

**UNIVERSIDADE FEDERAL DE MINAS GERAIS**  
**INSTITUTO DE CIÊNCIAS EXATAS**  
**Programa de Pós-graduação em Estatística Ênfase em Indústria e Mercado**

Robert Bento Florentino

**DETERMINAÇÃO DA JANELA ÓTIMA DE MANUTENÇÃO PARA SISTEMAS  
REPARÁVEIS NA PRESENÇA DE COVARIÁVEIS**

Belo Horizonte

2021

Robert Bento Florentino

**DETERMINAÇÃO DA JANELA ÓTIMA DE MANUTENÇÃO PARA SISTEMAS  
REPARÁVEIS NA PRESENÇA DE COVARIÁVEIS**

Versão Final

Monografia de especialização apresentada ao Programa de Pós-Graduação em Estatística do Instituto de Ciências Exatas da Universidade Federal de Minas Gerais, como requisito parcial à obtenção do título de Especialista em Estatística.

Orientador: Prof. Dr. Fábio Nogueira Demarqui

Belo Horizonte

2021

© 2021, Robert Bento Florentino.  
Todos os direitos reservados.

Florentino, Robert Bento

F633d     Determinação da janela ótima de manutenção para sistemas reparáveis na presença de covariáveis [manuscrito] / Robert Bento Florentino. Belo Horizonte - 2021.  
73 f. : il. ; 29 cm corrigir

Orientador: Fábio Nogueira Demarqui.

Monografia (especialização) - Universidade Federal de Minas Gerais, Instituto de Ciências Exatas, Departamento de Estatística.  
Referências : f. 50

1. Estatística. 2. Verossimilhança (Estatística).  
3. Análise de covariância. 4. Teoria do potencial (Matemática). 5. Equipamento industrial – Manutenção e reparos. I. Demarqui, Fábio Nogueira. II. Universidade Federal de Minas Gerais, Instituto de Ciências Exatas, Departamento de Estatística. IV. Título.

CDU 519.2(043)

Ficha catalográfica elaborada pela bibliotecária Irénquer Vismeg  
Lucas Cruz - CRB 6ª n° 819.



**Universidade Federal de Minas Gerais**  
**Instituto de Ciências Exatas**  
**Departamento de Estatística**  
**Programa de Pós-Graduação / Especialização**  
Av. Pres. Antônio Carlos, 6627 - Pampulha  
31270-901 – Belo Horizonte – MG

E-mail: [pgest@ufmg.br](mailto:pgest@ufmg.br)  
Tel: 3409-5923 – FAX: 3409-5924

### **ATA DO 234ª. TRABALHO DE FIM DE CURSO DE ESPECIALIZAÇÃO EM ESTATÍSTICA DE ROBERT BENTO FLORENTINO.**

Aos vinte e seis dias do mês de outubro de 2021, às 16:30 horas, com utilização de recursos de videoconferência a distância, reuniram-se os professores abaixo relacionados, formando a Comissão Examinadora homologada pela Comissão do Curso de Especialização em Estatística, para julgar a apresentação do trabalho de fim de curso do aluno **Robert Bento Florentino**, intitulado: “*Determinação da Janela Ótima de Manutenção para Sistemas Reparáveis na Presença de Covariáveis.*”, como requisito para obtenção do Grau de Especialista em Estatística. Abrindo a sessão, o Presidente da Comissão, Professor Fábio Nogueira Demarqui – Orientador, após dar conhecimento aos presentes do teor das normas regulamentares, passou a palavra ao candidato para apresentação de seu trabalho. Seguiu-se a arguição pelos examinadores com a respectiva defesa do candidato. Após a defesa, os membros da banca examinadora reuniram-se sem a presença do candidato e do público, para julgamento e expedição do resultado final. Foi atribuída a seguinte indicação: o candidato foi considerado Aprovado condicional às modificações sugeridas pela banca examinadora no prazo de 30 dias a partir da data de hoje por unanimidade. O resultado final foi comunicado publicamente ao candidato pelo Presidente da Comissão. Nada mais havendo a tratar, o Presidente encerrou a reunião e lavrou a presente Ata, que será assinada por todos os membros participantes da banca examinadora. Belo Horizonte, 26 de outubro de 2021.

Prof. Fábio Nogueira Demarqui (Orientador)  
Departamento de Estatística / ICEx / UFMG

Prof. Enrico Antônio Colosimo  
Departamento de Estatística / ICEx / UFMG

Prof.ª Maristela Dias de Oliveira  
Departamento de Estatística / UFBA



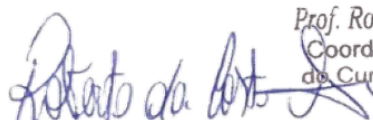
**Universidade Federal de Minas Gerais**  
**Instituto de Ciências Exatas**  
**Departamento de Estatística**  
**Programa de Pós-Graduação / Especialização**  
Av. Pres. Antônio Carlos, 6627 - Pampulha  
31270-901 – Belo Horizonte – MG

E-mail: [pgest@ufmg.br](mailto:pgest@ufmg.br)  
Tel: 3409-5923 – FAX: 3409-5924

## **DECLARAÇÃO DE CUMPRIMENTO DE REQUISITOS PARA CONCLUSÃO DO CURSO DE ESPECIALIZAÇÃO EM ESTATÍSTICA.**

Declaro para os devidos fins que **Robert Bento Florentino**, número de registro 2019705863, cumpriu todos os requisitos necessários para conclusão do curso de Especialização em Estatística, entregando a versão final do trabalho de conclusão de curso para seu orientador, o professor Fábio Nogueira Demarqui, que aprovou a versão final. O trabalho foi apresentado no dia 26 de outubro de 2021 com o título “*Determinação da Janela Ótima de Manutenção para Sistemas Reparáveis na Presença de Covariáveis.*”.

Belo Horizonte, 25 de novembro de 2021

  
Prof. Roberto da Costa Quinino  
Coordenador da Comissão  
de Curso de Especialização  
em Estatística  
Prof. Roberto da Costa Quinino  
Coordenador do curso de  
Especialização em Estatística  
Departamento de Estatística / UFMG

## **Agradecimentos**

Agradeço à minha família: Meu pai Adeildo, minha mãe Regina, meu irmão Roger e aos recém-chegados sobrinho Martim e cunhada Fabiana. Vocês me ajudaram neste trabalho, assim como em tudo que faço, mesmo que indiretamente e/ ou inconscientemente.

Aos professores dos cursos de Especialização em estatística da UFMG pelo excelente curso, didática e qualidade nos conteúdos altamente relevantes para o contexto profissional industrial cuja carreira escolhi.

Aos colegas do curso de especialização em estatística da UFMG, turma de 2019, especialmente às boas amizades que fiz aqui: Flora, Bernardo e Vinícios – Vocês tornaram essa jornada árdua mais fácil e prazerosa em meio aos trabalhos, discussões sobre estatística e tanto outros assuntos aleatórios (sic).

Ao professor Fábio ao qual dedicou dezenas de horas em reuniões quinzenais noturnas e se mostrou sempre extremamente disposto a transmitir conhecimento, discutir ideias, além da grande paciência em ensinar R, matemática e estatística relativamente avançadas à um leigo engenheiro. Obrigado por ser um excelente orientador e um excelente mentor! Espero que possamos continuar trabalhando em outros projetos futuros.

A meus colegas e amigos de trabalho... Neste período estive em diferentes cidades e diferentes empresas e conciliar vida pessoal, a rotina de trabalho e os estudos da especialização em estatística foi um desafio mais fácil graças ao agradável convívio e amizade de vocês.

Agradeço também aos vários acasos e àqueles que fizeram parte desta trajetória, mas que, de alguma forma, partiram...

*“The idea is to remain in a state of constant departure while always arriving. It saves on introductions and goodbyes. The ride does not require explanation - just occupants.”*

**Boat Car Guy, Waking Life (2001)**

*“Ever tried. Ever failed. No matter. Try again. Fail again. Fail better.”*

***Samuel Beckett***

## Resumo

Um importante campo de aplicação de confiabilidade é em dados de manutenção de sistemas e equipamentos industriais. Em grande parte das vezes, esses sistemas passam por várias manutenções preventivas e corretivas ao longo de sua vida útil e por esse motivo são classificados como sistemas reparáveis. Esta monografia apresenta uma revisão bibliográfica com ênfase prática da aplicação do Processo Lei de Potência para modelagem de sistemas reparáveis em dois *softwares* diferentes: *Minitab* por sua interface gráfica e vasta aplicação na indústria; e *R Studio*, *software* que utiliza linguagem R para as mais diversas aplicações em estatística. O trabalho apresenta como aprofundamento a proposição de um modelo que considera a inclusão de covariáveis na função intensidade do Processo Lei de Potência e assume o parâmetro de forma comum para diferentes níveis que podem representar em aplicações práticas lotes, modelos, fabricantes, tipos de componentes etc. Como primeiro estudo de caso, são utilizados dados reais de manutenção de transformadores elétricos apresentado por Gilardoni e Colosimo (2007) com a finalidade de replicar a mesma análise apresentada em seu artigo. No segundo estudo de caso, dados de falhas de componentes de freio de locomotivas de dois tipos distintos são analisados empregando as duas versões do modelo de Lei de Potência: o padrão e o com covariáveis.

Palavras-chave: *Sistemas Reparáveis; Processo Lei de Potência; Método de Máxima Verossimilhança; Manutenção Preventiva;*



## **Abstract**

An important field of application of reliability is in maintenance data for industrial systems and equipment. Most of the time, these systems undergo several preventive and corrective maintenance throughout their useful lifespan, and, for this reason, they are classified as repairable systems. This thesis presents a literature review with practical emphasis on the application of the Power Law Process for modeling repairable systems in two different software: *Minitab* for its graphical interface and wide application in industry; and *R Studio*, software that uses R language for the most diverse applications in statistics. The thesis presents as a deepening the proposal of a model that considers the inclusion of covariates in the Power Law Process's intensity function assuming a single shape parameter for different levels that can represent batches, models, manufacturers, types of components, etc. in practical applications. As a first case study, real maintenance data for electrical transformers presented by Gilardoni and Colosimo (2007) are used to replicate the same analysis presented in their article. In the second case study, failure data for locomotive brake components of two distinct types are analyzed using the two versions of the Power Law model: the standard and the one with covariates.

*Keywords: Repairable Systems; Power Law Process; Maximum Likelihood Method; Preventive Maintenance.*

## Sumário

1. Introdução .....	10
2. Objetivos .....	12
3. Revisão Bibliográfica .....	13
3.1. Confiabilidade .....	13
3.1.1. Características de dados de Confiabilidade .....	13
3.1.2. Sistemas Reparáveis e Não-Reparáveis .....	14
3.2. Análises de Sistemas reparáveis e outros dados de recorrência .....	15
3.2.1. Modelo Não-Paramétrico para dados de recorrências .....	16
3.2.2. Modelos Paramétricos para dados de recorrências .....	19
3.3. Métodos em Inferência Estatística – Método da Máxima Verossimilhança .....	22
3.3.1. Função de Log-Verossimilhança para o Processo Lei de Potência .....	22
3.4. Função de Custos de Manutenção e Determinação do Tempo Ótimo de Manutenção para Sistemas Reparáveis .....	25
4. Metodologia .....	29
4.1. Modelagem Estatística dos Dados de Falha .....	29
4.2. Implementação do Modelo de Lei de Potência considerando grupos como covariáveis .....	29
5. Resultados Numéricos .....	33
5.1. Dados de falhas de transformadores – Lei de Potência .....	33
5.1.1. Descrição dos dados .....	33
5.1.2. Modelagem utilizando <i>Software Minitab</i> .....	34
5.1.3. Modelagem utilizando o <i>Software R</i> .....	35
5.1.4. Análise Comparativa dos Resultados - Validação .....	36
5.2. Dados de falhas em Componentes de freio de Locomotivas – Lei de Potência com covariáveis .....	38
5.2.1. Descrição dos dados .....	38
5.2.2. Modelagem dos dados considerando Lei de Potência sem Covariáveis no <i>Software Minitab</i> .....	39
5.2.3. Modelagem dos dados considerando Lei de Potência com Covariáveis no <i>Software R</i> .....	41
5.2.4. Análise Comparativa dos Resultados aplicando os dois modelos de Lei de Potência .....	42
6. Discussão .....	44
7. Conclusão .....	46

8. Trabalhos Futuros.....	49
9. Referências Bibliográficas .....	50
Anexo I – Dados de Falhas de Transformadores Elétricos .....	51
Anexo II – Dados de Falhas de Componentes de Freio de Locomotivas.....	53
Anexo III – Configuração dos dados no <i>Minitab</i> .....	54
Anexo IV – Algoritmo para Cálculo da Função Acumulada Média – Método Não- Paramétrico (Nelson, 2003) .....	57
Anexo V - Código do script R para Maximização da Log-Verossimilhança da Lei de Potência .....	58

## 1. Introdução

Equipamentos e máquinas industriais são um grande exemplo de sistemas em projetados para serem reparados ao invés de substituídos quando falham. A engenharia destes sistemas é assim considerada em razão da complexidade destes, de seus elevados custos de produção, prazos de fabricação e elevados tempos de substituição. Grande parte destes equipamentos opera de forma contínua e sua parada não programada por causar acidentes e grandes perdas financeiras, seja por lucro cessante, multas ou danos de sistemas a jusante. Além disso, essas paradas não programadas podem interromper por longos períodos o fornecimento de recursos essenciais à sociedade, como é o caso de falhas em transformadores no sistema de distribuição de energia elétrica – abordado como um dos estudos de caso desta monografia. Em outros casos, falhas podem estar fortemente relacionadas à segurança de equipamentos e instalações, como é o caso de falhas em diferentes tipos de componentes de freio de locomotivas – abordado como segundo estudo de caso desta monografia.

Segundo Gilardoni e Colosimo (2007) - autores da principal referência bibliográfica deste estudo - o que motivou o trabalho inicialmente foi uma grande revisão do setor elétrico brasileiro no final da década de 1990 no Brasil, quando elevadas multas começaram a ser aplicadas em caso de indisponibilidade do sistema e manutenções não programadas. Com esta motivação, foi necessário adotar maior foco em manutenções preventivas do que anteriormente.

De acordo com o texto de apoio do *software Minitab* (2018), a análise de dados de manutenção em sistemas reparáveis permite:

- Obter e analisar a taxa de falhas, bem como interpretar sua tendência de se manter, diminuir ou aumentar com o tempo;
- Determinar o quão frequente um sistema irá requerer manutenção;
- Determinar o número de peças sobressalentes é necessário ter em estoque dado um determinado nível de serviço desejado;
- Determinar se um sistema está com performance em níveis aceitáveis;
- Determinar o provável custo de reparo durante o ciclo de vida do equipamento.

Com estas motivações, se torna de grande valia a utilização de uma metodologia que modele os dados de falhas e de manutenção em sistemas reparáveis para a determinação de uma

janela ótima para realização de manutenções preventivas, considerando a relação entre os custos envolvendo manutenções programadas e não programadas.

Esse trabalho propõe um modelo de Lei de Potência que inclui em um único modelo diferentes grupos que podem representar fornecedores, equipamentos, lotes etc. No contexto de confiabilidade aplicado ao meio industrial o modelo pode permitir a comparação do desempenho entre estes diferentes níveis de uma variável categórica.

## 2. Objetivos

São objetivos deste trabalho:

- Desenvolver e validar metodologia de análise estatística de confiabilidade em sistemas reparáveis utilizando Processos de Lei de Potência; Ajustar parâmetros através de método de máxima verossimilhança através de duas diferentes alternativas de *software* – *Minitab* e R e validar com dados e resultados reais apresentados em Gilardoni e Colosimo (2007) e Oliveira (2010).
- Discutir e apresentar conclusões sobre as análises dos dados utilizados bem como dos conceitos aplicados em formulações teóricas e gráficos.
- Desenvolver uma metodologia para modelagem teórica de dados de falhas de sistemas reparáveis considerando a inclusão de variáveis explicativas na função intensidade do Processo Lei de Potência; Aplicar a metodologia em um banco de dados real.

### **3. Revisão Bibliográfica**

#### **3.1. Confiabilidade**

De acordo com Meeker e Escobar (1998), há uma forte relação entre qualidade e confiabilidade, com consumidores exigindo cada vez mais produtos confiáveis e seguros. Sistemas, veículos, máquinas e dispositivos devem, com alta probabilidade, serem capazes de exercer suas funções em condições normais de uso por um determinado período. Esta probabilidade é a definição mais comum de confiabilidade.

Melhorar a qualidade de serviços e produtos tem como importante etapa a melhoria de confiabilidade e, para isto, métodos quantitativos são necessários para prever e avaliar vários aspectos de performance. Estes métodos quantitativos consistem em resultados de experimentos de laboratório, de materiais e componentes, testes de protótipos, acompanhamento de produção piloto, análises de garantia e rastreamento sistemático de longos prazos de produtos.

Coletar dados de confiabilidade pode trazer vários ganhos como alguns exemplos trazidos por Meeker e Escobar (1998):

- Avaliar características de materiais ao longo do período de garantia;
- Prever confiabilidade de produtos, sistemas e equipamentos;
- Prever custos de garantia em desenvolvimento de produtos;
- Gerar informações importantes para avaliações de risco de falha de sistemas;
- Comparar componentes de dois ou mais fornecedores, materiais, períodos de produção, ambientes de operação etc.;
- Rastrear produtos em aplicação provendo informações sobre causas de falhas e métodos para melhoria de confiabilidade.

