

UNIVERSIDADE FEDERAL DE MINAS GERAIS
INSTITUTO DE CIENCIAS EXATAS
DEPARTAMENTO DE ESTATÍSTICA

Cássio Azevedo de Lima

Avaliação Bayesiana de Falhas em Locomotivas
Modelo Dash9W –GE

Belo Horizonte

2021

Cassio Azevedo de Lima

Avaliação Bayesiana de Falhas em Locomotivas Modelo *Dash9W – GE*

Dissertação apresentada ao Programa de Pós-Graduação em Estatística do Instituto de Ciências Exatas da Universidade Federal de Minas Gerais como requisito parcial para obtenção do grau de Especialista em Estatística com Ênfase em Indústria e Mercado como requisito parcial para obtenção do grau de Especialista em Estatística.

Orientador: Prof. Dr. Roberto da Costa Quinino

Belo Horizonte

2021

© 2021, Cássio Azevedo de Lima. Todos os direitos reservados



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
Tel: 3409-5923 – FAX: 3409-5924

ATA DO 237º. TRABALHO DE FIM DE CURSO DE ESPECIALIZAÇÃO EM ESTATÍSTICA DE CÁSSIO AZEVEDO DE LIMA.

Aos quinze dias do mês de dezembro de 2021, às 16:00 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 Cássio Azevedo de Lima, intitulado: “*Avaliação Bayesiana de Falhas em Locomotivas Modelo Dash9W – GE*”, como requisito para obtenção do Grau de Especialista em Estatística. Abrindo a sessão, o Presidente da Comissão, Professor Roberto da Costa Quinino – 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, 15 de dezembro de 2021.

Prof. Roberto da Costa Quinino (Orientador)
DEST / ICEx / UFMG

Prof. Danilo Gilberto de Oliveira Valadares
DEST / ICEx / UFMG

Prof. Anderson Laécio Galindo Trindade
DEP / UFMG

AGRADECIMENTO

Autor agradece em primeiro lugar a Deus, pela vida e por ajudar a ultrapassar todos os obstáculos encontrado ao longo do curso e a todos aqueles que, direta ou indiretamente, colaboraram para a realização dele, em particular:

- À minha esposa e filha pela compreensão em virtude dos momentos de ausência devido aos estudos;
- Ao Professor Roberto da Costa Quinino pela coordenação do curso, orientação deste trabalho e inúmeras contribuições;
- Aos gestores Eric Bastos, Cristiano Barros e colegas de trabalho por potencializarem o desenvolvimento técnico-científico;
- Aos companheiros de jornada no curso de Especialização em Estatística: Karina de Aquino França e Alexandre de Freitas Gomes de Mendonça pelos momentos de estudo;
- Ao programa de Especialização do Departamento de Estatística da UFMG;
- À Universidade Federal de Minas Gerais, ao seu corpo docente e técnico e pela infraestrutura acadêmica para os estudos.

RESUMO

Neste trabalho avaliamos as falhas nos sistemas reparáveis das locomotivas modelo *Dash9W –GE* de acordo com um processo de Poisson Não Homogêneo, utilizando abordagem Bayesiana e lei potência para intensidade de falhas. Considerando a complexidade da função de verossimilhança a distribuição da posteriori dos parâmetros foi obtida com o algoritmo *Sampling Importance Resampling (SIR)*.

PALAVRAS-CHAVE: Processo de Poisson Não homogêneo, Sistemas Reparáveis, Lei Potência, Inferência Bayesiana, *Sampling Importance Resampling (SIR)*.

ABSTRACT

In this work we evaluate locomotive failures according to a non-homogeneous Poisson process, using a Bayesian approach and a power law for failure intensity. Considering the complexity of the likelihood function, the posterior distribution of the parameters was obtained with the SIR-Sampling Importance Resampling algorithm.

KEYWORDS: Non-homogeneous Poisson Process, Repairable Systems, Power Law, Bayesian Inference, Sampling Importance Resampling (SIR).

LISTA DE FIGURAS

Gráfico 1 – Gráfico de falhas por Locomotivas 1113 – 1132.....	19
Gráfico 2 – Gráfico de falhas por Locomotivas 1133 – 1154.....	20
Gráfico 3 – Gráfico de falhas por Locomotivas 1155 – 1174.....	20
Gráfico 4 – Gráfico de falhas por Locomotivas 1175 – 1195.....	20
Gráfico 5 – Gráfico de falhas por Locomotivas 1196 – 1243.....	21
Gráfico 6 – Gráfico de falhas por Locomotivas 1244 – 1264.....	21
Gráfico 7 – Gráfico de falhas por Locomotivas 1265 – 1288.....	21
Gráfico 8 – Gráfico de falhas por Locomotivas 1288 – 1309.....	22
Gráfico 9 – Histograma Distribuição Posteriori de β	23
Gráfico 10 –Histograma Distribuição Posteriori de $\theta/1000$	23
Gráfico 11 – Histograma Distribuição Posteriori de $\theta/1000$	24

LISTA DE TABELAS

Tabela 1 – Estimativa pontuais para β e θ	24
Tabela 2 – Intervalo Credibilidade para β e θ	24
Tabela 3 – Intervalo Credibilidade para θ	25
Tabela 4 – Intervalo Credibilidade para θ	26

LISTA DE ABREVIATURAS E SIGLAS

EMV Estimador de Máxima Verossimilhança

FMV Função de Máxima Verossimilhança G

GE *General Electric*

iid Distribuição Idêntica Independente

L_i Limite Inferior

L_s Limite Superior

MTBF Tempo Médio entre Falha

PLP *Power Law Process*

PPH Processo de Poisson Homogêneo

PPNH Processo de Poisson Não Homogêneo

SIR *Sampling Importance Resampling*

Sumário

1 INTRODUÇÃO	12
2 OBJETIVOS	14
2.1 Objetivo Geral	14
2.2 Objetivo Específico.....	14
3 METODOLOGIA.....	15
3.1 Importância dos Dados	15
3.2 Modelos Paramétricos para Dados de Recorrência	16
3.3 Processo de Poisson Homogêneo	16
3.4 Processo de Poisson Não Homogêneo.....	17
3.5 Função de Verossimilhança e Inferência Bayesiana.....	18
3.6 Sampling Importance Resampling (SIR)	19
4 APLICAÇÃO.....	20
5 CONCLUSÃO	26
ANEXO A: BASE DE DADOS	30
ANEXO B: PROGRAMAÇÃO SOFTWARE R.....	34

1 INTRODUÇÃO

O modal ferroviário no Brasil, possui diversas ligações nos estados da região Sudeste e Centro-Oeste, o que facilita o escoamento de diversos produtos para o litoral com finalidade de exportação. Uma das maiores mineradoras do nosso país, possui um provedor logístico estruturado para atender o escoamento do seu principal produto o minério de ferro, no prazo certo, quantidade certa com os melhores custos operacionais, mantendo sua competitividade no mercado mundial. Na sua malha ferroviária a máquina para tracionar os vagões é conhecida como locomotiva, no qual se trata de um veículo automotor que corre sobre os trilhos, utiliza formas diversas de energia para produzir movimento e serve para rebocar vagões de passageiro ou de carga tendo diversos modelos de locomotivas composta na sua frota. Uma das suas frotas mais moderna é composta por 196 locomotivas *Dash9W* da fabricante *General Electric-GE*, usualmente utilizada para tracionar as composições de vagões de minério em período pré-estabelecido, denominada frota de potência. Essa frota de locomotiva pode apresentar parada involuntária devido a possíveis falhas.

