

Estudo sobre o efeito da Autocorrelação de Modelos AR(1) no Controle Estatístico de Processo

Fernando de Jesus Moreira Junior (UFRGS) – fmjunior@ufrgs.br
Carla Schwengber ten Caten (UFRGS) – tencaten@producao.ufrgs.br

Resumo

Este trabalho apresenta um estudo sobre o efeito da autocorrelação em séries que apresentam um comportamento auto-regressivo de primeira ordem AR(1) no Controle Estatístico de Processo (CEP). Para tanto, foram realizadas simulações de séries temporais em determinados níveis de correlação em modelos AR(1), e analisados os resultados obtidos. Palavras chave: Autocorrelação, Modelos Auto-Regressivos de Primeira Ordem, Controle Estatístico de Processo.

1. Introdução

Os Gráficos de Controle (Gráficos de Shewhart), ferramentas surgidas no advento do Controle Estatístico de Processo (CEP), têm sido amplamente utilizados nos dias de hoje. Essa técnica, criada originalmente por Shewhart, na década de 20 nos Estados Unidos, foi sendo aperfeiçoada com o passar do tempo e, por sua eficácia na monitoração de processos, passou a ser utilizada nos mais diversos ramos. Para Deming (1990), essa é a maneira mais fácil de examinar os dados para questionar o estado de controle estatístico que produziu esses dados. As suposições necessárias para a implantação de um gráfico de controle (ou carta de controle) são que as observações a serem monitoradas devem ser independentes e normalmente distribuídas. Entretanto, muitas vezes na prática essas suposições não são atendidas ou verificadas. Segundo Montgomery (1997), quando a suposição de normalidade é violada leve ou moderadamente, esses gráficos de controle funcionam razoavelmente bem, entretanto quando a suposição de independência dos dados é violada, ou seja, a série apresenta autocorrelação (ou correlação serial), o gráfico de controle não funciona corretamente, podendo indicar muitos “falsos alarmes”, isto é, indicando possíveis causas especiais quando na verdade são causas comuns, e não identificando causas especiais reais. Para que esses dados possam ser utilizados no CEP, é necessário primeiro tratar da questão da autocorrelação e, depois, então, aplicar o CEP tradicional.

Uma forma conveniente para o tratamento da autocorrelação é a utilização de modelos ARIMA (Auto-Regressivo Integrado e de Média Móvel). O ajuste de um modelo ARIMA adequado gera resíduos independentes e normalmente distribuídos, que podem ser monitorados com as técnicas tradicionais de CEP. Moreira (2003) apresenta uma coletânea de diversas técnicas para o tratamento de dados autocorrelacionados.

Este trabalho apresenta um estudo sobre o efeito da autocorrelação de modelos Auto-Regressivos de Primeira Ordem, AR(1), no Controle Estatístico de Processos em gráficos de controle para valores individuais e para amplitude móvel. Para tanto, foram realizadas simulações de modelos AR(1) em CEP com autocorrelação positiva e negativa, forte e moderada. Os resultados são apresentados no item 3.

2. Modelos Auto-Regressivos de Primeira Ordem: AR(1)

Os modelos AR(1), Auto-Regressivos de Primeira Ordem, são relativamente simples e

bastante encontrados na prática, principalmente em indústrias químicas. Montgomery e Mastragelo (1991) apresentam um estudo onde um parâmetro crítico de um processo químico possui comportamento descrito por um modelo AR(1) com uma forte correlação positiva.

No CEP tradicional, as observações x_t podem ser modeladas pela equação

$$x_t = \mu + \varepsilon_t, \text{ onde} \quad (2.1)$$

μ é a média do processo (média constante), e

ε_t é um erro aleatório (independente e normalmente distribuído com média zero e desvio padrão σ).

Em um processo AR(1), as observações x_t são modeladas através da equação

$$x_t = \xi + \phi x_{t-1} + \varepsilon_t, \quad (2.2)$$

onde ξ e ϕ são constantes desconhecidas e $(-1 < \phi < 1)$.

Esse processo possui média $\xi/(1-\phi)$ e desvio padrão $\sigma/(1-\phi^2)^{1/2}$.

