



UNIVERSIDADE FEDERAL DE SANTA CATARINA  
CENTRO DE TECNOLÓGICO DE JOINVILLE  
CURSO DE GRADUAÇÃO EM ENGENHARIA NAVAL

Jacy Queiroz da Fonseca

**MÉTODO PARA ESTIMAÇÃO DA TAXA DE FALHA DE  
EQUIPAMENTOS POR MEIO DE ATUALIZAÇÃO BAYESIANA: UM  
ESTUDO COM BASE NOS DADOS DE CONFIABILIDADE ONSHORE E  
OFFSHORE DO OREDA**

Joinville  
2021

Jacy Queiroz da Fonseca

**MÉTODO PARA ESTIMAÇÃO DA TAXA DE FALHA DE  
EQUIPAMENTOS POR MEIO DE ATUALIZAÇÃO BAYESIANA: UM  
ESTUDO COM BASE NOS DADOS DE CONFIABILIDADE ONSHORE E  
OFFSHORE DO OREDA**

Trabalho de Conclusão de Curso do Curso de Graduação em Engenharia Naval do Centro de Tecnológico de Joinville da Universidade Federal de Santa Catarina para a obtenção do título de Bacharel em Engenharia Naval.  
Orientador: Prof. Dr. Luis Fernando Peres Calil

Joinville  
2021

Ficha de identificação da obra elaborada pelo autor,  
através do Programa de Geração Automática da Biblioteca Universitária da UFSC.

Fonseca, Jacy Queiroz da  
MÉTODO PARA ESTIMAÇÃO DA TAXA DE FALHA DE EQUIPAMENTOS  
POR MEIO DE ATUALIZAÇÃO BAYESIANA: UM ESTUDO COM BASE NOS  
DADOS DE CONFIABILIDADE ONSHORE E OFFSHORE DO OREDA / Jacy  
Queiroz da Fonseca ; orientador, Luis Fernando Peres  
Calil, 2021.  
47 p.

Trabalho de Conclusão de Curso (graduação) -  
Universidade Federal de Santa Catarina, Campus Joinville,  
Graduação em Engenharia Naval, Joinville, 2021.

Inclui referências.

1. Engenharia Naval. 2. Taxa de Falha. 3. Atualização  
Bayesiana. 4. OREDA. I. Peres Calil, Luis Fernando . II.  
Universidade Federal de Santa Catarina. Graduação em  
Engenharia Naval. III. Título.

Jacy Queiroz da Fonseca

**MÉTODO PARA ESTIMAÇÃO DA TAXA DE FALHA DE  
EQUIPAMENTOS POR MEIO DE ATUALIZAÇÃO BAYESIANA: UM  
ESTUDO COM BASE NOS DADOS DE CONFIABILIDADE ONSHORE E  
OFFSHORE DO OREDA**

Este Trabalho de Conclusão de Curso foi julgado adequado para obtenção do Título de “Bacharel em Engenharia Naval” e aprovado em sua forma final pelo Curso de Graduação em Engenharia Naval.

Joinville, 28 de setembro de 2021.

---

Prof. Dr. Luis Fernando Peres Calil  
Coordenador do Curso

**Banca Examinadora:**

---

Prof. Dr. Luis Fernando Peres Calil  
Orientador

---

Prof. Maurício de Campos Porath, Dr. Ing  
Avaliador  
Universidade Federal de Santa Catarina –  
UFSC

---

Prof. Eduardo Yuji Sakurada, Dr. Eng.  
Avaliador  
Instituto Federal de Educação, Ciência e  
Tecnologia de Santa Catarina – IFSC

Este trabalho é dedicado a Barbara e Joseci, meus pais.

## AGRADECIMENTOS

À Deus, pela dádiva da vida.

À minha mãe, Barbara, por toda a dedicação e amor imensuráveis durante toda a minha vida. Obrigada por sempre estar presente.

Ao meu pai, Joseci, pelo amor, pela inspiração e troca de experiências que orientaram minha formação como indivíduo e como profissional.

A toda minha família da Bahia, em especial meus avós Astério, Balbina e minha vó Zene.

Ao meu amado, Gabriel, pelo companheirismo e paciência inabalável.

A todos os meus companheiros de graduação, pela amizade e parceria. Em especial Leonardo Vinicius, Guilherme de Paula, Kallyne, Arthur, Kamila, Bruno Mamoru, João Victor Fabri e Pedro César.

Aos amigos do pedacinho do nordeste, Gabriel Leal, Antônio Carlos Júnior e Lucas Salviano, obrigada por todas as músicas, noites em claro estudando (ou em festas) e cuscuz saboreados.

Aos meus amigos do clube da comida: Mari, Pieper, Dani e Rudi obrigada por cada receita e parceria.

Aos professores, que guiaram minha jornada acadêmica.

Ao meu orientador, professor e amigo, Calil, por ter me ajudado em inúmeros momentos durante minha graduação, obrigada por me orientar e por todo suporte, dedicação e auxílio.

*“O acaso favorece  
a mente preparada”  
(Louis Pasteur)*

## RESUMO

Poucos setores apresentam um grau de criticidade na manutenção industrial tão elevado quanto o setor de óleo e gás. Indubitavelmente qualquer alteração que gere uma possível perda econômica deve ser criteriosamente avaliada. A proposta do presente trabalho é aplicar uma metodologia simples e iterativa para estimar a taxa de falha de equipamentos onshore e offshore, com base nos seus históricos de falhas. Para tal, é proposta uma abordagem orientada pela atualização bayesiana, a fim de gerar uma estimativa de taxa de falha pautada em dados a priori, bem como permitir atualizações iterativas a partir da aquisição de novos dados operacionais. O histórico de falhas são oriundos do projeto OREDA (Offshore and Onshore Reliability Data), uma base de dados que possui informações sobre operação e manutenção de diversos equipamentos da indústria do petróleo, utilizada como referência internacional para cálculos de confiabilidade. A metodologia aplicada se mostrou eficiente na obtenção de estatística para taxa de falha com menor variação, suprimindo uma carência do setor.

**Palavras-chave:** Taxa de Falha. Atualização Bayesiana. OREDA.



## **ABSTRACT**

Few sectors have a degree of criticality in industrial maintenance as high as the oil and gas field. Undoubtedly any change that generates a possible economic loss should be carefully evaluated. The main purpose of this work is to apply a simple and iterative methodology to estimate the failure rate of onshore and offshore equipment, based on fault history. So, an approach guided by Bayesian Update is proposed, in order to generate a prediction of failure rate based on a priori data, as well as allow iteratives update from the acquisition of new operational data. The failure history comes from the OREDA project (Offshore and Onshore Reliability Data), a database that has information on the operation and maintenance of various equipment in the petroleum industry, used as an international reference for reliability calculations. The applied methodology proved to be efficient in obtaining statistics for a failure rate with less variation, meeting the needs of the sector.

**Keywords:** Failure Rate. Bayesian Updating. OREDA.

## LISTA DE FIGURAS

Figura 1 – Espaço Amostral $S$ . . . . .	18
Figura 2 – Gráfico Função de Densidade de Probabilidade da Gama . . . . .	20
Figura 3 – Gráfico Função de Distribuição Acumulada da Gama . . . . .	21
Figura 4 – Gráfico da função densidade de probabilidade (esquerda) e função de distribuição (direita) de uma variável aleatória $\lambda$ . . . . .	22
Figura 5 – Fronteira Bombas . . . . .	27
Figura 6 – Variação Típica da Taxa de Falha em Componentes Mecânicos . . . . .	32
Figura 7 – Fluxograma Metodologia . . . . .	36

## LISTA DE QUADROS

Quadro 1 – Quadro utilizado na base de dados OREDA - Topside . . . . .	28
Quadro 2 – Dados de confiabilidade apresentados no OREDA relativos ao item Bomba. . . . .	37
Quadro 3 – Dados de confiabilidade da bomba, utilizando a distribuição Gama como base para o cálculo do intervalo de confiança. . . . .	38
Quadro 4 – Tempos fictícios até a falha gerados para ilustrar o método proposto. .	39
Quadro 5 – Comparação dos valores dos hiperparâmetros das Distribuições Gama a priori e a posteriori. . . . .	41

## **LISTA DE ABREVIATURAS E SIGLAS**

ABNT	Associação Brasileira de Normas Técnicas
ISO	International Organization for Standardization
MTTF	Mean Time To Failure
NBR	Norma Técnica brasileira
OREDA	Offshore and Onshore Reliability Data

## SUMÁRIO

<b>1</b>	<b>INTRODUÇÃO</b> . . . . .	<b>14</b>
1.1	OBJETIVOS . . . . .	15
1.1.1	<b>Objetivo geral</b> . . . . .	<b>15</b>
1.1.2	<b>Objetivos específicos</b> . . . . .	<b>15</b>
<b>2</b>	<b>ATUALIZAÇÃO BAYESIANA</b> . . . . .	<b>16</b>
2.1	PROBABILIDADE E SUAS PROPRIEDADES . . . . .	16
2.1.1	<b>Probabilidade condicional</b> . . . . .	<b>17</b>
2.1.2	<b>Eventos mutuamente exclusivos</b> . . . . .	<b>17</b>
2.1.3	<b>Regra do produto</b> . . . . .	<b>17</b>
2.1.4	<b>Independência de eventos</b> . . . . .	<b>18</b>
2.1.5	<b>Teorema da probabilidade total</b> . . . . .	<b>18</b>
2.1.6	<b>Teorema de bayes</b> . . . . .	<b>19</b>
2.2	DISTRIBUIÇÕES ESTATÍSTICAS . . . . .	20
2.2.1	<b>Distribuição gama</b> . . . . .	<b>20</b>
2.2.2	<b>Distribuição exponencial</b> . . . . .	<b>21</b>
2.2.3	<b>Distribuição qui-quadrado</b> . . . . .	<b>22</b>
2.3	FUNDAMENTOS DA ATUALIZAÇÃO BAYESIANA . . . . .	23
2.4	FUNÇÃO VEROSSIMILHANÇA . . . . .	23
2.5	DISTRIBUIÇÕES CONJUGADAS . . . . .	24
<b>3</b>	<b>BASE DE DADOS DE CONFIABILIDADE</b> . . . . .	<b>26</b>
3.1	CONFIABILIDADE . . . . .	26
3.2	COLETA DE DADOS . . . . .	26
3.3	OFFSHORE AND ONSHORE RELIABILITY DATA (OREDA) . . . . .	28
3.3.1	<b>Procedimentos de estimação</b> . . . . .	<b>30</b>
3.3.2	<b>Estimador da taxa de falha e intervalo de confiança</b> . . . . .	<b>32</b>
3.3.2.1	Amostras homogêneas . . . . .	33
3.3.2.2	Amostras não homogêneas . . . . .	33
<b>4</b>	<b>MÉTODO PROPOSTO PARA ATUALIZAÇÃO BAYESIANA</b> . . . . .	<b>36</b>
4.1	DEFINIR A DISTRIBUIÇÃO A PRIORI . . . . .	37
4.2	DEFINIR HIPERPARÂMETROS . . . . .	38
4.3	OBTER A VEROSSIMILHANÇA . . . . .	39
4.4	CALCULAR A POSTERIORI . . . . .	39
4.5	CALCULAR A ESTIMATIVA DA TAXA DE FALHA . . . . .	40
4.6	ANÁLISE DOS RESULTADOS . . . . .	41
<b>5</b>	<b>CONCLUSÕES E RECOMENDAÇÕES PARA TRABALHOS FUTU- ROS</b> . . . . .	<b>43</b>
	<b>REFERÊNCIAS</b> . . . . .	<b>44</b>

ANEXO A – VALORES PERCENTIS DA DISTRIBUIÇÃO QUI- QUADRADO . . . . .	47
--	----

## 1 INTRODUÇÃO

Desde o início da exploração de petróleo, grande atenção tem sido dada à segurança e confiabilidade dos processo e equipamentos, visto que a rentabilidade dessa indústria depende da confiabilidade, disponibilidade e manutenibilidade dos sistemas e componentes utilizados.(ABNT, 2017)

Por conta da elevada competitividade do mercado, eventos que possuem maior impacto econômico são objetos de estudo para diminuição do tempo de inatividade e reparo. Para estes, é importante conseguir estimar a probabilidade de falha dos equipamentos, sendo possível, desta forma, antever a disponibilidade esperada e, assim, obter uma estimativa melhor para o planejamento de manutenção.(SANDTORV; HOKSTAD; THOMPSON, 1996)

“O custo anual da indústria para a falta de confiabilidade de equipamentos é muito alto, embora recentemente, uma parcela maior da indústria tem dado maior atenção ao projeto e manutenção eficazes” (ABNT, 2011, p. XIV). Segundo Sandtorv, Hokstad e Thompson (1996), dados históricos sobre a frequência de falha e suas respectivas natureza, são essenciais para o desempenho das análises e avaliações.