Segundo [3] falha é a deterioração da acomodação de um item para realizar sua função específica. Podendo acontecer perda parcial ou de toda a capacidade de uma peça, parte constituinte de um sistema ou até mesmo da locomotiva de realizar sua função durante um determinado tempo, onde essa máquina poderá sofrer algum tipo de manutenção ou até mesmo ser substituído por outra. A falha na manutenção leva ao estado de indisponibilidade. Manter disponibilidade da frota de locomotivas requer grau de excelência, uma gestão de manutenção estruturada, pois essas máquinas apresentam nível de complexidade em seus sistemas e são altamente exigidos durante o transporte de composições de minério.

De acordo [5], o processo de manutenção é um dos processos mais custosos de qualquer organização, sendo que o custo de manutenção em média representa 4,11% do faturamento das organizações. Assim, a tendência das organizações é buscar o melhor custo para manutenção o que demanda entender o processo de geração de falhas das locomotivas

Locomotivas são ativos duráveis e realizar reparos é essencial para que seus sistemas possam desempenhar sua função nas características determinadas.

Usualmente as falhas de sistemas reparáveis podem ser avaliadas segundo um Processo de Poisson homogêneo – PPH ou um Processo de Poisson não homogêneo –PPNH.

Em sistemas reparáveis, [1] mencionam dois tipos de reparos: o reparo perfeito e o reparo mínimo. O reparo mínimo tem como propósito tão bom como estava antes da falha. Já o reparo perfeito tornaria o equipamento tão bom quanto novo.

Visando um modelo envolvendo probabilidade para reparar os sistemas, varia da quantidade do reparo e de como o sistema envolvido se comporta. Cita [10] sobre o fator importante na modelagem que é o truncamento, visando a coleta de dados. Truncamento por tempo, onde há uma coleta de dados em um determinado tempo em que o número de falhas é aleatório e há também o truncamento que ocorre através de falhas onde acontece coleta de dados que se trata de um número pré-determinado de falhas e esse tempo não é fixo.

Para tanto decidir o tipo de truncamento, [10] ainda usado para experimento varia do comportamento dos sistemas de um modo geral e onde acontece ou o meio onde ele está inserido, isso quer dizer se o equipamento tem muitas falhas em curto período ou não e por quanto tempo esse mesmo equipamento trabalha. No entanto os tipos reparo, truncamento e a variedade de sistemas em análise direcionam quais processos de distribuição devem serem aplicados.

O trabalho propõe avaliar falhas observadas em 196 locomotivas, com truncamento por tempo e realização de reparo mínimo (PPNH). A função intensidade de falhas será de acordo com a lei de potência.

2 OBJETIVOS

2.1 Objetivo Geral

Determinar um intervalo de credibilidade para o número de falhas em locomotivas modelo Dash9W utilizando a lei potência (*Power Law Processes*) via inferência bayesiana através do algoritmo *Sampling Importance Resampling*.

2.2 Objetivo Específico

Estimar os parâmetros da lei potência e determinar o número médio de falhas de locomotivas em um tempo determinado.

3 METODOLOGIA

Esta seção será dedicada a descrição dos métodos utilizados para os cálculos e as análises neste trabalho, são eles: Lei Potência para Intensidade de Falhas, Processo de Poisson Não Homogêneo (PPNH), Função de Máxima Verossimilhança (FMV), Estimador de Máxima Verossimilhança (EMV) e Inferência Bayesiana e *Sampling Importance Resampling*.

De maneira geral o objetivo é obter um modelo de falhas através da lei de potência, analisando dados de falhas de locomotivas por tempo de deterioração do equipamento e considerando um truncamento por tempo. Inicialmente adota-se um processo de Poisson Não Homogêneo (PPNH). A função de máxima verossimilhança depende do θ e do β , e os valores desses parâmetros que maximizam essa função são denominados estimadores de máxima verossimilhança.

Diante de tudo isso, a partir de um Processo de Poisson Não Homogêneo (PPNH) busca-se encontrar o número médio de falhas em período determinado.

3.1 Importância dos Dados

Para Filho, 2006, Qualquer elemento que possa transmitir ou representar informação são ditos dados. Representação convencional de fatos ou conceitos para comunicação, registros ou arquivos, seja para processamento de dados, transmissão, análise estatística ou qualquer outra finalidade.

Trata somente da natureza dos dados, não expõe o potencial nos negócios, podemos ter a seguinte definição nos negócios.

Dados são recursos da organização, assim os recursos devem ser vistos como:

- Investimento;
- Algo confiável;
- Algo progetivo

- Algo disponibilizado;
- Algo utilizável;
- Algo administrável.

O modelo que adotada para aos dados deste trabalho é sua confiabilidade .Para fazer a colheta de dados das locomotivas rebocadas com falhas (piores casos), o critério usado para observar os reboques (falhas), são:

- Tempo de 815 dias;
- 1 código para falha;
- 0 código sem ocorrência de falha;
- 196 locomotivas do modelo Dash9W.

3.2 Modelos Paramétricos para Dados de Recorrência

Contudo, modelos de processo de ponto paramétrico são utilizados em dados de recorrência. Os modelos habitualmente usados para dados de recorrência são processos de Poisson (homogêneos e não homogêneos), processos de renovação e versões sobrepostas desses processos. A aplicação mais comum é para verificar e modelar recorrências, como reparos, a partir de um único sistema.

3.3 Processo de Poisson Homogêneo

De acordo com [7] e [9] um Processo de Poisson homogêneo - PPH é um processo Poisson com uma taxa de recorrência constante, digamos, $\lambda(t) = 1/\theta$. Nesse caso:

1. $N(a, b)$ tem distribuição Poisson com parâmetro $\mu(a, b) = (b - a)/\theta$.
2. O número esperado de recorrências em $(a, b]$ é $\mu(a, b)$. Equivalentemente, o número esperado de recorrências por unidade de tempo ao longo $(a, b]$ é constante e igual a $1/\theta$. Essa propriedade é chamada "incrementos estacionários".
3. Os tempos de intercorrência (o tempo entre as recorrências), $\tau_j = T_j - T_{j-1}$, são independentes e distribuídos de forma idêntica (iid), cada um

com uma distribuição $Exp(\theta)$. Isso segue diretamente do relacionamento:

$$\Pr(\tau_j > t) = \Pr[N(T_{j-1}, T_{j-1} + t) = 0] = \exp(-t/\theta).$$

assim, o tempo médio no estado estacionário entre as recorrências de um HPP é igual a θ . Para um processo de falha, pode-se dizer que o tempo médio entre falhas é $MTBF = \theta$.

