} \Phi\left(\frac{2\alpha\omega}{\sqrt{1+\alpha^2}}\right) - 2 \left[ \Phi\left(\frac{\alpha\omega}{\sqrt{1+\alpha^2}}\right) \right]^2 \right]$ .
3.  $x_p = \exp\left\{ \Phi_{SN}^{-1}(p)\omega + \mu \right\}$  é o percentil de ordem  $p$ , em que  $\Phi_{SN}$  denota a fda de uma va  $X \sim \text{SN}(\alpha)$ .

A distribuição log skew normal também possui representação estocástica. Basta considerarmos as representações estocásticas em (2.4) e (2.5) e notarmos que, se  $Y \sim \text{LSN}(\mu, \omega^2, \alpha)$ , então  $\log(Y) \sim \text{SN}(\mu, \omega^2, \alpha)$

### 2.1.3 Famílias de distribuições de mistura na escala da skew normal univariada

Branco e Dey (2001) definem a família de distribuições skew-elípticas e apresentam

alguns de seus casos especiais. Aqui trabalharemos com um destes casos, conhecido como mistura na escala da distribuição skew normal (SMSN), a qual engloba uma grande variedade de distribuições assimétricas e de caudas pesadas.

Dizemos que a va  $X$  tem distribuição na família SMSN, denotada por  $X \sim \text{SMSN}(\mu, h^{-1}\omega^2, \alpha, \mathcal{H})$  se sua fdp é dada por

$$f(x|\mu, h^{-1}\omega^2, \alpha) = 2 \int_0^\infty \phi(x|\mu, h^{-1}\omega^2) \Phi\left(\alpha \frac{(x-\mu)}{h^{-1/2}\omega}\right) d\mathcal{H}(h|\boldsymbol{\nu}), \quad (2.7)$$

em que  $\mu \in \mathfrak{R}$ ,  $\omega \in \mathfrak{R}_+$  e  $\alpha \in \mathfrak{R}$  são parâmetros, respectivamente, de localização, escala e assimetria. A medida de probabilidade  $\mathcal{H}(h|\boldsymbol{\nu})$  é definida no espaço amostral de uma variável aleatória  $H$  e  $\boldsymbol{\nu}$  é um vetor de hiperparâmetros que indexam  $\mathcal{H}$ .

Algumas propriedades da família SMSN são:

- (i) Se  $E[H^{-1/2}] < \infty$ , então  $E[X] = \mu + \sqrt{\frac{2}{\pi}} k_1 \Delta$ .
- (ii) Se  $E[H^{-1}] < \infty$ , então  $\text{Var}[X] = \omega^2 K_2 - \frac{2}{\pi} k_1^2 \Delta^2$ .
- (iii) O percentil de ordem  $p$  de  $X$  é  $x_p = \mu + \omega \Phi_{\text{SMSN}}^{-1}(p|\alpha, h, \mathcal{H})$ ,

em que  $\Delta = \omega \delta$ ,  $\delta = \frac{\alpha}{\sqrt{(1+\alpha^2)}}$ ,  $k_m = E[H^{-m/2}]$  e  $\Phi_{\text{SMSN}}^{-1}(p|\alpha, h, \mathcal{H})$  denota o percentil de ordem  $p$  de uma distribuição SMSN padrão dados  $\alpha$ ,  $h$  and  $\mathcal{H}$ .

São exemplos de distribuições que pertencem a esta família SMSN as seguintes distribuições

- (i) Se  $\mathcal{H}$  é degenerada em 1 e  $\alpha = 0$ , então  $X \sim \text{Normal}(\mu, \omega^2)$  and  $k_m = 1$ .
- (ii) Se  $\mathcal{H}$  é degenerada em 1, então  $X \sim \text{Skew Normal}(\mu, \omega^2, \alpha)$  and  $k_m = 1$ .
- (iii) Se  $H \sim \text{Gamma}(\frac{\nu}{2}, \frac{\nu}{2})$  e  $\alpha = 0$ , então  $X \sim \text{T}(\mu, \omega^2, \nu)$  e  $k_m = \left(\frac{\nu}{2}\right)^{m/2} \frac{\left(\frac{\nu-m}{2}\right)}{\left(\frac{\nu}{2}\right)}$ .
- (iv) Se  $H \sim \text{Gamma}(\frac{\nu}{2}, \frac{\nu}{2})$ , então  $X \sim \text{Skew T}(\mu, \omega^2, \alpha, \nu)$  e  $k_m = \left(\frac{\nu}{2}\right)^{m/2} \frac{\left(\frac{\nu-m}{2}\right)}{\left(\frac{\nu}{2}\right)}$ .
- (v) Se  $H \sim \text{Beta}(\nu, 1)$  e  $\alpha = 0$ , então  $X \sim \text{Slash}(\mu, \omega^2, \nu)$  e  $k_m = \frac{2\nu}{2\nu-m}$  com  $\nu > \frac{m}{2}$ .
- (vi) Se  $H \sim \text{Beta}(\nu, 1)$ , então  $X \sim \text{Skew Slash}(\mu, \omega^2, \alpha, \nu)$  e  $k_m = \frac{2\nu}{2\nu-m}$  com  $\nu > \frac{m}{2}$ .



Assim como a família SN, Azzalini e Capitanio (2003) constrói as famílias SMSN via convolução de outras va's. Se  $X \sim \text{SMSN}(\mu, h^{-1}\omega^2, \alpha, \mathcal{H})$  então sua representação estocástica é dada por

$$X \stackrel{D}{=} \mu + \Delta U + H^{-1/2} \tau^{1/2} U_1 \quad (2.8)$$

em que  $\Delta = \omega\delta$ ,  $\delta = \frac{\alpha}{\sqrt{1+\alpha^2}}$ ,  $\tau = (1 - \delta^2)\omega^2$ ,  $U = H^{-1/2}|U_0|$ ,  $U_1 \perp U_0 \stackrel{iid}{\sim} \text{N}(0,1)$  e  $\mathcal{H}$  é a medida de probabilidade definida sobre o espaço amostral de  $H$  indexada pelo vetor  $\boldsymbol{\nu}$ .

Essa representação estocástica é muito útil tanto em inferência bayesiana quanto clássica, uma vez que torna as aproximações computacionais, tais como no algoritmo EM (*expectation maximization*), para obter os estimadores de máxima verossimilhança (EMV), e nos métodos MCMC para obter amostras *a posteriori* a partir de distribuições mais simples. No caso Bayesiano esta representação é mais conveniente pois conduz a distribuições condicionais completas *a posteriori* com formas fechadas.

### 2.1.4 Famílias de distribuições de mistura na escala da log skew normal univariada

Para modelarmos variáveis aleatórias não negativas é interessante considerarmos as famílias de mistura na escala da distribuição log skew normal (SMLSN), a qual é definida a seguir:

**Definição 1.** Dizemos que  $X \sim \text{SMLSN}(\mu, h^{-1}\omega^2, \alpha, \mathcal{H})$  se sua fdp é dada por:

$$f(x|\mu, h^{-1/2}\omega, \alpha) = \frac{2}{x} \int_0^\infty \phi(\ln(x)|\mu, h^{-1}\omega^2) \Phi\left(\alpha \frac{(\ln(x) - \mu)}{h^{-1/2}\omega}\right) d\mathcal{H}(h|\boldsymbol{\nu}), \quad (2.9)$$

em que  $x \in \mathfrak{R}_+$ ,  $\mu \in \mathfrak{R}$ ,  $\omega \in \mathfrak{R}_+$  e  $\alpha \in \mathfrak{R}$  são parâmetros, respectivamente, de localização, escala e assimetria. A medida de probabilidade  $\mathcal{H}(h|\boldsymbol{\nu})$  é definida no espaço amostral de uma variável aleatória  $H$  e  $\boldsymbol{\nu}$  é um vetor de hiperparâmetros que indexam  $\mathcal{H}$ .

Esta família de distribuições pode ser equivalentemente obtida como em Marchenko e Genton (2010), que definem a família de distribuições log-skew-elípticas assumindo

que  $X \stackrel{D}{=} \exp\{Y\}$ , em que  $Y$  tem distribuição skew elíptica. Particularmente, obtém-se que, se  $X \sim \text{LSMSN}(\mu, h^{-1}\omega^2, \alpha, \mathcal{H})$ , então  $\log(X) \sim \text{SMSN}(\mu, h^{-1}\omega^2, \alpha, \mathcal{H})$ .

Esta equivalência entre  $X \sim \text{LSMSN}(\mu, h^{-1}\omega^2, \alpha, \mathcal{H})$  e  $X \sim \text{SMLSN}(\mu, h^{-1}\omega^2, \alpha, \mathcal{H})$  é provada na seguinte proposição.

**Proposição 1.** *Se  $X \sim \text{LSMSN}(\mu, h^{-1}\omega^2, \alpha, \mathcal{H})$ , então  $X \sim \text{SMLSN}(\mu, h^{-1}\omega^2, \alpha, \mathcal{H})$*

*Demonstração.* Seja  $Y \sim \text{SMSN}(\mu, h^{-1}\omega^2, \alpha, \mathcal{H})$ , e defina  $X = e^Y$ .

Tomando-se a derivada da fda de  $X$  e aplicando-se a regra da cadeia em conjunto com o teorema fundamental do cálculo, segue que

$$\begin{aligned} f(x) &= \frac{\partial P(Y \leq \log(x))}{\partial y} \frac{\partial \log(x)}{\partial x} \\ &= 2 \int_0^\infty \phi(\log(x)|\mu, h^{-1}\omega^2) \Phi\left(\alpha \frac{(\log(x) - \mu)}{h^{-1/2}\omega}\right) d\mathcal{H}(h|\nu) \frac{1}{x} \\ &= \frac{2}{x} \int_0^\infty \phi(\log(x)|\mu, h^{-1}\omega^2) \Phi\left(\alpha \frac{(\log(x) - \mu)}{h^{-1/2}\omega}\right) d\mathcal{H}(h|\nu), \end{aligned} \quad (2.10)$$

o que conclui a prova.  $\square$

As propriedades dessa família como, por exemplo média e variância, não possuem forma fechada. Algumas das distribuições que pertencem à família de distribuições SMLSN são:

- (i) Se  $P(H = 1) = 1$  e  $\alpha = 0$ , então  $X \sim \text{Log Normal}(\mu, \omega^2)$ .
- (ii) Se  $\mathcal{H}$  é degenerada em 1, então  $X \sim \text{Log Skew Normal}(\mu, \omega^2, \alpha)$ .
- (iii)  $H \sim \text{Gamma}(\frac{\nu}{2}, \frac{\nu}{2})$  e  $\alpha = 0$ , então  $X \sim \text{Log T}(\mu, \omega^2, \nu)$ .
- (iv)  $H \sim \text{Gamma}(\frac{\nu}{2}, \frac{\nu}{2})$ , então  $X \sim \text{Log Skew T}(\mu, \omega^2, \alpha, \nu)$ .
- (v)  $H \sim \text{Beta}(\nu, 1)$  e  $\alpha = 0$ , então  $X \sim \text{Log Slash}(\mu, \omega^2, \nu)$ .
- (vi)  $H \sim \text{Beta}(\nu, 1)$ , então  $X \sim \text{Log Skew Slash}(\mu, \omega^2, \alpha, \nu)$ .

Além disso, pode-se obter uma representação estocástica para  $X \sim \text{SMLSN}(\mu, h^{-1}\omega^2, \alpha, \mathcal{H})$  por notarmos que  $\log(X) \sim \text{SMSN}(\mu, h^{-1}\omega^2, \alpha, \mathcal{H})$  cuja representação estocástica

é dada pela equação(2.8), ou seja, para fazer inferência nesta família basta utilizarmos uma transformação logarítmica de  $X$ .

## Capítulo 3

# Modelos gerais de degradação: inferência sobre a distribuição dos tempos de falha

Neste capítulo efetua-se uma revisão de literatura em métodos inferenciais sobre a distribuição dos tempos de falha a partir dos modelos gerais de degradação.

Alguns dos procedimentos baseados em inferência clássica são revisados na Seção 3.1. Retiramos a referida seção da dissertação de mestrado de Oliveira (2011), do mesmo autor do presente texto, com o intuito de situar o contexto de modelagem geral de degradação, a qual foi primeiramente abordada na literatura sob perspectiva da inferência clássica. Ou seja, para especialistas nesta área de estudo a leitura desta seção é dispensável, uma vez que não apresenta nenhuma contribuição científica gerada nesta texto.

Na Seção 3.2 apresentamos procedimentos baseados em inferência bayesiana. Esta seção também conta com partes do trabalho de mestrado de Oliveira (2011), porém aqui elaboramos uma discussão complementar sobre a validade teórica e interpretativa dos

procedimentos em inferência bayesiana, a qual ainda não foi explorada pela literatura vigente. A saber, fornecemos provas formais dos procedimentos sugeridos por (Hamada, 2005; Robinson e Crowder, 2000; Oliveira, 2011). Através de um exemplo simples provamos e exemplificamos um problema teórico no método proposto por Hamada (2005) para inferir sobre os percentis, e fornecemos uma estratégia baseada em aumento de dados para inferir sobre a distribuição do tempo de falha de objetos futuros da população utilizando a distribuição preditiva *a posteriori*.

### 3.1 Inferência clássica

Nesta seção serão brevemente apresentados os principais métodos de estimação e análise de dados de degradação baseados em inferência clássica. Embora tais métodos não sejam o foco do presente trabalho, a nomenclatura para modelos de degradação é introduzida através dos mesmos. Na Seção 3.1.1 apresenta-se o método aproximado (encontrado em Meeker e Escobar, 1998), na Seção 3.1.2 o método analítico (Lu e Meeker, 1993), na Seção 3.1.3 o método de dois estágios (Lu e Meeker, 1993) e finalmente o método numérico na Seção 3.1.4 (uma boa referência é Meeker e Escobar, 1998).

#### 3.1.1 O método aproximado

Considere o modelo de degradação da equação (1.3), o qual consiste de dois passos.

No primeiro passo uma análise separada para cada perfil é realizada de modo a definir (estimar) quando o limiar de degradação  $D_f$  é alcançado caracterizando a ocorrência de uma falha. Ao tempo de falha estimado, obtido pela solução da equação de degradação em relação a  $t$  quando a degradação é igual à degradação crítica, dá-se o nome “pseudo tempo de falha”.

No segundo passo os “pseudo tempos de falha” são considerados como uma amostra completa de tempos de falha e são analisados para a estimação de  $F_{T|\theta}(t|\theta)$ , em que  $\theta =$

$(\theta_1, \theta_2, \dots, \theta_q)'$  é o vetor de parâmetros da distribuição dos efeitos aleatórios  $\Lambda_{\beta|\theta}(\beta|\theta)$ , os quais no método aproximado são tratados como fixos. Formalmente, temos:

1. Para cada unidade  $i$  considere o modelo  $Y_{ij} = D(t_{ij}; \alpha; \beta_i) + \varepsilon_{ij}$  e a amostra observada  $(t_{i1}, y_{i1}), \dots, (t_{im_i}, y_{im_i})$  para encontrar as estimativas de máxima verossimilhança (condicionadas) de  $\alpha = (\alpha_1, \alpha_2, \dots, \alpha_r)'$  e  $\beta_i = (\beta_{i1}, \beta_{i2}, \dots, \beta_{ip})'$ . Note que fixando-se  $i$ , as estimativas são obtidas para uma unidade experimental específica. Portanto, o vetor  $\beta_i = (\beta_{i1}, \beta_{i2}, \dots, \beta_{ip})'$  é tratado como fixo. O vetor  $\alpha = (\alpha_1, \alpha_2, \dots, \alpha_r)'$  de efeitos fixos, que no modelo geral (1.3) é comum para todas as unidades, também será estimado para cada uma separadamente. Denote tais estimativas  $\hat{\alpha}_i$  e  $\hat{\beta}_i$ , as quais também podem ser obtidas através de mínimos quadrados.
2. Para cada unidade  $i$  resolva as equações  $D_i(t; \hat{\alpha}_i; \hat{\beta}_i) = D_f$  para  $t$  e obtenha a solução  $\hat{t}_1, \dots, \hat{t}_n$  (“pseudo tempos de falha”).
3. Execute o procedimento convencional de análise de tempo de falha (Nelson, 1990) utilizando os “pseudo tempos de falha”  $\hat{t}_1, \dots, \hat{t}_n$  para estimar  $F_{T|\theta}(t|\theta)$ .

Este método mostra-se adequado para caminhos de degradação relativamente simples e dados que sejam suficientes para estimar  $\alpha_i$  e  $\beta_i$ . O horizonte de extrapolação deve ser curto, uma vez que o modelo desconsidera erros de previsão e os erros inerentes aos valores de degradação observados.

### 3.1.2 O método analítico

Na primeira etapa do método utiliza-se a forma funcional imposta ao perfil de degradação para se obter a distribuição do tempo de falha a partir da transformação da distribuição dos efeitos fixos e aleatórios do modelo. Para modelos de degradação simples, a fda para o tempo até a falha  $F_{T|\theta}(t|\theta)$  pode ser expressa de forma fechada. Na

segunda etapa os parâmetros da distribuição dos efeitos fixos e aleatórios são estimados e avaliados nas quantidades de interesse da distribuição do tempo de falha obtida na primeira etapa. Os exemplos a seguir ilustram dois destes casos simples.

### Método analítico no modelo linear com efeito aleatório normal

Suponha que o perfil real de degradação de uma unidade particular seja dado por  $D(t) = \frac{1}{\beta}t$ , em que  $\frac{1}{\beta}$  é a taxa de degradação que varia de unidade para unidade.

Para o nível crítico, sob este modelo, tem-se a seguinte relação entre o nível crítico de degradação  $D_f$  e o tempo até a falha  $T$ :

$$D_f = \frac{1}{\beta}T \text{ e } T = D_f\beta.$$

Consequentemente a fda do tempo até a falha  $T$  é dada por

$$\begin{aligned} F_{T|\beta}(t|\boldsymbol{\theta}_\beta) &= P(D_f\beta \leq t|\boldsymbol{\theta}_\beta) \\ &= P\left(\beta \leq \frac{t}{D_f}|\boldsymbol{\theta}_\beta\right), \end{aligned} \quad (3.1)$$

em que  $\boldsymbol{\theta}_\beta$  é o vetor de parâmetros que indexa a distribuição  $\Lambda(\beta|\boldsymbol{\theta}_\beta)$  dos efeitos aleatórios.

Assim, se é assumido que  $\beta$  segue uma distribuição normal com parâmetros  $\mu_\beta \in \Re$  e  $\omega_\beta^2 > 0$ , isto é,  $\beta \sim N(\mu_\beta, \omega_\beta^2)$ , segue de resultados de transformação de variáveis em probabilidade que  $T \sim N(\mu_\beta D_f; \omega_\beta^2 D_f^2)$ .

Daí, as quantidades de interesse tais como, tempo esperado até a falha, percentil de ordem  $p$ , fração de falha e a confiabilidade do sistema são dados respectivamente por

1.  $E(T) = \mu_\beta D_f$ ,
2.  $t_p = \Phi^{-1}(p)\omega_\beta D_f + \mu_\beta D_f$ ,
3.  $F(t) = \Phi\left(\frac{t - \mu_\beta D_f}{\omega_\beta^2 D_f^2}\right)$ ,
4.  $R(t) = \Phi\left(-\frac{t - \mu_\beta D_f}{\omega_\beta^2 D_f^2}\right)$ ,

onde  $\Phi(a)$  denota a fda da distribuição normal padrão avaliada no ponto  $a$ . Utilizando-se os dados de degradação é possível estimar os parâmetros  $\mu_\beta$  e  $\omega_\beta^2$  via máxima verossimilhança ou mínimos quadrados e assim, por conta da propriedade de invariância destes estimadores, obtém-se as estimativas de máxima verossimilhança para  $E(T)$ ,  $t_p$ ,  $F(t)$  e  $R(t)$ .

### Método analítico no modelo linear com efeito aleatório lognormal

Suponha que o perfil real de degradação de uma unidade particular seja dado por  $D(t) = \alpha + \beta t$ , em que  $\alpha$  é efeito fixo e representa a degradação inicial de todas as unidades sob teste, ou seja,  $D(0) = \alpha$ ;  $\beta$  é a taxa de degradação que varia de unidade para unidade (corresponde ao efeito aleatório) de acordo com uma distribuição lognormal com parâmetros  $\mu_\beta \in \Re$  e  $\sigma_\beta > 0$ .

Para o nível crítico  $D_f$  tem-se:

$$D_f = \alpha + \beta T \text{ e } T = \frac{D_f - \alpha}{\beta}$$

A fda de T é dada por

$$\begin{aligned} F_{T|\alpha, \mu_\beta, \sigma_\beta}(t|\alpha, \mu_\beta, \sigma_\beta) &= P(T \leq t|\alpha, \mu_\beta, \sigma_\beta) \\ &= \Phi\left(\frac{\log t - [\log(D_f - \alpha) - \mu_\beta]}{\sigma_\beta}|\alpha, \mu_\beta, \sigma_\beta\right), t > 0. \end{aligned}$$

Consequentemente,  $T$  tem distribuição lognormal com parâmetro de escala  $\mu_T = [\log(D_f - \alpha) - \mu_\beta]$  e parâmetro de forma  $\sigma_T = \sigma_\beta > 0$ . Outras fdp podem ser utilizadas para estimar  $F_T(t)$  utilizando o mesmo procedimento. Resultados para outras distribuições, como por exemplo as distribuições weibull e bernstein, podem ser encontradas em Lu e Meeker (1993).

### 3.1.3 O método de dois estágios

Este método foi apresentado por Lu e Meeker (1993) e parte do pressuposto de que o vetor de efeitos aleatórios  $\beta$ , ou alguma reparametrização apropriada  $\varphi = H(\beta)$  destes efeitos, segue uma distribuição normal multivariada. A seguir são explicitados



os passos do método.

### Estágio 1

1. Para cada unidade amostral  $i$ , com  $i = 1, \dots, n$ , ajuste o modelo de degradação dado em (1.3) através do método de mínimos quadrados e obtenha as estimativas de estágio 1 para os parâmetros do modelo  $(\boldsymbol{\alpha}; \boldsymbol{\beta}_i)$ . Em outras palavras, para cada unidade  $i$ ,  $(\hat{\boldsymbol{\alpha}}_i; \hat{\boldsymbol{\beta}}_i)$  são, respectivamente, as estimativas de mínimos quadrados para os parâmetros de efeito fixo e aleatório do modelo. Ainda, um estimador da variância do erro  $\sigma_\varepsilon^2$ , obtida considerando-se a informação da unidade  $i$ , é dada pelo erro quadrático médio

$$\sigma_\varepsilon^2 = \left[ \sum_{i=1}^{m_i} \frac{\left\{ y_{ij} - D(t_{ij}; \hat{\boldsymbol{\alpha}}_i; \hat{\boldsymbol{\beta}}_i) \right\}^2}{(m_i - q)} \right], \text{ em que } q = p + r.$$

em que  $p$  e  $r$  são respectivamente as dimensões de  $\boldsymbol{\alpha}$  e  $\boldsymbol{\beta}$

2. Assuma que, para alguma reparametrização  $\phi_i$  apropriada (e.g, transformação de Box e Cox, 1964),  $\hat{\boldsymbol{\varphi}}_i = H(\hat{\boldsymbol{\beta}}_i)$  é aproximadamente distribuído segundo uma normal multivariada com média assintótica  $\boldsymbol{\mu}_\varphi$  e matriz de variância e covariância  $\boldsymbol{\Sigma}_\varphi$ .

### Estágio 2

No segundo estágio, os estimadores  $(\hat{\boldsymbol{\alpha}}_i; \hat{\boldsymbol{\varphi}}_i)$  são combinados para construir os estimadores dos parâmetros do modelo de degradação. Tais estimadores para  $\boldsymbol{\alpha}$ ,  $\boldsymbol{\mu}_\varphi$  e  $\boldsymbol{\Sigma}_\varphi$  são, respectivamente, dados por

$$\hat{\boldsymbol{\alpha}} = \sum_{i=1}^n \frac{\hat{\boldsymbol{\alpha}}_i}{n},$$

$$\hat{\boldsymbol{\mu}}_\varphi = \sum_{i=1}^n \frac{\hat{\boldsymbol{\varphi}}_i}{n},$$

$$\hat{\Sigma}_{\varphi} = \left[ \sum_{i=1}^n \frac{(\hat{\varphi}_i - \hat{\mu}_{\varphi})(\hat{\varphi}_i - \hat{\mu}_{\varphi})'}{(n-1)} \right] - \sum_{i=1}^n \frac{\hat{\sigma}_{\varphi_i}}{n}.$$

A estimativa pontual  $\hat{F}_{T|\alpha, \mu_{\varphi}, \Sigma_{\varphi}}(t|\alpha, \mu_{\varphi}, \Sigma_{\varphi})$  de  $F_{T|\alpha, \mu_{\varphi}, \Sigma_{\varphi}}(t|\alpha, \mu_{\varphi}, \Sigma_{\varphi})$  pode ser avaliada para qualquer nível de precisão desejada através de simulação Monte Carlo. Os passos básicos são:

1. Estime os parâmetros  $\alpha$ ,  $\mu_{\varphi}$  e  $\Sigma_{\varphi}$  de  $n$  perfis de degradação utilizando o método de dois estágios.
2. Gere  $N$  realizações  $\tilde{\varphi}$  de  $\varphi$  da distribuição  $N(\hat{\mu}_{\varphi}; \hat{\Sigma}_{\varphi})$  e obtenha as  $N$  correspondentes realizações  $\beta^*$  de  $\beta$  a partir de  $H^{-1}(\tilde{\varphi})$ , em que  $N$  é grande (e.g.,  $N = 100.000$ ) e  $H^{-1}$  é a transformada inversa de  $H$ . Em casos para os quais  $F_{\beta}(\cdot)$  de  $\beta$  é conhecida, as realizações  $\beta^*$  de  $\beta$  podem ser geradas diretamente desta distribuição. Tais valores podem ser aplicados nos passos 3 e 4.
3. Calcule os  $N$  tempos de falha  $t^*$  simulados substituindo  $\beta^*$  em  $D_f = D(\beta; \hat{\alpha}; t)$ .
4. Estime  $F_{T|\alpha, \mu_{\varphi}, \Sigma_{\varphi}}(t|\alpha, \mu_{\varphi}, \Sigma_{\varphi})$  da distribuição empírica simulada para qualquer valor desejado de  $t$  da seguinte maneira:

$$\hat{F}_{T|\alpha, \mu_{\varphi}, \Sigma_{\varphi}}(t|\alpha, \mu_{\varphi}, \Sigma_{\varphi}) = \frac{[\text{Número de tempos de falha } t^* \leq t]}{N}.$$

O erro Monte Carlo é calculado através da distribuição binomial. Este erro pode ser minimizado fazendo-se  $N$  suficientemente grande. Intervalos de confiança podem ser construídos através de procedimentos bootstrap (Efron, 1985).

O número de observações em cada nível afeta a acurácia da estimação dos parâmetros do modelo para cada um dos perfis individuais (Lu e Meeker, 1993). No desenvolvimento do método de dois estágios, Lu e Meeker (1993) assumiram que o número de observações para cada perfil era suficientemente grande para garantir

as propriedades assintóticas dos estimadores. Para o modelo utilizado no artigo, os autores concluíram através de um estudo por simulação (não incluído no artigo) que a normalidade assintótica era garantida com um número mínimo de  $m_i = 7$  observações ( $i = 1, 2, \dots, n$ ) para cada uma das  $n$  unidades amostrais.

### 3.1.4 O método numérico

Este método consiste na estimação dos parâmetros do modelo geral de degradação (1.3) pelo método de máxima verossimilhança e, posteriormente, a obtenção de  $F_T(t)$  através de simulação de Monte Carlo.

Primeiramente é necessário retomar o modelo geral (1.3) e introduzir uma notação adicional que será utilizada ao longo do texto nas próximas seções e capítulos. A referência básica para esta notação é o trabalho de Robinson e Crowder (2000).

Considere o modelo geral de degradação anteriormente definido em (1.3)

$$Y_{ij} = D(t_{ij}; \boldsymbol{\alpha}, \boldsymbol{\beta}_i) + \varepsilon_{ij} \quad i = 1, \dots, n; \quad j = 1, \dots, m_i,$$

em que,  $Y_{ij}$  representa a  $j$ -ésima medida de degradação da unidade amostral  $i$ , no  $j$ -ésimo tempo pré-especificado  $t_{ij}$ ,  $\boldsymbol{\alpha} = (\alpha_1, \dots, \alpha_r)'$  é o vetor de fixos, desconhecido e comum à todas as unidades,  $\boldsymbol{\beta}_i = (\beta_{i1}, \dots, \beta_{ip})'$  é o vetor de efeitos aleatórios do modelo o qual varia entre as unidades experimentais.

As suposições básicas deste modelo são:

1.  $\boldsymbol{\beta}_1, \dots, \boldsymbol{\beta}_n$  são *iid* e tal que  $\boldsymbol{\beta}_i$  segue uma distribuição  $\Lambda_{\boldsymbol{\beta}_i|\boldsymbol{\theta}}(\boldsymbol{\beta}_i|\boldsymbol{\theta})$  indexada por um vetor fixo de parâmetros  $\boldsymbol{\theta} = (\theta_1, \dots, \theta_q)'$  desconhecido. Denote por  $f_{\boldsymbol{\beta}_i|\boldsymbol{\theta}}(\boldsymbol{\beta}_i|\boldsymbol{\theta})$  a fdp do vetor de efeito aleatório  $\boldsymbol{\beta}_i$ .

2. Os erros são tais que  $\varepsilon_{ij} \stackrel{iid}{\sim} N(0, \sigma_\varepsilon^2)$ , em que  $\sigma_\varepsilon^2$  é fixo e desconhecido, cuja fdp é

$$f_{\varepsilon_{ij}|\sigma_\varepsilon^2}(\varepsilon_{ij}|\sigma_\varepsilon^2) = \frac{1}{\sqrt{2\pi\sigma_\varepsilon^2}} \exp\left\{-\frac{1}{2\sigma_\varepsilon^2}\varepsilon_{ij}^2\right\},$$

em que,  $i = 1, \dots, n$  e  $j = 1, \dots, m_i$ .

3. os vetores aleatórios  $\boldsymbol{\beta}_i$  e  $\boldsymbol{\varepsilon}_i = (\varepsilon_{i1}, \dots, \varepsilon_{im_i})'$  são independentes para todo  $i =$

$1, 2, \dots, n$ .

Seja  $\mathbf{Y}_i = (Y_{i1}, Y_{i2}, \dots, Y_{im_i})$  o vetor aleatório que denota as medidas de degradação da unidade  $i$ , e  $\mathbf{y}_i = (y_{i1}, y_{i2}, \dots, y_{im_i})$  o vetor de degradações observado na unidade  $i$ . De maneira análoga, seja  $\mathbf{Y} = (\mathbf{Y}_1, \dots, \mathbf{Y}_n)$  o vetor aleatório que representa o conjunto completo de medidas de degradação para as  $n$  unidades e  $\mathbf{y} = (\mathbf{y}_1, \dots, \mathbf{y}_n)$  os dados observados durante o período de ensaios. Seja  $\boldsymbol{\beta} = (\boldsymbol{\beta}_1, \boldsymbol{\beta}_2, \dots, \boldsymbol{\beta}_n) = (\beta_{11}, \beta_{12}, \dots, \beta_{1p}; \beta_{21}, \beta_{22}, \dots, \beta_{2p}; \dots; \beta_{n1}, \beta_{n2}, \dots, \beta_{np})$  que agrega todos os parâmetros aleatórios do modelo (para todas as  $n$  unidades experimentais).

Denote a fdp de  $\mathbf{Y}_i$  como  $f_{\mathbf{Y}_i|\boldsymbol{\alpha}, \boldsymbol{\beta}_i, \boldsymbol{\theta}, \sigma_\varepsilon^2}(\mathbf{y}_i|\boldsymbol{\alpha}, \boldsymbol{\beta}_i, \boldsymbol{\theta}, \sigma_\varepsilon^2)$ . Assim pela suposição de independência de  $\boldsymbol{\beta}_i$  e  $\varepsilon_{ij}$ ,  $\forall i, j$ , tem-se que as fdp de  $\mathbf{Y}$  e  $\boldsymbol{\beta}$  são dadas respectivamente por

$$f_{\mathbf{Y}|\boldsymbol{\alpha}, \boldsymbol{\beta}, \boldsymbol{\theta}, \sigma_\varepsilon^2}(\mathbf{y}|\boldsymbol{\alpha}, \boldsymbol{\beta}, \boldsymbol{\theta}, \sigma_\varepsilon^2) = \prod_i^n f_{\mathbf{Y}_i|\boldsymbol{\alpha}, \boldsymbol{\beta}_i, \boldsymbol{\theta}, \sigma_\varepsilon^2}(\mathbf{y}_i|\boldsymbol{\alpha}, \boldsymbol{\beta}_i, \boldsymbol{\theta}, \sigma_\varepsilon^2), \quad (3.2)$$

$$f_{\boldsymbol{\beta}|\boldsymbol{\theta}}(\boldsymbol{\beta}|\boldsymbol{\theta}) = \prod_i^n f_{\boldsymbol{\beta}_i|\boldsymbol{\theta}}(\boldsymbol{\beta}_i|\boldsymbol{\theta}). \quad (3.3)$$

Sejam  $\boldsymbol{\alpha}$ ,  $\boldsymbol{\theta}$  e  $\sigma_\varepsilon^2$  os parâmetros (fixos) a serem estimados. De (3.2) e (3.3), segue que a expressão geral da função de verossimilhança é dada por

$$\begin{aligned} f_{\mathbf{Y}|\boldsymbol{\alpha}, \boldsymbol{\theta}, \sigma_\varepsilon^2}(\mathbf{y}|\boldsymbol{\alpha}, \boldsymbol{\theta}, \sigma_\varepsilon^2) &= \int_{\Xi_{\boldsymbol{\beta}}} f_{\mathbf{Y}, \boldsymbol{\beta}|\boldsymbol{\alpha}, \boldsymbol{\theta}, \sigma_\varepsilon^2}(\mathbf{y}, \boldsymbol{\beta}|\boldsymbol{\alpha}, \boldsymbol{\theta}, \sigma_\varepsilon^2) d\boldsymbol{\beta} \\ &= \prod_{i=1}^n \int_{\Xi_{\boldsymbol{\beta}_i}} \left\{ f_{\mathbf{Y}_i|\boldsymbol{\alpha}, \boldsymbol{\beta}_i, \boldsymbol{\theta}, \sigma_\varepsilon^2}(\mathbf{y}_i|\boldsymbol{\alpha}, \boldsymbol{\beta}_i, \boldsymbol{\theta}, \sigma_\varepsilon^2) f_{\boldsymbol{\beta}_i|\boldsymbol{\theta}}(\boldsymbol{\beta}_i|\boldsymbol{\theta}) \right\} d\boldsymbol{\beta}_i, \end{aligned} \quad (3.4)$$

em que,  $\Xi_{\boldsymbol{\beta}} = (\Xi_{\boldsymbol{\beta}_1}, \dots, \Xi_{\boldsymbol{\beta}_n})$  é o espaço paramétrico associado a  $\boldsymbol{\beta}$ , e  $\Xi_{\boldsymbol{\alpha}}$  é o espaço paramétrico associado a  $\boldsymbol{\alpha}$ .

Se é assumido que  $\varepsilon_{ij} \stackrel{iid}{\sim} N(0, \sigma_\varepsilon^2)$ , então dado,  $\boldsymbol{\alpha}_i$ ,  $\boldsymbol{\beta}_i$ ,  $\boldsymbol{\theta}$  e  $\varepsilon_{ij}$ , segue que  $\mathbf{Y}_i \stackrel{iid}{\sim} N_{m_i}(\boldsymbol{\mu}_{\mathbf{Y}_i}; \boldsymbol{\Sigma}_{\mathbf{Y}_i})$ , em que  $\boldsymbol{\mu}_{\mathbf{Y}_i}$  é o vetor de médias (de dimensão  $m_i \times 1$  dado por  $\boldsymbol{\mu}_{\mathbf{Y}_i} = E(\mathbf{Y}_i|\boldsymbol{\alpha}, \boldsymbol{\beta}, \boldsymbol{\theta}) = [D(t_{i1}; \boldsymbol{\alpha}, \boldsymbol{\beta}_i), \dots, D(t_{im_i}; \boldsymbol{\alpha}, \boldsymbol{\beta}_i)]'$  e matriz de variância-covariância  $\boldsymbol{\Sigma}_{\mathbf{Y}_i}$

$=\sigma_\varepsilon^2 I_{m_i}$ , na qual  $I_{m_i}$  é a matriz identidade de ordem  $m_i$ .

Assim, a função de verossimilhança torna-se

$$f_{\mathbf{Y}|\boldsymbol{\alpha},\boldsymbol{\theta},\sigma_\varepsilon^2}(\mathbf{y}|\boldsymbol{\alpha},\boldsymbol{\theta},\sigma_\varepsilon^2) = \prod_{i=1}^n \int_{\boldsymbol{\beta}_i} \left\{ \left[ \prod_{j=1}^{m_i} N(D_i(t_{ij}; \boldsymbol{\alpha}, \boldsymbol{\beta}_i); \sigma_\varepsilon^2) \right] f_{\boldsymbol{\beta}_i|\boldsymbol{\theta}}(\boldsymbol{\beta}_i|\boldsymbol{\theta}) \right\} d\boldsymbol{\beta}_i. \quad (3.5)$$

A estimação de  $\boldsymbol{\alpha}$ ,  $\boldsymbol{\theta}$  e  $\sigma_\varepsilon^2$  é então obtida pela maximização da função (3.5). Entretanto a integração nos  $\boldsymbol{\beta}_i$ 's é em geral complexa e muitas vezes não pode ser obtida analiticamente. Uma solução possível é resolvê-las numericamente. Contudo, mesmo numericamente, este problema pode ser de difícil resolução se a dimensão de  $\boldsymbol{\beta}_i$  é grande.

Uma suposição adicional, em geral muito utilizada, é a de que os  $\boldsymbol{\beta}_i$ 's são *iid* segundo uma Normal Multivariada de ordem  $p$  com vetor de médias  $\boldsymbol{\mu}_\beta$  e matriz de variância-covariância  $\boldsymbol{\Sigma}_\beta$ , denotado por  $\boldsymbol{\theta} = \left( \boldsymbol{\mu}_\beta, \text{vech}(\boldsymbol{\Sigma}_\beta) \right)'$ . Conseqüentemente, a função de verossimilhança na equação (3.5) torna-se:

$$f_{\boldsymbol{\beta}_i|\boldsymbol{\mu}_\beta, \boldsymbol{\Sigma}_\beta}(\boldsymbol{\beta}_i|\boldsymbol{\mu}_\beta, \boldsymbol{\Sigma}_\beta) = \frac{1}{(2\pi)^{p/2} |\boldsymbol{\Sigma}_\beta|^{1/2}} \exp \left\{ -\frac{1}{2} (\boldsymbol{\beta}_i - \boldsymbol{\mu}_\beta)' \boldsymbol{\Sigma}_\beta^{-1} (\boldsymbol{\beta}_i - \boldsymbol{\mu}_\beta) \right\}.$$

Nesse caso em que os  $\boldsymbol{\beta}_i$ 's são normalmente distribuídos, as funções LME e NLME, implementadas no Software R (R Development Core Team, 2005) no pacote NLME Pinheiro *et al.* (2008), podem ser utilizadas para cálculo das estimativas de máxima verossimilhança aproximadas dos parâmetros  $\boldsymbol{\alpha}$ ,  $\boldsymbol{\theta} = \left( \boldsymbol{\mu}_\beta, \text{vech}(\boldsymbol{\Sigma}_\beta) \right)'$  e  $\sigma_\varepsilon^2$  (Pinheiro e Bates, 2000).

Assim, das estimativas de  $\boldsymbol{\alpha}$ ,  $\boldsymbol{\theta} = \left( \boldsymbol{\mu}_\beta, \text{vech}(\boldsymbol{\Sigma}_\beta) \right)'$  e  $\sigma_\varepsilon^2$ ,  $F_{T|\boldsymbol{\alpha}, \boldsymbol{\theta}, \boldsymbol{\Sigma}_\beta}(t|\boldsymbol{\alpha}, \boldsymbol{\theta}, \boldsymbol{\Sigma}_\beta)$  podem ser obtidos via integração direta.

Contudo, o tempo computacional para avaliação das integrais irá aumentar exponencialmente com o aumento das dimensões das integrais. Uma alternativa é calcular  $F_{T|\boldsymbol{\alpha}, \boldsymbol{\theta}, \boldsymbol{\Sigma}_\beta}(t|\boldsymbol{\alpha}, \boldsymbol{\theta}, \boldsymbol{\Sigma}_\beta)$  via simulação de Monte Carlo, de forma bastante similar àquela apresentada na etapa de estimação pontual de  $F_{T|\boldsymbol{\alpha}, \boldsymbol{\mu}_\varphi, \boldsymbol{\Sigma}_\varphi}(t|\boldsymbol{\alpha}, \boldsymbol{\mu}_\varphi, \boldsymbol{\Sigma}_\varphi)$ , na Seção 3.1.3.

### 3.2 Inferência bayesiana

A inferência bayesiana nos permite estimar parâmetros de um modelo e avaliar a incerteza sobre eles através da distribuição *a posteriori*. A inferência mais completa sobre os parâmetros é a distribuição de probabilidade *a posteriori* atribuída a eles. Esta distribuição é obtida combinando-se a informação inicial sobre  $\boldsymbol{\psi}$  obtida de outras fontes de informação tais como estudos anteriores, conhecimento do especialista sobre o processo, etc, com a informação contida nos dados amostrais, fruto do experimento realizado. A informação inicial é resumida pela distribuição *a priori* que, em geral, é uma função densidade de probabilidade  $f_{\boldsymbol{\Psi}}(\boldsymbol{\psi})$ . Esta distribuição é uma escolha subjetiva e descreve o conhecimento inicial (antes de se observar os dados) do especialista sobre  $\boldsymbol{\psi}$ . A informação fornecida pelos dados sobre  $\boldsymbol{\psi}$  é resumida pela função de verossimilhança  $f_{\mathbf{Y}|\boldsymbol{\Psi}}(\mathbf{y}|\boldsymbol{\psi}) = f_{Y_{11}, \dots, Y_{1m_1}, \dots, Y_{n1}, \dots, Y_{nm_n}|\boldsymbol{\Psi}}(y_{11}, \dots, y_{1m_1}, \dots, y_{n1}, \dots, y_{nm_n}|\boldsymbol{\psi})$ . A informação combinada é então descrita por outra distribuição de probabilidade  $f_{\boldsymbol{\Psi}|\mathbf{Y}}(\boldsymbol{\psi}|\mathbf{y})$ , chamada de distribuição *a posteriori* de  $\boldsymbol{\psi}$ . A distribuição *a posteriori* é uma atualização da distribuição *a priori* pela informação trazida pela amostra e é, em geral, obtida via teorema de Bayes, o qual fornece que a distribuição *a posteriori* é dada por

$$f_{\boldsymbol{\Psi}|\mathbf{Y}}(\boldsymbol{\psi}|\mathbf{y}) = \frac{f_{\mathbf{Y}|\boldsymbol{\Psi}}(\mathbf{y}|\boldsymbol{\psi})f_{\boldsymbol{\Psi}}(\boldsymbol{\psi})}{f_{\mathbf{Y}}(\mathbf{y})}, \quad (3.6)$$

em que  $f_{\mathbf{Y}}(\mathbf{y}) = \int_{\Xi_{\boldsymbol{\psi}}} f_{\mathbf{Y}|\boldsymbol{\Psi}}(\mathbf{y}|\boldsymbol{\psi})f_{\boldsymbol{\Psi}}(\boldsymbol{\psi})d\boldsymbol{\psi}$  é a distribuição preditiva *a priori* de  $\mathbf{y}$ , e  $\Xi_{\boldsymbol{\psi}}$  é o espaço paramétrico de  $\boldsymbol{\psi}$ .

O maior problema da inferência bayesiana reside em calcularmos a integral no denominador da equação (3.6). Todavia, com os avanços dos métodos computacionais tornou-se fácil obtermos amostras de  $f_{\boldsymbol{\Psi}|\mathbf{Y}}(\boldsymbol{\psi}|\mathbf{y})$ . Geman e Geman (1984), Gelfand e Smith (1990), Casella e George (1992) e Chib e Greenberg (1995) são exemplos de trabalhos que mostram como conseguirmos amostras da distribuição *a posteriori* através de simulação Monte Carlo via cadeias de Markov. Uma apresentação detalhada dos métodos MCMC podem ser encontrados, por exemplo, em Gamerman e Lopes (2006).

Para dados de degradação, a estimação da distribuição dos parâmetros do modelo, exclusivamente, é insuficiente para obtermos a distribuição dos tempos de falha das unidades sob teste e das unidades futuras que não se encontram no escopo da amostra, ou seja, alguma unidade qualquer da população. Para este fim, torna-se necessário a exploração da relação entre os parâmetros e tempos de medição, estabelecida pelo modelo (1.3). Usando essa relação, a distribuição dos tempos de falha poderá ser obtida através de  $f_{\Psi|\mathbf{Y}}(\psi|\mathbf{y})$ .

No que segue, brevemente, apresentamos alguns métodos sugeridos na literatura para inferirmos sobre o tempo até a falha em modelos de degradação. Começamos apresentando o método proposto por Hamada (2005) e discutimos suas limitações.

### 3.2.1 Método proposto por Hamada (2005)

O primeiro passo da análise efetuada por Hamada (2005) consiste em encontrarmos a relação existente entre a distribuição dos efeitos aleatórios e a distribuição dos tempos de falha a partir da forma funcional do modelo de degradação, assim como é feito no método analítico apresentado na Seção 3.1.2. A diferença na proposta de Hamada (2005) consiste na estimação dos parâmetros via inferência bayesiana.

Posteriormente, de posse das estimativas dos parâmetros, obtém-se as estimativas das funções fração de falha,  $F_{T_{n+1}|\mathbf{Y}}(t|\mathbf{y})$ , confiabilidade,  $R_{T_{n+1}|\mathbf{Y}}(t|\mathbf{y})$ , tempo médio até a falha,  $E[T_{n+1}|\mathbf{y}]$  e percentis de ordem  $p_{n+1}$ ,  $t_{p_{n+1}}$  para unidades futuras. Hamada (2005) apresenta como obter tais quantidades a partir das amostras *a posteriori* dos parâmetros, no entanto não apresenta quaisquer justificativas teóricas para a obtenção de tais quantidades. Neste texto a obtenção das mesmas é justificada com maior detalhamento técnico.

**Estimativas *a posteriori* das quantidades de interesse segundo Hamada (2005)**

Denote por  $QIPR_{T_{n+1}|\boldsymbol{\theta}}(t)$  a quantidade de interesse obtida a partir da distribuição *a priori* do tempo até a Falha, condicional no efeito aleatório, tais como percentis, função fração de falha, confiabilidade, tempo médio até a falha, etc.

Considere uma amostra  $\boldsymbol{\theta}_1, \dots, \boldsymbol{\theta}_s$  dos parâmetros da distribuição do efeito aleatório. Hamada (2005) propõe obtermos a estimativa *a posteriori* das quantidades de interesse sobre o tempo até a falha  $T_{n+1}$  de uma unidade futura ( $QIPOST_{T_{n+1}|\mathbf{y}}(\bullet)$ ) a partir do procedimento apresentado na Tabela 3.1 e na equação (3.7).

Tabela 3.1: Obtenção da amostra de  $QIPOST$  da distribuição preditiva *a posteriori* de T via Hamada (2005)

Amostra <i>a posteriori</i>	Quantidade de interesse
$\boldsymbol{\theta}_1$	$QIPR_{T_{n+1} \boldsymbol{\theta}_1}(t)$
$\vdots$	$\vdots$
$\boldsymbol{\theta}_s$	$QIPR_{T_{n+1} \boldsymbol{\theta}_s}(t)$

De posse da amostra de  $QIPR_{T_{n+1}|\boldsymbol{\theta}}(\bullet)$ , usando-se o método *plug-in* descrito na Tabela 3.1, a distribuição *a posteriori* da quantidade de interesse  $QIPOST_{T_{n+1}|\mathbf{y}}(t|\mathbf{y})$  é dada por

$$QIPOST_{T_{n+1}|\mathbf{y}}(t|\mathbf{y}) = \sum_{k=1}^s \frac{QIPR_{T_{n+1}|\boldsymbol{\theta}_k}(t)}{s} \dots \tag{3.7}$$

Na seguinte proposição apresentamos uma prova formal de que o procedimento sugerido por Hamada (2005) fornece corretamente o resumo da distribuição preditiva *a posteriori* para o tempo de falha  $T_{n+1}$  quando as quantidades de interesse são a fração de falha *a posteriori*, a confiabilidade *a posteriori*, e o tempo médio até a falha *a posteriori*.

**Proposição 2.** *Seja o modelo de degradação  $Y_{ij} = D(t_{ij}; \boldsymbol{\beta}_i) + \varepsilon_{ij}$ . Sejam  $\boldsymbol{\beta} = (\beta_1, \dots, \beta_n)$  e  $\beta_{n+1}$  os efeitos aleatórios em que  $\beta_i$  está associado ao tempo  $T_i$ . Assuma*



que, dado  $\boldsymbol{\theta}$ ,  $\beta_1, \dots, \beta_n, \beta_{n+1}$  são iid com fdp  $f(\beta_i|\boldsymbol{\theta})$  e seja  $f(\boldsymbol{\theta})$  a fdp da distribuição a priori para  $\boldsymbol{\theta}$ . Admita que é possível encontrar uma forma fechada para a distribuição do tempo de falha  $T_i$  a partir da de  $\beta_i$ , então

$$(i) F_{T_{n+1}|\mathbf{Y}}(t|\mathbf{y}) = E_{\boldsymbol{\theta}|\mathbf{Y}} \left( F_{T_{n+1}|\boldsymbol{\theta}}(t|\boldsymbol{\theta}) \right).$$

$$(ii) R_{T_{n+1}|\mathbf{Y}}(t|\mathbf{y}) = E_{\boldsymbol{\theta}|\mathbf{Y}} \left( R_{T_{n+1}|\boldsymbol{\theta}}(t|\boldsymbol{\theta}) \right).$$

$$(iii) E[T_{n+1}|\mathbf{y}] = E_{\boldsymbol{\theta}|\mathbf{Y}} \left[ E_{T_{n+1}|\boldsymbol{\theta}}[t|\boldsymbol{\theta}] \right].$$

*Demonstração.* (i) Usando as propriedades de medidas de probabilidade, e as hipóteses, segue que a fração de falha *a posteriori* é dada por

$$\begin{aligned} F_{T_{n+1}|\mathbf{Y}}(t|\mathbf{y}) &= \int_{-\infty}^t f_{T_{n+1}|\mathbf{Y}}(a|\mathbf{y}) da \\ &= \int_{-\infty}^t \int_{\Xi_{\boldsymbol{\beta}}} \int_{\Xi_{\beta_{n+1}}} \int_{\Xi_{\boldsymbol{\theta}}} \frac{f_{T_{n+1}, \mathbf{Y}, \boldsymbol{\beta}, \beta_{n+1}, \boldsymbol{\theta}}(a, \mathbf{y}, \boldsymbol{\beta}, \beta_{n+1}, \boldsymbol{\theta})}{f_{\mathbf{Y}}(\mathbf{y})} d\boldsymbol{\theta} d\beta_{n+1} d\boldsymbol{\beta} da \\ &= \int_{-\infty}^t \int_{\Xi_{\boldsymbol{\beta}}} \int_{\Xi_{\beta_{n+1}}} \int_{\Xi_{\boldsymbol{\theta}}} f_{T_{n+1}|\mathbf{Y}, \boldsymbol{\beta}, \beta_{n+1}, \boldsymbol{\theta}}(a|\mathbf{y}, \boldsymbol{\beta}, \beta_{n+1}, \boldsymbol{\theta}) \\ &\quad \times \frac{f_{\mathbf{Y}|\boldsymbol{\beta}, \beta_{n+1}, \boldsymbol{\theta}}(\mathbf{y}|\boldsymbol{\beta}, \beta_{n+1}, \boldsymbol{\theta}) f_{\boldsymbol{\beta}, \beta_{n+1}, \boldsymbol{\theta}}(\boldsymbol{\beta}, \beta_{n+1}, \boldsymbol{\theta})}{f_{\mathbf{Y}}(\mathbf{y})} d\boldsymbol{\theta} d\beta_{n+1} d\boldsymbol{\beta} da \\ &= \int_{-\infty}^t \int_{\Xi_{\boldsymbol{\beta}}} \int_{\Xi_{\beta_{n+1}}} \int_{\Xi_{\boldsymbol{\theta}}} f_{T_{n+1}|\beta_{n+1}, \boldsymbol{\theta}}(a|\beta_{n+1}, \boldsymbol{\theta}) \\ &\quad \times \frac{f_{\mathbf{Y}|\boldsymbol{\beta}, \boldsymbol{\theta}}(\mathbf{y}|\boldsymbol{\beta}, \boldsymbol{\theta}) f_{\boldsymbol{\beta}}(\boldsymbol{\beta}|\boldsymbol{\theta}) f_{\boldsymbol{\theta}}(\boldsymbol{\theta})}{f_{\mathbf{Y}}(\mathbf{y})} f_{\beta_{n+1}|\boldsymbol{\theta}}(\beta_{n+1}|\boldsymbol{\theta}) d\boldsymbol{\theta} d\beta_{n+1} d\boldsymbol{\beta} da \\ &= \int_{-\infty}^t \int_{\Xi_{\boldsymbol{\theta}}} \int_{\Xi_{\boldsymbol{\beta}}} f_{\boldsymbol{\beta}, \boldsymbol{\theta}|\mathbf{Y}}(\boldsymbol{\beta}, \boldsymbol{\theta}|\mathbf{y}) \\ &\quad \times \int_{\Xi_{\beta_{n+1}}} f_{T_{n+1}|\beta_{n+1}, \boldsymbol{\theta}}(a|\beta_{n+1}, \boldsymbol{\theta}) f_{\beta_{n+1}|\boldsymbol{\theta}}(\beta_{n+1}|\boldsymbol{\theta}) d\beta_{n+1} d\boldsymbol{\beta} d\boldsymbol{\theta} da \\ &= \int_{\Xi_{\boldsymbol{\theta}}} \int_{\Xi_{\boldsymbol{\beta}}} f_{\boldsymbol{\beta}, \boldsymbol{\theta}|\mathbf{Y}}(\boldsymbol{\beta}, \boldsymbol{\theta}|\mathbf{y}) \int_{-\infty}^t f_{T_{n+1}|\boldsymbol{\theta}}(a|\boldsymbol{\theta}) da d\boldsymbol{\beta} d\boldsymbol{\theta} \\ &= \int_{\Xi_{\boldsymbol{\theta}}} \int_{\Xi_{\boldsymbol{\beta}}} f_{\boldsymbol{\beta}, \boldsymbol{\theta}|\mathbf{Y}}(\boldsymbol{\beta}, \boldsymbol{\theta}|\mathbf{y}) F_{T_{n+1}|\boldsymbol{\theta}}(t|\boldsymbol{\theta}) d\boldsymbol{\beta} d\boldsymbol{\theta} \\ &= \int_{\Xi_{\boldsymbol{\theta}}} f_{\boldsymbol{\theta}|\mathbf{Y}}(\boldsymbol{\theta}|\mathbf{y}) F_{T_{n+1}|\boldsymbol{\theta}}(t|\boldsymbol{\theta}) d\boldsymbol{\theta}, \end{aligned}$$

$$= E_{\boldsymbol{\theta}|\mathbf{Y}} \left( F_{T_{n+1}|\boldsymbol{\theta}}(t|\boldsymbol{\theta}) \right).$$

o que conclui a prova do item (i). □

As provas dos itens (ii) e (iii) são muito similares à do item (i) e por esta razão encontram-se no Anexo A. É importante salientar que, até onde se tem conhecimento, a prova da Proposição 2 não consta em nenhum outro trabalho que não neste texto.

### **Crítica ao procedimento de Hamada (2005)**

Além das quantidades mencionadas na Proposição 2, Hamada (2005) sugere usarmos o método descrito na Seção 3.2.1 para obtermos os percentis da distribuição *a posteriori* de  $T_{n+1}$ .

Denote por  $t_p(\boldsymbol{\theta})$  o percentil de ordem  $p$  da distribuição  $T_{n+1}$  dado  $\boldsymbol{\theta}$ . Assim,  $t_p(\boldsymbol{\theta})$  satisfaz a relação

$$t_p(\boldsymbol{\theta}) = F_{T_{n+1}|\boldsymbol{\theta}}^{-1}(p).$$

Segundo Hamada (2005) o percentil de ordem  $p$  da distribuição *a posteriori* de  $T_{n+1}$ , denotado por  $t_p^*$ , é dado por

$$\begin{aligned} t_p^* &= E_{\boldsymbol{\theta}|\mathbf{Y}} \left( F_{T_{n+1}|\boldsymbol{\theta}}^{-1}(p) \right) \\ &= E_{\boldsymbol{\theta}|\mathbf{Y}} (t_p(\boldsymbol{\theta})). \end{aligned} \tag{3.8}$$

Note que uma vez que o percentil da distribuição de  $T_{n+1}$ , dado  $\boldsymbol{\theta}$ , depende de  $\boldsymbol{\theta}$ , a igualdade em (3.8) sugere que o percentil de ordem  $p$  da distribuição de  $T_{n+1}$ , dado  $\mathbf{y}$ , seja uma média ponderada, dos percentis de ordem  $p$  das distribuições de  $T_{n+1}$  dado  $\boldsymbol{\theta}$ , pela distribuição *a posteriori* de  $\boldsymbol{\theta}$ .

O que se pode provar, no entanto, é que  $t_p^*$  é tal que

$$\begin{aligned} p &= F_{T_{n+1}|\mathbf{y}}(t_p^*) \\ &= \int_{\boldsymbol{\theta}} f(\boldsymbol{\theta}|\mathbf{y}) F_{T_{n+1}|\boldsymbol{\theta}}(t_p^*) d\boldsymbol{\theta} \end{aligned}$$

$$= E_{\boldsymbol{\theta}|\mathbf{y}} \left( F_{T_{n+1}|\boldsymbol{\theta}}(t_p^*) \right). \quad (3.9)$$

O seguinte exemplo mostra que o método proposto por Hamada (2005) para o cálculo do percentil da distribuição *a posteriori* não é válido.

Suponha

$$T|\lambda \sim \exp(\lambda),$$

$$\lambda = \begin{cases} 0,5, & \text{com probabilidade } 1/2 \\ 2, & \text{com probabilidade } 1/2 \end{cases}.$$

A mediana de  $T|\lambda = 0,5$  é

$$0,5 = \int_0^{M_1} e^{-t} dt$$

$$0,5 = 1 - e^{-0,5M_1}$$

$$M_1 = \frac{\log(2)}{0,5} = 1,386,$$

e a mediana de  $T|\lambda = 2$  é

$$0,5 = 1 - e^{-2M_2}$$

$$M_2 = \frac{\log(2)}{2} = 0,347.$$

De (3.8), a combinação linear destas medianas é a estimativa proposta por Hamada (2005):

$$M^* = \frac{1}{2}(M_1 + M_2) = \frac{1}{2}(1,386 + 0,347) = 0,866. \quad (3.10)$$

Agora, de (3.9) segue que

$$F(T) = \frac{1}{2} [(1 - e^{-0,5t}) + (1 - e^{-2t})]$$

$$= \frac{1}{2} (2 - e^{-0,5t} - e^{-2t}). \quad (3.11)$$

Assim, a mediana é a solução para o valor de  $M$  tal que

$$0,5 = \frac{1}{2} (2 - e^{-0,5M} - e^{-2M})$$

$$e^{-0,001M} + e^{-2M} = 1$$

$$M \approx 0,645, \quad (3.12)$$

que não tem solução analítica única nem real, de modo que foi fornecida uma solução

numérica aproximada.

Observe que a solução em (3.10) é diferente da solução em (3.12), portanto a proposta de Hamada (2005) não é coerente, pois o verdadeiro valor da mediana, obtido a partir da distribuição preditiva *a priori* de  $T$ , é 0,645.

Note pela Figura 3.1 que o percentil estimado pela proposta de Hamada (2005) fornece uma estimativa de distribuição acumulada  $p = 0,587$  diferente da ordem do percentil requerido (figura à esquerda), quando o percentil deveria fornecer o verdadeiro valor 0.5 da fda sob a distribuição preditiva (figura à direita).

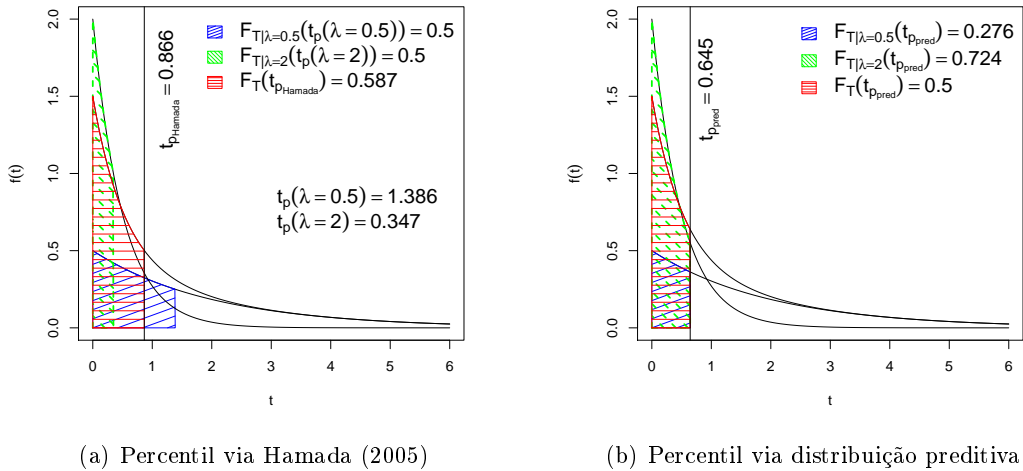


Figura 3.1: Estimativa proposta por Hamada (2005) *versus* estimativa via distribuição preditiva

Em outras palavras, a média dos percentis de ordem  $p$  não resulta no mesmo valor do percentil de mesma ordem sobre a distribuição preditiva dos tempos de falha. Isso só seria verdade caso a distribuição dos parâmetros do modelo bayesiano fosse degenerada em algum ponto.

A Seção 3.2.3 apresenta uma solução proposta por Oliveira (2011) para encontrar o percentil da distribuição preditiva dos tempos de falha a partir da distribuição *a posteriori* dos efeitos aleatórios.

Além do aspecto teórico, também são notadas algumas incoerências interpretativas na abordagem de Hamada (2005), as quais são comentados na Seção 3.2.5. A título de revisão literária também apresentamos brevemente o método bayesiano proposto por Robinson e Crowder (2000), que apesar de também apresentar uma incoerência interpretativa com relação ao paradigma bayesiano e ser limitado noutros aspectos (esclarecimentos nas Seções 3.2.2 e 3.2.5, respectivamente), permite maior flexibilidade na modelagem de degradação para obtenção da distribuição dos tempos de falha.

### 3.2.2 Método de Robinson e Crowder (2000)

Para obter as quantidades de interesse relacionadas ao tempo até a falha  $T_{n+1}$  de uma nova unidade, Robinson e Crowder (2000) assumem que a distribuição *a posteriori* de  $\boldsymbol{\beta}|\mathbf{y}$  é considerada como a distribuição *a posteriori* para  $\beta_{n+1}|\mathbf{y}$  de um objeto qualquer da população.

**Proposição 3.** *Considere a forma  $Y_{ij} = D(t_{ij}; \boldsymbol{\beta}_i) + \varepsilon_{ij}$  do modelo geral de degradação da equação (1.3) e assuma  $\varepsilon_{ij} \stackrel{iid}{\sim} N(0, \sigma_\varepsilon^2)$ . Sejam  $\boldsymbol{\beta} = (\beta_1, \dots, \beta_n)$  e  $\beta_{n+1}$  os efeitos aleatórios. Assuma que, dado  $\boldsymbol{\theta}$ ,  $\beta_1, \dots, \beta_n, \beta_{n+1}$  são iid com distribuição  $f(\beta_i|\boldsymbol{\theta})$  e seja  $f(\boldsymbol{\theta})$  a distribuição a priori para  $\boldsymbol{\theta}$ . Então a fda da distribuição preditiva a posteriori dos tempos de falha para uma unidade genérica da população é dada por*

$$F_{T_{n+1}|\mathbf{Y}}(t|\mathbf{y}) = E_{\beta_{n+1}, \sigma_\varepsilon^2|\mathbf{y}} \left[ 1 - \Phi \left( \frac{D_f - D(t_{n+1,j}; \beta_{n+1})}{\sigma_\varepsilon} \right) \right], \quad (3.13)$$

em que  $\Phi(\bullet)$  denota a fda da distribuição normal padrão.

*Demonstração.* Similar ao que ocorre na Proposição 2, como  $T_{n+1}$  é independente de  $T_1, \dots, T_n$ , dado  $\boldsymbol{\theta}$  e  $\sigma_\varepsilon^2$  segue que

$$F_{T_{n+1}|\mathbf{Y}}(t|\mathbf{y}) = E_{\beta_{n+1}, \sigma_\varepsilon^2|\mathbf{y}} [F_{T_{n+1}|\beta_{n+1}, \sigma_\varepsilon^2}(t|\beta_{n+1}, \sigma_\varepsilon^2)]. \quad (3.14)$$

Como  $\varepsilon_{ij} \stackrel{iid}{\sim} N(0, \sigma_\varepsilon^2)$ , segue que  $D(t_{n+1,j}; \beta_{n+1}) + \varepsilon_{n+1,j} \stackrel{iid}{\sim} N(D(t_{n+1,j}; \beta_{n+1}), \sigma_\varepsilon^2)$ .

Daí, temos que

$$\begin{aligned} F_{T_{n+1}|\beta_{n+1},\sigma_\varepsilon^2}(t|\beta_{n+1},\sigma_\varepsilon^2) &= P(D(t_{n+1,j};\beta_{n+1}) + \varepsilon_{n+1,j} \geq D_f|\beta_{n+1},\sigma_\varepsilon^2) \\ &= 1 - \Phi\left(\frac{D_f - D(t_{n+1,j};\beta_{n+1})}{\sigma_\varepsilon}\right). \end{aligned} \quad (3.15)$$

□

Usando-se o resultado da Proposição 3, uma amostra *a posteriori* da fda *a priori* para o tempo até a falha é obtida como descrito na Tabela 3.2:

Tabela 3.2: Amostra de  $F_{T_{n+1}|\beta_{n+1},\sigma_\varepsilon^2}(\bullet)$  pelo método de Robinson e Crowder (2000)

Amostra de $\beta, \theta, \sigma_\varepsilon^2   \mathbf{y}$					Amostra de $F_{T_{n+1} \beta_{n+1},\sigma_\varepsilon^2}(t \beta_{n+1},\sigma_\varepsilon^2)$			
$(\beta_{11}$	$\cdots$	$\beta_{n1}$	$\theta_1$	$\sigma_{\varepsilon 1}^2)$	$ \mathbf{y}$	$1 - \Phi\left(\frac{D_f - D(t;\beta_{11})}{\sigma_\varepsilon}\right)$	$\cdots$	$1 - \Phi\left(\frac{D_f - D(t;\beta_{n1})}{\sigma_\varepsilon}\right)$
$\vdots$	$\ddots$	$\vdots$	$\vdots$	$\vdots$		$\vdots$	$\ddots$	$\vdots$
$(\beta_{1s}$	$\cdots$	$\beta_{ns}$	$\theta_s$	$\sigma_{\varepsilon s}^2)$	$ \mathbf{y}$	$1 - \Phi\left(\frac{D_f - D(t;\beta_{n1})}{\sigma_\varepsilon}\right)$	$\cdots$	$1 - \Phi\left(\frac{D_f - D(t;\beta_{ns})}{\sigma_\varepsilon}\right)$

Considerando a amostra de  $F_{T_{n+1}|\beta_{n+1},\sigma_\varepsilon^2}(t|\beta_{n+1},\sigma_\varepsilon^2)$  obtida, a fdp da distribuição preditiva *a posteriori* para uma unidade futura, na equação (3.13) é estimada por

$$\hat{F}_{T_{n+1}|\mathbf{Y}}(t|\mathbf{y}) = \sum_{k=1}^{ns} \frac{1 - \Phi\left(\frac{D_f - D(t;\beta_{ns})}{\sigma_\varepsilon}\right)}{ns}. \quad (3.16)$$

Dada a amostra  $F_{T_{n+1}|\beta_{n+1},\sigma_\varepsilon^2}(t|\beta_{n+1},\sigma_\varepsilon^2)$  na tabela 3.2 Robinson e Crowder (2000) alega ser possível construir intervalos de credibilidade para  $F_{T_{n+1}|\mathbf{Y}}(t|\mathbf{y})$ , mas tal alegação recai no mesmo problema interpretativo de Hamada (2005), o qual será comentado na Seção 3.2.5.

Robinson e Crowder (2000) não apresentam uma maneira de estimar outras quantidades de interesse. A vantagem deste método, em relação ao método proposto por Hamada (2005), reside no fato de que utilizam a forma funcional do modelo de degradação e não a relação entre  $\beta$  e  $T$  estabelecida *a priori* para obterem a fda *a posteriori* de um objeto futuro. Além disso, é possível obtermos a distribuição do tempo de falha de cada perfil individual e os perfis podem assumir quaisquer formas funcionais, inclusive não lineares.

### 3.2.3 Distribuição preditiva *a posteriori* de $T_{n+1}$

Para obtermos as quantidades de interesse relacionadas ao tempo até a falha de uma nova unidade independente das demais já observadas, Oliveira (2011) obtém a distribuição preditiva *a posteriori* de  $T_{n+1}$  dadas as informações sobre a degradação das  $n$  unidades observadas.

