

APLICAÇÃO DE UM GRÁFICO CUSUM BINOMIAL NO MONITORAMENTO DE UM INDICADOR DE ÁGUA POTÁVEL

APPLICATION OF A BINOMIAL CUSUM CONTROL CHART TO MONITOR ONE DRINKING WATER INDICATOR

Elisa Henning* E-mail: elisa.henning@gmail.com
Andréa Cristina Konrath** E-mail: andreak@gmail.com
Olga Maria Formigoni Carvalho Walter** E-mail: olgaformigoni@gmail.com
Custodio da Cunha Alves*** E-mail: custodio.alves@gmail.com
Robert Wayne Samohyl* E-mail: samohyl@yahoo.com

*Universidade do Estado de Santa Catarina (UDESC), Florianópolis, SC

**Universidade Federal de Santa Catarina (UFSC), Florianópolis, SC

***Universidade da Região de Joinville (UNIVILLE), Joinville, SC

Resumo: A proposta deste artigo é analisar a aplicação de um gráfico de somas acumuladas (CUSUM) binomial no monitoramento da presença de coliformes totais, um indicador biológico para verificar a qualidade da água de abastecimento em um processo de tratamento de água. A série analisada é proveniente de uma estação de tratamento de água e corresponde ao período de 2007 a 2009, com frequência mensal. O tratamento estatístico dos dados foi realizado no GNU R, com desenvolvimento de rotinas para aproximar o limite superior do gráfico CUSUM binomial. Além disso, foi realizado um estudo comparativo para investigar se existe diferença significativa quanto à sensibilidade existente entre a aplicação do referido gráfico e o tradicional gráfico do tipo Shewhart, geralmente aplicado em monitoramento de processos. Os resultados demonstraram que este estudo foi fundamental para definir a melhor opção entre os gráficos para a análise estatística deste processo.

Palavras-chave: Controle estatístico de processos. CUSUM binomial. Coliformes Totais.

Abstract: The aim of this study is to analyze the use of a binomial cumulative sum chart (CUSUM) to monitor the presence of total coliforms, biological indicators of quality of water supplies in water treatment processes. The sample series were monthly taken from a water treatment plant and were analyzed from 2007 to 2009. The statistical treatment of the data was performed using GNU R, and routines were created for the approximation of the upper limit of the binomial CUSUM chart. Furthermore, a comparative study was conducted to investigate whether there is a significant difference in sensitivity between the use of CUSUM and the traditional Shewhart chart, the most commonly used chart in process monitoring. The results obtained demonstrate that this study was essential for making the right choice in selecting a chart for the statistical analysis of this process.

Keywords: Statistical process control. Binomial CUSUM. Total coliforms.

1 INTRODUÇÃO

O abastecimento de água, no que tange à quantidade e qualidade tem sido preocupação crescente, em função da escassez deste recurso e da deterioração da qualidade dos mananciais. A qualidade da água, desde o manancial, tem sido comprometida pelo lançamento de efluentes e resíduos, o que exige investimento nas estações de tratamento e alterações na dosagem de produtos para garantir a qualidade da água fornecida (BRASIL, 2006).

A água contém diversos componentes, provenientes do próprio ambiente natural ou que a ela são incorporados por meio de atividades humanas. A partir de suas características físicas, químicas e biológicas, é possível avaliar sua qualidade. Estas constituem impurezas quando as referidas características alcançam valores superiores aos estabelecidos por legislação, para determinado uso (UFV, 2011).

Para garantir a qualidade da água de abastecimento, além de atender os limites de especificação, também é importante avaliar a variabilidade da presença de coliformes totais. Como a qualidade é inversamente proporcional à variabilidade, sua contínua melhoria implica em sua redução. Pelo fato de variabilidade somente poder ser expressa estatisticamente, gráficos de controle estatístico de processos assumem importante papel para avaliar a melhoria da qualidade (SMETI *et al.*, 2007).

O principal objetivo do controle estatístico de um processo é eliminar a variabilidade e, nesta situação, os gráficos de controle são ferramentas eficientes que permitem a redução sistemática dessa variabilidade nas variáveis monitoradas (ALVES, 2003). Não há muito tempo, o paradigma tradicional no uso dos gráficos de controle eram os processos industriais mas, atualmente essa ferramenta atingiu outras áreas (SAMOHYL, 2009); entre elas, o monitoramento dos parâmetros da água (SMETI *et al.*, 2007; SMETI; KORONAKIS; GOLFINOPOULOS, 2007; AIZENCHTADT; INGMAN; FRIEDLER, 2008; ZHOU *et al.*, 2008; MESNIL; PETITGAS, 2009).

Entre os gráficos de controle estatístico, o mais usual ainda é o tradicional gráfico de Shewhart, cujo sucesso é devido a sua simplicidade. Pelo fato de, no

referido gráfico, as observações serem analisadas separadamente, é difícil detectar, nos dados, padrões cíclicos ou tendenciosos, o que é refletido na sua fraca sensibilidade cuja regra de decisão está fundamentada apenas na análise do último ponto considerado; ou seja, no caso de ele estar além dos limites de controle do gráfico, é necessário intervir no processo (ALVES, 2003). Este fato também constitui grande desvantagem, pelo fato de ignorar qualquer informação anterior, dada pela sequência anterior de pontos. Isto torna o gráfico de Shewhart relativamente insensível a pequenas mudanças no processo, da ordem de 1,5 erro padrão ou menos (MONTGOMERY, 2008).

