



UNIVERSIDADE ESTADUAL DA PARAÍBA
CENTRO DE CIÊNCIAS E TECNOLOGIA
DEPARTAMENTO DE ESTATÍSTICA

Isaac Pereira Santos

Introdução à análise de confiabilidade: Uma aplicação ao setor de Transportes.

Campina Grande - PB

Abril de 2016

Isaac Pereira Santos

**Introdução à análise de confiabilidade: Uma aplicação ao
setor de Transportes.**

Trabalho de Conclusão de Curso apresentado ao Curso de Especialização em Estatística Aplicada do Departamento de Estatística do Centro de Ciências e Tecnologia da Universidade Estadual da Paraíba em cumprimento às exigências legais para obtenção do título de especialista em Estatística.

Orientador: Tiago Almeida de Oliveira

Campina Grande - PB

Abril de 2016

É expressamente proibida a comercialização deste documento, tanto na forma impressa como eletrônica. Sua reprodução total ou parcial é permitida exclusivamente para fins acadêmicos e científicos, desde que na reprodução figure a identificação do autor, título, instituição e ano da dissertação.

S237i Santos, Isaac Pereira
Introdução à análise de confiabilidade: uma aplicação ao setor de transportes. [manuscrito] / Isaac Pereira Santos. - 2016.
46 p. : il. color.

Digitado.

Monografia (Estatística Aplicada) - Universidade Estadual da Paraíba, Centro de Ciências e Tecnologia, 2016.

"Orientação: Prof. Drº Tiago Almeida de Oliveira, Departamento de Estatística".

1.Tempo de Vida. 2.Modelos Paramétricos. 3. Análise de Confiabilidade. I. Título.

21. ed. CDD 519.544

Isaac Pereira Santos

Introdução à análise de confiabilidade: Uma aplicação ao setor de Transportes.

Trabalho de Conclusão de Curso apresentado ao Curso de Especialização em Estatística Aplicada do Departamento de Estatística do Centro de Ciências e Tecnologia da Universidade Estadual da Paraíba em cumprimento às exigências legais para obtenção do título de especialista em Estatística.

Trabalho aprovado em 20 de Abril de 2016.

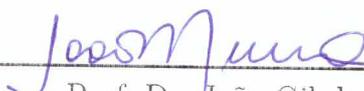
BANCA EXAMINADORA



Prof. Dr. Tiago Almeida de Oliveira
Universidade Estadual da Paraíba



Prof. Dr. Ricardo Alves de Olinda
Universidade Estadual da Paraíba



Prof. Dr. João Gil de Luna
Universidade Estadual da Paraíba

*As minhas filhas Ana Sophia Felix Santos e Isadora Felix Santos, A minha mãe Angel
Pereira Santos, Ao meu Pai Cícero dos Santos, A minha esposa Andreia Felix Da Cruz
com carinho.*

Agradecimentos

A Deus, por ter me dado força e discernimento para superar os momentos difíceis.

À minha esposa Andreia, pela compreensão por minha ausência com nossas filhas, nos momentos que me afastei para estudar.

À professora Ana Patricia Bastos Peixoto pelas leituras sugeridas ao longo dessa orientação e pela dedicação.

Ao professor Tiago Almeida de Oliveira, pela grande contribuição no desenvolvimento deste trabalho, com sugestões de leituras, esclarecimento de dúvidas e auxílio com as análises.

A todos os professores do Curso de Especialização da UEPB.

Aos colegas de classe pelos momentos de amizade e apoio.

Ao meu irmão Yranilson Rangel, por me ajudar e substituir no trabalho nos momentos em que tive que me ausentar para estudar.

A todos que acreditaram em mim e fizeram/fazem parte da minha vida, o meu obrigada a todos.

“Fé inabalável é somente aquela que pode encarar a razão, face a face, em todas as épocas da humanidade.”
(Allan Kardec)

Resumo

De acordo com a Confederação Nacional de Transportes, o transporte rodoviário detém a maior participação na matriz do transporte de cargas no Brasil, com 61,1% o que corresponde a 485,6 bilhões de TKU (tonelada por quilômetro útil), com cerca de 1,4 milhões de caminhões e ônibus, que consomem em torno 5,8 milhões de pneus anualmente. O presente trabalho tem por objetivo modelar o tempo de vida em quilômetros rodados do Pneu tipo 11.00R22, utilizando os conceitos de análise de confiabilidade paramétrica. De maneira geral, a confiabilidade pode ser definida como a operação bem sucedida de um item, unidade ou sistemas, sob condições de operação estabelecidas, por um determinado período de tempo, sem que ocorram quebras ou falhas. O tempo de uso foi medido por meio de computador de bordo acoplado nos caminhões, sendo este tempo de uso medido em quilômetros percorridos. Utilizou-se as funções de densidade, confiabilidade e risco. Foi utilizado o estimador produto-limite (Kaplan e Meier, 1958). Após a aplicação do estimador produto-limite verificou-se se os tempos de sobrevivência dos pneus em cada grupo de vida eram similares ou se apresentavam diferença estatística quanto a sua duração, por meio do uso do teste de log-rank. Após o uso das técnicas não paramétricas fez-se o ajuste dos modelos paramétricos, Normal, Log.Normal, Weibull e Exponencial, obtendo-se suas funções de densidade, sobrevivência e risco. Após este ajuste comparou-se por meio de teste de aderência de Anderson Darling, teste de razão de verossimilhança, critério de AIC e por meio de análises gráficas se as distribuições estavam bem ajustadas. Todas as análises foram realizadas utilizando o software R 3.2.2 (TEAM, 2015), por meio do pacote Survival versão 2.36 (THERNEAU; LUMLEY, 2014) e o software ActionEstat. As distribuições Lognormal e Weibull se apresentaram como fortes candidatas a modelar o tempo de vida de pneus 11.00R22. Todavia de acordo com os critérios de seleção utilizados, AIC (menor valor) e TRV (valor $P > 0,05$), chegou-se a conclusão de que a distribuição Lognormal é a mais adequada para modelar o tempo de vida de pneus 11.00R22.

Palavras-chaves: Tempo de Vida; Modelos Paramétricos; Análise de Confiabilidade.

Abstract

According to the National Transport Confederation, road transport has the largest share in cargo transportation matrix in Brazil, with 61.1 % corresponding to 485.6 billion RTK (ton per kilometer), with about 1.4 million trucks and buses that consume around 5.8 million tires annually. This study aims to model the lifetime of kilometers traveled 11.00R22 tire type, using the concepts of parametric reliability analysis. In general, reliability can be defined as the successful operation of an item, drive systems or under established operating conditions for a certain period of time without the occurrence of breakages or failures. The usage time was measured by on-board computer coupled to the trucks, and this usage time measured in kilometers traveled. the density functions, reliability and risk was used. It used the estimator product-limit (Kaplan and Meier, 1958). After applying the product-limit estimator it was found that the tires of survival time in each group were similar life or had a statistical difference in their length, by using the log-rank test. After using nonparametric techniques became adjusting parametric models, Normal, Log.Normal, exponential and Weibull, obtaining its density functions, survival risk. After this setting compared by means of adhesion Anderson Darling test, likelihood ratio test, and test the ICA by graphical analysis if the distributions were well adjusted. All analysis were performed using the R software 3.2.2 (TEAM, 2015), by means of Survival 2:36 package version (THERNEAU; LUMLEY, 2014) and ActionEstat software. The lognormal and Weibull presented themselves as strong candidates to shape the tire lifetime 11.00R22. However according to the selection criteria, AIC (lower value) and VRT (P value > 0, 05), came to the conclusion that the lognormal distribution is the most suitable for modeling the tire lifetime 11.00R22 .

Key-words: Lifetime; Parametric models; Reliability Analysis.

Lista de ilustrações

Figura 1 – Representação gráfica de três funções de risco.	17
Figura 2 – Representação gráfica dos mecanismo de censura, em que (●) representa falha e (○) representa censura.	21
Figura 3 – Função sobrevivência e função taxa de falha do modelo exponencial para diferentes valores de λ	26
Figura 4 – Função de confiabilidade, de risco e de risco acumulado com parâmetro escala $\lambda = 1$ e diferentes valores do parâmetro de forma s	28
Figura 5 – Funções de densidade de probabilidade, de confiabilidade e de risco da distribuição Log-Normal para diferentes valores dos parâmetros $(\mu; \sigma)$	29
Figura 6 – Gráfico do ajuste de Kaplan-Meier para os Pneus 11.00R22.	38
Figura 7 – Gráfico do ajuste de Kaplan-Meier versus as distribuições Normal, Exponencial, Weibull e Lognormal e Representação gráfica dos Papeis de probabilidade para as mesmas distribuições.	40
Figura 8 – Gráfico das funções estimadas para o modelo Weibull e Lognormal.	42
Figura 9 – Gráfico dos resíduos padronizados para as distribuições Weibull, Lognormal e para o estimador de Kaplan-Meier	43
Figura 10 – Gráfico dos resíduos Cox-Snell para as distribuições Weibull, Lognormal e para o estimador de Kaplan-Meier	43

Lista de tabelas

Tabela 1 – Distribuição das frequências dos eventos de falha e tempo mediano de acordo com o grupo de vida dos pneus 11.00R22.	38
Tabela 2 – Teste de log-rank aplicados aos grupos analisados.	39
Tabela 3 – Teste de Anderson - Darling para as diferentes distribuições analisadas	39
Tabela 4 – Estimativas dos parâmetros das distribuições Weibull e Lognormal. . .	40
Tabela 5 – Estimativas dos percentis obtidos por meio das distribuições Weibull e Lognormal com respectivos desvios padrão e intervalos de confiança. . .	41
Tabela 6 – Estimativas dos quartis, do tempo médio até a falha, da mediana e desvio padrão obtidos pelas distribuições Weibull e Lognormal.	41
Tabela 7 – Testes de razão de verossimilhanças - TRV e Critério de AIC para as distribuições estudadas.	42

Sumário

1	INTRODUÇÃO	12
2	FUNDAMENTAÇÃO TEÓRICA	13
2.1	Confiabilidade	13
2.2	Principais conceitos de confiabilidade	13
2.2.1	Gestão de confiabilidade	14
2.2.2	Medidas de confiabilidade	15
2.2.2.1	Tempo até a falha	15
2.2.2.2	Função de confiabilidade, $R(t)$	16
2.2.2.3	Função de risco $h(t)$	16
2.2.2.4	Tempo médio até falha, MTTF	18
2.2.3	Função de vida residual média, $L(t)$	19
2.3	Análise de dados censurados	19
2.4	Distribuições de probabilidade em confiabilidade	21
2.4.1	Modelos Não-Paramétricos	21
2.4.1.1	Estimador de Kaplan-Meier	22
2.4.1.2	Estimador de Nelson Aalen	23
2.4.1.3	Estimador da tabela de vida ou atuarial	23
2.4.2	Teste log-rank	24
2.5	Modelos Paramétricos	25
2.5.1	Modelo exponencial	25
2.5.2	Modelo Weibull	27
2.5.3	Modelo Log-Normal	28
2.5.4	Distribuição Gama e Gama Generalizada	30
2.6	Métodos de estimação dos parâmetros	30
2.6.1	Método da máxima verossimilhança	30
2.6.1.1	Função de verossimilhança para dados censurados	31
2.7	Teste de Anderson-Darling	32
2.8	Escolha do modelo paramétrico	32
2.8.1	Teste da razão de verossimilhanças (TRV)	33
2.8.2	Análise gráfica do ajuste do modelo	33
2.8.3	Medida global de qualidade do ajuste	34
2.8.4	Análise de resíduos	34
2.8.4.1	Resíduo de Cox-Snell	34
2.8.4.2	Resíduos Martingale	35

2.8.4.3	Resíduos escore: pontos influentes	35
2.8.5	Critérios de Informação	36
3	MATERIAL E MÉTODOS	37
4	RESULTADOS E DISCUSSÃO	38
5	CONCLUSÃO	44
	REFERÊNCIAS	45

1 Introdução

O transporte rodoviário é caracterizado pelo uso de veículos como caminhões, ônibus e carretas realizados em estradas de rodagem. O modal rodoviário corresponde ao percurso entre porto e embarcador ou consignatário. Nesse percurso o modal rodoviário, geralmente, é utilizado para o transporte de produtos industrializados por possuírem um maior valor agregado, e também em função da confiabilidade que apresenta. O transporte rodoviário também é bastante utilizado por produtos agrícolas, com maior frequência em épocas de safra.

De acordo com a Confederação Nacional de Transportes, o transporte rodoviário detém a maior participação na matriz do transporte de cargas no Brasil, com 61,1% o que corresponde a 485,6 bilhões de TKU (tonelada por quilômetro útil), com cerca de 1,4 milhões de caminhões e ônibus, que consomem em torno 5,8 milhões de pneus anualmente.