##### **3.1.1. Características de dados de Confiabilidade**

Dados em análises de confiabilidade possuem características especiais:

- São tipicamente censurados. Isso é, muitas vezes se faz necessário analisar dados cujos testes têm restrições de tempo de observação ou que são limitados quando determinado

número de falhas é atingido. As censuras são então os procedimentos para retirada de unidades sem falha de um teste em um tempo previsto, no caso truncamento por tempo e o procedimento de interromper um teste ao ser atingido um determinado número de falhas, no caso de truncamento por falhas;

- A maioria dos dados de confiabilidade são modelados utilizando distribuições para variáveis aleatórias (correspondem ao tempo até ocorrência/falha) positivas como exponencial, weibull, gamma e lognormal. Há poucas aplicações da distribuição normal em confiabilidade;

- Inferências envolvendo extrapolação são necessárias com frequência. Ex: Pode ser necessário estimar a proporção de falha após 900 horas de operação baseado em um experimento de 400 horas de duração ou estimar a proporção de falha que ocorre a 50 C de temperatura de operação baseado em experimentos a 85 C;

- É necessário o uso de experiências passadas ou outros julgamentos científicos ou de engenharia para prover informações de entrada na análise de dados para correto processo de tomada de decisão;

- Tipicamente, os parâmetros tradicionais de modelos estatísticos (médias, desvio padrão...) não são de interesse primário. Ao contrário, há um real interesse em métricas de confiabilidade de sistemas ou características de distribuição do tipo falha- tempo (probabilidades de falha, quantis da distribuição de vida – em que instante de tempo determinado percentual de uma amostra já tem falhas observadas, taxa de falhas etc.);

### **3.1.2. Sistemas Reparáveis e Não-Reparáveis**

Meeker e Escobar (1998) colocam ainda que, em análises de confiabilidade, se faz importante distinguir entre dados e modelos das duas seguintes situações:

- O tempo de falha para unidades não reparáveis ou componentes (incluindo dados de componentes não reparáveis dentro de um sistema reparável) ou o tempo até a primeira falha de um sistema (sendo ele reparável ou não);

- Uma sequência de tempos de falha reportados para um sistema reparável. De acordo com Oliveira (2010), estes são sistemas que podem experimentar mais de uma falha ao



longo de sua vida útil tendo como principal característica o fato de que o sistema pode ser restaurado e posto de volta em operação quando ocorre uma falha, sem necessariamente ser substituído.

São exemplos de dados de sistemas não reparáveis: Testes de laboratório para estudar durabilidade, desgaste ou outras propriedades de vida útil de um componente ou material; Dados de operação em campo de sistemas quando informações podem ser obtidas de componentes substituídos mediante falha.

Em sistemas reparáveis, por outro lado, o objetivo é descrever tendências de falhas e padrões para um sistema completo ou uma população de sistemas. Nestes casos, falhas destes sistemas são seguidas de reparos, e os dados consistem em uma sequência de tempos de falha para uma ou mais cópias do sistema em estudo. Quando um único componente ou subsistema em um sistema maior é reparado ou substituído após uma falha, a distribuição de tempo até a próxima falha vai depender do estado geral do sistema no momento imediatamente anterior à ocorrência da falha atual, e da natureza do reparo. Portanto, dados de sistemas reparáveis podem ser descritos através de modelos que permitam mudanças de estado do sistema ao longo do tempo ou que considerem a dependência entre falhas ao longo do tempo.

Oliveira (2010) define que o reparo em que sistemas reparáveis são submetidos quando ocorre uma falha são atividades de manutenção de dois possíveis tipos:

- Manutenção perfeita: Retorna o sistema à condição de um sistema novo e às condições de “tão bom quanto novo”;
- Reparo mínimo: Retorna o sistema à condição em que ele se encontrava no momento imediatamente anterior à falha, restaurando o sistema a um estado de “tão ruim quanto velho”.

### **3.2. Análises de Sistemas reparáveis e outros dados de recorrência**

De acordo com Meeker e Escobar (1998), métodos específicos são necessários à análise de dados de falhas/ eventos recorrentes em que os tempos de recorrências possam não ser estatisticamente independentes. Esses métodos são uteis para quantificar empiricamente a confiabilidade global de um sistema, para monitorar e prever custos de reparo, e para checar

se os tempos entre reparos de componentes individuais em um sistema podem ser tratados como sendo independentes ou não.

Nas análises de confiabilidade em sistemas reparáveis, reparos são tipicamente observados em um intervalo definido  $(t_0 = 0, t_a)$ . Um processo de reparo para um único sistema pode ser visto como uma sequência de tempos de reparo  $T_1, T_2, \dots$ . Nessa modelagem, o termo “reparo” pode ser entendido como falha, substituição, ajuste, manutenção etc. Pressupõe-se, portanto, que sempre que ocorrem falhas, estas são seguidas de ações de manutenção (cujos tempos são desprezíveis) que retornam o sistema ao funcionamento. Geralmente os tempos de reparo são medidos em função da idade do sistema ou tempo de operação, ou horímetros de máquinas.

Em casos em que há inspeções periódicas em tempos definidos, o número de reparos em cada subintervalo de tempo pode ser reportado, mas na maior parte dos casos, as informações de maior interesse são os tempos exatos das falhas.

A coleta de dados pode contemplar apenas um sistema ou dados de vários sistemas. Quase sempre a causa da falha e/ ou o custo de reparo são informações de interesse que devem ser consideradas na coleta de dados. Em casos mais complexos, é necessário também considerar condições de operação, ambientes, fornecedores, modelos e outras variáveis agrupadoras pois estas podem ter influência significativa na modelagem da coleção de sistemas considerado.

Os dados coletados podem ser utilizados para determinarmos:

- A distribuição de tempos entre reparos,  $\tau_j = T_j - T_{j-1}$ ,  $(j=1, 2, \dots)$ ;
- O tempo esperado entre falhas – MTBF (Mean Time Between Failures);
- O número esperado de reparos no intervalo  $(0,t]$  em função de  $t$ ;
- A taxa de reparos em função de  $t$ ;
- Custos médios de reparo em função de  $t$ .

### **3.2.1. Modelo Não-Paramétrico para dados de recorrências**

De acordo com Meeker e Escobar (1998), para um único sistema, dados de recorrências podem ser expressos como  $N(s,t)$  - o número de recorrências no intervalo de idade do sistema  $(s,t]$ . O modelo correspondente, usado para descrever a população de sistemas é baseado na

função média acumulada (Em inglês: *Mean Cumulative Function, ou MCF*) na idade  $t$  do sistema, que é definida como  $\mu(t) = E[N(t)]$ . Assumindo  $\mu(t)$  diferenciável, pode-se expressar a taxa de recorrência por sistema  $\nu(t)$  para a população conforme Equação (1):

$$\nu(t) = \frac{dE[N(t)]}{dt} = \frac{d\mu(t)}{dt}. \quad (1)$$

Meeker e Escobar (1998) afirmam que o modelo não paramétrico requer poucas considerações para sua utilização. São considerações básicas:

- Existência de uma população de funções média acumuladas (uma para cada sistema da população) do qual as amostras são observadas;
- A aleatoriedade na amostra é devida a amostragem de funções médias acumuladas da população de funções médias;
- O tempo em que a observação de um sistema termina não depende do histórico do sistema;
- A estimativa não paramétrica não requer que os sistemas amostrados sejam estatisticamente independentes.

A função média acumulada é uma função escada, com saltos a cada tempo de recorrência, mas com valores constantes entre tempos de recorrência. Um exemplo de visualização gráfica da função acumulada média é apresentado na Figura 1 que representa a função média acumulada para substituições de assentos de válvulas em 41 motores e intervalo de confiança de 95%.

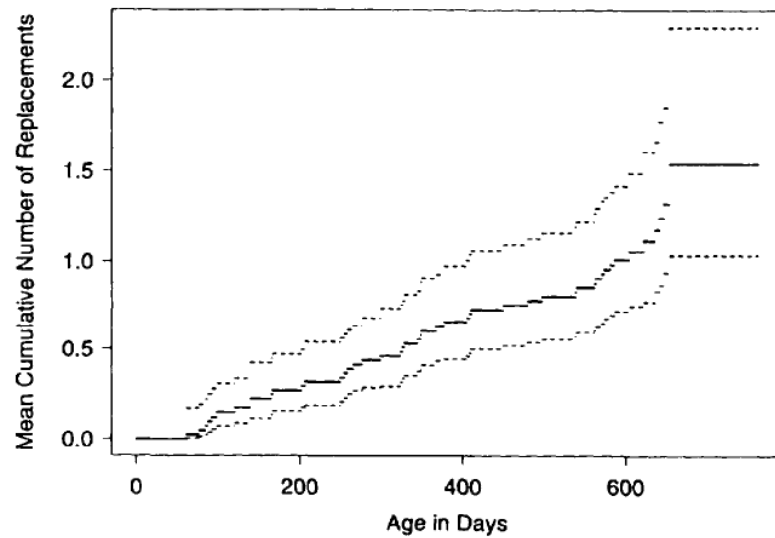


Figura 1 - Estimativa do número médio de substituições de assentos de válvulas para 41 motores e intervalo de confiança de 95%. FONTE: Meeker e Escobar (1998).

De acordo com Nelson (2003), em um determinado tempo  $i$ , uma fração da população de sistemas terá acumulado zero recorrências, outra fração terá acumulado uma recorrência e outra fração terá acumulado duas recorrências etc. Esta distribuição é diferente para sistema e tem uma função média chamada de MCF (*Mean Cumulative Function*) ou FAM (Função Acumulada Média). A Função média acumulada é a média pontual de todas as populações de curvas de todos os sistemas passando por uma linha vertical a cada idade  $i$ , conforme ilustrado na Figura 2. O algoritmo utilizado para cálculo da FAM segundo Nelson é apresentado no Anexo IV – Algoritmo para Cálculo da Função Acumulada Média – Método Não- Paramétrico (Nelson, 2003).

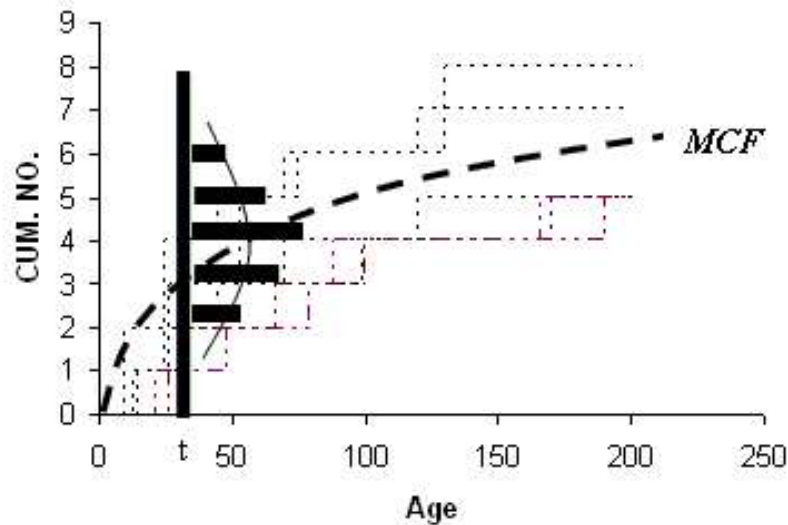


Figura 2 - Ilustração da Função Média Acumulada (FAM, ou MCF em inglês). FONTE: Nelson (2003)

### 3.2.2. Modelos Paramétricos para dados de recorrências

Modelos paramétricos também são utilizados na modelagem de dados recorrentes. Essa modelagem permite estimar número médio e taxa de ocorrência de falhas, conforme descrito por *Minitab* (2018) no texto de ajuda do *software*.

Os modelos paramétricos são utilizados para estabelecer:

- Quão frequente um sistema irá requerer manutenção;
- O número de peças sobressalentes é necessário em estoque para garantir determinado nível de serviço desejado;
- Se um sistema está com performance em níveis aceitáveis;
- Provável custo de reparo durante o ciclo de vida do equipamento.

De acordo com Meeker e Escobar (1998), o modelo mais comum para modelagem de dados de sistemas reparáveis é o processo de Poisson. Pode-se considerar dois tipos de processos de Poisson:

- Processo de Poisson Homogêneo: Utilizado para modelar os tempos de falha que ocorrem a uma taxa de recorrência que permanece constante ao longo do tempo.

- **Processo de Poisson Não Homogêneo:** Utilizado para modelar os tempos de falha que ocorrem a uma taxa que pode estar aumentando, diminuindo. Um exemplo bastante popular de processo de Poisson não homogêneo é o processo de Lei de Potências, que é utilizado neste trabalho. A taxa de recorrência para um processo de Lei de Potência é uma função do tempo.

### 3.2.2.1. Processo de Contagem e Processo de Poisson Homogêneo

De acordo com Colosimo (2007), se definirmos  $T_i$  como os tempos de falha de um determinado sistema, temos que  $N(t)$  é a variável aleatória que denota o número de falhas no intervalo  $[0, t]$ . Então a variável aleatória dada pela Equação (2) é chamada de variável de contagem.

$$N(t) = \sum_{k=1}^{\infty} I(T_k \leq t). \quad (2)$$

A Figura 3 apresenta uma visualização de uma variável de contagem.

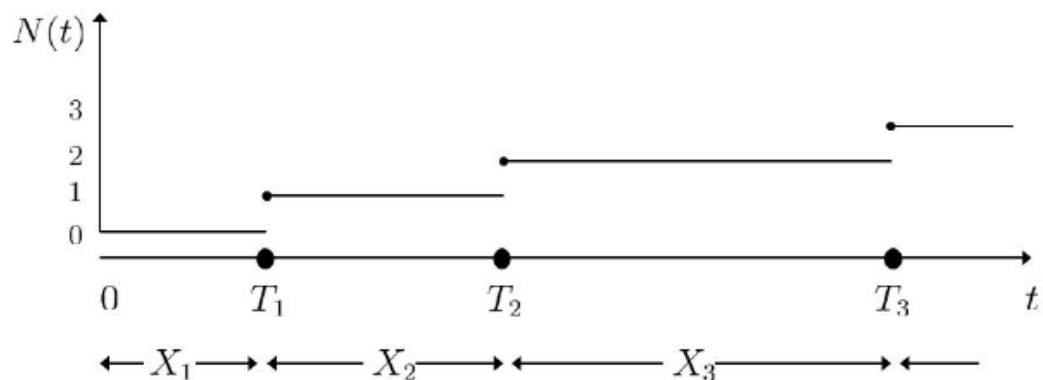


Figura 3 - Visualização gráfica de uma variável de contagem. FONTE: Colosimo (2007)

A função média de um processo de Poisson é definida como sendo esperança de  $N(t)$ , conforme Equação (3).

$$\Lambda(t) = E(N(t)) \quad (3)$$

Esta função representa o número esperado de falhas até o tempo  $t$ .

A função intensidade, interpretada como a taxa instantânea de mudança no número esperado de falhas é expressa, no caso de incrementos independentes pela Equação (4).

$$\lambda(t) = \lim_{\Delta t \rightarrow 0} \frac{P(N(t, t + \Delta t) = 1)}{\Delta t}. \quad (4)$$

De acordo com Colosimo (2007), um processo de contagem  $N(t)$  será um processo de Poisson se:

- $N(0) = 0$ ;
- Para qualquer  $a < b < c < d$ , as variáveis aleatórias  $N(a, b]$  e  $N(c, d]$  são independentes;
- Existe uma função  $\lambda$  descrita conforme (4);
- A probabilidade de falhas simultâneas em um mesmo instante de tempo  $t$  é zero.

O processo de Poisson homogêneo (PPH) é um processo de Poisson com uma função intensidade constante,  $\lambda(t) = \lambda$ ,  $t > 0$ . Os tempos entre falhas são variáveis aleatórias, independentes e identicamente distribuídas que seguem uma distribuição exponencial com média  $= 1/\lambda$ . A função intensidade é constante e este modelo somente tem utilidade prática adequada em sistemas que não apresentam taxa de falhas crescentes ou decrescentes em função de sua idade.

### 3.2.2.2. Processo de Poisson Não Homogêneo e Processo de Lei de Potência

Um processo de Poisson não homogêneo (PPNH) é um processo de Poisson cuja função intensidade não é constante. Um Processo de Poisson Não Homogêneo comum é o processo de Lei de Potência, cuja função intensidade é dada pela Equação (5), conforme indicado por Gilardoni e Colosimo (2007).

$$\lambda(t) = \lim_{\Delta t \rightarrow 0} \frac{P((N(t, t + \Delta t)) \geq 1)}{\Delta t} = \frac{d}{dt} E[N(t)] = \frac{\beta}{\theta} \left[ \frac{t}{\theta} \right]^{\beta-1}, \quad (5)$$

Em que,

$\beta$  é o parâmetro de forma ( $\beta > 0$ ),

$\theta$  é o parâmetro de escala ( $\theta > 0$ ),

$N$  é o número de falhas no intervalo  $(0,t]$ .

A função intensidade representa a taxa de recorrências. O valor do parâmetro de forma determina se o sistema está melhorando sua performance, deteriorando ou se permanece estável.

- Se  $0 < \beta < 1$ , a taxa de recorrências está decaindo. Portanto, seu sistema está melhorando com o passar do tempo;
- Se  $\beta = 1$ , a taxa de recorrências é constante. Portanto, seu sistema permanece estável ao longo do tempo;
- Se  $\beta > 1$ , a taxa de recorrências está aumentando. Portanto seu sistema está deteriorando ao passar do tempo.

A função média da Equação (5) é dada pela Equação (6):

$$\Lambda(t) = E[N(t)] = \int_0^t \lambda(u) du = \left(\frac{t}{\theta}\right)^\beta. \quad (6)$$

### 3.3. Métodos em Inferência Estatística – Método da Máxima Verossimilhança

A estimação dos parâmetros do Processo Lei de Potência é usualmente feita utilizando o Método da Máxima Verossimilhança, descrito em Ribeiro Jr (2012). Oliveira (2010) apresenta o desenvolvimento da função de Log-Verossimilhança da Função Lei de Potência conforme descrito em 3.3.1.

#### 3.3.1. Função de Log-Verossimilhança para o Processo Lei de Potência

Abaixo são desenvolvidos dois cenários, a) para dados de falhas de um único sistema, conforme Equações (7) a (9) e b) para dados de falhas de vários sistemas na Equações (10) a



(12). Em ambos os casos são comentadas as variações para modelos cujos dados possuem censura com truncamentos por tempo ou por falhas.

### (a) Dados de Falhas de um Único Sistema

A função de verossimilhança é expressa na Equação (7).

$$L(\lambda) = \prod_{i=1}^n \lambda(t_i) \cdot \exp \left[ - \int_0^y \lambda(x) dx \right]. \quad (7)$$

Nesta expressão, podem ser analisados dados truncados por falha ou por tempo. Em que  $y$  é o último tempo do intervalo observado dado por:

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

No caso de dados truncados por tempo, todos os sistemas são observados até um mesmo tempo  $T$ . Já em casos de dados truncados por falha, um ou mais sistemas são observados até que sejam registrados  $n$  falhas, cujo tempo de operação é representado por  $t_n$ .

