UM MODELO DE DEGRADAÇÃO BAYESIANO COM APLICAÇÕES PARA MANUTENÇÃO E DETERMINAÇÃO DE GARANTIA

Rívert Paulo Braga Oliveira
DEP-UFMG
Av. Presidente Antônio Carlos, 6627. Belo Horizonte - MG,Brasil.
lamabarista@gmail.com

Marta Afonso Freitas
DEP-UFMG
Av. Presidente Antônio Carlos, 6627. Belo Horizonte - MG,Brasil.
marta@dep.ufmg.br

Rosângela Helena Loschi
EST-UFMG
Av. Presidente Antônio Carlos, 6627. Belo Horizonte - MG,Brasil.
loschi@est.ufmg.br

Enrico Antônio Colosimo
EST-UFMG
Av. Presidente Antônio Carlos, 6627. Belo Horizonte - MG,Brasil.
enricoc@est.ufmg.br

RESUMO

Experimentos de degradação costumam ser utilizados com o objetivo de extrair a distribuição do tempo de falha de produtos que possuem elevada confiabilidade, os quais dificilmente falham nos testes de vida tradicionais ou testes de vida acelerados. Nestes casos, se existem características do produto as quais a degradação no tempo pode ser relacionada à confiabilidade, então a coleta de "dados de degradação" fornecerá informação sobre a confiabilidade do produto. No presente artigo, dados de degradação são analisados sob enfoque Bayesiano. A metodologia é aplicada a uma base de dados real advinda da literatura corrente sobre o assunto. A contribuição do artigo reside no fato da utilização da distribuição preditiva à *posteriori* dos tempos de falha para indicar adequados planos de manutenção e garantia do produto. Esta é uma informação importante tanto para fornecedores quanto para o usuário final.

PALAVRAS CHAVE. Modelos de Degradação. Estatística Bayesiana. Manutenção e Garantia. Estatística.

ABSTRACT

Degradation experiments are usually used to assess the lifetime distribution of highly reliable products which are not likely to fail under the traditional life tests or accelerated life tests. In such cases, if there exist product characteristics whose degradation over time can be related to reliability, then collecting "degradation data" can provide information about product reliability. Over this paper degradation data is analyzed through Bayesian approach. The paper contribution lies in using the time to failure predictive posterior distribution to indicate proper maintenance plans and product guarantee. This is an important information for the suppliers and final user.

KEYWORDS. Degradation Path Models. Bayesian Statistics. Maintenance and Guarantee. Statistics.

1. Introdução

Uma parte substancial da literatura em Confiabilidade tem como foco o uso de dados de tempo de falha (ou tempo de vida) em estudos nos quais o objetivo é a caracterização da confiabilidade de produtos. Esses dados são oriundos de ensaios de vida. Entretanto, para produtos que já têm um alto grau de confiabilidade, tais como componentes eletrônicos modernos, é comum encontrar-se diante de uma situação na qual, ao final do ensaio, poucas ou até mesmo nenhuma falha é observada, resultando em um 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 (Nelson, 1990) mas, mesmo nestes casos, o ensaio pode terminar com um alto índice de censuras. Dados com esta característica (alto índice de censuras) fornecem pouca informação a respeito da proporção de produtos que conseguem 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 (Chiao e Hamada, 1996, 2001; Lu e Meeker, 1993; Lu, Meeker e Escobar, 1996; Tseng ,Hamada e Chiao, 1995; Robinson e Crowder, 2000; Hamada, 2005) . Nos testes de degradação, a resposta de interesse 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 disso, a própria definição da falha deve estar associada à algum nível crítico pré-especificado (limiar de falha ou "failure threshold") da medida de degradação. 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. 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 "falha" durante o período de duração do ensaio ou observação.

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. Essa análise é, em geral, implementada em duas etapas. A primeira consiste na construção de um modelo não linear (ou linear) misto que explique o perfil de degradação ao longo do tempo e a estimação dos parâmetros desconhecidos deste modelo. Trata-se, portanto, da análise de dados longitudinais com resposta contínua, através de 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 ser vista como uma modelagem de dois estágios, no qual primeiramente 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 então considerados como sendo amostras (realizações) de distribuições Normais independentes, cujas médias são funções lineares (ou não lineares) destes efeitos aleatórios e possivelmente de outros parâmetros populacionais (efeitos fixos). Uma vez que uma família paramétrica é escolhida para os efeitos aleatórios é preciso estimar os parâmetros fixos do modelo, quais sejam os parâmetros da 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 e outras características de interesse. Quando o modelo de degradação não é suficientemente simples para que se obtenha uma forma fechada para a distribuição do tempo de falha F_T (t) então uma estimativa é obtida através de simulação de Monte Carlo.

