

**UNIVERSIDADE ESTADUAL PAULISTA - UNESP**  
**FACULDADE DE ENGENHARIA DE BAURU**

**WILLIAM BOSSAS PAULINO**

**DETERMINAÇÃO DA DISTRIBUIÇÃO DE LEI DE POTÊNCIA AOS TEMPOS DE  
REPARO DE UMA LINHA FÉRREA**

**Bauru**

**2012**

**WILLIAM BOSSAS PAULINO**

**DETERMINAÇÃO DA DISTRIBUIÇÃO DE LEI DE POTÊNCIA AOS TEMPOS DE  
REPARO DE UMA LINHA FÉRREA**

Dissertação apresentada ao Programa de  
Pós-Graduação em Engenharia de  
Produção da Universidade Estadual  
Paulista para obtenção do grau de mestre  
em Engenharia de Produção

Orientador: Prof. Dr. Antônio Fernando Crepaldi

Bauru

2012

Paulino, William Bossas.

Determinação da distribuição de Lei de Potência aos  
tempos de reparo de uma linha férrea. / William Bossas  
Paulino, 2012

64 f.

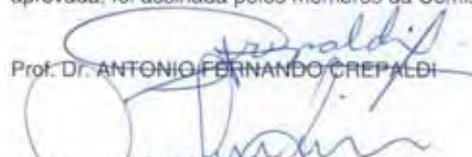
Orientador: Antônio Fernando Crepaldi

Dissertação (Mestrado)-Universidade Estadual Paulista.  
Faculdade de Engenharia, Bauru, 2012

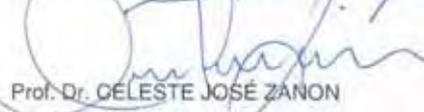
1. Lei de Potência. 2. Ferrovia. 3. Manutenção. I.  
Universidade Estadual Paulista. Faculdade de Engenharia.  
II. Título.

**ATA DA DEFESA PÚBLICA DA DISSERTAÇÃO DE MESTRADO DE WILLIAM BOSSAS PAULINO, DISCENTE DO PROGRAMA DE PÓS-GRADUAÇÃO EM ENGENHARIA DE PRODUÇÃO, DO(A) FACULDADE DE ENGENHARIA DE BAURU.**

Aos 27 dias do mês de agosto do ano de 2012, às 14:00 horas, no(a) Anfiteatro da Pós-graduação da Faculdade de Engenharia de Bauru, reuniu-se a Comissão Examinadora da Defesa Pública, composta pelos seguintes membros: Prof. Dr. ANTONIO FERNANDO CREPALDI do(a) Departamento de Engenharia de Produção / Faculdade de Engenharia de Bauru, Prof. Dr. CELESTE JOSÉ ZANON do(a) Departamento de Engenharia de Produção / Faculdade de Engenharia de Bauru, Prof. Dr. CAMILO RODRIGUES NETO do(a) Escola de Artes, Ciências e Humanidades - EACH/USP, sob a presidência do primeiro, a fim de proceder a arguição pública da DISSERTAÇÃO DE MESTRADO de WILLIAM BOSSAS PAULINO, intitulado "DETERMINAÇÃO DA DISTRIBUIÇÃO DE LEI DE POTÊNCIA AOS TEMPOS DE REPARO DE UMA LINHA FÉRREA". Após a exposição, o discente foi argüido oralmente pelos membros da Comissão Examinadora, tendo recebido o conceito final: APROVADO. Nada mais havendo, foi lavrada a presente ata, que, após lida e aprovada, foi assinada pelos membros da Comissão Examinadora.



Prof. Dr. ANTONIO FERNANDO CREPALDI



Prof. Dr. CELESTE JOSÉ ZANON



Prof. Dr. CAMILO RODRIGUES NETO

## **AGRADECIMENTOS**

Aos meus familiares pela compreensão.

Aos meus amigos pelos momentos de descontração.

A minha avó Ana.

Ao Prof. Dr. Antonio Fernando Crepaldi, pela orientação competente e oportunidade de conviver e aprender com um ótimo pesquisador e excelente pessoa.

Aos professores do Curso de Pós-Graduação da Faculdade de Engenharia de Bauru (FEB – UNESP) que colaboraram para a minha formação científica.

A Renato Lataliza e demais profissionais da Vale que colaboram com orientações e explicações sobre o processo.

A todos os que direta ou indiretamente contribuíram para a realização desta pesquisa.

## RESUMO

Com a privatização da malha ferroviária no meio dos anos de 1990, a produtividade do setor vem crescendo em decorrência do aumento de investimentos por parte das concessionárias, tanto na modernização das vias quanto na melhoria das práticas de manutenção. Dentro deste contexto, o uso de modelos probabilísticos é de fundamental importância para se determinar o comportamento de um dado sistema, não sendo diferente com as ferrovias. Ao se adequar os tempos de reparo ao modelo de distribuição de Lei de Potência, espera-se aproveitar uma de suas principais propriedades, que é a invariância de escala, que permite a generalização do modelo empregado em um trecho para toda malha. Os dados se adequaram e o modelo pôde ser utilizado como mais uma ferramenta de auxílio à resolução de problemas e antecipação de indisponibilidades da malha.

**Palavras-chave:** *Ferrovia, Lei de Potência, Manutenção.*

## **ABSTRACT**

The privatization of the railroad in the middle of the years of 1990 increases the segment productivity by the concessionary investments in modernization and maintenance practices improvement. Inside the context of the maintenance, the use of probability models has a fundamental importance for determine system behavior, not being different with the maintenance of the railroads. Determine the times of repair in to the Power Law distribution means takes advantage of their main properties like a free scale. That is the build model can be generalized for the entire railroad. The data were adapted at the model and can be used as one more tool of aid for the resolution of problems and anticipation of unavailability of the railroad.

**Keywords:** *Power Law, Maintenance, Railroad.*

## SUMÁRIO

1. Introdução .....	1
1.1. Justificativa .....	2
1.2. Objetivo .....	3
1.3. Organização do trabalho .....	3
2. Distribuições de probabilidade .....	4
2.1. Distribuição normal.....	5
2.2. Teorema do limite central.....	7
2.3. Distribuição exponencial .....	8
2.4. Distribuição lognormal .....	9
2.5. Distribuição de weibull .....	11
2.6. Distribuição de Lévy .....	12
3. Distribuições de leis de potência .....	15
3.1. Normalização .....	17
3.2. Momentos .....	18
3.3. Grandes valores .....	19
3.4. Invariância de escala.....	21
3.5. Variáveis discretas.....	22
3.6. Estimação de parâmetros .....	23
4. Conceitos de manutenção.....	27
4.1. Tipos de manutenção.....	28
4.2. Confiabilidade.....	30

4.3.	Modelagem em manutenção.....	32
5.	Linha férrea .....	34
5.1.	Elementos da superestrutura.....	34
5.2.	Inspeção da linha.....	36
5.3.	Reparos .....	38
5.4.	Estrada de ferro vitória a minas – EFVM.....	39
6.	Método .....	41
7.	Resultados .....	44
8.	Conclusão e considerações .....	50
9.	Referências .....	52
10.	Anexos - .....	56

## ÍNDICE DE FIGURAS

Figura 1 - Gráfico da Distribuição Normal.....	6
Figura 2 - Gráfico da Distribuição Exponencial.....	9
Figura 3 - Gráfico da Distribuição Lognormal.....	10
Figura 4 – Gráfico da distribuição de Weibull.....	11
Figura 5 – Curva da banheira.....	31
Figura 6 - Componentes da superestrutura.....	35
Figura 7 - Aparelho de mudança de via (AMV).....	35
Figura 8 - Dormente podre na via.....	36
Figura 9 - Bitola da via.....	37
Figura 10 - Processo de soldagem para eliminação de juntas.....	38
Figura 11 - Mapa geral da EFVM.....	39
Figura 12 - Ramal de Fábrica.....	41
Figura 13 – Teste de adequação dos dados à distribuição Normal.....	44
Figura 14 – Teste de adequação dos dados à distribuição Exponencial.....	45
Figura 15 – Teste de adequação dos dados à distribuição Lognormal.....	45
Figura 16 – Teste de adequação dos dados à distribuição de Weibull.....	46
Figura 17 – Gráfico Log-Log dos tempos de reparo.....	47

## ÍNDICE DE TABELAS

Tabela 1 - Carga transportada pelas ferrovias nacionais .....	40
Tabela 2 – Acidentes nas ferrovias .....	40
Tabela 3 – Base de dados dos tempos de reparo (Jan/2010 à Jun/2012). .....	42

## 1. INTRODUÇÃO

Existem diversos problemas que impactam diretamente a competitividade das companhias instaladas no Brasil. Dentre eles, pode-se destacar a falta de infraestrutura logística adequada, tanto para o recebimento quanto para o escoamento dos produtos. E, apesar dos investimentos do governo federal e da iniciativa privada, que vem melhorando aos poucos a situação do transporte, os modais não se encontram no estado ideal de conservação, o que impacta diretamente na eficiente distribuição de produtos e por consequência no desempenho das empresas. Esses fatores influenciam diretamente nas cadeias produtivas que utilizam este modal como meio de escoamento de seus produtos, um deles é o custo operacional envolvido.

Para Russo (2012), o transporte de cargas ferroviário é caracterizado por apresentar um custo fixo bastante elevado, em razão dos altos valores envolvidos no arrendamento da malha e dos terminais, e na aquisição e manutenção da via e do material rodante. Em contrapartida os custos variáveis como combustíveis são relativamente baixos, justificando o uso do trem para o transporte de mercadorias de baixo valor agregado e grande volume.

O setor ferroviário brasileiro tem experimentado um crescimento considerável nos últimos anos devido ao processo de concessão das malhas federais à iniciativa privada. Após a privatização o setor demandava altos investimentos e urgente reestruturação das operações para que se tornasse competitivo. Frente a essa necessidade, as concessionárias direcionaram recursos para equipamentos, materiais rodantes e em novas tecnologias, com o objetivo de atrair novos clientes e de oferecer novos serviços ao mercado. Como resultado desses investimentos houve um grande aumento da produtividade das malhas concedidas, contudo estima-se que apenas 20% da produção do país seja transportada pelas estradas de ferro, o que é considerado baixo, tendo em vista da extensão territorial do Brasil e do potencial de utilização desse modal (CNT, 2011).

Por outro lado, o aumento da capacidade de um sistema de transporte, seja qual for a sua natureza, não está diretamente relacionado a expansão de sua estrutura física, uma vez que políticas de intervenção e as decisões humanas envolvidas nos processos também influenciam o seu desempenho. A solução para os problemas dessa natureza é o resultado da

combinação do uso eficiente da infraestrutura disponível e de métodos adequados de gerenciamentos, tais como o gerenciamento da manutenção (RUSSO, 2012).

Assim, a manutenção da malha ferroviária acaba se tornando uma função estratégica do processo produtivo, garantindo a confiabilidade, disponibilidade e segurança das operações nas vias. A correta escolha do tipo de manutenção e a elaboração de um modelo que descreva o comportamento ao longo do tempo das falhas e do sistema tem impacto direto sobre o desempenho das operações, podendo ser decisivo para o bom funcionamento da ferrovia.

O uso de modelos matemáticos nas atividades de manutenção pressupõe que todas as características e comportamentos do sistema sejam representados. E para a determinação desses modelos as distribuições de probabilidade são amplamente utilizadas, possuindo um importante papel para a identificação do sistema como um todo. Essas distribuições são usadas para os mais variados tipos de dados, diferenciando-se em relação à função de probabilidade (caso os dados sejam discretos) ou função densidade de probabilidade (dados contínuos) (SIQUEIRA, 2005).

Dentre as diversas distribuições de probabilidade utilizadas para a modelagem de sistemas contínuos existe uma em particular que devido as suas características, pode se adequar muito bem a este contexto da manutenção, e particularmente da manutenção das vias férreas. Essa distribuição tem um conceito básico que a gravidade ou tamanho de um evento é inversamente proporcional a sua frequência. Essa distribuição chamada Lei de Potência é frequentemente encontrada nos mais diversos sistemas, tanto naturais quanto produzidos pelo homem (PINTO *et al.* 2012).

### 1.1. JUSTIFICATIVA

O trabalho se justifica por contribuir com mais uma ferramenta de análise para a determinação do tempo gasto para reparo de uma ferrovia. Tempo este que está diretamente ligado à disponibilidade da via, tendo grande importância para o gerenciamento da manutenção. O conhecimento do comportamento dos tempos de reparo é fundamental para o planejamento adequado das ações reativas à indisponibilidade da linha e para os principais tipos de manutenção, desde a corretiva até a preditiva.

O trecho estudado fica no interior de Minas Gerais, pertencente a EFVM (Estrada de Ferro Vitória-Minas), linha esta que transporta cerca de 40% de toda a carga ferroviária brasileira. Os tempos analisados foram referentes às falhas ocorridas no período de janeiro de 2010 a junho de 2012, contabilizando 4046 falhas ocorridas no trecho.

## 1.2. OBJETIVO

O presente trabalho tem por objetivo verificar de que modo existência da relação entre os tempos de reparo de uma linha férrea e a distribuição de Lei de Potência podem ser utilizados como auxílio ao planejamento da manutenção.

Espera-se que este modelo seja capaz de descrever o comportamento dos tempos de reparo da ferrovia e contribua para a tomada de decisão no gerenciamento da manutenção.

Ao se elaborar este modelo busca-se também mostrar as principais características matemáticas das Leis de Potência.

## 1.3. ORGANIZAÇÃO DO TRABALHO

O trabalho está organizado em oito capítulos, onde o primeiro é referente à introdução do trabalho. No capítulo 2 é apresentada a revisão das distribuições de probabilidade, contendo os conceitos matemáticos e suas principais características. O capítulo 3 abrange a distribuição de Lei de Potência e suas propriedades. O capítulo 4 é aborda os tipos de manutenção e sua classificação quanto a tomada de decisão. O capítulo 5 apresenta as definições e nomenclaturas pertencentes ao sistema estudado, que é a ferrovia. O método empregado para a obtenção do modelo é descrito no capítulo 6 e os resultados obtidos no capítulo 7. E por fim as conclusões e considerações são apresentadas no capítulo 8.

## 2. DISTRIBUIÇÕES DE PROBABILIDADE

A teoria da probabilidade, apesar de conhecida há bastante tempo, só foi vista como disciplina científica em 1654, com os estudos de Pascal e Fermat sobre as possibilidades dos possíveis resultados obtidos em um determinado jogo. A grande maioria dos conceitos que conhecemos hoje foi sendo determinada ao longo do tempo, tendo grande contribuição de pesquisadores como Bernoulli, De Moivre, Laplace, Gauss e Poisson (BOCCARA, 2004).

Ao se fazer  $X$  uma variável discreta, e medindo essa variável muitas vezes durante repetidas observações, pode-se determinar um valor específico de  $X$ , ( $X_i$ ), cuja tendência é de que ao longo do tempo, conforme se aumenta o número de observações também aumente o número de vezes em que a variável assuma o valor  $X_i$ . Depois de um número suficientemente grande de medidas, a frequência relativa com que  $X_i$  ocorre vai se estabilizando e tendendo a um valor bem determinado, ao qual chamamos de probabilidade da ocorrência de  $X_i$ . Indicamos essa probabilidade por  $P(X = X_i)$ , ou de maneira simplificada  $P(X_i)$ . Se considerarmos o conjunto formado pelos diversos valores  $X_i$  que a variável  $X$  pode assumir, e suas respectivas probabilidades, temos uma distribuição de probabilidades (DANTAS, 2008).