Como já provamos na Proposição 2, a fda associada à distribuição preditiva *a posteriori* de  $T_{n+1}$  é dada por

$$F_{T_{n+1}|\mathbf{Y},\mathbf{Y}_{n+1}}(t) = E_{\beta_{n+1}|\mathbf{Y},\mathbf{Y}_{n+1}} \left( F_{T_{n+1}|\boldsymbol{\theta}}(t|\boldsymbol{\theta}) \right).$$

Partindo deste fato e considerando que  $D = t/\beta$ , podemos provar que

$$\begin{aligned} F_{T_{n+1}|\mathbf{Y},\mathbf{Y}_{n+1}}(t) &= \int_{-\infty}^t f_{T_{n+1}|\mathbf{Y},\mathbf{Y}_{n+1}}(a|\mathbf{y},\mathbf{y}_{n+1}) da \\ &= \int_{-\infty}^t \frac{f_{T_{n+1},\mathbf{Y},\mathbf{Y}_{n+1}}(a,\mathbf{y},\mathbf{y}_{n+1})}{f_{\mathbf{Y},\mathbf{Y}_{n+1}}(\mathbf{y},\mathbf{y}_{n+1})} da \\ &= \int_{-\infty}^t \int_{\Xi_{\boldsymbol{\beta}}} \int_{\Xi_{\beta_{n+1}}} \frac{f_{T_{n+1},\mathbf{Y},\mathbf{Y}_{n+1},\boldsymbol{\beta},\beta_{n+1}}(a,\mathbf{y},\mathbf{y}_{n+1},\boldsymbol{\beta},\beta_{n+1})}{f_{\mathbf{Y},\mathbf{Y}_{n+1}}(\mathbf{y},\mathbf{y}_{n+1})} d\boldsymbol{\beta} d\beta_{n+1} da \\ &= \int_{-\infty}^t \int_{\Xi_{\boldsymbol{\beta}}} \int_{\Xi_{\beta_{n+1}}} \frac{f_{\boldsymbol{\beta},\beta_{n+1},\mathbf{Y},\mathbf{Y}_{n+1}}(\boldsymbol{\beta},\beta_{n+1},\mathbf{y},\mathbf{y}_{n+1})}{f_{\mathbf{Y},\mathbf{Y}_{n+1}}(\mathbf{y},\mathbf{y}_{n+1})} \\ &\quad \times f_{T_{n+1}|\mathbf{Y}_{n+1},\beta_{n+1}}(a|\mathbf{y}_{n+1},\beta_{n+1}) d\boldsymbol{\beta} d\beta_{n+1} da \\ &= \int_{\Xi_{\boldsymbol{\beta}}} \int_{\Xi_{\beta_{n+1}}} f_{\boldsymbol{\beta},\beta_{n+1}|\mathbf{Y},\mathbf{Y}_{n+1}}(\boldsymbol{\beta},\beta_{n+1}|\mathbf{y},\mathbf{y}_{n+1}) \\ &\quad \times \int_{-\infty}^t f_{T_{n+1}|\mathbf{Y}_{n+1},\beta_{n+1}}(a|\mathbf{y}_{n+1},\beta_{n+1}) da d\boldsymbol{\beta} d\beta_{n+1} \\ &= \int_{\Xi_{\boldsymbol{\beta}}} \int_{\Xi_{\beta_{n+1}}} f_{\boldsymbol{\beta},\beta_{n+1}|\mathbf{Y},\mathbf{Y}_{n+1}}(\boldsymbol{\beta},\beta_{n+1}|\mathbf{y},\mathbf{y}_{n+1}) P(\beta_{n+1} D_f \leq t | \mathbf{y}_{n+1}, \beta_{n+1}) d\boldsymbol{\beta} d\beta_{n+1} \end{aligned}$$

$$\begin{aligned}
 &= \int_{\Xi_{\beta_{n+1}}} P(\beta_{n+1} D_f \leq t | \mathbf{y}_{n+1}, \beta_{n+1}) f_{\beta_{n+1} | \mathbf{Y}, \mathbf{Y}_{n+1}}(\beta_{n+1} | \mathbf{y}, \mathbf{y}_{n+1}) d\beta_{n+1} \\
 &= \int_{\Xi_{\beta_{n+1}}} P\left(\beta_{n+1} \leq \frac{t}{D_f} \middle| \mathbf{y}_{n+1}, \beta_{n+1}\right) f_{\beta_{n+1} | \mathbf{Y}, \mathbf{Y}_{n+1}}(\beta_{n+1} | \mathbf{y}, \mathbf{y}_{n+1}) d\beta_{n+1} \\
 &= \int_{\Xi_{\beta_{n+1}}} I_{\left\{\beta_{n+1} \leq \frac{t}{D_f} \middle| \mathbf{y}_{n+1}, \beta_{n+1}\right\}} f_{\beta_{n+1} | \mathbf{Y}, \mathbf{Y}_{n+1}}(\beta_{n+1} | \mathbf{y}, \mathbf{y}_{n+1}) d\beta_{n+1} \\
 &= F_{\beta_{n+1} | \mathbf{Y}, \mathbf{Y}_{n+1}}\left(\frac{t}{D_f} \middle| \mathbf{y}, \mathbf{y}_{n+1}\right), \tag{3.17}
 \end{aligned}$$

de onde vem que  $f_{T_{n+1} | \mathbf{Y}, \mathbf{Y}_{n+1}}(t | \mathbf{y}, \mathbf{y}_{n+1}) = f_{\beta_{n+1} | \mathbf{Y}, \mathbf{Y}_{n+1}}\left(\frac{t}{D_f} \middle| \mathbf{y}, \mathbf{y}_{n+1}\right) \frac{1}{D_f}$ .

Sendo assim, de posse de uma amostra de tamanho  $s$  da distribuição *a posteriori* de  $\beta_{n+1}$ , podemos estimar  $F_{T_{n+1} | \mathbf{Y}, \mathbf{Y}_{n+1}}(t | \mathbf{y}, \mathbf{y}_{n+1})$  da seguinte forma

$$\hat{F}_{T_{n+1} | \mathbf{Y}, \mathbf{Y}_{n+1}}(t | \mathbf{y}, \mathbf{y}_{n+1}) = \frac{\sum_{k=1}^s I_{\left\{\beta_{n+1} \leq \frac{t}{D_f}\right\}}}{s}. \tag{3.18}$$

Esta nova proposta é bem atrativa por ser de fácil aplicação e analiticamente justificável. Por exemplo, o tempo médio até a falha *a posteriori*  $E_{T_{n+1} | \mathbf{Y}, \mathbf{Y}_{n+1}}(t | \mathbf{y}, \mathbf{y}_{n+1})$  é dado por

$$\begin{aligned}
 E_{T_{n+1} | \mathbf{Y}, \mathbf{Y}_{n+1}}(t | \mathbf{y}, \mathbf{y}_{n+1}) &= \int_{\Xi_{T_{n+1}}} t f_{T_{n+1} | \mathbf{Y}, \mathbf{Y}_{n+1}}(t | \mathbf{y}, \mathbf{y}_{n+1}) dt \\
 &= \int_{\Xi_{\beta_{n+1}}} t f_{\beta_{n+1} | \mathbf{Y}, \mathbf{Y}_{n+1}}\left(\frac{t}{D_f} \middle| \mathbf{y}, \mathbf{y}_{n+1}\right) \frac{1}{D_f} dt. \tag{3.19}
 \end{aligned}$$

A equação (3.19) pode ser aproximada pela equação (3.20) a partir de uma amostra *a posteriori* de  $\beta_{n+1}$ , de modo que se  $D = \beta/t$  então  $T_{n+1} = \beta_{n+1} D_f$ .

$$\hat{E}_{T_{n+1} | \mathbf{Y}, \mathbf{Y}_{n+1}}(t | \mathbf{y}, \mathbf{y}_{n+1}) = \frac{\sum_{k=1}^s T_{n+1,s}}{s}. \tag{3.20}$$

Sempre será possível aproximar o valor de uma quantidade de interesse com base na amostra de  $T_{n+1} | \mathbf{Y}_{n+1}$ , pois teoricamente foi gerada a partir da distribuição correta com base na inferência obtida para o modelo de degradação. Vale ressaltar que o procedimento de estimação da proposta desta seção é feita através de uma média aritmética, uma função de custo computacional muito baixo. Isto acarreta um tempo



computacional muito inferior aos métodos propostos por Hamada (2005) e Robinson e Crowder (2000) no momento da estimação das quantidades de interesse.

Para obtermos uma amostra da distribuição *a posteriori* de  $T_{n+1}$  é necessário obtermos uma amostra diretamente de  $\beta_{n+1}|\mathbf{y}, \mathbf{y}_{n+1}$ , um indivíduo da população que carregue consigo toda diversidade possível da mesma. Oliveira (2011) indica uma forma de obter essa amostra, entretanto para um caso muito particular, em que  $y_{n+1,1} = 0$ . Na próxima seção descreveremos um procedimento para amostrar da distribuição de  $\beta_{n+1}|\mathbf{y}, \mathbf{y}_{n+1}$  usando a técnica de aumento de dados.

### 3.2.4 Nova unidade na verossimilhança via aumento de dados

Para amostrar a partir de  $\beta_{n+1}|\mathbf{y}, \mathbf{y}_{n+1}$  seria necessário conhecer a degradação inicial de um indivíduo na população. Infelizmente em muitas situações esta informação não está disponível, mas é possível simulá-la via técnica de aumento de dados utilizando-se das ideias de Tanner e Wong (1987). Tanner e Wong (1987) propõem uma maneira de lidar com dados parcialmente ou não observados.

Considere o vetor de dados de degradação aumentado como sendo  $(\mathbf{Y}, \mathbf{Y}_{n+1})$ , em que  $m_{n+1} = 1$  e  $\mathbf{Y}_{n+1}$  é o conjunto de medições não observadas de um objeto futuro. Assuma que o real perfil de degradação é  $y_{ij} = \frac{t_{ij}}{\beta_i}$ . Neste caso a função de verossimilhança torna-se

$$f_{\mathbf{Y}, \mathbf{Y}_{n+1}|\boldsymbol{\beta}, \beta_{n+1}, \sigma_\varepsilon^2}(\mathbf{y}, \mathbf{y}_{n+1}|\boldsymbol{\beta}, \beta_{n+1}, \sigma_\varepsilon^2) = \prod_{i=1}^{n+1} \prod_{j=1}^{m_i} \left( \frac{1}{2\pi\sigma_\varepsilon^2} \right)^{1/2} \exp \left\{ -\frac{1}{2\sigma_\varepsilon^2} \left( y_{ij} - \frac{t_{ij}}{\beta_i} \right)^2 \right\}. \quad (3.21)$$

A degradação  $Y_{n+1,1}$  é não observada, portanto um objeto aleatório. Então,  $Y_{n+1,1}$  será gerado no procedimento inferencial como qualquer outro parâmetro. Para isto, no caso bayesiano, necessitamos obter a distribuição condicional completa *a posteriori*

para  $Y_{n+1,1}$ , a qual é

$$f_{Y_{n+1,1}}(\mathbf{Y}, \boldsymbol{\beta}, \beta_{n+1}, \sigma_\varepsilon^2)(y_{n+1,1} | \mathbf{y}, \boldsymbol{\beta}, \beta_{n+1}, \sigma_\varepsilon^2) = \left( \frac{1}{2\pi\sigma_\varepsilon^2} \right)^{1/2} \exp \left\{ -\frac{1}{2\sigma_\varepsilon^2} \left( y_{n+1,1} - \frac{t_{n+1,1}}{\beta_{n+1,1}} \right)^2 \right\}. \quad (3.22)$$

Observe que  $Y_{n+1,1} | \mathbf{Y}, \boldsymbol{\beta}, \beta_{n+1}, \sigma_\varepsilon^2 \stackrel{d}{=} Y_{n+1,1} \sim N \left( \frac{t_{n+1,1}}{\beta_{n+1,1}}, \sigma_\varepsilon^2 \right)$ .

Usando esta estratégia amostras da distribuição *a posteriori* de  $\beta_{n+1}$  são naturalmente geradas durante o processo inferencial. Assim, desta amostra e do fato que  $D = t/\beta$  obtém-se uma amostra da distribuição *a posteriori* de  $T_{n+1}$ . A inferência resultante utilizando o método proposto por Oliveira (2011) é tanto mais suavizada quanto maior for a amostra gerada de  $\beta_{n+1}$ .

É importante notar que se  $y_{i,1} = 0$  para qualquer unidade da população, não existe necessidade de se utilizar o aumento de dados para inserir um novo objeto na função de verossimilhança. Neste caso  $y_{n+1,1} = 0$ , portanto observado.

### 3.2.5 Interpretação nas abordagens bayesianas

Além das vantagens teóricas e computacionais do método apresentado na Seção 3.2.3, o mesmo possui interpretação mais simples e coerente em termos do paradigma bayesiano. Nos artigos de Hamada (2005) e Robinson e Crowder (2000) os autores deveriam interpretar apenas as estimativas pontuais das quantidades de interesse para manter a mesma coerência, porém ambos propõem erroneamente a construção de intervalos de credibilidade para estas estimativas pontuais.

Robinson e Crowder (2000) constroem amostras de  $F(t | \boldsymbol{\beta}_i, \sigma_\varepsilon^2)$  usando o método *plug-in* em cada valor amostrado da distribuição *a posteriori*  $\boldsymbol{\beta}_i, \sigma_\varepsilon^2 | \mathbf{y}$ . Hamada (2005) obtém uma amostra de  $F(t | \boldsymbol{\theta})$  avaliando cada valor amostrado da distribuição *a posteriori*  $\boldsymbol{\theta} | \mathbf{y}$  (Seção “Laser degradation example revisited” de seu artigo). Em ambos os casos as quantidades de interesse são corretamente tratadas como quantidades aleatórias, uma vez que para cada valor amostrado obtém-se valores diferentes de  $F(t | \bullet)$ , ou qualquer outra função que se queira. Observe que  $F(t | \boldsymbol{\beta}_i, \sigma_\varepsilon^2)$  e  $F(t | \boldsymbol{\theta})$  são apenas

quantidades *a priori*, dessa forma o método *plug-in* na verdade fornece uma amostra *a posteriori* destas quantidades *a priori*, não uma amostra da distribuição preditiva de  $T_{n+1}|\mathbf{Y}$ . As Proposições 2 e 3 deixam claro que  $F_{T_{n+1}|\mathbf{Y}}(t|\mathbf{y})$  e  $F_{T_{n+1}|\mathbf{Y}}(t|\mathbf{y})$  são médias das amostras *a posteriori* destas quantidades *a priori*. Portanto, ao construir intervalos de credibilidade para  $F_{T_{n+1}|\mathbf{Y}}(t|\mathbf{y})$  e  $F_{T_{n+1}|\mathbf{Y}}(t|\mathbf{y})$ , com base nas amostras *a posteriori* de  $F(t|\boldsymbol{\beta}_i, \boldsymbol{\sigma}_\varepsilon^2)$  e  $F(t|\boldsymbol{\theta})$ , Hamada (2005) e Robinson e Crowder (2000) o fazem erroneamente.

Para ilustrar a abordagem de Hamada (2005) e Robinson e Crowder (2000) suponha que se queira conhecer a probabilidade de que uma unidade não observada falhe antes do tempo “t”, isto é  $F_{T_{n+1}|\mathbf{Y}}(t|\mathbf{y})$ . Na interpretação de Hamada (2005) e Robinson e Crowder (2000), a resposta seria que a probabilidade  $F_{T_{n+1}|\mathbf{Y}}(t|\mathbf{y})$  para uma unidade não observada falhar antes do tempo “t” pertence ao intervalo  $(L_a, L_b)$  com algum nível  $\gamma$  de credibilidade, em que  $(L_a, L_b)$  é o intervalo percentílico ou de mais alta densidade com base nas amostras *a posteriori* de  $F(t|\boldsymbol{\beta}_i, \boldsymbol{\sigma}_\varepsilon^2)$  e  $F(t|\boldsymbol{\theta})$ , respectivamente.

Ora, se  $F(t|\boldsymbol{\beta}_i, \boldsymbol{\sigma}_\varepsilon^2)$  e  $F(t|\boldsymbol{\theta})$  são quantidades *a priori*, então a interpretação correta seria que a probabilidade *a priori*  $F(t|\boldsymbol{\beta}_i, \boldsymbol{\sigma}_\varepsilon^2)$  (Robinson e Crowder, 2000) ou  $F(t|\boldsymbol{\theta})$  (Hamada, 2005), avaliada na amostra *a posteriori* dos parâmetros, de uma unidade dentre as sob teste falhar antes do tempo “t” pertence ao intervalo  $(L_a, L_b)$  com algum nível  $\gamma$  de credibilidade. A interpretação fica ainda mais estranha se a tentativa é interpretar os quantis  $L_a$  e  $L_b$  do intervalo. Por exemplo,  $L_a$  é a probabilidade acumulada *a posteriori* da probabilidade *a priori*  $F(t|\boldsymbol{\beta}_i, \boldsymbol{\sigma}_\varepsilon^2)$  (Robinson e Crowder, 2000) ou  $F(t|\boldsymbol{\theta})$  (Hamada, 2005) de uma unidade dentre as sob teste falhar antes do tempo “t”. Atente para o fato de que na interpretação correta o intervalo é para unidades dentre as sob teste, já que apenas a média das amostras de  $F(t|\boldsymbol{\beta}_i, \boldsymbol{\sigma}_\varepsilon^2)$  (Robinson e Crowder, 2000) e  $F(t|\boldsymbol{\theta})$  (Hamada, 2005) é que fornecem  $F_{T_{n+1}|\mathbf{Y}}(t|\mathbf{y})$  de um item futuro.

O intervalo de credibilidade proposto por (Robinson e Crowder, 2000) e (Hamada, 2005) decorre de um pensamento coerente com inferência clássica e diz respeito a uma

quantidade de interesse *a priori*. É como se o princípio de repetibilidade da estatística clássica fosse transportado para o contexto bayesiano de uma maneira extremamente confusa e não informativa para tomada de decisão.

No mesmo contexto a interpretação das quantidades de interesse por Oliveira (2011) é muito mais simples por ser plenamente coerente com o paradigma bayesiano. A probabilidade de que uma unidade da população falhe antes do tempo “ $t$ ” é um valor único pertencente ao intervalo  $[0,1]$ . Que é a mesma interpretação fornecida para qualquer objeto aleatório na abordagem bayesiana, uma informação muito mais facilmente compreendida por qualquer pessoa que nunca tenha lidado com estatística. Vale lembrar que qualquer quantidade de interesse pode ser obtida sem prejuízo teórico metodológico e sem perder essa interpretabilidade simples. Portanto, Robinson e Crowder (2000) e Hamada (2005) deveriam se ater apenas às estimativas pontuais para uma interpretação correta dos resultados.

## Capítulo 4

# Modelos de degradação robustos

Este capítulo se destina à modelagem de degradação por um modelo linear robusto. Por modelo robusto entende-se aquele cujas estimativas dos parâmetros são menos influenciadas na presença de valores atípicos. Além dessa característica, os modelos propostos são flexíveis o suficiente para acomodarem os mais diferentes graus de assimetria.

Na Seção 4.1 apresentamos o método analítico para dois modelos lineares univariados sem intercepto, em que os efeitos aleatórios são descritos pelas distribuições nas famílias SMSN e SMLSN. Os procedimentos necessários para a estimação dos parâmetros via inferência bayesiana são apresentados na Seção 4.2. A Seção 4.3 apresenta alguns critérios de seleção e medidas de qualidade de ajuste correntes da literatura, além de contextualizá-los quando aplicados em modelos gerais de degradação. Na Seção 4.4 alguns bancos de dados artificialmente gerados são analisados a fim de estudar os modelos propostos e os critérios de seleção de modelos. Medidas de qualidade de ajuste são aplicadas para averiguarmos se são adequadas para avaliar este tipo de modelo. Finalmente uma base de dados de 3 posições de rodas de trem, e dados de emissores de laser, são analisados na Seção 4.5.

## 4.1 Método analítico

Nesta seção apresentaremos o método analítico assumindo um modelo linear sem intercepto para os caminhos de degradação. A novidade será aplicarmos o método analítico considerando-se que, para este modelo linear sem intercepto, os efeitos aleatórios são distribuídos segundo distribuições nas famílias SMSN e SMLSN. Estas distribuições podem acomodar formas diferenciadas para a distribuição dos efeitos aleatórios, o que justifica sua aplicabilidade para a análise de dados de degradação. Além disso, a possibilidade de introdução de distribuições com caudas pesadas em modelagem de degradação ainda não foi explorada na literatura.

### 4.1.1 Modelo linear com efeito aleatório SMSN e SMLSN

Suponha que o caminho de degradação  $D(t_{ij}; \beta_i)$  para a unidade  $i$  seja linear tal que

$$D(t_{ij}; \beta_i) = \frac{t_{ij}}{\beta_i}, \quad (4.1)$$

em que  $\frac{1}{\beta_i}$  representa a taxa de degradação de uma particular unidade  $i$  sob teste.

Em modelos de degradação consideramos que uma falha ocorre na unidade  $i$  quando a degradação desta unidade ultrapassa o limiar  $D_f$ . Assim, para o modelo linear em (4.1), o tempo de falha  $T_i$  da unidade  $i$  é dado por

$$T_i = D_f \beta_i. \quad (4.2)$$

A partir de (4.2) a fda do tempo até a falha  $T_i$  pode ser obtida a partir da distribuição de  $\beta_i$  da seguinte forma:

$$F(t|\boldsymbol{\theta}_\beta) = P\left(\beta_i \leq \frac{t}{D_f} \middle| \boldsymbol{\theta}_\beta\right), \quad (4.3)$$

em que  $\boldsymbol{\theta}_\beta$  é o vetor de parâmetros que indexa a distribuição  $\Lambda(\beta|\boldsymbol{\theta}_\beta)$  dos efeitos aleatórios.

**Proposição 4.** *Seja o modelo de degradação em (4.1). Se  $\beta_i \sim \text{SMSN}(\mu_\beta, h_i^{-1}\omega_\beta^2, \alpha_\beta, \mathcal{H})$ , então o tempo até a falha  $T_i \sim \text{SMSN}(\mu_\beta D_f, h_i^{-1}\omega_\beta^2 D_f^2, \alpha_\beta, \mathcal{H})$ .*

*Demonstração.* Segue diretamente das equações (2.7) e (4.3) que

$$F(t_i|\mu_\beta, \omega_\beta, \alpha_\beta, \nu_\beta) = \int_{-\infty}^{\frac{t_i}{D_f}} \int_0^\infty 2 \left( \frac{1}{2\pi h_i^{-1} \omega_\beta^2} \right)^{\frac{1}{2}} \exp \left\{ -\frac{1}{2h_i^{-1} \omega_\beta^2} (\beta_i - \mu_\beta)^2 \right\} \\ \times \Phi \left( \alpha_\beta \left( \frac{\beta_i - \mu_\beta}{h_i^{-\frac{1}{2}} \omega_\beta} \right) \right) d\mathcal{H}(h_i|\nu_\beta) d\beta_i.$$

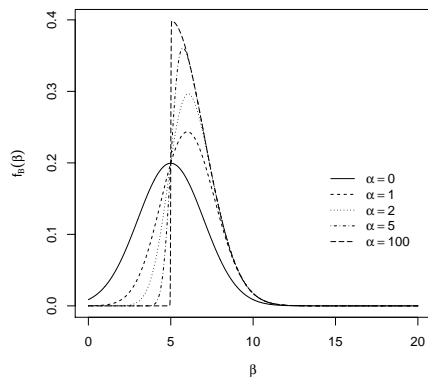
Fazendo  $t^* = \frac{t_i}{D_f}$ , como consequência segue que a fdp para  $T_i$  é

$$f(t_i|\mu_\beta, \omega_\beta, \alpha_\beta, \nu_\beta) = \frac{\partial F_{T_i|\mu_\beta, \omega_\beta, \alpha_\beta, \nu_\beta}(t_i|\mu_\beta, \omega_\beta, \alpha_\beta, \nu_\beta)}{\partial t^*} \frac{\partial t^*}{\partial t} \\ = \int_0^\infty 2 \left( \frac{1}{2\pi h_i^{-1} \omega_\beta^2} \right)^{\frac{1}{2}} \exp \left\{ -\frac{1}{2h_i^{-1} \omega_\beta^2} \left( \frac{t_i}{D_f} - \mu_\beta \right)^2 \right\} \times \\ \times \Phi \left( \alpha_\beta \left( \frac{\frac{t_i}{D_f} - \mu_\beta}{h_i^{-\frac{1}{2}} \omega_\beta} \right) \right) d\mathcal{H}(h_i|\nu_\beta) \left[ \frac{1}{D_f} \right] \\ = \int_0^\infty 2 \left( \frac{1}{2\pi h_i^{-1} \omega_\beta^2 D_f^2} \right)^{\frac{1}{2}} \exp \left\{ -\frac{1}{2h_i^{-1} \omega_\beta^2 D_f^2} (t_i - \mu_\beta D_f)^2 \right\} \times \\ \times \Phi \left( \alpha_\beta \left( \frac{t_i - \mu_\beta D_f}{h_i^{-\frac{1}{2}} \omega_\beta D_f} \right) \right) d\mathcal{H}(h_i|\nu_\beta), \quad (4.4)$$

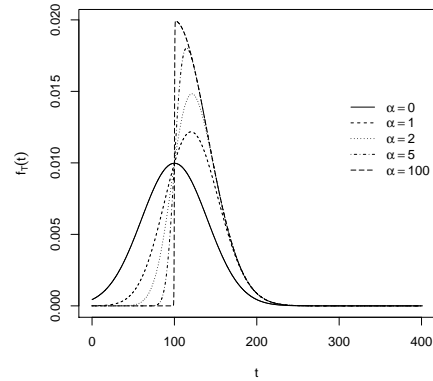
o que conclui a prova.  $\square$

É possível assumirmos diferentes distribuições de probabilidade  $\mathcal{H}$  com suporte em  $(0, \infty)$ . Como consequência, diferentes modelos de degradação podem ser definidos ao escolhermos diferentes medidas  $\mathcal{H}$ . Por exemplo, como apontado na Seção 2.1.3, se  $\mathcal{H}$  é tal que  $P(H = 1)$  recupera-se o modelo skew-normal. Se  $\mathcal{H}$  é a distribuição gamma  $Gamma(\nu/2, \nu/2)$  obtém-se o modelo skew- $t$  em que  $k_m = (\nu/2)^{m/2} ((\nu - m)/2) [(\nu/2)]^{-1}$ . Por outro lado, se  $\mathcal{H}$  é uma distribuição beta  $Beta(\nu, 1)$  retoma-se o modelo skew-slash, e  $k_m = 2\nu[2\nu - m]^{-1}$ , com  $\nu > m/2$ . Os casos normal,  $t$  e slash são obtidos se, além de tais suposições, for assumido que  $\alpha = 0$ .

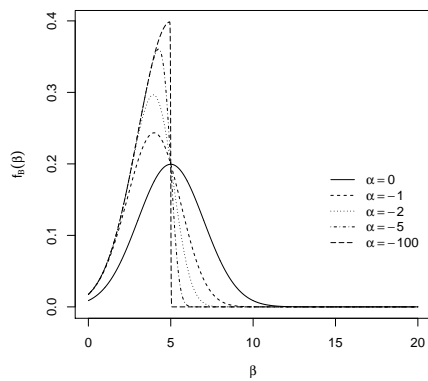
A Figura 4.1 mostra os gráficos para as distribuições de  $\beta$  e  $T$  no caso em que  $\beta$  tem distribuição skew-normal com locação  $\mu_\beta = 5$  e escala  $\omega_\beta^2 = 2$ .



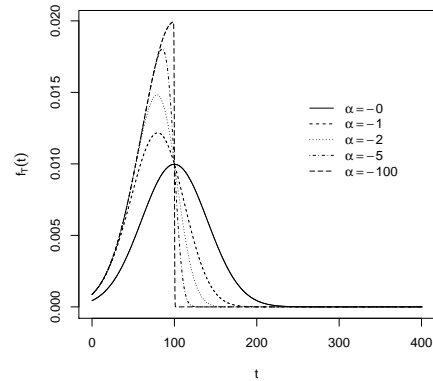
(a) FDP de  $\beta$  - Valores positivos de  $\alpha_\beta$



(b) FDP de  $T$  - Valores positivos de  $\alpha_\beta$



(c) FDP de  $\beta$  - Valores negativos de  $\alpha_\beta$



(d) FDP de  $T$  - Valores negativos de  $\alpha_\beta$

Figura 4.1: Funções densidade de probabilidade da distribuição skew normal com  $\mu_\beta = 5$  e  $\omega_\beta^2 = 4$  para efeito aleatório e tempo de falha, com limiar de falha para a degradação  $D_f = 20$ , assumindo-se diferentes valores de  $\alpha_\beta$  - modelo  $y_{ij} = \beta_i/t_{ij}$



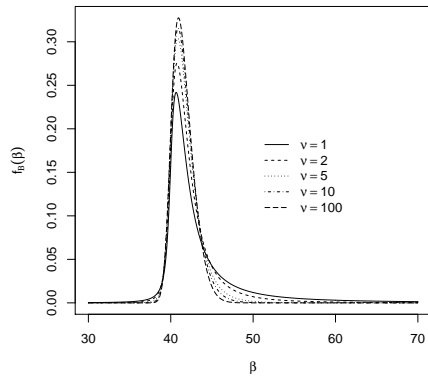
Observe pela figura 4.1, que a distribuição skew normal fornece flexibilidade maior que a distribuição normal para ajuste às distribuições possíveis para o tempo até a falha. Note que é possível evitar que seja colocada massa de probabilidade elevada para valores negativos tanto para  $\beta$  quanto  $t$ , basta que a assimetria apresente valores elevados ou que a distribuição esteja localizada à direita e distante de zero. Ou seja, apesar de ter suporte no conjunto dos números reais, a distribuição pode ser assimetrizada ou localizada de acordo com situações práticas reais, no sentido de que o tempo até a falha tem domínio positivo. No caso em que a distribuição assumida é skew-t as caudas da distribuição assimétrica se apresentam mais pesadas, como pode ser notado na Figura 4.2, o que permite lidar com essa característica de assimetria na presença de valores atípicos.

Vale notar também que o resultado da Proposição 4 é consequência direta do resultado obtido por (Branco e Dey, 2001), em que os autores provam que a família SMSN é fechada sob transformações lineares.

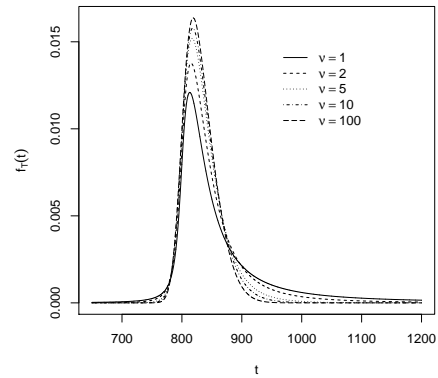
Esta propriedade de fechamento sob o modelo linear de degradação dado na equação (4.1) também é observada quando assumido que  $\beta_i \sim Weibull$  (Hamada, 2005).

Como a degradação é usualmente uma quantidade não negativa, é mais natural atribuímos uma distribuição com massa nula para valores negativos dos efeitos aleatórios. Assim, obtemos novos modelos lineares de degradação se assumirmos o modelo em (4.1) e considerarmos que  $\beta_i \sim SMLSN(\mu_\beta, h_i^{-1}\omega_\beta^2, \alpha_\beta, \mathcal{H})$ . A próxima proposição mostra a distribuição obtida para o tempo até a falha neste modelo.

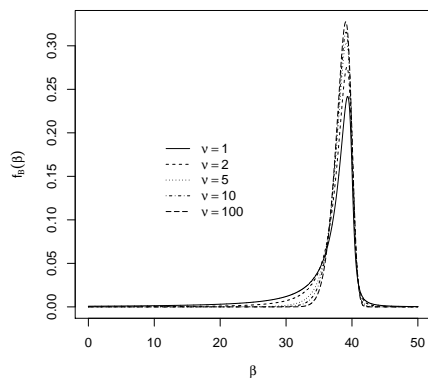
**Proposição 5.** *Seja o modelo de degradação na equação (4.1). Se  $\beta_i \sim SMLSN(\mu_\beta, h_i^{-1}\omega_\beta^2, \alpha_\beta, \mathcal{H})$ , então o tempo até a falha  $T_i \sim SMLSN(\mu_\beta + \ln(D_f), h_i^{-1}\omega_\beta^2, \alpha_\beta, \mathcal{H})$ .*



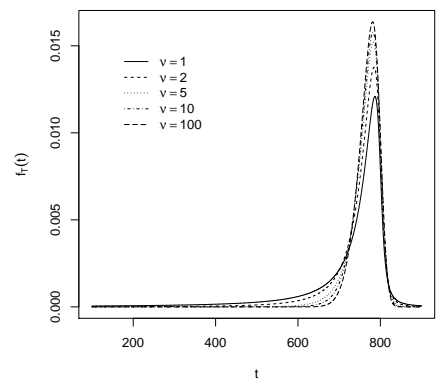
(a) FDP de  $\beta$  - Valor positivo de  $\alpha_\beta$



(b) FDP de  $T$  - Valor positivo de  $\alpha_\beta$



(c) FDP de  $\beta$  - Valor negativo de  $\alpha_\beta$



(d) FDP de  $T$  - Valor negativo de  $\alpha_\beta$

Figura 4.2: Funções densidade de probabilidade da distribuição skew-t com  $\mu_\beta = 40$ ,  $\omega_\beta^2 = 4$  e  $\alpha_\beta = \pm 3$  para efeito aleatório e tempo de falha, com limiar de falha para a degradação  $D_f = 20$ , assumindo-se diferentes valores de  $\nu_\beta$  - modelo  $y_{ij} = \beta_i/t_{ij}$

*Demonstração.* Das equações (2.9) e (4.3) segue que:

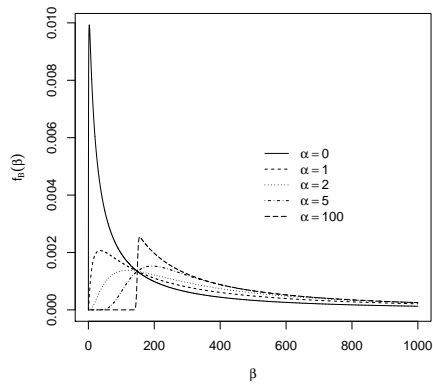
$$F(t_i|\mu_\beta, \omega_\beta, \alpha_\beta, \nu_\beta) = \int_0^{\frac{t_i}{D_f}} \int_0^\infty \frac{2}{\beta_i} \left( \frac{1}{2\pi h_i^{-1} \omega_\beta^2} \right)^{\frac{1}{2}} \exp \left\{ -\frac{1}{2h_i^{-1} \omega_\beta^2} (\ln(\beta_i) - \mu_\beta)^2 \right\} \\ \times \Phi \left( \alpha_\beta \left( \frac{\ln(\beta_i) - \mu_\beta}{h_i^{-\frac{1}{2}} \omega_\beta} \right) \right) d\mathcal{H}(h_i|\nu_\beta) d\beta_i.$$

Assim, fazendo  $t^* = \frac{t_i}{D_f}$ , a fdp para  $T_i$  é

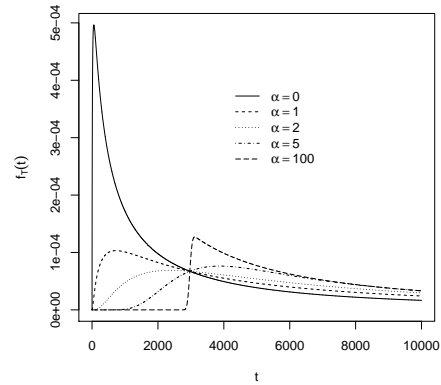
$$f(t_i|\mu_\beta, \omega_\beta, \alpha_\beta, \nu_\beta) = \int_0^\infty \frac{2}{\frac{t_i}{D_f}} \left( \frac{1}{2\pi h_i^{-1} \omega_\beta^2} \right)^{\frac{1}{2}} \exp \left\{ -\frac{1}{2h_i^{-1} \omega_\beta^2} \left( \ln \left( \frac{t_i}{D_f} \right) - \mu_\beta \right)^2 \right\} \times \\ \times \Phi \left( \alpha_\beta \left( \frac{\ln \left( \frac{t_i}{D_f} \right) - \mu_\beta}{h_i^{-\frac{1}{2}} \omega_\beta} \right) \right) d\mathcal{H}(h_i|\nu_\beta) \left[ \frac{1}{D_f} \right] \\ = \int_0^\infty \frac{2}{t_i} \left( \frac{1}{2\pi h_i^{-1} \omega_\beta^2} \right)^{\frac{1}{2}} \exp \left\{ -\frac{1}{2h_i^{-1} \omega_\beta^2} (\ln(t_i) - \mu_\beta - \ln(D_f))^2 \right\} \times \\ \times \Phi \left( \alpha_\beta \left( \frac{\ln(t_i) - \mu_\beta - \ln(D_f)}{h_i^{-\frac{1}{2}} \omega_\beta} \right) \right) d\mathcal{H}(h_i|\nu_\beta) \\ = \int_0^\infty \frac{2}{t_i} \left( \frac{1}{2\pi h_i^{-1} \omega_\beta^2} \right) \exp \left\{ -\frac{1}{2h_i^{-1} \omega_\beta^2} (\ln(t_i) - (\ln(D_f) + \mu_\beta))^2 \right\} \times \\ \times \Phi \left( \alpha_\beta \left( \frac{\ln(t_i) - (\ln(D_f) + \mu_\beta)}{h_i^{-\frac{1}{2}} \omega_\beta} \right) \right) d\mathcal{H}(h_i|\nu_\beta), \quad (4.5)$$

o que conclui a prova.  $\square$

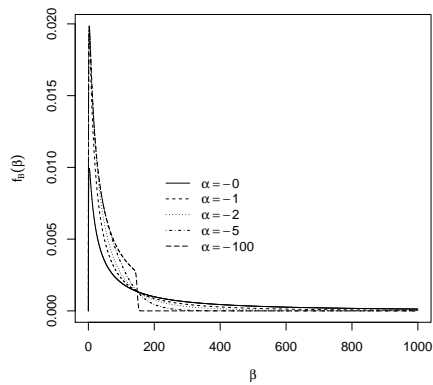
As boas propriedades apontadas da distribuição SMSN comentadas anteriormente também valem quando da utilização da distribuição SMLS<sub>N</sub> para modelar os efeitos aleatórios no modelo de degradação. Uma outra vantagem da família SMLS<sub>N</sub> reside no fato de que o suporte desta distribuição é sempre positivo e a riqueza de formas da distribuição é ainda mais abrangente, como pode ser evidenciado pelas distribuições log skew normal e log skew-t, respectivamente ilustradas nas Figuras 4.3 e 4.4.



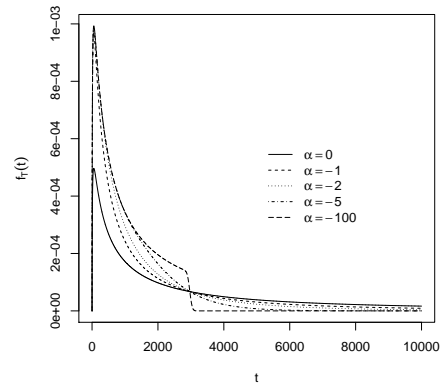
(a) FDP de  $\beta$  - Valores positivos de  $\alpha_\beta$



(b) FDP de  $T$  - Valores positivos de  $\alpha_\beta$



(c) FDP de  $\beta$  - Valores negativos de  $\alpha_\beta$



(d) FDP de  $T$  - Valores negativos de  $\alpha_\beta$

Figura 4.3: Funções densidade de probabilidade da distribuição log skew normal com  $\mu_\beta = 5$  e  $\omega_\beta^2 = 4$  para efeito aleatório e tempo de falha, com limiar de falha para a degradação  $D_f = 20$ , assumindo-se diferentes valores de  $\alpha_\beta$  - modelo  $y_{ij} = \beta_i/t_{ij}$

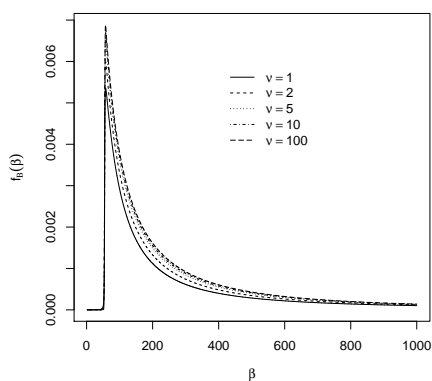
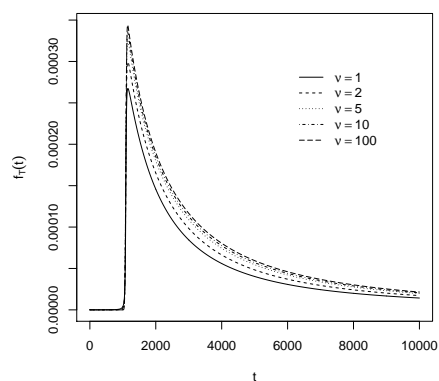
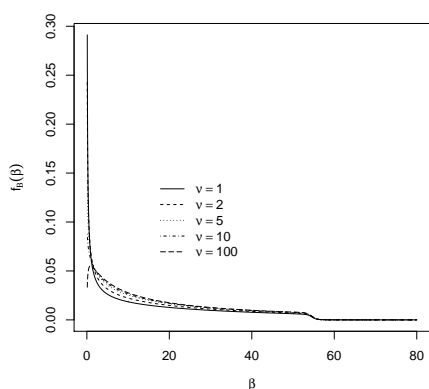
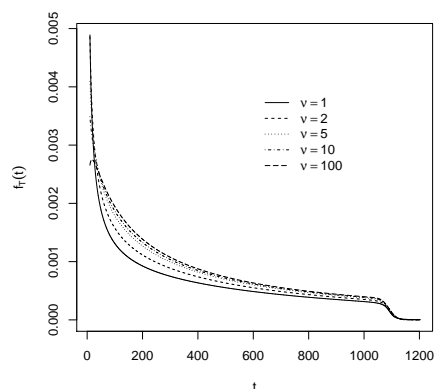
(a) FDP de  $\beta$  - Valor positivo de  $\alpha_\beta$ (b) FDP de  $T$  - Valor positivo de  $\alpha_\beta$ (c) FDP de  $\beta$  - Valor negativo de  $\alpha_\beta$ (d) FDP de  $T$  - Valor negativo de  $\alpha_\beta$ 

Figura 4.4: Funções densidade de probabilidade da distribuição log skew-t com  $\mu_\beta = 4$ ,  $\omega_\beta^2 = 4$  e  $\alpha_\beta = \pm 100$  para efeito aleatório e tempo de falha, com limiar de falha para a degradação  $D_f = 20$ , assumindo-se diferentes valores de  $\nu_\beta$  - modelo  $y_{ij} = \beta_i/t_{ij}$

É importante ressaltar que o resultado obtido na Proposição 5 fornece a transformação linear de variáveis aleatórias com distribuição SMLSN (ou LSMSN, como em Marchenko e Genton, 2010), a qual não se tem registro na literatura até então. Este resultado garante propriedade de fechamento sob o modelo linear proposto na equação (4.1) para a família SMLSN.

#### 4.1.2 Outro modelo robusto de degradação linear

Uma outra possibilidade para descrevermos o perfil linear real de degradação de uma unidade particular é considerarmos o modelo:

$$D(t_{ij}; \beta_i) = t_{ij}\beta_i, \quad (4.6)$$

em que  $\beta$  representa a taxa de degradação, a qual varia de unidade para unidade.

Para o nível crítico  $D_f$  obtemos que o tempo de falha pode ser obtido da seguinte relação:

$$T_i = \frac{D_f}{\beta_i}, \quad (4.7)$$

sempre que  $\beta_i \neq 0$ .

Sob este modelo a propriedade de fechamento por transformações lineares não é possível se assumirmos que  $\beta_i$  tem distribuição weibull. Isto é também observado se assumirmos que  $\beta_i$  tem distribuição na família SMSN, como pode ser visto na seguinte proposição.

**Proposição 6.** *Seja o modelo de degradação em (4.6). Se  $\beta_i \sim SMSN(\mu_\beta, h^{-1}\omega_\beta^2, \alpha_\beta, \mathcal{H})$ , então o tempo até a falha  $T_i$  tem distribuição cuja fdp é.*

$$\begin{aligned} f(t_i|\mu_\beta, \omega_\beta, \alpha_\beta, \nu_\beta) &= \int_0^\infty \frac{2}{t_i^2} \left( \frac{1}{2\pi h_i^{-1}\omega_\beta^2 D_f^{-2}} \right)^{\frac{1}{2}} \exp \left\{ -\frac{(t_i^{-1} - \mu_\beta D_f^{-1})^2}{2h_i^{-1}\omega_\beta^2 D_f^{-2}} \right\} \times \\ &\times \Phi \left( \alpha_\beta \left( \frac{t_i^{-1} - \mu_\beta D_f^{-1}}{h_i^{-\frac{1}{2}}\omega_\beta D_f^{-1}} \right) \right) d\mathcal{H}(h_i|\nu_\beta), \end{aligned} \quad (4.8)$$

em que  $\mu \in \mathfrak{R}$ ,  $\omega \in \mathfrak{R}_+$  e  $\alpha \in \mathfrak{R}$  são respectivamente parâmetros de localização, escala e assimetria. A medida de probabilidade  $\mathcal{H}(h|\nu)$  é definida no espaço amostral de uma variável aleatória  $H$ , e  $\nu$  é um vetor de hiperparâmetros que indexam  $\mathcal{H}$ .

*Demonstração.* Para provar a Proposição 6 basta notar que a fda do tempo até a falha  $T$  é dada por:

$$F(t|\theta_\beta) = 1 - P\left(\beta \leq \frac{D_f}{t} \middle| \theta_\beta\right), \quad (4.9)$$

em que  $\theta_\beta$  é o vetor de parâmetros que indexa a distribuição  $\Lambda(\beta|\theta_\beta)$  dos efeitos aleatórios.

Das equações (2.7) e (4.9) segue que

$$\begin{aligned} F(t_i|\mu_\beta, \omega_\beta, \alpha_\beta, \nu_\beta) &= 1 - \int_{-\infty}^{\frac{D_f}{t_i}} \int_0^\infty 2 \left( \frac{1}{2\pi h_i^{-1} \omega_\beta^2} \right)^{\frac{1}{2}} \exp \left\{ -\frac{(\beta_i - \mu_\beta)^2}{2h_i^{-1} \omega_\beta^2} \right\} \\ &\quad \times \Phi \left( \alpha_\beta \left( \frac{\beta_i - \mu_\beta}{h_i^{-\frac{1}{2}} \omega_\beta} \right) \right) d\mathcal{H}(h_i|\nu_\beta) d\beta_i. \end{aligned}$$

Como consequência, fazendo  $t^* = \frac{D_f}{t_i}$ , a fdp para  $T_i$  é

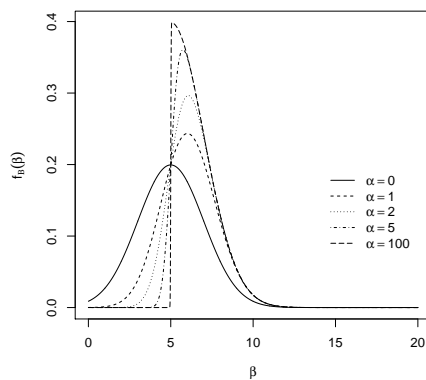
$$\begin{aligned} f(t_i|\mu_\beta, \omega_\beta, \alpha_\beta, \nu_\beta) &= - \int_0^\infty 2 \left( \frac{1}{2\pi h_i^{-1} \omega_\beta^2} \right)^{\frac{1}{2}} \exp \left\{ -\frac{\left(\frac{D_f}{t_i} - \mu_\beta\right)^2}{2h_i^{-1} \omega_\beta^2} \right\} \times \\ &\quad \times \Phi \left( \alpha_\beta \left( \frac{\frac{D_f}{t_i} - \mu_\beta}{h_i^{-\frac{1}{2}} \omega_\beta} \right) \right) d\mathcal{H}(h_i|\nu_\beta) \left[ \frac{-D_f}{t_i^2} \right] \\ &= \int_{-\infty}^\infty \frac{2}{t_i^2} \left( \frac{1}{2\pi h_i^{-1} \omega_\beta^2 D_f^{-2}} \right)^{\frac{1}{2}} \exp \left\{ -\frac{\left(t_i^{-1} - \mu_\beta D_f^{-1}\right)^2}{2h_i^{-1} \omega_\beta^2 D_f^{-2}} \right\} \times \\ &\quad \times \Phi \left( \alpha_\beta \left( \frac{t_i^{-1} - \mu_\beta D_f^{-1}}{h_i^{-\frac{1}{2}} \omega_\beta D_f^{-1}} \right) \right) d\mathcal{H}(h_i|\nu_\beta), \quad (4.10) \end{aligned}$$

em que  $t \neq 0$ , o que conclui a prova  $\square$

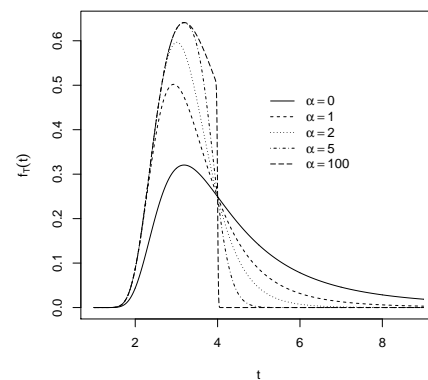
Observe que a distribuição obtida para o tempo de falha é uma transformação não linear de uma variável aleatória com distribuição SMSN, a qual ainda não foi explorada na literatura. Optou-se por nomeá-la como Mistura na Escala de uma distribuição Normal Assimetrizada Recíproca ( $SMRSN(\mu D_f^{-1}, h^{-1} \omega^2 D_f^{-2}, \alpha, \mathcal{H})$ ). Perceba que não

se observa a propriedade de fechamento para este caso.

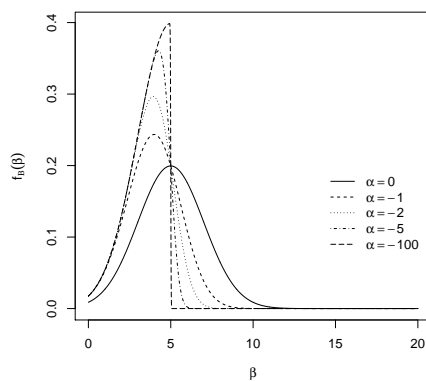
Pela Proposição 6 é possível notar, na Figura 4.5, a riqueza de formas resultante para os tempos de falha quando assumimos que  $\beta_i$  é distribuído segundo a família skew normal.



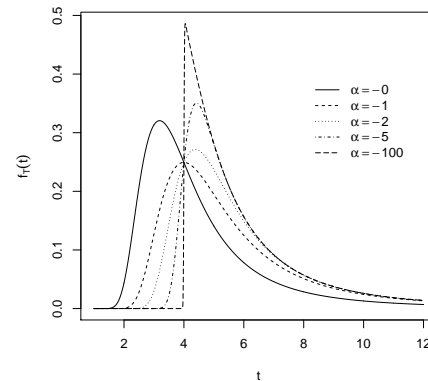
(a) FDP de  $\beta$  - Valores positivos de  $\alpha_\beta$



(b) FDP de  $T$  - Valores positivos de  $\alpha_\beta$



(c) FDP de  $\beta$  - Valores negativos de  $\alpha_\beta$



(d) FDP de  $T$  - Valores negativos de  $\alpha_\beta$

Figura 4.5: Funções densidade de probabilidade da distribuição log skew normal com  $\mu_\beta = 5$  e  $\omega_\beta^2 = 4$  para efeito aleatório e tempo de falha, com limiar de falha para a degradação  $D_f = 20$ , assumindo-se diferentes valores de  $\alpha_\beta$  - modelo  $y_{ij} = \beta_i t_{ij}$

Quando assumimos que  $\beta_i$  é distribuído segundo a família skew-t, a flexibilidade da assimetria e efeito no peso das caudas fica bem evidente pela Figura 4.6.



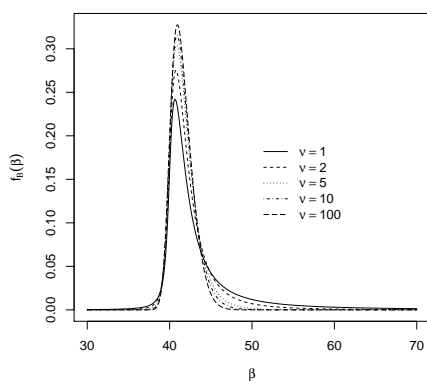
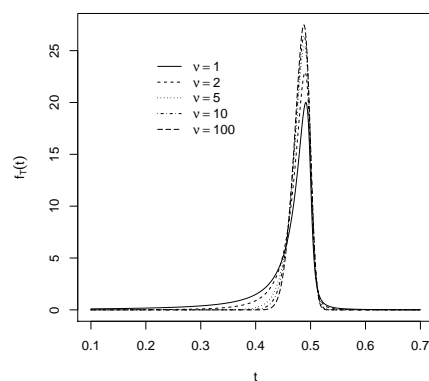
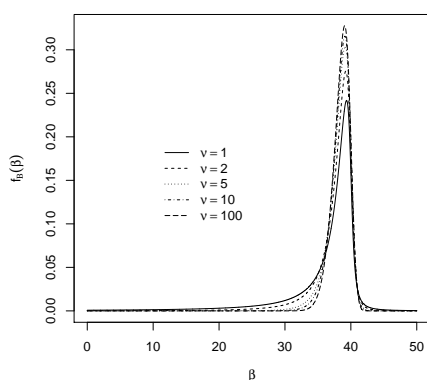
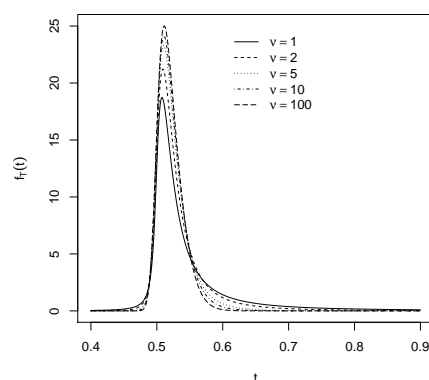
(a) FDP de  $\beta$  - Valor positivo de  $\alpha_\beta$ (b) FDP de  $T$  - Valor positivo de  $\alpha_\beta$ (c) FDP de  $\beta$  - Valor negativo de  $\alpha_\beta$ (d) FDP de  $T$  - Valor negativo de  $\alpha_\beta$ 

Figura 4.6: Funções densidade de probabilidade da distribuição log skew-t com  $\mu_\beta = 3$ ,  $\omega_\beta^2 = 0.04$  e  $\alpha_\beta = \pm 3$  para efeito aleatório e tempo de falha, com limiar de falha para a degradação  $D_f = 10$ , assumindo-se diferentes valores de  $\nu_\beta$  - modelo  $y_{ij} = \beta_i t_{ij}$

Considerando o modelo dado na equação (4.6), outros modelos robustos de degradação podem ser construídos se assumirmos que  $\beta_i \sim \text{SMLSN}(\mu_\beta, h_i^{-1}\omega_\beta^2, \alpha_\beta, \mathcal{H})$ . Neste caso, a distribuição do tempo até a falha é dada pela seguinte proposição

**Proposição 7.** *Considere o modelo de degradação em (4.6). Se  $\beta_i \sim \text{SMLSN}(\mu_\beta, h_i^{-1}\omega_\beta^2, \alpha_\beta, \mathcal{H})$ , então o tempo até a falha  $T_i \sim \text{SMLSN}(\mu_\beta - \ln(D_f), h_i^{-1}\omega_\beta^2, -\alpha_\beta, \mathcal{H})$ .*

*Demonstração.* Como por hipótese  $\beta_i \sim \text{SMLSN}(\mu_\beta, h_i^{-1}\omega_\beta^2, \alpha_\beta)$ , onde  $\mu_\beta \in \mathfrak{R}$ ,  $\omega_\beta^2 > 0$  e  $\alpha_\beta \in \mathfrak{R}$ . Segue que a fda de  $T_i$  é dada por:

$$\begin{aligned} F(t_i|\mu_\beta, \omega_\beta, \alpha_\beta, \nu_\beta) &= 1 - \int_0^{\frac{D_f}{t_i}} \int_0^\infty \frac{2}{\beta_i} \left( \frac{1}{2\pi h_i^{-1}\omega_\beta^2} \right)^{\frac{1}{2}} \exp \left\{ -\frac{(\ln(\beta_i) - \mu_\beta)^2}{2h_i^{-1}\omega_\beta^2} \right\} \\ &\quad \times \Phi \left( \alpha_\beta \left( \frac{\ln(\beta_i) - \mu_\beta}{h_i^{-\frac{1}{2}}\omega_\beta} \right) \right) d\mathcal{H}(h_i|\nu_\beta) d\beta_i. \end{aligned}$$

Conseqüentemente, fazendo  $t^* = \frac{D_f}{t_i}$ , temos que a fdp para  $T_i$  é

$$\begin{aligned} f(t_i|\mu_\beta, \omega_\beta, \alpha_\beta, \nu_\beta) &= - \int_0^\infty \frac{D_f}{t_i} \left( \frac{1}{2\pi h_i^{-1}\omega_\beta^2} \right)^{\frac{1}{2}} \exp \left\{ -\frac{(\ln(\frac{D_f}{t_i}) - \mu_\beta)^2}{2h_i^{-1}\omega_\beta^2} \right\} \\ &\quad \times \Phi \left( \alpha_\beta \left( \frac{\ln(\frac{D_f}{t_i}) - \mu_\beta}{h_i^{-\frac{1}{2}}\omega_\beta} \right) \right) d\mathcal{H}(h_i|\nu_\beta) \left[ -\frac{D_f}{t_i^2} \right] \\ &= \int_0^\infty \frac{2}{t_i} \left( \frac{1}{2\pi h_i^{-1}\omega_\beta^2} \right)^{\frac{1}{2}} \exp \left\{ -\frac{(-\ln(t_i) + \ln(D_f) - \mu_\beta)^2}{2h_i^{-1}\omega_\beta^2} \right\} \\ &\quad \times \Phi \left( \alpha_\beta \left( \frac{-\ln(t_i) + \ln(D_f) - \mu_\beta}{h_i^{-\frac{1}{2}}\omega_\beta} \right) \right) d\mathcal{H}(h_i|\nu_\beta) \\ &= \int_0^\infty \frac{2}{t_i} \left( \frac{1}{2\pi h_i^{-1}\omega_\beta^2} \right)^{\frac{1}{2}} \exp \left\{ -\frac{(-\ln(t_i) - [\mu_\beta - \ln(D_f)])^2}{2h_i^{-1}\omega_\beta^2} \right\} \\ &\quad \times \Phi \left( -\alpha_\beta \left( \frac{\ln(t_i) - [\mu_\beta - \ln(D_f)]}{h_i^{-\frac{1}{2}}\omega_\beta} \right) \right) d\mathcal{H}(h_i|\nu_\beta) \\ &= \int_0^\infty \frac{2}{t_i} \left( \frac{1}{2\pi h_i^{-1}\omega_\beta^2} \right)^{\frac{1}{2}} \exp \left\{ -\frac{(\ln(t_i) - [\mu_\beta - \ln(D_f)])^2}{2h_i^{-1}\omega_\beta^2} \right\} \times \end{aligned}$$

$$\times \Phi \left( -\alpha_{\beta} \left( \frac{\ln(t_i) - [\mu_{\beta} - \ln(D_f)]}{h_i^{-\frac{1}{2}} \omega_{\beta}} \right) \right) d\mathcal{H}(h_i | \nu_{\beta}), \quad (4.11)$$

o que conclui a prova.  $\square$

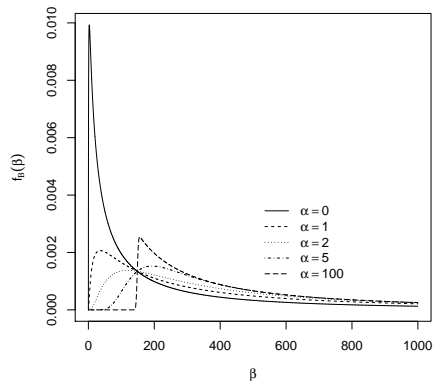
Assim, a propriedade de fechamento é obtida também para esta situação.

Observe pelas Figuras 4.7 e 4.8 que as distribuições log skew normal e log skew-t possuem formas que se mostram interessantes para modelar o comportamento do tempo até a falha. É importante notar que, escolhida uma assimetria para a distribuição os efeitos aleatórios, a assimetria é espelhada para a distribuição dos tempos de falha, isto é, se a assimetria for positiva (negativa) no efeito aleatório então será negativa (positiva) no tempo de falha.

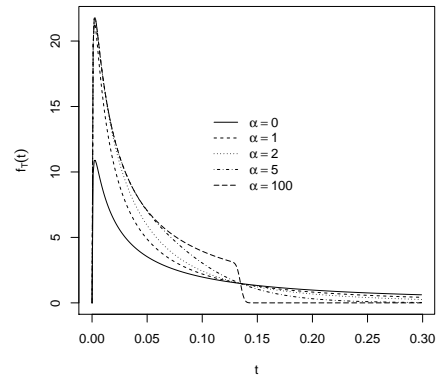
## 4.2 Inferência sobre os parâmetros do modelo linear de degradação

Neste texto a estimação dos objetos aleatórios do modelo de degradação é realizada sob o enfoque do paradigma bayesiano, o qual fornece uma forma natural para lidar com modelos de efeitos mistos. Sob a abordagem bayesiana, os efeitos aleatórios são tratados como parâmetros a serem estimados e, por causa dessa facilidade para lidar com o problema inferencial, torna-se possível considerarmos distribuições mais flexíveis e realistas para descrevermos a incerteza sobre o comportamento da taxa de degradação nas formas funcionais das equações (4.1) e (4.6).

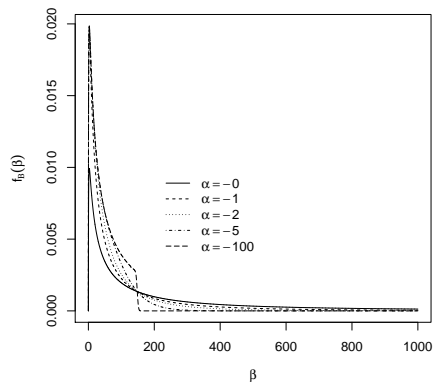
Seja  $\mathbf{Y} = (\mathbf{Y}_1, \dots, \mathbf{Y}_n)$  em que  $\mathbf{Y}_i = (Y_{i1}, \dots, Y_{im_i})$  é o vetor aleatório de medidas independentemente realizadas na unidade  $i$  sob teste, com  $i = 1, \dots, n$  caminhos de degradação. Seja  $\boldsymbol{\beta} = (\beta_1, \dots, \beta_n)^t$  o vetor de efeitos aleatórios das unidades experimentais.



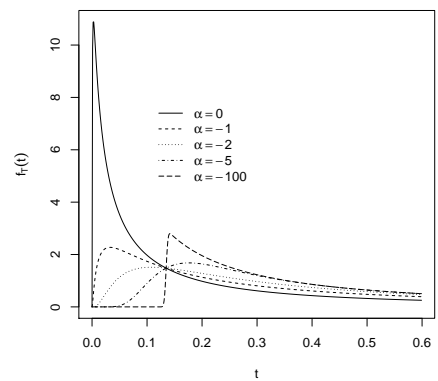
(a) FDP de  $\beta$  - Valores positivos de  $\alpha_\beta$



(b) FDP de  $T$  - Valores positivos de  $\alpha_\beta$



(c) FDP de  $\beta$  - Valores negativos de  $\alpha_\beta$



(d) FDP de  $T$  - Valores negativos de  $\alpha_\beta$

Figura 4.7: Funções densidade de probabilidade da distribuição log skew normal com  $\mu_\beta = 5$  e  $\omega_\beta^2 = 2$ , com limiar de falha para a degradação  $D_f = 20$ , assumindo-se diferentes valores de  $\alpha_\beta$  - modelo  $y_{ij} = \beta_i t_{ij}$

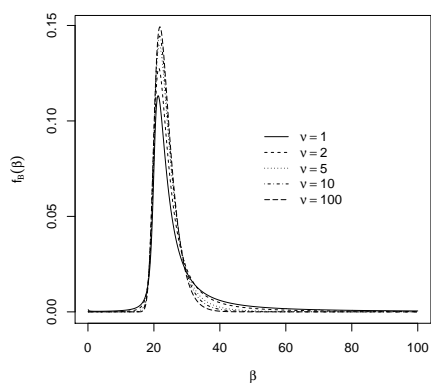
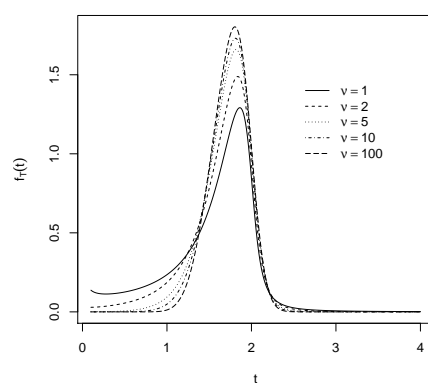
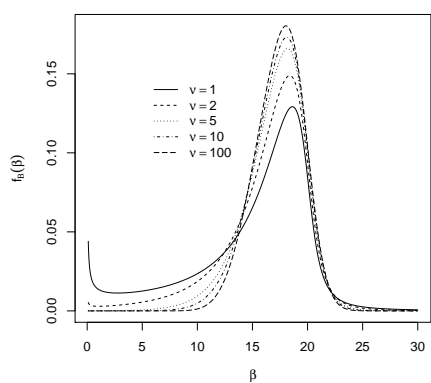
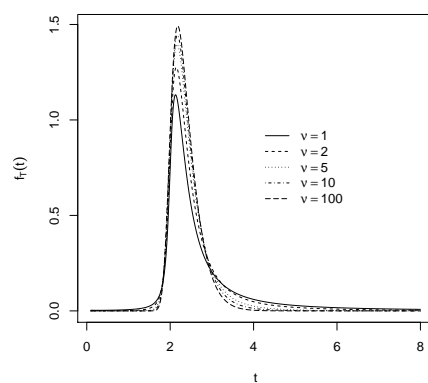
(a) FDP de  $\beta$  - Valor positivo de  $\alpha_\beta$ (b) FDP de  $T$  - Valor positivo de  $\alpha_\beta$ (c) FDP de  $\beta$  - Valor negativo de  $\alpha_\beta$ (d) FDP de  $T$  - Valor negativo de  $\alpha_\beta$ 

Figura 4.8: Funções densidade de probabilidade da distribuição log skew-t com  $\mu_\beta = 3$ ,  $\omega_\beta^2 = 0.04$  e  $\alpha_\beta = \pm 3$  para efeito aleatório e tempo de falha, com limiar de falha para a degradação  $D_f = 10$ , assumindo-se diferentes valores de  $\nu_\beta$  - modelo  $y_{ij} = \beta_i t_{ij}$

Assuma que a degradação real tem a forma linear da equação (4.1). Para o modelo geral de degradação na equação (1.3), dados  $\boldsymbol{\beta}$  e  $\sigma_\varepsilon^2$ , as medidas de degradação na unidade  $i$  são tais que

$$Y_{ij}|\beta_i, \sigma_\varepsilon^2 \stackrel{indep}{\sim} N\left(\frac{1}{\beta_i}t_{ij}; \sigma_\varepsilon^2\right). \quad (4.12)$$

Conseqüentemente, se assumimos, além disto, que as medidas de degradação  $\mathbf{Y}_i$  em unidades diferentes são independentes, dados  $\boldsymbol{\beta}$  e  $\sigma_\varepsilon^2$ , segue da equação 4.12 que

$$\mathbf{Y}_i|\beta_i, \sigma_\varepsilon^2 \stackrel{indep}{\sim} N_{m_i}(\boldsymbol{\mu}_i; \sigma_\varepsilon^2 I_{m_i}) = N_{m_i}\left(\frac{1}{\beta_i}\mathbf{t}_i^t; \sigma_\varepsilon^2 I_{m_i}\right), \quad (4.13)$$

em que  $I_{m_i}$  denota a matriz identidade de ordem  $m_i$ ,  $\mathbf{t}_i^t = (t_{i1}, \dots, t_{im_i})$  e  $\boldsymbol{\mu}_i = \left(\frac{1}{\beta_i}t_{i1}, \dots, \frac{1}{\beta_i}t_{im_i}\right)^t$ .

Sob estas suposições, se considerarmos  $n$  caminhos de degradação com distribuição dada na equação 4.13, temos que a função de verossimilhança de  $\mathbf{Y}, \boldsymbol{\beta}|\boldsymbol{\theta}, \sigma_\varepsilon^2$  torna-se a equação

$$f(\mathbf{y}, \boldsymbol{\beta}|\boldsymbol{\theta}, \sigma_\varepsilon^2) = \prod_{i=1}^n \left\{ \left[ \prod_{j=1}^{m_i} \left(\frac{1}{2\pi\sigma_\varepsilon^2}\right)^{1/2} \exp\left\{-\frac{1}{2\sigma_\varepsilon^2} \left(y_{ij} - \frac{t_{ij}}{\beta_i}\right)^2\right\} \right] f_{\beta_i|\boldsymbol{\theta}}(\beta_i|\boldsymbol{\theta}) \right\}. \quad (4.14)$$

A representação na equação (4.14) é conhecida como verossimilhança extendida. Se fôssemos utilizar uma abordagem clássica para obter inferências sobre os parâmetros deveríamos integrar nos efeitos aleatórios  $\boldsymbol{\beta}$ , o que poderia ser uma tarefa difícil para a estatística clássica.

Em inferência bayesiana não é necessário integrar a função de verossimilhança sob os efeitos aleatórios. Como  $\boldsymbol{\beta}$  é visto como um parâmetro, então a função de verossimilhança é escrita somente em termos da degradação, como na equação a seguir

$$f_{\mathbf{Y}|\boldsymbol{\beta}, \sigma_\varepsilon^2}(\mathbf{y}|\boldsymbol{\beta}, \sigma_\varepsilon^2) = \prod_{i=1}^n \prod_{j=1}^{m_i} \left(\frac{1}{2\pi\sigma_\varepsilon^2}\right)^{1/2} \exp\left\{-\frac{1}{2\sigma_\varepsilon^2} \left(y_{ij} - \frac{t_{ij}}{\beta_i}\right)^2\right\}. \quad (4.15)$$

Observe que no caso em que a degradação real tem forma linear da equação (4.6), as medidas de degradação em cada unidade são tais que

$$Y_{ij}|\beta_i, \sigma_\varepsilon^2 \stackrel{indep}{\sim} N(\beta_i t_{ij}; \sigma_\varepsilon^2). \quad (4.16)$$

Dessa maneira, a função de verossimilhança torna-se

$$f_{\mathbf{Y}|\boldsymbol{\beta},\sigma_\varepsilon^2}(\mathbf{y}|\boldsymbol{\beta},\sigma_\varepsilon^2) = \prod_{i=1}^n \prod_{j=1}^{m_i} \left( \frac{1}{2\pi\sigma_\varepsilon^2} \right)^{1/2} \exp \left\{ -\frac{1}{2\sigma_\varepsilon^2} (y_{ij} - \beta_i t_{ij})^2 \right\}. \quad (4.17)$$

A distribuição dos efeitos aleatórios é estabelecida como uma distribuição *a priori* no segundo estágio do modelo hierárquico (Bernardo e Smith, 1994).

No que segue, consideraremos ambos os modelos,  $y_{ij} = t_{ij}/\beta_i$  e  $y_{ij} = t_{ij}\beta_i$ , e assumiremos duas famílias de distribuições para descrever a incerteza sobre  $\beta_i$ . Primeiramente, assumamos que  $\beta_i \stackrel{iid}{\sim} \text{SMSN}(\mu_\beta, h^{-1}\omega_\beta^2, \alpha, \mathcal{H})$ .

Uma estratégia que facilita a inferência Bayesiana em modelos complexos é a técnica de aumento de dados. Essa técnica consiste na inclusão de variáveis latentes ou não observadas no modelo de maneira a simplificar procedimentos computacionais (veja Dyk e Meng, 2001, por exemplo). Se  $\beta_i \stackrel{iid}{\sim} \text{SMSN}(\mu_\beta, h^{-1}\omega_\beta^2, \alpha, \mathcal{H})$  um modelo equivalente pode ser construído considerando a sua representação estocástica (Azzalini e Capitanio, 2003)

$$\beta_i \stackrel{d}{=} \mu_\beta + \Delta_\beta U_i + H_i^{-1/2} \tau_\beta^{1/2} U_{1i},$$

em que  $\Delta_\beta = \omega_\beta \delta$ ,  $\delta = \frac{\alpha}{\sqrt{1+\alpha^2}}$ ,  $\tau_\beta = (1 - \delta^2) \omega_\beta^2$ ,  $U_i = H_i^{-1/2} |U_{0i}|$ ,  $U_{1i} \perp U_{0i} \stackrel{iid}{\sim} N(0,1)$  e  $H_i$  é uma variável aleatória com fda  $\mathcal{H}$  indexada pelo vetor  $\boldsymbol{\nu}$ . Considerando esta representação estocástica, a distribuição de  $\beta_i$  pode ser escrita hierarquicamente como a seguir

$$\left. \begin{array}{l} \beta_i | U_i = u_i, H_i = h_i \stackrel{iid}{\sim} N(\mu_\beta + \Delta_\beta u_i, h_i^{-1} \tau_\beta) \\ U_i | H_i = h_i \stackrel{iid}{\sim} N(0, h_i^{-1}) I(u_i)_{(0,\infty)} \\ H_i \stackrel{iid}{\sim} \mathcal{H}(h_i | \boldsymbol{\nu}) \end{array} \right\}. \quad (4.18)$$

A variável aleatória  $U_i$  é latente e, considerando a técnica de aumento de dados, também deve ser estimada. Consequentemente é necessário obtermos as distribuições *a posteriori* do vetor  $\boldsymbol{\psi} = (\boldsymbol{\beta}, \sigma_\varepsilon^2, \mathbf{U}, \mathbf{H}, \mu_\beta, \Delta_\beta, \tau_\beta)$ , em que  $\mathbf{U} = (U_1, \dots, U_n)$  e  $\mathbf{H} = (H_1, \dots, H_n)$ . Para completarmos as especificações do modelo assumamos que

$(\boldsymbol{\beta}, \sigma_\varepsilon^2, \mathbf{U}, \mathbf{H}, \mu_\beta, \Delta_\beta, \tau_\beta)$  são independentes *a priori*, e tais que

$$\left. \begin{aligned} \sigma_\varepsilon^2 &\sim \text{GI}(r_{\sigma_\varepsilon^2}, s_{\sigma_\varepsilon^2}) & \mu_\beta &\sim \text{N}(m, V) \\ \Delta_\beta &\sim \text{N}(\mu_{\Delta_\beta}, \sigma_{\Delta_\beta}^2) & \tau_\beta &\sim \text{GI}(r_{\tau_\beta}, s_{\tau_\beta}) \end{aligned} \right\}, \quad (4.19)$$

em que se  $X \sim \text{GI}(a, b)$ , com  $E(x) = b(a-1)^{-1}$  para  $a > 1$ , e  $V(x) = b(a-1)^{-2}(a-2)^{-1}$  para  $a > 2$ .