Substituindo-se as Equações desenvolvidas na Seção 3.2.2.2 e (6) na Equação (7) e realizando manipulações algébricas, obtém-se a expressão da verossimilhança, conforme (8).

$$L(\beta, \theta) = \frac{\beta^n}{\theta^{n\beta}} \left[ \prod_{i=1}^n t_i \right]^{\beta-1} \exp \left\{ - \left( \frac{y}{\theta} \right)^\beta \right\}. \quad (8)$$

O cálculo da log- verossimilhança para um único sistema leva então à (9),

$$l(\beta, \theta) = n \log \beta - n\beta \log \theta + (\beta - 1) \sum_{i=1}^n \log t_{ij} - \left( \frac{y}{\theta} \right)^\beta. \quad (9)$$

**(b) Dados de Falhas de um Vários Sistemas**

A função de verossimilhança é expressa na Equação (10).

$$L(\lambda) = \prod_{i=1}^K \prod_{j=1}^{n_i} \lambda(t_{ij}) \cdot \exp \left[ - \sum_{i=1}^K \int_0^{y_i} \lambda(x) dx \right], \quad (10)$$

Substituindo-se as Equações desenvolvidas na Seção 3.2.2.2 e (6) na Equação (10) e realizando manipulações algébricas, obtém-se a expressão da verossimilhança:

$$\begin{aligned} L(\beta, \theta) &= \prod_{i=1}^K \left\{ e^{-\left(\frac{y_i}{\theta}\right)^\beta} \frac{\beta^{n_i}}{\theta^{n_i \beta}} \prod_{j=1}^{n_i} t_{ij}^{\beta-1} \right\} \\ &= \frac{\beta^n}{\theta^{n\beta}} \left[ \prod_{i=1}^K \prod_{j=1}^{n_i} t_{ij} \right]^{\beta-1} \cdot \exp \left\{ - \sum_{i=1}^K \left(\frac{y_i}{\theta}\right)^\beta \right\}, \end{aligned} \quad (11)$$

Em que  $n = \sum_{i=1}^K n_i$ .

O cálculo da log-verossimilhança resultante é apresentado nas Equações (9) e (12) para um único sistema e para vários, respectivamente, e sua maximização para um dado conjunto de observações de falhas de vários sistemas fornece os valores dos parâmetros  $\beta$  e  $\theta$ . Este cálculo não tem solução analítica e necessita ser solucionado por métodos numéricos.

$$l(\beta, \theta) = n \log \beta - n\beta \log \theta + (\beta - 1) \sum_{i=1}^K \sum_{j=1}^{n_i} \log t_{ij} - \sum_{i=1}^K \left(\frac{y_i}{\theta}\right)^\beta. \quad (12)$$

### 3.4. Função de Custos de Manutenção e Determinação do Tempo Ótimo de Manutenção para Sistemas Reparáveis

Em seu artigo, Gilardoni e Colosimo (2007) desenvolvem uma função de custos de manutenção e define o intervalo de tempo  $\tau$  tal que minimiza os custos de manutenção de um equipamento reparável.

Neste modelo, as manutenções preventivas (MP), tem custo  $C_{MP}$ , e retornam o sistema a uma condição exatamente idêntica a um sistema novo. Entre intervalos de manutenção preventiva, podem, contudo, ocorrer falhas e o sistema ser reestabelecido por meio de manutenções corretivas mínimas (MC) que restauram o sistema à uma condição idêntica à do instante que antecede sua falha. As manutenções corretivas têm seu custo esperado definido pela variável  $C_{MC}$ . O modelo tem também como considerações a independência entre custos de reparo e tempos de falhas, e o tempo de reparo é desprezível. Este modelo é expresso matematicamente pela Equação (13).

$$C_{MP} + C_{MC}E(N(\tau)) = C_{MP} + C_{MC} \int_0^{\tau} \rho(u)du, \quad (13)$$

em que,

$C_{MP}$  é o custo de manutenções preventivas em um intervalo de amplitude  $\tau$ ,

$C_{MC}$  é o Custo de execução de manutenções corretivas em um intervalo de amplitude  $\tau$ ,

$E(N(\tau))$  é o número total esperado de manutenções preventivas em um intervalo de amplitude  $\tau$ ,

$\int_0^{\tau} \rho(u)du$  é a função intensidade acumulada que representa a contagem de média de eventos.

O Custo total no intervalo de  $(0, T]$  é então expresso pela Equação (14):

$$C_{(0,T]}(\tau) = m\{C_{MP} + C_{MC} \int_0^{\tau} \rho(u)du\} + R, \quad (14)$$

em que  $m$  é o maior número inteiro menor ou igual à  $T/\tau$  e representa o número total de intervalos,  $R$  é o custo esperado para o intervalo  $(m\tau, T]$ , dado por  $R = C_{MR}E(N(T - m\tau))$ .

É convenientemente definir também a expressão do custo de manutenção por unidade de tempo, dado pela Equação (15),

$$H(\tau) = \lim_{T \rightarrow \infty} C_{(0,T]}(\tau)/T = \frac{1}{\tau} [C_{MP} + C_{MC} \int_0^{\tau} \rho(u) du]. \quad (15)$$

O intervalo de periodicidade de execução de manutenções preventivas que minimiza o custo esperado é resolvido pela Equação (16),

$$\frac{dH}{d\tau}(\tau) = \frac{C_{MC}}{\tau^2} \left[ \tau \rho(\tau) - \int_0^{\tau} \rho(u) du - \frac{C_{MP}}{C_{MC}} \right]. \quad (16)$$

Usando integração por partes temos a Equação (17),

$$\frac{C_{MP}}{C_{MC}} = \tau \rho(\tau) - \int_0^{\tau} \rho(u) du = \int_0^{\tau} u \dot{\rho}(u) du. \quad (17)$$

A Expressão (17) é apresentada na Figura 4. Neste gráfico, a área sombreada deve ser igual ao quociente  $\frac{C_{MP}}{C_{MC}}$ .

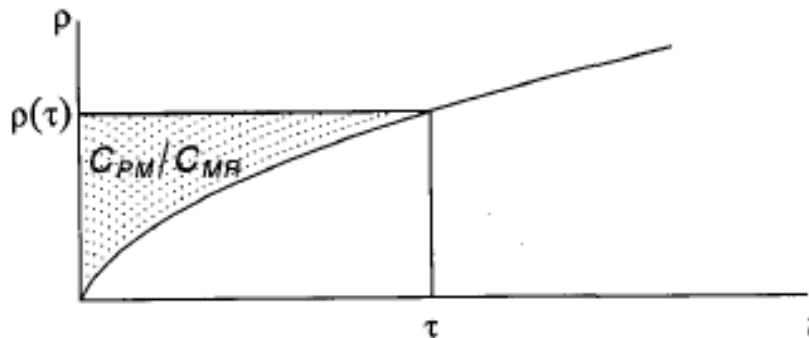


Figura 4 - Relação entre a razão entre custos de manutenções preventivas e corretivas e a função intensidade.  
 FONTE: Gilardoni e Colosimo (2007).

A Equação (17) não tem solução para taxa de falhas não crescente e  $C_{MP} > 0$  pois, neste caso, ações preventivas não trazem melhoria ao sistema. Para sistemas com função intensidade crescente e modelado pela Lei de Potência, e a expressão (17) se torna a equação (18),

$$(\beta - 1) \left( \frac{\tau}{\theta} \right)^\beta = \frac{C_{MP}}{C_{MC}}. \quad (18)$$

Conseqüentemente, pela propriedade da invariância dos estimadores de máxima verossimilhança, pode ser mostrado que o tempo ótimo de manutenção é dado pela Equação (19),

$$\tau = g(\beta, \theta, C_{MP}, C_{MC}) = \theta \left[ \frac{C_{MP}}{(\beta-1)C_{MC}} \right]^{1/\beta}. \quad (19)$$

O intervalo de confiança ao nível 100 (1- $\alpha$ ) % é feita a partir do cálculo da variância, pela aplicação do método delta através da Equação (20), conforme apresentado por Gilardoni e Colosimo (2007). Este estimador corresponde à variância assintótica de  $\hat{\tau}$ .

$$\sigma_{\hat{\tau}}^2 = [\nabla g(\hat{\beta}, \hat{\theta})]' \Sigma [\nabla g(\hat{\beta}, \hat{\theta})], \quad (20)$$

em que

$g(\hat{\beta}, \hat{\theta})$  são os estimadores de máxima verossimilhança dos parâmetros,

$\nabla g(\hat{\beta}, \hat{\theta})$  é o vetor gradiente de  $g$  avaliado em  $(\hat{\beta}, \hat{\theta})$ .

$\Sigma$  é a matriz de covariâncias assintótica de  $\hat{\beta}$  e  $\hat{\theta}$  e pode ser estimada por: menos o inverso da matriz hessiana da função de log-verossimilhança, avaliada em  $\hat{\beta}$  e  $\hat{\theta}$ .

As derivadas parciais de  $\tau$  em função de  $\beta$  e  $\theta$  são obtidas, respectivamente, pelas Equações (21) e (22).

$$\frac{\partial \tau}{\partial \beta} = -\theta \left[ \frac{C_{MP}}{(\beta-1)C_{MC}} \right]^{1/\beta} * \left( \frac{1}{(\beta-1)\beta} + \frac{\log\left(\frac{C_{MP}}{(\beta-1)C_{MC}}\right)}{\beta^2} \right). \quad (21)$$

$$\frac{\partial \tau}{\partial \theta} = \left[ \frac{C_{MP}}{(\beta-1)C_{MC}} \right]^{\frac{1}{\beta}}. \quad (22)$$

A aplicação do método delta leva então à Equação (23).

$$\sigma_{\tau}^2 = [\nabla g(\hat{\beta}, \hat{\theta})]' \Sigma [\nabla g(\hat{\beta}, \hat{\theta})] = \begin{bmatrix} \frac{\partial \tau}{\partial \beta} & \frac{\partial \tau}{\partial \theta} \end{bmatrix} * \begin{bmatrix} \sigma_{\beta}^2 & cov(\beta, \theta) \\ cov(\beta, \theta) & \sigma_{\theta}^2 \end{bmatrix} * \begin{bmatrix} \frac{\partial \tau}{\partial \beta} \\ \frac{\partial \tau}{\partial \theta} \end{bmatrix}. \quad (23)$$

## 4. Metodologia

### 4.1. Modelagem Estatística dos Dados de Falha

Para a modelagem estatística proposta foram utilizados dois bancos de dados e implementação em dois *softwares*: *Minitab* e *R*.

A solução do problema utilizando o *Minitab* utiliza a interface gráfica para importação de dados e obtenção dos resultados de forma simples, uma vez que o *software* já possui este tipo de modelagem implementada para o usuário final. Foi analisada a documentação técnica do *software* e foi possível verificar que os dados são modelados exatamente conforme expressão desenvolvidas nas Seções 3.2 e 3.3. A configuração dos dados é apresentada no “Anexo III – Configuração dos dados no *Minitab*”.

Foi utilizado também o *software R* para implementar na forma de um algoritmo as expressões da log-verossimilhança e a obtenção das estimativas dos parâmetros. O script desenvolvido é apresentado no “Anexo V - Código do script R para Maximização da Log-Verossimilhança da Lei de Potência”.

Após realizadas ambas as implementações e analisados os dados, é realizada uma comparação entre os resultados fornecidos por cada *software*, bem como os resultados apresentados no artigo original de Gilardoni e Colosimo (2007).

### 4.2. Implementação do Modelo de Lei de Potência considerando grupos como covariáveis

Uma aplicação típica para modelagem de dados de tempos de falha de equipamentos é a comparação entre diferentes grupos relacionados ao objeto de estudo. Em manutenção industrial, exemplos de grupos seriam: Fornecedor do componente, Lote, Local de aplicação, Tipo de componente, Tipo de Material entre outros.

Como premissas para desenvolvimento de um modelo que compare dois grupos assume-se que o parâmetro de forma  $\beta$  seja constante independente dos grupos e que o parâmetro de

escala  $\theta$ , que pode ser interpretado como a “velocidade” com que as falhas ocorrem, possa variar e indicar se diferentes componentes têm diferentes taxas de falhas.

Portanto, assume-se um modelo de  $\theta(x_i)$  para dados de falhas que possuam dois grupos, conforme a Equação (24).

$$\theta(x_i) = e^{(\alpha_0 + \alpha_1 x_i)}, \quad (24)$$

em que  $\alpha_0$  e  $\alpha_1 \in \mathbb{R}$  são parâmetros desconhecidos.

$x_i$  é uma variável *dummy* dada por

$$x_i = \begin{cases} 0, & \text{para componentes do grupo de referência;} \\ 1, & \text{para componentes do grupo alternativo.} \end{cases}$$

A função intensidade apresentada na Equação (5) considerando um único parâmetro de forma pode então ser reescrita na Equação (25),

$$\lambda_i(t, x_i) = \frac{\beta}{\theta(x)} \left[ \frac{t}{\theta(x_i)} \right]^{\beta-1} = \frac{\beta}{e^{(\alpha_0 + \alpha_1 x_i)}} \left[ \frac{t}{e^{(\alpha_0 + \alpha_1 x_i)}} \right]^{\beta-1}. \quad (25)$$

Consequentemente, a função média assumirá a forma dada pela Equação (26),

$$\Lambda(t, x) = E[N(t)] = \int_0^t \lambda(u) du = \left( \frac{t}{\theta(x_i)} \right)^\beta = \left( \frac{t}{e^{(\alpha_0 + \alpha_1 x_i)}} \right)^\beta. \quad (26)$$

A expressão de verossimilhança apresentada anteriormente na Equação (11) é reescrita na Equação (27).

$$\begin{aligned} L(\beta, \alpha_0, \alpha_1, x) &= \prod_{i=1}^K \left\{ e^{-\left(\frac{y_i}{\theta(x)}\right)^\beta} \frac{\beta^{n_i}}{\theta(x_i)^{n_i \beta}} \prod_{j=1}^{n_i} t_{ij}^{\beta-1} \right\} \\ &= \frac{\beta^n}{\theta(x_i)^{n\beta}} \left[ \prod_{i=1}^K \prod_{j=1}^{n_i} t_{ij} \right]^{\beta-1} * \exp \left\{ - \sum_{i=1}^K \left( \frac{y_i}{\theta(x)} \right)^\beta \right\}, \end{aligned} \quad (27)$$



em que  $n = \sum_{i=1}^K n_i$ .

O cálculo da log-verossimilhança resulta na Equação (28) e sua maximização para um dado conjunto de observações de falhas de vários sistemas fornece os valores dos parâmetros  $\beta$  e  $\theta$ .

$$l(\beta, \alpha_0, \alpha_1, x) = n \log \beta - \beta \sum_{i=1}^K \sum_{j=1}^{n_i} \log \theta(x_{ij}) + (\beta - 1) \sum_{i=1}^K \sum_{j=1}^{n_i} \log t_{ij} - \sum_{i=1}^K \left( \frac{y_i}{\theta(x_i^*)} \right)^\beta. \quad (28)$$

Substituindo a Equação (24) na Equação (28), obtém-se a Equação (29).

$$l(\beta, \alpha_0, \alpha_1, x) = n \log \beta - n\beta \alpha_1 + \alpha_2 \sum_{i=1}^K \sum_{j=1}^{n_i} x_{ij} + (\beta - 1) \sum_{i=1}^K \sum_{j=1}^{n_i} \log t_{ij} - \sum_{i=1}^K \left( \frac{y_i}{e^{(\alpha_0 + \alpha_1 x_i^*)}} \right)^\beta. \quad (29)$$

A adequação da Equação (19) para o tempo ótimo de manutenções é então obtido pela Equação (30) e deve ser calculado de forma independente para cada grupo através da utilização da variável *dummy*.

$$\tau(\beta, \alpha_0, \alpha_1, x) = \theta(x) \left[ \frac{C_{MP}}{(\beta-1)C_{MC}} \right]^{\frac{1}{\beta}} = e^{(\alpha_0 + \alpha_1 x)} \left[ \frac{C_{MP}}{(\beta-1)C_{MC}} \right]^{\frac{1}{\beta}}. \quad (30)$$

Para realizar os cálculos dos desvios padrão, é necessário novamente a aplicação do método delta, conforme Equação (20).

A derivadas parciais de  $\tau = g(\beta, \alpha_0, \alpha_1, x)$  com respeito a  $\beta, \alpha_0, \alpha_1$  são, respectivamente dadas pelas Equações (31), (32) e (33).

$$\frac{\partial \tau}{\partial \beta} = -e^{(\alpha_1 + \alpha_2 x)} \left[ \frac{C_{PM}}{(\beta-1)C_{MR}} \right]^{\frac{1}{\beta}} * \left( \frac{1}{(\beta-1)\beta} + \frac{\log\left(\frac{C_{PM}}{(\beta-1)C_{MR}}\right)}{\beta^2} \right). \quad (31)$$

$$\frac{\partial \tau}{\partial \alpha_0} = e^{(\alpha_0 + \alpha_1 x)} \left[ \frac{C_{PM}}{(\beta - 1)C_{MR}} \right]^{\frac{1}{\beta}}. \quad (32)$$

$$\frac{\partial \tau}{\partial \alpha_1} = x e^{(\alpha_0 + \alpha_1 x)} \left[ \frac{C_{PM}}{(\beta - 1)C_{MR}} \right]^{\frac{1}{\beta}}. \quad (33)$$

A aplicação do método delta leva então à Equação (34).

$$\sigma_\tau^2 = [\nabla g(\hat{\beta}, \hat{\theta})]' \Sigma [\nabla g(\hat{\beta}, \hat{\theta})] = \left[ \frac{\partial \tau}{\partial \beta}, \frac{\partial \tau}{\partial \alpha_0}, \frac{\partial \tau}{\partial \alpha_1} \right] * \begin{bmatrix} \sigma_\beta^2 & cov(\beta, \alpha_0) & cov(\beta, \alpha_1) \\ cov(\beta, \alpha_0) & \sigma_{\alpha_0}^2 & cov(\alpha_0, \alpha_1) \\ cov(\beta, \alpha_1) & cov(\alpha_0, \alpha_1) & \sigma_{\alpha_1}^2 \end{bmatrix} * \begin{bmatrix} \frac{\partial \tau}{\partial \beta} \\ \frac{\partial \tau}{\partial \alpha_0} \\ \frac{\partial \tau}{\partial \alpha_1} \end{bmatrix}. \quad (34)$$

## 5. Resultados Numéricos

### 5.1. Dados de falhas de transformadores – Lei de Potência

#### 5.1.1. Descrição dos dados

Gilardoni e Colosimo (2007) apresentam em seu artigo o estudo de caso feito com dados de falha de 30 transformadores – componentes básicos de um sistema de transmissão de eletricidade - entre 300 e 345 kV pertencentes a uma companhia de distribuição de energia elétrica. Nesses dados, observados entre janeiro de 1999 e julho de 2001:

- Foram observadas 21 falhas;
- Dez transformadores não apresentaram falhas e foram censurados após 21888 horas de acompanhamento;
- Foram registrados neste período a ocorrência de 11 manutenções preventivas – censura pela premissa de manutenção que retorna os transformadores ao estado de “tão bom quanto novo”;
- O custo de manutenção corretiva após falha é 15 vezes maior que o custo das manutenções preventivas.