A característica fundamental de um processo AR(1) se resume no fato da observação atual estar correlacionada com a primeira observação anterior, ou seja, há uma correlação significativa na primeira defasagem (ou no primeiro *lag*), isto é, entre x_t e x_{t-1} . Como x_{t-1} também está relacionada com x_{t-2} , há indiretamente uma correlação na segunda defasagem, entre x_t e x_{t-2} . Entretanto, no caso de séries AR(1), essa correlação está implícita na primeira defasagem. Essa seqüência de correlações em sucessivas k defasagens pode ser expressa graficamente pelo Gráfico de Correlograma da FAC (Função de Autocorrelação). A FAC é definida por

$$\rho_k = \frac{\gamma_k}{\gamma_0} = \frac{\text{cov}(x_t, x_{t-k})}{\text{var}(x_t)} \quad (2.3)$$

e estimada por

$$r_k = \frac{\hat{\gamma}_k}{\hat{\gamma}_0} = \frac{\sum_{t=1}^{n-k} (x_t - \bar{x})(x_{t-k} - \bar{x})}{\sum_{i=1}^n (x_i - \bar{x})^2}. \quad (2.4)$$

Esse coeficiente de correlação também pode ser obtido através da regressão simples

$$x_t = \rho_k x_{t-k} + \varepsilon_t. \quad (2.5)$$

Um outro Gráfico de Correlograma apresenta a Função de Autocorrelação Parcial (FACP). Denotada por ϕ_{kk} , a FACP, semelhantemente a FAC, mede a correlação entre as observações que estão k períodos afastados, entretanto remove o efeito das correlações intermediárias, ou seja, é a correlação entre x_t e x_{t-k} depois de remover a influência de $x_{t-1}, x_{t-2}, \dots, x_{t-k+1}$. Esse coeficiente de correlação é obtido através da regressão múltipla

$$x_t = \phi_{k1} x_{t-1} + \phi_{k2} x_{t-2} + \dots + \phi_{kk} x_{t-k} \quad (2.6)$$

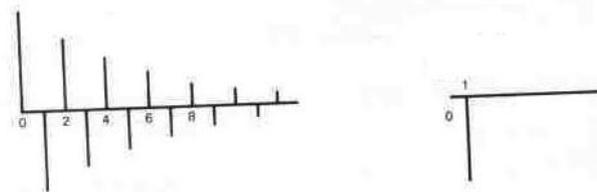
Em um processo AR(1), o comportamento da FAC é semelhante a uma queda exponencial, quando $\phi > 0$; e semelhante a uma queda exponencial alternando entre valores positivos e negativos, quando $\phi < 0$. Já a FACP apresenta apenas um pico significativo na primeira

defasagem: positivo se $\phi > 0$, e negativo se $\phi < 0$. As figuras 1 e 2 apresentam esses padrões de comportamento. Nota-se que esse valor significativo é o mesmo da FAC.



Fonte: (Adaptado de Fischer, 1982)

Figura 1 – Padrões de comportamento dos Gráficos dos Correlogramas da FAC e da FACP, respectivamente, para $\phi > 0$



Fonte: (Adaptado de Fischer, 1982)

Figura 1 – Padrões de comportamento dos Gráficos dos Correlogramas da FAC e da FACP, respectivamente, para $\phi < 0$

Fischer (1982) apresenta padrões de correlogramas da FAC e da FACP para diversos modelos ARIMA. Outras técnicas para a identificação dos Modelos ARIMA podem ser encontradas em Toloí e Morettin (1985).

Enquanto que as verdadeiras observações possuem o comportamento descrito pela equação 2.2, o modelo para o ajuste dos dados é dado pela equação

$$\hat{x}_t = \hat{\phi}_1 x_{t-1} + \hat{\xi}_t \quad (2.7)$$

Nota-se que a variável independente não é estimada, mas obtida das observações originais. Esse fato permite que as previsões só possam ser obtidas “um passo a frente”, ou seja, não é possível fazer uma previsão maior que um período.

Considerando que os estimadores $\hat{\phi}$ e $\hat{\xi}$ sejam adequados (não tendenciosos), das equações 2.2 e 2.7 obtemos os erros ou resíduos

$$\varepsilon_t = x_t - \hat{x}_t \quad (2.8)$$

Se o Modelo for adequado aos dados, esses resíduos serão independentes e normalmente distribuídos. Dessa forma as técnicas tradicionais de CEP podem ser aplicadas aos resíduos, e os gráficos de controle funcionarão corretamente para a detecção de causas especiais.