Desenvolvemos um algoritmo baseado nos métodos MCMC e na técnica de aumento de dados previamente descrita para obter amostras da distribuição *a posteriori* de  $\boldsymbol{\psi}$ . Para que isto seja possível, é necessário que obtenhamos as distribuições condicionais completas (dcc) *a posteriori* dos componentes de  $\boldsymbol{\psi}$ . Na Proposição 8, cuja prova encontra-se no Anexo A, apresentamos a dcc de todos estes parâmetros

**Proposição 8.** *Considere o modelo de degradação (1.3) e as especificações a priori dadas nas equações (4.18) e (4.19)*

(i) *Se o perfil de degradação degradação tem forma funcional:*

(a)  $D(t_{ij}; \beta_i) = \frac{t_{ij}}{\beta_i}$  como na equação (4.1), então as dcc a posteriori para  $\boldsymbol{\beta}$  e  $\sigma_\varepsilon^2$  são, respectivamente

$$\begin{aligned} f(\beta_i | \boldsymbol{\psi}^{[-\beta_i]}, \mathbf{y}) &\propto \exp \left\{ -\frac{1}{2\sigma_\varepsilon^2} \sum_{j=1}^{m_i} \left( y_{ij} - \frac{t_{ij}}{\beta_i} \right)^2 - \frac{1}{2h_i^{-1}\tau} [\beta_i - (\mu_\beta + \Delta_\beta u_i)]^2 \right\}, \\ \sigma_\varepsilon^2 | \boldsymbol{\psi}^{[-\sigma_\varepsilon^2]}, \mathbf{y} &\sim \text{GI} \left( \sum_{i=1}^n \frac{m_i}{2} + r_{\sigma_\varepsilon^2}, \frac{1}{2} \sum_{i=1}^n \sum_{j=1}^{m_i} \left( y_{ij} - \frac{t_{ij}}{\beta_i} \right)^2 + s_{\sigma_\varepsilon^2} \right), \end{aligned}$$

(b)  $D(t_{ij}; \beta_i) = t_{ij}\beta_i$  como na equação (4.6), então a dcc a posteriori para  $\boldsymbol{\beta}$  e  $\sigma_\varepsilon^2$  são, respectivamente

$$\begin{aligned} \beta_i | \boldsymbol{\psi}^{[-\beta_i]}, \mathbf{y} &\sim \text{N} \left( \frac{\sum_{j=1}^{m_i} y_{ij} t_{ij} + (\mu_\beta + \Delta_\beta u_i)}{h_i^{-1}\tau_\beta \sum_{j=1}^{m_i} t_{ij}^2 + \sigma_\varepsilon^2}, \frac{h_i^{-1}\tau_\beta \sigma_\varepsilon^2}{h_i^{-1}\tau_\beta \sum_{j=1}^{m_i} t_{ij}^2 + \sigma_\varepsilon^2} \right), \\ \sigma_\varepsilon^2 | \boldsymbol{\psi}^{[-\sigma_\varepsilon^2]}, \mathbf{y} &\sim \text{GI} \left( \sum_{i=1}^n \frac{m_i}{2} + r_{\sigma_\varepsilon^2}, \frac{1}{2} \sum_{i=1}^n \sum_{j=1}^{m_i} (y_{ij} - \beta_i t_{ij})^2 + s_{\sigma_\varepsilon^2} \right), \end{aligned}$$

(ii) *Para ambos os perfis de degradação mencionados em (i) as dcc a posteriori para*



$u_i, \mu_\beta$  e  $\Delta_\beta$  são dadas respectivamente por

$$\begin{aligned}
u_i | \boldsymbol{\psi}^{[-u_i]}, \mathbf{y} &\sim N \left( \frac{\Delta_\beta (\beta_i - \mu_\beta)}{\Delta_\beta^2 + \tau_\beta}, \frac{h_i^{-1} \tau_\beta}{\Delta_\beta^2 + \tau_\beta} \right) I(u_i)_{(0, \infty)}, \\
\mu_\beta | \boldsymbol{\psi}^{[-\mu_\beta]}, \mathbf{y} &\sim N \left( \frac{\sum_{i=1}^n \beta_i v_i^* + m \tau_\beta \prod_{i=1}^n h_i^{-1} - \Delta_\beta \sum_{i=1}^n u_i v_i^*}{\sum_{i=1}^n v_i^* + \tau_\beta \prod_{i=1}^n h_i^{-1}}, \frac{V \tau_\beta \prod_{i=1}^n h_i^{-1}}{\sum_{i=1}^n v_i^* + \tau_\beta \prod_{i=1}^n h_i^{-1}} \right), \\
\Delta_\beta | \boldsymbol{\psi}^{[-\Delta_\beta]}, \mathbf{y} &\sim N \left( \frac{\sigma_{\Delta_\beta}^2 \left[ \sum_{i=1}^n \beta_i u_i^* - \sum_{i=1}^n u_i^* \right] + \mu_{\Delta_\beta} \tau_\beta \prod_{i=1}^n h_i^{-1}}{\sigma_{\Delta_\beta}^2 \sum_{i=1}^n u_i^* + \tau_\beta \prod_{i=1}^n h_i^{-1}}, \frac{\sigma_{\Delta_\beta}^2 \tau_\beta \prod_{i=1}^n h_i^{-1}}{\sigma_{\Delta_\beta}^2 \sum_{i=1}^n u_i^* + \tau_\beta \prod_{i=1}^n h_i^{-1}} \right), \\
\tau_\beta | \boldsymbol{\psi}^{[-\tau_\beta]}, \mathbf{y} &\sim GI \left( \frac{n}{2} + r_{\tau_\beta}, \sum_{i=1}^n \frac{[\beta_i - (\mu_\beta + \Delta_\beta u_i)]^2}{2h_i^{-1}} + s_{\tau_\beta} \right), \\
\text{em que } k &= \{1, \dots, n\}, v_i^* = V \prod_{k \neq i} h_k^{-1} \text{ e } u_i^* = u_i \prod_{k \neq i} h_k^{-1}.
\end{aligned}$$

Alguns modelos particulares para descrever a degradação são definidos ao escolhermos a medida  $\mathcal{H}$ . Por exemplo, se  $\mathcal{H}$  é uma distribuição degenerada com toda massa de probabilidade sobre 1, então obtemos o modelo de degradação SN. Neste caso as dcc *a posteriori* são obtidas fazendo  $h_i = 1, \forall i = 1, \dots, n$  em todas as dcc apresentadas na Proposição 8. Se  $\mathcal{H}$  é a distribuição *Gama*( $\nu/2, \nu/2$ ), então o modelo de degradação ST é definido. Assumindo que  $\nu$  tem a distribuição *a priori* marginal independente de Jeffreys a seguir, tal como proposta por Fonseca *et al.* (2008)

$$f(\nu) \propto \left( \frac{\nu}{\nu+3} \right)^{\frac{1}{2}} \left\{ \frac{\partial^2 \log [\Gamma(\frac{\nu}{2})]}{\partial \nu^2} - \frac{\partial^2 \log [\Gamma(\frac{\nu+1}{2})]}{\partial \nu^2} - \frac{2(\nu+3)}{\nu(\nu+1)^2} \right\}^{\frac{1}{2}}. \quad (4.20)$$

Então as dcc *a posteriori* para  $h_i$  e  $\nu$ , em quaisquer dos modelos (4.1) e (4.6) são dadas por:

$$\begin{aligned}
h_i | \boldsymbol{\psi}^{[-h_i]}, \mathbf{y} &\sim \text{Gama} \left( \frac{\nu}{2} + 1, \frac{1}{2} \left\{ \tau_\beta^{-1} [\beta_i - (\mu_\beta + \Delta_\beta u_i)]^2 + u_i^2 + \nu \right\} \right), \\
f(\nu | \boldsymbol{\psi}^{[-\nu]}, \mathbf{y}) &\propto \left[ \frac{\frac{\nu}{2}}{\Gamma(\frac{\nu}{2})} \right]^n \left[ \prod_{i=1}^n h_i^{\frac{\nu}{2}-1} \right] \exp \left\{ - \sum_{i=1}^n h_i \frac{\nu}{2} \right\} \\
&\quad \times \left( \frac{\nu}{\nu+3} \right)^{\frac{1}{2}} \left\{ \frac{\partial^2 \log [\Gamma(\frac{\nu}{2})]}{\partial^2 \nu} - \frac{\partial^2 \log [\Gamma(\frac{\nu+1}{2})]}{\partial^2 \nu} - \frac{2(\nu+3)}{\nu(\nu+1)^2} \right\}^{\frac{1}{2}},
\end{aligned}$$

$$(4.21)$$

em que  $f\left(\nu|\boldsymbol{\psi}^{[-\nu]}\right)$  é própria pelo Corolário 2 do Teorema 2 em Fonseca *et al.* (2008).

A prova da Proposição 8 e das dcc em (4.21) encontram-se no Anexo A.

Como em geral os tempos de falha são estritamente positivos, também podemos considerar distribuições de probabilidade que colocam massa probabilística apenas sobre  $\mathfrak{R}^+$ . Assumimos que  $\beta_i \stackrel{iid}{\sim} \text{SMLSN}(\mu_\beta, h^{-1}\omega_\beta^2, \alpha, \mathcal{H})$ . Neste caso o procedimento inferencial pode ser elaborado de forma similar à realizada anteriormente considerando-se a transformação  $\log(\beta_i)$ , isto é, se  $\beta_i \stackrel{iid}{\sim} \text{SMLSN}(\mu_\beta, h^{-1}\omega_\beta^2, \alpha, \mathcal{H})$  então  $\log(\beta_i) \stackrel{iid}{\sim} \text{SMSN}(\mu_\beta, h^{-1}\omega_\beta^2, \alpha, \mathcal{H})$ . Desta forma as funções de verossimilhança e a representação estocástica, respectivamente nas equações (4.15), (4.17) e (4.18), podem ser utilizadas para reescrevermos as funções de verossimilhança e a representação estocástica em função de  $\log(\beta_i)$ . As funções de verossimilhança para os modelos nas (4.12) e (4.16) são agora dadas respectivamente por

$$f_{\mathbf{Y}|\log(\boldsymbol{\beta}), \sigma_\varepsilon^2}(\mathbf{y}|\log(\boldsymbol{\beta}), \sigma_\varepsilon^2) = \prod_{i=1}^n \prod_{j=1}^{m_i} \left( \frac{1}{2\pi\sigma_\varepsilon^2} \right)^{1/2} \exp \left\{ -\frac{1}{2\sigma_\varepsilon^2} \left( y_{ij} - \frac{1}{\log(\beta_i)} t_{ij} \right)^2 \right\}, \quad (4.22)$$

$$f_{\mathbf{Y}|\log(\boldsymbol{\beta}), \sigma_\varepsilon^2}(\mathbf{y}|\log(\boldsymbol{\beta}), \sigma_\varepsilon^2) = \prod_{i=1}^n \prod_{j=1}^{m_i} \left( \frac{1}{2\pi\sigma_\varepsilon^2} \right)^{1/2} \exp \left\{ -\frac{1}{2\sigma_\varepsilon^2} (y_{ij} - \log(\beta_i) t_{ij})^2 \right\}. \quad (4.23)$$

A representação estocástica para  $\log(\beta_i)$  é como em (4.18), isto é

$$\left. \begin{aligned} \log(\beta_i) | U_i = u_i, H_i = h_i &\stackrel{iid}{\sim} \text{N}(\mu_\beta + \Delta_\beta u_i, h_i^{-1} \tau_\beta) \\ U_i | H_i = h_i &\stackrel{iid}{\sim} \text{N}(0, h_i^{-1}) I(u_i)_{(0, \infty)} \\ H_i &\stackrel{iid}{\sim} \mathcal{H}(h_i | \boldsymbol{\nu}) \end{aligned} \right\} \quad (4.24)$$

As distribuições condicionais completas (dcc) *a posteriori* dos componentes de  $\boldsymbol{\psi}$  são dadas na Proposição 9 a seguir, cuja demonstração é análoga a da Proposição 8, e por isto será omitida.

**Proposição 9.** *Considere o modelo de degradação (1.3) e as especificações a priori dadas nas equações (4.24) e (4.19)*

(i) Se o perfil de degradação tem forma funcional:

(a)  $D(t_{ij}; \log(\beta_i)) = \frac{t_{ij}}{\log(\beta_i)}$  como na equação (4.1), então as dcc a posteriori para  $\beta$  e  $\sigma_\varepsilon^2$  são, respectivamente

$$f\left(\log(\beta_i) | \boldsymbol{\psi}^{[-\log(\beta_i)]}, \mathbf{y}\right) \propto \exp\left\{-\frac{\sum_{j=1}^{m_i} \left(y_{ij} - \frac{t_{ij}}{\log(\beta_i)}\right)^2}{2\sigma_\varepsilon^2} - \frac{[\log(\beta_i) - (\mu_\beta + \Delta_\beta u_i)]^2}{2h_i^{-1}\tau}\right\},$$

$$\sigma_\varepsilon^2 | \boldsymbol{\psi}^{[-\sigma_\varepsilon^2]}, \mathbf{y} \sim GI\left(\sum_{i=1}^n \frac{m_i}{2} + r_{\sigma_\varepsilon^2}, \frac{1}{2} \sum_{i=1}^n \sum_{j=1}^{m_i} \left(y_{ij} - \frac{t_{ij}}{\log(\beta_i)}\right)^2 + s_{\sigma_\varepsilon^2}\right),$$

(b)  $D(t_{ij}; \log(\beta_i)) = t_{ij} \log(\beta_i)$  como na equação (4.6), então a dcc a posteriori para  $\beta$  e  $\sigma_\varepsilon^2$  são, respectivamente

$$\log(\beta_i) | \boldsymbol{\psi}^{[-\log(\beta_i)]}, \mathbf{y} \sim N\left(\frac{\sum_{j=1}^{m_i} y_{ij} t_{ij} + (\mu_\beta + \Delta_\beta u_i)}{h_i^{-1}\tau_\beta \sum_{j=1}^{m_i} t_{ij}^2 + \sigma_\varepsilon^2}, \frac{h_i^{-1}\tau_\beta \sigma_\varepsilon^2}{h_i^{-1}\tau_\beta \sum_{j=1}^{m_i} t_{ij}^2 + \sigma_\varepsilon^2}\right),$$

$$\sigma_\varepsilon^2 | \boldsymbol{\psi}^{[-\sigma_\varepsilon^2]}, \mathbf{y} \sim GI\left(\sum_{i=1}^n \frac{m_i}{2} + r_{\sigma_\varepsilon^2}, \frac{1}{2} \sum_{i=1}^n \sum_{j=1}^{m_i} (y_{ij} - \log(\beta_i) t_{ij})^2 + s_{\sigma_\varepsilon^2}\right),$$

(ii) Para ambos os perfis de degradação mencionados em (i) as dcc a posteriori para  $u_i$ ,  $\mu_\beta$  e  $\Delta_\beta$  são dadas respectivamente por

$$u_i | \boldsymbol{\psi}^{[-u_i]}, \mathbf{y} \sim N\left(\frac{\Delta_\beta (\log(\beta_i) - \mu_\beta)}{\Delta_\beta^2 + \tau_\beta}, \frac{h_i^{-1}\tau_\beta}{\Delta_\beta^2 + \tau_\beta}\right) I(u_i)_{(0, \infty)},$$

$$\mu_\beta | \boldsymbol{\psi}^{[-\mu_\beta]}, \mathbf{y} \sim N\left(\frac{\sum_{i=1}^n \log(\beta_i) v_i^* + m\tau_\beta \prod_{i=1}^n h_i^{-1} - \Delta_\beta \sum_{i=1}^n u_i v_i^*}{\sum_{i=1}^n v_i^* + \tau_\beta \prod_{i=1}^n h_i^{-1}}, \frac{V\tau_\beta \prod_{i=1}^n h_i^{-1}}{\sum_{i=1}^n v_i^* + \tau_\beta \prod_{i=1}^n h_i^{-1}}\right),$$

$$\Delta_\beta | \boldsymbol{\psi}^{[-\Delta_\beta]}, \mathbf{y} \sim N\left(\frac{\sigma_{\Delta_\beta}^2 \left[\sum_{i=1}^n \log(\beta_i) u_i^* - \sum_{i=1}^n u_i^*\right] + \mu_{\Delta_\beta} \tau_\beta \prod_{i=1}^n h_i^{-1}}{\sigma_{\Delta_\beta}^2 \sum_{i=1}^n u_i^* + \tau_\beta \prod_{i=1}^n h_i^{-1}}, \frac{\sigma_{\Delta_\beta}^2 \tau_\beta \prod_{i=1}^n h_i^{-1}}{\sigma_{\Delta_\beta}^2 \sum_{i=1}^n u_i^* + \tau_\beta \prod_{i=1}^n h_i^{-1}}\right),$$

$$\tau_\beta | \boldsymbol{\psi}^{[-\tau_\beta]}, \mathbf{y} \sim GI\left(\frac{n}{2} + r_{\tau_\beta}, \sum_{i=1}^n \frac{[\log(\beta_i) - (\mu_\beta + \Delta_\beta u_i)]^2}{2h_i^{-1}} + s_{\tau_\beta}\right),$$

em que  $k = \{1, \dots, n\}$ ,  $v_i^* = V \prod_{k \neq i} h_k^{-1}$  e  $u_i^* = u_i \prod_{k \neq i} h_k^{-1}$ .

Analogamente à Proposição 8 alguns modelos particulares para descrever a degradação são definidos ao escolhermos a medida  $\mathcal{H}$ . Se  $\mathcal{H}$  é uma distribuição degenerada com toda massa de probabilidade sobre 1, então obtemos o modelo de degradação LSN. Neste caso as dcc *a posteriori* são obtidas fazendo  $h_i = 1, \forall i = 1, \dots, n$  em todas as dcc apresentadas na Proposição 9. Se  $\mathcal{H}$  é a distribuição  $Gama(\nu/2, \nu/2)$ , então o modelo de degradação LST é definido. Assumindo que  $\nu$  tem a distribuição *a priori* marginal independente de Jeffreys na equação (4.20), tal como proposta por Fonseca *et al.* (2008), então as dcc *a posteriori* para  $h_i$  e  $\nu$ , em quaisquer dos modelos (4.1) e (4.6) são dadas por

$$\begin{aligned} h_i | \boldsymbol{\psi}^{[-h_i]}, \mathbf{y} &\sim \text{Gama} \left( \frac{\nu}{2} + 1, \frac{1}{2} \left\{ \tau_\beta^{-1} [\log(\beta_i) - (\mu_\beta + \Delta_\beta u_i)]^2 + u_i^2 + \nu \right\} \right), \\ f(\nu | \boldsymbol{\psi}^{[-\nu]}, \mathbf{y}) &\propto \left[ \frac{\nu^{\frac{\nu}{2}}}{\Gamma(\frac{\nu}{2})} \right]^n \left[ \prod_{i=1}^n h_i^{\frac{\nu}{2}-1} \right] \exp \left\{ - \sum_{i=1}^n h_i \frac{\nu}{2} \right\} \\ &\quad \times \left( \frac{\nu}{\nu+3} \right)^{\frac{1}{2}} \left\{ \frac{\partial^2 \log [\Gamma(\frac{\nu}{2})]}{\partial^2 \nu} - \frac{\partial^2 \log [\Gamma(\frac{\nu+1}{2})]}{\partial^2 \nu} - \frac{2(\nu+3)}{\nu(\nu+1)^2} \right\}^{\frac{1}{2}}, \end{aligned} \quad (4.25)$$

em que  $f(\nu | \boldsymbol{\psi}^{[-\nu]})$  é própria pelo Corolário 2 do Teorema 2 em Fonseca *et al.* (2008).

As dcc *a posteriori* em (4.25) possuem demonstração similar à da equação (4.21), portanto será omitida.

Os resultados das Proposições 8 e 9 nos permitem fazer inferência sobre os objetos aleatórios em  $\boldsymbol{\psi}$ , uma vez que fornecem amostras da distribuição *a posteriori* de  $\boldsymbol{\psi}$  via métodos MCMC.

A representação estocástica das famílias SMSN e SMLSNS tornam a implementação do modelo uma tarefa menos árdua, pois pode ser realizada através de softwares disponibilizados livremente, como por exemplo o WinBUGS (Spiegelhalter *et al.*, 2000). A opção que se faz neste texto é por este último, através da integração com o software R (R Development Core Team, 2005). Os códigos principais utilizados também são disponibilizados no Anexo B. Uma dificuldade adicional no uso do WinBUGS surge

quando desejamos estimar  $\nu$  nos modelos T, Skew-t, Log-t e Log-skew-t, uma vez que o software não dispõe da função trigama  $\text{Trigamma}(\nu) = \frac{\partial^2}{\partial \nu^2} \ln \Gamma(\nu)$ , a qual faz parte da distribuição *a priori* Jeffreys independente proposta por Fonseca *et al.* (2008). A implementação da distribuição *a priori* para  $\nu$  será fornecida no Anexo B para subsidiar trabalhos futuros. Para esta implementação utilizamos a fórmula de recorrência da função trigamma e sua expansão em séries de Laurent, assim como em Abramowitz e Stegun (1965).

### 4.3 Critérios e medidas: seleção de modelos

Uma vez que a inferência sobre os objetos aleatórios dos modelos foi realizada, necessita-se avaliar a qualidade de ajuste dos mesmos através de métodos gráficos ou de medidas que permitem comparar diferentes modelos ajustados. Este capítulo explora seleção de modelos e algumas dessas maneiras de avaliar qualidade de ajuste.

#### 4.3.1 Ordenada preditiva condicional (CPO)

Uma maneira atraente de avaliação de qualidade de ajuste de um modelo é avaliando sua capacidade preditiva. Segundo Gelman *et al.* (2014), mesmo no contexto bayesiano é interessante resumirmos o ajuste do modelo aos dados sem que se considere diretamente a forma da distribuição *a priori* ou *a posteriori* dos objetos aleatórios (Gelman *et al.*, 2014).

Considerando este princípio, definiu-se a estatística CPO (do termo em inglês *conditional predictive ordinate*). Formalmente, a  $CPO_i$  associada a um componente  $x_i$  da amostra é definida como se segue

$$CPO_i = f(x_i | \mathbf{x}_{(-x_i)}) = \int_{\Xi_{\boldsymbol{\theta}}} f(x_i | \boldsymbol{\theta}, \mathbf{x}_{(-x_i)}) h(\boldsymbol{\theta} | \mathbf{x}_{(-x_i)}) d\boldsymbol{\theta}, \quad (4.26)$$

em que  $h(\boldsymbol{\theta} | \mathbf{x}_{(-x_i)})$  denota a distribuição *a posteriori* de  $\boldsymbol{\theta}$ ,  $\mathbf{x}_{(-x_i)}$  denota o vetor de observações de  $\mathbf{x}$  sem a componente  $i$  e  $\Xi_{\boldsymbol{\theta}}$  é o espaço paramétrico associado a  $\boldsymbol{\theta}$ .

Valores altos de  $CPO_i$  indicam bom ajuste do modelo.

Segundo Newton e Raftery (1994) é possível obtermos uma aproximação para a  $CPO_i$  considerando uma amostra da distribuição *a posteriori* de  $\boldsymbol{\theta}$  da seguinte forma

$$\widehat{CPO}_i = \left\{ \frac{1}{S} \sum_{s=1}^S \frac{1}{f(x_i | \boldsymbol{\theta}^s, \mathbf{x}_{(-i)})} \right\}^{-1}. \quad (4.27)$$

A expressão na equação (4.27) é simplificada quando  $x_1, \dots, x_n$  são condicionalmente independentes, pois daí temos  $f(x_i | \boldsymbol{\theta}, \mathbf{x}_{(-i)}) = f(x_i | \boldsymbol{\theta})$ .

Sendo a medida que resume a  $CPO_i$  considerada na comparação de modelos é dada por

$$LPML = \sum_{i=1}^n \log \{(CPO_i)\}, \quad (4.28)$$

em que um valor alto para a LPML indica melhor ajuste do modelo considerado.

### 4.3.2 Deviance information criterion (DIC)

Uma medida muito utilizada na comparação de modelos é o *Deviance Information Criterion* (DIC).

Spiegelhalter *et al.* (2002) apresentam a esperança do logaritmo da densidade preditiva ( $elpd_{DIC}$ ) conforme a equação (4.29):

$$\widehat{elpd}_{DIC} = \log f(\mathbf{x} | \hat{\boldsymbol{\theta}}_{Bayes}) - p_{DIC} \quad (4.29)$$

em que  $\hat{\boldsymbol{\theta}}_{Bayes} = E(\hat{\boldsymbol{\theta}}_{Bayes} | \mathbf{x})$  e a quantidade  $p_{DIC}$  regula a penalização devido à complexidade do modelo e é estimada por

$$\hat{p}_{DIC} = 2 \left[ \log f(\mathbf{x} | \hat{\boldsymbol{\theta}}_{Bayes}) - \frac{1}{S} \sum_{s=1}^S \log f(\mathbf{x} | \boldsymbol{\theta}^s) \right]. \quad (4.30)$$

Quanto maior  $\widehat{elpd}_{DIC}$  melhor é o ajuste do modelo. O critério de informação “Deviance” (DIC) é calculado como  $DIC = -2\widehat{elpd}_{DIC}$ , de modo que quanto menor o valor do DIC melhor é o ajuste do modelo.

### 4.3.3 Watanabe-Akaike information criterion (WAIC)

Watanabe (2010) propôs o critério de informação amplamente aplicável (*WAIC*, do inglês *Widely Applicable Information Criterion*), em que primeiramente toma-se a Esperança do logaritmo da densidade preditiva ponto a ponto ( $elppd_{WAIC}$ ) dada pela equação

$$\widehat{elppd}_{WAIC} = \sum_{i=1}^n \left[ \frac{1}{S} \sum_{s=1}^S f(x_i | \boldsymbol{\theta}^s) \right] - p_{WAIC}. \quad (4.31)$$

A penalização  $p_{WAIC}$  do critério de seleção é estimada pela equação (4.32):

$$\hat{p}_{WAIC} = \sum_{i=1}^n V_{s=1}^S \log [f(x_i | \boldsymbol{\theta}^s)], \quad (4.32)$$

em que  $V_{s=1}^S$  representa a variância com respeito a amostra do argumento.

Quanto maior  $\widehat{elppd}_{WAIC}$  na equação 4.31 melhor é o ajuste do modelo. Gelman *et al.* (2014) sugere uma mudança na escala de  $\widehat{elppd}_{WAIC}$  para que seja comparável com o critério DIC, então define a estatística  $WAIC = -2\widehat{elppd}_{WAIC}$ , segundo a qual quanto menor for seu valor, melhor.

### 4.3.4 Fator de bayes

O Fator de Bayes para comparar dois modelos  $\mathcal{M}_1$  e  $\mathcal{M}_2$  é dado pela razão das distribuições preditivas *a priori* sob os dois modelos, isto é

$$FB(\mathcal{M}_1, \mathcal{M}_2) = \frac{f(\mathbf{x} | \mathcal{M}_1)}{f(\mathbf{x} | \mathcal{M}_2)} = \frac{\int_{\Xi_{\boldsymbol{\theta}_1}} f(\mathbf{x} | \boldsymbol{\theta}_1, \mathcal{M}_1) h(\boldsymbol{\theta}_1 | \mathcal{M}_1) d\boldsymbol{\theta}_1}{\int_{\Xi_{\boldsymbol{\theta}_2}} f(\mathbf{x} | \boldsymbol{\theta}_2, \mathcal{M}_2) h(\boldsymbol{\theta}_2 | \mathcal{M}_2) d\boldsymbol{\theta}_2}. \quad (4.33)$$

Se  $FB(\mathcal{M}_1, \mathcal{M}_2) \geq 1$  temos evidência em favor do modelo  $\mathcal{M}_1$ , o caso contrário desta estatística indica evidência em favor do modelo  $\mathcal{M}_2$ .

Jeffreys (1961) introduz uma tabela de referência para decisões sobre  $\mathcal{M}_1$  usando o FB, baseada na interpretação da força da evidência observada a favor de  $\mathcal{M}_1$  trazida pelo Fator de Bayes. Posteriormente Kass e Raftery (1995) modifica estes valores de referência estabelecendo o critério exibido na Tabela 4.1.

Segundo Newton e Raftery (1994), a distribuição preditiva *a priori* sob os modelos

$FB(\mathcal{M}_1, \mathcal{M}_2)$	Evidência sobre $\mathcal{M}_1$
$\leq (1/150)$	muito forte contra
$(1/150) < FB \leq (1/20)$	forte contra
$(1/20) < FB \leq (1/3)$	substancial contra
$(1/3) < FB < 1$	fraca contra
$1 \leq FB < 3,2$	fraca a favor
$3,2 \leq FB < 10$	substancial a favor
$10 \leq FB < 100$	forte a favor
$\geq 100$	decisiva a favor

Tabela 4.1: Força da evidência sobre  $FB(\mathcal{M}_1, \mathcal{M}_2)$

$M_k, f(\mathbf{x}|\mathcal{M}_k)$ , pode ser aproximada por integração Monte Carlo como

$$f(\widehat{\mathbf{x}}|\mathcal{M}_k) = \left\{ \frac{1}{S} \sum_{s=1}^S \frac{1}{f(\mathbf{x}|\boldsymbol{\theta}^s)} \right\}^{-1}, \quad (4.34)$$

em que  $\boldsymbol{\theta}^s$  é a  $s$ -ésima amostra da distribuição *a posteriori* para  $\boldsymbol{\theta}$ .

Esta aproximação em (4.34) pode ser muito instável, entretanto para alguns modelos mais complexos é a única da qual se dispõe e é bem atrativa por ser facilmente calculada.

### 4.3.5 Teste de Kolmogorov-Smirnov (KS)

O teste de Kolmogorov-Smirnov é um teste clássico para ajuste de modelos e se baseia na diferença das fda empírica  $\hat{F}_e(\bullet)$  e a do modelo sob comparação, denotado por  $\hat{F}_M(\bullet)$ . Os seguintes passos são considerados em sua construção:

Passo (i) Ordene a amostra  $X_{(1)} \leq X_{(2)} \leq \dots \leq X_{(n)}$ ;

Passo (ii) Calcule a estatística de teste conforme a equação (4.35)

$$D_n = \max_{j=\{1, \dots, n\}} \left\{ \hat{F}_e(j) - \hat{F}_M(x_{(j)}), \hat{F}_M(x_{(j)}) - \hat{F}_e(j-1) \right\}; \quad (4.35)$$

em que  $\hat{F}_e(j) = \frac{j}{n}$  é a f.d.a empírica no ponto  $j$ , e  $\hat{F}_M(x_{(j)})$  é a fda estimada sob o modelo de interesse aplicada no ponto  $x_{(j)}$ ;

Passo (iii) Gere  $n$  valores  $u_1, \dots, u_n$  aleatoriamente de uma distribuição uniforme  $U(0,1)$  e ordene-os  $u_{(1)}^{(i)} \leq u_{(2)}^{(i)} \leq \dots \leq u_{(n)}^{(i)}$ ;



Passo (iv) Calcule a estatística de teste para estes valores simulados como a equação (4.36);

$$d^{(i)} = \max_{j=\{1,\dots,n\}} \left\{ \hat{F}_e(j) - u_{(j)}^{(i)}, u_{(j)}^{(i)} - \hat{F}_e(j-1) \right\}; \quad (4.36)$$

Passo (v) Faça  $I_i = \begin{cases} 1, & \text{se } d^{(i)} \geq D_n \\ 0, & \text{caso contrário.} \end{cases}$

Repita os passos (iii) e (iv) para  $i = 1, \dots, N$  obtendo  $I_1, \dots, I_N$ . O valor p do teste é estimado por  $\hat{p} = \sum_{i=1}^N \frac{I_i}{N}$ .

Se o interesse está em compararmos os modelos  $M_1$  e  $M_2$ , a estatística  $D_n$  pode ser reescrita como segue

$$D_n = \max_{j=\{1,\dots,n\}} \left| \hat{F}_{M_1}(y_{(j)}) - \hat{F}_{M_2}(y_{(j)}) \right|, \quad (4.37)$$

em que  $Y_{(1)} \leq Y_{(2)} \leq \dots \leq Y_{(n)}$  é um sequência de valores ordenados, advindos ou não de uma amostra, e  $\hat{F}_{M_2}$  é a fda estimada sob o modelo  $M_i$ ,  $i = 1, 2$ .

#### 4.3.6 Diagnóstico de influência via divergência de Kullback-Leibler (KL)

Uma outra estatística bastante útil na comparação de modelos é a divergência de Kullback-Leibler (KL), a qual nos permite identificar observações influentes no modelo e, conseqüentemente, avaliar a robustez na estimação dos parâmetros pelo modelo proposto. Em contexto bayesiano, para análise da presença de observações influentes McCulloch (1989) propõe considerar a divergência de KL entre as distribuições *a posteriori* baseadas nos dados completos e em  $\mathbf{x}_{(-x_i)}$ , a qual é dada por

$$KL_i = \int_{\Xi_{\boldsymbol{\theta}}} f(\boldsymbol{\theta}|\mathbf{x}) \frac{f(\boldsymbol{\theta}|\mathbf{x})}{f(\boldsymbol{\theta}|\mathbf{x}_{(-x_i)})} d\boldsymbol{\theta}. \quad (4.38)$$

A estatística  $KL_i$  mede o efeito, na distribuição *a posteriori* de  $\boldsymbol{\theta}$ , de removermos dos dados a  $i$ -ésima observação. Uma estimativa desta estatística, em função de  $CPO_i$ , que é recorrentemente utilizada na literatura (por exemplo, Cancho *et al.*, 2011) é dada

por

$$\widehat{KL}_i = -\log\left(\widehat{CPO}_i\right) + \frac{1}{S} \sum_{s=1}^S \log f(\mathbf{x}_i|\boldsymbol{\theta}_s), \quad (4.39)$$

em que  $\widehat{CPO}_i$  é dada na equação (4.26) e  $\boldsymbol{\theta}_s$  denota a  $s$ -ésima amostra da distribuição *a posteriori* de  $\boldsymbol{\theta}$ .

Uma observação é influente quando  $p_i = 0,5 \left[1 + \sqrt{1 - \exp\{-2KL_i\}}\right]$  é muito maior que 0,5 (McCulloch, 1989).

### 4.3.7 Seleção de modelos em modelos de degradação

A modelagem dos dados de degradação é feita considerando a degradação  $Y_{ij}$  da  $i$ -ésima unidade sob teste, no  $j$ -ésimo tempo de medição. Deste modo, para avaliarmos a qualidade de ajuste do modelo segundo os critérios nas Seções 4.3.1 a 4.3.4, é possível definirmos duas maneiras distintas para a deleção das observações necessárias no cálculo das densidades preditivas do modelo ajustado usadas em tais critérios. Podemos remover apenas uma degradação  $y_{ij}$  obtendo

$$f(x_i|\boldsymbol{\theta}) = f\left(y_{ij}|\boldsymbol{\theta}, \mathbf{y}^{[-y_{ij}]}\right) = N\left(D(t_{ij}; \beta_i), \sigma_\varepsilon^2\right), \quad (4.40)$$

ou ainda, reconhecendo o fato de que cada perfil de degradação corresponde a uma unidade amostral que incorreria num tempo de falha, podemos avaliar o efeito completo desse perfil na inferência removendo-o da análise. Neste caso, consideramos

$$f(x_i|\boldsymbol{\theta}) = f\left(\mathbf{y}_i|\boldsymbol{\theta}, \mathbf{y}^{[-\mathbf{y}_i]}\right) = \prod_{j=1}^{m_i} N\left(D(t_{ij}; \beta_i), \sigma_\varepsilon^2\right). \quad (4.41)$$

Como o interesse final da modelagem de dados de degradação reside na obtenção de inferências sobre os tempos de falha, pode ser de interesse trabalharmos com a avaliação de qualidade de ajuste e diagnóstico admitindo esta última alternativa. Desse modo, considera-se

$$f(x_i|\boldsymbol{\theta}) = f\left(t_i|\boldsymbol{\theta}, \mathbf{t}^{[-t_i]}\right), \quad (4.42)$$

em que  $f(t_i|\boldsymbol{\theta}, \mathbf{t}^{[-t_i]})$  é definida a partir da escolha da distribuição dos efeitos aleatórios como vimos na Seção 4. Observe que  $f(t_i|\boldsymbol{\theta}, \mathbf{t}^{[-t_i]})$  é a densidade *a priori* para o tempo de falha, dado que é conhecida somente se o método analítico é aplicável de modo a

fornecer distribuições fechadas para o tempo de falha. Na prática, como os tempos de falha não são observados, podemos considerar os pseudotempos de falha.

## 4.4 Análise de dados de degradação artificiais

Esta seção se destina à aplicação e avaliação dos modelos desenvolvidos neste texto em alguns conjuntos de dados artificialmente gerados. Os modelos são comparados entre si e também ao modelo weibull, que é amplamente utilizado para análises deste tipo de dados devido à sua flexibilidade.

Ao longo das análises o método proposto em Oliveira (2011) para inferir sobre os tempos até a falha, o qual é mais formalmente discutido neste trabalho, será referenciado como “proposta” pela abreviação “prop” e por “método via distribuição preditiva *a posteriori*”, indistintamente.

Ajustaremos apenas o modelo linear de degradação cuja forma funcional é  $D = t/\beta$

### 4.4.1 Geração dos perfis artificiais e especificações para inferência bayesiana

Serão gerados quatro bancos de dados para os quais se conhece previamente a distribuição dos tempos de falha. Aqui fornecemos o procedimento genérico para geração dos dados assumindo-se o modelo de degradação com perfil linear apresentado na equação (4.1). A escolha dessa forma funcional deve-se ao fato de termos observado, para a distribuição do tempo até a falha, a propriedade de fechamento nas famílias SMSN e SMLSN quando utilizado o método analítico. Além disso, Hamada (2005) também observou esse fechamento no caso weibull ( $W$ ) que é outro modelo considerado aqui. Outro ganho é que a comparação dos nove modelos ajustados aos bancos de dados será feita à luz da mesma forma funcional, unificada.<sup>1</sup>

---

<sup>1</sup>Seria perfeitamente possível utilizar a forma funcional da equação (4.6) para realizar as análises. Porém a comparação dos resultados para o caso SMSN e SMLSN não seria justa com o caso  $W$ . O ganho de se utilizar a forma da equação (4.6), nos casos SMSN e SMLSN, está ligado à velocidade do

Passos da geração dos dados de degradação:

1. Gere uma amostra de  $n$  tempos de falha a partir da distribuição de probabilidade de interesse. Denote por  $t_i$  os tempos de falha gerados,  $i = 1, \dots, n$ .
2. Defina um limiar de degradação  $D_f$ .
3. Utilize a equação  $\beta_i = T_i/D_f$  para encontrar a amostra de efeitos aleatórios  $\beta_i$  que gerarão cada perfil.
4. Defina os tempos  $t_j$  nos quais serão simuladas as medidas de degradação  $y_{ij}$ ,  $j = 1, \dots, m_i$ , em que  $m_i$  denota o número de medidas de degradação para um perfil específico.
5. Defina o desvio padrão  $\sigma_\varepsilon$  do erro associado ao modelo de degradação  $\varepsilon_{ij}$  na equação (1.3).
6. Gere as medidas de degradação  $y_{ij}$  utilizando a equação (1.3) do modelo de degradação. Os erros associados a cada  $y_{ij}$  são gerados de  $\varepsilon_{ij} \sim \text{Normal}(0, \sigma_\varepsilon^2)$ .

Diferentes distribuições foram assumidas para os efeitos aleatórios. O quadro 4.1 mostra os diferentes cenários considerados

Cenário	Distribuição de $T_i$	Distribuição de $\beta_i$
B-SN	$T_i \stackrel{iid}{\sim} \text{SN}(350 \times 10, 300^2 \times 10, 10)$	$\beta_i \stackrel{iid}{\sim} \text{SN}(350, 300^2, 10)$
B-ST	$T_i \stackrel{iid}{\sim} \text{ST}(355 \times 10; 100^2 \times 10; 10; 2, 1)$	$\beta_i \stackrel{iid}{\sim} \text{ST}(355; 100^2; 10; 2, 1)$
B-LST	$T_i \stackrel{iid}{\sim} \text{LST}(6 + \ln(10); 0, 2^2; 10; 4)$	$\beta_i \stackrel{iid}{\sim} \text{LST}(6; 0, 2^2; 10; 4)$
B-W	$T_i \stackrel{iid}{\sim} \text{Weibull}(6; 550 \times 10)$	$\beta_i \stackrel{iid}{\sim} \text{Weibull}(6; \frac{1}{550^6})$

Quadro 4.1: Cenários dos dados artificiais de degradação

em que a fdp de  $\beta_i \sim \text{Weibull}(\varsigma_\beta, \kappa_\beta)$  é dada pela equação

$$f(\beta_i | \varsigma_\beta, \kappa_\beta) = \varsigma_\beta \kappa_\beta \beta_i^{\varsigma_\beta - 1} \exp\{-\kappa_\beta \beta_i^{\varsigma_\beta}\},$$

processo de amostragem no algoritmo do MCMC, pois para esta forma funcional o mesmo conta com mais distribuições condicionais completas de forma fechada e conhecida.

com esperança  $E(\beta_i) = \kappa_\beta^{-1/\varsigma_\beta} \Gamma(1 + 1/\varsigma_\beta)$  e variância  $V(\beta_i) = \kappa_\beta^{-2/\varsigma_\beta} (\Gamma(1 + 2/\varsigma_\beta) - (\Gamma(1 + 1/\varsigma_\beta))^2)$ .

Todos os bancos de dados gerados são formados por  $n = 15$  perfis e em cada um deles  $m_i = 17$  medidas de degradação, todas feitas nos mesmos instantes de tempo, os quais foram igualmente espaçados. A degradação inicial foi estipulada como  $y_{i1} = 0$ , o limiar de falha como  $D_f = 10$  e o tempo de total de estudo  $L = 4000$ .

Para a análise dos cenários descritos no Quadro 4.1 assumimos que as distribuições *a priori* dos parâmetros pertencem às famílias mencionadas nas Proposições 8 e 9 e na equação (4.20).

Nas referidas distribuições *a priori* é necessário que se estabeleçam os hiperparâmetros dos níveis mais elevados da árvore do modelo hierárquico bayesiano, os quais afetam diretamente a incerteza sobre os parâmetros do modelo de degradação e, conseqüentemente, sobre a distribuição do tempo até a falha. Apesar de se tratar de um estudo em dados simulados, como usualmente ocorre em situações práticas, atribuiremos valores para os hiperparâmetros que reflitam total desconhecimento sobre os parâmetros do modelo de degradação, mais especificamente, utilizamos as seguintes distribuições *a priori* vagas para todos os parâmetros por eliciar

$$\left. \begin{array}{ll} \sigma_\varepsilon^2 \sim \text{GI}(10^{-3}, 10^3) & \mu_\beta \sim \text{N}(0, 10^6) \\ \Delta_\beta \sim \text{N}(0, 10^6) & \tau_\beta \sim \text{GI}(10^{-3}, 10^3) \\ \varsigma_\beta \sim \text{Gama}(10^{-3}, 10^{-3}) & \kappa_\beta \sim \text{Gama}(10^{-3}, 10^{-3}) \end{array} \right\}, \quad (4.43)$$

em que se  $X \sim \text{Gamma}(a, b)$ ,  $E(x) = ab^{-1}$  para  $a > 1$  e  $V(x) = ab^{-2}$ ; e se  $X \sim \text{GI}(a, b)$ ,  $E(x) = b(a-1)^{-1}$  para  $a > 1$  e  $V(x) = b(a-1)^{-2}(a-2)^{-1}$  para  $a > 2$ .

Cada um dos bancos de dados foi analisado usando o modelo SMSN. Para o modelo SMSN, vale lembrar que se  $h_i = 1$ , então  $\beta_i \stackrel{iid}{\sim} \text{SN}(\mu_\beta, \omega_\beta^2, \alpha_\beta)$ , e se adicionalmente  $\alpha_\beta = 0$  temos o modelo normal. O modelo skew-t com  $\beta_i \stackrel{iid}{\sim} \text{ST}(\mu_\beta, \omega_\beta^2, \alpha_\beta, \nu_\beta)$  é obtido se assumimos  $h_i \stackrel{iid}{\sim} \text{Gama}(\nu_\beta/2, \nu_\beta/2)$ , e retomamos o modelo t-student se  $\alpha_\beta = 0$ .

Também analisamos as bases de dados com o modelo SMLSN, onde tem-se que que

se  $h_i = 1$ , então  $\beta_i \stackrel{iid}{\sim} \text{LSN}(\mu_\beta, \omega_\beta^2, \alpha_\beta)$ , e se adicionalmente  $\alpha_\beta = 0$  temos o modelo log-normal. O modelo log skew-t com  $\beta_i \stackrel{iid}{\sim} \text{LST}(\mu_\beta, \omega_\beta^2, \alpha_\beta, \nu_\beta)$  é obtido se assumimos  $h_i \stackrel{iid}{\sim} \text{Gama}(\nu_\beta/2, \nu_\beta/2)$ , e retomamos o modelo log t-student se  $\alpha_\beta = 0$

A análise dos dados em todos os casos foi feita utilizando-se os softwares WinBUGS 1.4 (Spiegelhalter *et al.*, 2000)<sup>2</sup> e R 3.1.0 (R Development Core Team, 2005). A fim de garantirmos a convergência das cadeias, também minimizar a autocorrelação das amostras obtidas, foram realizadas 220.000.000 iterações, com um período de aquecimento (*burn-in*) de 100.000.000, e saltos de 60.000 para a amostragem. O resultado foi uma amostra *a posteriori* final de tamanho  $s = 2.000$  para cada parâmetro ( $\psi|\mathbf{y}$ ). Estas especificações garantiram a convergência das cadeias em todos os nove modelos ajustados para todos os bancos de dados analisados aqui, a saber, ajustes pelos modelos de degradação normal (N), log normal (LN), t-Student (T), log t-Student (LT), skew normal (SN), log skew-normal (LSN), skew-t (ST), log skew-t (LST) e weibull (W), cujos nomes estão relacionados com a distribuição adotada para os efeitos aleatórios. As inferências realizadas a partir das amostras *a posteriori* obtidas para os objetos aleatórios serão apresentados nas próximas seções.

É importante observar que o comprimento excessivo da cadeia se deu pelo grande número de variáveis latentes dos modelos propostos, os quais geram elevada autocorrelação das amostras obtidas para os parâmetros e diminuem a velocidade de convergência para a distribuição invariante. Por esta razão, para evitar autocorrelação das amostras, os saltos devem ser grandes, e para uma amostra representativa da distribuição *a posteriori* a cadeia deve ser longa. Pelo mesmo motivo o período de *burn-in* também foi elevado. Este padrão de cadeias longas se repete para todos os modelos deste texto, porém são um pouco menores para os modelos do Capítulo 5, pois contam com menos

---

<sup>2</sup>Para implementação dos modelos no WinBUGS 1.4 atente para o fato de que o software dispõe da distribuição  $N(\text{med}, \text{prec})$ , em que *med* é a média e *prec* é a precisão (variância recíproca). Esse detalhe impõe mudanças sobre distribuições *a priori* assumidas para a variância. Por exemplo, se  $\sigma_\varepsilon^2 \sim \text{GI}(r_{\sigma_\varepsilon^2}, s_{\sigma_\varepsilon^2})$ , no WinBUGS 1.4 fazemos  $1/\sigma_\varepsilon^2 \sim \text{Gamma}(r_{\sigma_\varepsilon^2}, s_{\sigma_\varepsilon^2})$

variáveis latentes.

#### 4.4.2 Resultados

O gráfico dos perfis de degradação simulados encontram-se na Figura 4.9. Note que os perfis são bem diversos e, como esperado, observamos perfis atípicos nos bancos de dados gerados a partir de distribuições de caudas pesadas.

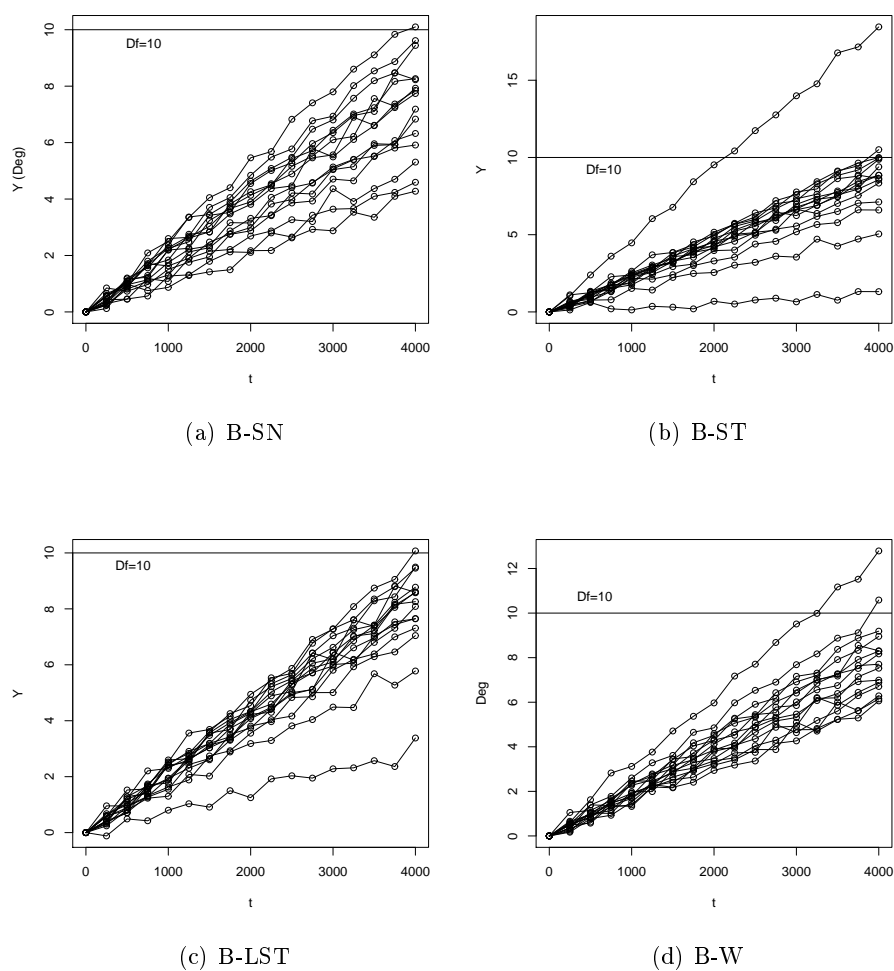


Figura 4.9: Perfis de degradação dos dados artificiais

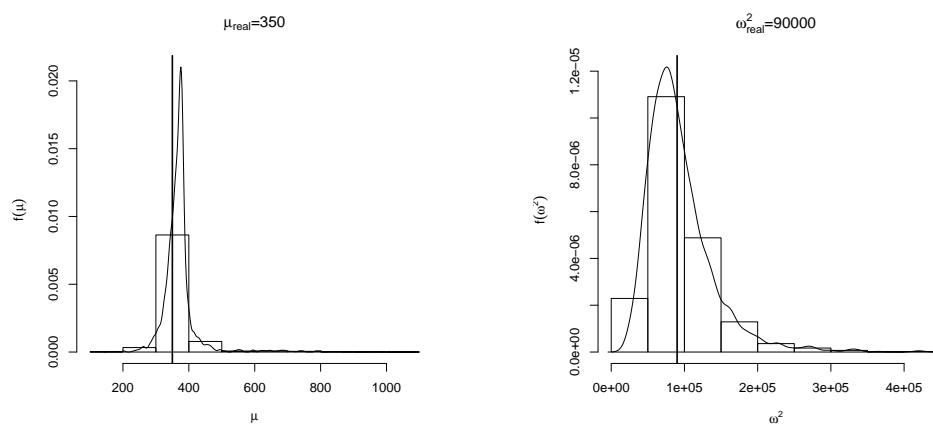
Dada a construção do banco de dados, espera-se que seja possível recuperar os

valores reais, e a respectiva distribuição dos tempos de falha apresentados no Quadro 4.1, quando os ajustes forem os mesmos da distribuição geradora dos dados. Quanto aos demais ajustes, a intenção é avaliar o quanto a flexibilidade de cada um consegue se aproximar da distribuição real. Em todas as distribuições assimétricas faremos a opção de estimar o valor real  $\delta_\beta = 0,9950$  em vez do correspondente valor real de  $\alpha_\beta = 10$ .

As Figuras 4.10 a 4.13 apresentam a distribuição *a posteriori* sobre os parâmetros quando o ajuste foi o mesmo correspondente à distribuição geradora dos dados. Nota-se que tais distribuições colocam maior massa probabilística na região dos valores reais apesar de termos utilizado distribuições *a priori* pouco informativas (exceto para  $\nu_\beta$  no ajuste ST, que está muito próximo da cauda superior da respectiva distribuição *a posteriori*). Então, em termos da inferência sobre os efeitos aleatórios os modelos parecem conseguir recuperar os parâmetros verdadeiros.

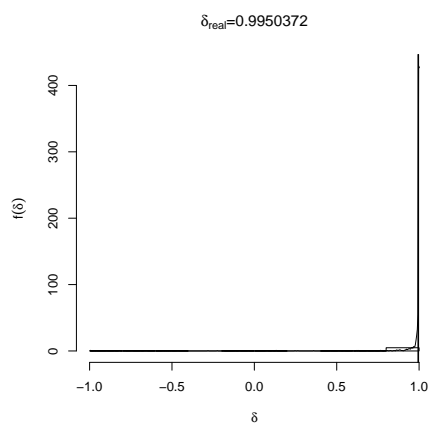
A Tabela 4.2 contém a média, a mediana, a moda e os intervalos de 95% de credibilidade de mais alta densidade (HPD, do inglês *highest probability density*) para os parâmetros. Observe que o único valor real de parâmetro que não está contido nos limites inferior ( $HPD_{inf}$ ) e superior ( $HPD_{sup}$ ) dos intervalos HPD do respectivo ajuste é  $\nu_\beta$  no ajuste ST, corroborando o que foi observado nas Figuras 4.10 a 4.13. Entretanto, pela literatura para inferência em distribuições assimétricas o valor de  $\nu_\beta$  é reconhecidamente de difícil estimação, tanto o é que geralmente é assumido como fixo, como por exemplo em Branco *et al.* (2013).





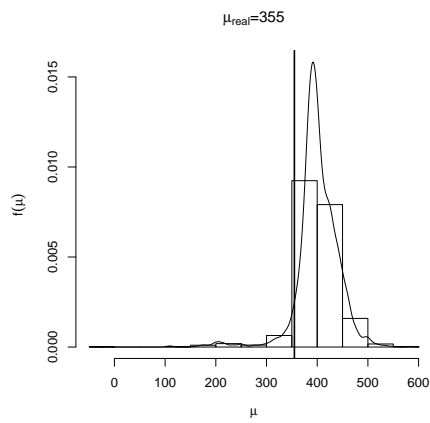
(a) Histograma para  $\mu_\beta$ : valor real  $\mu_\beta = 350$  na reta vertical espessa

(b) Histograma para  $\omega_\beta^2$ : valor real  $\omega_\beta^2 = 300^2$  na reta vertical espessa

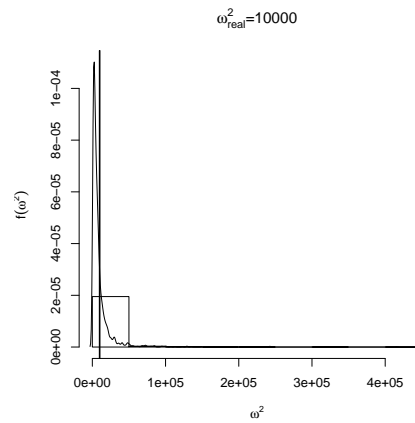


(c) Histograma para  $\delta_\beta$ : valor real  $\delta_\beta = 0,9950$  na reta vertical espessa

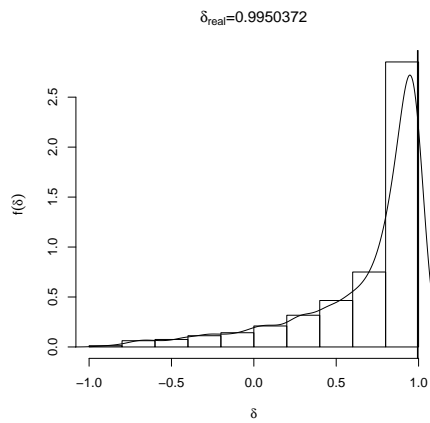
Figura 4.10: Comparação da distribuição *a posteriori* dos parâmetros com os valores reais no ajuste SN: dados artificiais com efeito aleatório Skew Normal



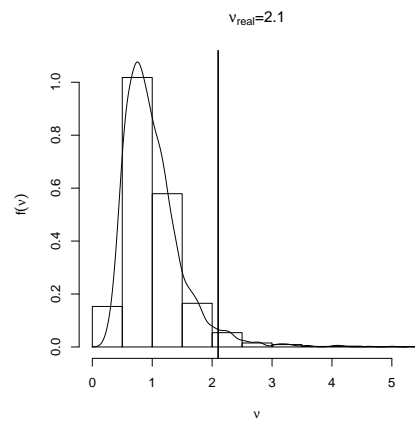
(a) Histograma para  $\mu_\beta$ : valor real  $\mu_\beta = 355$  na reta vertical espessa



(b) Histograma para  $\omega_\beta^2$ : valor real  $\omega_\beta^2 = 100^2$  na reta vertical espessa

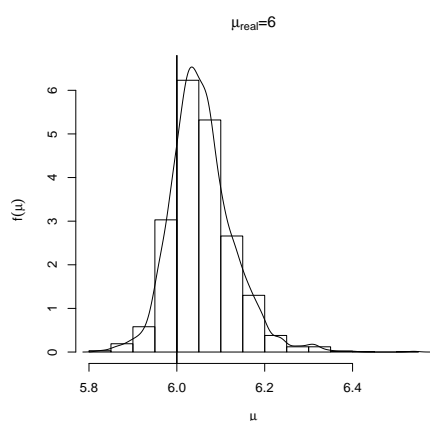


(c) Histograma para  $\delta_\beta$ : valor real  $\delta_\beta = 0,9950$  na reta vertical espessa

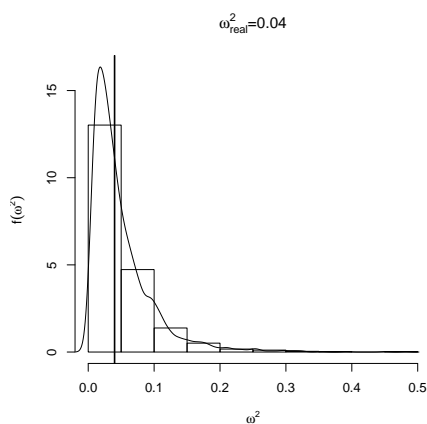


(d) Histograma para  $\nu_\beta$ : valor real  $\nu_\beta = 2,1$  na reta vertical espessa

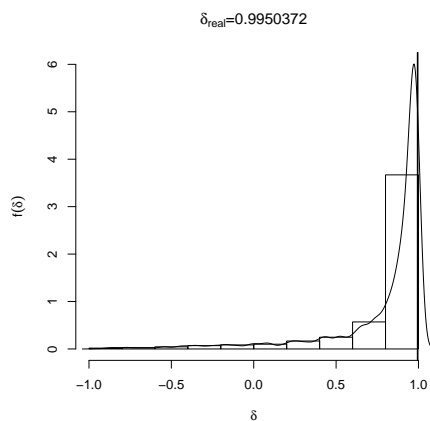
Figura 4.11: Comparação da distribuição *a posteriori* dos parâmetros com os valores reais no ajuste ST: dados artificiais com efeito aleatório Skew T



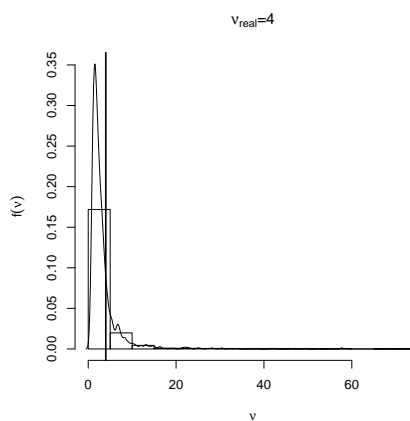
(a) Histograma para  $\mu_\beta$ : valor real  $\mu_\beta = 6$  na reta vertical espessa



(b) Histograma para  $\omega_\beta^2$ : valor real  $\omega_\beta^2 = 0,2^2$  na reta vertical espessa

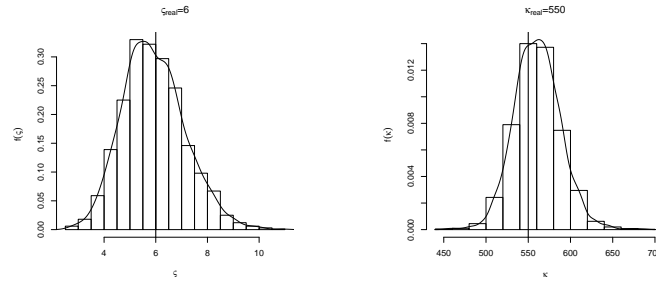


(c) Histograma para  $\delta_\beta$ : valor real  $\delta_\beta = 0,9950$  na reta vertical espessa



(d) Histograma para  $\nu_\beta$ : valor real  $\nu_\beta = 4$  na reta vertical espessa

Figura 4.12: Comparação da distribuição *a posteriori* dos parâmetros com os valores reais no ajuste LST: dados artificiais com efeito aleatório Log Skew T



(a) Histograma para  $\zeta_\beta$ : valor real  $\zeta_\beta = 6$  na reta vertical espessa

(b) Histograma para  $\kappa_\beta$ : valor real  $\kappa_\beta = 550$  na reta vertical espessa

Figura 4.13: Comparação da distribuição *a posteriori* dos parâmetros com os valores reais no ajuste W: dados artificiais com efeito aleatório Weibull

Tabela 4.2: Resumos *a posteriori* nos ajustes de mesma distribuição das degradações geradas

Gerado/Ajuste	Parâmetro	Real	Média	Mediana	Moda	$HPD_{inf}$	$HPD_{sup}$
SN	$\mu_\beta$	350,0000	371,8525	369,1000	371,1000	278,7000	456,6000
	$\omega_\beta^2$	90.000,0000	93.676,2800	84.020,0000	74.730,0000	27.470,0000	183.700,0000
	$\delta_\beta$	0,9950	0,9582	1,0000	1,0000	0,9199	1,0000
ST	$\mu_\beta$	355,0000	400,7520	398,9000	396,9500	328,0000	482,7000
	$\omega_\beta^2$	10.000,0000	9.967,2782	4.929,0000	2.926,1000	105,5000	30.700,0000
	$\delta_\beta$	0,9950	0,6807	0,8531	0,9267	-0,2305	1,0000
	$\nu_\beta$	<b>2,1000</b>	1,0117	0,9153	0,7584	<b>0,3081</b>	<b>1,9700</b>
LST	$\mu_\beta$	6,0000	6,0574	6,0500	6,0430	5,9240	6,2020
	$\omega_\beta^2$	0,0400	0,0500	0,0352	0,0217	0,0021	0,1393
	$\delta_\beta$	0,9950	0,7937	0,9324	0,9659	0,0085	0,9993
	$\nu_\beta$	4,0000	3,4498	2,2245	1,6933	0,4101	8,9770
W	$\zeta_\beta$	6,0000	5,9290	5,8250	5,6995	3,6300	8,2640
	$\kappa_\beta$	550,0000	560,7426	560,1167	560,4391	508,4262	613,5035

As Figuras 4.14 e 4.15 apresentam as fda e fdp *a posteriori* estimadas dos tempos de falha para uma unidade futura obtidas pelos nove modelos de degradação ajustados utilizando-se a abordagem proposta por Hamada (2005) (em conjunto com as Proposições 4 e 5) e por Oliveira (2011) (via nova unidade na verossimilhança, Seção 3.2.4). No caso das fda a linha tracejada cinza vertical ( $L$ ) indica o tempo de observação da degradação. A linha azul mais fina ( $sn.X$ ) corresponde à fda empírica baseada na amostra gerada sob a distribuição dos tempos de falha real. A linha preta cheia ( $sn.real$ ) indica a fda real. As demais linhas, a distribuição acumulada preditiva *a posteriori* dos tempos

de falha sob cada um dos nove ajustes. Para o caso das fdp, na parte inferior do gráfico adicionaram-se estrelas para indicar os tempos de falha gerados com intuito de auxiliar a leitura das densidades.

Para o banco de dados B-SN, nas Figuras 4.14 e 4.15 nota-se que as fda e fdp *a posteriori* que mais se aproximam da fda real são aquelas fornecidas pelos modelos SN, LSN e ST respectivamente. Os demais ajustes se distanciam bastante do real em algum ponto, com destaque especial para a cauda esquerda, em que os modelos T, W, N e LT fornecem estimativas ruins da cauda verdadeira. Isso é preocupante pois a cauda esquerda é de bastante interesse prático, o que justifica a necessidade da utilização de distribuições mais flexíveis para inferir sobre os tempos de falha. Note que as fda e fdp fornecidas pelos métodos propostos por Hamada (2005) e Oliveira (2011) são muito similares, o que era esperado já que se trata apenas de duas maneiras diferentes de obtermos a distribuição preditiva *a posteriori* dos tempos de falha para uma unidade da população.

No caso do banco de dados B-ST, pelas Figuras 4.14 e 4.15 podemos concluir que os modelos ST, LST, T e LT fornecem as melhores estimativas das fda e fdp *a posteriori*, em que ST e LST estimam melhor a cauda esquerda da distribuição real, enquanto T e LT estimam melhor a cauda direita. É interessante observar que estes modelos foram os menos afetados pelo valor atípico observado na amostra. Sobre o método proposto em Oliveira (2011), a amostra *a posteriori* via distribuição preditiva dos tempos de falha para uma unidade futura no ajuste T apresentou um valor extremamente atípico, o que dificultou a estimação da respectiva fdp (note pela Figuras 4.14 que a estimativa do modelo T não consta no gráfico). A Figura 4.16 apresenta o histograma do ajuste T que permite constatar que o problema da distorção se deve ao método de estimação da fdp disponível no software R Development Core Team (2005), e não ao ajuste no método de Oliveira (2011) (a estimação da densidade exige um espaço não disponível de armazenamento dos objetos do R na memória).

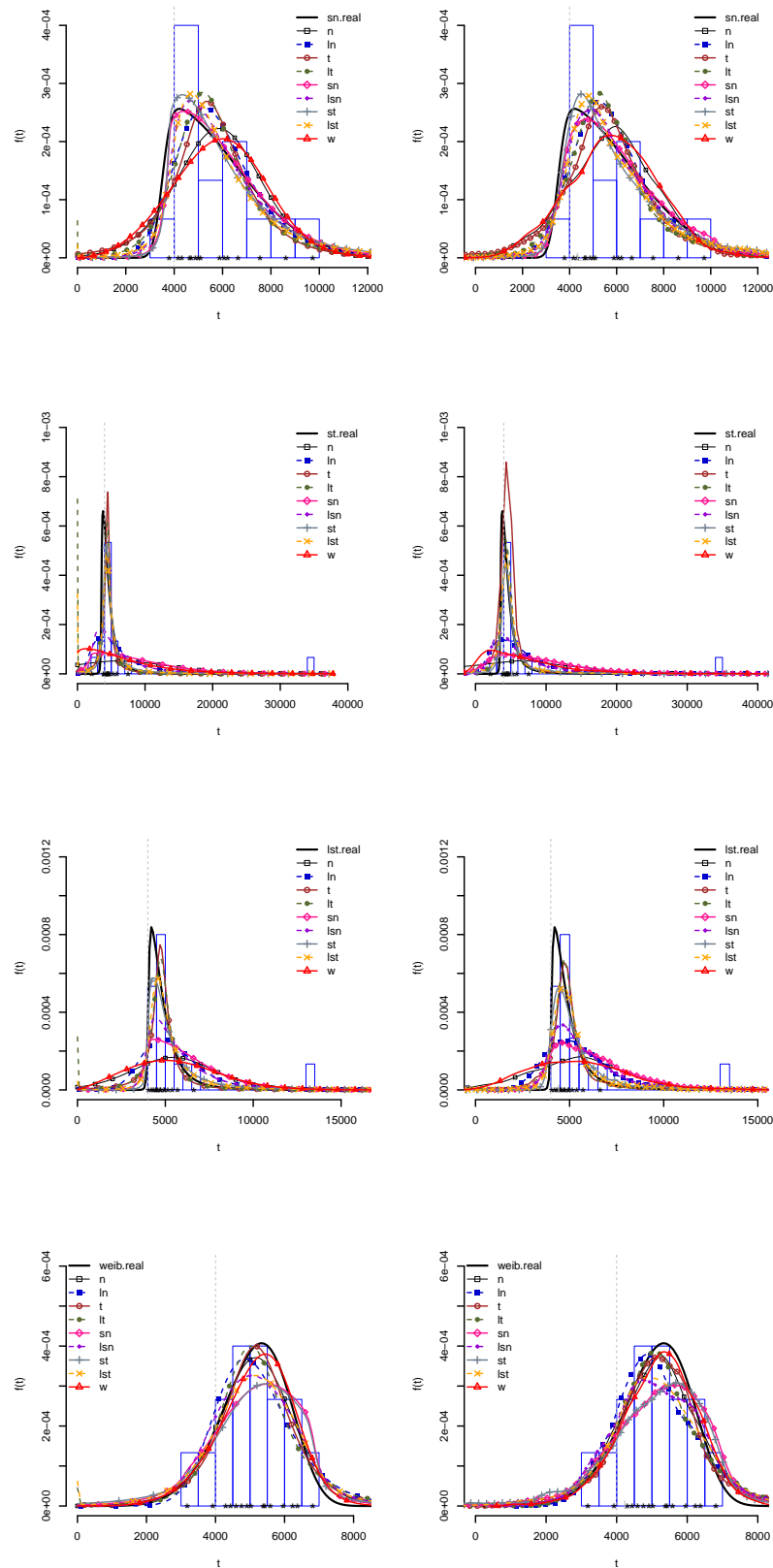


Figura 4.14: Função densidade de probabilidade *a posteriori* dos tempos de falha para uma unidade futura, pelos métodos de Hamada (2005) (esquerda) e proposta pela preditiva a posteriori (direita), para todos os modelos ajustados aos dados B-SN (primeira linha), B-ST (segunda linha), B-LST (terceira linha) e B-W (quarta linha)

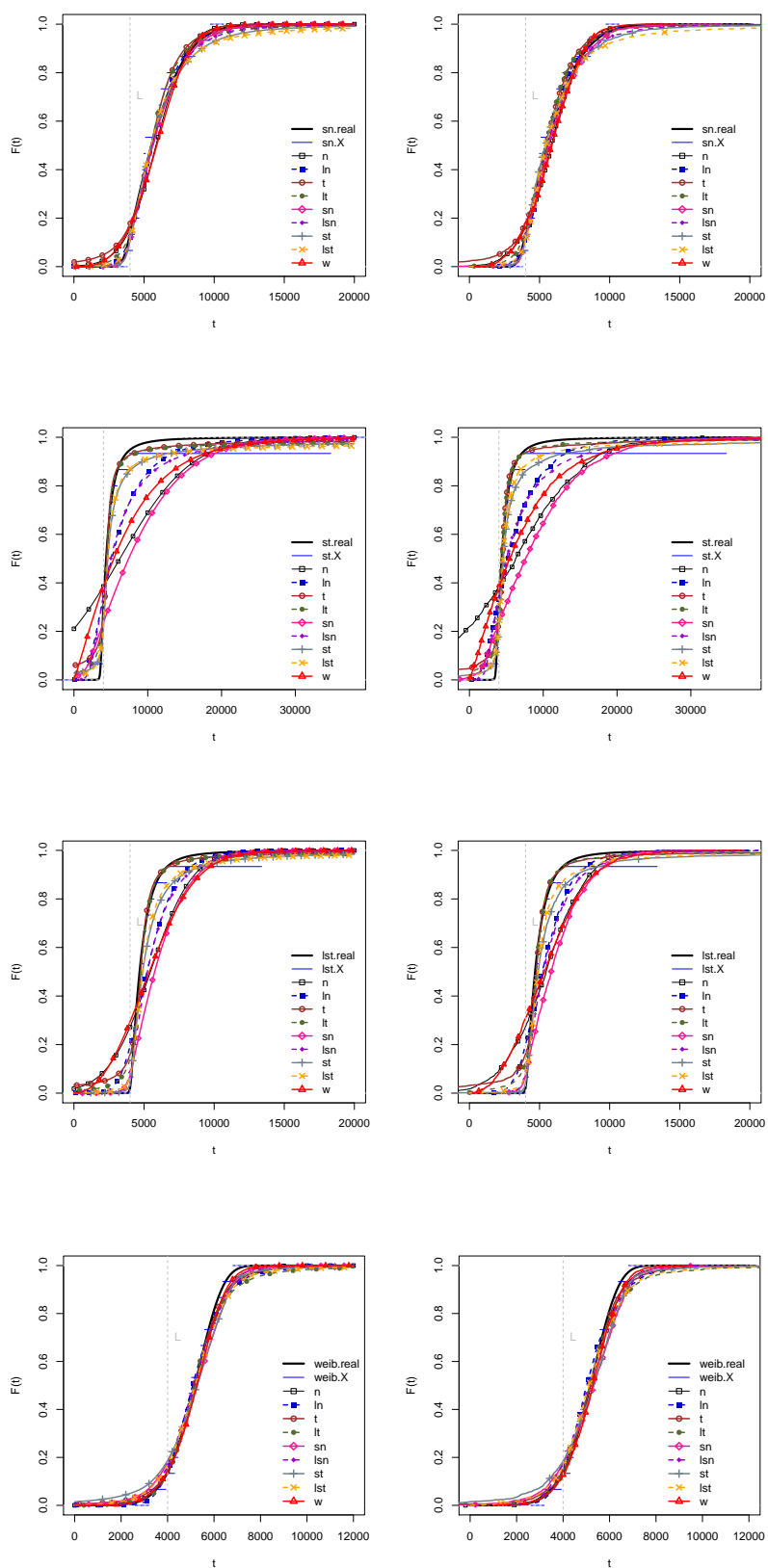


Figura 4.15: Função distribuição acumulada *a posteriori* dos tempos de falha para uma unidade futura, pelos métodos de Hamada (2005) (esquerda) e proposta pela preditiva *a posteriori* (direita), para todos os modelos ajustados aos dados B-SN (primeira linha), B-ST (segunda linha), B-LST (terceira linha) e B-W (quarta linha)

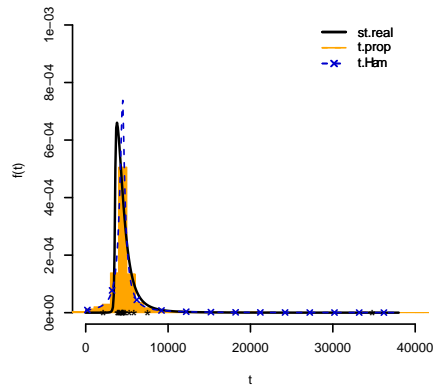


Figura 4.16: Função densidade de probabilidade dos tempos de falha para uma unidade futura pelo método proposto via preditiva a posteriori do modelo ajustado T, dados B-ST

Considerando o banco de dados B-LST, observamos pelas Figuras 4.14 e 4.15 que os melhores modelos para inferir sobre a fda e fdp de uma unidade futura foram o ST e LST. Note que essa capacidade de melhor se ajustar aos dados é seguida pelos modelos capazes de incorporar caudas pesadas ou assimetria. Os modelos W, LN e LT, estimam muito mal o comportamento da fda e fdp reais para este banco de dados.