Segundo [Haviaras \(2005\)](#), depois de custos com combustível e mão de obra, os custos relativos a pneus é o de maior destaque em transportes de carga. De acordo com [Costa \(2000\)](#), 53,5% das paradas de caminhões em estradas, com falhas que causam sua imobilização, foram causadas por ocorrências relacionadas a pneus. Este fato não só acarreta prejuízo associado ao custo do pneu, mas também gera perda de tempo de viagem com o conseqüente custo associado a atrasos nas entregas, comprometendo a confiabilidade do serviço de transporte.

Segundo [Fogliatto e Ribeiro \(2009\)](#), com o advento da economia globalizada, observou-se um aumento na demanda por produtos e sistemas de melhor desempenhos e custos competitivos. Conseqüentemente, surgiu a necessidade de redução na probabilidade de falhas em produtos, o que resultou numa ênfase crescente em sua confiabilidade. É importante utilizar ferramentas que descrevam o comportamento dessas falhas, uma metodologia capaz de obter essas informações é a análise de confiabilidade.

De maneira geral, confiabilidade pode ser definida como a operação bem sucedida de um item, unidade ou sistemas, sob condições de operação estabelecidas, por um determinado período de tempo, sem que ocorram quebras ou falhas. No entanto, em análise de confiabilidade é necessária uma definição quantitativa em termos de uma probabilidade. Algumas distribuições podem ser utilizadas para representar a probabilidade de falha, possibilitando determinar a partir de que momento o item, unidade ou sistema sob análise é considerado com desempenho abaixo daquele apontado como aceitável, e portanto considerando a ocorrência de falha ([HAVIARAS, 2005](#)).

O presente trabalho tem por objetivo modelar o tempo de vida em quilômetros rodados do Pneu tipo 11.00R22, utilizando os conceitos de análise de confiabilidade paramétrica.

2 Fundamentação Teórica

O conteúdo desta seção aborda os principais conceitos em análise de sobrevivência, mais especificamente enfocando a análise de confiabilidade, utilizando-se de livros práticos e teóricos, com o objetivo de obter-se informações sobre o assunto abordado.

2.1 Confiabilidade

De acordo com [Fogliatto e Ribeiro \(2009\)](#) após o término da Primeira Guerra Mundial o conceito de confiabilidade adquiriu um significado tecnológico, quando foi utilizado para descrever estudos comparativos feitos em aviões com um, dois ou quatro motores. Naquele contexto a confiabilidade era medida como número de acidentes por hora de voo. Outro fato de grande relevância para o avanço na área da confiabilidade foi a corrida para ser a primeira nação a enviar uma missão tripulada à lua, tendo em vista os riscos humanos envolvidos. Em 1963, surgiu, nos Estados Unidos, a primeira associação que reunia engenheiros de confiabilidade e o primeiro periódico para divulgação de trabalhos na área, o IEEE - *Transaction on Reliability*.

Na década de 1970, o estudo da confiabilidade esteve centrado na análise dos riscos associados a construção e operação de usinas nucleares. A partir daí, aplicações da confiabilidade nas mais diversas áreas se consolidaram. Algumas dessas áreas de aplicação a engenharia de produção, foram elencadas por [Rausand e Hoyland \(2003\)](#), sendo listadas como análises de riscos e segurança, qualidade, otimização da manutenção, proteção ambiental, proteção ambiental e projeto de produtos ([FOGLIATTO; RIBEIRO, 2009](#)).

2.2 Principais conceitos de confiabilidade

Segundo [Fogliatto e Ribeiro \(2009\)](#), por definição usual, em seu sentido mais amplo, o termo confiabilidade está associado à operação bem sucedida de um produto ou sistema, na ausência de quebras ou falha sob condições de uso previamente determinadas, em um período de tempo estabelecido. A confiabilidade de um item equivale à sua probabilidade de executar adequadamente sua função específica, por um determinado período de tempo e sob condições ambientais predeterminadas. Confiabilidade é definida como uma probabilidade, isso significa que todas as confiabilidades devem apresentar valores entre 0 e 1 e que os axiomas clássicos da probabilidade podem ser aplicados em cálculos de confiabilidade ([LEEMIS, 1995](#)).

Os principais conceitos associados à confiabilidade são: qualidade, disponibilidade,

manutenibilidade, segurança e confiança. Tais conceitos são definidos a seguir tendo como base principal o texto das normas NBR ISO - 8402 (1994) e 5462 (1994).

- i) Qualidade: Define-se como cumprimento a especificações de projeto e manufatura com menor variabilidade possível.
- ii) Disponibilidade: é definida como a capacidade de um item, mediante manutenção apropriada, desempenhar sua função requerida em um determinado instante de tempo ou em um período de tempo predeterminado. Esse conceito varia conforme a capacidade de reparo de uma unidade. Em unidades não-reparáveis, os conceitos de disponibilidade e confiabilidade se equivalem. Em unidades reparáveis, os possíveis estados de tempo da unidade em um tempo t de análise são funcionando ou em manutenção. Nestes casos, costuma-se supor que reparos devolvam a unidade à condição de nova e trabalha-se com um valor médio de disponibilidade para a unidade dada por

$$A = \frac{MTTF}{MTTF + MTTR}, \quad (2.1)$$

em que, A denota a disponibilidade média da unidade, $MTTF$ é o tempo médio entre as falhas, ou seja, o tempo médio de funcionamento da unidade e $MTTR$ é o tempo médio até a conclusão de reparos feitos na unidade.

- iii) Manutenibilidade: é a capacidade de um item, quando submetido a manutenção, ser mantido ou recolocado em condições de desenvolver suas funções requeridas, mediante condições preestabelecidas de uso. A manutenibilidade é um fator de extrema importância no estabelecimento de uma unidade.
- iv) Segurança: é definido como a ausência (ou nível aceitável de risco) de condições que possam causar morte, dano ou doenças comportamentais a pessoas, assim como dano ou perda de equipamentos ou propriedade.
- v) Confiança: O termo confiança ou dependabilidade é utilizado para designar um coletivo que inclui a disponibilidade e seus fatores: o desempenho da confiabilidade, da manutenibilidade e do suporte técnico. Pode-se considerar os conceitos de confiança e confiabilidade como análogos; o termo confiança, todavia, estaria associado a uma definição mais ampla, mas não estritamente probabilística de confiabilidade.

2.2.1 Gestão de confiabilidade

Fogliatto e Ribeiro (2009) afirmam que um programa integrado de confiabilidade compreende o estabelecimento de práticas e procedimentos para gerir a confiabilidade nas seguintes fases da vida de um produto:

- i) projeto e desenvolvimento,

- ii) manufatura e instalação,
- iii) operação e manutenção,
- iv) descarte, quando encerra a vida operacional do produto.

Um programa de confiabilidade define a estrutura organizacional, responsabilidades, procedimentos, processos e recursos utilizados na gestão da confiabilidade. Os elementos de um programa de confiabilidade incluem uma tarefa ou conjunto de tarefas realizadas por um indivíduo ou uma equipe, assim como a implantação de um sistema eficiente de coleta e análise de dados, que permita a construção de uma base histórica de dados de confiabilidade na empresa.

2.2.2 Medidas de confiabilidade

Serão apresentadas a seguir as diversas medidas de confiabilidade para uma unidade não-reparável.

2.2.2.1 Tempo até a falha

Segundo Colosimo e Giolo (2006), para realizar uma análise de confiabilidade, é preciso possuir informações sobre o tempo de falha e censura, estes dois componentes constituem a resposta na análise de confiabilidade. O tempo de início do estudo deve ser bem definido, permitindo comparar os indivíduos na origem do estudo, com exceção de diferenças medidas pelas covariáveis.

Por tempo até a falha, entende-se o tempo transcorrido desde o momento em que a unidade é colocada em operação até que ocorra sua primeira falha. Convencionalmente, $t = 0$ é dado como o início da operação do sistema. Por estar sujeito a variações aleatórias, o tempo até a falha é definido como uma variável aleatória designada por T . O estado da unidade em um tempo t pode ser descrito por uma variável de estado $X(t)$, que é uma variável aleatória definida por dois estados: $X(t) = 1$, no caso de uma unidade operando no tempo t , e $X(t) = 0$, no caso de a unidade estar não-operacional no tempo t . O tempo de falha nem sempre é medido de forma contínua, podendo assumir valores discretos, como o número de ciclos até a falha.

Seja uma variável T distribuída continuamente, com densidade de probabilidade dada por $f(t)$ e função de distribuição dada por

$$F(t) = P(T \leq t) = \int_0^t f(u)du, \quad t > 0. \quad (2.2)$$

A função $F(t)$ denota a probabilidade de falha da unidade numa missão com duração menor ou igual a t .

A densidade de probabilidade $f(t)$ é definida como

$$f(t) = F'(t) = \frac{d}{dt}F(t) = \lim_{\Delta t \rightarrow 0} \frac{F(t + \Delta t) - F(t)}{\Delta t} = \lim_{\Delta t \rightarrow 0} \frac{P(t < T \leq t + \Delta t)}{\Delta t}. \quad (2.3)$$

Para valores pequenos de Δt , a seguinte aproximação pode ser usada:

$$P(t < T \leq t + \Delta t) \approx f(t) \times \Delta t.$$

2.2.2.2 Função de confiabilidade, $R(t)$

Suponha n_0 unidades idênticas submetidas a teste em condições predefinidas. Transcorrido o intervalo $(t - \Delta t, t)$, $n_f(t)$ unidades falharam e $n_s(t)$ sobreviveram, tal que $n_f(t) + n_s(t) = n_0$. A confiabilidade da unidade é definida como a sua probabilidade acumulada de sucesso; assim, em um tempo t , a função de confiabilidade $R(t)$ é

$$R(t) = \frac{n_s(t)}{n_s(t) + n_f(t)} = \frac{n_s(t)}{n_0}. \quad (2.4)$$

Conforme anteriormente definido, sendo, T uma variável aleatória, a função de confiabilidade em um tempo t pode ser expressa como:

$$R(t) = P(T > t). \quad (2.5)$$

A função de distribuição de T , $F(t)$, é o complemento de $R(t)$, ou seja:

$$R(t) = 1 - F(t) = 1 - \int_0^t f(u)du = \int_t^{+\infty} f(u)du. \quad (2.6)$$

Logo, a função de confiabilidade (ou função de sobrevivência) $R(t)$ indica a probabilidade de uma unidade apresentar sucesso em seu funcionamento (ou seja, ausência de falhas) no intervalo de tempo $(0, t)$ e ainda estar funcionando no tempo t (FOGLIATTO; RIBEIRO, 2009, p.10).

2.2.2.3 Função de risco $h(t)$

A função de risco, ou taxa de falha é definida como a probabilidade de uma unidade venha a falhar durante um intervalo $(t; t + \Delta t]$, sabendo-se que o item está funcionando no instante de tempo t é dada pela expressão

$$P(t \leq T \leq t + \Delta t) = \int_t^{t+\Delta t} f(u)du = R(t) - R(t + \Delta t). \quad (2.7)$$

Condicionando no evento de a unidade estar operando no tempo t , chega-se a seguinte expressão

$$P(t \leq T \leq t + \Delta t | T \geq t) = \frac{P(t \leq T \leq t + \Delta t)}{P(T \geq t)} = \frac{R(t) - R(t + \Delta t)}{R(t)}. \quad (2.8)$$

Uma função de risco média no intervalo $(t, t + \Delta t)$ pode ser obtida dividindo a Equação (2.8) por Δt . Supondo $\Delta t \rightarrow 0$, obtém-se a taxa de risco instantânea, que é a função de risco, dada por:

$$h(t) = \lim_{\Delta t \rightarrow 0} \frac{R(t) - R(t + \Delta t)}{R(t)\Delta t} = \frac{-R'(t)}{R(t)} = \frac{f(t)}{R(t)}, \quad t \geq 0. \quad (2.9)$$

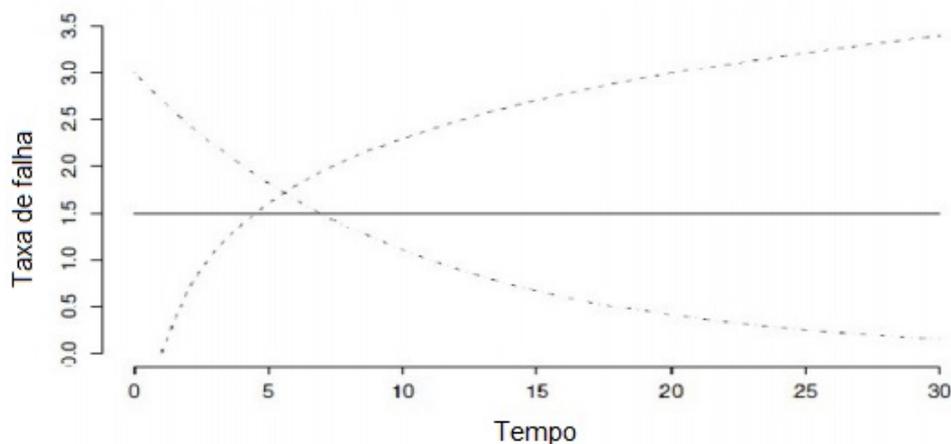
Funções de risco devem satisfazer as seguintes condições

- i) $\int_0^{+\infty} h(t)dt = +\infty$
- ii) $h(t) \geq 0$

Segundo Ramires (2013) a função de risco é um indicativo da forma como uma unidade envelhece, assim como, a quantidade de risco a que uma unidade está exposta em um tempo t . A função de risco pode ser classificadas de três formas básicas

- i) Crescente, **FRC**, em que a incidência de risco cresce com o tempo;
- ii) Decrescente, **FRD**, em que a incidência de risco diminui a medida que o tempo passa;
- iii) Constante ou estacionária, **FRE**, em que a incidência de risco não sofre alterações com o passar do tempo.