Métodos de estimação baseados em inferência clássica podem ser empregados para obter estimativas dos parâmetros desconhecidos nos modelos de degradação e, a partir delas, estimar a distribuição do tempo de falha. Lu e Meeker (1993) sugeriram um modelo geral para dados de degradação bem como procedimentos de inferência para estimação dos parâmetros. O procedimento foi ilustrado com dados de evolução de trincas em ensaios de fadiga. Yacout *et al.* (1996) utilizaram estes métodos para modelagem da degradação de elementos de combustível. Lu *et al.* (1997) estenderam o modelo de degradação e sugeriram métodos de estimação baseados na verossimilhança; entretanto, estes não são adequados para todos os tipos de dados de degradação. Dentre alguns métodos disponíveis na literatura estão o método analítico, numérico e o método aproximado (Meeker e Escobar, 1998; capítulos 13 e 21) e todos baseiam-se em estimação por máxima verossimilhança. Os métodos analítico e aproximado aplicam-se melhor a situações nas quais o modelo de degradação é simples (por exemplo linear) 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.

As rotinas disponíveis para a estimação dos parâmetros por máxima verossimilhança 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). Entretanto, essa suposição nem sempre é verdadeira em situações práticas e dada a dificuldade de verificação de sua validade, muitos trabalhos na literatura têm tentado acessar 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, Caffo e Ohman-Strickland (2004), Litière, Alonso e Molenberghs (2008), 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 e da confiabilidade em um tempo t, Freitas et al.(2009) compararam os métodos numérico e aproximado através de um estudo por simulação. Os autores utilizaram um modelo de degradação linear simples onde o efeito aleatório associado à cada perfil de degradação era oriundo das distribuições Normal, Weibull e Lognormal. Nos três casos a análise foi implementada sob a suposição de Normalidade dos efeitos aleatórios. Os autores constataram que os dois métodos são afetados por desvios na suposição de Normalidade, sendo que o impacto no método numérico é ainda maior.

Em função da grande dificuldade de estimação dos parâmetros do modelo em situações nas quais a suposição de Normalidade não se aplica e dos efeitos desta má especificação, uma outra frente de pesquisa foi aberta cujo foco é o desenvolvimento de métodos computacionais alternativos, que permitam a obtenção de estimadores de máxima verossimilhança para modelos de efeitos mistos (utilizados na modelagem de dados de degradação) com efeitos aleatórios não-normais. Pinheiro *et al*. (2001) propuseram um modelo linear misto com efeitos aleatórios de uma distribuição T multivariada. Para modelos gerais com efeitos aleatórios não-normais, alguns métodos de estimação que aparecem na literatura incluem a verossimilhança hierárquica (Lee e Nelder, 1996) e a maximização por partes (Song, Fang e Kalbfleish, 1995). Entretanto, a implementação destes métodos para utilização intensiva em análise de dados mostrou-se uma tarefa desafiadora. A quadratura Gaussiana é uma ferramenta que tem sido utilizada com sucesso na estimação de parâmetros em modelos com efeitos aleatórios Normais. Para modelos mistos com efeitos aleatórios não-normais, Nelson et al.(2006) propôs um método computacional simples que utiliza o método PIT ("Probability Integral Transformation") para obter estimativas de máxima verossimilhança. Entretanto essa abordagem só consegue lidar com efeitos aleatórios para os quais a inversa da função distribuição acumulada (f.d.a) tem forma fechada, o que limita a sua aplicação. Liu e Yu (2007) propuseram um novo método de estimação que reformula a verossimilhança condicional de efeitos aleatórios não-normais. Os autores reformulam a verossimilhança, dividindo e multiplicando a verossimilhança condicional por uma função

densidade 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 lidar com distribuições de efeitos aleatórios para as quais não há forma fechada para a f.d.a.

Portanto a abordagem do problema através da inferência clássica ainda esbarra em problemas relacionados aos métodos que possam acomodar a suposição de não-normalidade dos efeitos aleatórios.