Ao analisarmos o banco de dados B-W, observamos pelas Figuras 4.14 e 4.15 que os piores modelos para inferir sobre a fda e fdp de uma unidade futura foram o SN e ST. Os demais modelos foram muito similares, em que o melhor ajuste se deu pelo modelo W, como já era esperado. Vale ressaltar que os modelos LSN, T, LT, LST cumpriram um bom papel mesmo neste cenário cuja distribuição geradora (W) dos dados diferiu dos modelos ajustados.

Algumas quantidades de interesse podem ser calculadas. Para o método de Hamada (2005) é necessário que se utilize os resultados das proposições 4 e 5, e os procedimentos da Seção 3.2.1. Já o método via distribuição preditiva proposto por Oliveira (2011) pode ser computado segundo a Seção 3.2.3.

As Tabelas 4.3 a 4.6 apresentam os percentis de ordem 1%, 5%, 95% e 99% para



todos os nove modelos ajustados nos quatro bancos de dados artificiais. A coluna da média corresponde ao percentil *a posteriori* de uma unidade futura da segundo a proposta de Hamada (2005),  $HPD_{INF}$  e  $HPD_{SUP}$  aos intervalos de mais alta densidade com 95% credibilidade para os percentis sugeridos pelo mesmo. O método proposto por Oliveira (2011) fornece os referidos percentis obtidos diretamente da amostra da distribuição preditiva *a posteriori* de  $T_{n+1}$ . Em negrito estão destacadas as estimativas cujos modelos foram indicados pelas conclusões sobre as Figuras 4.14 e 4.15, e sublinhadas estão as estimativas mais distantes das reais.

Para o cenário B-SN, a Tabela 4.3 fornece estimativas pontuais bem parecidas na maioria dos ajustes, com destaque para os ajustes  $SN$ ,  $LSN$  e  $ST$ . Por exemplo, para o ajuste  $SN$  o tempo em que 1% dos itens futuros falharão é de 3156,00 via Oliveira (2011) e 3418,00 via Hamada (2005), ou seja, o método considerado em Oliveira (2011) fornece uma estimativa melhor para o percentil 1% uma vez que o percentil real é de 3269,00 (na distribuição empírica amostral de 3849,00). A interpretação do  $HPD$  proposto por Hamada (2005) é confusa: com 95% de credibilidade o percentil 1% *a priori* dos tempos de falha, avaliado nos parâmetros *a posteriori*, encontra-se no intervalo (2252,60 ; 4084,80). No modelo LT note que o intervalo HPD (7642,50 ; 102714,00) de 95% de credibilidade não contém a estimativa pontual  $\hat{t}_{99} = 1 \times 10^{68}$ , o que não é nenhuma surpresa, pois foi apontado na Seção 3.2.5 que este intervalo é referente à distribuição *a posteriori* do percentil *a priori* dos tempos de falha, e não *a posteriori*. Observe também que a estimativa do percentil é absurdamente distante da real, 11227.

Tabela 4.3: Percentis da distribuição *a posteriori* dos tempos de falha usando a Distribuição Preditiva *a posteriori* e o método de Hamada (2005) em todos os modelos ajustados para o cenário B-SN

Ajuste	$t_{p\%}$	Real	Empírica	Oliveira (2011)		Hamada (2005)		
				Pred.	Post.	Média	$HPD_{INF}$	$HPD_{SUP}$
N	$t_1$			1296,70		1646,60	-319,99	3389,80
	$t_5$			2663,90		2877,00	1328,90	4257,00
	$t_{95}$			8929,30		8816,20	7381,50	10334,00
	$t_{99}$			10522,00		10047,00	8308,70	12032,00
LN	$t_1$			2797,00		2935,20	2023,40	3681,50
	$t_5$			3554,90		3552,40	2656,20	4251,50
	$t_{95}$			9159,40		9130,30	7158,80	11402,00
	$t_{99}$			11400,00		11161,00	8176,60	14657,00
T	$t_1$			-3236,00		-226698,00	-29064,00	3770,80
	$t_5$			2057,70		1180,00	-1704,00	4149,60
	$t_{95}$			9077,10		9888,90	6642,60	12269,00
	$t_{99}$			12500,00		237767,00	7814,50	40158,00
LT	$t_1$			1675,20		2165,30	1,00E-64	3289,20
	$t_5$			3218,90		3222,80	2032,10	4317,70
	$t_{95}$			9513,30		10508,00	6789,10	14291,00
	$t_{99}$			13517,00		1,00E+68	7642,50	102714,00
SN	$t_1$	<b>3269,00</b>	<b>3849,00</b>	<b>3156,00</b>		<b>3418,00</b>	2252,60	4084,80
	$t_5$	<b>3619,00</b>	<b>4063,00</b>	<b>3744,00</b>		<b>3724,00</b>	2967,70	4284,90
	$t_{95}$	<b>9380,00</b>	<b>8956,00</b>	<b>9734,00</b>		<b>9511,00</b>	7476,10	11928,00
	$t_{99}$	<b>11227,00</b>	<b>9572,00</b>	<b>11961,00</b>		<b>11343,00</b>	8473,20	14420,00
LSN	$t_1$			2840,80		3202,50	2076,20	4077,80
	$t_5$			3585,80		3661,80	2729,70	4335,90
	$t_{95}$			10092,00		10132,00	7069,60	13788,00
	$t_{99}$			14220,00		13450,00	8239,90	20724,00
ST	$t_1$			3154,70		3396,40	2056,20	4079,50
	$t_5$			3731,00		3720,80	2957,00	4299,00
	$t_{95}$			10721,00		11702,00	7087,70	18277,00
	$t_{99}$			16801,00		22820,00	8375,40	45010,00
LST	$t_1$			2365,00		2962,00	1186,20	4166,90
	$t_5$			3544,90		3563,10	2210,30	4393,60
	$t_{95}$			11732,00		3,00E+12	6813,70	28821,00
	$t_{99}$			31142,00		6,00E+54	8297,80	414527,00
W	$t_1$			1678,40		1820,30	900,65	2866,70
	$t_5$			2645,40		2843,80	1770,90	3942,20
	$t_{95}$			8897,70		8985,60	7621,00	10635,00
	$t_{99}$			10440,00		10183,00	8406,10	12297,00

Tabela 4.4: Percentis da distribuição *a posteriori* dos tempos de falha usando a Distribuição Preditiva *a posteriori* e o método de Hamada (2005) em todos os modelos ajustados para o cenário B-ST

Ajuste	$t_{p\%}$	Real	Empírica	Oliveira (2011)		Hamada (2005)	
				Pred.	Post.	Média	$HPD_{INF}$
N	$t_1$			-13601,00	-11735,00	-21076,00	-4478,40
	$t_5$			-7563,60	-6477,20	-13151,00	-319,98
	$t_{95}$			19222,00	18903,00	13023,00	25692,00
	$t_{99}$			27332,00	24160,00	16692,00	32653,00
LN	$t_1$			977,82	1279,70	500,55	1996,20
	$t_5$			1766,80	1915,90	999,04	2824,10
	$t_{95}$			14252,00	14837,00	8591,10	23330,00
	$t_{99}$			23351,00	23276,00	11101,00	40365,00
T	$t_1$			-61642,00	-9,00E+06	-3,00E+06	2326,50
	$t_5$			359,88	-11201,00	-34336,00	3735,20
	$t_{95}$			8851,10	20103,00	5274,40	43351,00
	$t_{99}$			47775,00	9,00E+06	7031,50	3,00E+06
LT	$t_1$			0,00	455,31	0,00	1846,70
	$t_5$			1321,80	1934,90	3,00E-73	3191,30
	$t_{95}$			8235,00	3,00E+76	5304,60	253272,00
	$t_{99}$			70300,00	$\infty$	7340,10	3,00E+51
SN	$t_1$			-170,37	1090,90	-1528,30	2483,60
	$t_5$			1635,30	1731,60	-604,68	3092,50
	$t_{95}$			21161,00	19764,00	12625,00	27928,00
	$t_{99}$			29561,00	25596,00	15718,00	36043,00
LSN	$t_1$			1549,90	1847,40	884,86	2664,80
	$t_5$			2251,00	2320,90	1372,70	3194,10
	$t_{95}$			17925,00	18920,00	8435,50	37436,00
	$t_{99}$			33924,00	37465,00	11586,00	87284,00
ST	$t_1$	<b>3449,10</b>	<b>2380,80</b>	<b>-4532,40</b>	<b>-5292,70</b>	<b>-35617,00</b>	<b>3959,90</b>
	$t_5$	<b>3593,30</b>	<b>3356,00</b>	<b>2832,50</b>	<b>1576,80</b>	<b>-4050,10</b>	<b>4153,80</b>
	$t_{95}$	<b>7661,80</b>	<b>15707,00</b>	<b>16510,00</b>	<b>32843,00</b>	<b>5037,30</b>	<b>113871,00</b>
	$t_{99}$	<b>12715,00</b>	<b>30978,00</b>	<b>97900,00</b>	<b>1,00E+06</b>	<b>6216,20</b>	<b>330094,00</b>
LST	$t_1$			90,10	1291,50	2,00E-13	3058,70
	$t_5$			2601,00	2601,50	404,27	3906,10
	$t_{95}$			13441,00	1,00E+77	5027,10	2,00E+07
	$t_{99}$			465716,00	2,00E+95	6702,40	1,00E+26
W	$t_1$			138,79	192,30	4,98	461,71
	$t_5$			547,89	676,69	82,34	1338,40
	$t_{95}$			19594,00	19270,00	11123,00	29897,00
	$t_{99}$			34043,00	28085,00	15137,00	45792,00

Tabela 4.5: Percentis da distribuição *a posteriori* dos tempos de falha usando a Distribuição Preditiva *a posteriori* e o método de Hamada (2005) em todos os modelos ajustados para o cenário B-LST

Ajuste	$t_p\%$	Real	Empírica	Oliveira (2011)		Hamada (2005)	
				Pred.	Post.	Média	$HPD_{INF}$
N	$t_1$			-772,00	-128,03	-2846,60	2078,80
	$t_5$			1203,00	1507,30	-682,75	3214,40
	$t_{95}$			9467,30	9401,40	7664,40	11619,00
	$t_{99}$			11722,00	11037,00	8736,90	13687,00
LN	$t_1$			2463,00	2603,00	1698,80	3303,50
	$t_5$			3081,40	3186,50	2312,20	3847,60
	$t_{95}$			8745,20	8664,40	6794,50	10972,00
	$t_{99}$			11630,00	10718,00	7900,30	14551,00
T	$t_1$			-8859,10	-4,00E+06	-217976,00	3470,70
	$t_5$			2258,40	-1802,20	-5869,80	4219,60
	$t_{95}$			7425,00	11328,00	5218,00	15365,00
	$t_{99}$			21326,00	4,00E+06	6174,10	227386,00
LT	$t_1$			677,53	1461,60	0,00	3082,90
	$t_5$			3278,60	3013,40	1009,50	4149,20
	$t_{95}$			7093,90	9,00E+19	5301,00	20261,00
	$t_{99}$			26175,00	$\infty$	5815,60	2,00E+11
SN	$t_1$			3369,00	3693,80	2901,10	4168,80
	$t_5$			3892,80	3897,90	3172,50	4356,80
	$t_{95}$			10001,00	9757,30	7674,90	12156,00
	$t_{99}$			12050,00	11652,00	8895,30	14880,00
LSN	$t_1$			3290,00	3611,10	2843,50	4165,70
	$t_5$			3823,90	3882,30	3251,50	4336,20
	$t_{95}$			9672,10	9391,90	6875,40	13463,00
	$t_{99}$			13263,00	12454,00	7478,70	19136,00
ST	$t_1$			<b>3758,00</b>	<b>3690,70</b>	3213,50	4176,10
	$t_5$			<b>4035,00</b>	<b>3972,60</b>	3525,30	4354,20
	$t_{95}$			<b>10253,00</b>	<b>16855,00</b>	5638,00	26247,00
	$t_{99}$			<b>28965,00</b>	<b>1,00E+06</b>	6642,10	109397,00
LST	$t_1$	<b>3963,70</b>	<b>4071,10</b>	2572,20	3357,90	1378,00	4213,40
	$t_5$	<b>4068,00</b>	<b>4162,00</b>	3832,90	3805,40	2788,10	4375,20
	$t_{95}$	<b>7029,50</b>	<b>8648,00</b>	9405,30	5,00E+31	5524,70	43809,00
	$t_{99}$	<b>10131,00</b>	<b>12435,00</b>	28807,00	4,00E+41	6570,30	1,00E+09
W	$t_1$			698,78	956,52	299,92	1635,70
	$t_5$			1710,90	1843,10	906,92	2835,60
	$t_{95}$			10191,00	10039,00	7942,60	12802,00
	$t_{99}$			13282,00	12078,00	9427,60	15879,00

Tabela 4.6: Percentis da distribuição *a posteriori* dos tempos de falha usando a Distribuição Preditiva *a posteriori* e o método de Hamada (2005) em todos os modelos ajustados para o cenário B-W

Ajuste	$t_{p\%}$	Real	Empírica	Oliveira (2011)		Hamada (2005)	
				Pred.	Post.	Média	$HPD_{INF}$
N	$t_1$			2417,97	2656,36	1466,90	3665,22
	$t_5$			3259,50	3393,12	2421,74	4175,21
	$t_{95}$			7009,50	6949,57	6074,74	7837,42
	$t_{99}$			8023,28	7686,33	6623,55	8853,68
LN	$t_1$			2803,70	3067,16	2356,44	3693,31
	$t_5$			3413,70	3553,57	2937,36	4159,74
	$t_{95}$			7373,45	7330,12	6092,00	8849,56
	$t_{99}$			9280,32	8542,20	6806,31	10873,37
T	$t_1$			1506,72	-662,42	-5144,43	3829,75
	$t_5$			3104,00	3033,24	1514,52	4299,45
	$t_{95}$			7277,65	7314,49	6040,64	8825,67
	$t_{99}$			8676,15	11010,16	6538,31	15740,43
LT	$t_1$			2571,99	2449,27	521,20	3722,14
	$t_5$			3334,00	3406,13	2413,08	4232,01
	$t_{95}$			7610,20	9152,03	6143,18	10756,90
	$t_{99}$			10980,20	6,49491E+33	6466,47	43114,96
SN	$t_1$			1372,99	1955,85	-380,78	3564,29
	$t_5$			2917,35	2966,25	1365,86	4132,50
	$t_{95}$			7046,15	7031,55	6175,70	8321,99
	$t_{99}$			8220,25	7649,36	6532,87	9823,35
LSN	$t_1$			2483,86	2752,27	1730,51	3630,76
	$t_5$			3210,90	3327,21	2393,43	4090,11
	$t_{95}$			7217,10	7211,28	6134,34	8846,56
	$t_{99}$			8556,43	8189,99	6514,28	10886,10
ST	$t_1$			-654,69	-1139,12	-9712,48	3818,21
	$t_5$			2415,25	2231,09	-1091,07	4365,76
	$t_{95}$			7058,05	7486,56	6189,02	9669,37
	$t_{99}$			9300,45	8791,98	6506,99	14324,38
LST	$t_1$			1420,90	2158,19	0,00	3399,65
	$t_5$			3003,75	3050,89	1522,26	4193,78
	$t_{95}$			7522,00	16872289,80	6066,04	11031,30
	$t_{99}$			10620,40	5,91213E+23	6491,92	25734,81
W	$t_1$	<b>2555,01</b>	<b>3290,52</b>	<b>2410,69</b>	<b>2536,56</b>	1676,04	3431,10
	$t_5$	<b>3352,53</b>	<b>3703,24</b>	<b>3386,90</b>	<b>3349,26</b>	2474,64	4097,51
	$t_{95}$	<b>6603,59</b>	<b>6507,93</b>	<b>6890,05</b>	<b>6804,39</b>	6149,78	7554,67
	$t_{99}$	<b>7094,20</b>	<b>6753,31</b>	<b>7608,52</b>	<b>7343,94</b>	6562,31	8332,71

No cenário B-ST observamos que as estimativas fornecidas pelos modelos ajustados, na Tabela 4.4, são bastante afetadas principalmente nas caudas (percentil 1% e 99%). Por exemplo, para o ajuste  $ST$  via distribuição preditiva proposta por Oliveira (2011), que foi o melhor, o tempo em que 1% dos itens futuros falharão é de  $-4532,40$  (sequer é positivo), distante do valor real que é  $3449,10$ . O percentil 99% é  $97900,00$  enquanto o real é  $12715,00$ . Os percentis 5% e 95% estão mais próximos dos respectivos valores reais. Se observarmos a curva da fdp na Figura 4.14 isso não parece razoável, uma vez que a fdp do ajuste  $ST$  parece ter sido a mais próxima da real.

Para o cenário B-LST percebemos pela Tabela 4.5 que, no ajuste  $ST$  pelo método proposto por Oliveira (2011), o tempo em que 1% dos itens futuros falharão é de  $3758,00$ , e  $3963,70$  para a distribuição real. Já para o percentil 99% o tempo correspondente é de  $28965,00$ , mais que o dobro do real que é de  $12435,00$ .

A Tabela 4.6 apresenta os percentis estimados no cenário B-W para todos os nove modelos. As estimativas pontuais são mais próximas das reais para o ajuste  $W$ . Por exemplo, para o ajuste  $W$  o tempo em que 1% dos itens futuros falharão é de  $2410,69$  via proposta distribuição preditiva *a posteriori* de  $T_{n+1}$ , e  $2536,56$  via Hamada (2005), enquanto na distribuição real é de  $2555,01$  e na empírica amostral de  $3290,52$ . As estimativas fornecidas pelos demais modelos fornecem piores estimativas que o modelo  $W$ , porém não tão distante dos valores reais como ocorreu nas análises dos outros bancos de dados.

Um comentário importante diz respeito às dificuldades numéricas observadas na abordagem de Hamada (2005). Nos ajustes  $T$ ,  $LT$ ,  $ST$  e  $LST$ , para alguns valores da distribuição *a posteriori* dos parâmetros não é possível obter o respectivo percentil<sup>3</sup>,

---

<sup>3</sup>Para obter percentis das distribuições  $ST$ ,  $LST$ ,  $T$  e  $LT$  podemos usar o pacote “sn” de Azzalini (2015), no software R. Porém, para alguns valores de parâmetros a função `qst()` deste pacote não é capaz de retornar o valor do percentil. Uma solução é utilizar a função `uniroot()` do pacote “stats” (R Development Core Team, 2005) para construir uma função de busca numérica combinada com a função `pst()`. Ainda assim, se o percentil estiver localizado nos extremos das caudas da correspondente distribuição acumulada, a função `pst()` pode entrar em *looping* infinito e não fornecerá o percentil requerido

o que empobrece a estimativa do mesmo neste método. Por este motivo, para alguns dos ajustes pelos modelos de caudas pesadas, as estimativas foram muito distantes dos valores reais no que se refere aos percentis  $t_1$  ou  $t_{99}$ . Esse padrão se repete para todos os bancos de dados.

Portanto, além de ser mais coerente com o paradigma bayesiano e de interpretação mais simples, a abordagem proposta por Oliveira (2011) mostrou-se menos suscetível a problemas numéricos. Ainda, o tempo computacional para extrair as quantidades de interesse é quase que instantâneo, basta que se aplique uma função simples que permita extrair quantis da amostra *a posteriori* dos tempos de falha.

Vale lembrar que o método de Hamada (2005) para a estimação dos percentis possuía uma inconsistência teórica, para estimação dos percentis, e deficiências na interpretação dos intervalos de credibilidade.

As Figuras 4.14 e 4.15 parecem fornecer evidências que a verdadeira distribuição dos tempos de falha nos cenários SN, ST e W resultam em estimativas próximas das reais. Já para o cenário LST o modelo ST parece ser o mais adequado, contudo ao menos o modelo mais indicado pelos gráficos é assimétrico e de cauda pesada.

Finalmente, é necessário avaliarmos a qualidade de ajuste dos modelos usando critérios científicos. As medidas de qualidade de ajuste e diagnóstico, e critérios de seleção de modelo da Seção 4.3 podem ser utilizadas para este fim. A pergunta que se faz é se a melhor alternativa para fazermos esta avaliação é a partir do modelo ajustado para as observações individuais de degradação  $y_{ij}$  (Deg), através dos perfis de degradação  $\mathbf{y}_i$ , pelos tempos de falha  $t_i$  (TF). Ou ainda, como em situações práticas não se dispõe dos tempos de falha, medir a qualidade de ajuste utilizando os pseudo tempos de falha  $\hat{t}_i$  (PTF), estimados conforme a Seção 3.1.1. A Seção 4.3.7 indica como direcionar as medidas de qualidade de ajuste e critério de seleção para cada tipo de foco, em degradação, em perfis de degradação, em tempos de falha e pseudo tempos de falha.

A Tabela 4.7 apresenta os melhores (no topo e em negrito) e piores (pé) modelos se-

leccionados segundo os critérios LPML, DIC, WAIC, FB, baseados na degradação (Deg), tempos reais (T) e pseudo tempos de falha (PTF), assim como o teste de Kolmogorov-Smirnov sobre a fda empírica (KS-FDA EMP), fda real (KS-FDA REAL) e com particionamento (“gradeamento”) do eixo do tempo (KSG). No caso do teste KS também foram utilizados os tempos reais (T) e pseudo tempos de falha (PTF), exceto quando se gradeou o domínio dos tempos de falha para varrer ao máximo as distâncias dos modelos ajustados para a fda real. A ordenação de LPML, DIC e WAIC foi a mesma para TF e PTF em todos os casos e por esta razão foram condensadas na tabela.

A seguir apresentaremos as conclusões sobre os critérios de seleção e ajuste presentes na Tabela 4.7. Os comentários referenciarão as tabelas que contêm os valores numéricos para tais critérios, as quais encontram-se no Anexo C. Por se tratarem de muitas tabelas, recomendamos a leitura pela Tabela 4.7, a qual foi construída exclusivamente para facilitar a interpretação condensada dos resultados.

As Tabelas C.1 a C.3 (Anexo C) apresentam, respectivamente as medidas de qualidade de ajuste (LPML, DIC, WAIC) baseadas nas observações individuais  $y_{ij}$  e nos perfis  $\mathbf{y}_i$ , nos tempos de falha  $t_i$  (TF) e nos pseudo tempos de falha  $\hat{t}_i$  (PTF).





A partir da Tabela C.1 (Anexo C), correspondente a medidas de qualidade de ajuste baseadas na degradação, e com auxílio da Tabela 4.7

- LPML, DIC e WAIC selecionaram o modelo LSN para o cenário B-SN, que na análise das Figuras 4.14 e 4.15 aparentava ser o segundo melhor. Já  $LPML_{perfil}$  seleciona o modelo LST. Observe, entretanto, que esta conclusão é válida somente se a análise gráfica for voltada para a distribuição real. Se observarmos os histogramas dos tempos de falha, o modelo LST acompanha bem o mesmo, assim como o modelo LSN. Portanto, as medidas com foco na degradação não se mostram capazes de selecionar o modelo da fda real, contudo selecionam um modelo que acompanha bem o histograma dos tempos de falha gerados. Uma observação importante diz respeito a  $LPML_{perfil}$ , para qual SN é o pior modelo dentre todos, conclusão bastante incoerente;
- no cenário B-ST, LPML, DIC e WAIC indicam o modelo ST. Já  $LPML_{perfil}$  indica o modelo T. Isto é, a princípio as medidas com foco na degradação (LPML, DIC e WAIC) conseguem apontar a distribuição real dos dados gerados artificialmente. Note que a segunda melhor escolha, segundo LPML, DIC e WAIC, é pelo modelo LN. Pelas Figuras 4.14 e 4.15 o modelo LN é um dos piores e novamente a degradação sugere bem o modelo real mas no restante da ordenação dos modelos (melhor até o pior) carrega certa incoerência;
- para o cenário B-LST, LPML, DIC e WAIC selecionam o modelo LST,  $LPML_{perfil}$  o modelo T. Novamente, as medidas com foco na degradação tendem a apontar o modelo real dos dados gerados. Contudo, observamos que o pior modelo apontado por  $LPML_{perfil}$  é o LT, que na análise das Figuras 4.14 e 4.15 está entre os melhores. LPML, DIC e WAIC também apresentam certa incoerência colocando o modelo ST como um dos piores, quando na análise gráfica tal modelo estava dentre os mais próximos da distribuição real;

- considerando o cenário B-W, o modelo selecionado por todas as medidas é o LSN, diferente da distribuição geradora do cenário. Todavia, pelas Figuras 4.14 e 4.15 os modelos ajustados acompanham bem a curva real e os dados empíricos, o que dificultaria a discriminação de modelos independentemente do critério de seleção.

A partir da Tabela C.2 (Anexo C), correspondente a medidas de qualidade de ajuste baseadas nos tempos de falha reais, e pela Tabela 4.7

- os critérios indicam que melhor modelo ajustado ao cenário B-SN é fornecido pela distribuição SN, ou seja, indicam a distribuição verdadeira;
- no cenário B-ST, LPML, DIC selecionam o modelo LT. Já WAIC indica o modelo LSN. Isto é, a princípio as medidas baseadas nos tempos de falha direcionaram a seleção de modelos de caudas pesadas ou com assimetria, características do cenário gerado. Ainda, pela análise das Figuras 4.14 e 4.15, nota-se que o modelo LT é dos que melhor se ajustam à distribuição real e à distribuição empírica;
- para o cenário B-LST, LPML selecionou o modelo LST, enquanto DIC e WAIC o modelo ST, o que significa que os critérios apontaram na direção da distribuição verdadeira;
- considerando o cenário B-W, o modelo selecionado por LPML e DIC é o W, já WAIC seleciona o modelo LSN.

Observe que o desempenho das medidas de qualidade de ajuste, em termos de seleção da distribuição geradora real, quando o foco são os tempos de falha reais é tão satisfatório quanto quando o foco é a degradação, e até melhor se considerarmos a ordenação do melhor para o pior modelo. Porém, em situações práticas não dispomos dos tempos de falha, apenas das medidas de degradação. Para contornar este problema podemos utilizar os pseudo tempos de falha. Pelas Tabelas C.3 (Anexo C), correspondente a medidas de qualidade de ajuste baseadas nos pseudo tempos de falha, e 4.7, as

conclusões são exatamente as mesmas daquelas obtidas a partir dos tempos de falha. Ou seja, os pseudotempo de falha podem ser utilizados para calcular os critérios de qualidade de ajuste.

Outra forma para auxiliar na escolha do melhor modelo é através do Fator de Bayes apresentado nas Tabelas C.4 a C.6 (Anexo C). Um modelo (linha) é melhor que outro (coluna) quando FB maior que 1, e quando isso ocorre destacamos o FB em negrito.

Da ordenação do FB, na Tabela 4.7, notamos que o Fator de Bayes não seleciona o modelo real para a maioria dos cenários (acertam somente no cenário SN). Contudo os FB baseados em TF e PTF ao menos apresentam uma ordenação de modelos em direção coerente com a dos dados gerados e com as análises das Figuras 4.14 e 4.15. Por exemplo, para o cenário B-LST o FB baseado em TF e PTF aponta que os melhores modelos são os de caudas pesadas e assimétricos, e os piores são W e N, enquanto para FB baseado na degradação indica W como o melhor modelo e LT como o pior. Portanto para estes bancos de dados artificiais critério FB, quando baseado nas degradações, não se mostrou adequado para ser aplicado. Quando o FB se baseia nos tempos e pseudo tempos de falha poderá ser útil para indicar a direção do melhor ajuste.

No contexto de avaliação da qualidade de ajuste através da distribuição dos tempos de falha é possível aplicar a estatística de Kolmogorov-Smirnov (KS). No caso, tal estatística pode ser aplicada tanto para os tempos de falha quanto para os pseudo tempos de falha. Utiliza-se o método da Seção 4.3.5 em que queremos avaliar a distância do modelo real para os dados empíricos, do modelo estimado pelo método de Hamada (2005) para os dados empíricos, e do modelo estimado com base na amostra *a posteriori* na distribuição preditiva de  $T_{n+1}$  para os dados empíricos. Nessa mesma linha repete-se a avaliação em termos dos pseudo tempos de falha.

As Tabelas C.7 a C.7 (Anexo C) apresentam as referidas estatísticas K-S dos nove modelos ajustados aos cenários B-SN, B-ST, B-LST e B-W.

Com auxílio da Tabela 4.7, no cenário

- B-SN notamos pela Tabela C.7 que segundo o teste K-S o modelo que melhor se ajusta aos dados gerados é o ST, se usamos os tempos reais até a falha, ou os pseudotempos de falha somente para o método de Hamada (2005). O modelo ST é o selecionado se usamos o método proposto por Oliveira (2011). O pior desempenho é o do modelo N, seguido do modelo W e LN. É importante notar que o modelo real (aquele dos dados simulados e em que se conhece os verdadeiros valores dos parâmetros) quando comparado com os tempos de falha reais e os pseudo tempos de falha apresentou um p-valor elevado para a estatística  $D_n$ , tanto para os tempos de falha (0,93749) quanto para os pseudo tempos de falha (0,97806). Se formos considerar este p-valor, percebe-se que os p-valores dos ajustes pelos modelos T, LT, LSN e LST também fornecem ajustes satisfatórios. Se observarmos a ordenação do melhor para o pior modelo, verifica-se que o modelo SN é apontado como um dos piores, o que leva a crer que o critério não foi capaz de apontar o melhor modelo.
- B-ST a Tabela C.8 o melhor ajuste pela estatística K-S para o modelo LST. O modelo ST é o segundo melhor ajuste se observamos a estatística K-S pelos pseudo tempos de falha em ambos os métodos de estimação, e tempos de falha para o método de Hamada (2005). Os melhores ajustes subsequentes são fornecidos pelos modelos T e LT. Os piores resultados são dados pelos modelos SN, N, W, LN e LSN. Ou seja, novamente os modelos de cauda pesada se destacam como melhores ajustes. Vale ressaltar que quando comparamos a distribuição real com os tempos e pseudo tempos empíricos, o valor da estatística K-S não foi tão elevado quando comparado com o valor dos melhores modelos ajustados.
- B-LST percebemos pela Tabela C.9 que o melhor modelo ajustado para os tempos reais é o LT, já para os pseudo tempos o modelo LST. O segundo melhor modelo pelos pseudo tempos é o LT, já para os tempos é o LST no caso de considerarmos o

método de Hamada (2005) e T se utilizarmos a proposta via distribuição preditiva *a posteriori* de  $T_{n+1}$ . Os piores ajustes são os referentes aos modelos SN, N, W, LSN e LN.

- B-W, pela Tabela C.10 constata-se que todos os modelos ficaram bem ajustados, de modo que a diferença dos modelos pelo teste K-S foi pequena. Pelo teste K-S o melhores modelos pelo métodos de Hamada (2005) e o proposto foram, respectivamente, N e T. Novamente, como para o cenário B-SN, pela ordenação do melhor para o pior modelo o critério não foi capaz de apontar o melhor modelo, uma vez que W se encontra entre os piores.

Pelos resultados constatamos que o teste K-S baseado nos tempos e pseudo tempos empíricos não apontou sempre o modelo real dos bancos de dados simulados, entretanto permitiu escolher modelos condizentes com as características de assimetria e caudas pesadas quando estas estavam presentes.

Uma vez que dispomos do conhecimento da fda real para os cenários simulados, podemos considerar a distância  $D_n$  entre a curva real e os modelos estimados. A Tabela C.11 (Anexo C) fornece os resultados destas distâncias para todos os cenários.

Pela ordenação dos modelos na Tabela 4.7, no cenário

- B-SN, considerando a distância entre a fda real e a estimada, observamos que o método de Hamada (2005) fornece as menores distâncias no modelo SN. Pelo método proposto em Oliveira (2011) as menores distâncias são referentes ao modelo LSN. Este resultado independe se foi utilizado o tempo de falha ou o pseudo tempo de falha como pontos referenciais para comparação das fda. Perceba também que as distâncias não são muito diferentes para os modelos.
- B-ST, o modelo LT é o que fornece a menor distância para a fda real em todos os métodos de estimação. Note que as distâncias são mais discriminantes que

no caso dos modelos ajustados para o cenário B-SN. É possível perceber que os modelos SN, N, W e LSN fornecem as maiores distâncias da curva real.

- B-LST, a menor distância é referente ao modelo T para todos os métodos de estimação. Os modelos mais distantes da fda real são o SN, N, W e LSN. Observe que as diferenças entre as distâncias dos modelos é grande, assim como no cenário B-ST.
- B-W a distância menor é dada pelo modelo W, se consideramos os método de Hamada (2005). Pelo método proposto a menor distância é referente ao modelo N. Repare que as distâncias são muito pequenas quando se compara os modelos, menores até que para o cenário B-SN, o que leva a crer que a maior parte dos modelos se ajusta bem à fda real.

Assim como ocorreu para todos os critérios de seleção e ajuste dos modelos apresentados anteriormente, o teste K-S também apresentou deficiências para apontar o modelo real dos dados gerados, contudo forneceu apontamentos na direção de modelos com características similares às reais quando os dados possuíam valores atípicos.

Os cenários foram gerados com pequenos tamanhos de amostra a fim de representar situações práticas realísticas. Por esta razão o tamanho da amostra pode influenciar a inferência e a seleção de modelos pelos critérios utilizados. Uma boa maneira de avaliar se os critérios permitiram selecionar as distribuições mais próximas das geradoras do cenário é particionar (“gradear”) o espaço amostral dos tempos de falha e calcular a distância máxima das estimativas obtidas em cada modelo ajustado para a fda real. Isso só é possível porque estamos lidando com cenários simulados, o que também nos permite avaliar melhor a adequabilidade dos critérios de seleção.

A Tabela C.12 (Anexo C) fornece as distâncias obtidas a partir desse particionamento. Os melhores e piores modelos na Tabela 4.7 são, em sua maioria, aqueles mesmos selecionados pelos critérios LPML, DIC, WAIC e KS EMP baseados nos tempos de fa-

lha e pseudo tempos de falha, com destaque para LPML e DIC. Por exemplo, para o cenário ST os quatro melhores modelos por KSG são LST, ST, LT e T, e por LPML e DIC também. Isto leva a crer que estes critérios conseguem apontar os melhores ajustes.

A capacidade preditiva dos modelos (para  $y_{ij}$ ,  $\mathbf{y}_i$ ,  $t_i$  e  $\hat{t}_i$ ) pode ser medida através da estatística  $CPO_i$  na Figura 4.17, a qual mostra que os modelos não são muito diferentes segundo sua capacidade preditiva, tanto para foco nas medidas de degradação quanto para o foco nos perfis. Note que a maioria dos pontos se sobrepõem em ambos os gráficos, além disso nenhum ajuste apresenta valores de  $CPO_i$  persistentemente maiores para todas as observações. Ou seja, não existe indicação explícita do melhor ajuste quando a  $CPO_i$  é calculada com base nas medidas de degradação, ou nos perfis de degradação.

Por outro lado, na Figura 4.18, quando a  $CPO_i$  é baseada nos tempos, ou pseudo tempos de falha, é possível distinguir os modelos com melhor e pior capacidade preditiva. Por exemplo, para a maioria dos cenários os modelos W e N apresentaram capacidade preditiva predominantemente inferior aos demais. Em geral os modelos T, ST, LT e LST forneceram boa capacidade preditiva, principalmente para os cenários B-ST e B-LST. Os modelos LSN e SN apresentam baixa  $CPO_i$  na presença dos valores atípicos dos cenários B-ST e B-LST, o que aconteceu também com o modelo LN. Para o cenário B-SN os modelos SN, LSN, ST e LST apresentaram os melhores resultados de  $CPO_i$ , seguidos pelos de caudas pesadas T e LT, e finalmente pelos modelos N e W.

As medidas de diagnóstico de influência indicam as observações influentes para estimação dos parâmetros do modelo. As Figuras 4.19 a 4.22 apresentam os gráficos das estatísticas  $KL_i$  e  $p_i$ , utilizando a degradação  $y_{ij}$ , os perfis  $\mathbf{y}_i$ , os tempos de falha  $t_i$  e os pseudo tempos de falha  $\hat{t}_i$ , dos nove modelos ajustados para os cenários B-SN, B-ST, B-LST e B-W.



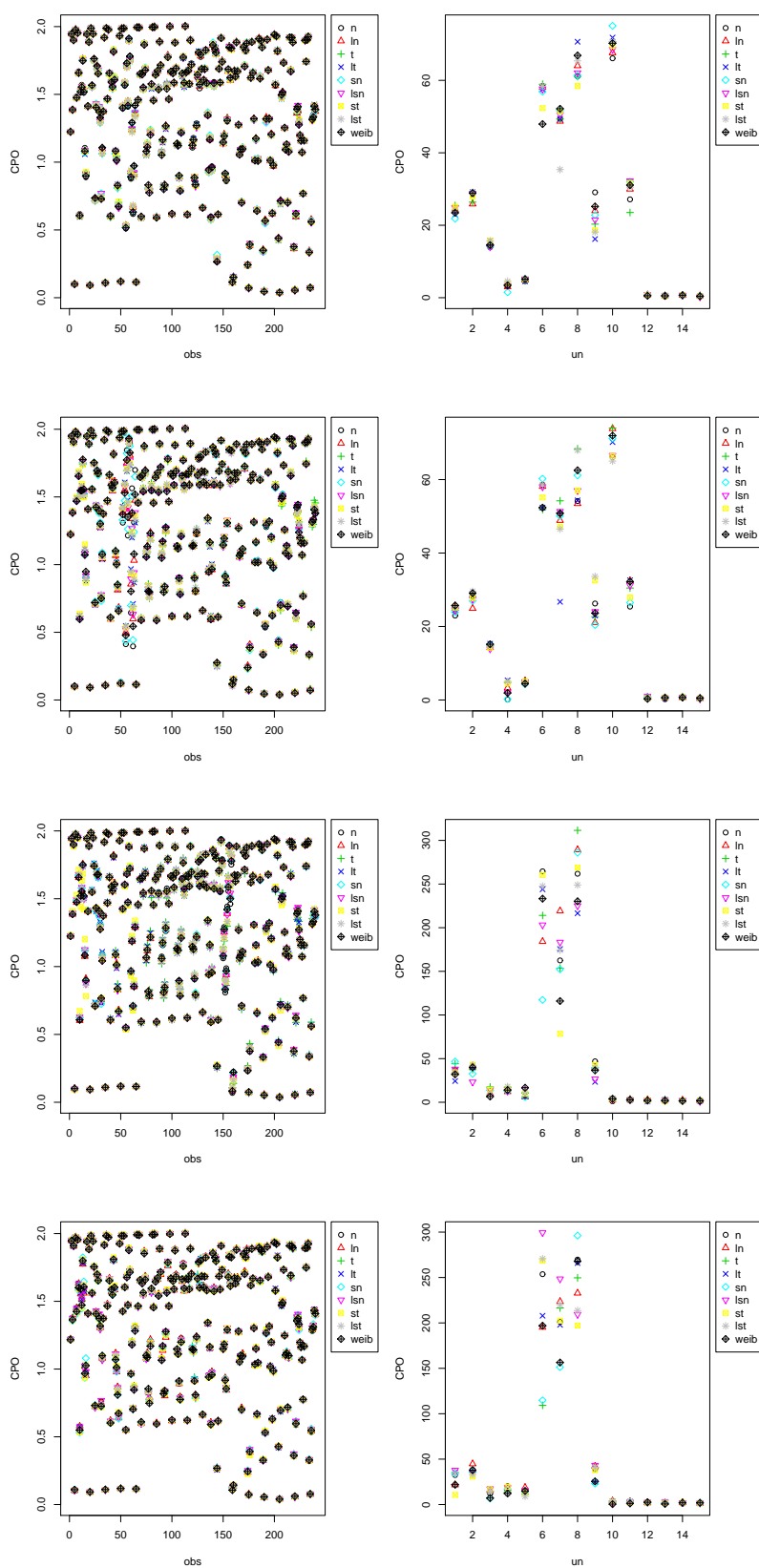


Figura 4.17:  $CPO_i$  por degradação  $y_{ij}$  (esquerda) e perfil  $y_i$  (direita), nos cenários B-SN (primeira linha), B-ST (segunda linha), B-LST (terceira linha) e B-W (quarta linha).

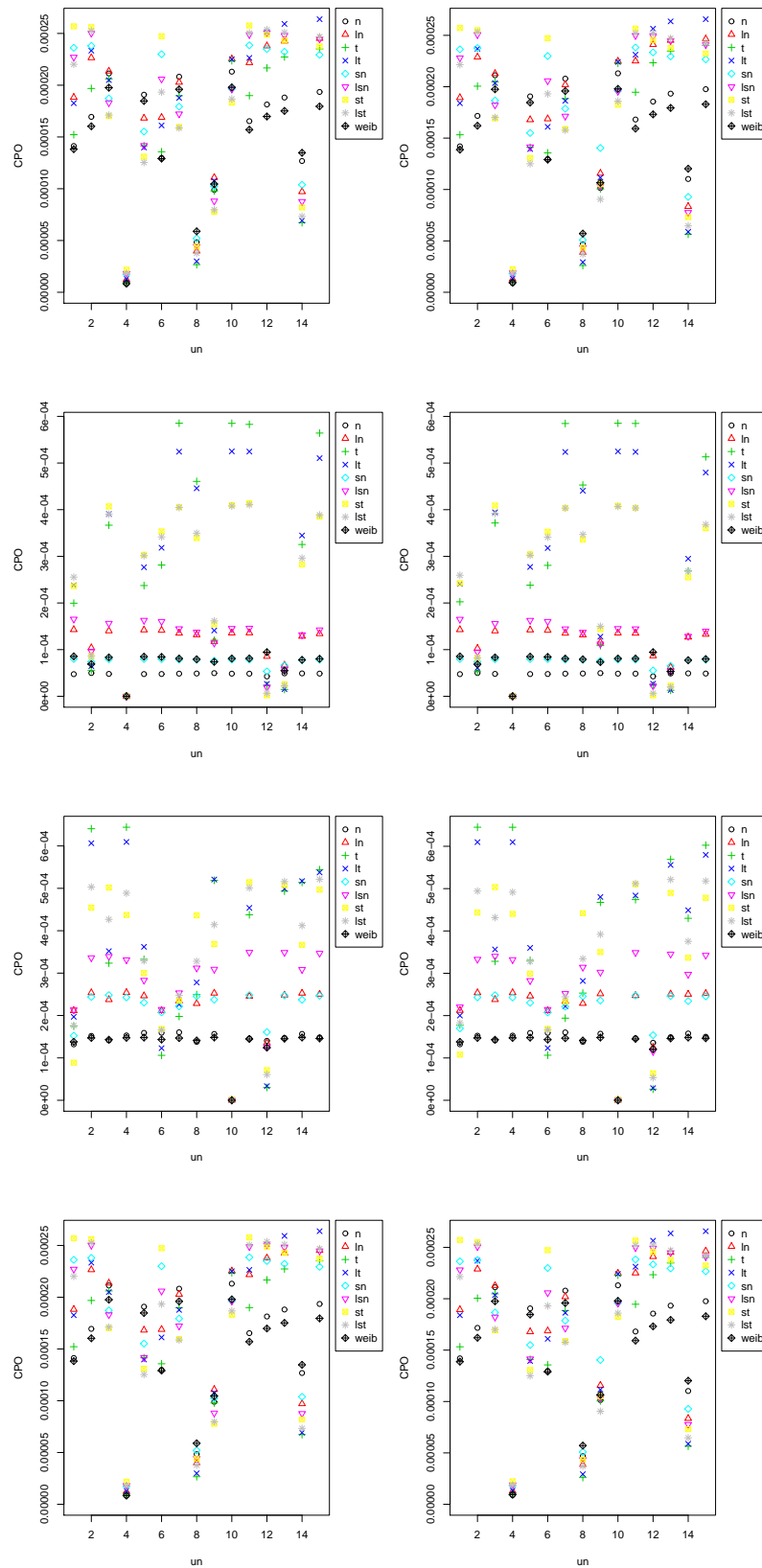


Figura 4.18:  $CPO_i$  por tempo de falha (esquerda) e pseudo tempo de falha (direita), nos cenários B-SN (primeira linha), B-ST (segunda linha), B-LST (terceira linha) e B-W (quarta linha).

Pelas Figuras 4.19 e 4.20 notamos que as medidas de diagnóstico de influência para as observações  $y_{ij}$  e  $\mathbf{y}_i$  praticamente não diferem de modelo para modelo. A medida  $KL_i$  aponta vários pontos influentes na estimação dos parâmetros da maioria dos modelos ajustados em todos os cenários. Se for observada a versão calibrada  $p_i$ , verificamos que todos os perfis são influentes para a inferência dos objetos aleatórios do modelo, pois todos distam bastante do valor desejado (0,5) para esta medida. Noutras palavras, basicamente as observações e perfis afetam a inferência da mesma maneira em todos os modelos, ou esta medida não é adequada se aplicada às degradações a fim de concluir sobre a robustez em termos da distribuição dos tempos de falha.

Quando consideramos a medida de divergência KL e  $p_i$ , respectivamente nas Figuras 4.21 e 4.22, para os tempos de falha e pseudo tempos de falha, então para o cenário

- B-SN as divergências de Kullback Leibler indicam que as unidades 4 e 9 são pontos influentes para grande parte dos modelos ajustados, e os modelos LT, T, LST LSN e W acomodam melhor estes pontos influentes. Considerando  $p_i$  verifica-se que todos as inferências são influenciadas pelas unidades 4, 8 e 9, pois todos distam bastante do valor (0,5). O modelo LT é o menos suscetível à influência destas observações.
- B-ST a divergência KL indica a unidade 4 como influente para todos os ajustes. Observando  $p_i$  verifica-se as observações 4 e 12 como pontos influentes. Os modelos T e LT são os menos afetados por estas observações.
- B-LST os pontos 1 e 10 são os mais influentes, em que o ponto 1 afeta especialmente os modelos SN, LSN, ST, LST.
- B-W fornece  $KL_i$  e  $p_i$  elevado para as observações 1 e 3. Tais observações afetam a inferência em todos os modelos de forma bem parecida.

Por estes resultados não se pode dizer que algum modelo tenha sido robusto o

suficiente para modelar todas as quatro base de dados geradas. Em todos os casos sempre que ajustamos os modelos reais os resultados foram insatisfatórios, ou seja, os modelos reais também são afetados pelos pontos influentes. Desse modo a medida KL não se apresenta muito conclusiva.

### 4.4.3 Conclusões

Pelos resultados encontrados nas análises dos quatro cenários simulados, podemos concluir que a modelagem de degradação e os procedimentos inferenciais desenvolvidos nas Seções 4, 3.2 e 4.2 fornecem boas estimativas da distribuição dos tempos de falha para unidades futuras. Isso pôde ser notado por conta da inferência sobre os parâmetros de cada ajuste quando a distribuição correspondia aos dados gerados (Figura 4.10 a 4.13), e também pela análise gráfica das fda e fdp preditivas *a posteriori* dos tempos de falha nas Figuras 4.14 e 4.15.

Durante o procedimento inferencial notou-se também que, uma vez de posse das estimativas *a posteriori* dos objetos aleatórios do modelo de degradação, o método proposto em Oliveira (2011) é computacionalmente muito menos intensivo que o de Hamada (2005), principalmente para as distribuições assimétricas e ou de cauda pesada. Isso ocorre porque o método de Hamada (2005) exige avaliar a amostra *a posteriori* de todos parâmetros no estimador da quantidade de interesse, o que significa 2000 avaliações se a amostra *a posteriori* tem tamanho 2000. Em contrapartida, o método proposto por Oliveira (2011) permite obter a quantidade de interesse diretamente da amostra da distribuição preditiva *a posteriori* dos tempos de falha executando apenas uma única operação. Portanto fornece respostas mais rápidas para a análise dos dados.

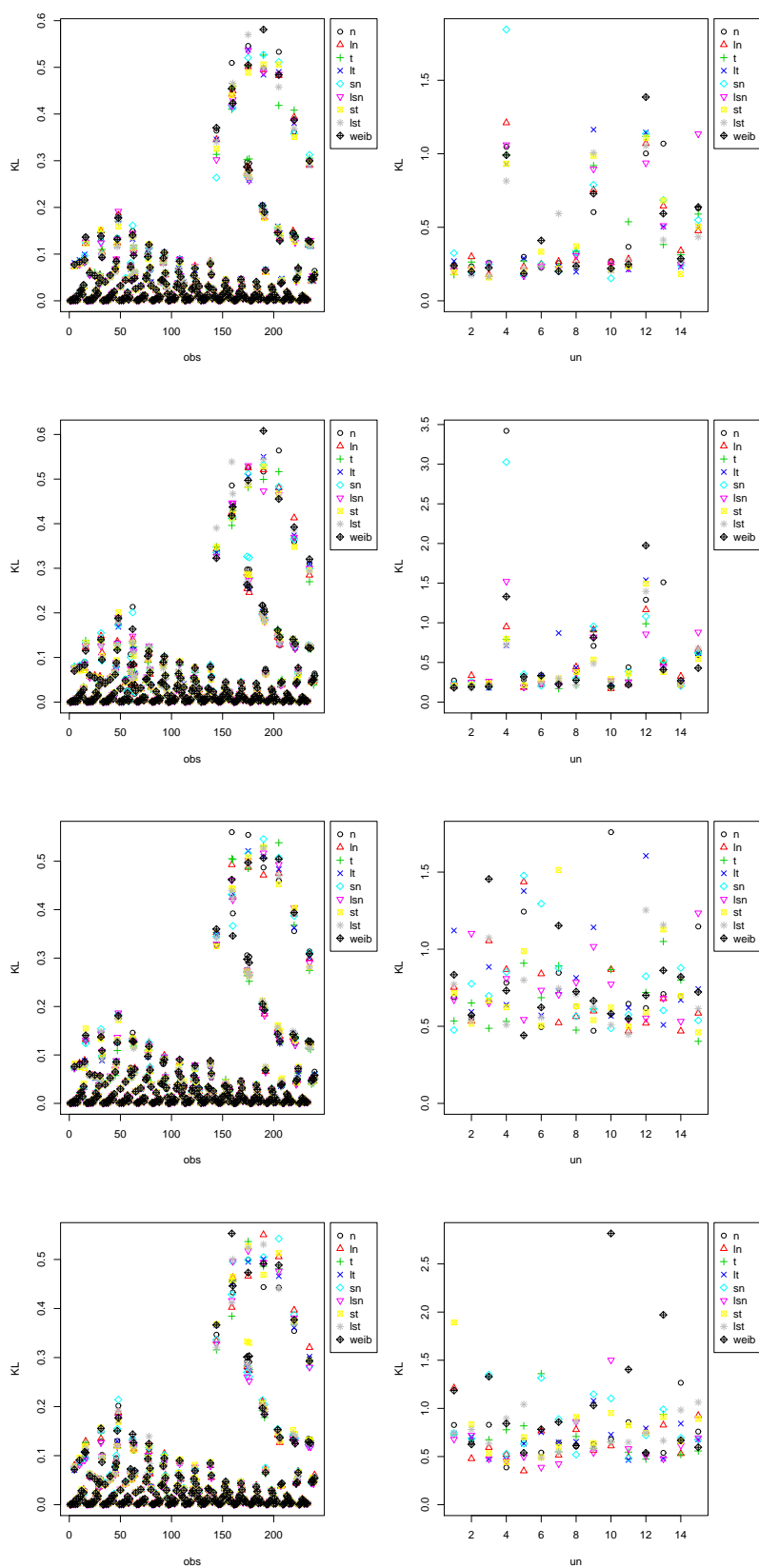


Figura 4.19:  $KL_i$  por degradação  $y_{ij}$  (esquerda) e por perfil  $y_i$  (direita), nos cenários B-SN (primeira linha), B-ST (segunda linha), B-LST (terceira linha) e B-W (quarta linha).

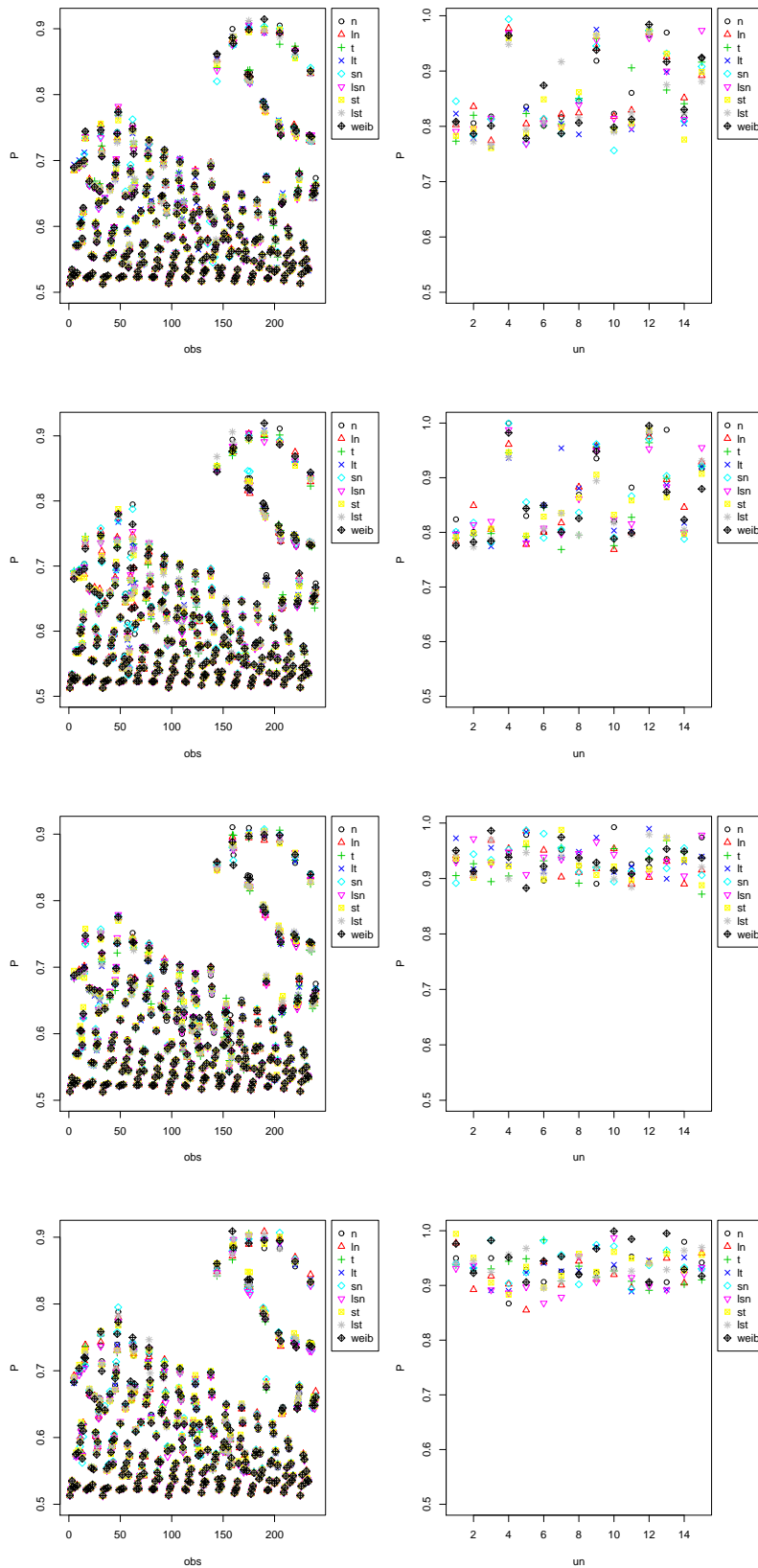


Figura 4.20:  $p_i$  por degradação  $y_{ij}$  (esquerda) e por perfil  $y_i$  (direita), nos cenários B-SN (primeira linha), B-ST (segunda linha), B-LST (terceira linha) e B-W (quarta linha).

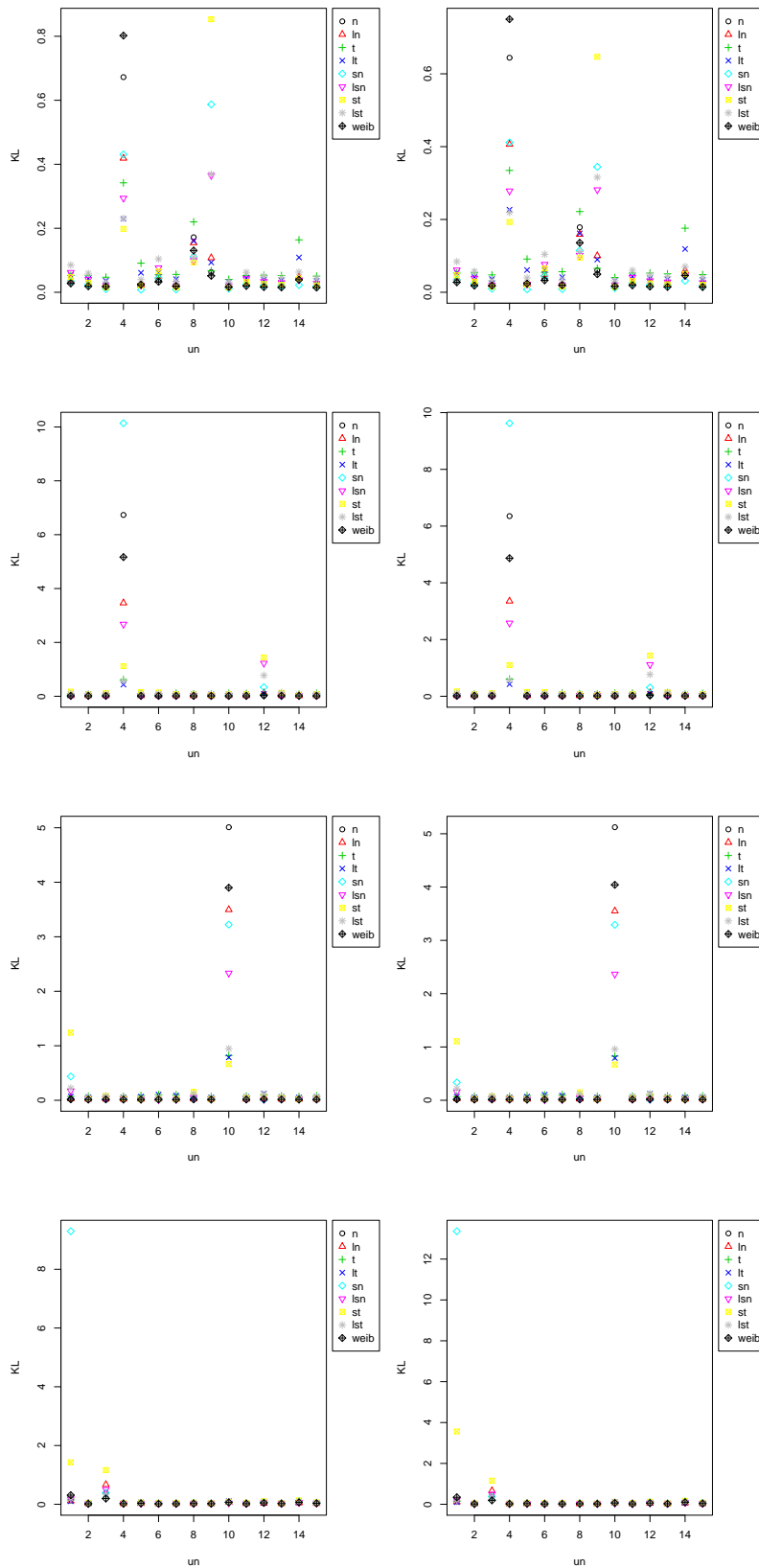


Figura 4.21:  $KL_i$  por tempo de falha  $t_i$  (esquerda) e por pseudotempo de falha  $\hat{t}_i$  (direita), nos cenários B-SN (primeira linha), B-ST (segunda linha), B-LST (terceira linha) e B-W (quarta linha).

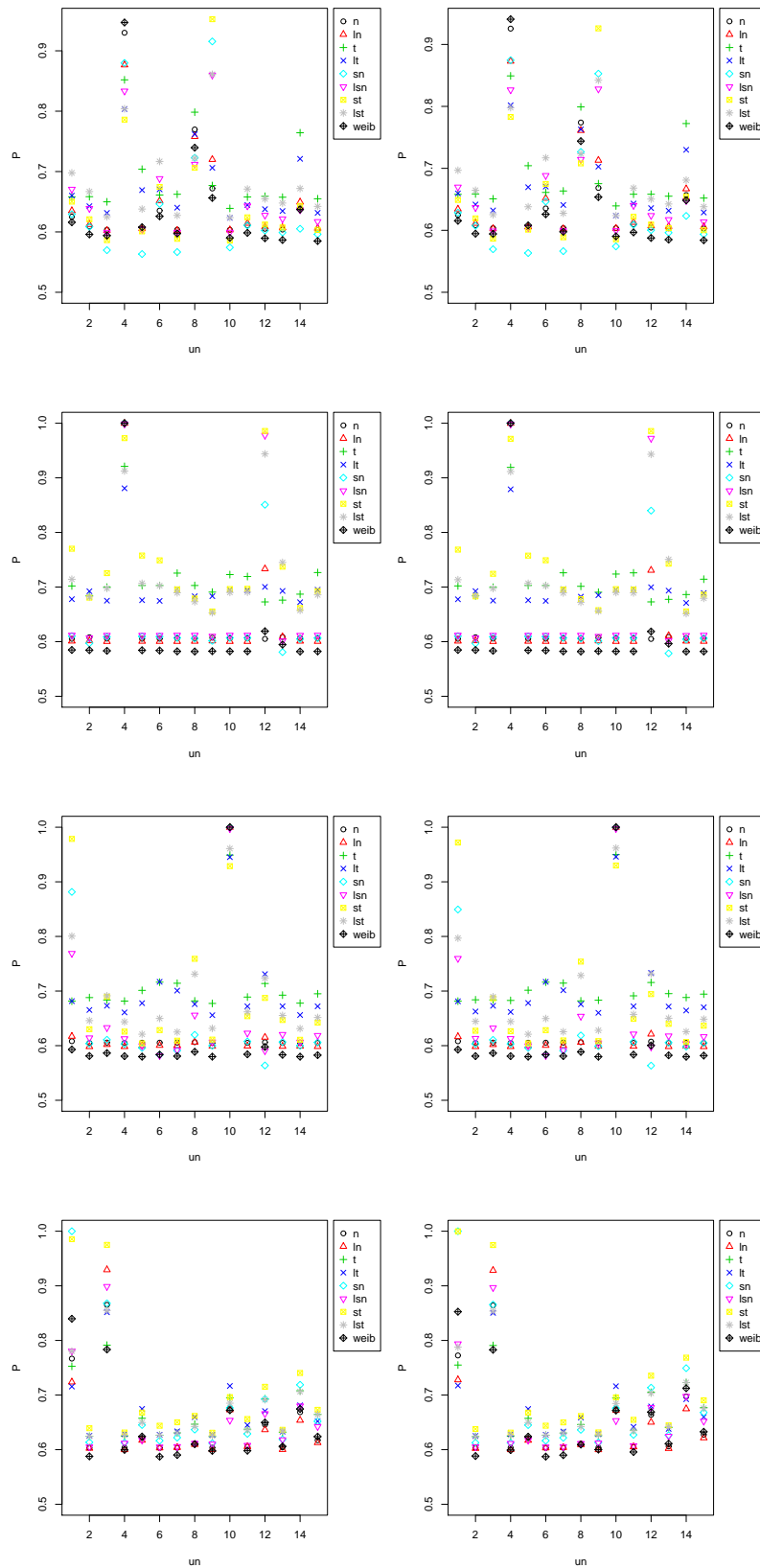


Figura 4.22:  $p_i$  por tempo de falha  $t_i$  (esquerda) e por pseudotempo de falha  $\hat{t}_i$  (direita), nos cenários B-SN (primeira linha), B-ST (segunda linha), B-LST (terceira linha) e B-W (quarta linha).



Num contexto de valores atípicos, ou presença de assimetria, as famílias de distribuições SMSN e SMLSN podem desempenhar um papel importante para inferir sobre os tempos de falha, uma vez que acomodam distribuições assimétricas e de caudas pesadas. Neste mesmo contexto, como verificado na análise dos cenários SN, ST e LST, observamos que as distribuições tradicionalmente utilizadas (N, LN e W) para modelar o efeito aleatório do modelo linear não são flexíveis o suficiente para capturar adequadamente a distribuição do tempo até a falha. Verificamos também que nos cenários ST e LST, os modelos SN e LSN foram bastante afetados pela presença de dados atípicos, assim como os modelos N, LN e W. Para cenário B-W, os modelos tradicionais N, LN e W realizaram boa inferência sobre os tempos de falha, já os modelos SN, LSN, T, LT, ST, e LST também realizaram boa inferência, porém um pouco pior que os modelos tradicionais. Portanto, num caso de má especificação dos efeitos aleatórios, o impacto negativo é menor ao assumirmos distribuições com caudas pesadas e assimétricas.

Sobre o cálculo das quantidades de interesse, os percentis geralmente são os mais utilizados para descrever o tempo até a falha. A análise dos cenários simulados procurou entender o comportamento das estimativas dos percentis e observou que os de ordem 5% e 95% geralmente são bem estimados tanto no método proposto por Oliveira (2011) quanto no de Hamada (2005), todavia o método de Hamada (2005) enfrenta dificuldades computacionais no caso dos modelos T, LT, ST e LST (para determinados valores de parâmetros e tempos de falha as funções de cálculo de percentis do software R podem não retornar as quantidades requeridas). Quanto aos percentis 1% e 99% observou-se que para ambos os métodos (especialmente o de Hamada (2005)) as estimativas podem ficar um pouco comprometidas principalmente nas distribuições de caudas pesadas. Por conta dos problemas interpretativos, inconsistências teóricas do método de Hamada (2005) para estimação de percentis, e pelos entraves e tempo computacionais, sugere-se a utilização da amostra da distribuição preditiva *a posteriori* de  $T_{n+1}$  Oliveira (2011) para estimar as quantidades de interesse.

Com respeito aos critérios de seleção baseados em qualidade de ajuste e métodos de avaliação de robustez, pode-se dizer que nenhum deles é útil se o foco da avaliação é direcionado às medidas de degradação por perfil. Os demais critérios baseados na degradação parecem ser úteis para selecionar o melhor modelo quando os dados possuem características de caudas pesadas. Contudo, fazemos uma ressalva sobre este aparente melhor desempenho, pois LPML, DIC e WAIC são calculados sob a distribuição de  $y_{ij}|\beta_i$ , e todos os modelos são capazes de estimar bem os efeitos aleatórios correspondentes a cada perfil de degradação. Observe pelas Figuras 4.23 e 4.26, que as estimativas dos efeitos aleatórios são praticamente idênticas, naturalmente um critério de seleção baseado na degradação irá fornecer estatísticas e conclusões muito parecidas, conseqüentemente terá dificuldade em permitir organizar os modelos entre melhores e piores. Noutras palavras, quando observamos as degradações, ou os perfis individualmente, estes critérios não conseguem avaliar se o modelo ajustado captura bem a característica relacionada à diferença entre perfis.

Se o foco é a distribuição do tempo até a falha todos os critérios auxiliam na escolha do melhor modelo em contexto de caudas leves para a distribuição dos efeitos aleatórios. Maior atenção deve ser dada aos gráficos baseados na fda e fdp preditiva *a posteriori* dos tempos de falha para uma unidade futura (os quais passam apenas por um crivo visual), e para os critérios LPML, DIC, que ajudam a indicar o melhor ajuste. O teste de Kolmogorov-Smirnov e o Fator de Bayes (baseado nos tempos e pseudo tempos de falha) apresentaram resultados razoáveis e apontaram na direção dos modelos com a mesma característica dos cenários simulados. A capacidade preditiva pode ser avaliada pela estatística  $CPO_i$  e a robustez pela divergência de Kullback-Leibler, em que ambos podem auxiliar na escolha do melhor modelo para os dados, porém não são muito conclusivos. Como os tempos de falha não são observáveis em situações práticas, concluímos que substituí-los pelos pseudo tempos de falha não compromete a seleção dos modelos (mesma estratégia utilizada por Peng e Tseng, 2013).