Nesse contexto, surgiu o projeto OREDA (Offshore and Onshore Reliability Data – Base de dados de confiabilidade de equipamentos offshore e onshore), uma base de dados de equipamentos da indústria do petróleo, com informações operacionais e ambientais, além de dados sobre manutenção preventiva e corretiva. O projeto é comumente utilizado como referência para os cálculos de probabilidade de falha de equipamentos onshore e offshore. De acordo com Langseth, Haugen e Sandtorv (1998) a análise dos dados de confiabilidade é um dos vários fatores-chave na escolha de soluções econômicas.

De acordo com Silva (2015, p. 12), “Os crescentes desafios que a indústria do óleo e gás têm se deparado no decorrer dos últimos anos influenciaram o desenvolvimento e o estudo de técnicas, numa visão probabilística, que quantificam e qualificam as incertezas”. A respeito da utilização do raciocínio probabilístico, uma distribuição de probabilidade relacionada às incertezas existentes, possibilita uma estimativa mais realista do tempo total necessário de vida de um equipamento, assim como sua probabilidade de falha.(SILVA, 2015)

Com o objetivo de obter a confiabilidade de equipamentos e tirar conclusões sobre a população de interesse, utiliza-se principalmente o modelo de probabilidade clássica e Bayesiana. Silva e Mattos (2001, p. 2) afirmam que “A abordagem clássica utiliza somente os dados de operação ou de teste do item para avaliar o seu desempenho, enquanto a abordagem Bayesiana permite combinar esses dados com outras informações relevantes já disponíveis.”

Como ferramenta para avaliar confiabilidade de equipamentos, a teoria bayesiana tem se destacado na indústria do petróleo por conta da dificuldade de obtenção de dados

de falha, seja por conta do elevado período de vida útil do equipamento ou dificuldade em reproduzir as condições de uso operacionais.

## 1.1 OBJETIVOS

Nas seções abaixo estão descritos os objetivos geral e específicos deste TCC.

### 1.1.1 Objetivo geral

O objetivo do presente trabalho é criar uma metodologia de estimativa da taxa de falha de equipamentos baseada em Atualização Bayesianas e na base de dados do OREDA.

### 1.1.2 Objetivos específicos

Os objetivos específicos desse trabalho são:

- Apresentar a base de dados OREDA numa linguagem acessível de modo a estimular seu uso nas universidades;
- Desenvolver um método simples e iterativo para obtenção da taxa de falha de equipamentos;
- Encontrar um intervalo de confiança com menor amplitude comparado com os disponíveis na base de dados OREDA;
- Exemplificar a aplicação do método com equipamentos amplamente utilizados na indústria do óleo e gás.



## 2 ATUALIZAÇÃO BAYESIANA

O raciocínio bayesiano é o mecanismo que tem a função de descrever a probabilidade de um evento acontecer a partir do conhecimento anterior (a priori). A fim de elucidar a teoria Bayesiana, algumas propriedades de probabilidades serão apresentadas a seguir.

### 2.1 PROBABILIDADE E SUAS PROPRIEDADES

Segundo Oliveira Costa Neto (2002) a probabilidade é um número associado a um evento destinado a medir a sua possibilidade de ocorrência.

No entanto, o conceito de probabilidade não é único e tem sido normalmente definido de diferentes pontos de vista filosóficos: a probabilidade clássica (a priori), geométrica, frequentista (a posteriori), subjetiva, entre outros.

A definição de probabilidade clássica segundo Magalhães (2006), se refere a subconjuntos unitários e equiprováveis. Desta forma a probabilidade de ocorrer o evento  $A$  é dado por:

$$P(A) = \frac{\text{Número de elementos } A}{\text{Número total de elementos}} \quad (1)$$

Um possível exemplo desta abordagem é com o lançamento de dados. Joga-se um dado, a probabilidade de se obter o número 2 será igual a  $\frac{1}{6}$ . Outro exemplo é o lançamento de uma moeda, neste caso a probabilidade de se obter o resultado “cara” é  $\frac{1}{2}$ .

A abordagem geométrica ocorre quando os elementos são dimensões: intervalos, medidas de áreas ou similares. Dantas (2013) afirma, que a probabilidade de que um ponto selecionado ao acaso, a partir de  $S$ , localize-se na região  $A$ , nele incluída, é dada pela razão:

$$P(A) = \frac{\text{Medida } A}{\text{Medida } S} \quad | \quad A \in S \quad (2)$$

A abordagem frequentista, considera um experimento que possa realizar repetições independentes, nas mesmas condições um número total  $N$  de vezes. De acordo com Gonçalves (2004) apud Dantas (2013), um evento  $A$  ocorre  $N(A)$  vezes ( $0 \leq N(A) \leq N$ ), então, a probabilidade da ocorrência de  $A$  é dada por:

$$P(A) = \frac{N(A)}{N} \quad (3)$$

A interpretação de probabilidade subjetiva ou Bayesiana é sugerida por Dantas (2013), como o grau de convicção ou crença que um indivíduo atribui à determinada ocorrência ou não acontecimento de um evento. Um exemplo de probabilidade subjetiva é

a previsão do tempo. Matematicamente a abordagem bayesiana está relacionada à ideia de verossimilhança.

### 2.1.1 Probabilidade condicional

A probabilidade condicional refere-se ao fato que um determinado evento já ocorreu, com isso, se altera a probabilidade atribuída a outro evento de interesse, possivelmente relacionado ao evento anterior (OLIVEIRA COSTA NETO, 2002). O conceito de probabilidade condicional permite considerar as novas informações a fim de calcular as novas probabilidades.

Denotamos como  $P(A|B)$  em que a barra vertical é interpretada como *dado* e a expressão completa exprime a probabilidade de ocorrência do evento  $A$ , sabendo que o evento  $B$  ocorreu, ou simplesmente, a probabilidade de  $A$  *dado*  $B$ . Demonstrado a seguir:

$$P(A|B) = \frac{P(A \cap B)}{P(B)} \quad (4)$$

### 2.1.2 Eventos mutuamente exclusivos

Dizemos que dois eventos são mutuamente exclusivos quando eles são disjuntos, ou seja, não possuem interseção.

$$(A \cap B) = \emptyset \quad (5)$$

Logo, o evento  $A$  e  $B$  não podem ocorrer em simultâneo. Assim, a probabilidade de uma união de dois eventos, que é dada por:

$$P(A \cup B) = P(A) + P(B) - P(A \cap B) \quad (6)$$

Pode ser simplificada para:

$$P(A \cup B) = P(A) + P(B) \quad (7)$$

### 2.1.3 Regra do produto

Uma das consequências da expressão da probabilidade condicional, apresentada na Equação 4, é a regra do produto. O cálculo da probabilidade interseção de dois evento,  $P(A \cap B)$  é dado por:

$$P(A \cap B) = P(A).P(B|A) \quad (8)$$

Para o caso de três ou mais eventos, a probabilidade conjunta seguiria de forma equivalente, a exemplo de:

$$P(A \cap B) = P(A).P(B|A).P(C|A \cap B) \quad (9)$$

#### 2.1.4 Independência de eventos

A independência de eventos é uma das propriedades probabilísticas utilizadas na teoria de Bayes. Denotada conforme equação abaixo:

$$P(A|B) = P(A) \text{ ou } P(B|A) = P(B) \quad (10)$$

Se os eventos  $A$  e  $B$  forem independentes, isso significa que saber da ocorrência de um evento não interfere na probabilidade de outro evento acontecer, a fórmula para o cálculo da probabilidade da intersecção (Equação 8) será simplificada para:

$$P(A \cap B) = P(A).P(B) \quad (11)$$

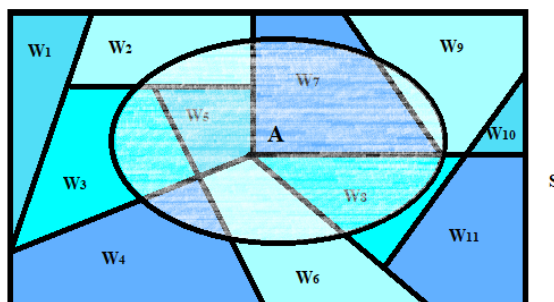
Podemos considerar que os eventos  $A$  e  $B$  são independentes quando a relação 10 é satisfeita.

Caso as variáveis  $A$  e  $B$  fossem dependentes, a probabilidade da intersecção entre os eventos poderia ser calculada através da Equação 8.

#### 2.1.5 Teorema da probabilidade total

Seja o espaço amostral  $S$  repartido em múltiplas partes  $W_n$  mutuamente exclusivos, conforme Figura 1.

Figura 1 – Espaço Amostral  $S$



Fonte: Figura elaborada para este trabalho.

Então, a probabilidade do evento de interesse  $A$ , dependente dos eventos repartidos  $W$  é obtida a partir da Equação 12.

$$P(A) = P(A \cap W_1) + P(A \cap W_2) + P(A \cap W_3) + P(A \cap W_4) + \dots + P(A \cap W_{11}) \quad (12)$$

Note que ela permite calcular a probabilidade de um evento  $A$  quando se conhece as probabilidades de um conjunto de eventos distintos, cuja reunião é o espaço amostral e as probabilidades condicionais de  $A$  dado cada um deles (DANTAS, 2013).

Assim, pode-se desenvolver as probabilidades conjuntas da Equação 12 e obter a representação clássica do Teorema da Probabilidade Total, apresentada na Equação 13.

$$P(A) = \sum_i^n P(W_i) \cdot P(A|W_i) \quad (13)$$

### 2.1.6 Teorema de bayes

O Teorema de Bayes é uma formulação para a avaliação de probabilidades, com base em nova informação amostral (KOEHLER *et al.*, 1998). É um mecanismo importante que permite relacionar diversas probabilidades conhecidas e deduzir uma probabilidade posterior, com melhor eficácia na estimativa dos resultados (VERONEZ *et al.*, 2000).