4. O tempo $T_k = \tau_1 + \dots + \tau_k$ até a i -ésima recorrência tem uma distribuição $GAM(\beta, \theta)$.

A forma como as falhas das Locomotivas acontece nesse tipo de sistema não é constante, sendo assim o Processo de Poisson Não Homogêneo é muito utilizado para modelar este tipo de processo, uma vez que ele se adequa muito bem a forma com que a aleatoriedade das falhas ocorre [7], [8] e [9].

3.4 Processo de Poisson Não Homogêneo

O Processo de Poisson Não Homogêneo - PPNH é modelado por uma função intensidade não constante $\lambda(t)$, os tempos entre falhas não são independentes e nem distribuído de forma harmônica, em que o número esperado de falhas em $(0; t]$ é dado por :

$$\lambda(t) = \int_a^b \lambda(t) dt$$

Um modelo PPNH é continuamente caracterizado para a taxa de recorrência $\lambda(t)$. Para diferenciar um modelo PPNH que usamos $\lambda(t) = \lambda(t; \theta)$, uma função de um vetor incógnita de parâmetros θ . A taxa de recorrência do modelo de potência (ou lei potência) é

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

O número acumulado médio de recorrências ao longo de $(0, t]$ é $\mu(t; \beta, \theta) = (t/\theta)^\beta$. Quando $\beta = 1$, este modelo se reduz ao modelo PPH.

3.5 Função de Verossimilhança e Inferência Bayesiana

Seguindo [4] e [8] vamos denotar t_{ij} a j -ésima falha no i -ésimo sistema. Suponha que o sistema i seja observado até o tempo T_i , onde $T_i = t_{i,n_i}$ se o sistema i tiver uma falha truncada e $T_i \geq t_{i,n_i}$, se os sistemas i eram truncados por tempo. Seja n_i o número de falhas observadas para o sistema i , e $N = n_1, n_2, \dots, n_k$. Como as falhas em sistemas separados são independentes, a função de verossimilhança é dada por:

$$\begin{aligned} L(\theta, \beta) &= \prod_{i=1}^k \left\{ \left[\prod_{j=1}^{n_i} \lambda(t_{i,j}) \right] \exp[-\Lambda(T_i)] \right\} \\ &= \prod_{i=1}^k \left\{ \left[\prod_{j=1}^{n_i} \frac{\beta}{\theta} \left(\frac{t_{i,j}}{\theta} \right)^{\beta-1} \right] \exp[-(T_i/\theta)^\beta] \right\} \\ &= \beta \sum_{i=1}^k n_i \theta^{-\beta} \sum_{i=1}^k n_i \left(\prod_{i=1}^k \prod_{j=1}^{n_i} (t_{i,j}) \right)^{\beta-1} \exp \left[- \sum_{i=1}^k (T_i/\theta)^\beta \right] \end{aligned}$$

[1].

Se considerarmos que a distribuição a priori para β e θ $[\Phi(\beta; \theta)]$ são independentes e podem ser representadas respectivamente por distribuições uniformes nos intervalos $(\rho_\beta, \alpha_\beta)$ e $(\rho_\theta, \alpha_\theta)$ então a distribuição a posteriori $\Phi(\beta; \theta)$ será dada por:

$$\Phi(\beta, \theta) = K \frac{L(\theta, \beta)}{(\alpha_\beta - \rho_\beta)(\alpha_\theta - \rho_\theta)} \quad [2]$$

Em que $K = \left(\int_{\rho_\beta}^{\alpha_\beta} \int_{\rho_\theta}^{\alpha_\theta} \frac{L(\theta, \beta)}{(\alpha_\beta - \rho_\beta)(\alpha_\theta - \rho_\theta)} \right)^{-1}$.

Considerando a dificuldade para obtenção da constante K (devido a solução de integrais não triviais) é recomendado utilizar métodos aproximados para obtenção da distribuição a posteriori e respectivo intervalo de credibilidade. Neste trabalho utilizaremos o algoritmo *Sampling Importance Resampling (SIR)*.

3.6 Sampling Importance Resampling (SIR)

Para obter a distribuição de probabilidade a posteriori de β e θ dada pela Eq. (2) utilizaremos um método não iterativo de Monte Carlo chamado *Sampling Importance Resampling (SIR)* e discutido em detalhes por *Rubin (1988)* e *Smith & Gelfand (1992)*.

Uma amostra aleatória de tamanho k é retirada da densidade a priori $\Phi(\beta; \theta)$ em que k é recomendado ser 5000 ou mais. Os pontos são indexados por $i = 1, 2, \dots, k$. Uma função de peso $w_i = 1, 2, \dots, k$, é associada a cada ponto amostrado, onde w_i é proporcional ao lado direito Eq.1. Na última etapa, uma amostra aleatória de tamanho k é retirada da densidade a priori com pesos w_i . A amostra gerada com esse esquema constitui-se aproximadamente em uma amostra da distribuição a posteriori dada pela equação Eq. 2. Assim, podemos calcular aproximadamente as medidas usuais de inferência bayesiana. Por exemplo, o estimador pontual bayesiano pode ser a média, mediana ou moda da amostra gerada da distribuição a posteriori. O intervalo de credibilidade com probabilidade $1 - \alpha$, com limite superior L_s e limite inferior L_i , é expresso pelos percentis $100 (\alpha/ 2) \%$ e $100 (1-\alpha / 2) \%$ ou pelo intervalo de credibilidade de amplitude mínima (com credibilidade $1-\alpha$).

Para o intervalo de credibilidade de mínima amplitude podemos utilizar o algoritmo de Chen-Shao descrito em [6]:

1. Obter uma amostra para SIR $\{\theta_i, i = 1, 2, \dots, n\}$ para 3.3.
2. Classificar $\{\theta_i, i = 1, 2, \dots, n\}$ para obter os valores ordenados:

$$\theta_{(1)} \leq \theta_{(2)} \leq \dots \leq \theta_{1(n)}$$
3. Calcular o intervalo de credibilidade $100(1 - \alpha)\%$.

$$R_j(n) = (\theta_j), \theta_{\{j+[(1-\alpha)n]\}} \text{ para } j = 1, 2, \dots, n - [(1 - \alpha)n].$$

4. O intervalo de mínima amplitude $100(1 - \alpha)\%$ é o intervalo denominado por $R_j. (n)$, com menor largura de intervalos entre todos os intervalos possíveis com credibilidade $100(1 - \alpha)\%$.

4 APLICAÇÃO

Durante a ocorrência de alguma anomalia apresentada pela locomotiva que estiver operando em trem ou manobra. A central atendimento dotada de nome Help-Desk realiza o atendimento emergencial com o propósito de restabelecer completa ou parcialmente as funcionalidades da locomotiva que apresentou a falha; caso não seja possível a correção ou atenuação do problema, o Help - Desk determina o reboque da locomotiva e recomenda que a mesma seja direcionada para oficina. É realizado registro dos detalhes da falha no sistema informatizado conforme tabulado no anexo A.

Quando a locomotiva chega nas oficinas as informações do atendimento é consultado e realizado o reparo mínimo necessário para retornar as funcionalidades da locomotiva.

