MÉTODO COMPUTACIONALMENTE INTENSIVO PARA O CÁLCULO DOS INTERVALOS DE CONFIANÇA PARA VALORES DA FUNÇÃO DE CONFIABILIDADE

André Luiz Emidio de Abreu
Programa de Pós Graduação em Métodos Numéricos em Engenharia – PPGMNE – UFPR
ametodosest@gmail.com

Anselmo Chaves Neto
Departamento de Estatística – UFPR
Programa de Pós Graduação em Métodos Numéricos em Engenharia – PPGMNE – UFPR
anselmo@ufpr.br

RESUMO

Este trabalho apresenta a aplicação do método computacionalmente intensivo *bootstrap* para calcular intervalos de confiança para valores da função de confiabilidade estimados pelo método de Kaplan-Meier. Tal aplicação se fez necessária, pois o estimador de Kaplan-Meier tem característica assintótica, impondo a obtenção de amostras de tempos de falha extensas, porem, para alguns tipos de produtos, por exemplo, produtos projetados para durar anos, as amostras de tempos de falha, ou são insuficientes, ou demandam muito tempo, tornando o estudo muito demorado, ou em certos casos inviável. Assim, utilizando o método *bootstrap*, corrigi-se a alta variabilidade, obtendo-se intervalos mais ajustados em torno das estimativas calculadas pelo método de Kaplan-Meier. Ao final do trabalho são apresentados os resultados, que se demonstraram satisfatórios tanto quantitativamente, tempo de processamento dos estudos, quanto qualitativamente, os valores dos intervalos *bootstrap* tiveram amplitudes menores que os intervalos calculados pelo método de Kaplan-Meier.

Palavras Chave: Método *bootstrap*, Kaplan-Meier, Intervalos de Confiança.

ABSTRACT

This paper presents the application of computationally intensive bootstrap method to calculate confidence intervals for values of the reliability function estimated by the Kaplan-Meier. This application was necessary because the Kaplan-Meier estimator has asymptotic character, requiring the obtaining of samples of failure times extensive, however, for some types of products, eg products designed to last for years, samples of failure times or are insufficient, or take a long time, making the study too long, or in some cases impossible. Thus, using the bootstrap method, to correct high variability, resulting in longer intervals around the adjusted estimates calculated by the Kaplan-Meier. At the end of the paper are presented the results, which showed satisfactory quantitatively, processing time of the studies, and qualitatively, the values of the bootstrap intervals had smaller amplitudes than the intervals calculated by the Kaplan-Meier.

Key-words: Bootstrap Method, Kaplan-Meier, Confidence Intervals.

1. INTRODUÇÃO

A análise de confiabilidade é uma das áreas da estatística que mais cresceu nas últimas duas décadas do século passado. Isso se deve ao desenvolvimento e ao aperfeiçoamento de técnicas estatísticas em conjunto com computadores cada vez mais velozes. Essa é a área da estatística que mais se destacou no período de 1979 até 1989 (Bailar III e Mostteler, 1992), tanto na área da engenharia, quanto na área da medicina onde é conhecida como análise de sobrevivência. Os dois artigos mais citados em toda literatura estatística no período de 1987 a 1989 foram, o do estimador de Kaplan-Meier para a função de confiabilidade (1958) (sobrevivência) e o modelo de Cox (1972) (Stigler, 1994). Essa área está muito relacionada com a garantia de um produto. Para auxiliar esta análise várias técnicas matemáticas e estatísticas foram criadas, um exemplo é o Estimador de Kaplan-Meier ou Estimador Limite-Produto, que estima os valores da Função de Confiabilidade $R(t)$. Esse método de estimação dos valores da função de confiabilidade vem sendo aplicado por estatísticos e engenheiros. Embora seja um estimador muito confiável, ele é assintótico, logo necessita de uma amostra grande de dados de falha que geralmente é difícil de obter. O principal problema é encontrar estimativas razoáveis, perto dos valores verdadeiros da função de confiabilidade, mesmo com amostras de falha de tamanho pequeno. É de extrema importância a utilização de técnicas estatísticas auxiliares junto aos estimadores não-paramétricos para se obter estimativas cada vez melhores e mais confiáveis, além também do ganho de tempo que isso pode proporcionar com o auxilio dos modernos computadores que existem nos dias de hoje. O objetivo principal deste trabalho é aplicando a técnica de reamostragem, computacionalmente intensiva chamada *bootstrap*, construir intervalos de confiança *bootstrap*, intervalo *bootstrap-t* que será definido nas secções a seguir, para os valores da função de confiabilidade, que são estimados pelo método de Kaplan-Meier, e estimar o erro padrão *bootstrap*, que é utilizado no calculo do intervalo *bootstrap-t*. E, ainda, comparar os intervalos *bootstrap* com o intervalo clássico. Assim, apresentar uma aplicação que se revelou eficiente no estudo da função de confiabilidade, primeiramente aplicada a produtos e sistemas, mas podendo ser ampliada para outras áreas da ciência moderna, como na Medicina, Economia, no estudo da confiabilidade estrutural, entre outros. Assim como os estimadores não-paramétricos da função de confiabilidade são todos assintóticos, necessita-se de técnicas alternativas a esses estimadores para se obter respostas mais confiáveis.