Uma alternativa para solucionar ou minimizar este problema é o gráfico de controle de somas acumuladas (CUSUM), indicado para a detecção de pequenas e persistentes mudanças na característica da qualidade monitorada em determinado processo. Este gráfico incorpora, diretamente, toda a sequência de informações e demarca as somas acumuladas dos desvios em relação ao valor-alvo (ALVES; SAMOBYL, 2004; SOUZA, DOMINGOS FILHO, SAMOBYL, 2007). A decisão a respeito do estado do controle estatístico do processo é baseada na informação acumulada das diversas amostras anteriores, o que torna possível sinalizar, com maior rapidez, os pequenos desajustes, bem como identificar, no tempo, o momento em que ocorre mudança no processo.

Gráficos CUSUM para dados discretos, ou atributos, vão de encontro ao enfoque filosófico propagado pela teoria da Gestão da Qualidade: “fazer certo na primeira vez”. Por suas próprias características, o gráfico CUSUM sinaliza quando acontece mudança no processo, além de permitir a indicação de não conformidade (LUCAS, 1985).

Iniciando com Page (1954), existem alguns trabalhos que abrangem gráficos CUSUM para atributos: Lucas (1985, 1989), Gan (1993), Hawkins e Olwell (1998), Bourke (2001), entre outros. Um gráfico CUSUM para atributos (variáveis discretas) tem os mesmos conceitos que um CUSUM para variáveis mensuráveis. Algumas particularidades (valor nominal e parâmetros) são decorrentes das distribuições subjacentes à variável de interesse. Deste modo, os gráficos são caracterizados em função da distribuição da variável, sendo os mais comuns: o CUSUM de Poisson:

para os casos de variáveis com distribuição de Poisson (não conformidades, contagens em um intervalo de tempo); e CUSUM binomial: para os casos de variáveis com distribuição binomial (itens não conformes, proporções).

Embora o gráfico CUSUM seja mais efetivo que o gráfico de Shewhart para pequenas e moderadas alterações na proporção, o gráfico de controle np (ou seu equivalente p) do tipo Shewhart ainda é utilizado com maior frequência no monitoramento de atributos (WU; LUO, 2003). O projeto de um gráfico CUSUM não é trivial, ou seja, seu uso e interpretação comparados ao projeto de um gráfico de Shewhart são mais trabalhosos (LUCAS, 1985; WU; JIAO; LIU, 2008). Gráficos de controle de soma acumulada são mais comuns para controle de variáveis (contínuas) que de atributos. Infelizmente, mesmo para variáveis mensuráveis, o gráfico CUSUM ainda está ausente na maioria dos processos, inclusive nos processos industriais (SAMOHYL, 2009).

Para Albers (2011) em virtude dos contínuos esforços para melhorar os padrões de produção, tem sido mais frequente a ocorrência de processos de alta qualidade. Lucas, Darwin e Saniga (2006) referem que os tradicionais gráficos p , np e c de Shewhart podem não ser efetivos quando a proporção p for muito pequena. Também de acordo com Albers (2011), na área da saúde são bastante comuns pequenas frações ou itens não conformes. Falhas em equipamentos, recidivas de algumas doenças e erros em procedimentos cirurgicos, são exemplos de situações de rara ocorrência e devem ser evitadas tanto quanto possível. Em tais situações o uso do gráfico de controle para atributos é fortemente recomendado (SONESSON; BOCK, 2003; WOODALL, 2006; THOR et al., 2007; ALBERS, 2011).

Neste sentido, o monitoramento ambiental também implica em controle mais rigoroso, face a sua estreita ligação com a Saúde Pública. O abastecimento de água – no que tange à quantidade e qualidade – é preocupação crescente, em função da escassez deste recurso e da deterioração da qualidade dos mananciais.

Por esses motivos, esta pesquisa tem como objetivo principal analisar a aplicação de um gráfico CUSUM binomial como ferramenta para monitoramento da presença de coliformes totais, que constituem indicador da qualidade da água de abastecimento. Em nosso país, na literatura referente ao Controle Estatístico da

Qualidade, não é usual a aplicação de gráficos CUSUM binomiais. Foi encontrado apenas um registro, em que esse gráfico é aplicado a um processo industrial e seu desempenho é comparado a um gráfico combinado Shewhart-CUSUM binomial (HENNING et al., 2012).

A proposta deste trabalho é, contribuir para maior inserção deste tipo de gráfico – o CUSUM binomial – tanto no desenvolvimento de pesquisas quanto em aplicações práticas. No que respeita a aplicações práticas, neste documento, está estruturada uma sequência de passos e, também é empregada uma abordagem alternativa para cálculo do limite superior.