As distribuições teóricas são modelos ideais, aos quais procuramos adaptar as variáveis reais, que podem ser discretas ou contínuas. A grande vantagem de se adaptar um fenômeno real a uma distribuição teórica é o conhecimento prévio das propriedades de determinada distribuição, o que permite a retirada de importantes conclusões sobre o fenômeno real estudado (MOREIRA, 2008).

Uma distribuição de probabilidade é definida por Montgomery e Runger (2009) como sendo um modelo matemático que relaciona o valor da variável com a ocorrência daquele valor na população. A média de uma distribuição de probabilidade é uma medida da tendência central da distribuição, ou de sua posição, sendo pela Eq. (2.1):

$$\mu = \begin{cases} \int_{-\infty}^{\infty} xf(x)dx, x \text{ contínua} \\ \sum_{i=1}^{\infty} x_i p(x_i), x \text{ discreta} \end{cases} . \quad 2.1$$

A dispersão ou variabilidade da distribuição é expressa pela variância  $\sigma^2$ . Sendo definida pela Eq. 2.2:

$$\sigma^2 = \begin{cases} \int_{-\infty}^{\infty} (x - \mu)^2 f(x)dx, & \text{se } x \text{ é contínua} \\ \sum_{i=1}^{\infty} (x_i - \mu)^2 p(x_i), & \text{se } x \text{ é discreta} \end{cases} . \quad 2.2$$

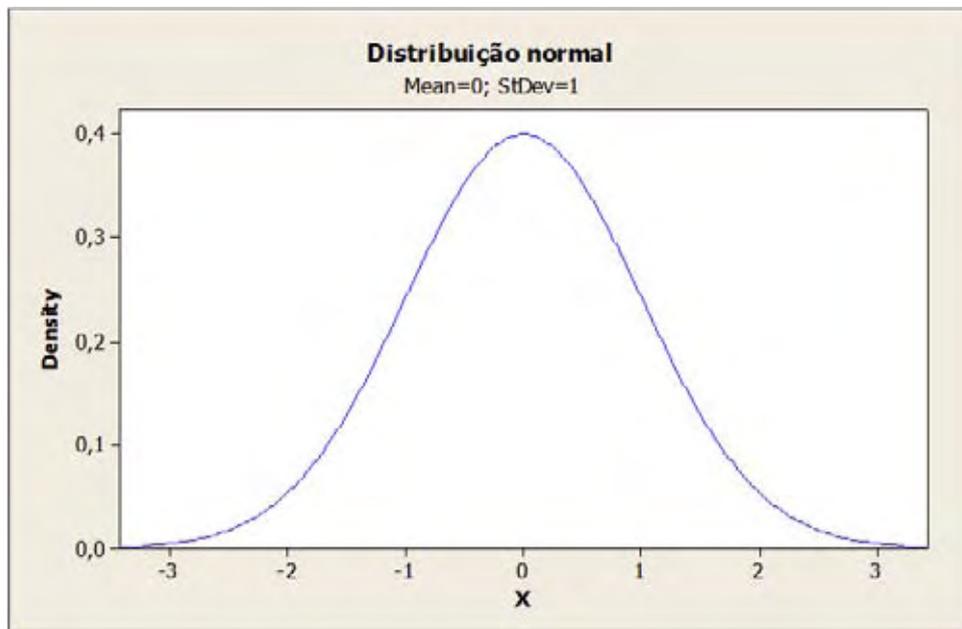
## 2.1. DISTRIBUIÇÃO NORMAL

Indubitavelmente, o modelo mais largamente utilizado para a distribuição de uma variável aleatória é a distribuição normal. Toda vez que um experimento aleatório for replicado, o valor média das observações se aproximará à uma distribuição normal. Este resultado foi apresentado por De Moivre em 1733. Infelizmente seu trabalho ficou perdido por algum tempo e Gauss, independentemente, desenvolveu uma distribuição normal cerca de 100 anos depois. Embora De Moivre tivesse recebido, posteriormente, o crédito pela dedução, uma distribuição normal é também referida como distribuição gaussiana (MEYER, 1983).

A distribuição normal aparece no estudo de numerosos fenômenos físicos básicos. Um exemplo foi a distribuição normal desenvolvida por Maxwell a partir de suposições simples, sobre a velocidade das moléculas. O valor  $E(x) = \mu$  conforme visto anteriormente na Eq. (2.1) determina o centro da função de densidade de probabilidade e o valor de  $V(X) = \sigma^2$  Eq. (2.2) determina a largura. Se uma variável aleatória normal apresentar  $\mu = 0$  e  $\sigma^2 = 1$  ela é chamada de variável aleatória normal padrão, sendo denotada por  $Z$  (BOCCARA, 2004).

A distribuição normal é representada graficamente por uma curva no formato de sino, simétrica, cujo maior valor é dado pela média. Isso quer dizer que a grande maioria dos dados amostrais está presente ao redor deste resultado, e que os valores vão diminuindo conforme se avança para as extremidades da curva, o que pode ser visto na Fig. 1. Nota-se ainda que  $f$  depende de  $x$  somente por intermédio da expressão  $(x - \mu)^2$ , o que torna evidente que o gráfico de  $f$  é simétrico em relação à  $\mu = x$ . Outra observação é que quando  $x \rightarrow \pm\infty, f(x) \rightarrow 0$ , assintoticamente (DANTAS, 2008).

Figura 1 - Gráfico da Distribuição Normal.



Fonte: Autor.

De acordo com Meyer (1983), dada uma variável aleatória  $X$ , que tome todos os valores reais  $-\infty < x < \infty$ , essa variável terá uma distribuição normal (ou Gaussiana) se possuir uma função de densidade de probabilidade da forma da Eq. (2.3):

$$f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{1}{2}\left[\frac{x - \mu}{\sigma}\right]^2\right). \quad 2.3$$

A importância da distribuição normal deve-se ao fato da grande maioria das variáveis que ocorrem na natureza, sejam elas físicas, biométricas, econômicas ou sociais, apresentarem esse mesmo comportamento.

## 2.2. TEOREMA DO LIMITE CENTRAL

De acordo com Meyer (1983), se uma variável aleatória  $X$  puder ser representada pela soma de quaisquer  $n$  variáveis aleatórias independentes, então, para  $n$  suficientemente grande,  $X$  terá distribuição aproximadamente normal.

Assim, considerando  $X_1, X_2, \dots, X_n$  uma sequência de variáveis aleatórias independentes, com  $E(X_i) = \mu$  e  $V(X_i) = \sigma^2$  (para  $i = 1, 2, \dots$ ) e fazendo  $X = X_1 + X_2 + \dots + X_n$ , temos que sob condições gerais  $Z_n$  tem distribuição normal com  $N(0,1)$ , Eq. (2.4).

$$Z_n = \frac{X - \sum_{i=1}^n \mu_i}{\sqrt{\sum_{i=1}^n \sigma_i^2}}. \quad 2.4$$

Este teorema constitui uma generalização da aproximação de De-Moivre-Laplace, devido ao fato das variáveis aleatórias independentes de  $X_i$ , assumirem os valores de 1 e 0 poderem ser substituídas por outras que apresentem quaisquer espécie de distribuição (desde que tenham expectância finita e variância finita). O fato das variáveis  $X_i$  poderem ter qualquer espécie de distribuição, e de sua soma ( $X = X_1 + X_2 + \dots + X_n$ ) poder ser aproximada por uma variável aleatória normalmente distribuída constitui a razão fundamental da importância da distribuição normal na teoria da probabilidade.

De modo geral, o teorema do limite central pode ser definido como: cada parcela na soma das variáveis aleatórias contribui com um valor sem importância para a variação da soma, e é muito improvável que qualquer parcela isolada dê uma contribuição muito grande para a soma; a soma de um número finito qualquer de variáveis aleatórias independentes normalmente distribuídas é também normalmente

distribuída. O teorema do limite central afirma que as parcelas não necessitam ser normalmente distribuídas, para que a soma seja aproximada por uma normal (MEYER, 1983).

### 2.3. DISTRIBUIÇÃO EXPONENCIAL

Segundo Dantas (2008), a distribuição exponencial desempenha um importante papel na descrição de uma grande classe de fenômenos, particularmente nos assuntos ligados à manutenção e disponibilidade de equipamentos. Pode ser definida como uma variável aleatória  $X$  que tome todos os valores não negativos com parâmetros  $\alpha > 0$ , se sua função de distribuição de probabilidade for dada por  $f(x) = \alpha e^{-\alpha x}, x > 0$ .

A função acumulada da distribuição exponencial é dada por

$$F(x) = P(X \leq x) = \int_0^x \alpha e^{-\alpha t} dt = 1 - e^{-\alpha x}, x \geq 0 . \quad 2.5$$

*F(x) = 0 para quaisquer outros valores*

Generalizando, a Eq.(2.5) pode ser definida por  $P(X > x) = e^{-\alpha x}$ . Os valores para as medidas de posição  $E(X)$  (valor esperado) e dispersão  $V(X)$  (variância) são dadas respectivamente pelas Eqs. (2.7) e (2.8), à partir da Eq. (2.6):

$$E(X) = \int_0^{\infty} x \alpha e^{-\alpha x} dx , \quad 2.6$$

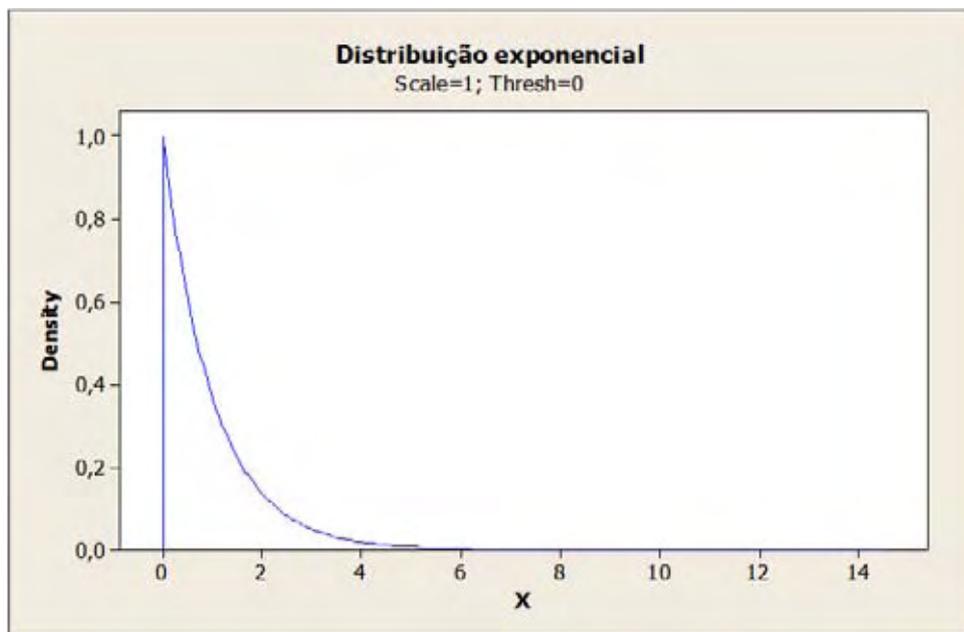
onde após a integração por partes têm-se finalmente:

$$E(X) = \frac{1}{\alpha} . \quad 2.7$$

$$V(X) = E(X^2) - [E(X)]^2 = \frac{1}{\alpha^2}. \quad 2.8$$

O gráfico da distribuição exponencial é mostrado na Fig. 2:

Figura 2 - Gráfico da Distribuição Exponencial



Fonte: Autor.

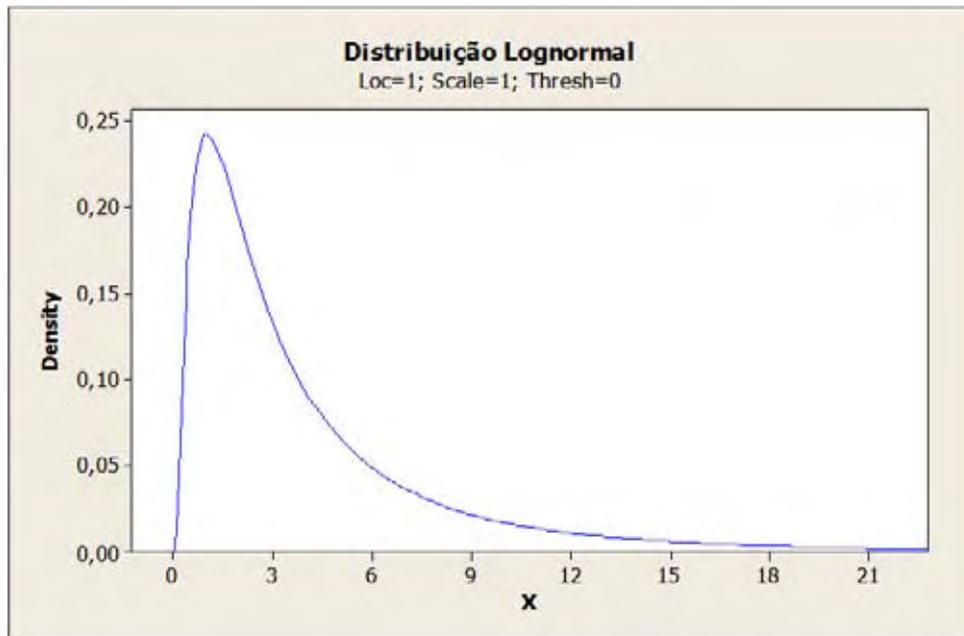
#### 2.4. DISTRIBUIÇÃO LOGNORMAL

Boccaro (2004) define a função de densidade de probabilidade da distribuição Lognormal, dependente dos parâmetros  $x_0$  e  $\sigma$  dada pela Eq. (2.9) seguinte:

$$f_{lognormal}(x) = \frac{1}{x\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(\log(x/x_0))^2}{2\sigma^2}\right). \quad 2.9$$

Isto é, se  $X$  é uma variável aleatória Lognormal,  $\log X$  é normalmente distribuída com  $\langle \log X \rangle = \log x_0$  e  $\langle (\log X)^2 \rangle - \langle \log X \rangle^2 = \sigma^2$ , que pode ser representada pelo gráfico da Fig. 3.

Figura 3 - Gráfico da Distribuição Lognormal



Fonte: Autor.

Uma distribuição Lognormal pode resultar da multiplicação de várias variáveis aleatórias, de modo que fazendo cada  $X_i (i = 1, 2, \dots, n)$  como uma variável aleatória independente têm-se:

$$\log \prod_{i=1}^n X_i = \sum_{i=1}^n \log X_i . \quad 2.10$$

A expressão representada pela Eq. (2.10) acima é válida desde que as variáveis independentes  $\log X_i (i = 1, 2, \dots, n)$  satisfaçam as condições referentes ao Teorema do Limite Central visto anteriormente. A distribuição Lognormal é muito usada na biologia para se medir a sensibilidade de determinado indivíduo a certas drogas,

fatores estes que podem depender de muitas variáveis como peso, frequência do pulso, pressão sistólica, nível de colesterol, dentre outros.

