

Procedimentos inferenciais em índices de capacidade para dados autocorrelacionados via *bootstrap*

ALBERTO W. RAMOS

Departamento de Engenharia de Produção - EPUSP
E-mail: awramos@usp.br

LINDA LEE HO

Departamento de Engenharia de Produção - EPUSP
Email: lindalee@usp.br

Resumo

Este artigo apresenta procedimentos inferenciais para construir intervalos de confiança em índices de capacidade C_p e C_{pk} através de técnicas de reamostragem (*bootstrap*) quando os dados provenientes de processos normais são autocorrelacionados.

Palavras-chave

Autocorrelação; *bootstrap*; índices de capacidade C_p e C_{pk} ; séries temporais.

Inferential procedures for capability indices in autocorrelated data by bootstrap

Abstract

This paper presents procedures to build confidence intervals for capability indices C_p and C_{pk} based on bootstrap (resampling) methodology when data from the normal process are autocorrelated.

Key words

Autocorrelation; bootstrap; capability index C_p and C_{pk} ; time series.

INTRODUÇÃO

Os índices de capacidade têm por objetivo avaliar se um processo gera produtos que atendam às especificações de engenharia. Uma grande variedade de índices de capacidade pode ser encontrada na literatura, porém, dois índices são mais freqüentemente utilizados: C_p e C_{pk} . O primeiro é definido como sendo a razão entre a tolerância de engenharia e a dispersão do processo, isto é:

$$C_p = \frac{LSE - LIE}{6\sigma}$$

onde: LSE é o limite superior da especificação; LIE é o limite inferior da especificação e σ é o desvio-padrão do processo. O índice C_p compara a tolerância da especificação do produto com a variação do processo (também chamada de tolerância natural). Intrinsecamente, este índice admite que a média do processo pode ser facilmente ajustada e, portanto, somente a tolerância de engenharia (a distância entre o limite superior e o inferior da especificação) é comparada com a dispersão total. Esta é sempre a melhor condição possível para estudo, daí o motivo do índice ser habitualmente chamado de capacidade potencial.

O outro índice (C_{pk}), é definido como

$$C_{pk} = \min\left(\frac{\mu - LIE}{3\sigma}, \frac{LSE - \mu}{3\sigma}\right)$$

onde μ é a média do processo. Este índice além de avaliar a variabilidade total permitida às peças com a tolerância natural de fabricação, verifica, também, a posição do processo em relação aos limites (superior e inferior) da especificação. Este pode, também, ser empregado nos casos de processos com especificações unilaterais, quando inexistente LIE ou LSE , ou quando a média do processo não pode ser centralizada na especificação devido a um problema de engenharia ou a um custo elevado de alteração ou ajuste. Em casos de processos em que a média é centralizada na especificação e a característica de interesse obedece a uma distribuição normal, é válida a igualdade entre C_p e C_{pk} .

A avaliação da capacidade do processo costuma ser feita mediante a comparação dos resultados dos índices de capacidade com valores mínimos estipulados em normas ou fixados pelos próprios clientes. Por exemplo, com o lançamento dos requisitos QS-9000, foi estabelecido o critério $C_{pk} > 1,67$ para a decisão da capacidade preliminar do processo para características críticas. Este valor é superior ao estabelecido em outras metodologias

tal como em American Society for Quality Control (ASQC) 1986, que adota $C_{pk} > 1,33$. Os critérios para decisão sobre a aceitabilidade do processo variam enormemente de indústria para indústria, não havendo uma padronização. Caso os índices tenham revelado uma inadequação do processo quanto ao atendimento das especificações de engenharia, ações corretivas adequadas devem ser tomadas, ou seja, deve-se combater a variação excessiva.

A avaliação dos índices de capacidade, embora simples, apresenta alguns problemas associados. É necessário deixar claro que, em situações práticas, μ e σ são geralmente desconhecidos e suas estimativas são utilizadas, obtendo-se deste modo estimativas de C_p e C_{pk} que estão sujeitas a um erro de estimação em função da quantidade de dados disponíveis e, por vezes, do conhecimento limitado sobre o comportamento do processo. De outra forma, estimativas de C_p e C_{pk} superiores a 1 não asseguram necessariamente a capacidade do processo em atender às especificações. As metodologias para realizar estudos de capacidade de curto-prazo em geral propõem valores estimados superiores a 1,33 para C_p e C_{pk} (ou até 1,67, dependendo da situação) para um processo ser considerado capaz. Tais valores mínimos não possuem embasamento estatístico, mas são apenas uma margem de segurança que busca assegurar que os valores de C_p ou C_{pk} , que são desconhecidos, tenham uma alta probabilidade de serem superiores a 1,0. Conforme Cheng (1994), dependendo do tamanho da amostra, estas margens de segurança (33% ou 67%) dadas aos índices podem ser insuficientes, conduzindo a interpretações incorretas sobre a real capacidade do processo. Ainda que C_p e C_{pk} possam ser relacionados com a porcentagem de produtos não conformes gerados pelo processo, tal sistemática necessita do emprego de uma distribuição de probabilidade (normal ou outra qualquer adequada ao caso em estudo) para o seu cálculo. Caso seja conhecida a distribuição de probabilidade das estimativas destes índices, a construção de intervalos de confiança e a avaliação da chance de um certo índice ser menor do que 1,0 (processo incapaz) torna-se viável.

Aspectos inferenciais do índice C_p já foram estudados por vários autores, considerando as observações normalmente distribuídas. Cryer; Ryan (1990) demonstraram que a estimação do desvio-padrão baseada na amplitude móvel média de um gráfico para valores individuais é ineficiente quando a distribuição dos valores individuais é normal e o processo é estável. O desvio-padrão amostral s também é um estimador viesado de σ e, portanto, s deve ser corrigido através do coeficiente, c_p , cujas tabelas podem ser encontradas em Montgomery (1996, p. A-15). Quando o tamanho da

amostra $n > 30$ tal correção se torna desnecessária, já que $c_4 \cong 1$. Se os valores individuais possuem distribuição normal, s^2 terá uma distribuição proporcional a uma distribuição qui-quadrado, com $n-1$ graus de liberdade. O limite inferior de um intervalo de confiança construído desta forma está na expressão (1).

$$P\left(Cp \geq \hat{Cp} \sqrt{\frac{\chi^2_{(n-1); 1-\alpha}}{n-1}}\right) = 1 - \alpha \quad (1)$$

No caso de índice Cpk , Bissel (1990) construiu um intervalo de confiança, cujo limite inferior está expresso em (2).

$$P\left\{Cpk \geq \hat{Cpk} \left[1 - z_\alpha \sqrt{\frac{1}{9n\hat{Cpk}} + \frac{1}{2(n-1)}}\right]\right\} = 1 - \alpha \quad (2)$$

Levinson (1997) construiu limites de confiança para Cpk , com o emprego da distribuição t-Student não centrada. Maiores detalhes dos aspectos inferenciais deste índice ver Kotz; Johnson (1993) e Kotz & Lovelace (1998).

Embora de uso comum no meio industrial, os índices convencionais costumam ser somente válidos quando os dados obtidos são independentes e com distribuição normal. Este não é o caso de diversas características dimensionais, químicas ou físicas, habitualmente mensuradas em processos produtivos.