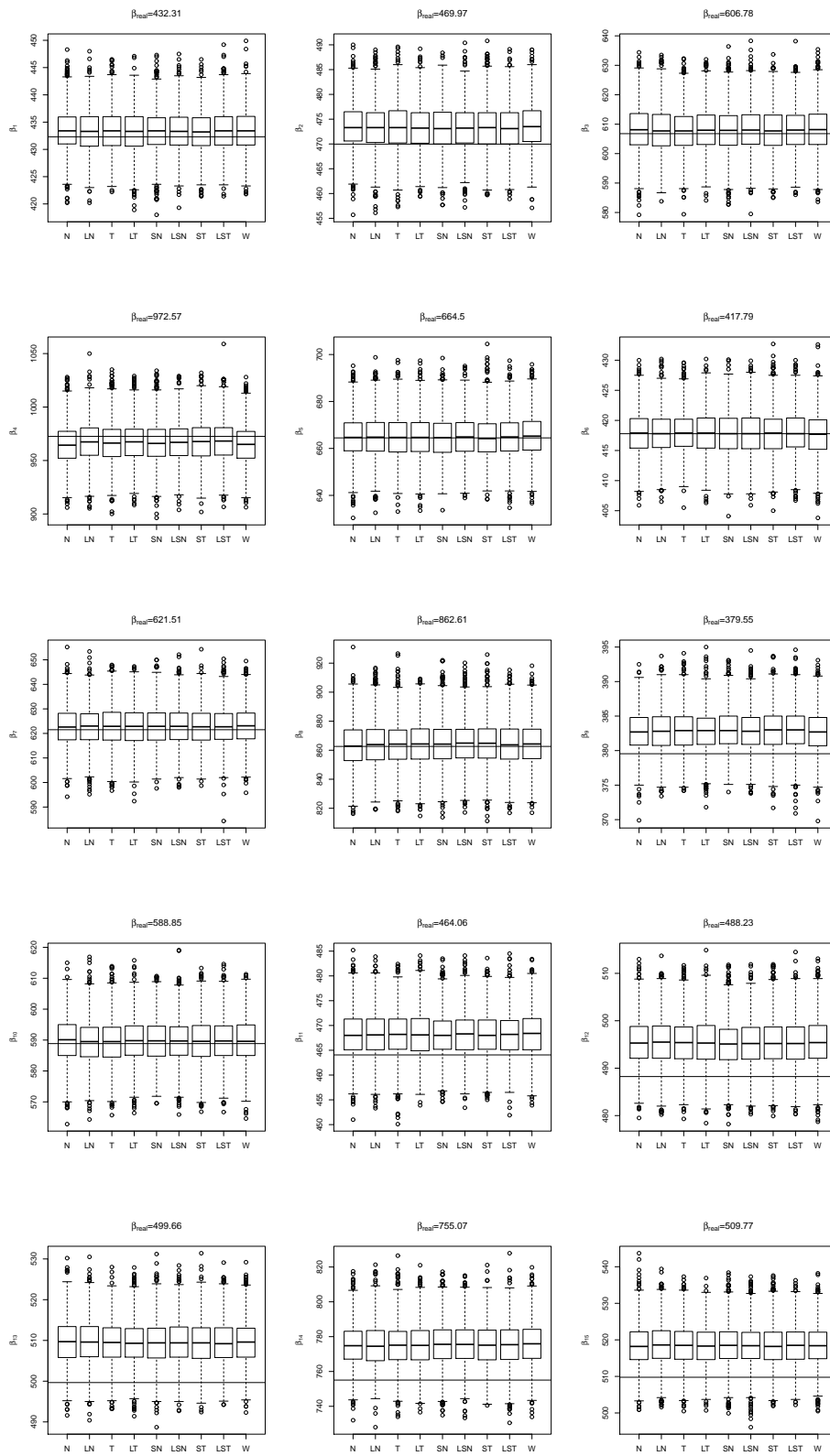


Figura 4.23: Distribuição *a posteriori*  $\beta_i|\mathbf{y}$  sob todos os modelos, cenário B-SN. As linhas pretas horizontais indicam os valores reais  $\beta_i$

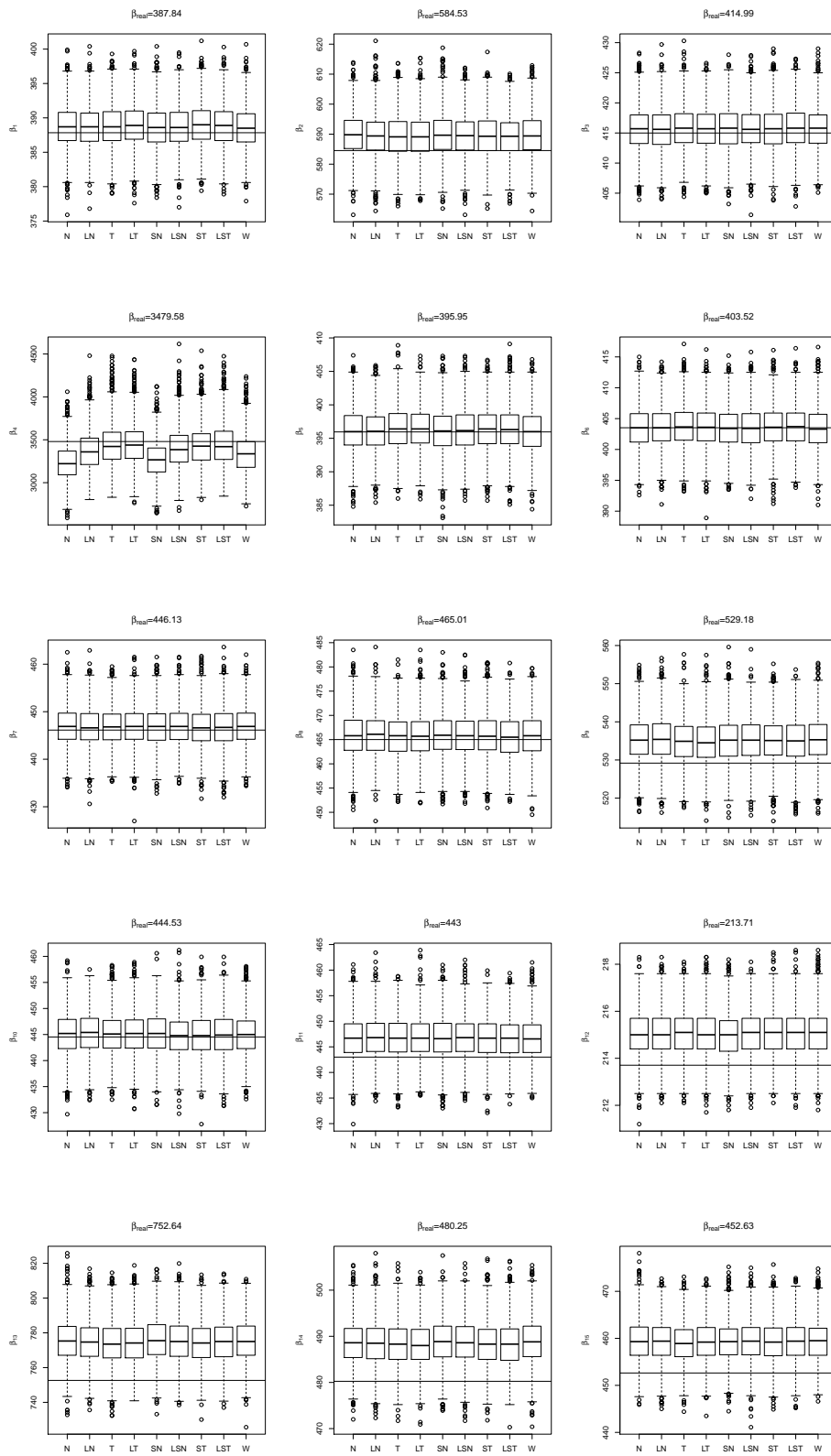


Figura 4.24: Distribuição *a posteriori*  $\beta_i|\mathbf{y}$  sob todos os modelos, cenário B-ST. As linhas pretas horizontais indicam os valores reais  $\beta_i$

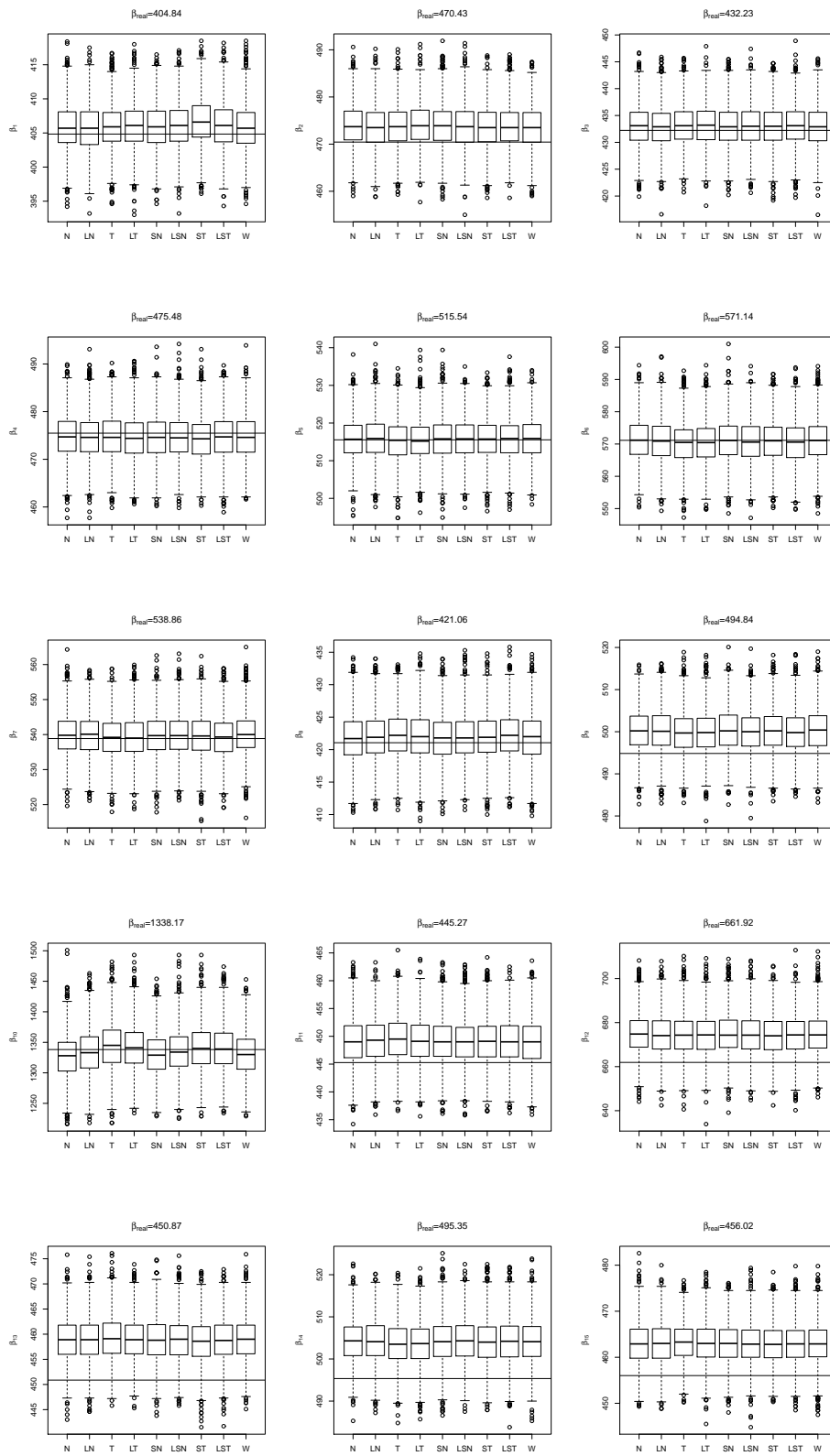


Figura 4.25: Distribuição *a posteriori*  $\beta_i|\mathbf{y}$  sob todos os modelos, cenário B-LST. As linhas pretas horizontais indicam os valores reais  $\beta_i$

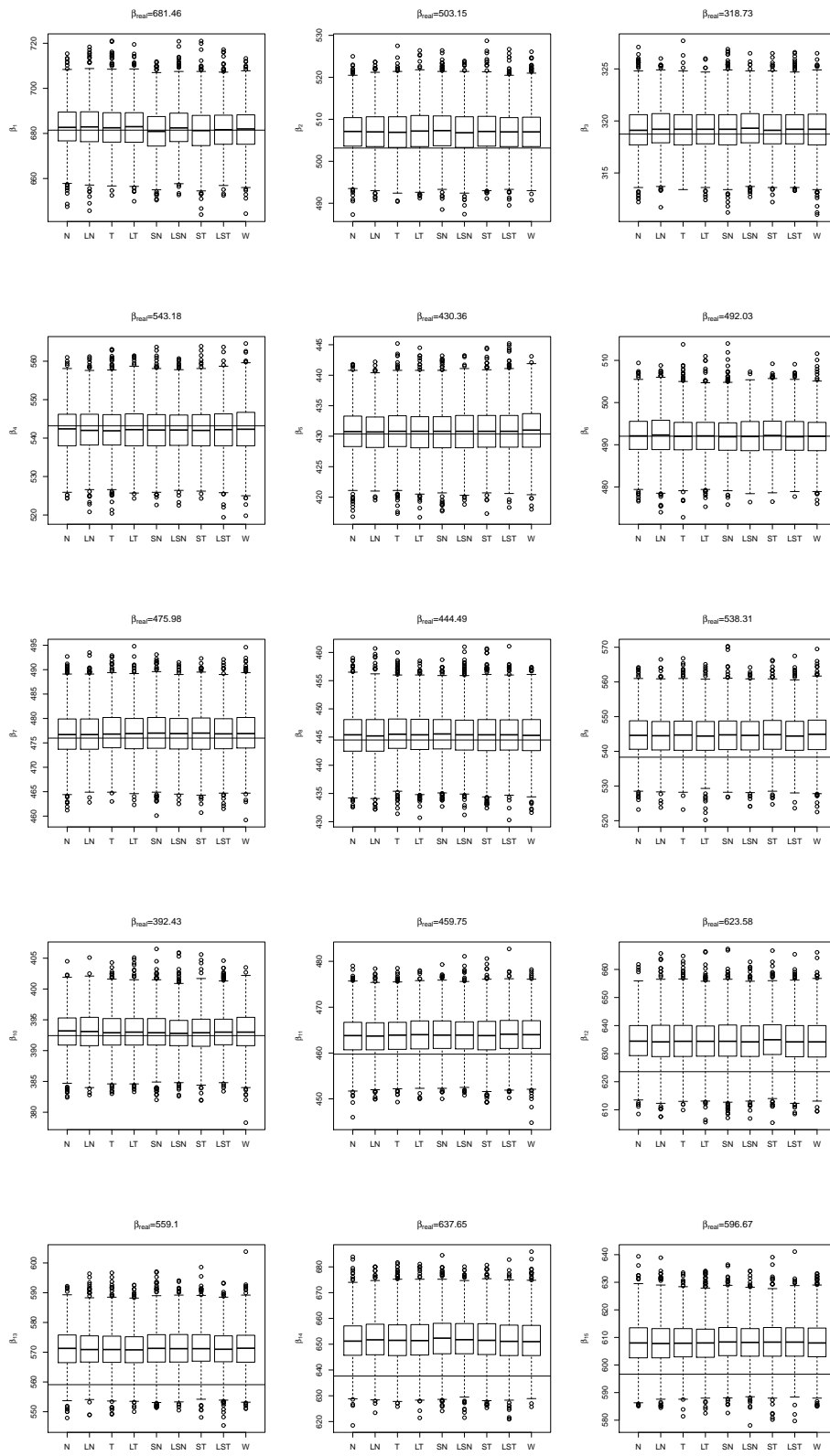


Figura 4.26: Distribuição *a posteriori*  $\beta_i|\mathbf{y}$  sob todos os modelos, cenário B-W. As linhas pretas horizontais indicam os valores reais  $\beta_i$

É importante observar que estes critérios de seleção e avaliação de robustez baseados nos tempos de falha, se utilizada a proposta de Hamada (2005), só serão aplicáveis caso o método analítico forneça uma forma fechada para a distribuição dos tempos até a falha. Quando isso não é possível somente o método proposto por Oliveira (2011) será capaz de fornecer subsídios para selecionar o melhor modelo através do teste de Kolmogorov-Smirnov, o qual não é muito conclusivo e por este motivo deve ser acompanhado de uma análise crítica da fda e fdp da distribuição preditiva *a posteriori* de  $T_{n+1}$ , assim como da avaliação de LPML e DIC. Isso se deve ao fato de que método baseado na distribuição preditiva *a posteriori* de  $T_{n+1}$  é livre em qualquer forma funcional que se estabeleça para o perfil de degradação, independente se é verificado o fechamento entre distribuição dos efeitos aleatórios do modelo de degradação e dos tempos de falha.

#### 4.5 Dados reais: modelos de degradação

Nesta seção os modelos desenvolvidos na Seção 4.1 foram considerados para a modelagem de degradação em dois estudos de caso. O primeiro estudo de caso diz respeito à degradação de rodas de trem (nomeado Dados de Rodas de Trem). Estes dados foram inicialmente analisados por Freitas *et al.* (2009) e, posteriormente, analisado em Freitas *et al.* (2010), em que neste último artigo uma das análises contou com a abordagem bayesiana conforme o método proposto por Hamada (2005). O segundo estudo de caso é bastante conhecido da literatura, trata-se dos dados de degradação de emissores de laser (Meeker e Escobar, 1998) nomeados neste texto por Dados de Laser.

Em ambos os estudos assumimos o modelo de degradação dado na expressão (4.12) e consideramos as distribuições N, LN, T, LT, SN, LSN, ST, LST e W para descrever o comportamento dos efeitos aleatórios. As suposições sobre a degradação e as distribuições *a priori* para os objetos aleatórios, foram as mesmas assumidas na Seção 4.4. A saber, assumimos que as distribuições *a priori* dos parâmetros pertencem às famílias mencionadas nas Proposições 8 e 9 (também  $\beta_i \stackrel{iid}{\sim} W(\kappa_\beta, \varsigma_\beta)$ ), em que atribuímos para

os hiperparâmetros distribuições *a priori* que refletem total desconhecimento sobre os parâmetros do modelo de degradação conforme as expressões em (4.43). A distribuição *a priori* para  $\nu$  é dada como em Fonseca *et al.* (2008) conforme equação (4.20).

A análise dos dados em todos os casos foi feita utilizando-se os softwares WinBUGS 1.4 (Spiegelhalter *et al.*, 2000) e R 3.1.0 (R Development Core Team, 2005). A fim de garantirmos a convergência das cadeias, também minimizar a autocorrelação das amostras obtidas, foram realizadas 220.000.000 iterações, com um período de aquecimento (*burn-in*) de 100.000.000, e saltos de 60.000 para a amostragem. O resultado foi uma amostra *a posteriori* final de tamanho  $s = 2.000$  para cada parâmetro ( $\psi|\mathbf{y}$ ). Estas especificações garantiram a convergência das cadeias em todos os nove modelos ajustados para todos os bancos de dados analisados aqui, a saber, ajustes pelos modelos de degradação Normal (N), Log-normal (LN), t-Student (T), Log-t-Student (LT), Skew-normal (SN), Log-skew-normal (LSN), Skew-t (ST), Log-skew-t (LST) e Weibull (W), cujos nomes estão relacionados com a distribuição adotada para os efeitos aleatórios. As inferências realizadas a partir das amostras *a posteriori* obtidas para os objetos aleatórios serão apresentados nas próximas seções.

A estatística  $CPO_i$  e a medida de divergência de Kullback Leibler para os pseudotempos de falha não forneceram resultados conclusivos para auxiliar na seleção do melhor modelo nestes estudos de caso, por esta razão serão omitidos.

#### 4.5.1 Estudo de caso 1: dados de desgaste de rodas de trem

Os dados de desgaste de rodas de trens analisados anteriormente em Freitas *et al.* (2009) serão analisados usando os modelos de degradação robustos aqui introduzidos. Estes dados foram obtidos num estudo conduzido por uma empresa do setor ferroviário e inclui, entre outras informações, as medidas de diâmetros de rodas tomadas em 13 igualmente espaçados tempos de inspeção:  $t_0 = 0km, t_1 = 50.000Km, \dots, t_{13} = 600.000Km$ . Estas medidas foram registradas para 14 trens, cada um composto por quatro carros



(vagões), sendo um motor (MA) e três reboques (R1,R2,R3). Os carros reboque não possuem tração própria, por isso consideramos apenas a degradação no carro motor. Cada vagão em um dado trem, possui dois truques, cada truque com dois eixos, cada eixo com duas rodas, totalizando assim oito rodas por carro (vagão). A Figura 4.27 apresenta o croqui para o vagão motor.

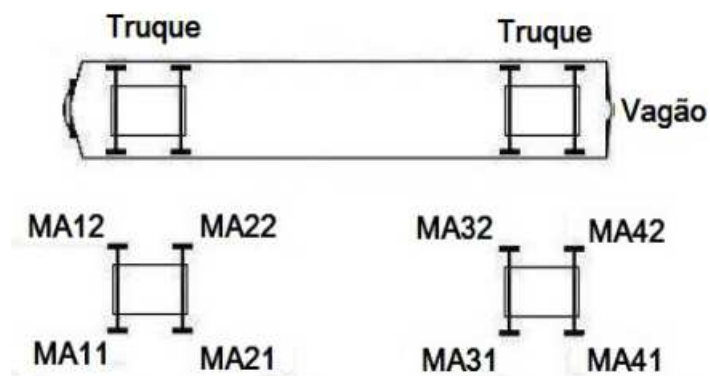
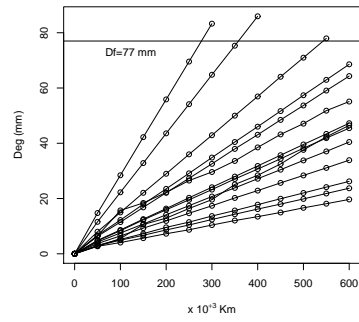
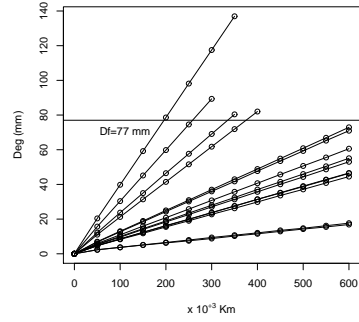


Figura 4.27: Croqui - vista principal inferior de um vagão motor.

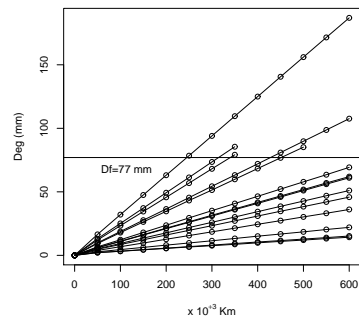
Os vagões motor foram selecionados porque as características do modo de operação aceleram o desgaste das rodas destes carros. O objetivo do estudo é a caracterização da função fração de falha das rodas. As rodas estudadas são de aço forjado e laminado e cada uma possui um diâmetro inicial de  $966\text{ mm}$ . A equipe de engenharia determinou que a redução crítica de diâmetro ocorre quando este atinge  $889\text{ mm}$ , ocasião em que a roda deve ser substituída. Ao final obtiveram-se 14 medidas de diâmetro para cada uma das oito posições de rodas do vagão motor (MA11, MA12, MA21, MA22, MA31, MA32, MA41, MA42), as quais são exibidas na figura 4.28.



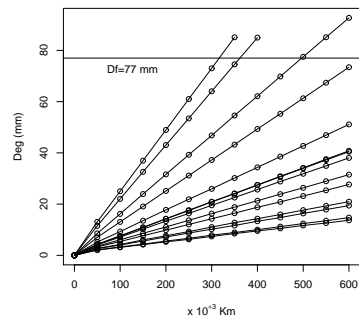
(a) MA11



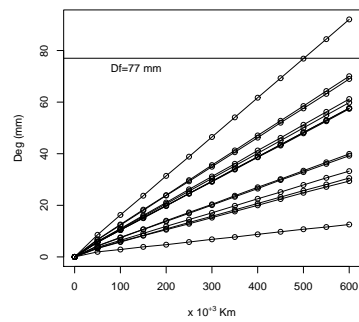
(b) MA12



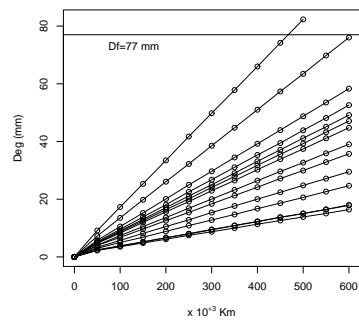
(c) MA21



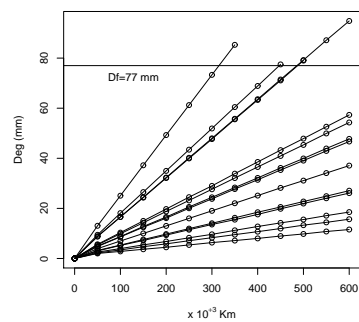
(d) MA22



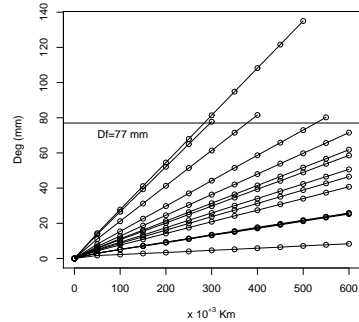
(e) MA31



(f) MA32



(g) MA41



(h) MA42

Figura 4.28: Perfis de desgaste

Freitas *et al.* (2009) analisam os dados das quatorze (14) rodas MA11 apenas e para tal utilizaram como informação o desgaste das rodas, que é dado por

$$Y_{ij} = 966\text{mm} - \text{Diam}_{ij}, \quad (i = 1, 2, \dots, 14; j = 1, 2, \dots, 13),$$

em que  $Y_{ij}$  é a variável aleatória que define o desgaste da  $i$ -ésima roda no  $j$ -ésimo tempo (quilometragem) de medição.

O limiar crítico de degradação  $D_f$  (“Failure Threshold”) é dado por

$$D_f = 966\text{mm} - \text{Diam}_{\text{crítico}} = 966\text{mm} - 889\text{mm} = 77\text{mm},$$

em que considera-se que a “falha” de uma determinada roda ocorreu quando o desgaste atinge este limiar crítico. Observe da Figura 4.28 que os perfis das rodas em todas as posições são lineares e apresentam características bem diferentes. Por exemplo, para rodas da posição MA31 notamos que, em dois trens, o perfil de degradação se distancia dos demais, indicando que uma distribuição com caudas pesadas pode descrever melhor a incerteza sobre a velocidade de degradação. Isto ocorre também para as rodas MA12 e MA41. Note ainda que os perfis de degradação nas rodas MA12 e MA41 indicam que a velocidade de degradação aparenta vir de uma mistura de distribuições.

A análise feita em Freitas *et al.* (2010) utilizou métodos clássicos e bayesianos para estimação dos parâmetros do modelo de degradação, em particular o método de máxima verossimilhança e amostragem via métodos MCMC. Para obterem a distribuição do tempo até falha das rodas MA11 via inferência clássica utilizaram o método analítico, aproximado e de dois estágios. Já para a abordagem bayesiana utilizaram o método analítico no contexto bayesiano como em Hamada (2005). No método analítico e aproximado foram consideradas as distribuições lognormal e weibull.

Neste trabalho serão analisados os dados das rodas MA31, MA11 e MA12, em que utilizaremos a abordagem bayesiana. Os métodos empregados serão aqueles utilizados por Hamada (2005) e Oliveira (2011) combinados com os procedimentos inferenciais para distribuições de caudas pesadas apresentados nesta tese.

**Análise e Discussão**

As fda e fdp preditiva *a posteriori* para um perfil futuro via Hamada (2005) e Oliveira (2011) encontram-se nas Figuras 4.29 a 4.31. A distribuição acumulada empírica foi baseada nas estimativas procedentes dos pseudo tempos de falha (Meeker e Escobar, 1998).

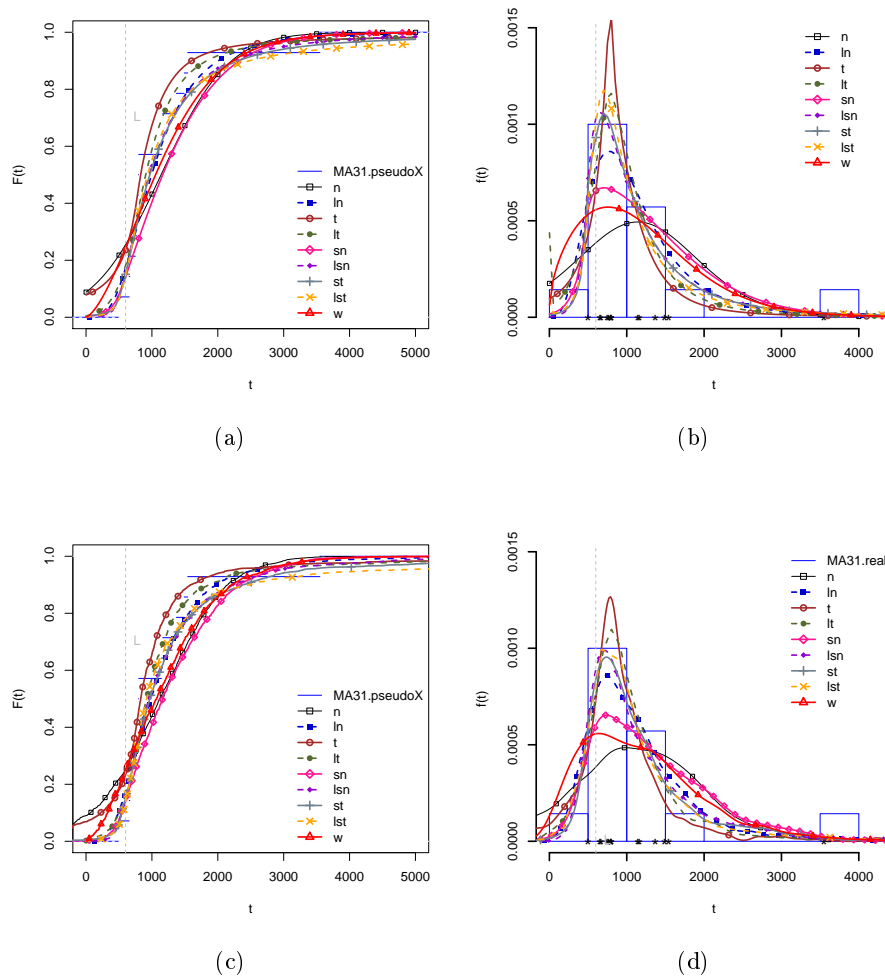


Figura 4.29: Função distribuição acumulada (esquerda) e densidade de probabilidade (direita) *a posteriori* dos tempos de falha para uma unidade futura, pelos métodos de Hamada (2005) (acima) e proposta pela preditiva *a posteriori* (embaixo), para todos os modelos ajustados aos dados da roda de trem MA31

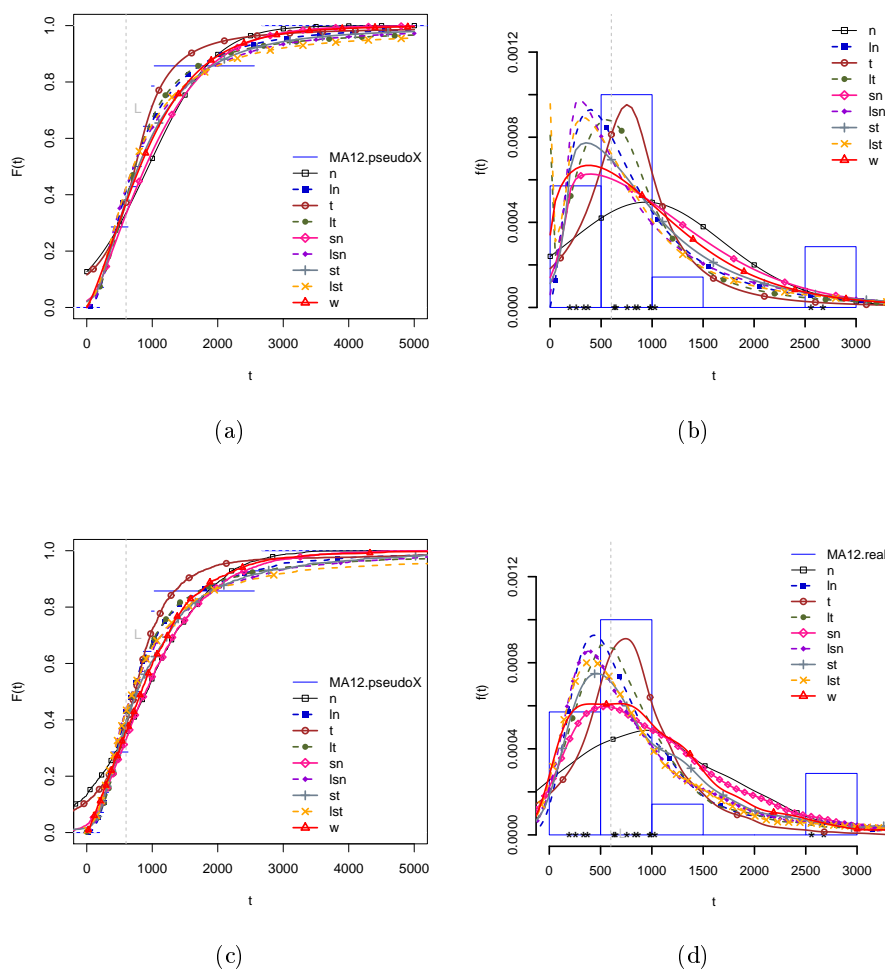


Figura 4.30: Função distribuição acumulada (esquerda) e densidade de probabilidade (direita) *a posteriori* dos tempos de falha para uma unidade futura, pelos métodos de Hamada (2005) (acima) e proposta pela preditiva *a posteriori* (embaixo), para todos os modelos ajustados aos dados da roda de trem MA12

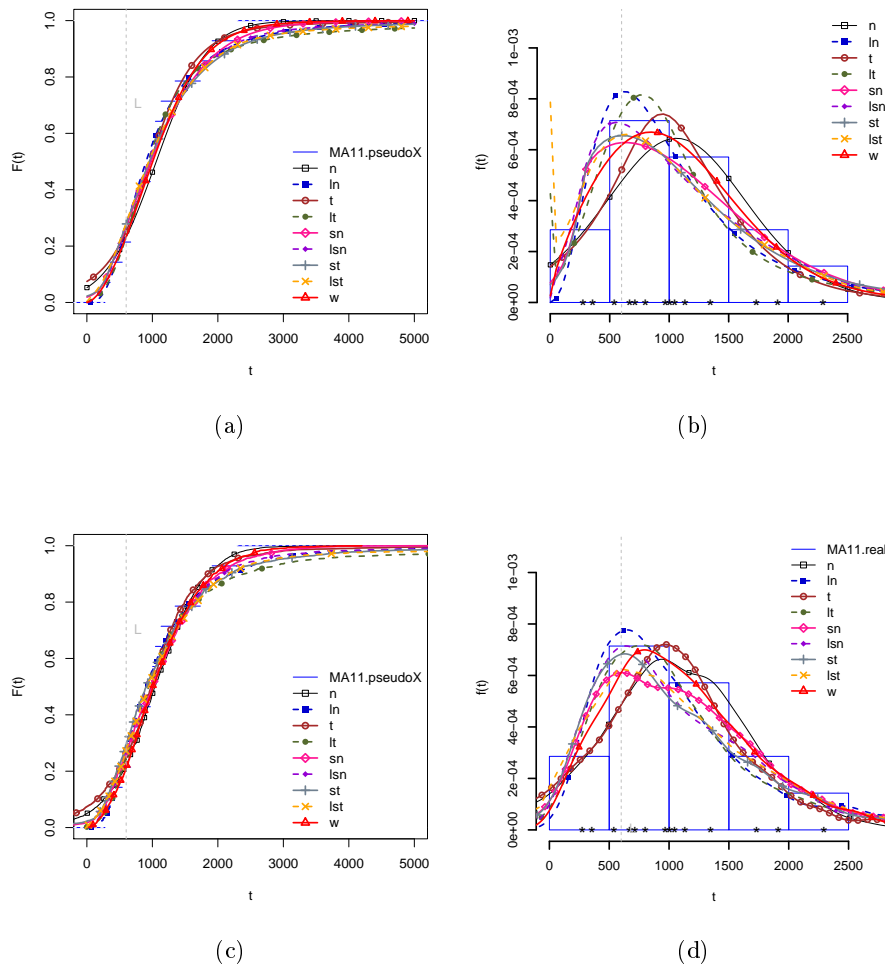


Figura 4.31: Função distribuição acumulada (esquerda) e densidade de probabilidade (direita) *a posteriori* dos tempos de falha para uma unidade futura, pelos métodos de Hamada (2005) (acima) e proposta pela preditiva a posteriori (embaixo), para todos os modelos ajustados aos dados da roda de trem MA11

Para as rodas da posição

- MA31, percebemos pela Figura 4.29 que os modelos W, N e T não parecem adequados para modelar os tempos de falha. Observe pelas fdp estimadas que as caudas desses ajustes não conseguem acompanhar bem as caudas do histograma e nem a região onde se concentram os pseudo tempos de falha (asteriscos na base do gráfico). Os modelos ST, LSN, LST, LT e LN identificam bem a posição da moda relativa e as caudas da distribuição empírica dos pseudo tempos de falha.
- MA12, a Figura 4.30 os modelos LT, LN, LST e LSN são os que melhor descrevem o comportamento das caudas e da posição da moda relativa considerando-se o histograma dos pseudo tempos de falha. As versões não logarítmicas destes modelos tendem a colocar massa de probabilidade em valores negativos do desgaste das rodas, portanto não são interessantes para modelar os tempos de falha. Já o modelo W coloca muita massa de probabilidade nas proximidades de zero e superestima a cauda superior dos pseudo tempos de falha.
- MA11, observamos na Figura 4.31 que o modelo W se destaca entre os demais, porém todos os modelos conseguem se ajustar bem à distribuição empírica dos pseudo tempos de falha.

Nas Tabelas C.13 a C.17 (Anexo C) apresentamos os critérios de seleção LPML, DIC, WAIC, Fator de Bayes (FB) e o teste de Kolmogorov Smirnov (KS), baseados em degradação (quando aplicável) e pseudo tempos de falha, para as rodas MA31, MA12 e MA11. A Tabela 4.8 resume os critérios de seleção das referidas tabelas e fornece os modelos ordenados do melhor (no topo e em negrito) para o pior (pé) em cada um dos bancos de dados das Rodas de Trem analisadas. A tabela 4.8

- sugere que LSN é o melhor ajuste para modelar os tempos de falha das rodas na posição MA31 (segundo os critérios LPML, DIC, WAIC e KS na proposta pela

distribuição preditiva) e LT (segundo FB e KS via Hamada (2005)), seguido pelos modelos LST (segundo LPML, DIC, FB e KS via Hamada (2005)), ST (WAIC) e T (proposta pela distribuição preditiva), quando baseamo-nos nos pseudo tempos de falha. Note que a indicação do modelo T como segundo melhor, na estatística KS via proposta pela distribuição preditiva, se dá porque na amostra não existem pseudo tempos na cauda esquerda da fdp da Figura 4.29, portanto é preciso cautela ao avaliarmos a escolha do modelo segundo este critério. Apesar deste fato, a escolha do melhor modelo para todos estes critérios foi condizente com os modelos visualmente selecionados ao avaliarmos a Figura 4.29.

Se observarmos os critérios baseados nas medidas de degradação apenas o FB permite escolher um modelo coerente com a análise gráfica da Figura 4.29, entretanto por este mesmo critério o segundo melhor modelo é o N, um modelo que pela análise gráfica distoa bastante da distribuição empírica dos pseudo tempos de falha. Os demais critérios selecionam como melhor ajuste o do modelo T (LPML, DIC e WAIC) e N (LPML). Ou seja, os critérios baseados na degradação contradizem a análise da Figura 4.29.

- aponta os melhores modelos como LN (segundo LPML, DIC e FB), SN (critério WAIC) e T (teste KS) para as rodas na posição MA12; seguidos pelo modelos LSN (segundo LPML e DIC), ST (segundo WAIC), LT (teste KS) e W (FB), quando consideramos os pseudo tempos de falha para calcular os critérios de seleção. Note que FB indica o modelo W como segundo melhor ajuste, uma escolha incoerente com o que se percebe da Figura análise da Figura 4.30.



Tabela 4.8: Modelos ordenados do melhor (topo) para o pior (pé) segundo LPML, DIC, WAIC, Fator de Bayes (FB), Teste de Kolmogorov-Smirnov (KS), baseados na degradação (Deg) e pseudo tempos de falha (PTF), Rodas de Trem

Roda	Deg				PTF			FB		KS	
	LPML	$LPML_{perfil}$	DIC	WAIC	LPML	DIC	WAIC	Deg	PTF	$Hamada_{pseudo}$	$Prop_{pseudo}$
MA31	<b>T</b>	<b>N</b>	<b>T</b>	<b>T</b>	<b>LSN</b>	<b>LSN</b>	<b>LSN</b>	<b>LST</b>	<b>LT</b>	<b>LT</b>	<b>LSN</b>
	LT	T	N	LT	LST	LST	ST	N	LST	LST	T
	W	ST	LSN	LSN	LN	ST	LST	SN	LN	T	LT
	N	LST	SN	N	ST	LN	SN	W	ST	LSN	LST
	LSN	LT	W	W	LT	LT	LN	LSN	LSN	W	W
	LST	LN	LT	LST	SN	SN	W	ST	T	ST	LN
	SN	SN	LST	SN	T	T	LT	T	SN	LN	ST
	LN	W	ST	LN	W	W	N	LT	W	N	N
	ST	LSN	LN	ST	N	N	T	LN	N	SN	SN
MA12	<b>T</b>	<b>N</b>	<b>N</b>	<b>N</b>	<b>LN</b>	<b>LN</b>	<b>SN</b>	<b>N</b>	<b>LN</b>	<b>T</b>	<b>T</b>
	N	LN	LN	LN	LSN	LSN	ST	LSN	W	LT	LT
	LN	LT	T	LT	LT	LT	W	SN	LSN	LN	LN
	ST	W	LT	T	ST	ST	LSN	LT	LT	LST	LST
	W	T	W	ST	W	W	LN	ST	ST	LSN	LSN
	LT	LST	SN	W	SN	SN	LST	T	LST	ST	ST
	LSN	ST	LSN	SN	LST	LST	N	W	SN	W	W
	SN	LSN	ST	LSN	T	T	LT	LN	T	SN	SN
	LST	SN	LST	LST	N	N	T	LST	N	N	N
MA11	<b>LT</b>	<b>LT</b>	<b>LT</b>	<b>LT</b>	<b>W</b>	<b>W</b>	<b>SN</b>	<b>LT</b>	<b>W</b>	<b>LT</b>	<b>LN</b>
	ST	W	LN	LN	SN	SN	LSN	ST	LN	T	LT
	W	ST	ST	ST	LN	LN	W	T	SN	LST	T
	LSN	LST	W	W	LSN	LSN	ST	SN	N	LSN	LST
	SN	T	LST	LST	LT	ST	LST	W	LT	LN	LSN
	LN	LSN	N	LSN	ST	LT	N	N	LSN	ST	ST
	LST	LN	LSN	N	N	LST	LN	LN	T	SN	W
	T	N	SN	SN	LST	N	LT	LST	LST	W	SN
	N	SN	T	T	T	T	T	LSN	ST	N	N

Sob as medidas de degradação, os critérios de seleção e ajuste apontam os piores modelos possíveis como os melhores modelos. A saber, o modelo N (segundo  $LPML_{perfil}$ , DIC, WAIC e WAIC) e T (segundo LPML). Note pela Figura 4.30 que ambos os modelos colocam grande massa probabilística nas proximidades dos tempos zero. Novamente, os critérios baseados na degradação fornecem escolhas incoerentes com a análise gráfica.

- indica W (segundo LPML, DIC e FB), SN (segundo WAIC), LT (teste KS via Hamada (2005)) e LN (teste KS via proposta pela distribuição preditiva) como o melhores modelos para as rodas na posição MA11, quando consideramos os pseudo tempos de falha para calcular os critérios de seleção. Pela Figura 4.30 os modelos W, LT e LN aparentam ser bem razoáveis, o que indica que os critérios

não levam à escolhas contraditórias.

Se são consideradas as medidas de degradação, todos os critérios selecionam o modelo LT.

Os critérios de seleção baseados nos pseudo tempos de falha forneceram resultados condizentes com a análise das Figuras 4.29 a 4.31, em que LPML e DIC são especialmente coerentes. Já os critérios baseados na degradação levaram a escolhas contraditórias com a análise gráfica, o que já era esperado por conta das conclusões obtidas na Seção 4.4.

A Tabela 4.9 fornece os percentis de ordem 1%, 5%, 95% e 99%, *a posteriori*, dos tempos de falha para uma roda futura da população de rodas MA31. Sob LSN, selecionado como o melhor modelo, concluímos que 1% das rodas da população degradam-se além de  $77mm$  antes de completarem  $338,4 \times 10^3 Km$  rodados, 5% antes de  $480,95 \times 10^3 Km$  rodados, 95% antes de  $2870,6 \times 10^3 Km$  rodados e 99% antes de  $5257,3 \times 10^3 Km$  rodados, considerando-se o método proposto por Oliveira (2011) com uma nova unidade inclusa na verossimilhança.

Já para a roda MA12, na Tabela 4.10 verificamos que sob o modelo LN, selecionado como o melhor modelo pelos critérios LPML e DIC, que 1% das rodas da população degrada-se além de  $77mm$  antes de completar  $88 \times 10^3 Km$  rodados, 5% antes de  $181,02 \times 10^3 Km$  rodados, 95% antes de  $2919,15 \times 10^3 Km$  rodados e 99% antes de  $6350,02 \times 10^3 Km$  rodados, considerando-se o método proposto por Oliveira (2011). O modelo SN, selecionado como melhor modelo pelo critério WAIC, forneceu uma estimativa negativa para o percentil 1% e por esse motivo seus resultados não serão considerados.

Tabela 4.9: Percentis da distribuição *a posteriori* dos tempos de falha para a Roda MA31

Ajuste	$t_p\%$	Pseudo tempos $\times 10^3$		Oliveira (2011)		Hamada (2005)		
		Empírico		Pred.	Post.	Média	$HPD_{INF}$	$HPD_{SUP}$
N	$t_1$	519,91		-1059,70		-759,17	-1750,21	6,93
	$t_5$	600,99		-307,75		-203,44	-978,74	427,55
	$t_{95}$	2245,13		2513,36		2479,16	1816,39	3171,62
	$t_{99}$	3288,65		3119,49		3034,89	2243,44	3961,98
LN	$t_1$	519,91		256,14		300,82	134,83	444,60
	$t_5$	600,99		385,67		425,85	234,80	590,78
	$t_{95}$	2245,13		2559,36		2476,50	1529,96	3819,78
	$t_{99}$	3288,65		4256,20		3643,96	2001,65	6259,28
T	$t_1$	519,91		-5255,78		-1,59E+09	-431057,30	256,01
	$t_5$	600,99		-262,66		-26301,31	-8469,30	528,54
	$t_{95}$	2245,13		2177,37		28022,94	1088,54	10119,96
	$t_{99}$	3288,65		10463,99		1,588E+09	1384,09	432556,34
LT	$t_1$	519,91		86,39		174,30	0,00	368,83
	$t_5$	600,99		392,22		356,79	70,81	605,46
	$t_{95}$	2245,13		2551,90		7,997E+09	1190,51	8901,58
	$t_{99}$	3288,65		12483,47		$\infty$	1689,22	1,91E+09
SN	$t_1$	519,91		205,57		322,33	-136,19	567,49
	$t_5$	600,99		444,24		443,01	111,96	668,29
	$t_{95}$	2245,13		2798,76		2672,57	1825,98	3615,20
	$t_{99}$	3288,65		3475,88		3386,37	2285,80	4645,77
LSN	$t_1$	519,91		<b>338,40</b>		<b>401,67</b>	193,39	558,95
	$t_5$	600,99		<b>480,95</b>		<b>486,97</b>	312,27	645,87
	$t_{95}$	2245,13		<b>2870,60</b>		<b>3095,60</b>	1522,59	5781,23
	$t_{99}$	3288,65		<b>5257,29</b>		<b>5591,74</b>	1919,72	12167,18
ST	$t_1$	519,91		202,17		271,37	-211,17	667,31
	$t_5$	600,99		483,28		445,96	148,51	687,16
	$t_{95}$	2245,13		3253,44		4462,31	1410,52	10014,22
	$t_{99}$	3288,65		11485,55		15586,02	1989,39	61985,79
LST	$t_1$	519,91		274,32		374,21	50,80	578,65
	$t_5$	600,99		483,71		483,20	233,02	668,13
	$t_{95}$	2245,13		4317,54		2,903E+59	1281,43	38495,61
	$t_{99}$	3288,65		39868,21		8,33E+152	1635,58	1,012E+09
W	$t_1$	519,91		50,96		93,10	10,46	194,85
	$t_5$	600,99		207,03		232,33	55,64	411,54
	$t_{95}$	2245,13		2716,37		2615,16	1761,51	3646,06
	$t_{99}$	3288,65		3680,62		3428,02	2188,41	5086,18

Tabela 4.10: Percentis da distribuição *a posteriori* dos tempos de falha para a Roda MA12

Ajuste	$t_p\%$	Pseudo tempos $\times 10^3$		Oliveira (2011)			Hamada (2005)		
		Empírico		Pred.	Post.	Média	$HPD_{INF}$	$HPD_{SUP}$	
N	$t_1$	204,23		-1352,94		-948,47	-1961,00	-218,58	
	$t_5$	236,13		-568,90		-394,09	-1157,04	215,83	
	$t_{95}$	2599,76		2406,37		2282,00	1630,98	3029,13	
	$t_{99}$	2660,99		3194,03		2836,37	2069,55	3813,51	
LN	$t_1$	204,23		<b>88,00</b>		<b>119,48</b>	31,54	211,29	
	$t_5$	236,13		<b>181,02</b>		<b>200,50</b>	73,80	321,66	
	$t_{95}$	2599,76		<b>2919,15</b>		<b>2939,11</b>	1381,18	5502,05	
	$t_{99}$	2660,99		<b>6350,02</b>		<b>5404,20</b>	2010,75	11716,97	
T	$t_1$	204,23		-4747,47		-172334,70	-53544,32	44,34	
	$t_5$	236,13		-626,61		-1589,32	-3733,68	370,74	
	$t_{95}$	2599,76		2014,59		3071,50	1076,06	5169,78	
	$t_{99}$	2660,99		7435,44		173816,87	1513,98	54881,50	
LT	$t_1$	204,23		28,64		64,62	0,00	169,08	
	$t_5$	236,13		160,30		171,12	0,00	316,60	
	$t_{95}$	2599,76		3407,52		2,266E+16	1179,20	22010,92	
	$t_{99}$	2660,99		25225,20		$\infty$	1764,48	7,25E+10	
SN	$t_1$	204,23		-175,51		-27,57	-572,02	254,00	
	$t_5$	236,13		91,44		121,46	-281,48	327,11	
	$t_{95}$	2599,76		2658,89		2496,04	1673,42	3577,87	
	$t_{99}$	2660,99		3739,13		3250,15	2136,69	4691,95	
LSN	$t_1$	204,23		56,27		114,58	4,88	209,45	
	$t_5$	236,13		161,90		183,98	35,94	303,47	
	$t_{95}$	2599,76		3550,55		4441,23	1247,32	10672,96	
	$t_{99}$	2660,99		9585,58		13848,93	1994,95	36938,65	
ST	$t_1$	204,23		-219,93		-155,74	-871,75	248,82	
	$t_5$	236,13		154,23		87,12	-285,94	361,63	
	$t_{95}$	2599,76		3361,28		4292,42	1173,69	7628,53	
	$t_{99}$	2660,99		8752,05		10881,66	1611,28	33890,23	
LST	$t_1$	204,23		10,96		85,01	0,00	199,91	
	$t_5$	236,13		129,21		159,37	0,00	296,00	
	$t_{95}$	2599,76		4626,85		2,25E+201	1047,60	35950,47	
	$t_{99}$	2660,99		28121,02		4,54E+201	1403,20	9261246	
W	$t_1$	204,23		26,37		43,06	0,82	104,38	
	$t_5$	236,13		98,28		129,45	23,68	263,61	
	$t_{95}$	2599,76		2478,67		2525,13	1455,11	3976,36	
	$t_{99}$	2660,99		4096,32		3543,77	2018,09	6108,97	

Na Tabela 4.11 observamos que, para a roda MA11 sob o modelo W, selecionado como o melhor modelo pelos critérios LPML e DIC, que 1% das rodas da população degrada-se além de  $77mm$  antes de completar  $107,11 \times 10^3 Km$  rodados, 5% antes de  $257,25 \times 10^3 Km$  rodados, 95% antes de  $2260,30 \times 10^3 Km$  rodados e 99% antes de  $2927,77 \times 10^3 Km$  rodados, considerando-se o método proposto por Oliveira (2011) com uma nova unidade inclusa na verossimilhança. O modelo SN, selecionado como melhor modelo pelo critério WAIC, forneceu uma estimativa negativa para o percentil 1%, portanto será desconsiderado para interpretação.

O método proposto por Hamada (2005) apresentou resultados similares (apesar da inconsistência teórica apontada na Seção 3.2.1), e seus intervalos HPD de 95% credibilidade são disponibilizados nas Tabelas 4.9 a Tabelas 4.11 apenas a título ilustrativo, uma vez que não possuem interpretabilidade que traga alguma informação útil para a tomada de decisão.

#### 4.5.2 Estudo de caso 2: corrente de operação de emissores de laser

Meeker e Escobar (1998) apresentaram dados de quinze (15) emissores de laser, acompanhados por um tempo de estudo de 4000 horas de operação. A cada 250 h foi registrado para cada emissor, o percentual de aumento na corrente de operação, calculado em relação à corrente nominal de operação, isto é, aquela do início do ensaio. Esta informação é importante visto que a luz dos lasers se degrada ao longo do tempo se a corrente de operação é mantida constante. Consequentemente, para que seja gerada uma saída de luz constante, a corrente de operação precisa ser corrigida ao longo do tempo de uso. A medida de degradação, neste caso, é o percentual de aumento na corrente de operação. A “falha” do laser é estabelecida para o momento no qual o percentual de aumento na corrente excede 10%, isto é, a partir do limiar  $D_f = 10\%$  considera-se que o laser falhou, embora ainda esteja em operação.

Tabela 4.11: Percentis da distribuição *a posteriori* dos tempos de falha para a Roda MA11

Ajuste	$t_{p\%}$	Pseudo tempos $\times 10^3$		Hamada (2005)		
		Empírico	Oliveira (2011) Pred. Post.	Média	$HPD_{INF}$	$HPD_{SUP}$
N	$t_1$	286,49	-410,42	-395,13	-1127,03	227,47
	$t_5$	328,33	5,39	31,12	-525,53	525,96
	$t_{95}$	2046,03	2103,87	2088,69	1596,05	2674,81
	$t_{99}$	2244,72	2596,78	2514,94	1893,37	3253,11
LN	$t_1$	286,49	168,75	212,62	77,29	348,50
	$t_5$	328,33	309,05	323,02	160,21	491,07
	$t_{95}$	2046,03	2827,94	2774,15	1500,85	4650,26
	$t_{99}$	2244,72	5060,56	4491,16	2005,81	8475,83
T	$t_1$	286,49	-1317,65	-12402,81	-9698,66	275,74
	$t_5$	328,33	-186,51	-396,19	-1519,21	512,18
	$t_{95}$	2046,03	2247,86	2360,95	1358,81	3465,67
	$t_{99}$	2244,72	3879,54	14367,57	1759,91	11471,23
LT	$t_1$	286,49	80,07	134,79	0,00	290,57
	$t_5$	328,33	261,95	291,56	36,43	494,44
	$t_{95}$	2046,03	3394,97	107017,01	1228,91	10958,46
	$t_{99}$	2244,72	15881,56	5,8587E+95	1750,19	11997706,51
SN	$t_1$	286,49	-176,91	-22,04	-793,08	401,66
	$t_5$	328,33	200,30	190,44	-238,85	514,06
	$t_{95}$	2046,03	2398,20	2385,74	1641,82	3444,27
	$t_{99}$	2244,72	3257,95	3039,86	2015,91	4423,64
LSN	$t_1$	286,49	116,08	164,41	11,54	300,41
	$t_5$	328,33	226,68	272,04	62,64	447,13
	$t_{95}$	2046,03	2635,98	3040,69	1590,82	5570,54
	$t_{99}$	2244,72	5109,00	6190,49	1838,64	11798,03
ST	$t_1$	286,49	-429,29	-195,51	-1496,44	389,20
	$t_5$	328,33	186,78	143,73	-425,63	578,46
	$t_{95}$	2046,03	2906,79	3289,95	1551,19	5563,53
	$t_{99}$	2244,72	6860,53	6966,01	1871,74	16476,47
LST	$t_1$	286,49	34,85	103,88	0,00	278,65
	$t_5$	328,33	188,80	220,05	0,00	428,02
	$t_{95}$	2046,03	2804,11	2,6407E+21	1277,41	10129,71
	$t_{99}$	2244,72	8447,44	2,657E+134	1754,67	146854,12
W	$t_1$	286,49	<b>107,11</b>	<b>122,10</b>	16,17	246,58
	$t_5$	328,33	<b>257,25</b>	<b>269,51</b>	93,45	468,33
	$t_{95}$	2046,03	<b>2260,30</b>	<b>2201,82</b>	1576,07	3030,40
	$t_{99}$	2244,72	<b>2927,77</b>	<b>2784,68</b>	1900,19	3984,88

A Figura 4.32 apresenta o gráfico dos perfis de degradação para cada uma das 15 unidades emissoras de laser.

Notamos que os perfis são lineares e que 3 das 15 unidades (20%) ultrapassaram o limiar crítico  $D_f$  durante o período de coleta de dados. O objetivo aqui é obter a distribuição do tempo até a falha dos lasers bem como outras características tais como percentis da distribuição.

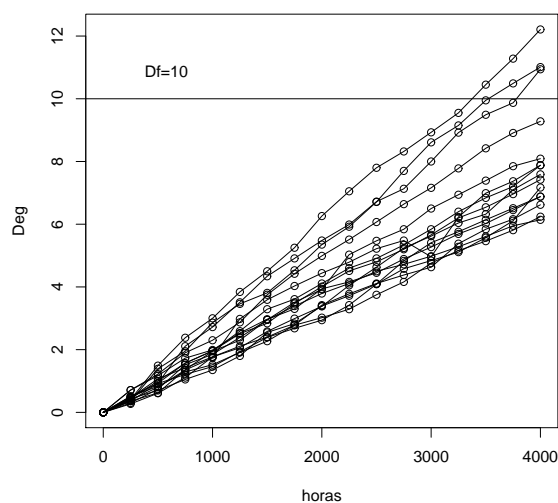


Figura 4.32: Perfil de degradação das unidades emissoras laser.

Hamada (2005) analisou estes dados utilizando distribuição weibull para os efeitos aleatórios do modelo de degradação pela abordagem bayesiana. Peng e Tseng (2013) analisou este mesmo conjunto de dados via abordagem clássica através de um processo estocástico por um incremento linear aleatório com *drift* browniano, e constatou que a distribuição Skew Normal cumpria bem o papel de descrever o incremento linear aleatório.

Neste trabalho os dados dos lasers serão re-analisados utilizando também uma abordagem bayesiana. A motivação para reanalisarmos estes dados é que, diferentemente do trabalho de Peng e Tseng (2013) que assumia uma degradação linear no tempo e no efeito aleatório, aqui a forma funcional do modelo é linear no tempo, porém é não linear no efeito aleatório. Por exemplo, um modelo W pode descrever bem o efeito aleatório  $\beta$  para um modelo do tipo  $D = t/\beta$ , porém não necessariamente descreverá bem o efeito aleatório no modelo  $D = \beta t$ .

**Análise e Discussão**

As respectivas fda e fdp encontram-se na Figura 4.33. A distribuição acumulada empírica foi baseada nas estimativas procedentes dos pseudo tempos de falha (Meeker e Escobar, 1998).

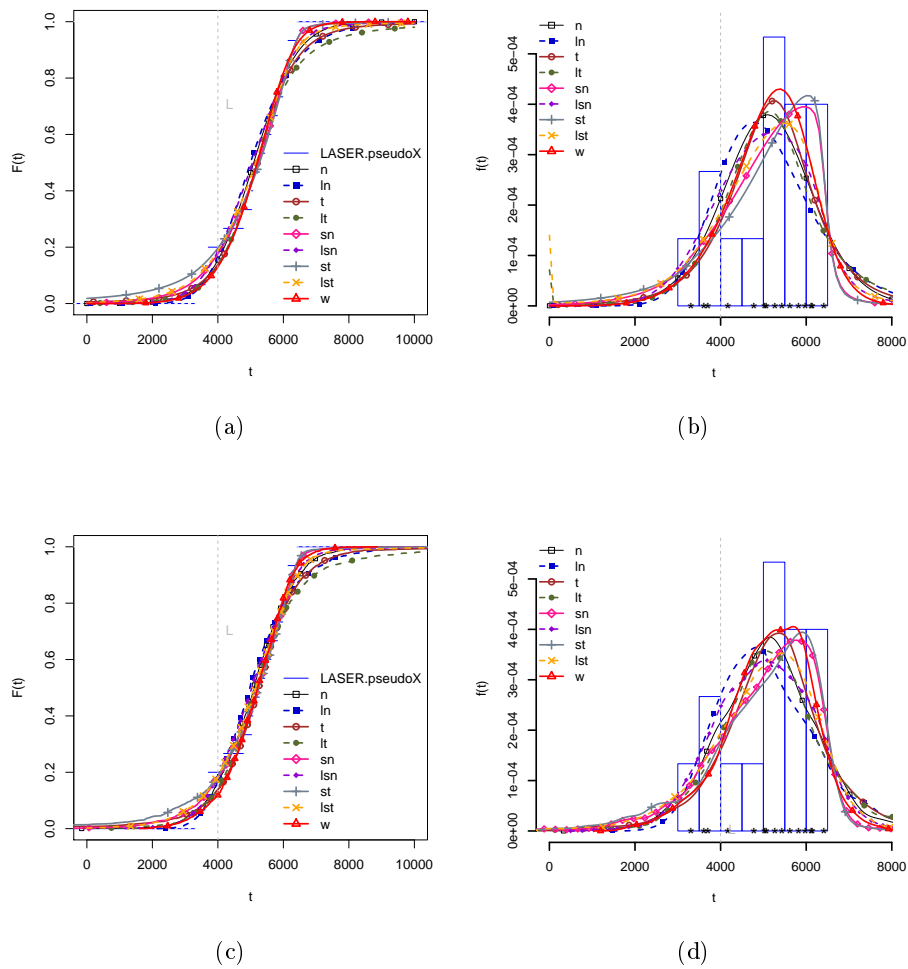


Figura 4.33: Função distribuição acumulada (esquerda) e densidade de probabilidade (direita) *a posteriori* dos tempos de falha para uma unidade futura, pelos métodos de Hamada (2005) (acima) e proposta pela preditiva a posteriori (embaixo), para todos os modelos ajustados aos dados de laser

Pelas Figuras 4.33 observamos que modelo W aparenta ser o mais adequado para



modelar os tempos de falha. Observe pela fdp dos tempos de falha que o modelo W indica bem a posição relativa da moda e as caudas da distribuição empírica dos pseudo tempos de falha. Outro modelo que desempenha um bom papel é o SN, note que ele também acompanha a moda do histograma e a cauda direita, apesar de superestimar um pouco a cauda esquerda.

A Tabela 4.12 fornece os modelos ordenados (melhores no topo e piores no pé) segundo os critérios de seleção de modelos apresentados nas Tabelas C.18 a C.20 do Anexo C. Verificamos que o critério LPML, e FB baseado nos pseudo tempos de falha, indicam que o melhor ajuste é fornecido pelo modelo W, os demais critérios selecionam modelos que superestimam excessivamente alguma das caudas da distribuição empírica. Como os critérios de seleção não concordam, optamos pela escolha do modelo W, indicado pelo critério LPML, por conta do melhor comportamento da análise gráfica com base na Figura 4.33.

Tabela 4.12: Modelos ordenados do melhor (topo) para o pior (pé) segundo LPML, DIC, WAIC, Fator de Bayes (FB), Teste de Kolmogorov-Smirnov (KS), baseados na degradação (Deg) e pseudo tempos de falha (PTF), Dados de Laser

Deg				PTF			FB		KS	
LPML	$LPML_{perfil}$	DIC	WAIC	LPML	DIC	WAIC	Deg	PTF	$Hamada_{pseudo}$	$Prop_{pseudo}$
<b>W</b>	<b>T</b>	<b>T</b>	<b>T</b>	<b>W</b>	<b>SN</b>	<b>LSN</b>	<b>ST</b>	<b>W</b>	<b>ST</b>	<b>SN</b>
T	LN	LSN	W	LSN	W	LST	LSN	N	SN	ST
ST	ST	ST	N	LST	LSN	N	SN	LSN	W	W
LSN	N	N	LT	N	ST	LN	N	LN	LST	LST
N	W	W	ST	LN	LST	W	LT	T	T	N
LT	LST	LT	LSN	T	N	ST	LST	LST	N	T
LST	SN	LN	LN	LT	T	SN	T	LT	LSN	LT
LN	LSN	LST	LST	ST	LN	T	W	ST	LT	LSN
SN	LT	SN	SN	SN	LT	LT	LN	SN	LN	LN

A Tabela 4.13 fornece os percentis *a posteriori* (1%, 5%, 95% e 99%) dos tempos de falha para um laser futuro da população. Segundo o modelo W, 1% das unidades emissoras de laser atingem o limiar crítico antes de 2501,9 horas de uso, 5% antes de 3313,9 horas, 95% antes de 6548,6 horas e 99% antes de 7174,1 horas para o método proposto por Oliveira (2011).

O método proposto por Hamada (2005) apresentou resultados similares. Os res-

pectivos intervalos HPD de 95% de credibilidade são disponibilizados na Tabela 4.13 apenas a título ilustrativo, uma vez que não possui interpretabilidade que traga alguma informação útil para a tomada de decisão.

Tabela 4.13: Percentis da distribuição *a posteriori* dos tempos de falha para os emissores de laser

Ajuste	$t_{p\%}$	pseudo tempos Empírico $\times 10^3$	Oliveira (2011)		Hamada (2005)	
			Pred.	Post.	Média	$HPD_{INF}$
N	$t_1$	3350,38	2386,00	2631,10	1449,44	3610,17
	$t_5$	3521,14	3217,95	3352,79	2464,04	4167,41
	$t_{95}$	6224,33	6881,60	6836,54	6014,40	7759,81
	$t_{99}$	6376,85	7893,08	7558,24	6467,47	8686,81
LN	$t_1$	3350,38	2961,99	2976,70	2257,42	3640,42
	$t_5$	3521,14	3456,95	3461,64	2765,57	4040,73
	$t_{95}$	6224,33	7421,45	7272,10	6146,05	8747,92
	$t_{99}$	6376,85	9001,65	8507,16	6896,24	10842,38
T	$t_1$	3350,38	735,68	-16112,17	-6416,85	3739,71
	$t_5$	3521,14	3015,85	2974,73	1542,91	4386,59
	$t_{95}$	6224,33	7246,35	7421,58	6011,64	9052,23
	$t_{99}$	6376,85	9341,99	26508,48	6499,14	17077,72
LT	$t_1$	3350,38	2143,34	2319,90	0,00	3391,80
	$t_5$	3521,14	3251,40	3352,83	2207,20	4245,97
	$t_{95}$	6224,33	8075,90	9402,59	6055,35	11809,69
	$t_{99}$	6376,85	13724,30	2,91E+73	6458,51	134836,28
SN	$t_1$	3350,38	1472,80	1890,94	-130,95	3417,53
	$t_5$	3521,14	2824,65	2947,83	1465,31	4094,27
	$t_{95}$	6224,33	6457,00	6523,80	6049,47	7430,58
	$t_{99}$	6376,85	7071,34	6833,84	6170,69	8661,51
LSN	$t_1$	3350,38	2469,67	2638,61	1745,26	3448,49
	$t_5$	3521,14	3217,95	3232,91	2416,25	3962,31
	$t_{95}$	6224,33	6715,55	6798,86	5839,11	8088,48
	$t_{99}$	6376,85	7745,46	7494,66	6206,67	9985,42
ST	$t_1$	3350,38	-836,03	-2294,10	-15549,68	3633,96
	$t_5$	3521,14	2344,65	1691,00	-1826,96	4233,01
	$t_{95}$	6224,33	6438,00	6557,10	6037,77	7463,81
	$t_{99}$	6376,85	6937,83	6946,85	6159,41	9063,34
LST	$t_1$	3350,38	1524,96	1836,32	0,00	3147,54
	$t_5$	3521,14	2813,55	2827,37	1073,57	4189,95
	$t_{95}$	6224,33	6761,10	6938,56	5946,02	8584,14
	$t_{99}$	6376,85	7899,94	3,61E+14	6195,93	13082,95
W	$t_1$	3350,38	<b>2501,85</b>	<b>2685,34</b>	1787,64	3563,54
	$t_5$	3521,14	<b>3313,85</b>	<b>3452,39</b>	2652,62	4246,41
	$t_{95}$	6224,33	<b>6548,60</b>	<b>6557,80</b>	5977,53	7227,45
	$t_{99}$	6376,85	<b>7174,10</b>	<b>7027,36</b>	6339,70	7929,71

## Capítulo 5

# Modelagem robusta para análise convencional de tempo de falha

Por se tratarem de distribuições relativamente novas, as distribuições assimétricas e assimétricas de caudas pesadas nas famílias SMSN e SMLSN, tanto quanto sabemos, ainda não foram utilizadas para modelar tempo até a falha no contexto convencional. Contudo, alguns artigos já utilizaram distribuições de cauda pesada para dados de tempo de falha.

Antes de discorrermos sobre a literatura, ressaltamos que utilizaremos o termo tempo de vida, em substituição para tempo de falha, sempre que nos referirmos a dados em saúde. Além disso construiremos procedimentos inferenciais apenas para casos de censura à direita, isto é, quando a falha ocorre à direita do tempo registrado. Dessa maneira utilizaremos o termo censura como sinônimo de censura à direita.

Em Barros *et al.* (2008) e Cancho *et al.* (2010) os autores assumem um modelo linear para um dos parâmetros da distribuição associada ao logaritmo do tempo de vida, o qual é distribuído segundo a distribuição Generalizada Birnbaum Saunders Student-t proposta por Díaz-García e Leiva-Sánchez (2005). Num trabalho bastante recente desenvolvido por Vallejos e Steel (2015) os autores utilizam uma mistura na forma

da distribuição Lognormal (*Shape Mixture of Lognormal – SMLN*) para modelar o tempo de vida na presença de valores atípicos. O objetivo é robustecer a inferência sobre os coeficientes de covariáveis inseridas no modelo através de um preditor linear para o parâmetro de escala.

Neste capítulo serão abordados os aspectos inferenciais da obtenção da distribuição do tempo até a falha no contexto das distribuições assimétricas de cauda pesada, com presença ou não de censura, em que a modelagem adotada engloba a não presença de censura como caso particular. O processo inferencial será baseado na técnica de aumento de dados, que também foi útil na modelagem dos dados de degradação apresentadas no Capítulo 4. Esta técnica tende a facilitar a inferência, a qual nem sempre é trivialmente obtida, se distribuições assimétricas são consideradas.