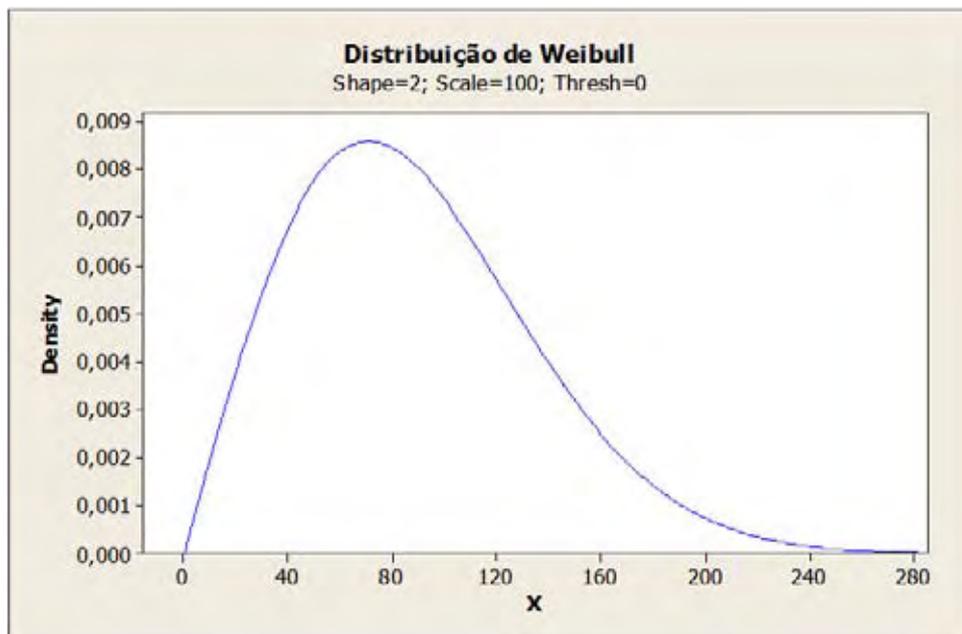
## 2.5. DISTRIBUIÇÃO DE WEIBULL

A distribuição de Weibull é conhecida como uma extensão da distribuição Gama, sendo definida como na Eq. (2.11):

$$f(x) = \frac{\beta}{\theta} \left(\frac{x}{\theta}\right)^{\beta-1} \exp\left[-\left(\frac{x}{\theta}\right)^\beta\right], \quad 2.11$$

onde  $\theta > 0$  é o parâmetro de escala e  $\beta > 0$  é o parâmetro de forma. Sendo representada pela Fig. 4:

Figura 4 – Gráfico da distribuição de Weibull



Fonte: Autor.

A média e a variância da distribuição de Weibull são dadas pelas Eq.(2.12) e (2.13) seguintes:

$$\mu = \theta \Gamma\left(1 + \frac{1}{\beta}\right). \quad 2.12$$

$$\sigma^2 = \theta^2 \left[ \Gamma\left(1 + \frac{2}{\beta}\right) - \Gamma\left(1 + \frac{1}{\beta}\right)^2 \right]. \quad 2.13$$

A distribuição de Weibull tem sido usada extensivamente na engenharia de confiabilidade como modelo do tempo de falha para componentes tanto mecânicos quanto elétricos. Uma de suas principais características é que ela consegue sintetizar dados discretos e contínuos em uma mesma função densidade de probabilidade (MONTGOMERY e RUNGER, 2009).

## 2.6. DISTRIBUIÇÃO DE LÉVY

Segundo Mantegna e Stanley (2000), Lévy e Khitchine foram responsáveis pela resolução do problema de determinação de distribuições estáveis. Eles encontraram a forma da maioria das funções de um processo estável, dada pela Eq. (2.14).

$$\ln \varphi(x) = \begin{cases} i\mu x - \gamma |x|^\alpha \left[ 1 - i\beta \frac{x}{|x|} \tan\left(\frac{\pi}{2}\alpha\right) \right] & [\alpha \neq 1] \\ i\mu x - \gamma |x| \left[ 1 + i\beta \frac{x}{|x|} \frac{2}{\pi} \ln|x| \right] & [\alpha = 1] \end{cases}, \quad 2.14$$

onde  $0 < \alpha \leq 2$ ,  $\gamma$  é o fator positivo de escala,  $\mu$  é um número real qualquer, e  $\beta$  é o parâmetro de assimetria que varia de -1 a 1.

A forma analítica da distribuição estável de Lévy é conhecida somente para poucos valores de  $\alpha$  e  $\beta$ :

- $\alpha = 1/2, \beta = 1$

- $\alpha = 1, \beta = 0$
- $\alpha = 2$

As distribuições consideradas serão somente as que apresentarem comportamentos estáveis simétricos ( $\beta = 0$ ) e com média igual a zero ( $\mu = 0$ ). De acordo com essas suposições a função característica assume a forma da Eq. (2.15):

$$\varphi(x) = e^{-\gamma|x|^\alpha}. \quad 2.15$$

A distribuição estável simétrica com índice  $\alpha$  e fator de escala  $\gamma$  ( $P_L(x)$ ) pode ser definida de acordo com a Eq. (2.16) que é a transformada inversa de Fourier e Eq.(2.15) vista anteriormente:

$$P(q) \equiv \frac{1}{2\pi} \int_{-\infty}^{\infty} \varphi(x) e^{-iqx} dx. \quad 2.16$$

$$P_L(x) \equiv \frac{1}{\pi} \int_0^{\infty} e^{-\gamma|x|^\alpha} \cos(qx) dq. \quad 2.17$$

Para  $\gamma = 1$ , a expansão da série é válida para grandes argumentos ( $|x| \gg 0$ ) é:

$$P_L(|x|) = -\frac{1}{\pi} \sum_{k=1}^n \frac{(-1)^k \Gamma(\alpha k + 1)}{k! |x|^{\alpha k + 1}} \sin\left[\frac{k\pi\alpha}{2}\right] + R(|x|), \quad 2.18$$

onde  $\Gamma(x)$  é a função  $\Gamma$  de Euler determinada pela Eq. (2.19):

$$\Gamma(\alpha) = \int_0^{\infty} x^{\alpha-1} e^{-x} dx. \quad 2.19$$

Da equação  $P_L(|x|)$  encontra-se a aproximação assintótica da distribuição estável de índice  $\alpha$ , Eq.(2.20):

$$P_L(|x|) \sim \frac{\Gamma(1 + \alpha) \sin(\pi \alpha/2)}{\pi |x|^{1+\alpha}} \sim |x|^{-(1+\alpha)}. \quad 2.20$$

O comportamento assintótico para grandes valores de  $x$  pode ser considerado como pertencente a uma Lei de Potência, que interfere ativamente na definição dos momentos da distribuição.

### 3. DISTRIBUIÇÕES DE LEIS DE POTÊNCIA

A maioria das variáveis inerentes aos mais diversos processos apresentam um comportamento de agrupamento em torno de um dado valor central (Teorema do Limite Central). Embora este seja um comportamento comum para a maioria dos casos, existem algumas variáveis que não possuem esse comportamento, apresentando grande variabilidade, às vezes de muitas ordens de magnitude, como o tamanho das cidades (NEWMAN, 2005).

Dentre as distribuições que não seguem o comportamento do Teorema do Limite Central, as Leis de Potência tem atraído um pouco mais de atenção no decorrer dos últimos anos, isso se deve à algumas de suas propriedades matemáticas que podem conduzir a diversos resultados práticos (CLAUSET *et al.*, 2009).

Muitas análises empíricas sugerem que o comportamento das Leis de Potência na distribuição de valores é muito frequente na natureza. Esse comportamento pode ser encontrado na distribuição dos tamanhos dos incêndios florestais, da intensidade de terremotos, do número de mortos em ataques terroristas até da riqueza de pessoas e nações (PINTO *et al.*, 2012).

Entretanto, deve-se manter em mente que esse comportamento pode ser puramente aparente ou apenas uma transição de fase. A Lei de Potência pode se manifestar de diversas maneiras, onde uma grande quantidade de elementos interage entre si para produzir uma estrutura de nível superior. São descritos por expressões matemáticas do tipo  $Y = cX^b$ , onde  $X$  e  $Y$  são duas variáveis, quantidades ou observações,  $c$  é uma constante e  $b$  é expoente de escala (GLERIA *et al.*, 2004).

Os estudos dos mecanismos responsáveis pelo aparecimento das distribuições de Leis de Potência representam um ativo campo de pesquisa. Esse tipo de distribuição foi primeiramente encontrado em áreas distantes da física tradicional, como a Lei de Pareto da distribuição de riqueza na economia, e a lei de Zipf para a taxa de ocorrência de palavras no campo da linguística, sugerindo a existência de características gerais para a dinâmica básica. Leis de Potência também são encontradas em sistemas em desequilíbrio, como avalanches ou processos estocásticos multiplicativos. Apresentam uma peculiar característica, tendo

notavelmente uma cauda diferente das outras distribuições definidas pelo equilíbrio da mecânica estatística, como a distribuição exponencial (MITZENMACHER, 2003).

Uma variável aleatória não negativa  $X$  é dita ter uma distribuição de Lei de Potência se apresentar a seguinte forma dada pela Eq. (3.1) de maneira mais simplificada:

$$P(X \geq x) \sim Cx^{-\alpha}, \quad 3.1$$

onde as constantes são definidas como  $C > 0$  e  $\alpha > 0$ . De modo geral, em uma distribuição de Lei de Potência, as caudas assintóticas decaem de acordo com a potência  $\alpha$ . Tal distribuição nos conduz para uma cauda muito mais “pesada” que alguns modelos comuns como o da distribuição exponencial por exemplo. O formato da distribuição de Lei de Potência para a lei de Pareto é dada pela Eq. (3.2):

$$P(X \geq x) = \left(\frac{x}{k}\right)^{-\alpha}, \quad 3.2$$

para qualquer  $k > 0$  e  $\alpha > 0$ . A distribuição de Pareto requer  $X \geq k$ . A função de densidade para a distribuição de Pareto é  $f(x) = \alpha k^\alpha x^{-\alpha-1}$ . Em cada caso  $X$  possui variância infinita. Se  $\alpha < 1$ , então  $X$  também possui média infinita (MITZENMACHER, 2003).

Se  $X$  for uma distribuição de Lei de Potência, então, em um gráfico Log-Log de  $P(X \geq x)$  (também conhecido como *função de distribuição cumulativa complementar*) assintoticamente o comportamento da função será o de uma linha reta. O que proporciona um teste empírico simples para determinar se a amostra é aderente a uma Lei de Potência. Para o caso específico da distribuição de Pareto, o comportamento é exatamente linear, como:

$$\ln(P(X \geq x)) = -\alpha(\ln x - \ln k). \quad 3.3$$

Similarmente, em um gráfico Log-Log a função de densidade para a distribuição de Pareto é também uma linha reta (PINTO *et al.*, 2012):

$$\ln f(x) = (-\alpha - 1) \ln x + \alpha \ln k + \ln \alpha . \quad 3.4$$

### 3.1. NORMALIZAÇÃO

De acordo com Newman (2005), a distribuição de Lei de Potência pode ser dividida segundo as características das variáveis envolvidas: distribuições contínuas governadas por números reais e distribuições discretas, normalmente representadas por inteiros positivos. Sendo  $x$  a quantidade de interesse, temos para o caso contínuo a função densidade de probabilidade dada por:

$$p(x)dx = P(x \leq X \leq x + dx) = Cx^{-\alpha}dx, \quad 3.5$$

onde  $X$  é o valor observado e  $C$  é a constante de normalização. A função anterior diverge para todo  $x \rightarrow 0$ , deste modo para que a função se mantenha, existe um valor mínimo que delimita o comportamento da Lei de Potência, denotado por  $x_{min}$ . Então, determinando  $\alpha > 1$ , a constante de normalização pode ser calculada conforme as Eqs. (3.5) e (3.6):

A constante  $C$  na Eq.(3.5) é dada por um tratamento de normalização onde:

$$1 = \int_{x_{min}}^{\infty} p(x)dx = C \int_{x_{min}}^{\infty} x^{-\alpha}dx = \frac{C}{1-\alpha} [x^{-\alpha+1}]_{x_{min}}^{\infty} . \quad 3.6$$

A equação acima só faz sentido se  $\alpha > 1$ , caso contrário, o lado direito da equação divergiria e os expoentes das Leis de Potência menores que 1 de não poderiam ser normalizados. Se  $\alpha > 1$  então a Eq. (3.6) resulta em:

$$C = (\alpha - 1)x_{min}^{\alpha-1} , \quad 3.7$$

e a expressão da Lei de Potência torna-se:

$$p(x) = \frac{\alpha - 1}{x_{min}} \left( \frac{x}{x_{min}} \right)^{-\alpha} . \quad 3.8$$

Algumas distribuições seguem uma Lei de Potência somente para parte dos dados, desconsiderando os grandes valores de  $x$ . Ou seja, para esses dados, a função diverge e decai rapidamente para zero. Se isso acontece, então a distribuição pode ser normalizada não importando o valor do expoente.

### 3.2. MOMENTOS

O valor da média de  $x$  de uma distribuição de Lei de Potência é dado pela Eq. (3.9):

$$\langle x \rangle = \int_{x_{min}}^{\infty} xp(x)dx = C \int_{x_{min}}^{\infty} x^{-\alpha+1} dx = \frac{C}{2 - \alpha} [x^{-\alpha+2}]_{x_{min}}^{\infty} . \quad 3.9$$

A expressão se torna infinita se  $\alpha \leq 2$ . Leis de Potência com valores abaixo de  $\alpha \leq 2$  não apresentam uma média finita, portanto não se pode precisar seus momentos. Entretanto, pode-se obter um valor bem definido a obtenção desses momentos, desde que cada medida individual de  $x$  seja um número finito. Porém, ao se calcular a média desses valores repetidas vezes, o resultado tende a ser igual à média geral, como se todas as repetições fossem combinadas em um grande experimento. Isto implica que a média pode assumir um valor relativamente pequeno em qualquer repetição particular do experimento, deve-se pegar ocasionalmente um grande valor, para que a média global divirja à medida que se faz o número de repetições.

A divergência formal de  $\langle x \rangle$  é um sinal de que esse valor não é confiável para a obtenção do tamanho típico das amostras de outros exemplos do mesmo experimento. Para  $\alpha > 2$ , entretanto, a média é perfeitamente bem definida, com o valor dado pela Eq. (3.10):

$$\langle x \rangle = \frac{\alpha - 1}{\alpha - 2} x_{min} . \quad 3.10$$

Pode-se também calcular momentos mais altos para a distribuição  $p(x)$ . Por exemplo, o segundo momento, é dado por:

$$\langle x^2 \rangle = \frac{C}{3 - \alpha} [x^{-\alpha+3}]_{x_{min}}^{\infty} , \quad 3.11$$

que diverge se  $\alpha \leq 3$ . Para este caso a distribuição de Lei de Potência não apresenta significativas médias quadradas, variância ou desvio padrão. Se  $\alpha > 3$ , então o segundo momento é finito e bem definido, tendo como valor a Eq. (3.12):

$$\langle x^2 \rangle = \frac{\alpha - 1}{\alpha - 3} x_{min}^2 . \quad 3.12$$

Estes resultados podem ser facilmente expandidos para mostrar que em geral todos os momentos  $\langle x^m \rangle$  existem para  $m < \alpha - 1$  e todos os momentos maiores divergem. De modo geral para Leis de Potência os momentos existentes são dados de acordo com a Eq. (3.13) (NEWMAN, 2005):

$$\langle x^m \rangle = \frac{\alpha - 1}{\alpha - 1 - m} x_{min}^m . \quad 3.13$$

### 3.3. GRANDES VALORES

Tomando-se  $n$  medições de uma distribuição de Lei de Potência, deve-se determinar qual o maior valor dessas medições que provavelmente será retirado, ou de um modo mais preciso a probabilidade  $P(x)dx$  de se encontrar o maior valor no intervalo entre  $x$  e  $x+dx$ . A propriedade definitiva do maior valor em uma amostra é de que não há nenhum outro valor maior do que o que já foi retirado. Assim para se determinar essa probabilidade  $x$  que é  $P(x)$ , pode-se usar a Eq. (3.14):

$$P(x) = \int_x^{\infty} p(x')dx' = \frac{C}{\alpha - 1} x^{-\alpha+1} = \left(\frac{x}{x_{min}}\right)^{-\alpha+1}, \quad 3.14$$

válida para  $\alpha > 1$ . As chances de uma amostra qualquer não ser maior do que  $x$  é  $1 - P(x)$ , de ficar entre  $x$  e  $x+dx$ , incluindo todos os outros valores é  $p(x)dx \times [1 - P(x)]^{n-1}$ . Dessa forma existem  $n$  maneiras de se escolher  $i$ , para se encontrar o valor da probabilidade total.