Isso significa que o Teorema de Bayes, na Equação 14 a seguir, permite calcular a probabilidade das possíveis causas de um evento, após se ter a informação de que outro evento  $B$  ocorreu.

$$P(A|B) = \frac{P(B|A) \cdot P(A)}{P(B)} \quad (14)$$

Em que:

- $P(A|B)$ : probabilidade de ocorrer o evento  $A$  dado a ocorrência do evento  $B$ ;
- $P(B|A)$ : probabilidade de ocorrer o evento  $B$  dado a ocorrência do evento  $A$ ;
- $P(A)$ : probabilidade de ocorrência do evento  $A$ ;
- $P(B)$ : probabilidade de ocorrência do evento  $B$ .

A probabilidade  $P(A)$  é chamada de probabilidade a priori e representa o conhecimento prévio (LOUREIRO, 2018), já  $P(A|B)$  é a probabilidade a posteriori e “[...] representa o que é conhecido a respeito de  $A$ , dado o conhecimento dos dados.” (BARROS, 2008, p. 29), subsequente à observação do evento  $B$ .

## 2.2 DISTRIBUIÇÕES ESTATÍSTICAS

Distribuições estatísticas ou de probabilidade descrevem como os valores de probabilidade são distribuídos associadas a todos os possíveis valores dessa variável (ASSIS; SOUZA; DIAS, 2019). Nesta seção, apresenta-se algumas das distribuições estatísticas mais utilizadas em confiabilidade, e as suas principais características.

### 2.2.1 Distribuição gama

A distribuição gama, de acordo com Barros (2008), encontra o tempo necessário para obter o número específico de ocorrências de um evento.

Uma variável aleatória  $X$  tem uma distribuição gama de parâmetros  $\alpha > 0$  e  $\beta > 0$ , se sua função densidade de probabilidade for dada por:

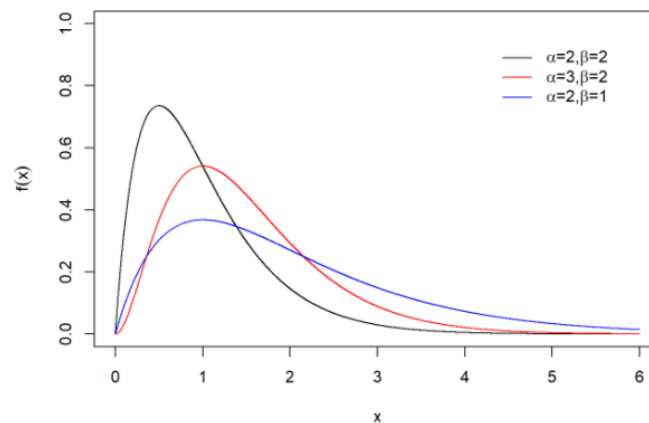
$$f(x) = \begin{cases} 0, & \text{para } x \leq 0 \\ \frac{\beta^\alpha}{\Gamma(\alpha)} x^{\alpha-1} e^{-\beta x}, & \text{para } x > 0 \end{cases} \quad (15)$$

Sendo:

- $\alpha$  parâmetro de forma;
- $\beta$  parâmetro de escala;
- $\Gamma(\alpha)$  a função gama definida pela integral  $\int_0^\infty x^{\alpha-1} e^{-x} dx$ .

A Figura 2 a seguir ilustra a função de densidade de probabilidade da distribuição Gama, para valores definidos de  $\alpha$  e  $\beta$ .

Figura 2 – Gráfico Função de Densidade de Probabilidade da Gama

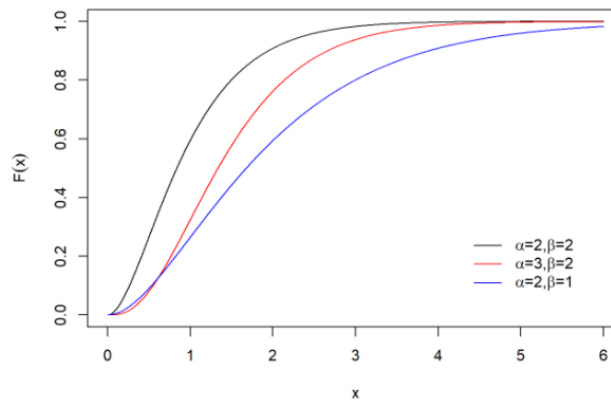


A sua função de distribuição acumulada é dada por:

$$F(x) = \begin{cases} 0, & \text{para } x \leq 0 \\ \int_0^x \frac{\beta^\alpha}{\Gamma(\alpha)} x^{\alpha-1} e^{-\beta x} dx, & \text{para } x > 0 \end{cases} \quad (16)$$

A Figura 3 seguir ilustra a função de distribuição acumulada da distribuição Gama, para valores determinados de  $\alpha$  e  $\beta$ .

Figura 3 – Gráfico Função de Distribuição Acumulada da Gama



Fonte: Mattos (2018, p.1)

A distribuição Gama possui a distribuição exponencial como caso especial, quando  $\alpha$  é igual a 1. Essa distribuição tem sido usada em problemas de confiabilidade por conta do seu fácil ajuste. (COLOSIMO; GIOLO, 2006)

### 2.2.2 Distribuição exponencial

De acordo com Pesce (2005), considera-se que uma variável aleatória  $x$  tem distribuição exponencial de parâmetro  $\lambda > 0$ , se a sua função densidade de probabilidade for dada por:

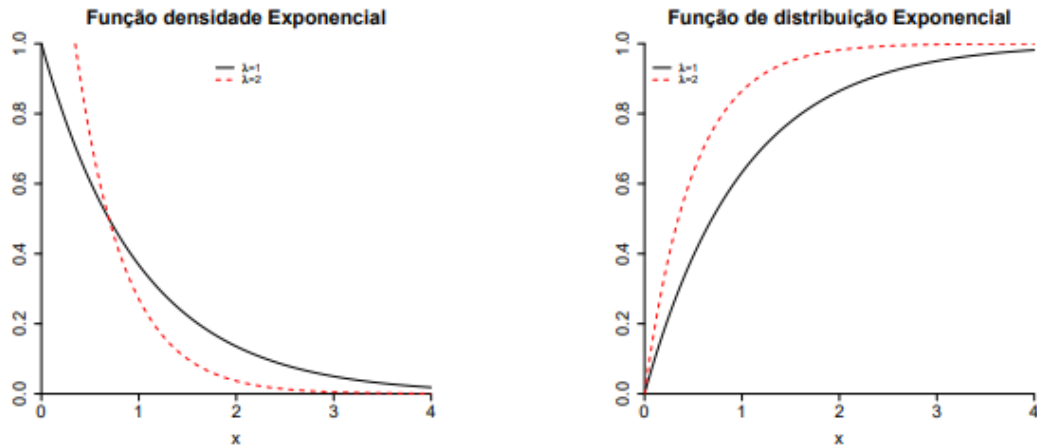
$$f(x) = \begin{cases} 0, & \text{para } x < 0 \\ \lambda \cdot e^{(-\lambda \cdot x)}, & \text{para } x \geq 0 \end{cases} \quad (17)$$

A sua função de distribuição é dada por:

$$F(x) = \begin{cases} 0, & \text{para } x < 0 \\ 1 - e^{(-\lambda \cdot x)}, & \text{para } x \geq 0 \end{cases} \quad (18)$$

A representação gráfica da função densidade de probabilidade da exponencial e sua função de distribuição pode ser observada na Figura 4 a seguir.

Figura 4 – Gráfico da função densidade de probabilidade (esquerda) e função de distribuição (direita) de uma variável aleatória  $\lambda$ .



Fonte: Caeiro (2009, p. 33)

A distribuição exponencial apresenta um único parâmetro e é a única que possui a função taxa de falha constante, por conta disso, é um dos modelos probabilísticos mais simples usados para descrever o tempo de falha (COLOSIMO; GIOLO, 2006). Além disso, “esta distribuição pode descrever inúmeros fenômenos físicos, como o tempo  $t$  para o decaimento de um núcleo radioativo” (PESCE, 2005, p. 24).

### 2.2.3 Distribuição qui-quadrado

A distribuição Qui-quadrado está definida apenas para valores positivos de uma variável  $x$  e depende dos graus de liberdade  $v$ . Uma variável aleatória  $x$  tem uma distribuição Qui-Quadrado se sua função de densidade de probabilidade for:

$$f(x) = \begin{cases} \chi_v^2(x) = \frac{x^{\frac{v}{2}-1} e^{-\frac{x}{2}}}{2^{\frac{v}{2}} \Gamma(\frac{v}{2})} & \text{se } x \geq 0 \text{ e } v > 0 \\ 0, & \text{para qualquer outro caso} \end{cases} \quad (19)$$

Essa distribuição é uma particularidade da distribuição Gama, para valores de  $\alpha = \frac{1}{2}$  e  $\beta = \frac{v}{2}$ . A tabela com os valores percentis da distribuição  $\chi^2$  se encontra no Anexo A do presente trabalho.

### 2.3 FUNDAMENTOS DA ATUALIZAÇÃO BAYESIANA

A fim de elucidar o processo da atualização bayesiana, o Teorema de Bayes será novamente apresentado a seguir.

$$P(A|B) = \frac{P(B|A).P(A)}{P(B)} \quad (20)$$

Nota-se que  $\frac{1}{P(B)}$  funciona como uma constante normalizadora de  $P(B|A)$  pois  $B$  não depende de  $A$  (EHLERS, 2003). Assim, a Equação 20 pode ser escrita como:

$$P(A|B) \propto P(B|A).P(A) \quad (21)$$

No caso de se aplicar a Equação 20 para estimação de parâmetros, a equação pode ser reescrita considerando como eventos de interesse as distribuições de probabilidade do parâmetro de interesse. De fato, o termo  $\frac{1}{P(B)}$  “garante uma área sob a distribuição de probabilidades da posteriori igual a 1 ou que a integral da posteriori tem valor 1” (LOUREIRO, 2018, p. 20).

Assim, a Equação 20 pode ser escrita como:

$$\pi(\theta|x_1, \dots, x_n) \propto L(x_1, \dots, x_n|\theta).\pi(\theta) \quad (22)$$

A Equação 22 pode ser traduzida, segundo Ehlers (2003) da seguinte forma: a distribuição a posteriori do parâmetro  $\theta$  é proporcional ao produto da função de verossimilhança com a distribuição a priori do parâmetro  $\theta$ . “Assim, a abordagem bayesiana usa informações prévias, e até julgamentos subjetivos, para construir um modelo de distribuição a priori do parâmetro” (CALIL; HIRANO; DIAS, 2005, p. 4, tradução nossa) e posteriormente atualiza o conhecimento sobre o parâmetro  $\theta$  com base em novos dados obtidos por uma amostra  $x_1, \dots, x_n$ .

De acordo com Barros (2008), a verossimilhança, então, permite atualizar iterativamente a informação a priori à medida que mais dados são coletados. Desta forma, a Atualização Bayesiana pode ser considerado um método para atualização de um modelo conhecido. A cada nova evidência amostral se calcula uma nova distribuição posteriori e esta distribuição a posteriori calculada se tornará a próxima distribuição a priori (HOFF, 2009 apud LOUREIRO, 2018).