Em situações práticas, espera-se que itens amostrais possam experimentar falhas prematuras ou tardias (valores atípicos) e as mais variadas formas de assimetria, portanto a consideração dessas distribuições assimétricas e de caudas pesadas desempenhará um papel importante em termos de robustez e ampliação de possibilidades para a análise de tempo de falha. Neste trabalho assumiremos as distribuições nas famílias SMSN e SMLSN, também utilizadas no Capítulo 4, para modelarmos o comportamento dos tempos de falha. Diferentemente dos trabalhos de Barros *et al.* (2008); Cancho *et al.* (2010) e Vallejos e Steel (2015), nosso objetivo principal é modelar os tempos de vida sem a introdução de covariáveis. A inserção de preditores lineares na modelagem será realizada em trabalhos futuros. Os modelos desenvolvidos serão aplicados a bancos de dados artificialmente gerados e também a dados de câncer de pulmão (Kalbfleisch e Prentice, 2002).

Este capítulo está assim organizado. Na Seção 5.1 apresentamos a modelagem para tempos de falha na presença de censura. Para facilitar o processo de inferência nos modelos propostos utilizaremos as técnicas de aumento de dados para dados censurados apresentada em Chib (1992), a qual nos permite reescrever a função de verossimilhança

de forma mais tratável. Na Seção 5.2.1 introduzimos os modelos robustos SMSN e SMLSN para tempos de falha e apresentamos um método para amostrar das distribuições *a posteriori* dos parâmetros destes modelos. Na Seção 5.2.2 trabalhamos com a parametrização centrada do modelo SN (Azzalini, 1985), a fim de permitir a modelagem de tempos de falha através das distribuições Skew Normal e Log Skew normal em casos de menores graus de assimetria. Na Seção 5.3 geramos três bancos de dados e os analisamos para estudar os modelos propostos neste capítulo. Finalmente, na Seção 5.4 aplicamos os modelos desenvolvidos nesse capítulo à mesma base de dados de câncer de pulmão (Kalbfleisch e Prentice, 2002) também utilizada por Barros *et al.* (2008); Cancho *et al.* (2010); Vallejos e Steel (2015).

## 5.1 Modelo para tempos de falha e técnica de aumento de dados

Defina por  $T_i$  o tempo até a falha do item  $i$  sob teste e por  $C_i$  como tempo de censura, ou seja, último instante de tempo em que o item amostral  $i$  foi observado,  $i = 1, \dots, n$ . O item  $i$  é considerado censurado se saiu do estudo antes que falhasse por algum modo de falha alheio ao de interesse. Seja  $\ell_i$  a função indicadora de falha da unidade amostral  $i$ , assumindo valor 1 se o item  $i$  falhou e zero caso contrário.

Assuma que  $\boldsymbol{\theta}$  é o vetor de parâmetros que indexa a distribuição do tempo até a falha, denote por  $f(t|\boldsymbol{\theta})$  a fdp de  $T$  e por  $R(t|\boldsymbol{\theta}) = \int_t^\infty f(a|\boldsymbol{\theta}) da$  a função de sobrevivência de  $T$  para um item  $i$  sob teste que sobreviveu até o tempo  $t$ . Assuma que os tempos de falha são iid e seja  $X_i = \min\{T_i, C_i\}$ . Conseqüentemente, a função de verossimilhança associada a  $X_i$  para  $n$  itens sob teste, com  $i = 1, \dots, n$ , pode ser escrita como segue

$$f(\mathbf{x}, \boldsymbol{\ell}|\boldsymbol{\theta}) = \prod_{i=1}^n f(x_i|\boldsymbol{\theta})^{\ell_i} R(x_i|\boldsymbol{\theta})^{1-\ell_i}$$

$$= \prod_{i=1}^n f(x_i|\boldsymbol{\theta})^{\ell_i} \left\{ \int_{x_i}^{\infty} f(t_i|\boldsymbol{\theta}) dt_i \right\}^{1-\ell_i}. \quad (5.1)$$

Existem casos em que  $R(t|\boldsymbol{\theta})$  não possui forma fechada. Neste contexto as funções de verossimilhança e log-verossimilhança são de difícil otimização via métodos de busca baseados no vetor gradiente, então opta-se pela utilização de outras estratégias para facilitar o procedimento inferencial.

Tanner e Wong (1987) fornecem uma estratégia computacionalmente menos dispendiosa para amostrar da distribuição dos parâmetros do modelo sem que seja necessária a integração, a qual referenciam como aumento de dados.

A ideia é realizar a inferência bayesiana a partir de amostras obtidas via métodos MCMC propostos por Gelfand e Smith (1990), num procedimento muito similar ao adotado por Chib (1992).

Defina  $\mathbf{T} = \{T_1, \dots, T_n\}$  e  $\mathbf{C} = \{C_1, \dots, C_n\}$ , respectivamente, como vetores aumentados com os tempos de falha e tempos de censura (não observados). Se  $\ell_i = 1$  então  $T_i \leq C_i$  e observamos um tempo de falha. Se  $\ell_i = 0$  então  $T_i > C_i$  e observamos um tempo de censura, o que torna o tempo até a falha  $T_i$  uma variável latente. Por conveniência, denote por  $\mathbf{t}_{Obs}$  o vetor de tempos observados e por  $\mathbf{t}_{Lat}$  o vetor de tempos latentes.

Usando a técnica proposta por Tanner e Wong (1987) temos que inferir sobre  $\mathbf{t}_{Lat}$ . Isto é possível a partir de uma reformulação da função de verossimilhança que inclua os tempos latentes. Considerando a notação anteriormente definida observamos que, para os tempos realmente observados, a fdp  $f(t_i|\boldsymbol{\theta}, \ell_i = 1, c_i) = f(t_i|\boldsymbol{\theta})$ . É necessário encontrarmos também a fdp  $f(t_i|\boldsymbol{\theta}, \ell_i = 0, c_i)$ , correspondente aos tempos de falha latentes. Se para um item amostral  $T_i$  é um tempo de falha latente (censura), e o tempo censurado observado para este mesmo item foi  $C_i = c_i$ , a fda associada a este tempo latente  $T_i$  é

$$F_{T_i|\boldsymbol{\theta}, \ell=0, c_i}(t_i) = \frac{P(T_i \leq t_i, T_i > c_i|\boldsymbol{\theta})}{P(T_i > c_i|\boldsymbol{\theta})}$$

$$\begin{aligned}
&= \frac{P(c_i \leq T_i \leq t_i | \boldsymbol{\theta})}{P(T_i > c_i | \boldsymbol{\theta})} \\
&= \frac{F(t_i | \boldsymbol{\theta}) - F(c_i | \boldsymbol{\theta})}{P(T_i > c_i | \boldsymbol{\theta})}. \tag{5.2}
\end{aligned}$$

Derivando a expressão em (5.2) obtemos que a fdp de  $T_i$ , dados  $\boldsymbol{\theta}$  e  $\ell = 0, c_i$ , é

$$\begin{aligned}
f(t_i | \boldsymbol{\theta}, \ell = 0, c_i) &= \frac{f(t_i | \boldsymbol{\theta})}{P(T_i > c_i | \boldsymbol{\theta})} \\
&= f(t_i | \boldsymbol{\theta}) I(t_i)_{(c_i, \infty)}, \tag{5.3}
\end{aligned}$$

em que  $I(t)_A$  denota a função indicadora assumindo 1 se  $t \in A$  e 0 caso contrário.

Note da equação (5.3) que  $f(t_i | \boldsymbol{\theta}, \ell = 0, c_i)$  é a função taxa de falha se  $t_i = c_i$ .

Do resultado na equação (5.3) a função de verossimilhança completa é dada pela equação

$$f(\mathbf{t}, \boldsymbol{\ell} | \boldsymbol{\theta}) \propto \prod_{i=1}^n f(t_i | \boldsymbol{\theta})^{\ell_i} \{f(t_i | \boldsymbol{\theta}) I(t_i)_{(c_i, \infty)}\}^{1-\ell_i} \tag{5.4}$$

Observe que a função de verossimilhança na equação (5.4) não é equivalente àquela dada em (5.1). A função de verossimilhança em (5.1) é referente à variável aleatória  $X_i = \min\{T_i, C_i\}$ , já a função de verossimilhança em (5.4) foi escrita considerando a variável aleatória  $T_i$ , em que  $\mathbf{t}$  é composto por tempos de falha observados e latentes.

Assuma que, *a priori*,  $\boldsymbol{\theta} \sim f(\boldsymbol{\theta})$ . Como consequência, temos que a distribuição à posteriori para  $(\boldsymbol{\theta}, \mathbf{t}_{lat})$ , dados  $\mathbf{t}_{obs}$ ,  $\boldsymbol{\ell}$  e  $\mathbf{c}$  é

$$f(\boldsymbol{\theta}, \mathbf{t}_{Lat} | \mathbf{t}_{Obs}, \boldsymbol{\ell}, \mathbf{c}) \propto \prod_{i=1}^n \left\{ f(t_i | \boldsymbol{\theta}, \ell_i = 1, c_i)^{\ell_i} f(t_i | \boldsymbol{\theta}, \ell_i = 0, c_i)^{1-\ell_i} \right\} f(\boldsymbol{\theta}). \tag{5.5}$$

Na próxima seção aplicaremos esta técnica em modelos robustos para o tempo até a falha.

## 5.2 Modelos robustos para tempo até a falha

Considerando distribuições para o tempo até a falha nas famílias SMSN e SMLSN apresentadas no Capítulo 2 podemos construir os modelos robustos para tempo até a falha. Dada a complexidade da forma destas distribuições, a técnica de aumento de dados apresentada na Seção 5.1 facilita o processo de inferência, pois não considera na função de verossimilhança a distribuição acumulada de tais distribuições.

Assuma que os tempos de falha  $T_i \stackrel{iid}{\sim} \text{SMSN}(\mu, h_i^{-1}\omega^2, \alpha, \mathcal{H})$ . A partir da equação 5.4 podemos escrever a função de verossimilhança como segue

$$f(\mathbf{t}, \boldsymbol{\ell} | \boldsymbol{\theta}) = \prod_{i=1}^n \left[ 2 \int_0^\infty \phi(t_i | \mu, h_i^{-1}\omega^2) \Phi\left(\alpha \frac{(t_i - \mu)}{h_i^{-1/2}\omega}\right) d\mathcal{H}(h_i | \boldsymbol{\nu}) \right]^{\ell_i} \\ \times \left[ 2 \int_0^\infty \phi(t_i | \mu, h_i^{-1}\omega^2) \Phi\left(\alpha \frac{(t_i - \mu)}{h_i^{-1/2}\omega}\right) d\mathcal{H}(h_i | \boldsymbol{\nu}) I(t_i)_{(c_i, \infty)} \right]^{1-\ell_i}. \quad (5.6)$$

Outra extensão que pode ser feita para modelarmos o comportamento do tempo até a falha, para acomodar tanto assimetria quanto caudas pesadas, é assumirmos que  $T_i \stackrel{iid}{\sim} \text{SMLSNS}(\mu, h^{-1}\omega^2, \alpha, \mathcal{H})$ . Este caso é uma extensão natural do trabalho de Vallejos e Steel (2015), o qual visava tornar robusta a estimativa de coeficientes do preditor linear para o parâmetro de escala na modelagem de tempos de falha que tinham seu comportamento descrito por distribuições na família de distribuições com mistura na forma da distribuição log normal. Além da robustez, adicionamos maior flexibilidade ao modelo através da incorporação do parâmetro de assimetria ao assumirmos distribuições na família SMLSNS. No caso considerado aqui, segue que a função de verossimilhança é dada por

$$f(\mathbf{t}, \boldsymbol{\ell} | \boldsymbol{\theta}) = \prod_{i=1}^n \left[ \frac{2}{t_i} \int_0^\infty \phi(\log(t_i) | \mu, h_i^{-1}\omega^2) \Phi\left(\alpha \frac{(\log(t_i) - \mu)}{h_i^{-1/2}\omega}\right) d\mathcal{H}(h_i | \boldsymbol{\nu}) \right]^{\ell_i} \\ \times \left[ \frac{2}{t_i} \int_0^\infty \phi(\log(t_i) | \mu, h_i^{-1}\omega^2) \Phi\left(\alpha \frac{(\log(t_i) - \mu)}{h_i^{-1/2}\omega}\right) d\mathcal{H}(h_i | \boldsymbol{\nu}) I(t_i)_{(c_i, \infty)} \right]^{1-\ell_i}. \quad (5.7)$$

O modelo proposto por Vallejos e Steel (2015), adaptado ao contexto aqui discutido, surge como caso particular de (5.7) ao assumirmos  $\alpha = 0$ .

As funções de verossimilhança em (5.6) e (5.7) são de difícil otimização, apesar de suas formas já terem sido simplificadas pelo tratamento de dados aumentados para lidar com as censuras. Contudo, a técnica de aumento de dados não é utilizada somente para lidar com dados censurados. Seu uso é também amplamente considerado visando simplificar procedimentos computacionais (veja Dyk e Meng, 2001, por exemplo). Neste caso, variáveis latentes são artificialmente incluídas no modelo, visando no contexto



bayesiano obtermos formas mais tratáveis computacionalmente para as distribuições condicionais completas. Consideremos primeiramente o modelo dado em (5.6), onde assumimos que  $T_i \stackrel{iid}{\sim} \text{SMSN}(\mu, h_i^{-1}\omega^2, \alpha, \mathcal{H})$ . Neste caso, considerando a representação estocástica dada na equação (2.8), isto é

$$T_i = \mu + \Delta U_i + H_i^{-1/2} \tau^{1/2} U_1$$

em que  $\Delta = \omega\delta$ ,  $\delta = \frac{\alpha}{\sqrt{1+\alpha^2}}$ ,  $\tau = (1 - \delta^2)\omega^2$ ,  $U_i = H_i^{-1/2} |U_0|$ ,  $U_1 \perp U_0 \stackrel{iid}{\sim} \text{N}(0,1)$  e  $\mathcal{H}$  é a medida de probabilidade definida sobre o espaço amostral de  $H$  indexada pelo vetor  $\boldsymbol{\nu}$ , o modelo em (5.6) pode ser equivalentemente construído considerando a seguinte representação hierárquica para a distribuição de  $T_i$

$$\left. \begin{aligned} T_i | U_i = u_i, H_i = h_i &\stackrel{iid}{\sim} \text{N}(\mu + \Delta u_i, h_i^{-1}\tau) \\ U_i | H_i = h_i &\stackrel{iid}{\sim} \text{N}(0, h_i^{-1}) I(u_i)_{(0,\infty)} \\ H_i &\stackrel{iid}{\sim} \mathcal{H}(h_i | \boldsymbol{\nu}) \end{aligned} \right\} \quad (5.8)$$

Assim, assumindo a representação hierárquica em (5.8), a função de verossimilhança em (5.6) pode ser reescrita da seguinte forma

$$f(\mathbf{t}, \boldsymbol{\ell} | \boldsymbol{\theta}) = \prod_{i=1}^n \phi(t_i | \mu + \Delta u_i, h_i^{-1}\tau)^{\ell_i} \{ \phi(t_i | \mu + \Delta u_i, h_i^{-1}\tau) I(t_i)_{(c_i, \infty)} \}^{1-\ell_i}. \quad (5.9)$$

Sabe-se que se  $T_i \stackrel{iid}{\sim} \text{SMLSN}(\mu, h_i^{-1}\omega^2, \alpha, \mathcal{H})$ , então  $\log(T_i) \stackrel{iid}{\sim} \text{SMSN}(\mu, h_i^{-1}\omega^2, \alpha, \mathcal{H})$ . Analogamente a função de verossimilhança em (5.7) pode ser reescrita em função de  $\log(T_i)$  como sendo

$$f(\log(\mathbf{t}), \boldsymbol{\ell} | \boldsymbol{\theta}) = \prod_{i=1}^n \phi(\log(t_i) | \mu + \Delta u_i, h_i^{-1}\tau)^{\ell_i} \{ \phi(\log(t_i) | \mu + \Delta u_i, h_i^{-1}\tau) I(t_i)_{(c_i, \infty)} \}^{1-\ell_i}. \quad (5.10)$$

Considerando essa nova estrutura, faz-se necessário obtermos as distribuições *a posteriori* para o vetor de objetos aleatórios  $\boldsymbol{\psi} = (\mathbf{t}_{\text{lat}}, \mathbf{U}, \mathbf{H}, \mu, \Delta, \tau)$ . Assuma que *a priori*  $\mathbf{U}, \mathbf{H}, \mu, \Delta$  e  $\tau$  são independentes e considere as seguintes distribuições *a priori* para  $\mu, \Delta$  e  $\tau$ . Para completarmos a especificação do modelo assumamos que *a priori*

$$\left. \begin{aligned} \mu &\sim \text{N}(m, V) \\ \Delta &\sim \text{N}(\mu_\Delta, \sigma_\Delta^2) \\ \tau &\sim \text{GI}(r_\tau, s_\tau) \end{aligned} \right\}, \quad (5.11)$$

em que  $m$  e  $\mu_\Delta$  são números reais e  $V$ ,  $\sigma_\Delta^2$ ,  $r_\tau$  e  $s_\tau$  são números reais positivos. Além disso, se  $X \sim \text{GI}(a, b)$ , com  $E(x) = b(a-1)^{-1}$  para  $a > 1$ , e  $V(x) = b(a-1)^{-2}(a-2)^{-1}$  para  $a > 2$ .

Dada a função de verossimilhança em (5.9) e as especificações *a priori* para os objetos aleatórios em (5.8) e (5.11), a distribuição *a posteriori* do modelo SMSN para os tempos de falha é dada por

$$\begin{aligned} f(\boldsymbol{\psi}|\mathbf{t}, \boldsymbol{\ell}) \propto & \left\{ \prod_{i=1}^n \phi(t_i|\mu + \Delta u_i, h_i^{-1}\tau)^{\ell_i} [\phi(t_i|\mu + \Delta u_i, h_i^{-1}\tau) I(t_i)_{(c_i, \infty)}]^{1-\ell_i} \right. \\ & \left. \times \phi(u_i|0, h_i^{-1}) I(u_i)_{(0, \infty)} \mathcal{H}(h_i|\boldsymbol{\nu}) \right\} \phi(\mu|m, V) \phi(\Delta|\mu_\Delta, \sigma_\Delta^2) GI(\tau|r_\tau, s_\tau). \end{aligned} \quad (5.12)$$

Similarmente, de posse da função de verossimilhança em (5.10) e a especificação *a priori* para os objetos aleatórios em (5.8) e (5.11), tem-se que a distribuição *a posteriori* do modelo SMLS<sub>N</sub> para os tempos de falha é escrita como

$$\begin{aligned} f(\boldsymbol{\psi}|\log(\mathbf{t}), \boldsymbol{\ell}) \propto & \left\{ \prod_{i=1}^n \phi(\log(t_i)|\mu + \Delta u_i, h_i^{-1}\tau)^{\ell_i} [\phi(\log(t_i)|\mu + \Delta u_i, h_i^{-1}\tau) I(t_i)_{(c_i, \infty)}]^{1-\ell_i} \right. \\ & \left. \times \phi(u_i|0, h_i^{-1}) I(u_i)_{(0, \infty)} \mathcal{H}(h_i|\boldsymbol{\nu}) \right\} \phi(\mu|m, V) \phi(\Delta|\mu_\Delta, \sigma_\Delta^2) GI(\tau|r_\tau, s_\tau). \end{aligned} \quad (5.13)$$

Ambas as distribuições *a posteriori* em (5.12) e (5.13) não possuem forma funcional fechada. Para amostrar da distribuição *a posteriori* construiremos um procedimento computacional baseado em métodos MCMC, o qual é descrito na Seção 5.2.1.

### 5.2.1 Amostragem via MCMC

Para obtermos amostras da distribuição *a posteriori* via amostrador de Gibbs é necessário obtermos as dcc dos componentes de  $\boldsymbol{\psi} = (\mathbf{t}_{lat}, \mathbf{U}, \mathbf{H}, \mu, \Delta, \tau)$ . Para o modelo (5.12) tais distribuições são fornecidas na seguinte proposição.

**Proposição 10.** *Assuma que os tempos de falha são tais que  $T_i \stackrel{iid}{\sim} \text{SMSN}(\mu, h^{-1}\omega^2, \alpha, \mathcal{H})$  e considere as especificações a priori dadas em (4.19). Então, segue que:*

i A dcc a posteriori para o componente  $t_i$  de  $\mathbf{t}_{lat}$

$$t_i | \boldsymbol{\psi}^{[-t_i]}, \mathbf{D} \stackrel{d}{=} t_i | \boldsymbol{\psi}^{[-t_i]}, \ell_i = 0, c_i \sim N(\mu + \Delta u_i, h_i^{-1} \tau) I(t_i)_{(c_i, \infty)},$$

ii As dcc a posteriori para  $u_i$ ,  $\mu$ ,  $\Delta$  e  $\tau$  são dadas, respectivamente por:

$$u_i | \boldsymbol{\psi}^{[-u_i]}, \mathbf{D} \sim N\left(\frac{\Delta(t_i - \mu)}{\Delta^2 + \tau}, \frac{h_i^{-1} \tau}{\Delta^2 + \tau}\right) I(u_i)_{(0, \infty)},$$

$$\mu | \boldsymbol{\psi}^{[-\mu]}, \mathbf{D} \sim N\left(\frac{\sum_{i=1}^n t_i v_i^* + m\tau \prod_{i=1}^n h_i^{-1} - \Delta \sum_{i=1}^n u_i v_i^*}{V \sum_{i=1}^n v_i^* + \tau \prod_{i=1}^n h_i^{-1}}, \frac{V\tau \prod_{i=1}^n h_i^{-1}}{\sum_{i=1}^n v_i^* + \tau \prod_{i=1}^n h_i^{-1}}\right),$$

$$\Delta | \boldsymbol{\psi}^{[-\Delta]}, \mathbf{D} \sim N\left(\frac{\sigma_\Delta^2 \left[\sum_{i=1}^n t_i u_i^* - \sum_{i=1}^n u_i^*\right] + \mu_\Delta \tau \prod_{i=1}^n h_i^{-1}}{\sigma_\Delta^2 \sum_{i=1}^n u_i^* + \tau \prod_{i=1}^n h_i^{-1}}, \frac{\sigma_\Delta^2 \tau \prod_{i=1}^n h_i^{-1}}{\sigma_\Delta^2 \sum_{i=1}^n u_i^* + \tau \prod_{i=1}^n h_i^{-1}}\right),$$

$$\tau | \boldsymbol{\psi}^{[-\tau]}, \mathbf{D} \sim GI\left(\frac{n}{2} + r_\tau, \sum_{i=1}^n \frac{[t_i - (\mu + \Delta u_i)]^2}{2h_i^{-1}} + s_\tau\right),$$

em que  $v_i^* = V \prod_{k \neq i} h_k^{-1}$ ,  $u_i^* = u_i \prod_{k \neq i} h_k^{-1}$ ,  $k = \{1, \dots, n\}$  e  $\boldsymbol{\psi}^{[-a]}$  denota o vetor  $\boldsymbol{\psi}$  sem a componente  $a$ .

A prova da Proposição 10 é obtida de resultados usuais de cálculo de probabilidade e pode ser encontrada no apêndice A.

Note que ao utilizarmos a representação estocástica da família de distribuições SMSN (aumento de dados) obtivemos formas fechadas para todas as dcc *a posteriori*. Isto nos permite utilizar o amostrador de Gibbs para gerarmos amostras das distribuições *a posteriori* para  $\mu$ ,  $\Delta$  e  $\tau$ , e também para as variáveis latentes  $t_i \in \mathbf{t}_{lat}$  e  $u_i$ .

Se  $\mathcal{H}$  é uma distribuição degenerada com toda massa de probabilidade sobre 1, então  $T_i$  tem distribuição skew normal e as dcc *a posteriori* para todos os parâmetros são obtidas de (i) e (ii) fazendo-se  $h_i = 1$ ,  $\forall i = 1, \dots, n$ . Por outro lado, se  $\mathcal{H}$  é a distribuição Gama( $\nu/2, \nu/2$ ), então o modelo de tempos de falha ST é definido. Assumindo que *a priori*  $\nu$  tem a distribuição *a priori* Jeffreys independente (Fonseca

*et al.*, 2008) dada por

$$f(\nu) \propto \left( \frac{\nu}{\nu+3} \right)^{\frac{1}{2}} \left\{ \frac{\partial^2 \log [\Gamma(\frac{\nu}{2})]}{\partial \nu^2} - \frac{\partial^2 \log [\Gamma(\frac{\nu+1}{2})]}{\partial \nu^2} - \frac{2(\nu+3)}{\nu(\nu+1)^2} \right\}^{\frac{1}{2}}, \quad (5.14)$$

segue que as dcc *a posteriori* para  $t_i \in \mathbf{t}_{\text{lat}}$ ,  $u_i$ ,  $\mu$ ,  $\Delta$  e  $\tau$ , sob o modelo ST, são dadas na Proposição 10 e as dcc *a posteriori* para  $h_i$  e  $\nu$  são, respectivamente, dadas por

$$\begin{aligned} h_i | \boldsymbol{\psi}^{[-h_i]}, \mathbf{D} &\sim \text{Gama} \left( \frac{\nu}{2} + 1, \frac{1}{2} \left\{ \tau^{-1} [t_i - (\mu + \Delta u_i)]^2 + u_i^2 + \nu \right\} \right), \\ f(\nu | \boldsymbol{\psi}^{[-\nu]}, \mathbf{D}) &\propto \left[ \frac{\frac{\nu}{2}}{\Gamma(\frac{\nu}{2})} \right]^n \left[ \prod_{i=1}^n h_i^{\frac{\nu}{2}-1} \right] \exp \left\{ - \sum_{i=1}^n h_i \frac{\nu}{2} \right\} \\ &\times \left( \frac{\nu}{\nu+3} \right)^{\frac{1}{2}} \left\{ \frac{\partial^2 \log [\Gamma(\frac{\nu}{2})]}{\partial^2 \nu} - \frac{\partial^2 \log [\Gamma(\frac{\nu+1}{2})]}{\partial^2 \nu} - \frac{2(\nu+3)}{\nu(\nu+1)^2} \right\}^{\frac{1}{2}}. \end{aligned} \quad (5.15)$$

Segue do Corolário 2 em Fonseca *et al.* (2008) que  $f(\nu | \boldsymbol{\psi}^{[-\nu]})$  é própria. A prova das equações em (5.15) é fornecida no Anexo A.

Se assumimos o modelo em (5.7) em que  $T_i \stackrel{iid}{\sim} \text{SMLSN}(\mu, h^{-1}\omega^2, \alpha, \mathcal{H})$ , considerando a representação estocástica em (5.8) e a respectiva função de verossimilhança para  $\log(T_i)$  em (5.10) podemos inferir sobre  $\boldsymbol{\psi} = (\log(\mathbf{t}_{\text{lat}}), \mathbf{U}, \mathbf{H}, \mu, \Delta, \tau)$  obtendo sua distribuição *a posteriori*. A proposição a seguir fornece as dcc *a posteriori* para todos os componentes de  $\boldsymbol{\psi}$

**Proposição 11.** *Considere que os tempos de falha são tais que  $T_i \stackrel{iid}{\sim} \text{SMLSN}(\mu, h^{-1/2}\omega, \alpha, \mathcal{H})$ , em que  $\log(T_i) \stackrel{iid}{\sim} \text{SMSN}(\mu, h^{-1/2}\omega, \alpha, \mathcal{H})$  tem representação estocástica como na equação (5.8) e as especificações a priori conforme (4.19), então:*

*i* A dcc *a posteriori* para o componente de  $\log(t_i)$  em  $\log(\mathbf{t}_{\text{lat}})$  é dada por

$$\log(t_i) | \boldsymbol{\psi}^{[-\log(t_i)]}, \mathbf{D} \stackrel{d}{=} \log(t_i) | \boldsymbol{\psi}^{[-\log(t_i)]}, \ell_i = 0, c_i \sim N(\mu + \Delta u_i, h_i^{-1}\tau) I(t_i)_{(c_i, \infty)},$$

*ii* As dcc *a posteriori* para  $u_i$ ,  $\mu$ ,  $\Delta$  e  $\tau$  são dadas, respectivamente, por:

$$u_i | \boldsymbol{\psi}^{[-u_i]}, \mathbf{D} \sim N \left( \frac{\Delta (\log(t_i) - \mu)}{\Delta^2 + \tau}, \frac{h_i^{-1}\tau}{\Delta^2 + \tau} \right) I(u_i)_{(0, \infty)},$$

$$\begin{aligned} \mu|\boldsymbol{\psi}^{[-\mu]}, \mathbf{D} &\sim N \left( \frac{\sum_{i=1}^n \log(t_i) v_i^* + m\tau \prod_{i=1}^n h_i^{-1} - \Delta \sum_{i=1}^n u_i v_i^*}{V \sum_{i=1}^n v_i^* + \tau \prod_{i=1}^n h_i^{-1}}, \frac{V\tau \prod_{i=1}^n h_i^{-1}}{\sum_{i=1}^n v_i^* + \tau \prod_{i=1}^n h_i^{-1}} \right), \\ \Delta|\boldsymbol{\psi}^{[-\Delta]}, \mathbf{D} &\sim N \left( \frac{\sigma_\Delta^2 \left[ \sum_{i=1}^n \log(t_i) u_i^* - \sum_{i=1}^n u_i^* \right] + \mu\Delta\tau \prod_{i=1}^n h_i^{-1}}{\sigma_\Delta^2 \sum_{i=1}^n u_i^* + \tau \prod_{i=1}^n h_i^{-1}}, \frac{\sigma_\Delta^2 \tau \prod_{i=1}^n h_i^{-1}}{\sigma_\Delta^2 \sum_{i=1}^n u_i^* + \tau \prod_{i=1}^n h_i^{-1}} \right), \\ \tau|\boldsymbol{\psi}^{[-\tau]}, \mathbf{D} &\sim GI \left( \frac{n}{2} + r_\tau, \sum_{i=1}^n \frac{[\log(t_i) - (\mu + \Delta u_i)]^2}{2h_i^{-1}} + s_\tau \right), \end{aligned}$$

em que  $v_i^* = V \prod_{k \neq i} h_k^{-1}$ ,  $u_i^* = u_i \prod_{k \neq i} h_k^{-1}$ ,  $k = \{1, \dots, n\}$  e  $\boldsymbol{\psi}^{[-a]}$  denota o vetor  $\boldsymbol{\psi}$  sem a componente  $a$ .

A prova da Proposição 11 é análoga à da Proposição 10 e por esta razão será omitida.

Note que ao utilizarmos a representação estocástica da família de distribuições SMSN (aumento de dados) para  $\log(T_i)$  obtivemos formas fechadas para todas as dcc *a posteriori*. Isto nos permite utilizar o amostrador de Gibbs para gerarmos amostras das distribuições *a posteriori* para  $\mu$ ,  $\Delta$  e  $\tau$ , e também para as variáveis latentes  $\log(t_i) \in \log(\mathbf{t}_{lat})$  e  $u_i$ .

Do modo análogo ao resultado obtido na Proposição 10 se  $\mathcal{H}$  é uma distribuição degenerada colocando toda massa de probabilidade em 1, então  $\log(T_i)$  tem distribuição skew normal e  $T_i$  tem distribuição log skew normal, em que as dcc *a posteriori* para todos os parâmetros são obtidas de (i) e (ii) assumindo  $h_i = 1$ . Se  $\mathcal{H}$  é a distribuição Gama( $\nu/2, \nu/2$ ), então o modelo de tempos de falha ST é definido para  $\log(T_i)$ , isto é,  $T_i$  segue um modelo LST. Assumindo que  $\nu$  tem a distribuição *a priori* de Jeffreys dada em (5.14), segue que as dcc's *a posteriori* para  $h_i$  e  $\nu$  são, respectivamente, dadas por

$$h_i|\boldsymbol{\psi}^{[-h_i]}, \mathbf{D} \sim \text{Gama} \left( \frac{\nu}{2} + 1, \frac{1}{2} \left\{ \tau^{-1} [\log(t_i) - (\mu + \Delta u_i)]^2 + u_i^2 + \nu \right\} \right),$$

$$\begin{aligned}
 f\left(\nu|\boldsymbol{\psi}^{[-\nu]},\mathbf{D}\right) &\propto \left[\frac{\nu^{\frac{\nu}{2}}}{\Gamma(\frac{\nu}{2})}\right]^n \left[\prod_{i=1}^n h_i^{\frac{\nu}{2}-1}\right] \exp\left\{-\sum_{i=1}^n h_i \frac{\nu}{2}\right\} \\
 &\times \left(\frac{\nu}{\nu+3}\right)^{\frac{1}{2}} \left\{\frac{\partial^2 \log [\Gamma(\frac{\nu}{2})]}{\partial^2 \nu} - \frac{\partial^2 \log [\Gamma(\frac{\nu+1}{2})]}{\partial^2 \nu} - \frac{2(\nu+3)}{\nu(\nu+1)^2}\right\}^{\frac{1}{2}},
 \end{aligned} \tag{5.16}$$

a qual é uma fdp própria. A prova das equações em (5.16) será omitida por ser análoga àquelas nas equações em (5.15).

As amostras da distribuição *a posteriori* dos parâmetros originais são obtidos a partir das relações  $\omega^2 = \tau + \Delta^2$  e  $\delta = \Delta\omega^{-1}$ .

Para calcularmos as quantidades de interesse para um item futuro da população utilizamos a média *a posteriori* (similarmente à Hamada (2005)) e a distribuição preditiva *a posteriori*. Por média *a posteriori* entenda o procedimento apresentado na Seção 3.2.1 justificado na Proposição 3.7, e pode ser calculado conforme a Proposição (2). Ressaltamos que a inconsistência teórica em relação aos percentis permanece a mesma. As quantidades de interesse sob a distribuição preditiva *a posteriori* dos tempos de falha são calculadas seguindo o procedimento descrito na Seção 3.2.3, em que a amostra dos tempos de falha para um novo item populacional é obtida de acordo com a Seção 3.2.4.

### 5.2.2 Modelo SN centrado para o tempo até a falha

No caso específico em que os tempos de falha são tais que  $T_i \stackrel{iid}{\sim} \text{SN}(\mu, \omega^2, \alpha)$ , é possível utilizarmos a parametrização centrada (Azzalini, 1985) da distribuição SN. Segundo consta na literatura (Arellano-Valle e Azzalini, 2008), sob esta parametrização a função de verossimilhança é mais bem comportada e permite realizar inferência razoável sobre os parâmetros mesmo quando o parâmetro de assimetria assume valores nas proximidades de  $\alpha = 0$ . A parametrização centrada da distribuição SN centrada é obtida se

assumimos que  $\mu$ ,  $\omega^2$  e  $\alpha$  são tais que

$$\left. \begin{aligned} \mu_o &= \mu + \omega_o c \\ \omega_o &= \omega^2 \sigma_z^2 \\ \alpha_o &= \frac{4 - \pi}{2} \frac{\mu_z^3}{(\sigma_z^2)^{3/2}} \end{aligned} \right\}, \quad (5.17)$$

em que  $\mu_z = b\delta$ ,  $\sigma_z^2 = 1 - \mu_z^2$ ,  $b = \sqrt{\frac{2}{\pi i}}$ ,  $\mu_z = \frac{\rho}{(1+\rho^2)^{1/2}}$ ,  $\rho = \left(\frac{2\alpha_o}{4-\pi}\right)^{1/3}$  e  $\delta = \frac{\alpha}{\sqrt{(1-\alpha^2)}}$ .

Considerando esta reparametrização do modelo, a representação hierárquica para  $T_i \stackrel{iid}{\sim} \text{SNC}(\mu_o, \omega_o^2, \alpha_o)$  toma a seguinte forma

$$\left. \begin{aligned} T_i | U_i &\stackrel{iid}{\sim} \text{N}(\mu_o + \omega_o \rho (1 - b^{-1} U_i), \tau_o) \\ U_i &\stackrel{iid}{\sim} \text{N}(0, 1) I(u_i)_{(0, \infty)} \end{aligned} \right\}, \quad (5.18)$$

em que  $\tau_o = \omega_o^2 K_o$  e  $K_o = [b^2 + \rho^{2/3} (b^2 - 1)] b^{-2}$  e  $I(u)_A$  denota a função indicadora de  $u$ , que assume o valor 1 se  $u > 0$  e 0 caso contrário.

Assumimos como distribuições *a priori* para os parâmetros  $\mu_o$ ,  $\omega_o^2$  e  $\alpha_o$  as seguintes distribuições

$$\left. \begin{aligned} \mu_o &\sim \text{N}(m_o, V_o) \\ \omega_o^2 &\sim \text{GI}(r_{\omega_o}, s_{\omega_o}) \\ \alpha_o &\sim \text{Unif}(-d, d) \end{aligned} \right\}, \quad (5.19)$$

em que  $d = \frac{b^3(4-\pi)}{2(1-b^2)^{(3/2)}}$ .

Considerando estas especificações temos que as distribuições condicionais completas (dcc) *a posteriori* dos componentes de  $\boldsymbol{\psi} = (\mu_o, \omega_o^2, \alpha_o)$  são dadas na Proposição 12.

**Proposição 12.** *Assuma que os tempos de falha são tais que  $T_i \stackrel{iid}{\sim} \text{SNC}(\mu_o, \omega_o^2, \alpha_o)$ , em que  $\mu_o$ ,  $\omega_o^2$  e  $\alpha_o$  são dados na expressão (5.17), e considere as especificações a priori dadas em (5.19). Então*

*i A dcc a posteriori para cada componente  $t_i$  de  $\mathbf{t}_{\text{lat}}$  é dada por:*

$$t_i | \boldsymbol{\psi}^{[-t_i]}, \mathbf{D} \stackrel{d}{=} t_i | \boldsymbol{\psi}^{[-t_i]}, \ell_i = 0, c_i \sim \text{N}(\mu_o + \omega_o \rho (1 - b^{-1} u_i), \tau_o) I(t_i)_{(c_i, \infty)},$$

ii As dcc a posteriori para  $u_i$ ,  $\mu_o$ ,  $\omega_o^2$  e  $\alpha_o$  são dadas respectivamente por

$$\begin{aligned} u_i | \boldsymbol{\psi}^{[-u_i]}, \mathbf{D} &\sim N\left(\frac{b\omega_o\rho(\mu_o + \omega_o\rho - t_i)}{\omega_o^2\rho^2 + \tau_o b^2}, \frac{\tau_o b^2}{\omega_o^2\rho^2 + \tau_o b^2}\right) I(u_i)_{(0,\infty)}, \\ \mu_o | \boldsymbol{\psi}^{[-\mu_o]}, \mathbf{D} &\sim N\left(\frac{m_o\tau_o + V_o \left[ \sum_{i=1}^n t_i - \omega_o\rho \left( n - b^{-1} \sum_{i=1}^n u_i \right) \right]}{nV_o + \tau_o}, \frac{\tau_o}{nV_o + \tau_o}\right), \\ \omega_o^2 | \boldsymbol{\psi}^{[-\omega_o^2]}, \mathbf{D} &\propto \left(\frac{1}{\omega_o^2}\right)^{\frac{n}{2} + r_{\omega_o} + 1} \exp\left\{s_{\omega_o} - \frac{1}{\omega_o} (2K_o)^{-1} \sum_{i=1}^n [t_i - (\mu_o + \omega_o\rho(1 - b^{-1}u_i))]^2\right\}, \\ f(\alpha_o | \boldsymbol{\psi}^{[-\alpha_o]}, \mathbf{D}) &\propto \left[ \prod_{i=1}^n N(\mu_o + \omega_o\rho(1 - b^{-1}U_i), \tau_o) \right] I(\alpha_o)_{(-d,d)}, \end{aligned}$$

em que  $\tau_o = \omega_o^2 K_o$ ,  $K_o = [b^2 + \rho^{2/3}(b^2 - 1)]b^{-2}$ , e  $I(\alpha_o)_A$  denota a função indicadora que assume valor 1 se  $a \in A$  e valor 0 caso contrário.

A prova da Proposição 12 é apresentada no Anexo A.

Para va's estritamente positivas, que é sempre o caso de tempos de falha, pode-se assumir que  $T_i \stackrel{iid}{\sim} \text{LSN}(\mu, \omega, \alpha, \mathcal{H})$ . Consequentemente segue que  $\log(T_i) \stackrel{iid}{\sim} \text{SN}(\mu, \omega, \alpha, \mathcal{H})$ , cuja parametrização centrada é  $\log(T_i) \stackrel{iid}{\sim} \text{SNC}(\mu_o, \omega_o^2, \alpha_o)$ . Neste caso basta aplicarmos a Proposição 12 ao logaritmo dos tempos de falha para obtermos amostras de  $\mu_o$ ,  $\omega_o^2$  e  $\alpha_o$ .

Após alguns algebrismos, obtemos das igualdades na equação (5.17), que os parâmetros originais são recuperados conforme:

$$\left. \begin{aligned} \mu &= \mu_o - \omega_o \left( \frac{2\alpha_c}{4 - \pi} \right)^{1/3} \\ \omega^2 &= \omega_o^2 \left( 1 + \left( \frac{2\alpha_c}{4 - \pi} \right)^{2/3} \right) \\ \delta &= \frac{\left( \frac{2\alpha_o}{4 - \pi} \right)^{1/3}}{\sqrt{\frac{2}{\pi}} \sqrt{1 + \left( \frac{2\alpha_c}{4 - \pi} \right)^{2/3}}} \end{aligned} \right\}, \quad (5.20)$$



Nas próximas seções os modelos desenvolvidos são aplicados para a análise de dados artificiais e reais, visando ilustrar os procedimentos propostos.

### 5.3 Análise de dados artificiais de tempos de falha

Nesta seção o avaliaremos o desempenho do modelo SN Centrado para modelar tempos até a falha. Nossa meta é avaliar se há algum ganho em utilizarmos tal parametrização quando a assimetria real está na vizinhança de  $\alpha = 0$  e também para valores distantes de  $\alpha = 0$ . Para este fim apenas um banco de dados artificial é gerado em cada cenário e verifica-se a capacidade de recuperação dos valores reais dos parâmetros do modelo SN pelos ajustes usando as parametrizações centrada e não centrada.

Num segundo momento, uma vez que os modelos robustos são a principal contribuição deste trabalho, avalia-se as famílias de distribuições SMSN, SMLSN e a distribuição Weibul para modelar os tempos de falha de um conjunto de dados artificiais gerados a partir de uma distribuição LST.

Nos estudos realizados nas próximas seções foram utilizadas as mesmas distribuições *a priori* vagas consideradas nos modelos de degradação SMSN e SMLSN, a saber

$$\left. \begin{array}{ll} \sigma_{\varepsilon}^2 \sim \text{GI}(10^{-3}, 10^3) & \mu \sim \text{N}(0, 10^6) \\ \Delta \sim \text{N}(0, 10^6) & \tau \sim \text{GI}(10^{-3}, 10^3) \\ \varsigma \sim \text{Gama}(10^{-3}, 10^{-3}) & \kappa \sim \text{Gama}(10^{-3}, 10^{-3}); \end{array} \right\} \quad (5.21)$$

em que se  $X \sim \text{Gamma}(a, b)$ ,  $E(x) = ab^{-1}$  para  $a > 1$  e  $V(x) = ab^{-2}$ ; e se  $X \sim \text{GI}(a, b)$ ,  $E(x) = b(a-1)^{-1}$  para  $a > 1$  e  $V(x) = b(a-1)^{-2}(a-2)^{-1}$  para  $a > 2$ . Para  $\nu$  consideramos a distribuição *a priori* Jeffreys Independente proposta por Fonseca *et al.* (2008) (ver equação (5.14)).

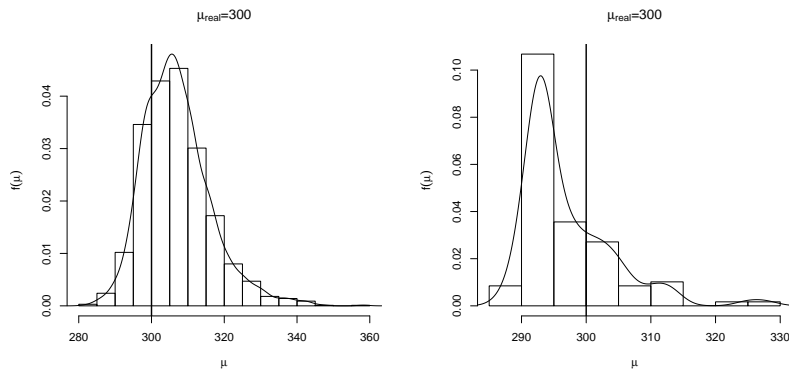
Para o modelo SNC assumimos que, *a priori*,  $\mu_o \sim \text{N}(0, 10^6)$ ,  $\omega_o^2 \sim \text{GI}(10^{-3}, 10^3)$  e  $\alpha_o \sim \text{Unif}(-0,995272; 0,995272)$ , as quais também são vagas. A análise dos dados foi efetuada utilizando-se os softwares WinBUGS 1.4 (Spiegelhalter *et al.*, 2000) e R 3.1.0 (R Development Core Team, 2005). A fim de garantirmos a convergência das

cadeias e também minimizar a autocorrelação das amostras obtidas, foram realizadas 101.000.000 iterações da simulação MCMC, com um período de aquecimento (*burn-in*) de 1.000.000, e saltos de 50.000. O resultado foi uma amostra *a posteriori* final de tamanho  $s = 2000$ . Estas especificações garantiram a convergência nos modelos ajustados.

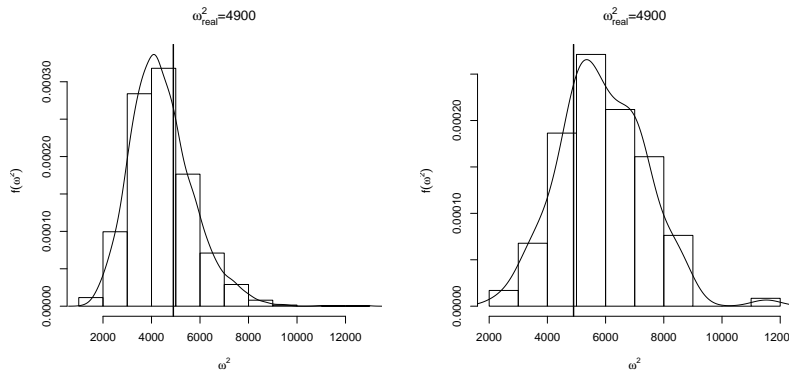
### 5.3.1 Dados artificiais: skew normal centrada *versus* skew normal não centrada

Como a intenção nesta seção é avaliarmos a capacidade de recuperação dos valores reais da distribuição dos tempos de falha através dos ajustes pelas distribuições SN Centrada e SN Não Centrada, optamos por gerar dois bancos de dados de tempo de falha com  $n = 50$  itens sob teste. Para o primeiro banco de dados (BT-SN1), a distribuição dos tempos de falha gerados é  $T_i \stackrel{iid}{\sim} \text{SN}(\mu = 300, \omega^2 = 70^2, \alpha = 10)$ . Para o segundo banco de dados (BT-SN2), a distribuição dos tempos de falha gerados é  $T_i \stackrel{iid}{\sim} \text{SN}(\mu = 300, \omega^2 = 70^2, \alpha = 0,1)$ . Assume-se que o tempo de estudo é de 440 unidades de medida (*u.m.*). Os bancos de dados BT-SN1 e BT-SN2 apresentam, respectivamente, 4% e 2% de censura. Ajustamos a estes dois conjuntos de dados os modelos SN Centrado e SN não centrado.

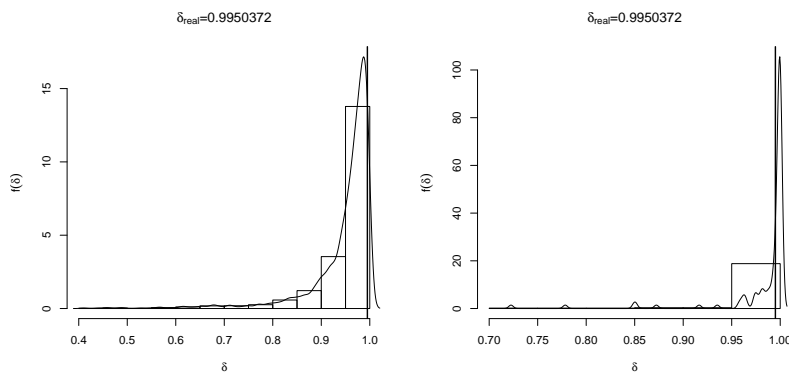
Dos gráficos na Figura 5.1 percebe-se que, independentemente da parametrização utilizada, para o banco de dados BT-SN1, ambos modelos forneceram estimativas para os parâmetros muito próximas dos reais valores utilizados para gerar os dados artificiais. Em todos os casos a distribuição *a posteriori* de cada parâmetro tem maior massa probabilística concentrada no entorno do valor real.



(a) SN Centrada: histograma para  $\mu$ : (b) SN Não Centrada: histograma para  $\mu$ : valor real  $\mu = 300$  na reta vertical es- para  $\mu$ : valor real  $\mu = 300$  na reta vertical espessa



(c) SN Centrada: histograma para  $\omega^2$ : valor real  $\omega^2 = 70^2$  na reta verti- para  $\omega^2$ : valor real  $\omega^2 = 70^2$  na reta cal espessa vertical espessa



(e) SN Centrada: histograma para  $\delta$ : (f) SN Não Centrada: histograma para  $\delta$ : valor real  $\delta = 0,9950372$  na reta ver- para  $\delta$ : valor real  $\delta = 0,9950372$  na reta vertical espessa tical espessa

Figura 5.1: Distribuições *a posteriori* dos parâmetros sob os modelos SN centrado (esquerda) e SN não centrado (direita), dados BT-SN1.

Dos gráficos na Figura 5.2 percebemos que, para o banco de dados BT-SN2 com assimetria nas proximidades de  $\alpha = 0$ , apenas o modelo com parametrização centrada fornece boas estimativas para os parâmetros, ou seja, para bancos de dados com pouco ou nenhum grau de assimetria o modelo SNC parece fornecer melhores estimativas.

Vale ressaltar que a implementação do modelo usando a parametrização centrada da distribuição SN é mais trabalhosa se utilizarmos o software WinBUGS 1.4 (Spiegelhalter *et al.*, 2000). Além disto o tempo computacional demandado é superior ao observado quando utilizamos o modelo SN com parametrização não centrada, o qual é de mais simples implementação. Outra vantagem de utilizarmos a parametrização não centrada é que os procedimentos de amostragem via MCMC são consequência imediata da Proposição 10, uma vez que a distribuição SN é um caso particular da SMSN.

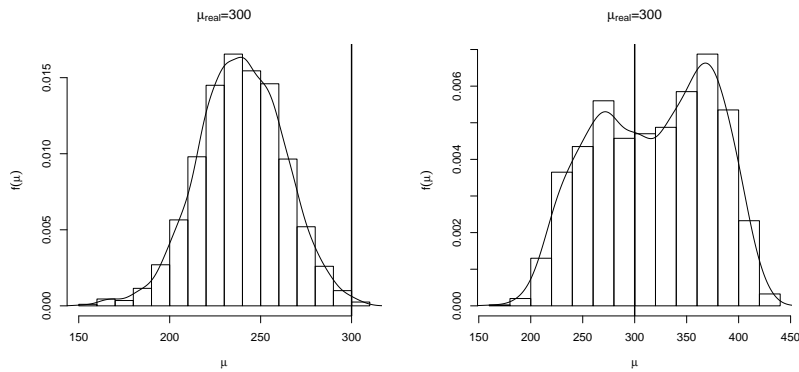
Na próxima seção analisaremos um banco de dados artificial com elevado grau de assimetria, razão pela qual consideramos a parametrização não centrada para a distribuição SN.

### 5.3.2 Dados artificiais log skew-t

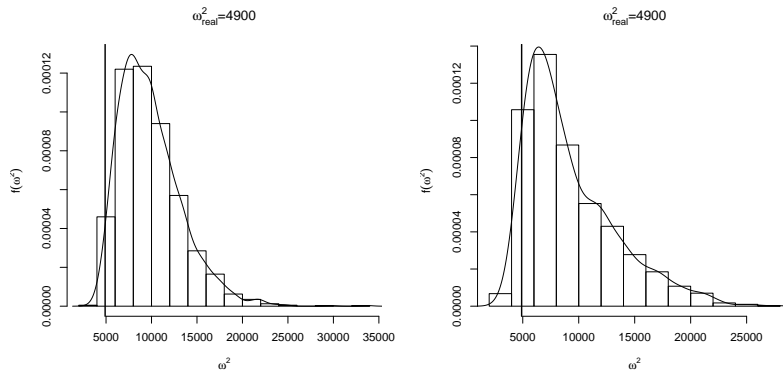
Nesta seção geramos de  $n = 50$  tempos de falha tal que  $T_i \stackrel{iid}{\sim} LST(\mu = 5; \omega^2 = 0,3^2; \alpha = 10; \nu = 2,1)$ . Assumimos que o tempo de estudo é de 300 unidades de medida (*u.m.*), de modo que o banco de dados apresenta 10% de censura.

Para analisar os dados ajustaremos os modelos nas famílias SMSN e SMLSN e avaliaremos, por exemplo, estimativas para a função fração de falha, função de sobrevivência (confiabilidade) e alguns percentis de interesse. Além destes modelos também ajustaremos o modelo weibull (W), o qual é amplamente utilizado na literatura para modelar o comportamento do tempos até a falha.

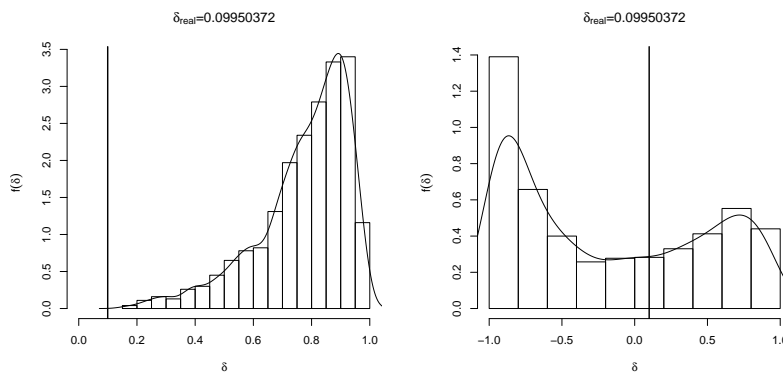
Da Figura 5.3 notamos que as distribuições *a posteriori* dos parâmetros colocam maior massa de probabilidade sempre no entorno dos verdadeiros valores que geraram os tempos de falha.



(a) SN Centrada: histograma para  $\mu$ : (b) SN Não Centrada: histograma para valor real  $\mu = 300$  na reta vertical es- para  $\mu$ : valor real  $\mu = 300$  na reta vertical espessa

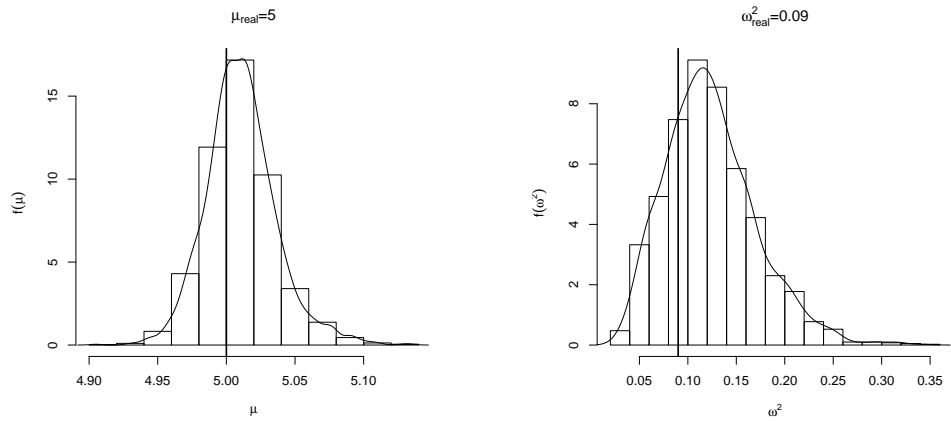


(c) SN Centrada: histograma para  $\omega^2$ : valor real  $\omega^2 = 70^2$  na reta verti- para  $\omega^2$ : valor real  $\omega^2 = 70^2$  na reta cal espessa vertical espessa



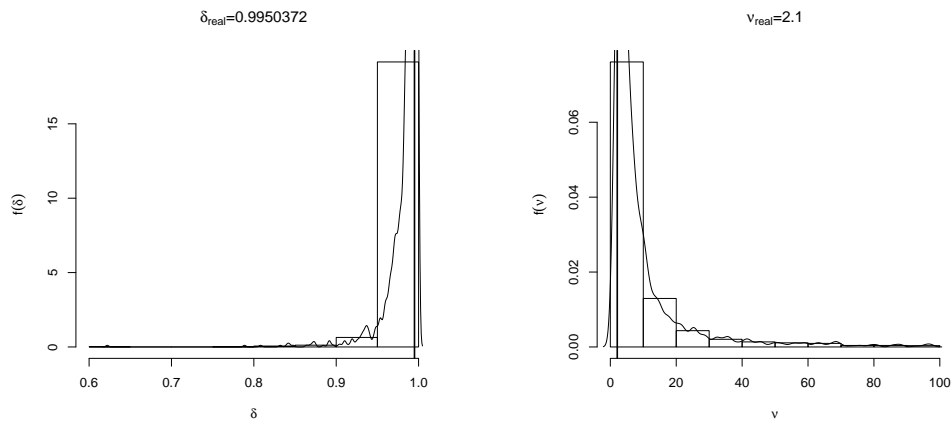
(e) SN Centrada: histograma para  $\delta$ : (f) SN Não Centrada: histograma para valor real  $\delta = 0,09950372$  na reta ver- para  $\delta = 0,09950372$  na reta vertical espessa

Figura 5.2: Distribuições *a posteriori* dos parâmetros sob os modelos SN centrado (esquerda) e SN não centrado (direita), dados BT-SN2.



(a) Histograma para  $\mu$ : valor real  $\mu = 5$  na reta vertical espessa

(b) Histograma para  $\omega^2$ : valor real  $\omega^2 = 0,3^2$  na reta vertical espessa



(c) Histograma para  $\delta$ : valor real  $\delta = 0,9950$  na reta vertical espessa

(d) Histograma para  $\nu$ : valor real  $\nu = 2,1$  na reta vertical espessa

Figura 5.3: Comparação da distribuição *a posteriori* dos parâmetros com os valores reais no ajuste LST: dados simulados log skew-t

As Figuras 5.4 e 5.5 fornecem uma comparação da função de confiabilidade empírica obtida via Kaplan-Meyer a obtida ajustando os modelos propostos para os tempos de falha observados, as quais foram obtidas usando respectivamente o método de Hamada (2005) e distribuição preditiva *a posteriori*. Quando a maioria dos pontos se aproxima da reta  $x = y$ , maior a evidência de que o modelo está bem ajustado. Considerando este critério o modelo LST é o mais adequado para modelar os tempos de falha. Outros dois modelos que apresentaram bons ajustes foram os modelos ST e LSN.

Na Figura 5.6 apresentamos as funções de sobrevivência real, empírica obtida via Kaplan-Meyer, e distribuição preditiva *a posteriori* sob os modelo LST, ST, LSN e W, para os tempos de falha de uma unidade futura da população. Notamos que a estimativa fornecida pelo modelo LST é a que mais se aproxima das funções de sobrevivência empírica fornecida pelo estimador de Kaplan-Meyer, e da real, seguidas pelas obtidas ajustando os modelos ST e LSN. O modelo W, bastante usado em situações práticas, não consegue captar o comportamento dos dados gerados. Observe que o ajuste W subestima a confiabilidade de tempos de falha menores e maiores, além de superestimar os valores centrais próximos à mediana dos tempos até a falha. isto significa que, para este banco de dados com características de assimetria e caudas pesadas, a escolha do modelo W levaria a inferências errôneas sobre a confiabilidade para a grande maioria dos tempos de falha nas proximidades de ocorrências dos mesmos.

A Tabela 5.1 fornece os percentis de ordem 5%, 10%, 25%, 50%, 75%, 90% para os tempos de falha, aqueles tempos para os quais confiabilidade é, respectivamente, 0,95, 0,90, 0,75, 0,50, 0,25 e 0,10. A partir da tabela notamos que o modelo LST fornece as estimativas mais próximas da distribuição empírica obtida via Kaplan-Meyer. Portanto, a melhor inferência realizada foi aquela que correspondia ao modelo que originou os dados de tempo de falha.

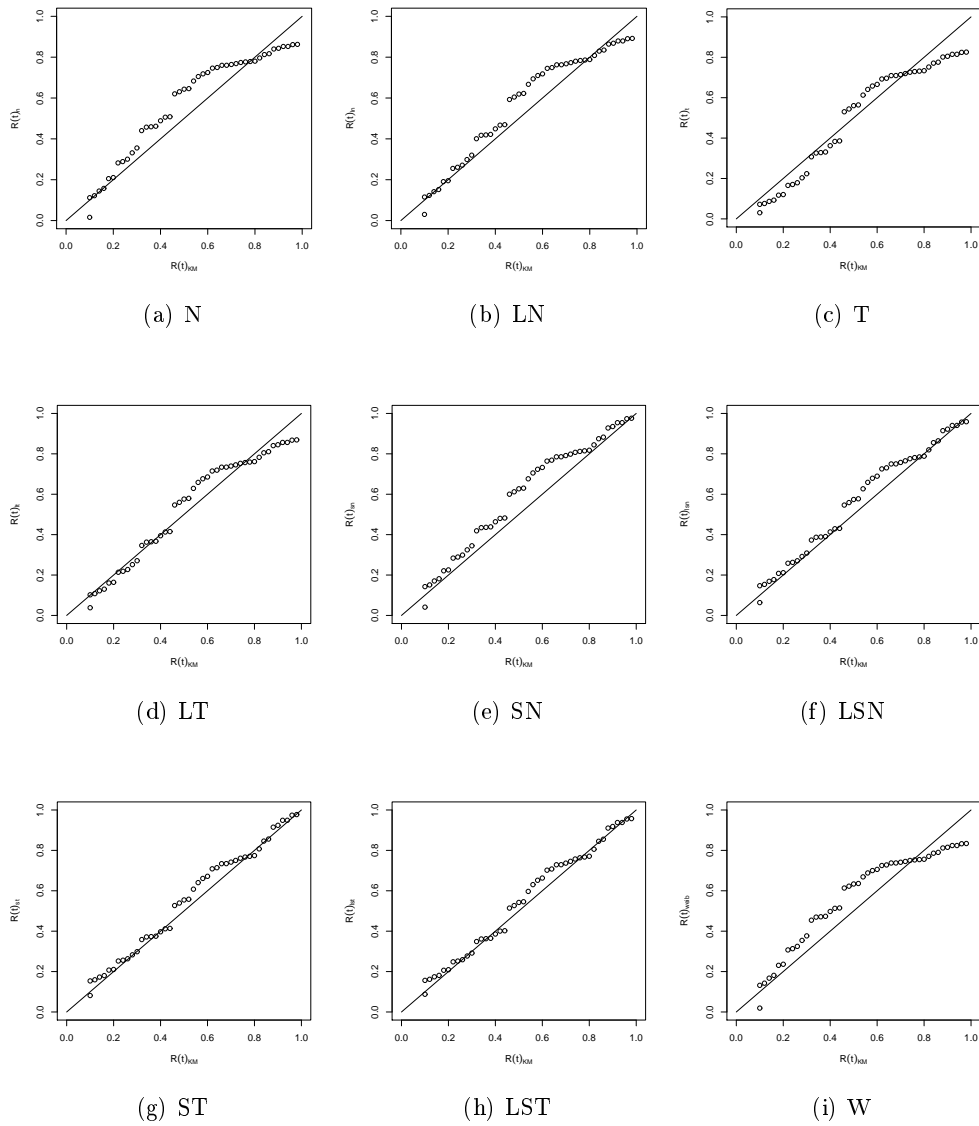


Figura 5.4: Confiabilidade empírica via Kaplan-Meier (abscissa) versus modelos ajustados pela média *a posteriori* (ordenada) para os tempos de falha. N, LN e T (topo). Tempos LST.



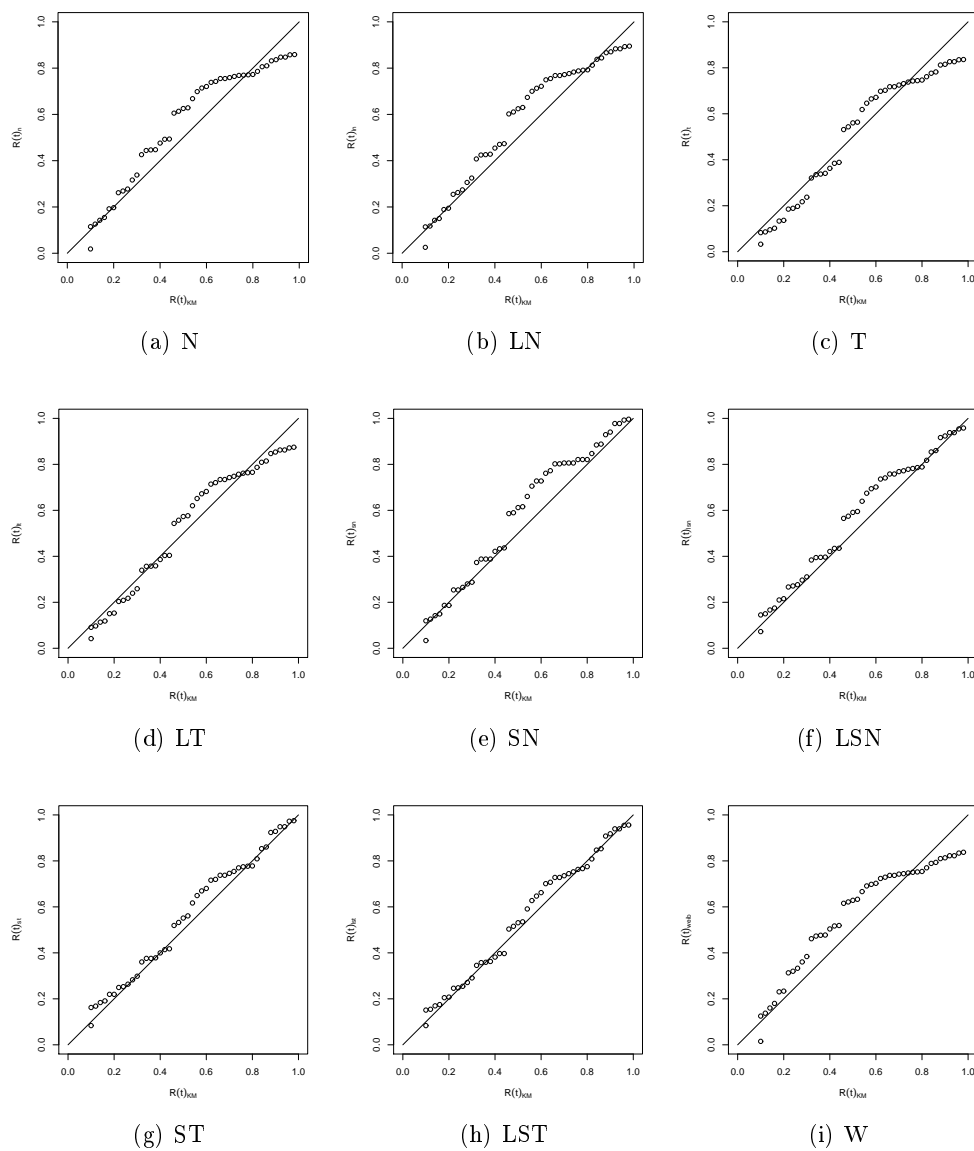


Figura 5.5: Confiabilidade empírica via Kaplan-Meier (abscissa) versus modelos ajustados pela distribuição preditiva *a posteriori* (ordenada) para os tempos de falha. N, LN e T (topo). Tempos LST.

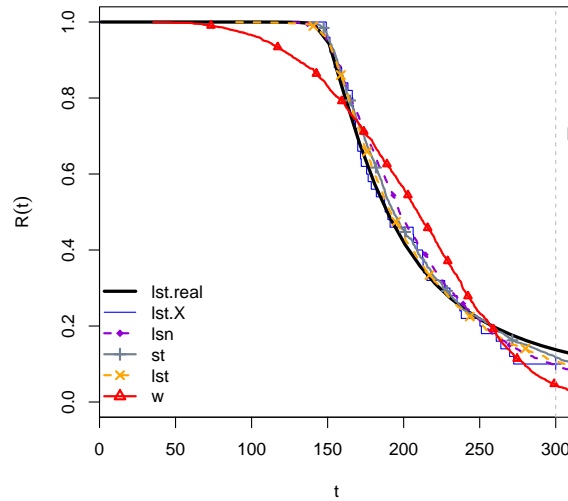


Figura 5.6: Confiabilidade Real, empírica via Kaplan-Meier, e distribuição preditiva *a posteriori* sob os modelos LSN, ST, LST e W para os tempos de falha de uma unidade futura da população

## 5.4 Dados de câncer: análise de tempo de vida

Como situação prática escolhemos o banco de dados de um tratamento experimental aplicado a 137 pacientes com um tipo câncer de pulmão inoperável, originalmente apresentado por Kalbfleisch e Prentice (2002), e o mesmo analisado por Vallejos e Steel (2015). Neste estudo selecionaram-se aleatoriamente os pacientes que receberiam um novo tratamento quimioterápico ou o tratamento padrão, em que o banco de dados conta ainda com outras covariáveis. O tempo de vida dos pacientes foi medido em dias.

As distribuições *a priori* para os parâmetros e variáveis latentes dos modelos considerados para analisar os dados de tempo de vida, a saber SMSN, SMLSN e W, são os mesmos considerados na análise de dados artificiais da Seção 5.3.