Esta pesquisa está estruturada da seguinte forma: a seção 2 é destinada aos fundamentos teóricos e gráficos de controle de Somas Acumuladas para variáveis discretas com distribuição binomial. A seção 3 descreve os procedimentos metodológicos adotados nesta pesquisa. Na seção 4, são verificadas as suposições de autocorrelação, estabilidade estatística e aderência à distribuição normal, além de ser apresentada a aplicação do gráfico CUSUM binomial no processo de tratamento de água. Finalmente, na seção 5, estão as conclusões e considerações finais.

2 REFERENCIAL TEÓRICO

2.1 CUSUM para variáveis discretas com distribuição binomial

Conceitualmente, um gráfico CUSUM binomial examina o número de não conformidades acumuladas em uma sequência de amostras. Este gráfico é denominado np e tem como objetivo a detecção de aumentos ou reduções do número de não conformidades np . O referido gráfico está associado à distribuição binomial, em que n corresponde ao tamanho das amostras e é constante a probabilidade p de que um item qualquer seja classificado como não conforme (DAVIS, MCWILLIANS, 1995; WOODALL, 1997; HAWKINS; OLWELL, 1998; SANIGA; BOURKE, 2001; WU; YEO; SPEDING, 2001; KHOO, 2004; WU; LUO; ZHANG, 2006; WU; JIAO; LIU, 2008).

Neste gráfico é fácil observar, a base do conceito do gráfico CUSUM (SAMOHYL, 2009), resultando na soma acumulada dos desvios das observações ao redor da média (np), como ponto de referência. Assim, um gráfico CUSUM binomial unilateral superior pode ser aplicado para detectar aumento no valor esperado de itens não conformes, do valor nominal np_0 (média sob controle estatístico) para np_1 (média considerada fora de controle estatístico). Então, seja uma série de amostras Y_i com $i \in Z^+$ independentes, de tamanho n , com distribuição binomial. O gráfico de controle CUSUM unilateral superior é obtido, plotando-se, em um gráfico, a estatística C_i (Equação 1),

$$C_i = \max(0, C_{i-1} + X_i - k), \quad (1)$$

em relação a amostra Y_i , onde:

- a) C_i é a estatística CUSUM com $C_0 = u$, $0 \leq u \leq h$;
- b) X_i é o número de itens não conformes na amostra Y_i com $i = 1, 2, \dots$,
- c) k é uma constante denominada valor de referência do CUSUM e depende da magnitude da mudança que se deseja detectar;
- d) h é limite superior de controle do CUSUM.

Se o processo estiver sob controle estatístico, as somas acumuladas definidas em (1) descreverão um percurso aleatório com média zero. Porém, se a média assumir algum valor acima, então uma tendência ascendente será revelada na soma acumulada C_i . Se C_i ultrapassa o limite superior de controle h , diz-se que o processo está fora de controle estatístico.

2.1.1 Parâmetros de um CUSUM binomial

O desenvolvimento de um gráfico CUSUM binomial depende da escolha dos valores de k , h e, conseqüentemente, do número médio de amostras até a emissão de um sinal, ou seja, o *Average Run Length* (ARL), considerado o estado inicial. Estes parâmetros são calculados por meio de tabelas, gráficos ou *softwares*. Um

CUSUM binomial completo envolve dois procedimentos distintos: primeiro a seleção de k e h para determinar um ARL, enquanto o segundo envolve avaliar quão rapidamente o CUSUM (sob controle estatístico) detecta mudança no parâmetro de interesse (HAWKINS; OLWELL, 1998). Em resumo, o desenvolvimento de um CUSUM binomial leva às seguintes situações: dados h e k , encontrar o ARL ou dados o ARL e k , encontrar h .

Nesta pesquisa é abordada a segunda situação. Significa que o plano é utilizar um gráfico CUSUM binomial para detectar uma mudança pré-especificada na proporção sob controle, com um valor tolerável mínimo para o ARL_0 (ou taxa máxima de alarmes falsos). O limite superior do CUSUM é determinado para estes valores de k e ARL_0 (GAN, 1993; HAWKINS; OLWELL, 1998).

2.1.2 Valor de referência do CUSUM binomial

A escolha do valor de referência k implica que o gráfico CUSUM será delineado para detectar uma particular magnitude desejada de mudança. Consideradas as variáveis discretas, o parâmetro k é determinado pela taxa de contagem aceitável e a taxa de contagem de não conformidades que se quer detectar. Neste sentido, em um esquema CUSUM para variáveis com distribuição binomial, seriam a média sob controle (np_0) e a média fora de controle estatístico (np_1), respectivamente.

Um primeiro procedimento para a escolha dos parâmetros de um CUSUM binomial é baseado na aproximação à Distribuição de Poisson. Este procedimento foi reproduzido em livros-texto da época (KEMP, 1962; DUNCAN, 1986; BOURKE, 2001).

