1º EINEPRO
ENCONTRO INTERESTADUAL DE ENGENHARIA DE PRODUÇÃO
28, 29 e 30 de Abril de 2015
SESC Mineiro de Grussaí | São João da Barra - RJ

Estimação de Intervalos de Confiança via Reamostragem *Bootstrap*

Kátia Maria Domingues (UFOP) katiamd2009@hotmail.com

Fernando Luiz Pereira de Oliveira (UFOP) fernandoluiz@iceb.ufop.br

Frederico R. B. Cruz (UFMG) fcruz@est.ufmg.br

Lupércio França Bessegato (UFJF) bessegato@gmail.com

Resumo: Nas últimas décadas, os métodos de reamostragem têm sido muito úteis, principalmente devido ao avanço da tecnologia computacional, na estimação de parâmetros para uma distribuição de interesse. Há diversas técnicas de reamostragem, dentre elas a técnica bootstrap. Neste trabalho, o bootstrap será utilizado para a construção de intervalos de confiança tanto em situações simuladas quanto reais. Será utilizado software R para obtenção dos resultados. Os intervalos abordados nesse trabalho foram o intervalo de confiança bootstrap t, o intervalo de confiança bootstrap percentílico e o intervalo de confiança bootstrap BCa, que ajusta o intervalo de confiança em relação à assimetria, que foram comparados ao intervalo de confiança assintótico. Inicialmente, foram utilizados dados simulados, em que foram testadas amostras de várias distribuições (simétricas e assimétricas). Em seguida, para melhor ilustrar a aplicação dos intervalos de confiança utilizando o método de reamostragem bootstrap, foram utilizados dados reais em dois exemplos. Verificamos pelos resultados que os intervalos de confiança estimados foram coerentes e confirmaram a confiabilidade do método para o processo de estimação em situações em que as distribuições de probabilidade são conhecidas e desconhecidas. Além disso, mostramos que a estimativa do vício do intervalo de confiança bootstrap diminui quando os tamanhos das reamostragens aumentam.

Palavras-chave: Reamostragem; bootstrap; Intervalos de Confiança.

1. Introdução

De acordo com Efron e Tibshirani (1986), a reamostragem consiste em sortear dados de uma amostra, anteriormente retirada de uma população de interesse, para formar uma nova amostra. As técnicas de reamostragem têm sido muito úteis, por exemplo, quando se visa estimar parâmetros para uma distribuição de interesse e o cálculo dos estimadores, por métodos analíticos, é muito complexo ou impossível. Além disso, a reamostragem é flexível e permite diferentes alternativas para se encontrar desvios padrões e intervalos de confiança. Não por acaso, nas últimas décadas, os métodos baseados em reamostragem experimentaram um grande desenvolvimento e uma crescente popularidade, principalmente devido ao avanço e barateamento da tecnologia computacional. Criando múltiplas amostras da amostra original, a reamostragem precisa apenas do poder computacional para estimar um valor de uma estatística para cada amostra. Logo que eles estejam todos calculados, pode-se realizar o teste de normalidade dos valores e até mesmo construir intervalos de confiança e realizar testes de hipóteses. Há diversas técnicas de reamostragens e nesse artigo será abordado de forma rápida o método *jackknife* e, em maior profundidade, o método *bootstrap*.

O objetivo desse artigo é apresentar uma descrição acessível do método de reamostragem *bootstrap* e apresentar uma avalição empírica do seu desempenho na construção de intervalos de confiança. Inicialmente, são analisadas situações simuladas, com distribuições de probabilidades conhecidas, tanto simétricas (a distribuição Normal), quanto assimétricas (as distribuições de Poisson e a Exponencial). Em seguida, aplica-se o método a dados reais, quando a distribuição de probabilidade não é conhecida. No intuito de alcançar tal fim, foram desenvolvidos programas em R (R CORE TEAM, 2013), disponíveis para fins de pesquisa a pedido, para geração dos dados simulados, provenientes das distribuições conhecidas, para o cálculo dos intervalos de confiança e para a geração das tabelas com os resultados de desempenho. Os resultados experimentais são apresentados a seguir, após uma breve revisão bibliográfica sobre o assunto.

2. Revisão

Segundo Costa (2006), a reamostragem descarta a distribuição amostral assumida de uma estatística e calcula uma distribuição empírica, a real distribuição da estatística ao longo de centenas ou milhares de amostras. Além disso, a reamostragem não usa a distribuição de probabilidades assumida. Ao invés disso ela calcula uma distribuição empírica de estatísticas estimadas. Para Hair *et. al.* (2005), com a reamostragem, não é preciso confiar na distribuição assumida nem preocupar-se quanto à violação das suposições inerentes. Pode-se calcular uma real distribuição de estatísticas da amostra e pode-se agora ver onde realmente estão os percentis 95 ou 99, por exemplo, assumindo-se que a amostra original seja confiável.