Nos dias de hoje, o fenômeno de autocorrelação nas observações tem sido relatado cada vez mais por diversos autores, tais como Montgomery (1996), Pyzdek (1992b) e Woodall; Faltin (1994), Neto e Souza (1996), por exemplo. A existência de autocorrelação pode sempre ser explicada em termos físicos. Dito de outra forma, a existência de dados autocorrelacionados pode ser criada por algum mecanismo que faz com que estes não mais sejam independentes entre si ao longo do tempo. Para medir o grau e intensidade da autocorrelação existente nos dados, utiliza-se o coeficiente de autocorrelação, definido como:

$$\rho_L = \frac{COV(x_t, x_{t-L})}{\sigma^2(x_t)} \quad L = 1, 2, 3, \dots, n-1$$

em que, L é o retardo (do inglês, *lag*) existente entre os dados no cálculo de ρ_L , $COV(x_t, x_{t-L})$, a covariância,

$\sigma^2(x_t)$, a variância populacional e n é a quantidade de dados disponíveis. Na prática, ρ_L é estimado por

$$\hat{\rho}_L = \frac{\sum_{t=1}^{n-L} (x_t - \bar{x})(x_{t+L} - \bar{x})}{\sum_{t=1}^n (x_t - \bar{x})^2} \quad L = 0, 1, 2, 3, \dots, n, \text{ que}$$

pode ser calculado de forma aproximada para um tamanho de amostra grande ($n > 30$), como habitualmente ocorre em estudos de índice de capacidade. Os valores de $\hat{\rho}_L$ estão sempre entre -1 e $+1$ (inclusive) e quanto maior o seu valor absoluto, maior a possibilidade de existência de dependência estatística serial entre os dados. Valores negativos de $\hat{\rho}_L$, por exemplo, indicam autocorrelação negativa, ou seja, quando um valor da série temporal se apresenta alto, o seguinte tem tendência a apresentar valores baixos e vice-versa.

Na interpretação de índices tradicionais de capacidade (Cp e Cpk), há a restrição de que os valores individuais sigam uma distribuição normal. Entretanto, embora a distribuição de dados autocorrelacionados possa parecer normal, alguns problemas inexistentes, quando as observações são independentes, são introduzidos. O mais relevante é, provavelmente, o fato de que o estimador s^2 deixa de ser justo (ou não-viesado) e, conseqüentemente, seu valor esperado será menor do que a variância real do processo σ^2 , quando a autocorrelação for positiva (SHORE, 1997, p. 619). Isto decorre do fato de que, quando os dados estão positivamente autocorrelacionados, um ponto acima da linha média num gráfico de controle tende a ser sucedido por pontos também superiores à média, formando várias seqüências (*runs*) de pontos consecutivos. Esta tendência de proximidade entre os pontos faz com que a estimação de σ^2 seja afetada. Por outro lado, a média da amostra pode também ser substancialmente diferente da média da população, afetando a sua estimação. Portanto, as estimativas dos índices Cp e Cpk estarão viesadas.

Em relação à avaliação da capacidade de processos com dados autocorrelacionados, pouco tem sido escrito e pesquisado. O interesse pelo assunto ainda é recente, segundo Zhang (1996, p.49), e mesmo alguns autores de renome, tais como Kotz; Johnson (1993) referem-se pouco a este problema específico. Diversos autores, tais como Rodriguez (1992), Franklin; Wasserman (1992), e Constable; Yen (1991), têm sugerido a utilização de técnicas de simulação, para avaliação da capacidade em situações especiais. Dentre os métodos de simulação, um parece ser de particu-

lar interesse: o método de reamostragem (Bootstrap). Embora o método já seja conhecido há algum tempo, sua aplicação a dados autocorrelacionados tem sido incipiente e não têm sido feitas comparações quanto a sua eficiência e robustez nesta situação.

Este artigo apresenta procedimentos inferenciais para os índices de capacidade para características de qualidade autocorrelacionadas normalmente distribuídos via *bootstrap*. Na primeira parte do artigo são apresentados os procedimentos bootstrap para alguns modelos de séries temporais mais realistas na indústria; em seguida, estão os resultados obtidos via *bootstrap* e ilustrados através de exemplos numéricos, e, finalmente, estão as conclusões.

REAMOSTRAGEM EM ESTUDOS DE CAPACIDADE COM AUTOCORRELAÇÃO

O método de reamostragem (*Bootstrap*) tem, em sua origem, os trabalhos independentes de Efron (1979) e Simon *et al.* (1976). Sua idéia consiste em que, por detrás de todo procedimento estatístico, existe sempre um processo físico que gera os dados em análise e, portanto, permite que se trabalhe diretamente com este modelo físico, através de simulação, ao invés da forma teórica da estatística indutiva clássica.

É uma técnica geral que permite avaliar a incerteza na estimação de parâmetros, porém substitui a análise matemática pela amostragem (com reposição) dos dados amostrais originais. Na realidade, este método não deixa de ser uma forma diferente de simulação Monte Carlo, porém utilizando a própria distribuição empírica dos dados.

O método de reamostragem é constituído dos seguintes passos:

- Passo 1: considera-se uma amostra constituída de n elementos da população em estudo ($x_i, i=1, 2, 3, \dots, n$).
- Passo 2: mediante sorteio com reposição, extraem-se novas amostras (chamadas de reamostras) de r valores ($r \leq n$) a partir da amostra original ($x_{ij}^*, i=1, 2, 3, \dots, r; j=1, 2, \dots, B$).
- Passo 3: calcula-se uma estimativa pontual $\hat{\psi}^*(j)$ do parâmetro de interesse com base nas observações reamostradas no passo 2.
- Passo 4: repetem-se os passos 1 a 3 B vezes.
- Passo 5: calculam-se a média amostral ($\bar{\psi}^*$) e o desvio-padrão amostral s_{ψ}^* das estimativas.

Existem três formas de construir intervalos de confiança para o parâmetro de interesse do procedimento Bootstrap. A primeira é denominada de Método Padrão,

dada por:

$$\bar{\psi}^* \pm z_{\alpha/2} s_{\psi}^* \quad (3)$$

A segunda forma, denominada de Método Percentil, é obtida a partir das B estimativas de $\hat{\psi}^*(j)$ ordenadas em ordem crescente e utilizam-se os percentis $(\alpha/2)$ e $(1-\alpha/2)$ e, como limites, os intervalos de confiança, isto é:

$$[\hat{\psi}^*(\alpha/2); \hat{\psi}^*(1-\alpha/2)] \quad (4)$$

E a terceira forma é denominada de Método Percentil Corrigido. Ele foi desenvolvido visando corrigir alguns problemas que surgem com o método anterior, quando o estimador é viesado em relação ao parâmetro estimado. Os limites do intervalo de confiança são:

$$[\hat{\psi}^*(p_i B); \hat{\psi}^*(p_s B)] \quad (5)$$

onde $p_i = \Phi(2z_0 - z_{\alpha/2})$; $p_s = \Phi(2z_0 + z_{\alpha/2})$; $z_0 = \Phi^{-1}(p_0)$ e p_0 é a probabilidade de uma estimativa de $\hat{\psi}^*(j)$ ser inferior à estimativa pontual da amostra original.

Existem basicamente dois tipos de séries temporais: estacionárias e não-estacionárias (BOX; JENKINS; REINSEL, 1994). Por estacionária, entende-se uma série em que sua média (e dispersão) está sempre em torno de um certo valor fixo, ou seja, $X_t = \mu + e_t$, onde μ é a média do processo, suposta constante, e e_t é a variação aleatória em torno da média, assumida como sendo $N(0; \sigma^2)$, com σ^2 constante. Já no outro caso, o mesmo não ocorre, ou seja, sua média oscila livremente, não havendo sentido em falar em um determinado valor para esta. Segundo os conceitos adotados em controle estatístico de processo, uma série não-estacionária não seria previsível, já que não é possível antever seu comportamento em futuro próximo. Em outras palavras, não há sentido em tentar determinar a capacidade do processo em séries não-estacionárias.