2. REVISÃO DE LITERATURA

2.1. Confiabilidade de produtos e sistemas

Nos últimos anos houve uma ênfase muito grande na Gestão de Qualidade Total, mostrando que a qualidade é tarefa de todos os setores da indústria. Houve tempos em que nas empresas a qualidade era apenas tarefa do departamento do Controle de Qualidade. Nos dias de hoje, sob o guarda-chuva da Gestão da Qualidade Total, são incluídas as atitudes da alta liderança, do planejamento estratégico, da gestão de recursos humanos, da gestão dos processos, dos próprios resultados obtidos do negócio e da satisfação do cliente.

2.2. Aplicação da confiabilidade no desenvolvimento de produtos

Num ambiente de altíssima competição, o que é hoje muito comum nos mercados, é importante que as empresas sejam capazes de determinar e controlar a confiabilidade dos seus produtos. A confiabilidade dos produtos pode ser definida como a capacidade que estes produtos têm de desempenhar suas atividades por certo período de tempo em determinadas condições de uso. Sabendo-se o nível de confiabilidade dos produtos, pode-se saber de antemão se as expectativas dos clientes quanto a este produto serão atingidas, bem como o nível de qualidade do mesmo. Nesta definição, produtos podem ser considerados componentes, subsistemas ou sistemas que constituem um produto ou serviço (Hoyland e Rausand, 1994).

2.3. Definição de Confiabilidade

Conceito de Confiabilidade: Confiabilidade é um atributo inerente ao projeto do produto e representa a capacidade potencial que deveria ser atingida em condições habituais, desde que fabricado exatamente conforme projetado e operado e mantido exatamente nas condições prescritas.

2.4. Função de Confiabilidade

Definição: A função de confiabilidade é definida como a probabilidade de um produto funcionar sem falhar até um dado tempo t.
Esta é a principal função probabilística usada para descrever dados de durabilidade e, expressando o tempo até falhar como a variável aleatória T, sua expressão é dada por:

$$R(t) = P(T > t) \tag{1}$$

2.4.1. Falha

A definição de falha é um conceito fundamental para qualquer estudo em análise de confiabilidade. De maneira geral, uma falha consiste na interrupção ou alteração da capacidade de um item em desempenhar uma função requerida ou esperada, sendo que item corresponde a qualquer parte, componente, dispositivo, subsistema, unidade funcional, equipamento ou sistema que possa ser considerado individualmente (Guzzon, 2009).
A definição de falha no contexto da confiabilidade diz que: Define-se falha como a impossibilidade de um componente ou sistema cumprir sua função dentro das especificações (níveis) de seu projeto.

2.4.2. Censura

Nos estudos de análise de sobrevivência e confiabilidade é comum situações em que o experimento termina antes que todos os equipamentos, ou produtos, venham a falhar.
Uma característica marcante destes estudos é a presença de observações incompletas ou parciais dos tempos de falha dos produtos observados. Dessa forma existe a necessidade de introduzir uma variável dicotômica na análise que indique se o valor do tempo de sobrevida para um determinado produto foi ou não observado até a sua falha. Essa variável é conhecida na literatura como variável indicadora de censura, e é definida como sendo igual a um, se o tempo de sobrevida é observado, e igual a zero, caso o tempo de sobrevida seja censurado antes da falha acontecer.