Entretanto, uma questão importante é saber de onde vêm as múltiplas amostras, uma vez que, se for necessário reunir amostras separadas da população de interesse, aumenta-se sensivelmente o custo de coleta de dados. Felizmente, ao longo dos anos, os estatísticos desenvolveram diversos procedimentos para criar as múltiplas amostras necessárias. Utilizando-se reamostragem, a amostra original pode gerar um grande número de outras amostras, que podem então ser empregadas para gerar a distribuição amostral empírica de uma estatística de interesse. A reamostragem engloba diversos métodos, dentre eles o *bootstrap* e o *jackknife*.

A principal diferença entre os métodos de reamostragem está na forma de extração das amostras, que podem ser com reposição (*bootstrap*), ou sem reposição (*jackknife*). A amostragem com reposição obtém uma observação a partir da amostra e então a coloca de volta na amostra para poder ser usada novamente. A amostragem sem reposição obtém observações da amostra, mas uma vez obtidas, elas são retiradas da amostra. O verdadeiro poder da reamostragem, entretanto, vem da amostragem com reposição, pois pesquisas tem mostrado que esse método fornece estimativas diretas para os intervalos de confiança.

Apesar de os procedimentos de reamostragens não serem restritos a quaisquer suposições paramétricas, eles ainda possuem algumas limitações:

1. A amostra precisa ser grande o bastante e obtida (de maneira aleatória) de forma a ser representativa de toda a população. As técnicas de reamostragem não podem conter enviezamentos, pois pode gerar uma amostra não representativa.

2. Os métodos paramétricos, em muitos casos, são melhores para fazer estimativas pontuais. Os procedimentos de reamostragens podem completar as estimativas pontuais de métodos paramétricos para fornecer as estimativas de intervalos de confiança.
3. As técnicas de reamostragens são inadequadas para identificar parâmetros que apresentam um domínio amostral muito estreito, como os valores de máximo e mínimo. A reamostragem funciona melhor quando a distribuição inteira é considerada para obter o parâmetro em análise.

2.1 Método *Jaccknife*

Esse método de reamostragem, computacionalmente intensivo, foi introduzido por Quenouille (1949), retomado por Tukey (1958), e desenvolvido na última década. De acordo com Costa (2006), o *jackknife* é um método não paramétrico destinado a estimar o viés, e, portanto reduzi-lo, e a variância de estimadores em condições teoricamente complexas ou em que não há confiança no modelo especificado.

O método *jackknife* baseia-se na remoção de 1 amostra (podendo ser mais) do conjunto total observado, recalculando-se o estimador a partir dos valores que restaram. É de fácil implementação e possui número fixo de iterações (n caso retire apenas uma amostra por vez).

O método *bootstrap* obtém as amostras através da amostragem com reposição da amostra original. Já o método *jackknife* obtém as amostras através da amostragem sem reposição da amostra original, isto é, cada amostra tem um tamanho n – 1 e difere apenas pelo caso omitido em cada amostra.

2.2 Método *Bootstrap*

De acordo com Liu e Tang (1996) o método de *bootstrap*, introduzido por Efron no final da década de 70, é uma ferramenta poderosa na estimativa da distribuição de amostragem de uma determinada estatística. A estimativa de inicialização da distribuição de amostragem é geralmente melhor do que a aproximação normal baseada no Teorema do Limite Central, mesmo que a estatística não seja padronizada.

De acordo com Martinez-Espinosa, Sandanielo e Louzada-Neto (2006), o procedimento *bootstrap* é uma técnica de reamostragem, muito utilizada em diferentes situações estatísticas. A base da técnica é a obtenção de um novo conjunto de dados, por reamostragem do conjunto de dados original. Um ponto importante da técnica *bootstrap* não é somente avaliar as estimativas dos parâmetros, mas também obter boas estimativas dos erros padrão da distribuição gerada pelas estimativas dos parâmetros nas iterações de reamostragem.

Segundo Cymrot e Rizzo (2006), uma vantagem em utilizar a técnica de reamostragem *bootstrap* é a generalidade com que pode ser aplicada, pois requer que menos suposições sejam feitas. Outras vantagens são que geralmente fornece respostas mais precisas, além de favorecer o entendimento. Muitas vezes a distribuição de probabilidade da estatística de interesse é desconhecida. Nesse caso o *bootstrap* é muito útil, pois é uma técnica que não exige diferentes fórmulas para cada problema e pode ser utilizada em casos gerais, não dependendo da distribuição original da estatística do parâmetro estudado.

Dessa forma, o *bootstrap* possui maiores possibilidades para obter as respostas, pois não há uma fórmula para cada problema, facilitando o trabalho. Também não depende da distribuição original e não precisa de tantas suposições para estimação de parâmetros das distribuições de interesse (Manteiga, Sánches e Romo, 1994).