Figura 1 – Representação gráfica de três funções de risco.



Fonte: Colosimo e Giolo (2006)

A função de confiabilidade $R(t)$ e a função de densidade $f(t)$ podem ser obtidas a partir da função de risco, conforme demonstrado abaixo. Usando como ponto de partida a Equação (2.9), tem-se

$$h(t) = -\frac{R'(t)}{R(t)} = -\frac{d}{dt} \ln R(t) \quad (2.10)$$

Como, por definição, $R(0) = 1$, tem-se:

$$\int_0^t h(u)du = -\ln R(t) \quad (2.11)$$

$$R(t) = e^{-\int_0^t h(u)du} \quad (2.12)$$

A partir da Equação (2.9) e do resultado na Equação (2.12), é possível estabelecer a seguinte relação entre $f(t)$ e a função de risco:

$$f(t) = h(t)e^{-\int_0^t h(u)du}, \quad t \geq 0 \quad (2.13)$$

Integrando-se a função de risco sobre um período de tempo, obtém-se a função acumulada de risco, $H(t)$, dada por:

$$H(t) = \int_0^t h(u)du, \quad t \geq 0. \quad (2.14)$$

A função acumulada de risco oferece uma representação alternativa da função de risco, na forma de uma função não-decrescente no tempo. Entretanto, de maneira análoga às funções de distribuição $F(t)$, a maioria das funções acumuladas de risco se assemelha na forma, independente da distribuição que modela os tempos até a falha (FOGLIATTO; RIBEIRO, 2009, p.12).

2.2.2.4 Tempo médio até falha, MTTF

O tempo médio até falha de uma unidade, designado por MTTF, pode ser definido como

$$MTTF = E(T) = \int_0^{+\infty} tf(t)dt. \quad (2.15)$$

ou seja, trata-se do valor esperado da variável T . Como, a partir da equação (2.3), $f(t) = -R'(t)$, uma expressão alternativa para MTTF pode ser assim obtida, tal qual

$$MTTF = -\int_0^{\infty} tR'(t)dt. \quad (2.16)$$

Integrando-se por partes, obtém-se

$$MTTF = -[tR(t)]_0^{\infty} + \int_0^{\infty} R(t)dt. \quad (2.17)$$

Se o $MTTF < \infty$, pode-se demonstrar que $[tR(t)]_0^{\infty} = 0$. Nesse caso, obtém-se a expressão alternativa para o MTTF, dada por

$$MTTF = \int_0^{\infty} R(t)dt. \quad (2.18)$$

Para a maioria das funções de distribuição que a variável T pode apresentar, a determinação do MTTF a partir da expressão na Equação (2.18) costuma ser mais fácil, se comparada à expressão na Equação (2.15).

2.2.3 Função de vida residual média, $L(t)$

Segundo Fogliatto e Ribeiro (2009), a função de vida residual média corresponde à vida remanescente esperada da unidade, dado que ela sobreviveu até o tempo t . Se T designar a duração da vida da unidade, a vida residual média corresponde ao intervalo $T - t$. A vida residual média é designada por $L(t)$ e dada por

$$L(t) = E[T - t | T \geq t], \quad t \geq 0. \quad (2.19)$$

O valor esperado não-condicional da distribuição de T , $E(T)$, é um caso especial da função $L(t)$, quando $t = 0$. A fórmula para cálculo da expectância na Equação (2.19) é dada por

$$L(t) = \int_t^\infty u \frac{f(u)}{R(t)} du - t = \frac{1}{R(t)} \int_t^\infty u f(u) du - t. \quad (2.20)$$

(FOGLIATTO; RIBEIRO, 2009, p.13).

2.3 Análise de dados censurados

Em estudos de confiabilidade, normalmente desenvolvem-se testes de vida com uma unidade (ou componente), para obter informações sobre a distribuição de probabilidade da unidade em estudo, com o objetivo de registrar seus tempos até a falha, nesses estudos, se o teste se desenvolver de maneira que venha a fornecer o tempo de falha de todas as unidades, o conjunto dos dados de tempo de falha obtido é dito *completo*. Na prática, em muitas situações, análises de confiabilidade não podem ser conduzidas com conjuntos completos de dados. Um conjunto de dados incompletos de tempos de falha é denominado **censurado**.

De acordo com Colosimo e Giolo (2006), censura é a observação parcial da resposta que foi interrompida por algum motivo, impedindo a observação do tempo de falha da unidade. Uma censura pode acontecer por várias razões, tais como: critérios de ordem prática ou econômica não permitirem que o estudo se desenvolva até que todas as unidades falhem, perda ou dano de algumas unidades durante o teste, ou não foi possível registrar o exato momento de ocorrência da falha, entre outras. Existem três tipos de censuras.

- i) Censura à direita: é aquela em que o tempo de ocorrência do evento está à direita do tempo de interesse, ou seja, o tempo entre o início do estudo e o evento é maior do que o tempo observado.

Segundo Lawless (2011), em ensaios de confiabilidade, três tipos de censura a direita ocorrem com maior frequência, são eles: censura tipo I, quando o teste de vida é interrompido em um tempo predeterminado; censura tipo II, o teste é interrompido após ter ocorrido o evento de interesse em um número predeterminado de unidades;

e censura tipo III, ocorre se a observação for interrompida no decorrer do estudo sem ter ocorrido o evento de interesse ou se o evento de interesse ocorrer por uma razão diferente da estudada.

- ii) Censura à esquerda: é aquela que acontece quando não conhecemos o momento da ocorrência do evento, mas sabemos que a duração do evento é menor que a observada.
- iii) Censura intervalar: Nesse caso, dados de tempos até a falha são agrupados em intervalos, sendo assim, não é possível determinar o momento exato da falha, somente será possível afirmar que o evento de interesse (falha) ocorreu em um certo intervalo de tempo.

Colosimo e Giolo (2006), representam de maneira simples o mecanismo de censura aleatória por meio de duas variáveis aleatórias, em que, suponha T uma variável aleatória representando o tempo até a falha de uma unidade, e S uma outra variável aleatória independente de T , representando o tempo de censura associado à esta unidade. Observa-se, portanto,

$$t = \min(T; S)$$

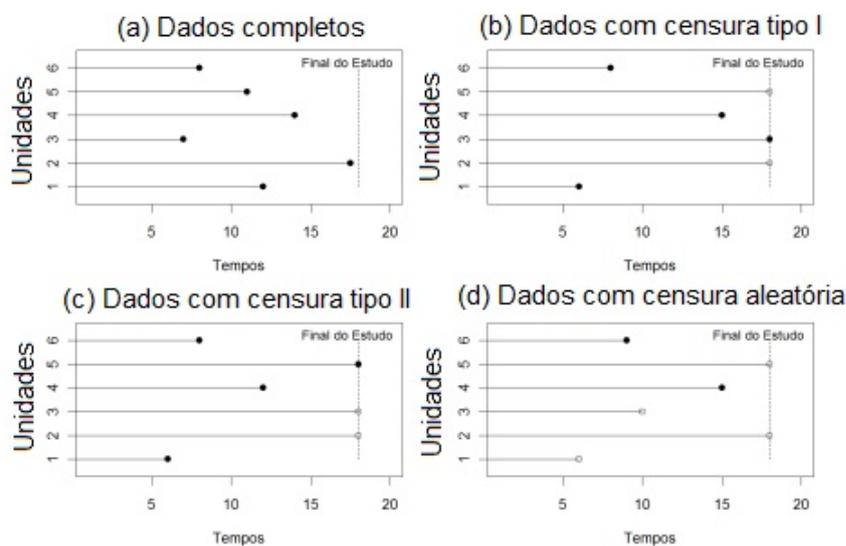
$$\delta = \begin{cases} 1, & \text{quando } T \leq S, \\ 0, & \text{quando } T > S. \end{cases} \quad (2.21)$$

em que, δ representa a variável indicadora de falha ou censura.

A Figura 2 apresenta os mecanismos de censura à direita, onde o tempo de ocorrência do evento de interesse está à direita do tempo observado.

De acordo com Leemis (1995), existem pelo menos três maneiras distintas de estudo de dados censurados, no entanto, ressalta, que uma abordagem correta consiste em tratar os dados censurados *probabilisticamente* incluindo os valores censurados na função de verossimilhança utilizada para estimar os parâmetros da distribuição que melhor caracteriza os tempos até a falha.

Figura 2 – Representação gráfica dos mecanismo de censura, em que (●) representa falha e (○) representa censura.



Fonte: Colosimo e Giolo (2006)

2.4 Distribuições de probabilidade em confiabilidade

A definição mais usual de confiabilidade de uma unidade (componente ou sistema) é dada em termos de sua probabilidade de sobrevivência até um tempo t de interesse. A determinação de tal probabilidade é possível através de modelagem dos tempos até falha da unidade em estudo. Conhecendo-se a distribuição de probabilidade que melhor se ajusta a esses tempos, é possível estimar a probabilidade de sobrevivência da unidade para qualquer tempo t . A modelagem dos tempos até a falha é, portanto, central em estudos de confiabilidade (FOGLIATTO; RIBEIRO, 2009, p.12).

2.4.1 Modelos Não-Paramétricos

Segundo Miranda (2012), em análise de confiabilidade, métodos não-paramétricos são utilizados quando os modelos paramétricos (probabilístico) não estão se adequando aos dados, estes estudos normalmente apresentam censuras. Algumas técnicas estatísticas podem ser utilizadas para analisar dados de tempo de confiabilidade na presença de censura.

De acordo com Hermeto (2014), os estimadores não-paramétricos, usam os próprios dados para estimar as quantidades necessárias da análise, ou seja, processo de estimação é realizado sem que se faça nenhuma suposição sobre a distribuição de probabilidade do tempo de confiabilidade. Podem ser citados três estimadores não-paramétricos usados para estimação da função de confiabilidade sendo eles:

- i) Estimador de Kaplan-Meier,
- ii) Estimador de Nelson Aalen,
- iii) Estimador da Tabela de Vida.

2.4.1.1 Estimador de Kaplan-Meier

Segundo [Herrmann \(2011\)](#), o estimador Kaplan-Meier ou estimador limite-produto é a técnica estatística mais importante e aplicada em análise de confiabilidade. A expressão, limite-produto, refere-se ao fato de que a ocorrência dos eventos são independentes e conseqüentemente a função de confiabilidade é estimada utilizando o produto das probabilidades de confiabilidade até o tempo t ([CARVALHO, 2011](#)).

O estimador de Kaplan-Meier estima uma curva de confiabilidade em que incorpora a informação da censura. Para incorporar a informação de censura a este estimador, tem no denominador o número de itens sob operação em determinado período. Ou seja, o estimador de Kaplan-Meier considera apenas os itens que estão operacionais no instante de tempo t . A expressão geral do estimador de Kaplan-Meier na forma adaptada a partir da função de confiabilidade empírica é dada por:

$$\hat{R}(t) = \frac{\text{número total de observações que não falharam até o tempo } t}{\text{número total de observações no estudo}},$$

em que, $\hat{R}(t)$ é uma função escada com degraus nos tempos observados até falha de tamanho $\frac{1}{n}$, onde n é o tamanho da amostra: serão tantos intervalos quanto forem a quantidade de falhas distintas. O tamanho dos degraus pode variar: será $\frac{1}{n}$ se não houver empates e será n° de empates/ n se houverem empates. Considera-se como empate duas ou mais falhas ocorridas no mesmo instante de tempo ([HERRMANN, 2011](#)).

De acordo com [Carvalho \(2011\)](#), a construção do estimador de Kaplan-Meier considera o número de intervalos iguais ao número de falhas distintas e os limites dos intervalos são os próprios tempos até falhas da amostra. Sejam $t_1 < t_2 < \dots < t_m$ os m diferentes tempos onde ocorreram os eventos (falha ou censura) em uma amostra com n indivíduos. O estimador da função de confiabilidade de Kaplan-Meier é definido como,

$$\hat{R}_{km}(t) = \prod_{j:t_j < t} \left(\frac{n_j - d_j}{n_j} \right) = \prod_{j:t_j < t} \left(1 - \frac{d_j}{n_j} \right), \quad (2.22)$$

em que d_j é o número de falhas no tempo t_j , e n_j é o número de itens sob risco em t_j , ou seja, o número de itens que não falharam e não foram censurados até o instante imediatamente anterior a t_j .