Agora, para se calcular o valor da média  $\langle x_{max} \rangle$  para grandes amostras têm-se:

$$\langle x_{max} \rangle = \int_{x_{min}}^{\infty} xp(x)dx = n \int_{x_{min}}^{\infty} xp(x)[1 - P(x)]^{n-1} dx. \quad 3.15$$

Usando as Eq. (3.9) e (3.15):

$$\begin{aligned} \langle x_{max} \rangle &= n(\alpha - 1) \int_{x_{min}}^{\infty} \left(\frac{x}{x_{min}}\right)^{-\alpha+1} \left[1 - \left(\frac{x}{x_{min}}\right)^{-\alpha+1}\right]^{n-1} dx \\ &= nx_{min} \int_0^1 \frac{y^{n-1}}{(1-y)^{1/(\alpha-1)}} dy = nx_{min} B(n, (\alpha - 2)/(\alpha - 1)), \end{aligned} \quad 3.16$$

quando fazemos a substituição  $y = 1 - (x/x_{min})^{-\alpha+1}$  e  $B(a,b)$  é a função-beta de Legendre, cuja definição é dada por

$$B(a,b) = \frac{\Gamma(a)\Gamma(b)}{\Gamma(a+b)}. \quad 3.17$$

Na maioria dos casos, o número  $n$  de amostras da distribuição de Lei de Potência se tornará grande, quando:

$$B(n(\alpha - 2)/(\alpha - 1)) \sim n^{(\alpha-2)/(\alpha-1)} = \langle x_{max} \rangle \sim n^{1/(\alpha-1)}. \quad 3.18$$

Então quanto maior for  $\alpha > 1$ , pode-se dizer que o valor de  $\langle x_{max} \rangle$  será sempre maior quanto maior for o valor de  $n$  (NEWMAN, 2005).

### 3.4. INVARIÂNCIA DE ESCALA

Clauset *et al.* (2009) definem a distribuição de Lei de Potência como sendo uma distribuição livre de escala. Isto se deve ao fato dela ser a única distribuição que apresenta a mesma escala independentemente do ponto analisado. Suponha uma distribuição de probabilidade  $p(x)$  para uma quantidade  $x$ , e que ela satisfaz a seguinte propriedade:

$$p(bx) = g(b)p(x), \quad 3.19$$

para qualquer que seja  $b$ . Quer dizer, ao se aumentar a escala ou unidade pelas quais  $x$  é medido por um fator de  $b$ , a forma da distribuição  $p(x)$  fica inalterada, a não ser que essa distribuição seja multiplicada por uma constante.

Esta propriedade só é válida para a distribuição de Lei de Potência. Isso faz com que ela tenha uma grande importância para a generalização de diversos tipos de sistemas, podendo se obter conclusões sobre todo o sistema conhecendo somente parte deste. Assim, matematicamente pode-se chegar à propriedade de invariância de escala, partir da Eq. (3.18), fazendo  $x = 1$ , e  $p(b) = g(b)p(1)$ . Deste modo  $g(b) = p(b)/p(1)$  e a Eq. (3.18) pode ser escrita como:

$$p(bx) = \frac{p(b)p(x)}{p(1)}. \quad 3.20$$

Considerando que a Eq. (3.20) seja verdadeira para qualquer  $b$ , após processos de reorganização e diferenciação chega-se a Eq. (3.21):

$$p(x) = p(1)x^{-\alpha}. \quad 3.21$$

A equação Eq. (3.21) satisfaz o critério de invariância de escala.

Existem alguns sistemas que se tornam livres de escala para certos valores de seus parâmetros. Este valor é chamado de “transição de fase contínua”, e o argumento dado acima mostra que a partir deste ponto as quantidades observáveis no sistema podem apresentar uma distribuição de Lei de Potência. Isso pode ser determinado de modo experimental, sendo um dos fatores que motivou o estudo das Leis de Potência na física (as Leis de Potência não são somente o provável resultado de uma transição de fase – a variação de outros sistemas pode produzir o comportamento de Lei de Potência) (GLERIA *et al*, 2004).

### 3.5. VARIÁVEIS DISCRETAS

A Lei de Potência só é obedecida na cauda da distribuição onde os valores medidos são tão grandes que podem ser considerados contínuos. Tecnicamente, porém, as distribuições de Lei de Potência deveriam ser definidas de maneiras diferentes para valores inteiros (NEWMAN, 2005).

Se  $k$  for uma variável inteira, uma alternativa é declarar que ela segue uma Lei de Potência se a probabilidade  $p_k$  dos valores de  $k$  medidos obedece:

$$p_k = Ck^{-\alpha}, \quad 3.22$$

para alguma constante no expoente  $\alpha$ . Para  $k \neq 0$  a constante  $C$  é então definida pela condição de normalização:

$$1 = \sum_{k=1}^{\infty} p_k = C \sum_{k=1}^{\infty} k^{-\alpha} = C\zeta(\alpha), \quad 3.23$$

onde  $\zeta(\alpha)$  é a função  $\zeta$  de Riemann. Rearranjando, a Eq. (3.23) e identificando o comportamento da Lei de Potência na cauda da distribuição para valores de  $k \geq k_{min}$ , a expressão equivalente é:

$$p_k = \frac{k^{-\alpha}}{\zeta(\alpha, k_{min})}, \quad 3.24$$

onde  $\zeta(\alpha, k_{min}) = \sum_{k=k_{min}}^{\infty} k^{-\alpha}$  é a generalização da função  $\zeta$ .

Para Clauset *et al.* (2009), uma alternativa melhor e frequentemente mais conveniente para a distribuição de Lei de Potência para valores discretos é dada pela Eq.(3.25):

$$p_k = C \frac{\Gamma(k)\Gamma(\alpha)}{\Gamma(k + \alpha)} = CB(k, \alpha). \quad 3.25$$

Essa função se comporta como uma Lei de Potência ( $B(k, \alpha) \sim k^{-\alpha}$ ), para grandes valores de  $k$ . Essa equação também é conhecida como distribuição de Yule. O primeiro e segundo momento são dados de acordo com a Eq. (3.26) seguinte:

$$\langle k \rangle = \frac{\alpha - 1}{\alpha - 2}, \quad \langle k^2 \rangle = \frac{(\alpha - 1)^2}{(\alpha - 2)(\alpha - 3)}. \quad 3.26$$

### 3.6. ESTIMAÇÃO DE PARÂMETROS

O estudo de distribuições empíricas que apresentam um comportamento ajustado à Lei de Potência normalmente fornece uma estimativa do parâmetro de escala  $\alpha$  e ocasionalmente também o limite inferior da região de escala  $x_{min}$ . Conforme visto anteriormente, a ferramenta usualmente empregada para este trabalho é um simples histograma. Infelizmente este método e outras variações do mesmo

tema costumam gerar alguns erros sistemáticos, o que faz com que os resultados iniciais dados pela curva Log-Log não sejam tão confiáveis (CLAUSET *et al.*, 2009).

O primeiro parâmetro a ser estimado é  $\alpha$ , considerando que o valor do limite inferior  $x_{min}$  é conhecido. O método de escolha para ajuste dos parâmetros referentes aos dados observados é o método da máxima verossimilhança, que costuma fornecer valores bem ajustados para o limite da grande maioria das amostras simples. Assumindo que os dados vieram de uma distribuição que segue a Lei de Potência para  $x > x_{min}$  podemos derivar os estimadores por máxima verossimilhança do parâmetro de escala:

$$\hat{\alpha} = 1 + n \left[ \sum_{i=1}^n \ln \frac{x_i}{x_{min}} \right]^{-1}, \quad 3.27$$

onde  $x_i, i = 1, \dots, n$  são os valores observados de  $x$  tais que  $x \geq x_{min}$ . A Eq. (3.27) é equivalente ao estimador de Hill, que é assintoticamente normal e consistente (ou seja,  $\hat{\alpha} \rightarrow \alpha$  no limite para grandes valores de  $n$ ). O erro padrão para  $\hat{\alpha}$  é derivado da “largura” de máxima verossimilhança, dada pela Eq. (3.28) (NEWMAN, 2005):

$$\sigma = \frac{\hat{\alpha} - 1}{\sqrt{n}} \quad 3.28$$

Alguns os dados empíricos costumam seguir uma distribuição de Lei de Potência somente para valores acima de um determinado limite inferior denominado  $x_{min}$ . Antes do cálculo da estimativa do parâmetro de escala  $\alpha$ , primeiro é necessário descartar todas as amostras abaixo deste ponto  $x_{min}$  para que o restante dos dados esteja adequado ao modelo de distribuição de Lei de Potência. Então, para a obtenção de um modelo assertivo da estimativa de  $\alpha$ , deve-se ter também um método preciso de estimativa para o valor de  $x_{min}$ . Uma das principais preocupações com a estimação do valor de  $x_{min}$  é que caso o valor escolhido seja muito baixo o parâmetro de escala pode ser parcialmente influenciado. Caso o valor de  $x_{min}$  seja muito alto corre-se o risco de excluir pontos legítimos da distribuição de Lei de Potência,  $x_i < \hat{x}_{min}$ , que aumenta tanto o erro estatístico do parâmetro de escala quanto à influência de efeitos de tamanhos finitos (NEWMAN, 2005).

O modo mais comum de escolher  $x_{min}$  é calcular o ponto visualmente onde a função se torna uma linha reta no gráfico Log-Log, ou construir um gráfico de  $\hat{\alpha}$  em função de  $x_{min}$  e determinar o ponto de estabilidade da curva resultante. Essas aproximações são claramente subjetivas e sensíveis a ruídos na cauda da distribuição. A ideia fundamental da aproximação do cálculo de  $x_{min}$  é a de escolher o valor determinante para que as distribuições de probabilidade dos dados medidos e o modelo de Lei de Potência escolhido sejam ajustados da mesma forma. De modo geral ao se escolher  $x_{min}$  maior do que o verdadeiro valor de  $x_{min}$  tem-se uma redução efetiva do tamanho do conjunto dos dados, e de modo inverso, ao se escolher  $x_{min}$  menor do que o verdadeiro  $x_{min}$  as distribuições do modelo e dos dados empíricos serão diferentes (CLAUSET *et al.*, 2009).

O método para se determinar o ajuste entre a distribuição dos dados e a distribuição de Lei de Potência para dados não normais é dado pela quantificação da distância entre as duas distribuições de probabilidade. Essa distância é definida de acordo com o teste de Kolmogorov-Smirnov que determina que a distância máxima entre os dados da função de densidade de probabilidade e modelo escolhido pode ser determinado de acordo com a Eq. (3.29)

$$D = \max_{x > x_{min}} |S(x) - P(x)| \quad 3.29$$

Aqui  $S(x)$  é a função para os dados das observações com valores de no mínimo  $x_{min}$  e  $P(x)$  é a função para o modelo de Lei de Potência que melhor se adapta aos dados na região onde  $x \geq x_{min}$ . A estimativa para o valor de  $x_{min}$  é o valor que minimiza a função  $D$  (CLEMENTI, F. *et al.*, 2006)

Uma aproximação padrão para verificar se certo conjunto de dados segue uma distribuição de Lei de Potência é usar um teste de melhoria de ajuste que gera um valor  $p$ . A hipótese nula é de que os dados se adequam ao tipo de função testada. Este teste é baseado na medida distância entre a distribuição dos dados empíricos e o modelo hipotético. A distância é comparada com as dimensões do conjunto de dados sintéticos retirados do modelo, e o  $p$ -valor é definido como sendo a fração das distâncias sintéticas que são maiores do que a distância empírica. Se  $p$ -valor é muito grande (próximo de 1) as diferenças entre os dados

empíricos e o modelo podem ser consideradas somente flutuações estatísticas (ruídos), se por outro lado *p-valor* for muito pequeno o modelo não é ajustável aos dados. (GOLDSTEIN *et al.*, 2004).

Ao se calcular o *p-valor* é necessário decidir o quão pequeno será seu valor para que os dados não sigam uma Lei de Potência, ou de modo semelhante se os dados confirmam a hipótese em questão. Desta forma para valores de *p* menores que 0,1 não se deve afirmar que os dados seguem uma distribuição de Lei de Potência. É importante ressaltar que um *p-valor* grande não significa necessariamente que a distribuição de Lei de Potência é a distribuição correta dos dados, pois pode haver outras distribuições que correspondem igualmente aos dados, e é possível que para tamanhos pequenos de amostra (*n*), a distribuição aproximadamente semelhante à Lei de Potência e então o *p-valor* será alto, mesmo quando a distribuição não for adequada aos dados (CLAUSET *et al.*, 2009).

#### 4. CONCEITOS DE MANUTENÇÃO

Todos os sistemas produtivos estão sujeitos ao desgaste e deterioração ao longo do tempo. Sendo assim, esses sistemas devem ser submetidos a um programa de manutenção constante para que seja mantida a capacidade produtiva e garantida a segurança de todos os envolvidos direta ou indiretamente nos processos. Com a ferrovia não é diferente, e por isso a manutenção é uma atividade permanente e decisiva para o seu contínuo funcionamento (KARDEC, 2001).

A manutenção pode ser definida como o conjunto de ações para detectar, prevenir, ou corrigir falhas e defeitos com o objetivo de manter as condições operacionais e de segurança dos itens, sistemas ou ativos. O que faz da manutenção uma função estratégica do processo produtivo e influi diretamente na competitividade das organizações, uma vez que é através dela que é garantida a confiabilidade, disponibilidade e a segurança das operações das plantas industriais (SIQUEIRA, 2005).

Uma das maneiras de se otimizar o faturamento de qualquer sistema é a melhoria da disponibilidade e confiabilidade dos processos de forma geral. Isto pode ser traduzido em aumento de produtividade, minimização dos prazos de parada dos sistemas e outros fatores que só podem ser obtidos com o empenho e eficácia da manutenção da organização em questão (SANTOS *et al.*, 2007).

Um sistema de produção pode apresentar um grande número de componentes técnicos, cada qual com uma função bem definida e inter-relacionada. Cada parte desse sistema contribui para alcançar o objetivo global proposto pela organização. A manutenção contribui para que a essas metas sejam alcançadas por meio da integração total dos processos. Dentre as principais metas podemos citar a redução do custo total (maior tempo de funcionamento), a melhoria do equipamento, a segurança das pessoas e do ambiente. Todas elas impõem cargas de trabalho maiores para as equipes envolvidas, pressionando para o aumento da eficiência e capacidade de resolução de problemas (FUENTES, 2006).