De acordo com Davison e Hinkley (1997), repetir um procedimento de análise original com muitas réplicas de dados é denominado método intensivo computadorizado. Para realizar uma estimação através da utilização de *bootstrap* é necessária a realização de um número grande de reamostragens e o cálculo de diversas estatísticas para cada uma destas reamostragens. Isso exige o auxílio de programas computacionais para realizar as reamostras e os cálculos de forma mais rápida e eficaz. Um dos motivos do grande avanço dessa técnica nos últimos anos é ao grande avanço tecnológico dos softwares.

Segundo Cymrot e Rizzo (2006), a técnica *bootstrap* se encaixa na solução de problemas complexos. Através do uso da técnica de *bootstrap* os parâmetros como a média, a variância, a proporção e até mesmo parâmetros menos utilizados como o máximo, mínimo, etc. de uma população podem ser estimados pontualmente e por intervalo.

A técnica de *bootstrap* é denominada não paramétrica, pois a distribuição de probabilidades da estatística do parâmetro a ser estimado é desconhecida. A partir dessa técnica é possível obter a distribuição amostral de um parâmetro a partir da amostra original.

O método *bootstrap* pode ser implementado tanto de forma paramétrica quanto não paramétrica (Hesterberg, 2003), podendo ser utilizado em situações nas quais a teoria de distribuição necessária para apoiar métodos paramétricos não é satisfeita. O que difere os dois métodos é a forma de obtenção da amostra. No caso paramétrico, desde que a distribuição dos dados seja conhecida, a amostra será composta realizando-se a amostragem diretamente dessa distribuição e os parâmetros desconhecidos serão substituídos por estimativas paramétricas. Segundo Cymrot e Rizzo (2006) nesse caso o interesse será estimar o vício das estimativas dos parâmetros e depois efetuar as correções necessárias. Por outro lado, no caso não paramétrico, a amostra de tamanho *n* será composta por extrações -dos elementos da amostra original com reposição.

Segundo Cymrot e Rizzo (2006), a forma não paramétrica é a mais utilizada. Para realizar o teste utilizando a técnica de *bootstrap* é necessário colher uma amostra de tamanho *n* que será denominada de amostra mestre. Essa amostra deve ser coletada de maneira bem planejada para que possa representar a população e que não haja problemas futuros na hora da análise dos dados.

De acordo com Cymrot e Rizzo (2006), a amostra mestre deve representar a população da qual foi retirada. As reamostras desta amostra mestre representam o que se deve obter quando são retiradas muitas amostras da população original. A distribuição *bootstrap* da estatística, baseada em muitas reamostras, representa uma distribuição amostral desta estatística. Esta característica faz com que uma das utilidades da técnica *bootstrap* seja checar a normalidade da distribuição original da estatística em estudo.

Para que a aplicação da técnica resulte em valores confiáveis devem ser feitas, a partir da amostra mestre, centenas ou até milhares de reamostras do mesmo tamanho n. A maioria dos autores recomenda a utilização de 1000 reamostras. Segundo Montgomery, Peck e Vining

(2001) o número de reamostragens pode ser estipulado verificando a variação do desvio padrão para a estimativa do parâmetro em questão calculado para as reamostras à medida que estas são realizadas. No momento em que esse valor se estabilizar o tamanho da reamostra *bootstrap* estará adequado.

Segundo Costa (2006), o *bootstrap* aborda o cálculo do intervalo de confiança de parâmetros e cálculos de valores-p, em circunstâncias em que outras técnicas não são aplicáveis, principalmente quando o número de amostras é reduzido. Essa técnica tenta realizar o que seria desejável na prática, se fosse possível: repetir a experiência.

Conforme feito por Xanchão e Cruz (2010), as observações são escolhidas de forma aleatória e as estimativas recalculadas. Assim uma vez que não se dispõe de toda a população, mas tão somente de uma amostra (observações), faz-se o melhor com o que se tem disponível, isto é, a amostra $x = (x_1, x_2, \dots, x_n)$. Assim, a técnica de *bootstrap* trata a amostra observada como se esta representasse exatamente toda a população (conjunto de experiências, de realizações).

Segundo Cymrot e Rizzo (2006), a distribuição *bootstrap* usualmente tem aproximadamente a mesma forma e amplitude que a distribuição amostral, porém está centrada na estatística dos dados originais, a amostra mestre, enquanto a distribuição amostral está centrada no parâmetro da população.

De acordo com Costa (2010), o intervalo de confiança *bootstrap* também pode ser diretamente calculado, com as duas abordagens simples:

1. Calcula o erro padrão simplesmente como o desvio padrão das estimativas estimadas;
2. Literalmente ordenam as estimativas e definem os valores que contém os 5% extremos dos valores estimados.