2.4.2.1. Tipos de censura

A censura pode ser de três tipos:

- **Censura por Tempo ou do Tipo I**: o teste termina após um período pré-estabelecido de tempo;
- **Censura por Falha ou do Tipo II**: o teste termina após ter ocorrido falha em um número pré-determinado de itens sob teste.
- **Censura do Tipo Aleatório**: o item é retirado do teste antes de ocorrer a falha; é o caso do item falhar por uma causa diferente da que foi definida.

2.5. Técnicas não-paramétricas

2.5.1. Estimador de Kaplan-Meier

O estimador de Kaplan-Meier, ou estimador limite-produto, é um estimador não-paramétrico para a função de confiabilidade. O estimador utiliza os conceitos de independência de eventos e de probabilidade condicional para desdobrar a condição sobreviver até o tempo *t* em uma seqüência de elementos independentes que caracterizam a sobrevivência em cada intervalo de tempo anterior a *t* e cuja probabilidade é condicional aos que estão em risco em cada período.
Por se tratar de um estimador não-paramétrico assintótico, a veracidade das suas estimativas esta garantida apenas para amostras de falhas e tempos de falhas extensas. Sendo assim as estimativas calculadas pelo estimado de Kaplan-Meier ficam mais próximas dos valores reais da confiabilidade conforme o crescimento da amostra utilizada, caso a amostra seja grande, as estimativas tendem para os valores reais da confiabilidade.
Esse estimador é obtido da seguinte forma: suponha que existam n itens sob teste e $k \leq n$ falhas distintas nos tempos $t_1 < t_2 < ... < t_k$. Pode ocorrer mais de uma falha ao mesmo tempo, o que é chamado de empate (Chaves Neto, 2010).
O estimador da função de confiabilidade, para amostras censuradas ou não,

$$\hat{R}(t) \doteq \left[\left(\frac{n_1 - d_1}{n_1} \right) \left(\frac{n_2 - d_2}{n_2} \right) \cdots \left(\frac{n_n - d_n}{n_n} \right) \right] \tag{2}$$

onde,
d_i é o número de falhas no tempo t_i;
n_i é o número de itens sob risco (em teste) no tempo t_i;
t_n é o maior tempo de falha menor que *t*.
Sua variância é calculada a partir da fórmula de Greenwood:

$$\hat{V}[\hat{R}(t)] \doteq \hat{R}(t)^2 \left[\frac{d_1}{n_1(n_1 - d_1)} + \frac{d_2}{n_2(n_2 - d_2)} + ... + \frac{d_n}{n_n(n_n - d_n)} \right] \tag{3}$$

onde,
d_i é o número de falhas no tempo t_i;
n_i é o número de itens sob risco (em teste) no tempo t_i;
t_n é o maior tempo de falha menor que *t*.
E o estimador do erro padrão dessa estatística, $\hat{V}[\hat{R}(t)]$ é:

$$\hat{ep}[\hat{R}(t)] = \sqrt{\hat{V}[\hat{R}(t)]} \tag{4}$$

E assim o intervalo de confiança de nível $(1-\alpha)$ é:

$$P\left(\hat{R}(t) - z_{1-\alpha/2}\hat{ep}[\hat{R}(t)] \leq R(t) \leq \hat{R}(t) + z_{1-1/\alpha}\hat{ep}[\hat{R}(t)]\right) = 1 - \alpha \tag{5}$$

Quando se tem o valor real de R(t) próximo de um ou de zero (valores extremos) pode acontecer do limite superior do intervalo ultrapassar o limite 1 ou do limite inferior ser negativo. Então, deve-se construir um novo intervalo baseado na transformação descrita a seguir (Chaves Neto, 2010).

Seja a estatística $\hat{U}(t)$ função de $\hat{R}(t)$

$$\hat{U}(t) \doteq \ln(-\ln[\hat{R}(t)]) \tag{6}$$

Cuja variância é estimada por:

$$\hat{V}[\hat{U}(t)] \doteq \frac{\hat{V}[\hat{R}(t)] / \hat{R}(t)^2}{\{\ln[(n_1 - d_1)/n_1] + \ln[(n_2 - d_2)/n_2] + ... + \ln[(n_0 - d_0)/n_0]\}} \tag{7}$$

e o erro padrão é dado por

$$\hat{e}p[\hat{R}(t)] = \sqrt{\hat{V}[\hat{U}(t)]} \tag{8}$$

Logo o intervalo de nível $(1-\alpha)$ para R(t) é

$$P\left(\hat{R}(t)^{\exp\left[z.\hat{e}p[\hat{U}(t)]\right]} \le R(t)] \le R(t) \le \hat{R}(t)^{\exp\left[-z.\hat{e}p[\hat{U}(t)]\right]}\right) = 1-\alpha \tag{9}$$