### 2.4 FUNÇÃO VEROSSIMILHANÇA

A função de verossimilhança é uma função dos parâmetros de uma distribuição de probabilidade que permite inferir sobre o seu valor a partir de um conjunto de observações.



A função de verossimilhança de um parâmetro  $\theta$  é definida como a probabilidade de se observar os valores  $x_1, \dots, x_n$  de uma amostra, dados os possíveis valores de  $\theta$ . O que pode ser escrito como:  $L(x_1, x_2, \dots, x_n|\theta)$

Desenvolvendo a probabilidade conjunta ( $x_1 \cap x_2 \cap \dots \cap x_n$ ), considerando os elementos da amostra como independentes, temos (MONTGOMERY; RUNGER, 2009):

$$L(x_1, \dots, x_n|\theta) = f(x_1|\theta) \cdot f(x_2|\theta) \dots f(x_n|\theta) = \prod_{i=1}^n f(x_i|\theta) \quad (23)$$

Por exemplo, se  $x$  é uma variável aleatória contínua que pode ser modelada por uma Distribuição Exponencial, então:

$$f(x) = \begin{cases} 0, & \text{para } x < 0 \\ \lambda \cdot e^{(-\lambda \cdot x)}, & \text{para } x \geq 0 \end{cases} \quad (24)$$

Substituindo a Equação 24 na Equação 23 pode-se obter a Função Verossimilhança para a Distribuição Exponencial (MONTGOMERY; RUNGER; HUBELE, 2009):

$$\begin{aligned} L(x_1, \dots, x_n|\lambda) &= \prod_{i=1}^n f(x_i|\lambda) \\ &= \prod_{i=1}^n \lambda \cdot e^{(-\lambda \cdot x_i)} \\ &= \lambda^n \cdot e^{(-\lambda \cdot \sum_{i=1}^n x_i)} \end{aligned} \quad (25)$$

## 2.5 DISTRIBUIÇÕES CONJUGADAS

As distribuições conjugadas possuem grande interesse prático na inferência bayesiana, visto que a distribuição a posteriori pode ser obtida a partir da combinação da distribuição a priori, com a função de verossimilhança. A distribuição a priori representa probabilisticamente o conhecimento prévio que se tem sobre um dado evento, anterior ao experimento, já a função de verossimilhança possui as informações provenientes dos dados, após o experimento. (JESUS, 2014)

Uma forma de encontrar a função posteriori é fazendo o uso das distribuições conjugadas. Dado um evento com distribuição com parâmetro de interesse  $\theta$ , é importante salientar que:

A distribuição a priori é representada por uma forma funcional, cujos parâmetros devem ser especificados de acordo com este conhecimento. Estes parâmetros indexadores da família de distribuições a priori são chamados de hiperparâmetros para distinguí-los dos parâmetros de interesse  $\theta$ . (EHLERS, 2003, p. 13)

Diferentes escolhas de prioris podem dificultar o cálculo da integral, muitas vezes, impossibilitando uma solução analítica. Por conta disso, Ehlers (2003, p. 13) explica ser desejável que “as distribuições a priori e a posteriori pertençam a mesma classe de distribuições, desta forma, a atualização do conhecimento que se tem de  $\theta$  envolve apenas uma mudança nos hiperparâmetros”. Nos casos em que não é possível utilizar o recurso das distribuições conjugadas se efetua o uso de soluções numéricas.

### 3 BASE DE DADOS DE CONFIABILIDADE

Este capítulo apresenta uma revisão bibliográfica sobre os conceitos de confiabilidade e OREDA tendo como tópicos as definições necessárias para o entendimento da análise, procedimentos e as aplicações.

#### 3.1 CONFIABILIDADE

Segundo a norma ABNT NBR 5.462:1994, a confiabilidade pode ser definida como a capacidade de um item desempenhar uma função requerida sob condições especificadas, durante um dado intervalo de tempo.

A quantificação da capacidade do item desempenhar uma função, atrelado a um tempo pode ser entendido como uma probabilidade. Logo, a confiabilidade é uma grandeza que pode ser medida, para isso utilizam-se algumas métricas.

Segundo Signorini Filho (2019) essas métricas são:

- Confiabilidade,  $R(t)$ : Probabilidade de funcionamento do item até um dado tempo “ $t$ ”.
- Probabilidade Acumulada de Falha,  $F(t)$ : Probabilidade do item falhar até um determinado tempo “ $t$ ”;
- Tempo Médio Até a Falha (Mean Time To Failure – MTTF): Tempo médio que o item irá funcionar antes de ocorrer a falha;
- Taxa de Falha,  $\lambda(t)$ : Quantidade instantânea da falha por unidade de tempo.

Signorini Filho (2019, p. 38) destaca, ainda, que “Para que essas métricas agreguem valor a análise dos dados é fundamental que uma referência ou até mesmo uma base de dados seja utilizada para comparação de modo a gerar tomadas de decisão efetivas.” A base de dados que será utilizada no presente trabalho será a Offshore and Onshore Reliability Data (OREDA) que possui dados de 278 instalações, 17.000 unidades de equipamentos com 39.000 falhas e 73.000 registros de manutenção.

#### 3.2 COLETA DE DADOS

Considerando que “A coleta de dados exige grandes investimentos e padronização combinados a um efetivo gerenciamento eletrônico dos mesmos” (SIGNORINI FILHO, 2019, p.62), as indústrias de petróleo e gás natural dispõem de uma norma internacional, a ABNT NBR ISO 14.224, cujo objetivo é padronizar a coleta e intercâmbio de dados de confiabilidade e manutenção para equipamentos.

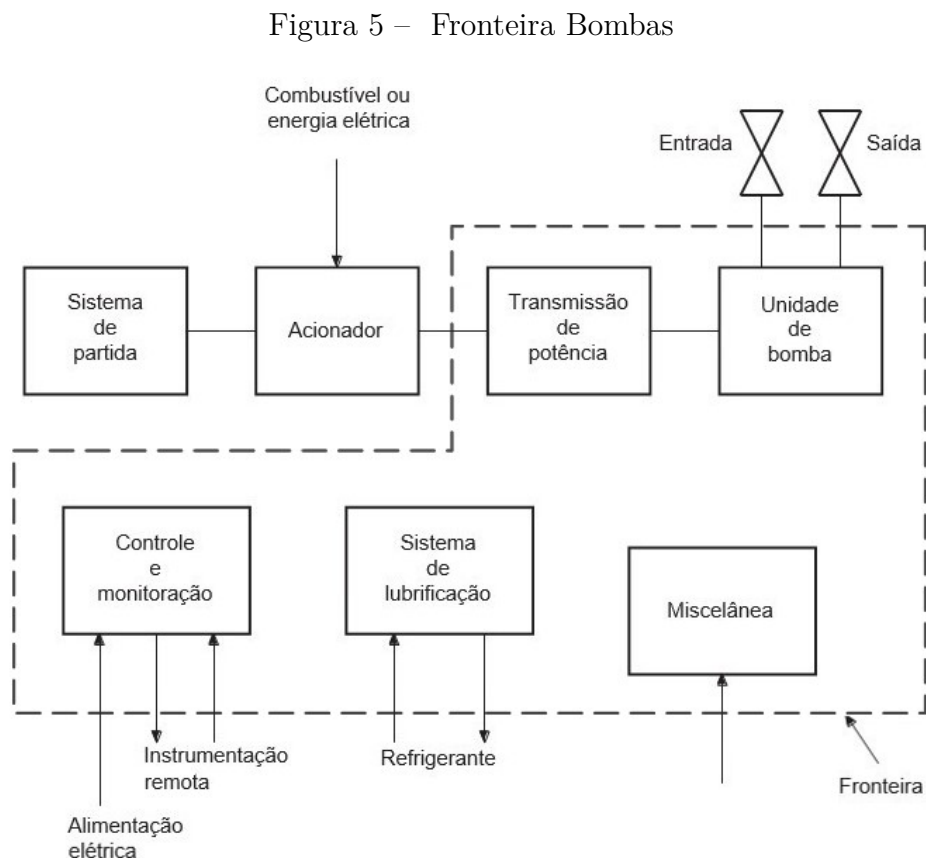
A norma não detalha os métodos de análise de dados, mas especifica quais devem ser adquiridos de modo a viabilizar os estudos de confiabilidade. Segundo o documento, todos

os equipamentos usados na indústria de petróleo, gás natural e petroquímica, são contemplados; assim, inclui equipamentos de processo, de segurança, submarinos, carga/descarga, e todo maquinário que pertençam às instalações, permanentemente ou usados durante as fases de instalação, manutenção ou modificação.

A ABNT NBR ISO 14.224 possui algumas limitações de uso, os dados de custo, por exemplo, apesar de usualmente utilizados nas análises de desempenho de confiabilidade e serem parâmetros importantes para a definição de prioridades, não são contemplados.

Bons resultados de análise dependem diretamente da qualidade dos dados coletados. A norma não detalha as medidas de qualidade, mas apresenta práticas de controle e garantia de qualidade de dados. Dados de boa qualidade são caracterizados por serem completos em relação à especificação, estarem em conformidade com definições de parâmetros de confiabilidade, possuírem entrada, transferência, manuseio e armazenamento de dados precisamente e possua população suficiente e período de observação adequado.

Uma das exigências da norma é a definição de fronteira, que inclui identificação das interfaces e da subdivisão interna do equipamento. A Figura 5 ilustra um exemplo de fronteiras para bombas.



Fonte: ABNT (2011, p.64)

Os dados coletados são divididos por nove níveis taxonômicos hierárquicos, classificados em itens em grupos genéricos. Os cinco primeiros níveis representam uma cate-

gorização de alto nível que está relacionada às indústrias e à configuração de planta, os quatro níveis restantes descrevem o equipamento e subdivisões inferiores.

O período de coleta de dados é definido conforme o equipamento e a fase analisada (juventude, estabilidade ou velhice), a fase de um equipamento será explicada na Seção 3.3.1. A extensão do período de coleta de dados deve considerar a taxa de falha prevista o tamanho da população e o acesso aos dados. Caso durante o período de observação, não ocorra nenhuma falha é possível estimar a taxa de falha.

A estimativa da taxa de falha contemplada na norma é apresentada na seção Estimation Procedures do OREDA, pois a norma é oriunda dessa base de dados, logo os procedimentos adotados são semelhantes.

### 3.3 OFFSHORE AND ONSHORE RELIABILITY DATA (OREDA)

O Projeto OREDA teve início em 1981 e o seu principal objetivo é coletar dados da operação de equipamentos de superfície e submarinos de operações offshore e onshore de produção de petróleo e gás, com o intuito de melhorar os dados básicos em estudos de confiabilidade de segurança. Os dados coletados dos equipamentos topside - parte da plataforma de petróleo que fica acima da água - são tabulados conforme ilustrado no Quadro 1.