Por outro lado, uma outra abordagem para o tratamento de dados de degradação é utilização de métodos Bayesianos . Esta abordagem aparece de maneira natural visto que utiliza a informação contida na distribuição a *priori* dos parâmetros do modelo de degradação. Em fármaco-cinética, métodos Bayesianos já foram utilizados (Wakefield *et al.,*1994; Gelman *et al.,*1996; Wakefield, 1996), entretanto só recentemente essa abordagem começou a ser utilizada em dados de degradação. Hamada (2005) utilizou métodos Bayesianos em dados de degradação associados a falhas em lasers. O autor baseou-se em um modelo linear de degradação com estrutura hierárquica, no qual uma distribuição à *priori* Weibull foi escolhida para os efeitos aleatórios e distribuições à *priori* Gamma foram utilizadas para os dois hiperparâmetros. Amostras da distribuição à *posteriori* dos parâmetros do modelo foram obtidas através de simulação MCMC (Markov Chain Monte Carlo; Dani e Lopes, 2006). As distribuições à *posteriori* de quantis da distribuição do tempo de falha T, bem como da confiabilidade à *posteriori* em 4000 h de uso também foram obtidas. Para isso, as expressões para o cálculo do p-ésimo quantil e de R(t) (a confiabilidade em um tempo t) de uma distribuição Weibull foram aplicadas às amostras da distribuição à *posteriori* dos parâmetros do modelo.

O objetivo deste artigo é mostrar que a análise conduzida com a abordagem Bayesiana pode fornecer subsídios importantes para a manutenção de equipamentos. Isto será evidenciado através da re-análise dos dados contidos no trabalho de Hamada (2005). Para atingir este objetivo, a uma abordagem um pouco diferente daquela utilizada pelo autor será empregada aqui.

O presente texto organiza-se da seguinte forma. Na Seção 2 são apresentados os dados de degradação do laser, descrito em Hamada (2005). Na Seção 3, expõe-se a forma geral do modelo de degradação. Na Seção 4 é fornecida a especificação do modelo sob o enfoque Bayesiano, seguida da análise dos dados dos emissores de laser. Na Seção 5 são apresentadas as conclusões.

2. Dados de degradação do Laser

Hamada (2005) apresentou dados de degradação de lasers. Estes dados foram originalmente analisados por Meeker e Escobar (1998), utilizando a modelagem baseada em inferência clássica descrita na introdução (método numérico).

Quinze (15) lasers foram acompanhados por 4000 h de operação. Registrou-se para cada um, a cada 250 h, 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. Conseqüentemente, para que seja gerada uma saída de luz constante, a corrente de operação precisa ser aumentada ao longo do tempo. Portanto a medida de degradação neste caso é o percentual de aumento na corrente de operação. A "falha" do laser é definida como tendo ocorrido quando o percentual de aumento exceder 10%. Em outras palavras, no limiar $D_f = 10\%$ (denominado limiar de falha ou "failure threshold") considera-se que o laser falhou, embora ainda esteja em operação.

A Figura 1 apresenta o gráfico dos perfis de degradação para cada uma das 15 unidades emissoras de laser (os dados estão disponíveis na página http://www.stat.lanl.gob/staff/MHamada/degradation_data.htm). Por exemplo, para a 10ª. Unidade, cuja degradação é a mais rápida, os dados são (0,0000; 0,4136; 1,4880;2,3810; 2,9950;

3,8350; 4,5010; 5,2510; 6,2560; 7,0510; 7,8030; 8,3210; 8,9300; 9,5540; 10,4500; 11.2800; 12,2100).

O questionamento que se faz é a respeito de qual o melhor momento para efetuar a manutenção de cada unidade laser. Ainda, atentando ao ponto de vista de um fabricante, qual tempo de garantia um fabricante deveria estipular para seus emissores de laser?

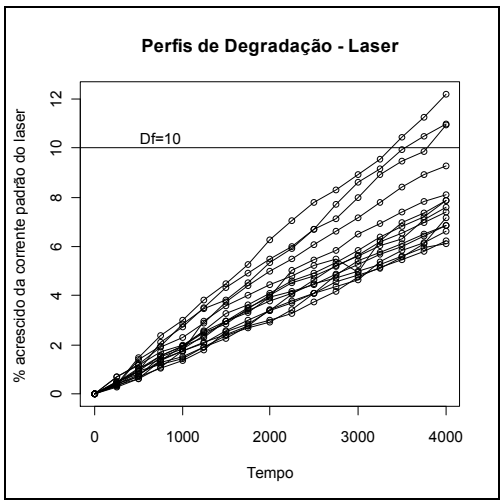

Figura 1 – Perfis de degradação dos emissores de laser

3. O Modelo geral de degradação.

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 i-ésima unidade, a verdadeira degradação no tempo t é dada por:

$$y_{it} = D_i(t) = D(t;\boldsymbol{\alpha};\boldsymbol{\beta}_i) \qquad (1)$$

em que $D(t_{ij};\boldsymbol{\alpha};\boldsymbol{\beta}_i)$ assume uma forma funcional que depende de parâmetros (cuja natureza será explicitada mais adiante). Para uma dada unidade i, a falha é considerada como tendo ocorrido quando a degradação atinge o limiar D_f. Assim o tempo de falha da i-ésima unidade é o tempo T_i no qual:

$$y_{iT_i} = D_f = D_i(T_i) = D(T_i;\boldsymbol{\alpha};\boldsymbol{\beta}_i) \qquad (2)$$

Entretanto, os dados observados nos ensaios de degradação para uma dada unidade i são na verdade amostras do seu verdadeiro perfil (dado por (1)), obtidas em m_i tempos pré-especificados t_{ij} (j = 1,2,...,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 ε_{ij}.