Santos *et al.* (2007) afirmam que a evolução da manutenção está ligada à própria evolução humana, pois a criação e conservação dos objetos permite um

conhecimento e domínio maior sobre o ambiente. Mesmo com o constante avanço tecnológico, tanto produtos quanto equipamentos de produção, têm uma duração limitada, e certamente falharão em algum momento de suas vidas, tendo isso em vista, a importância da manutenção para manter ou recuperar a funcionalidade desses equipamentos é cada vez mais frequente. O estudo da manutenção engloba alguns termos práticos que são frequentemente usados no estudo de sistemas de manutenção. A seguir têm-se algumas definições:

**Falha:** é o término da habilidade ou a impossibilidade de um item para desempenhar sua função requerida. Leva o item, invariavelmente, ao estado indisponível.

**Defeito:** é a imperfeição do estado de uma instalação ou equipamento, embora não chegue a ponto de causar o fim da habilidade do mesmo de desempenhar a função requerida, já que a instalação ou equipamento pode operar com restrições.

**Manutenção perfeita:** quando no ato da manutenção, além de reparar componentes do equipamento falho ou com iminência de falha, atua-se também nos componentes que apresentam potencialidade de falha.

**Reparo mínimo:** restaura o equipamento ao estado em que se encontrava imediatamente antes da falha. Atua-se somente na parte defeituosa do equipamento, substituindo-a ou restaurando a sua condição original de funcionamento.

#### 4.1. TIPOS DE MANUTENÇÃO

De acordo com Branco Filho (2004), existem várias classificações para os tipos de manutenção existentes, de forma geral, ela é dividida em três tipos básicos: manutenção corretiva, preventiva e preditiva.

A manutenção corretiva pode ser definida pelo conjunto de tarefas que são realizadas para a correção de uma falha que provoca a parada ou desempenho abaixo do esperado. Geralmente, este tipo de manutenção ocorre após a parada ou baixo rendimento de um determinado equipamento ou sistema que causa distúrbios à atividade produtiva (FUENTES, 2006).

A manutenção corretiva é a execução das tarefas de manutenção não planejadas para restaurar as capacidades funcionais de equipamentos ou sistemas falhos. É uma das mais antigas e caras formas de manutenção, apesar disso tem uma grande importância, pois não se pode prever com exatidão qual o momento em que uma falha ocorrerá, obrigando uma manutenção corretiva. Como vantagem deste tipo de manutenção tem-se a não obrigatoriedade de planejamento estratégico para a execução das tarefas. Mas, certo nível de planejamento no que diz respeito ao uso de ferramentas e equipamentos, peças de reposição e técnicos especializados é altamente recomendado. Devido às próprias características, este tipo de manutenção pode causar estoque de peças não usadas, desconhecimento do estado técnico do sistema e paradas caras e demoradas (LUCATELLI, 2002).

A manutenção preventiva é caracterizada pelas tarefas realizadas com o objetivo de reduzir ou evitar falhas e quedas no nível produtivo, obedecendo a um plano previamente elaborado, baseado em intervalos definidos de tempo ou produção. Sendo assim, este tipo de manutenção tem o objetivo de antecipar às falhas, apresentando um sistema bem definido para realização das manutenções independentemente da condição do equipamento. Os prazos de intervenção são definidos pelos fabricantes dos equipamentos industriais, experiência dos mantenedores e por estudos das áreas de engenharia de manutenção local (SIQUEIRA, 2005).

Na manutenção preventiva o elemento mais importante é o tempo médio entre falhas, que pode ser obtido por meio da análise dos dados referentes ao tempo em que o sistema se manteve parado ou com o funcionando abaixo do desejado. Para a substituição de itens deteriorados com uma frequência determinada, deve-se estabelecer a programação das ações que garantam a realização dessas atividades em seu devido momento. Devem ser observados alguns preparativos como: um manual de procedimentos, programação de peças sobressalentes, frequência de realização das inspeções, dimensionamento da equipe técnica e registro das atividades. Esse tipo de manutenção é utilizado para manter certo item em condições satisfatórias de operação, por meio de inspeções periódicas e, detecção e prevenção de falhas incipientes. Pode se basear no tempo de uso ou na condição do equipamento. É

tomada como referência a data em que as atividades de reparo foram planejadas para serem realizadas nos equipamentos em operação. (SANTOS *et al*, 2007).

A manutenção preditiva é realizada com base em modificações de parâmetros de condição ou desempenho, cujo acompanhamento obedece a uma sistemática de medições que são realizadas com os equipamentos em funcionamento e que, após análises estatísticas dos dados coletados é possível estimar uma quebra ou parada de determinado equipamento. Tal tipo de manutenção possui metodologias, procedimentos e equipamentos específicos que permitem parar o equipamento antes que a falha ocorra e corrigi-la de forma programada, causando assim o mínimo de impacto no processo produtivo (KARDEC, 2001).

Este tipo de manutenção busca determinar antecipadamente a necessidade de reparos para que seja possível a eliminação de montagens desnecessárias, o aumento da vida útil e da disponibilidade de equipamentos. Esta política de manutenção apesar de ser extremamente eficaz quando aplicada ao problema correto, não pode ser utilizada em todo e qualquer item. Para se utilizar este tipo de manutenção, deve-se monitorar e acompanhar a condição das máquinas por meio de técnicas de controle estatístico de processo e equipamentos especializados. Como desvantagens este tipo de manutenção apresenta: demanda de grande investimento inicial, mão de obra altamente qualificada e impossibilidade de se realizar para sistemas produtivos complexos (LUCATELLI, 2002).

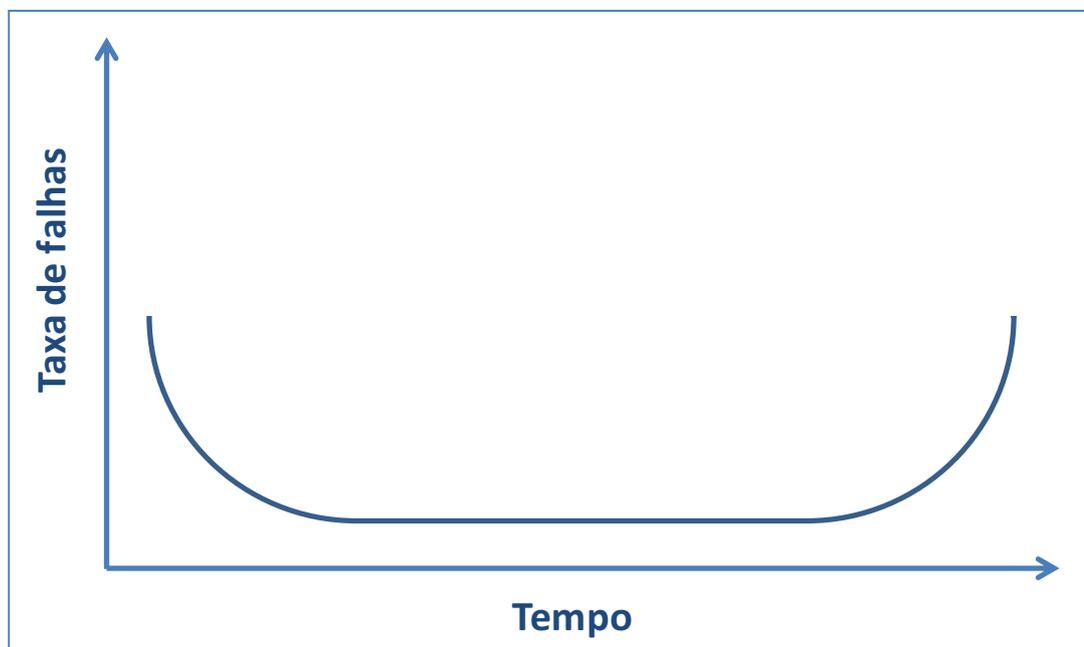
#### 4.2. CONFIABILIDADE

A manutenção centrada na confiabilidade teve sua origem ligada aos processos tecnológicos e sociais desenvolvidos após a Segunda Guerra Mundial. Os conceitos e técnicas da manutenção centrada em confiabilidade podem ser aplicados em quaisquer sistemas, independentemente da tecnologia empregada, pois a intenção é de que a funcionalidade dos processos se mantenha. Ao se compará-la com outras técnicas, a manutenção centrada na confiabilidade é considerada ainda em desenvolvimento, pois está constantemente incorporando novos métodos de análise,

absorvendo principalmente os conhecimentos modernos de otimização estatística (SIQUEIRA, 2005).

A confiabilidade de equipamentos e sistemas é uma das principais preocupações do gerenciamento da manutenção por concentrar os riscos à segurança operacional, ao meio ambiente e a otimização dos recursos. Na época em que as políticas de manutenção eram puramente corretivas, as falhas eram caracterizadas pela curva da banheira (Fig. 5), que é definida uma convenção onde os equipamentos se tornam mais propensos às falhas na medida em que envelhecem (LUCATELLI, 2002).

Figura 5 – Curva da banheira



Fonte: Adaptado de LUCATELLI (2002).

Os equipamentos de interesse da manutenção são todos os aqueles que formam sistemas reparáveis, ou seja, aqueles que podem ser plenamente restaurados após a perda de desempenho total ou parcial de suas funções. A confiabilidade de um sistema é a probabilidade de que este sistema opere adequadamente e sem falhas sob as condições de projeto, durante um intervalo de tempo especificado. Dentre as vantagens da manutenção centrada na confiabilidade podemos citar a preservação das funções do sistema, a identificação das falhas funcionais, classificação e priorização

das falhas funcionais de acordo com suas consequências, dentre outras. (SELLITTO, 2005).

#### 4.3. MODELAGEM EM MANUTENÇÃO

A manutenção centrada na confiabilidade produz um conjunto de atividades aplicáveis e efetivas, que são apropriadas para prevenir ou reduzir as consequências das falhas. Em condições ideais, a definição dessa frequência de manutenção resulta em uma melhoria do desempenho do processo como um todo. O uso de modelos em atividades preventivas pressupõe que essas atividades representem plenamente as características do sistema. Os modelos mais adequados são aqueles que possuem estados discretos e contínuos no tempo (a transição ocorre em uma escala contínua de tempo em um número finito e discreto de estados). Existem vários modelos e teorias que são propostos para a tomada de decisão sobre a frequência da manutenção. De acordo com Siqueira (2005), alguns dos métodos mais utilizados são:

- a) Processo markoviano: a principal característica de um processo markoviano é a capacidade de não armazenar informações, ou seja, apenas o estado atual do processo é importante para a determinação de estados futuros, isto simplifica a análise e a predição do comportamento futuro de uma cadeia de Markov, com suas probabilidades e valores esperados que quantifiquem esse comportamento.
- b) Exploração da idade: é usada normalmente quando há a presença de novos itens e a total ausência de mecanismos de falha. Neste caso usa-se um mecanismo empírico de falha baseando na tentativa e erro. Inicialmente determina-se um intervalo qualquer de manutenção como base, com os defeitos desse primeiro nível encontrados, se os níveis forem reduzidos, aumenta-se o intervalo de manutenção do próximo item da população.
- c) Teoria dos jogos: utilizada para se determinar o modelo da periodicidade da manutenção, juntamente com a análise dos riscos de adiamento dessa manutenção. De um modo geral pode ser descrita como um duelo, onde quanto menor a distância entre os oponentes, maiores são as chances de acertar, mas

por outro lado as chances de ser acertado também são maiores, de modo mais simples pode-se dizer que acertar o oponente significa encontrar e corrigir o defeito e ser acertado significa o defeito danificar o equipamento.

Outro tipo de modelo também usado é a árvore de eventos, que é uma maneira de se abordar o problema da periodicidade da manutenção de acordo com a evolução dos acontecimentos que dão origem a uma falha ou acidente. Ela mostra os possíveis resultados dessa ocorrência, levando-se em consideração os fatores condicionantes do acontecimento dessas falhas, de acordo com suas chances de sucesso ou fracasso (GUEVARA CARAZAS, 2011).

E por fim, as funções de confiabilidade dizem respeito a variáveis aleatórias, também conhecidas como estocásticas. Essas variáveis não possuem valores constantes: eles variam sob a influência de diversos fatores casuais. Conhecer uma variável aleatória não é conhecer seu valor numérico nem enumerar seus valores possíveis. Conhecer uma variável aleatória é conhecer as probabilidades de que cada resultado possível da saída de um experimento associado tenha um determinado valor (SELLITTO, 2005).

## 5. LINHA FÉRREA

A via permanente é um dos constituintes da linha férrea e pode ser definida como a parte que está sujeita à ação de desgaste das rodas dos veículos. Ela é feita para que possa ser renovada quando o seu desgaste atingir o limite de tolerância exigido pela segurança ou comodidade da circulação, e ser substituída em seus principais constituintes, quando assim o exigir a intensidade do tráfego ou o aumento de peso do material rodante. Esta parte da linha, por sua vez se divide em: infraestrutura e superestrutura. A infraestrutura é constituída pela terraplanagem (cortes, aterros), pontes, viadutos, túneis, elementos de drenagem e todos os elementos que dão suporte a via. Enquanto a superestrutura é representada pela linha propriamente dita, ou seja, é a parte da ferrovia que sofre o contato direto dos trens e sofre desgaste devido ao trafego. Ela é composta por grades (trilhos, dormentes, fixações) e aparelhos de mudança de via, ou simplesmente AMV (BRINA, 1983).

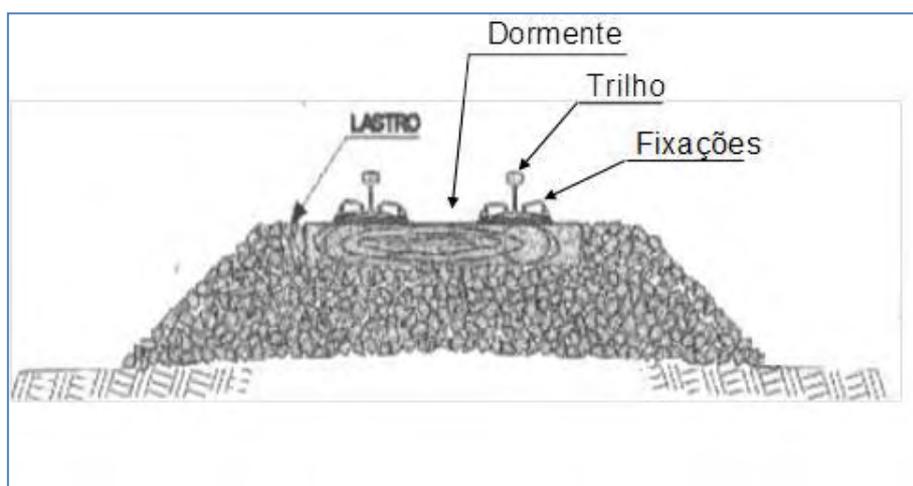
### 5.1. ELEMENTOS DA SUPERESTRUTURA

A superestrutura é uma estrutura contínua, mas para facilitar a localização dos problemas, rastreabilidade e controle, é dividida em equipamentos, de acordo com Stopatto (1987), como visto na Fig. 6:

- a) AMV (Aparelhos de Mudança de Via) – com características especiais, diferentes do resto da via, é um equipamento visível e fácil de distinguir dos demais (Fig. 7). Sua função é permitir que os veículos possam mudar de linha ao longo do trajeto.
- b) Os trilhos são elementos da grade ferroviária que tem a função de receber os esforços provenientes do material rodante e servem de guia para os frisos dos veículos ferroviários. Ficam dispostos em duas filas paralelas separados por uma distancia constante conhecida como bitola.
- c) Os dormentes são responsáveis por receber os esforços transmitidos pelos trilhos e garantem a bitola da via. São fixados aos trilhos por grampos e placas de apoio.