Quadro 1 – Quadro utilizado na base de dados OREDA - Topside

Taxonomy no		Item											
Population	Installations	Aggregated time in service (10 <sup>6</sup> hours)					No of demands						
		Calendar time *		Operational time †									
Failure mode	No of failures	Failure rate (per 10 <sup>6</sup> hours)					Active rep.hrs	Repair (manhours)					
		Lower	Mean	Upper	SD	n/t		Min	Mean	Max			
<b>Comments</b>													

Os campos apresentados no Quadro 1 possuem identificação do item e os parâmetros de confiabilidade estimados e podem ser descritos da seguinte forma (OREDA, 2002) :

- Taxonomy nº e Item (Número de Taxonomia e Item) – Código do item que terá os dados apresentados. A descrição do item se dá de forma hierárquica.
- Population (População) – Número total de itens da base da estimativa.
- Installations (Instalações) – Número total de instalações/amostras, neste caso plataformas em que as informações foram coletadas.
- Aggregated Time in Service (Tempo Agregado em Serviço) - Dois tipos de escalas de tempo são utilizadas: tempo calendário e tempo operacional. O tempo agregado em serviço para a população total é fornecido para ambas as escalas de tempo.
- Number of demands (Número de demandas) – número acumulado de demandas/ciclos para a população total é fornecido quando disponível.
- Failure mode (Modo de Falha) – descrição de como a falha ocorreu, quando esse dado está disponível.
- Number of failures (Número de falhas) – O número total de eventos de falha é apresentado para cada modo de falha. O número acumulado de falhas é apresentado como “all modes” (Todos os modos).
- Failure rate (Taxa de Falha) – Exibe estimativas da taxa de falha para cada modo de falha.
- Mean (média) – é uma estimativa da taxa de falha média obtida usando o estimador OREDA.
- Lower/Upper (Inferior, Superior) – intervalo de incerteza de 90% para a taxa de falha.
- SD (Desvio padrão) – variação entre as várias amostras.
- $\frac{n}{t}$  – para amostras homogêneas é estimada nesta coluna.
- A taxa de falha medida por  $10^6$  horas refere-se ao tempo calendário (marcado com \*) ou tempo operacional (marcado com †).
- Active repair time (Tempo de reparo ativo) - Contém o tempo médio de calendário (horas) necessário para reparar e devolver o item para um estado em que está pronto para retomar suas funções. O tempo de reparo ativo é o momento em que o trabalho de reparo está sendo feito. Não inclui tempo para desligar a

unidade, emitir ordem de serviço, espere por peças de reposição, inicialização após reparo, etc. O tempo de reparo ativo é, portanto, normalmente menor do que o tempo de inatividade onde algumas das atividades indicadas acima podem ser incluídas.

- Repair (Reparo) -As colunas de reparo apresentam três valores de tempo de reparo (horas-homem). O valor médio é o número médio de horas-homem registrado para reparar a falha e restaurar a função. Os valores mínimo e máximo são o menor e o maior número de horas de trabalho registrados para o reparo do item.
- Comments (Comentários) – Contém informação a respeito da probabilidade de falha sob demanda, quando disponível

Ao longo dos últimos 35 anos, o projeto OREDA tem sido patrocinado por cerca de onze diferentes companhias e hoje conta com 7 membros ativos. A sua execução é realizada em fases, normalmente com duração de 2 a 3 anos e, segundo o documento, alguns marcos foram alcançados em cada fase, podendo citar como principais:

- Criação da norma ABNT NBR ISO 14.224: Indústrias de petróleo, petroquímica e gás natural, coleta e troca de dados de confiabilidade e manutenção de equipamentos, baseada no conceito OREDA, emitida em 2006;
- Diretrizes e softwares para coleta e análise de dados foram estabelecidas;
- Seis edições de bases de dados de confiabilidade foram publicadas em mais de cinquenta países;
- A base de dados foi o berço da criação da ABNT NBR ISO 20.815: Indústrias de petróleo, petroquímica e gás natural – Garantia da produção e gestão da confiabilidade.

Em 2018 o projeto disponibilizou a base de dados digitalmente por meio do @OREDA-Cloud, para não membros, é necessário adquirir o acesso pelo site.

### 3.3.1 Procedimentos de estimação

O projeto OREDA apresenta estimativas de taxa média de falha para prever tempos de reparo. Para isso é importante estabelecer os métodos estatísticos utilizados e suas considerações.

A função taxa de falha  $\lambda(t)$ , segundo Signorini Filho (2019), informa a probabilidade de falha de um item por unidade de tempo, logo é possível estimar a probabilidade de falha futura. Para isso, a função deve ser modelada conforme o comportamento típico de um produto, equipamento e/ou sistema e sua aplicação, considerando as diferentes

distribuições estatísticas que podem ser utilizadas para caracterizar a curva característica da vida destes. (KARDEC; NASCIF, 2009)

Colosimo e Giolo (2006) destacam que a taxa instantânea de falha muda com o tempo, podendo ser definida pela equação a seguir:

$$\lambda(t) = \lim_{\Delta t \rightarrow 0} \frac{P(t \leq T < t + \Delta t)}{\Delta t} \quad (26)$$

Existem três comportamentos muito presentes na literatura que a taxa de falha pode assumir, do ponto de vista do uso, “três períodos de vidas distintos: juventude, estabilidade ou normal e velhice ou senilidade” (DIAS, 2009, p. 7).

As falhas são denominadas falhas precoces de juventude, por ocorrerem no período inicial. Nesse período a função taxa de falha se apresenta decrescente. “Essas falhas estão, geralmente, associadas a erros oriundos do projeto, da manufatura, da má qualidade dos materiais, não detectadas no controle de qualidade.”(DIAS, 2009, p. 8).

Na fase de estabilidade ou normal, a curva da taxa de falha  $\lambda(t)$  possui um comportamento constante. Segundo Germano (2018) é nessa fase que o índice de falhas é menor, pois estão associadas a falhas aleatórias, oriundas de cargas externas inevitáveis e imprevisíveis, e não tendem a variar à medida que o maquinário envelhece, desta forma não desenvolve uma memória associada a falha. Dias (2009) afirma em Confiabilidade na manutenção industrial, que esse período pode ser modelado pela distribuição Exponencial.

A função taxa de falha no período da velhice ou senilidade é crescente e aponta que o item tem maior probabilidade de falhar com o tempo, indicando o fim da vida útil do equipamento. Essas falhas representam o desgaste ou a fase da fadiga (DIAS, 2009).

A combinação das três fases supracitadas em diferentes intervalos de tempo formam a curva mostrada na Figura 6 que, por conta do formato em que a mesma se apresenta, ficou conhecida como Curva da Banheira. Esta curva é válida para diversos sistemas e componentes mecânicos, determinada a partir de estudos estatísticos.(KARDEC; NASCIF, 2009)

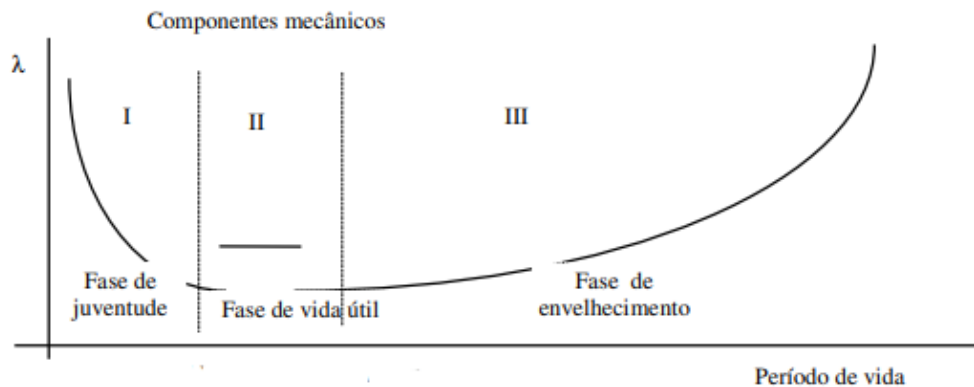
Na coleta de dados do OREDA, problemas de instalação e mecanismos de desgaste foram desconsiderados, analisando apenas a fase de estabilidade do item em que as falhas são aleatórias e a função taxa de falha é constante, podendo ser modelado pela distribuição exponencial.

Essa aproximação para distribuição exponencial e razoável na Fase II da Figura 6 e implica em considerar que o componente está sempre na mesma condição – em termos de probabilidade de falha – independentemente do tempo de operação. Isto pode ser apresentado matematicamente pela Equação 27.

$$P(T > s + t \mid T > s) = P(T > t) \quad (27)$$



Figura 6 – Variação Típica da Taxa de Falha em Componentes Mecânicos



Fonte: Dias (2009, p. 8).

Essa propriedade é conhecida como falta de memória – pois caso a falha não tenha acontecido até o tempo  $s$  (na Equação 27 é o evento conhecido da probabilidade condicional,  $T > s$ ), a probabilidade de falhar no próximo período  $t$  (i.e.,  $T > s + t$ ) é a mesma probabilidade de falha no início da observação (i.e.,  $T > t$ ).

Isto pode ser demonstrado matematicamente desenvolvendo a probabilidade condicional apresentada à esquerda da Equação 27 e substituindo as probabilidades para o caso da distribuição exponencial (que foi apresentada na Equação 18).

$$\begin{aligned}
 P(T > s + t | T > s) &= \frac{P(T > s + t \cap T > s)}{P(T > s)} \\
 &= \frac{P(T > s + t)}{P(T > s)} \\
 &= e^{-\lambda \cdot (s+t)} \cdot e^{+\lambda \cdot s} \\
 &= e^{-\lambda \cdot t}
 \end{aligned} \tag{28}$$

Uma implicação significativa ao assumir a taxa de falha constante é que as falhas são puramente causais e são independentes da idade do item. Neste caso, o tempo médio até a falha (Mean Time To Failure – MTTF) pode ser calculado conforme a Equação 29.

$$MTTF = \frac{1}{\lambda} \tag{29}$$

### 3.3.2 Estimador da taxa de falha e intervalo de confiança

O OREDA (2002) utiliza dois métodos distintos para estimar a taxa de falha e o intervalo de confiança, inicialmente os dados são separados em duas categorias de amostras, as homogêneas e não homogêneas.

### 3.3.2.1 Amostras homogêneas

As amostras são consideradas homogêneas, pela base de dados, quando existem dados de falha de itens idênticos que trabalham sob as mesmas condições operacionais e ambientais. Os dados utilizados para estimar a taxa de falha  $\lambda$ , são o número observado de falhas  $n$ , e o tempo em serviço  $t$ . O valor estimado da taxa de falha pode ser calculado conforme a Equação 30.

$$\lambda = \frac{n}{t} \quad (30)$$

Vale ressaltar, que, por conta da dificuldade em determinar o tempo total que um componente estará em funcionamento e exposto ao estresse, no OREDA, o tempo em serviço pode ser medido como tempo calendário ou como tempo operacional, e ambos são apresentados nas tabelas de dados.

No que se refere ao nível de confiança no intervalo para a estimativa da taxa de falha, o OREDA define o valor de 90%. Assim a probabilidade do valor verdadeiro da taxa de falha estar entre os limites máximos e mínimos do intervalo de confiança e de 90%, conforme ilustrado na Equação 31.

$$P(\lambda_{\min} \leq \lambda \leq \lambda_{\max}) = 90\% \quad (31)$$

Considerando um número  $n$  de falhas durante o tempo em serviço  $t$ , este intervalo de confiança de 90% é dado por:

$$\left( \frac{1}{2t} \cdot Z_{[0.95;2n]} ; \frac{1}{2t} \cdot Z_{[0.05;2(n+1)]} \right) \quad (32)$$

Sendo a estatística  $Z$  obtida pela distribuição qui-quadrado ( $\chi^2$ ) para os percentis de 95% e 5% com graus de liberdade  $2.n$  e  $2(n+1)$ , respectivamente. A tabela da distribuição qui-quadrado se encontra no Anexo A do presente documento.