Dentre as várias categorias existentes, são estacionárias as séries auto-regressivas de ordem p : $AR(p)$, as de média móvel de ordem q : $MA(q)$, e as auto-regressivas com média móvel de ordens (p, q) : $ARMA(p, q)$, que podem ser representadas pela equação: $X_t = \xi + \phi_1 X_{t-1} + \dots + \phi_p X_{t-p} + \varepsilon_t - \theta_1 \varepsilon_{t-1} - \dots - \theta_q \varepsilon_{t-q}$, onde ξ , ϕ_i e θ_j , $i=1, \dots, p$; $j=1, \dots, q$, são parâmetros das séries temporais. Segundo Montgomery; Runger (1994), das várias séries temporais estacionárias, merecem destaque as séries: $AR(1)$, $AR(2)$, $MA(1)$, $MA(2)$ e $ARMA(1,1)$ por serem aquelas

mais comumente encontradas no campo da engenharia e, portanto, as mais importantes no estudo de dados autocorrelacionados. Assim, em decorrência deste fato, optou-se por avaliar somente estas. Por outro lado, deve-se lembrar que uma das restrições dos modelos ARMA é a exigência de que os resíduos sejam independentes e identicamente distribuídos com média zero e variância constante, ou seja, $\varepsilon_t \sim \text{IID}(0, \sigma_\varepsilon^2)$.

Neste trabalho, somente foram analisadas as séries temporais com resíduos normais. O método originalmente proposto por Efron; Tibshirani (1986) necessita ser modificado quando aplicado a estruturas de dados mais complexas, em virtude da existência de dependência serial. É necessário eliminar a estrutura interna existente entre os dados (autocorrelação) através do ajuste de um modelo adequado ARIMA. Se tal modelo for eficiente na extração da informação contida na série temporal, os resíduos devem mostrar-se estatisticamente independentes (ruído branco). Assim sendo, o procedimento de reamostragem para dados autocorrelacionados é constituído das seguintes etapas:

- a) Gerar uma amostra com n valores segundo NIID (0,1);
- b) Gerar uma série temporal segundo o modelo ARIMA (p,q) selecionado;
- c) Obter a(s) estimativa(s) do(s) parâmetro(s) do modelo ARIMA (p,q) selecionado, pelo método dos mínimos quadrados;
- d) Ajustar o modelo aos dados, obtendo os resíduos estimados $\hat{\varepsilon}_i = X_i - \hat{X}_i$, e definir como amostra original

- ($\hat{\varepsilon}_1, \hat{\varepsilon}_2, \hat{\varepsilon}_3, \dots, \hat{\varepsilon}_n$) a estes;
- e) Obter uma reamostra $\hat{\varepsilon}_1^*, \hat{\varepsilon}_2^*, \hat{\varepsilon}_3^*, \dots, \hat{\varepsilon}_n^*$, mediante a amostragem (com reposição) dos resíduos da etapa d;
- f) Reconstituir a série temporal ARIMA (p,q), utilizando a reamostra dos resíduos $\hat{\varepsilon}_1^*, \hat{\varepsilon}_2^*, \hat{\varepsilon}_3^*, \dots, \hat{\varepsilon}_n^*$ e a(s) estimativa(s) do(s) parâmetro(s) obtida(s);
- g) Calcular as estimativas para os índices de capacidade $Cp^{(i)}$ e $Cpk^{(i)}$;
- h) Repetir as etapas a até g B vezes;
- i) Obter as médias amostrais de Cp e Cpk e os respectivos desvios-padrões s_{Cp}^* e s_{Cpk}^* ;
- j) Construir intervalos de confiança para Cp e Cpk , segundo as expressões (3) a (5).

Para ilustrar o procedimento proposto, três tamanhos de amostra foram utilizados: $n=50, 100$ e 250 , uma vez que é reconhecida a importância de tamanho da amostra na exatidão e precisão das estimativas. Tamanhos superiores a estes não costumam ser comuns na execução de estudos de capacidade de curto-prazo, em função do custo de sua obtenção e interferência na produção normal. Nos casos simulados, a quantidade de repetições (B) foi mantida constante e igual a 1000. Segundo Efron; Tibshirani (1993, p. 274-6) estimativas empregadas na construção dos intervalos de confiança via reamostragem são satisfatórias para valores de $B > 500$. Os limites inferior e superior de especificação foram fixados em -3 e $+3$, respectivamente. Na Tabela 1 estão

Tabela 1: Valor dos parâmetros das séries temporais utilizadas nas simulações.

SÉRIE	VALOR DOS PARÂMETROS	VARIÂNCIA DA CARACTERÍSTICA (σ^2)	
AR(1)	$\phi_1=0,1; 0,3; 0,5; 0,7$ e $0,9$	$\sigma_\varepsilon^2/(1-\phi_1^2)$	(6)
AR(2)	$\phi_1=0,1; 0,2; \dots$ e $0,8$ $\phi_2=0,9-\phi_1$	$\sigma_\varepsilon^2(1-\phi_2^2)(1+\phi_2^2)^{-1}[(1-\phi_2)^2-\phi_1]^{-1}$	(7)
MA(1)	$\theta_1=-1,5; -1,3; -0,9; -0,5$ e $-0,1$	$\sigma_\varepsilon^2(1+\theta_1^2)$	(8)
MA(2)	$(\theta_1, \theta_2)=(-1,5; -1,3); (-1,5; -0,9);$ $(-1,1; -0,9); (-1,1; -0,5); (-0,7; -1,3);$ $(-0,7; -0,1); (-0,3; -1,5)$ e $(-0,3; -0,1)$	$\sigma_\varepsilon^2(1+\theta_1^2+\theta_2^2)$	(9)
ARMA(1,1)	$\theta_1=0,1; 0,5$ e $0,9$ combinados com $\phi_1=0,1; 0,5$ e $0,9$	$\sigma_\varepsilon^2(1-2\phi_1\theta_1+\theta_1^2)(1-\phi_1^2)^{-1}$	(10)

os valores dos parâmetros utilizados na simulação. Foram escolhidos de modo que as séries fossem estacionárias, as autocorrelações positivas e $\xi = 0$.

Nas simulações realizadas, foram empregados dois *softwares*: MINITAB e RESAMPLING STATS. O primeiro para geração de amostras e análise estatística, tal como na obtenção de estimativas dos parâmetros dos modelos *ARIMA*. O outro *software* (RESAMPLING STATS) foi útil na fase de reamostragem dos resíduos, na recomposição da série temporal e na obtenção das distribuições de \hat{C}_p e \hat{C}_{pk} .

ANÁLISE DOS RESULTADOS OBTIDOS

Para avaliar os resultados obtidos na simulação via Bootstrap, algumas medidas foram calculadas para fazer comparações com valores teóricos e avaliar as estimativas obtidas, entre eles o erro quadrático médio (EQM) dado por

$$EQM = \sum_{i=1}^B \frac{(\hat{\psi}_i - \psi)^2}{B-1} = \sum_{i=1}^B \frac{(\hat{\psi}_i - \bar{\psi})^2}{B-1} + \frac{B(\bar{\psi} - \psi)^2}{B-1} = s_{\psi}^2 + \frac{B(\bar{\psi} - \psi)^2}{B-1}$$

onde B é o número de repetições feitas na simulação e $\hat{\psi}_i$ é o valor da i -ésima estimativa de ψ . É uma medida que combina o viés (diferença entre a estimativa média $\bar{\psi}$ e o valor verdadeiro do parâmetro ψ) e a variância de um conjunto de estimativas. Ela permitirá uma comparação do desempenho do método de Bootstrap em função do modelo selecionado [AR(p), MA(q) ou ARMA(p,q)], do valor do seu parâmetro e do tamanho da amostra.