Portanto o modelo geral de degradação é dado por:

$$y_{ij} = D_{ij} = D(t_{ij};\boldsymbol{\alpha};\boldsymbol{\beta}_i) + \varepsilon_{ij}, \qquad (3)$$

em que $D(t_{ij};\boldsymbol{\alpha};\boldsymbol{\beta}_i)$ é o real perfil de degradação da unidade i em um tempo pré-especificado t_{ij} (j = (i = 1,...,n; j = 1,...,m_i); $\boldsymbol{\alpha} = (\alpha_1,\alpha_2,...,\alpha_p)^t$ é um vetor de efeitos fixos que descreve as características da população (este se mantém constante para todas as unidades); $\boldsymbol{\beta}_i = (\beta_{i1}, \beta_{i2}, ..., \beta_{i3})^t$ é um vetor de efeitos aleatórios associado à i-ésima unidade, representando características individuais de cada unidade (por exemplo, variações em propriedades da material prima, no

processo de produção, nas dimensões do component, etc.) e ε_{ij} é o erro aleatório (de medida) associado à i-ésima unidade no tempo t_{ij}.

A forma funcional de $D(t_{ij};\boldsymbol{\alpha};\boldsymbol{\beta}_i)$ pode ser baseada em análise empírica dos perfis dos de degradação do processo sob estudo, mas sempre que possível deve ser baseada em algum fenômeno físico-químico associado à ele. Assume-se que os ε_{ij}'s (i = 1,...,n; j = 1,...,m_i) são independents e identicamente distribuídos *(iid)*, segundo uma Normal com media zero e variância desconhecida σ^2_ε.

Os $\boldsymbol{\beta}_i$'s (i = 1,2,...,n) são independentes, com distribuição $\Lambda_\beta(\boldsymbol{\theta})$, que depende de um vetor de parâmetros desconhecidos $\boldsymbol{\theta} = (\theta_1,...,\theta_q)^t$, o qual precisa ser estimado a partir dos dados de degradação. Além disso, $\{\varepsilon_{ij}\}$ e $\{\boldsymbol{\beta}_i\}$ são, por suposição do modelo, independentes. 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 até o tempo t. Portanto, é possível definir a distribuição do tempo de falha T para o modelo (1) como:

$$F_T(t) = F_T(t;\boldsymbol{\alpha};\Lambda_{\boldsymbol{\beta}}(.);D_f;D) = P(T \le t) = P[D(t;\boldsymbol{\alpha};\boldsymbol{\beta}) \ge D_f]$$

quando as medidas de degradação são crescentes com o tempo ou

$$F_T(t) = F_T(t;\boldsymbol{\alpha};\Lambda_{\boldsymbol{\beta}}(.);;D) = P(T \le t) = P[D(t;\boldsymbol{\alpha};\boldsymbol{\beta}) \le D_f]$$

quando as medidas são decrescentes com o tempo.

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 $\boldsymbol{\alpha}$ (o vetor de efeitos fixos) e $\boldsymbol{\theta} = (\theta_1,...,\theta_q)^t$ o vetor de parâmetros (também fixos) da distribuição do efeitos aleatórios $\Lambda_\beta(\boldsymbol{\theta})$.

Conforme já mencionado anteriormente, para modelos com formas funcionais simples, a função distribuição $F_T(t)$ pode ser expressa em forma fechada. Entretanto, para muitos perfis, isto não é possível. Quando a forma functional de $D(t_{ij};\boldsymbol{\alpha};\boldsymbol{\beta})$ é não linear e o modelo tem mais de um parâmetro aleatório (ou seja, quando o vetor de parâmetros $\boldsymbol{\beta}$ tem dimensão $k > 1$), a especificação de $F_T(t)$ 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 $\boldsymbol{\alpha}$, $\Lambda_\beta(\boldsymbol{\theta})$, D_f e D especificados, utilizando simulação de Monte Carlo. Todavia, esse procedimento só pode ser implementado se os parâmetros fixos $\boldsymbol{\alpha}$ e o vetor de parâmetros $\boldsymbol{\theta}$ da distribuição dos efeitos aleatórios $\Lambda_\beta(\boldsymbol{\theta})$ puderem ser estimados de alguma forma.