3. Simulação dos Modelos AR(1)

Para analisar o efeito da autocorrelação de Modelos AR(1) no CEP, foram realizadas 8 simulações descritas na tabela 1. O erro utilizado nas simulações possuía distribuição normal com média igual a zero e desvio padrão igual a 1 unidade. O valor do termo constante foi escolhido de tal forma que a média teórica das séries fosse igual a 10 unidades.

Para cada um dos modelos simulados, obteve-se o Gráfico de Controle dos valores individuais e da amplitude móvel, para as observações originais e para os resíduos, para tamanhos de amostra de amplitude móvel igual a 2. Entretanto, esse trabalho apresenta apenas os gráficos de controle dos casos considerados críticos que são os modelos 1-a e 2-a (autocorrelação

positiva sem a presença de causas especiais), e 3-b e 4-b (autocorrelação negativa com a presença de causas especiais). Esses gráficos são apresentados nas figuras 3, 4, 5 e 6. Para o ajuste dos modelos AR(1), foi utilizado o método dos mínimos quadrados através de regressão simples. Os gráficos de correlograma da FAC e da FACP são apresentados nas figuras 7, 8 e 9 e 10.

Modelo Simulado AR(1)	Presença de Causas Especiais	Grau da Autocorrelação
1-a) $x_t = 0,9x_{t-1} + 1 + \varepsilon_t$	Não	Positiva e forte
1-b) $x_t = 0,9x_{t-1} + 1 + \varepsilon_t$	Sim	Positiva e forte
2-a) $x_t = 0,5x_{t-1} + 5 + \varepsilon_t$	Não	Positiva e moderada
2-b) $x_t = 0,5x_{t-1} + 5 + \varepsilon_t$	Sim	Positiva e moderada
3-a) $x_t = -0,5x_{t-1} + 15 + \varepsilon_t$	Não	Negativa e moderada
3-b) $x_t = -0,5x_{t-1} + 15 + \varepsilon_t$	Sim	Negativa e moderada
4-a) $x_t = -0,9x_{t-1} + 19 + \varepsilon_t$	Não	Negativa e forte
4-b) $x_t = -0,9x_{t-1} + 19 + \varepsilon_t$	Sim	Negativa e forte

Tabela 1 – Modelos AR(1) Simulados

Segundo Ribeiro e ten Caten (2001), quando monitora-se os valores individuais não pode-se levar em consideração o teorema do limite central. Uma das consequências disso é a presença de correlação nos pontos da carta de amplitude móvel. Devido a esse fato, os gráficos de controle para a amplitude móvel não serão analisados.

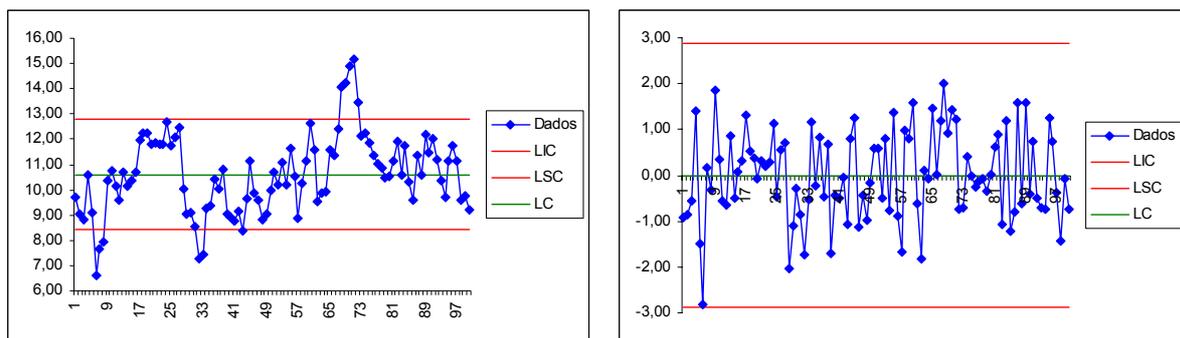


Figura 3 – Gráficos de Controle para as observações e os resíduos do Modelo 1-a, respectivamente

