Modelos Estatísticos em Processos de Poisson Não Homogêneos: Aplicação em Sistemas Reparáveis

Maristela Dias de Oliveira

Belo Horizonte - MG
Setembro / 2010

Maristela Dias de Oliveira

Modelos Estatísticos em Processos de Poisson Não Homogêneos: Aplicação em Sistemas Reparáveis

Tese de doutorado apresentada ao Departamento
de Estatística do Instituto de Ciências Exatas
da Universidade Federal de Minas Gerais,
como requisito parcial à obtenção do título de Doutor em Estatística.

Orientador: Enrico A. Colosimo

Co-orientador: Gustavo L. Gilardoni

Doutorado em Estatística
Departamento Estatística
Instituto de Ciências Exatas
Universidade Federal Minas Gerais

Belo Horizonte - MG
Setembro / 2010

A minha princesa

Inaê

Agradecimentos

A Reinaldo, companheiro de todas as horas. Até mesmo nas ausentes. Agradeço principalmente pela compreensão, sem a qual as coisas teriam sido bem mais difíceis...

A toda a minha família. Cada um teve uma participação fundamental na retaguarda dessa trajetória.

A Inaê, que mesmo sem entender o que estava acontecendo, acompanhou literalmente cada peça colocada na montagem desse quebra-cabeças.

Ao Enrico, por acreditar na possibilidade e por apostar nesses mais quatro anos de jornada juntos. Mais do que orientador, você se mostrou um grande amigo. Depois de tanto tempo, sinto-me à vontade para chamar-lhe não só de parceiro, mas de cúmplice.

Ao Gustavo, que topou se juntar mais recentemente, e só veio agregar valores. Muitos valores. Sua contribuição e orientação foram fundamentais.

Aos professores do Departamento de Estatística da UFMG, em especial a Denise, Rosângela e Adrian, bem como aos seus funcionários, em especial a Rogéria.

Aos colegas professores do Departamento de Estatística da UFBa, por compreender a importância que esse investimento representa.

Agora, alguns agradecimentos mais que especiais:

"Há pessoas que nos falam e nem as escutamos, há pessoas que nos ferem e nem cicatrizes deixam, mas há pessoas que simplesmente aparecem em nossas vidas e nos marcam para sempre."
Cecília Meireles

Ao amigo Fábio. Não preciso justificar esse agradecimento. Apenas fico feliz em constatar que nossas divergências (e essas não foram poucas) nunca superaram nossa amizade. O melhor: com uma solução sempre muito 'Original'.

À amiga Marcinha. Seu apoio em muitas questões nada acadêmicas, certamente tornou minha estadia em BH bem mais agradável.

Algumas pessoas, não menos importantes, fazem parte de uma imensa lista de participações especiais. Citarei apenas alguns: Elias, Michelle, Magda, Marcia, Rodrigo, Ricardo,...

Este trabalho contou ainda com o apoio financeiro da CAPES, através de Universidade Federal da Bahia.

"Ser um empreendedor é executar os sonhos, mesmo que haja riscos. É enfrentar os problemas, mesmo não tendo forças. É caminhar por lugares desconhecidos, mesmo sem bússola. É tomar atitudes que ninguém tomou. É ter consciência de que quem vence sem obstáculos triunfa sem glória. É não esperar uma herança, mas construir uma história..."

Augusto Cury

1 Introdução

Modelos estatísticos para eventos recorrentes têm sido amplamente estudados na literatura de confiabilidade e manutenção de sistemas reparáveis (Ascher and Feingold (1984), Bain and Engelhardt (1991), Rigdon and Basu (2000)).

Um sistema reparável é um sistema, podendo ser um sistema mecânico, ao qual é permitido experimentar mais de uma falha ao longo de sua vida útil. A principal característica de um sistema reparável é que, quando ocorre uma falha, o sistema pode ser restaurado e posto de volta em operação, sem necessariamente ser substituído. O reparo a que é submetido o sistema que experimentou uma falha representa uma atividade de manutenção. Essa manutenção pode ser de dois tipos. O primeiro tipo é a Manutenção Perfeita, que retorna o sistema à condição de um sistema novo, popularmente conhecida como "tão bom quanto o novo". Um outro tipo de manutenção é o que se chama de Reparo Mínimo, que retorna o sistema à condição em que o mesmo se encontrava no momento imediatamente anterior à falha. Esse tipo de manutenção pode ser referida como "tão ruim quanto o velho". Estocasticamente, a manutenção perfeita é um processo de renovação. O interesse na combinação manutenção perfeita e reparo mínimo pode ser encontrado desde o trabalho de Barlow and Hunter (1960). Essa situação pode ser encontrada ainda em Gerstack (1977), Block et al. (1990) and Park et al. (2000).

Além das atividades de reparo, é indicado que sejam realizadas ações de Manutenção Preventiva (MP) com o objetivo de minimizar a ocorrência de falhas no sistema. Essa é uma atividade planejada com o objetivo de melhorar a confiabilidade de um sistema reparável e é necessariamente uma manutenção perfeita.

Um programa de MP é geralmente realizado em intervalos periódicos de tempo para cada sistema ou conjunto de sistemas. Em geral, a MP envolve tarefas como inspeção, limpeza, lubrificação, ajuste, alinhamento e/ou a recolocação de sub-componentes. O ideal é definir uma política de MP de maneira que o custo total envolvendo falhas do sistema, manutenção e a recolocação durante seu ciclo de vida esperado seja minimizado. Frequentemente o modelo adotado sob reparo mínimo é um Processo de Poisson Não-Homogêneo (PPNH), $\{N(t) : t \geq 0\}$, em que $N(t)$ é o número de falhas a partir do início da observação até o tempo t (Barlow and Hunter, 1960). Em geral, $N(t)$ é modelada por uma distribuição de Poisson com média $\Lambda(t) = E(N(t)) = \int_0^t \lambda(u)\, du$. A função intensidade de falhas, $\lambda(t)$, é definida como:

$$\lambda(t) = \lim_{\Delta t \to 0} \frac{P(N(t, t + \Delta t] = 1)}{\Delta t}.$$

1.1 Esquemas de truncamento

Existem basicamente duas formas de se observar dados de sistemas reparáveis. Numa delas a coleta desses dados é interrompida depois que um número específico de falhas é observado. Na outra, essa interrupção ocorre em um tempo fixo predeterminado, T. Esses esquemas

de amostragem são conhecidos como *Truncamento por Falha* ou *Truncamento por Tempo*, respectivamente. Do ponto de vista estatístico, a diferença de se considerar um ou outro esquema de amostragem, está no procedimento de estimação. Por exemplo, suponha que um sistema é observado até um tempo T, e que, ao final desse tempo n falhas tenham sido observadas. No esquema de truncamento por tempo, o número de falhas obtidas, n, é uma variável aleatória. Já no truncamento por falha, o tempo final de acompanhamento, $T = t_{(n)}$, é que é aleatório e corresponde ao tempo da última falha observada. Essa diferença, ainda que sutil, afeta a forma de se construir a função de verossimilhança, já que os esquemas não podem ser tratados da mesma maneira.

1.2 Inferência

No caso de um único sistema, suponha que até o tempo de observação y, seja esse um tempo de falha ou de truncamento, tenham ocorrido n falhas, respectivamente nos tempos $0 \leq t_1 < t_2 < \cdots < t_n \leq y$. O procedimento de estimação da função intensidade é baseado na função de verossimilhança, que assume a seguinte forma (Rigdon and Basu, 2000):

$$L(\underset{\sim}{t}, \lambda) = \left(\prod_{i=1}^{n} \lambda(t_i) \right) \exp \left(- \int_0^y \lambda(x) dx \right), \tag{1}$$

em que

$$y = \begin{cases} T, & \text{se trucamento por tempo;} \\ t_n, & \text{se truncamento por falha.} \end{cases}$$

No contexto de múltiplos sistemas reparáveis, a metodologia mais comumente utilizada é a de considerar que tais sistemas são independentes e identicamente distribuídos. Ou seja, os sistemas são tratados como sendo K realizações independentes de um mesmo processo, com função intensidade λ.

Se os processos $N_1(t), N_2(t), \ldots, N_K(t)$ são todos abservados até um mesmo tempo T, o processo de Poisson não homogêneo resultante da superposição é dado por $N_+(t) = \sum_{i=1}^{K} N_i(t)$ e tem função intensidade dada por $\lambda_+(t) = K\lambda(t)$.

Inferências em modelos propostos para essa situação ainda podem ser realizadas através da verossimilhança dada em (1).

Se, por outro lado, as K realizações independentes do mesmo processo são observadas respectivamente até os tempos $T_1, T_2, \ldots, T_K$, seja t_{ij} o tempo do $j-$ésimo evento para a $i-$ésima realização, $i = 1, \cdots, K$; $j = 1, \cdots, n_i$. Dessa forma a função de verossimilhança é dada por (2), (Rigdon and Basu, 2000):

$$L(\lambda) = \left(\prod_{i=1}^{K} \prod_{j=1}^{n_i} \lambda(t_{ij}) \right) \exp \left(- \sum_{i=1}^{K} \int_0^{y_i} \lambda(x) dx \right), \tag{2}$$

em que

$$y_i = \begin{cases} T_i, & \text{se trucamento por tempo;} \\ t_{n_i}, & \text{se truncamento por falha.} \end{cases}$$

Nas duas situações acima assume-se que os sistemas são réplicas de um mesmo sistema. Ou seja, considera-se que os K sistemas são idênticos e, portanto, possuem a mesma função intensidade e assim teríamos uma amostra aleatória de sistemas.

Por outro lado, essa suposição pode não ser verdadeira. Ou seja, pode haver alguma heterogeineidade entre os sistemas, sem que nenhuma variável capaz de avaliá-la tenha sido efetivamente medida. Nesse caso, há que se propor um modelo estatístico capaz de captar essa heterogeneidade e finalmente distinguir, se for o caso, as respectivas funções intensidade da amostra de sistemas.

1.3 Escopo do Trabalho

Este trabalho estuda o histórico de falhas e reparos de um conjunto de sistemas que foram acompanhados por tempos parcialmente superpostos. O trabalho é formado por três textos. Nos dois primeiros os sistemas são considerados como sendo réplicas de um mesmo sistema e portanto formam uma amostra aleatória de sistemas. Em ambos, o objetivo é estimar o tempo ótimo para a realização de manutenções preventivas perfeitas. É considerado como tempo ótimo de manutenção, o tempo τ que minimiza a função custo, e esta é dada por:

$$H(\tau) = \frac{1}{\tau}[C_{MP} + C_{RM}\Lambda(\tau)], \tag{3}$$

em que C_{MP} é o custo associado à manutenção preventiva e C_{RM} é o custo associado ao reparo mínimo. Pode ser mostrado (Barlow and Hunter, 1960; Gilardoni and Colosimo, 2007) que a periodicidade τ que minimiza $H(\tau)$ satisfaz:

$$\tau\lambda(\tau) - \Lambda(\tau) = C_{MP}/C_{RM}. \tag{4}$$

Nesse caso, a função intensidade de falhas deve necessariamente ser crescente. Isto porque manutenção preventiva só tem sentido para um processo de desgaste do sistema. O primeiro texto propõe uma abordagem não paramétrica para estimar τ, enquanto o segundo propõe uma abordagem Bayesiana para estimar τ através de um modelo paramétrico para λ.

O último dos três textos deste trabalho tem como foco a identificação de modelos, levando-se em consideração a heterogeneidade entre os múltiplos sistemas. Nesse texto a proposta é usar uma abordagem Bayesiana hierárquica na identificação dos modelos. Para isso, vários cenários são considerados e critérios de seleção auxiliam na escolha do modelo que melhor se ajusta aos dados analisados.

2 Resumo dos Textos

Como citado anteriormente, esta tese é formada por três trabalhos distintos, todos sobre modelagem e/ou estimação em múltiplos processos de Poisson não homogêneos acompanhados por tempos parcialmente superpostos. Nos dois primeiros os processos são considerados independentes e identicamente distribuídos. No terceiro deles é levada em consideração a heterogeneidade entre os processos. Como o primeiro texto utiliza fundamentalmente resultados de trabalho ainda não disponibilizado para consulta, reproduziremos aqui as principais contribuições do trabalho em questão, reservando o mérito pela produção aos autores do texto original: Gustavo Gilardoni e Enrico Colosimo.

2.1 Uma abordgaem não paramétrica e intevalos de confiança bootstrap para o tempo ótimo de manutenção de sistemas reparáveis

Título original: "*Optimal Maintenance Time for Repairable Systems: A Nonparametric Approach and Bootstrap Confidence Intervals*"

A abordagem não paramétrica do primeiro texto é baseada no trabalho intitulado *On the superposition of overlapping Poisson processes and nonparametric estimation of their intensity function*, de Gilardoni e Colosimo que propõe usar a escala *Tempo Total sob Teste* (TTT) para transformar vários processos de Poisson parcialmente superpostos em um único processo. Dessa forma o processo de estimação da função intensidade pode ser realizado a partir de resultados conhecidos na literatura. O trabalho citado acima é motivado por um conjunto de dados consistindo da história de falhas de 40 transformadores de potência elétrica, cujas evi-dências e considerações na área de engenharia, sugerem serem menos confiáveis com o passar do tempo, no sentido de que a frequência das falhas aumenta com a idade desses transformadores. Daí ser razoável modelar a história de falhas como 40 realizações independentes de um mesmo PPNH com uma função intensidade crescente, e calcular o Estimador de Máxima Verossimilhança Não Paramétrico (EMVNP) de λ sob essa restrição de monotonicidade. O termo "crescente" citado aqui é usado num sentido amplo, ou seja, dizemos que $\hat{\lambda}$ é crescente se $t_1 < t_2$ implica que $\hat{\lambda}(t_1) \leq \hat{\lambda}(t_2)$. Quando apenas uma realização do PPNH é observada, com eventos ocorrendo nos tempos $0 < t_1 < \cdots < t_n \leq T$, Boswell (1966) mostrou que o EMVNP, sujeito ao crescimento da função intensidade do processo, é dado pela derivada à direita da máxima minorante convexa do número acumulado de eventos, que é o estimador de Nelson-Aalen (Aalen, 1978) de Λ. Essa abordagem é rapidamente estendida ao caso de múltiplos PPNHs, observados todos até o mesmo tempo T. Por outro lado, o estimador não paramétrico de máxima verossimilhança de uma função média convexa Λ não é a máxima minorante convexa do estimador de Nelson-Aalen, quando os processos são acompanhados por tempos distintos, como é o caso dos dados dos transformadores. Assim, para fazer inferências nesse caso, propõe-se reduzir os processos

parcialmente superpostos a apenas uma realização e usar o resultado contido em Boswell (1966). A transformação TTT é a base de todo o trabalho e as proposições reproduzidas a seguir, dão o suporte à abordagem proposta. Antes algumas definições são necessárias:

Considere K realizações independentes $N_i(t)$, truncadas respectivamente nos tempos T_i, cada uma com intensidade $\lambda_i(t) = \lambda(t)I(t \leq T_i)$ e seja t_{ij} o tempo do j-ésimo evento da i-ésima ralização, $j = 1, \ldots, n_i$, $i = 1, \ldots, K$, em que $I(t \leq T_i) = 1$, para $t \leq T_i$ e 0 caso contrário.

Defina a transformação TTT como sendo $R(t) = \int_0^t r(u)\, du = \sum_{i=1}^K \min(t, T_i)$
e seja $r(t) = \sum_{i=1}^K I(t \leq T_i)$ o número de realizações sob risco no tempo t.

Proposição 1 *O processo* $\{N_S(y) = \sum_{i=1}^K N_i[R^{-1}(y)] : y > 0\}$ *é um PPNH com intensidade* $\lambda_S(y) = \lambda[R^{-1}(y)]\, I(y \leq \sum_{i=1}^K T_i)$. *Além disso,* $\{N_S(y)\}$ *é suficiente e completa.*

Proposição 2 *Sejam* t_{ij} , $i = 1, \ldots, K$, $j = 1, \ldots, n_i$, *os tempos dos eventos de* K *realizações independentes de um PPNH, truncadas respectivamente nos tempos* $T_1, \ldots, T_K$. *Sejam* $0 < y_1 < \cdots < y_n < \sum_{i=1}^K T_i$ *os tempos dos eventos ordenados do processo superposto* $N_S(y)$ *e faça* $y_{n+1} = \sum_{i=1}^K T_i$. *O estimador não paramétrico restrito,* $\hat{\lambda}$, *de* λ *é uma função escada com saltos em um subconjunto dos* $t_{ij}s$, $\hat{\lambda}[R^{-1}(y_1) - 0] = 0$ *e, para* $j = 1, \ldots, n$,

$$\hat{\lambda}[R^{-1}(y_j)] = \max_{1 \leq h \leq j} \min_{j \leq k \leq n+1} \frac{k-h}{y_k - y_h}\,. \tag{5}$$

Com base nas proposições acima, τ em 4 é estimado a partir da estimação de λ em 5. Intervalos de confiança bootstrap para τ são obtidos a partir de reamostragem feita nos sistemas, e posterior cálculo de $\hat{\lambda}$ em cada reamostra, como proposto em Field and Welsh (2007). Esses intervalos são comparados aos obtidos quando a reamostragem é feita nos tempos de falha do processo superposto. Veja metodologia em Cowling et al. (1996).

2.2 Inferência Bayesiana para o processo lei de potências com aplicação em sistemas reparáveis

Título original: "*Bayesian Inference for Power Law Processes with Applications in Repairable Systems*"

Nesse texto o foco é também a estimação da função intensidade de falhas em processos de Poisson não homogêneos, com objetivo de estimar o tempo ótimo de manutenção de sistemas reparáveis. Considerando um único PPNH, é proposta uma função paramétrica para modelar λ e o Processo Lei de Potências (PLP) é o modelo assumido. O PLP é caracterizado pela seguinte função intensidade de falhas (Crow, 1974): $\lambda(t) = \frac{\beta}{\theta}\left(\frac{t}{\theta}\right)^{\beta-1}$, e sua função média é dada por $\Lambda(t) = EN(t) = \int_0^t \lambda(u)\, du = \left(\frac{t}{\theta}\right)^{\beta}$, em que $\theta > 0$ e $\beta > 0$. β é a elasticidade do número médio de eventos com respeito ao tempo, enquanto θ é tempo para o qual esperamos observar um único evento. O PLP é sem dúvida o modelo

mais atrativo no contexto de modelagem em sistemas reparáveis, por comportar igualmente funções intensidade crescentes ($\beta > 1$), decrescentes ($\beta < 1$) ou constantes ($\beta = 1$). A abordagem Bayesiana é o principal diferencial desse trabalho, que inclui uma vasta discussão sobre diversas distribuições a *priori* para os parâmteros do processo lei de potências. Na proposta Bayesiana, a interpretação operacional dos parâmetros do modelo é de fundamental importância, para se extrair do pesquisador a informação necessária ao estabelecimento apropriado de distribuições de probabilidade a *priori*. Nesse sentido, é proposto que a intensidade PLP seja reparametrizada em termos de (β, η), em que $\eta = \left(\frac{T}{\theta}\right)^{\beta}$ é o número médio de eventos para o período no qual o sistema foi efetivamente observado. É mostrado que β e η são ortogonais e que a função de verossimilhança fica proporcional ao produto de funções densidade de variáveis aletórias com distribuição gama, nessa nova parametrização. Portanto, a família conjugada natural de distribuições a *priori* é também um produto de funções densidade de variáveis aletórias com distribuição gama. A ideia é estendida ao caso em que várias realizações de um mesmo PLP são observadas por períodos de tempo parcialmente superpostos.