Para a exemplificação numérica foram coletadas as informações de falhas de 196 locomotivas modelo *Dash9W* da fabricante *GE* com truncamento por tempo em 815 dias, conforme gráficos:

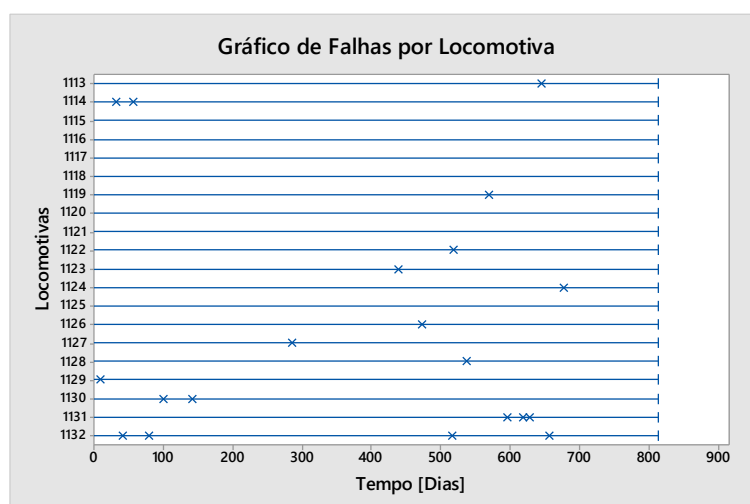


Gráfico 1: Gráfico de falhas por Locomotivas 1113 - 1132

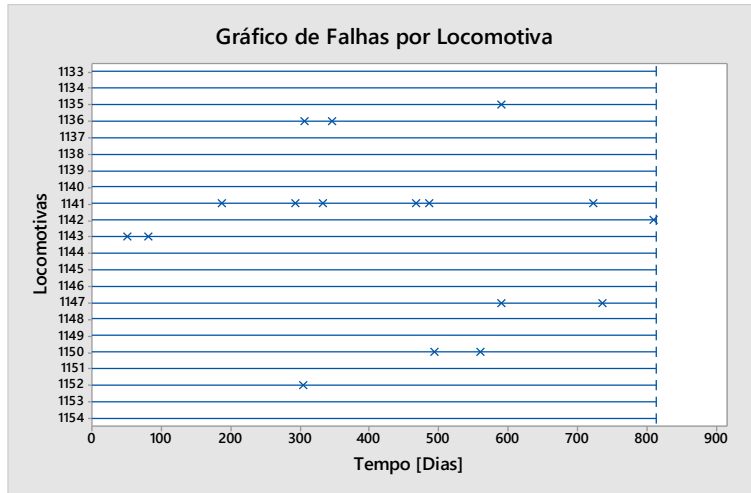


Gráfico 2: Gráfico de falhas por Locomotivas 1133 - 1154

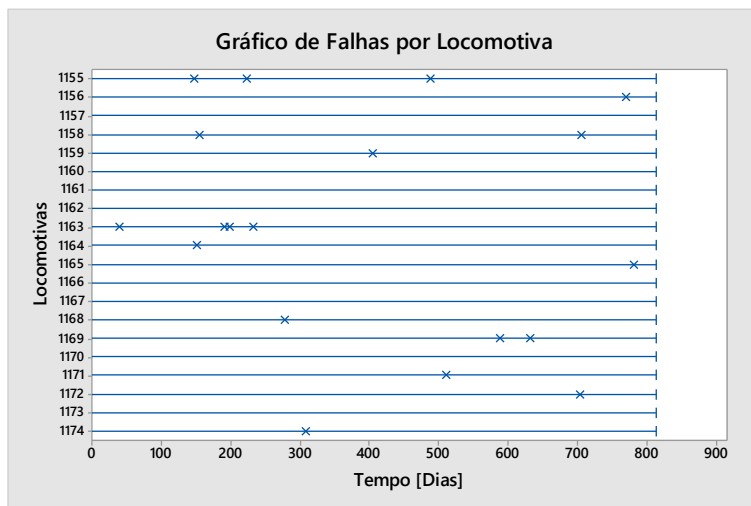


Gráfico 3: Gráfico de falhas por Locomotivas 1155 - 1174

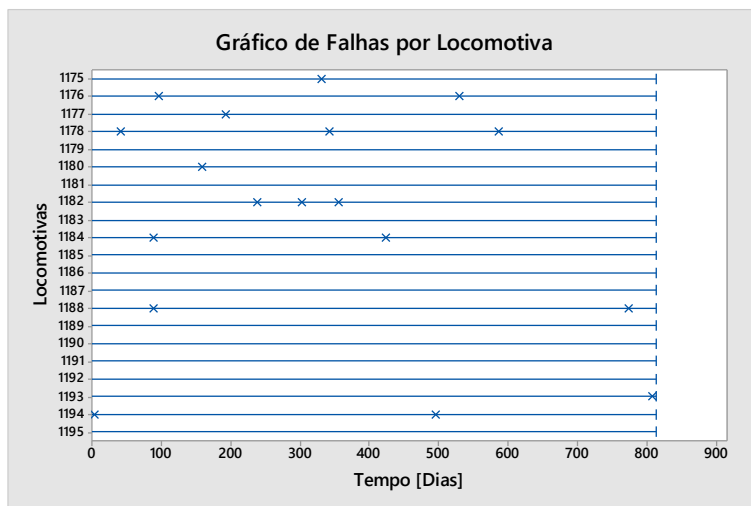


Gráfico 4: Gráfico de falhas por Locomotivas 1175 - 1195

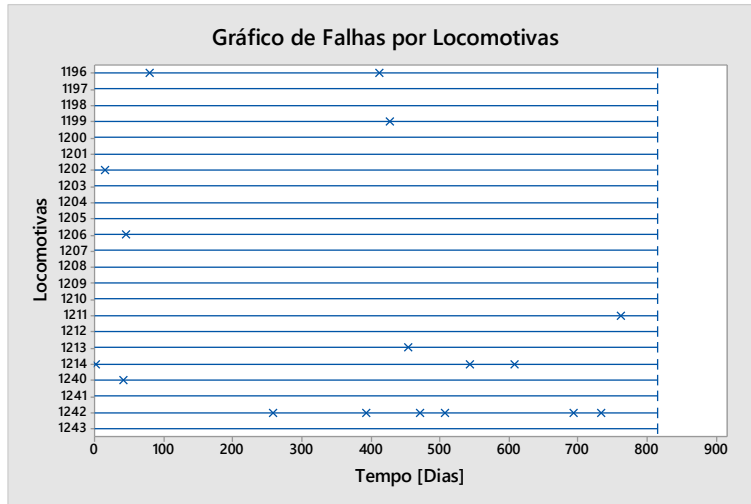


Gráfico 5: Gráfico de falhas por Locomotivas 1196 - 1243

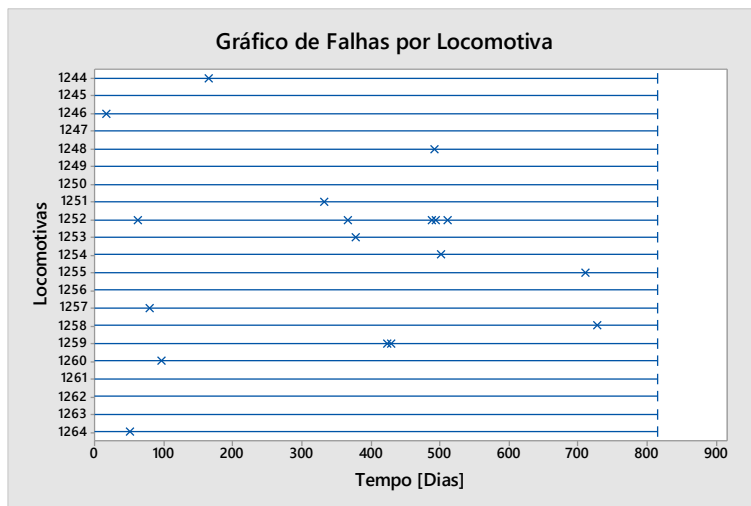