Segundo [Carvalho \(2011\)](#), o estimador Kaplan-Meier apresenta-se como não-viciado para grandes amostras, altamente consistente, e é estimador de máxima verossimilhança para a função de confiabilidade $R(t)$. A estimativa não muda nos tempos censurados, o

efeito dos tempos censurados é, entretanto, sentido nos valores de n_j , conseqüentemente, nos tamanhos dos degraus em $R(t)$.

2.4.1.2 Estimador de Nelson Aalen

Uma alternativa para estimar a função de taxa de risco acumulado é utilizar o estimador de Nelson-Aalen, dado por:

$$\hat{H}_{na}(t) = \sum_{j:t_j < t} \left(\frac{d_j}{n_j} \right). \quad (2.23)$$

Observe que, por meio da relação $H(t) = -\ln(S(t))$, podemos, a partir do estimador de Nelson-Aalen para $H(t)$, obter outra estimativa não paramétrica para confiabilidade dada por:

$$\hat{R}_{na}(t) = \exp(-\hat{H}_{na}(t)). \quad (2.24)$$

Para amostras suficientemente grandes, a estimativa de Kaplan-Meier e Nelson-Aalen para função de confiabilidade são bem próximas. Mas quando a amostra é pequena, o estimador Nelson-Aalen possui propriedades estatísticas que o tornam preferível (CARVALHO, 2011).

2.4.1.3 Estimador da tabela de vida ou atuarial

Segundo Herrmann (2011), o método não-paramétrico da tábua de vida, ou método atuarial, é um método que pode ser utilizado para estimar curvas de confiabilidade. Através da tábua de vida é possível estimar o tempo de confiabilidade dos itens (ou unidades) de amostras homogêneas. Normalmente, para que se possa organizar os tempos de confiabilidade em intervalos utiliza-se amostras de tamanho $n \geq 30$. O método atuarial calcula as probabilidades de confiabilidade em intervalos previamente fixados.

De acordo com Lee e Wang (2003) apud Herrmann (2011), o primeiro passo para criar uma tabela de vida consiste em dividir-se o tempo total de observação em determinado número de intervalos e para cada intervalo estima-se o valor da taxa de falha e a partir da obtenção desses valores estima-se a função de confiabilidade. O estimador da tabela de vida apresenta a mesma forma do Kaplan-Meier, mas utiliza q_j uma vez que se tenha d_j e n_j , em que

- i) $d_j =$ número de falhas no intervalo $[t_{j-1}, t_j)$ e
- ii) $n_j =$ [número sob risco em $t_{j-1}] - [\frac{1}{2} \times$ número de censuras em $[t_{j-1}, t_j]$];

O estimador da função de confiabilidade por meio da tábua de vida é definido como,

$$\hat{R}(t_j) = \prod_{i=1}^j (1 - \hat{q}_{i-1}), \quad j = 1, \dots, m. \quad (2.25)$$

Segundo [Herrmann \(2011\)](#), graficamente a função de confiabilidade apresenta uma forma de escada, com valor constante em cada intervalo de tempo. A função de confiabilidade é interpretada como a probabilidade de um item não falhar até o tempo t_j .

2.4.2 Teste log-rank

De acordo com [Carvalho \(2011\)](#), para comparar curvas de confiabilidade mais formalmente, deve-se recorrer a testes de hipóteses, dentre eles, o mais simples é o teste de Mantel-Haenzel, ou Log-rank. [Colosimo e Giolo \(2006\)](#), afirmam que um dos testes mais utilizados para estudar diferenças entre grupos, curvas de confiabilidade é o Teste log-rank. O teste compara os valores observados e esperados de cada estrato sob a hipótese de que o risco é o mesmo em todos os grupos.

Testar se curvas de confiabilidades são iguais equivale a testar se a incidência de eventos é semelhante em cada estrato. Se for semelhante, a curva de confiabilidade será a mesma ([CARVALHO, 2011](#)). Assim, a hipótese nula é definida como:

$$H_0 : \lambda_1(t) = \lambda_2(t) = \dots = \lambda_k(t),$$

em que, k é o número de estratos. Note que esta é uma hipótese que se refere à função de risco em todo tempo de observação e não às diferenças/semelhanças em trechos da curva. Rejeitar a hipótese nula significa que pelo menos uma curva difere das outras, significativamente, em algum momento do tempo ([CARVALHO, 2011](#)).

O teste log-rank compara a distribuição da ocorrência dos eventos observados em cada estrato com a distribuição que seria esperada se a incidência fosse igual em todos os estratos. Para calcular a distribuição esperada de eventos, o total de eventos precisamente no tempo t , $\Delta N(t)$, é redistribuído pelos k estratos, proporcionalmente ao número de itens presentes em cada estrato. Assim, para o estrato k , tem-se

$$E_k = N(t) \frac{R_k(t)}{R(t)}, \quad (2.26)$$

em que $\Delta N(t)$ é o número total de eventos observados em t , $R_k(t)$ é o número de itens em risco no estrato k no tempo t e $R(t)$ é o número total de itens em risco no estudo no tempo t . Quando apenas dois estratos estão sendo comparados ($k = 2$), a estatística log-rank é calculada utilizando-se os dados de apenas um dos estratos. O resultado do teste para um estrato se estende ao outro estrato por simetria.

A estatística log-rank é calculada a partir da diferença entre o número total de eventos observados e o número total de eventos esperados;

$$\text{Log-rank} = \frac{(O_1 - E_1)^2}{\text{Var}(O_1 - E_1)}, \quad (2.27)$$

em que, E_1 é o total de eventos esperados no estrato 1 e O_1 o total de eventos observados no estrato 1.

A variância, que entra no cálculo como um fator de padronização, tem a fórmula (para $k = 2$) dada por

$$\text{Var}(O_1 - E_1) = \sum_t \frac{R_1(t)R_2(t)\Delta N(t)[R(t) - \Delta N(t)]}{R(t)^2[R(t) - 1]}. \quad (2.28)$$

A estatística log-rank, sob a hipótese nula, segue uma distribuição χ^2 , com $k - 1$ graus de liberdade (CARVALHO, 2011).

2.5 Modelos Paramétricos

De acordo com Herrmann (2011), modelos paramétricos, também conhecidos por modelos probabilísticos, são baseados na suposição de uma distribuição de probabilidade para o tempo de confiabilidade por meio dos parâmetros estimados das amostras em estudo. Para Miranda (2012), o uso de modelos paramétricos têm se mostrado bastante adequado na análise estatística de dados de confiabilidade, com destaque para os modelos exponencial, weibull e log-normal, por suas comprovadas adequações a várias situações práticas, sendo estes considerados como principais modelos paramétricos para dados de confiabilidade.

2.5.1 Modelo exponencial

Segundo Carvalho (2011), o modelo exponencial é um modelo adequado para estudos onde, o tempo até falha é bem detalhado através de uma distribuição de probabilidade exponencial, foi o mais amplamente utilizado em estudos de confiabilidade em sistemas eletrônicos e posteriormente em estudos de tempo de vida, e em remissão de doenças crônicas e infecciosas no final dos anos 1940. Este modelo paramétrico é apontado como o método mais simples em termos matemáticos, e é também tido como um dos mais importantes entre os modelos paramétricos. Sua importância na análise de sobrevivência é comparada com a importância de uma distribuição normal nas diversas análises da área estatística (LEE; WANG, 2003) apud (HERRMANN, 2011).

De acordo com Herrmann (2011), a função de densidade de probabilidade para a variável aleatória tempo até falha T , com distribuição Exponencial é dada pela expressão

$$f(t) = \lambda \exp\{-\lambda t\}, \quad t \geq 0 \quad e \quad \lambda > 0. \quad (2.29)$$

O modelo Exponencial possui apenas um parâmetro λ . Este parâmetro representa o inverso do tempo médio de sobrevivência, isto é, o tempo médio de sobrevivência é dado por $1/\lambda$. A função de sobrevivência do modelo Exponencial é dada pela expressão:

$$R(t) = \exp\{-\lambda t\}, \quad t \geq 0 \quad e \quad \lambda > 0, \quad (2.30)$$

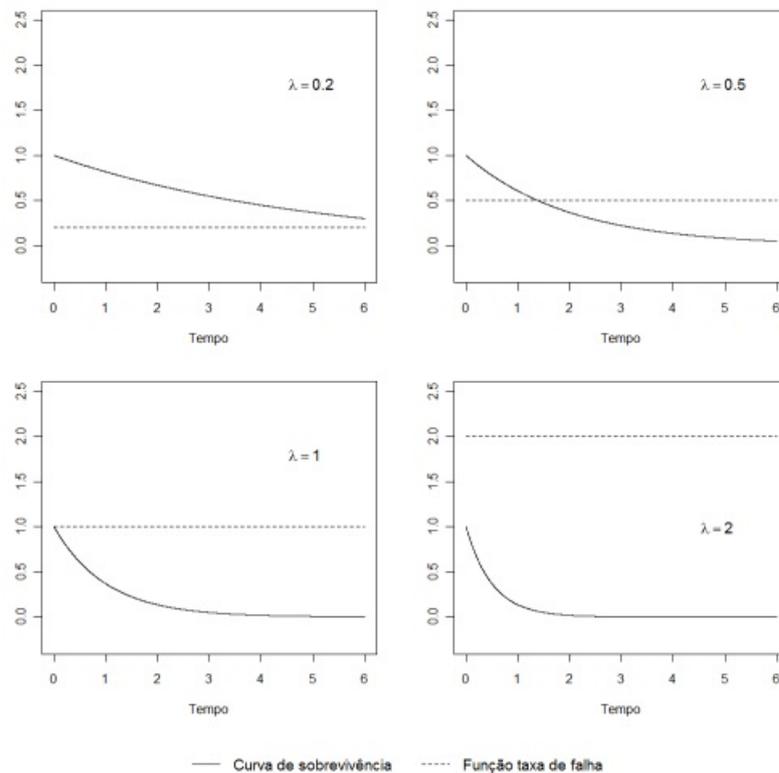
e a função de taxa de falha por:

$$h(t) = \lambda, \quad t \geq 0 \quad e \quad \lambda > 0. \quad (2.31)$$

Uma característica fundamental do modelo exponencial é que ele apresenta a função de taxa de falha constante ao longo do tempo, ou seja, o risco de falha não muda com o tempo, é sempre o mesmo para qualquer tempo t . O valor da função de risco é igual ao valor do parâmetro da distribuição. Essa propriedade é conhecida como falta de memória da distribuição exponencial (HERRMANN, 2011). Se o modelo exponencial for adequado para analisar os dados, sabe-se automaticamente, que o risco de falha para as unidades que estão em estudo há pouco tempo ou há muito tempo, dado que ainda não tenham falhado é o mesmo (LEE; WANG, 2003).

Na Figura 3 é perceptível a influência dos valores do parâmetro λ na função de sobrevivência e na taxa de falha. Em que, para valores grandes de λ o risco será alto e a confiabilidade baixa, para valores pequenos de λ o risco será baixo e a confiabilidade alta (LEE; WANG, 2003) apud Herrmann (2011).

Figura 3 – Função sobrevivência e função taxa de falha do modelo exponencial para diferentes valores de λ .



Fonte: Herrmann (2011, p. 38)

2.5.2 Modelo Weibull

O modelo Weibull é indicado para situações em que o evento de interesse (tempo até falha) é bem definido por meio da distribuição de probabilidade Weibull, pois ele possui uma grande variedade de formas, todas com propriedades básicas: função de taxa de falha é monótona, isto é, crescente, decrescente ou constante e, por isso, consegue se adaptar a diversas situações práticas (HERRMANN, 2011).