Nas Figuras 1 a 3, estão os valores de EQMs de C_p e C_{pk} dos vários modelos de séries em função do tamanho da amostra. Nas Tabelas 2 a 6, encontram-se os resultados obtidos, no tocante às estimativas de C_p (\bar{C}_p^*) e C_{pk} (\bar{C}_{pk}^*). Constam, ainda, os valores teóricos de $C_p=C_{pk}$, os valores de σ^2 conforme as expressões (6) a (10), dependendo da série simulada e dos limites inferiores dos intervalos de confiança (IC) a 95%, segundo os três métodos: padrão, percentil e corrigido conforme expressões (3) a (5). Para fins comparativos, os limites inferiores de intervalo de confiança de C_p e C_{pk} dados, respectivamente (1) e (2), também foram calculados e identificados como Bissel nas Tabelas 2-6.

A Figura 1 mostra uma comparação entre os EQMs das estimativas de C_p e C_{pk} em função dos vários tamanhos de amostra (n) dos modelos AR(1) e AR(2). O EQM

diminui sensivelmente com o aumento do tamanho da amostra nos dois modelos auto-regressivos. Em geral, a queda é mais acentuada quando se aumenta o tamanho da amostra de 50 para 100, especialmente para valores altos de ϕ nos modelos AR(1). No modelo AR(2), os valores de EQM para C_{pk} são menores do que para C_p .

Das Tabelas 2 e 3, pode-se depreender que:

- Nos dois modelos autoregressivos, as três diferentes formas para construir IC fornecem valores muito próximos entre si, demonstrando coerência dos 3 métodos.
- No modelo AR(1), para valores baixos de ϕ (0,1 ou 0,3) o método padrão fornece valores mais próximos do IC de Bissel, no caso de C_p . Contudo, à medida que ϕ vai aumentando, o método do percentil corrigido torna-se preferível. No entanto, o método corrigido tende a fornecer ICs mais próximos dos de Bissel, com exceção de $\phi = 0,1$, quando os métodos padrão ou percentil

se revelaram melhores no caso de C_{pk} .

- No modelo AR(2), o método percentil corrigido revelou-se superior a todos os demais no caso do C_{pk} , contudo, os valores determinados por este

método indicaram sempre uma cobertura maior do que os 95% de confiança de Bissel.

Em relação aos modelos MA(q), a Figura 2 traz os EQMs de C_p e C_{pk} , em função de n e θ . Similar aos modelos AR(p), nos modelos MA(q), há um decréscimo dos EQMs com o aumento de n , mais acentuado quando esta vai de 50 para 100 nos dois índices. Estes EQMs são similares em valor tanto para C_p como C_{pk} .

Os intervalos unilaterais com 95% de confiança para C_p e C_{pk} , obtidos, via Bootstrap, dos modelos MA(q), encontram-se nas Tabelas 4 e 5, cuja análise revela os seguintes pontos de importância:

- Nos modelos MA(q), não há um método (padrão, percentil ou corrigido) que seja superior aos demais na obtenção dos ICs para C_p , já que este varia em função do n e θ selecionados. Contudo, quando $n > 50$, os valores dos ICs são muito próximos entre si para que se possa dizer que haja uma diferença prática entre estes resultados. As mesmas considerações são também válidas para o caso de C_{pk} ;
- No modelo MA(1), os ICs obtidos contiveram os valores reais dos parâmetros estimados em todos os casos analisados; no modelo MA(2), a mesma observação é válida, contudo apresentaram uma cobertura ligeiramente

Figura 1: Modelos AR(1) e AR(2) - Valores de EQM em função do tamanho da amostra: a.

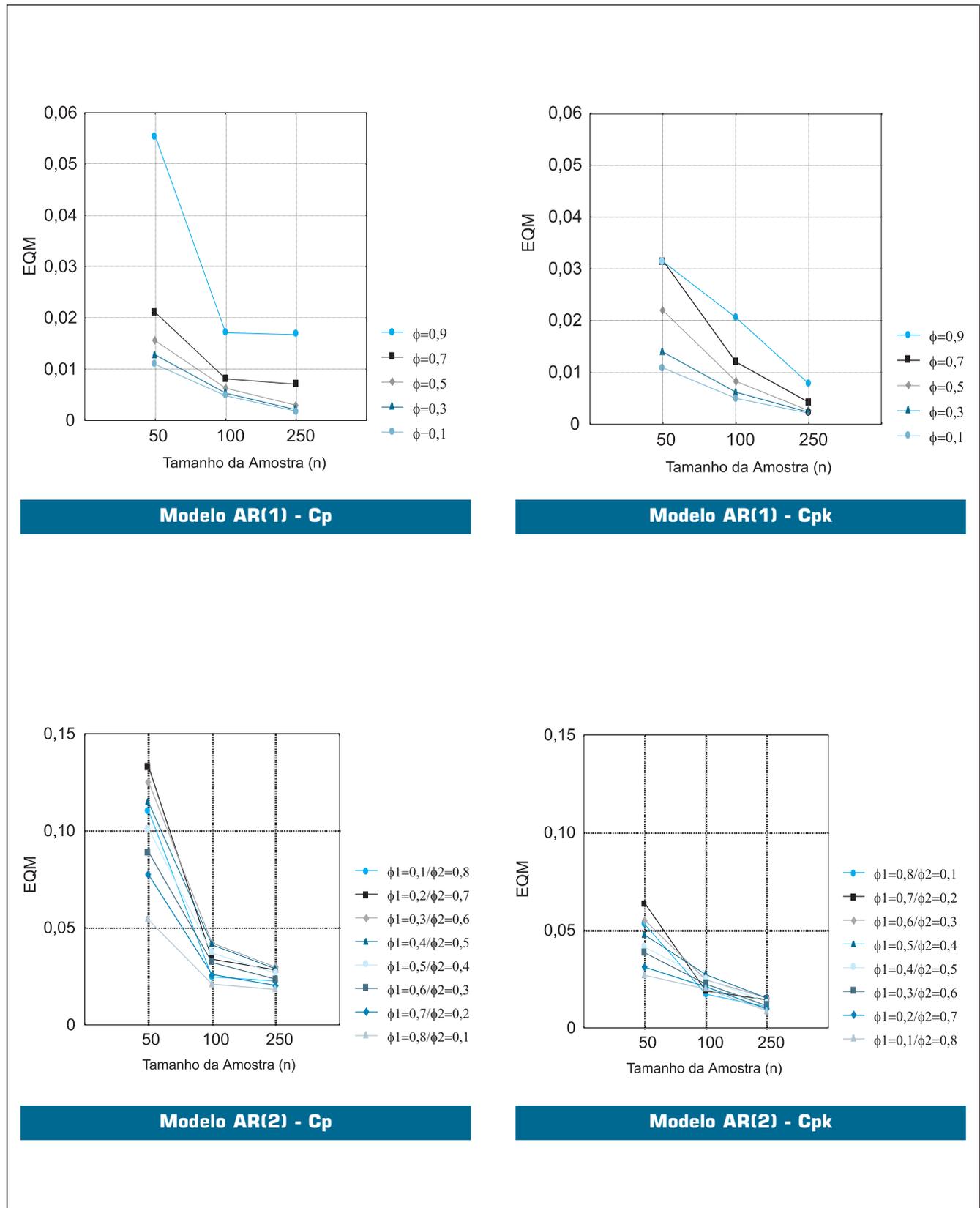


Tabela 2: Modelo AR(1) - Limite inferior dos intervalos de confiança.