### 3.3.2.2 Amostras não homogêneas

Em muitos casos os dados podem ser oriundos de diferentes instalações com condições operacionais e ambientais distintas, com isso a amostra não pode ser considerada homogênea (OREDA, 2002). Nestes casos, a base de dados chama o problema de multi-amostras e utiliza uma metodologia diferente da já apresentada para estimar a taxa de falha e o intervalo de confiança.

Para estes casos a base de dados se baseia nos procedimentos adotados na Sociedade Escandinava de Engenheiros de Confiabilidade, apresentado por Spjøtvoll (1985). A estimativa da taxa de falha pode ser modelada, assumindo que  $\lambda$  é uma variável aleatória,

de alguma distribuição ainda não determinada, com a função densidade de probabilidade dada por  $f(\lambda)$ .

O modelo da estimativa segue as seguintes considerações:

- O número de amostras é dado por “ $k$ ”, ou seja, pode-se ter dados de  $k$  diferentes instalações;
- As falhas são representadas por “ $n$ ”. É possível observar um número  $n_i$  falhas durante um tempo total  $t$ , para  $i = 1, 2, \dots, k$ ;
- A taxa de falha é constante para cada instalação;
- O valor da taxa de falha pode variar entre as amostras devido a diferentes condições operacionais e ambientais em cada instalação.

A média  $\theta$ , dos valores da taxa de falha é dada por:

$$E(\lambda) = \theta = \int_0^{\infty} \lambda \cdot f(\lambda) \cdot d\lambda \quad (33)$$

A variância  $\sigma^2$ , pode ser descrita da seguinte forma:

$$V(\lambda) = \sigma^2 = \int_0^{\infty} (\lambda - \theta)^2 \cdot f(\lambda) \cdot d\lambda \quad (34)$$

Para calcular a taxa de falha de problemas multi-amostras, inicialmente o OREDA faz uma estimativa inicial da média  $\theta_1$  da taxa de falha.

$$\theta_1 = \frac{\text{número total de falhas}}{\text{tempo total em operação}} = \frac{\sum_{i=1}^k n_i}{\sum_{i=1}^k t_i} \quad (35)$$

Em seguida os seguintes valores foram calculados:

- $S1 = \sum_{i=1}^k t_i$
- $S2 = \sum_{i=1}^k t_i^2$
- $V = \sum_{i=1}^k \frac{(n_i - \theta_1)^2}{t_i} = \sum_{i=1}^k \frac{n_i^2}{t_i} - \theta_1^2 \cdot S1$

Estimar a variância média é o próximo passo:

$$\sigma^2 = \frac{V - (k - 1)\theta_1}{S1^2 - S2} \cdot V \quad (36)$$

Por fim, a média da taxa de falha pode ser estimada pela Equação 37:

$$\hat{\theta} = \frac{1}{\sum_{i=1}^k \left( \frac{1}{\frac{\theta_1}{t_i} + \sigma^2} \right)} \cdot \sum_{i=1}^k \left( \frac{1}{\frac{\theta_1}{t_i} + \sigma^2} \cdot \frac{n_i}{t_i} \right) \quad (37)$$

O intervalo de confiança determinado nessa análise é de 90%. Para obter os valores máximos e mínimos do intervalo de confiança, a função de densidade de probabilidade é considerada uma distribuição Gama e os parâmetros  $\alpha$  e  $\beta$ , são calculados conforme as equações a seguir:

$$\beta = \frac{\hat{\theta}}{\sigma^2} \quad (38)$$

$$\alpha = \beta \cdot \hat{\theta} \quad (39)$$

Os valores máximos e mínimos do intervalo de confiança são obtidos utilizando a distribuição qui-quadrado.

$$\theta_{\min} = \frac{1}{2 \cdot \beta} \cdot z_{(0,95; 2 \cdot \alpha)}; \quad \theta_{\max} = \frac{1}{2 \cdot \beta} \cdot z_{(0,05; 2 \cdot \alpha)} \quad (40)$$

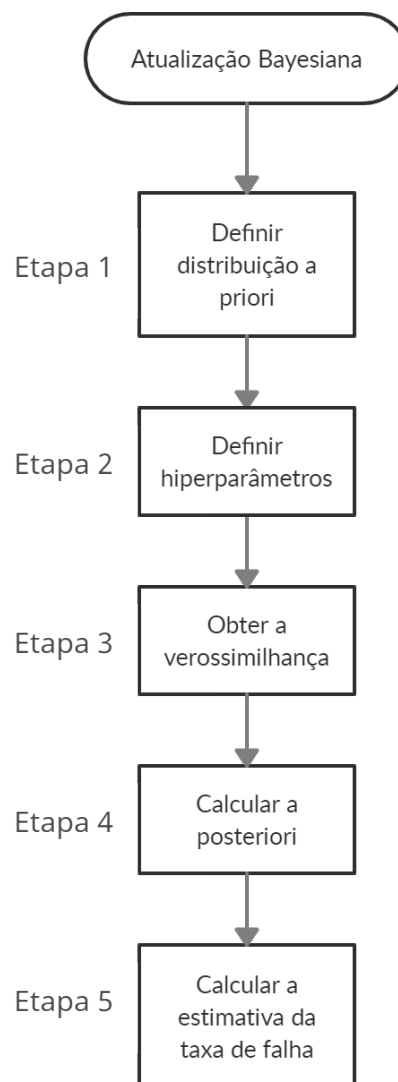
Em que  $z_{(0,95; 2 \cdot \alpha)}$  e  $z_{(0,05; 2 \cdot \alpha)}$  denotam os percentis de 95% e 5%, respectivamente, da distribuição  $\chi^2$  com  $2 \cdot \alpha$  graus de liberdade. Todas as equações da Seção 3.3 foram retiradas da base de dados (OREDA, 2002).

#### 4 MÉTODO PROPOSTO PARA ATUALIZAÇÃO BAYESIANA

O método proposto no presente trabalho, ilustrada na Figura 7, consiste em uma sequência de passos para atualizar as estimativas de valores de taxa de falha apresentadas na base de dados OREDA com os dados obtidos em campo – na planta em que o sistema está instalado.

Com a aplicação deste método, tem-se uma estimativa da taxa de falha mais assertiva e com menor variância.

Figura 7 – Fluxograma Metodologia



Fonte: Fluxograma elaborado para este trabalho.

Com o intuito de melhorar a compreensão dos passos apresentados na Figura 7, será realizada uma implementação fictícia do método proposto para uma bomba.

#### 4.1 DEFINIR A DISTRIBUIÇÃO A PRIORI

O artifício matemático das distribuições conjugadas de mesma família, foi utilizado na solução deste trabalho, desta forma, a solução analítica se fez possível. Diferentemente do OREDA, que adota uma aproximação utilizando a distribuição Qui-quadrado, o presente trabalho adota a distribuição Gama para obter o intervalo de confiança da taxa de falha. Essa distribuição foi adotada após cálculos em planilhas eletrônicas e verificação que os valores estavam semelhantes a do OREDA.

Os valores disponíveis na tabela de Bombas do OREDA na página 175, serviram como base para calcular os parâmetros da equação a priori, conforme apresentado no Quadro 2.

Quadro 2 – Dados de confiabilidade apresentados no OREDA relativos ao item Bomba.

Failure mode	n° of failures	Failure rate per 10 <sup>6</sup> hours				
		SD	n/t	Lower [5%]	Mean	Upper [95%]
Critical	524 *	49,34	4,83	0,00	20,52	108,44
	524 †	72,93	2,56	1,14	65,40	204,64
Degraded	754 *	86,32	6,95	0,00	44,20	210,34
	754 †	239,40	3,68	11,39	238,41	714,72
Incipient	1.124 *	86,32	10,37	0,08	55,97	228,31
	1.124 †	552,65	5,49	127,07	760,29	1.833,66
Unknown	21 *	4,16	0,19	0,00	2,04	9,97
	21 †	8,12	0,10	0,09	6,97	22,41
All modes	2.423 *	196,22	22,34	0,12	123,75	515,05
	2.423 †	810,62	11,84	230,16	1.184,82	2.751,06

Fonte: Adaptado de OREDA (2002, p. 173 a 175).

O Quadro 2 apresenta os dados obtidos do OREDA referentes a uma bomba, que diferencia os modos de falha (Failure mode) em: críticos (Critical), degradados (Degraded), incipientes (Incipient) e desconhecidos (Unknown). Nas últimas linhas, então, é apresentada a soma de todos os modos de falha (All modes). Para cada situação dessas, o quadro divide em duas linhas: na primeira apresentam-se os valores para o tempo calendário (\*) e na segunda para o tempo que os itens operaram (†).

As colunas apresentam o número de falhas para cada situação (n° of failures); o desvio padrão (SD) obtido conforme a Equação 36; a estimativa da taxa de falha em 10<sup>6</sup> horas (Failure rate per 10<sup>6</sup> hours) considerando a amostra homogênea ( $n/t$ ), vide Equação 30; a estimativa da taxa de falha considerando a amostra homogênea (Mean), conforme Equação 37, e os limites inferior (Lower) e superior (Upper) do intervalo de confiança, conforme Equação 40.

## 4.2 DEFINIR HIPERPARÂMETROS

Para dar início a atualização bayesiana, após definir a distribuição Gama como distribuição a priori do parâmetro taxa de falha ( $\lambda$ ), seus hiperparâmetros iniciais foram calculados. Para tanto, utilizou-se a penúltima linha do Quadro 2 para demonstração dos cálculos. Em uma distribuição Gama, O parâmetro  $\beta$  pode ser obtido dividindo o valor do desvio padrão (SD), pela média elevada ao quadrado ( $\bar{x}^2$ ), conforme a Equação 41 a seguir:

$$\beta = \frac{SD}{\bar{x}^2} = \frac{196,22}{(123,75)^2} = 0,003214 \quad (41)$$

Note-se que a média obtida no Quadro 2 é a estimativa a priori da taxa de falha, portanto  $\lambda = 123,75$ .

O parâmetro  $\alpha$  foi obtido ao multiplicar o valor da taxa de falha ( $\lambda$ ) pelo valor de beta ( $\beta$ ) encontrado:

$$\alpha = \lambda \cdot \beta = 123,75 \cdot 0,003214 = 0,39774 \quad (42)$$

Por fim, foram obtidos os valores máximo e mínimo (para 95% e 5%) do intervalo de confiança, utilizando a distribuição Gama com os hiperparâmetros  $\alpha$  e  $\beta$  calculados:

$$Lower [5\%] = 0,12 \quad e \quad Upper [95\%] = 515,18 \quad (43)$$

O Quadro 3 apresenta os valores equivalentes aos do Quadro 2, mas obtidos utilizando a distribuição Gama.

Quadro 3 – Dados de confiabilidade da bomba, utilizando a distribuição Gama como base para o cálculo do intervalo de confiança.