Gráfico 6: Gráfico de falhas por Locomotivas 1244 - 1264

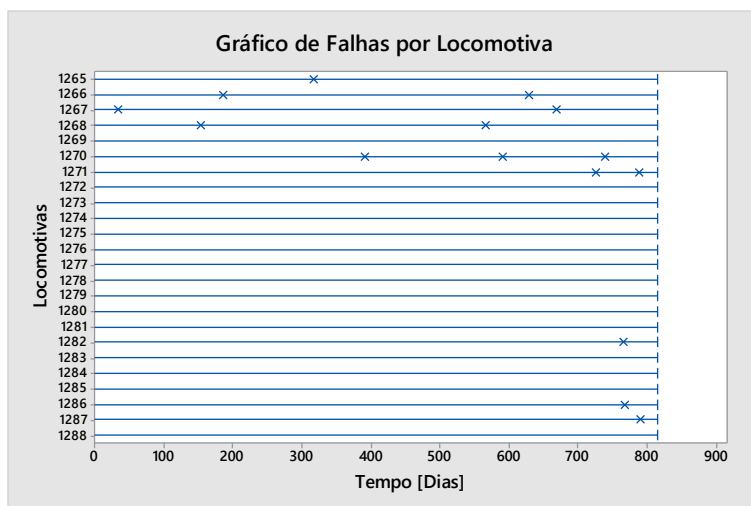


Gráfico 7: Gráfico de falhas por Locomotivas 1265 - 1288

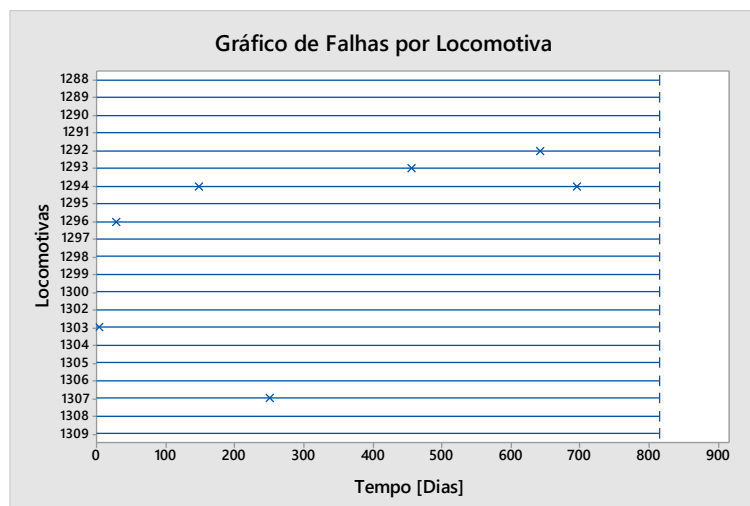
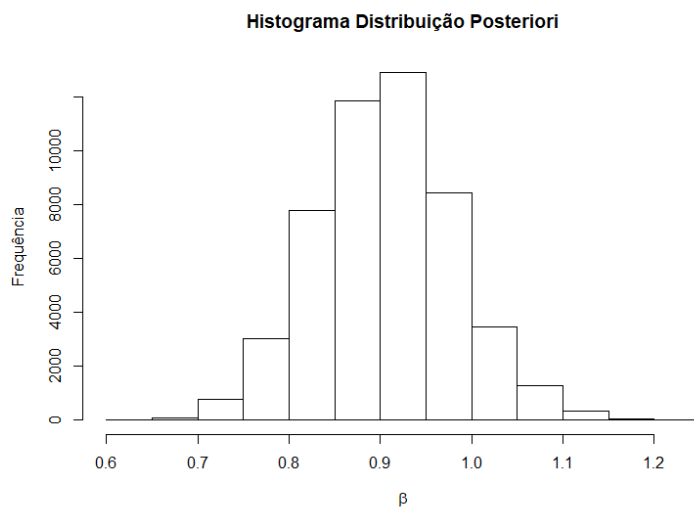
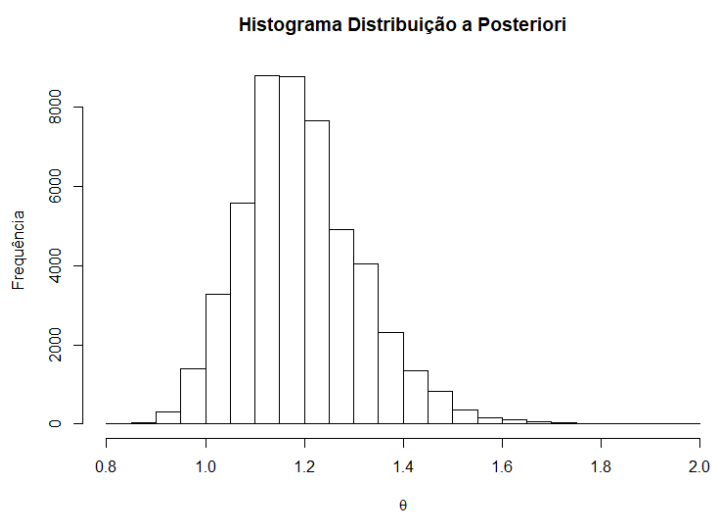


Gráfico 8: Gráfico de falhas por Locomotivas 1288 - 1309

Contudo, o gráfico de falhas por locomotivas exibe falhas iniciais em que ocorre em intervalos pequenos e falhas posteriores ocorrem em intervalos maiores, o que indicaria como melhoria (redução) da taxa de falha, ou início de falhas ocorrendo em intervalos grandes e falhas posteriores ocorrendo em intervalos menores, o que indicaria uma taxa de falha de degradação (crescente).