2.3 Identificação de heterogeneidade de sistemas reparáveis

Título original: "*Heterogeneity Identification of Repairable Systems* "

A abordagem tradicional dos dois textos anteriores, considera que os sistemas são independentes e identicamente distribuídos. Sendo assim, conduz uma análise estatística baseada em uma única amostra de sistemas. Contudo os sistemas podem ser heterogêneos devido a variáveis não medidas. A fim de verificar essa suposição, uma abordagem clássica e uma Bayesiana são propostas nesse texto. Alguns possíveis modelos, considerando a heterogeneidade de diferentes sistemas, são comparados usando modelos frequentistas e modelos Bayesianos hierárquicos. Critérios de informação e testes da razão de verossimilhanças são usados para selecionar o modelo mais provável para um particular conjunto de dados. A abordagem é essencialmente paramétrica e novamente o modelo adotado é o processo lei de potências. Os cenários propostos foram idealizados de maneira a comportar a máxima heterogeneidade possível entre os sistemas, em que os β's e os θ's dos PLPs são todos diferentes entre si, e a mínima heterogeneidade, em que todos os β's são iguais a 1 e todos os θ's são iguais. Além desses, alguns cenários intermediários são levados em consideração. Na proposta Bayesiana hierárquica, novamente uma reparametrização da função intensidade é considerada, visando garantir tanto a proposição de uma distribuição a *priori* adequada para ambos os parâmetros do PLP, bem como a permutabilidade destes em sua distribuição conjunta de probabilidades a *priori*. Assim, alternativamente à versão anterior, é proposto um parâmetro auxiliar $\xi = \Lambda(1) = \left(\frac{1}{\theta}\right)^{\beta}$. O parâmetro ξ é utilizado na construção de uma distribuição a *priori* para θ em cada PLP, condicional ao parâmetro β. A metodologia é aplicada a dados de registros de falhas de 11 transformadores de potência elétrica.

Optimal Maintenance Time for Repairable Systems: A Nonparametric Approach and Bootstrap Confidence Intervals

Abstract

Statistical models are of great interest in the context of repairable systems maintenance. Usually, a repairable system operates under a maintenance strategy that calls for complete preventive repair actions at pre-scheduled times and minimal repair actions whenever a failure occurs. Nonhomogeneous Poisson Process (NHPP) is frequently adopted to model minimum repair maintenance history of failure times. In general, papers in the literature assume a parametric form for the NHPP intensity function. Therefore, statistical inference for the unknown quantities can proceed by using frequentist or Bayesian approaches. This paper proposes a nonparametric inference approach to obtain the optimal preventive policy as a function of the intensity function of the process. Superposition of many overlapping similar NHPPs in a time scale allows us to obtain a nonparametric estimate of the intensity function. Optimal preventive maintenance time is estimated in this scale and transformed back to the original one. Paper main goal is related to the comparison of bootstrap confidence intervals. Two bootstrap methods proposed in the literature are used in order to build confidence intervals for the optimal maintenance time. The first one is based on samples with replacement in the original scale and the second one considers the data in the transformed scale. The methods are illustrated using a real data set consisting of the failure histories of 40 electrical power transformers.

1 Introduction

Statistical models for recurrent events are of great interest in the context of repairable systems maintenance (Ascher and Feingold, 1984; Bain and Engelhardt, 1991; Meeker and Escobar, 1998; Rigdon and Basu, 2000)). Repairable systems are allowed to experience more than one failure throughout theirs life. After each failure, a maintenance activity is necessary for the system return to the operation condition. In general, such activity only replace the damaged part of the equipment, leaving it in the same condition as it was just before the failure. This kind of repair is usually called Minimum Repair (MR). On the other hand, Perfect Repair or Preventive Maintenance (PM) renews the entire system. The combination of PM and MR has been of interest since the work of Barlow and Hunter (1960). Some of the developments for this situation can be found in Gerstack (1977), Block et al. (1990) and Park et al. (2000).

PM is performed at periodic intervals of time. This policy specifies the periodicity with which a system is maintained. Optimal PM check points that minimize expected cost is a fundamental aspect for any repairable system.

Statistical inference procedures for the maintenance stochastic models developed in the literature are still under consideration. In general, papers in the literature takes a parametric form for the NHPP intensity function. Power law intensity function is the most used one (Crow, 1974). Statistical inference for the unknown quantities can proceed by using frequentist (Gilardoni and Colosimo, 2007) or Bayesian (Guida et al., 1989; Pan and Rigdon, 2009) approaches. This paper, on the other hand, takes a nonparametric inference approach in order to estimate the intensity function under growing restriction. It follows the steps proposed by Gilardoni and Colosimo (2010). Total Time on Test (TTT) scale is used to transform several superimposed NHPPs into only one. Intensity function is, therefore, estimated in the TTT scale. Optimal preventive maintenance time is obtained in this scale and transformed back to the original one.

Bootstrap confidence intervals are used to estimate the optimal maintenance time. Two bootstrap methods proposed in the literature can be used to build confidence intervals for this quantity. The first one is based on the original scale (Field and Welsh, 2007). It is a simple cluster bootstrap based on drawing sampling of clusters independently with replacement. The second one is done over the transformed TTT scale. There are some bootstrap methods available in the literature for just one NHPP (Cowling et al., 1996). In general, they produce similar results. Therefore, we decided to use in this paper the simplest one in computation terms.

This paper is motivated by a data set consisting of the failure histories of 40 electrical power transformers. The data set is shown in Figure 1. Company interest was in finding an optimal PM check points that minimize expected cost.

The paper is organized as follows. In Section 2 we obtain the expected cost per unit of time for each PM policy. This expected cost is then minimized to obtain the optimal policy. Optimal time is obtained in terms of the process intensity function. A nonparametric estimate for this function is obtained in Section 3 under growing restriction. In Section 3.1, a known estimator for this situation is presented for just one system. This estimator is known as the Greatest Convex Minorant (GCM). It is an estimator of a convex cumulative intensity function for only one NHPP. In Section 3.2 we propose a nonparametric estimator of the intensity function based on several overlapping realizations from NHPP. After transforming several NHPP realizations into only one by using the TTT time scale, the methodology established for only one NHPP can be used. In Section 5, the proposed methodology is

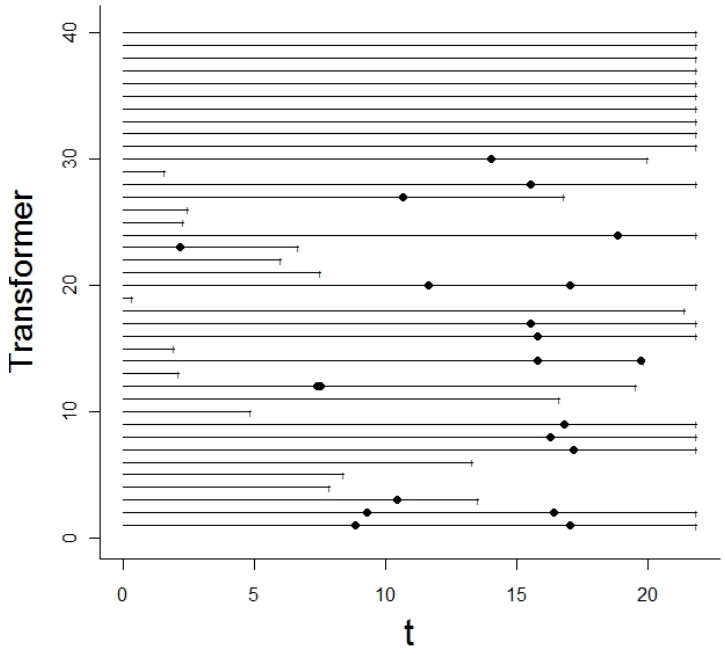

Figure 1: Failure Histories of Electrical Power Transformers.
Dots represents failure times. Time unit is 1000 hours.

applied to the transformers data set displayed in Figure 1. Confidence intervals based on systems and on the transformed scale are obtained and compared for the optimal preventive time. Paper ends with some final remarks in Section 6.

2 Optimal Maintenance Time

Usually, the model adopted under MR is a NHPP, $\{N(t)\colon t \geq 0\}$, where $N(t)$ is the number of failures in the time interval $(0, t]$. In general, $N(t)$ model is determined by the mean function

$$E(N(t)) = \Lambda(t) = \int_0^t \lambda(u)\, du, \tag{1}$$

where the intensity function, $\lambda(t)$, is defined as:

$$\lambda(t) = \lim_{\Delta t \to 0} \frac{P(N(t, t + \Delta t] = 1)}{\Delta t}.$$

Consider a repairable system, modeled by a PPNH, under an increasing intensity function. Preventive maintenance can only be justified under system degration over time. That is, under an increasing intensity function. System is observed in the interval $(0, T]$. The expected total cost of maintenance policy, $H_{(0,T]}(\tau)$, is obtained by considering that a perfect PM will be held every τ time units. Time interval $(0, T]$ is decomposed as $(0, \tau] \cup (\tau, 2\tau] \cup \cdots \cup ((m-1)\tau, m\tau] \cup (m\tau, T]$, where m is the largest integer smaller than or equal to T/τ. Figure 2 illustrates this situation.

Figure 2: Decomposing the time interval $(0, T]$.

The limiting expected cost per unit of time is given by

$$H_{(0,T]}(\tau) = \lim_{T\to\infty} \frac{C_{(0,\tau]}(\tau)}{T} = \frac{1}{\tau}[C_{MP} + C_{RM}\Lambda(\tau),$$

where C_{MP} and C_{RM} are maintenance PM and MR costs, respectively. It can be shown (Barlow and Hunter, 1960; Gilardoni and Colosimo, 2007) that the periodicity τ that minimizes $H(\tau)$ satisfies:

$$\tau\lambda(\tau) - \Lambda(\tau) = C_{PM}/C_{MR}. \tag{2}$$

In order to estimate τ it is necessary to estimate Λ, and consequently λ. C_{PM} and C_{MR} are fixed values supplied by the power company.

3 Nonparametric Inference for a Convex Intensity Function

There are basically two ways to observe data from a repairable system, depending on whether data collection is ceased after a specified number of failures k or at some predetermined time T. These sampling schemes are said to be *failure truncated* or *time truncated*, respectively. There are slight differences in the inference procedures depending on the sampling scheme considered. However, for large samples inference, differences between them vanish and it is considered just time truncation in the inference developments in this paper.

In Section 3.1, it is presented inference procedures for just one system and the extension for many systems is considered in Section 3.2. In the latter case, TTT transformation makes possible the superposition of many similar NHPPs in just one system. Therefore, the results obtained in Section 3.1 can be used for this situation. Optimal preventive maintenance time is obtained in this new scale and transformed back to the original one.

3.1 Just one System

Likelihood function for only one NHPP is given by (Rigdon and Basu, 2000):

$$L(\lambda) = \left(\prod_{i=1}^{n}\lambda(t_i)\right)\exp\left(-\int_0^T \lambda(x)dx\right), \tag{3}$$

where $0 \le t_1 < t_2 < \cdots < t_n \le T$ are the observed failure times.

Nonparametric maximum likelihood estimate (NMLE) of λ is obtained by maximizing (3) among all growing intensities. Maximum value is necessarily a step function with jumps at $t_i's$. This estimator for only one realization $\{N(t) : 0 \le t \le T\}$ is well known in the literature (Barlow et al., 1972). In this case, restricted to increasing λ, one can obtain $\hat{\Lambda}(t)$ by computing first the unrestricted NMLE for $\Lambda(t)$, $\widetilde{\Lambda}(t) = \sum_{j=1}^{n} I(t_j \le T)$, where $I(t_j \le T) = 1$, if $t_j \le T$ and 0 other case. The restricted NMLE is $\hat{\Lambda} = \sup\{f : f \text{ is convex and } f \le \widetilde{\Lambda}\}$.

That is, the Greatest Minorant Convex (GMC) of $\widetilde{\Lambda}$. $\hat{\lambda}$, the restricted NMLE for λ, is obtained by differentiating $\hat{\Lambda}$.

In a general case, this work is motivated by a situation involving several independent realizations of NHPPs. They are observed over overlapping time intervals. See, for example, Figure 1. It is therefore necessary to propose a nonparametric restricted estimator for λ based on K independent systems.

3.2 K Independent Systems

Suppose K independent realizations of the same NHPP, let say $N_1(t), \ldots, N_K(t)$, observed, respectively up to times $T_1, \ldots, T_K$. Let t_{ij} the time of j-th event of i-th realization, $i = 1, \cdots, K$; $j = 1, \cdots, n_i$, of the process. Likelihood function became:

$$L(\lambda) = \prod_{i=1}^{K}\prod_{j=1}^{n_i}\lambda(t_{ij}) \times \exp\left[-\sum_{i=1}^{K}\int_0^{T_i}\lambda(x)dx\right]. \tag{4}$$

In the especial case that $T_1 = T_2 = \cdots = T_K = T$, the superimposed process, $N_+(t) = \sum_{i=1}^{K} N_i(t)$, is also a NHPP with intensity $\lambda_+(.) = K\lambda(.)$, observed up to time T. Therefore, an estimate for $\lambda_+(.)$ can be obtained as in Section 3.1 by $\hat{\lambda} = \hat{\lambda}_+(.)/K$.

In a general situation, as our electrical company data set of interest, $T_i \neq T_j$ for any $i, j = 1, \ldots, K$. In this situation Gilardoni and Colosimo (2010) proposed to use the TTT transformation:

$$R(t) = \int_0^t r(u)\,du = \sum_{i=1}^{K}\min(t, T_i).$$

They also defined $R^{-1}(y)$ to be the usual inverse of R when $y < \sum_{i=1}^{K} T_i$ and equal to $\max\{T_1, \ldots, T_K\}$ when $y \geq \sum_{i=1}^{K} T_i$. It can be observed that the inner term of (4) can be reexpressed as

$$\sum_{i=1}^{K}\int_0^{T_i}\lambda(u)\,du = \int_0^{\sum_{i=1}^{K} T_i}\lambda[R^{-1}(y)]\,dy\,, \tag{5}$$

and therefore the likelihood (4) becomes

$$\begin{aligned} L(\lambda) &= \prod_{i=1}^{K}\prod_{j=1}^{n_i}\lambda(t_{ij}) \times \exp\left[-\int_0^{\sum_{i=1}^{K} T_i}\lambda[R^{-1}(y)]\,dy\right] = \\ &= \prod_{i=1}^{K}\prod_{j=1}^{n_i}\lambda_S(y_{ij}) \times \exp\left[-\int_0^{\sum_{i=1}^{K} T_i}\lambda_S(y)\,dy\right], \end{aligned} \tag{6}$$

where $\lambda_S(y) = \lambda[R^{-1}(y)]$ and $y_{ij} = R(t_{ij})$.

The main result proved by Gilardoni and Colosimo (2010) established that the resulting process $N_S(y) = \sum_{i=1}^{K} N_i[R^{-1}(y)]$ is also a NHPP, and is sufficient for inferences about λ. According to their results, estimating of Λ can obtained by the process, $N_S(y) = \sum_{i=1}^{K} N_i[R^{-1}(y)]$ following the steps:

- (i) get the restricted NMLE, $\hat{\Lambda}$, as the GMC of the unrestricted estimate $\widetilde{\Lambda}$ as described in Section 3.1;
- (ii) return $\hat{\Lambda}$ to the original scale of the data set;

- (iii) differentiating $\hat{\Lambda}$ in order to get restricted $\hat{\lambda}_{NP}$;
- (iv) substitute $\hat{\Lambda}$ and $\hat{\lambda}$ on (2) to get $\hat{\tau}$, the estimate of the optimal time.

The four steps described above are illustrated in Figure 3 using our data set of interest that is displayed in Figure 1. In Figure 3, C in step (iv) is the costs ratio. That is, $C = C_{PM}/C_{MR}$.

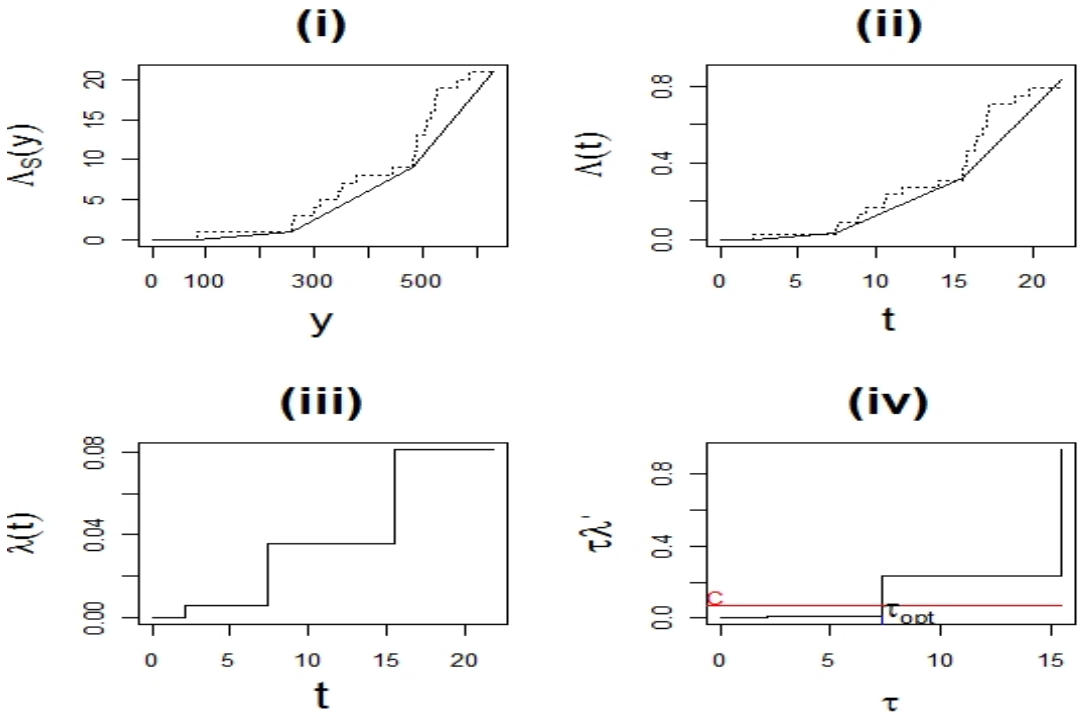

Figure 3: (i) $\widetilde{\Lambda}$ (dashed) and $\hat{\Lambda}$ (solid) on TTT scale, (ii) $\widetilde{\Lambda}$ (dashed) and $\hat{\Lambda}$ (solid) on original scale, (iii) $\hat{\lambda}$ and (iv) τ_{opt}, the instant in which τ times the differential of λ overcomes C.

4 Bootstrap Methods

Interval estimation for τ is fundamental. Two bootstrap methods are used in order to obtain confidence intervals for τ. The first one, Method 1, resamples are obtained straight from the sampling units, the transformers. This method is supported by Field and Welsh (2007). Transformers are treated as independent and resampled with replacement to form each one of the bootstrap samples. For each sample, TTT transformation is applied and then following the four steps described in Section 3.2 in order to get τ.

In Method 2, the data are aggregated before by applying the TTT transformation. Conditional on the observed failure times, number of failures, N, has a Poisson distribution with parameter $\hat{\Lambda}(1)$. Firstly, generate a value of $N = n$ from the Poisson distribution and next draw a bootstrap sample by sampling randomly with replacement from the transformed failure times. After that, one proceeds with steps from (i) to (iv) described in Section 3.2. This is supported by Chiang et al. (2005). In both situations we adopted the percentile confidence interval, which takes the values of τ, ranging between 2.5 and 97.5 percentile. Gilardoni and

Colosimo (2010) used Method 2 in order to estimate the intensity function in the same situation. However, they used kernel estimate to obtain $\hat{\tau}$ and follow the methodology proposed by Cowling et al. (1996) to obtain bootstrap confidence bands for it.