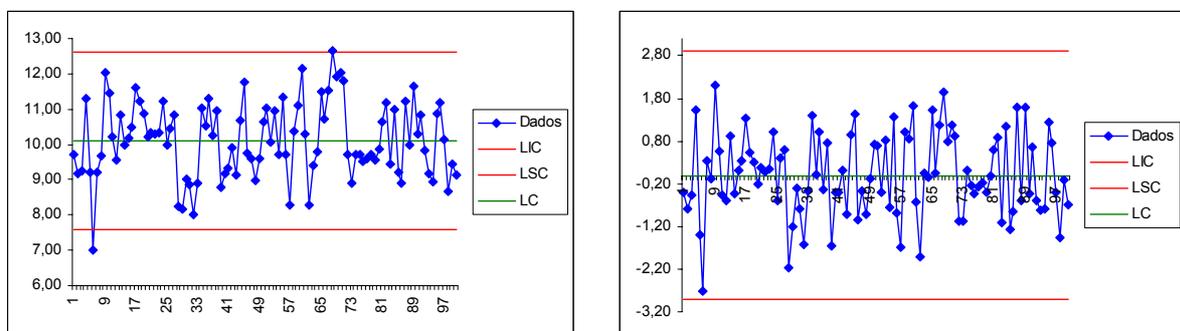


Figura 4 – Gráficos de Controle para as observações e os resíduos do Modelo 2-a, respectivamente

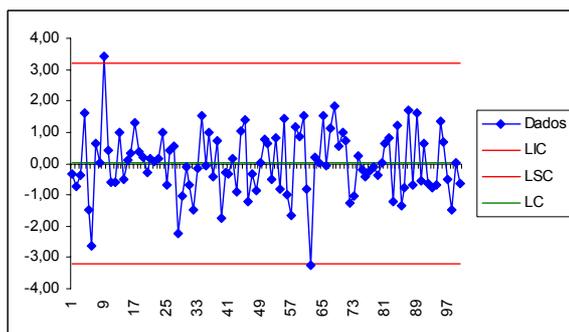
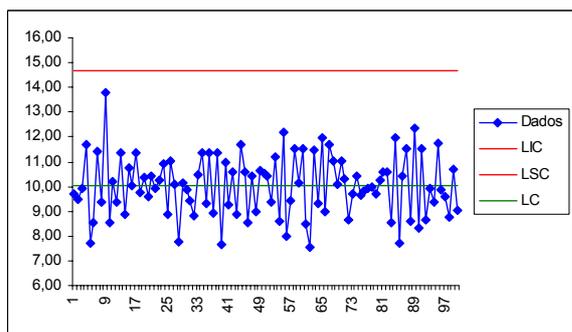


Figura 5 – Gráficos de Controle para as observações e os resíduos do Modelo 3-b, respectivamente

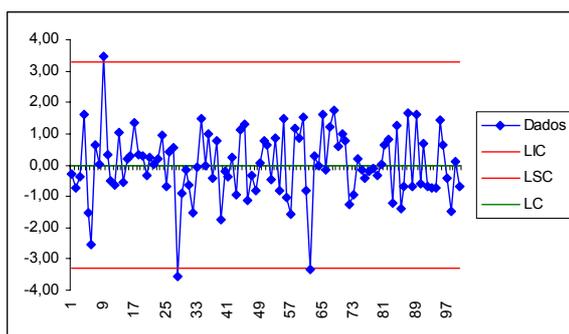
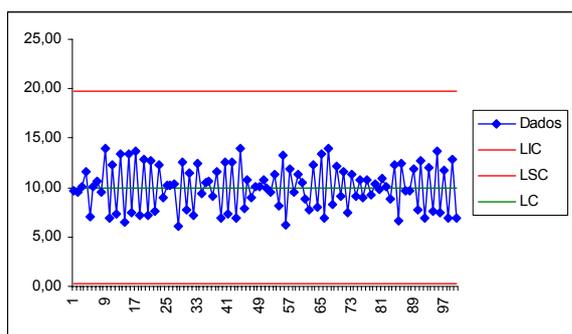


Figura 6 – Gráficos de Controle para as observações e os resíduos do Modelo 4-b, respectivamente