A função densidade de probabilidade do modelo Weibull para variável aleatória tempo até falha T é dada pela expressão:

$$f(t) = \lambda s t^{s-1} \exp\{-\lambda t^s\} \quad t \geq 0 \quad e \quad s, \lambda > 0, \quad (2.32)$$

e sua função de confiabilidade por:

$$R(t) = \exp\{-\lambda t^s\} \quad t \geq 0 \quad e \quad s, \lambda > 0, \quad (2.33)$$

e a função taxa de falha por:

$$h(t) = \lambda s t^{s-1} \quad t \geq 0 \quad e \quad s, \lambda > 0 \quad (2.34)$$

O parâmetro s determina a forma da função taxa de falha da seguinte maneira:

- i) $s < 1 \rightarrow$ a função da taxa de falha decresce,
- ii) $s > 1 \rightarrow$ a função da taxa de falha cresce,
- iii) $s = 1 \rightarrow$ a função da taxa de falha é constante,

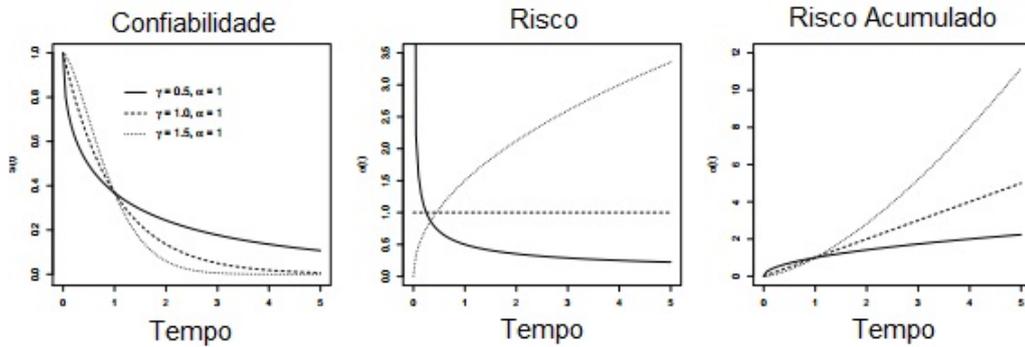
Note que quando $s = 1$, obtêm-se a distribuição Exponencial como caso particular da distribuição Weibull. Algumas formas das funções de densidade de confiabilidade, de risco e de risco acumulado de uma variável T com distribuição de Weibull são apresentadas Figura 4 em que observa-se a flexibilidade da distribuição Weibull, quando diferentes valores do parâmetro s são utilizados.

Para encontrar o tempo mediano de confiabilidade, de acordo com a distribuição Weibull, iguala-se a probabilidade de confiabilidade a 0,5:

$$\begin{aligned} R(t) &= \exp\{-\lambda t^s\} = 0,5 \\ -(\lambda t)^s &= \ln(1/2)^{1/s} \\ \lambda t &= \ln(2)^{1/s} \\ T_{\text{mediano}} &= \frac{\ln(2)^{1/s}}{\lambda} \end{aligned} \quad (2.35)$$

Quando o tempo até a falha segue uma distribuição Weibull, o logaritmo do tempo até falha segue uma distribuição Gumbel, também chamada de distribuição do valor

Figura 4 – Função de confiabilidade, de risco e de risco acumulado com parâmetro escala $\lambda = 1$ e diferentes valores do parâmetro de forma s .



Fonte: [Carvalho \(2011, p.142\)](#)

extremo. Ou seja, se T segue distribuição Weibull, $Y = \log(T)$ segue distribuição do valor extremo.

A função densidade de probabilidade, a função de confiabilidade e a função taxa de falha de Y são dadas por:

$$f(y) = \frac{1}{\sigma} \exp \left\{ \left(\frac{y - \mu}{\sigma} \right) - \exp \left\{ \frac{y - \mu}{\sigma} \right\} \right\} \quad y, \mu \in \mathbb{R}, \sigma > 0 \quad (2.36)$$

$$R(y) = \exp \left\{ -\exp \left\{ \frac{y - \mu}{\sigma} \right\} \right\} \quad y, \mu \in \mathbb{R}, \sigma > 0 \quad (2.37)$$

$$h(y) = \frac{1}{\sigma} \exp \left\{ \frac{y - \mu}{\sigma} \right\} \quad y, \mu \in \mathbb{R}, \sigma > 0 \quad (2.38)$$

Segundo [Herrmann \(2011\)](#), as relações entre as distribuições Weibull e Gumbel também pode ser expressa por uma relação entre os parâmetros dessas duas distribuições, para tanto, tem-se que $s = 1/\sigma$ e $\lambda = (1/\exp\{\mu\})^s$. Em que Y e $\mu \in \mathbb{R}$ e $\sigma > 0$. Se $\mu = 0$ e $\sigma = 1$, tem-se a distribuição do valor extremo padrão. Os parâmetros μ e σ são denominados parâmetros de locação e escala, respectivamente.

2.5.3 Modelo Log-Normal

[Herrmann \(2011\)](#), afirma que o modelo Log-Normal se adequa muito bem em estudos de confiabilidade em que o tempo até a falha é bem definido através de uma distribuição de probabilidades Log-Normal. Uma característica valiosa para o modelo Log-Normal é que, o logaritmo de uma variável com distribuição Log-Normal com parâmetros μ e σ , tem distribuição normal com média μ e desvio padrão σ .

A distribuição Log-Normal é muito utilizada para modelagem de tempos até reparos em unidades reparáveis. Nesse caso, a função de densidade para a variável aleatória tempo

até a falha T é dada por:

$$f(t) = \frac{1}{t\sigma\sqrt{2\pi}} \exp \left\{ -\frac{1}{2} \left(\frac{\log(t) - \mu}{\sigma} \right)^2 \right\}, \quad t > 0, \mu \in \mathbb{R}, \sigma > 0, \quad (2.39)$$

em que, μ , é a média do logaritmo do tempo até falha, assim como σ é o desvio padrão (COLOSIMO; GIOLO, 2006).

A função de confiabilidade do modelo Log-Normal é dada por:

$$R(t) = \phi \left(\frac{-\log(t) + \mu}{\sigma} \right), \quad t > 0, \mu \in \mathbb{R}, \sigma > 0, \quad (2.40)$$

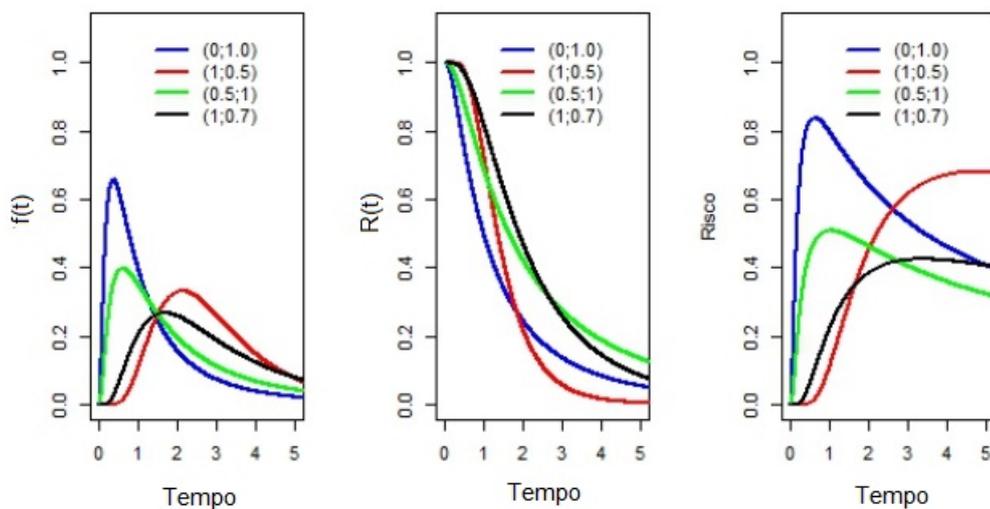
e a função taxa de falha por:

$$h(t) = \frac{f(t)}{R(t)}, \quad t > 0, \mu \in \mathbb{R}, \sigma > 0, \quad (2.41)$$

em que $\phi(\cdot)$, é a função de distribuição acumulada de uma normal padrão (LEE; WANG, 2003).

As funções de confiabilidade e taxa de falha de uma variável Log-Normal não possuem uma forma analítica explícita. A característica da função taxa de falha do modelo paramétrico Log-Normal é ser inicialmente crescente e, quando atingir o ponto de máximo passar a decrescer, ou seja, preserva a relação de ordem, portanto, não é uma função monótona como pode ser observado na Figura 5.

Figura 5 – Funções de densidade de probabilidade, de confiabilidade e de risco da distribuição Log-Normal para diferentes valores dos parâmetros $(\mu; \sigma)$.



Na Figura 5 também pode ser observado a influência dos parâmetros μ e σ . Em que, valores grandes para μ produzem curvas de confiabilidade com confiabilidade maior que valores pequenos para μ . Logo, quanto maior for o parâmetro μ maior será o tempo

de confiabilidade do item ou unidade em estudo. Por consequência, quanto menor for o valor de μ , mais ressaltado será o pico da função taxa de falha. O parâmetro σ influencia na oscilação das curvas, ou seja, curvas de confiabilidade que tem valores mais altos para σ terão maior probabilidade de confiabilidade para tempos maiores do que curvas com valores baixo para σ (HERRMANN, 2011).

2.5.4 Distribuição Gama e Gama Generalizada

A Distribuição Gama é bastante utilizada para descrever dados relacionados com tempo, porém isso é graças aos pioneiros Brown e Flood (1947) que exploraram essa distribuição para explicar o tempo de vida de copos de vidro em uma cafeteria. Na área medica, o uso dessa distribuição é recente, principalmente quando existem fatores aleatórios. A função de densidade da distribuição Gama com π o parâmetro de forma e δ o de escala, para $\pi, \delta > 0$ é expresso, por

$$f(t) = \frac{1}{\Gamma(\pi)\delta^\pi} t^{\pi-1} \exp\left\{-\left(\frac{t}{\delta}\right)\right\}, t > 0$$

e sua função de sobrevivência $S(t)$ expressada da seguinte maneira:

$$R(t) = \int_t^\infty \frac{1}{\Gamma(\pi)\delta^\pi} v^{\pi-1} \exp\left\{-\left(\frac{v}{\delta}\right)\right\} dv$$

e acrescentando mais um parâmetro γ de forma, essa distribuição se torna mais flexível para mais informações ver Stacy (1962).

2.6 Métodos de estimação dos parâmetros

De acordo com Colosimo e Giolo (2006) os parâmetros são quantidades desconhecidas que descrevem os modelos de probabilidade e que precisam ser estimados a partir das observações amostrais. O método de máxima verossimilhança é uma opção apropriada para dados censurados, incorporando as censuras é relativamente simples, e possui propriedades para grandes amostras.

2.6.1 Método da máxima verossimilhança

De acordo com Carvalho (2011), a função de verossimilhança é uma medida de concordância entre os dados e os possíveis valores do parâmetro pela qual se avalia o quanto os dados apoiam, concordam ou suportam cada valor possível do parâmetro a ser estimado. Ainda por Carvalho (2011), o método de estimação da máxima verossimilhança (MMV) é exatamente o que o nome diz, ou seja, considera que a melhor estimativa possível para o parâmetro populacional é o valor que maximiza a função de verossimilhança obtida a partir dos dados amostrais.

2.6.1.1 Função de verossimilhança para dados censurados

Segundo Fogliatto e Ribeiro (2009, p.55), para um conjunto de dados censurados a função de verossimilhança pode ser escrita de várias formas, uma das quais é apresentada a seguir. Sejam T_1, \dots, T_n variáveis aleatórias que seguem uma distribuição de probabilidade $f(t, \theta)$, em que θ é um parâmetro desconhecido. Em alguns casos, $f(t)$ pode ser caracterizado por um vetor de parâmetros θ . A seguir temos, os r tempos até a falha observados são designados, em ordem crescente de ocorrência, por t_1, \dots, t_r ; os dados censurados são designados por t_1^+, \dots, t_{n-r}^+ . No caso de censura à direita do tipo I, $t_1^+ = \dots = t_{n-r}^+ = t_0$. Um total de n unidades foram colocadas em estudo.

Para um conjunto completo de dados (ou seja, sem censura, tal que $r = n$). A função de verossimilhança associada a amostra é dada por:

$$L(\theta) = \prod_{i=1}^n f(t_i, \theta). \quad (2.42)$$

A função na Equação (2.42) informa sobre a verossimilhança (possibilidade) de as variáveis T_1, \dots, T_n assumirem os valores t_1, \dots, t_n ; tal possibilidade é dada pelo valor da função de densidade. Para o caso de variáveis discretas, a verossimilhança corresponde a um valor de probabilidade.

A expressão na Equação (2.42) é função apenas do parâmetro desconhecido θ . O estimador de máxima verossimilhança de θ corresponde ao valor de θ que maximiza $L(\theta)$; tal valor é obtido derivando a Equação (2.42) com relação a θ , igualando o resultado a 0 e isolando θ . Pode-se demonstrar que $L(\theta)$ e $l(\theta) = \ln[L(\theta)]$ apresentam seus máximos no mesmo valor de θ . Nesse trabalho será adotado que o estimador de parâmetro θ é designado por $\hat{\Theta}$ e suas estimativas por $\hat{\theta}$.

No caso de uma amostra contendo dados censurados, a Equação (2.42) é assim reescrita para incluir a porção censurada da amostra:

$$L(\theta) = \prod_{i=1}^r f(t_i, \theta) \prod_{i=1}^{n-r} R(t_i^+, \theta) \quad (2.43)$$

em que, $R(t_i^+, \theta)$ é a função de confiabilidade com parâmetro θ avaliado no tempo censurado t_i^+ . Observa-se que a função de confiabilidade $R(t_i^+, \theta)$ utilizada na expressão da função de verossimilhança para uma observação censurada à direita corresponde à probabilidade de a unidade i sobreviver a um tempo t_i^+ .