Section 5 applies the methodology presented in Sections 3 and 4 for the power transformers data set.

5 Application

Data set supplied by the electrical company, consists of 21 failure times observed in 40 transformers. This data set is displayed in Figure 1. The company also provided the costs ratio, $C = C_{MR}/C_{PM} = 15$, necessary to obtain τ in (2). Figure 3(i) shows the convex form of the cumulative intensity function. That means, the intensity function is increasing. Restricted intensity function is estimated in Figure 3(iii). Finally the optimal τ is estimated in Figure 3(iv). Numeric results are presented in Table 1. 5000 bootstrap samples were generated for each method and they were used to obtain respective confidence intervals. Maximum likelihood estimates (MLE) obtained by Gilardoni and Colosimo (2007) for τ are also reported in Table 1. They adopted the Power Law Process parametric form for the intensity function. Results were obtained by using the statistical software R.

Table 1: τ estimates.

	MLE	NMLE	
Estimate	6.285	7.396	
Interval	Asymptotic	BM 1	BM 2
	[4.870;7.701]	[1.231;15.550]	[1.254;11.664]
Range	2.831	14.319	10.410

BM 1 - Bootstrap in the transformers. BM 2 - Bootstrap in the TTT scale.

As expected, the range of nonparametric confidence intervals are larger than MLE's one. It can be noticed that bootstrap Method 1 has larger range than the one obtained by using Method 2. This fact might be an evidence that the former method is subject to variability between and intra systems while Method 2 just takes into consideration the variability between systems.

In some samples the product of the first failure time by differential of λ already exceeds costs ratio, C. It was necessary to make a slight adjustment in the optimal τ estimation, otherwise $\hat{\tau}$ would be the first failure time. We proposed an interpolation so that τ became C divided by the differential of λ in such situations. In Method 1, it was necessary to interpolate 2297 time while in Method 2 it happened 1821 times.

6 Final Remarks

This paper presents a nonparametric approach to estimate the intensity function of a NHPP, under growing restriction. The TTT time transformation was used in order to turn several NHPPs into only one and then apply methodologies well established for just one system. The aim of this paper was to estimate the preventive maintenance optimal time, τ. Confidence intervals for τ were obtained by using bootstrap samples in the original scale and in the TTT one. The last one showed better performance than the former one.

We have to take care when getting bootstrap samples in this context. Resampling with replacement implies in getting the same system more than once. So the final results, using the TTT transformation, might not be a PPNH any more.

References

H. Ascher and H. W. Feingold, *Repairable Systems Reliability: Modeling, Inference, Misconception and their Causes*, Marcel Dekker: New York, 1984.

L. J. Bain and M. Engelhardt, *Statistical Analysis of Reliability and Life-Testing Models, Theory and Methods* Marcel Dekker: New York, 1991.

R. E. Barlow and L. C. Hunter, Optimum Preventive Maintenance Policies. *Operations Research* 8, pp. 90-100, 1960.

R. E. Barlow, D. J. Bartholomew, J. M. Bremner, and H. D. Brunk, *Statistical Inference Under Order Restrictions: The Theory and Application Of Isotonic Regression*, J. Wiley: London, 1972.

H. W. Block, W. S. Borges and T. H. Savits, A General Age Replacement with Minimal Repair. *Naval Research Logistics* 35, pp. 365-372, 1990.

C. T. Chiang, M. C. Wang, and C. Y. Huang. Kernel estimation of rate function for recurrent event data. *Scandinavian Journal of Statistics* 32, pp.77-91, 2005.

A. Cowling, P. Hall, and M. J. Phillips, Bootstrap Confidence Regions for the Intensity of a Poisson Point Process. *Journal Of The American Statistical Association 91*, 436, pp.1516-1524, 1996.

L. R. Crow, Reliability Analysis for Complex Systems. *Reliability and Biometry*, F. Proschan and J. Serfling (Eds), pp. 379-410, 1974.

C. A. Field and A. H. Welsh, Bootstrapping clustered data. *Journal of Royal Statistical Society*, 69, pp. 369-390, 2007.

I. B. Gerstack, Models of Preventive Maintenance. North Holand: Amsterdam, 1977.

G. L. Gilardoni and E. A. Colosimo, Optimal maintenance time for repairable systems. *Journal of Quality Technology* 39, pp. 48-53, 2007.

G. L. Gilardoni and E. A. Colosimo, On the Superposition of Overlapping Poisson Processes and Nonparametric Estimation of their Intensity Functions. 2010.

M. Guida, R. Calabria and G. Pulcini, Bayes Inference for a Nonhomogeneous Poisson Process with Power Intensity Law. *IEEE Transactions on Reliability* 38, pp. 603-609, 1989.

W. Q. Meeker and L. A. Escobar, *Statistical Methods for Reliability Data*. John Wiley: New York, 1998.

R. Pan and S. E. Rigdon, Bayes Inference for General Repairable Systems. *Journal of Quality Technology* 41, pp. 82-94, 2009.

D. H. Park, G. M. Jung and J. K. Yum, Cost Minimization for Periodic Maintenance Policy of a System Subject to Slow Degrada- tion. *Reliability Engineering & System Safety* 68, pp. 105-112, 2000.

S. E. Rigdon and A. P. Basu, *Statistical Methods for the Reliability of Repairable Systems*. John Wiley: New York, 2000.

Bayesian Inference for Power Law Processes with Applications in Repairable Systems

Maristela Dias de Oliveira[1], Enrico A. Colosimo[2] and Gustavo L. Gilardoni[3]
Universidade Federal da Bahia, Universidade Federal de Minas Gerais and Universidade de Brasília

Abstract. Statistical models for recurrent events are of great interest in repairable systems reliability and maintenance. The adopted model under minimal repair maintenance is frequently a Nonhomogeneous Poisson Process with the Power Law Process (PLP) intensity function. Although inference for the PLP is generally based on maximum likelihood theory, some advantages of the Bayesian approach have been reported in the literature. In this paper it is proposed that the PLP intensity be reparametrized in terms of (β, η), where β is the elasticity of the mean number of events with respect to time and η is the mean number of events for the period in which the sistem was actually observed. It is shown that β and η are orthogonal and that the likelihood becomes proportional to a product of Gamma densities. Therefore, the family of natural conjugate priors is also a product of Gammas. The idea is extended to the case that several realizations of the same PLP are observed along overlapping periods of time. The results are applied to a real problem concerning the determination of the optimal periodicity of preventive maintenance for a set of power transformers. Some Monte Carlo simulations are provided to study the frequentist behavior of the Bayesian estimates and to compare them with the maximum likelihood estimates.

Keywords. Conjugate prior, optimal maintenance, Poisson Process, posterior distribution, reference priors.

[1]Departamento de Estatística, Universidade Federal da Bahia, Salvador, Brazil. *e-mail*: maridias@est.dout.ufmg.br. Research partially supported by CAPES grants.
[2]Departamento de Estatística, Universidade Federal de Minas Gerais, Belo Horizonte, MG, 31270–901, Brazil. *e-mail*: enricoc@est.ufmg.br. Research partially supported by FAPEMIG, CAPES and CNPq grants.
[3]Departamento de Estatística, Universidade de Brasília, Brasília, DF 70910–900, Brazil. *e-mail*: gilardon@unb.br. Research partially supported by CAPES, CNPq and FINATEC grants.

1 Introduction

Statistical models for recurrent events have been investigated in many papers in the literature. Such models are of great interest to study the reliability and maintenance policies for repairable systems (Ascher and Feingold (1984), Bain and Engelhardt (1991), Rigdon and Basu (2000)). Frequently, the adopted model under minimal repair maintenance is a Nonhomogeneous Poisson Process (NHPP), $\{N(t)\colon t \geq 0\}$, where $N(t)$ is the number of failures from the beginning of the follow-up until time t (Barlow and Hunter, 1960). A flexible parametric form for the intensity function of the NHPP is

$$\lambda(t) = \frac{\beta}{\theta}\left(\frac{t}{\theta}\right)^{\beta-1}, \tag{1}$$

with mean function

$$\Lambda(t) = EN(t) = \int_0^t \lambda(u)\,du = \left(\frac{t}{\theta}\right)^{\beta},$$

where $\theta > 0$ and $\beta > 0$. This model, known as the Power Law Process (PLP), was proposed by Crow (1974) and since then it has become the most popular parametric intensity in the repairable systems literature. The intensity function is increasing for $\beta > 1$, decreasing for $\beta < 1$ and constant (i.e. the NHPP is actually a Homogeneous Poisson Process) for $\beta = 1$. Adequacy of the PLP for a particular data set can be diagnosed graphically using either Duane plots (Duane, 1964) or some modified Total Time on Tests plots (Klefsjö and Kumar, 1992). More formal hypotheses tests are considered by Baker (1996) and Bhattacharjee *et al.* (2004).

Statistical inference for the PLP is generally based on the maximum likelihood estimator (MLE) and its asymptotic properties (Berman and Turner (1992), Zhao and Xie (1996)). However, some papers appeared in the literature using the Bayesian approach for the PLP model (Sen (2002), Guida *et al.* (1989)). The Bayesian approach deals with the uncertainty of the parameters in the model used to describe a recurrent system. A prior distribution is assumed to represent the uncertainty in the model parameters before the current data is observed. Reference prior distributions have been used in the Bayesian context by Guida *et al.* (1989), Sen (2002) and Yu *et al.* (2006), among others. On the other hand, identifying a family of conjugate prior distributions will often result in mathematical and computational simplifications. Moreover, a conjugate prior distribution can be interpreted as additional data, hence making prior elicitation easier (Raiffa and Schlaifer, 1961; Gelman *et al.*, 2003). Huang and Bier (1998), Huang (2001) and Kim *et al.* (2008) have proposed a conjugate prior distribution for the parameters of the PLP looking at the functional form of the likelihood function. More precisely, suppose that we observe n events at times $t_1 < \cdots < t_n$ and that the process has been either *time truncated* (T is fixed, n is random and $t_n < T$) or *failure truncated* (n is fixed, $T = t_n$ is random). In both cases the likelihood function is

$$L(\beta,\theta) = [\prod_{i=1}^{n} \lambda(t_i)]\, e^{-\Lambda(T)} = \beta^n \theta^{-n\beta}\, [\prod_{i=1}^{n} t_i]^{\beta-1}\, \exp\{-(T/\theta)^{\beta}\} \tag{2}$$

(Berman and Turner, 1992; Rigdon and Basu, 2000). Huang and Bier (1998) parametrize the intensity (1) in terms of β and $\lambda_0 = \theta^{-\beta}$ and, by analogy with the resulting likelihood, find the four parameters (i.e. m, c, α and y below) conjugate family

$$\pi(\beta, \lambda_0) \propto \lambda_0^{m-1} \beta^{m-1} [y^m e^{-\alpha}]^{\beta-1} \exp\{-\lambda_0 c y^\beta\} . \tag{3}$$

They go on by finding the normalizing constant and some moments and discussing properties of the resulting posterior. However, the whole approach becomes somewhat difficult because the parameter λ_0 lacks an operational interpretation and the distributions in the family (3) do not belong to any known class. Motivated by this, and also partly by results obtained by Sen and Khattree (1998) and Sen (2002) for the posterior analysis under noninformative priors of the form $(\theta\, \beta^\delta)^{-1}$, we propose here to parametrize the problem in terms of β and $\eta = \Lambda(T) = (T/\theta)^\beta$. On one side, β and η have simple operational definitions which will often make prior elicitation easier. On the other side, in the (β, η) parametrization the likelihood (2) becomes

$$L(\beta, \eta) = c\, [\beta^n e^{-n\beta/\hat{\beta}}]\, [\eta^n e^{-\eta}] \propto \gamma(\beta|\, n+1, n/\hat{\beta})\, \gamma(\eta|\, n+1, 1) , \tag{4}$$

where $c = \prod_{j=1}^n t_j^{-1}$, $\hat{\beta} = n/\sum_{j=1}^n \log(T/t_j)$ is the MLE of β and $\gamma(x|\, a, b) = b^a x^{a-1} e^{-bx}/\Gamma(a)$ $(x, a, b > 0)$ is the density of the Gamma distribution with shape and scale parameters equal to a and b, respectively. It follows then that β and η are orthogonal and the natural conjugate family has densities of the form

$$\pi(\beta, \eta) = \gamma(\beta|\, a_\beta, b_\beta) \times \gamma(\eta|\, a_\eta, b_\eta) , \tag{5}$$

where the prior parameters a_β, b_β, a_η and b_η must all be positive if we want $\pi(\beta, \eta)$ to be proper, although non positive values can also be entertained as long as the posterior becomes proper. The posterior density is

$$\pi(\beta, \eta|\, t_1, \ldots, t_n, T) \propto L(\beta, \eta)\, \pi(\beta, \eta) \propto \gamma(\beta|\, a_\beta + n, b_\beta + n/\hat{\beta}) \times \gamma(\eta|\, a_\eta + n, b_\eta + 1) , \tag{6}$$

so that both a priori and a posteriori β and η are independent, each following a Gamma distribution. Hence, the form of the posterior is quite tractable and, even when dealing with parameters whose exact expectations are difficult to attain, it is quite easy to obtain accurate approximations based on an *i.i.d.* posterior sample.

There are some advantages of using the Bayesian approach in this situation. First, even if the likelihood takes essentially the same form, for both *time* and *failure truncation*, the sampling distributions of the MLEs are different and hence a different analysis is required for each case when using the frequentist approach. On the contrary, those two situations can be considered simultaneously in the Bayesian approach, since the posterior distribution will be the same (provided, of course, that we use the same prior). Second, the use of MLEs is justified on asymptotic grounds and may require somewhat sophisticated arguments such as appropriate reparametrizations to avoid extremely skewed sampling distributions, while

the Bayesian approach deals quite naturally with small sample sizes and skewed posterior distributions (for instance, the MLE $\hat{\beta}$ is only defined when there has been observed at least one event). Third, the Bayesian approach will typically produce results similar to those based on MLEs when using appropriate reference priors. However, it also allows for the introduction of external information to the data through the use of informative priors. Moreover, the elicitation of such kind of prior distributions is facilitated in our approach due to the operational interpretation attached to the parameters β and η (cf. subsection 2.2). Fourth, even when using approximate inferences based on a posterior sample, the Bayesian approach deals quite naturally when dealing with several quantities of interest. On the other hand, obtaining estimates of standard errors for several MLEs may require the algebraic calculation of gradients and hessians for each one of them. Finally, the Bayesian approach also deals naturally with restrictions on the parameter space. For instance, inference about the optimal maintenance time mentioned below requires the intensity to be increasing (i.e. $\beta > 1$). In the Bayesian approach one simply specifies a prior distribution truncated for $\beta > 1$ and the analysis proceeds in about the same manner as for the unrestricted case. However, the frequentist approach will have trouble dealing with this situation when the MLE $\hat{\beta}$ lies close to $\beta = 1$, which is indeed the case for the transformers data set discussed in Section 5.

The parametrization (β, η) suggests rather easily how to treat the case when several realizations of the PLP are observed along overlapping time intervals. Although this case appears frequently in practice, because repairable systems are usually observed in different time intervals (truncation times), methodological developments have been somewhat lacking in the literature, especially in the Bayesian setting. More precisely, suppose that K realizations of the same PLP have been observed and let t_{ij} denote the j-th event time for the i-th realization ($j = 1, \ldots, n_i$ and $i = 1, \ldots, K$). Let T_i be the truncation time corresponding to the $i - th$ realization. Then we show in Section 3 that the parameters β and $\eta = \sum_{i=1}^{K} \Lambda(T_i) = \sum_{i=1}^{K} (T_i/\theta)^{\beta}$ are orthogonal and that, under the prior specification (5), the posterior has the same form (6) but with an additional factor which does not depend on η and is proportional to $\exp\{KL[(\frac{T_1^{\hat{\beta}}}{\sum_{h=1}^{K} T_h^{\hat{\beta}}}, \cdots, \frac{T_K^{\hat{\beta}}}{\sum_{h=1}^{K} T_h^{\hat{\beta}}}) || (\frac{T_1^{\beta}}{\sum_{h=1}^{K} T_h^{\beta}}, \cdots, \frac{T_K^{\beta}}{\sum_{h=1}^{K} T_h^{\beta}})]\}$, where $KL[\cdot||\cdot]$ is the Kullback-Leibler divergence. Hence, the form of the posterior lends itself to an easy i.i.d. simulation using for instance the rejection sampling algorithm.

Our interest in the case of many overlapping realizations stems mainly from a real application concerning the estimation of the optimal maintenance time for a set of power transformers, which we discuss in Section 5. In short, consider a repairable system modeled by a NHPP with an increasing intensity function subject to two types of repairs: either a *minimal repair* after a failure which restores the system (i.e. the intensity) to exactly the same level it was immediately before the failure or a *preventive maintenance* which restores the system to "as good as new" condition. If the preventive maintenances are performed every τ units of time, the expected cost per unit of time is

$$H(\tau) = [C_{PM} + C_{MR} EN(\tau)]/\tau = [C_{PM} + C_{MR}\Lambda(\tau)]/\tau\,, \tag{7}$$

where C_{MR} and C_{PM} are the expected costs associated to the two types of repair actions. It can be shown (Barlow and Hunter, 1960; Gilardoni and Colosimo, 2007) that the periodicity τ which minimizes $H(\tau)$ satisfies that $\tau\lambda(\tau) - \Lambda(\tau) = C_{PM}/C_{MR}$. In the special case of the PLP, τ becomes

$$\tau = \theta \left[\frac{C_{PM}}{(\beta - 1) C_{MR}} \right]^{1/\beta} . \tag{8}$$

However, inference about τ only makes sense when $\beta > 1$, leading to the necessity of truncating the prior density for β. This can be done preserving conjugacy by truncating the prior (5) to the set $\beta > 1$, because then the posterior density would be the same as (6) but also truncated for $\beta > 1$. However, because of the term $(\beta - 1)^{1/\beta}$ in the denominator of τ and the fact that the posterior density is non null near $\beta = 1$, the posterior expectation of τ will be infinite. Still, under the truncated prior, one can use for instance a *maximum a posteriori estimate* for the optimal time. An alternative, non-conjugate formulation, which puts less weight to values of β close to one and hence will make the posterior expectation of τ finite, is to consider a priori that $(\beta - 1)$ follows a Gamma distribution.

Besides Sections 3 and 5, which deal respectively with the many realizations setting and with the inference for the optimal periodicity for the power transformers data set, the rest of the paper is organized as follows. In Section 2 we make some additional considerations regarding inference for a single realization of the PLP. It also includes a discussion of reference and informative priors and some computational aspects when the interest is centered in a function of the parameters whose posterior expectation cannot be computed explicitly. Section 4 shows some Monte Carlo simulations that help to understand the frequentist behavior of the Bayes estimates under different prior specifications and to compare them to the MLE estimates in the case of several realizations. The simulation scenarios and prior distributions are motivated from the real case discussed in Section 5. Finally, some concluding remarks end the paper in Section 6.