De acordo com Costa (2006), matematicamente, a obtenção da amostra *bootstrap* e suas estimativas do erro padrão são obtidas da seguinte maneira:

Seja uma amostra original e a estatística de interesse abaixo:

$$x = \{x_1, x_2, x_3, \dots, x_{n-1}, x_n\},$$
$$\hat{\theta} = F(x);$$

1º) Geram-se as amostras Bootstrap $x_1, x_2, x_3, \dots, x_{n^*}$ com reposição de x.

2º) Calculam-se as estimativas da estatística de interesse:

$$\hat{\theta}_{(b)} = F\left[x_{(b)}\right], b = 1, \dots, B;$$

3º) Calcula-se o erro padrão Bootstrap, S boot, dado por:

$$\hat{S}_{boot} = \sqrt{\frac{1}{B-1}\sum_{b=1}^{B}(\hat{\theta}_{(b)} - \hat{\theta}_{(.)})^2},$$

em que $\hat{\theta}_{(.)} = \frac{1}{B}\sum_{b=1}^{B}\hat{\theta}_{(b)}$

Com $\hat{\theta}_{(b)}$ igual ao valor da estatística para cada reamostra e B igual ao número de reamostras realizadas. Algumas literaturas utilizam apenas B ao invés de (B – 1), pois como o número de reamostras é muito grande, essa alteração torna-se praticamente insignificante.

Segundo Montgomery e Runger (2003), uma estatística utilizada para estimar um parâmetro é viciada quando a distribuição amostral não estiver centrada no verdadeiro valor do parâmetro. A técnica Boostrap nos permite verificar o vício olhando se a distribuição Bootstrap da estatística está centrada na estatística da amostra mestre.

O estimador do vício da distribuição Bootstrap é dado por:

$$\text{vício}_{\text{bootstrap}} = \overline{\hat{\theta}_{(b)}} - \hat{\theta}.$$

Para Costa (2006), se B → ∞, então as estimativas do erro-padrão, do enviesamento e do EMQ se igualam às estimativas de máxima verossimilhança (Efron, 1983). Para o cálculo das estimativas Bootstrap geralmente é suficiente um valor de B=100. Contudo, para se determinar à distribuição por amostragem com precisão deve considerar-se um valor para B substancialmente mais elevado. Segundo Efron (1983), geralmente B = 1000 proporciona bons resultados. E em ambos os casos, convém ensaiar diferentes valores para B até se verificar a convergência dos resultados.

No processo de inferência estatística de um determinado parâmetro, em geral, não se pode valer apenas da utilização de uma estimativa pontual θ, porque esta estimativa não apresenta medidas de precisão e confiança decorrentes do processo de estimação. Para a estimação dos intervalos de confiança, a precisão do estimador e o erro de estimação são levados em consideração, possibilitando assim a obtenção de estimativas mais confiáveis.

Segundo Efron e Tibshirani (1983), o método *bootstrap*, juntamente com o seu erro padrão, pode ser uma técnica robusta na hora de construir um intervalo de confiança para as estimativas do parâmetro de interesse. Além do intervalo de confiança *bootstrap* convencional existem vários outros métodos de cálculo do intervalo de confiança, que eliminam alguns problemas que geralmente aparecem quando se trabalha com amostras "estranhas", ou seja, amostras que não tenham distribuições parecidas com a distribuição normal.

Segundo alguns autores, uma das aplicações do método *bootstrap* é obter intervalos de confiança que são confiáveis. Há várias técnicas distintas para o cálculo de intervalos de confiança *bootstrap* e nesse trabalho serão utilizados o método *Bootstrap t*, o método Percentil, que pode ser obtido pelo Método de Correção de Vício Acelerado, Biased Corrected Accelerated (BCa).

3. Simulações e Resultados

Serão apresentados os resultados das simulações feitas. Os programas foram desenvolvidos no software R (R CORE TEAM, 2013). O método *bootstrap* é considerado uma ferramenta muito útil para as estatísticas e pode ser facilmente implementado no software R. O pacote boot do R oferece algumas opções para desenvolver a técnica de *bootstrap* e métodos de reamostragem. Nas simulações mostradas na Tabela 1, foram utilizados os pacotes boot e *bootstrap* do software R. Nas simulações foram testadas amostras de várias distribuições (simétricas e assimétricas). Também foram utilizados vários tamanhos de amostras. A seção seguinte mostra os resultados para estes cenários.