O logaritmo da função na Equação (2.43) é dado por:

$$l(\theta) = \sum_{i=1}^r \ln f(t_i, \theta) + \sum_{i=1}^{n-r} \ln R(t_i^+, \theta). \quad (2.44)$$

2.7 Teste de Anderson-Darling

Segundo [Ferreira, Souza e Silva \(2013\)](#), o teste de Anderson-Darling é um teste estatístico utilizado para verificar se um conjunto de dados é proveniente de uma determinada distribuição de probabilidade. De acordo com [Morais, Ferreira e Balestrassi \(2005\)](#), no teste de Anderson-Darling considera-se normal a distribuição que apresentar p-valor maior que 0,05, o que significaria uma probabilidade maior que 5% em cometer erro, ao rejeitar a hipótese de normalidade da distribuição em análise.

As duas hipóteses para o teste de Anderson-Darling para uma distribuição normal são:

H_0 : os dados seguem uma distribuição de probabilidade específica.

H_1 : os dados não seguem uma distribuição de probabilidade específica.

A hipótese nula é a que os dados seguem uma determinada distribuição e hipótese alternativa é a que os dados não seguem uma distribuição específica. Se o nível de significância (p-valor) for pequeno ($p \leq 0,05$), a hipótese nula é rejeitada e conclui-se que os dados analisados não seguem uma determinada distribuição. A estatística teste de Anderson-Darling é expressa por:

$$AD = -n - \frac{1}{n} \sum_{i=1}^n (2i - 1) [\ln F(X_i) + \ln(1 - F(X_{n-i+1}))], \quad (2.45)$$

em que n é tamanho da amostra, F é a função de distribuição acumulada para a distribuição específica e i é a i -ésima amostra quando os dados estão ordenados em ordem crescente.

Para pequenas amostras ($n \leq 200$), o valor de AD deve ser ajustado para:

$$AD^* = AD \left(1 + \frac{0,75}{n} + \frac{2,25}{n^2} \right) \quad (2.46)$$

2.8 Escolha do modelo paramétrico

A escolha do modelo paramétrico adequado para analisar um conjunto de dados de sobrevivência é uma etapa muito importante da análise. Se o modelo utilizado não é adequado aos dados os resultados podem não retratar corretamente a realidade, invalidando toda análise realizada ([HERRMANN, 2011](#)).

Existem métodos para auxiliar na escolha do modelo paramétrico que melhor se ajusta aos dados, ou seja, indicam qual a distribuição probabilística se adequa melhor as características específicas dos dados. Entre estes métodos estão alguns métodos gráficos, o teste da razão de verossimilhança e os critérios de informação.

2.8.1 Teste da razão de verossimilhanças (TRV)

Segundo [Carvalho \(2011\)](#), o teste da razão de verossimilhança é um teste de hipóteses, adequado para comparar modelos aninhados, isto é, o modelo com maior número de parâmetros estimados (modelo completo) tem que conter todos os parâmetros do modelo menor (modelo reduzido), para determinar qual explica melhor o comportamento dos dados. Baseado na função de verossimilhança, esse teste envolve a comparação dos valores do logaritmo da função de verossimilhança maximizada sem restrição e sob H_0 . De acordo com [Lehmann \(1993\)](#), o TRV é baseado no log da razão entre as duas verossimilhanças, ou seja, na diferença entre o $\log L(\tilde{\theta})$ e $\log L(\hat{\theta})$. A estatística é dada por

$$TRV = -2[\log L(\tilde{\theta}) - L(\hat{\theta})] \quad (2.47)$$

sendo $\log L(\hat{\theta})$ igual ao logaritmo da função de verossimilhança do modelo reduzido e $\log L(\tilde{\theta})$ o do modelo completo. Essa estatística segue assintoticamente uma distribuição χ^2 com número de graus de liberdade igual à diferença no número de covariáveis dos modelos ([CARVALHO, 2011](#)).

O modelo proposto não será adequado quando rejeita-se H_0 , ou seja, quando o p – valor for menor que o nível de significância estabelecido, geralmente 5%, geralmente se p – valor $> 0,05$, então não há evidências estatísticas de que o modelo proposto não seja adequado ([HERRMANN, 2011](#)).

2.8.2 Análise gráfica do ajuste do modelo

De acordo com [Carvalho \(2011\)](#), o método gráfico é uma técnica auxiliar na escolha da família de distribuições paramétricas. É portanto uma ferramenta exploratória para escolha de possíveis modelos. [Colosimo e Giolo \(2006\)](#), sugerem dois métodos gráficos para auxiliar na escolha dos modelos candidatos. Segundo [Herrmann \(2011\)](#), um desses métodos gráficos baseia-se na ideia de linearização da função de confiabilidade. De modo que, se o modelo paramétrico proposto for adequado o resultado será uma reta. Alterações na linearidade são facilmente observadas, indicando que o modelo paramétrico sugerido não se ajusta adequadamente aos dados. No caso do modelo paramétrico de Weibull, espera-se encontrar retas paralelas, e no caso do modelo paramétrico exponencial espera-se que a inclinação das retas seja igual a um. Pode-se também fazer um gráfico dos valores estimados de confiabilidade via Kaplan-Meier versus os valores estimados de confiabilidade pelas distribuições paramétricas. Uma reta traçada no ângulo de 45 graus indica um bom ajuste entre os modelos.

Outra análise gráfica útil consiste em comparar as curvas de sobrevivência geradas pelo estimador não paramétrico de Kaplan-Meier com as estimadas parametricamente. Quanto mais próximo o modelo paramétrico estiver da curva de sobrevivência do estimador de Kaplan-Meier, melhor o ajuste do modelo. Os métodos gráficos apresentados nesta

seção, no entanto devem ser usados para descartar modelos claramente inapropriados, e não para demonstrar que um modelo paramétrico é melhor, pois para isso a análise de resíduos seria mais apropriada (CARVALHO, 2011).

2.8.3 Medida global de qualidade do ajuste

Segundo Carvalho (2011, p.162), para avaliar a qualidade do ajuste de um modelo paramétrico, pode-se utilizar uma medida global de qualidade do ajuste denominada **função desvio** (ou *deviance*). Em que, a hipótese nula H_0 de que o modelo se ajusta aos dados pode ser testada pela estatística:

$$D = 2 \left[l(\tilde{\theta})_{\text{completo}} - l(\hat{\theta})_{\text{reduzido}} \right], \quad (2.48)$$

em que $l(\hat{\theta})_{\text{reduzido}}$ e $l(\tilde{\theta})_{\text{completo}}$ são, o logaritmo da função de verossimilhança do modelo avaliada nas estimativas de máxima verossimilhança e o logaritmo da função de verossimilhança do modelo completo.

Um modelo completo é aquele que o número de parâmetros ajustados é exatamente igual ao número de observações. Para calcular a *deviance*, os parâmetros de perturbação são mantidos fixos em ambos modelos. Por parâmetros de perturbação entende-se, aqueles comuns aos dois ajustes, e nos quais não há interesse no estimador. Quanto maior for a *deviance* do modelo, pior a qualidade do ajuste.

A estatística D segue assintoticamente, uma distribuição qui-quadrado (χ^2) com $n - p - 1$ graus de liberdade, ou seja, o número de observações menos o número de covariáveis do modelo menos um. Nesse caso, quanto menor a estatística, maior o *p-valor* e melhor a qualidade de ajuste do modelo (CARVALHO, 2011).

2.8.4 Análise de resíduos

De acordo com Carvalho (2011), em estudos de confiabilidade existem alguns resíduos que são adequados para avaliação global de modelos paramétricos. Técnicas gráficas, que fazem uso dos diferentes resíduos propostos, são, em particular, bastante utilizadas para examinar diferentes aspectos do modelo. Em todos os casos, o objetivo é identificar observações que são pontos influentes no modelo, e seu impacto sobre.

2.8.4.1 Resíduo de Cox-Snell

De acordo com Colosimo e Giolo (2006), os resíduos de Cox-Snell é uma medida útil para examinar o ajuste global do modelo. É definido por:

$$\hat{e}_i = \hat{H}(t_i|x_i), \quad (2.49)$$

em que para a i -ésima unidade, se tem que t_i é o instante, x_i é o vetor de covariáveis e \hat{H} é a função de risco acumulado estimado para o modelo. Para os modelos vistos anteriormente os resíduos de CoxSnell são dados por:

- i) Exponencial: $\hat{e}_i = [t_i \exp\{-x_i' \hat{\beta}\}]$,
- ii) Weibull: $\hat{e}_i = [t_i \exp\{-x_i' \hat{\beta}\}]^{\hat{\gamma}}$,
- iii) Log-Normal: $\hat{e}_i = -\log \left[1 - \phi \left(\frac{\log t_i - x_i' \hat{\beta}}{\sigma} \right) \right]$.

Se o modelo for adequado e as estimativas dos parâmetros estiverem próximas dos verdadeiros valores, estes resíduos devem parecer como uma amostra censurada de uma distribuição exponencial padrão ($\lambda = 1$).

O gráfico das curvas de sobrevivência desses resíduos, obtidas por Kaplan-Meier e pelo modelo ajustado, também auxilia na verificação da qualidade do modelo. Quanto mais próximas elas se apresentarem, melhor é considerado o ajuste do modelo aos dados.

2.8.4.2 Resíduos Martingale

Segundo [Velho \(2015\)](#), os resíduos martingale m_i , para modelos paramétricos onde todas as covariáveis são fixas, são baseados no processo de contagem individual e definidos por:

$$\hat{m}_i = \hat{\delta}_i - \hat{e}_i, \quad (2.50)$$

em que $\hat{\delta}_i$ é a variável indicadora de censura e \hat{e}_i os resíduos de Cox-Snell. Esses resíduos, são vistos como uma estimativa do número de falhas em excesso observada nos dados mas não previsto pelo modelo.

São usados, em geral, para examinar a melhor forma funcional (linear, quadrática, etc.) para uma dada covariável em um modelo de regressão assumido para os dados sob estudo. Se a curva suavizada obtida no gráfico: resíduo \times variável for linear, nenhuma transformação na variável é necessária ([COLOSIMO; GIOLO, 2006](#)).

2.8.4.3 Resíduos score: pontos influentes

Os resíduos score, também conhecidos como *dfbeta*, são adequados para analisar a influência de cada observação no ajustamento do modelo e para estimar a variância dos coeficientes de regressão. Este resíduo, é definido pela diferença entre o vetor de parâmetros estimado pelo modelo e o mesmo estimado sem a unidade i . Assim o resíduo score é dado por:

$$\Delta\beta = \hat{\beta} - \hat{\beta}_{(-i)}. \quad (2.51)$$

Se essa diferença for igual a zero, então a i -ésima observação tem pouca influência na estimativa. Esse resíduo representa a mudança aproximada que o vetor de parâmetros β sofreria se a observação fosse eliminada; quanto maior o valor de $\Delta\beta_k$, maior a influência da observação (CARVALHO, 2011).

2.8.5 Critérios de Informação

Na validação do desempenho dos modelos, o Critério de informação de Akaike - Akaike's Information Criterion (AIC) tem sido bastante utilizado para diferentes estruturas de covariâncias. Akaike (1973), relacionou a discrepância, medida que existe entre o modelo verdadeiro e o modelo aproximado, com a máxima verossimilhança, que é o que possui melhor ajuste para os dados observados.

Segundo Akaike (1973), o princípio do Critério de informação de Akaike é selecionar uma combinação de variáveis exploratórias a modelos para a função de correlação que minimize o valor do AIC. Ou seja, quanto menor o valor de AIC melhor será o ajuste do modelo.

De acordo com Velho (2015), este critério irá penalizar a log verossimilhança de cada modelo, de forma a refletir o número de parâmetros que estão a ser estimados, permitindo a comparação de modelos.

A estimativa do AIC para um determinado modelo é dada por:

$$AIC = -2l(\hat{\theta}) + 2k \quad (2.52)$$

Em que, $l(\hat{\theta})$ é o logaritmo da função verossimilhança do modelo e k é número de parâmetros do modelo .

3 Material e Métodos

Neste trabalho foram estudados 552 pneus com especificação 11.00R22 de caminhões transportadores de safras, dos quais foram avaliados em três vidas úteis. O estudo acompanhou a durabilidade do pneu até ocorrer interrupção para troca ou recauchutagem do mesmo. O tempo de uso foi medido por meio de computador de bordo acoplado nos caminhões, sendo este tempo de uso medido em quilômetros percorridos. Os pneus podiam voltar ao estudo para uma segunda ou terceira vida útil, até que ocorresse novamente a falha ou censura (até que o mesmo não tivesse características adequadas para a rodagem).

Utilizou-se as funções de densidade, sobrevivência e risco, em que, a função de densidade de probabilidade é caracterizada pelo evento de interesse ao observar um indivíduo no intervalo de tempo $[t, t+\Delta t]$. $f(t)$; Enquanto a função de sobrevivência, denotada por $R(t)$, é definida como a probabilidade de um indivíduo sobreviver até um certo tempo t , sem experimentar o evento. E a função de risco, ou taxa de falha que descreve a forma com que a taxa de falha muda com o tempo, $h(t)$.