2 A single PLP realization

Suppose a *time truncated* process observed in $(0, T)$, and let $\ell(\beta, \eta) = \log L(\beta, \eta) = \log c + n \log \beta - n\beta/\hat{\beta} + n \log \eta - \eta$ be the log-likelihood in the (β, η) parametrization. Since $\nabla \ell = (\partial\ell/\partial\beta, \partial\ell/\partial\eta)' = (n/\beta - n/\hat{\beta}, n/\eta - 1)'$, the maximum likelihood estimates (MLE) of β and η are $\hat{\beta}$ and n respectively. Hence, the MLE of $\theta = T\,\eta^{-1/\beta}$ is $\hat{\theta} = T\,\hat{\eta}^{-1/\hat{\beta}} = T\,n^{-1/\hat{\beta}}$. From the Fisher information matrix

$$I(\beta, \eta) = - \begin{pmatrix} E\frac{\partial^2 \ell}{\partial \beta^2} & E\frac{\partial^2 \ell}{\partial \beta\, \partial \eta} \\ E\frac{\partial^2 \ell}{\partial \beta\, \partial \eta} & E\frac{\partial^2 \ell}{\partial \eta^2} \end{pmatrix} = n \begin{pmatrix} \frac{1}{\beta^2} & 0 \\ 0 & \frac{1}{\eta^2} \end{pmatrix} , \tag{9}$$

it follows that the asymptotic covariance matrix of $(\hat{\beta}, \hat{\eta})$ is $\mathrm{Var}(\hat{\beta}, \hat{\eta}) \approx \mathrm{n}^{-1}\,\mathrm{Diag}(\beta^2, \eta^2)$.

2.1 Posterior Analysis

Let $a_\beta > -n$, $b_\beta > -n/\hat{\beta}$, $a_\eta > -n$ and $b_\eta > -1$ in (5) so that the posterior density (6) is proper. Suppose that the interest is centered in a function $\phi(\beta, \eta)$ such as $\theta = T/\eta^{1/\beta}$ or, perhaps, as in Sen (2002), the current intensity $\lambda(T) = \beta T^{\beta-1}/\theta^\beta = \beta\,\eta/T$. Under squared error loss, the Bayes estimate of ϕ is $E[\phi(\beta, \eta)|t_1, \ldots, t_n]$. For instance, the posterior expectation of the current intensity is

$$E[\lambda(T)|t_1, \ldots, t_n] = E[\frac{\beta\eta}{T}|t_1, \ldots, t_n] = \frac{1}{T}\,E[\beta|t_1, \ldots, t_n]\,E[\eta|t_1, \ldots, t_n] = \frac{1}{T}\,\frac{a_\beta + n}{b_\beta + n/\hat{\beta}}\,\frac{a_\eta + n}{b_\eta + 1}\,.$$

Credible intervals can be obtained from the posterior quantiles of ϕ. An alternative that we consider in Section 5 is to use *Maximum a Posteriori* estimates. In this case the mode of the posterior density (6) is attained for $\tilde{\beta} = (a_\beta + n - 1)/(b_\beta + n/\hat{\beta})$ and $\tilde{\eta} = (a_\eta + n - 1)/(b_\eta + 1)$. Hence, an alternative estimate for $\lambda(T) = \beta\,\eta/T$ is

$$\widetilde{\lambda}(T) = \frac{1}{T}\,\tilde{\beta}\,\tilde{\eta} = \frac{1}{T}\,\frac{a_\beta + n - 1}{b_\beta + n/\hat{\beta}}\,\frac{a_\eta + n - 1}{b_\eta + 1}\,.$$

When integration of moments or quantiles of ϕ with respect to the posterior distribution (6) is difficult, one can easily generate Monte Carlo samples $(\beta_1, \eta_1), \ldots, (\beta_m, \eta_m)$ from the posterior and approximate, for instance, $E[\phi(\beta, \eta)|t_1, \ldots, t_n]$ by $m^{-1}\sum_{h=1}^{m}\phi(\beta_h, \eta_h)$.

2.2 Prior Elicitation

It follows from (9) that the noninformative Jeffrey's prior is

$$\pi(\beta, \eta) \propto [\det \mathrm{I}(\beta, \eta)]^{\frac{1}{2}} \propto (\beta\eta)^{-1}. \tag{10}$$

In the original (β, θ) parametrization this is equivalent to $\pi(\beta, \theta) \propto \theta^{-1}$ (see (11) below).

The improper reference priors $\pi(\beta, \theta) \propto (\theta\,\beta^\delta)^{-1}$ $(\delta < n)$, considered by Bar-Lev *et al.* (1992) and Sen (2002), which generalize the noninformative priors $\pi(\beta, \theta) \propto \theta^{-1}$ and $\pi(\beta, \theta) \propto (\theta\,\beta)^{-1}$ (Lingham and Sivaganesan, 1997; Guida *et al.*, 1989; Box and Tiao, 1973), are special cases of (5) when $a_\eta = b_\eta = b_\beta = 0$ and $a_\beta = -\delta$. To see this, note that

$$\begin{aligned}\pi(\beta, \eta) &= \pi(\beta, \theta)\big|_{\theta = T/\eta^{1/\beta}} \times |J| \\ &\propto (\theta\,\beta^\delta)^{-1}\Big|_{\theta = T/\eta^{1/\beta}} \times T\,\beta^{-1}\,\eta^{-1-1/\beta} \propto \beta^{-\delta-1}\,\eta^{-1}\,,\end{aligned} \tag{11}$$

where $J = -T^\beta\,\beta^{-1}\,\eta^{-1-1/\beta}$ is the Jacobian of the transformation $(\beta, \theta) \mapsto (\beta, \eta)$.

To finish this section we note that the elicitation of proper informative priors in the (β, η) parametrization may be facilitated in view that both β and η have clear operational

interpretations. In this sense, since

$$\frac{d\,\Lambda(t)/\Lambda(t)}{d\,t/t} = t\,\frac{\Lambda'(t)}{\Lambda(t)} = t\,\frac{\lambda(t)}{\Lambda(t)} = t\,\frac{(\beta/\theta)(t/\theta)^{\beta-1}}{(t/\theta)^{\beta}} = \beta\,,$$

β is the elasticity of the mean number of events $\Lambda(t)$ with respect to time, i.e. the relative change in Λ due to relative change in t. Indeed, the PLP is characterized by the fact that this elasticity is constant over time. On the other hand, $\eta = (T/\theta)^{\beta} = \Lambda(T) = EN(T)$ is the expected number of events during the period that the process has been observed.

3 Several overlapping realizations

The methods established in Section 2 can be easily extended to the case that K independent realizations of the same PLP, say $N_1(t), \ldots, N_K(t)$, are observed all up to the same time T. This follows from the well known fact that the superposition of NHPPs is also a NHPP whose intensity function is the sum of the individual intensities (Thompson, 1998). In other words, $N_+(t) = \sum_{i=1}^{K} N_i(t)$ has intensity $\lambda_+(t) = K\lambda(t) = K\,\beta\,t^{\beta-1}/\theta^{\beta}$ and hence is also a PLP with parameters $\beta_+ = \beta$ and $\theta_+ = \theta/K^{1/\beta}$. Therefore, one can use the ideas in Section 2 to draw inferences about β_+ and θ_+ and these are equivalent to inferences about the original parameters $\beta = \beta_+$ and $\theta = \theta_+\,K^{1/\beta}$. However, it is not clear how to proceed when the K realizations have been observed along different time intervals.

3.1 Overlapping realizations of a PLP

Suppose that $N_1(t), \ldots, N_K(t)$ are independent realizations of the same PLP observed respectively up to times $T_1, \ldots, T_K$. Let t_{ij} be the j-th event time for the i-th realization, $i = 1, \cdots, K$; $j = 1, \cdots, n_i$. According to equation (2), the likelihood in the original (β, θ) parametrization is

$$L(\beta,\theta) = \prod_{i=1}^{K}\left\{ e^{-(T_i/\theta)^{\beta}}\,\frac{\beta^{n_i}}{\theta^{n_i\beta}}\,\prod_{j=1}^{n_i} t_{ij}^{\beta-1}\right\} = \frac{\beta^{n}}{\theta^{n\beta}}\,[\prod_{i=1}^{K}\prod_{j=1}^{n_i} t_{ij}]^{\beta-1}\,\exp\{-\sum_{i=1}^{K}(T_i/\theta)^{\beta}\}\,, \qquad (12)$$

where $n = \sum_{i=1}^{K} n_i$ is the total number of events. If for some of the realizations no event has been observed, take the corresponding $n_i = 0$ and set in equation (12) empty sums and products equal to 0 and 1, respectively. Hence, the MLE satisfies that $\hat{\theta} = [\sum_{i=1}^{K} T_i^{\hat{\beta}}/n]^{1/\hat{\beta}}$ and

$$\frac{1}{n}\sum_{i=1}^{K}\sum_{j=1}^{n_i}\log t_{ij} = \frac{\sum_{i=1}^{K} T_i^{\hat{\beta}}\,\log T_i}{\sum_{i=1}^{K} T_i^{\hat{\beta}}} - \frac{1}{\hat{\beta}}, \qquad (13)$$

and must be obtained numerically (Rigdon and Basu, 2000).

If we reparametrize the problem in terms of β and $\eta = \sum_{i=1}^{K}(T_i/\theta)^{\beta}$, it follows after some

algebra that the likelihood (12) becomes

$$L(\beta,\eta) = c \times \eta^n e^{-\eta} \times \beta^n e^{-n\beta/\hat{\beta}} \times e^{nF(\beta)} \propto \gamma(\eta|\, n+1, 1)\, \gamma(\beta|\, n+1, n/\hat{\beta})\, e^{nF(\beta)} \qquad (14)$$

where now $c = \prod_{i=1}^K \prod_{j=1}^{n_i} t_{ij}^{-1}$, $\hat{\beta}$ satisfies (13) and

$$F(\beta) = \frac{\sum_{i=1}^K T_i^{\hat{\beta}} \log T_i}{\sum_{i=1}^K T_i^{\hat{\beta}}}\, \beta - \log \sum_{i=1}^K T_i^{\beta}\,.$$

Note that β and η are still orthogonal. The log-likelihood is $\ell(\beta,\eta) = \log c + n\log\eta - \eta + n\log\beta - n\beta/\hat{\beta} + nF(\beta)$. Therefore, the MLE are obtained solving $\partial\ell/\partial\beta = n/\beta - n/\hat{\beta} + nF'(\beta) = 0$ and $\partial\ell/\partial\eta = n/\eta - 1 = 0$, which gives $\hat{\eta} = n$ and, of course, $\hat{\beta}$ given by (13). In order to compute asymptotic variances note that $\frac{\partial^2\ell}{\partial\beta^2} = -n/\beta^2 + nF''(\beta)$, $\frac{\partial^2\ell}{\partial\eta^2} = -n/\eta^2$ and $\frac{\partial^2\ell}{\partial\beta\,\partial\eta} = 0$. Hence, the Fisher information matrix is $I(\beta,\eta) = n\ \mathrm{Diag}(\beta^{-2} - F''(\beta), \eta^{-2})$, where

$$F''(\beta) = -\sum_{i=1}^K \frac{T_i^{\beta}}{\sum_{h=1}^K T_h^{\beta}} [\log T_i]^2 + \left(\sum_{i=1}^K \frac{T_i^{\beta}}{\sum_{h=1}^K T_h^{\beta}}\ \log T_i\right)^2$$

is formally the same as minus the variance of a random variable taking values $\log T_i$ with probabilities proportional to T_i^{β} (i = 1 , . . . , K). The asymptotic covariance matrix of $(\hat{\beta},\hat{\eta})$ is then $[I(\beta,\eta)]^{-1} = n^{-1}\,\mathrm{Diag}([1/\beta^2 - F''(\beta)]^{-1}, \eta^2)$. Asymptotic variances for functions of the parameters can be obtained using the Delta Method.

3.2 Posterior analysis

Under the prior specification (5), the posterior density becomes

$$\pi(\beta,\eta|D) \propto \gamma(\eta|\, a_\eta + n, b_\eta + 1) \times \gamma(\beta|\, a_\beta + n, b_\beta + n/\hat{\beta}) \times e^{nF(\beta)}\,, \qquad (15)$$

where $D = \{t_{ij} : i = 1, \ldots, K; j = 1, \ldots, n_i\}$. It should be immediate from the comparison of (15) with (6) that the behavior of $F(\beta)$ is crucial to understand the difference between the one and the many realizations settings. Now, if $\hat{\beta}$ is the solution of (13), it follows that

$$
\begin{aligned}
&F(\hat{\beta}) - F(\beta) \\
&\quad = \frac{\sum_{i=1}^K T_i^{\hat{\beta}} \log T_i}{\sum_{i=1}^K T_i^{\hat{\beta}}}\,\hat{\beta} - \log \sum_{i=1}^K T_i^{\hat{\beta}} - \frac{\sum_{i=1}^K T_i^{\hat{\beta}} \log T_i}{\sum_{i=1}^K T_i^{\hat{\beta}}}\,\beta + \log \sum_{i=1}^K T_i^{\beta} \\
&\quad = \frac{\sum_{i=1}^K T_i^{\hat{\beta}} \log(T_i^{\hat{\beta}}/T_i^{\beta})}{\sum_{i=1}^K T_i^{\hat{\beta}}}\,\beta + \log \frac{\sum_{i=1}^K T_i^{\beta}}{\sum_{i=1}^K T_i^{\hat{\beta}}} \\
&\quad = \sum_{i=1}^K \frac{T_i^{\hat{\beta}}}{\sum_{h=1}^K T_h^{\hat{\beta}}} \log \frac{T_i^{\hat{\beta}}/\sum_{h=1}^K T_h^{\hat{\beta}}}{T_i^{\beta}/\sum_{h=1}^K T_h^{\beta}} \\
&\quad = KL[(\frac{T_1^{\hat{\beta}}}{\sum_{h=1}^K T_h^{\hat{\beta}}}, \cdots, \frac{T_K^{\hat{\beta}}}{\sum_{h=1}^K T_h^{\hat{\beta}}})||(\frac{T_1^{\beta}}{\sum_{h=1}^K T_h^{\beta}}, \cdots, \frac{T_K^{\beta}}{\sum_{h=1}^K T_h^{\beta}})] \geq 0\,,
\end{aligned}
$$

where $KL[(p_1, \ldots, p_K)||(q_1, \ldots, q_K)) = \sum_{i=1}^K p_i \log \frac{p_i}{q_i}$ is the Kullback-Leibler divergence. Hence, $F(\beta)$ attains a maximum when $\beta = \hat{\beta}$. Moreover, $F(\beta)$ is constant and equal to $F(\hat{\beta})$ if and only if $T_1 = T_2 = \cdots = T_K$.

In order to sample from the posterior distribution (15) we use the independence between η and β and obtain first $\eta_1, \ldots, \eta_m \overset{i.i.d.}{\sim} \text{Gamma}(a_\eta + n, b_\eta + 1)$. Simulation from the posterior distribution of β becomes easy by using, for instance, the rejection or importance sampling algorithms (see Gelman *et al.* (2003) or Devroye (1986)). For instance, the rejection algorithm produces an observation from the posterior of β by sampling repeatedly $\beta \sim \gamma(\beta|a_\beta + n, b_\beta + n/\hat{\beta})$ and $u \sim \text{Uniform(0,1)}$ until $u \leq \exp\{n[F(\beta) - F(\hat{\beta})]\}$. Repeating the rejection algorithm m times we obtain an i.i.d. sample $\beta_1, \ldots, \beta_m$. Once that an i.i.d. sample from the posterior $\pi(\beta, \eta|D)$ has been obtained we proceed essentially as in Section 2. In our practice the rejection sampling method has been quite efficient, in the sense that even for problems with few failures the rejection rate is below 10%.

The rejection algorithm can also be used when the prior for β is a Gamma distribution truncated to the right of $\beta = 1$. In this case, one just changes the proposal distribution above to be also a truncated Gamma. In other words, to obtain an observation from $\pi(\beta|D)$ one samples repeatedly $\beta \sim \gamma(\beta|a_\beta + n, b_\beta + n/\hat{\beta})$ and $u \sim \text{Uniform(0,1)}$ until both $\beta > 1$ and $u \leq \exp\{n[F(\beta) - F(\hat{\beta})]\}$. However, the rejection algorithm would need some major adaptation if one wants to consider a prior distribution for β which is restricted to have support in $(1, \infty)$ and which is not a truncated Gamma. For instance, we argue in Sections 4 and 5 that for the power transformers problem it may be better to consider a shifted Gamma prior, i.e. $\beta - 1 \sim \text{Gamma}(a_\beta, b_\beta)$. In this case one could use the Metropolis algorithm to obtain an approximate sample from the posterior of β. Briefly, we set a starting value β_0 (e.g. $\beta_0 = \hat{\beta}$) and proceed iteratively as follows. At step $(i+1)$ we generate $z \sim \text{Normal(0,1)}$ and $u \sim \text{Uniform(0,1)}$ and let $\beta_{cand} = \beta_i + Z$. Now if $u < \min\{\pi(\beta_{cand}|D)/\pi(\beta_i|D), 1\}$ we let $\beta_{i+1} = \beta_{cand}$, otherwise let $\beta_{i+1} = \beta_i$. In general the Metropolis algorithm produce correlated observations which may be unduly influenced by the starting value. If one wants to avoid

this, the algorithm can be run for $M = B + ml$ cycles, the first B observations discarded (the "burn-in") and then every other l of the remaining simulated values can be kept to end up with a size m approximate i.i.d. sample.

4 Monte Carlo Simulation

In this section we describe some Monte Carlo simulations in order to compare Bayes estimates under different prior specifications in the case of overlapping realizations of a PLP. The Bayes estimates were also compared to the ones obtained by maximum likelihood. As described in Section 1, the optimal preventive maintenance policy that minimizes expected cost per unit of time is the value τ defined by (8). τ was the quantity of interest in the simulations. The prior (and hence also the posterior) distribution must satisfy $Pr(\beta > 1) = 1$ since the intensity function of failures must be increasing as discussed in Section 1. This information has been incorporated in all of the prior specifications in simulations.

Different prior distributions for β and η were used in the simulations. The following notations and definitions were used in the simulation runs:

- MLE - Maximum likelihood estimate;
- BayesE_1 - Bayes estimator by considering a reference prior distribution (11) for $\delta = 1$ truncated at $\beta = 1$;
- BayesE_2 - Bayes estimator by considering Jeffrey's prior distribution (10) truncated at $\beta = 1$;
- BayesE_3 - Bayes estimator by considering gamma prior distributions truncated at $\beta = 1$. That is $\pi(\beta, \eta) \propto \gamma(\beta|\, a_\beta, b_\beta) \times \gamma(\eta|\, a_\eta, b_\eta)\, A_{(1,\infty)}(\beta)$, where $A_{(1,\infty)}(\beta) = 1$, if $\beta \in (1, \infty)$ and 0, otherwise.
- BayesE_4 - Bayes estimator by considering a gamma prior distribution shifted to 1 for β and gamma for η. That means, $\beta - 1 \sim \gamma(a_\beta, b_\beta)$.;
- CP - Interval Coverage Percentage.

Prior hyperparameters for BayesE_3 and BayesE_4 were set to 10^{-4} except for a_β in the latter that must be shifted by 1 and was set to $1 + 10^{-4}$.

In the likelihood approach, asymptotic confidence intervals for τ were obtained by using the delta method (Gilardoni and Colosimo, 2007). In the Bayesian approach, we used the highest posterior density (HPD) intervals. The Bayesian estimates were the posterior distribution mode. That is, the value that maximize the posterior distribution of τ.

Throughout the Monte Carlo study we consider $\beta = 2$, $\theta = 24$ and $C_{MR}/C_{PM} = 16$, so that it follows from (8) that $\tau = 6$. The number of systems K and truncation times T_i's were set to study three different situations. The first two achieve a large number of failures by considering respectively many systems and large truncation times. More precisely, we

have in situation 1 $K = 500$ systems all truncated at $T = 100$, so that the expected number of failures per system equals 17.361. Situation 2 considers $K = 50$ systems all truncated at $T = 320$, resulting in 178 expected failures per system. Finally, the third situation has $K = 50$ systems truncated at T =30 and hence only 1.5625 expected failures per system, so that this situation is probably closer to the real example considered in the next section.