ϕ	n	σ	$Cp=$ Cpk	Intervalo de confiança de Cp				Intervalo de confiança de Cpk			
				Bissel	padrão	percentil	corrigido	Bissel	padrão	percentil	corrigido
0,1	50	1,0050	0,9950	0,8239	0,8565	0,8755	0,9059	0,8082	0,8059	0,8232	0,8508
	100			0,8813	0,8976	0,9086	0,9228	0,8700	0,8655	0,8767	0,8944
	250			0,9197	0,9290	0,9322	0,9386	0,9122	0,9079	0,9133	0,9213
0,3	50	1,0483	0,9539	0,7770	0,7893	0,8118	0,8326	0,7603	0,7308	0,7394	0,7674
	100			0,8394	0,8452	0,8532	0,8661	0,8275	0,8056	0,8106	0,8186
	250			0,8948	0,8899	0,8951	0,9022	0,8771	0,8644	0,8679	0,8753
0,5	50	1,1547	0,8660	0,6952	0,6683	0,6905	0,7084	0,6764	0,5811	0,5923	0,6655
	100			0,7553	0,7452	0,7537	0,7686	0,7420	0,6922	0,6960	0,7250
	250			0,8181	0,8143	0,8170	0,8299	0,8098	0,7821	0,7835	0,7964
0,7	50	1,4003	0,7141	0,5827	0,5085	0,5362	0,5594	0,5604	0,3751	0,3976	0,4023
	100			0,6169	0,5822	0,5890	0,6081	0,6009	0,5032	0,5048	0,5227
	250			0,7060	0,6897	0,6902	0,7000	0,6965	0,6463	0,6478	0,6576
0,9	50	2,2942	0,4359	0,4683	0,3692	0,4034	0,4266	0,4413	0,1285	0,1396	0,1928
	100			0,3883	0,3373	0,3592	0,3717	0,3650	0,1631	0,1549	0,1805
	250			0,4840	0,4519	0,4558	0,4639	0,4709	0,3740	0,3668	0,3867

Tabela 3: Modelo AR(2) - Limite inferior dos intervalos de confiança.

ϕ_1	ϕ_2	n	σ	$Cp=$ Cpk	Intervalo de confiança de Cp				Intervalo de confiança de Cpk			
					Bissel	padrão	percentil	corrigido	Bissel	padrão	percentil	corrigido
0,1	0,8	50	1,9245	0,5196	0,6243	0,6127	0,6301	0,6569	0,6034	0,4301	0,4356	0,4639
		100			0,4847	0,4610	0,4778	0,4972	0,4650	0,2997	0,2980	0,3110
		250			0,5895	0,5503	0,5560	0,5671	0,5783	0,4675	0,4628	0,4986
0,2	0,7	50	1,8787	0,5323	0,6401	0,6224	0,6342	0,6617	0,6197	0,4183	0,4312	0,4529
		100			0,5038	0,4883	0,4902	0,5167	0,4847	0,3153	0,3097	0,3114
		250			0,6122	0,5708	0,5757	0,5833	0,6013	0,4915	0,4931	0,5021
0,3	0,6	50	1,8898	0,5291	0,6304	0,6009	0,6147	0,6538	0,6097	0,4192	0,4192	0,4402
		100			0,5012	0,4826	0,4947	0,5166	0,4821	0,2898	0,2990	0,3163
		250			0,6104	0,5679	0,5692	0,5904	0,5996	0,4832	0,4841	0,5065
0,4	0,5	50	1,9245	0,5196	0,6119	0,5785	0,5935	0,6144	0,5906	0,3767	0,3770	0,4220
		100			0,4893	0,4563	0,4587	0,4859	0,4697	0,2521	0,2607	0,2854
		250			0,5981	0,5520	0,5566	0,5646	0,5871	0,4693	0,4685	0,4858
0,5	0,4	50	1,9739	0,5066	0,5881	0,5263	0,5404	0,5612	0,5660	0,3348	0,3385	0,3687
		100			0,4725	0,4456	0,4627	0,4806	0,4524	0,2386	0,2357	0,2648
		250			0,5801	0,5353	0,5435	0,5495	0,5688	0,4573	0,4572	0,4821
0,6	0,3	50	2,0352	0,4914	0,5615	0,4857	0,5010	0,5359	0,4325	0,2972	0,3071	0,3419
		100			0,4533	0,4106	0,4238	0,4566	0,5471	0,2304	0,2378	0,2643
		250			0,5588	0,5212	0,5251	0,5409	0,5082	0,4384	0,4364	0,4444
0,7	0,2	50	2,1082	0,4743	0,5324	0,4592	0,4897	0,5202	0,5082	0,2648	0,2730	0,3104
		100			0,4324	0,3913	0,3971	0,4212	0,4109	0,2128	0,2100	0,2141
		250			0,5784	0,4936	0,5019	0,5120	0,5234	0,4188	0,4198	0,4454
0,8	0,1	50	2,1938	0,4558	0,5013	0,4231	0,4354	0,4729	0,4758	0,2001	0,1984	0,2548
		100			0,4107	0,3684	0,3789	0,4028	0,3883	0,1907	0,1770	0,2214
		250			0,5104	0,4725	0,4797	0,4876	0,4979	0,3999	0,4034	0,4151

Figura 2: Modelos MA(1) e MA(2) - Valores de EQM em função do tamanho da amostra.