Tabela 5.1: Percentis de ordem 5%, 10%, 25%, 50%, 75%, 90% para os tempos de falha reais simulados do modelo LST, as estimativas segundo o estimador de Kaplan-Meyer, distribuição preditiva *a posteriori* e média *a posteriori*

Modelo	$t_p$	$R(t_p)$	Real	KM	Pred. Post	Média	Modelo	$t_p$	$R(t_p)$	Real	KM	Pred. Post.	Média
N	$t_{0,05}$	0,95	150,44	118,39	119,98		LN	$t_{0,05}$	0,95	150,44	133,88	135,34	
	$t_{0,10}$	0,90	153,59	136,69	139,28			$t_{0,10}$	0,90	153,59	148,56	147,92	
	$t_{0,25}$	0,75	167,11	170,58	171,53			$t_{0,25}$	0,75	167,11	172,09	171,65	
	$t_{0,50}$	0,50	188,56	206,00	207,35			$t_{0,50}$	0,50	188,56	203,77	202,56	
	$t_{0,75}$	0,25	236,00	239,60	243,18			$t_{0,75}$	0,25	236,00	238,65	239,13	
	$t_{0,90}$	0,10	272,36	278,88	275,43			$t_{0,90}$	0,10	272,36	276,75	277,74	
T	$t_{0,05}$	0,95	150,44	91,13	85,23		LT	$t_{0,05}$	0,95	150,44	129,48	125,74	
	$t_{0,10}$	0,90	153,59	129,26	124,43			$t_{0,10}$	0,90	153,59	143,44	142,39	
	$t_{0,25}$	0,75	167,11	165,40	163,27			$t_{0,25}$	0,75	167,11	167,96	168,09	
	$t_{0,50}$	0,50	188,56	194,20	193,72			$t_{0,50}$	0,50	188,56	195,59	196,81	
	$t_{0,75}$	0,25	236,00	225,30	224,17			$t_{0,75}$	0,25	236,00	228,15	230,54	
	$t_{0,90}$	0,10	272,36	261,42	263,01			$t_{0,90}$	0,10	272,36	268,27	272,61	
SN	$t_{0,05}$	0,95	150,44	152,54	151,99		LSN	$t_{0,05}$	0,95	150,44	150,05	150,54	
	$t_{0,10}$	0,90	153,59	156,57	157,44			$t_{0,10}$	0,90	153,59	155,87	155,98	
	$t_{0,25}$	0,75	167,11	173,33	173,99			$t_{0,25}$	0,75	167,11	170,72	169,96	
	$t_{0,50}$	0,50	188,56	199,10	204,47			$t_{0,50}$	0,50	188,56	197,35	197,30	
	$t_{0,75}$	0,25	236,00	238,78	245,23			$t_{0,75}$	0,25	236,00	240,81	240,75	
	$t_{0,90}$	0,10	272,36	278,57	287,58			$t_{0,90}$	0,10	272,36	297,44	296,23	
ST	$t_{0,05}$	0,95	150,44	151,34	151,50		LST	$t_{0,05}$	0,95	150,35	150,44	<b>150,20</b>	<b>150,27</b>
	$t_{0,10}$	0,90	153,59	156,37	155,82			$t_{0,10}$	0,90	154,36	153,59	<b>155,23</b>	<b>155,54</b>
	$t_{0,25}$	0,75	167,11	168,70	168,95			$t_{0,25}$	0,75	165,38	167,11	<b>167,80</b>	<b>168,30</b>
	$t_{0,50}$	0,50	188,56	193,50	195,04			$t_{0,50}$	0,50	189,14	188,56	<b>191,52</b>	<b>193,39</b>
	$t_{0,75}$	0,25	236,00	237,60	239,33			$t_{0,75}$	0,25	238,07	236,00	<b>235,92</b>	<b>238,54</b>
	$t_{0,90}$	0,10	272,36	311,13	313,46			$t_{0,90}$	0,10	346,58	272,36	<b>306,19</b>	<b>319,09</b>
W	$t_{0,05}$	0,95	150,44	109,70	109,57		W	$t_{0,05}$	0,95	150,44	109,70	109,57	
	$t_{0,10}$	0,90	153,59	130,65	130,81			$t_{0,10}$	0,90	153,59	130,65	130,81	
	$t_{0,25}$	0,75	167,11	166,98	167,63			$t_{0,25}$	0,75	167,11	166,98	167,63	
	$t_{0,50}$	0,50	188,56	209,20	208,43			$t_{0,50}$	0,50	188,56	209,20	208,43	
	$t_{0,75}$	0,25	236,00	247,00	247,58			$t_{0,75}$	0,25	236,00	247,00	247,58	
	$t_{0,90}$	0,10	272,36	279,51	280,91			$t_{0,90}$	0,10	272,36	279,51	280,91	

As Figuras 5.7 e 5.8 fornecem as funções de sobrevivência empírica estimada via Kaplan-Meyer e obtidas sobre os modelos ajustados para os tempos de vida observados, respectivamente via método de Hamada (2005) e distribuição preditiva *a posteriori*. O melhor ajuste é aquele cuja maioria dos pontos se aproxima da reta  $x = y$ . Segundo este critério os modelos LST e LSN mostram-se adequados para modelar os tempos de vida.

Na Figura 5.9 apresentamos a função de sobrevivência empírica via Kaplan-Meyer, e distribuição preditiva *a posteriori* sob os modelos LSN, LST e W, para os tempos de vida de um futuro paciente da população. Percebemos que os modelos LSN e LST fornecem estimativas para a função de sobrevivência muito próximas das fornecidas pelo estimador de Kaplan-Meyer. Sob o modelo W percebemos que as estimativas da sobrevivências em tempos de vida mais centrais e também na cauda direita são piores que nos modelos LSN e LST.

A Tabela 5.2 fornece os percentis de ordem 5%, 10%, 25%, 50%, 75%, 90% para os tempos de vida, que são os mesmos tempos para os quais a probabilidade de sobrevivência dos pacientes é, respectivamente, 0,95, 0,90, 0,75, 0,50, 0,75 e 0,10. Os modelos LST (negrito) e LSN (sublinhado) fornecem as estimativas mais próximas da distribuição empírica obtida via Kaplan-Meyer.

Em resumo, os modelos LST e LSN são os dois que melhor se adequam para modelar os tempos de vida da base de dados de pacientes de câncer pulmonar, indicando que os modelos propostos mostram-se úteis para modelar tempos de vida. A flexibilidade trazida pela modelagem da assimetria em conjunto com características de caudas pesadas dos modelos SMSN e SMLSN poderá ser útil em diversas situações práticas.

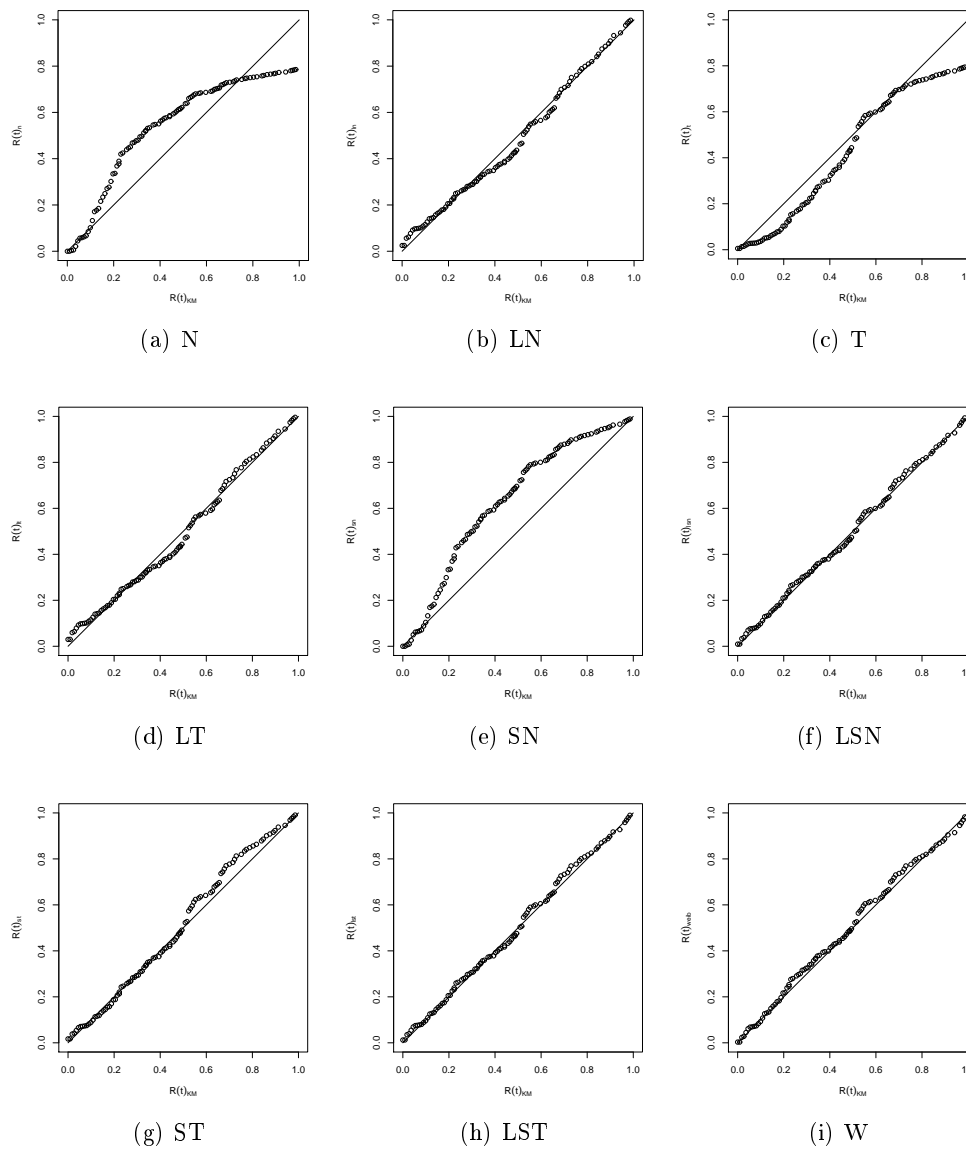


Figura 5.7: Sobrevivência empírica via Kaplan-Meier (abscissa) versus modelos ajustados pela média *a posteriori* (ordenada) para os tempos de vida. Dados de câncer de pulmão.

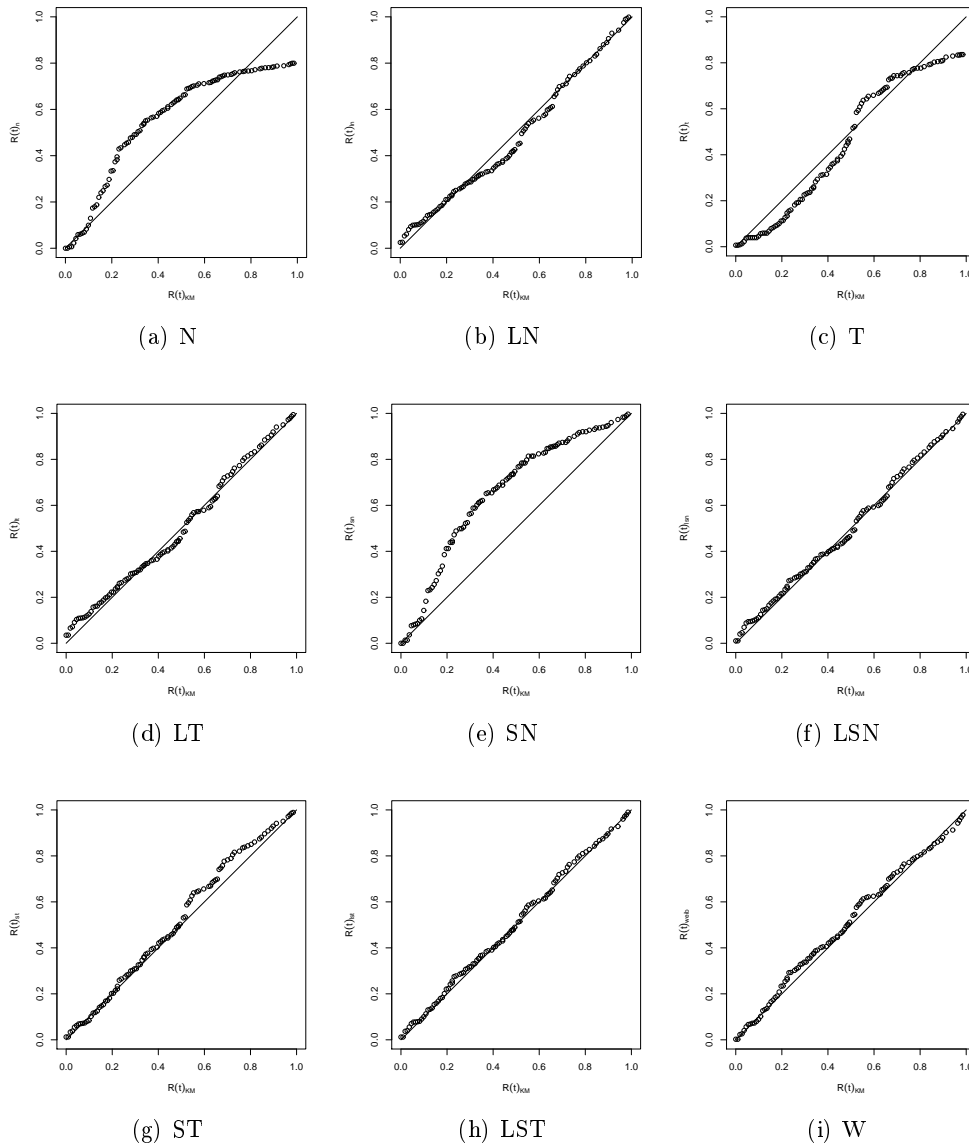


Figura 5.8: Sobrevivência empírica via Kaplan-Meier (abscissa) versus modelos ajustados pela distribuição preditiva *a posteriori* (ordenada) para os tempos de vida. Dados de câncer de pulmão.

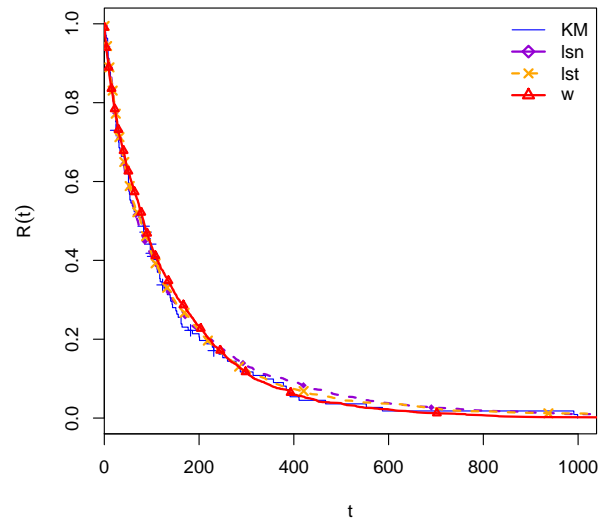


Figura 5.9: Sobrevivência empírica via Kaplan-Meier, e distribuição preditiva *a posteriori* sob os modelos LSN, LST e W, para os tempos de falha de uma unidade futura da população

O próximo passo para estudos futuros será incorporar preditores lineares aos parâmetros dos modelos de maneira a robustecer a estimativa dos coeficientes, assim como fizeram Barros *et al.* (2008); Cancho *et al.* (2010) e Vallejos e Steel (2015). Outra extensão subsequente é a necessidade de construirmos procedimentos para análise de diagnóstico de influência assim como em Cancho *et al.* (2010), os quais não são trivialmente obtidos quando lidamos com dados censurados.

Tabela 5.2: Percentis de ordem 5%, 10%, 25%, 50%, 75%, 90% para os pacientes de câncer de pulmão inoperável, segundo o estimador de Kaplan-Meyer, distribuição preditiva *a posteriori* e média *a posteriori*

Modelo	$t_p$	$R(t_p)$	KM	Pred. Post.	Média	Modelo	$t_p$	$R(t_p)$	KM	Pred. Post.	Média
N	$t_{0,05}$	0,95	5,85	-135,31	-138,98	LN	$t_{0,05}$	0,95	5,85	6,41	6,65
	$t_{0,10}$	0,90	9,70	-84,23	-79,39		$t_{0,10}$	0,90	9,70	10,26	10,97
	$t_{0,25}$	0,75	24,08	29,49	20,18		$t_{0,25}$	0,75	24,08	24,06	25,37
	$t_{0,50}$	0,50	77,15	134,60	130,81		$t_{0,50}$	0,50	77,15	62,18	64,62
	$t_{0,75}$	0,25	157,95	240,68	241,44		$t_{0,75}$	0,25	157,95	161,94	165,14
	$t_{0,90}$	0,10	337,26	339,12	341,01		$t_{0,90}$	0,10	337,26	391,19	385,46
T	$t_{0,05}$	0,95	5,85	-105,04	-149,92	LT	$t_{0,05}$	0,95	5,85	7,01	6,73
	$t_{0,10}$	0,90	9,70	-42,46	-62,82		$t_{0,10}$	0,90	9,70	11,43	11,68
	$t_{0,25}$	0,75	24,08	28,33	15,95		$t_{0,25}$	0,75	24,08	26,66	27,32
	$t_{0,50}$	0,50	77,15	75,01	70,03		$t_{0,50}$	0,50	77,15	68,34	66,78
	$t_{0,75}$	0,25	157,95	126,40	124,10		$t_{0,75}$	0,25	157,95	174,21	163,96
	$t_{0,90}$	0,10	337,26	217,20	202,87		$t_{0,90}$	0,10	337,26	424,24	387,85
SN	$t_{0,05}$	0,95	5,85	9,94	11,22	LSN	$t_{0,05}$	0,95	5,85	5,68	5,26
	$t_{0,10}$	0,90	9,70	24,15	24,43		$t_{0,10}$	0,90	9,70	10,57	9,88
	$t_{0,25}$	0,75	24,08	77,69	64,91		$t_{0,25}$	0,75	24,08	26,24	27,02
	$t_{0,50}$	0,50	77,15	152,10	139,58		$t_{0,50}$	0,50	77,15	70,88	73,73
	$t_{0,75}$	0,25	157,95	264,00	239,42		$t_{0,75}$	0,25	157,95	174,78	173,04
	$t_{0,90}$	0,10	337,26	365,50	343,17		$t_{0,90}$	0,10	337,26	378,49	337,93
ST	$t_{0,05}$	0,95	5,85	7,08	6,55	LST	$t_{0,05}$	0,95	5,85	<b>5,15</b>	<b>5,04</b>
	$t_{0,10}$	0,90	9,70	12,66	13,32		$t_{0,10}$	0,90	9,70	<b>9,71</b>	<b>10,00</b>
	$t_{0,25}$	0,75	24,08	34,52	34,45		$t_{0,25}$	0,75	24,08	<b>27,49</b>	<b>27,95</b>
	$t_{0,50}$	0,50	77,15	80,91	78,28		$t_{0,50}$	0,50	77,15	<b>76,67</b>	<b>74,38</b>
	$t_{0,75}$	0,25	157,95	168,50	160,45		$t_{0,75}$	0,25	157,95	<b>180,32</b>	<b>170,62</b>
	$t_{0,90}$	0,10	337,26	317,40	317,69		$t_{0,90}$	0,10	337,26	<b>340,36</b>	<b>335,47</b>
W	$t_{0,05}$	0,95	5,85			W	$t_{0,05}$	0,95	5,85	3,41	3,79
	$t_{0,10}$	0,90	9,70				$t_{0,10}$	0,90	9,70	8,09	8,75
	$t_{0,25}$	0,75	24,08				$t_{0,25}$	0,75	24,08	27,23	28,28
	$t_{0,50}$	0,50	77,15				$t_{0,50}$	0,50	77,15	82,72	79,35
	$t_{0,75}$	0,25	157,95				$t_{0,75}$	0,25	157,95	188,38	179,53
	$t_{0,90}$	0,10	337,26				$t_{0,90}$	0,10	337,26	316,90	326,99



## Capítulo 6

# Conclusões

O objetivo desta tese residiu em construir modelos gerais de degradação e tempo de falha que fossem capazes de capturar características de assimetria e caudas pesadas. Iniciamos o trabalho apresentando uma revisão da literatura em modelos gerais de degradação, apontamos algumas de suas virtudes e limitações. Discorremos sobre a abordagem Bayesiana para inferir sobre modelos de degradação. Fornecemos provas formais para alguns dos procedimentos inferenciais vigentes, assim como uma contra-prova para procedimentos inferenciais sobre os percentis do tempo até a falha propostos por Hamada (2005), de maneira que ficou evidente que o cálculo dos percentis e a interpretação de intervalos de credibilidade das quantidades de interesse apresentam falhas metodológicas, exceto para a proposta em Oliveira (2011). Introduziu-se também a técnica de aumento de dados para possibilitar a generalização dos procedimentos propostos em Oliveira (2011).

Num segundo momento, fizemos uma breve revisão sobre a família de distribuições mistura de escala da distribuição Skew Normal (Branco e Dey, 2001) e obtivemos uma nova construção para sua versão logarítima (Marchenko e Genton, 2010), a partir de uma mistura na escala de distribuições Log Skew Normal extendendo a família considerada por Vallejos e Steel (2015), que consideram a família mistura na forma da

distribuição Log Normal (caso particular da família SMLSN).

Posteriormente, introduzimos modelos robustos para perfis lineares de degradação considerando duas formas funcionais para o modelo. Numa delas os efeitos aleatórios que capturavam as diferenças entre perfis de degradação representavam a taxa de degradação. Na outra estes efeitos aleatórios representavam a taxa recíproca de degradação. Para ambos os casos assumimos que a distribuição dos efeitos aleatórios seguia distribuições nas famílias SMSN e SMLSN. Aplicamos o método analítico para obtermos a distribuição dos tempos de falha e observamos que, em vários casos, a propriedade de fechamento entre as distribuições dos efeitos aleatórios e tempo de falha era observada. Vale ressaltar que os resultados obtidos são interessantes também do ponto de vista teórico pois, até onde sabemos, alguns não haviam sido explorados na literatura, como por exemplo a transformação linear de variáveis com distribuição nas famílias SMLSN e transformações recíprocas de variáveis SMSN e LSMSN. A fim de obtermos inferências sobre os tempos de falha nos novos modelos robustos de degradação, foram construídos algoritmos MCMC para geração de amostras da distribuição *a posteriori* dos parâmetros dos modelos. Para a elaboração destes, a técnica de aumento de dados foi utilizada, permitindo-nos obter distribuições condicionais completas *a posteriori* de forma fechada na maioria dos casos.

Os novos modelos gerais de degradação foram então aplicados a bancos de dados simulados, bancos de dados de rodas de trem e bancos de dados de emissores de laser. Observamos que a má especificação dos efeitos aleatórios dos modelos de degradação provoca inferências piores, para os percentis e fração de falha, quando assumidos os modelos N, LN e W, modelos amplamente utilizados na literatura. Noutras palavras, os modelos robustos propostos neste texto tendem a realizar inferência razoável inclusive quando a distribuição dos dados não é proveniente de distribuições assimétricas ou de caudas pesadas.

Vários critérios de seleção de modelos e avaliação de qualidade de ajuste foram

explorados. Sempre que possível, estes critérios foram calculados usando dados de degradação e pseudo tempos de falha que são, na prática, as informações disponíveis. Percebemos que os modelos robustos são úteis para a modelagem de dados de degradação, entretanto notamos que a avaliação de qualidade de ajuste e os critérios de seleção ainda fornecem evidências pobres na direção de encontrar o modelo mais adequado. Os critérios LMPL e DIC baseados nas medidas individuais de degradação tenderam a selecionar modelos de caudas pesadas quando os dados eram provenientes destes, já os critérios baseados em pseudo tempos de falha tenderam a selecionar corretamente modelos advindos de distribuições com caudas leves. As medidas *CPO* e *KL* não ofereceram decisões conclusivas que permitiram recomendar o seu uso na seleção de modelos de degradação da maneira como foram aqui abordadas.

Na segunda parte do trabalho construímos modelos de análise de tempo de falha (vida) convencional que também se aproveitam da flexibilidade e robustez oferecidas pelas famílias de distribuições SMSN e SMLSN. A técnica de aumento de dados foi utilizada em duas direções, uma delas para nos permitir construir a função de verossimilhança (que envolve censura) de forma mais tratável e a outra para elaboração dos procedimentos de amostragem para a distribuição *a posteriori* dos parâmetros dos novos modelos. Os modelos propostos foram ajustados para modelarmos os tempos de falha para bancos de dados simulados, e para analisar os tempos de vida de um banco de dados de pacientes com câncer de pulmão. Os modelos robustos para tempos de falha foram adequados para modelar os bancos de dados com características de assimetria e caudas pesadas. Para ajustar o modelo SN a dados com características de caudas leves e pequeno grau de assimetria recomendamos que seja utilizada sua parametrização centrada. Critérios numéricos que levem em conta presença de censura para a seleção dos modelos e avaliação de qualidade de ajuste ainda carecem ser elaborados.

Amostras da distribuição *a posteriori* dos parâmetros para análise dos bancos de dados simulados e reais, tanto para de degradação quanto para tempos de falha, foram

obtidas com uso do software WinBUGS 1.4 (Spiegelhalter *et al.*, 2000). As implementações foram disponibilizadas no apêndice deste texto para uso em trabalhos futuros e aplicação imediata dos modelos propostos. Algumas distribuições implementadas como, por exemplo, a distribuição *a priori* de Fonseca *et al.* (2008) para os graus de liberdade das distribuições de caudas pesadas, poderá servir inclusive a outros contextos que não somente exclusivos das famílias de distribuições SMSN e SMLSN. Apesar dos ganhos da simplicidade ao trabalharmos com as amostragens via WinBUGS, pontuamos que necessitamos de descartar muitas amostras para alcançar a convergência dos parâmetros, e dar saltos muito grandes para evitarmos amostras autocorrelacionadas dos mesmos. Por exemplo para a degradação fizemos 220.000.000 iterações, com *burn-in* de 100.000.000, e saltos de 60.000, para conseguir uma amostra não autocorrelacionada da distribuição *a posteriori* dos parâmetros de tamanho  $s = 2.000$ , o que demandou sempre um elevado tempo computacional.

Diante das conclusões apresentadas podemos vislumbrar como próximos passos deste trabalho

- o estudo das propriedades da família de distribuições de mistura de escala da distribuição Skew Normal Recíproca (SMRSN), obtida como resultado da Proposição 6;
- a aplicação do método proposto por Oliveira (2011) para modelos gerais de degradação com formas funcionais não lineares e que contenham mais de um parâmetro de efeitos fixos e aleatórios;
- a elaboração métodos mais efetivos para seleção de modelos e avaliação qualidade de ajuste dos modelos de degradação sob a abordagem bayesiana. Para esta tarefa devemos ter em mente que poucos bancos de dados terão disponíveis uma quantidade grande de perfis de degradação;

- a proposição de distribuições ainda mais flexíveis para modelar perfis de degradação incluindo, por exemplo, características de multimodalidade;
- a modelagem de dados de degradação multivariados que possibilitem determinarmos a distribuição conjunta do tempo de falha das rodas, a fim de otimizarmos o tempo de parada ideal do sistema;
- a estimação do limiar de falha  $D_f$  do modelo de degradação em lugar de sua estipulação determinística;
- a introdução de covariáveis para modelar a degradação;
- a elaboração de métodos para a seleção e avaliação de qualidade de ajuste dos modelos robustos para tempo de falha (vida) na presença de censura;
- a criação de um procedimento mais eficiente, em termos de tempo computacional, para amostrar da distribuição *a posteriori* dos modelos robustos de degradação e tempo de falha (vida);
- a análise de sensibilidade sobre as distribuições *a priori* assumidas para os modelos propostos.

Motivado pelos dados de câncer de pulmão e pelos trabalhos de Barros *et al.* (2008), Cancho *et al.* (2010) e Vallejos e Steel (2015), seria também um ponto a ser explorado a construção de modelos que inserirem um preditor linear no parâmetro de locação dos modelos robustos para análise convencional de tempo de falha (vida) aqui propostos, com objetivo de averiguar a robustez nas estimativas dos coeficientes do preditor linear.

# Referências Bibliográficas

- Abramowitz, M. e Stegun, I. (1965). *Handbook of Mathematical Functions with Formulas Graphs and Mathematical Tables*. Dover Publications.
- 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.
- Alonso, A. e Molenberghs, G. (2008). Evaluating time to cancer recurrence as a surrogate marker for survival from an information theory perspective. *Statistical Methods in Medical Research*, 17:497–504.
- Arellano-Valle, R. B. e Azzalini, A. (2008). The centred parametrization for the multivariate skew-normal distribution. *Journal of Multivariate Analysis*, 99(7):1362 – 1382. Special Issue: Multivariate Distributions, Inference and Applications in Memory of Norman L. Johnson.
- Azzalini, A. (1985). A class of distributions which includes the normal ones. *Scandinavian Journal of Statistics*, 12(2):171–178.
- Azzalini, A. (2015). *The R package sn: The skew-normal and skew-t distributions (version 1.2-4)*. Università di Padova, Italia.
- Azzalini, A. e Capitanio, A. (2003). Distributions generated by perturbation of sym-

- metry with emphasis on a multivariate skew t-distribution. *Journal of the Royal Statistical Society, B (Statistical Methodology)*, 65(2):367–389.
- Azzalini, A., Cappello, T. d., e Kotz, S. (2003). Log-skew-normal and log-skew-t distributions as models for family income data. *Journal of Income Distribution, An International Quarterly*, 11:12–20.
- Bagdonavicius, V. e Nikulin, M. (2001). Estimation in degradation models with explanatory variables. *Lifetime Data Analysis*, 7(1):85–103.
- Bagdonavicius, V. ius, V., Masiulaityte, I., e Nikulin, M. (2010). Reliability estimation from failure-degradation data with covariates. In Nikulin, M., Limnios, N., Balakrishnan, N., Kahle, W., e Huber-Carol, C., editors, *Advances in Degradation Modeling*, Statistics for Industry and Technology, pages 275–291. Birkhäuser, Boston.
- Barros, M., Paula, G., e Leiva-Sánchez, V. (2008). A new class of survival regression models with heavy-tailed errors: robustness and diagnostics. *Lifetime Data Analysis*, 14(3):316–332.
- Bernardo, J. M. e Smith, A. F. M. (1994). *Bayesian Theory*. John Wiley & Sons.
- Box, G. E. P. e Cox, D. R. (1964). An analysis of transformations. *Journal of the Royal Statistical Society, Series B*, 26(2):211–252.
- Branco, M. D. e Dey, D. K. (2001). A general class of multivariate skew-elliptical distributions. *Journal of Multivariate Analysis*, 79(1):99 – 113.
- Branco, M. D., M. D., Genton, M. G., M. G., e Liseo, B. (2013). Objective bayesian analysis of skew-t distributions. *Scandinavian Journal of Statistics*, 40(1):63–85.
- Cancho, V. G., Dey, D. K., Lachos, V. H., e Andrade, M. G. (2011). Bayesian nonlinear regression models with scale mixtures of skew-normal distributions: Estimation and

- case influence diagnostics. *Computational Statistics & Data Analysis*, 55(1):588 – 602.
- Cancho, V. G., Ortega, E. M., e Paula, G. A. (2010). On estimation and influence diagnostics for log-birnbaum-saunders student-t regression models: Full bayesian analysis. *Journal of Statistical Planning and Inference*, 140(9):2486 – 2496.
- Casella, G. e George, E. (1992). Explaining the gibbs sampler. *The American Statistician*, 46(3):167–174.
- Chiao, C. e Hamada, M. (1996). Using degradation data from an experiment to achieve robust reliability for light emitting diodes. *Quality and Reliability Engineering International*, 12:89–94.
- Chiao, C. e Hamada, M. (2000). Experiments with degradation data for improving reliability and for achieving robust reliability. *Quality and Reliability Engineering International*, 17:333–344.
- Chib, S. (1992). Bayes inference in the tobit censored regression model. *Journal of Econometrics*, 51(12):79 – 99.
- Chib, S. e Greenberg, E. (1995). Understanding the metropolis-hastings algorithm. *The American Statistician*, 49(4):327–335.
- Crk, V. (2000). Reliability assessment from degradation data. In *Proceedings of the Annual Reliability and Maintainability Symposium*, pages 155 –161.
- Díaz-García, J. A. e Leiva-Sánchez, V. (2005). A new family of life distributions based on the elliptically contoured distributions. *Journal of Statistical Planning and Inference*, 128(2):445 – 457.
- de 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(4):363–373.
- Doksum, K. (1991). Degradation rate models for failure time and survival data. *CWI Quarterly*, 4:195–201.
- Dyk, D. A. v. e Meng, X.-L. (2001). The art of data augmentation. *Journal of Computational and Graphical Statistics*, 10(1):pp. 1–50.
- Efron, B. (1985). Bootstrap confidence intervals for a class of parametric problems. *Biometrika*, 72(1):45–58.
- Fitzmaurice, G. M., Laird, N. M., e Ware, J. H. (2004). *Applied Longitudinal Analysis*. New York:Wiley.
- Fonseca, T. C. O., Ferreira, M. A. R., e Migon, H. S. (2008). Objective bayesian analysis for the student-t regression model. *Biometrika*, 95(2):325–333.
- Freitas, M., Toledo, M., Colosimo, E., e Pires, M. (2009). Using degradation data to assess reliability: A case study on train wheel degradation. *quality and reliability engineering international*. *Quality and Reliability Engineering International*, 25:607–629.
- Freitas, M. A., Colosimo, E. A., Santos, T. R. d., e Pires, M. C. (2010). Reliability assessment using degradation models: bayesian and classical approaches. *Pesquisa Operacional*, 30:194 – 219.
- Gamerman, D. e Lopes, H. (2006). *Markov Chain Monte Carlo*. Texts in Statistical Science. Chapman & Hall, 2 edition.
- Gelfand, A. E. e Smith, A. F. M. (1990). Sampling-based approaches to calculating marginal densities. *Journal of the American Statistical Association*, 85:398–409.

- Gelman, A., Bois, F., e Jiang, J. (1996). Physiological pharmacokinetic analysis using population modeling and informative prior distributions. *Journal of the American Statistical Association*, 91:1400–1412.
- Gelman, A., Hwang, J., e Vehtari, A. (2014). Understanding predictive information criteria for bayesian models. *Statistics and Computing*, 24(6):997–1016.
- Geman, S. e Geman, D. (1984). Stochastic relaxation, gibbs distributions, and the bayesian restoration of images. *IEEE Trans. Pattern Analysis and Machine Intelligence*, 6:721–741.
- Gertsbackh, I. e Kordonskiy, K. (1969). *Models of Failure*. Springer: New York.
- Hamada, M. (2005). Using degradation data to assess reliability. *Quality Engineering*, 17:615–620.
- Henze, N. (1986). A probabilistic representation of the 'Skew-Normal' distribution. *Scandinavian Journal of Statistics*, 13(4):271–275.
- Hong, Y., Duan, Y., Meeker, W. Q., Stanley, D. L., e Gu, X. (2015). Statistical methods for degradation data with dynamic covariates information and an application to outdoor weathering data. *Technometrics*, 57(2):180–193.
- Hudak, S. J., Saxena, A., Bucci, R. J., e Malcolm, R. (1978). Development of standard methods of testing and analyzing fatigue crack growth rate data. Technical report afml-tr-78-40, Westinghouse R & D Center, Westinghouse Electric Corporation, Pittsburgh, PA.
- Jeffreys, H. (1961). *Theory of Probability*. Oxford, Oxford, England, third edition.
- Jiang, M. e Zhang, Y. (2002). Dynamic modeling of degradation data. In *Proceedings of the Annual Reliability and Maintainability Symposium*, pages 607–611.

- Kalbfleisch, J. e Prentice, R. (2002). *The Statistical Analysis of Failure Time Data*. John Wiley & Sons, New York, 2 edition.
- Kass, R. E. e Raftery, A. E. (1995). Bayes factors. *Journal of the American Statistical Association*, 90(430):773–795.
- Klaassen, C. A. J. (1985). Strong unimodality. *Advances in Applied Probability*, 17(4):pp. 905–907.
- Lawless, J. e Crowder, M. (2004). Covariates and random effects in a gamma process model with application to degradation and failure. *Lifetime Data Analysis*, 10(3):213–227.
- Lee, E. e Nelder, J. (1996). Hierarchical generalized linear models (with discussion). *Journal of the Royal Statistical Society, Series B*, 58:619–678.
- Lindstrom, M. J. e Bates, D. M. (1990). Nonlinear mixed effects models for repeated measures data. *Biometrics*, 46(3):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.
- Lu, C. e Meeker, W. (1993). Using degradation measurements to estimate a time-to-failure distribution. *Technometrics*, 35:161–174.
- Lu, C., Meeker, W., e Escobar, L. (1996). Using degradation measurements to estimate a time-to-failure distribution. *Statistica Sinica*, 6:531–546.
- 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.

- Marchenko, Y. V. e Genton, M. G. (2010). Multivariate log-skew-elliptical distributions with applications to precipitation data. *Environmetrics*, 21 (3-4):318–340.
- McCulloch, R. E. (1989). Local model influence. *Journal of the American Statistical Association*, 84(406):pp. 473–478.
- 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*, R-30(2):149 –155.
- Nelson, W. (1990). *Accelerated Testing: Statistical Models, Test Plans and Data Analyses*. John Wiley: New York.
- Newton, M. A. e Raftery, A. E. (1994). Approximate bayesian inference with the weighted likelihood bootstrap. *Journal of the Royal Statistical Society, B (Statistical Methodology)*, 56:3 –48.
- Oliveira, R. P. B. (2011). Uma abordagem bayesiana para modelos de degradação: a obtenção da distribuição preditiva a posteriori dos tempos de falha de unidades amostrais futuras e sob teste. Dissertação de Mestrado, Universidade Federal de Minas Gerais.
- Peng, C.-Y. e Tseng, S.-T. (2013). Statistical lifetime inference with skew-wiener linear degradation models. *IEEE Transactions on Reliability*, 62(2):338–350.

- Pettit, L. e Young, K. (1999). Bayesian analysis for inverse gaussian lifetime data with measures of degradation. *Journal of Statistical Computation and Simulation*, 63(3):217–234.
- Pinheiro, J., Bates, D., DebRoy, S., Sarkar, D., e the R Core team (2008). *NLME: Linear and Nonlinear Mixed Effects Models*. R package version 3.1-89.
- 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. M. (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.
- R Development Core Team (2005). R: A language and environment for statistical computing. R Foundation for Statistical Computing. Vienna, Austria.
- Robinson, M. e Crowder, M. (2000). Bayesian methods for a growth-curve degradation model with repeated measures. *Lifetime Data Analysis*, 6:357–374.
- Sobczyk, K. e Spencer, B. (1992). *Random Fatigue: From Data to Theory*. Academic Press: 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.
- Spiegelhalter, D., Thomas, A., e Best, N. (2000). Winbugs – a bayesian modelling framework: concepts, structure, and extensibility. *Statistics and Computing*, 10:325–337.

- Spiegelhalter, D. J., Best, N. G., Carlin, B. P., e Van Der Linde, A. (2002). Bayesian measures of model complexity and fit. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 64(4):583–639.
- Sturtz, S., Ligges, U., e Gelman, A. (2005). R2winbugs: A package for running winbugs from r. *Journal of Statistical Software*, 12(3):1–16.
- Su, C., Lu, J., Chen, D., e Hughes-Oliver, J. (1999). A random coefficient degradation model with random sample size. *Lifetime Data Analysis*, 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(5148439):562 – 566.
- Tanner, M. A. e Wong, W. H. (1987). The calculation of posterior distributions by data augmentation. *Journal of the American Statistical Association*, 82(398):528–540.
- Thomas, A., O’Hara, B., Ligges, U., e Sturtz, S. (2006). Making bugs open. *R News*, 6(1):12–17.
- Toledo, M. L. G. (2007). Ensaio de degradação: estudo comparativo de métodos de análise de dados. Dissertação de Mestrado, Universidade Federal de Minas Gerais.
- 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.
- Vallejos, C. A. e Steel, M. F. J. (2015). Objective bayesian survival analysis using shape mixtures of log-normal distributions. *Journal of the American Statistical Association*, 110(510):697–710.

- 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.
- Wakefield, J. (1996). The bayesian analysis of population pharmacokinetic models. *Journal of the American Statistical Association*, 91:62–75.
- Wakefield, J. C., Smith, A. F. M., Racine-Poon, A., e Gelfand, A. E. (1994). Bayesian analysis of linear and non-linear population models by using the gibbs sampler. *Journal of the Royal Statistical Society. Series C (Applied Statistics)*, 43(1):201–221.
- Watanabe, S. (2010). Asymptotic equivalence of bayes cross validation and widely applicable information criterion in singular learning theory. *Journal of Machine Learning Research*, pages 3571–3594.
- Whitmore, G. e Schenkelberg, F. (1997). Modelling accelerated degradation data using wiener diffusion with a time scale transformation. *Lifetime Data Analysis*, 3:27–45. 10.1023/A:1009664101413.
- Wu, S. e Shao, J. (1999). Reliability anaysis using the least squares method in nonlinear mixed effects degradation models. *Statistica Sinica*, 9:855–877.
- Wu, S. e Shao, J. (2000). Estimation of time-to-failure distribution derived from a degradation model using fuzzy clustering. *Quality and Reliability Engineering International*, 16:261267.
- 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.

# Anexos



## Anexo A

# Provas de Proposições

Em seguida apresentamos a prova dos itens (ii) e (iii) da Proposição 2 da Seção 3.2.1.

*Demonstração.* Usando as propriedades de medidas de probabilidade e as hipóteses da Proposição 2 obtemos que a confiabilidade *a posteriori* é dada por

$$\begin{aligned} R_{T_{n+1}|\mathbf{Y}}(t|\mathbf{y}) &= 1 - \int_{-\infty}^t f_{T_{n+1}|\mathbf{Y}}(a|\mathbf{y}) da \\ &= 1 - \int_{-\infty}^t \int_{\Xi_{\beta}} \int_{\Xi_{\beta_{n+1}}} \int_{\Xi_{\theta}} \frac{f_{T_{n+1},\mathbf{Y},\beta,\beta_{n+1},\theta}(a,\mathbf{y},\beta,\beta_{n+1},\theta)}{f_{\mathbf{Y}}(\mathbf{y})} d\theta d\beta_{n+1} d\beta da \\ &= 1 - \int_{-\infty}^t \int_{\Xi_{\beta}} \int_{\Xi_{\beta_{n+1}}} \int_{\Xi_{\theta}} f_{T_{n+1}|\mathbf{Y},\beta,\beta_{n+1},\theta}(a|\mathbf{y},\beta,\beta_{n+1},\theta) \\ &\quad \times \frac{f_{\mathbf{Y}|\beta,\beta_{n+1},\theta}(\mathbf{y}|\beta,\beta_{n+1},\theta) f_{\beta,\beta_{n+1},\theta}(\beta,\beta_{n+1},\theta)}{f_{\mathbf{Y}}(\mathbf{y})} d\theta d\beta_{n+1} d\beta da \\ &= 1 - \int_{-\infty}^t \int_{\Xi_{\beta}} \int_{\Xi_{\beta_{n+1}}} \int_{\Xi_{\theta}} f_{T_{n+1}|\beta_{n+1},\theta}(a|\beta_{n+1},\theta) \\ &\quad \times f_{\beta,\theta|\mathbf{Y}}(\beta,\theta|\mathbf{y}) f_{\beta_{n+1}|\theta}(\beta_{n+1}|\theta) d\theta d\beta_{n+1} d\beta da \\ &= 1 - \int_{-\infty}^t \int_{\Xi_{\theta}} \int_{\Xi_{\beta}} f_{\beta,\theta|\mathbf{Y}}(\beta,\theta|\mathbf{y}) \\ &\quad \times \int_{\Xi_{\beta_{n+1}}} f_{T_{n+1}|\beta_{n+1},\theta}(a|\beta_{n+1},\theta) f_{\beta_{n+1}|\theta}(\beta_{n+1}|\theta) d\beta_{n+1} d\beta d\theta da \end{aligned}$$

$$\begin{aligned}
&= 1 - \int_{\Xi_{\boldsymbol{\theta}}} \int_{\Xi_{\boldsymbol{\beta}}} f_{\boldsymbol{\beta}, \boldsymbol{\theta} | \mathbf{Y}}(\boldsymbol{\beta}, \boldsymbol{\theta} | \mathbf{y}) \int_{-\infty}^t f_{T_{n+1} | \boldsymbol{\theta}}(a | \boldsymbol{\theta}) da d\boldsymbol{\beta} d\boldsymbol{\theta} \\
&= 1 - \int_{\Xi_{\boldsymbol{\theta}}} \int_{\Xi_{\boldsymbol{\beta}}} f_{\boldsymbol{\beta}, \boldsymbol{\theta} | \mathbf{Y}}(\boldsymbol{\beta}, \boldsymbol{\theta} | \mathbf{y}) F_{T_{n+1} | \boldsymbol{\theta}}(t | \boldsymbol{\theta}) d\boldsymbol{\beta} d\boldsymbol{\theta} \\
&= 1 - \int_{\Xi_{\boldsymbol{\theta}}} f_{\boldsymbol{\theta} | \mathbf{Y}}(\boldsymbol{\theta} | \mathbf{y}) \left\{ 1 - R_{T_{n+1} | \boldsymbol{\theta}}(t | \boldsymbol{\theta}) \right\} d\boldsymbol{\theta} \\
&= \int_{\Xi_{\boldsymbol{\theta}}} f_{\boldsymbol{\theta} | \mathbf{Y}}(\boldsymbol{\theta} | \mathbf{y}) R_{T_{n+1} | \boldsymbol{\theta}}(t | \boldsymbol{\theta}) d\boldsymbol{\theta} \\
&= E_{\boldsymbol{\theta} | \mathbf{Y}} \left( R_{T_{n+1} | \boldsymbol{\theta}}(t | \boldsymbol{\theta}) \right).
\end{aligned}$$

o que conclui a prova do item (ii).

Usando a definição de esperança e argumentos similares aos usados para os itens (i) e (ii), temos que o tempo médio até a falha *a posteriori* é:

$$\begin{aligned}
E[T_{n+1} | \mathbf{y}] &= \int_{-\infty}^{\infty} t f_{T_{n+1} | \mathbf{Y}}(t | \mathbf{y}) dt \\
&\stackrel{(i)}{=} \int_{-\infty}^{\infty} \int_{\Xi_{\boldsymbol{\theta}}} t f_{T_{n+1} | \boldsymbol{\theta}}(t | \boldsymbol{\theta}) f_{\boldsymbol{\theta} | \mathbf{Y}}(\boldsymbol{\theta} | \mathbf{y}) d\boldsymbol{\theta} dt \\
&= \int_{\Xi_{\boldsymbol{\theta}}} f_{\boldsymbol{\theta} | \mathbf{Y}}(\boldsymbol{\theta} | \mathbf{y}) \int_{-\infty}^{\infty} t f_{T_{n+1} | \boldsymbol{\theta}}(t | \boldsymbol{\theta}) dt d\boldsymbol{\theta} \\
&= \int_{\Xi_{\boldsymbol{\theta}}} E_{T_{n+1} | \boldsymbol{\theta}}[t | \boldsymbol{\theta}] f_{\boldsymbol{\theta} | \mathbf{Y}}(\boldsymbol{\theta} | \mathbf{y}) d\boldsymbol{\theta} \\
&= E_{\boldsymbol{\theta} | \mathbf{Y}} \left[ E_{T_{n+1} | \boldsymbol{\theta}}[t | \boldsymbol{\theta}] \right].
\end{aligned}$$

o que conclui a prova do item (iii). □

A seguir apresentamos a prova da Proposição 8 da Seção 4.2.

*Demonstração.* Considerando a função de verossimilhança na equação (4.15) e a forma hierárquica da representação estocástica da distribuição SMSN em (4.18), segue que a dcc *a posteriori* de cada  $\beta_i$  é dada por

$$\begin{aligned}
f(\beta_i | \boldsymbol{\psi}^{[-\beta_i]}, \mathbf{y}) &\propto \left[ \prod_{j=1}^{m_i} N\left(\frac{1}{\beta_i} t_{ij}; \sigma_{\varepsilon}^2\right) \right] N(\mu_{\beta} + \Delta_{\beta} u_i, h_i^{-1} \tau_{\beta}) \\
&\propto \exp \left\{ -\frac{1}{2\sigma_{\varepsilon}^2} \sum_{j=1}^{m_i} \left( y_{ij} - \frac{t_{ij}}{\beta_i} \right)^2 - \frac{1}{2h_i^{-1} \tau_{\beta}} [\beta_i - (\mu_{\beta} + \Delta_{\beta} u_i)]^2 \right\},
\end{aligned}$$

o que conclui a prova.

Da mesma forma, considerando a distribuição *a priori* para  $\sigma_\varepsilon^2$  em (4.19), sua dcc *a posteriori* torna-se

$$\begin{aligned}
f\left(\sigma_\varepsilon^2 | \boldsymbol{\psi}^{[-\sigma_\varepsilon^2]}, \mathbf{y}\right) &\propto \left[ \prod_{i=1}^n \prod_{j=1}^{m_i} N\left(\frac{1}{\beta_i} t_{ij}; \sigma_\varepsilon^2\right) \right] GI(r_{\sigma_\varepsilon^2}, s_{\sigma_\varepsilon^2}) \\
&\propto (\sigma_\varepsilon^2)^{-\sum_{i=1}^n \frac{m_i}{2}} \exp\left\{-\frac{1}{2\sigma_\varepsilon^2} \sum_{i=1}^n \sum_{j=1}^{m_i} \left(y_{ij} - \frac{t_{ij}}{\beta_i}\right)^2\right\} \\
&\quad \times (\sigma_\varepsilon^2)^{-r_{\sigma_\varepsilon^2}-1} \exp\left\{-\frac{s_{\sigma_\varepsilon^2}}{\sigma_\varepsilon^2}\right\} \\
&\propto (\sigma_\varepsilon^2)^{-\left(\sum_{i=1}^n \frac{m_i}{2} + r_{\sigma_\varepsilon^2}\right)-1} \exp\left\{-\frac{1}{\sigma_\varepsilon^2} \left[\frac{1}{2} \sum_{i=1}^n \sum_{j=1}^{m_i} \left(y_{ij} - \frac{t_{ij}}{\beta_i}\right)^2 - s_{\sigma_\varepsilon^2}\right]\right\},
\end{aligned}$$

o que conclui a prova, e portanto a prova do item (i-a).  $\square$

*Demonstração.* Similarmente, considerando a função de verossimilhança na equação (4.17) e a forma hierárquica da representação estocástica da distribuição SMSN em (4.18), segue que a dcc *a posteriori* de cada  $\beta_i$  é dada por

$$\begin{aligned}
f\left(\beta_i | \boldsymbol{\psi}^{[-\beta_i]}, \mathbf{y}\right) &\propto \left[ \prod_{j=1}^{m_i} N(\beta_i t_{ij}; \sigma_\varepsilon^2) \right] N(\mu_\beta + \Delta_\beta u_i, h_i^{-1} \tau_\beta) \\
&\propto \exp\left\{-\frac{1}{2\sigma_\varepsilon^2} \sum_{j=1}^{m_i} (y_{ij} - \beta_i t_{ij})^2 - \frac{1}{2h_i^{-1} \tau_\beta} [\beta_i - (\mu_\beta + \Delta_\beta u_i)]^2\right\} \\
&\propto \exp\left\{-\frac{1}{2\sigma_\varepsilon^2 h_i^{-1} \tau_\beta} \left[\beta_i^2 \left(h_i^{-1} \tau_\beta \sum_{j=1}^{m_i} t_{ij}^2 + \sigma_\varepsilon^2\right) \right. \right. \\
&\quad \left. \left. - 2\beta_i \left(\sum_{j=1}^{m_i} y_{ij} t_{ij} + (\mu_\beta + \Delta_\beta U_i)\right)\right]\right\} \\
&\propto \exp\left\{-\frac{h_i^{-1} \tau_\beta \sum_{j=1}^{m_i} t_{ij}^2 + \sigma_\varepsilon^2}{2\sigma_\varepsilon^2 h_i^{-1} \tau_\beta} \left[\beta_i^2 - 2\beta_i \left(\frac{\sum_{j=1}^{m_i} y_{ij} t_{ij} + (\mu_\beta + \Delta_\beta U_i)}{h_i^{-1} \tau_\beta \sum_{j=1}^{m_i} t_{ij}^2 + \sigma_\varepsilon^2}\right)\right]\right\},
\end{aligned}$$

o que conclui a prova.

Da mesma forma, considerando a distribuição *a priori* para  $\sigma_\varepsilon^2$  em (4.19), segue que

sua dcc *a posteriori* é

$$\begin{aligned}
 f\left(\sigma_\varepsilon^2 | \boldsymbol{\psi}^{[-\sigma_\varepsilon^2]}, \mathbf{y}\right) &\propto \left[ \prod_{i=1}^n \prod_{j=1}^{m_i} N(\beta_i t_{ij}; \sigma_\varepsilon^2) \right] GI(r_{\sigma_\varepsilon^2}, s_{\sigma_\varepsilon^2}) \\
 &\propto (\sigma_\varepsilon^2)^{-\sum_{i=1}^n \frac{m_i}{2}} \exp \left\{ -\frac{1}{2\sigma_\varepsilon^2} \sum_{i=1}^n \sum_{j=1}^{m_i} (y_{ij} - \beta_i t_{ij})^2 \right\} \\
 &\quad \times (\sigma_\varepsilon^2)^{-r_{\sigma_\varepsilon^2} - 1} \exp \left\{ -\frac{s_{\sigma_\varepsilon^2}}{\sigma_\varepsilon^2} \right\} \\
 &\propto (\sigma_\varepsilon^2)^{-\left(\sum_{i=1}^n \frac{m_i}{2} + r_{\sigma_\varepsilon^2}\right) - 1} \exp \left\{ -\frac{1}{\sigma_\varepsilon^2} \left[ \frac{1}{2} \sum_{i=1}^n \sum_{j=1}^{m_i} (y_{ij} - \beta_i t_{ij})^2 - s_{\sigma_\varepsilon^2} \right] \right\},
 \end{aligned}$$

o que conclui a prova, e portanto a prova do item (i-b).  $\square$

*Demonstração.* Considerando a forma hierárquica da representação estocástica da distribuição SMSN em (4.18), segue que a dcc *a posteriori* de cada  $u_i$  é dada por

$$\begin{aligned}
 f\left(u_i | \boldsymbol{\psi}^{[-u_i]}, \mathbf{y}\right) &\propto N(\mu_\beta + \Delta_\beta u_i, h_i^{-1} \tau_\beta) N(0, h^{-1}) I(u_i)_{(0, \infty)} \\
 &\propto \exp \left\{ -\frac{u_i^2}{2h_i^{-1}} - \frac{1}{2h_i^{-1} \tau_\beta} [\beta_i - (\mu_\beta + \Delta_\beta u_i)]^2 \right\} I(u_i)_{(0, \infty)} \\
 &\propto \exp \left\{ -\frac{1}{2h_i^{-1} \tau_\beta} [u_i^2 (\tau_\beta + \Delta_\beta^2) - 2u_i \Delta_\beta (\beta_i - \mu_\beta)] \right\} I(u_i)_{(0, \infty)} \\
 &\propto \exp \left\{ -\frac{\tau_\beta + \Delta_\beta^2}{2h_i^{-1} \tau_\beta} \left[ u_i^2 - 2u_i \frac{\Delta_\beta (\beta_i - \mu_\beta)}{\tau_\beta + \Delta_\beta^2} \right] \right\} I(u_i)_{(0, \infty)},
 \end{aligned}$$

o que conclui a prova.

Analogamente, a dcc *a posteriori* para  $\mu_\beta$  é

$$\begin{aligned}
 f\left(\mu_\beta | \boldsymbol{\psi}^{[-\mu_\beta]}, \mathbf{y}\right) &\propto \left[ \prod_{i=1}^n N(\mu_\beta + \Delta_\beta u_i, h_i^{-1} \tau_\beta) \right] N(m, V) \\
 &\propto \exp \left\{ -\sum_{i=1}^n \left\{ \frac{1}{2h_i^{-1} \tau_\beta} [\beta_i - (\mu_\beta + \Delta_\beta u_i)]^2 \right\} - \frac{1}{2V} (\mu_\beta - m)^2 \right\} \\
 &\propto \exp \left\{ -\frac{1}{2V \prod_{i=1}^n h_i^{-1}} \left[ -2\mu_\beta V \sum_{i=1}^n \beta_i \prod_{k \neq i} h_k^{-1} + \mu_\beta^2 \sum_{i=1}^n V \prod_{k \neq i} h_k^{-1} \right. \right. \\
 &\quad \left. \left. + 2\mu_\beta \Delta_\beta \sum_{i=1}^n u_i V \prod_{k \neq i} h_k^{-1} + \mu_\beta^2 \tau_\beta \prod_{i=1}^n h_i^{-1} - 2\mu_\beta m \tau_\beta \prod_{i=1}^n h_i^{-1} \right] \right\}
 \end{aligned}$$

$$\begin{aligned}
& \propto \exp \left\{ -\frac{1}{2V \prod_{i=1}^n h_i^{-1}} \left[ \mu_\beta^2 \left( \sum_{i=1}^n V_i^* + \tau_\beta \prod_{i=1}^n h_i^{-1} \right) \right. \right. \\
& \quad \left. \left. - 2\mu_\beta \left( \sum_{i=1}^n \beta_i V_i^* + m\tau_\beta \prod_{i=1}^n h_i^{-1} - \Delta_\beta \sum_{i=1}^n u_i V_i^* \right) \right] \right\} \\
& \propto \exp \left\{ -\frac{\sum_{i=1}^n V_i^* + \tau_\beta \prod_{i=1}^n h_i^{-1}}{2V \prod_{i=1}^n h_i^{-1}} [\mu_\beta^2 \right. \\
& \quad \left. - 2\mu_\beta \frac{\sum_{i=1}^n \beta_i V_i^* + m\tau_\beta \prod_{i=1}^n h_i^{-1} - \Delta_\beta \sum_{i=1}^n u_i V_i^*}{\sum_{i=1}^n V_i^* + \tau_\beta \prod_{i=1}^n h_i^{-1}}] \right\},
\end{aligned}$$

o que conclui a prova.

Sobre os mesmos pressupostos, a dcc *a posteriori* para  $\Delta_\beta$  é

$$\begin{aligned}
f(\Delta_\beta | \boldsymbol{\psi}^{[-\Delta_\beta]}, \mathbf{y}) & \propto \left[ \prod_{i=1}^n N(\mu_\beta + \Delta_\beta u_i, h_i^{-1} \tau_\beta) \right] N(\mu_\Delta, \sigma_\Delta^2) \\
& \propto \exp \left\{ -\sum_{i=1}^n \left\{ \frac{1}{2h_i^{-1} \tau_\beta} [\beta_i - (\mu_\beta + \Delta_\beta u_i)]^2 \right\} - \frac{1}{2\sigma_\Delta^2} (\Delta_\beta - \mu_\Delta)^2 \right\} \\
& \propto \exp \left\{ -\frac{1}{2\sigma_\Delta^2 \prod_{i=1}^n h_i^{-1}} \left[ -2\Delta_\beta \sigma_\Delta^2 \sum_{i=1}^n \beta_i u_i \prod_{k \neq i} h_k^{-1} + 2\Delta_\beta \sigma_\Delta^2 \sum_{i=1}^n u_i \prod_{k \neq i} h_k^{-1} \right. \right. \\
& \quad \left. \left. + \Delta^2 \sigma_\Delta^2 \sum_{i=1}^n u_i \prod_{k \neq i} h_k^{-1} + \Delta^2 \tau_\beta \prod_{i=1}^n h_i^{-1} - 2\Delta_\beta \mu_\Delta \tau_\beta \prod_{i=1}^n h_i^{-1} \right] \right\} \\
& \propto \exp \left\{ -\frac{1}{2\sigma_\Delta^2 \prod_{i=1}^n h_i^{-1}} \left[ \Delta^2 \left( \sigma_\Delta^2 \sum_{i=1}^n U_i^* + \tau_\beta \prod_{i=1}^n h_i^{-1} \right) \right. \right. \\
& \quad \left. \left. - 2\Delta_\beta \left( \sigma_\Delta^2 \left( \sum_{i=1}^n \beta_i U_i^* - \sum_{i=1}^n U_i^* \right) + \mu_\Delta \tau_\beta \prod_{i=1}^n h_i^{-1} \right) \right] \right\} \\
& \propto \exp \left\{ -\frac{\sigma_\Delta^2 \sum_{i=1}^n U_i^* + \tau_\beta \prod_{i=1}^n h_i^{-1}}{2\sigma_\Delta^2 \prod_{i=1}^n h_i^{-1}} [\Delta^2 \right. \\
& \quad \left. - 2\Delta_\beta \frac{\sigma_\Delta^2 (\sum_{i=1}^n \beta_i U_i^* - \sum_{i=1}^n U_i^*) + \mu_\Delta \tau_\beta \prod_{i=1}^n h_i^{-1}}{\sigma_\Delta^2 \sum_{i=1}^n U_i^* + \tau_\beta \prod_{i=1}^n h_i^{-1}}] \right\},
\end{aligned}$$

o que conclui a prova.

Ainda sob as mesmas suposições, a dcc *a posteriori* para  $\tau_\beta$  é dada por

$$\begin{aligned}
f(\tau_\beta | \boldsymbol{\psi}^{[-\tau_\beta]}, \mathbf{y}) & \propto \left[ \prod_{i=1}^n N(\mu_\beta + \Delta_\beta U_i, H_i^{-1} \tau_\beta) \right] GI(r_{\tau_\beta}, s_{\tau_\beta}) \\
& \propto \left( \frac{1}{\tau_\beta} \right)^{\frac{n}{2}} \exp \left\{ -\sum_{i=1}^n \left\{ \frac{1}{2h_i^{-1} \tau_\beta} [\beta_i - (\mu_\beta + \Delta_\beta U_i)]^2 \right\} \right\} \frac{1}{\tau_\beta} r_{\tau_\beta}^{+1} \exp \left\{ -\frac{s_{\tau_\beta}}{\tau_\beta} \right\}
\end{aligned}$$

$$\propto \left( \frac{1}{\tau_\beta} \right)^{\left(\frac{n}{2} + r_{\tau_\beta}\right) + 1} \exp \left\{ -\frac{1}{\tau_\beta} \left\{ \sum_{i=1}^n \left\{ \frac{1}{2h_i^{-1}} [\beta_i - (\mu_\beta + \Delta_\beta U_i)]^2 \right\} + s_{\tau_\beta} \right\} \right\},$$

em que  $k = \{1, \dots, n\}$ ,  $V_i^* = V \prod_{k \neq i} h_k^{-1}$  e  $U_i^* = u_i \prod_{k \neq i} h_k^{-1}$ .

O que conclui a prova, e portanto a prova do item (ii).  $\square$

A seguir apresentamos a prova das equações em (4.21) da Seção 4.2.

*Demonstração.* Como  $h_i \stackrel{iid}{\sim} \text{Gama}(\nu/2, \nu/2)$  e  $\nu$  tem a distribuição *a priori* marginal independente de Jeffreys tal como proposta por Fonseca *et al.* (2008) na equação (4.20), sob os modelos de degradação com forma funcional nas equações (4.1) e (4.6), a dcc *a posteriori* para  $h_i$  é dada por

$$\begin{aligned} f(h_i | \boldsymbol{\psi}^{[-h_i]}, \mathbf{y}) &\propto N(\mu_\beta + \Delta_\beta u_i, h_i^{-1} \tau_\beta) N(0, H_i^{-1}) \text{Gama}\left(\frac{\nu}{2}, \frac{\nu}{2}\right) \\ &\propto \left(\frac{1}{h_i^{-1}}\right)^{\frac{1}{2}} \exp\left\{-\frac{1}{h_i^{-1} \tau_\beta} [\beta_i - (\mu_\beta + \Delta_\beta U_i)]^2\right\} \\ &\quad \times \left(\frac{1}{h_i^{-1}}\right)^{\frac{1}{2}} \exp\left\{-\frac{u_i^2}{2h_i^{-1}}\right\} \\ &\quad \times h_i^{\frac{\nu}{2}-1} \exp\left\{-h_i \frac{\nu}{2}\right\} \\ &\propto h_i^{\left(\frac{\nu}{2}+1\right)-1} \exp\left\{-h_i \frac{1}{2} \left\{ \tau_\beta^{-1} [\beta_i - (\mu_\beta + \Delta_\beta U_i)]^2 + u_i^2 + \nu \right\}\right\}, \end{aligned}$$

o que conclui a prova.

Sob as mesmas suposições a dcc *a posteriori* para  $\nu$  será dada por

$$\begin{aligned} f(\nu | \boldsymbol{\psi}^{[-\nu]}, \mathbf{y}) &\propto GI\left(\frac{\nu}{2}, \frac{\nu}{2}\right) f(\nu) \\ &\propto \left[ \frac{\nu^{\frac{\nu}{2}}}{\Gamma(\frac{\nu}{2})} \right]^n \left[ \prod_{i=1}^n h_i^{\frac{\nu}{2}-1} \right] \exp\left\{-\sum_{i=1}^n h_i \frac{\nu}{2}\right\} \\ &\quad \times \left(\frac{\nu}{\nu+3}\right)^{\frac{1}{2}} \left\{ \frac{\partial^2 \log [\Gamma(\frac{\nu}{2})]}{\partial^2 \nu} - \frac{\partial^2 \log [\Gamma(\frac{\nu+1}{2})]}{\partial^2 \nu} \right. \\ &\quad \left. - \frac{2(\nu+3)}{\nu(\nu+1)^2} \right\}^{\frac{1}{2}}, \end{aligned}$$

o que conclui a prova.  $\square$

Apresentamos agora a prova da Proposição 10, Seção 5.2.1.

*Demonstração.* Considere a representação estocástica em (5.8) e as distribuições *a priori* em (5.11). Como  $T_i \stackrel{iid}{\sim} SMSN(\mu, h_i^{-1}\omega^2, \alpha, \mathcal{H})$ , usando a técnica de aumento de dados descrita na Seção 5.1, temos que a dcc *a posteriori* de qualquer tempo latente  $t_i \in \mathbf{t}_{lat}$  segue diretamente de (5.3), o que conclui a prova do item (i);  $\square$

*Demonstração.* Como  $T_i$  são iid SMSN, as va's  $U_i$  também são iid com distribuição Normal Truncada, com locação 0 e escala  $h_i^{-1}$ , colocando massa apenas em valores não negativos. Segue que a dcc *a posteriori* de  $u_i$  é:

$$\begin{aligned} f(u_i | \boldsymbol{\psi}^{[-\mathbf{u}_i]}, \mathbf{D}) &\propto N(\mu + \Delta u_i, h_i^{-1}\tau) N(0, h^{-1}) I(u_i)_{(0, \infty)} \\ &\propto \exp\left\{-\frac{u_i^2}{2h_i^{-1}} - \frac{1}{2h_i^{-1}\tau} [t_i - (\mu + \Delta U_i)]^2\right\} I(u_i)_{(0, \infty)} \\ &\propto \exp\left\{-\frac{1}{2h_i^{-1}\tau} [u_i^2 (\tau_\beta + \Delta_\beta^2) - 2u_i \Delta (t_i - \mu)]\right\} I(u_i)_{(0, \infty)}, \\ &\propto \exp\left\{-\frac{\tau_\beta + \Delta_\beta^2}{2h_i^{-1}\tau} \left[u_i^2 - 2u_i \frac{\Delta (t_i - \mu)}{\tau_\beta + \Delta_\beta^2}\right]\right\} I(u_i)_{(0, \infty)} \end{aligned}$$

de onde segue o resultado.

Da mesma forma obtemos que

$$\begin{aligned} f(\mu | \boldsymbol{\psi}^{[-\boldsymbol{\mu}]}, \mathbf{D}) &\propto \left[ \prod_{i=1}^n N(\mu + \Delta u_i, h_i^{-1}\tau) \right] N(m, V) \\ &\propto \exp\left\{-\sum_{i=1}^n \left\{ \frac{1}{2h_i^{-1}\tau} [t_i - (\mu + \Delta u_i)]^2 \right\} - \frac{1}{2V} (\mu - m)^2 \right\} \\ &\propto \exp\left\{-\frac{1}{2V \prod_{i=1}^n h_i^{-1}} \left[ -2\mu V \sum_{i=1}^n t_i \prod_{k \neq i} h_k^{-1} + \mu^2 \sum_{i=1}^n V \prod_{k \neq i} h_k^{-1} \right. \right. \\ &\quad \left. \left. + 2\mu \Delta \sum_{i=1}^n u_i V \prod_{k \neq i} h_k^{-1} + \mu^2 \tau \prod_{i=1}^n h_i^{-1} - 2\mu m \tau \prod_{i=1}^n h_i^{-1} \right] \right\} \\ &\propto \exp\left\{-\frac{1}{2V \prod_{i=1}^n h_i^{-1}} \left[ \mu^2 \left( \sum_{i=1}^n V_i^* + \tau \prod_{i=1}^n h_i^{-1} \right) \right. \right. \\ &\quad \left. \left. - 2\mu \left( \sum_{i=1}^n t_i V_i^* + m \tau \prod_{i=1}^n h_i^{-1} - \Delta \sum_{i=1}^n u_i V_i^* \right) \right] \right\} \end{aligned}$$

$$\propto \exp \left\{ -\frac{\sum_{i=1}^n V_i^* + \tau \prod_{i=1}^n h_i^{-1}}{2V \prod_{i=1}^n h_i^{-1}} [\mu^2 - 2\mu \frac{\sum_{i=1}^n t_i V_i^* + m\tau \prod_{i=1}^n h_i^{-1} - \Delta \sum_{i=1}^n u_i V_i^*}{\sum_{i=1}^n V_i^* + \tau \prod_{i=1}^n h_i^{-1}}] \right\},$$

em que  $V_i^* = V \prod_{k \neq i} h_k^{-1}$  e de onde obtemos o resultado.

A dcc *a posteriori* para  $\Delta$  é tal que

$$\begin{aligned} f(\Delta | \boldsymbol{\psi}^{[-\Delta]}, \mathbf{D}) &\propto \left[ \prod_{i=1}^n N(\mu + \Delta u_i, h_i^{-1} \tau) \right] N(\mu_\Delta, \sigma_\Delta^2) \\ &\propto \exp \left\{ -\sum_{i=1}^n \left\{ \frac{1}{2h_i^{-1} \tau} [t_i - (\mu + \Delta u_i)]^2 \right\} - \frac{1}{2\sigma_\Delta^2} (\Delta - \mu_\Delta)^2 \right\} \\ &\propto \exp \left\{ -\frac{1}{2\sigma_\Delta^2 \prod_{i=1}^n h_i^{-1}} \left[ -2\Delta \sigma_\Delta^2 \sum_{i=1}^n t_i u_i \prod_{k \neq i} h_k^{-1} + 2\Delta \sigma_\Delta^2 \sum_{i=1}^n u_i \prod_{k \neq i} h_k^{-1} \right. \right. \\ &\quad \left. \left. + \Delta^2 \sigma_\Delta^2 \sum_{i=1}^n u_i \prod_{k \neq i} h_k^{-1} + \Delta^2 \tau \prod_{i=1}^n h_i^{-1} - 2\Delta \mu_\Delta \tau \prod_{i=1}^n h_i^{-1} \right] \right\} \\ &\propto \exp \left\{ -\frac{1}{2\sigma_\Delta^2 \prod_{i=1}^n h_i^{-1}} \left[ \Delta^2 \left( \sigma_\Delta^2 \sum_{i=1}^n U_i^* + \tau \prod_{i=1}^n h_i^{-1} \right) \right. \right. \\ &\quad \left. \left. - 2\Delta \left( \sigma_\Delta^2 \left( \sum_{i=1}^n t_i U_i^* - \sum_{i=1}^n U_i^* \right) + \mu_\Delta \tau \prod_{i=1}^n h_i^{-1} \right) \right] \right\} \\ &\propto \exp \left\{ -\frac{\sigma_\Delta^2 \sum_{i=1}^n U_i^* + \tau \prod_{i=1}^n h_i^{-1}}{2\sigma_\Delta^2 \prod_{i=1}^n h_i^{-1}} [\Delta^2 \right. \\ &\quad \left. - 2\Delta \frac{\sigma_\Delta^2 (\sum_{i=1}^n t_i U_i^* - \sum_{i=1}^n U_i^*) + \mu_\Delta \tau \prod_{i=1}^n h_i^{-1}}{\sigma_\Delta^2 \sum_{i=1}^n U_i^* + \tau \prod_{i=1}^n h_i^{-1}}] \right\}, \end{aligned}$$

em que  $U_i^* = u_i \prod_{k \neq i} h_k^{-1}$ , o que conclui a prova.

De forma análoga temos que a dcc *a posteriori* de  $\tau$  é

$$\begin{aligned} f(\tau | \boldsymbol{\psi}^{[-\tau]}, \mathbf{D}) &\propto \left[ \prod_{i=1}^n N(\mu + \Delta u_i, h_i^{-1} \tau) \right] GI(r_\tau, s_\tau) \\ &\propto \left( \frac{1}{\tau} \right)^{\frac{n}{2}} \exp \left\{ -\sum_{i=1}^n \left\{ \frac{1}{2h_i^{-1} \tau} [t_i - (\mu + \Delta u_i)]^2 \right\} \right\} \frac{1}{\tau} \exp \left\{ -\frac{s_\tau}{\tau} \right\} \\ &\propto \left( \frac{1}{\tau} \right)^{\left(\frac{n}{2} + r_\tau\right) + 1} \exp \left\{ -\frac{1}{\tau} \left\{ \sum_{i=1}^n \left\{ \frac{1}{2h_i^{-1}} [t_i - (\mu + \Delta u_i)]^2 \right\} + s_\tau \right\} \right\}, \end{aligned}$$

de onde segue o resultado e conclui a prova do item (ii).  $\square$

A seguir apresentamos a prova das distribuições condicionais completas para  $h_i$  e



$\nu$ , dadas pelas equações em (5.15) da Seção 5.2.1.