3.1 Dados Simulados

A partir dos métodos descritos nas seções anteriores, podem-se obter vários intervalos de confiança *bootstrap*. Os abordados nesse trabalho foram o intervalo de confiança *bootstrap* t, o intervalo de confiança *bootstrap* percentil e o intervalo de confiança *bootstrap* BCa, que ajusta o intervalo de confiança em relação à assimetria. Os intervalos de confiança serão calculados para a média das distribuições de probabilidades simuladas. O valor de B para o número de replicações foi crescendo gradativamente durante as análises, sendo que para cada amostra os valores de n e B foram sempre os mesmos. De acordo com os resultados dos dados simulados, quando aumentaram os valores de B os resultados começaram a ficar melhores.

Nas simulações apresentadas na Tabela 1 foram utilizados quatro tamanhos (n) diferentes para as amostras aleatórias e também os mesmos tamanhos B para a reamostragem *bootstrap*. Nas simulações com a distribuição de Poisson, o parâmetro para a média é λ e para a variância é λ. Para a distribuição exponencial, os parâmetros para a média e avariância são, respectivamente, $\frac{1}{\lambda}$ e $\frac{1}{\lambda^2}$.

Pode-se perceber que os intervalos de confiança calculados pelo método clássico e pelos métodos *bootstrap* não diferem significativamente quando as amostras são provenientes da distribuição Normal. Os intervalos de confiança ficam praticamente iguais quando utiliza-se $n = B = 500$. Isto comprova que o intervalo de confiança *bootstrap* fica melhor quando se aumenta o valor de B, conforme assegurado em diversos artigos da literatura.

Para a distribuição de Poisson não há diferenças significativas entre os intervalos de confiança pelos métodos clássico e *bootstrap*, mas pode-se perceber que a amplitude dos intervalos *bootstrap* é maior que do método clássico, quando n e B são pequenos. Entretanto, quando aumenta n e B para 500, essa amplitude fica praticamente igual entre os métodos.

Já para a distribuição exponencial, também não há diferenças significativas entre os intervalos de confiança pelos métodos clássico e *bootstrap*, mas amplitude fica praticamente igual para valores de n e B iguais a 1000.

Entretanto, para as simulações realizadas, os métodos *bootstrap* fornecem intervalos de confiança com menores amplitudes, sendo que essa diferença é quase imperceptível quando a amostra é grande. Isto é, os dois métodos parecem concordar nos intervalos de confiança nas grandes amostras. A relação de melhor ou pior intervalo será feita em questão da amplitude, dessa forma o intervalo que ficar mais extenso é considerado pior, ou menos ajustado em torno da estimativa, que o outro de amplitude menor.

Pode-se perceber pelos resultados obtidos que, conforme Cymrot e Rizzo (2006), quando a distribuição do parâmetro a ser estimado é conhecida, há coincidência entre o intervalo paramétrico e o intervalo *bootstrap* reforçando a hipótese de veracidade sobre as suposições do modelo paramétrico.

TABELA 1 – Simulação de algumas distribuições de probabilidade.