Posteriormente, de acordo com alguns autores, o valor de referência k para o CUSUM binomial pode ser escolhido com uso da representação do gráfico CUSUM como uma sequência de testes *Sequential Probability Ratio Test* (SPRT) (HAWKINS, 1992, HAWKINS; GAN, 1993 OLWELL, 1998; REYNOLDS; STOUMBOS, 1998). Uma primeira aproximação do valor de referência com o mesmo k do SPRT foi proposta por Lucas (1985) para um CUSUM com distribuição de Poisson. A escolha

de parâmetros ótimos para um CUSUM binomial foi estudada por Gan (1993). O autor demonstrou que o k do SPRT pode ser considerado o valor ótimo para o k do gráfico CUSUM. Assim, o cálculo do valor de referência k pode ser resumido por:

$$k = \frac{n \ln((1-p_0)/(1-p_1))}{\ln((1-p_0)/(1-p_1)) - \ln(p_0/p_1)}, \quad (2)$$

Em que p_0 corresponde à proporção sob controle estatístico; p_1 é a proporção fora de controle estatístico que se deseja detectar e n é o tamanho da amostra. A relação entre o valor de referência k do CUSUM e o SPRT pode ser consultada no Anexo A.

Hawkins (1992) e Hawkins e Olwell (1998) reforçaram a ligação entre o k do CUSUM e do SPRT, ao desenvolverem uma expressão aplicável a distribuições contínuas e discretas pertencentes à família exponencial.

O valor de k é escolhido para uma resposta ótima para mudança de uma magnitude específica. Isto implica que não terá o mesmo o desempenho para mudanças cujo tamanho se afaste substancialmente da mudança desejada. Um gráfico planejado para detectar mudanças em um patamar de três desvios-padrão, pode não ser adequado para a detecção de alterações em um desvio-padrão. Esta escolha vai estar diretamente relacionada à natureza do processo que se deseja controlar (HAWKINS; OLWELL, 1998).

2.1.3 ARL e limite superior de controle

Como exposto anteriormente, para determinar o limite superior de controle (h) do CUSUM, é necessário especificar ARL_0 , ou vice-versa. A escolha do valor do ARL_0 , ou do limite superior, depende fundamentalmente do processo a ser analisado. Pelo fato de ser necessário escolher um deles para calcular o outro, a opção foi abranger toda a teoria, em conjunto.

Em um gráfico de controle o ARL_0 indica o número médio de amostras coletadas até a emissão de um sinal durante o período sob controle, isto é, está relacionado à emissão de um alarme falso (erro tipo I), com probabilidade igual ao

nível de significância α desejado. Em contraponto ARL_1 representa o número médio de amostras até a emissão de um alarme verdadeiro, ou seja, de um sinal que indique uma situação fora de controle, quando esta realmente ocorreu (ALVES; SAMOHYL, 2004).

A determinação das medidas de desempenho de um gráfico de controle, entre as quais o ARL_0 , é fundamental para a avaliação de sua eficiência. Um gráfico é mais eficiente que outro se, para um mesmo valor de ARL_0 , ele tiver um valor menor que RL_1 , detectando, mais rapidamente, uma alteração, quando esta realmente existir.

Há vários procedimentos na literatura para o cálculo do ARL_0 e do ARL_1 de um CUSUM binomial. Uma abordagem para determinar o ARL de um gráfico CUSUM para variáveis discretas é mediante cadeias de Markov. Esta abordagem foi originalmente proposta por Brook e Evans (1972). Permite determinar, além do ARL, os demais momentos e percentis da distribuição. Esse método pode ser utilizado para distribuições discretas, como acurada aproximação (BROOK; EVANS, 1972; MORAIS, 2002) ou (considerada) exata (REYNOLDS; STOUMBOS, 2000). Algumas desvantagens são apontadas principalmente do ponto de vista operacional. Este é um procedimento sofisticado do ponto de vista de programação e alguns detalhes podem resultar dificuldades na resolução. Por exemplo, quando p_0 for muito pequeno, a matriz de transição de probabilidades da cadeia de Markov pode ser muito grande, resultando “grande esforço computacional” (BROOK; EVANS, 1972; HAWKINS, 1992; REYNOLDS; STOUMBOS, 1999). Todavia, muitos autores aplicam a abordagem markoviana para procedimentos CUSUM binomiais, como Hawkins (1992); Gan (1993); Reynolds e Stoumbos (2000); Lucas, Davis e Saniga (2006), além de Wu, Jiao e Liu (2008).

Outra abordagem para determinar o ARL é simular processos produzindo o CUSUM, guardar os valores de RL obtidos e, então, calcular a média deles. A coleta de valores de RL também pode ser obtida para derivar sua distribuição empírica. Quando é utilizado este método, é gerada uma sequência de observações de uma distribuição pré-determinada e, então, é calculado o valor da estatística de interesse, até que ela fique fora dos limites de controle estatístico. O número médio

de observações necessárias até a emissão de um sinal, isto é, o valor do ARL, varia de uma sequência para outra. A estimativa para o ARL é obtida repetindo o procedimento um elevado número de vezes (10.000 ou mais). Para facilitar, o ARL é determinado assumindo que as mudanças no valor médio, em relação ao valor pretendido, ocorrem no instante em que tem início o procedimento (ALVES, 2009).

Esta técnica pode ser computacionalmente trabalhosa e não produzir os resultados desejados no caso de processos com valores maiores de ARL (HAWKINS; OLWELL, 1998). Técnicas de redução da variância são sugeridas para melhorar a estimativa de valores de ARL por simulação (ROSS, 1990; JUN; CHOI, 1993).