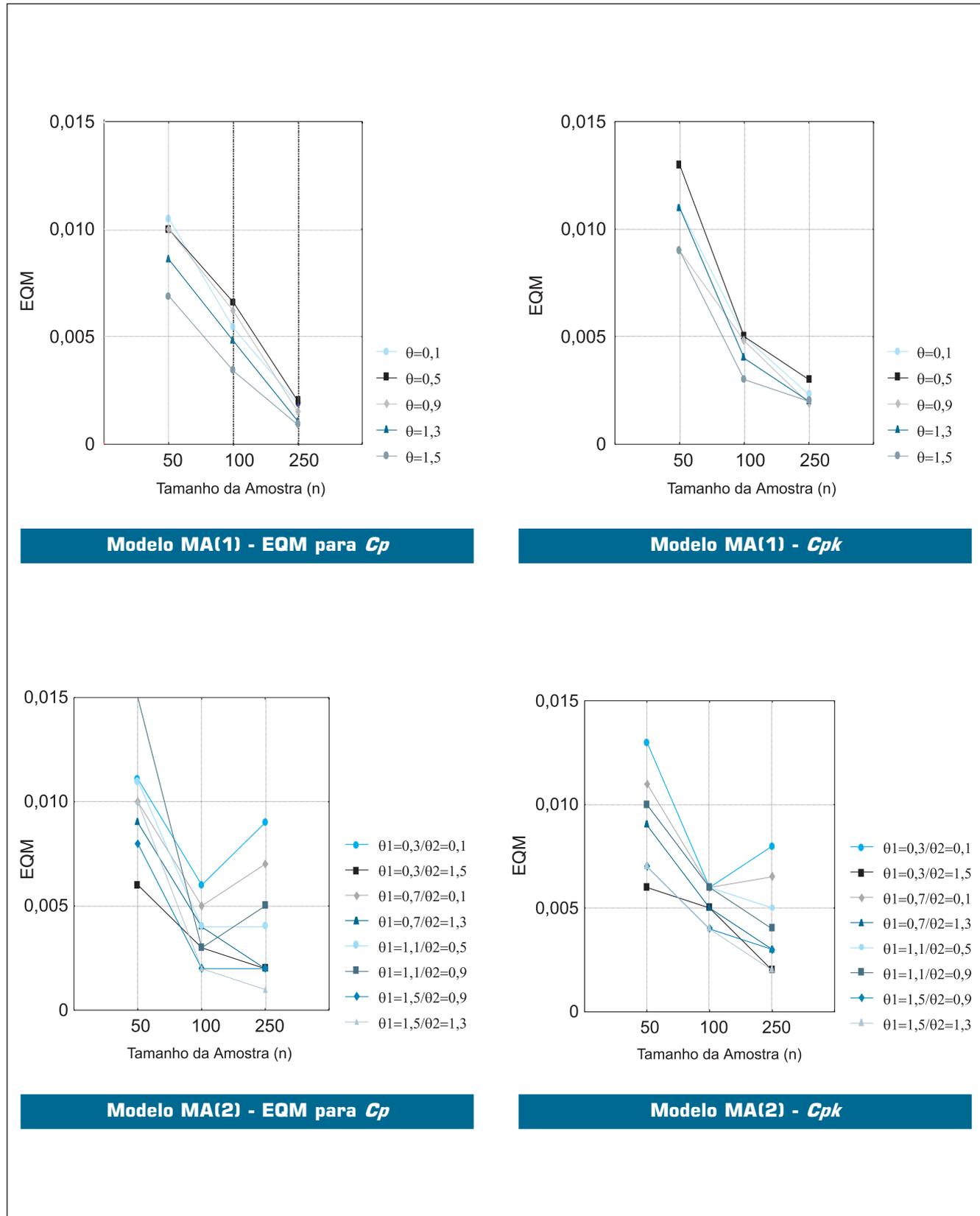


Tabela 4: Modelo MA(1) - Limite inferior dos intervalos de confiança

θ	n	σ	$Cp=$ Cpk	Intervalo de confiança de Cp				Intervalo de confiança de Cpk			
				Bissel	padrão	percentil	corrigido	Bissel	padrão	percentil	corrigido
0,1	50	1,0050	0,9950	0,8239	0,8517	0,8693	0,8983	0,8082	0,8064	0,8229	0,8475
	100			0,8822	0,8997	0,9056	0,9213	0,8709	0,8727	0,8787	0,8925
	250			0,9190	0,9253	0,9296	0,9387	0,9116	0,9037	0,9060	0,9176
0,5	50	1,1180	0,8944	0,7319	0,7341	0,7428	0,7761	0,7141	0,6788	0,6917	0,7165
	100			0,7933	0,8142	0,8210	0,8305	0,7807	0,7737	0,7801	0,7920
	250			0,8203	0,8180	0,8183	0,8264	0,8120	0,7918	0,7911	0,8027
0,9	50	1,3454	0,7433	0,6123	0,6320	0,6488	0,6791	0,5910	0,5717	0,5849	0,6092
	100			0,6593	0,6722	0,6819	0,7040	0,6442	0,6312	0,6386	0,6601
	250			0,6804	0,6861	0,6925	0,6959	0,6705	0,6585	0,6622	0,6695
1,3	50	1,6401	0,6097	0,5080	0,4704	0,4912	0,5178	0,4828	0,4103	0,4270	0,4562
	100			0,5412	0,5437	0,5517	0,5723	0,5232	0,5013	0,5066	0,5224
	250			0,5584	0,5567	0,5594	0,5689	0,5467	0,5280	0,5289	0,5388
1,5	50	1,8028	0,5547	0,4649	0,4351	0,4528	0,4686	0,4377	0,3745	0,3894	0,4079
	100			0,4923	0,5003	0,5055	0,5127	0,4728	0,4539	0,4586	0,4645
	250			0,5085	0,5053	0,5045	0,5107	0,4958	0,4765	0,4751	0,4793

Tabela 5: Modelo MA(2) - Limite inferior dos intervalos de confiança

θ_1	θ_2	n	σ	$Cp=$ Cpk	Intervalo de confiança de Cp				Intervalo de confiança de Cpk			
					Bissel	padrão	percentil	corrigido	Bissel	padrão	percentil	corrigido
0,3	0,1	50	1,0488	0,9535	0,7784	0,7980	0,8161	0,8354	0,7602	0,7413	0,7538	0,7803
		100			0,8386	0,8509	0,8600	0,8762	0,8267	0,8082	0,8170	0,8341
		250			0,8804	0,8146	0,8224	0,8516	0,8726	0,7955	0,7906	0,8147
	1,5	50	1,8276	0,5472	0,4615	0,4674	0,4774	0,4941	0,4342	0,4004	0,4047	0,4184
		100			0,4668	0,4633	0,4692	0,4809	0,4464	0,4127	0,4169	0,4257
		250			0,5133	0,4931	0,4970	0,5089	0,5007	0,4622	0,4642	0,4750
0,7	0,1	50	1,2247	0,8165	0,6662	0,6625	0,6830	0,7070	0,6467	0,6073	0,6226	0,6456
		100			0,7201	0,7171	0,7255	0,7272	0,7063	0,6709	0,6769	0,7011
		250			0,7505	0,6926	0,7020	0,7112	0,7414	0,6709	0,6700	0,7112
	1,3	50	1,7833	1,7860	0,4659	0,4530	0,4685	0,4958	0,4388	0,3800	0,3900	0,4958
		100			0,4802	0,4730	0,4835	0,4933	0,4604	0,4198	0,4189	0,4368
		250			0,5219	0,4792	0,4830	0,4910	0,5096	0,4473	0,4423	0,4557
1,1	0,5	50	1,5684	0,6376	0,5227	0,5339	0,5511	0,5956	0,4981	0,4624	0,4764	0,5000
		100			0,5559	0,5400	0,5477	0,5591	0,5383	0,4916	0,4965	0,5090
		250			0,5870	0,5377	0,5430	0,5520	0,5758	0,5102	0,4995	0,5190
	0,9	50	1,7378	0,5754	0,4736	0,4948	0,5116	0,5394	0,4469	0,4140	0,4261	0,4506
		100			0,4976	0,4861	0,4946	0,5007	0,4783	0,4306	0,4349	0,4549
		250			0,5312	0,4944	0,5049	0,5229	0,5190	0,4678	0,4571	0,4742
1,5	0,9	50	2,0149	0,4963	0,4115	0,4210	0,4323	0,4521	0,3816	0,3382	0,3444	0,3649
		100			0,4320	0,4284	0,4353	0,4427	0,4104	0,3736	0,3757	0,3849
		250			0,4570	0,4134	0,4203	0,4276	0,4433	0,3845	0,3784	0,3867
	1,3	50	2,2226	0,4499	0,3750	0,3928	0,4062	0,4275	0,3429	0,3080	0,3101	0,3449
		100			0,3903	0,3808	0,3877	0,3968	0,3671	0,3229	0,3227	0,3346
		250			0,4150	0,4002	0,4042	0,4111	0,4003	0,3634	0,3613	0,3713

superior àquela fornecida pelos ICs de Bissel (95% de confiança).