De acordo com o estado de conservação, os pneus foram classificados em três grupos: G_1 : pneus novos (1ª vida útil), G_2 : pneus que falharam e foram consertados (2ª vida útil) e G_3 : pneus que falharam pela segunda vez (3ª vida útil). Foi utilizado o estimador de Kaplan e Meier (produto-limite). Este estimador permite realizar testes de hipóteses que não requerem pressupostos sobre a forma da distribuição subjativa dos dados. Após a aplicação do estimador produto-limite verificou-se se os tempos de sobrevivência dos pneus em cada grupo de vida, eram similares ou se apresentavam diferença estatística quanto a sua duração, por meio do uso do teste de log-rank.

Após o uso das técnicas não paramétricas fez-se o ajuste modelos paramétricos, normal, log-normal, weibull e exponencial, obtendo-se suas funções de densidade, sobrevivência e risco. Após este ajuste comparou-se por meio de teste de aderência de Anderson Darling, teste de razão de verossimilhança (comparando a distribuição gama generalizada com as demais), critério de AIC e por meio de análises gráficas se as distribuições estavam bem ajustadas. Foram realizadas análises de resíduos (resíduos padronizados, Cox-Snell e deviance), para melhor compreensão do ajuste das distribuições aos dados. Todas as análises foram feitas utilizando o software R 3.2.2 (TEAM, 2015), por meio do pacote *survival* versão 2.36 (THERNEAU; LUMLEY, 2014) e o software *ActionEstat*.

4 Resultados e Discussão

A seguir apresenta-se os principais resultados obtidos a partir da análise realizada com os dados dos pneus 11.00R22. Primeiramente foram obtidas as estimativas de Sobrevivência utilizando-se o estimador de Kaplan-Meier aos quais são apresentadas na Tabela 1 e no gráfico da Figura 6.

Tabela 1 – Distribuição das frequências dos eventos de falha e tempo mediano de acordo com o grupo de vida dos pneus 11.00R22.

	n	Evento\Falha	Tempo mediano
Vida 1	552	449	25953
Vida 2	416	195	36781
Vida 3	138	5	NA
Geral	1106	649	32229

Como pode ser observado na Tabela 1, das 1106 observações realizadas com os pneus avaliados na para rodagem, 649 (58,67 %) deles obtiveram alguma falha, as demais censuraram em algum período. É importante notar que na primeira vida útil se tinha 552 pneus, estes perfazem o total de pneumáticos avaliados. A segunda vida útil são os pneus que não censuraram até este momento, assim como na terceira vida útil, deste modo na terceira vida útil como era de se esperar apenas 138 pneus estava aptos dos quais deste total 5 experimentaram a falha, os demais censuraram com pouca quilometragem. O tempo mediano das vidas úteis foi por volta de 32000 quilômetros.

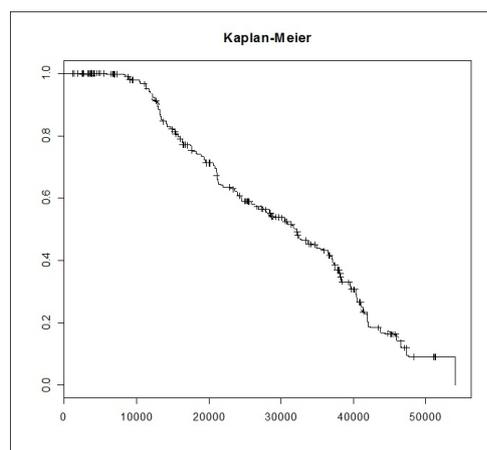


Figura 6 – Gráfico do ajuste de Kaplan-Meier para os Pneus 11.00R22.

A estimativa de Kaplan-Meier para a função de sobrevivência ($\hat{R}(t)$) está representada na Figura 6. Sendo esta uma função escada, com valor constante para cada intervalo de tempo. Por definição, a função de sobrevivência estimada no primeiro intervalo $[0, t_1)$

é igual a 1. Por outro lado, a função de probabilidade estimada para o último intervalo, $[t_k, \infty)$, é zero, se o maior tempo observado for uma falha, e não atingirá o zero se for uma censura. Neste gráfico pode-se observar a ocorrência de todas as falhas e censuras ocorridas no estudo, ocorrendo a presença de uma censura no último intervalo observado. Após aplicação do estimador de Kaplan-Meier aos dados, procedeu-se a análise com o teste de log-rank. O teste log-rank foi aplicado com o interesse de verificar estatisticamente se existe diferença entre os três tipos de vida dos pneus.

Tabela 2 – Teste de log-rank aplicados aos grupos analisados.

Log-Rank	N	Informação	Valor numérico
grupos=1	552	Qui-quadrado	73,84
grupos=2	416	p-valor	0
Log-Rank	N	Informação	Valor numérico
grupos=2	416	Qui-quadrado	34,79
grupos=3	138	p-valor	$3,66 \times 10^{-09}$
Log-Rank	N	Informação	Valor numérico
grupos=1	552	Qui-quadrado	72,73
grupos=3	138	p-valor	0

De acordo com a Tabela 2, obser-se que houve diferença significativa entre os grupos 1, 2 e 3, indicando que existe diferença no tempo de vida entre os pneus desses grupos. Pneus da primeira vida útil apresentaram tempo média de falha maior que estes mesmos pneus na segunda e/ou terceira vida útil.

O teste de Anderson - Darling é utilizado para testar a hipótese de que uma dada amostra tenha sido retirada de uma dada população com função de distribuição acumulada contínua $F(x)$, pode-se verificar a adequabilidade a uma determinada distribuição.

Tabela 3 – Teste de Anderson - Darling para as diferentes distribuições analisadas

Distribuição	Anderson-Darling	p-Valor
Normal	13,87	$4,86 \times 10^{-33}$
Exponencial	71,34	0,005
Weibull	8,28	0,01
Log.Normal	25,90	$5,10 \times 10^{-59}$

De acordo com a Tabela 3, o teste de Anderson - Darling, com um p-Valor = 0,01, indica que a distribuição Weibull é a que melhor se adequa aos dados do estudo, sendo portanto a principal candidata para modelar os dados em estudo.

A Figura 7 corrobora esse fato, pois através da representação gráfica observa-se que a distribuição Weibull melhor se ajustou a reta, com maior valor de correlação com o Kaplan Meier e a que se comportou mais próximo de uma reta no gráfico Papel de probabilidade.

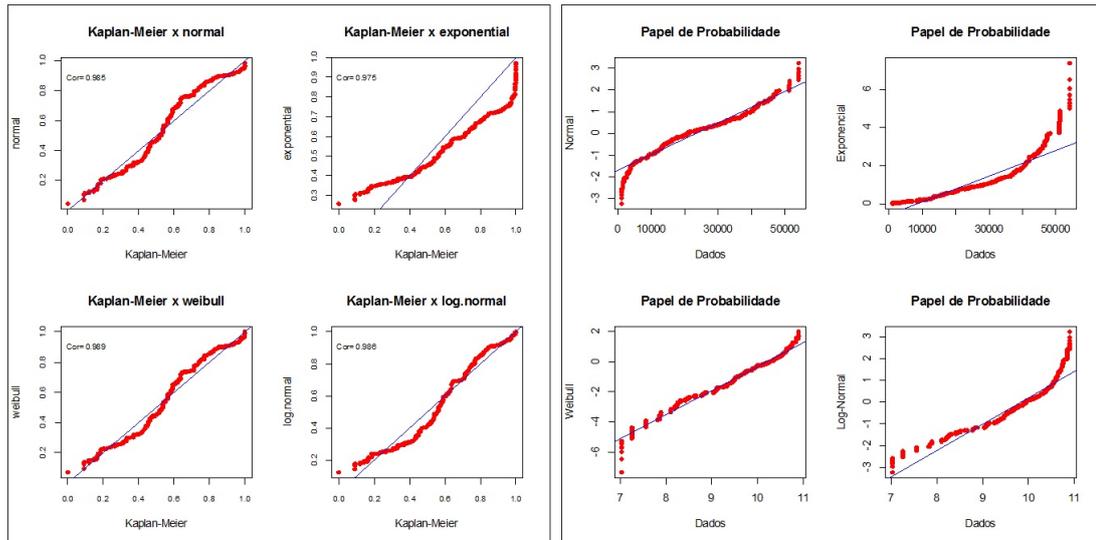


Figura 7 – Gráfico do ajuste de Kaplan-Meier versus as distribuições Normal, Exponencial, Weibull e Lognormal e Representação gráfica dos Papeis de probabilidade para as mesmas distribuições.

Na Tabela 4 tem-se as estimativas dos parâmetros das distribuições Weibull e Lognormal, em que o parâmetro de forma determina a forma da função da distribuição. O parâmetro de forma define como os dados são distribuídos, mas não afeta a escala da distribuição. Maiores valores de forma darão curvas assimétricas para a esquerda, enquanto que valores de forma menores dão uma curva assimétrica para direita. Apesar de necessitar de outros valores para se comparar com estas distribuições, pode-se afirmar que os valores estimados de forma condizem com distribuições aproximadamente simétricas e que estes valores foram bem estimados devido a magnitude de seus desvios padrão serem mais de 10 vezes menor que o valor estimado para o parâmetro de forma em ambas distribuições e os intervalos de confiança serem estreitos.

Tabela 4 – Estimativas dos parâmetros das distribuições Weibull e Lognormal.

Distribuições	Parâmetros	Estimativas	Desvio Padrão	Lim. Inf.	Lim. Sup.
Weibull	Forma	2,30	0,07	2,16	2,44
	Escala	35231,23	$4,70 \times 10^{-08}$	35231,22	35231,23
Lognormal	Forma	10,24	0,02	10,20	10,28
	Escala	0,56	0,01	0,53	0,59

O parâmetro de escala de uma distribuição determina a escala da função de distribuição. Um valor maior para a estimativa deste parâmetros amplia a sua distribuição deixando-a leptocúrtica, o que caracteriza grande concentração dos dados.

Tabela 5 – Estimativas dos percentis obtidos por meio das distribuições Weibull e Lognormal com respectivos desvios padrão e intervalos de confiança.

Distribuições	Percentis (%)	Estimativas	Desvio Padrão	Lim. Inf.	Lim. Sup.
Weibull	0,01	4783,15	294,77	4205,40	5360,90
	0,05	9704,89	386,17	8948,01	10461,77
	0,1	13264,53	399,90	12480,73	14048,31
	0,5	30049,23	147,54	29760,04	30338,41
	0,9	50600,92	565,39	49492,79	51709,07
Lognormal	0,01	7585,33	287,32	7022,17	8148,48
	0,05	11134,85	325,28	10497,30	11772,40
	0,1	13663,34	345,11	12986,92	14339,76
	0,5	28122,61	564,67	27015,86	29229,36
	0,9	57883,41	1818,04	54320,11	61446,72

Na Tabela 5 percebe-se que o tempo mediano de vida dos pneus estudados é de 30049,23 Km para distribuição Weibull e de 28122,61, que 10% dos pneus falharam quando estavam por volta de 13264,53 à 13663,34 e que apenas 10% dos pneus continuavam sendo utilizados após 50600,92 Km e 57883,41, para as distribuições Weibull e Lognormal, respectivamente.

Tabela 6 – Estimativas dos quartis, do tempo médio até a falha, da mediana e desvio padrão obtidos pelas distribuições Weibull e Lognormal.

Distribuições	Índices	Valor
Weibull	MTTF	31212,45
	Desvio Padrão	14370,71
	Mediana	30049,23
	1° Quartil	20514,10
	3° Quartil	40597,98
Lognormal	MTTF	32957,18
	Desvio Padrão	20138,28
	Mediana	28122,61
	1° Quartil	19233,53
	3° Quartil	41119,91

Pela Tabela 6, o tempo médio de falha (MTTF), foi registrado em 31212,45 km para a distribuição weibull e 32957,18 para a distribuição lognormal. 25% dos pneus falharam antes de 20514,10 e 19233,53 para as distribuições weibull e lognormal, respectivamente depois de 40597,98 e 41119,91 km restavam 25% dos pneus aptos para uso. Na Figura 8, percebe-se que a distribuição weibull modelada para estes dados apresenta risco crescente, função de densidade levemente assimétrica a direita e o ajuste do modelo weibull em conjunto com o Kaplan-Meier satisfatório, pois a curva de sobrevivência da distribuição modela bem as estimativas obtidas pelo Kaplan-Meier. O mesmo acontece com a distribuição lognormal, com excessão que o risco da distribuição lognormal assume inicialmente risco crescente e posteriormente o risco decresce. O teste de razão de verossimilhança e

o critério AIC foi utilizado para ajudar na escolha da melhor distribuição dos dados. Os resultados apresentados na Tabela 7, indicam a adequação do modelo lognormal pelos critérios de AIC (menor valor) e TRV (p-valor > 0,05). Porém pelas técnicas gráficas apresentadas nas Figuras 9 e 10, a distribuição weibull foi a que melhor se ajustou.