The results of the Monte Carlo simulations based on 3000 replicas are shown in Table 1. In the first two situations there were no significant differences among methods and prior distributions. Probably, sample sizes were large enough to overcome differences in the prior specifications. In the third situation, Bayes estimates have similar results. BayesE_4 has the worst interval covarage. In general, all estimates have a small bias, the MLE being the least biased.

Table 1: Summary of simulation results

		MLE	BayesE_1	BayesE_2	BayesE_3	BayesE_4
Situation 1	Mean of $\hat{\tau}$	6.00	6.00	6.00	6.00	6.02
	CP	94.9	94.3	94.4	94.4	95.0
	Mean length	0.475	0.475	0.474	0.474	0.482
Situation 2	Mean of $\hat{\tau}$	6.00	6.00	6.00	6.00	6.03
	CP	94.7	94.5	94.6	94.6	95.0
	Mean length	0.753	0.753	0.752	0.753	0.765
Situation 3	Mean of $\hat{\tau}$	6.11	6.17	6.12	6.13	6.16
	CP	95.4	95.2	95.9	95.1	93.1
	Mean length	2.125	2.149	2.125	2.124	1.971

5 Example: Maintenance of electrical power transformers

The data in Figure 1(a) shows the failure history of 40 electrical power transformers (Gilardoni and Colosimo, 2007). The usual nonparametric estimate of Λ (Meeker and Escobar, 1998) is shown in Figure 1(b). The convex form of this plot provides some evidence that the intensity function is increasing.

The same prior distributions used in Section 4 were considered here. According to the electrical power company, the ratio between minimal repair and preventive maintenance costs is $C_{MR}/C_{PM} = 15$. Table 4 presents the results. The interval based on the ML estimates is the shortest one. Point estimates are in agreement among Bayesian methods taking a value around 6400 hours, although the ML estimate is 6290 hours. Among the Bayesian intervals, those considering the Jeffrey's and the translated gamma prior are shorter. Posterior density function of τ appears to be slightly skewed to the right (see Figure 2 for the Jeffrey's prior case).

In addition to point and interval estimates for the optimal maintenance time τ, it is useful to gain an idea of the size of the difference between estimated and true expected costs

Figure 1: (a) Power transformer data. Each horizontal line represents a transformer and dots are observed failure times; (b) Mean Cumulative Failure (MCF) type estimate for Λ.

Table 2: τ estimates for the power transformers data (in 1000 hours).

	MLE	BayesE_1	BayesE_2	BayesE_3	BayesE_4
Estimate	6.29	6.44	6.41	6.41	6.56
Interval	[4.87;7.70]	[5.06;8.48]	[5.00;8.72]	[5.02;8.74]	[5.04;8.44]
Length	2.83	3.42	3.72	3.72	3.40

(Gilardoni and Colosimo, 2007). This difference can be obtained from (7) as

$$\begin{aligned} H(\hat{\tau}) - H(\tau) &= \frac{1}{\hat{\tau}}\left[C_{PM} + C_{MR}\left(\frac{\hat{\tau}}{\theta}\right)^{\beta}\right] - \frac{1}{\tau}\left[C_{PM} + C_{MR}\left(\frac{\tau}{\theta}\right)^{\beta}\right] \\ &= C_{PM}\left[\frac{1}{\hat{\tau}}\left(1 + \frac{C_{MR}}{C_{PM}}\frac{\hat{\tau}^{\beta}}{\sum_{i=1}^{K} T_i^{\beta}}\eta\right) - \frac{1}{\tau}\left(1 + \frac{C_{MR}}{C_{PM}}\frac{\tau^{\beta}}{\sum_{i=1}^{K} T_i^{\beta}}\eta\right)\right]. \end{aligned}$$

$H(\hat{\tau}) - H(\tau)$ measures the difference in the cost attained for the present state of information and that which could be attained if one had perfect information (i.e. if sampling could be continued forever). Observe that because of the definition of the optimal τ one must have that $H(\hat{\tau}) - H(\tau) \geq 0$. Hence, we usually compute a credible upper limit for the difference.

To compute a rao-blackwellized (Gelfand and Smith , 1990) approximation to the posterior density of the optimal τ, note that

$$\pi(\tau|D) = \int \pi(\tau|\beta, D)\pi(\beta|D)\, d\beta \approx \frac{1}{m}\sum_{j=1}^{m} \pi(\tau|\beta_j, D)\,,$$

Figure 2: Approximate (a) posterior density of τ, (b) cost per unit of time and (c) posterior density of the difference between costs, all for the Jeffrey's prior.

where we use that $\psi = \frac{C_{MR}}{C_{PM}} \frac{(\beta-1)\tau^\beta}{\sum_{i=1}^K T_i^\beta}$ to obtain that

$$\pi(\tau|\beta, D) = \pi(\psi|\beta, D) \left|\frac{\partial\psi}{\partial\tau}\right| = \gamma\left(\frac{C_{MR}}{C_{PM}} \frac{(\beta-1)\tau^\beta}{\sum_{i=1}^K T_i^\beta} \middle| a_\psi + n, b_\psi + 1\right) \frac{C_{MR}}{C_{PM}} \frac{\beta(\beta-1)\tau^{\beta-1}}{\sum_{i=1}^K T_i^\beta}.$$

Figure 2(c) presents the posterior density of the difference between costs.

6 Final Remarks

In this paper a conjugate prior distribution was derived for the PLP model in the one system case. The proposed conjugate prior is a product of gamma distributions for the parameters of the PLP in an alternative parametrization. The results are extended for overlapping realizations of the same PLP. Although in the many realization case the product of gamma prior is no longer conjugate, it was showed that posterior sampling is easy to implement.

Monte Carlo simulations are used in order to compare some proposed prior distributions in the context of a real application. Three different situations and four prior distributions are considered in the simulations. It can be observed no significant differences among prior distributions in the considered scenarios. They are also very close to the MLE results. Some Monte Carlo simulations were carried out for small sample sizes. They are not shown in the paper. It was difficult to summarize the results since, by chance, a small number of samples were obtained such that MLE of β was smaller than one. Under this condition, τ is not defined since there is no optimal time when the intensity function is decreasing. This situation is easily handled in the Bayesian approach by making prior distributions truncated for $\beta > 1$.

In the real case analysis in Section 5, point estimates are similar among the methods.

Bayesian intervals considering Jeffreys and gamma prior distributions are shorter than the one using prior of reference. Maximum likelihood confidence interval is the shortest one.

We considered just the time truncation situation in this paper. That is T is fixed by design. This is basically the way that most of the practical situations collect data from repairable systems. Another possible situation is the case in which data collection is ceased after a fixed number of failures. This sampling scheme is said to be failure truncated. Since the likelihood is essentially the same as for the time truncated case, the Bayesian analysis takes the same form irrespective of the experimental design. However, a cautionary note regarding the transformation $\eta = (T/\theta)^\beta$ or $\eta = \sum_{i=1}^{K}(T_i/\theta)^\beta$ is in order when one or more of the realizations are failure truncated. In this case, one or more of the T_i's are random and η would depend on data. Hence, the prior density (5) depends indirectly on the observed data, which is not allowed from a strict Bayesian viewpoint. In our opinion this fact has little, if any, practical importance. Moreover, we stress that the problem appears only in the case of failure truncation and even then, disappears if one uses a non-informative prior $\pi(\eta) \propto \eta^{-1}$ (i.e. $a_\eta = b_\eta = 0$), because such a prior would be equivalent to $\pi(\theta) \propto \theta^{-1}$ (see equation (11)), which of course does not depend on data.

Acknowledgements

We thank the editor and a referee for suggestions that led to improvements in our paper. The partial support of CNPq, CAPES/Brazil and FAPEMIG is also gratefully acknowledged.

References

H. Ascher and H. W. Feingold, *Repairable Systems Reliability: Modeling, Inference, Misconception and their Causes*, Marcel Dekker: New York, 1984.

L. J. Bain and M. Engelhardt, *Statistical Analysis of Reliability and Life-Testing Models, Theory and Methods* Marcel Dekker: New York, 1991.

R. D. Baker, Some new tests of the power law process. *Technometrics* 38, pp. 256-265, 1996.

S. K. Bar-Lev, I. Lavi, and B. Reiser, Bayesian Inference for the Power Law Process. *Annals Institute Statistical Mathematics* 44, pp. 623-639, 1992.

R. E. Barlow and L. C. Hunter, Optimum Preventive Maintenance Policies. *Operations Research* 8, pp. 90-100, 1960.

M. Berman and T. R. Turner, Approximating Point Process Likelihoods with GLIM. *Applied Statistics* 41, pp. 31-38, 1992.

M. Bhattacharjee, J. V. Deshpande and U. V. Naik-Nimbalkar, Unconditional Tests of Goodness of Fit for the Intensity of Time-Truncated Nonhomogeneous Poisson Processes. *Technometrics* 46, pp. 330-338, 2004.

G. E. P. Box and G. C. Tiao, *Bayesian Inference in Statistical Analysis*, Addison-Wesley: Reading, 1973.

L. R. Crow, Reliability Analysis for Complex Systems. *Reliability and Biometry*, F. Proschan and J. Serfling (Eds), pp. 379-410, 1974.

L. Devroye, Non-Uniform Random Variate Generation. New York, Springer-Verlag, 1986.

J. T. Duane, Learning Curve approach to Reliability Monitoring. *IEEE Transactions on Aerospace* 2, pp. 563-566, 1964.

A. E. Gelfand and A. F. M. Smith, Sampling-based Approaches to Calculating Marginal Densities. *Journal of the American Stasticial Association* 85, pp. 398-409, 1990.

A. Gelman, J. B. Carlin, H. S. Stern and D. B. Rubin *Bayesian Data Analysis*. Chapman and Hall: New York, 2003.

I. B. Gerstack, *Models of Preventive Maintenance*. North Holand: Amsterdam, 1977.

G. L. Gilardoni and E. A. Colosimo, Optimal maintenance time for repairable systems. *Journal of Quality Technology* 39, pp. 48-53, 2007.

M. Guida, R. Calabria, and G. Pulcini, Bayes Inference for a Non-Homogeneous Poisson Process with Power Intensity Law. *IEEE Transactions on Reliability* 38, pp. 603-609, 1989.

Y. S. Huang, A Decision Model for Deteriorating Repairable Systems. *IIE Transactions* 33, pp. 479-485, 2001.

Y. S. Huang and V. B. Bier, A Natural Conjugate Prior for the Non-Homogeneous Poisson Process with a Power Law Intensity Function. *Communications in Statistics - Simulation and Computation* 27, pp. 525-551, 1998.

T. S. Kim, C. S. Park and S. E. Ahn, Determinig the Optimal Maintenance Action for a Deteriorating Repairable System. *Probabilistic Engineering Mechanics* 23, pp. 95-101, 2008.

B. Klefsjö and U. Kumar, Goodness of Fit Tests for the Power-Law Process Based on the TTT-Plot. *IEEE Transactions on Reliability* 41, pp. 593-598, 1992.

R. T. Lingham and S. Sivaganesan, Testing Hypotheses About the Power Law Process Under Failure Truncation Using Intrinsic Bayes Factors. *Annals of the Institute of Statistical Mathematics*, 49, pp. 693-710.

W. Q. Meeker and L. A. Escobar, *Statistical Methods for Reliability Data*. John Wiley: New York, 1998.

H. Raiffa and R. Schlaifer, *Applied Statistical Decision Theory*. Harvard University Press: Cambridge, Massachussets, 1961.

S. E. Rigdon and A. P. Basu, *Statistical Methods for the Reliability of Repairable Systems*. John Wiley: New York, 2000.

A. Sen, Bayesian Estimation and Prediction of the Intensity of the Power Law Process. *Journal of Statistical Computation and Simulation* 72, pp. 613-631, 2002.

A. Sen and R. Khattree, On estimating the current intensity of failure for the power-law process. *Journal of Statistical Planning and Inference*, 74, pp. 253-272, 1998.

W. A. Thompson, *Point Process Models with Applications to Safety and Reliability*. Chapman and Hall: London, 1988.

J. W. Yu, G. L. Tian and M. L. Tang, Predictive Analyses for Non-Homogeneuous Poisson Process with Power Law using Bayesian Approach. *Computational Statistics and Data Analysis* 51, 4254-4268, 2007.

M. Zhao and M. Xie, On Maximum Likelihood Estimation for a General Non-homogeneous Poisson Process. *Scandinavian Journal of Statistics* 23, 597-607, 1996.

Heterogeneity Identification of Repairable Systems

Abstract

A repairable system, under minimal repair, is usually modeled according to a Non-Homogeneous Poissson Process (NHPP) assuming a Power Law intensity function. A traditional approach considers iid NHPPs in order to conduct a statistical analysis based on a sample of systems. However, systems might be heterogeneous due to unmeasured variables such as age, suppliers and so on. In order to verify this assumption a classical and a Bayesian approaches are proposed in this paper. Some possible model scenarios considering different systems heterogeneity are compared using likelihood ratio tests and hierarchical Bayesian model. Information criteria are used in order to select the most likely model for a particular data set. Real data sets illustrate the proposed methodology.

1 Introduction

Statistical models for recurrent events are of great interest in reliability and maintenance of repairable systems (Ascher and Feingold (1984), Bain and Engelhardt (1991), Rigdon and Basu (2000)). Repairable system is allowed to experience more than one failure throughout its lifetime. The system is restored and put back in operation whenever a failure occurs, without necessarily being replaced by a new one. After each failure, a maintenance activity is necessary to put the system back into operation. This activity can usually be of two types: a) Minimal Repair (MR), popularly known by *as bad as old*, identified for only replace the damaged part of the equipment, keeping it in the situation it was immediately before the failure; b) Perfect Maintenance or major overhaul (PM), popularly known by *as good as new*, identified for renew the entire system. The combination of PM and MR has been of interest since the work of Barlow and Hunter (1960). Some of the developments for this situation can be found in Gerstack (1977), Block et al (1990) and Park et al (2000). However, statistical inference procedures for maintenance stochastic models developed in the literature are still under consideration.

A repairable system, under MR, is usually modeled according to a Non-Homogeneous Poissson Process (NHPP), $\{N(t): t \geq 0\}$ where $N(t)$ counts the system number of failures from the beginning of the follow-up until time t (Barlow and Hunter, 1960). $N(t)$, for fixed t, is modeled by a Poisson distribution with mean function $\Lambda(t) = E(N(t)) = \int_0^t \lambda(u)\, du$, where $\lambda(t)$, the intensity function of the process, is defined as:

$$\lambda(t) = \lim_{\Delta t \to 0} \frac{P(N(t, t + \Delta t] = 1)}{\Delta t}.$$

This paper is motivated by a data set concerned to failure histories of eleven electrical power transformers. Figure 1 presents failure records of these transformers between 1999 and 2009. A traditional approach would consider iid NHPPs in order to conduct a statistical analysis based on this sample of systems. However, systems might be heterogeneous due to unmeasured variables such as age, suppliers and so on. Indeed, it seems the case, according to the maintenance engineers, for the power transformers described in Figure 1. In order to verify this assumption a classical and a Bayesian approaches are proposed in this paper.

The start point to answer questions related to repairable systems behavior consists in modeling $\lambda(t)$. The most popular parametric model for $\lambda(t)$ is the Power Law Process (PLP), whose analytical form is given by (Crow, 1974)

$$\lambda(t) = \frac{\beta}{\theta} \left(\frac{t}{\theta}\right)^{\beta-1}, \beta,\ \theta\ > 0, \tag{1}$$

where λ is increasing if $\beta > 1$ and it is decreasing if $\beta < 1$.

The goal of this work is to identify the best model to fit data from several systems under

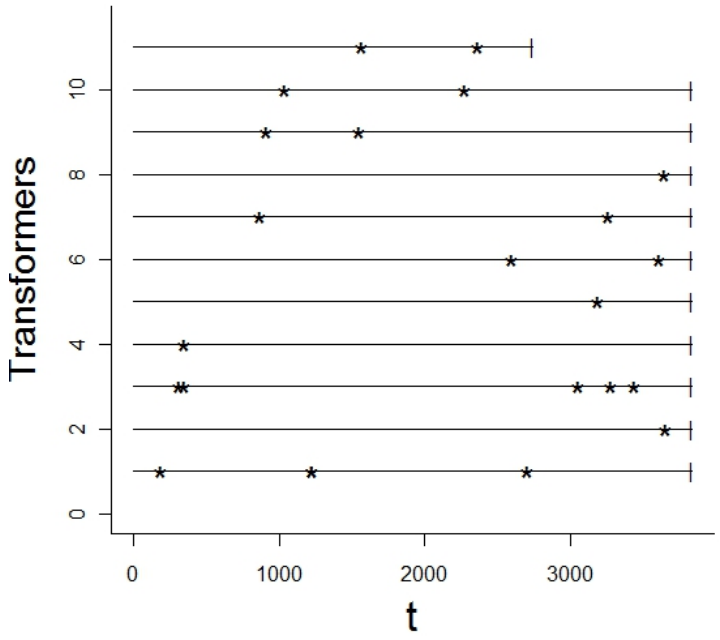

Figure 1: Failure Histories of Electrical Power Transformers
Each horizontal line represents a transformer and '*' are observed failure times.

MR, assuming a Power Law intensity function for each one. It is natural to model such a problem hierarchically, with observable outcomes modeled conditionally on certain parameters, which themselves are given a probabilistic specification in terms of further parameters (Gelman *et al.*, 2003). A strong tool for adjusting multiples systems is the Hierarchical Bayesian approach. According to Gelman *et al.* (2003), hierarchical modeling techniques can improve each individual estimate of $N(t)$. In the other hand, frequentist approach can be also considered in this work for the same purpose.

The paper is organized as follows. In Section 2, we present six possible scenarios considering different systems heterogeneities. Likelihood functions and ratio tests, associated with each one of the scenarios, and their respective maximum likelihood estimators (MLE) are presented Section 3. In Section 4, we model PLPs hierarchically. Information criteria are used in order to select the most likely model for a particular data set. In Section 5 two real data sets illustrate the proposed methodology.

2 Scenarios of interest

Some possible model scenarios considering different systems heterogeneities are presented in this section. These scenarios are described next considering a sample of K systems.

1. All systems are different. They are essentially K separate systems. Therefore, it is

necessary to estimate $2K$ parameters.

2. λ functions have different β's but they have the same θ. In this case, it is necessary to estimate $K+1$ parameters.

3. λ functions have different θ but they have the same β. Again, it is necessary to estimate $K+1$ parameters.

4. All PLPs are identical. It means, that there is only one θ and β to be estimated.

5. Each $N_i(\cdot)$ represents a different PPH ($\beta = 1$). Therefore, it is necessary to estimate K θ's scale parameters.

6. All systems are identical, following a PPH. It remains just one scale parameter to be estimated in the modeling structure.

The relationship among the scenarios are summarized in Table 1.

Table 1: Comparative restrictions among the scenarios of interest for K systems.

	Number of Parameters	Restriction on β	Restriction on θ
Scenario 1	2K	$\beta_i \neq \beta_j\,, \quad \forall i \neq j$	$\theta_i \neq \theta_j\,, \quad \forall i \neq j$
Scenario 2	K+1	$\beta_i \neq \beta_j\,, \quad \forall i \neq j$	$\theta_i = \theta\ \ , \quad \forall i$
Scenario 3	1+K	$\beta_i = \beta\ \ , \quad \forall i$	$\theta_i \neq \theta_j\,, \quad \forall i \neq j$
Scenario 4	2	$\beta_i = \beta\,, \quad \forall i$	$\theta_i = \theta\ \ , \quad \forall i$
Scenario 5	K	$\beta_i = 1\ \ , \quad \forall i$	$\theta_i \neq \theta_j\,, \quad \forall i \neq j$
Scenario 6	1	$\beta_i = 1\ \ , \quad \forall i$	$\theta_i = \theta\ \ , \quad \forall i$

Likelihood functions related to each scenario are presented in the next section.