		Intervalo de Confiança			
Distribuição	Tam. (B)	Clássico	*Bootstrap t*	*Bootstrap* Percentil	BCA
Normal (10,2)	50	[9,4; 10,5]	[9,5; 10,6]	[9,2; 10,3]	[9,2; 10,3]
	100	[9,2; 10,2]	[9,3; 10,2]	[9,3; 10,2]	[9,3; 10,2]
	500	[9,3; 10,2]	[9,9; 10,3]	[9,9; 10,3]	[9,9; 10,3]
	1000	[9,9; 10.1]	[9,9; 10,1]	[9,9; 10,1]	[9,9; 10,1]
Normal(20,1)	50	[19,6; 20,4]	[19,7; 20,4]	[19,5; 20,3]	[19,5; 20,3]
	100	[19,7; 20,2]	[19,8; 20,3]	[19,7; 20,2]	[19,7; 20,2]
	500	[19,9; 20,1]	[19,9; 20,1]	[19,9; 20,1]	[19,9; 20,1]
	1000	[19,9; 20,0]	[19,9; 20,0]	[19,9; 20,0]	[19,9; 20,0]
Normal(20,8)	50	[18,3; 22,7]	[17,9; 22,7]	[18,3; 23,0]	[18,3; 23,0]
	100	[18,9; 22,5]	[18,6; 22,5]	[18,9; 22,8]	[18,9; 23,1]
	500	[18,9; 20,4]	[18,9; 20,4]	[18,9; 20,4]	[19,0; 20,4]
	1000	[19,4; 20,3]	[19,4; 20,3]	[19,4; 20,3]	[19,3; 20,3]
Poisson(5)	50	[4,4; 5,4]	[4,2; 5,4]	[4,2; 5,4]	[4,5; 5,5]
	100	[4,8; 5,8]	[4,8; 5,7]	[4,9; 5,7]	[4,8; 5,7]
	500	[4,9; 5,3]	[4,9; 5,3]	[4,9; 5,3]	[4,9; 5,3]
	1000	[4,9; 5,1]	[4,9; 5,1]	[4,9; 5,1]	[4,9; 5,1]
Poisson(10)	50	[8,8; 10,8]	[8,8; 10,9]	[8,7; 10,8]	[8,7; 10,7]
	100	[9,5; 10,8]	[9,2; 10,8]	[9,4; 10,9]	[9,6; 11,7]
	500	[9,8; 10,3]	[9,8; 10,3]	[9,8; 10,3]	[9,8; 10,3]
	1000	[9,8; 10,2]	[9,8; 10,2]	[9,8; 10,2]	[9,8; 10,2]
Poisson(20)	50	[19,4; 21,5]	[19,0; 21,5]	[19,5; 22,0]	[19,4; 21,8]
	100	[19,7; 21,1]	[19,7; 21,1]	[19,7; 21,1]	[19,7; 21,0]
	500	[19,6; 20,4]	[19,6; 20,4]	[19,6; 20,4]	[19,6; 20,4]
	1000	[19,6; 20,1]	[19,6; 20,1]	[19,6; 20,1]	[19,6; 20,1]
Exp. (10)	50	[0,06; 0,13]	[0,04; 0,12]	[0,08; 0,16]	[0,08; 0,16]
	100	[0,08; 0,12]	[0,08; 0,12]	[0,08; 0,12]	[0,08; 0,11]
	500	[0,09; 0,11]	[0,09; 0,11]	[0,09; 0,11]	[0,09; 0,11]
	1000	[0,08; 0,10]	[0,09; 0,10]	[0,09; 0,10]	[0,09; 0,10]
Exp. (20)	50	[0,041; 0,065]	[0,037; 0,066]	[0,041; 0,069]	[0,045; 0,071]
	100	[0,035; 0,049]	[0,035; 0,048]	[0,035; 0,048]	[0,036; 0,049]
	500	[0,045; 0,054]	[0,046; 0,055]	[0,046; 0,059]	[0,046; 0,056]
	1000	[0,048; 0,053]	[0,047; 0,053]	[0,048; 0,053]	[0,048; 0,053]
Exp. (0,5)	50	[1,22; 2,09]	[1,05; 2,02]	[1,36; 2,33]	[1,36; 2,33]
	100	[1,79; 2,63]	[1,79; 2,65]	[1,72; 2,58]	[1,79; 2,59]
	500	[1,75; 2,06]	[1,74; 2,05]	[1,75; 2,05]	[1,76; 2,08]
	1000	[1,91; 2,15]	[1,90; 2,15]	[1,91; 2,15]	[1,91; 2,15]

3.2 Dados Reais

Para melhor ilustrar a aplicação dos intervalos de confiança utilizando o método de reamostragem *bootstrap*, serão utilizados dados reais em dois exemplos extraídos de Montgomery (2004).

Exemplo 1: Um artigo no Journal of the Electrochemical Society (Vol.139, nº 2, 1992, pp.524-532) descreve um experimento para investigar a deposição de vapor de baixa pressão de polissilício. O experimento foi realizado em um reator de alta capacidade na fábrica Sematech em Austin, Texas. O reator tem várias posições de placas, e selecionam-se quatro dessas posições aleatoriamente. A variável resposta é a uniformidade da espessura do filme. Foram feitas três replicações do experimento, e os resultados são os seguintes.

TABELA 2: Exemplo 1.

Posição da Placa	Uniformidade		
1	2,76	5,67	4,49
2	1,43	1,70	2,19
3	2,34	1,97	1,47
4	0,94	1,36	1,65

[Fonte: Montgomery, 2004)

Os resultados dos intervalos de confiança pelo método clássico e pelos métodos *bootstrap* estão apresentados na tabela 3.

TABELA 3: Intervalos de Confiança para os dados do exemplo 1.

		Intervalo de Confiança			
Distribuição	Tam. [B]	Clássico	*Bootstrap t*	*Bootstrap* Percentil	BC_A
Amostra [12]	50	[1,427; 3,121]	[1,339; 2,947]	[1,715; 3,322]	[1,761; 3,351]
	100	[1,666; 3,106]	[1,676; 3,054]	[1,607; 2,985]	[1,816; 3,608]
	500	[1,497; 3,130]	[1,450; 3,046]	[1,616; 3,212]	[1,700; 3,493]
	1000	[1,604; 3,082]	[1,573; 3,002]	[1,659; 3,088]	[1,749; 3,435]

A Tabela 3 mostra os intervalos de confiança pelo método clássico e pelos métodos *bootstrap* com a amostra n = 12 e vários tamanhos para a reamostragem *bootstrap* B. Pode-se perceber que os intervalos de confiança diferem muito pouco e quando aumenta o valor de B, à partir de 500, os intervalos de confiança Clássico e bootstrap t praticamente ficam iguais e não há diferenças significativas na amplitude desses intervalos. Já o intervalo de confiança *bootstrap* percentil apresenta valores um pouco diferentes, mas a amplitude não apresenta diferenças significativas, o mesmo acontece com o intervalo de confiança BCa, que corrigi o vício do intervalo *bootstrap* percentil. Nesse caso, o pesquisador poderá escolher qualquer um dos intervalos.