2.6. Técnicas computacionalmente intensivas

2.6.1. Método *Jackknife*

O antecessor ao método *bootstrap* foi o método *Jackknife*, que também é um método computacionalmente intensivo não-paramétrico. O método *Jackknife* foi criado algumas dezenas de anos antes do método *bootstrap*, mas não recebeu tanta atenção dos pesquisadores quanto o método *bootstrap* (Chaves Neto, 1991). A primeira abordagem feita do método *Jackknife* foi apresentada em 1949 por Maurice Quenouille (Quenouille, 1949), e que tem a proposta de reduzir o vício de um estimador de correlação serial com base na divisão da amostra original em duas semi-amostras. Posteriormente, em 1956, Quenouille completou o trabalho com a generalização do método (Chaves Neto, 1991). Tal como o *bootstrap* é um método de reamostragem, pois se baseia na construção de sub-amostras geradas de uma amostra original inicial.

2.6.2. Método *Bootstrap*

2.6.2.1. Definições e propriedades

O *bootstrap* é uma técnica estatística computacionalmente intensiva não-paramétrica que permite a avaliação da variabilidade de estatísticas com base nos dados de uma amostra existente inicial. É indicado para problemas onde o procedimento estatístico padrão não exista o seja de difícil aplicação. É vantajoso se usado em problemas com amostra finita ou com amostra grande, desde que forneça resultado semelhante ao método assintótico usual em grande amostra e superior em amostra pequena (Chaves Neto, 1991).

O *bootstrap* surgiu quando Efron em 1979 estudava o problema da estimação da distribuição amostral de uma estatística $T_n(\underline{x}, F)$ com base nos dados da amostra $\underline{x}' = [x_1, x_2, ..., x_n]$ de tamanho n, de uma distribuição de probabilidade desconhecida $F, X_i \sim i.i.d. F$. O procedimento *bootstrap* é feito a partir de uma amostra original $\underline{x}' = [x_1, x_2, ..., x_n]$ criando-se B novas amostras *bootstrap*, ou reamostras, com tamanho igual ao da amostra inicial, com reposição dos valores. Então, calcula-se a estatística de interesse para cada amostra *bootstrap*. Estas amostras *bootstrap* são conhecidas como pseudo-dados. Assim o conjunto de valores *bootstrap* da estatística corresponde a uma estimativa da verdadeira distribuição amostral da estatística em questão. O fluxograma da Figura 1 mostra o algoritmo de construção da distribuição *bootstrap*.