4. Os dados do Laser revisitados

4.1 Especificação do modelo sob a abordagem Bayesiana.

Como as curvas da Figura 1 são aproximadamente lineares, Hamada (2005) utilizou um modelo no qual o verdadeiro perfil de degradação no tempo t, para a i-ésima unidade do laser é dado por:

$$D_i(t) = \left(\frac{1}{\beta_i}\right)t \qquad (4)$$

Utilizando o mesmo argumento da seção anterior, temos que o tempo de falha T_i da i-ésima unidade é o tempo no qual $D_f = D_i(T_i)$, ou seja:

$$T_i = D_f \beta_i . \qquad (5)$$

O modelo utilizado descreve a degradação observada y_{ij} isto é, a degradação observada do *i*-ésimo laser no tempo t_{ij} toma a seguinte forma:

$$y_{ij}=D_{ij}=D(t_{ij};\beta_i)+\varepsilon_{ij}=\frac{1}{\beta_i}t_{ij}+\varepsilon_{ij} \quad i=1,\ldots,15 \text{ (unidades)}; \; j=1,\ldots,17 \text{ (medidas)} \qquad \textbf{(6)}$$

Análises anteriores indicaram que os tempos de falha dos lasers poderiam ser bem explicados pela distribuição de Weibull (Meeker e Escobar ,1998), e devido à relação (5), para que os tempos de falha T_i possam ter distribuição Weibull, é preciso que os β_i´s também tenham distribuição Weibull.

De fato, pode-se demonstrar que $\beta_i \sim Weibull(\delta,\lambda)$ com função de densidade dada por:

$$p(\beta_i \mid \delta,\lambda)=\delta\lambda(\beta_i)^{\delta-1}\exp\{-\lambda(\beta_i)^{\delta}\} \; (\delta>0;\lambda>0) \text{ então, } T_i \mid \delta,\lambda \sim Weibull\left(\delta\,,\,\lambda/D_f^{\delta}\right)$$

Portanto a especificação completa do modelo (6) é a seguinte:

- $\varepsilon_{ij} \mid \sigma_\varepsilon^2 \sim N(0;\sigma_\varepsilon^2)$ (iid), ou seja, $p(\varepsilon_{ij} \mid \sigma_\varepsilon^2) = N(0;\sigma_\varepsilon^2)$
- $y_{ij} \mid \beta_i \sim N(t_{ij}/\beta_i\,;\sigma_\varepsilon^2)$
- Distribuições à *priori* para os parâmetros (β_i´s e σ^2_ε) e hiperparâmetros dos β_i´s:
  - $\beta_i \mid \delta,\lambda \overset{iid}{\sim} Weibull(\delta,\lambda)$
  - $\lambda \sim$ Gamma (0,01; 0,01)
  - $\delta \sim$ Gamma (0,01; 0,01)
  - $\sigma^2_\varepsilon \sim$ Gamma-Inversa (0,01; 0,01)

Assim, de acordo com a abordagem Bayesiana, $\eta=\left(\beta_1,\beta_2,\ldots,\beta_{15},\lambda,\delta,\sigma_\varepsilon^2\right)^t$ é o vetor de parâmetros desconhecidos cujas distribuições à *posteriori* serão obtidas.

4.2. Análise dos dados

Como apresentado no item 2, dispomos de quinze (15) emissores de laser sob teste.

Para conseguirmos elaborar um plano de manutenção para os emissores de laser utilizamos a distribuição preditiva à *posteriori* dos tempos de falha para cada unidade experimental, da qual podemos extrair sua respectiva função distribuição acumulada (f.d.a) e ou densidade de probabilidade (f.d.p.) empírica. Para obter essa distribuição preditiva à *posteriori* basta aplicar a equação 5 em $\eta \mid Dados$ (aqui entende-se Dados pelos perfis de degradação). A idéia é utilizar as medidas de degradação no tempo, correspondentes a um específico emissor de laser, e estipular o melhor momento para efetuar sua manutenção.