3 Likelihood Functions and Ratio Tests

In this section, likelihood functions are established for each model associated with the corresponding scenario described in Section 2. In reality, only the likelihood function for scenario 1 is derived next. The others ones are easily obtained from it by algebraic manipulation since they are special cases of scenario 1. Analytic expressions of the Maximum Likelihood Estimators (MLE) for each one of the scenarios are also presented in this section. It is considered for each scenario that there are K processes with time truncations, respectively, at $T_1, T_2, \ldots, T_K$.

It was observed n_i failure times for the i-th system, indexed by failures at times $t_{ij}\,, i = 1, \ldots, K\,, j = 1, \ldots, n_i$. The likelihood function is given by: (Rigdon and Basu, 2000):

$$L(\lambda) = \prod_{i=1}^{K}\prod_{j=1}^{n_i}\lambda(t_{ij}) \times \exp\left[-\sum_{i=1}^{K}\int_0^{T_i}\lambda(x)dx\right]. \tag{2}$$

Likelihood function for the model associated with Scenario 1, is one that allows for maximum heterogeneity among systems. It can be obtained by plugging the PLP intensity function in (2). That is

$$L(N_1, N_2, \ldots, N_K|\beta_1, \ldots, \beta_K, \theta_1, \ldots \theta_K) = \prod_{i=1}^{K} \frac{\beta_i^{n_i}}{\theta_i^{n_i\beta_i}} \prod_{i=1}^{K}\prod_{j=1}^{n_i} t_{ij}^{\beta_i - 1} \exp\left[-\sum_{i=1}^{K}\left(\frac{T_i}{\theta_i}\right)^{\beta_i}\right], \quad (3)$$

and the corresponding log-likelihood function is

$$\ell = \sum_{i=1}^{K} n_i \log \beta_i - \sum_{i=1}^{K} n_i\beta_i \log \theta_i + \sum_{i=1}^{K}(\beta_i - 1)\sum_{j=1}^{n_i} \log t_{ij} - \sum_{i=1}^{K}\left(\frac{T_i}{\theta_i}\right)^{\beta_i}. \quad (4)$$

MLEs are obtained, by solving the system of equations obtained by the partial derivatives of (4) with respect to each of the parameters. That is, for the scenario 1, is given by

$$\hat{\beta}_i = \frac{n_i}{\sum_{j=1}^{n_i} \log \frac{T_i}{t_{ij}}} \qquad \text{and} \qquad \hat{\theta}_i = \frac{T_i}{n_i^{1/\hat{\beta}_i}}, \quad i = 1, \ldots, K.$$

Scenarios restrictions presented in Table 1 can be used in (3) in a way to obtain the likelihood functions for models associated to the remaining scenarios. Table 2 presents the analytical form or the equations in order to get MLEs of the parameters under each of the scenarios. In order to simplify the notation, the total failures number $\sum_{i=1}^{K} n_i$, has been replaced by N. Each model in Table 2 refers to each respective scenario.

Likelihood ratio test (LRT) is used in order to compare the scenarios of interest. This test can be used because the scenarios are nested within each other. A model is said to be nested within another one if it represents a special case of it. For example, models from 2 to 6 above are special cases of Model 1. LRT, under the restricted model, has large sample chi-square distribution with the degrees of freedom equal to the restricted number of parameters. It can observed that Scenarios 1 and 2 require at least one failure per system for the existence of MLEs.

4 Hierarchical Modeling for PLPs

Inference for hierarchical modeling is fundamentally Bayesian, in terms that population unknown quantities have a probabilistic specification as hyperparameters. The aim is to find a general model, sufficiently flexible to comport several scenarios, but very simple to data analysis and result interpretation. Nonhierarchical models are usually inappropriate for hierarchical data: with few parameters, they generally cannot fit large data sets accurately, whereas with many parameters, they tend to 'overfit' such data in the sense of produce

Table 2: MLE or Normal Equations for the six models of interest.

	Nr of β's	Nr of θ's	$\hat{\beta}_i$	$\hat{\theta}_i$
Model 1	K	K	$n_i \,/\, \sum_{j=1}^{n_i} \log \frac{T_i}{t_{ij}}$	$T_i \,/ n_i^{1/\hat{\beta}_i}$
Model 2	K	1	$\frac{n_i}{\hat{\beta}_i} + \sum_{j=1}^{n_i} \log t_{ij} = n_i \log \hat{\theta} + \left(\frac{T_i}{\hat{\theta}}\right)^{\hat{\beta}_i} \log \frac{T_i}{\hat{\theta}}$ e $\sum_{i=1}^{K} n_i \hat{\beta}_i = \sum_{i=1}^{K} \hat{\beta}_i \left(\frac{T_i}{\hat{\theta}}\right)^{\hat{\beta}_i}$	
Model 3	1	K	$N \,/ \sum_{i=1}^{K} \sum_{j=1}^{n_i} \log \frac{T_i}{t_{ij}}$	$T_i \,/\, n_i^{1/\hat{\beta}}$
Model 4	1	1	$N \,/ \hat{\beta} + \sum_{i=1}^{K} \sum_{j=1}^{n_i} \log t_{ij} = \frac{N \sum_{i=1}^{K} T_i^{\hat{\beta}} \log T_i}{\sum_{i=1}^{K} T_i^{\hat{\beta}}}$	$\left(\sum_{i=1}^{K} T_i^{\hat{\beta}} \,/\, N\right)^{1/\hat{\beta}}$
Model 5	0	K	1	$T_i \,/\, n_i$
Model 6	0	1	1	$\sum_{i=1}^{K} T_i \,/\, N$

models that fit the existing data well but lead to inferior predictions for new data (Gelman *et al.*, 2003). Bayesian hierarchical modeling might work with more parameters than data. This fact is a nice advantage of this approach. Hierarchical models with three levels are presented in Section 4.1

4.1 The Fully Bayesian Treatment of Hierarchical Modeling

A hierarchical model structure with three levels based on Scenario 1 is presented in Figure 2. In this figure, $\eta_i = (\theta_i, \beta_i)$; $i = 1, \ldots, K$, is treated as a sample from a common population distribution, indexed by ϕ.

In the hierarchical modeling, not only $\eta = (\eta_1, \ldots, \eta_K)$ is unknown, but so is ϕ, with probability distribution $\pi(\phi)$. The goal is get a joint probability distribution for the vector (η, ϕ). Thus, prior distribution for the unknown quantities is given by:

$$\pi(\eta, \phi) = \pi(\phi)\pi(\eta|\phi), \tag{5}$$

and the joint posterior distribution is:

$$\begin{aligned} p(\eta, \phi | N_1(t), \ldots, N_K(t)) &\propto \pi(\eta, \phi) L(N_1(t), \ldots, N_K(t) | \eta, \phi) && (6) \\ &= \pi(\eta, \phi) L(N_1(t), \ldots, N_K(t) | \eta), && (7) \end{aligned}$$

where $L(N_1(t), N_2(t), \ldots, N_K(t)|\eta)$ is the likelihood function defined in Section 3. Classical approach developed in Section 2 takes η as unknown and fixed. In the usual Bayesian

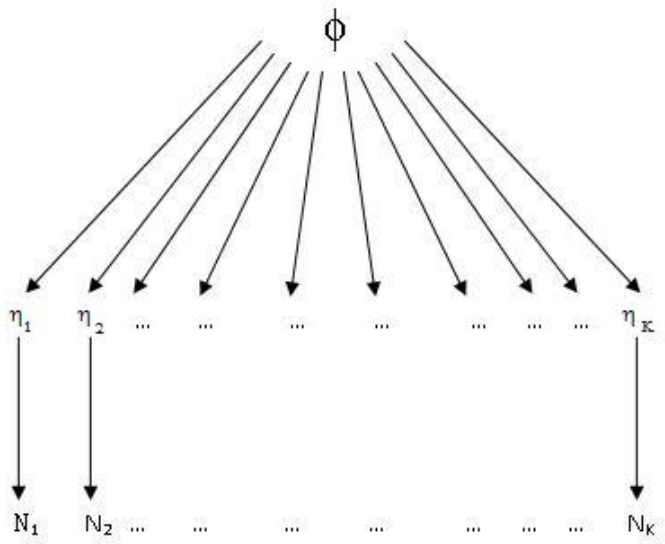

Figure 2: Structure of a hierarchical model with three levels

approach, η is random and ϕ is treated as known. Now, in hierarchical Bayesian approach, uncertainty in the model is included also in ϕ.

4.2 Prior and Hyperprior Elicitation

In order to specify a joint probability model for $\eta = (\eta_1, \cdots, \eta_K)$, it is necessary to use the crucial idea of exchangeability. It represents probabilistically the symmetry among the parameters under their joint prior distribution. The pairs of parameters $(\beta_1, \theta_1), (\beta_2, \theta_2), \ldots,$ (β_K, θ_K) are exchangeable in their joint distribution if $\pi((\beta_1, \theta_1), (\beta_2, \theta_2), \ldots, (\beta_K, \theta_K))$ is invariant to permutations of the indexes $(1, ..., K)$. See Gelman *et al.* (2003) for details about exchangeability and setting up hierarchical models.

A useful reparametrization $\xi_T = (T/\theta)^\beta$ was considered by Oliveira, Colosimo and Gilardoni (2010) and Huang and Bier (1998). ξ_T has a nice interpretation as the expected number of events during the period that the process has been observed T and it is orthogonal to β. However, this parametrization does not satisfy the exchangeability principle since each parameter depends on T_i and they are different for different systems. In order to overcome this drawback, consider $\xi_i = (1/\theta_i)^{\beta_i}$ as an alternative parametrization to get an exchangeable model. That is, let's take $T = 1$. ξ_i still has an interpretation as the expected number of events in one unit of time. Let's consider gamma prior distributions for ξ_i and β_i.

Let $\boldsymbol{\phi} = (a_\beta, a_\xi, b_\beta, b_\xi)$ be the hyperparameters vector of the third level. As established above, for $i = 1, \ldots, K$,

$$\pi(\xi_i|\boldsymbol{\phi}) = \frac{b_\xi^{a_\xi} \xi_i^{a_\xi - 1} e^{-b_\xi \xi_i}}{\Gamma(a_\xi)} \qquad \pi(\beta_i|\boldsymbol{\phi}) = \frac{b_\beta^{a_\beta} \beta_i^{a_\beta - 1} e^{-b_\beta \beta_i}}{\Gamma(a_\beta)}, \tag{8}$$

where $a_\beta, a_\xi, b_\beta, b_\xi > 0$.

Using the jacobian for transformation of variables, it follows that

$$\pi(\theta_i|\beta_i, \boldsymbol{\phi}) = \frac{\beta_i b_\xi^{a_\xi} \exp(-b_\xi/\theta_i^{\beta_i})}{\Gamma(a_\xi)\theta_i^{1+\beta_i a_\xi}}, \tag{9}$$

and, therefore, $\pi(\boldsymbol{\beta}, \theta|\boldsymbol{\phi}) = \pi(\theta|\beta, \boldsymbol{\phi})\pi(\boldsymbol{\beta}|\boldsymbol{\phi})$.

In the third level, consider also a Gamma hyperprior distribution for each of the four components of ϕ

$$\begin{aligned} a_\beta &\sim \gamma(a_\beta|a_{a_\beta}, b_{a_\beta}); \\ a_\xi &\sim \gamma(a_\xi|a_{a_\xi}, b_{a_\xi}); \\ b_\beta &\sim \gamma(b_\beta|a_{b_\beta}, b_{b_\beta}); \\ b_\xi &\sim \gamma(b_\xi|a_{b_\xi}, b_{b_\xi}), \end{aligned} \tag{10}$$

where $\gamma(x \mid , a\ ,\ b) = b^a x^{a-1} e^{-bx}/\Gamma(a)$ $(x, a, b > 0)$ is the density of the Gamma distribution with shape and scale parameters equal to a and b, respectively.

It follows then that the prior distributions are specified by (8) and (9) for the second level and (10) for the third level. This developments are the ones for the most complex scenario. That is, for scenario 1. In order words, in order to obtain prior specifications for the other scenarios is just a matter of exclude some pieces of prior specifications. The set of restrictions for each scenario is presented in Table 3.

4.3 Posterior Inference and Computation

Joint posterior distribution can be expressed by:

$$\pi(\phi, \beta, \theta|N_1, \ldots, N_K) \propto \pi(\boldsymbol{\phi})\pi(\beta, \theta|\boldsymbol{\phi})L(N_1, \ldots, N_K|\beta_1, \ldots, \beta_K, \theta_1, \ldots \theta_K). \tag{11}$$

Under prior and hyperprior specifications (8), (9) and (10), respectively, conditional posterior densities for Model 1 are shown next. Conditional posterior densities functions for other models can be easily found by considering restrictions in Tables 1 and 3.

These conditional posterior densities might be useful for posterior computations.

$$
\begin{aligned}
&\beta_i|\theta_i, N_i \sim \gamma\left(\beta_i|a_\beta + n_i + 1\,,\ b_\beta + \log\theta_i^{(n_i+a_\xi)} - \sum_{j=1}^{n_i}\log t_{ij}\right)\exp\left[-\tfrac{T_i^{\beta_i}+b_\xi}{\theta_i^{\beta_i}}\right],\\
&\pi(\theta_i|\beta_i, N_i) \propto \tfrac{1}{\theta_i^{1+\beta_i(n_i+a_\xi)}}\exp\left[-\tfrac{1}{\theta_i^{\beta_i}}T_i^{\beta_i+b_\xi}\right],\\
&\pi(a_\beta|\text{ remainder }) \propto \tfrac{a_\beta^{a_{a_\beta}-1}}{[\Gamma(a_\beta)]^K}b_\beta^{Ka_\beta}\prod_{i=1}^{K}\beta_i^{a_\beta-1}\exp\left[-b_{a_\beta}a_\beta\right],\\
&b_\beta|\text{ remainder }\ \sim \gamma\left(b_\beta \mid Ka_\beta + a_{b_\beta}\ ,\ b_{b_\beta} + \textstyle\sum_{i=1}^{K}\beta_i\right),\\
&\pi(a_\xi|\text{ remainder }) \propto \tfrac{a_\xi^{a_{a_\xi}-1}}{[\Gamma(a_\xi)]^K}b_\xi^{Ka_\xi}\textstyle\prod_{i=1}^{K}\theta_i^{-\beta_i a_\xi-1}\exp\left[-b_{a_\xi}a_\xi\right],\\
&b_\xi|\text{ remainder }\ \sim \gamma\left(b_\xi \mid Ka_\xi + a_{b_\xi}\ ,\ b_{b_\xi} + \textstyle\sum_{i=1}^{K}\frac{1}{\theta_i^{\beta_i}}\right).
\end{aligned}
\tag{12}
$$

Table 3: Restrictions for hierarchical models with at most three levels

Scenario	$Var(\beta)$	$Var(\xi)$	$Var(b_\beta)$	$Var(b_\xi)$	Additional Restriction
Model 1	-	-	∞	∞	No restriction
Model 2	-	$a_\xi/\,h_1^2$	∞	0	$E(b_\xi) = h_1$ and $a_{b_\xi} = h_1\, b_{b_\xi}$
Model 3	$a_\beta/\,h_2^2$	-	0	∞	$E(b_\beta) = h_2$ and $a_{b_\beta} = h_2 b_{b_\beta}$
Model 4	$a_\beta/\,h_2^2$	$a_\xi/\,h_1^2$	0	0	Idem Models 2 and 3
Model 5	0	-	0	∞	$\beta_i \sim \gamma(a_\beta\ ,\ a_\beta)$, with $a_\beta \approx \infty$
Model 6	0	$a_\xi/\,h_1^2$	0	0	$\beta_i \sim \gamma(a_\beta\ ,\ a_\beta)$, with $a_\beta \approx \infty$

h_1 e h_2 are fixed positive constants.

Samples from the quantities of interest were obtained by using MCMC techniques. Adaptive Rejection Metropolis sampling (ARMS) algorithm, introduced by Gilks *et al.* (1995), was used to sample from the posterior distribution of β , θ , a_β and a_ξ, whereas samples from the posterior distribution of b_β and b_ξ were obtained straightforwardly. All of third level prior distributions were assumed $Gamma(0.001; 0.001)$. For all scenarios considered in our analysis, it was considered single chains of size 300,000. Samples of size 25,000 were obtained after considering a burn-in period of 50,000 iterations and a lag of 10 to eliminate correlations. All computational procedures were implemented in the Object-oriented Matrix Programming Language Ox version 5.1 (Doornik , 2007), and are available from the first author upon request.

In Section 5, two real data set illustrate the metodology proposed in Sections 3 and 4.

5 Numerical Examples

In this section, the proposed methodology is applied in two data sets. The first one comes from Proschan (1963) and it is analyzed by some papers in the literature. The second one is the motivating example of this work and it was previously introduced in Section 1. In both, the graphs of the estimates were constructed as follows: internal points of the vertical lines and the internal horizontal line represent the exact estimate, that in Bayesian approach is the mean of posterior distribution. Bayesian approach intervals are Highest Posterior Density (HPD) intervals.

5.1 Proschan's air conditioning data set

This data set firstly appeared in Proschan (1963). Records were kept for the time of successive failures of the air conditioning system of each member of a fleet of Boeing 720 jet airplanes. The intervals between successive failures were recorded. Failure records from thirteen planes are described in Figure 3. This data set were analized by some papers in literature, such as Rigdon and Basu (1990), Pan and Rigdon (2009) and Gaver and O'Muircheartaigh (1987) among others.

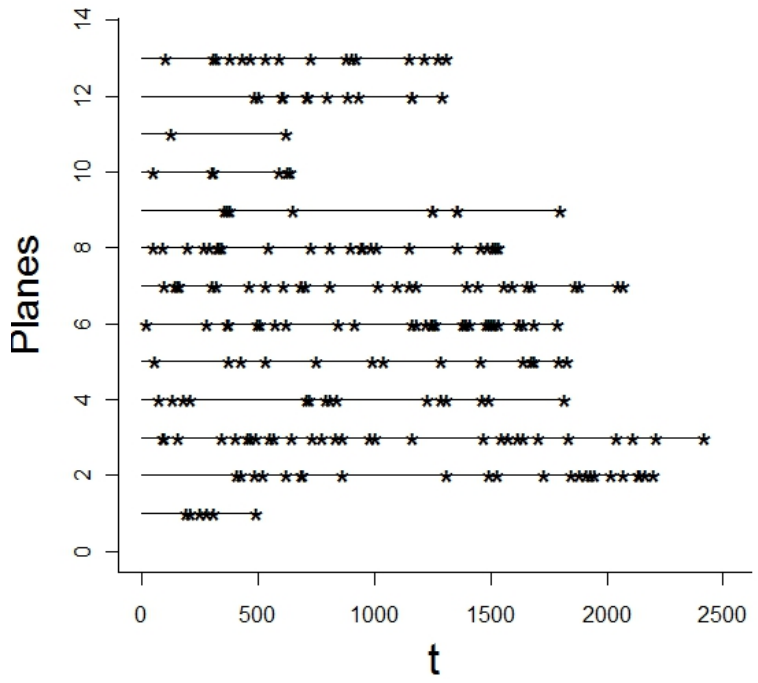

Figure 3: Failure Times of 13 Aircraft Air Conditioners ('$*$' represents failure times).

At first, some preliminary plots are presented in Figures 4, 5, 6 and 7. In each one of these figures, two models are compared according to their difference. Comparisons are made

by using MLE confidence intervals and HPD intervals. In general HPD intervals are smaller than the asymptotic MLE ones. Figures 4 and 5 indicate a intensity functions for planes with same β but larger than one. That means, data set should not be treated as a PPH.