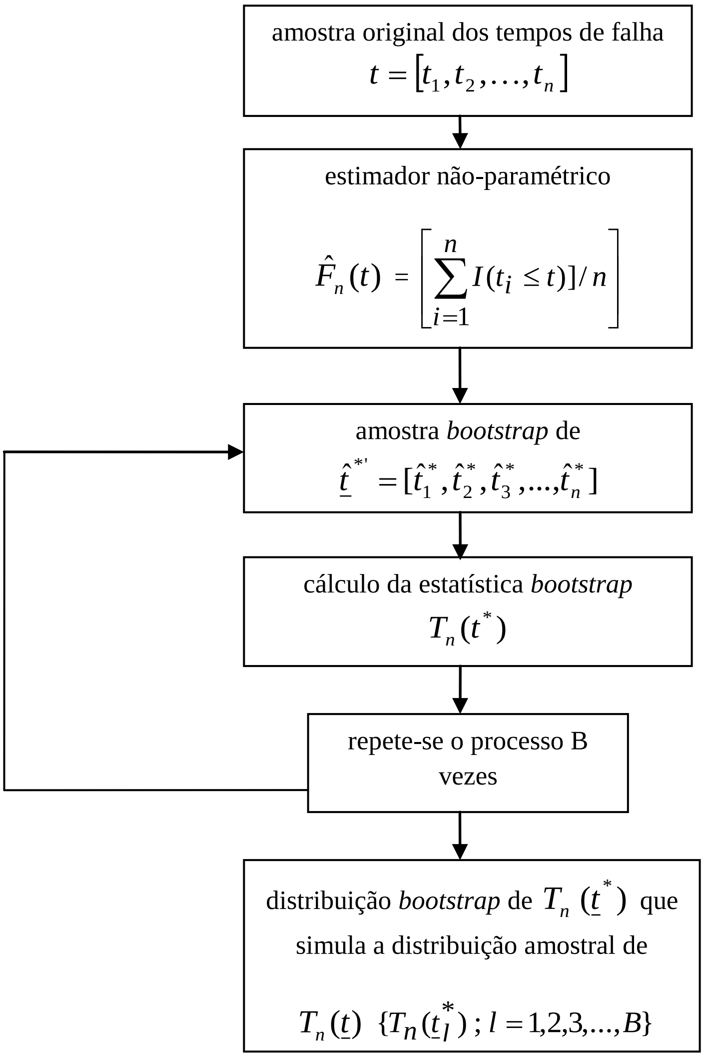

Figura 1. Algoritmo da distribuição *bootstrap* da estatística.

3. MATERIAL E MÉTODO

Para calcular alguns tipos de intervalo de confiança *bootstrap* é necessário calcular primeiramente o erro padrão *bootstrap*, já que é utilizado no lugar do erro padrão da amostra original para o intervalo clássico do estimador de Kaplan-Meier, na expressão do intervalo de confiança do estimador.

O erro padrão *bootstrap* é calculado para cada categoria de tempo de falha, utilizando todas as B amostras *bootstrap*.

Sua expressão apenas depende dos valores da confiabilidade para cada categoria de tempos de falha e a média de cada categoria:

Erro padrão *bootstrap*: $\hat{ep}_i^* = \sqrt{\frac{1}{B-1}\sum\left(\hat{R}_j^{i\,*} - \frac{1}{B}\sum \hat{R}_j^{i\,*}\right)^2}$,

onde B é o número de amostras *bootstrap* e $\hat{R}_j^{i\,*}$ é o valor da confiabilidade estimada para cada ponto do vetor de categorias de tempos de falha, com $i = 1,\ldots,k$ e $j = 1,\ldots,B$ (Abreu, 2011).

3.2. Intervalo de confiança *bootstrap-t*

O intervalo de confiança *bootstrap-t* é calculado da mesma maneira que o intervalo de confiança do estimador de Kaplan-Meier, apenas substituindo o erro padrão clássico do estimador de Kaplan-Meier pelo erro padrão *bootstrap*:

$$P\left(\hat{R}(t) - z_{1-\frac{\alpha}{2}}\hat{ep}^{*}[\hat{R}(t)] \leq R(t) \leq \hat{R}(t) + z_{1-\frac{\alpha}{2}}\hat{ep}^{*}[\hat{R}(t)]\right) = 1-\alpha$$

com n o tamanho da amostra original e $\hat{ep}^{*}$ é o erro padrão *bootstrap* da estatística.

Embora seja flexível e praticamente automático, o cálculo do intervalo de confiança possui um problema que pode afetar a sua eficácia. O método funciona bem quando a distribuição *bootstrap* da estatística é aproximadamente Gaussiana e a estatística pouco viciada. Sendo respeitadas estas condições o intervalo de confiança *bootstrap-t* pode ser calculado na estimação de vários parâmetros.

3.2. Sequência da aplicação do método *bootstrap*

A aplicação do método *bootstrap* segue a seguinte sequência (Abreu, 2011):

- Efetua-se o cálculo da estatística desejada para a amostra inicial $\underline{x}' = [x_1, x_2, ..., x_n]$
- Após isso, o inicia-se o método *bootstrap*, construindo B novas amostras a partir da amostra inicial, com reposição;
- Terminada a construção das B reamostras, B geralmente muito grande, na casa das 10000, calcula-se os valores da estatística para cada reamostra;
- Assim calcula-se a média *bootstrap* para cada posição e o erro padrão *bootstrap*;
- Por fim constrói-se os intervalos de confiança *bootstrap*.

4. RESULTADOS

4.1. Apresentação dos dados

Os resultados obtidos estão apresentados nesta secção. É apresentada uma aplicação do método, para uma amostra com 10 categorias de tempos de falha.