- d) O lastro é o componente que recebe os esforços provenientes dos dormentes e os transmite a infraestrutura ferroviária. Além disso, ele também é o elemento que garante a correta geometria da via, uma vez que a grade é acoplada nele. Outra característica importante do lastro é a permeabilidade, cuja função é proteger os dormentes do contato prolongado com a umidade do solo e outras variáveis climáticas.

Figura 6 - Componentes da superestrutura



Fonte: Adaptado de Stopatto (1987).

Figura 7 - Aparelho de mudança de via (AMV)



Fonte: VALE S.A. (2011).

## 5.2. INSPEÇÃO DA LINHA

O processo de inspeção na linha consiste na verificação das condições para que se possa garantir simultaneamente o aumento da segurança e da velocidade às composições ferroviárias. A inspeção pode ser visual, quando é realizada de acordo com a sensibilidade do operador, ou sistemática, onde são executadas várias medidas nos diversos pontos da linha. De acordo com Carvalho (1973), os principais componentes inspecionados são:

- a) Dormentes: é obtido o percentual de dormentes podres (Fig.8) da ferrovia que é obtido através da contagem dos dormentes defeituosos. Esta contagem é realizada *in loco*, onde inspetores avaliam um a um, as condições dos dormentes na linha. O resultado destas inspeções é comparado com o número total de dormentes do segmento analisado, com a intenção de se determinar quais dormentes devem ser reparados. Com os valores do número de dormentes danificados, as prioridades de substituição desses equipamentos podem ser definidas. Este índice é importante para a manutenção, uma vez que a má conservação dos dormentes pode levar a sérias consequências, como abertura de bitola, perda da estabilidade da via e acidentes ferroviários.

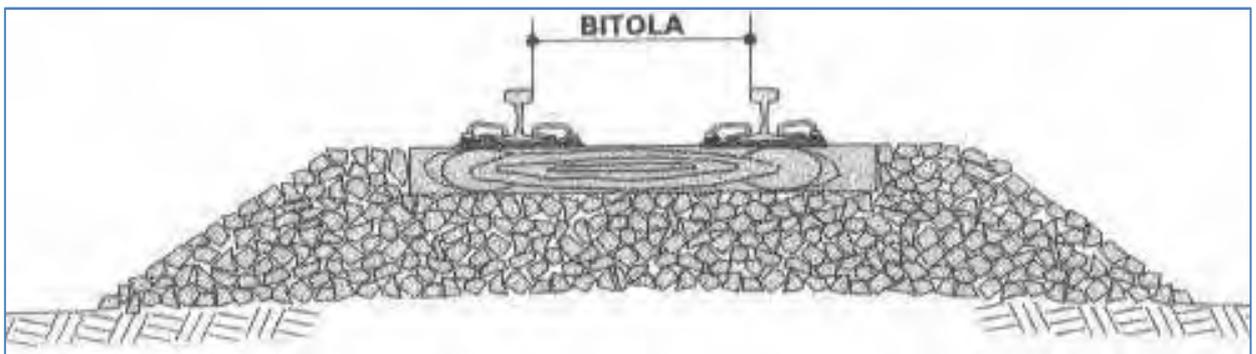
Figura 8 - Dormente podre na via.



Fonte: VALE S.A. (2011).

- b) Bitola: é a distância entre as faces internas das duas filas de trilhos (Fig.9). Este é o principal critério de manutenção da via permanente, pois caso este item esteja fora do estado ideal de funcionamento, podem acontecer sérios acidentes, além de comprometer a vida útil dos trilhos e dormentes e aumento do consumo de combustível dos equipamentos de tração ferroviária. A inspeção da bitola na via pode ser feita de forma mecanizada ou manual. Quando feita de modo manual é utilizada uma régua apropriada para se obter precisão.

Figura 9 - Bitola da via.



Fonte: Adaptado de Stopatto (1987).

- c) Trilhos: o percentual de desgaste dos trilhos é uma medida obtida através da comparação de um trilho com certa utilização com a área de um trilho novo. Para se determinar este desgaste são realizadas medidas em campo com um equipamento específico para a obtenção do percentual de desgaste. Este problema ocorre principalmente em curvas e acontece com um intervalo de tempo menor nas curvas cujo raio é menor. O desgaste acentuado dos trilhos leva ao aumento na bitola da via e risco de fratura dos trilhos. O percentual de achatamento dos trilhos é uma medida obtida por meio da comparação das dimensões de um trilho em serviço com as dimensões de um trilho novo. No achatamento dos trilhos a ação mais significativa é da componente vertical, sendo descartada a componente lateral, pois não afeta o processo em si. As medidas são realizadas no campo por equipamento específico. Tal medida não

deve ultrapassar os 30% de achatamento do trilho usado em relação ao trilho novo.

- d) AMV: A inspeção de AMV consiste na medição dos pontos identificados como defeituosos no equipamento. Estas medidas, tais como o desgaste e achatamento dos trilhos de encosto, dependem do tipo da bitola da via nos AMV's e variam conforme a ferrovia. São de suma importância para a segurança da circulação dos trens ferroviários, pois desvios de poucos milímetros podem levar a descarrilamentos.

### 5.3. REPAROS

Os serviços de reparo da linha são divididos por componente, de acordo com seus respectivos defeitos, conforme Brina (1983):

- a) Trilhos: substituição (troca de um trilho com problema por outro novo ou em melhor estado), substituição das juntas isoladas, alívio de tensões térmicas (garantir que tensões nos trilhos provocadas por variações térmicas sejam as mínimas possíveis) e eliminação das juntas (Fig.10: soldagem das juntas existentes na ferrovia para garantir o trilho continuamente soldado – TCS).

Figura 10 - Processo de soldagem para eliminação de juntas.



- b) Dormentes: substituição (trocar dormentes podres por dormentes novos, podendo haver mudança no tipo do dormente, como madeira por aço ou vice versa) e quadramento (manter a distancia uniforme entre os dormentes, deslocando àqueles que tenham saído de sua posição original).
- c) Demais componentes: manutenção de fixações (substituição de chapas de apoio e correções de bitola), reparo do lastro (desguarnecimento de lastro, recomposição de lastro) e AMV (substituições de componentes metálicos, dormentes especiais, correção e manutenção do lastro do AMV).

#### 5.4. ESTRADA DE FERRO VITÓRIA A MINAS – EFVM

A Estrada de Ferro Vitória a Minas (EFVM) (Fig. 11) foi criada em 1904, e é considerada uma das mais eficientes e produtivas do país, liga a região produtora de minério do estado de Minas Gerais ao complexo portuário de Tubarão, no estado do Espírito Santo. Possui 905 km de extensão e transporta cerca de 40% de toda a carga ferroviária brasileira. O principal produto transportado é o minério de ferro, que corresponde a 80% de tudo o que é transportado. É administrada desde 1997 pela Companhia Vale do Rio Doce, que obtém a concessão para a exploração dos serviços de transporte de cargas e de passageiros (CNT, 2011).

Figura 11 - Mapa geral da EFVM



. Fonte: CNT. (2011).

As tabelas 1 e 2 seguintes mostram o evolutivo da carga transportada em tonelada por km útil e a quantidade de acidentes respectivamente. Pode-se comprovar a importância da operação da EFVM, pois além de possuir uma grande quantidade de carga transportada apresenta um baixo índice de acidentes quando comparada com as outras ferrovias,

Tabela 1 - Carga transportada pelas ferrovias nacionais

Carga transportada em tonelada por km útil										
Concessionárias	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012*
ALLMN	2.103	2.259	7.957	7.446	9.394	11.297	13.887	14.618	16.073	8.848
ALLMO	1.232	1.191	1.312	1.431	1.202	1.345	1.312	1.783	1.760	866
ALLMP	9.221	9.473	2.286	2.240	1.920	3.054	3.019	4.004	4.689	1.838
ALLMS	13.850	14.175	15.415	18.423	17.147	17.378	17.196	17.474	18.121	7.651
EFC	52.411	63.622	69.525	76.691	83.334	87.513	83.945	91.044	98.897	46.851
FERROESTE	406	323	349	1.005	620	747	469	273	209	112
<b>EFVM</b>	<b>60.487</b>	<b>64.773</b>	<b>68.648</b>	<b>73.442</b>	<b>75.511</b>	<b>72.783</b>	<b>57.929</b>	<b>73.480</b>	<b>74.554</b>	<b>35.461</b>
FCA	7.477	9.523	10.712	9.143	14.231	15.060	14.198	15.320	13.606	8.325
FNS	—	—	—	—	—	1.026	1.155	1.524	1.884	996
FTC	152	169	170	183	189	213	202	185	173	98
MRS	34.515	39.355	44.445	47.662	52.590	55.621	51.273	57.490	61.259	31.032
TLSA	790	848	814	678	963	920	730	728	681	376
Total	182.644	205.711	221.633	238.344	257.101	266.957	245.315	277.923	291.906	142.454

\*Dados até Junho de 2012

(Fonte: ANTT, 2012).

Tabela 2 – Acidentes nas ferrovias

Acidentes por milhão de trens e km										
Concessionárias	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012*
ALLMN	7,00	10,00	84,60	68,49	19,95	11,80	21,56	10,62	5,81	5,22
ALLMO	151,00	188,00	302,30	261,35	67,53	46,66	27,20	26,42	23,69	30,19
ALLMP	27,00	33,00	24,80	33,67	26,05	26,90	27,51	23,58	20,51	28,09
ALLMS	19,00	15,00	16,20	13,82	10,35	15,60	16,96	17,40	15,70	17,07
EFC	6,00	6,00	5,20	7,43	4,39	6,62	5,50	4,98	4,20	2,91
<b>EFVM</b>	<b>15,00</b>	<b>10,00</b>	<b>10,50</b>	<b>5,78</b>	<b>4,07</b>	<b>2,84</b>	<b>3,82</b>	<b>3,56</b>	<b>2,82</b>	<b>1,86</b>
FERROESTE	3,00	5,00	0,00	0,00	5,91	22,81	4,02	4,47	4,95	8,68
FCA	67,00	43,00	26,20	20,54	17,89	18,06	22,76	23,86	24,13	22,20
FNS	-	-	-	-	-	14,07	2,71	8,13	15,38	0,00
FTC	13,00	13,00	10,00	10,16	15,27	10,00	11,51	9,66	13,78	0,00
MRS	23,00	16,00	8,50	6,98	6,33	4,26	4,60	5,41	7,58	10,55
TLSA	307,00	354,00	328,10	149,72	114,46	174,12	167,82	196,65	196,43	104,04
Total	638	693	816	578	292	354	316	335	335	231

\*Dados até Junho de 2012

(Fonte: ANTT, 2012).

## 6. MÉTODO

O método de investigação proposto consiste na adequação e uso dos dados relativos aos tempos de reparo da linha férrea para a elaboração de um modelo de distribuição de probabilidade de Lei de Potência. Os dados devem atender às principais características desse tipo de distribuição, como a presença de pelo menos um dos momentos, representação gráfica linear e presença do expoente de escala.

Primeiramente foi delimitado o trecho da ferrovia que será estudado. Isso foi feito com base na facilidade de obtenção dos dados e importância do trecho. Com a definição do local a ser analisado, obteve-se os respectivos valores dos tempos de reparo para finalmente verificar a adequação dos mesmos ao modelo de distribuição de probabilidade de Lei de Potência.

O trecho estudado é o que se encontra sob Supervisão de Costa Lacerda, tem início na região de João Monlevade e finda no município de Congonhas, mais exatamente na mina de Fábrica. O Ramal de Fábrica se refere ao trecho que liga a linha tronco, na altura da cidade de Nova Era - MG, a mina de Fábrica que fica próximo ao município de Congonhas. Este trecho possui 168 km de linhas, onde 62 km são compostos por linha dupla e o restante por linha singela. Neste ramal estão localizadas as minas de Água Limpa (Rio Piracicaba), Fazendão, Alegria (Mariana), Timbopeba, Fábrica (Ouro Preto), é responsável pelo transporte principalmente de insumos e de produtos siderúrgicos, o que pode ser visto na Fig.12 seguinte:

Figura 12 - Ramal de Fábrica.



Fonte: Adaptado de VALE S.A. (2011).

Para a obtenção dos dados foi utilizado um relatório de falhas (Tabela 3) ocorridas entre Janeiro de 2010 à Junho de 2012, onde para cada falha que gera uma indisponibilidade na via, seja por meio de ocupação indevida ou detectada automaticamente pelo painel no Centro de Controle Operacional (CCO), é gerada uma OS (Ordem de Serviço) e um número de ocorrência do CCM (Centro de Controle e Manutenção). A falha é considerada somente após a averiguação da mesma. Caso seja considerada ela é então registrada, solucionada e encaminhada para o banco de dados de falhas.

Tabela 3 – Base de dados dos tempos de reparo (Jan/2010 à Jun/2012).

DESCRIÇÃO FALHA	CAUSA	COMPONENTE	MÊS	ANO	DATA/HORA FALHA	DATA/HORA AÇIONAMENTO	DATA/HORA CHEGADA	DATA/HORA FIM	Tempo total
Ocupação	Curto Devido A Limalha	Junta Isolante	1	2012	01/01/2012 06:45	01/01/2012 06:50	01/01/2012 07:22	01/01/2012 07:45	1:00:00
Deslizamento/Barreira	Chuva em Excesso	Talude	1	2012	01/01/2012 13:33	01/01/2012 13:40	01/01/2012 15:25	01/01/2012 19:30	5:57:00
Defeito em Acessórios	Couce De Lança Sem Parafuso	Agulha	1	2012	01/01/2012 16:20	01/01/2012 16:25	01/01/2012 18:00	01/01/2012 18:20	2:00:00
Erosão	Chuva em Excesso	Plataforma	1	2012	01/01/2012 23:50	01/01/2012 23:58			*
Deslizamento/Barreira	Chuva em Excesso	Talude	1	2012	02/01/2012 00:20	02/01/2012 00:35			*
Deslizamento/Barreira	Chuva em Excesso	Talude	1	2012	02/01/2012 05:15	02/01/2012 05:23			*
Ocupação	Tala Em Curto	Junta Isolante	1	2012	02/01/2012 06:45	02/01/2012 07:00	03/01/2012 07:00	03/01/2012 07:30	24:45:00
Deslizamento/Barreira	Chuva em Excesso	Talude	1	2012	02/01/2012 07:10	02/01/2012 07:13			*
Erosão	Chuva em Excesso	Plataforma	1	2012	02/01/2012 11:38	02/01/2012 11:57			*
Ocupação	Tala Em Curto	Junta Isolante	1	2012	02/01/2012 12:00	02/01/2012 12:10	02/01/2012 16:50	02/01/2012 17:43	5:43:00
Pancada	Jacaré Com Desgaste Excessivo	care simples Arer	1	2012	02/01/2012 12:40	02/01/2012 12:45	02/01/2012 15:15		*
Ávore Na Linha	Chuva em Excesso	Plataforma	1	2012	02/01/2012 12:50	02/01/2012 12:51	02/01/2012 14:20	02/01/2012 14:30	1:40:00

Os dados utilizados para a elaboração do modelo foram retirados da coluna de tempo total da tabela 3, onde este tempo é dado pela diferença entre a data e hora da falha registrada e a data e hora do fim do reparo. Nota-se que alguns valores não

apresentaram o tempo total, pois os mesmos não haviam sido reparados até o momento de extração dos dados. Foram desconsideradas as falhas que apresentaram um tempo de reparo nulo ou iguais a 0.