Outra questão que tem apelo comum em confiabilidade é estimar qual seria o comportamento (f.d.a. e ou f.d.p.) de um emissor de laser qualquer, que não se encontra no escopo da amostra. Para obtermos tal estimativa, na abordagem da Estatística Bayesiana, precisamos encontrar a distribuição preditiva à *posteriori* dos tempos de falha para um novo emissor de laser. Como no modelo de degradação utilizado a degradação de todas as unidades experimentais é nula, basta inserir a nova unidade no modelo com a medida de degradação nula no tempo zero. Daí, o vetor de parâmetros a ser estimado torna-se $\eta=\left(\beta_1,\beta_2,\ldots,\beta_{15},\beta_{16},\lambda,\delta,\sigma_\varepsilon^2\right)^t$, então a distribuição preditiva à *posteriori* dos tempos de falha

da nova unidade e sua respectiva f.d.a. e ou f.d.p. são obtidas da mesma maneira que para os emissores de laser que possuíam medidas de degradação durante o tempo de estudo. Com a f.d.a. e a f.d.p. à *posteriori* do novo emissor de laser o fabricante desse equipamento poderia estipular um período de garantia para qualquer equipamento que fosse produzir.

A análise dos dados foi efetuada utilizando-se os softwares WinBUGS 1.4 e R 2.8.1 (pacote R2WinBUGS), uma vez que $\eta \mid Dados$ não tem forma analítica fechada e necessita de algorítimos numéricos (Markov Chain Monte Carlo Methods – MCMC presentes no WinBUGS) para obtenção de sua estimativa. O código escrito no WinBugs, utilizado para obtenção das estimativas dos parâmetros, encontra-se disponível na página (http://www.stat.lanl.gov/staff/MHamada/degradation_code.htm).

A fim de garantir a convergência das estimativas dos parâmetros, assim como minimizar a autocorrelação das mesmas, foram realizadas 10.001.000 iterações do algoritmo MCMC, com um período de *burn-in* de 1000, e saltos para de 1000 para a amostragem, o que resultou em 10.000 amostras para cada parâmetro. Todos os parâmetros do modelo convergiram e apresentaram decaimento rápido dos autocorrelogramas.

Verificam-se pelo Gráfico 1 as f.d.p.'s da distribuição preditiva *à posteriori* dos tempos de falha para cada emissor de laser, incluindo a f.d.p. da nova unidade de emissão. As unidades 10, 6 e 1 são as que falharam, por isso apresentam um alto grau de afunilamento da curva no entorno do mais provável tempo de falha, e esse afunilamento não se torna uma linha pois não se conhece o tempo exato da falha dessas unidades. A curva das unidades laser que não falharam apresenta maior dispersão, indicando maior grau de incerteza com relação aos seus tempos de falha (quanto menos inclinado o perfil da Figura 1, maior a incerteza evidenciada no Gráfico 1).

No Gráfico 1 apresentamos também as f.d.a.'s da distribuição preditiva *à posteriori* dos tempos de falha para cada emissor de laser, incluindo a f.d.a. da nova unidade de emissão. É notável que a probabilidade de falha acumulada da nova unidade começa a crescer no limiar da f.d.a. da primeira unidade a falhar, e se estende até um pouco além da f.d.a. da última unidade que falharia. Esse resultado é perfeitamente razoável, uma vez que utilizamos distribuições a *priori* não informativas, no sentido de colocar massas de probabilidade aproximadamente iguais para o espaço paramétrico, para todos os parâmetros.