*Demonstração.* Se  $H$  é a lei de probabilidade associada à distribuição  $Gamma(\nu/2, \nu/2)$ , e assumindo a distribuição *a priori* para  $\nu$  dada em (4.20), segue que as dcc para  $h_i$ ,  $i = 1, \dots, n$ , e  $\nu$  são respectivamente

$$\begin{aligned}
f(h_i | \boldsymbol{\psi}^{[-h_i]}, \mathbf{D}) &\propto N(\mu + \Delta U_i, H_i^{-1} \tau) N(0, H_i^{-1}) Gama\left(\frac{\nu}{2}, \frac{\nu}{2}\right) \\
&\propto \left(\frac{1}{h_i^{-1}}\right)^{\frac{1}{2}} \exp\left\{-\frac{1}{h_i^{-1} \tau} [t_i - (\mu + \Delta U_i)]^2\right\} \\
&\quad \times \left(\frac{1}{h_i^{-1}}\right)^{\frac{1}{2}} \exp\left\{-\frac{u_i^2}{2h_i^{-1}}\right\} \\
&\quad \times h_i^{\frac{\nu}{2}-1} \exp\left\{-h_i \frac{\nu}{2}\right\} \\
&\propto h_i^{(\frac{\nu}{2}+1)-1} \exp\left\{-h_i \frac{1}{2} \left\{\tau^{-1} [t_i - (\mu + \Delta U_i)]^2 + u_i^2 + \nu\right\}\right\}, \\
f(\nu | \boldsymbol{\psi}^{[-\nu]}, \mathbf{D}) &\propto GI\left(\frac{\nu}{2}, \frac{\nu}{2}\right) f(\nu) \\
&\propto \left[\frac{\frac{\nu}{2}}{\Gamma(\frac{\nu}{2})}\right]^n \left[\prod_{i=1}^n h_i^{\frac{\nu}{2}-1}\right] \exp\left\{-\sum_{i=1}^n h_i \frac{\nu}{2}\right\} \times \\
&\quad \times \left(\frac{\nu}{\nu+3}\right)^{\frac{1}{2}} \left\{\frac{\partial^2 \log[\Gamma(\frac{\nu}{2})]}{\partial^2 \nu} - \frac{\partial^2 \log[\Gamma(\frac{\nu+1}{2})]}{\partial^2 \nu} - \right. \\
&\quad \left. - \frac{2(\nu+3)}{\nu(\nu+1)^2}\right\}^{\frac{1}{2}},
\end{aligned}$$

o que encerra a prova. □

Apresentamos agora a prova da Proposição 12, Seção 5.2.2.

*Demonstração.* Considere a representação estocástica em (5.18) e as distribuições *a priori* em (5.17). Como  $T_i \stackrel{iid}{\sim} SNC(\mu_o, \omega_o^2, \alpha_o)$ , usando a técnica de aumento de dados descrita na Seção 5.1, temos que a dcc *a posteriori* de qualquer tempo latente  $t_i \in \mathbf{t}_{lat}$  segue diretamente de (5.3), o que conclui a prova do item (i); □

*Demonstração.* Como  $T_i$  são iid SMC, as va's  $u_i$  também são iid com distribuição Normal Truncada com locação 0 e escala 1, colocando massa apenas em valores não

negativos. Segue que a dcc *a posteriori* de  $u_i$  é:

$$\begin{aligned}
f(u_i | \boldsymbol{\psi}^{[-u_i]}, \mathbf{D}) &\propto N(\mu_o + \omega_o \rho (1 - b^{-1} u_i), \tau_o) N(0, 1) I(u_i)_{(0, \infty)} \\
&\propto \exp \left\{ -\frac{[t_i - (\mu_o + \omega_o \rho (1 - b^{-1} u_i))]^2}{2\tau_o} - \frac{u_i^2}{2} \right\} I(u_i)_{(0, \infty)} \\
&\propto \exp \left\{ -\frac{1}{2\tau_o} [-2t_i \omega_o \rho (1 - b^{-1} u_i) + 2\mu_o \omega_o \rho (1 - b^{-1} u_i) + \right. \\
&\quad \left. + \omega_o^2 \rho^2 (1 - b^{-1} u_i)^2] + \frac{u_i^2}{2} \right\} I(u_i)_{(0, \infty)} \\
&\propto \exp \left\{ -\frac{1}{2\tau_o} [2t_i \omega_o \rho b^{-1} u_i - 2\mu_o \omega_o \rho b^{-1} u_i + \right. \\
&\quad \left. + \omega_o^2 \rho^2 (1 - 2b^{-1} u_i + b^{-2} u_i^2) + \tau_o u_i^2] \right\} I(u_i)_{(0, \infty)} \\
&\propto \exp \left\{ -\frac{1}{2\tau_o} [u_i^2 (\omega_o^2 \rho^2 b^{-2} + \tau_o) - \right. \\
&\quad \left. - 2u_i b^{-1} \omega_o \rho (\mu_o + \omega_o \rho - t_i)] \right\} I(u_i)_{(0, \infty)} \\
&\propto \exp \left\{ -\frac{\omega_o^2 \rho^2 + b^2 \tau_o}{2\tau_o b^2} \left[ u_i^2 - 2u_i \frac{b \omega_o \rho (\mu_o + \omega_o \rho - t_i)}{\omega_o^2 \rho^2 + b^2 \tau_o} \right] \right\} I(u_i)_{(0, \infty)};
\end{aligned}$$

o que conclui a prova.

Segue também que a dcc *a posteriori* para  $\mu_o$  é

$$\begin{aligned}
f(\mu_o | \boldsymbol{\psi}^{[-\mu_o]}, \mathbf{D}) &\propto \left[ \prod_{i=1}^n N(\mu_o + \omega_o \rho (1 - b^{-1} u_i), \tau_o) \right] N(m_o, v_o) \\
&\propto \exp \left\{ -\frac{\sum_{i=1}^n [t_i - (\mu_o + \omega_o \rho (1 - b^{-1} u_i))]^2}{2\tau_o} - \frac{(\mu_o - m_o)^2}{2v_o} \right\} \\
&\propto \exp \left\{ -\frac{\sum_{i=1}^n [-2t_i \mu_o + \mu_o^2 + 2\mu_o \omega_o \rho (1 - b^{-1} u_i)]}{2\tau_o} - \right. \\
&\quad \left. - \frac{(\mu_o^2 - 2\mu_o m_o)}{2v_o} \right\} \\
&\propto \exp \left\{ -\frac{1}{2v_o \tau_o} [\mu_o^2 (nv_o + \tau_o) - \right. \\
&\quad \left. - 2\mu_o \left( v_o \sum_{i=1}^n t_i + m_o \tau_o + v_o \omega_o \rho \left( n - b^{-1} \sum_{i=1}^n u_i \right) \right)] \right\} \\
&\propto \exp \left\{ -\frac{nv_o + \tau_o}{2v_o \tau_o} [\mu_o^2 - \right. \\
&\quad \left. - 2\mu_o \frac{(m_o \tau_o + v_o (\sum_{i=1}^n t_i + \omega_o \rho (n - b^{-1} \sum_{i=1}^n u_i)))}{nv_o + \tau_o}] \right\};
\end{aligned}$$

como segue o resultado.

A dcc *a posteriori* para  $\omega_o^2$  é tal que

$$\begin{aligned}
 f\left(\omega_o^2 | \boldsymbol{\psi}^{[-\omega_o^2]}, \mathbf{D}\right) &\propto \left[ \prod_{i=1}^n N\left(\mu_o + \omega_o \rho (1 - b^{-1} u_i), \tau_o\right) \right] GI\left(r_{\omega_o^2}, s_{\omega_o^2}\right) \\
 &\propto \left(\frac{1}{\omega_o^2}\right)^{\frac{n}{2}} \exp\left\{-\frac{\sum_{i=1}^n [t_i - (\mu_o + \omega_o \rho (1 - b^{-1} u_i))]^2}{2\omega_o^2 K_o}\right\} \times \\
 &\quad \times \left(\frac{1}{\omega_o^2}\right)^{r_{\omega_o^2}+1} \exp\left\{-\frac{s_{\omega_o^2}}{\omega_o}\right\} \\
 &\propto \left(\frac{1}{\omega_o^2}\right)^{\left(\frac{n}{2}+r_{\omega_o^2}\right)+1} \exp\left\{-\frac{1}{\omega_o} (2K_o)^{-1} \sum_{i=1}^n [t_i - \right. \\
 &\quad \left. - (\mu_o + \omega_o \rho (1 - b^{-1} u_i))]^2 + s_{\omega_o^2}\right\};
 \end{aligned}$$

o que conclui a prova do item (ii)

□

## Anexo B

# Códigos no R/WinBUGS

### B.1 Modelo LST para degradação

Os códigos para implementação do modelo LST no contexto de degradação são apresentados a seguir. Os códigos dos demais modelos serão omitidos, porém são facilmente obtidos pela particularização dos códigos aqui fornecidos.

```
#-----#  
require(sn) # pacote skew normal/skew normal package (Azzalini, 2015)  
setwd("Diretório de trabalho do R/R working directory")  
gerada<-"nome/name"#nome do arquivo que contém os dados de degradação/name  
of the file that contains the degradations data  
dados<-read.table(paste("dados_",gerada,".txt",sep=), sep=, dec=".",head=TRUE)  
#lendo os dados/readind the data (arquivo .txt/.txt file). O arquivo deve conter três  
colunas, a coluna "Unidade"contendo um indicador numérico da unidade sob teste,  
a coluna "Tempo"com os tempos de medição e a coluna "Desgaste"com a respectiva  
degradação./The file must contain three columns, the column "Unidade"contains a nu-  
meric indicator of the unit under test, the column "Tempo"the time measurements and
```

the column "Desgaste" the following degradation.

```

head(dados)
dim(dados)
nperfis <- dados[dim(dados)[1],1] #número de perfis de degradação na base de
dados/number of degradation paths in data base
nmed <- length(dados$Unidade[dados$Unidade == 1]) #número máximo de
observações de degradação em cada perfil/maximum number of measurements for each
degradation path
ntotal <- nperfis*nmed #número total de observações de degradação/total num-
ber of observed degradation measurements
nperfis <- nperfis+1 #inserindo nova unidade na verossimilhança como proposto
em Oliveira (2011)/adding a new unit in likelihood path as proposed by Oliveira (2011)
dados <- rbind(dados,c(nperfis,0,0))
library("R2WinBUGS") #pacote R2WinBugs/ R2WinBugs package (Sturtz et al.,
2005)
require(R2WinBUGS)
require(BRugs) #pacote BRugs/Brugs package (Thomas et al., 2006)
#Ajuste com efeito aleatório  $\beta_i$  log-skew-t/ set up log-skew-t distribution for  $\beta_i$ 
random effects
#Preparação dos dados para o WinBUGS/Preparing data for WinBUGS
bk<-c(-1/2,1/6,0,-1/30,0,1/42,0,-1/30,0,5/66,0,-691/2730,0,7/6,0,-3617/510,0,
43867/798,0,-174611/330)#números de bernoulli de 1 até  $n = 20$  para aproximar
a função trigama através da função de recorrência por expansão em séries de lau-
rent/bernoulli numbers from 1 to  $n = 20$  to approximate the trigamma function by
recurrence using the laurent series expansion (Abramowitz e Stegun, 1965)
indice<-seq(2,2*10,2)# indicador dos números de bernoulli selecionandos para com-
por a série de laurent/indicator of selected bernoulli numbers to be used in laurent series

```

```

data <- list(y = dados$Desgaste, hours = dados$Tempo, units = dados$Unidade,
bk = bk, indice = indice, n1 = ntotal, n2 = nperfis, m = 0, A = 1/(106), p =
0.001, q = 0.001, r = 0.001, s = 0.001, e = 0, f = 200, PI = pi) # dados para entrada no
WinBUGS/input data for WinBUGS

source(file="modelo dlstl.R", local = FALSE) #leitura da sintaxe do modelo/reading
model syntax

#Atribua valores iniciais aos parâmetros/set up initial values of the parameters
inits <- function(){
list(tau1 = tau, mu = log(mu), DELTA = DELTA, TAU = TAU, nu = nu, lbeta =
c(log(betas), mean(log(betas))), w = c(w, w[1]))
} #TAU corresponde a  $\tau_{\beta}^{-1}$  na equação (4.18)/TAU equals  $\tau_{\beta}^{-1}$  in equation (4.18)
# Determinação dos parâmetros a serem monitorados/determination of monitored
parameters

parameters <- c("mu", "omega2", "delta", "nu", "beta", "sigma2")
semente = 1000

dlst.sim <- bugs(data, inits, parameters, model.file,
n.chains = 1, n.iter = 220000000, n.burnin = 100000000, n.thin = 60000, codaPkg =
FALSE, debug = TRUE,

bugs.directory = "C : \\ProgramFiles(x86)\\WinBUGS14", bugs.seed = semente,
working.directory = getwd(), clearWD = FALSE, over.relax = FALSE)

print(dlst.sim)

#-----#

Apresentamos a seguir o arquivo que contém a sintaxe do modelo LST conside-
rando a forma funcional de taxa recíproca para os perfis de degradação degradação,
que corresponde ao arquivo "modelo dlstl.R"./The syntax of the model for the recip-
cal degradation following LST model is presented straight ahead.

#-----#

```

```

modelo <- function(){
for( i in 1 : n1 ) {
y[i] ~ dnorm(media[i],tau1)# verossimilhança/likelihood
media[i] < -hours[i]/(beta[units[i]])# forma funcional do modelo de degrada-
ção/functional form for degradation path
}
for( i in 1 : n2 ) {
lbeta[i] ~ dnorm(mediab[i],tau2[i])# priori para log( $\beta_i$ )/prior for log( $\beta_i$ )
beta[i] < -exp(lbeta[i])
mediab[i] < -mu + DELTA * u[i]/sqrt(w[i])
a[i] ~ dnorm(0,1)
u[i] < -abs(a[i])# priori para  $u_i$ /prior for  $u_i$ 
w[i] ~ dgamma(nuh,nuh)#priori para  $u_i$ /prior for  $u_i$ 
tau2[i] < -w[i] * TAU
}
mu dnorm(m, A)#priori para  $\mu$ /prior for  $\mu$ 
tau1 dgamma(r,s)#priori para  $\tau_1 = (\sigma^2)^{-1}$ /prior for  $\tau_1 = (\sigma^2)^{-1}$ 
sigma2 < -1/tau1
DELTA dnorm(m, A)#priori para  $\Delta_\beta$ /prior for  $\Delta_\beta$ 
TAU dgamma(p,q)#priori para  $TAU = \tau_\beta^{-1}$ /prior for  $TAU = \tau_\beta^{-1}$ 
OMEGA2 < -1/TAU
omega2 < -DELTA * DELTA + OMEGA2
alpha < -DELTA/sqrt(OMEGA2)
delta < -alpha/sqrt(1 + alpha * alpha)
nuh < -nu/2
nug < -(nu + 1)/2

```

*for*(*in1* : 10){*parc*[*i*] < -*bk*[*indice*[*i*]]/*pow*((*nuh* + 3 + 1),(2 \* *i* + 1))}#termos da série de laurent/laurent series terms

*tig4* < -1/(*nuh* + 3 + 1) + 1/(2 \* *pow*(*nuh* + 3 + 1,2)) + *sum*(*parc*[])

#acima foi feita a expansão por série de laurent que fornece ótima aproximação para valores grandes do argumento

#como o argumento pode ser próximo de 0 optei por somar 4 unidades a x e usar a forma recursiva

#*trigama*(*y*)=*trigamma*(*y*+1)+1/(*y*<sup>2</sup>) para encontrar a função digama no ponto do argumento com melhor aproximação

#above it's been set up a laurent series expansion that provides a good approximation for high values of the argument

#since the argument can be a near 0 value, we sum 4 units to it and use the recursive form

#*trigama*(*y*)=*trigamma*(*y*+1)+1/(*y*<sup>2</sup>) to get the digamma function evaluated at the argument point with better approximation

*tig3* < -*tig4* + 1/*pow*((*nuh* + 2 + 1),2)

*tig2* < -*tig3* + 1/*pow*((*nuh* + 1 + 1),2)

*tig1* < -*tig2* + 1/*pow*((*nuh* + 1),2)

*tig* < -*tig1* + 1/*pow*(*nuh*,2)#valor da função trigama no ponto  $\frac{v}{2}$ /trigamma function evaluated at point  $\frac{v}{2}$

*for*(*i* in 1:10){*parcg*[*i*] < -*bk*[*indice*[*i*]]/*pow*((*nug*+3+1),(2\**i*+1))}#termos da série de laurent/laurent series terms

*tig4g* < -1/(*nug*+3+1)+1/(2\**pow*(*nug*+3+1,2))+*sum*(*parcg*[])

#acima foi feita a expansão por série de laurent que fornece ótima aproximação para valores grandes do argumento

#como o argumento pode ser próximo de 0 optei por somar 4 unidades a x e usar a forma recursiva



#trigama(y)=trigamma(y+1)+1/(y<sup>2</sup>) para encontrar a função digama no ponto do argumento com melhor aproximação

#above it's been set up a laurent series expansion that provides a good approximation for high values of the argument

#since the argument can be a near 0 value, we sum 4 units to it and use the recursive form

#trigama(y)=trigamma(y+1)+1/(y<sup>2</sup>) to get the digamma function evaluated at the argument point with better approximation

$$tig3g < -tig4g + 1/pow((nug + 2 + 1),2)$$

$$tig2g < -tig3g + 1/pow((nug + 1 + 1),2)$$

$$tig1g < -tig2g + 1/pow((nug + 1),2)$$

tigg < -tig1g + 1/pow(nug,2)#valor da função trigama no ponto  $\frac{\nu+1}{2}$ /trigamma function evaluated at point  $\frac{\nu+1}{2}$

#zero trick para a priori de nu por Fonseca *et al.* (2008)/zero trick for nu prior in Fonseca *et al.* (2008)

$$zero < -0$$

C < -10000 # grande o bastante para assegurar que  $\phi > 0$ / big enough to ensure  $\phi > 0$

$$L < -pow(nu/(nu + 3),1/2) * sqrt(tig - tigg - 2 * (nu + 3)/(nu * pow(nu + 1,2)))$$

$$lver < - - log(L) + C$$

$$zero \text{ dpois}(lver)$$

$$nu \text{ duni}f(e,f) \text{ # priori 'flat'}$$

}

# algum nome temporário de arquivo/some temporary filename

$$model.file < -file.path(tempdir(),"modelo.bug")$$

# escreve o arquivo do modelo/write model file:

$$write.model(modelo,model.file)$$

```
# mostra o arquivo/shows the file
file.show(model.file)
```

## B.2 Modelo LST para tempos de falha

Os códigos para implementação do modelo LST no contexto de análise convencional de tempo de falha são apresentados a seguir. Os códigos dos demais modelos serão omitidos, porém são facilmente obtidos pela particularização dos códigos aqui fornecidos. Analisamos os dados de câncer de pulmão.

```
# -----#
require(sn)
setwd("Diretório de trabalho do R/R working directory")
require(survival)
attach(veteran)
tempos<-veteran$time
tempos<-c(tempos,-100000)
indicadora<-veteran$status
indicadora<-c(indicadora,0)
n=length(tempos)
t<-tempos
pcensura<-1-mean(indicadora)
t.f<-t
t.f[indicadora==0]<-NA
lt.f<-log(t.f)
t.cen<-t
lt.cen<-log(t.cen)
lt.cen[n]<-100000
lt.cen[indicadora==1]<-100000
```

```

#-----#
#mu~N(m,v)
m<-0
v<-1000000
#tau~rigamma(r,s)
r<-0.001
s<-0.001
#nu~unif(e,f)
e<-0
f<-100
#-----#
u<-rep(0.01,1,length(t))
mu<- mean(log(t[1:(n-1)]))
omega2<- var(log(t[1:(n-1)]))
omega<- sqrt(omega2)
delta<- 0.1
tau<-1/((1-delta2)*omega2)
DELTA<-omega*delta
w<-rep(1,n)
nu<-4
#-----#
bk<-c(-1/2,1/6,0,-1/30,0,1/42,0,-1/30,0,5/66,0,-691/2730,0,7/6,0,-3617/510,0,43867/798,0,-
174611/330)
indice<-seq(2,2*10,2)
data <- list(lt = lt.f,
lt.cen = lt.cen,
n = n,m=m,A=1/v,r=r,s=s,e=e,f=f,bk=bk, indice=indice)

```

```

inits <- function(){
list(mu=mu,DELTA=DELTA,tau=tau,nu=nu,a=abs(u),w=w) }
library("R2WinBUGS")
require(R2WinBUGS)
require(BRugs)
source(file="modelo tvlst.R", local = FALSE)
parameters <- c("mu", "omega2", "delta","nu","lt[138]")
semente=2000

lst.sim <- bugs(data, inits, parameters, model.file,
n.chains = 1, n.iter = 101000000,n.burnin=1000000,n.thin=50000,codaPkg=FALSE,
debug = TRUE,
bugs.directory = "C:\\ProgramFiles(x86)\\WinBUGS14",bugs.seed = semente,
working.directory =getwd() , clearWD = FALSE)
print(lst.sim)
#-----#
Apresentamos a seguir o arquivo que contém a sintaxe do modelo LST para os
tempos de falha, que corresponde ao arquivo "modelo dlstl.R"./The syntax of the model
for the failure times following LST model is presented straight ahead.
#-----#
modelo <- function(){
for( i in 1 : n ) {
lt[i]~dnorm(media[i],tau2[i])
media[i]<-mu+DELTA * u[i]/sqrt(w[i])
a[i] ~ dnorm(0, 1)
u[i] <- abs(a[i])
w[i] ~ dgamma(nuh,nuh)
tau2[i]<-w[i]*tau

```

```

}
mu ~ dnorm(m, A)
DELTA ~ dnorm(m, A)
tau ~ dgamma(r,s)
TAU<-1/tau
omega2<-TAU+DELTA*DELTA
delta<-DELTA/sqrt(omega2)
nuh<-nu/2
nug<-(nu+1)/2
for(i in 1:10){parc[i]<-bk[indice[i]]/pow((nuh+3+1),(2*i+1))}
tig4<-1/(nuh+3+1)+1/(2*pow(nuh+3+1,2))+sum(parc[])
tig3<-tig4+1/pow((nuh+2+1),2)
tig2<-tig3+1/pow((nuh+1+1),2)
tig1<-tig2+1/pow((nuh+1),2)
tig<-tig1+1/pow(nuh,2)
for(i in 1:10){parcg[i]<-bk[indice[i]]/pow((nug+3+1),(2*i+1))}
tig4g<-1/(nug+3+1)+1/(2*pow(nug+3+1,2))+sum(parcg[])
tig3g<-tig4g+1/pow((nug+2+1),2)
tig2g<-tig3g+1/pow((nug+1+1),2)
tig1g<-tig2g+1/pow((nug+1),2)
tigg<-tig1g+1/pow(nug,2)
zero <- 0
C <- 10000
L<-pow(nu/(nu+3),1/2)*sqrt(tig-tigg-2*(nu+3)/(nu*pow(nu+1,2)))
lver <- -log(L) + C
L<- 1/2*(nu/(nu+3))+1/2*(tig-tigg-2*(nu+3)/(nu*pow(nu+1,2)))
lver <- -L + C

```

```

zero ~ dpois(lver)
nu ~ dunif(e,f)
}
model.file <- file.path(tempdir(), "modelo.bug")
model.file <- "modelo.bug"
write.model(modelo, model.file)
file.show(model.file)

```

### B.2.1 Modelo SN Centrado no WinBUGS

A seguir apresentamos o código do WinBUGS do o modelo Skew Normal Centrado para tempos de falha.

```

modelo <- function(){
for( i in 1 : n ) { t[i]~dnorm(media[i],Oc)
#para alfc<0
tbranch[i,1]<-muc+sqrt(wc2)*(-pow(x,(1/3)))*(1-u[i]/sqrt(2/PI))
#para alfc>=0
tbranch[i,2]<-muc+sqrt(wc2)*pow(x,(1/3))*(1-u[i]/sqrt(2/PI))
#condicional 1 para alfc>=0 e 2 para alfc<0
iftbranch[i] < -1 + step(alfc)1parag1 < 0, 2parag1 >= 0
media[i]<-tbranch[i,iftbranch[i]]
a[i] ~ dnorm(0, 1)
u[i] <- abs(a[i])
}
x<-abs(2*alfc/(4-PI))
Oc <- (1/wc2)*((2/PI)/ ((2/PI)+pow(x,(2/3))*((2/PI)-1)))
muc ~ dnorm(m, A)
tauc ~ dgamma(r, s)

```

```
alfc ~ dunif(c, d)
wc2<-1/tauc
mu<-muc-sqrt(wc2)*pow(x,(1/3))
omega2<-wc2*(1+pow(x,(2/3)))
delta<-pow(x,(1/3))/(sqrt(2/PI)*sqrt(1+pow(x,(2/3))))
}
model.file <- file.path(tempdir(), "modelo.bug")
write.model(modelo, model.file)
file.show(model.file)
```

# Anexo C

## Tabelas

Nas Tabelas C.1 à C.12 fornecemos as medidas de qualidade de ajuste referentes a Seção 4.4.

Tabela C.1: Medidas de qualidade de ajuste de modelos baseadas nas degradações  $y_{ij}$  e nos perfis  $y_i$ , dados artificiais

Modelo	B-SN				B-ST			
	LPML	$LPML_{perfil}$	DIC	WAIC	LPML	$LPML_{perfil}$	DIC	WAIC
N	0,1211	2,1457	-63,3535	-58,8737	0,1193	1,8795	-62,5329	-58,1296
LN	0,1217	2,1724	-63,2845	-58,9640	0,1227	2,1652	-63,6129	-59,4080
T	0,1218	2,1685	-63,3758	-59,0986	0,1225	<b>2,2311</b>	-63,5745	-59,2838
LT	0,1219	2,1941	-63,3495	-58,9897	0,1220	2,1462	-63,4997	-59,0700
SN	0,1225	2,1336	-63,5035	-59,2499	0,1208	1,9974	-62,8390	-58,5145
LSN	<b>0,1231</b>	2,1717	<b>-63,7768</b>	<b>-59,5664</b>	0,1226	2,1574	-63,5731	-59,3775
ST	0,1228	2,1906	-63,5918	-59,3714	<b>0,1229</b>	2,2031	<b>-63,7984</b>	<b>-59,5528</b>
LST	0,1225	<b>2,2060</b>	-63,6007	-59,3612	0,1216	2,2299	-63,3328	-58,8992
W	0,1212	2,1638	-63,1697	-58,7845	0,1213	2,1119	-63,1070	-58,7549

Ajuste	B-LST				B-W			
	LPML	$LPML_{perfil}$	DIC	WAIC	LPML	$LPML_{perfil}$	DIC	WAIC
N	0,1210	2,5405	-63,2091	-58,7607	0,1225	2,6524	-63,5684	-59,2638
LN	0,1213	2,6178	-63,0195	-58,6755	0,1219	2,6803	-63,3650	-59,0409
T	0,1218	<b>2,6639</b>	-63,4785	-59,0355	0,1227	2,6402	-63,6354	-59,3845
LT	0,1220	2,5016	-63,3755	-59,0566	0,1222	2,6681	-63,4814	-59,1538
SN	0,1216	2,5692	-63,2438	-58,9213	0,1203	2,4998	-62,8704	-58,4133
LSN	0,1216	2,5827	-63,2389	-58,9481	<b>0,1234</b>	<b>2,7224</b>	<b>-63,8897</b>	<b>-59,7411</b>
ST	0,1212	2,6222	-63,0496	-58,6887	0,1201	2,5318	-62,8116	-58,2644
LST	<b>0,1223</b>	2,6036	<b>-63,6137</b>	<b>-59,2530</b>	0,1217	2,5872	-63,4102	-59,0020
W	0,1209	2,5690	-62,8910	-58,5031	0,1206	2,2783	-62,9392	-58,6322



Tabela C.2: Medidas de qualidade de ajuste de modelos baseadas nos tempos de falha (PTF) reais, dados artificiais

Modelo	B-SN			B-ST		
	LPML	DIC	WAIC	LPML	DIC	WAIC
N	-8,9826	534,0697	282,4785	-10,9079	626,2453	441,0301
LN	-8,8592	527,4625	280,8611	-9,7259	568,1159	389,7868
T	-9,0148	535,3093	289,5755	-8,9939	532,1807	394,9618
LT	-8,9070	530,1250	285,9273	<b>-8,9533</b>	<b>531,2741</b>	393,8657
SN	<b>-8,8143</b>	<b>523,6340</b>	<b>275,0983</b>	-10,6465	595,7661	409,3599
LSN	-8,8446	525,9056	278,6815	-9,6436	561,7001	<b>355,4042</b>
ST	-8,8448	524,9650	278,0805	-9,1519	533,7326	383,1838
LST	-8,8989	528,5043	282,2210	-9,0510	533,4785	380,9957
W	-8,9994	534,9831	280,2596	-10,1891	589,8157	368,1006
Ajuste	B-LST			B-W		
	LPML	DIC	WAIC	LPML	DIC	WAIC
N	-9,5511	551,7562	354,2747	-8,4181	501,2622	260,6508
LN	-9,0251	526,3243	343,1260	-8,4590	502,9074	263,1730
T	-8,5499	505,2040	332,8158	-8,4656	503,9144	263,6992
LT	-8,5083	503,4263	332,4536	-8,4863	504,7932	265,3587
SN	-8,9778	523,0999	322,9940	-9,0902	504,4282	259,8556
LSN	-8,6793	509,3661	312,8246	-8,4549	502,7386	<b>259,2077</b>
ST	-8,4583	<b>497,6399</b>	<b>304,7936</b>	-8,6758	507,0431	261,5920
LST	<b>-8,4445</b>	499,0590	309,1387	-8,4981	505,1979	260,7756
W	-9,4146	548,4669	318,8850	<b>-8,4059</b>	<b>500,7770</b>	259,7861

Tabela C.3: Medidas de qualidade de ajuste de modelos baseadas nos pseudotempos de falha  $\hat{t}_i$ , dados artificiais

Ajuste	B-SN			B-ST		
	LPML	DIC	WAIC	LPML	DIC	WAIC
N	-8,9804	534,0018	282,2347	-10,8647	625,1808	431,9699
LN	-8,8621	527,6598	281,0950	-9,7196	568,1861	387,1167
T	-9,0162	535,3835	290,0863	-9,0274	534,2539	395,3182
LT	-8,9106	530,3399	286,4916	<b>-8,9836</b>	<b>533,1309</b>	394,4154
SN	<b>-8,8019</b>	<b>523,7457</b>	<b>274,9967</b>	-10,5982	595,0245	402,4525
LSN	-8,8432	526,2199	278,7557	-9,6339	561,9405	<b>353,8696</b>
ST	-8,8371	525,1909	278,1316	-9,1750	535,2339	383,0486
LST	-8,9007	528,8884	282,4183	-9,0752	534,9790	381,0978
W	-8,9947	534,8832	279,8786	-10,1658	589,6113	364,9605
Ajuste	B-LST			B-W		
	LPML	DIC	WAIC	LPML	DIC	WAIC
N	-9,5617	551,9403	356,2024	-8,4518	503,0943	261,9678
LN	-9,0353	526,7135	344,4547	-8,4910	504,7288	264,5145
T	-8,5555	505,4361	334,6485	-8,5032	506,0028	265,2583
LT	-8,5158	503,8110	334,2192	-8,5228	506,8982	266,9720
SN	-8,9837	523,5626	324,2889	-9,3817	505,4883	261,5477
LSN	-8,6921	510,0756	313,7699	-8,4835	504,2636	<b>260,3017</b>
ST	-8,4693	<b>498,7989</b>	<b>305,4941</b>	-8,8393	508,1281	262,5317
LST	<b>-8,4599</b>	500,0034	310,1739	-8,5276	506,7626	261,8846
W	-9,4283	548,7190	320,0286	<b>-8,4418</b>	<b>502,5802</b>	261,0373

Tabela C.4: Fator de Bayes baseado nas degradações  $y_{ij}$  e nos perfis  $y_i$ , dados artificiais

Cenário	Numerador	Denominador								
		N	LN	T	LT	SN	LSN	ST	LST	W
B-SN	N	1,00	0,08	0,06	0,26	0,05	0,17	0,48	1,57	0,22
	LN	12,14	1,00	0,75	3,17	0,55	2,10	5,81	19,04	2,73
	T	16,23	1,34	1,00	4,24	0,74	2,81	7,76	25,46	3,65
	LT	3,83	0,32	0,24	1,00	0,17	0,66	1,83	6,00	0,86
	SN	<b>22,04</b>	<b>1,82</b>	<b>1,36</b>	<b>5,76</b>	<b>1,00</b>	<b>3,82</b>	<b>10,54</b>	<b>34,57</b>	<b>4,95</b>
	LSN	5,77	0,48	0,36	1,51	0,26	1,00	2,76	9,05	1,30
	ST	2,09	0,17	0,13	0,55	0,09	0,36	1,00	3,28	0,47
	LST	0,64	0,05	0,04	0,17	0,03	0,11	0,30	1,00	0,14
	W	4,45	0,37	0,27	1,16	0,20	0,77	2,13	6,98	1,00
B-ST	N	1,00	0,16	0,48	0,14	0,70	3,81	1,05	0,11	0,33
	LN	6,38	1,00	3,08	0,92	4,49	24,30	6,72	0,70	2,12
	T	2,07	0,32	1,00	0,30	1,46	7,89	2,18	0,23	0,69
	LT	6,96	1,09	3,36	1,00	4,90	26,49	7,33	0,76	2,31
	SN	1,42	0,22	0,69	0,20	1,00	5,41	1,50	0,16	0,47
	LSN	0,26	0,04	0,13	0,04	0,18	1,00	0,28	0,03	0,09
	ST	0,95	0,15	0,46	0,14	0,67	3,61	1,00	0,10	0,31
	LST	<b>9,13</b>	<b>1,43</b>	<b>4,40</b>	<b>1,31</b>	<b>6,42</b>	<b>34,76</b>	<b>9,62</b>	<b>1,00</b>	<b>3,03</b>
	W	3,02	0,47	1,46	0,43	2,12	11,49	3,18	0,33	1,00
B-LST	N	1,00	0,57	0,55	55,28	6,38	0,49	0,87	0,42	0,29
	LN	1,76	1,00	0,96	97,33	11,24	0,86	1,54	0,74	0,50
	T	1,83	1,04	1,00	101,37	11,71	0,89	1,60	0,77	0,53
	LT	0,02	0,01	0,01	1,00	0,12	0,01	0,02	0,01	0,01
	SN	0,16	0,09	0,09	8,66	1,00	0,08	0,14	0,07	0,04
	LSN	2,05	1,16	1,12	113,35	13,09	1,00	1,79	0,86	0,59
	ST	1,14	0,65	0,62	63,27	7,31	0,56	1,00	0,48	0,33
	LST	2,38	1,35	1,30	131,47	15,18	1,16	2,08	1,00	0,68
	W	<b>3,49</b>	<b>1,98</b>	<b>1,90</b>	<b>192,79</b>	<b>22,26</b>	<b>1,70</b>	<b>3,05</b>	<b>1,47</b>	<b>1,00</b>
B-W	N	1,00	0,98	0,65	1,32	2,63	0,71	4,48	0,54	0,44
	LN	1,02	1,00	0,66	1,34	2,68	0,73	4,56	0,55	0,45
	T	1,55	1,52	1,00	2,04	4,08	1,10	6,94	0,83	0,68
	LT	0,76	0,75	0,49	1,00	2,00	0,54	3,40	0,41	0,33
	SN	0,38	0,37	0,25	0,50	1,00	0,27	1,70	0,20	0,17
	LSN	1,40	1,38	0,91	1,85	3,69	1,00	6,28	0,76	0,61
	ST	0,22	0,22	0,14	0,29	0,59	0,16	1,00	0,12	0,10
	LST	1,86	1,82	1,20	2,44	4,89	1,32	8,32	1,00	0,81
	W	<b>2,29</b>	<b>2,24</b>	<b>1,47</b>	<b>3,01</b>	<b>6,02</b>	<b>1,63</b>	<b>10,23</b>	<b>1,23</b>	<b>1,00</b>

Tabela C.5: Fator de Bayes baseado nos tempos de falha  $t_i$ , dados artificiais

Cenário	Numerador	Denominador								
		N	LN	T	LT	SN	LSN	ST	LST	W
B-SN	N	1,00	1,35	2,01	0,68	0,19	0,35	0,42	4,47	1,47
	LN	0,74	1,00	1,49	0,51	0,14	0,26	0,31	3,32	1,10
	T	0,50	0,67	1,00	0,34	0,09	0,17	0,21	2,23	0,73
	LT	1,47	1,97	2,94	1,00	0,27	0,51	0,61	6,55	2,16
	SN	<b>5,34</b>	<b>7,18</b>	<b>10,71</b>	<b>3,64</b>	<b>1,00</b>	<b>1,86</b>	<b>2,23</b>	<b>23,87</b>	<b>7,87</b>
	LSN	2,87	3,86	5,75	1,96	0,54	1,00	1,20	12,81	4,23
	ST	2,40	3,22	4,81	1,64	0,45	0,84	1,00	10,71	3,53
	LST	0,22	0,30	0,45	0,15	0,04	0,08	0,09	1,00	0,33
	W	0,68	0,91	1,36	0,46	0,13	0,24	0,28	3,03	1,00
B-ST	N	1,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00
	LN	822740,13	1,00	0,00	0,00	2470,26	0,12	0,00	0,00	167,04
	T	8002931032,74	9727,17	1,00	0,57	24028662,67	1125,15	28,28	48,59	1624781,99
	LT	<b>1408788829,99</b>	<b>17123,13</b>	<b>1,76</b>	<b>1,00</b>	<b>42298643,72</b>	<b>1980,65</b>	<b>49,78</b>	<b>85,54</b>	<b>2860170,59</b>
	SN	333,06	0,00	0,00	0,00	1,00	0,00	0,00	0,00	0,07
	LSN	7112772,66	8,65	0,00	0,00	21355,98	1,00	0,03	0,04	1444,06
	ST	283028552,75	344,01	0,04	0,02	849788,36	39,79	1,00	1,72	57461,41
	LST	164699398,57	200,18	0,02	0,01	494507,11	23,16	0,58	1,00	33437,83
	W	4925,54	0,01	0,00	0,00	14,79	0,00	0,00	0,00	1,00
B-LST	N	1,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,64
	LN	702,82	1,00	0,01	0,00	1,29	0,39	0,00	0,00	447,90
	T	56789,31	80,80	1,00	0,35	104,34	31,23	0,38	0,38	36190,71
	LT	<b>160532,51</b>	<b>228,41</b>	<b>2,83</b>	<b>1,00</b>	<b>294,96</b>	<b>88,27</b>	<b>1,07</b>	<b>1,07</b>	<b>102304,20</b>
	SN	544,25	0,77	0,01	0,00	1,00	0,30	0,00	0,00	346,84
	LSN	1818,66	2,59	0,03	0,01	3,34	1,00	0,01	0,01	1159,00
	ST	149675,46	212,96	2,64	0,93	275,01	82,30	1,00	1,00	95385,21
	LST	149657,27	212,94	2,64	0,93	274,98	82,29	1,00	1,00	95373,63
	W	1,57	0,00	0,00	0,00	0,00	0,00	0,00	0,00	1,00
B-W	N	<b>1,00</b>	<b>1,52</b>	<b>6,58</b>	<b>4,11</b>	<b>972,65</b>	<b>2,16</b>	<b>161,01</b>	<b>6,05</b>	<b>1,66</b>
	LN	0,66	1,00	4,32	2,70	638,58	1,41	105,71	3,97	1,09
	T	0,15	0,23	1,00	0,62	147,74	0,33	24,46	0,92	0,25
	LT	0,24	0,37	1,60	1,00	236,40	0,52	39,13	1,47	0,40
	SN	0,00	0,00	0,01	0,00	1,00	0,00	0,17	0,01	0,00
	LSN	0,46	0,71	3,05	1,91	451,34	1,00	74,71	2,81	0,77
	ST	0,01	0,01	0,04	0,03	6,04	0,01	1,00	0,04	0,01
	LST	0,17	0,25	1,09	0,68	160,66	0,36	26,60	1,00	0,27
	W	0,60	0,92	3,96	2,48	585,79	1,30	96,97	3,65	1,00

Tabela C.6: Fator de Bayes baseado nos pseudo tempos de falha  $\hat{t}_i$ , dados artificiais

Cenário	Numerador	Denominador								
		N	LN	T	LT	SN	LSN	ST	LST	W
B-SN	N	1,00	1,22	1,77	0,63	0,18	0,37	0,38	4,63	1,46
	LN	0,82	1,00	1,45	0,52	0,15	0,30	0,31	3,80	1,20
	T	0,56	0,69	1,00	0,36	0,10	0,21	0,21	2,61	0,83
	LT	1,58	1,93	2,81	1,00	0,29	0,58	0,60	7,33	2,32
	SN	<b>5,47</b>	<b>6,67</b>	<b>9,69</b>	<b>3,46</b>	<b>1,00</b>	<b>2,01</b>	<b>2,06</b>	<b>25,34</b>	<b>8,01</b>
	LSN	2,72	3,31	4,82	1,72	0,50	1,00	1,02	12,59	3,98
	ST	2,65	3,23	4,70	1,68	0,49	0,98	1,00	12,29	3,89
	LST	0,22	0,26	0,38	0,14	0,04	0,08	0,08	1,00	0,32
	W	0,68	0,83	1,21	0,43	0,12	0,25	0,26	3,16	1,00
B-ST	N	1,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00
	LN	511918,89	1,00	0,00	0,00	1185,01	0,11	0,00	0,00	148,01
	T	4641016486,90	9065,92	1,00	0,62	10743172,87	989,02	25,31	38,68	1341876,43
	LT	<b>7458990456,04</b>	<b>14570,65</b>	<b>1,61</b>	<b>1,00</b>	<b>17266308,82</b>	<b>1589,55</b>	<b>40,69</b>	<b>62,16</b>	<b>2156648,98</b>
	SN	432,00	0,00	0,00	0,00	1,00	0,00	0,00	0,00	0,12
	LSN	4692524,42	9,17	0,00	0,00	10862,41	1,00	0,03	0,04	1356,77
	ST	183333408,08	358,13	0,04	0,02	424386,02	39,07	1,00	1,53	53007,95
	LST	119987468,41	234,39	0,03	0,02	277750,82	25,57	0,65	1,00	34692,48
	W	3458,60	0,01	0,00	0,00	8,01	0,00	0,00	0,00	1,00
B-LST	N	1,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,64
	LN	662,92	1,00	0,01	0,00	1,22	0,39	0,00	0,00	422,74
	T	67519,36	101,85	1,00	0,38	123,92	39,47	0,42	0,48	43057,10
	LT	<b>176999,48</b>	<b>267,00</b>	<b>2,62</b>	<b>1,00</b>	<b>324,86</b>	<b>103,47</b>	<b>1,10</b>	<b>1,26</b>	<b>112872,58</b>
	SN	544,85	0,82	0,01	0,00	1,00	0,32	0,00	0,00	347,45
	LSN	1710,69	2,58	0,03	0,01	3,14	1,00	0,01	0,01	1090,91
	ST	161141,82	243,08	2,39	0,91	295,76	94,20	1,00	1,15	102760,15
	LST	140417,62	211,82	2,08	0,79	257,72	82,08	0,87	1,00	89544,33
	W	1,57	0,00	0,00	0,00	0,00	0,00	0,00	0,00	1,00
B-W	N	<b>1,00</b>	<b>1,45</b>	<b>6,73</b>	<b>4,33</b>	<b>40778,96</b>	<b>1,93</b>	<b>131,18</b>	<b>5,62</b>	<b>1,51</b>
	LN	0,69	1,00	4,65	3,00	28188,81	1,34	90,68	3,88	1,05
	T	0,15	0,22	1,00	0,64	6060,96	0,29	19,50	0,83	0,22
	LT	0,23	0,33	1,55	1,00	9411,31	0,45	30,28	1,30	0,35
	SN	0,00	0,00	0,00	0,00	1,00	0,00	0,00	0,00	0,00
	LSN	0,52	0,75	3,48	2,24	21084,06	1,00	67,82	2,90	0,78
	ST	0,01	0,01	0,05	0,03	310,86	0,01	1,00	0,04	0,01
	LST	0,18	0,26	1,20	0,77	7262,09	0,34	23,36	1,00	0,27
	W	0,66	0,96	4,45	2,86	26949,07	1,28	86,69	3,71	1,00

Tabela C.7: Teste de Kolmogorov-Smirnov baseado nos tempos  $t_i$  e pseudotempos de falha  $\hat{t}_i$ : comparação com a distribuição empírica no cenário B-SN

Ajuste		$t_i$			$\hat{t}_i$		
		Real	Hamada	Prop	$Real_{pseudo}$	$Hamada_{pseudo}$	$Prop_{pseudo}$
N	$D_n$		0,19471	0,19483		0,17719	0,17833
	p		0,54592	0,54821		0,66258	0,65666
LN	$D_n$		0,17319	0,17783		0,15112	0,15033
	p		0,69020	0,65870		0,82830	0,83263
T	$D_n$		0,13950	0,13633		0,13396	0,13633
	p		0,88762	0,90390		0,91332	0,90343
LT	$D_n$		0,14541	0,14333		0,12101	0,12333
	p		0,85956	0,87185		0,95799	0,95067
SN	$D_n$	0,12767	0,15354	0,15983	0,11215	0,13326	0,13733
	p	0,93749	0,81387	0,77608	0,97806	0,91570	0,89820
LSN	$D_n$		0,14173	0,14233		0,11974	0,12133
	p		0,87644	0,87571		0,96172	0,95706
ST	$D_n$		<b>0,12664</b>	<b>0,10733</b>		<b>0,10561</b>	0,11900
	p		<b>0,94102</b>	<b>0,98642</b>		<b>0,98837</b>	0,96386
LST	$D_n$		0,13208	0,13033		0,10937	<b>0,10983</b>
	p		0,92120	0,92872		0,98327	<b>0,98271</b>
W	$D_n$		0,19254	0,19933		0,17653	0,18433
	p		0,56301	0,51952		0,66807	0,61653

Tabela C.8: Teste de Kolmogorov-Smirnov baseado nos tempos  $t_i$  e pseudotempos de falha  $\hat{t}_i$ : comparação com a distribuição empírica no cenário B-ST

Ajuste		$t_i$			$\hat{t}_i$		
		Real	Hamada	Prop	$Real_{pseudo}$	$Hamada_{pseudo}$	$Prop_{pseudo}$
N	$D_n$		0,38670	0,39767		0,38409	0,39367
	p		0,01566	0,01188		0,01655	0,01320
LN	$D_n$		0,28438	0,29650		0,27902	0,29067
	p		0,14345	0,11480		0,15673	0,12655
T	$D_n$		0,16461	0,15083		0,16633	0,15233
	p		0,74620	0,83088		0,73753	0,81997
LT	$D_n$		0,15874	0,16833		0,16069	0,17133
	p		0,78523	0,72294		0,77068	0,70129
SN	$D_n$		0,48117	0,48817		0,47735	0,48417
	p		0,00111	0,00074		0,00119	0,00078
LSN	$D_n$		0,26861	0,27467		0,26363	0,27267
	p		0,18810	0,16861		0,20354	0,17765
ST	$D_n$	0,19904	0,15343	0,16183	0,20777	0,13711	0,14367
	p	0,52162	0,81457	0,76306	0,46617	0,89862	0,86761
LST	$D_n$		<b>0,14812</b>	<b>0,13133</b>		<b>0,13340</b>	<b>0,12717</b>
	p		<b>0,84622</b>	<b>0,92370</b>		<b>0,91533</b>	<b>0,93897</b>
W	$D_n$		0,33386	0,33767		0,33033	0,33517
	p		0,05362	0,04867		0,05727	0,05272

Tabela C.9: Teste de Kolmogorov-Smirnov baseado nos tempos  $t_i$  e pseudotempos de falha  $\hat{t}_i$ : comparação com a distribuição empírica no cenário B-LST

Ajuste		$t_i$			$\hat{t}_i$		
		Real	Hamada	Prop	$Real_{pseudo}$	$Hamada_{pseudo}$	$Prop_{pseudo}$
N	$D_n$		0,32346	0,31950		0,32379	0,31867
	p		0,06728	0,07196		0,06573	0,07317
LN	$D_n$		0,25646	0,25450		0,25423	0,25200
	p		0,22976	0,23935		0,23964	0,24742
T	$D_n$		0,14095	0,13483		0,14285	0,13633
	p		0,88212	0,90971		0,87097	0,90362
LT	$D_n$		<b>0,12797</b>	<b>0,11833</b>		0,13006	0,11983
	p		<b>0,93568</b>	<b>0,96591</b>		0,92773	0,96259
SN	$D_n$		0,38123	0,39717		0,37913	0,39717
	p		0,01745	0,01166		0,01900	0,01221
LSN	$D_n$		0,25985	0,26450		0,25746	0,26400
	p		0,21829	0,19923		0,22824	0,20193
ST	$D_n$		0,14584	0,14467		0,14257	0,14217
	p		0,85873	0,86163		0,87620	0,87528
LST	$D_n$	0,15257	0,13959	0,13917	0,17888	<b>0,12544</b>	<b>0,11933</b>
	p	0,81947	0,88937	0,89079	0,65564	<b>0,94439</b>	<b>0,96399</b>
W	$D_n$		0,31696	0,31267		0,31726	0,31267
	p		0,07492	0,08218		0,07494	0,08414

Tabela C.10: Teste de Kolmogorov-Smirnov baseado nos tempos  $t_i$  e pseudotempos de falha  $\hat{t}_i$ : comparação com a distribuição empírica no cenário B-W

Ajuste		$t_i$			$\hat{t}_i$		
		Real	Hamada	Prop	$Real_{pseudo}$	$Hamada_{pseudo}$	$Prop_{pseudo}$
N	$D_n$		<b>0,084946</b>	0,089833		<b>0,079707</b>	0,090167
	p		<b>0,99936</b>	0,99845		<b>0,99973</b>	0,99837
LN	$D_n$		0,096668	0,101667		0,094398	0,110167
	p		0,99551	0,9924		0,99707	0,9832
T	$D_n$		0,088903	<b>0,085833</b>		0,080146	<b>0,075667</b>
	p		0,99876	<b>0,99917</b>		0,99972	<b>0,9999</b>
LT	$D_n$		0,10545	0,106		0,103455	0,1035
	p		0,98923	0,98788		0,99053	0,9908
SN	$D_n$		0,134161	0,155333		0,111676	0,128167
	p		0,91383	0,80223		0,97827	0,93583
LSN	$D_n$		0,110217	0,111667		0,110717	0,111667
	p		0,98138	0,97975		0,98162	0,97984
ST	$D_n$		0,132479	0,144333		0,112021	0,120167
	p		0,92012	0,86517		0,97918	0,96202
LST	$D_n$		0,104653	0,106333		0,105104	0,100667
	p		0,98959	0,98792		0,98873	0,99306
W	$D_n$	0,089781	0,118906	0,129833	0,106803	0,105836	0,117833
	p	0,99861	0,96492	0,93188	0,98744	0,98808	0,96656

As medidas de qualidade de ajuste da análise das rodas de trem na Seção 4.5.1 são apresentadas nas Tabelas C.13 à C.17.

Tabela C.11: Teste de K-S comparando as distribuições estimadas e a real baseado nos tempos de falha e pseudotempos de falha, todos os cenários

Cenário	Ajuste	Hamada	Prop	Hamada <sub>pseudo</sub>	Prop <sub>pseudo</sub>	Cenário	Ajuste	Hamada	Prop	Hamada <sub>pseudo</sub>	Prop <sub>pseudo</sub>
B-SN	N	0,067041	0,067572	0,069186	0,070327	B-ST	N	0,379048	0,388441	0,379808	0,389386
	LN	0,049706	0,061774	0,049896	0,062591		LN	0,275153	0,286941	0,274737	0,286386
	T	0,049874	0,045991	0,049928	0,045318		T	0,073495	0,084585	0,074146	0,08525
	LT	0,045061	0,059491	0,045072	0,059318		LT	<b>0,072545</b>	<b>0,063585</b>	<b>0,073152</b>	<b>0,06375</b>
	SN	<b>0,028316</b>	0,034774	<b>0,028311</b>	0,037159		SN	0,471947	0,478941	0,473071	0,479886
	LSN	0,030633	<b>0,026255</b>	0,030625	<b>0,026022</b>		LSN	0,259388	0,265441	0,259335	0,268386
	ST	0,041496	0,032534	0,041473	0,033838		ST	0,106518	0,107085	0,106789	0,107608
B-LST	LST	0,046252	0,048034	0,046073	0,049338	LST	0,099672	0,092085	0,098788	0,09025	
	W	0,077619	0,071663	0,077525	0,076327	W	0,324635	0,328441	0,326052	0,330886	
	N	0,299627	0,298134	0,299638	0,298342	N	0,048342	0,050818	0,048972	0,05327	
	LN	0,235095	0,233134	0,235075	0,232842	LN	0,069835	0,068667	0,070038	0,068506	
	T	<b>0,063529</b>	<b>0,073829</b>	<b>0,063271</b>	<b>0,073453</b>	T	0,05544	<b>0,049667</b>	0,055152	<b>0,050506</b>	
B-LST	LT	0,072477	0,085329	0,073207	0,079453	LT	0,078618	0,079167	0,07846	0,078506	
	SN	0,35986	0,373335	0,359977	0,373012	SN	0,092602	0,109613	0,094451	0,110636	
	LSN	0,238482	0,243134	0,2383	0,244842	LSN	0,066939	0,075952	0,066757	0,074136	
	ST	0,122325	0,1201	0,122285	0,120963	ST	0,093428	0,097452	0,094309	0,099136	
	LST	0,112944	0,124094	0,112929	0,123682	LST	0,077594	0,084818	0,077456	0,08477	
	W	0,29313	0,288835	0,293102	0,288512	W	<b>0,047639</b>	0,054952	<b>0,046991</b>	0,058636	

Tabela C.12: Distância KS com particionamento dos espaço amostral dos tempos de falha para comparação dos modelos ajustados com a distribuição real, todos os cenários

Modelo	B-SN		B-ST		B-LST		B-W	
	Hamada	Prop	Hamada	Prop	Hamada	Prop	Hamada	Prop
N	0,07388	0,082504	0,393199	0,396701	0,299709	0,30129	0,049015	0,053336
LN	0,049984	0,062947	0,275438	0,290333	0,235096	0,234031	0,071331	0,070041
T	0,093863	0,093745	0,151522	0,131592	0,142878	0,145658	0,056248	<b>0,051594</b>
LT	0,045862	0,062025	0,132965	0,13894	0,12355	<b>0,114492</b>	0,07877	0,080056
SN	<b>0,028554</b>	0,038203	0,475674	0,486898	0,36024	0,374508	0,094498	0,114302
LSN	0,030633	<b>0,027906</b>	0,25939	0,274499	0,239451	0,248575	0,067032	0,077595
ST	0,041514	0,036881	0,108842	0,110945	0,122452	0,123858	0,094556	0,102007
LST	0,04764	0,050564	<b>0,102597</b>	<b>0,105338</b>	<b>0,112945</b>	0,124391	0,077642	0,08537
W	0,082086	0,0831	0,330847	0,336246	0,293639	0,296333	<b>0,047839</b>	0,061436



Tabela C.13: Qualidade de ajuste baseado na degradação (Deg) e pseudotempos de falha (PTF), Rodas de Trem

Obs	Modelo	MA31				MA12				MA11			
		LPML	LPML <sub>perfil</sub>	DIC	WAIC	LPML	LPML <sub>perfil</sub>	DIC	WAIC	LPML	LPML <sub>perfil</sub>	DIC	WAIC
Deg	N	-0,7439	<b>-9,3484</b>	255,6032	249,5636	-0,7874	<b>-9,5782</b>	<b>268,4404</b>	<b>250,7277</b>	-1,5263	-19,8679	516,1642	497,5350
	LN	-0,7448	-9,5426	256,0184	249,8359	-0,7950	-9,6077	268,5703	250,7785	-1,5094	-19,8599	515,6351	496,1909
	T	<b>-0,7433</b>	-9,3578	<b>255,5585</b>	<b>249,4012</b>	<b>-0,7795</b>	-9,7395	268,8255	251,3791	-1,5163	-19,7678	516,6397	498,1032
	LT	-0,7437	-9,5296	255,7401	249,5129	-0,7998	-9,6783	268,8872	251,2814	<b>-1,4935</b>	<b>-19,4444</b>	<b>515,5063</b>	<b>495,9683</b>
	SN	-0,7442	-9,5923	255,6970	249,6137	-0,8002	-10,0950	268,9764	251,6846	-1,5081	-20,0559	516,4764	497,8016
	LSN	-0,7440	-9,6461	255,6245	249,5599	-0,8000	-10,0585	269,0174	251,6863	-1,5036	-19,7706	516,3173	497,4901
	ST	-0,7450	-9,4938	255,9823	249,8455	-0,7984	-9,9774	269,1815	251,4600	-1,5009	-19,6327	515,6925	496,3006
	LST	-0,7441	-9,4994	255,7498	249,5844	-0,8025	-9,9592	269,4032	251,7311	-1,5151	-19,6911	516,0843	497,0523
	W	-0,7438	-9,6452	255,7214	249,5738	-0,7984	-9,7191	268,9378	251,6743	-1,5010	-19,6189	515,7940	496,9094
	PTF	N	-8,4016	-	453,8406	276,8087	-8,2048	-	453,4203	248,0447	-7,9117	-	438,9808
LN		-7,7700	-	429,2944	255,2812	<b>-7,8378</b>	-	<b>434,7077</b>	245,1622	-7,8411	-	435,0261	232,2823
T		-7,9908	-	436,3538	277,6596	-8,0572	-	445,3126	267,8679	-7,9696	-	441,3401	237,6883
LT		-7,8002	-	429,4459	261,0900	-7,8907	-	437,1197	249,9080	-7,8915	-	437,6296	235,9206
SN		-7,9437	-	434,0069	251,8414	-7,9204	-	438,6348	<b>238,5702</b>	-7,8386	-	434,9659	<b>226,5135</b>
LSN		<b>-7,7260</b>	-	<b>425,6686</b>	<b>246,9074</b>	-7,8818	-	436,7402	242,4806	-7,8772	-	435,8439	228,5815
ST		-7,7847	-	426,7041	247,9472	-7,8962	-	437,1459	241,6820	-7,9040	-	437,0685	228,9553
LST		-7,7596	-	426,5063	249,8785	-7,9661	-	439,3449	246,8971	-7,9168	-	438,3373	230,5342
W		-8,0723	-	441,5259	255,3836	-7,9144	-	438,5667	242,3704	<b>-7,8221</b>	-	<b>434,1471</b>	228,8892

Tabela C.14: Fator de Bayes baseado na degradação (Deg) e pseudo tempos de falha (PTF), Roda MA31

Obs	Numerador	Denominador								
		N	LN	T	LT	SN	LSN	ST	LST	W
Deg	N	1,00	15,24	1,70	11,44	1,03	1,60	1,64	0,90	1,24
	LN	0,07	1,00	0,11	0,75	0,07	0,10	0,11	0,06	0,08
	T	0,59	8,99	1,00	6,75	0,61	0,94	0,97	0,53	0,73
	LT	0,09	1,33	0,15	1,00	0,09	0,14	0,14	0,08	0,11
	SN	0,97	14,75	1,64	11,07	1,00	1,54	1,59	0,87	1,20
	LSN	0,63	9,56	1,06	7,17	0,65	1,00	1,03	0,57	0,78
	ST	0,61	9,28	1,03	6,97	0,63	0,97	1,00	0,55	0,76
	LST	<b>1,11</b>	<b>16,88</b>	<b>1,88</b>	<b>12,67</b>	<b>1,14</b>	<b>1,77</b>	<b>1,82</b>	<b>1,00</b>	<b>1,38</b>
	W	0,81	12,27	1,36	9,21	0,83	1,28	1,32	0,73	1,00
PTF	N	1,00	0,00	0,01	0,00	0,02	0,01	0,00	0,00	0,03
	LN	264,15	1,00	3,40	0,48	5,95	3,14	1,16	0,58	7,83
	T	77,65	0,29	1,00	0,14	1,75	0,92	0,34	0,17	2,30
	LT	<b>547,45</b>	<b>2,07</b>	<b>7,05</b>	<b>1,00</b>	<b>12,33</b>	<b>6,51</b>	<b>2,40</b>	<b>1,20</b>	<b>16,23</b>
	SN	44,38	0,17	0,57	0,08	1,00	0,53	0,19	0,10	1,32
	LSN	84,14	0,32	1,08	0,15	1,90	1,00	0,37	0,19	2,49
	ST	227,77	0,86	2,93	0,42	5,13	2,71	1,00	0,50	6,75
	LST	454,38	1,72	5,85	0,83	10,24	5,40	1,99	1,00	13,47
	W	33,73	0,13	0,43	0,06	0,76	0,40	0,15	0,07	1,00

Tabela C.15: Fator de Bayes baseado na degradação (Deg) e pseudo tempos de falha (PTF), Roda MA12

Obs	Numerador	Denominador								
		N	LN	T	LT	SN	LSN	ST	LST	W
Deg	N	<b>1,00</b>	<b>9,89</b>	<b>4,09</b>	<b>2,28</b>	<b>2,23</b>	<b>2,07</b>	<b>3,09</b>	<b>221,59</b>	<b>5,81</b>
	LN	0,10	1,00	0,41	0,23	0,23	0,21	0,31	22,41	0,59
	T	0,24	2,42	1,00	0,56	0,55	0,51	0,76	54,20	1,42
	LT	0,44	4,33	1,79	1,00	0,98	0,91	1,35	97,01	2,54
	SN	0,45	4,43	1,83	1,02	1,00	0,93	1,38	99,22	2,60
	LSN	0,48	4,78	1,98	1,10	1,08	1,00	1,49	107,05	2,81
	ST	0,32	3,20	1,32	0,74	0,72	0,67	1,00	71,65	1,88
	LST	0,00	0,04	0,02	0,01	0,01	0,01	0,01	1,00	0,03
	W	0,17	1,70	0,70	0,39	0,38	0,36	0,53	38,15	1,00
PTF	N	1,00	0,01	0,15	0,03	0,10	0,01	0,06	0,08	0,01
	LN	<b>90,00</b>	<b>1,00</b>	<b>13,20</b>	<b>2,62</b>	<b>8,90</b>	<b>1,30</b>	<b>5,45</b>	<b>7,10</b>	<b>1,29</b>
	T	6,82	0,08	1,00	0,20	0,67	0,10	0,41	0,54	0,10
	LT	34,34	0,38	5,04	1,00	3,40	0,49	2,08	2,71	0,49
	SN	10,11	0,11	1,48	0,29	1,00	0,15	0,61	0,80	0,14
	LSN	69,49	0,77	10,19	2,02	6,87	1,00	4,21	5,48	0,99
	ST	16,52	0,18	2,42	0,48	1,63	0,24	1,00	1,30	0,24
	LST	12,68	0,14	1,86	0,37	1,25	0,18	0,77	1,00	0,18
	W	70,01	0,78	10,27	2,04	6,92	1,01	4,24	5,52	1,00

Tabela C.16: Fator de Bayes baseado na degradação (Deg) e pseudo tempos de falha (PTF), Roda MA11

Obs	Numerador	Denominador								
		N	LN	T	LT	SN	LSN	ST	LST	W
Deg	N	1,00	1,76	0,06	0,03	0,62	5,23	0,04	2,23	0,71
	LN	0,57	1,00	0,04	0,02	0,35	2,98	0,02	1,27	0,41
	T	15,44	27,13	1,00	0,46	9,57	80,76	0,67	34,46	11,04
	LT	<b>33,67</b>	<b>59,13</b>	<b>2,18</b>	<b>1,00</b>	<b>20,86</b>	<b>176,06</b>	<b>1,46</b>	<b>75,11</b>	<b>24,06</b>
	SN	1,61	2,83	0,10	0,05	1,00	8,44	0,07	3,60	1,15
	LSN	0,19	0,34	0,01	0,01	0,12	1,00	0,01	0,43	0,14
	ST	23,10	40,57	1,50	0,69	14,31	120,77	1,00	51,53	16,50
	LST	0,45	0,79	0,03	0,01	0,28	2,34	0,02	1,00	0,32
	W	1,40	2,46	0,09	0,04	0,87	7,32	0,06	3,12	1,00
PTF	N	1,00	0,38	2,01	1,06	0,80	1,27	6,31	2,23	0,14
	LN	2,65	1,00	5,31	2,81	2,13	3,37	16,70	5,89	0,37
	T	0,50	0,19	1,00	0,53	0,40	0,63	3,14	1,11	0,07
	LT	0,94	0,36	1,89	1,00	0,76	1,20	5,95	2,10	0,13
	SN	1,24	0,47	2,50	1,32	1,00	1,58	7,85	2,77	0,17
	LSN	0,79	0,30	1,58	0,83	0,63	1,00	4,96	1,75	0,11
	ST	0,16	0,06	0,32	0,17	0,13	0,20	1,00	0,35	0,02
	LST	0,45	0,17	0,90	0,48	0,36	0,57	2,83	1,00	0,06
	W	<b>7,24</b>	<b>2,74</b>	<b>14,53</b>	<b>7,67</b>	<b>5,82</b>	<b>9,21</b>	<b>45,67</b>	<b>16,12</b>	<b>1,00</b>

Tabela C.17: Teste de Kolmogorov-Smirnov baseado nos pseudotempos de falha  $\hat{t}_i$  para as Rodas MA31, MA12 e MA11

Ajuste		MA31		MA12		MA11	
		Hamada <sub>pseudo</sub>	Prop <sub>pseudo</sub>	Hamada <sub>pseudo</sub>	Prop <sub>pseudo</sub>	Hamada <sub>pseudo</sub>	Prop <sub>pseudo</sub>
N	$D_n$	0,23806	0,23393	0,31387	0,29414	0,16704	0,15479
	p-value	0,34435	0,36395	0,09789	0,14348	0,76226	0,83520
LN	$D_n$	0,23168	0,21693	0,18331	0,17364	0,11547	<b>0,10843</b>
	p-value	0,37420	0,45474	0,66148	0,72087	0,97971	<b>0,98865</b>
T	$D_n$	0,20518	0,20557	<b>0,12388</b>	<b>0,12564</b>	0,10047	0,11029
	p-value	0,52831	0,52222	<b>0,96009</b>	<b>0,95688</b>	0,99542	0,98687
LT	$D_n$	<b>0,18691</b>	0,20843	0,17087	0,17064	<b>0,09610</b>	0,11000
	p-value	<b>0,63791</b>	0,50520	0,74023	0,74491	<b>0,99726</b>	0,98720
SN	$D_n$	0,29559	0,28693	0,29022	0,29214	0,12901	0,14179
	p-value	0,13938	0,16043	0,15291	0,14605	0,94719	0,89712
LSN	$D_n$	0,20634	<b>0,20393</b>	0,20617	0,19814	0,11369	0,12021
	p-value	0,51696	<b>0,53070</b>	0,51639	0,56539	0,98165	0,97112
ST	$D_n$	0,21755	0,21993	0,23281	0,23664	0,11798	0,12971
	p-value	0,45396	0,43854	0,37132	0,35068	0,97469	0,94402
LST	$D_n$	0,20005	0,20893	0,19563	0,19064	0,11304	0,11429
	p-value	0,55518	0,50486	0,58657	0,61326	0,98252	0,98091
W	$D_n$	0,21336	0,21507	0,24565	0,24464	0,13379	0,13879
	p-value	0,47811	0,46534	0,30392	0,31368	0,93076	0,91082

A seguir apresentamos as tabelas referentes à análise dos Dados de Laser na Seção 4.5.2.

Tabela C.18: Qualidade de ajuste baseado na degradação (Deg) e pseudo tempos de falha (PTF), Dados de Laser

Obs	Modelo	LPML	$LPML_{perfil}$	DIC	WAIC
Deg	N	0,0644	1,3829	-34,5006	-31,6743
	LN	0,0641	1,4289	-34,4078	-31,2783
	T	0,0653	<b>1,4426</b>	<b>-34,6361</b>	<b>-31,8135</b>
	LT	0,0643	1,2338	-34,4585	-31,6693
	SN	0,0638	1,3535	-34,1303	-31,0569
	LSN	0,0645	1,3466	-34,5351	-31,3113
	ST	0,0648	1,4252	-34,5091	-31,3827
	LST	0,0642	1,3541	-34,1834	-31,1532
	W	<b>0,0655</b>	1,3603	-34,4792	-31,7725
PTF	N	-8,4225	-	501,8075	259,0910
	LN	-8,4796	-	505,1447	259,3002
	T	-8,4910	-	505,1057	264,7177
	LT	-8,5480	-	507,9920	265,2726
	SN	-11,4198	-	<b>497,4119</b>	264,1025
	LSN	-8,3753	-	499,2792	<b>254,9972</b>
	ST	-8,9038	-	499,6285	261,8798
	LST	-8,4207	-	501,1123	258,1654
	W	<b>-8,3716</b>	-	498,6327	261,0349

Tabela C.19: Fator de Bayes baseado na degradação (Deg) e pseudo tempos de falha (PTF), Dados de Laser

Obs	Numerador	Denominador								
		N	LN	T	LT	SN	LSN	ST	LST	W
Deg	N	1,00	6,79	1,57	1,31	0,89	0,74	0,29	1,34	2,13
	LN	0,15	1,00	0,23	0,19	0,13	0,11	0,04	0,20	0,31
	T	0,64	4,32	1,00	0,84	0,56	0,47	0,19	0,85	1,36
	LT	0,76	5,17	1,20	1,00	0,67	0,56	0,22	1,02	1,62
	SN	1,13	7,66	1,77	1,48	1,00	0,84	0,33	1,51	2,41
	LSN	1,35	9,16	2,12	1,77	1,20	1,00	0,40	1,81	2,88
	ST	<b>3,41</b>	<b>23,13</b>	<b>5,35</b>	<b>4,48</b>	<b>3,02</b>	<b>2,53</b>	<b>1,00</b>	<b>4,57</b>	<b>7,27</b>
	LST	0,75	5,06	1,17	0,98	0,66	0,55	0,22	1,00	1,59
	W	0,47	3,18	0,74	0,62	0,42	0,35	0,14	0,63	1,00
PTF	N	1,00	2,51	3,53	4,92	5,64E+17	1,57	26,72	3,77	0,38
	LN	0,40	1,00	1,41	1,96	2,25E+17	0,63	10,65	1,50	0,15
	T	0,28	0,71	1,00	1,39	1,6E+17	0,45	7,58	1,07	0,11
	LT	0,20	0,51	0,72	1,00	1,15E+17	0,32	5,43	0,77	0,08
	SN	1,77E-18	4,45E-18	6,25E-18	8,72E-18	1,00	2,79E-18	4,74E-17	6,69E-18	6,73E-19
	LSN	0,64	1,59	2,24	3,12	3,58E+17	1,00	16,97	2,40	0,24
	ST	0,04	0,09	0,13	0,18	2,11E+16	0,06	1,00	0,14	0,01
	LST	0,26	0,66	0,93	1,30	1,49E+17	0,42	7,08	1,00	0,10
	W	<b>2,63</b>	<b>6,61</b>	<b>9,28</b>	<b>12,94</b>	<b>1,49E+18</b>	<b>4,14</b>	<b>70,34</b>	<b>9,94</b>	<b>1,00</b>

Tabela C.20: Teste de Kolmogorov-Smirnov baseado nos pseudo tempos de falha  $\hat{t}_i$  para os Dados de Laser

Modelo		$Hamada_{pseudo}$	$Prop_{pseudo}$	Modelo		$Hamada_{pseudo}$	$Prop_{pseudo}$
N	$D_n$	0,14171	0,13367	LN	$D_n$	0,17838	0,18917
	p-value	0,87960	0,91511		p-value	0,65481	0,58654
T	$D_n$	0,13120	0,13850	LT	$D_n$	0,16093	0,16250
	p-value	0,92462	0,89286		p-value	0,77140	0,75882
SN	$D_n$	0,09579	<b>0,10317</b>	LSN	$D_n$	0,15681	0,16517
	p-value	0,99627	<b>0,99119</b>		p-value	0,79635	0,73938
ST	$D_n$	<b>0,09521</b>	0,10617	LST	$D_n$	0,12433	0,12317
	p-value	<b>0,99690</b>	0,98807		p-value	0,94851	0,95253
W	$D_n$	0,11444	0,11550				
	p-value	0,97476	0,97156				