Para gráficos CUSUM binomiais, até o momento do desenvolvimento deste artigo não foram encontrados trabalhos em que a simulação fosse aplicada para determinar especificamente parâmetros de desempenho. Todavia, a simulação pode ser útil para checar a implementação das cadeias de Markov ou outros métodos de cálculo de ARL (HAWKINS; OLWELL, 1998).

Outra linha está baseada nas aproximações de Wald (1947) e nas extensões deste trabalho propostas por Siegmund (1985), Reynolds e Stoumbos (1999), e Henning, Samohyl e Alves (2010). O princípio desta aproximação foi substituir a condição de passeio aleatório da estatística de controle por um processo de movimento browniano contínuo, o qual somente é válido para distribuições fortemente não aritméticas. Nesse sentido, a distribuição binomial não atende aos pressupostos básicos que devem ser considerados para a correção. Para resolver este problema, Reynolds e Stoumbos (1999) propuseram a padronização da variável Y_i , – para aproximá-la a um processo de movimento browniano – e aplicaram um argumento de ajuste de curvas. O limite superior h corrigido é denotado por b (Equação 3).

$$b = h + \frac{1}{3}(1 - 2p_0), \quad 0 < p_0 \leq 0,5 \quad (3)$$

Reynolds e Stoumbos (1999) incluíram esta correção nas aproximações de Wald e, posteriormente, estenderam esta teoria para determinar as medidas de

desempenho de um CUSUM binomial. A Equação 4 apresenta o cálculo do ARL_0 com o limite h corrigido pela Equação 3 (REYNOLDS; STOUMBOS, 2001).

$$ARL_0 \approx \left(\frac{r_2}{n(r_2 p_0 - r_1)} \right) \left(b - \frac{n\gamma(e^{br_2} - 1)}{1 - e^{-n\gamma r_2}} \right), \quad (4)$$

onde $r_1 = -\ln\left(\frac{1-p_1}{1-p_0}\right)$; $r_2 = -\ln\left(\frac{p_1(1-p_0)}{p_0(1-p_1)}\right)$ e $\gamma = \frac{r_1}{r_2}$.

Pode ser facilmente verificado que a quantidade $n\gamma$ corresponde ao valor de referência k (Equação 2). A Equação 4 permite que seja diretamente calculado o ARL_0 , uma vez conhecidos a proporção sob controle (p_0), a proporção considerada fora de controle que se deseja detectar rapidamente (p_1), o tamanho da amostra (n) e o limite superior (h).

O parâmetro h pode ser encontrado por meio da Equação 4 para um valor pré-determinado de ARL_0 mediante um método numérico para solução de equações. Métodos iterativos, envolvendo ou não a utilização de derivadas, podem ser utilizados. Embora os recursos computacionais atuais tornem simples a aplicação desses métodos, a escolha de adequados valores iniciais é necessária para a sua convergência. Assim, foi buscada uma forma de aproximação que permitisse o cálculo direto do parâmetro h , sem necessidade de empregar métodos iterativos.

Henning, Samohyl e Alves (2010) propuseram uma alternativa para contornar os cálculos iterativos necessários para encontrar o limite de controle h . Para atingir esse objetivo, foi empregada a função LambertW e seu desenvolvimento em série, de acordo com a metodologia utilizada por Rogerson (2006). Assim, o limite superior do CUSUM pode ser aproximado por:

$$h \approx \frac{1}{r_2} \left(- \left(\frac{n\gamma_2}{e^{-n\gamma_2} - 1} \right) + nARL_0(p_0r_2 - r_1) - \left(L_1 - L_2 + \frac{L_2}{L_1} + \frac{L_2(L_2 - 2)}{2L_1^2} \right) \right) - \frac{(1 - 2p_0)}{3} \quad (5)$$

onde $L_1 = \ln \left(- \left(- \left(\frac{1}{e^{-n\gamma_2} - 1} \right) \left(e^{n\gamma_2 + n(r_1 - p_0r_2)} \right)^{-ARL_0 - (r_2 / (1 - e^{-n\gamma_2}) \times (r_1 - p_0r_2))} \right) n\gamma_2 \right) \right)$ e $L_2 = \ln(-L_1)$.

Maiores detalhes sobre a Equação 5 e a teoria subjacente podem ser encontrados em Rogerson (2006), Henning (2010), e Henning, Samohyl e Alves (2010).

2.2 Gráficos de controle de Shewhart com as regras suplementares

Os gráficos de controle “np” de Shewhart, conforme dito anteriormente, são amplamente utilizados devido à sua simplicidade. Como serão feitas comparações com este gráfico, incluindo o uso de regras suplementares, entendeu-se ser necessário uma breve descrição dessas regras.

As regras suplementares aplicadas aos gráficos do tipo Shewhart foram criadas com o objetivo de acelerar a detecção de alterações no processo. Elas procuram formalizar a percepção que algo, no processo, possa estar fora de ordem. Todavia, a incorporação dessas regras implica o aumento do risco α , ou seja, há maior número de alarmes falsos (COSTA; EPPRECHT; CARPINETTI, 2005).