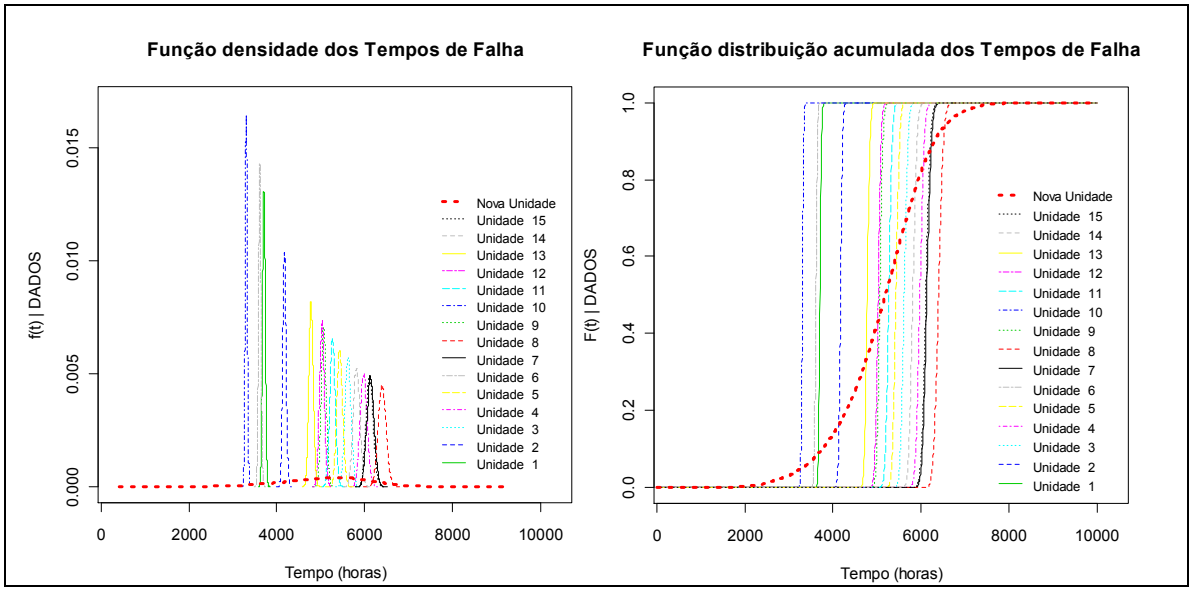

Gráfico 1 – Funções densidade de probabilidade e acumulada da distribuição preditiva *à posteriori* dos tempos de falha para todos os emissores de laser e para um novo emissor

A Tabela 1 apresenta as informações de resumo da distribuição preditiva à *posteriori* dos tempos de falha para os 15 emissores de laser assim como para uma nova unidade de emissão.

Tabela 1 – Estimativas das quantidades de interesse para a distribuição preditiva à posteriori dos tempos de falha

Unidade de Emissão de Laser	média(t|*Dados*) (em horas)	$t_{0,01}$ | *Dados* *(em horas)*	$t_{0,025}$ | *Dados* *(em horas)*	$t_{0,5}$ | *Dados* *(em horas)*	$t_{0,975}$ | *Dados* *(em horas)*
1	3710	3641	3651	3709	3768
2	4176	4089	4103	4176	4251
3	5623	5464	5490	5622	5760
4	5981	5808	5831	5981	6138
5	5435	5289	5312	5434	5565
6	3614	3550	3560	3614	3672
7	6137	5952	5979	6135	6302
8	6402	6199	6233	6401	6580
9	5070	4939	4962	5070	5182
10	3309	3254	3263	3309	3357
11	5271	5130	5151	5271	5394
12	5031	4904	4925	5031	5140
13	4782	4669	4687	4782	4882
14	5818	5653	5676	5818	5965
15	6116	5936	5964	6115	6279
16 (Nova unidade)	5110	2425	2859	5194	6938

A interpretação da Tabela 1 é a mesma para todas as unidades laser, inclusive para a nova (unidade 16). Dessa maneira é possível estabelecermos um plano de manutenção das unidades sob teste e fornecer a garantia para qualquer um emissor de laser.

Por exemplo, se o percentil 1 ("$t_{0,01}$ | *Dados*" - tempo no qual a probabilidade de falha é 0,01) é tomado como referência, verifica-se que os tempos relativos a esta coluna da Tabela 1 são os tempos para os quais a manutenção deveria ser planejada para cada respectivo emissor de laser.

Se o fabricante desejasse estabelecer a garantia para um novo emissor de laser, ele poderia indicar o tempo de 2425 horas sabendo que a probabilidade de falha de qualquer nova unidade antes desse tempo seria de apenas 1%.

O estabelecimento desses tempos também poderia ser associado a uma função de custo, que, se otimizada, forneceria os tempos com menor percentual de falha possível para obter o maior retorno financeiro ao fabricante ou ao usuário dos emissores de laser.

Note que para estabelecimento do plano de manutenção e da garantia não se utiliza intervalos de confiança como na Estatística Clássica, uma vez que a medida de incerteza sobre os tempos de falha à *posteriori* é a própria probabilidade de falha para o tempo estipulado. Diferentemente da Estatística Clássica, na Estatística Bayesiana os intervalos de credibilidade referem-se apenas aos parâmetros do modelo estatístico.

5. Conclusões

Neste artigo uma análise Bayesiana foi realizada para os dados de emissores de laser. Os resultados obtidos foram úteis para planejar o tempo de manutenção dos equipamentos, assim como determinar o tempo de garantia de uma nova unidade de emissão.

Pôde-se observar a utilidade do modelo Bayesiano nesta situação. Através dele é possível fazer a previsão futura para cada unidade assim como para uma nova unidade laser.

Em especial, foi possível predizer o tempo de falha para 15 unidades laser. Neste caso, mesmo para unidades que haviam falhado, o modelo foi capaz predizer este tempo, dado que as medições foram realizadas em momentos fixos de tempo e não é possível obter de forma determinística o verdadeiro tempo de falha.

Foram obtidas também a f.d.a. e f.d.p. para uma unidade típica de laser.

Agradecimentos

Os autores agradecem à Capes e Cnpq pelo financiamento da pesquisa.

Referências

Agresti, A., Caffo, B., Strickland-Ohman, P. (2004). Examples in which misspecification of random effects distribution reduces efficiency, and possible remedies. *Computational Statistics & Data Analysis,* v.47, 639-653.