Foi realizado o teste de normalidade (Shapiro-Wilk) para verificar se os dados são oriundos de uma população com distribuição Normal e obteve-se o p-valor igual a 0,009121 comprovando que a suposição de normalidade dos dados não é rejeitada. Também foi realizado o teste de Anderson-Darling para normalidade e obteve-se p-valor igual a 0,006501, também rejeitando a normalidade.

Exemplo 2: Conduziu-se um experimento para investigar a capacidade de enchimento de um equipamento de embalagem em uma vinícola em Newberg, Oregon. Vinte garrafas de Pinot Gris foram selecionadas aleatoriamente e o volume de enchimento (em ml) foi medido. Suponha que o volume de enchimento tenha distribuição normal. Os dados são os seguintes: 753; 751; 752; 753; 753; 753; 752; 753; 754; 754; 752; 752; 752; 750; 753; 755; 753; 756; 751; 750.

Os resultados dos intervalos de confiança pelo método Clássico e pelo método *bootstrap* estão apresentados na Tabela 5.

A Tabela 5 mostra os intervalos de confiança pelo método Clássico e pelos métodos *bootstrap* com a amostra n = 20 e vários tamanhos para a reamostragem *Bootstrap B*. Pode-se perceber que os intervalos de confiança diferem muito pouco e quando aumenta o valor de *B*, à partir de 100, os intervalos de confiança Clássico e *bootstrap* praticamente ficam iguais e não há diferenças significativas na amplitude desses intervalos. Novamente, nesse caso, o pesquisador poderá escolher qualquer um dos intervalos.

TABELA 5: Intervalos de Confiança para os dados do exemplo 2.

		Intervalo de Confiança			
Distribuição	Tam. [B]	Clássico	*Bootstrap t*	*Bootstrap* Percentil	BC_A
Amostra [20]	50	[752,0; 753,2]	[751,9; 753,3]	[751,9; 753,3]	[751,8; 753,1]
	100	[751,9; 753,2]	[751,9; 753,2]	[752,0; 753,3]	[751,8; 753,2]
	500	[752,0; 753,2]	[752,0; 753,3]	[751,9; 753,2]	[751,9; 753,1]
	1000	[752,0; 753,3]	[752,0; 753,3]	[751,9; 753,2]	[751,9; 753,2]

Foi realizado o teste de normalidade (Shapiro-Wilk) para verificar se os dados são oriundos de uma população com distribuição Normal e obteve-se o p-valor igual a 0,2915 comprovando a suposição de normalidade dos dados. Também foi realizado o teste de Anderson-Darling para normalidade e obteve-se p-valor igual a 0,1266, comprovando normalidade dos dados.

4. Considerações Finais

Após a aplicação dos intervalos de confiança utilizando o método de reamostragem *bootstrap* foi verificado que os intervalos de confiança estimados foram coerentes e confirmaram a confiabilidade do método para estimação de parâmetros em situações em que as distribuições das estimativas dos parâmetros eram conhecidas.

Foi observado que, quando a estatística do parâmetro estudado tinha distribuição Normal, os intervalos de confiança *bootstrap* estimados pelos métodos t e Percentil forneceram resultados muito próximos, e assim, ambos foram considerados adequados. O mesmo foi verificado quando a estatística do parâmetro estudado tinha distribuição Poisson ou Exponencial.

Analisando os exemplos (dados reais) apresentados, foi possível verificar a generalidade de aplicação da técnica de estimação por meio da reamostragem. Concluiu-se que o método *bootstrap* permite que o cálculo do intervalo de confiança seja realizado de modo mais simples e abrangente para diversas estatísticas, mesmo quando as distribuições de probabilidades das mesmas são desconhecidas. Foi visto que quando o tamanho da amostra é pequeno, como no caso do exemplo 1, o método de reamostragem *bootstrap* é o mais indicado para a construção de intervalos de confiança.

Para o exemplo 2, em que a suposição de normalidade foi comprovada, pode-se utilizar qualquer um dos intervalos de confiança, pois a amplitude dos intervalos clássicos e *bootstrap* coincidem.

Outra observação importante está relacionada com o valor pequeno do viés nos exemplos 1 e 2. Tais resultados sugerem que os valores estimados devem estar próximos dos verdadeiros valores. Como foi mostrado, o valor absoluto do vício do intervalo de confiança bootstrap diminui quando o tamanho de B aumenta. O menor valor absoluto do vício foi verificado quando B = 1000, o que corrobora com as teorias de muitos autores que afirmam que, para que a técnica resulte em valores confiáveis, o tamanho de B deve ser igual a 1000.