Os resultados do modelo $ARMA(1,1)$ estão representados na Figura 3. Há queda nos EQMs com o aumento de n e esta é mais acentuada quando o tamanho da amostra vai de 50 para 100, mas não em todas as combinações de ϕ e θ .

Os intervalos unilaterais para Cp e Cpk , com confiança de 95%, segundo estes três métodos de *bootstrap*, estão na Tabela 6, cuja análise permite depreender que:

- Em geral, no caso de Cp , o método padrão é aquele que proporciona resultados mais próximos do IC unilateral de Bissel (exceto para $\phi = 0,1$ e $\theta = 0,1$) e, quando n aumenta, os três diferentes métodos fornecem resultados semelhantes. Contudo, para ϕ maiores, os valores dos ICs tendem a ficar muito diferentes do valor teórico, proporcionando uma cobertura inferior aos 95% de confiança;
- No caso de Cpk , é também o método padrão que fornece valores mais próximos de Bissel para o IC, embora haja

algumas situações em que o obtido pelo método corrigido se revela ligeiramente superior a este;

- Por várias vezes os ICs obtidos não contiveram o valor teórico do parâmetro estimado.

CONCLUSÕES

Os métodos tradicionais existentes para avaliação de capacidade de processo falham quando os dados se encontram autocorrelacionados. Os procedimentos para realizar inferências com os índices de capacidade devem ser modificados de forma a compensar não somente a variabilidade do processo onde os dados são coletados, mas também a dispersão adicional introduzida pela correlação serial.

A utilização da técnica Bootstrap é uma alternativa viável para contornar a dificuldade de realizar inferências com os índices Cp e Cpk . As simulações revelaram que há uma sensível redução do erro quadrático médio (EQM)

Figura 3: Modelo $ARMA(1,1)$:EQM para Cp e Cpk .

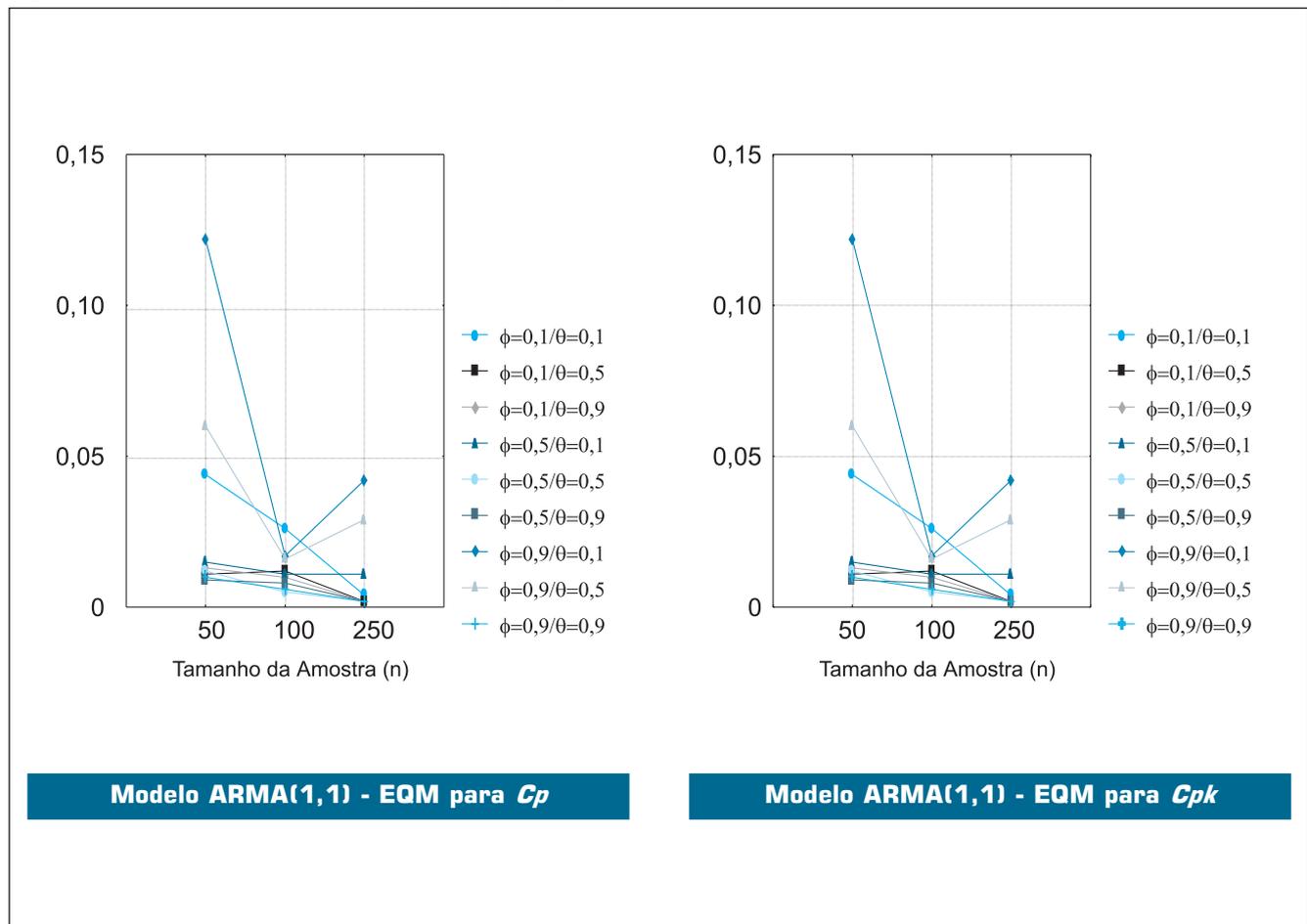


Tabela 6: Modelo ARMA(1,1) - Limite inferior dos intervalos de confiança

ϕ	θ	n	σ	$Cp=$ Cpk	Intervalo de confiança de Cp				Intervalo de confiança de Cpk			
					Bissel	padrão	percentil	corrigido	Bissel	padrão	percentil	corrigido
0,1	0,1	50	1,0000	1,0000	0,8063	0,6367	0,6532	0,6820	0,7903	0,5561	0,5804	0,6009
		100			0,8674	0,7316	0,7399	0,7519	0,8559	0,6839	0,6919	0,7049
		250			0,9047	0,8849	0,8913	0,9004	0,8972	0,8740	0,8817	0,8869
	0,5	50	1,0778	0,9278	0,6987	0,7387	0,7577	0,7753	0,6800	0,7273	0,7423	0,7700
		100			0,7566	0,9037	0,9096	0,9299	0,7434	0,8805	0,8902	0,9056
		250			0,7871	0,8468	0,8490	0,8618	0,7785	0,8390	0,8390	0,8457
	0,9	50	1,2831	0,7793	0,5799	0,6813	0,7000	0,7200	0,5575	0,6756	0,6933	0,7151
		100			0,6230	0,7366	0,7492	0,7601	0,6071	0,7322	0,7452	0,7569
		250			0,6478	0,7227	0,7252	0,7401	0,6375	0,7211	0,7237	0,7326
0,5	0,1	50	1,1015	0,9078	0,6589	0,6946	0,7199	0,7445	0,6391	0,6038	0,6154	0,6530
		100			0,7155	0,7048	0,7146	0,7259	0,7015	0,6458	0,6505	0,6730
		250			0,7788	0,9309	0,9331	0,9449	0,7701	0,9076	0,9084	0,9207
	0,5	50	1,0000	1,0000	0,5270	0,8901	0,9036	0,9246	0,5026	0,8513	0,8640	0,8880
		100			0,5681	0,8941	0,8993	0,9187	0,5508	0,8517	0,8549	0,8714
		250			0,6243	0,9289	0,9334	0,9408	0,6137	0,9148	0,9173	0,9261
	0,9	50	1,1015	0,9078	0,4276	0,8080	0,8180	0,8456	0,3985	0,7974	0,8101	0,8301
		100			0,4569	0,8524	0,8601	0,8763	0,4362	0,8437	0,8519	0,8727
		250			0,5027	0,8297	0,8316	0,8442	0,4899	0,8270	0,8282	0,8422
0,9	0,1	50	2,0901	0,4785	0,4341	0,5468	0,5594	0,5813	0,4053	0,3657	0,3585	0,3850
		100			0,3560	0,3696	0,3972	0,4150	0,3312	0,1998	0,2027	0,2343
		250			0,4454	0,5704	0,5753	0,5841	0,4314	0,5056	0,5092	0,5160
	0,5	50	1,3572	0,7368	0,3314	0,7644	0,7796	0,8153	0,2963	0,6248	0,6380	0,6586
		100			0,2653	0,5711	0,5861	0,6042	0,2356	0,4364	0,4508	0,4713
		250			0,3337	0,8055	0,8055	0,8155	0,3167	0,7490	0,7497	0,7597
	0,9	50	1,0000	1,0000	0,2654	0,8117	0,8209	0,8385	0,2251	0,7223	0,7335	0,7571
		100			0,2102	0,9039	0,9129	0,9222	0,1768	0,8585	0,8642	0,8753
		250			0,2645	0,9311	0,9347	0,9405	0,2451	0,9204	0,9226	0,9284