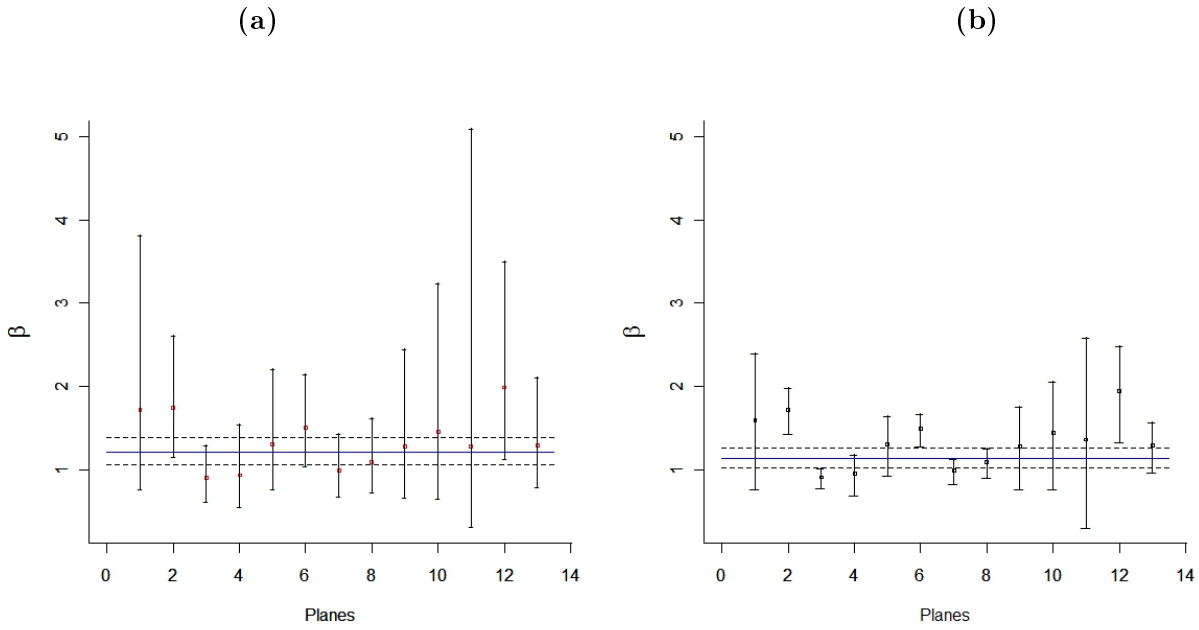

Figure 4: Aircraft β's comparisons between Models 1 and 3: (a) MLE confidence intervals and (b) HPD intervals. Vertical lines represent Model 1 and horizontal lines represent Model 3.

Table 4 presents LRT p-values comparing the six models under consideration. Notice that they confirm some indications from Figures 4 and 5. That is, there is a statistical evidence in favor Model 3. However, this fact does not confirm indications of Figures 6 and 7.

Table 4: LRT Aircraft data

Scenario	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6
Model 1	–	0.3051 (12)	0.3716 (12)	0.0472 (24)	0.0907 (13)	0.0089 (25)
Model 2		–	Not applicable	0.0301 (12)	Not applicable	0.0036 (13)
Model 3			–	0.0223 (12)	0.0072 (1)	0.0026 (13)
Model 4				–	Not applicable	0.0044 (1)
Model 5					–	0.0170 (12)

Degrees of freedom for chi square distribution in parenthesis.

Bayesian hierarchical analysis was also performed. Some information criteria were calculated in order to pick the best model. Results are presented in Table 5. According to AIC and BIC criteria, Model 4 is the best one. Model 4 considers same PLP for the systems. This result is in agreement of those of (Proschan , 1963). On the other hand, DIC indicates the same model as the LRT.

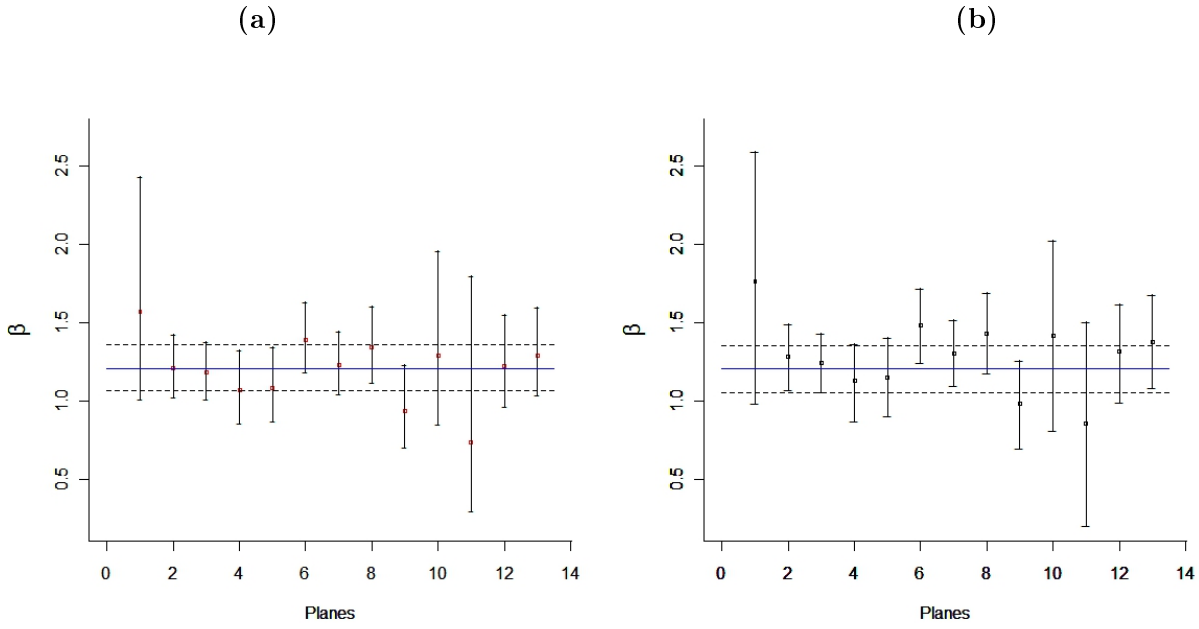

Figure 5: Aircraft β's comparisons between Models 2 and 4: (a) MLE confidence intervals and (b) HPD intervals. Vertical lines represent Model 2 and horizontal lines represent Model 4.

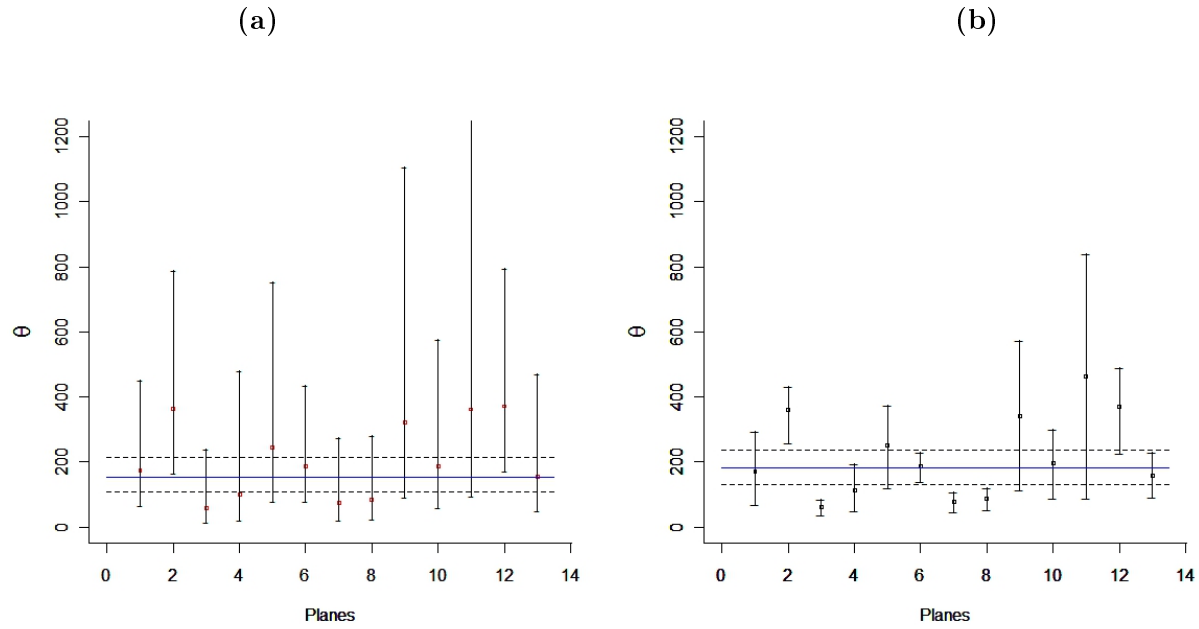

Figure 6: Aircraft θ's comparisons between Models 1 and 2: (a) MLE confidence intervals and (b) HPD intervals. Vertical lines represent Model 1 and horizontal lines represent Model 2.

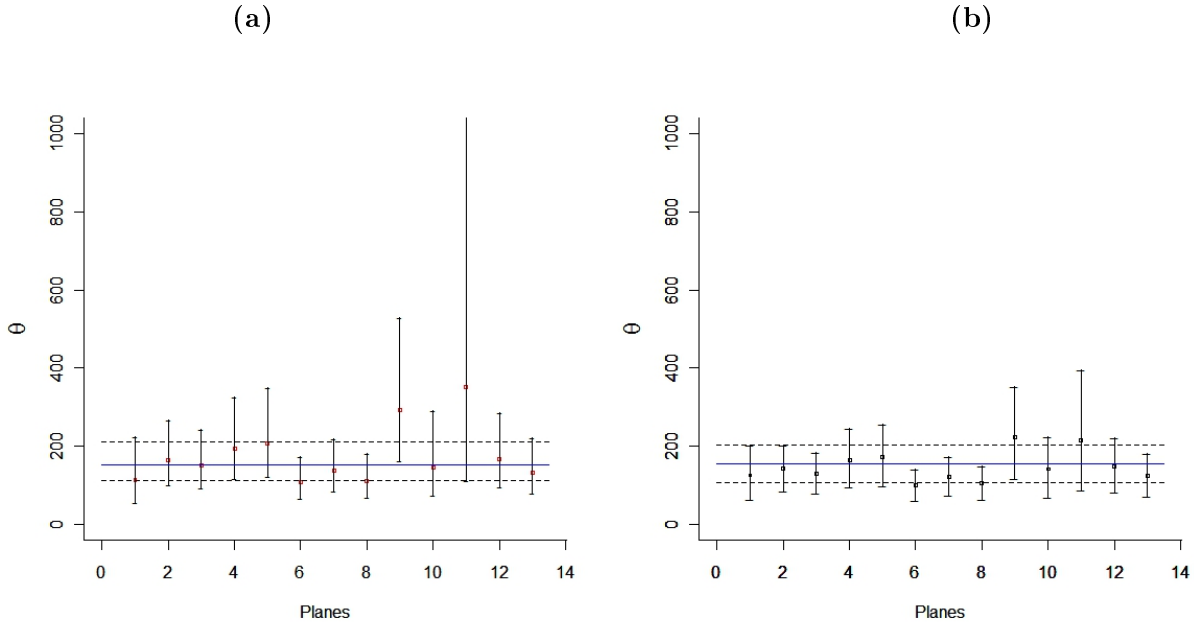

Figure 7: Aircraft θ's comparisons between Models 3 and 4: (a) MLE confidence intervals and (b) HPD intervals. Vertical lines represent Model 3 and horizontal lines represent Model 4.

Table 5: Decision criteria for Aircraft data

Scenario	AIC	BIC	DIC
Model 1	2364.774	2379.463	1177.515
Model 2	2354.706	2362.616	1177.659
Model 3	2353.739	2361.649	**1174.863**
Model 4	**2353.440**	**2354.570**	1176.731
Model 5	2358.959	2366.304	1176.595
Model 6	2359.532	2360.097	1179.757

5.2 Example Electrical Power Transformers

Data set in Figure 1 presents the failure histories of eleven 500 kvolt transformers belonging to a state electrical power company in Brazil, between 1999 and 2009. Electric power transformers are the basic components of an electrical power transmission system. They are complex and most of their repairs involve the replacement of only a small fraction of their parts. Therefore, its reasonable to suppose that the system's reliability after a repair is essentially the same as before the failure. This fact characterizes MR and, hence, justifies the statistical analysis by using a NHPP for these data.

LRT in Table 7 indicates the most simple is the best fitted model. Model 6 consists of only one parameter. Figures 8, 9, 10 and 11 describe these results. These figures compares two scenarios where one of them is nested in the other. Asymptotic confidence intervals are graphed in these figures. Some intervals lower limits were negatives. Asymmetrical intervals, calculated by exponential of the interval of the logarithm of paramete have greater range

Table 6: Parameter Estimates for Aircraft Data under Model 4

Parameter	MLE		Bayes	
	Exact	Interval	Exact	Interval
β	1.205	[1.055 , 1.354]	1.205	[1.055 , 1.353]
θ	152.5	[103.7 , 201.2]	153.172	[106.55 , 203.45]

than the original intervals.

Models 1 and 3 (just one β) were compared in Figure 8 and 9. Models 1 and 2 (just one θ) were compared in In Figure 10 and models 3 and 4 in Figure 11. Almost of the intervals of the larger model intercepts the confidence interval of the smaller model. This is in agreement to the adequacy of model 6.

Table 7: LRT Transformer data

Scenario	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6
Model 1	–	0.416 (12)	0.102 (12)	0.278 (24)	0.135 (13)	0.322 (25)
Model 2		–	Not applicable	0.228 (12)	Not applicable	0.284 (13)
Model 3			–	0.698 (12)	0.620 (1)	0.757 (13)
Model 4				–	Not applicable	0.647 (1)
Model 5					–	0.701 (12)

Degrees of freedom for chi square distribution in parenthesis

In hierarchical Bayesian approach same prior distributions used in Section 4.2 were considered here. DIC criterion agrees with AIC and BIC criteria for this case. They are in Table 8. Estimate for β and θ form Model 6 are in Table 9.

Table 8: Decision criteria for Transformer data

Scenario	AIC	BIC	DIC
Model 1	411.013	419.767	200.940
Model 2	409.414	414.189	200.434
Model 3	406.937	411.712	198.623
Model 4	394.226	395.021	197.113
Model 5	405.182	409.559	197.998
Model 6	**392.436**	**392.833**	**196.170**

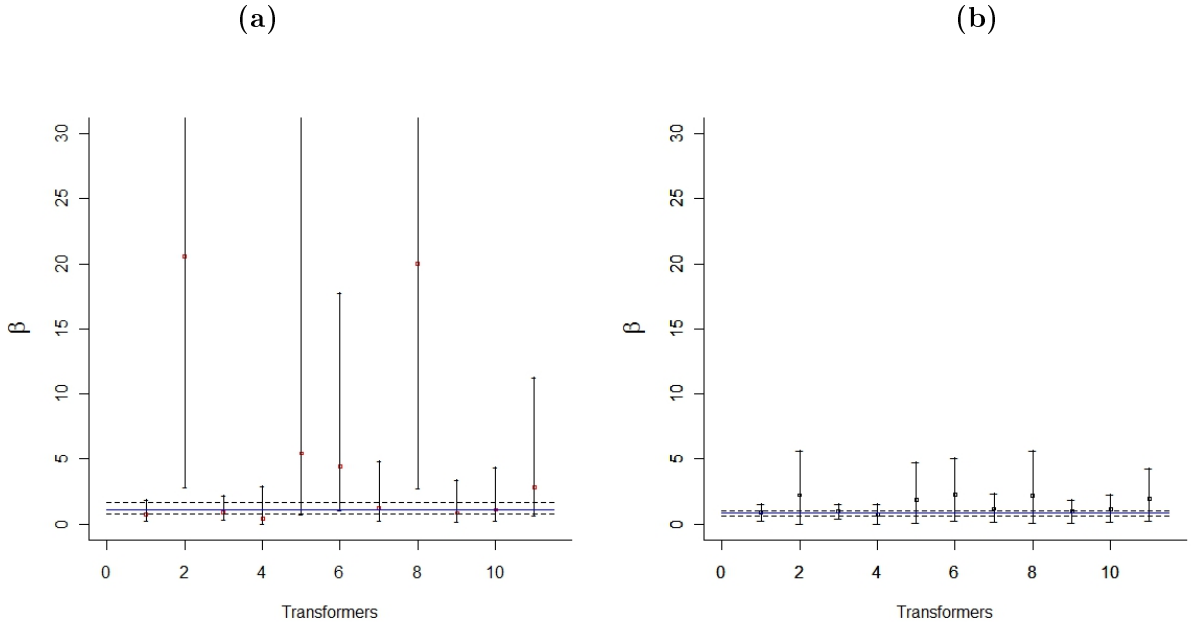

Figure 8: Transformers β's comparisons between Models 1 and 3: (a) MLE confidence intervals and (b) HPD intervals. Vertical lines represent Model 1 and horizontal lines represent Model 3.

Table 9: Parameter Estimates for Transformer Data under Model 6

E_{MV}	Exact	1785.997
	Interval	[1055.94 , 2516.06]
E_{Bayes}	Exact	1951.058
	Interval	[1054.84 , 2873.85]

References

H. Ascher and H. W. Feingold, *Repairable Systems Reliability: Modeling, Inference, Misconception and their Causes*, Marcel Dekker: New York, 1984.

L. J. Bain and M. Engelhardt, *Statistical Analysis of Reliability and Life-Testing Models, Theory and Methods* Marcel Dekker: New York, 1991.

R. D. Baker, Some new tests of the power law process. *Technometrics* 38, pp. 256-265, 1996.

R. E. Barlow and L. C. Hunter, Optimum Preventive Maintenance Policies. *Operations Research* 8, pp. 90-100, 1960.

M. Berman and T. R. Turner, Approximate Point Process Likelihoods with GLIM. *Applied Statistics* 41, pp. 31-38, 1992.

M. Bhattacharjee, J. V. Deshpande and U. V. Naik-Nimbalkar, Unconditional Tests of Goodness of Fit for the Intensity of Time-Truncated Nonhomogeneous Poisson Processes. *Technometrics* 46, pp. 330-338, 2004.

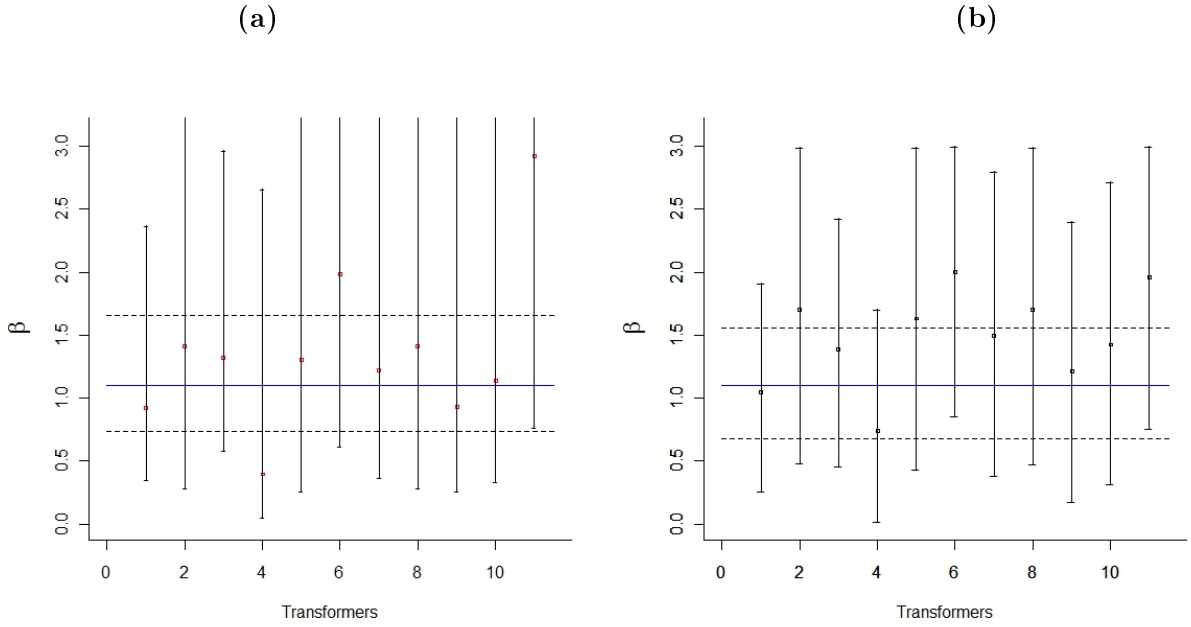

Figure 9: Transformers β's comparisons between Models 2 and 4: (a) MLE confidence intervals and (b) HPD intervals. Vertical lines represent Model 2 and horizontal lines represent Model 4.