Os dados da amostra são apresentados na seguinte ordem: primeiro é apresentado uma tabela que exibe as categorias de tempos de falha, o número de falha em cada categoria, a confiabilidade estimada por Kaplan-Meier para cada categoria, a média *bootstrap* e o erro padrão *bootstrap*. Uma segunda tabela apresenta os limites dos intervalos de confiança de Kaplan-Meier e os limites do intervalo *bootstrap-t*, todos para um valor de 95% de confiança, e por fim o gráfico contendo os intervalos de confiança do estimador de Kaplan-Meier e o intervalo *bootstrap-t*.

4.2. Aplicação do método *bootstrap*

A amostra apresentada possui censura, contendo 10 categorias de tempos de falha, com um total de $n = 33$ componentes em teste, mas com apenas 18 falhas, sendo os 15 restantes censurados. A censura nesta aplicação é assumida como censura do tipo I ou II, não influenciando diretamente as estimativas da confiabilidade, uma vez que são consideradas como ausência de falha.

Categoria	Nº de falhas	$\hat{R}_i(t)\ i=1,..,10$	Média das categorias	Erro padrão *bootstrap*
$t=1$	1	0.9696970	0.9698923	0.0284697
$t=2$	1	0.9393940	0.9395501	0.0384284
$t=3$	2	0.8787879	0.8793879	0.0632171
$t=4$	3	0.7878789	0.7887975	0.0962398
$t=5$	2	0.7272728	0.7281400	0.0997467
$t=6$	2	0.6666667	0.6675317	0.0999279
$t=7$	1	0.6363637	0.6373060	0.0939728
$t=8$	2	0.5757576	0.5765691	0.0881435
$t=9$	2	0.5151516	0.5154868	0.0769828
$t=10$	2	0.4545455	0.4544869	0.0581672

Tabela 1. Valores da confiabilidade estimada e estatísticas *bootstrap*.

Categoria	IC Kaplan-Meier		IC *bootstrap-t*	
$t=1$	0.8037434	0.9956753	0.9679861	0.9713178
$t=2$	0.7787644	0.9844893	0.9348108	0.9436648
$t=3$	0.7085648	0.9526940	0.8639367	0.8921208
$t=4$	0.6059378	0.8927420	0.7498338	0.8208406
$t=5$	0.5413215	0.8476930	0.6789450	0.7695858
$t=6$	0.4794410	0.7996069	0.6106756	0.7165235
$t=7$	0.4494479	0.7745697	0.5807734	0.6866351
$t=8$	0.3912513	0.7226830	0.5188301	0.6284640
$t=9$	0.3353672	0.6685141	0.4624017	0.5653012
$t=10$	0.2818899	0.6121349	0.4132626	0.4948494

Tabela 2. Intervalos de confiança do estimador de Kaplan-Meier e *bootstrap-t*.

Categoria	Amplitude K-Meier	Amplitude *bootstrap-t*
$t=1$	0.1919	0.0033 (1,72%)
$t=2$	0.2057	0.0089 (4,32%)
$t=3$	0.2441	0.0282 (11,6%)
$t=4$	0.2868	0.0710 (24,8%)
$t=5$	0.3064	0.0906 (29,6%)
$t=6$	0.3202	0.1058 (33,1%)
$t=7$	0.3251	0.1059 (32,6%)
$t=8$	0.3314	0.1096 (33,1%)
$t=9$	0.3331	0.1029 (30,9%)
$t=10$	0.3302	0.0816 (24,7%)

Tabela 3. Amplitude dos intervalos de confiança.

As porcentagens ao lado dos valores das amplitudes na Tabela 3 são referentes ao intervalo de estimador de Kaplan-Meier.