Utilizando a macro construída no programa R, disponibilizada no Anexo B, a distribuição a posteriori para β e θ , com distribuição a priori respectivamente $(\rho_\beta = 0,2, \alpha_\beta = 1,25)$ e $(\rho_\theta = 500, \alpha_\theta = 2000)$ foi gerada. Para evitar valores da função de verossimilhança muito pequenos (quando substituímos nesta função os valores da priori) trabalhamos com $\theta/1000$ (ou seja, mudamos a unidade do tempo das falhas). Com os valores das falhas construímos a função de verossimilhança e aplicamos o algoritmo SIR com uma amostra a priori de tamanho 50 mil para β e θ para assim obtermos a distribuição a posteriori. As Figuras 1 e 2 ilustram respectivamente a distribuição a posteriori para β e θ .

Gráfico 9: Histograma Posteriori de β Gráfico 10: Histograma Posteriori de $\theta/1000$

A Tabela 1 apresenta as estimativas a posteriori para os parâmetros $\hat{\beta}$ e $\hat{\theta}/1000$ (média, mediana e moda) enquanto a Tabela 2 apresenta os intervalos de credibilidade.

	Média	Mediana	Moda
$\hat{\beta}$	0,905128	0,903063	0,9125
$\hat{\theta}/1000$	1,196874	1,185967	1,147

Tabela 1: Estimativas pontuais para $\hat{\beta}$ e $\hat{\theta}/1000$.

	Percentílico		Mínima Amplitude	
	L_i	L_s	L_i	L_s
$\hat{\beta}$	0,761	1,060	0,757	1,056
$\hat{\theta}/1000$	0,990	1,465	0,973	1,442

Tabela 2: Intervalos de Credibilidade para $\hat{\beta}$ e $\hat{\theta}/1000$.

Observando a Tabela 2 não podemos rejeitar a hipótese de que $H_0:\beta = 1$ uma vez que o intervalo de credibilidade inclui o valor 1. Assim, utilizaremos o modelo PPH ($\beta = 1$) o que nos leva a estimar apenas o parâmetro θ . Utilizando os dados do Anexo A e o modelo PPH a distribuição a posteriori de θ é ilustrada na Figura 3.

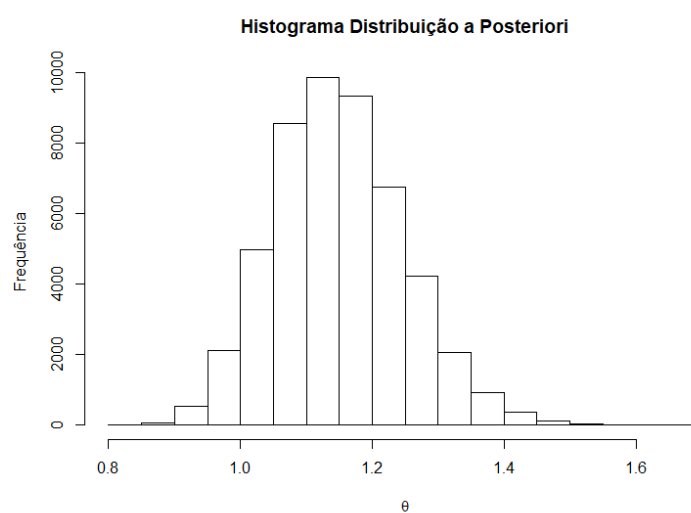


Gráfico 11: Histograma a posteriori de $\theta/1000$

A Tabela 3 apresenta as estimativas a posteriori para os parâmetros β e $\hat{\theta}/1000$ (média, mediana e moda) enquanto a Tabela 4 apresenta os intervalos de credibilidade.

	Média	Mediana	Moda
$\hat{\theta}/1000$	1,149	1,144	1,133

Tabela 3: Estimativa pontual para $\theta/1000$

	Percentílico		Mínima Amplitude	
	L_1	L_5	L_1	L_5
$\hat{\theta}/1000$	0,972	1,355	0,964	1,344

Tabela 4: Intervalos de Credibilidade para $\theta/1000$

Como trata-se de um processo PPH (Processo de Poisson Homogêneo com função intensidade de falhas constante) o número esperado de falhas no tempo $T=30$ dias para as 196 locomotivas será dado por $196 \times \frac{T}{\theta} = 196 \times \frac{30}{1133} \cong 5,19$ falhas. Optamos em usar a estimativa para θ como sendo a moda a posteriori uma vez que foi o resultado que mais aproximou o número de falhas estimado com o observado (5 falhas estimado para as 196 locomotivas com no período de 30 dias. Com os dados obtidos o trabalho futuro seria avaliar economicamente uma política de manutenção.

5 CONCLUSÃO

Neste trabalho consideramos inicialmente um modelo de Poisson não homogêneo (PPNH) para explicar as falhas de 196 locomotivas das marcas *Dash9W* da fabricante *General Electric-GE*. Após sua avaliação por meio de inferência Bayesiana constatamos que o modelo mais adequado/parcimonioso seria o processo de Poisson Homogêneo (PPH). Com os dados obtidos o trabalho futuro seria avaliar economicamente uma política de manutenção.

Outra possibilidade de trabalhos futuros seria adotar um modelo que considere reparos imperfeitos pois o modelo PPH observado pode ser na realidade uma consequência dos reparos imperfeitos que não são observáveis. Entre as classes de reparos imperfeitos sugerimos os modelos de idade virtual como proposto por Doyen&Gaudoin (2004).

REFERÊNCIAS

1. BARLOW, R; HUNTER,L. **Optimum preventive maintenance policies**. Operations Research, v.8, n. 1, p.90-100,1960.
2. DOYEN, L. and GAUDOIN O. (2004). Classes of imperfect repair models based on reduction of failure intensity or virtual age. Reliability Engineering & System Safety, 84:45–56.
3. FILHO, GIL BRANCO. **Dicionário de Termos de Manutenção, Confiabilidade e Qualidade**. Rio de Janeiro:Editora Ciência Moderna Ltda. p. 48 ,2006.
4. GILARDONI, Gustavo.; DE OLIVEIRA, Maristela D.; COLOSIMO, Enrico A. **Optimal maintenance time for repairable systems. Journal of quality Technology, v39, n.1, p. 48-53, 2007**
5. KARDEC, ALAN. **Manutenção-Função estratégia/** Alan Kardec, Júlio Nascif. – 4° ed. – Rio Janeiro: Qualitymark Editora, p. 80—88, 2013.
6. KRUSCHKE, J. K. 2011. **Doing Bayesian data analysis: a tutorial with R and BUGS**. Elsevier, Amsterdam, section 3.3.5.
7. MEEKER, W. Q.; ESCOBAR, Luis A. **Statistical Methods for Reliability dData**. A. Wiley Interscience Publications, 1998
8. RIGDON, S.E.; BASU, A. P. **Statiscal methods for thr reliability of repairable systems**. 2000.
9. STEPHENS,KENNETH S. **Reliability data analysisi With Excel and Minitab/** Kenneth S.Stephens. p.cm. Includes bibliographical references and index.ISBN 978-0-87389-831-7 (hard cover:alk.paper) 1-Reliability

(Engineering)-Statiscal Methods. 2.Realiability (Engineering)-
Dataprocesssing. 3.Quality control-Data processing.4.Microsoft Excel
(computer file).5.Minitab;1.Tide, Capther 11

10. VERSSANI, Bruna Aparecida Wruck. Modelo de regressão para sistemas repáreveis: um estado da confiabilidade de colhedoras de cana-de-açúcar/Bruna Aparecida Wruck Verssani.—versão revisada de acordo com a resolução CoPGr 6018 de 2011.—Piracicaba, 2018.