Após o tratamento dos dados foram então realizados os testes de adequação dos mesmos às distribuições de probabilidade contínuas descritas anteriormente. A verificação da Lei de Potência foi feita de maneira gráfica e por intermédio da interpretação dos parâmetros estimados. Para os testes de adequação e estimação de parâmetros foram utilizados os softwares MINITAB16 e MATLAB 2009 respectivamente. Depois de determinados os parâmetros da Lei de Potência ( $\hat{\alpha}$  e  $\hat{x}_{min}$ ) foi definida então, a função densidade de probabilidade dos tempos de reparo da linha férrea estudada.

Para se chegar à decisão de aceitar ou não os dados como pertencentes à distribuição de Lei de Potência foi realizado o teste proposto por Clauset *et al.*, (2009), que tem como base o teste de Kolmogorv-Smirnov, onde:

*H<sub>0</sub>: Os dados seguem a distribuição testada*

*H<sub>1</sub>: Os dados não seguem a distribuição testada*

O nível de significância  $\alpha$  para o teste é de  $0,1$ . O teste é realizado de maneira ajustada, pois os parâmetros da distribuição são calculados de acordo com os dados da amostra. Deste modo para valores de  $p$  (*coeficiente de Pearson*) menores do que  $0,1$  (*nível de significância do teste*) existe pouca evidência da adequação à distribuição testada, alternativamente se o valor do coeficiente  $p$  for maior do que  $0,1$  pode-se dizer que os dados se ajustam a distribuição.

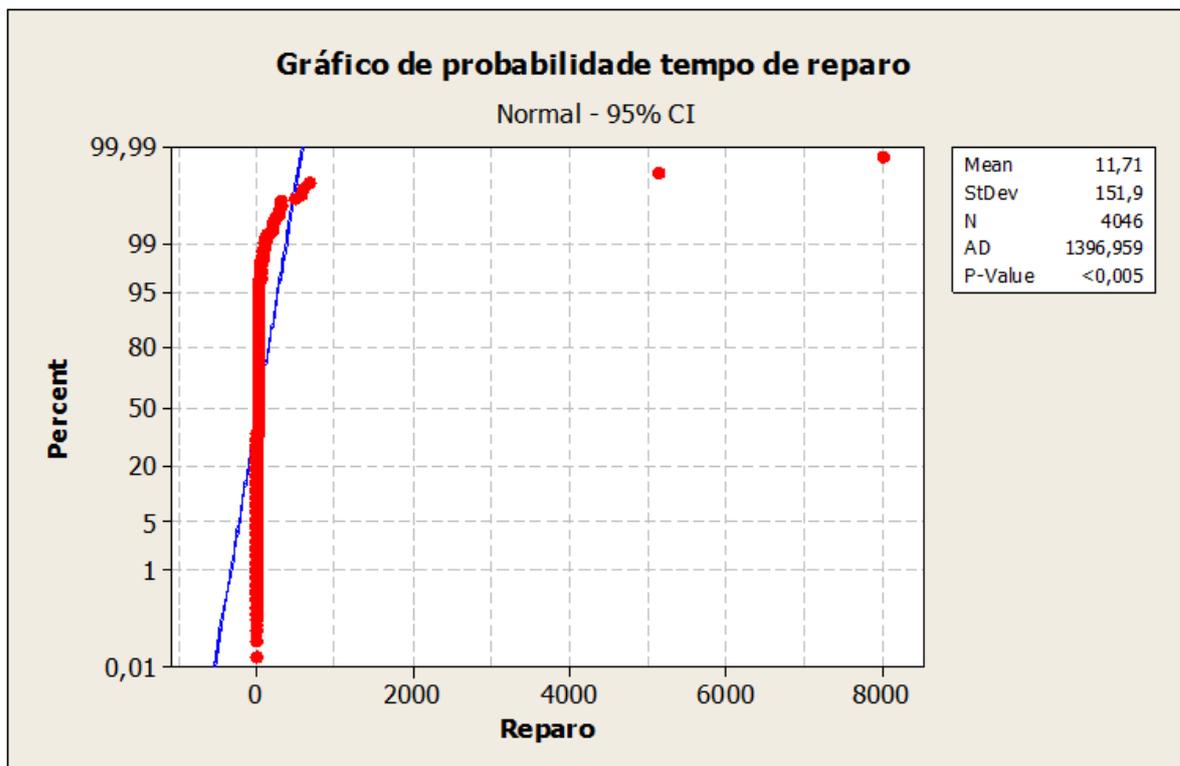
Com a função densidade de probabilidade e os respectivos parâmetros calculados, resta então a confirmação de que a distribuição realmente é resultante de uma Lei de Potência. Isso é feito por meio do cálculo ajustado de  $p$ , já que os parâmetros são provenientes da amostra.

## 7. RESULTADOS

Depois de delimitado o trecho a ser estudado e feito o tratamento dos dados, chegou-se a uma quantidade de 4046 (de um total de 4869) tempos de reparo ou falhas ocorridas no intervalo de tempo determinado (Janeiro de 2010 à Junho de 2012).

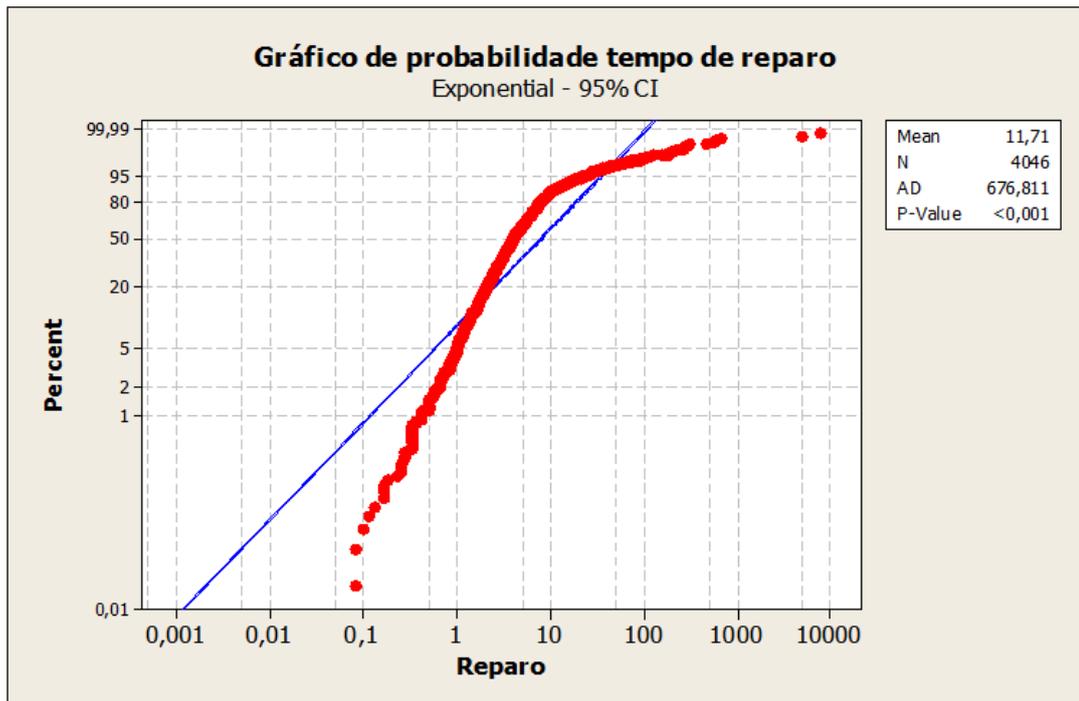
Os gráficos representados pelas figuras seguintes mostram os testes realizados para as distribuições normal, exponencial, lognormal e Weibull respectivamente. Nota-se que os dados não se adequaram a nenhuma das distribuições. Apresentando um valor de  $p$  abaixo de 0,05 em todas as situações, ou seja, rejeita-se a hipótese nula dos dados se adequarem aos modelos testados Fig. 13 à Fig.16.

Figura 13 – Teste de adequação dos dados à distribuição Normal.



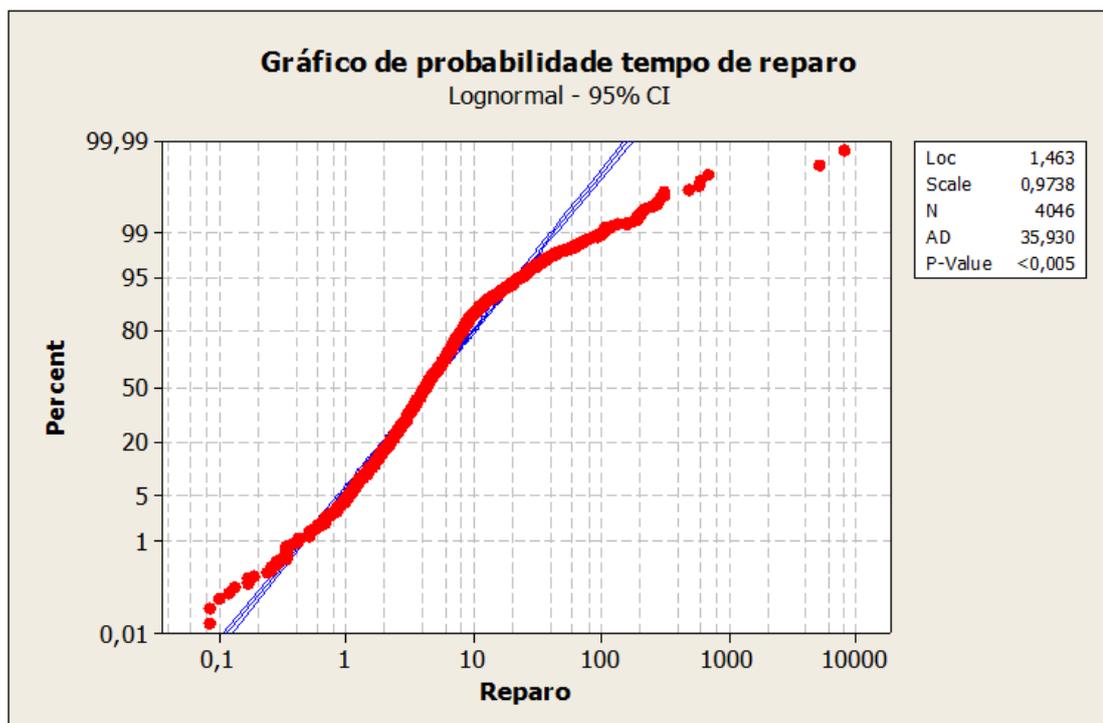
Fonte: Autor.

Figura 14 – Teste de adequação dos dados à distribuição Exponencial.



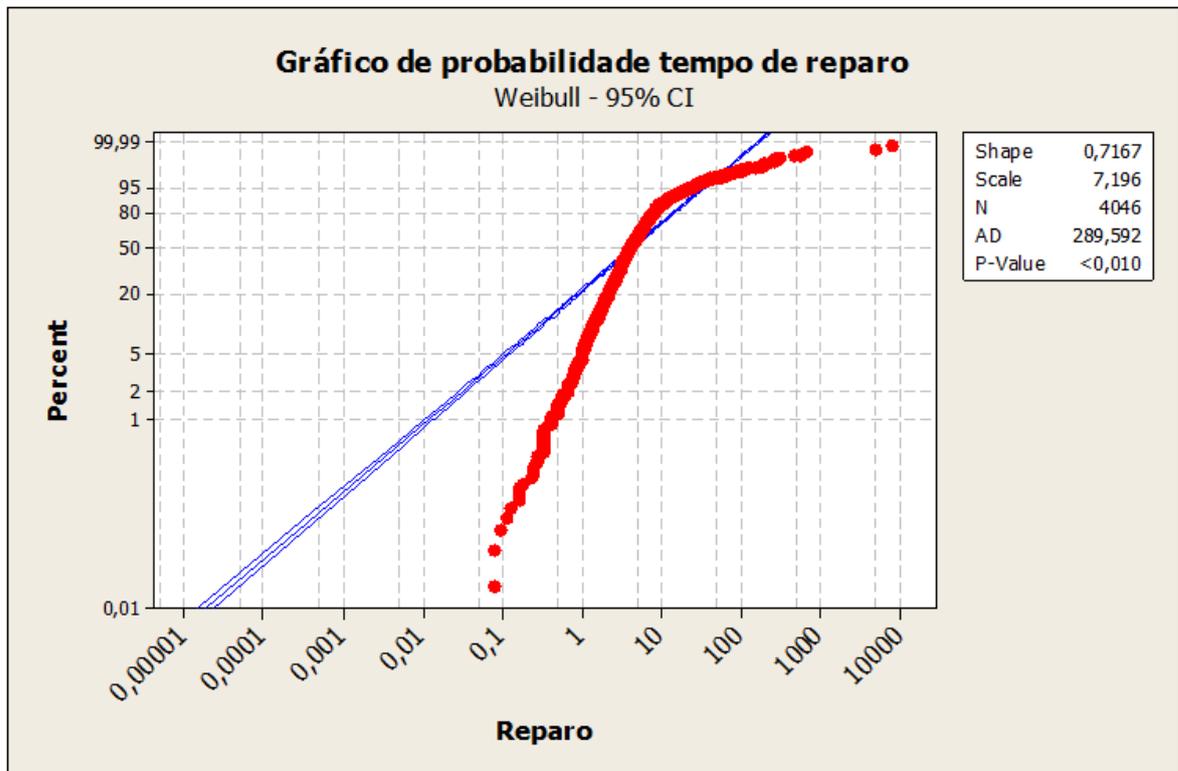
Fonte: Autor.

Figura 15 – Teste de adequação dos dados à distribuição Lognormal.



Fonte: Autor.

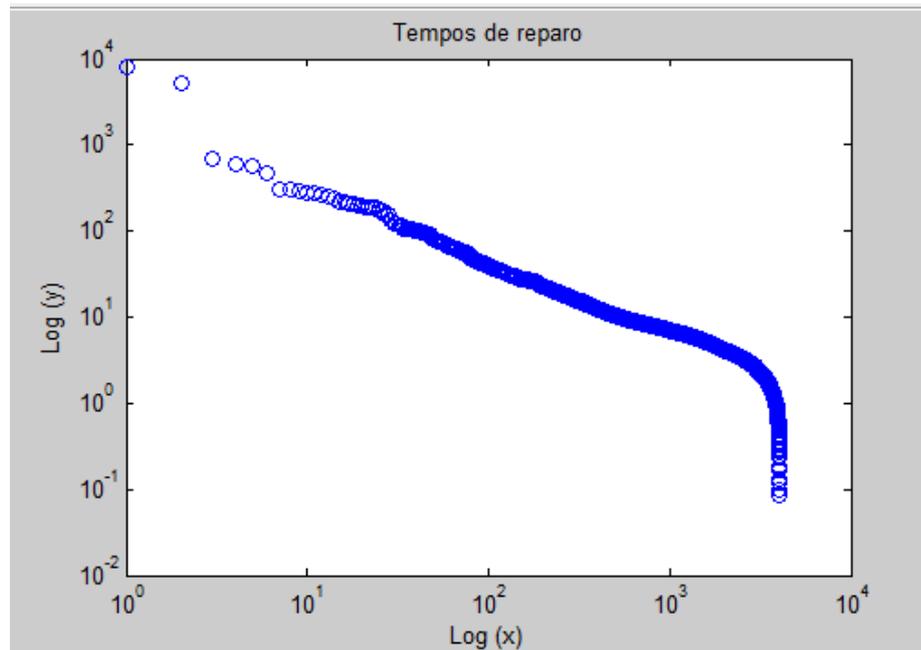
Figura 16 – Teste de adequação dos dados à distribuição de Weibull.