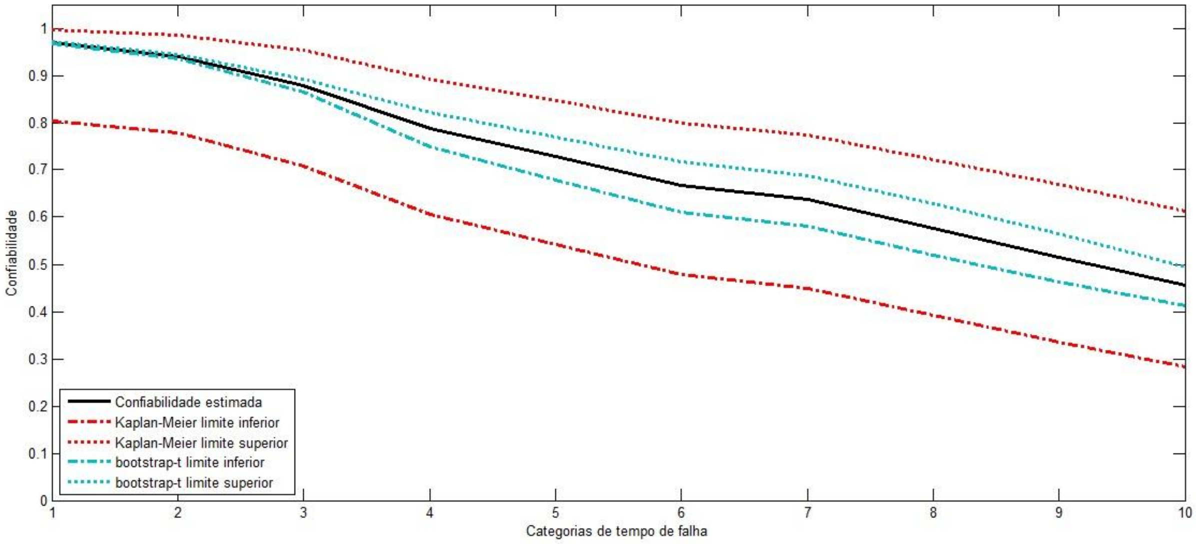

Figura 2. Intervalos de confiança do estimador de Kaplan-Meier e *bootstrap-t*.

A Tabela 3 mostra a eficácia da aplicação para uma amostra hipotética censurada, ficando evidente que o intervalo *bootstrap-t* ficou mais curto, ou seja, amplitude menor, que o intervalo clássico para o estimador de Kaplan-Meier, demonstrando que o objetivo foi alcançado. Para os valores iniciais, o intervalo *bootstrap-t* obteve amplitudes muito pequenas, aumentando nos demais intervalos, mas sempre se mantendo abaixo de um terço da amplitude do intervalo clássico do estimador. Na Figura 2 fica evidente, visivelmente, a eficácia da aplicação.

5. CONCLUSÃO

O objetivo principal do trabalho, a construção de intervalos de confiança para os valores da função confiabilidade estimados pelo método de Kaplan-Meier, foi alcançado usando-se técnica computacionalmente intensiva *bootstap*. Assim foram construídos intervalos de confiança utilizando o método *bootstrap*, além do próprio intervalo de confiança do estimador de Kaplan-Meier. O intervalo de confiança abordado foi o intervalo *bootstrap-t*, comparado com o intervalo do estimador de Kaplan-Meier.
Foi apresentada a análise feita em uma amostra de tempo de falha censurada e ficou evidente a eficácia da aplicação para o intervalo *bootstrap*. Na aplicação do método *bootstrap* foram utilizadas 10000 replicações.

REFERÊNCIAS

ABREU, A. L. E. **Intervalo de confiança *bootstrap* para valores da função de confiabilidade estimados pelo método de Kaplan-Meier** (Dissertação de Mestrado). Programa de Pós-Graduação em Métodos Numéricos em Engenharia, Universidade Federal do Paraná, Curitiba, 2011.

CHAVES NETO, A., **"Bootstrap" em Séries Temporais.** (Tese de Doutorado, 1991). Pontifícia Universidade Católica do Rio de Janeiro. PUC-RJ.

BAILAR III, J. C., MOSTTELER, F. **Medical Uses of Statistics**. 2nd edition, NEJM Book, Boston. 1992.

CHAVES NETO, A. **Confiabilidade e Métodos Estatísticos Aplicados a Sistemas de Engenharia.** Notas de aula. Departamento de Estatística, Universidade Federal do Paraná, Curitiba, 2010.

GUZZON, S. O. **Proposta de análise quantitativa de confiabilidade a partir de dados quantitativos provenientes da FMEA** (Dissertação de Mestrado, 2009). Universidade Federal do Rio Grande do Sul, UFRGS.

HOYLAND, A.; RAUSAND, M. **System reliability theory: models and statistical methods.** New York: Wiley-Interscience, 1994.

QUENOUILLE, M. H. **Aproximate tests of correlation in time-series.** Journal os Statistical Computation and Simulation, 8:75-80, 1949.

STIGLER, S. M. **Citation Patterns in the Jounals of Statistics and Probability**. Statistical Science, 94-108, 1994.