ANEXO A: Base de Dados

n	Ativo	Trucamento	Tempo [Dias]	n	Ativo	Trucamento	Tempo [Dias]
1	1113	1	646	94	1202	1	16
	1113	0	815		1202	0	815
	1114	1	32		95	1203	1
2	1114	1	57	96	1204	0	815
	1114	0	815	97	1205	0	815
3	1115	0	815	98	1206	1	45
4	1116	0	815	98	1206	0	815
5	1117	0	815		99	1207	0
6	1118	0	815	100	1208	0	815
7	1119	1	570	101	1209	0	815
	1119	0	815	102	1210	0	815
8	1120	0	815	103	1211	1	762
9	1121	0	815		1211	0	815
10	1122	1	518	104	1212	0	815
	1122	0	815	105	1213	1	454
11	1123	1	439		1213	0	815
	1123	0	815	1214	1	2	
12	1124	1	677	106	1214	1	544
13	1124	0	815		1214	1	608
14	1125	0	815	1214	1	815	
	1126	1	473	1214	0	815	
15	1126	0	815	107	1215	0	815
	1127	1	285	108	1216	0	815
16	1127	0	815		1217	1	290
	17	1128	1	538	109	1217	1
1128		0	815	1217		0	815
18	1129	1	9	110	1218	0	815
	1129	0	815	111	1219	0	815
19	1130	1	101	112	1220	0	815
	1130	1	142	113	1221	1	666
	1130	0	815		1221	0	815
20	1131	1	597	114	1222	0	815
	1131	1	620	115	1223	1	300
	1131	1	629		1223	0	815
	1131	0	815	116	1224	0	815
	1132	1	42	117	1225	1	815
1132	1	80	1225		0	815	
21	1132	1	517	1226	1	475	
	1132	1	657	118	1226	1	588
	1132	0	815		1226	1	619

22	1133	0	815		1226	0	815
23	1134	0	815	119	1227	0	815
24	1135	1	590	120	1228	0	815
25	1136	1	307	121	1229	0	815
26	1136	1	347	122	1230	0	815
27	1136	0	815	123	1231	0	815
28	1137	0	815	124	1232	0	815
29	1138	0	815	125	1233	0	815
30	1139	0	815	126	1234	0	815
31	1140	0	815	127	1235	0	815
32	1141	1	188		1236	1	601
	1141	1	293	128	1236	0	815
	1141	1	334		1237	1	610
	1141	1	467	129	1237	0	815
33	1141	1	486	130	1238	0	815
	1141	1	724		1239	1	276
	1141	0	815	131	1239	0	815
	1142	1	810		1240	1	41
34	1142	0	815	132	1240	0	815
	1143	1	51	133	1241	0	815
35	1143	1	82		1242	1	259
	1143	0	815		1242	1	394
36	1144	0	815		1242	1	471
37	1145	0	815	134	1242	1	507
38	1146	0	815		1242	1	694
	1147	1	591		1242	1	733
39	1147	1	737		1242	0	815
	1147	0	815	135	1243	0	815
40	1148	0	815		1244	1	165
41	1149	0	815	136	1244	0	815
	1150	1	494	137	1245	0	815
42	1150	1	560		1246	1	18
	1150	0	815	138	1246	0	815
43	1151	0	815	139	1247	0	815
44	1152	1	305		1248	1	493
	1152	0	815	140	1248	0	815
45	1153	0	815	141	1249	0	815
46	1154	0	815	142	1250	0	815
	1155	1	147		1251	1	333
	1155	1	224	143	1251	0	815
47	1155	1	489		1252	1	63
	1155	0	815		1252	1	366
	1156	1	770		1252	1	489
48	1156	0	815	144	1252	1	494
49	1157	0	815		1252	1	511
50	1158	1	155		1252	0	815

	1158	1	706		145	1253	1	378
	1158	0	815			1253	0	815
51	1159	1	406		146	1254	1	501
	1159	0	815			1254	0	815
52	1160	0	815		147	1255	1	711
53	1161	0	815			1255	0	815
54	1162	0	815		148	1256	0	815
	1163	1	40		149	1257	1	80
	1163	1	191			1257	0	815
55	1163	1	199		150	1258	1	728
	1163	1	233			1258	0	815
	1163	0	815			1259	1	424
56	1164	1	151		151	1259	1	429
	1164	0	815			1259	0	815
57	1165	1	781		152	1260	1	97
	1165	0	815			1260	0	815
58	1166	0	815		153	1261	0	815
59	1167	0	815		154	1262	0	815
60	1168	1	279		155	1263	0	815
	1168	0	815		156	1264	1	51
	1169	1	589			1264	0	815
61	1169	1	632		157	1265	1	317
	1169	0	815			1265	0	815
62	1170	0	815		158	1266	1	186
63	1171	1	511			1266	1	630
	1171	0	815			1266	0	815
64	1172	1	705			1267	1	34
	1172	0	815		159	1267	1	670
65	1173	0	815			1267	0	815
66	1174	1	309			1268	1	154
	1174	0	815		160	1268	1	566
67	1175	1	331			1268	0	815
	1175	0	815		161	1269	0	815
	1176	1	96			1270	1	391
68	1176	1	530		162	1270	1	591
	1176	0	815			1270	1	739
69	1177	1	194			1270	0	815
	1177	0	815			1271	1	727
	1178	1	42		163	1271	1	788
70	1178	1	343			1271	0	815
	1178	1	587		164	1272	0	815
	1178	0	815		165	1273	0	815
71	1179	0	815		166	1274	0	815
72	1180	1	159		167	1275	0	815
	1180	0	815		168	1276	0	815
73	1181	0	815		169	1277	0	815

	1182	1	238	170	1278	0	815
74	1182	1	302	171	1279	0	815
	1182	1	356	172	1280	0	815
	1182	0	815	173	1281	0	815
75	1183	0	815	174	1282	1	766
	1184	1	89		1282	0	815
76	1184	1	425	175	1283	0	815
	1184	0	815	176	1284	0	815
77	1185	0	815	177	1285	0	815
78	1186	0	815	178	1286	1	768
79	1187	0	815		1286	0	815
	1188	1	89	179	1287	1	791
80	1188	1	774		1287	0	815
	1188	0	815	180	1288	0	815
81	1189	0	815	181	1289	0	815
82	1190	0	815	182	1290	0	815
83	1191	0	815	183	1291	0	815
84	1192	0	815	184	1292	1	642
	1193	1	808		1292	0	815
85	1193	0	815	185	1293	1	457
	1194	1	3		1293	0	815
86	1194	1	496		1294	1	148
	1194	0	815	186	1294	1	695
87	1195	0	815		1294	0	815
	1196	1	80	187	1295	0	815
88	1196	1	413	188	1296	1	29
	1196	0	815		1296	0	815
89	1197	0	815	189	1297	0	815
90	1198	0	815	190	1298	0	815
91	1199	1	427	191	1299	0	815
	1199	0	815	192	1300	0	815
92	1200	0	815	193	1302	0	815
93	1201	0	815	194	1303	1	3
					1303	0	815
				195	1304	0	815
				196	1305	0	815

ANEXO B: Programação Software R

Abaixo tem-se a programação completa da linguagem R, utilizado para encontrar os resultados e breves comentários sobre algumas funções utilizadas nos cálculos. Código abaixo foi simulado um Notebook Windows 10 Pro, processador Intel® Core™ i5-2450M [CPU@2.55GHz](#), Sistema Operacional de 64bits, com o aplicativo RStudio versão 3.6.3 (2020-02-29).