Algumas regras suplementares para os gráficos de controle de Shewhart são listadas a seguir. As quatro primeiras são conhecidas como regras da *Western Electric*. A primeira regra da lista corresponde ao sinal de ação padrão, não é uma regra suplementar. Desta forma, um processo é considerado fora de controle estatístico se forem verificados um ou mais dos critérios abaixo (MONTGOMERY, 2008):

- 1) um ou mais pontos fora do limite de controle;
- 2) dois ou três pontos sucessivos fora dos limites de alerta de dois σ (desvios-padrão);
- 3) quatro ou cinco pontos consecutivos além dos limites de um σ ;

- 4) uma sequência de oito pontos consecutivos de um mesmo lado da linha central;
- 5) seis pontos em uma sequência sempre crescente ou decrescente;
- 6) quinze pontos em sequência dentro dos limites de um σ , tanto acima como abaixo da linha central;
- 7) quatorze pontos em uma sequência alternadamente para cima e para baixo;
- 8) oito pontos em uma sequência de ambos os lados da linha central com nenhum dentro dos limites de um σ ;
- 9) um padrão não usual ou não aleatório dos pontos;
- 10) um ou mais pontos perto de um limite de alerta ou de controle.

Vale complementar, de acordo com Werkema (1995), AIAG (2004) e Nomelini, Ferreira e Oliveira (2009) quando são observados sete ou mais pontos consecutivos (ascendentes ou descendentes); o processo também pode ser considerado fora de controle estatístico. No entanto, é preciso ter cautela no uso dessas regras, pois o grande número de alarmes falsos pode ser ruim para o Controle Estatístico de Processo, tornando o gráfico de Shewhart mais complexo (MONTGOMERY, 2008).

3 ASPECTOS METODOLÓGICOS

3.1 Caracterização e condução da pesquisa

Esta pesquisa é classificada como aplicada, pelo fato de utilizar conceitos já estabelecidos, ou seja, gráficos de controle CUSUM binomial; ela tem como finalidade aprimorar o conhecimento com a aplicação de seus resultados. A pesquisa aplicada tem como motivação adquirir conhecimento em uma situação específica da realidade (BARROS; LEHFELD, 2000; APPOLINÁRIO, 2004; GIL 2010). No caso desta pesquisa, a situação em questão trata do monitoramento dos coliformes totais, um importante indicador da água qualidade da água e, por consequência de sua potabilidade.

A aplicação desta pesquisa abrange a análise do monitoramento de coliformes totais na rede de distribuição de água potável da Companhia Águas de Joinville. Este parâmetro (coliformes totais) é um indicador da possibilidade da existência de micro-organismos patogênicos (causadores de doenças).

Segundo a Portaria N^o 518/2004 do Ministério da Saúde (BRASIL, 2004), os coliformes totais, bactérias do grupo coliforme, são bacilos gram-negativos, aeróbios ou anaeróbios facultativos, não formadores de esporos, oxidase-negativos, capazes de desenvolver na presença de sais biliares ou agentes tensoativos que fermentam a lactose com produção de ácido, gás e aldeído a $35,0 \pm 0,5^{\circ}\text{C}$, em 24-48 horas, e que podem apresentar atividade da enzima β -galactosidase. A maioria das bactérias do grupo coliforme pertence aos gêneros Escherichia, Citrobacter, Klebsiella e Enterobacter, embora vários outros gêneros e espécies pertençam ao grupo.

A Tabela 1 apresenta as quantidades máximas permitidas de coliformes totais para água potável destinada ao consumo humano.