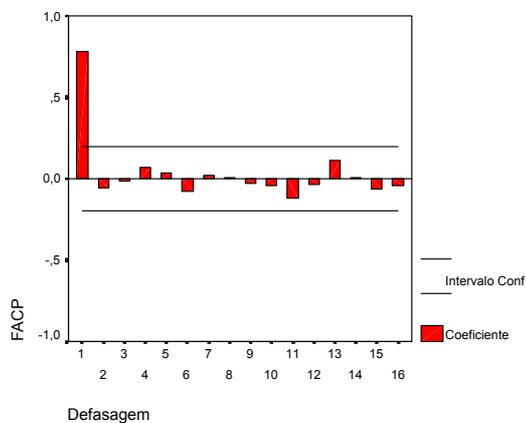
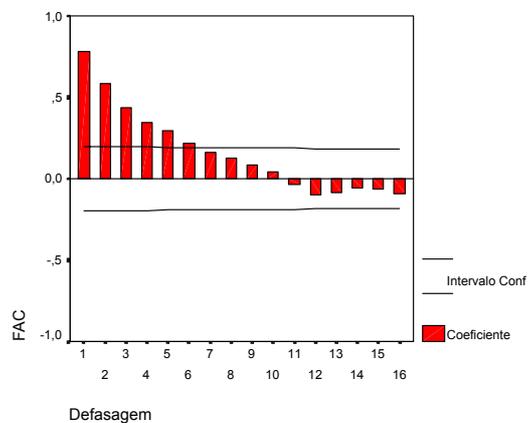


Figura 7 – Gráficos de Correlograma da FAC e da FACP respectivamente para o Modelo 1-a

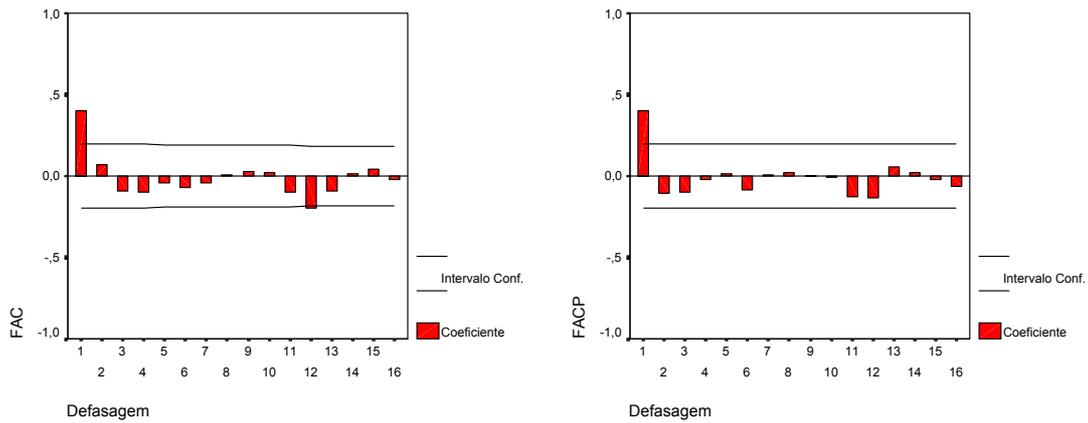


Figura 8 – Gráficos de Correlograma da FAC e da FACP respectivamente para o Modelo 2-a