Failure mode	n° of failures	Failure rate per 10 <sup>6</sup> hours				
		$\beta$	$\alpha$	Lower [5%]	Mean	Upper [95%]
Critical	524 *	0,008429	0,172964	0,00	20,52	109,57
	524 †	0,012296	0,804161	1,82	65,40	211,76
Degraded	754 *	0,005932	0,262193	0,00	44,20	210,90
	754 †	0,004160	0,991746	11,98	238,41	716,31
Incipient	1.124 *	0,007512	0,420424	0,08	55,97	228,56
	1.124 †	0,002489	1,892597	125,80	760,29	1.835,31
Unknown	21 *	0,117881	0,240477	0,00	2,04	9,99
	21 †	0,105711	0,736807	0,15	6,97	23,29
All modes	<b>2.423 *</b>	<b>0,003214</b>	<b>0,397744</b>	<b>0,12</b>	<b>123,75</b>	<b>515,18</b>
	2.423 †	0,001803	2,136339	228,16	1.184,82	2.752,70

Fonte: Quadro elaborado para este trabalho.

Note-se que os valores no Quadro 3 são próximos aos apresentados no Quadro 2; o que evidencia que existe pouca diferença por usar a distribuição Gama (como adotada pelo presente trabalho) ou a distribuição qui-quadrado (como adotada pelo OREDA) para obter o intervalo de confiança da taxa de falha.

### 4.3 OBTER A VEROSSIMILHANÇA

Para dar seguimento nos cálculos se faz necessário encontrar a função verossimilhança. Os tempos até a falha se comportam como uma distribuição exponencial, logo pode-se assumir que a taxa de falha  $\lambda$  é constante; mas, de valor desconhecido.

No caso da Distribuição Exponencial, a Função Verossimilhança foi obtida na Seção 2.4 e está apresentada na Equação 25, novamente ilustrada a seguir.

$$L(t_1, \dots, t_n | \lambda) = \lambda^n \cdot e^{-\lambda \cdot \sum_{i=1}^n t_i} \quad (44)$$

Para realizar a atualização bayesiana, tempos até a falha fictícios foram gerados de forma randômica e estão dispostos no Quadro 4.

Quadro 4 – Tempos fictícios até a falha gerados para ilustrar o método proposto.

Tempos até a falha	
1	698 horas
2	23.125 horas
3	3.534 horas
4	1.720 horas
5	10.842 horas

Fonte: Quadro elaborado para este trabalho.

Quando os tempos até a falham forem coletados em campo, deve-se atentar as orientações presentes na ABNT NBR ISO 14.224.

### 4.4 CALCULAR A POSTERIORI

A Equação 22, então, pode ser escrita para o caso do parâmetro de interesse ser a taxa de falha ( $\lambda$ ), sendo  $t_i$  os tempos até a falha observados na amostra:

$$\pi(\lambda | t_1, \dots, t_n) \propto L(t_1, \dots, t_n | \lambda) \cdot \pi(\lambda) \quad (45)$$

Dado que a família Gama de distribuições corresponde a família conjugada natural da Distribuição Exponencial e o resultado da sua multiplicação é uma nova distribuição Gama. Assim, considerando que o parâmetro  $\lambda$  da Distribuição Exponencial se comporta



como uma Distribuição Gama de parâmetros  $\alpha$  e  $\beta$ , a distribuição a priori de  $\lambda$  pode ser escrita como  $\pi(\lambda|\alpha, \beta)$ .

Substituindo a Equação 44 na Equação 45, tem-se que:

$$\begin{aligned}\pi(\lambda|t_1, \dots, t_n) &\propto L(t_1, \dots, t_n|\lambda) \cdot \pi(\lambda|\alpha, \beta) \\ &\propto \left[ \lambda^n \cdot e^{-\lambda \sum_{i=1}^n t_i} \right] \cdot [\lambda^{\alpha-1} \cdot e^{(-\beta \cdot \lambda)}] \\ &\propto \lambda^{\alpha+n-1} \cdot e^{-\lambda(\beta + \sum_{i=1}^n t_i)}\end{aligned}\tag{46}$$

Nota-se que a função densidade de probabilidade da Distribuição Gama, apresentado na Seção 2.2.1, pode ser rescrita substituindo a variável aleatória “ $x$ ” pela variável aleatória “ $\lambda$ ” (para valores de  $\lambda$  positivos):

$$f(\lambda) \propto \lambda^{\alpha-1} e^{-\lambda \cdot \beta}\tag{47}$$

Pois, o fator  $\frac{\beta^\alpha}{\Gamma(\alpha)}$  não depende da variável aleatória “ $\lambda$ ” e, portanto, é uma constante.

Confrontando o resultado da Equação 46 com a Equação 47, é possível notar que a distribuição a posteriori é uma Gama de parâmetros “ $\alpha + n$ ” e “ $\beta + \sum_{i=1}^n t_i$ ”, conforme apresentado na Equação 48.

$$\pi(\lambda|t_1, \dots, t_n) \approx \text{Gama}(\alpha + n; \beta + \sum_{i=1}^n t_i)\tag{48}$$

#### 4.5 CALCULAR A ESTIMATIVA DA TAXA DE FALHA

Substituindo os dados em destaque do Quadro 2 na Equação 48, assim como os tempos até a falha do Quadro 4, os seguintes valores foram encontrados para os hiperparâmetros  $\alpha$  e  $\beta$ :

- $\alpha = 5,397744$
- $\beta = 0,043134$

A estimativa da taxa de falha pode ser obtida para o valor central – média – da Distribuição Gama a posteriori, conforme apresentado a seguir:

$$\lambda = \frac{\alpha}{\beta} = \frac{5,397744}{0,043134} = 125,14\tag{49}$$

Por fim, é importante destacar que a base de dados considera tanto tempo até a falha operacional quanto o tempo de calendário. Para o exemplo apresentado no presente trabalho foi utilizado o tempo calendário.

## 4.6 ANÁLISE DOS RESULTADOS

O Quadro 5 traz a comparação dos valores dos hiperparâmetros das Distribuições Gama a priori e a posteriori. Com base nestes valores é possível obter os intervalos com 90% de confiança para os valores da taxa de falha,  $\lambda$ .

Quadro 5 – Comparação dos valores dos hiperparâmetros das Distribuições Gama a priori e a posteriori.

	$\beta$	$\alpha$	IC Inferior	Média	IC Superior	Desvio Padrão
A priori	0,003214	0,397744	0,12	123,75	515,18	196,22
A posteriori	0,043134	5,397744	51,51	125,14	224,84	53,86

Fonte: Quadro elaborado para este trabalho.

Sendo:

- Média: a estimativa da taxa de falha;
- IC Inferior: o limite inferior do Intervalo de Confiança do parâmetro  $\lambda$ ;
- IC Superior: o limite superior do Intervalo de Confiança do parâmetro  $\lambda$ ; e
- Desvio Padrão: o desvio padrão da Distribuição Gama com os respectivos hiperparâmetros  $\alpha$  e  $\beta$ .

O aumento significativo do parâmetro de forma  $\alpha$  em relação ao parâmetro de escala, terá influência direta no formato do curva, que ficará menos achatada, resultando num desvio padrão menor. Já o novo parâmetro de escala  $\beta$  irá deslocar toda a curva para a direita, aumentando o valor da média em relação ao resultado obtido na priori.

O desvio padrão é uma medida que indica a dispersão dos dados dentro de uma amostra com relação à média, quanto menor o desvio padrão, mais homogênea será amostra. No Quadro 5 o valor encontrado para o desvio padrão na posteriori é significativamente menor que o da priori, sendo assim, a posteriori traz resultados mais homogêneos.

Os limites inferiores e superiores do intervalo de confiança da priori, possui um intervalo de aproximadamente zero a quinhentos. Já na distribuição a posteriori a sua faixa é de cerca de cinquenta a duzentos e vinte, desta forma o valor da média se encontra num intervalo menor (que é uma consequência da redução do desvio padrão).

Os valores obtidos a partir da atualização bayesiana, apesar de parecidos, indicam que a atualização traz resultados aperfeiçoados e com maior precisão. Logo, o uso iterativo da técnica permitirá uma melhor estimativa da taxa de falha e, por consequência, valores de probabilidades de falha mais adequados para análise de confiabilidade dos equipamentos.

A utilização do intervalo de confiança, ao invés de apenas utilizar o valor médio, é importante para fazer uma análise de sensibilidade e risco, variando o valor da taxa de

falha (dentro de cada intervalo de confiança) dos vários componentes de um determinado sistema. Em situações que é preciso ser mais conservador, pode-se usar o pior caso do intervalo de confiança.

## 5 CONCLUSÕES E RECOMENDAÇÕES PARA TRABALHOS FUTUROS

Neste trabalho foram expostas definições e implementações da técnica de Atualização Bayesiana, de modo a contribuir à falta de textos científicos introdutórios sobre o assunto e de fácil entendimento para alunos de graduação. Além disso, foi apresentado os principais pontos da ABNT NBR ISO 14.224 assim como a base de dados para cálculos de confiabilidade OREDA.

O conhecimento de distribuições conjugadas, permitiu simplificar os cálculos, visto que as distribuições pertenciam as mesmas famílias. Tal conhecimento pouco explorado na graduação em engenharia, foi a base para o desenvolvimento do método proposto, desta forma todos os cálculos podem ser realizados em planilhas de forma simples, sem utilizar ferramentas de simulação numéricas.

A aplicação do método permite gerar estimativas de taxa de falha de equipamentos mais adequadas. Com isso, é possível obter previsões realísticas do comportamento do equipamento, em termos de confiabilidade e disponibilidade; assim, evitando o desperdício de recursos. Para sua aplicação, a maior dificuldade é a coleta de dados, visto que os equipamentos operam em unidades com restrito acesso e dificuldade de acesso.

A atualização bayesiana se mostrou eficaz para obtenção dos objetivos propostos neste trabalho, os resultados de taxa de falha obtidos se encontram num intervalo menor, frente aos intervalos disponíveis na base. Destaca-se que este processo pode ser repetido periodicamente e, com isso, gerar novos intervalos de confiança – cada vez mais ajustados ao equipamento em análise.

Por fim, através do presente estudo, identificou-se que futuros trabalhos podem ser realizados para verificar a eficácia do método com dados colhidos em campo, assim como classificação das falhas ocorridas e predição da taxa de falha de acordo com cada modo de falha.

## REFERÊNCIAS

ASSIS, Janilson Pinheiro de; SOUZA, Roberto Pequeno de; DIAS, Carlos Tadeu dos Santos. Glossário de Estatística. **Mossoró/RN, EdUFERSA, 901f**, 2019.

ASSOCIAÇÃO BRASILEIRA DE NORMAS TÉCNICAS. **NBR 14224**: Indústrias de petróleo e gás natural — Coleta e intercâmbio de dados de confiabilidade e manutenção para equipamentos. Rio de Janeiro, out. 2011.

\_\_\_\_\_. **NBR 20815**: Indústrias de petróleo, petroquímica e gás natural — Garantia da produção e gestão da confiabilidade. Rio de Janeiro, jan. 2017.