quando o tamanho da amostra (n) passa de 50 para 100. Esta queda já não é tão acentuada quando n aumenta de 100 para 250 elementos, quando a distribuição normal adere bem aos dados.

Dos três diferentes métodos avaliados na construção de intervalos de confiança (ICs): padrão, percentil e percentil corrigido, há uma tendência de o primeiro gerar ICs mais fechados e do último, mais abertos. Nenhum dos três se revelou sistematicamente superior aos demais para qualquer um dos modelos ARMA (p,q) simulados, em termos de obtenção de limites próximos àqueles fornecidos pelos ICs teóricos.

Nos modelos AR(1), com $\phi < 0,5$, o método padrão demonstrou ser ligeiramente superior ao método corrigido; no modelo AR(2), os métodos percentil e corrigido são superiores ao método padrão.

Nos modelos MA(1) e MA(2), não há método que tenha demonstrado ser melhor em todas as situações analisadas. Nos modelos ARMA (1,1), quando ϕ e θ aumentam, qualquer um dos métodos avaliados conduz a ICs muito mais fechados do que o IC teórico. Embora haja uma vantagem do método padrão sobre os demais, a simulação bootstrap demonstrou um fraco desempenho nesta última situação.

Artigo recebido em 21/11/2002

Aprovado para publicação em 10/09/2003

■ Referências Bibliográficas

- AMERICAN SOCIETY FOR QUALITY CONTROL *Statistical process control manual*. Milwaukee, ASQC, 1986.
- BISSELL, A. F. How reliable is your capability index? *Applied Statistics*, London, v. 39, n. 3, p. 331-40, 1990.
- BOX, G. E. P.; HUNTER, W. G.; HUNTER, J. S. *Statistics for experimenters*. New York: Wiley, 1978.
- BOX, G. E. P.; JENKINS, G. M.; REINSEL, G. C. *Time series analysis, forecasting and control*. 3 ed. Englewood Cliffs: Prentice Hall, 1994.
- CHENG, S. W. Practical implementation of the process capability indices. *Quality Engineering*, Monticello, v.7, n.2, p.239-59, 1994.
- CONSTABLE, G. K.; YEN, V. Autocorrelation and control. In: ASQC QUALITY CONGRESS TRANSACTIONS, 45th, Milwaukee, 1991. *Proceedings*. Milwaukee, ASQC, p.375-82, 1991.
- CRYER, J. D.; RYAN, T. P. The estimation of sigma for an x chart: \bar{MR}/d_2 or s/c_4 ? *Journal of Quality Technology*, Milwaukee, v.22, n.3, p.187-92, 1990.
- EFRON, B. Bootstrap methods: another look at the jackknife. *The Annals of Statistics*, Hayward, v.7, n.1, p.1-26, 1979.
- EFRON, B.; TIBSHIRANI, R. Bootstrap methods for standard errors, confidence intervals, and other measures of statistical accuracy. *Statistical Science*, Hayward, v.1, n.1, p. 54-77, 1986.
- EFRON, B.; TIBSHIRANI, R. *An introduction to the bootstrap*. New York, Chapman & Hall, 1993.
- FRANKLIN, L. A.; WASSERMAN, G. S. Bootstrap lower confidence limits for capability indices. *Journal of Quality Technology*, Milwaukee, v. 24, n. 4, p. 196-210, 1992.
- KOTZ, S.; JOHNSON, N. L. *Process capability indices*. London: Chapman & Hall, 1993.
- KOTZ, S.; LOVELACE, C. R. *Process Capability indices in theory and practice*, London, Arnold, 1998.
- LEVINSON, W.A. Exact confidence limits for process capabilities. *Quality Engineering*, Monticello, vol. 9, n. 3, p. 521-8, 1997a.
- LEVINSON, W. A. Approximate confidence limits for Cpk and confidence limits for non-normal process capabilities. *Quality Engineering*, Monticello, vol. 9, n. 4, p. 635-40, 1997b.
- MINITAB, INC. *Minitab statistical software, version 12*. Pennsylvania, Minitab, s.d.
- MONTGOMERY, D. C. *Introduction to statistical quality control*. 3 ed. New York: Wiley, 1996
- MONTGOMERY, D. C.; RUNGER, G. C. *Statistical process control for the chemical and process industries*. Kirkland, Statistical Productivity, 1994.
- NETO, A. C.; SOUZA, R. C. A bootstrap simulation in ARMA (p, q) structures. *J. of Forecasting*, vol. 15, n. 4, p. 343-353.
- PYZDEK, T. *Pyzdek's guide to SPC*. Milwaukee, ASQC Quality, v. 2, 1992.
- RESAMPLING STATS *Resampling stats software, version 4*. Arlington, Resampling Stats, 1997.
- RÓDRIGUEZ, R. N. Recent developments in process capability analysis. *Journal of Quality Technology*, Milwaukee, v. 24, n. 2, p. 176-87, 1992.
- SCHNEIDER, H.; PRUETT, J. M. Control charting issues in the process industries. *Quality Engineering*, Monticello, v. 6, n. 2, p. 347-73, 1994.
- SHORE, H. Process capability analysis when data are autocorrelated. *Quality Engineering*, Monticello, v. 9, n. 4, p. 615-26, 1997.
- SIMON, J. L.; ATKINSON, D. T.; SHEVOKAS, C. Probability and Statistics: Experimental results of a radical different teaching method. *The American Mathematical Monthly*, Washington, v. 83, n. 9, p. 733-9, 1976.
- WOODALL, W. H.; FALTIN, F. W. Autocorrelated data and SPC. *ASQC Statistics Division Newsletter*, Milwaukee, v. 13, n. 4, p. 18-21, 1994.
- ZHANG, N.F. Estimating process capability indices for autocorrelated processes. *Proceedings of the Section on Quality and Productivity*, Alexandria, p. 49-54, 1996.

■ Agradecimentos

Os autores agradecem as valiosas sugestões dadas pelos avaliadores anônimos.