No artigo, são considerados dois tipos de manutenção:

- a. Corretivas e aquelas em que apenas uma pequena parte de suas peças constituintes é substituída – Neste caso, é assumido que a confiabilidade do equipamento após o reparo é a mesma daquela imediatamente antes da falha;
- b. Preventivas: Manutenções realizadas sem que tivesse ocorrido uma falha precedente e que se assume um escopo de revisão geral do equipamento – Neste caso, é assumido que a confiabilidade do equipamento após o reparo é a mesma de um equipamento novo.

Os dados de falhas dos transformadores são apresentados no Anexo I – Dados de e podem também ser visualizados na Figura 5 em que os transformadores analisados são apresentados ao longo de todas as 21888 horas do período de observação das falhas e manutenções.

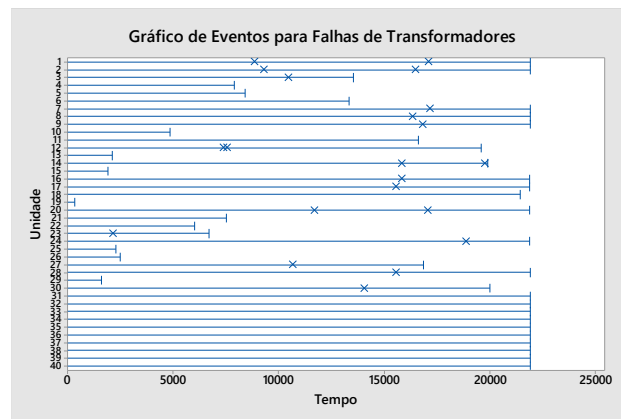


Figura 5 - Gráfico de Eventos de falhas em Transformadores

### 5.1.2. Modelagem utilizando *Software Minitab*

Os comandos utilizados e resultados completos da modelagem no *software Minitab* são apresentados no Anexo III – Configuração dos dados no *Minitab*.

A Função Média Acumulada para as falhas é apresentada na Figura 6, bem como as estimativas de máxima verossimilhança dos parâmetros de forma e escala associado ao modelo Lei de Potência. A Figura 7 apresenta também a função média acumulada de falhas.

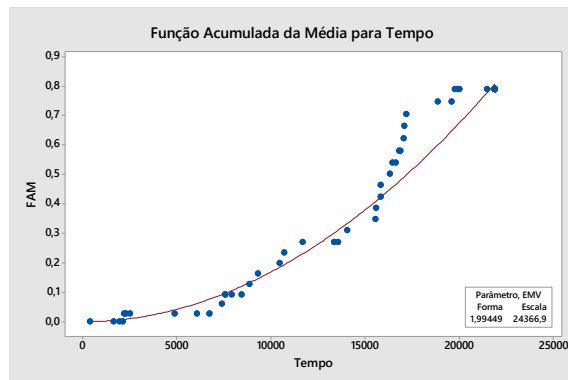


Figura 6 - Função acumulada da média para falhas de transformadores.

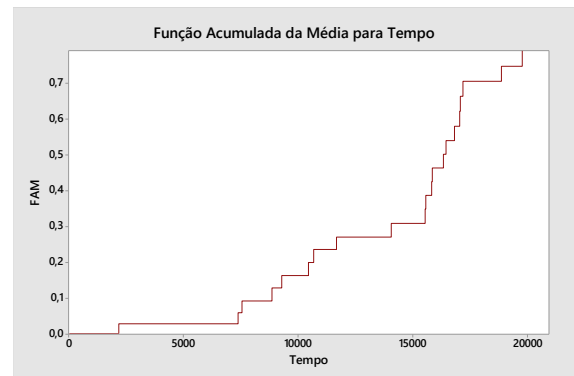


Figura 7 - Visualização sem o ajuste do modelo paramétrico para a função acumulada de média para falha de transformadores.

As estimativas dos parâmetros obtidos pelo método da máxima verossimilhança são apresentadas na Tabela 1.

Tabela 1 - Estimativa de parâmetros do modelo de Lei de Potência

Parâmetro	Estimativa	Erro Padrão	IC (95%)	
			Inferior	Superior
<b>Forma</b>	1,99449	0,400	1,34651	2,95431
<b>Escala</b>	24366,9	2798,891	19454,8	30519,2

### 5.1.3. Modelagem utilizando o *Software R*

A modelagem do problema no *software R* é apresentada de forma completa através do script no Anexo V - Código do script R para Maximização da Log-Verossimilhança da Lei de Potência.

Os modelos paramétricos e não paramétricos são apresentados nas Figura 8 e Figura 9, respectivamente.

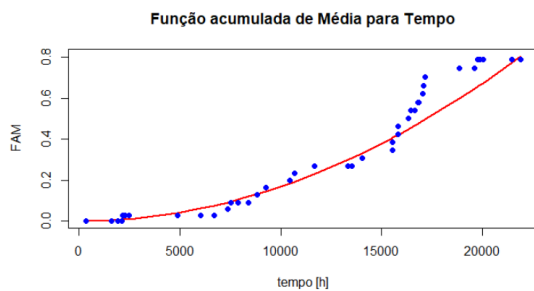


Figura 8 - Função acumulada da média para falhas de transformadores.

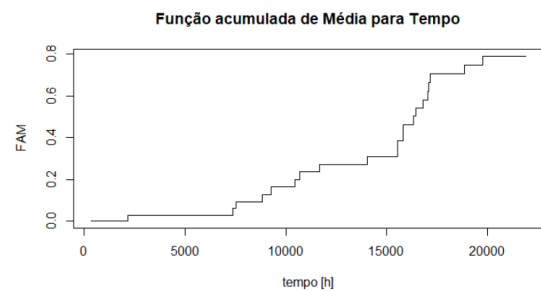


Figura 9 - Visualização sem o ajuste do modelo não paramétrico para a função acumulada de média para falha de transformadores.

As estimativas dos parâmetros obtidos pelo método da máxima verossimilhança são apresentadas na Tabela 2.

Tabela 2 - Estimativas de parâmetros do modelo de Lei de Potência

Parâmetro	Estimativa	Erro Padrão	IC (95%)	
			Inferior	Superior
<b>Forma</b>	1,99449	0,4032054	1,204211	2,784776
<b>Escala</b>	24366,9	3046,465	18395,85	30337,99

#### 5.1.4. Análise Comparativa dos Resultados - Validação

Os resultados obtidos nesta análise são comparados com aqueles obtidos originalmente pelo autor do artigo na Tabela 3.

Tabela 3 - Comparação dos valores estimados dos parâmetros do Modelo de Lei de Potência

Método/ Software	Parâmetro de forma ( $\beta$ )	Parâmetro de escala ( $\theta$ )	Tempo ótimo para Manutenções Preventivas ( $\tau$ )
<b>Gilardoni e Colosimo (2007)</b>	1,9880 [1,2020; 2,7740]	24844,0 [19016,7; 30671,3]	6400 [3021; 9780]
<b>Oliveira (2010)</b>	1,9945* <i>OBS1</i>	24366,9* <i>OBS1</i>	6285 [4870; 7701]
<b>Minitab</b>	1,9945 [1,3465; 2,9543]	24366,9 [19545,8; 30519,2]	6285* <i>OBS2</i>
<b>R</b>	1,9945 [1,2042; 2,7848]	24366,9 [19395,9; 30338,0]	6285 [4786; 7785]

*OBS1: Intervalos de confiança não apresentados na bibliografia;*

*OBS2:  $\tau$  calculado analiticamente a partir da expressão (19) e não fornecido diretamente pelo Minitab.*

Observou-se uma diferença de -0,3269 % no parâmetro de forma, de -1,9204% no parâmetro de escala e de 1,7904% no tempo ótimo para manutenções preventivas em comparação aos resultados apresentados por Gilardoni e Colosimo (2007). Os resultados da modelagem nos dois *softwares* utilizados foram exatamente iguais.

Na Figura 10 e Figura 11 são visualizados os gráficos da função média acumulada de falhas em função do tempo no artigo de referência de Gilardoni e Colosimo (2007) e nos 2

*softwares* em que o modelo foi desenvolvido neste trabalho. O gráfico da Modelagem Paramétrica não está presente no trabalho de Gilardoni e Colosimo (2007) e foi obtido através de Oliveira (2010) em trabalho que utilizou os mesmos dados de falha dos transformadores.

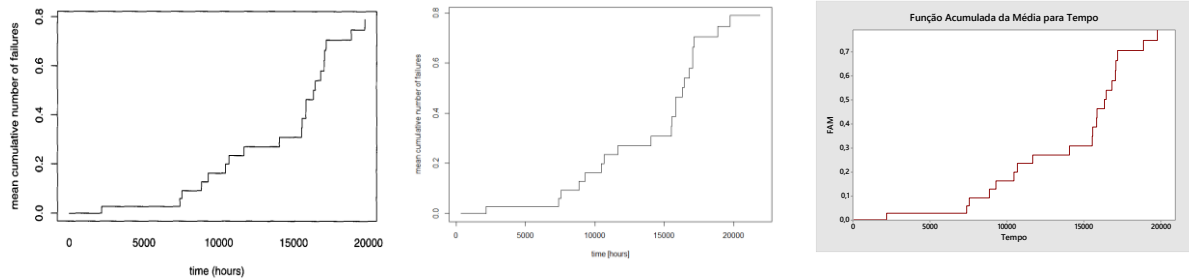


Figura 10 - Comparação da Modelagem Não Paramétrica

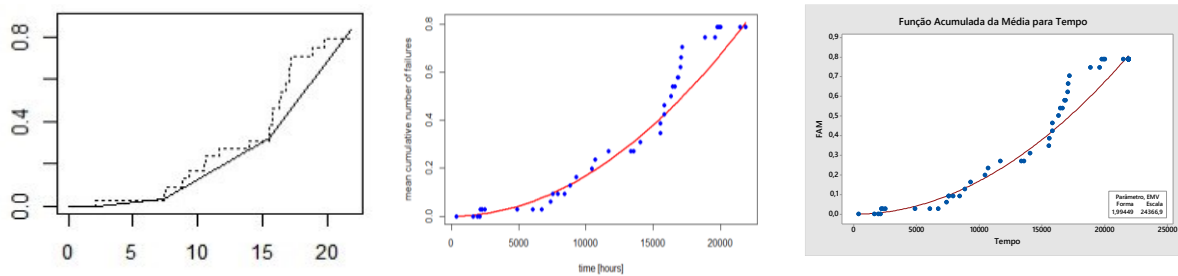


Figura 11 - Comparação da Modelagem Paramétrica

Uma visualização com a comparação sobreposta dos modelos ajustados pelo neste trabalho e por Gilardoni e Colosimo (2007) é apresentado na Figura 12. O gráfico reflete pouca diferença, assim como os resultados numéricos das diferenças, apresentados no início desta Seção. Observa-se, contudo, que há um maior distanciamento do modelo paramétrico em relação ao não paramétrico a partir eixo x de tempos próximos a 16000 horas de operação.

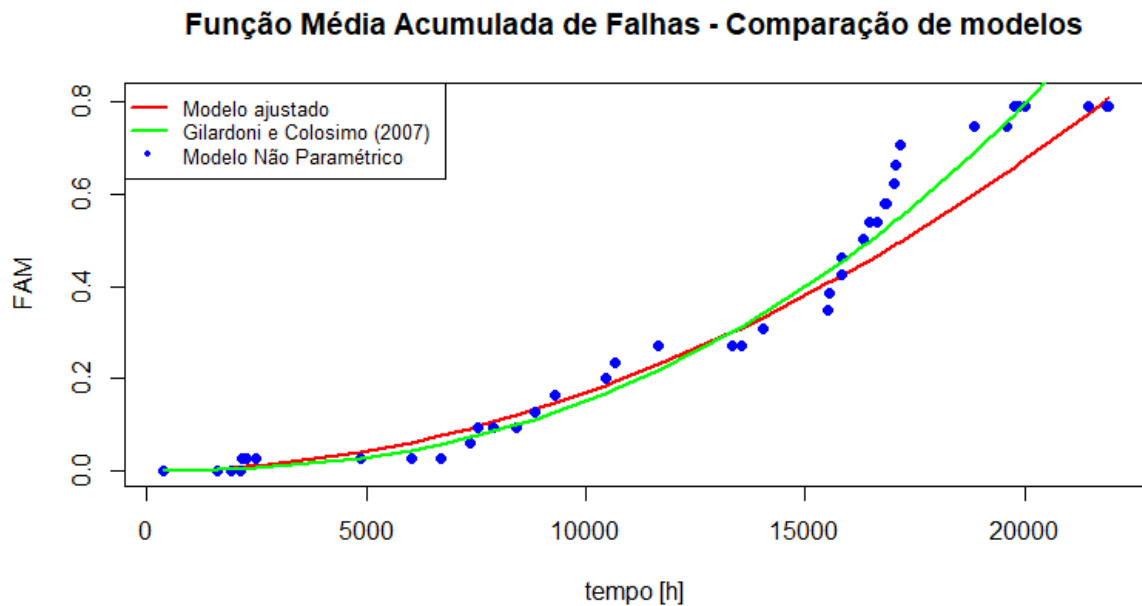


Figura 12 - Comparação dos modelos paramétricos entre si e entre o modelo não paramétrico.

## 5.2. Dados de falhas em Componentes de freio de Locomotivas – Lei de Potência com covariáveis

### 5.2.1. Descrição dos dados

O banco de dados utilizado para aplicação do modelo com covariáveis apresenta o momento de reposição de componentes de freio de 29 trens. A cada vez que uma unidade falhava, esta era reparada e retornava à operação. Além disso, os dados não contêm informações de manutenções preventivas que pudessem ser assumidas como *reset* dos componentes a um status de “tão bom quanto novo”, dessa forma, a modelagem dos dados deve considerar truncamento por falhas apenas. Os dados foram obtidos de um exemplo de aplicação do *software Minitab* e são apresentados na íntegra no Anexo II – Dados de Falhas de Componentes de Freio de Locomotivas.



## 5.2.2. Modelagem dos dados considerando Lei de Potência sem Covariáveis no *Software Minitab*

### Modelo sem Distinção de tipos de freios

Uma primeira análise dos dados considera o Processo de Lei de Potência pelo método da Máxima Verossimilhança sem a distinção da variável que identifica o tipo de componente de freio aplicado. Os resultados do modelo, ajustado pelo *software Minitab*, são apresentados nas Figura 13 e Figura 14 e na Tabela 4, em que os intervalos de confiança dos parâmetros ajustados são exibidos.

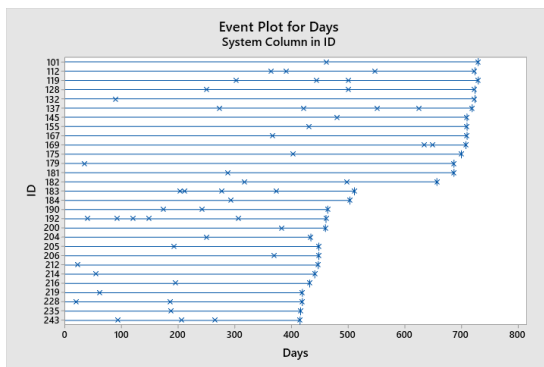


Figura 13 - Gráfico de eventos de falhas

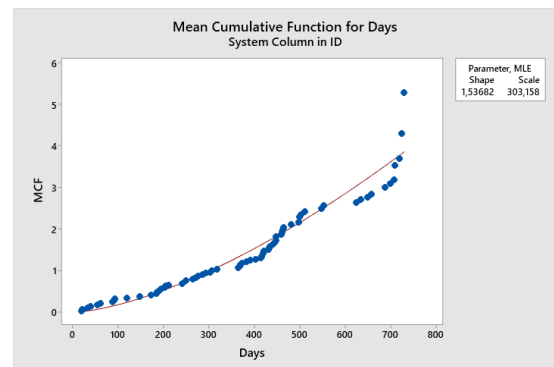


Figura 14 - Função média acumulada sem distinção de tipos de freio

Uma análise visual do Gráfico da função média acumulada de falhas deste modelo, apresentado na Figura 14, permite verificar que há um ajuste de menor qualidade para tempos de operação maiores.

Tabela 4 - Estimativa dos parâmetros da Lei de Potência sem considerar tipo de freios como covariáveis

### Parameter Estimates

Parameter	Estimate	Standard 95% Normal CI		
		Error	Lower	Upper
Shape	1,53682	0,145	1,27691	1,84963
Scale	303,158	32,131	246,293	373,151

A aplicação da Equação (19) permite estimar o tempo ótimo para manutenções preventivas. O resultado para esta modelagem seria de 78,02 dias, assumindo  $CPM/CMR = 1/15$ . Estes valores de custo para estes dados de freio de locomotivas não é conhecido e assume-se, a título de ilustração, a mesma relação dos dados dos transformadores.

Modelo com Distinção de tipos de freios

A segunda análise feita utilizou as informações do tipo de componentes de freio. Na Figura 15 é possível visualizar os eventos de falhas para componentes de freios de locomotivas do tipo 1 e do tipo 2. A Figura 16 apresenta os modelos de Lei de Potência para os dados isolados de componentes de freio de cada tipo.

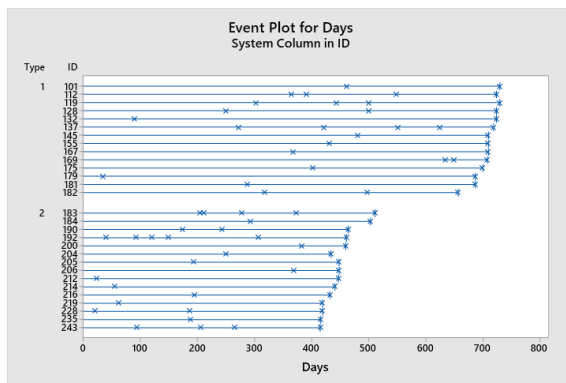


Figura 15 - Gráfico de eventos de falhas considerando tipos de componentes de freio.