BARROS, Angela Alves de Araújo. Distribuições Exponencializadas e Estendidas: uma abordagem clássica e bayesiana. Universidade Federal Rural de Pernambuco, 2008. Disponível em: [http://www.ppgbea.ufrpe.br/sites/www.ppgbea.ufrpe.br/files/documentos/dissertacao\\_final\\_angela\\_alves\\_de\\_araujo\\_barros.pdf](http://www.ppgbea.ufrpe.br/sites/www.ppgbea.ufrpe.br/files/documentos/dissertacao_final_angela_alves_de_araujo_barros.pdf).

CAEIRO, Frederico. Probabilidade e Estatística. Faculdades de Ciências e Tecnologia da Universidade de Lisboa, 2009.

CALIL, Lus Fernando Peres; HIRANO, Eduardo Wulff; DIAS, Acires. A RISKS QUANTIFICATION PROCEDURE BASED ON BAYESIAN INFERENCE. *In*: PROCEEDINGS of the COBEM 2005: 18 th International Congress of Mechanical Engineering. [S.l.: s.n.], 2005. Disponível em: <https://www.abcm.org.br/anais/cobem/2005/PDF/COBEM2005-1142.pdf>.

COLOSIMO, Enrico Antônio; GIOLO, Suely Ruiz. **Análise de Sobrevida Aplicada**. 1. ed. São Paulo: [s.n.], 2006.

DANTAS, Carlos Alberto Barbosa. **Probabilidade: Um Curso Introdutório Vol. 10**. [S.l.]: Edusp, 2013.

DIAS, Acires. Confiabilidade na manutenção industrial. **Florianópolis: Departamento de Engenharia Mecânica, Universidade Federal de Santa Catarina**, 2009.

EHLERS, Ricardo Sandes. Introdução à Inferência bayesiana. URL: <http://www.leg.ufpr.br/%7Eepaulojus/CE227/ce227.pdf>, 2003.

GERMANO, Arthur Oliveira. Efeitos da implantação de manutenções preventivas na produtividade - Indústrias Becker. Universidade Federal do Rio Grande do Norte, 2018.

GONÇALVES, Mauro César. **Concepções de professores e o ensino de probabilidade na escola básica. 2004**. 2004. Tese (Doutorado) – Dissertação (Mestrado)-PUC-SP, São Paulo.

HOFF, Peter D. **A first course in Bayesian statistical methods**. [S.l.]: Springer, 2009. v. 580.

JESUS, Sandra Rêgo de. Análise bayesiana objetiva para as distribuições normal generalizada e lognormal generalizada. Universidade Federal de São Carlos, 2014.

KARDEC, Alan; NASCIF, Júlio. Manutenção: função estratégica. 3ª edição. **Rio de Janeiro: Qualitymark: Petrobrás**, 2009.

KOEHLER, Cristiane *et al.* Uma abordagem probabilística para sistemas especialistas. Florianópolis, SC, 1998.

LANGSETH, Helge; HAUGEN, Knut; SANDTORV, Helge. Analysis of OREDA data for maintenance optimisation. **Reliability Engineering & System Safety**, Elsevier, v. 60, n. 2, p. 103–110, 1998.

LOUREIRO, Alexandre José Figueiredo. Análise de propagação em vegetação utilizando Bayes e UT, 2018.

MAGALHÃES, Marcos Nascimento. **Probabilidade e variáveis aleatórias**. [S.l.]: Edusp, 2006.

MATTOS, Thalita do Bem. Noções de Inferência no R. **IMECC/UNICAMP**, 2018.

MONTGOMERY, Douglas C; RUNGER, George C; HUBELE, Norma F. **Engineering statistics**. [S.l.]: John Wiley & Sons, 2009.

OLIVEIRA COSTA NETO, Pedro Luiz de. **Estatística**. 2. ed. São Paulo: [s.n.], 2002.

PESCE, Rosana Angélica Gonçalves. Distribuições estatísticas: citações de publicações científicas e de cientistas. Universidade Estadual Paulista (UNESP), 2005.

SANDTORV, Helge A; HOKSTAD, Per; THOMPSON, David W. Practical experience with a data collection project: the OREDA project. **Reliability Engineering & System Safety**, Elsevier, v. 51, n. 2, p. 159–167, 1996.

SIGNORINI FILHO, Geraldo Antonio. Estudo das métricas de confiabilidade de um módulo de controle submarino aplicado em produção de petróleo no pré-sal. Universidade Tecnológica Federal do Paraná, 2019.

SILVA, Rodolfo Romualdo da; MATTOS, Néli Maria Costa. Análise Bayesiana da Confiabilidade de Itens Submetidos a Testes de Degradação. Abepro, 2001. Disponível em: [http://www.abepro.org.br/biblioteca/enegep2001\\_tr25\\_0637.pdf](http://www.abepro.org.br/biblioteca/enegep2001_tr25_0637.pdf).

SILVA, Sebastião Andreino da. Análise de Riscos Aplicada às Atividades de Exploração e Produção de Petróleo. Universidade Estácio de Sá - Departamento de Pós-Graduação em Tecnologia e Desenvolvimento, 2015.

SINTEF INDUSTRIAL MANAGEMENT. **OREDA 2002**. Norway, 2002.

SPIEGEL, Murray R. Manual de fórmulas e tabelas matemáticas, 2019.

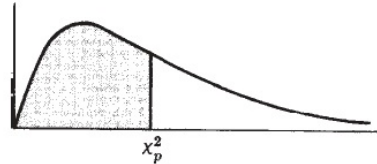
SPJØTVOLL, E. Estimation of failure rate from reliability data bases. *In: PRESENTATION at the SRE Symposium*. [S.l.: s.n.], 1985.

VERONEZ, Cleverson Alessandro *et al.* Gerência de desempenho do tráfego em redes utilizando baseline bayesiana. Florianópolis, SC, 2000.

## ANEXO A – VALORES PERCENTIS DA DISTRIBUIÇÃO QUI-QUADRADO

## Valores Percentis $\chi_p^2$ da Distribuição $\chi^2$ (Qui-Quadrado)

com  $n$  graus de liberdade (área sombreada =  $p$ )



$n$	$\chi_{.995}^2$	$\chi_{.99}^2$	$\chi_{.975}^2$	$\chi_{.95}^2$	$\chi_{.90}^2$	$\chi_{.75}^2$	$\chi_{.50}^2$	$\chi_{.25}^2$	$\chi_{.10}^2$	$\chi_{.05}^2$	$\chi_{.025}^2$	$\chi_{.01}^2$	$\chi_{.005}^2$
1	7.88	6.63	5.02	3.84	2.71	1.32	0,455	0,102	0,0158	0,0039	0,0010	0,0002	0,0000
2	10.6	9.21	7.38	5.99	4.61	2.77	1.39	0,575	0,211	0,103	0,0506	0,0201	0,0100
3	12.8	11.3	9.35	7.81	6.25	4.11	2.37	1.21	0,584	0,352	0,216	0,115	0,072
4	14.9	13.3	11.1	9.49	7.78	5.39	3.36	1.92	1.06	0,711	0,484	0,297	0,207
5	16.7	15.1	12.8	11.1	9.24	6.63	4.35	2.67	1.61	1.15	0,831	0,554	0,412
6	18.5	16.8	14.4	12.6	10.6	7.84	5.35	3.45	2.20	1.64	1.24	0,872	0,676
7	20.3	18.5	16.0	14.1	12.0	9.04	6.35	4.25	2.83	2.17	1.69	1.24	0,989
8	22.0	20.1	17.5	15.5	13.4	10.2	7.34	5.07	3.49	2.73	2.18	1.65	1.34
9	23.6	21.7	19.0	16.9	14.7	11.4	8.34	5.90	4.17	3.33	2.70	2.09	1.73
10	25.2	23.2	20.5	18.3	16.0	12.5	9.34	6.74	4.87	3.94	3.25	2.56	2.16
11	26.8	24.7	21.9	19.7	17.3	13.7	10.3	7.58	5.58	4.57	3.82	3.05	2.60
12	28.3	26.2	23.3	21.0	18.5	14.8	11,3	8.44	6.30	5.23	4.40	3.57	3.07
13	29.8	27.7	24.7	22.4	19.8	16.0	12.3	9.30	7.04	5.89	5.01	4.11	3.57
14	31.3	29.1	26.1	23.7	21.1	17.1	13,3	10.2	7.79	6.57	5.63	4.66	4.07
15	32.8	30.6	27.5	25.0	22.3	18.2	14.3	11.0	8.55	7.26	6.26	5.23	4.60
16	34.3	32.0	28.8	26.3	23.5	19.4	15.3	11.9	9.31	7.96	6.91	5.81	5.14
17	35.7	33.4	30.2	27.6	24.8	20.5	16.3	12.8	10.1	8.67	7.56	6.41	5.70
18	37.2	34.8	31.5	28.9	26.0	21.6	17.3	13.7	10.9	9.39	8.23	7.01	6.26
19	38.6	36.2	32.9	30.1	27.2	22.7	18.3	14.6	11.7	10.1	8.91	7.63	6.84
20	40.0	37.6	34.2	31.4	28.4	23.8	19.3	15.5	12.4	10.9	9.59	8.26	7.43
21	41.4	38.9	35.5	32.7	29.6	24.9	20.3	16.3	13.2	11.6	10.3	8.90	8.03
22	42.8	40.3	36.8	33.9	30.8	26.0	21.3	17.2	14.0	12.3	11.0	9.54	8.64
23	44.2	41.6	38.1	35.2	32.0	27.1	22.3	18.1	14.8	13.1	11.7	10.2	9.26
24	45.6	43.0	39.4	36.4	33.2	28.2	23.3	19.0	15.7	13.8	12.4	10.9	9.89
25	46.9	44.3	40.6	37.7	34.4	29.3	24.3	19.9	16.5	14.6	13.1	11.5	10.5
26	48.3	45.6	41.9	38.9	35.6	30.4	25.3	20.8	17.3	15.4	13.8	12.2	11.2
27	49.6	47.0	43.2	40.1	36.7	31.5	26.3	21.7	18.1	16.2	14.6	12.9	11.8
28	51.0	48.3	44.5	41.3	37.9	32.6	27.3	22.7	18.9	16.9	15.3	13.6	12.5
29	52.3	49.6	45.7	42.6	39.1	33.7	28.3	23.6	19.8	17.7	16.0	14.3	13.1
30	53.7	50.9	47.0	43.8	40.3	34.8	29.3	24.5	20.6	18.5	16.8	15.0	13.8
40	66.8	63.7	59.3	55.8	51.8	45.6	39.3	33.7	29.1	26.5	24.4	22.2	20.7
50	79.5	76.2	71.4	67.5	63.2	56.3	49.3	42.9	37.7	34.8	32.4	29.7	28.0
60	92.0	88.4	83.3	79.1	74.4	67.0	59.3	52.3	46.5	43.2	40.5	37.5	35.5
70	104.2	100.4	95.0	90.5	85.5	77.6	69.3	61.7	55.3	51.7	48.8	45.4	43.3
80	116.3	112.3	106.6	101.9	96.6	88.1	79.3	71.1	64.3	60.4	57.2	53.5	51.2
90	128.3	124.1	118.1	113.1	107.6	98.6	89.3	80.6	73.3	69.1	65.6	61.8	59.2
100	140.2	135.8	129.6	124.3	118.5	109.1	99.3	90.1	82.4	77.9	74.2	70.1	67.3

Fonte: Adaptado de Spiegel (2019)