Chiao, C.H. & Hamada, M. (1996). Using Degradation Data from an Experiment to Achieve Robust Reliability for Light Emitting Diodes. *Quality And Reliability Engineering International* v.12, 89-94.

Chiao, C.H. & Hamada, M. (2000). Experiments with Degradation Data for Improving Reliability and for Achieving Robust Reliability . *Quality And Reliability Engineering International*, v.17, 333-344.

Freitas, M.A, Toledo, M.L.G, Colosimo, E.A., Pires, M.C. (2009). Using Degradation Data do Assess Reliability: A Case Study on Train Wheel Degradation. Quality and Reliability Engineering International, v.25, 607-629..

Gamerman, D., Lopes, H.F.(2006). *Markov Chain Monte Carlo*, 2nd edition. Texts in Statistical Science. Chapman & Hall.

Gelfand, A. E. & Smith, A. F. M. (1990), Sampling-based approaches to calculating marginal densities, *Journal of the American Statistical Association*, v.85, 398–409.

Gelman, A., **Bois, F.Y, Jiang, J**.(1996). Physiological pharmacokinetic analysis using population modeling and informative prior distributions. *Journal of the American Statistical Association,* v.91, 1400-1412.

Hamada, M.(2005). Using Degradation Data to Assess Reliability.*Quality Engineering*, v.17, 615-620.

Lee, E., Nelder, J.A.(1996) Hierarchical generalized linear models (with Discussion). *Journal of the Royal Statistical Society*, *Series B*, v.58, 619-678.

Litière, S. Alonso, S., 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. *Statitics in Medicine*, v.27, 3125-3144.

Liu, L., Yu, Z. (2008). A likelihood reformulation method in non-normal random effects models. *Statistics in Medicine*, v.27, 3105-3124.

Lu, C.J. & Meeker, W.Q. (1993).Using Degradation Measurements to Estimate a Time-to-Failure Distribution. *Technometrics*, v.35: 161-174.

Lu, C.J., Meeker, W.Q. & Escobar, L.A. (1996).Using Degradation Measurements to Estimate a Time-to-Failure Distribution. *Statistica Sinica*, v.6: 531-546.

Meeker, W. & Escobar, L. A. (1998).*Statistical Methods for Reliability Data.*Wiley Series in Probability and Statistics: New York.

Nelson, W. (1990)**.** *Accelerated Testing: Statistical Models, Test Plans and Data Analyses*. John Wiley: New York.

Pinheiro, J. C., Bates D. C.(1995). Approximations to the log-likelihood function in The nonlinear mixed-effects model. *Journal of Computational and Graphical Statistics*, v.1,no.4, 12-35.

Pinheiro, J.C., Liu C., Wu, Y.N. (2001). Efficient algorithms for robust estimation in linear mixed0effects models using the multivariate t distribution. *Journal of Computational and Graphical Statistics*, v.10, 249-276.

Plummer, M., Best, N., Cowles, K. & Vines, K. (2005). Output analysis and diagnostics for MCMC. The Coda Package. Software R. Avaliable at http://www.fis.iarc.fr/coda/

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 Anaysis, v.6**,** 357-374.

Song, PX, Fan, Y., Kalbfleisch, J.D. (2005). Maximization by parts in likelihood inference (with Discussion). *Journal of the American Statistical Association*, v.100, 1145-1167.

Spiegelhalter, D., Thomas, A., Best, N. (2000). WinBUGS Version 1.3 User Manual, avaliable at http://www.mrc-bsu.cam.ac.uk/bugs/.

Tseng, S.T., Hamada, M. & Chiao, C.H. (1995). Using degradation data from a factorial experiment to improve fluorescent lamp reliability. *Journal of Quality Technology*, v.27: 363-369.

Verbeke, G., Lesaffre, E. (1997). The effect of misspecifying the random-effects distgribution in linear mixed effects distribution in linear mixed models for longitudinal data. *Computational Statistics & Data Analysis*, v.23, 541-556.

Wakefield, J.C. (1996). The Bayesian analysis of population pharmacokinetic models. *Journal of the American Statistical Association*, v.91, 62-75.

Wakefield, J.C., Smith, A.F.M., Racine-Poon, A., Gelfand, A.E. (1994). Bayesian analysis of linear and non-linear population models by using the Gibbs sampler. *Applied Statistics*, v.43, 201-221.

Yacout, A. M., Salvatores, S., Orechwa, Y. (1996).Degradation Analysis Estimates of the Time-to-Failure Distribution of Irradiated Fuel Elements. *Nuclear Technology*, v.113, 177-189.