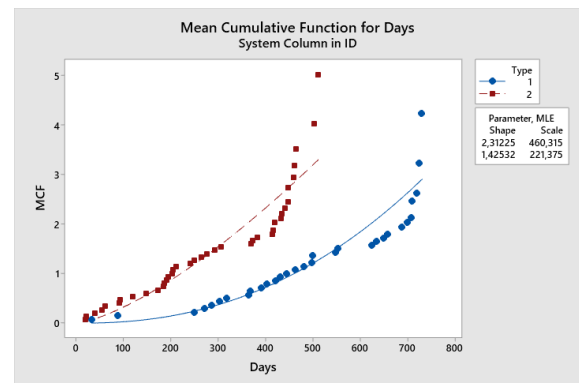


Figura 16 - Função Média Acumulada considerando tipos de componentes de freio.

Os resultados apresentados na Tabela 5 fornecem os parâmetros ajustados do modelo para componentes de freio do tipo 1 e seu intervalo de confiança. A Tabela 6 apresenta os mesmos resultados para os componentes de freio do tipo 2.

Tabela 5 - Estimativas de parâmetros da Lei de Potência para o Tipo 1.

Parameter Estimate	Parameter Estimate	Standard 95% Normal CI		
		Error	Lower	Upper
Shape	2,31225	0,318	1,76564	3,02807
Scale	460,315	46,625	377,431	561,400

Tabela 6 - Teste de igualdade do parâmetro de forma para o Tipo 2.

Parameter Estimate	Parameter Estimate	Standard 95% Normal CI		
		Error	Lower	Upper
Shape	1,42532	0,185	1,10451	1,83932
Scale	221,375	35,210	162,085	302,351

### 5.2.3. Modelagem dos dados considerando Lei de Potência com Covariáveis no *Software R*

O *software R* foi utilizado para desenvolver um código adaptado daquele desenvolvido em 5.1 para calcular os tempos ótimos de manutenção considerando um parâmetro de forma  $\beta$  único compartilhado entre os dois tipos de freios de locomotivas. Esta consideração foi assumida a partir dos resultados apresentados em 5.2.2 em que o pode ser observado pela Tabela 5 e pela Tabela 6 que os intervalos de confiança dos parâmetros de forma  $\beta$  para componentes de freio do Tipo 1 e Tipo 2 possuem intervalos coincidentes, não havendo assim evidências para se rejeitar a hipótese de igualdade entre os valores de  $\beta$  para cada subconjunto dos dados.

A Figura 17 indica o resultado do modelo de Lei de Potência com parâmetros determinados pela maximização da verossimilhança conforme modelagem presente na Seção 3.3. Na Figura 18 apresenta-se o modelo de Lei de Potência ajustado para os dois tipos de componentes de freio de locomotivas considerando o modelo em que são consideradas as covariáveis conforme modelagem apresentada na seção 4.2.

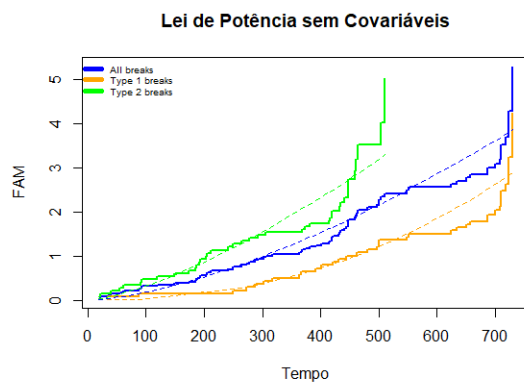


Figura 17 - Ajuste da Lei de Potência sem covariáveis para o conjunto de dados completo e para os componentes de freio de forma independente.

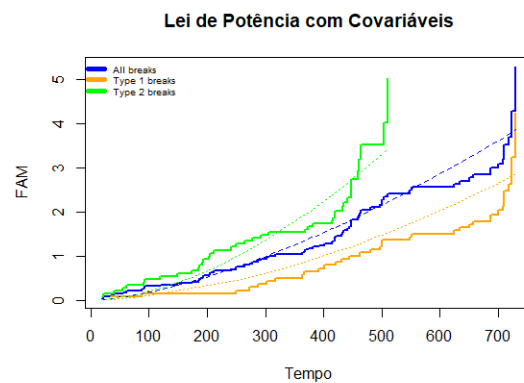


Figura 18 - Ajuste da Lei de Potência com covariáveis para cada tipo de componente de freio e para o conjunto de dados completo.

Os intervalos de confiança para o tempo ótimo de manutenções preventivas foram também calculados aplicando o método delta, conforme Equações (23) e (34) e são apresentados na 5.1.4. Os resultados dos parâmetros ajustados dos modelos são apresentados na Tabela 7 para

o modelo de Lei de Potência sem covariáveis e na Tabela 8 para o modelo de Lei de Potência com covariáveis.

Tabela 7 - Parâmetros ajustados utilização maximização da verossimilhança e Lei de Potência sem covariáveis

Freio	Parâmetro	Estimativa	Erro Padrão	IC (95%)	
				Inferior	Superior
Tipo 1	$\beta$	2,31225	0,37437	1,57848	3,04602
	$\theta$	460,3152	45,63993	370,86094	549,76946
Tipo 2	$\beta$	1,42532	0,22170	0,99079	1,859852
	$\theta$	221,37439	34,3965	153,95725	288,79153

Tabela 8 - Parâmetros ajustados utilizando maximização da verossimilhança da Lei de Potência com covariáveis

Freio	Parâmetro	Estimativa	Erro Padrão	IC (95%)	
				Inferior	Superior
Tipo 1	$\beta$	1,74637	0,19575	1,3823	2,13004
	$\theta$	400,11972	45,17221	311,5822	488,6572
Tipo 2	$\theta$	252,1893	27,86963	197,5849	306,8138

#### 5.2.4. Análise Comparativa dos Resultados aplicando os dois modelos de Lei de Potência

Na seção 5.2.3 foram apresentados os resultados gráficos dos cálculos implementados no *software* R em que foram definidos os modelos de Lei de Potência para cada tipo de componente de freio de locomotivas. A Figura 19 apresenta os modelos ajustados por ambas as metodologias sobrepostos em uma mesma visualização.

### Gráfico de Comparação de Modelos

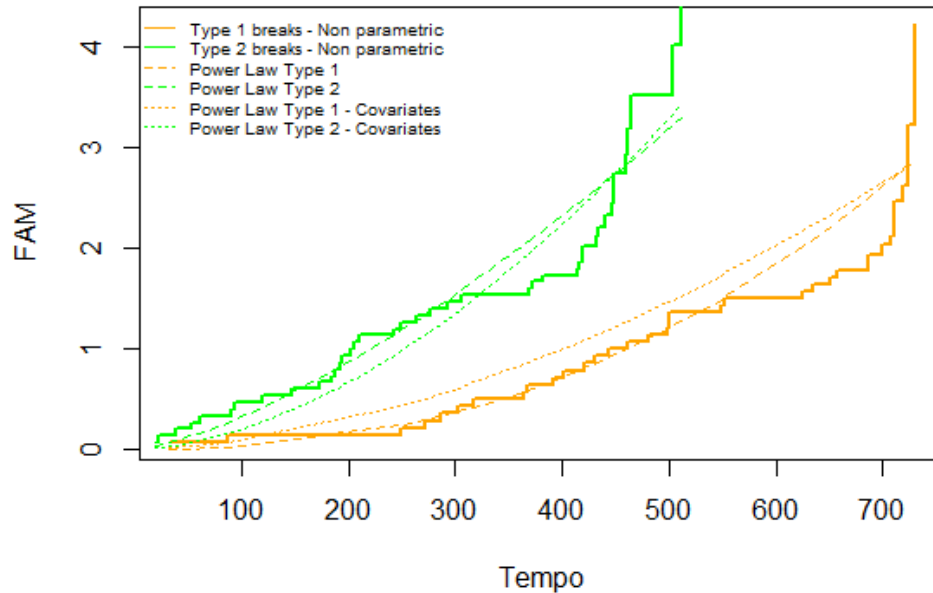


Figura 19 - Modelos de Lei de Potência com e sem Covariáveis

A Tabela 9 apresenta os resultados de tempo ótimo de manutenções empregando as metodologias apresentadas nas Seções 3.3.1 e 4.2, isto é, o modelo de Lei de Potência padrão e aquele desenvolvido considerando as covariáveis de tipo de componentes de freio.

Tabela 9 - Comparação dos resultados entre metodologias utilizados para Cálculo do Tempo Ótimo para manutenções preventivas

Método	$\tau_0$ (Tempo ótimo de manutenções preventivas para o grupo referência)	$\tau_0$ (Tempo ótimo de manutenções preventivas para o grupo alternativo)
1. Modelos de Lei de Potência independentes ( <i>Minitab</i> )	126,88	60,32
	IC 95% = [84,49; 169,26] Amplitude IC= 84,77	IC 95% = [45,92; 74,73] Amplitude IC=28,81
2. Modelo Único com parâmetro de forma compartilhado	100,34	63,25
	IC 95% = [77,58; 123,11] Amplitude IC= 45,52	IC 95% = [49,18; 77,31] Amplitude IC= 28,12

## 6. Discussão

Os resultados apresentados na seção 5 foram divididos em duas etapas. Na primeira delas o Processo Lei de Potência foi aplicado utilizando um banco de dados conhecido de falhas em transformadores elétricos com a finalidade de comparação e validação dos resultados numéricos do modelo. Foi desenvolvida a metodologia no *software* R e validado com os resultados obtidos no *software Minitab* e com os resultados presentes na bibliografia. Houve uma pequena diferença nos desvios padrões dos parâmetros do modelo originados provavelmente por pequenas diferenças de métodos de cálculo e ajustes nos bancos de dados.

Quanto aos resultados numéricos dos dados de falhas dos transformadores, conforme discutido na Seção 0, o parâmetro de forma indica se o sistema está melhorando, piorando ou permanece estável. Uma vez que o valor do parâmetro de forma encontrado foi de 1,9945, a conclusão é de que a taxa de falha está aumentando e conseqüentemente os transformadores estão deteriorando ao longo do tempo. Tal conclusão é reforçada ao analisar o intervalo de confiança de 95% do parâmetro de forma e os gráficos de função média acumulada de falhas por tempo apresentados na Seção 0.

Uma avaliação visual do modelo gerado através do método da máxima verossimilhança, avaliado através dos gráficos apresentados na Seção 5, cujos modelos não paramétrico e paramétrico são sobrepostos, permite afirmar que foi possível atingir uma boa aderência do modelo aos dados observados. Uma análise mais detalhada permite dizer que o modelo se ajusta melhor para tempos de operação de até 16000 horas, aproximadamente. Depois deste período, o modelo subestima a taxa de falhas real dos sistemas avaliados, uma vez que a curva do modelo se situa abaixo da função média não paramétrica.

Na segunda parte do trabalho, os modelos de Lei de Potência desenvolvidos na Seção 4.2, considerando covariáveis nominais e um parâmetro de forma constante, foram aplicados aos dados de falhas de componentes de freio de locomotivas. A aplicação do Processo de Lei de Potência se mostrou adequado uma vez que o parâmetro de forma foi maior que 1, indicando degradação dos sistemas justificando a abordagem de determinação de tempos ótimos de manutenção preventiva.

Neste modelo desenvolvido, a consideração de parâmetro de forma comum foi testada a partir de um teste de hipóteses e pode ser verificada também pelos intervalos de confiança sobrepostos de  $\beta$  para os componentes do tipo 1 e tipo 2. Conforme esperado, o desvio padrão de todos os parâmetros estimados foi menor aplicando o modelo com covariáveis, o que pode ser justificado principalmente pela redução do número de parâmetro do modelo que passa de 4 ( $\beta$  e  $\theta$  para o tipo 1 e  $\beta$  e  $\theta$  para o tipo 2) para 3 ( $\beta, \alpha_0, \alpha_1$ ) e pela utilização dos dados em sua totalidade para o ajuste do modelo, enquanto o modelo padrão ajusta 2 modelos independentes através do particionamento dos dados. Nos modelos em que foi empregada a validação comparando a modelagem nos dois *softwares*, observou-se pequenas diferenças nos intervalos de confiança dos parâmetros de forma e de escala em cada *software*, tal diferença pode ser justificada provavelmente por diferentes métodos computacionais empregados pelo *Minitab* e pela função `optim()` utilizada no R e estimação pela matriz hessiana.

Os tempos ótimos de manutenção também foram determinados e observou-se redução do desvio padrão dessa estimativa para o grupo referência em 46,3%. A Análise visual através da Figura 19 também permite verificar a adequação dos modelos desta metodologia em comparação aos modelos não paramétricos. Observa-se, contudo, que há um maior distanciamento deste modelo quando comparado ao método não paramétrico de referência em relação ao modelo sem covariáveis. Além disso, o modelo para o grupo referência aparentemente superestima a função média acumulada de falhas na maior parte do intervalo de tempo analisado.

## 7. Conclusão

Foi apresentado uma revisão de análise de sistemas reparáveis e seu contexto dentro da área de conhecimento de confiabilidade na Seção 3, bem como seus modelos matemáticos utilizados. Nessa seção, foram apresentados o Processo de Poisson não homogêneo, o método de máxima verossimilhança, método delta e a formulação para cálculo dos tempos ótimos de manutenções preventivas considerando premissas e considerações, sendo a principal delas a de que a manutenção preventiva reestabelece o sistema a um estado de “tão bom quanto novo” e a manutenção corretiva reestabelece o sistema para um estado de “tão ruim quanto velho”.

Na Seção 4, foi desenvolvida a metodologia de análise de dados de falhas e de recuperações para sistemas reparáveis utilizando diferentes métodos e *softwares* disponíveis – um *software* com interface gráfica e ampla aplicação em indústria (*Minitab*) e um que utiliza a linguagem R para programação e resolução de problemas por meio do *software R Studio*. A Aplicação do modelo de lei de potência permite interpretar várias tendências dos sistemas avaliados, como taxa de falhas e previsibilidade de comportamento futuro – informações de grande utilidade na tomada de decisão em cenário de gestão de manutenção industrial. Além disso, é aplicada análise de otimização de tempo de manutenção preventiva considerando os custos de manutenção.

Na Seção 4 também se apresentou um aprofundamento teórico no modelo ao propor a modelagem da Lei de Potência considerando covariáveis nominais no modelo. Essa abordagem assume igualdade dos parâmetros de forma e o emprego de um parâmetro de escala como uma função de uma variável nominal. Uma adaptação do modelo computacional desenvolvido previamente foi realizada e implementada no *software R Studio*. A mesma implementação não foi possível no *software Minitab* devido a este ser um modelo adaptado nas modelagens de sistemas reparáveis.

Na Seção 5, as teorias e modelos desenvolvidos nas seções 3 e 4 foram aplicados a dois bancos de dados reais. Na primeira aplicação, objetivou-se a modelagem de dados de falhas de transformadores elétricos e a determinação do seu tempo ótimo de manutenções preventivas. Considera-se que houve sucesso na validação da implementação do modelo no *Minitab* atingindo os mesmos resultados de parâmetros do modelo daqueles obtidos por Oliveira (2010).



No *software* R foi também possível o desenvolvimento do código e replicar os dados dos modelos não paramétricos e paramétricos para a função média acumulada de falhas.

Na segunda parte da Seção 5, um banco de dados de falhas em componentes de freio de locomotivas foi utilizado para aplicação do modelo de Lei de Potência com covariáveis. Estes dados tiveram uma amostragem truncada por falhas, enquanto os dados de falhas em transformadores eram truncados por tempo. Neste caso, foi possível adequar o código desenvolvido no *software* R para implementar o novo modelo e compará-lo com a modelagem de Lei de Potência padrão. Os resultados numéricos obtidos tanto para parâmetros do modelo quanto para tempo ótimo para manutenções preventivas foram comparados entre modelos com e sem covariáveis.

Na Seção 6 os resultados numéricos e toda a modelagem foram discutidos. Como principais pontos a partir de todo o trabalho destaca-se:

- Foi realizado uma revisão bibliográfica revisando a importância de modelos matemáticos estatísticos e sua implementação computacional em sistemas reparáveis e seu contexto em confiabilidade;
- Foi possível replicar os resultados apresentados na bibliografia para dados de falhas em transformadores elétricos em um *software* comercial e interface gráfica. Como grande ganho didático, destaca-se a validação do mesmo modelo obtido a partir do desenvolvimento de um código na linguagem R;
- Através de aprofundamento teórico e conceitual, foi possível propor uma alteração no modelo de Lei de Potência para que este considerasse covariáveis nominais que pudessem representar grupos, tipos, lotes, fabricantes etc.;
- O modelo teórico para Lei de Potência com covariáveis foi implementado no *software* R através de adaptação do código desenvolvido. Os resultados deste modelo foram comparados com o modelo de Lei de Potência padrão implementado tanto no *software* R quanto no *Minitab*.
- O *software* R através do desenvolvimento de códigos customizados permitiu a implementação de modelo não convencional e não disponível no *software* comercial *Minitab*, demonstrando assim grande importância científica na implantação de novos conceitos.

- O Processo Lei de Potência com covariáveis apresentou como vantagem a diminuição da complexidade do modelo e a utilização de todos os dados em um equacionamento único. Essas vantagens se traduziram em uma redução dos desvios padrões dos parâmetros estimados e também do desvio padrão do tempo ótimo de manutenções preventivas.
- Em análise visual, verificou-se que houve certo distanciamento de um dos grupos de tipo de freios do modelo não paramétrico. Este fenômeno teria de ser investigado com maiores detalhes neste e em outros bancos de dados. Observou-se também que, em análise visual, há menor aderência do modelo com covariáveis do que o modelo sem covariáveis quando comparados aos modelos não paramétricos de referência. Este fato pode sugerir que a premissa de um  $\beta$  único seja uma consideração não suficientemente adequada a aplicação do Processo Lei de Potência.

## 8. Trabalhos Futuros

A conclusão desse trabalho abre oportunidade para diversos outros estudos e aprofundamentos os quais podem ser exemplificados:

- Análise da adequação dos modelos ajustados;
- Aplicação em um banco de dados real atual e ainda não apresentado em nenhuma bibliografia;
- Aplicação do modelo com covariáveis para casos mais gerais com mais níveis de uma mesma covariável e/ ou outros tipos de covariáveis (contínuas, por exemplo);
- Aplicação de outras funções intensidade além da Função Lei de Potência.

## 9. Referências Bibliográficas

E. A. Colosimo, Métodos Estatísticos em Confiabilidade de Sistemas Reparáveis e Não Reparáveis, Notas de Aula, 2012.

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

*Minitab*, *Minitab* 18 Support Documentation, *Minitab* LLC, 2019.

Nelson, Wayne, Recurrent Events Data Analysis for Product Repairs, Disease Recurrences, and Other Applications, ASA-SIAM, 2003.