Tabela 1 - Padrão microbiológico de potabilidade da água para consumo humano

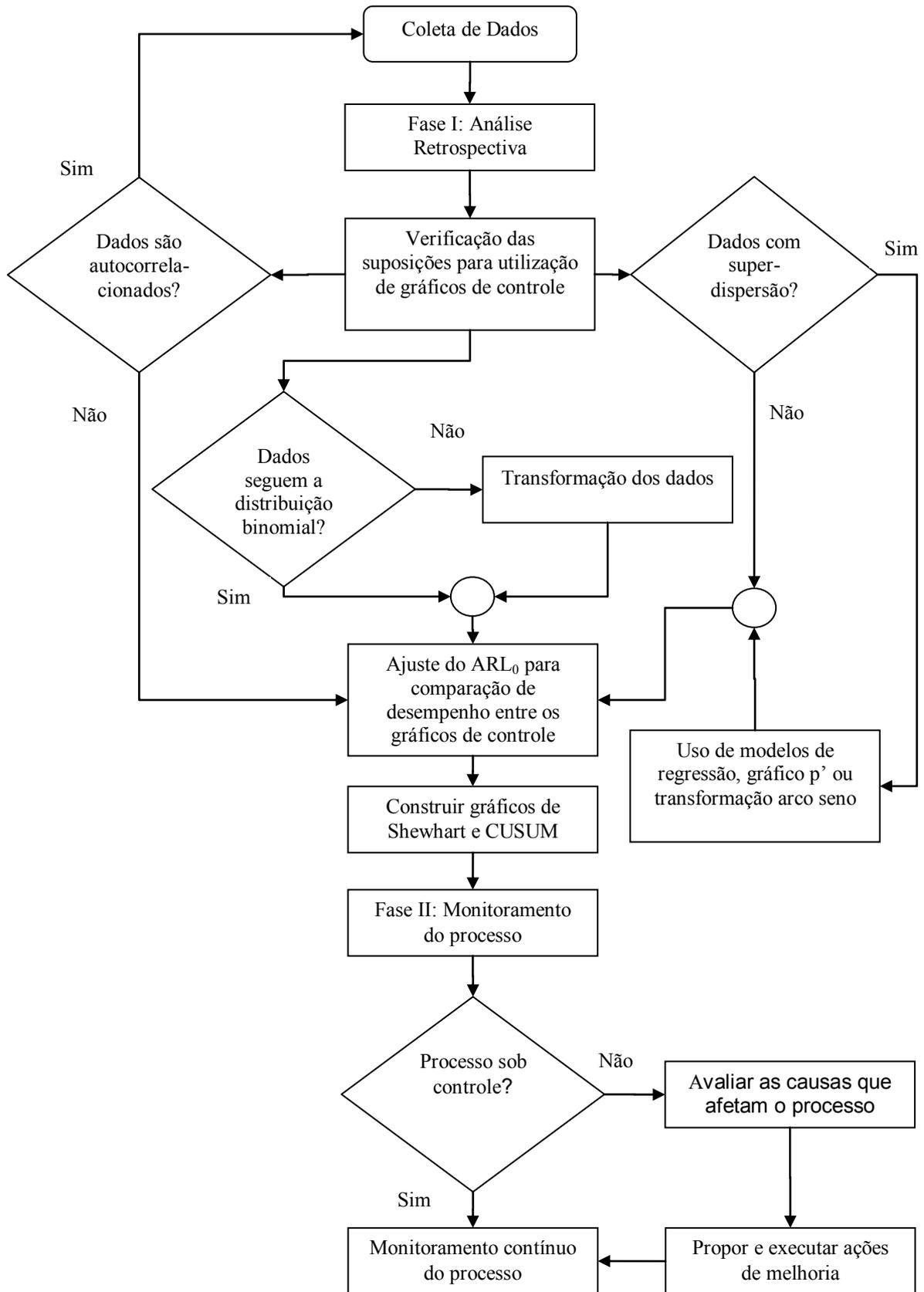
Parâmetro	Valor máximo!
Coliformes Totais	Ausência em 100 ml
	Sistemas que analisam 40 ou mais amostras por mês: ausência em 100 ml em 95% das amostras examinadas no mês.
	Sistemas que analisam menos de 40 amostras por mês: apenas uma amostra poderá apresentar mensalmente resultado positivo em 100 ml.

Fonte: Portaria MS N.º 518/2004 (BRASIL, 2004)

De acordo com a Portaria N^o 518/2004, do Ministério da Saúde (BRASIL, 2004), é tolerada a presença de coliformes totais em até 5% das amostras analisadas.

A Figura 1 apresenta os procedimentos utilizados para condução desta pesquisa. A sequência de passos apresentada também pode ser aplicada em outros processos, incluindo produtivos, para dados com distribuição binomial.

Figura 1 – Procedimentos de condução da pesquisa



De acordo com a Figura 1, após a coleta de dados, foram verificadas durante a fase I – que corresponde à análise retrospectiva dos dados (PEDRINI; SANT'ANA; TEN CATEN, 2011) – as suposições para implantação de gráficos de controle. As suposições foram aceitas (ver seção 4) e, na sequência, foi realizado o ajuste do ARL_0 , e foi mantido o mesmo desempenho para que, desta forma, os gráficos de Shewhart e CUSUM pudessem ser comparados. Na fase II, ou seja, a de monitoramento do processo, foram avaliadas as condições apresentadas pelos dois gráficos de controle.

Quando da presença de autocorrelação, é recomendado utilizar técnicas que capturem a autocorrelação serial. Assim, é possível avaliar a aplicação de gráficos de controle para os resíduos obtidos por meio de modelos de séries temporais (WOODALL; ADAMS; BENNEYAN, 2012) ou modelos de regressão não linear (RUSSO; CAMARGO; SAMOHYL, 2008). Outra opção – quando possível – é aumentar o espaço de tempo entre as retiradas das amostras (MONTGOMERY, 2008).

Quando os dados apresentarem superdispersão, devem ser aplicadas técnicas apropriadas para contornar este problema. Como sugestões estão o uso de modelos de regressão antes da aplicação dos gráficos de controle (RUSSO; CAMARGO; SAMOHYL, 2008) ou de gráficos específicos para esta classe de problemas, como o gráfico p' (LANEY, 2002). Transformações nos dados, como o arco seno, também são alternativas (JONES; GOVINDARAJU, 2000).

Para a falta de aderência à distribuição binomial, a sugestão é realizar a transformação dos dados. Jones e Govindaraju (2000) ressaltaram que, se houver superdispersão, a suposição de aderência à distribuição binomial não se aplica.