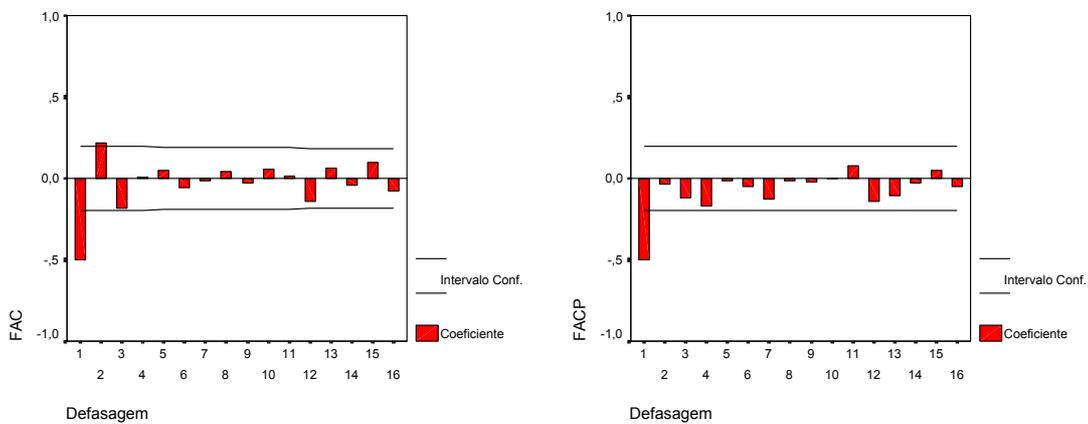


Figura 9 – Gráficos de Correlograma da FAC e da FACP respectivamente para o Modelo 3-b

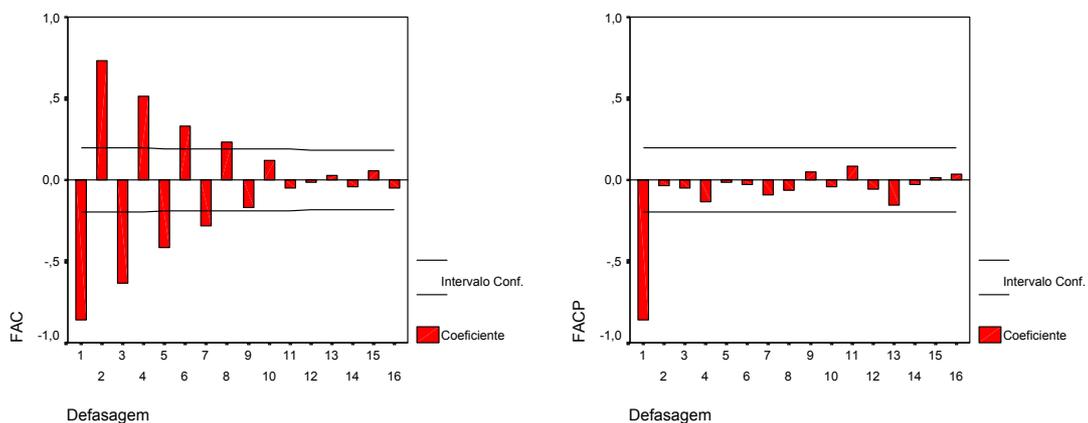


Figura 10 – Gráficos de Correlograma da FAC e da FACP respectivamente para o Modelo 4-b

4. Análise dos Resultados

Considerando o modelo simulado 1, correlação positiva e forte (figura 3), pode-se observar uma forte tendência de comportamento do modelo em assinalar diversas causas especiais. Entretanto, muitas dessas causas (ou até mesmo todas) podem não ser causas especiais, como mostra o gráfico de resíduos da figura 3. O modelo pode também não assinalar verdadeiras

causas especiais. Os gráficos de correlograma da FAC e da FACP (figura 7) mostram que o modelo ideal para o ajuste do modelo simulado 1 realmente é um AR(1) com $\phi > 0$.

Observando o modelo simulado 2, onde a correlação é positiva e moderada (figura 4), também percebe-se uma tendência de comportamento em assinalar várias causas especiais, entretanto em menor quantidade que o modelo 1. Da mesma forma todas ou quase todas podem não ser causas especiais, como verificado nos gráficos de resíduos da figura 4. Um fato interessante constatado pelo gráfico de correlograma da FAC (figura 8) mostra que a partir do *lag* 2 os coeficientes não são significativos. Isso indica que o modelo também pode ser ajustado adequadamente por um modelo de Média Móvel de Primeira Ordem, MA(1), que não é abordado nesse estudo. Pode-se observar que o correlograma da FACP também possui apenas o primeiro *lag* significativo. Isso permite também a utilização do CEP tradicional, sem problemas de autocorrelação, se o intervalo de amostragem fosse duplicado, ou seja, em vez de coletar a cada período de tamanho t , coletar a cada período $2t$.