Oliveira, Maristela Dias de Modelos Estatísticos em Processos de Poisson Não Homogêneos: Aplicação em Sistemas Reparáveis. Departamento Estatística, Instituto de Ciências Exatas, Universidade Federal Minas Gerais, 2010.

Ribeiro Jr, Paulo Justiniano et al, MINICURSO Métodos Computacionais em Inferência Estatística. Laboratório de Estatística e Geoinformação (LEG) - Departamento de Estatística, Universidade Federal do Paraná (UFPR), 2012.

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

## Anexo I – Dados de Falhas de Transformadores Elétricos

Uma vez que as manutenções preventivas são revisões completas e se assume que a confiabilidade do equipamento retorne “tão bom quanto novo”, essas manutenções geraram registros de outras unidades de análise – identificadas na coluna “Unit” como sistemas diferentes. A coluna “MP” assume valores 0 para registros de tempo que não foram censurados por manutenção preventiva e de 1 caso contrário. A coluna “Tempo” indica o tempo em horas para ocorrência de falhas, manutenções ou censuras. Por último, a coluna “Cens” indica para valor 1 informações não censuradas e 0 informações censuradas, seja por tempo ou por manutenção preventiva.

<b>Trans</b>	<b>Unit</b>	<b>MP</b>	<b>Tempo</b>	<b>Cens</b>	<b>Trans</b>	<b>Unit</b>	<b>MP</b>	<b>Tempo</b>	<b>Cens</b>
1	1	0	8839	1	13	18	1	21440	0
1	1	0	17057	1	13	19	0	369	0
1	1	0	21887	0	14	20	0	11664	1
2	2	0	9280	1	14	20	0	17031	1
2	2	0	16442	1	14	20	0	21857	0
2	2	0	21887	0	15	21	1	7544	0
3	3	0	10445	1	15	22	1	6039	0
3	3	1	13533	0	15	23	0	2168	1
3	4	0	7902	0	15	23	0	6698	0
4	5	1	8414	0	16	24	0	18840	1
4	6	0	13331	0	16	24	0	21879	0
5	7	0	17156	1	17	25	1	2288	0
5	7	0	21887	0	17	26	1	2499	0
6	8	0	16305	1	18	27	0	10668	1
6	8	0	21887	0	18	27	0	16838	0
7	9	0	16802	1	19	28	0	15550	1
7	9	0	21887	0	19	28	0	21887	0
8	10	1	4881	0	20	29	1	1616	0
8	11	0	16625	0	20	30	0	14041	1
9	12	0	7369	1	20	30	0	20004	0
9	12	0	7541	1	21	31	0	21888	0
9	12	1	19590	0	22	32	0	21888	0
9	13	0	2121	0	23	33	0	21888	0
10	14	0	15821	1	24	34	0	21888	0
10	14	0	19746	1	25	35	0	21888	0
10	14	1	19877	0	26	36	0	21888	0
10	15	0	1927	0	27	37	0	21888	0
11	16	0	15813	1	28	38	0	21888	0
11	16	0	21886	0	29	39	0	21888	0

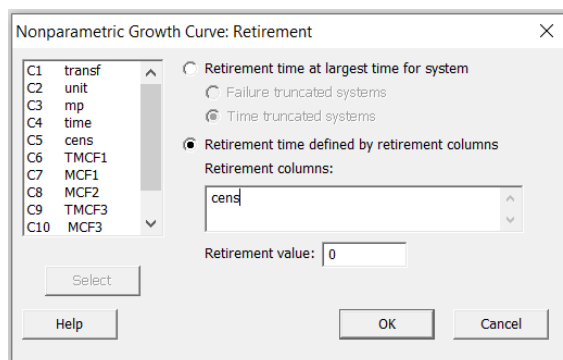
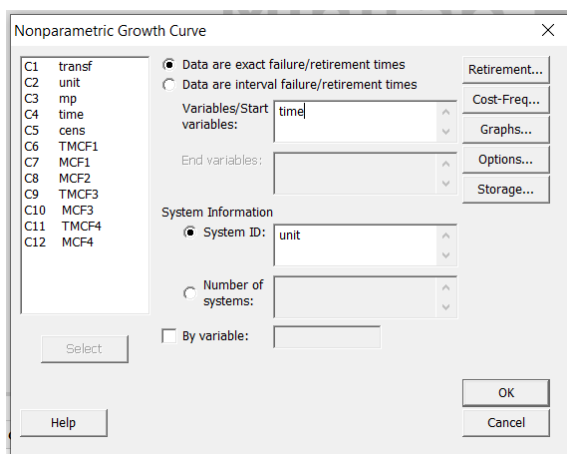
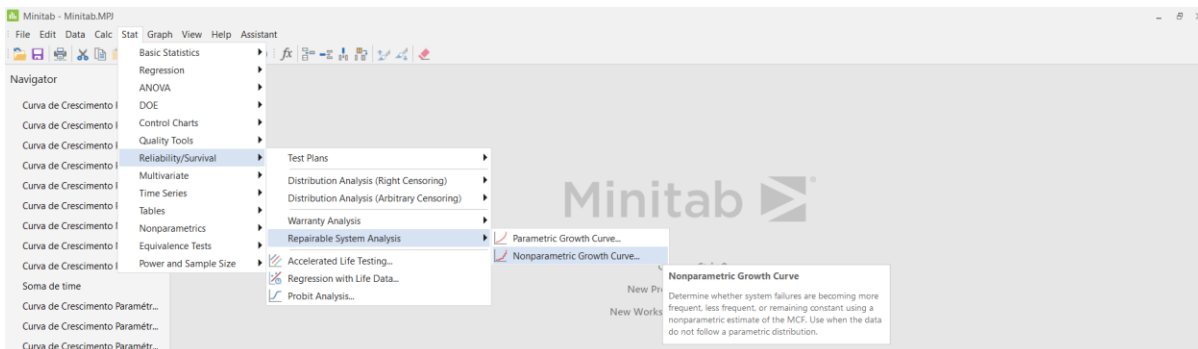
<b>12</b>	17	0	15524	1	30	40	0	21888	0
<b>12</b>	17	0	21886	0					

## Anexo II – Dados de Falhas de Componentes de Freio de Locomotivas

System	Days	Type	x	cens	System	Days	Type	x	cens
101	462	1	0	0	183	277	2	1	0
101	730	1	0	1	183	373	2	1	0
112	364	1	0	0	183	511	2	1	1
112	391	1	0	0	184	293	2	1	0
112	548	1	0	0	184	503	2	1	1
112	724	1	0	1	190	173	2	1	0
119	302	1	0	0	190	242	2	1	0
119	444	1	0	0	190	464	2	1	1
119	500	1	0	0	192	39	2	1	0
119	730	1	0	1	192	91	2	1	0
128	250	1	0	0	192	119	2	1	0
128	500	1	0	0	192	148	2	1	0
128	724	1	0	1	192	306	2	1	0
132	88	1	0	0	192	461	2	1	1
132	724	1	0	1	200	382	2	1	0
137	272	1	0	0	200	460	2	1	1
137	421	1	0	0	204	250	2	1	0
137	552	1	0	0	204	434	2	1	1
137	625	1	0	0	205	192	2	1	0
137	719	1	0	1	205	448	2	1	1
145	481	1	0	0	206	369	2	1	0
145	710	1	0	1	206	448	2	1	1
155	431	1	0	0	212	22	2	1	0
155	710	1	0	1	212	447	2	1	1
167	367	1	0	0	214	54	2	1	0
167	710	1	0	1	214	441	2	1	1
169	635	1	0	0	216	194	2	1	0
169	650	1	0	0	216	432	2	1	1
169	708	1	0	1	219	61	2	1	0
175	402	1	0	0	219	419	2	1	1
175	700	1	0	1	228	19	2	1	0
179	33	1	0	0	228	185	2	1	0
179	687	1	0	1	228	419	2	1	1
181	287	1	0	0	235	187	2	1	0
181	687	1	0	1	235	416	2	1	1
182	317	1	0	0	243	93	2	1	0
182	498	1	0	0	243	205	2	1	0
182	657	1	0	1	243	264	2	1	0
183	203	2	1	0	243	415	2	1	1
183	211	2	1	0					

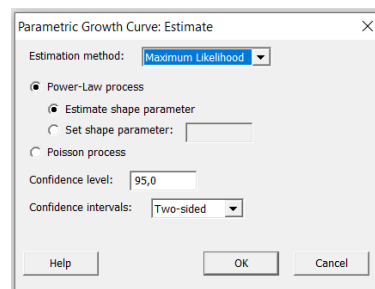
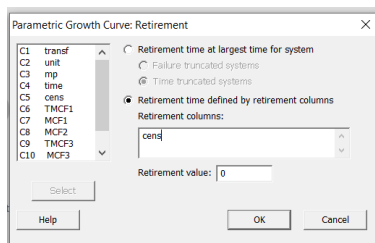
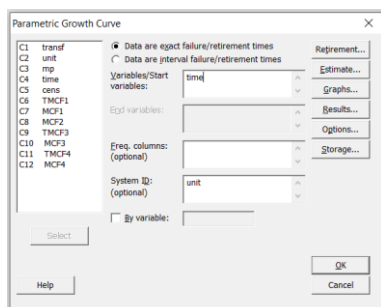
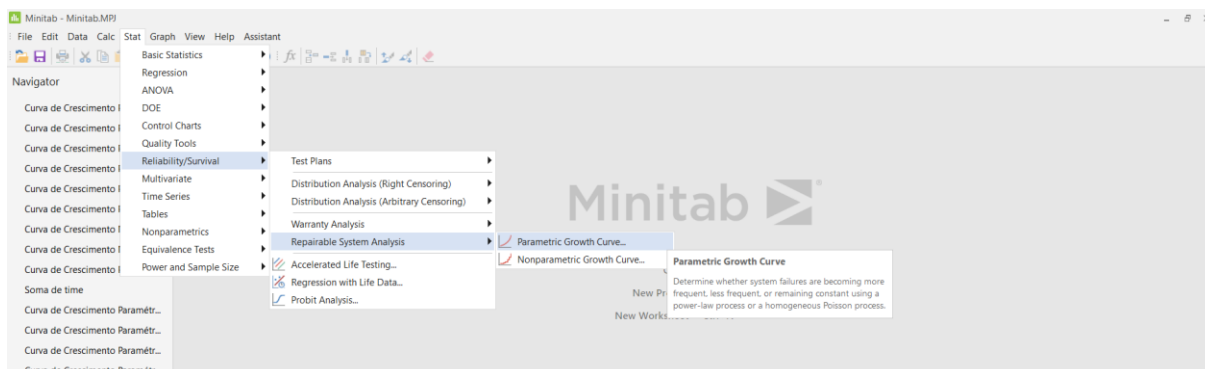
## Anexo III – Configuração dos dados no *Minitab*

### A) Configurações do e Resultados do Modelo Não Paramétrico





## B) Configurações do e Resultados do Modelo Paramétrico



### C) Resultados dos Modelos Não Paramétrico e Paramétrico

TRANSFORMADORES

#### Nonparametric Growth Curve: time

System: unit  
Retirement value: cens = 0  
Nonparametric Estimates

Table of Mean Cumulative Function

Time	Mean Cumulative Standard Error 95% Normal CI				System
	Function	Error	Lower	Upper	
2168	0.027778	0.027389	0.004022	0.19186	23
7369	0.060036	0.041919	0.015278	0.23591	12
7541	0.092294	0.069125	0.021264	0.40058	12
8839	0.128008	0.076443	0.039712	0.41262	1
9280	0.163722	0.082570	0.060928	0.43994	2
10445	0.199437	0.087756	0.084190	0.47244	3
10668	0.235151	0.092158	0.109082	0.50692	27
11664	0.270865	0.095885	0.135341	0.54210	20
14041	0.309327	0.100014	0.164134	0.58296	30
15524	0.347788	0.103431	0.194165	0.62296	17
15550	0.386250	0.106204	0.225331	0.66209	28
15813	0.424711	0.108382	0.257558	0.70035	16
15821	0.463173	0.110002	0.290794	0.73774	14
16305	0.501635	0.111087	0.325004	0.77426	8
16442	0.540096	0.123344	0.345206	0.84501	2
16802	0.580096	0.123388	0.382338	0.88014	9
17031	0.621763	0.134888	0.406407	0.95124	20
17057	0.663429	0.144984	0.432291	1.01815	1
17156	0.705096	0.143962	0.472559	1.05206	7
18840	0.746763	0.142425	0.513852	1.08524	24
19746	0.790241	0.152425	0.541473	1.15330	14

TRANSFORMADORES

#### Parametric Growth Curve: time

System: unit  
Retirement value: cens = 0  
Model: Power-Law Process  
Estimation Method: Maximum Likelihood

#### Parameter Estimates

Parameter Estimate	Standard 95% Normal CI			
	Error	Lower	Upper	
Shape	1,99449	0,400	1,34651	2,95431
Scale	24366,9	2798,891	19454,8	30519,2

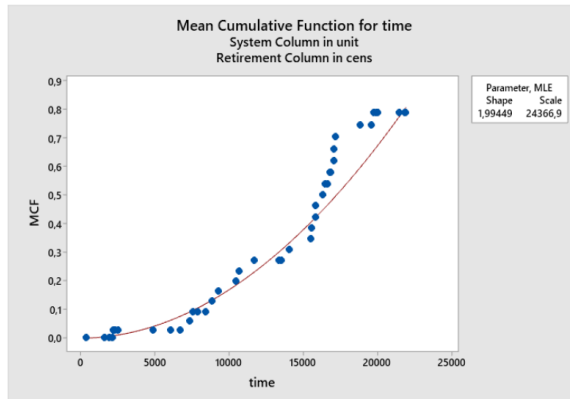
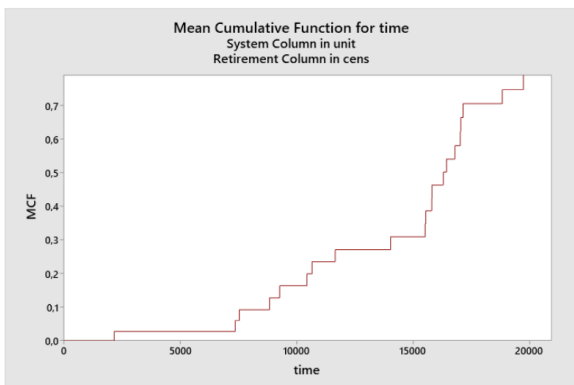
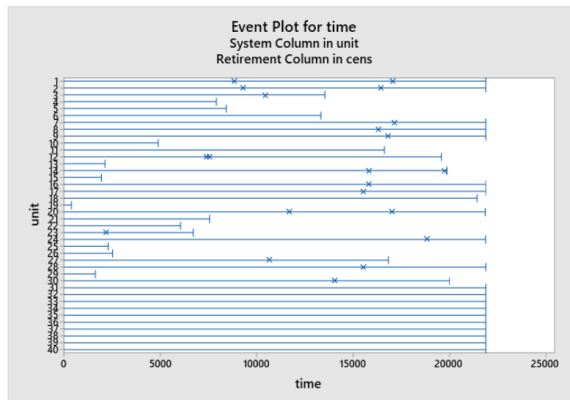
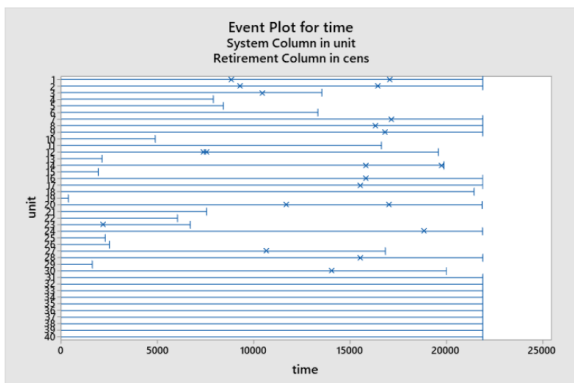
#### Test for Equal Shape Parameters

Bartlett's Modified Likelihood Ratio Chi-Square

\* NOTE \* Test skipped - There must be at least one other failure which is not at the time the system is retired: Check unit = 4.