Por fim, vale ressaltar que, com o avanço tecnológico, os métodos computacionais são muito rápidos. Assim, pode-se fazer simulações com diferentes valores de B com baixo custo computacional.

Para trabalhos futuros sugerimos uma avaliação inferencial mais detalhada a respeito dos intervalos de confiança utilizados, uma comparação com outros métodos de estimação intervalar e uma análise com dados reais de áreas afins.

5. Agradecimentos

Este trabalho foi parcialmente financiado pelo Conselho Nacional de Desenvolvimento Científico e Tecnológico - CNPq e pela Fundação de Amparo à Pesquisa do estado de Minas Gerais - FAPEMIG.

6. Referências

BADGWELL, T.A.; EDGAR, T.F.; TRACHTENBERG, I.; ELLIOTT, J.K. Experimental verification of a fundamental model for multi-wafer low-pressure chemical vapor deposition of polysilicon. *Journal of the Electrochemical Society*, v. 139, n. 2, p. 524-532, 1992.

COSTA, G.G.O. Intervalo de confiança e teste de significância bootstrap para coeficiente de correlação linear referente à hipótese de um valor não nulo. *Revista GEPROS (Gestão de Produção, Operações e Sistemas)*, Bauru, n.2, p.177-186, 2010.

COSTA, G.G.O. *Um procedimento inferencial para análise fatorial utilizando as técnicas bootstrap e jackknife: Construção de intervalos de confiança e testes de hipóteses*. 189 f. Tese (Doutorado em Engenharia Elétrica) – Pontifícia Universidade Católica do Rio de Janeiro – PUC-RIO, Rio de Janeiro, 2006.

CYMROT, R.; RIZZO, A.L.T. Aplicação da técnica de reamostragem bootstrap na estimação da probabilidade dos alunos serem usuários de transporte público. *In: Environmental and Health Word Congress*, Santos, SP, p. 292-296, 2006.

DAVISON, A.C.; HINKLEY, D.V. *Bootstrap methods and their application* (Cambridge Series in Statistical and Probabilistic Mathematics). 5a ed., Cambridge: University of Cambridge, 1997.

EFRON, B.; TIBSHIRANI, R. *An introduction to the bootstrap*. Chapman and Hall, New York, 1983.

EFRON, B.; TIBSHIRANI, R. Bootstrap methods for standard errors, confidence intervals, and other measures of statistical accuracy. *Statistical Science*, v. 1, n. 1, p. 55–77, 1986.

HAIR, J.F.; ANDERSON, R.E.; TATHAM, R.L.; BLACK, W.C. *Análise multivariada de dados*. Bookman, Porto Alegre, 2005.

HESTERBERG, T. Bootstrap methods and permutation tests. *In: The practice of business statistics: using data for decisions*. New York, NY, cap 18, 2003.

LIU, R.Y.; TANG, J. Control charts for dependent and independent measurements based on bootstrap methods. *Journal of American Statistical Association*, v. 91, n. 436, p. 1694-1700, 1996.

MANTEIGA, W.G.; SÁNCHEZ, J.M.P.; ROMO, J. The bootstrap - A review. *Computational Statistics*, v. 9, n. 4, p. 165-176, 1994.

MARTINEZ-ESPINOSA, M.; SANDANIELO, V.L.M.; LOUZADA-NETO, F. O método de bootstrap para o estudo de dados de fadiga dos materiais. *Revista de Matemática e Estatística*, v.24, n.2, p.37-50, 2006.

MONTGOMERY, D.C. *Introdução ao controle estatístico da qualidade*. 4a ed. Rio de Janeiro: LTC, 2004.

MONTGOMERY, D.C.; PECK, E.A.; VINING, G.G. *Introduction to linear regression analysis*. 3a ed. New York: Wiley, 2001.

MONTGOMERY, D.C.; RUNGER, G.C. *Estatística aplicada e probabilidade para engenheiros*. 2a ed. Rio de Janeiro: LTC, 2003.

QUENOUILLE, M.H. Approximate tests of correlation in time-series. *Journal of the Royal Statistical Society. Series B*. v. 11, n. 1, p. 68-84, 1949.

R CORE TEAM. *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria, 2013. URL http://www.R-project.org/

TUKEY, J.W. Bias and confidence in not-quite large samples. *Annals of Mathematical Statistics*, v. 29, p. 614, 1958.

XANCHÃO, R.C.; CRUZ, F.R.B. *Construção de intervalos de confiança via bootstrap*. Relatório Técnico RTE-04/2010, Departamento de Estatística – ICEx – UFMG, Belo Horizonte, MG, 2010.