Tabela 7 – Testes de razão de verossimilhanças - TRV e Critério de AIC para as distribuições estudadas.

Distribuições	Loglik	AIC	Parâmetros	TRV	p-Valor
Exponencial	-7526,09	15054,18	1	534,7	0
Weibull	-7267,64	14538,38	2	17,8	$2,45 \times 10^{-4}$
Log-Normal	-7259,70	14523,37	2	1,92	0,16
Normal	-7310,02	14624,04	2	102,56	0
Gama Generalizada	-7258,74	14523,48	3	-	-

Após aplicação dos testes inferenciais aos dados de Pneus 11.00R22, procedeu-se com uma análise de resíduos (Figuras 9 e 10). O que difere as figuras 9 e 10, são o tipo de resíduos utilizados, na primeira figura são utilizados os resíduos padronizados e na segunda os resíduos Cox-Snell. Ambas distribuições candidatas obtidas pela Tabela 7, foram utilizadas para comparação dos resíduos (Weibull e Lognormal).

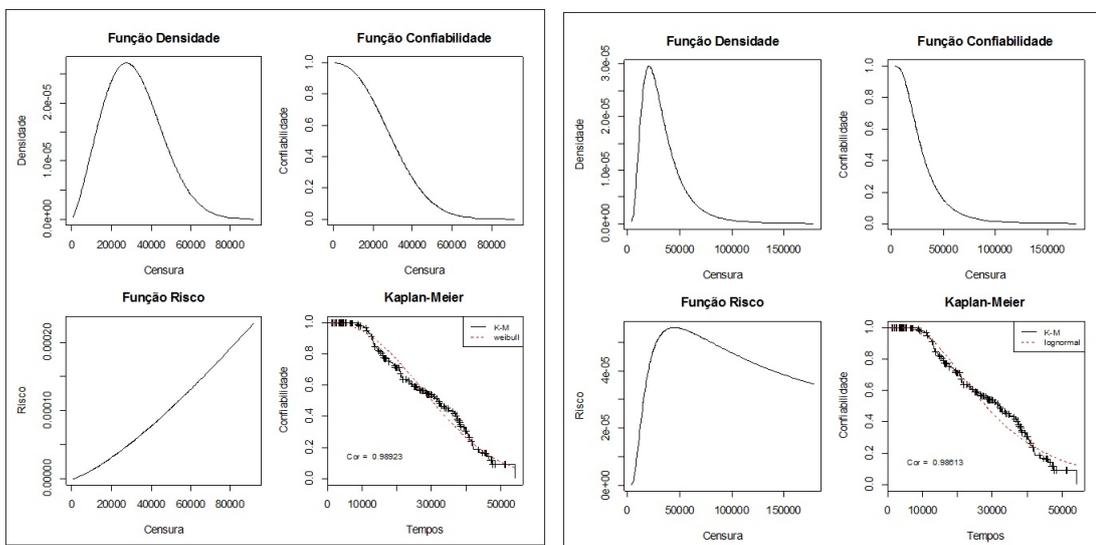


Figura 8 – Gráfico das funções estimadas para o modelo Weibull e Lognormal.

Em ambas as Figuras 9 e 10, pode-se perceber que os resíduos padronizados e coxsnell da distribuição weibull apresentam comportamento mais próximo de uma reta de 45 graus, indicando melhor ajustamento desta distribuição em relação a Lognormal que apresenta-se com uma curvatura maior.

Diferentes critérios tem sido utilizados para mensurar a qualidade do ajuste dos modelos aos dados na análise de confiabilidade (sobrevivência). Eles são todos obtidos por meio de diferentes procedimentos e mostram grande discrepância quando calculados

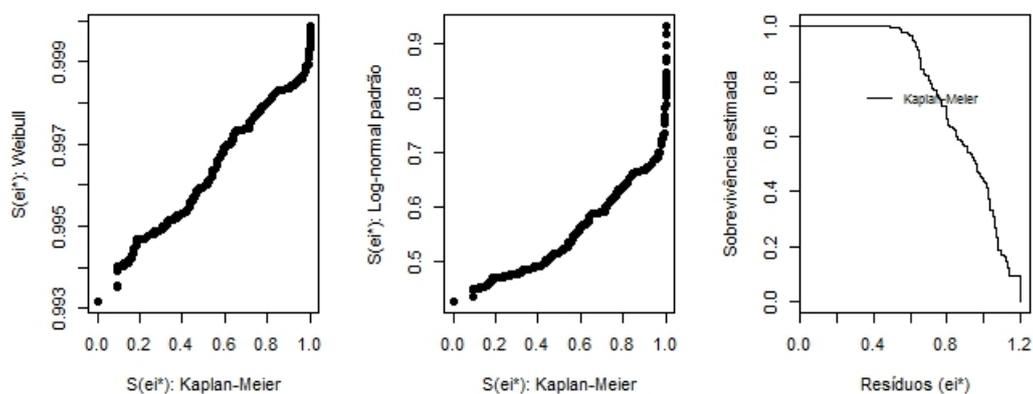


Figura 9 – Gráfico dos resíduos padronizados para as distribuições Weibull, Lognormal e para o estimador de Kaplan-Meier

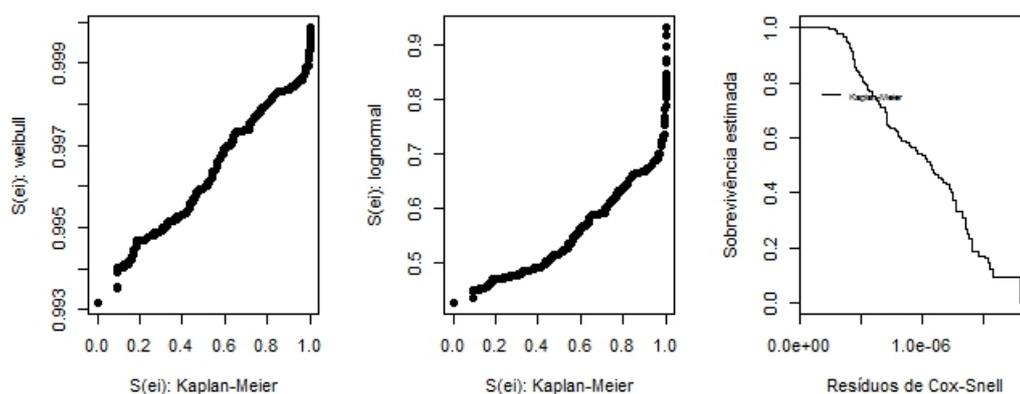


Figura 10 – Gráfico dos resíduos Cox-Snell para as distribuições Weibull, Lognormal e para o estimador de Kaplan-Meier

simultaneamente, além disso dependendo da distribuição escolhida sem te diferentes recomendações a serem dadas quanto ao tempo e estatísticas de interesse para a durabilidade dos Pneus 11.00R22.

5 Considerações finais

Falhas que ocorrem em pneus de veículos de transporte de carga no Brasil são de capital importância nos custos das empresas envolvidas no transporte rodoviário. A distribuição Lognormal e Weibull se apresentaram como fortes candidatas a modelar o tempo de vida de pneus 11.00R22, ajudando a estabelecer o tempo de troca de pneus, baixando assim o custo com tempo perdido e danos ao transporte. Foi apontado que o tempo mediano de troca dos Pneus 11.00R22 é de aproximadamente 30000 quilômetros ao se utilizar a distribuição Weibull para descrever os dados e 28122,61 ao se utilizar a Lognormal. Embora ambas as distribuições Lognormal e weibull tenham se mostrado como fortes candidatas a modelar o tempo de vida dos pneus 11.00R22; a distribuição Weibull apresentou um melhor ajuste através das análises gráficas, como mostram as figuras 9, 10. No entanto os resultados apresentados na Tabela 7, indicam a adequação do modelo Lognormal pelos critérios de AIC (menor valor) e TRV (valor $P > 0, 05$). De acordo com [Carvalho \(2011\)](#), os métodos gráficos devem ser usados para descartar modelos claramente inapropriados, e não para demonstrar que um modelo paramétrico é melhor, pois para isso os critérios de seleção seriam mais apropriados. Dessa forma, concluímos que a distribuição Lognormal é a mais adequada para modelar o tempo de vida de pneus 11.00R22.

Referências

- AKAIKE, H. *Information theory as an extension of the maximum likelihood principle*. [S.l.]: Budapest, Akadémia Kiadó, 1973. Citado na página 36.
- BROWN, G. W.; FLOOD, M. M. Tumbler mortality. *Journal of the American Statistical Association*, Taylor & Francis, v. 42, n. 240, p. 562–574, 1947. Citado na página 30.
- CARVALHO, M. S. *Análise de Sobrevivência: teoria e aplicações em saúde*. [S.l.]: Rio de Janeiro, 2011. Citado 10 vezes nas páginas 22, 23, 24, 25, 28, 30, 33, 34, 36 e 44.
- COLOSIMO, E. A.; GIOLO, S. R. *Análise de Sobrevivência Aplicada*. [S.l.]: São Paulo, 2006. v. 1. Citado 11 vezes nas páginas 15, 17, 19, 20, 21, 24, 29, 30, 33, 34 e 35.
- COSTA, A. L. de Aragão. *Estudo de desgaste de pneus de caminhões e ônibus utilizando-se o Método dos Elementos Finitos*. Dissertação (Mestrado) — Escola Poletécnica, Universidade de São Paulo, 2000. Citado na página 12.
- FERREIRA, A. R.; SOUZA, A. M. M. de; SILVA, D. T. da. Testes estatísticos para diagnóstico da normalidade de uma distribuição de dados. *Congresso Nacional de iniciação científica*, v. 1, 2013. Citado na página 32.
- FOGLIATTO, F. S.; RIBEIRO, J. L. D. *Confiabilidade e Manutenção Industrial*. [S.l.]: Rio de Janeiro, 2009. Citado 8 vezes nas páginas 12, 13, 14, 16, 18, 19, 21 e 31.
- HAVIARAS, G. J. *Metodologia para análise de confiabilidade de pneus radiais em frota de caminhões de longa distância*. Dissertação (Mestrado) — Escola Politécnica da Universidade de São Paulo, 2005. Citado na página 12.
- HERMETO, R. T. *Análise de Sobrevivência na Modelagem do Tempo de Vida de Redes de Sensores sem Fio*. Dissertação (Mestrado) — Universidade Federal do Ceará, 2014. Citado na página 21.
- HERRMANN, L. *Estimação de curvas de sobrevivência para estudos de custo-efetividade*. Dissertação (Mestrado) — Universidade Federal do Rio Grande do Sul, 2011. Citado 10 vezes nas páginas 22, 23, 24, 25, 26, 27, 28, 30, 32 e 33.
- LAWLESS, J. F. *Statistical models and methods for lifetime data*. [S.l.]: John Wiley & Sons, 2011. Citado na página 19.
- LEE, E. T.; WANG, J. W. *Statistical Methods for Survival Data Analysis*. [S.l.]: New Jersey, 2003. Citado 4 vezes nas páginas 23, 25, 26 e 29.
- LEEMIS, L. M. *Reliability: probabilistic models and statistical methods*. [S.l.]: Nova York, 1995. Citado 2 vezes nas páginas 13 e 20.
- MIRANDA, M. S. *Técnicas não-paramétricas e paramétricas usadas na análise de Sobrevivência de Chrysoperla externa (Neuroptera: Crysoptidae)*. Dissertação (Mestrado) — Universidade Federal de Viçosa, 2012. Citado 2 vezes nas páginas 21 e 25.

MORAIS, C. F. de; FERREIRA, J. R.; BALESTRASSI, P. P. Análise crítica da aplicação de métodos estatísticos em processos definidos por dados que não apresentam distribuição normal . *Simpósio de Engenharia de produção*, 12 2005. Anais... Bauru: UNESP. Citado na página 32.

RAMIRES, T. G. *A distribuição beta semi-normal generalizada geométrica*. Dissertação (Mestrado) — Escola Superior de Agricultura Luiz de Queiroz., 2013. Citado na página 17.

RAUSAND, S. E.; HOYLAND, A. *System Reliability Theory: Models, Statistical Methods, and Applications*. [S.l.]: Nova York, 2003. Citado na página 13.

STACY, E. Generalization of the gamma distribution. *Ann. Math. Stat.*, v:33,, p. 1187–1192, 1962. Citado na página 30.

TEAM, R. C. *R: a language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing; 2012*. 2015. Citado 3 vezes nas páginas 6, 7 e 37.

THERNEAU, T.; LUMLEY, T. *Survival. r package version 2.36-12*. 2014. Citado 3 vezes nas páginas 6, 7 e 37.

VELHO, S. D. F. *Modelos de sobrevivência para estudo do tempo até a ocorrência de excesso de peso em indivíduos adultos submetidos a Transplante Alogénico de Células Progenitoras Hematopoiéticas*. Dissertação (Mestrado) — Universidade de Lisboa, 2015. Citado 2 vezes nas páginas 35 e 36.