#### Trend Tests

Test Statistic	MIL-Hdbk-189	Laplace's		Anderson-Darling	
	TTT-based Pooled	TTT-based Pooled	Anderson-Darling	Anderson-Darling	Anderson-Darling
P-Value	0,002	0,002	0,006	0,006	0,004
DF	42	42			



## Anexo IV – Algoritmo para Cálculo da Função Acumulada Média – Método Não- Paramétrico (Nelson, 2003)

### Algoritmo de Cálculo da Função Acumulada Média, de acordo com Nelson (2003):

1. Ordenar tempos de falhas: Todos os dados de falha e censura são ordenados do maior para o menor.
2. Calcular o número r observado:

$$r_i = r_{i-1}, \text{ se } i \text{ é um tempo de falha,}$$

$$r_i = r_{i-1} - 1, \text{ se } i \text{ é um tempo de censura.}$$

Em que,

r é o número de sistemas em risco,

i são os índices dos tempos de falha ou censura de todos os sistemas observados,

N é o número total de sistemas,

$r_0 = N$  no primeiro tempo observado (falha ou censura)

3. Calcular a Função Média Acumulada (Ou *MCF*, *Mean Cumulative Function*) através da Equação abaixo.

$$MCF_i = \frac{1}{r_i} + MCF_{i-1} \text{ e}$$

$$MCF_0 = \frac{1}{r_0} \text{ no primeiro evento observado.}$$

## Anexo V - Código do script R para Maximização da Log-Verossimilhança da Lei de Potência

### A) Análise dos Dados de Falhas dos Transformadores Elétricos: “Power Law\_script.R”

```

### Script para Obtenção de parâmetros da Lei de Potência através do Método da
### Máxima Verossimilhança para dados de Falhas de Sistemas reparáveis ###

# Desenvolvido por: Robert Bento Florentino (38) 99943-9356
# Trabalho completo intitulado: "Determinação do Período Ótimo de Manutenção em Sistemas Re
paráveis através de Processo de Poisson Não Homogêneo com Parâmetros de Modelo de Lei de Po
tência determinados pelo Método da Máxima Verossimilhança - Metodologia e Estudo de Caso"
# Trabalho desenvolvido como Trabalho de Conclusão de Curso da Especialização em Estatístic
a com Ênfase em Indústria e Mercados da UFMG
# Orientador do trabalho: Prof. Fábio Nogueira Demarqui

rm(list=ls(all=TRUE)) #Remove os objetos salvos no ambiente
library(tidyverse) #Lê pacote tidyverse

#### Leitura dos dados ####

dados <- read.csv("Dados/transformadores.txt", header=TRUE, sep="") #Lê os dados dos tran
sformadores
summary(dados) #Sumário dos dados
sapply(dados, class) #Classe dos dados

#Função da Log-verossimilhança para vários sistemas

#Definição da Função de Log-verossimilhança da Lei de Potência para vários sistemas

llpowersystem <- function(par, t, T, n){
  beta <- par[1]
  theta <- par[2]
  ll <- n*log(beta) - n*beta*log(theta) +(beta-1)*sum(log(t)) - sum(exp( beta*(log(T) - 1
og(theta)))) #Função da Log-verossimilhança para Lei de Potência com vários sistemas
  return(-ll)
}

### Execução da Log-verossimilhança para a Lei de Potência para vários sistemas ###

#Dados dos tempos de falhas (cens=1)

t <- dados %>%
  filter(cens==1) %>%
  select(time)
t <- t$time

#Dados dos tempos de censura (cens=0)

T <- dados %>%
  filter(cens==0) %>%
  select(time)
T <- T$time

```

```

#Número total de falhas

n <- length(t)

# Definição dos parâmetros de forma e de escala pela maximização da Log-verossimilhança da
Lei de Potência

# Utilização da função Optim
saida <- optim(par = c(1,1), fn = llpowersystem,
              t=t, T=T, n=n, method="L-BFGS-B",
              lower = 1e-16, upper = Inf,
              hessian = TRUE)

saida$par #Parâmetros de escala e de forma obtidos

#Cálculo do desvio padrão dos parâmetro do modelo

V <- solve(saida$hessian) #Matriz inversa da matriz hessiana
sqrt(V[1,1])
sqrt(V[2,2])

#Intervalo de confiança 95% para Beta
saida$par[1]-1.96*sqrt(V[1,1])
saida$par[1]+1.96*sqrt(V[1,1])

#Intervalo de confiança 95% para Theta
saida$par[2]-1.96*sqrt(V[2,2])
saida$par[2]+1.96*sqrt(V[2,2])

#Aplicação do método delta para cálculo do desvio padrão de Tau - tempo ótimo de manutenções
preventivas
#var = vetor de derivadas parciais * Inversa da Hessiana * vetor de derivadas parciais

derivadas <- 0 #Criação do vetor

#Parâmetro estimados pelo método da máxima verossimilhança
beta <- saida$par[1] #Atribuição dos resultados ao parâmetro beta
theta <- saida$par[2] #Atribuição dos resultados ao parâmetro theta

#Definição de variáveis auxiliares para cálculo das derivadas parciais
CPM <- 1
CMR <- 15
A <- CPM/CMR
B <- 1/(beta-1)
C <- A*B
D <- C ^ (1/beta)

#Derivadas parciais
derivadas[1] <- theta*D*(-B/beta-((log(C))/(beta^2))) #Derivada de tau com respeito a beta
derivadas[2] <- D #Derivada de tau com respeito a theta

#Intervalo de confiança 95% para Tau - tempo ótimo para manutenções preventivas
lowerbound <- D*theta -1.96*sqrt(derivadas %*% V %*% derivadas)
upperbound <- D*theta +1.96*sqrt(derivadas %*% V %*% derivadas)

lowerbound

```

```

upperbound

### Gráfico 3D do ponto mínimo da função Log-Verossimilhança ###

x <- seq(1.04,5,by=0.04) #Eixo x do gráfico, correspondente aos valores do parâmetro beta
y <- seq(10200,30000, by=200) #Eixo y do gráfico, correspondente aos valores do parâmetro
theta
z<-0 #Criação da variável z: Log-Verossimilhança
z <- as.data.frame(z) #Transformação de z em data.frame

#Loop for que calcula os valores de log-verossimilhança para cada combinação de x e y.
for (i in 1:100){
  for (j in 1:100) {
    z[i,j] <- llpowersystem(par=c(x[i],y[j]),t=t,T=T,n=n)
  }
}

z <- as.matrix(z) #Transforma z em uma matriz

#Função do base R que faz gráfico 3D. Parâmetros theta e phi rotacionam a visualização.
persp(x=x,y=y,z=z, ylim=c(10000,30000),
      main = "Log-Verossimilhança Lei de Potência - Transformadores",
      theta = 360,phi=0, r=1,
      col="blue",ticktype = "detailed",
      xlab = "Beta",
      ylab="Theta",
      zlab = "Log-Verossimilhança")

### Cálculos da Função Média Acumulada de Falhas no Tempo NÃO-PARAMÉTRICO ###

tempos_ordenados <- (dados %>% arrange(-desc(time)))$time #Cria um vetor com os tempos d
e operação ordenados do menor ao maior

mcf_minitab <- dados %>% arrange(-desc(time)) #Cria um dataframe com todos os dados dos t
ransformadores ordenados do menor ao maior tempo de operação

#Loop que implementa o algoritmo de cálculo de MCF, conforme NELSON 2003:

for (i in 1:nrow(mcf_minitab)){
  mcf_minitab$total_falhas[i] <- sum(mcf_minitab$cens[1:i]) #Coluna criada no dataframe q
ue armazena a quantidade acumulada de falhas de todos os sistemas ao longo do tempo
  mcf_minitab$total_sistemas[i] <- max(mcf_minitab$unit) - nrow(mcf_minitab[1:i,] %>% f
ilter(cens==0)) #Coluna criada no dataframe que armazena a quantidade total de sistemas em
operação ao longo do tempo
  if (i == 1) { #0 primeiro registro da tabela de transformadores é um dado de censura, p
ortanto, é atribuído 0 à primeira linha
    mcf_minitab$mcf[i] = 0}
  else if (mcf_minitab$cens[i]==1){ #Se o registro for de uma falha, Calcula-se 1/Total d
e sistemas naquele tempo de operação somado ao resultado do mesmo calculo da linha anterior
    mcf_minitab$mcf[i] <- 1/ mcf_minitab$total_sistemas[i] + mcf_minitab$mcf[i-1]
  }
  else if (mcf_minitab$cens[i]==0){ #Caso o registro seja de uma censura, o valor de mcf
será igual ao valor da linha anterior
    mcf_minitab$mcf[i] <- mcf_minitab$mcf[i-1]
  }
}

# Linhas de código que exportam os resultados para o clipboard #

```

```

mcf_minibab %>% filter(cens==1)
writeClipboard(as.character(mcf_minibab %>% filter(cens==1) %>% select(mcf)))

### Função que faz o gráfico MCF ###

# Foi criada uma função genérica que permite informar as informações de tempo de operação,
censura e ID do sistema para cálculo do MCF não paramétrico em qualquer base de dados
MCF_function <- function(time, cens, unit){
  mcf_dataframe <- data.frame(time, cens, unit) #Cria um dataframe com os argumentos pass
ados pelo usuário
  names(mcf_dataframe) <- c("time","cens","unit") #Atribui ao dataframe criado os nomes c
orretos para as colunas
  mcf_dataframe <- mcf_dataframe %>% arrange(-desc(time)) #Ordem de forma crescente os te
mpos de operação no dataframe criado

  mcf_dataframe$total_falhas <- 0 #Cria a coluna de total de falhas no dataframe
  mcf_dataframe$total_sistemas <- 0 #Cria a coluna de total de sistemas no dataframe
  mcf_dataframe$mcf <- 0 #Cria a coluna de função média de falhas acumuladas no dataframe

  for (i in 1:nrow(mcf_dataframe)){ #Loop que implementa o algoritmo de cálculo de MCF, c
onforme NELSON 2003:
    mcf_dataframe$total_falhas[i] <- sum(mcf_dataframe$cens[1:i]) #Total de falhas acumul
adas
    mcf_dataframe$total_sistemas[i] <- max(mcf_dataframe$unit) - nrow(mcf_dataframe[1:i
,] %>% filter(cens==0)) #Total de sistemas acumulados
    if (i == 1) { #Valor inicial
      if(mcf_dataframe$cens[1]==0){ #Se for censura, mcf=0
        mcf_dataframe$mcf[i] = 0
      }
      else {
        mcf_dataframe$mcf[i] = 1/ mcf_dataframe$total_sistemas[i] #Se for falha mcf=1/Tot
al de sistemas
      }
    }
    else if (mcf_dataframe$cens[i]==1){ #Se o registro for de uma falha, Calcula-se 1/Tot
al de sistemas naquele tempo de operação somado ao resultado do mesmo calculo da linha ante
rior
      mcf_dataframe$mcf[i] <- 1/ mcf_dataframe$total_sistemas[i] + mcf_dataframe$mcf[i-1]
    }
    else if (mcf_dataframe$cens[i]==0){ #Caso o registro seja de uma censura, o valor de
mcf será igual ao valor da linha anterior
      mcf_dataframe$mcf[i] <- mcf_dataframe$mcf[i-1]
    }
  }

  #Cria o gráfico da Função Acumulada Média de falhas
grafico <- plot(mcf_dataframe$time, mcf_dataframe$mcf, type="s",pch =4, cex=2,
               main = "Função acumulada de Média para Tempo",
               xlab="tempo [h]", ylab="FAM")

  return(c(mcf_dataframe, grafico)) #Função retorna o dataframe calculado e o gráfico
}

MCF_function(dados$time, dados$cens, dados$unit) #Aplicação da função aos dados de falhas d
e transformadores

```

### ### Gráfico MCF Não Paramétrico das Análises ###

```
plot(mcf_minibab$time, mcf_minibab$mcf, type="s", pch =4, cex=0.8, lheight=5,
     main = "Função acumulada de Média para Tempo",
     xlab="time [hours]", ylab="mean cumulative number of failures")
```

### ### Função que faz Gráfico MCF Paramétrico ###

```
mcf_par <- function(t,theta,beta){ #Cria a função que calcula a função intensidade paramétrica através da Lei de Potência
  mcf <- (t/theta)^(beta)
}
```

```
# Linhas que visualizam e exportam os resultados do cálculo para o clipboard #
#view(mcf_par(t=dados$time, beta = saida$par[1], theta = saida$par[2]))
writeClipboard(as.character(mcf_par(t=dados$time, beta = saida$par[1], theta = saida$par[2])))
writeClipboard(as.character(dados %>% select(time)))
```

#Cria gráfico da Função Média Acumulada de Falhas Paramétrica pela Lei de Potência

```
plot(x = tempos_ordenados, y = mcf_par(t=tempos_ordenados, beta = saida$par[1], theta = saida$par[2]),
     main = "Função acumulada de Média para Tempo",
     xlab="tempo [h]", ylab="FAM",
     type="l", lwd =2, col="red")
points(mcf_minibab$time, mcf_minibab$mcf, pch =20, cex=1.5, col="blue") #Sobrepõe os dados de mcf Não Paramétrico
points(x = tempos_ordenados, y = mcf_par(t=tempos_ordenados, beta = 2.4, theta = 22000),
       type="l",
       lwd=2,col="green") #Sobrepõe os dados de paramétrico com outras combinações de parâmetros
```

#Legendas

```
legend(x=14000, 0.22, legend=c("Modelo ajustado", "COLOSIMO, 2007", "Modelo Não Paramétrico"),
       col=c("red", "green", "blue"), lty=c(1,1,0), lwd=c(2,2,0),pch=c(NA_integer_,NA_integer_,20), cex=0.8)
```

```
legend(10000, 0.12, legend=c("Máxima Verossimilhança, Beta=1,9945, Theta=24366,9", "Modelo arbitrário, Beta=2.4, Theta=22000", "Modelo Não Paramétrico"),
       col=c("red", "green", "blue"), lty=c(1,1,0), lwd=c(2,2,0),pch=c(NA_integer_,NA_integer_,20), cex=0.8) #Legenda do gráfico
```

#Análise visual com segmentos que indicam a distância entre o modelo paramétrico e o não paramétrico

```
for (i in 1:(nrow(mcf_minibab))){ #Adiciona segmentos de linha entre os pontos do modelo paramétrico e do modelo não paramétrico
  segments(x0 = mcf_minibab$time[i], y0 = mcf_minibab$mcf[i], x1= mcf_minibab$time[i], y1=mcf_par(t=tempos_ordenados, beta = 2.4, theta = 22000)[i],col="green")
  #segments(x0 = mcf_minibab$time[i], y0 = mcf_minibab$mcf[i], x1= mcf_minibab$time[i], y1=mcf_par(t=tempos_ordenados, beta = saida$par[1], theta = saida$par[2])[i],col="red")
}
```

### ### Gráfico Eventos ###

```
# Cria o gráfico de visualização de eventos de falha e censura para todos os 40 transformadores avaliados
```



```

plot(0,0, type="l",
     main="Gráfico de Eventos para Falhas de Transformadores",
     xlim=c(0,22000), xlab="Time[hours]",
     ylim=c(40,1), ylab="Transformers")

for (i in 1:max(dados$unit)){
  segments(0,i,max(dados %>% filter(unit==i) %>% select(time)),i, col="blue") #Cria segmentos de reta com o comprimento igual a duração máxima de vida de cada sistema
  points(max(dados %>% filter(unit==i) %>% select(time)),i, pch=20, col="blue") #Cria um círculo sólido na censura final de cada sistema
  points((dados %>% filter(unit==i & cens==1))$time, rep(i,length((dados %>% filter(unit==i & cens==1))$time)), pch=4, col="red") #Cria um x vermelho em cada ocorrência de falha e manutenção corretiva
}

### Tempo Ótimo de manutenções preventivas ###

# Cálculo do tempo ótimo para manutenções preventivas, conforme COLOSIMO, 2007:
CPM <- 1 #Custo de manutenção preventiva que restaure o sistema a uma condição de "tão bom quanto novo"
CMR <- 15 #Custo de falha/ manutenção corretiva que restaura o sistema a uma condição de "tão ruim quanto imediatamente antes da falha"

theta <- saida$par[2] #Parâmetro de forma da Lei de potência
beta <- saida$par[1] #Parâmetro de escala da Lei de potência

thau <- theta*((CPM/((beta-1)*CMR))^(1/beta)) #Expressão do tempo ótimo, conforme COLOSIMO, 2007

#####

```

## B) Funções para Análise dos Dados de Falhas com Lei de Potência com Covariáveis: "Function.R"

```

### Script para Obtenção de parâmetros da Lei de Potência com covariáveis categóricas ###

# Desenvolvido por: Robert Bento Florentino (38) 99943-9356
# Trabalho completo intitulado: "Determinação do Período Ótimo de Manutenção em Sistemas Reparáveis através de Processo de Poisson Não Homogêneo com Parâmetros de Modelo de Lei de Potência determinados pelo Método da Máxima Verossimilhança - Metodologia e Estudo de Caso"
# Trabalho desenvolvido como Trabalho de Conclusão de Curso da Especialização em Estatística com Ênfase em Indústria e Mercados da UFMG
# Orientador do trabalho: Prof. Fábio Nogueira Demarqui

library(tidyverse)

#Script contendo apenas as funções

#função Lei de potência para truncamento por falhas SEM grupos como covariáveis

```

```

llpowersystemFailureT <- function(par, t, T, n){
  beta <- par[1] #Parâmetro beta da Lei de Potência
  theta <- par[2] #Parâmetro theta da Lei de Potência
  ll <- n*log(beta) - n*beta*log(theta) +(beta-1)*sum(log(t)) - sum(exp( beta*(log(T) - 1
og(theta)))) #Função da Log-verossimilhança para Lei de Potência com vários sistemas
  return(-ll)
}

#função de Lei de potência considerando grupos como covariáveis
llpowersystemGroups <- function(par, t, T, n, x, x1){
  #Em que
  #t são os tempos de falhas
  #T são os últimos tempos de falha de cada sistema
  #n é o número total de falhas
  #x é a variável dummy para todos os dados de falhas
  #x1 é a variável dummy para os últimos registros de falha (correspondente à T)

  beta <- par[1] #Parâmetro beta da função Lei de potência
  theta1 <- par[2] #Parâmetro theta da função de Lei de potência para o grupo 0
  theta2 <- par[3] #Parâmetro theta da função de Lei de potência para o grupo 1

  llp2 <- 0 #2a parte da função de Log-verossimilhança para Lei de potência
  llp1 <- 0 #1a parte da função de Log-verossimilhança para Lei de potência

  #Cálculo da 2a parte da função, referente aos últimos tempos de falha por sistema
  for (i in 1:length(x1)){
    llp2[i] <- - exp( beta*(log(T[i]) - (theta1+theta2*x1[i])))
  }
  llp2 <- sum(llp2)

  #Cálculo da 1a parte da função, referente à todos os tempos de falha
  llp1 <- n*log(beta) - n*beta*theta1 - beta*theta2*sum(x) +(beta-1)*sum(log(t))

  #Resultado final da função Log-verossimilhança
  ll <- llp1+llp2

  return(-ll)
}

#Função que calcula a função média acumulada para dados SEM separação de grupos
mcf_par <- function(t,theta,beta){ #Cria a função que calcula a função intensidade paramé
trica através da Lei de Potência
  #Em que
  #t são todos os tempos de falhas ordenados de forma crescente
  #theta é o parâmetro de escala da função Lei de potência
  #Beta é o parâmetro de forma

  t <- t[order(t)]
  tempos_ordenados <- t

  mcf <- (t/theta)^(beta)
  grafico <- plot(x=tempos_ordenados,y=mcf, type="l", col="red",
                 xlab = "Time", ylab = "Mean Cumulative Function",
                 main = "Parametric Growth Curve")

  return(c(mcf, grafico))
}

```

```

# MCF não paramétrico para dados censurados por falhas

#Função que cria a MCF não paramétrica dos dados
mcf_n_param_groups <- function(system, time, x){

  #Cria um dataframe com os argumentos passados
  dados <- data.frame(system, time, x)

  #Ordena os dados de falhas de forma crescente
  dados <- dados %>%
    arrange(-desc(time))

  #Loop que insere duas colunas calculadas de acordo com metodologia no texto
  for (i in 1:nrow(dados)){
    if(i==1){
      dados$systemas[i] <- length(unique(dados$system))
      dados$mcf[i] <- 1/dados$systemas[i]
    }
    else{
      if(dados$time[i]==dados$time[i-1]){dados$systemas[i] <- dados$systemas[i-1]}
      else{dados$systemas[i] <- length(unique(dados$system[i:nrow(dados)]))}
      dados$mcf[i] <- dados$mcf[i-1]+1/dados$systemas[i]
    }
  }

  #gráfico da função média acumulada não paramétrica
  grafico <- plot(x=dados$time, y=dados$mcf, type="s", pch=18, col="blue",
                 xlab = "Time", ylab = "Mean Cumulative Function",
                 main = "Non parametric Growth Curve")

  #Função retorna o dataframe calculado e o gráfico
  return(c(grafico,dados))
}

#Função que cria a MCF paramétrica dos dados COM separação de grupos

mcf_param_groups <- function(t,theta1,theta2,beta,x){
  #Em que
  #t são os tempos de falhas
  #theta1 é o primeiro argumento da função linear de theta
  #theta2 é o 2o argumento da função linear de theta
  #x é a variável dummy que indica os grupos utilizados como covariáveis de theta

  dados <- data.frame(t,x)
  dados <- dados %>% arrange(-desc(t))

  dados$mcfpg <- (t/exp(theta1+theta2*x))^(beta)

  #grafico <- plot(x=dados$t,y=dados$mcfpg, type="p", col="red",
  #               xlab = "Time", ylab = "Mean Cumulative Function",
  #               main = "Parametric Growth Curve with groups")

  return(dados$mcfpg)
}