H. W. Block, W. S. Borges and T. H. Savits, A General Age Replacement with Minimal Repair. *Naval Research Logistics* 35, pp. 365-372, 1990.

G. E. P. Box and G. C. Tiao, *Bayesian Inference in Statistical Analysis*, Addison-Wesley: Reading, 1973.

L. R. Crow, Reliability Analysis for Complex Systems. *Reliability and Biometry*, F. Proschan and J. Serfling (Eds), pp. 379-410, 1974.

L. Devroye, Non-Uniform Random Variate Generation. New York, Springer-Verlag, 1986.

J. A. Doornik, Ox 5 - An Object-oriented Matrix Programming Language, Timberlake Consultants Ltd, 2007.

Gaver, Donald P. and O'Muircheartaigh, I. G., Robust Empirical Bayes Analyses of Event Rates. *Technometrics*, Vol. 29, No. 1, pp. 1-15, 1987

A. Gelman, J. B. Carlin, H. S. Stern and D. B. Rubin *Bayesian Data Analysis*. Chapman and Hall: New York, 2003.

I. B. Gerstack, Models of Preventive Maintenance. North Holand: Amsterdam, 1977.

G. L. Gilardoni and E. A. Colosimo, Optimal maintenance time for repairable systems. *Journal of Quality Technology* 39, pp. 48-53, 2007.

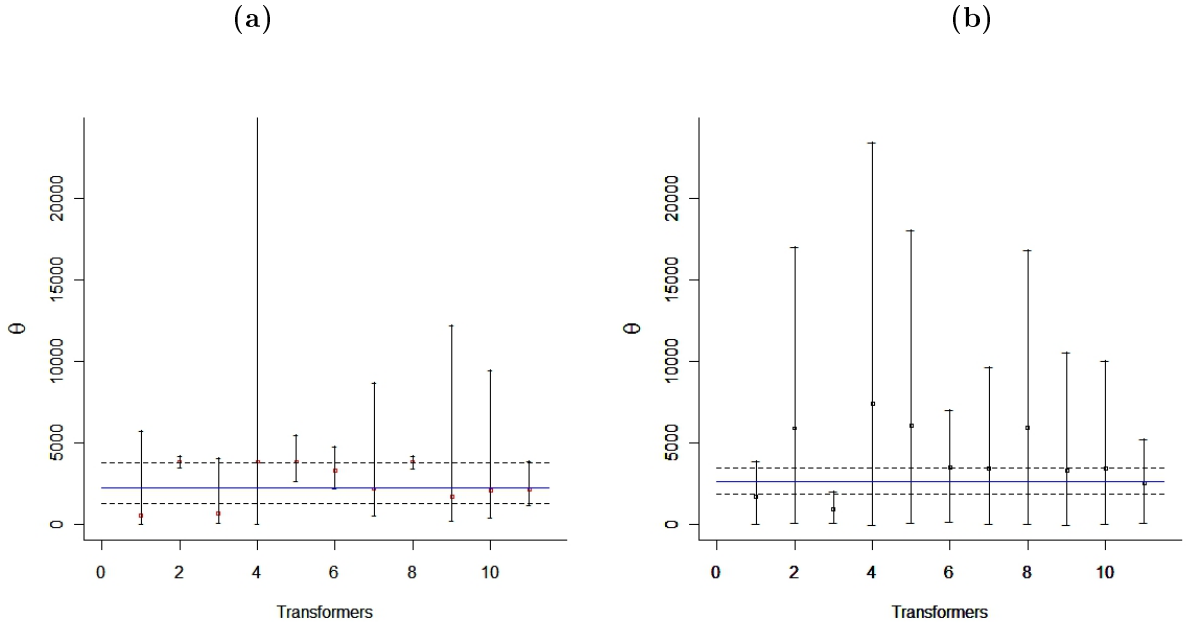

Figure 10: Transformers θ's comparisons between Models 1 and 2: (a) MLE confidence intervals and (b) HPD intervals. Vertical lines represent Model 1 and horizontal lines represent Model 2.

W. R. Gilks and P. Wild, Adaptive rejection sampling for gibbs sampling. *Applyed Statistical* 41, pp. 337-348, 1992.

W. R. Gilks, N. G. Best and K. K. C. Tan, Adaptive rejection metropolis sampling within gibbs sampling. *Applyed Statistical* 44, pp. 455-472, 1995.

W. R. Gilks, S. Richardson and D. J. Spiegelhalter *Markov Chain Mont Carlo in Practice: Interdisciplinary Statistics*. Chapman & Hall, 1996.

M. Guida, R. Calabria and G. Pulcini, Bayes Inference for a Non-Homogeneous Poisson Process with Power Intensity Law. *IEEE Transactions on Reliability* Vol. 38, pp. 603-609, 1989.

Y. S. Huang and V. B. Bier, A Natural Conjugate Prior for the Non-Homogeneous Poisson Process with a Power Law Intensity Function. *Communications in Statistics - Simulation and Computation* 27, pp. 525-551, 1998.

E. L. Lehmann, Testing Statistical Hypotheses. Wiley, New York, 1986.

M. Oliveira, E. Colosimo and G. Gilardoni, Bayesian Inference for Power Law Processes with Applications in Repairable Systems. *Journal of Statistical Planning and Inference*. Submetido para publicação, 2010

R. Pan and S. E. Rigdon, Bayes Inference for General Repairable Systems. *Journal of Quality Technology* Vol. 41, No. 1, pp. 82-94, January 2009.

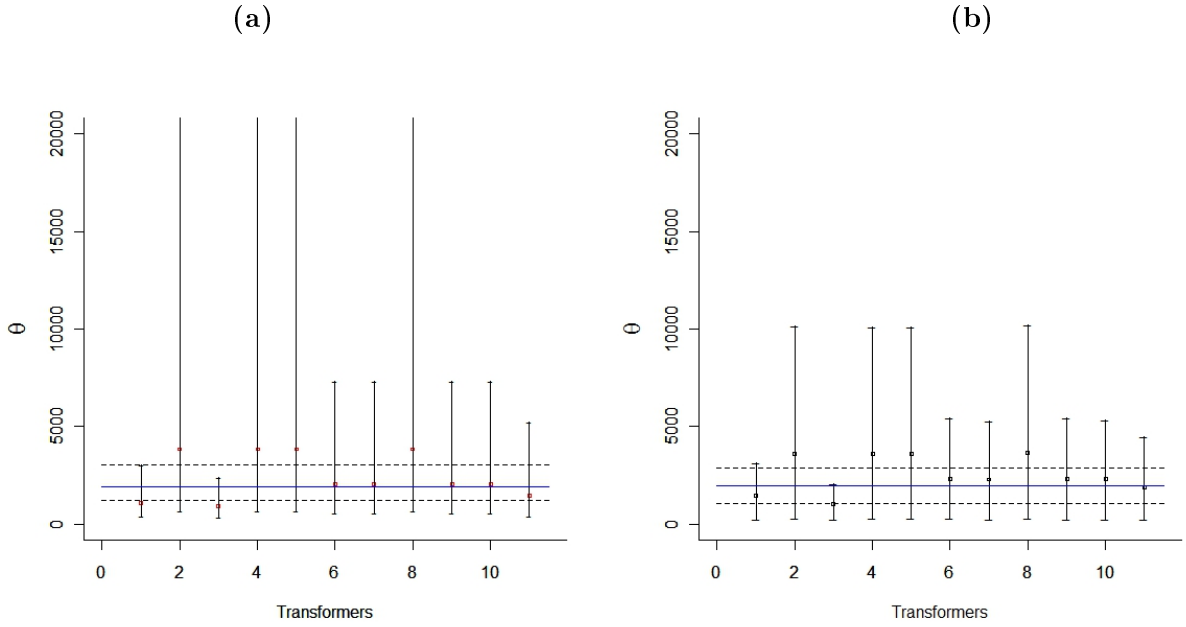

Figure 11: Transformers θ's comparisons between Models 3 and 4: (a) MLE confidence intervals and (b) HPD intervals. Vertical lines represent Model 3 and horizontal lines represent Model 4.

D. H. Park, G. M. Jung and J. K. Yum, Cost Minimization for Periodic Maintenance Policy of a System Subject to Slow Degrada- tion. *Reliability Engineering & System Safety* 68, pp. 105-112, 2000.

F. Proschan, Theoretical Explanation of Observed Decreasing Failure Rate. *Technometrics* Vol. 5, No. 3, pp. 375-383, 1963.

S. E. Rigdon and A. P. Basu, *Statistical Methods for the Reliability of Repairable Systems*. John Wiley: New York, 2000.

S. E. Rigdon and A. P. Basu, *The Effect os Assuming a Homogeneus Poisson Process when the True Process Is a Power Law Process*. Journal of Quality Tachnology 11, pp. 111-117, 1990.

A. Sen, Bayesian Estimation and Prediction of the Intensity of the Power Law Process. *Journal of Statistical Computation and Simulation* Vol. 72, pp. 613-631, 2002.

M. Zhao and M. Xie, On Maximum Likelihood Estimation for a General Non-homogeneous Poisson Process. *Scandinavian Journal of Statistics* 23, 597-607, 1996.

Conclusões

Nesta tese de doutorado foram apresentadas algumas propostas de modelagem para história de falhas de sistemas reparáveis, sujeitos a reparo mínimo e manutenção perfeita. O Processo de Poisson Não Homogêneo norteou todas as abordagens consideradas. Conjuntos de dados referentes a história de falhas de transformadores elétricos, foram o grande motivador no desenvolvimento de todo o trabalho, apesar de que em algum momento, um conjunto de dados tido como auxiliar, exaustivamente estudado na literatura de sistemas reparáveis, foi adotado e reanalisado através da modelagem bayesiana hierárquica.

A seguir, são discutidas resumidamente as principais contribuições e resultados associados a cada texto apresentado nesta tese.

Primeiro texto

Esse texto apresentou uma abordagem não paramétrica para estimar a função intensidade de um processo de Poisson não homogêneo, sob restrição de crescimento. Foi mostrado como a transformação *tempo total sob teste* - TTT pode ser usada para agregar tempos de eventos da superposição de processos de Poisson não homogêneos. A partir daí, metodologias conhecidas para um único processo, como por exemplo a contida em Boswell (1966) pode ser adotada para vários processos parcialmente superpostos. Intervalos de confiança bootstrap foram construídos a partir de dois métodos. No primeiro método foram consideradas reamostras de sistemas, considerando a suposição de independência entre esses sistemas. Esse método apresentou precisão menor do que os intervalos obtidos pelo segundo método, no qual a reamostragem é realizada nos tempos transformados, ou seja, nos tempos dos eventos do processo resultante da superposição. Uma hipótese levantada para a diferença na precisão é que a reamostragem nos sistemas considera os tempos finais de acompanhamento T_i, de cada sistema, como sendo aleatórios, enquanto que a reamostragem nos tempos transformados (via a transformação TTT) os T_i's ficam fixos. Ou seja, como o primeiro método amostra os transformadores, estes são considerados como vindo da mesma população. O que não acontece quando os tempos são agregados antes do procedimento de reamostragem. Quando todos os sistemas são truncados num mesmo tempo T fixo, essa diferença entre as precisões dos dois métodos fica consideravelmente reduzida.

Segundo texto

Uma nova parametrização para a função intensidade de um processo lei de potências foi proposta nesse trabalho. A parametrização permite identificar uma família de distribuiçãoes a *priori* conjugada como um produto de funções densidade de variáveis aletórias com distribuição gama para os parâmetros do PLP. A metodologia é estendida ao caso de muitas

realizações independentes de um mesmo PLP, no entanto a conjugação não mais se verifica. Apesar disso, simulações da distribuição a *posteriori* é de fácil implementação.

Restrito à estimação de uma função intensidade crescente, a abordagem Bayesiana considerou apenas distribuições a *priori* para β cujo suporte era o intervalo $(1, \infty)$. Isso foi possível, ou truncando a distribuição, ou propondo uma distribuição deslocada do valor 1. As distribuições truncadas em $\beta = 1$, consideradas, foram a distribuição a *priori* de referência $\pi(\beta, \eta) \propto \beta^{-1}\,\eta^{-1}$, a de Jeffreys $\pi(\beta, \eta) \propto (\beta\eta)^{-1}$ e o produto de funções de distribuição gama $\pi(\beta, \eta) \propto \gamma(\beta|\, a_\beta, b_\beta) \times \gamma(\eta|\, a_\eta, b_\eta)\, I_{(1,\infty)}(\beta)$, como sugerido pela reparametrização adotada. Por fim, foi considerada para β uma distribuição gama deslocada de 1, mantendo-se para η a distrbuição a *priori* gama usual. Simulações de Monte Carlo mostraram a vantagem da abordagem Bayesiana sobre a de máxima veossimilhança, principalmente no caso de amostras com poucas falhas. As simulações foram realizadas sob vários cenários distintos de tamanho de amostra, todos com truncamento temporal, seguindo o esquema amostral de truncamente por tempo, que é o mesmo dos dados reais sobre a história de falhas de transformadores de potência elétrica, analisados. Não foi observada diferença significativa entre as distribuições a *priori* propostas através dos cenários avaliados.

Terceiro texto

Nesse trabalho modelamos hierarquicamente o problema de ajustar dados de vários processos de Poisson não homogêneos, admitindo um processo lei de potências diferente para cada sistema. São consideradas as abordagens frequentista e Bayesiana para identificar o melhor modelo para se ajustar aos dados. Modelos com três níveis foram propostos na abordagem Bayesiana hierárquica, estabelecendo como primeiro nível o nível dos PLPs, no segundo nível ficaram os parâmetros (β, θ) dos PLPs sendo o terceiro nível, o nível dos parâmetros das distribuições a *priori* propostas para os parâmetros β e θ do segundo nível. A reparametrização da função intensidade do PLP do segundo texto, serviu como base para estabelecermos uma distribuição a *priori* para os parâmetros do segundo nível do modelo hierárquico. Nesse sentido, foram propostas distribuição gama para β e para $\xi = \Lambda(1)$, e a distribuição a *priori* para θ foi deduzida a partir destas. Essa parametrização respeita a permutabilidade das distribuições a *priori*. Os valores dos hiperparâmetros do segundo nível de um modelo hierárquico com três níveis, surgem naturalmente como uma medida resumo das distriuições a *posteriori* dos parâmteros do terceiro nível. Distribuições gama com ambos os parâmteros iguais a 10^{-3} foram propostas para cada um dos quatro hiperparâmetros do terceiro nível.

Os resultados obtidos a partir da análise do conjunto de dados sobre ar condicionado de aviões, de Proschan (1963) mostraram uma heterogeneidade entre aviões que não havia sido detectada na análise original para esses dados. Outros autores analisaram esses mesmos dados considerando os sistemas sempre como réplicas de um mesmo. A abordagem apresentada aqui, mostrou ser esta uma prática inadequada para esses dados. Já a probreza de

informação dos dados dos transformadores, fez com que a aplicação da metodologia a estes dados fosse pouco conclusiva, na detecção da heterogeneidade entre sistemas para esses dados, indicando sempre o modelo mais simples. Apesar de observamos intervalos de confiança assintóticos mais estreitos do que os intervalos HPD, para sistemas com poucas falhas, a teoria assintótica na abordagem por máxima verossimilhança pode 'falhar' com tão poucos dados.

Uma observação importante sobre a quantidade de falhas em cada sistema, é que a teoria por máxima verossimilhança só pode ser aplicada, se for observada pelo menos uma falha por sistema. O que não é necessário na abordagem Bayesiana.

Trabalhos Futuros

Este trabalho trata da modelagem em processos de Poisson não homogêneos. São abordadas metodologias não paramétrica, Bayesiana e hierárquica associadas a esse tema. Por outro lado, nenhuma das situações discutidas considera a inclusão de covariáveis, o que pode ser tratado num trabalho futuro. Outro tema que pode ser levantado, e que pode atuar em conjunto com a abordagem hierárquica, é a introdução de fragilidades (efeitos aleatórios) na modelagem. Modelos de fragilidade tem recebido crescente atenção na literatura estatística, sobretudo nas duas últimas décadas, com especial destaque às abordagens Bayesianas. Outra possível proposta de trabalho é a elaboração de um artigo de revisão cobrindo as principais abordagens existentes para PPNHs, e o desenvolvimento de um pacote estatístico para ajustar esses modelos, para ser disponibilizado em R.

Referências

Aalen, O. Nonparametric inference for a family of counting processes. *The Annals of Statistics 6*(4), 701–726, 1978.

H. Ascher and H. W. Feingold, *Repairable Systems Reliability: Modeling, Inference, Misconception and their Causes*, Marcel Dekker: New York, 1984.

L. J. Bain and M. Engelhardt, *Statistical Analysis of Reliability and Life-Testing Models, Theory and Methods* Marcel Dekker: New York, 1991.

R. E. Barlow and L. C. Hunter, Optimum Preventive Maintenance Policies. *Operations Research* 8, pp. 90-100, 1960.

H. W. Block, W. S. Borges and T. H. Savits, A General Age Replacement with Minimal Repair. *Naval Research Logistics* 35, pp. 365-372, 1990.

Boswell, M. T. Estimating and testing trend in a stochastic process of Poisson type. *The Annals of Mathematical Statistics* 37(6), pp. 1564-1573, 1966.

L. R. Crow, Reliability Analysis for Complex Systems. *Reliability and Biometry*, F. Proschan and J. Serfling (Eds), pp. 379-410, 1974.

A. Cowling, P. Hall, and M. J. Phillips, Bootstrap Confidence Regions for the Intensity of a Poisson Point Process. *Journal Of The American Statistical Association 91*, 436, pp.1516-1524, 1996.

C. A. Field and A. H. Welsh, Bootstrapping clustered data. *Journal of Royal Statistical Society*, 69, pp. 369-390, 2007.

I. B. Gerstack, Models of Preventive Maintenance. North Holand: Amsterdam, 1977.

G. L. Gilardoni and E. A. Colosimo, Optimal maintenance time for repairable systems. *Journal of Quality Technology* 39, pp. 48-53, 2007.

D. H. Park, G. M. Jung and J. K. Yum, Cost Minimization for Periodic Maintenance Policy of a System Subject to Slow Degrada- tion. *Reliability Engineering & System Safety* 68, pp. 105-112, 2000.

F. Proschan, Theoretical Explanation of Observed Decreasing Failure Rate. *Technometrics* Vol. 5, No. 3, pp. 375-383, 1963.

S. E. Rigdon and A. P. Basu, *Statistical Methods for the Reliability of Repairable Systems*. John Wiley: New York, 2000.