Intervalo de Credibilidade para o Número Médio de Falhas em Sistemas Reparáveis Utilizando a Lei Potência

```
library(pracma)      # Para simular comandos do Matlab
library(latex2exp)  # Permitir escrever letras Gregas, etc
library(HDInterval) # Intervalo de Credibilidade de mínima amplitude
tic()                # A função tic inicia o temporizador, função Matlab (tic e
toc);
clear()              # Remova esses ou todos os itens do espaço de
trabalho, ou seja, o ambiente global e liberando a memória do sistema.
```

#Lendos dados - Adapte para o seu diretório.#

```
library('readxl')
R <- read_excel ("C:/Users/Cassio/Desktop/Simulação/R.xlsx")
n=337                # Número de Informações banco de dados, 337
eventos.
corridas=50000      # Quantidade de simulação;
R$T1=R$T1/1000      # Para Evitar números muito pequenos na Vero;
R$T2=R$T2/1000      # Para Evitar números muito pequenos na Vero.

#Verossimilhanca
VERO<-matrix(0,n,1) # Local para armazenar os resultados da Vero
VERO
```

```

VE <- function(delta){ .....#Uma função para calcular a Vero
depois. Dellta=(Beta,Teta)

  for (i in 1:n) {

    j=R$C[i]

    if (j==1) {

VERO[i,1] <- (delta[1]/delta[2]) *(R$T2[i]/delta[2])^(delta[1]-1)*
      exp(-(R$T2[i]/delta[2])^delta[1] + (R$T1[i]/delta[2])^delta[1])
    }else{

      VERO[i,1] <- exp(-(R$T2[i]/delta[2])^delta[1] + (R$T1[i]/delta[2])^delta[1])
    }

  }

  SV <- sum(log(VERO))

  return(SV)

}

PostBT=matrix(0,corridas,6) #mariz de valor 0, n corridas, colunas=6)

PostBT

for (k in 1:corridas){

  k1=runif(1,0.2,1.25) # Beta

  # k1=1          # Em caso de PHH (ou seja, beta=1) deixar ativo.

  k2=runif(1,0.5,2)  # Teta

  PostBT[k,1]=k1    # Coluna [,1] Matrix PostBT, recebe os n°s gerados de
Beta 0.2~1.25 em caso PHH Beta=1.

  PostBT[k,2]=k2    # Coluna [,2] Matrix PostBT, recebe os n°s gerados de
Teta 0.5~ 2.

  PostBT[k,3]=exp(VE(c(k1,k2)))
#Função da VEROSSIMILHANÇA - VE- Função Delta. Coluna[,3] matrix
PostBT, recebe VERO(L).

}

PostBT

```

```

PostBT[,4]= PostBT[,3]/sum(PostBT[,3]) #Pesos:Valor de cada
VERO/SOMA

PostBT

Pesosoma<-sum(PostBT[,4])          # Conferido se a soma é 1, ok.

Amostra <-sample(1:corridas,corridas, TRUE, prob=PostBT[,4]) # Para
amostrar simultaneamente Beta e Teta

Amostra

for (y in 1:corridas) {

  kk=Amostra[y]

  PostBT[y,5]=PostBT[kk,1]

  PostBT[y,6]=PostBT[kk,2]

}

PostBT[y,5]

PostBT[y,6]

PostBT[,5]=sample(PostBT[,1],corridas,TRUE,prob=PostBT[,4])
#Posteriorioi de Beta

PostBT[,6]=sample(PostBT[,2],corridas,TRUE,prob=PostBT[,4])
#Posteriorioi de Teta

PostBT[,5]          # Valores da Posteriori de Beta

PostBT[,6]          # Valores da Posteriori de Teta

Betaest1=mean(PostBT[,5])    # média da Posteriori BETA

Betaest2=median(PostBT[,5])  # mediana da Posteriori BETA

h <- hist(PostBT[,5], breaks=512, main = "Distribuição a Posteoriori", xlab =
expression(beta), ylab=TeX(r'($\Beta | x)$'))

Betaest3 <- h$mids[h$counts == max(h$counts)]    # moda da Posteriori
BETA

cat("\n", "Betaest1/mean=",Betaest1,"\n", "Betaest1/median=",Betaest2,"\n", "Bet
aest3/mode=",Betaest3)

ICBeta=quantile(PostBT[,5],c(0.025,0.975))    # Intervalo de credibilidade .

```

```

ICBeta=hdi(PostBT[,5])      # Intervalo de Credibilidade de Mínima
Amplitude

cat("\n", "InterCredBetaLI=", ICBeta[[1]], "\n", "InterCredBetaLS=", ICBeta[[2]], "\n"
)

Tetaest1=mean(PostBT[,6])      # média da Posteriori TETA

Tetaest2=median(PostBT[,6])    # mediana da Posteriori TETA

h<- hist(PostBT[,6], breaks=512, main = "Distribuição a Posteoriori", xlab =
expression(theta), ylab=TeX(r'($f(\Teta | x)$'))

Tetaest3 <- h$mids[h$counts == max(h$counts)]      #moda da
Posteriori TETA

cat("\n", "Tetaest1/mean=", Tetaest1, "\n", "Tetaest2/median=", Tetaest2, "\n", "Tetaest3/mode=", Tetaest3)

ICTeta=quantile(PostBT[,6],c(0.025,0.975))
#Intervalo de credibilidade

ICTeta=hdi(PostBT[,6])
#Intervalo de Credibilidade de Mínima Amplitude

cat("\n", "InterCredTetaLI=", ICTeta[[1]], "\n", "InterCredTetaLS=", ICTeta[[2]], "\n")

write.csv(PostBT, "PostBT.csv")

T=30      # Calculo do número médio de defeitos até Tempo desejado -
NMDT

nL=196    #Número de Locomotivas

NMDT=(T/(PostBT[,6]*1000))^(PostBT[,5])*nL #Número médio de Falhas até
T

NMDTest=mean(NMDT)

ICNMDT=quantile(NMDT,c(0.025,0.975))

cat("\n", "NMDTest=", NMDTest,
"\n", "InterCredNMDTLI=", ICNMDT[[1]], "\n", "InterCredNMDTLS=", ICNMDT[[2]],
"\n")

toc() #TOC atualiza o tempo decorrido desde que o temporizador foi iniciado

```