```

```

# Função que calcula os intervalos de confiança de tau para Lei de Potência com Covariáveis
: Cálculo da variância
varTau <- function(saida, CPM, CMR, x){ #Entradas da função <- saída da função optim com
  todos seus argumentos, CPM, CMR e x, variável dummy da covariável categórica

  #Parâmetro da Lei de potência com covariável categórica
  beta <- saida$par[1]
  alpha_0 <- saida$par[2]
  alpha_1 <- saida$par[3]

  V <- solve(saida$hessian) #Inversa da Hessiana

  #Cálculos intermediários auxiliares do cálculo das derivadas parciais
  A <- CPM/CMR
  B <- 1/(beta-1)
  C <- A*B
  D <- C ^(1/beta)

  derivadas_3 <- 0 #Define o vetor de derivadas parciais

  #Cálculo das derivadas parciais
  derivadas_3[1] <- - exp(alpha_0 + alpha_1*x)*D*(B/beta+((log(C))/(beta^2))) #Derivada d
e Tau com respeito à beta
  derivadas_3[2] <- exp(alpha_0 + alpha_1*x)*D #Derivada de Tau com respeito à alpha 0
  derivadas_3[3] <- x*exp(alpha_0 + alpha_1*x)*D #Derivada de Tau com respeito à alpha 1

  return(derivadas_3 %*% V %*% derivadas_3) #Retorna a variância de Tau
}

```

### C) Análise dos Dados de Falhas de Componentes de Freio de Locomotivas: “Analisys.R”

```

### Script para Análise dos dados de falhas de componentes de freio de locomotiva ###
### pela Lei de Potência com covariáveis categóricas ###

# Desenvolvido por: Robert Bento Florentino (38) 99943-9356
# Trabalho completo intitulado: "Determinação do Período Ótimo de Manutenção em Sistemas Re
paráveis através de Processo de Poisson Não Homogêneo com Parâmetros de Modelo de Lei de Po
tência determinados pelo Método da Máxima Verossimilhança – Metodologia e Estudo de Caso"
# Trabalho desenvolvido como Trabalho de Conclusão de Curso da Especialização em Estatístic
a com Ênfase em Indústria e Mercados da UFMG
# Orientador do trabalho: Prof. Fábio Nogueira Demarqui

# Otimização da função para 2 grupos
library(tidyverse) #Leitura do tidyverse
source('Functions.R') #Leitura do arquivo com as funções

dadosgrupos <- read.csv("Dados/brakereliability.csv", header=TRUE, sep=";") #Leitura dos da
dos
names(dadosgrupos) <- c("system","days","type","x","cens") #Muda nome das colunas

### Execução da Log-verossimilhança para a Lei de Potência para vários sistemas ###

#Dados dos tempos de falhas

```

*#OBSERVAÇÃO IMPORTANTE PARA TRUNCAMENTO POR FALHA: DEVEM SER CONSIDERADOS TODOS OS TEMPOS D E FALHA, INCLUSIVE O ÚLTIMO (Na variável "t")*

```
t <- dadosgrupos %>%
  select(days)
t <- t$days
t
```

*#Dados dos tempos de censura (cens=1) - Neste caso, foi definido a coluna cens, em que cens =1 indica a última falha de cada sistema*

```
T <- dadosgrupos %>%
  filter(cens==1) %>%
  select(days, x)
T <- T$days
T
```

*#A variável x é uma variável dummy que permite comparar grupos. Foi definida no problema do s freio como x=0 para tipo=1 e x=1 para tipo=2.*

```
x <- dadosgrupos %>%
  select(days, x, cens)
x <- x$x
x
```

*#É necessário também a definição da variável x exclusivamente para os últimos tempos de falha*

```
x1 <- dadosgrupos %>% filter(cens==1) %>% select(x)
x1 <-x1$x
x1
```

*#Número total de falhas*

```
n <- nrow(dadosgrupos )
n
```

*# Utilização da função Optim para a função de Lei de Potência truncada por Falhas SEM segregação de grupos*

```
saida <- optim(par = c(1,1), fn = llpowersystemFailureT,
              t=t, T=T, n=n, method="L-BFGS-B",
              lower = 1e-16, upper = Inf,
              hessian = TRUE)
```

```
saida
saida$par #Parâmetros de escala e de forma obtidos
```

*# Utilização da função Optim para a função de Lei de Potência truncada por Falhas COM segregação de grupos*

```
saida <- optim(par = c(1,1,1), fn = llpowersystemGroups,
              t=t, T=T, n=n, x=x, x1=x1, method="L-BFGS-B",
              lower = c(1e-16,-Inf,-Inf), upper = c(Inf,Inf,Inf),
              hessian = TRUE)
```

```
saida
exp(saida$par) #Parâmetros de escala e de forma obtidos
saida$value #Valor da função Log-verossimilhança
```

*#Gráfico da MCF paramétrico para Lei de Potência truncada por falha SEM grupos*

```
beta <- saida$par[1] #Atribuição dos resultados ao parâmetro beta
theta <- saida$par[2] #Atribuição dos resultados ao parâmetro theta
```

```

#Gráfico do modelo paramétrico SEM grupos como covariáveis
mcf_par(t=dadosgrupos$days,theta=saida$par[2], beta=saida$par[1])

# MCF não paramétrico para dados censurados por falhas

#Modelo sem distinção de grupos
mcf_n_param_groups(system = dadosgrupos$system, time = dadosgrupos$days, x = dadosgrupos$x)

#Modelo para dados do grupo 1
mcf_n_param_groups(system = dadosgrupos$system[dadosgrupos$type==1], time = dadosgrupos$days[dadosgrupos$type==1], x = dadosgrupos$x[dadosgrupos$type==1])

#Modelo para dados do grupo 2
mcf_n_param_groups(system = dadosgrupos$system[dadosgrupos$type==2], time = dadosgrupos$days[dadosgrupos$type==2], x = dadosgrupos$x[dadosgrupos$type==2])

# MCF paramétrico utilizando os grupos como covariáveis
mcf_param_groups(t=dadosgrupos$days, theta1 = saida$par[2], theta2 = saida$par[3], beta = saida$par[1], x=dadosgrupos$x)

#Análise Gráfica comparativa entre modelos

#Armazenamento intermediário de função média acumulada não paramétrica
dadostodos <- mcf_n_param_groups(system = dadosgrupos$system, time = dadosgrupos$days, x = dadosgrupos$x)

dadostipo1 <- mcf_n_param_groups(system = dadosgrupos$system[dadosgrupos$type==1], time = dadosgrupos$days[dadosgrupos$type==1], x = dadosgrupos$x[dadosgrupos$type==1])

dadostipo2 <- mcf_n_param_groups(system = dadosgrupos$system[dadosgrupos$type==2], time = dadosgrupos$days[dadosgrupos$type==2], x = dadosgrupos$x[dadosgrupos$type==2])

#Visualização sobreposta dos modelos não paramétricos
plot(x=dadostodos$time, y=dadostodos$mcf, type="s", pch=18, col="blue", lwd=2,
      xlab = "Tempo", ylab = "FAM",
      main = "Lei de Potência com Covariáveis") #Modelo não paramétrico geral

plot(x=dadostipo1$time, y=dadostipo1$mcf, type="s", pch=18, lwd=2, col="orange",
      xlab = "Tempo", ylab = "FAM",
      main = "Gráfico de Comparação de Modelos") # Modelo não paramétrico para tipo 1
# xlab = "Time", ylab = "Mean Cumulative Function",
# main = "Lei de Potência com Covariáveis")
points(x=dadostipo2$time, y=dadostipo2$mcf, type="s", pch=18, lwd=2, col="green") # Modelo não paramétrico para tipo 2

#Visualização sobreposta dos modelos paramétricos SEM metodologia de separação de grupos por covariáveis
points(x=dadostodos$time, y=(dadostodos$time/(303.157702))^1.536818, col="blue", type="l", lty=2)

```

```

points(x=dadostodos$time[dadostodos$x==0], y=(dadostodos$time[dadostodos$x==0]/(461.315))
^2.31225, col="orange", type="l", lty=2)
points(x=dadostodos$time[dadostodos$x==1], y=(dadostodos$time[dadostodos$x==1]/(221.375))
^1.42532, col="green", type="l", lty=2)
#Legend(x=550, y=1.8, legend=c("All breaks - Non parametric", "Type 1 breaks - Non paramet
ric", "Type 2 breaks - Non parametric", "Power Law All", "Power Law Type 1", "Power Law Type
2"), col=c("blue", "orange", "green", "blue", "orange", "green"), pch=c(NA_integer_, NA_integer
_, NA_integer_, NA_integer_, NA_integer_, NA_integer_), lty=c(1, 1, 1, 2, 2, 2), cex=0.6, bty="n")
legend(x=600, y=1.3, legend=c("All breaks", "Type 1 breaks", "Type 2 breaks"), col=c("blue
", "orange", "green"), pch=c(NA_integer_, NA_integer_, NA_integer_), lty=c(1, 1, 1), cex=0.6, bty=
"n", lwd=5, box.lwd=2, box.lty=2, box.col = "black")

```

*#Visualização sobreposta dos modelos paramétricos COM metodologia de separação de grupos po  
r covariáveis*

```

dadoscov <- mcf_param_groups(dadostodos$time, saida$par[2], saida$par[3], saida$par[1], dad
ostodos$x)

```

*#Gráfico com grupo 1*

```

points(x=dadostodos$time[dadostodos$x==0], y=mcf_param_groups(t=dadostodos$time[dadostodo
s$x==0], theta1=saida$par[2], theta2=saida$par[3], beta=saida$par[1], x=dadostodos$x[dadosto
dos$x==0]), col="orange", type="l", lty=3)

```

*#Gráfico com grupo 2*

```

points(x=dadostodos$time[dadostodos$x==1], y=mcf_param_groups(t=dadostodos$time[dadostodo
s$x==1], theta1=saida$par[2], theta2=saida$par[3], beta=saida$par[1], x=dadostodos$x[dadosto
dos$x==1]), col="green", type="l", lty=3)

```

```

legend(x=550, y=1.8, legend=c("All breaks - Non parametric", "Type 1 breaks - Non paramet
ric", "Type 2 breaks - Non parametric", "Power Law All", "Power Law Type 1 - Covariates", "Pow
er Law Type 2 - Covariates"), col=c("blue", "orange", "green", "blue", "orange", "green"), pch
=c(NA_integer_, NA_integer_, NA_integer_, NA_integer_, NA_integer_, NA_integer_), lty=c(1, 1, 1, 3,
3, 3), cex=0.6, bty="n")

```

```

legend(x=530, y=1.4, legend=c("Type 1 breaks - Non parametric", "Type 2 breaks - Non param
etric", "Power Law Type 1", "Power Law Type 2", "Power Law Type 1 - Covariates", "Power Law
Type 2 - Covariates"), col=c("orange", "green", "orange", "green", "orange", "green"), pch=c(N
A_integer_, NA_integer_, NA_integer_, NA_integer_, NA_integer_, NA_integer_), lty=c(1, 1, 2, 2, 3, 3)
, cex=0.6, bty="n")
title("Gráfico de Comparação de Modelos")

```

*#Calculo dos intervalos de confiança de tau - tempo ótimo para manutenções preventivas #*

```

#          método          delta          com          3          variaveis          #

```

```

CPM      <-      1      #Definição do custo de manutenções preventivas

```

```

CMR      <-      15     #Definição do custo de manutenções corretivas

```

```

#Parâmetros obtidos pelo método da máxima verossimilhança.

```

```

beta                                           <-      saida$par[1]

```

```

alpha_0                                       <-      saida$par[2]

```

```

alpha_1                                       <-      saida$par[3]

```

```

x <- 1 #Variável dummy. Utilizar valor 0 para tipos de freio referência, 1 para tipo de
freios alternativos.

```

```

#Definições intermediárias para intervalos e derivadas parciais
A <- CPM/CMR
B <- 1/(beta-1)
C <- A*B
D <- C^(1/beta)

```

```

#Intervalo de confiança 95% para Tau (Se x=0, tipo de freio =1, se x=1, tipo de freio=2)
Lowerbound_3 <- D*exp(alpha_0 + alpha_1*x) -1.96*sqrt(varTau(saida=saida, CPM=CPM, CMR=CMR,
x=x))
upperbound_3 <- D*exp(alpha_0 + alpha_1*x) +1.96*sqrt(varTau(saida=saida, CPM=CPM, CMR=CMR,
x=x))

```

```

#Amplitude do intervalo de confiança
upperbound_3-Lowerbound_3

```

```

#Cálculo do Intervalo de Confiança dos parâmetros beta, theta 0 e theta 1 para Lei de Potência
com covariáveis
V_param <- 0 #
V <- solve(saida$hessian)

```

```

#Desvio padrão de Beta > Obtido diretamente pela matriz hessiana
sqrt(V[1,1])

```

```

#Eq.44: Método Delta para Variâncias de Theta 0 e Theta 1 à partir de alpha0 e alpha1

```

```

#Matriz gradiente
V_param <- matrix(nrow = 2, ncol = 2)
V_param[1,1] <- V[2,2] #Variância de alpha0
V_param[1,2] <- V[2,3] #Covariância de alpha0, alpha1
V_param[2,1] <- V[3,2] #Covariância de alpha1, alpha0
V_param[2,2] <- V[3,3] #Variância de alpha1

```



```

#Vetor de derivadas

x <- 1 #Dummy que indica o grupo da covariável categórica (0 ou 1 no exemplo de componentes
de freio de locomotivas)
derivadas_param <- 0
derivadas_param[1] <- exp(alpha_0+alpha_1*x)#Derivada de Theta com respeito à alpha0
derivadas_param[2] <- x*exp(alpha_0+alpha_1*x)#Derivada de Theta com respeito à alpha1

#Desvio padrão de Theta
sqrt(derivadas_param %>% V_param %>% derivadas_param)

#Intervalo de confiança 95% para Theta (Se x=0, tipo de freio =1, se x=1, tipo de freio=2)
lowerbound_theta <- exp(alpha_0 + alpha_1*x) -1.96*sqrt(derivadas_param %>% V_param %>%
derivadas_param)
upperbound_theta <- exp(alpha_0 + alpha_1*x) +1.96*sqrt(derivadas_param %>% V_param %>%
derivadas_param)

# #Intervalo de confiança 95% para Beta (Se x=0, tipo de freio =1, se x=1, tipo de freio=2)
lowerbound_beta <- beta -1.96*sqrt(V[1,1])
upperbound_beta <- beta +1.96*sqrt(V[1,1])

#Comparação com os valores pela metodologia MINITAB (sem covariáveis)
grupo <- 0 #Filtra dos dados. 0 para componentes de freio do tipo 1, 1 para componentes de
freio do tipo 2
t <- dadosgrupos %>%
  filter(x==grupo) %>%
  select(days)
t <- t$days

#Dados dos tempos de censura (cens=1) - Neste caso, foi definido a coluna cens, em que cens=1
indica a última falha de cada sistema

T <- dadosgrupos %>%
  filter(x==grupo) %>%
  filter(cens==1) %>%

```

```

select(days,
T
T
x)
T$days
T

#A variável x é uma variável dummy que permite comparar grupos. Foi definida no problema dos
freio como x=0 para tipo=1 e x=1 para tipo=2.
x <- dadosgrupos %>%
select(days, x, cens)
x <- x$x
x

#É necessário também a definição da variável x exclusivamente para os últimos tempos de falha
x1 <- dadosgrupos %>% filter(cens==1) %>% select(x)
x1 <- x1$x
x1

#Número total de falhas
n <- nrow(dadosgrupos %>% filter(x==grupo))
n

# Utilização da função Optim para a função de Lei de Potência truncada por Falhas SEM
segregação de grupos
saida <- optim(par = c(1,1), fn = llpowersystemFailureT,
t=t, T=T, n=n, method="L-BFGS-B",
lower = 1e-16, upper = Inf,
hessian = TRUE)

theta <- saida$par[2] #Parâmetro de escala
beta <- saida$par[1] #Parâmetro de forma

#Cálculo do tempo ótimo para manutenções preventivas
thau_0_ind <- theta*((CPM/((beta-1)*CMR))^(1/beta)) #Expressão do tempo ótimo, conforme
COLOSIMO, 2007
thau_0_ind

#Cálculo do Intervalo de Confiança para os parâmetros de escala Theta 0 e Theta 1 - SEM
covariáveis

```

```

saida$par      #Parâmetros de escala e de forma obtidos
V              <- solve(saida$hessian)
V
sqrt(V[1,1])   #Desvio padrão de Theta 0
sqrt(V[2,2])   #Desvio padrão de Theta 1

derivadas     <- 0 #Criação do vetor de derivadas parciais

CPM           <- 1 #Custo de manutenções preventivas
CMR           <- 15 #Custo de manutenções corretivas

#Cálculos intermediários para vetor de derivadas parciais
A             <- CPM/CMR
B             <- 1/(beta-1)
C             <- A*B
D             <- C^(1/beta)
D*theta       #Corresponde ao resultado de Tau

#Vetor de derivadas parciais
derivadas[1] <- theta*D*(-B/beta-((Log(C))/(beta^2))) #Derivada de Tau com respeito à beta
derivadas[2] <- D #Derivada de Tau com respeito à theta
derivadas

#Intervalo de Confiança de 95% para Tau - SEM covariáveis
Lowerbound   <- D*theta -1.96*sqrt(derivadas%% V %% derivadas)
upperbound   <- D*theta +1.96*sqrt(derivadas%% V %% derivadas)
Lowerbound
upperbound

#Amplitude do intervalo de confiança
upperbound-Lowerbound

```