Fonte: Autor

Após a comprovação de que os dados não se adequam a nenhuma das distribuições testadas, iniciou-se o procedimento de verificação da distribuição à Lei de Potência. O primeiro procedimento é a representação gráfica log-log dos dados, com a intenção de se obter uma reta, comportamento característico das Leis de Potência. Isso pode ser visto na Fig. 17.

Figura 17 – Gráfico Log-Log dos tempos de reparo.



Fonte: Autor.

É necessário retirar os menores valores que se encontram abaixo de  $x_{min}$ , com a finalidade de se determinar para quais valores a partir deste ponto a distribuição segue uma Lei de Potência. O valor estimado de  $x_{min}$  foi de 10,7 horas, ou de 10 horas e 42 minutos. O expoente escala encontrado foi  $\hat{\alpha} = 2,16$ , valor este que satisfaz a condição de  $2 < \alpha < 3$  que é característica da Lei de Potência, com um erro padrão de  $\pm 0,02$ , assim o valor do expoente de escala está entre 2,14 e 2,18.

Após a determinação dos parâmetros  $\hat{\alpha}$  e  $\hat{x}_{min}$  pode-se então calcular a expressão da distribuição de Lei de Potência, de acordo com a Eq. (3.8). Assim ao se substituir os valores tem-se a Eq. (7.1):

$$p(x) = \frac{2,16 - 1}{10,7} \left( \frac{x}{10,7} \right)^{-2,16} = 0,11(0,09x)^{-2,16} . \quad 7.1$$

O valor do primeiro momento pode então ser calculado através da Eq. (3.9):

$$\langle x \rangle = \frac{2,16 - 1}{2,16 - 2} 10,7 = 77,58. \quad 7.2$$

A média dos maiores valores, isto é, aqueles que podem ser encontrados ao se retirar uma amostra qualquer dos dados é calculada de acordo com a Eq.(3.18), vale notar que quanto maior a quantidade de valores observados ( $n$ ), maiores serão os valores encontrados:

$$\langle x_{max} \rangle \sim n^{1/(\alpha-1)} = 4046^{1/(2,16-1)} = 1264,84 \quad 7.3$$

Para a confirmação dos dados à distribuição foi realizado o seguinte teste de hipótese:

$H_0$ : *Os dados seguem a distribuição de Lei de Potência*

$H_1$ : *Os dados não seguem a distribuição de Lei de Potência*

Como os valores dos parâmetros já são conhecidos pode-se então calcular o valor de  $p$ , de acordo com as hipóteses definidas anteriormente. Para os dados dos tempos de reparo o valor encontrado foi de 0,57. Este resultado determina a aceitação da hipótese nula, onde dados apresentam um comportamento característico das Leis de Potência.

No âmbito da manutenção os resultados encontrados permitem dizer que as maiores indisponibilidades do sistema (representada pelas falhas) apresentam uma distribuição de Lei de Potência, com um tempo mínimo estimado de 10,7 horas. Ou seja, para os tempos superiores a este valor mínimo, as chances do componente ser reparado em um dado tempo  $x$  é definida de acordo com função densidade de probabilidade encontrada.

O conhecimento prévio (por meio de simulações e testes) do provável tempo de reparo gasto é importante para a manutenção corretiva, preditiva e preventiva. Para a primeira, verificam-se quais os tempos que estão fora do esperado para então realizar a adequação dos procedimentos e processos de reparo, buscando uma

melhoria e redução do tempo gasto. Para a preventiva e preditiva possibilita a criação de cenários de indisponibilidade, onde é possível determinar com antecedência os impactos das grandes paradas (programadas ou não) no sistema ferroviário.

Uma grande vantagem deste tipo de distribuição de probabilidade em relação às demais é que, para o mesmo sistema, ou seja, para outros trechos da linha com características semelhantes, pode-se utilizar estes resultados para encontrar os respectivos tempos de reparo, não sendo necessário um novo cálculo dos parâmetros da função densidade de probabilidade. Isso se deve à propriedade que a Lei de Potência tem de ser livre de escala.

## 8. CONCLUSÃO E CONSIDERAÇÕES

O presente trabalho apresentou inicialmente algumas das principais distribuições de probabilidade contínuas que, de alguma forma podem ser relacionadas com a distribuição de Leis de Potência ou com os tempos de reparo. Essas distribuições são amplamente utilizadas para a construção de modelos probabilísticos com a intenção de descrever as chances de acontecimento de certos eventos.

A distribuição de Lei de Potência foi estudada à parte por se tratar do foco do trabalho. Primeiramente foram levantadas suas origens e algumas de suas aplicações. Na sequência foi feita uma exploração matemática das suas características. Primeiramente foi levantada a expressão referente à função densidade de probabilidade com o valor da constante C normalizada. Foram determinados os momentos e algumas das principais propriedades como os grandes valores e a invariância de escala. Foram também mostradas as mesmas definições para dados discretos. É importante ressaltar que para uma amostra suficientemente grande as propriedades e definições não apresentam significativas diferenças entre os tipos de dados.

Dentro do contexto do gerenciamento da manutenção foram determinadas as características das principais práticas de manutenção, como a preditiva, preventiva e a corretiva. Foram também abordados os principais elementos da manutenção centrada na confiabilidade, que são definidos por técnicas probabilísticas de modelagem de sistemas.

O gerenciamento da manutenção foi visto aplicado à estrutura ferroviária, onde foram abordadas suas características e principais nomenclaturas do sistema. Foram determinados os tempos de reparo de cada falha causadora de indisponibilidade no sistema, tempos estes calculados como a diferença entre o instante em que a falha é descoberta até o instante em que o sistema volta a operar novamente. O trecho estudado faz parte de uma linha que corresponde a cerca de 40% da carga ferroviária no país, transportando tanto produtos como passageiros, sendo de fundamental importância econômica principalmente para mineradoras e siderúrgicas.

Sob o ponto de vista teórico os resultados encontrados permitem afirmar que os dados referentes aos tempos de reparo de uma ferrovia seguem a distribuição de Lei de Potência, apresentando características condizentes com este tipo de distribuição. Pode-se afirmar graficamente por meio da Fig. 12, e matematicamente por meio das propriedades dos parâmetros encontrados. Foi elaborada a função densidade de probabilidade e testada à hipótese dos dados realmente seguirem uma distribuição de Lei de Potência. Como o valor do  $p$  encontrado foi superior ao nível de significância, a hipótese nula foi aceita, confirmando a adequação dos dados à Lei de Potência.

A contribuição prática do trabalho está ligada diretamente ao gerenciamento e planejamento da manutenção, pois ao se obter um modelo probabilístico que determina as chances dos tempos gastos para o reparo da linha, ou de outro modo, a duração das indisponibilidades, é possível determinar antecipadamente os impactos dessas falhas no negócio. Ao se conhecer os prováveis cenários gerados por estas falhas, ações de melhoria e contorno podem ser tomadas de modo mais efetivo, minimizando os efeitos das paradas. Isso é válido para os 3 tipos de manutenção (corretiva, preventiva e preditiva).

Uma manutenção bem programada e executada pode contribuir para a redução dos custos de transporte e minimizar os impactos que uma parada pode provocar em todo o sistema logístico. No atual contexto ferroviário brasileiro, que apresenta somente 20% da carga transportada e que conta com apenas 30.000 km de linhas (sendo estas antigas e com bitolas diferentes) o bom planejamento da manutenção colabora para que o escoamento da produção possa ser realizado de maneira eficiente, contornando os problemas apresentados e aumentando a produtividade da linha.

## 9. REFERÊNCIAS

ANTT (2012). Evolução do transporte ferroviário, Brasília, DF. Disponível em: <[http://www.antt.gov.br/index.php/content/view/15884/Evolucao\\_do\\_Transporte\\_Ferrovuario.html](http://www.antt.gov.br/index.php/content/view/15884/Evolucao_do_Transporte_Ferrovuario.html)>. Acesso em: 15 ago. 2012.

BOCCARA, N. **Modeling complex systems**. New York: Springer-Verlag, 2004.

BRANCO FILHO, G. **Dicionário de Termos de Manutenção, Confiabilidade e Qualidade**. 1ª. Ed. Rio de Janeiro: Editora Ciência Moderna, 2004.

BRINA, H. L. **Estradas de Ferro**. 1ª. Ed. Rio de Janeiro: LTC, 1983.

GUEVARA CARAZAS, F. J. **Decisões baseadas em riscos: método aplicado na indústria de geração de energia para a seleção de equipamentos críticos e políticas de manutenção**. 2011. Tese (Doutorado em Engenharia) – Programa de Pós-Graduação em Engenharia de Mecânica – Escola Politécnica. USP, São Paulo.

CARVALHO, M. P. **Curso de Estradas: Estudos, projetos e Locação de Ferrovias e Rodovias**. Rio de Janeiro: Editora Cientifica, 1973.

CLEMENTI, F.; DI MATTEO, T.; GALLEGATI, M. The Power-law tail exponent of income distributions. **Physica A: Statistical and Theoretical Physics**, v.370, n.1, p. 49-53, 2006.

CLAUSET, A.; SHALIZI, C. R.; NEWMAN, M.E.J. Power-law distributions in empirical data. **SIAM Review**, v.4, n.51, p. 661-703, 2009.

CNT (2011). Pesquisa CNT de ferrovias 2011, Brasília, DF. Disponível em <[http://www.cnt.org.br/Paginas/Pesquisas\\_Detalhes.aspx?p=7](http://www.cnt.org.br/Paginas/Pesquisas_Detalhes.aspx?p=7)>. Acesso em: 09 abr 2012.

CRESSONI FILHO, L. **Estudo de lei de potência em distribuições estatísticas de consumo de água**. 2002. Dissertação (Mestrado em Física Aplicada) – Programa de Pós-Graduação em Física, UNESP, Rio Claro.

DANTAS, C. A. B. **Probabilidade: um curso introdutório**. 3ª Ed. São Paulo: Editora da Universidade de São Paulo, 2008.

FUENTES, F. F. E. **Metodologia para inovação da gestão da manutenção industrial**. 2006. Tese (Doutorado em Engenharia de Mecânica) – Programa de Pós-Graduação em Engenharia de Mecânica. UFSC, Florianópolis.

GIL, A.C. **Como Elaborar Projetos de Pesquisa**. 4ª. Ed. São Paulo: Atlas, 2006.

GOLDSTEIN, M.; MORRIS, S.; YEN, G. Problems with fitting to the power-law distribution. **The European Physical Journal B - Condensed Matter and Complex Systems.**, v.41, n.2. 2004.

GLERIA, I. ; MATSUSHITA, R.; SILVA, S. Sistemas complexos, criticalidade e leis de potência. **Revista brasileira de ensino de física**, v.26, n.2, p. 29-108, 2004.

KARDEC, A. N. J. **Manutenção Função Estratégica**. 2ª. Ed. Rio de Janeiro: Qualitymark, 2001.

LUCATELLI, M. V. **Proposta de aplicação da manutenção centrada em confiabilidade em equipamentos médico-hospitalares**. 2002. Tese (Doutorado em Engenharia de Elétrica) – Programa de Pós-Graduação em Engenharia de Elétrica. UFSC, Florianópolis.

MANTEGNA, R. N., STANLEY, H. E. **An Introduction to Econophysics: Correlations and Complexity in Finance**. Cambridge University Press, USA, 2000.

MEYER, P. L. **Probabilidade: aplicações à estatística**. 2ª Ed. Rio de Janeiro: LTC, 1983.

MITZENMACHER, M. A brief history of generative models for Power Law and Lognormal distributions. **Internet Mathematics**, v.1, n.2, p.226-251, 2003.

MONTGOMERY, D.C.; RUNGER, G.C. **Estatística Aplicada e Probabilidade para Engenheiros**. 4ª Ed. Rio de Janeiro: LTC, 2009.

MOREIRA, D. A. **Administração da Produção e Operações**. 2º Ed. São Paulo: Cengage Learning, 2008.

NEWMAN, M.E.J. Power laws, Pareto distributions and Zipf's law. **Contemporary Physics**, v.46, p. 323–351, 2005.

PINTO, C.M.A.; MENDES LOPES, A.; TENREIRO MACHADO, J.A. **A review on Power Law in real life phenomena**. *Commun Nonlinear Sci Number Simulat*, v.17, p. 3558-35788, 2012.

RUSSO, L. E. A. **Contribuição ao processo de avaliação técnica e seleção dos componentes da grade ferroviária para a implantação em ferrovias de transporte de carga**. 2012. Dissertação (Mestrado em Engenharia) – Programa de Pós-Graduação em Engenharia de Transportes. USP, São Paulo.

SANTOS, W. B.; COLOSIMO, E. A.; MOTTA, S. B. Tempo ótimo entre manutenções preventivas para sistemas sujeitos a mais de um tipo de evento aleatório. **Gestão e Produção**, São Carlos, v. 14, n. 1, p. 193 – 202 Jan./Abr. 2007.

SCHINAIDER, S. J. **Distribuição de Lei de Potência gradualmente truncada aplicada na educação: vestibular da academia da força aérea**. 2006. Dissertação (Mestrado em Física) – Programa de Pós-Graduação em Física, UNESP, Rio Claro.

SELLITTO, M. A. Formulação estratégica da manutenção industrial com base na confiabilidade dos equipamentos. **Revista Produção**, São Paulo, v. 15, n. 1, p. 044 – 059 Jan./Abr, 2005.

SILVA, E. L., MENEZES, E. M. **Metodologia da pesquisa e elaboração de dissertação**. 4ª Ed. Florianópolis: UFSC, 2005.

SIQUEIRA, I. P. **Manutenção centrada na confiabilidade: manual de implementação**. 1ª Ed. Rio de Janeiro: Qualitymark Editora, 2005.

STOPATTO, S. **Via Permanente: Conceitos e Aplicações**. 1ª. Ed. São Paulo: T.A Queiroz, 1987.

VALE S.A. - Figuras obtidas por meio de contato direto com a empresa. Belo Horizonte, MG, 2011.

## 10. ANEXOS - ALGORITMO

### ESTIMAÇÃO DE PARÂMETROS

Para a estimação dos parâmetros da Lei de Potência dos dados referentes aos tempos de reparo foi utilizado o seguinte algoritmo adaptado de CLAUSET *et al.* ( 2009)

```
load Teste.txt;
x=(Teste);
xmins = unique(x);
xmins = xmins(1:end-1);
dat = zeros(size(xmins));
z = sort(x);
for xm=1:length(xmins)
    xmin = xmins(xm);
    z = z(z>=xmin);
    n = length(z);
    a = n ./ sum( log(z./xmin) );
    cx = (0:n-1)'./n;
    cf = 1-(xmin./z).^a;
    dat(xm) = max( abs(cf-cx) );
end;
D = min(dat);
xmin = xmins(find(dat<=D,1,'first'));
z = x(x>=xmin);
n = length(z);
alpha = 1 + n ./ sum( log(z./xmin) );
```