Quanto ao modelo simulado 3, cuja correlação é negativa e moderada (figura 5), nota-se uma tendência de comportamento em não assinalar causas especiais, mesmo que elas existam. O gráfico de resíduos da figura 5 aponta causas especiais existentes que não foram assinaladas pelo modelo. Semelhantemente ao modelo simulado 2, o gráfico de correlograma da FAC (figura 9) pode indicar que um modelo MA(1) também seja adequado, pois a correlação no *lag* 2 é levemente significativa.

Finalmente, para o modelo simulado 4, onde a correlação é negativa e forte (figura 6), observa-se uma tendência clara em não assinalar causas especiais. Como pode-se verificar no gráfico de resíduos da figura 8, esse modelo não consegue assinalar as causas especiais, tratando-as como causas comuns. Outra característica marcante desse modelo é o fato de oscilar em torno da média, geralmente alternando dentre valores acima e abaixo da média, também observado no modelo simulado 3, mas não tão evidente quanto neste. Os gráficos de correlogramas desse modelo (figura 10) sugerem claramente que a série pode ser ajustada por um modelo AR(1) com $\phi < 0$.

4. Conclusões

Através desse estudo, baseado em simulações de modelos AR(1) com o intuito de analisar o efeito da autocorrelação desses modelos no desempenho do CEP, foi possível obter diversas conclusões.

Quando um modelo AR(1) possui autocorrelação positiva, quanto mais forte essa correlação, maior a tendência do modelo em apontar supostas causas especiais que são “falsos alarmes”. À medida que essa correlação for diminuindo, continuarão surgindo “alarmes falsos”, entretanto com menor intensidade. Quando a correlação é próxima de zero, geralmente ela não será significativa, isto é, os dados poderão ser considerados independentes e o CEP poderá ser aplicado normalmente.

Quando um modelo AR(1) possui autocorrelação negativa, quanto mais forte essa correlação, maior a tendência do modelo em não apontar causas especiais, ou seja, misturar as causas especiais com causas comuns. À medida que a intensidade dessa correlação for diminuindo, essa tendência vai enfraquecendo. Semelhantemente ao AR(1) com $\phi > 0$, quando a correlação é próxima de zero, geralmente ela não será significativa.

À medida que a correlação na primeira defasagem for se aproximando de zero, tanto para correlação positiva quanto negativa, o comportamento do correlograma da FAC tende a cair exponencialmente mais rapidamente. Em muitos casos, somente a correlação da primeira defasagem torna-se significativa, o que implica que o modelo também pode ser ajustado adequadamente por um MA(1). Poderia se pensar em um ARMA(1,1), entretanto esse ajuste

não é viável, pois nesse caso o efeito dos dois termos (do termo AR e do termo MA) se anula e eles passam a não serem significativos. No caso em que ambos os correlogramas da FAC e da FACP possuem somente o primeiro *lag*, poderia se utilizar CEP tradicional sem precisar ajustar por algum modelo se o intervalo de amostragem fosse duplicado. Dessa forma, a autocorrelação não seria constatada nos gráficos de controle.

Quando a autocorrelação é fraca, próxima de zero, geralmente ela não é significativa, e as técnicas de CEP tradicional podem ser utilizadas.

Referências

- DEMING, W.E. (1990) - *Qualidade: A revolução da Administração*. 1ª Edição. Editora Marques Saraiva. Rio de Janeiro.
- FISCHER, S. (1982) - *Séries Univariantes de Tempo – metodologia de Box & Jenkins*. FEE. Porto Alegre.
- MONTGOMERY, D.C. (1997) - *Introduction to Statistical Quality Control*. 3ª Edição. John Wiley & Sons. New York.
- MONTGOMERY, D.C. & MASTRAGELO, C.M. (1991) - Some Statistical Process Control Methods for Autocorrelated Data. *Journal of Quality Technology*, v. 23, n. 3, p. 179-193.
- MOREIRA, F.J.J. (2003) - Monografia (Graduação em Estatística) *Controle Estatístico de Processo para Dados Autocorrelacionados*. Departamento de Estatística, Instituto de Matemática, Universidade Federal do Rio Grande do Sul.
- RIBEIRO, J.L.D. & CATEN, C.S.T. (2001) - *Controle Estatístico do Processo*. FEENG/UFRGS. Porto Alegre.
- TOLOI, C.M.C. & MORETTIN, P.A. (1985) - *Previsão de Séries Temporais*. Atual Editora. São Paulo.