3.2 Caracterização do processo e coleta dos dados

A coleta de dados ocorreu nas duas estações de tratamento de água (ETA) de Joinville – ambas possuem rigoroso monitoramento – foram analisadas amostras de controle operacional e de qualidade nas águas brutas (água do rio), decantada e filtrada (fases intermediárias), e tratadas. A água de abastecimento que sai da ETA

vai para a rede de distribuição. As análises realizadas na rede de distribuição de água devem estar de acordo com a Portaria N^o 518/2004, do Ministério da Saúde (BRASIL, 2004). O plano de amostragem – também é definido pela portaria, que exige um número mínimo de análises mensais a serem realizadas em função da população abastecida. No caso de Joinville, são exigidas 200 análises para diversos parâmetros, entre eles, o de coliformes totais. Para tanto há 205 pontos de coleta, situados em locais diversos da cidade. As coletas são realizadas em dias e horários pré-definidos, as amostras são classificadas como conformes (não apresentam coliformes totais) e não conformes (apresentam coliformes totais). Ao final, é contado o número de amostras não conformes.

A série analisada corresponde ao período entre 2007 a 2009, com frequência mensal. O conjunto de dados consta do Anexo B. Os dois primeiros anos, 2007 e 2008, correspondem à fase I do processo. O valor de p , considerado sob controle, é $p_0 = 0,022$, que corresponde ao valor médio da proporção de amostras não conformes observadas na fase I. O monitoramento – segunda fase do processo – será feito com as observações do ano de 2009.

Nesta pesquisa, os dois métodos, Markov e o método fundamentado na aproximação de Siegmund (Equação 5), foram utilizados para o cálculo do limite de controle h . O tratamento estatístico dos dados foi feito no ambiente GNU R 2.13.0 (R DEVELOPMENT CORE TEAM, 2011) com auxílio dos pacotes qcc (SCRUCCA, 2004), para os gráficos do tipo Shewhart e *surveillance* (HÖHLE, 2007), para determinar h , a partir de cadeias de Markov. A rotina desenvolvida para o gráfico CUSUM binomial consta do Apêndice A, assim como as funções elaboradas para aproximar o limite superior (Apêndice B).

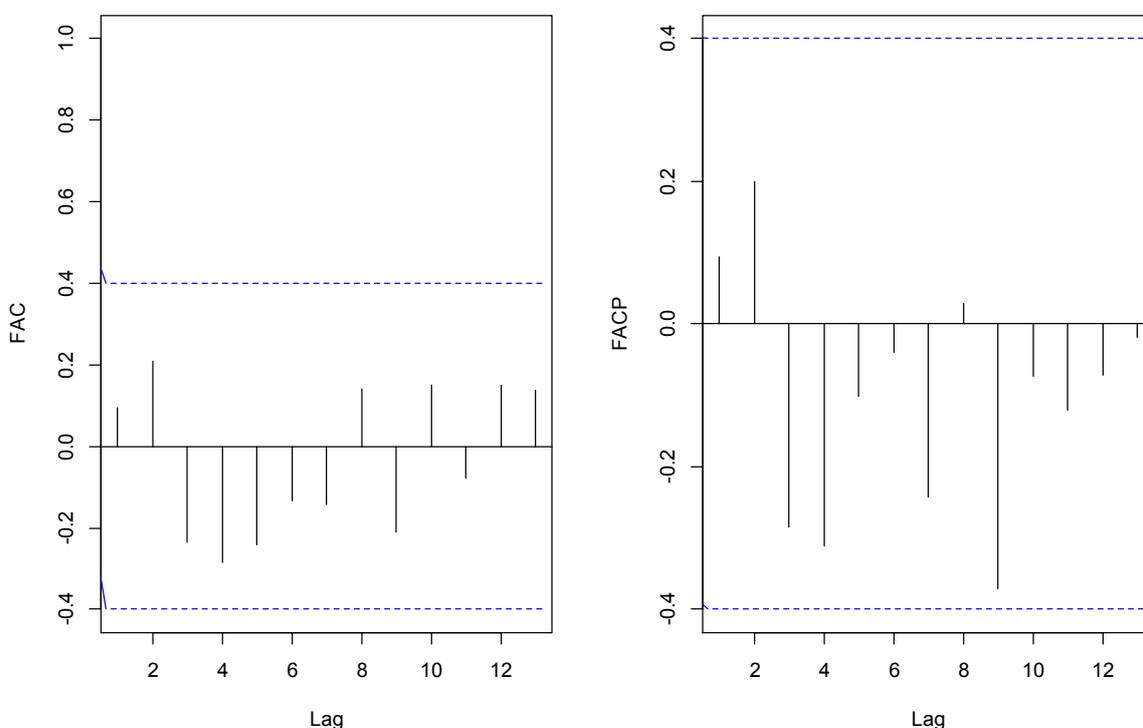
4 APLICAÇÃO DO GRÁFICO CUSUM BINOMIAL NO PROCESSO DE TRATAMENTO DE ÁGUA

Para aplicar os gráficos CUSUM e Shewhart é necessário verificar as suposições de independência, aderência à distribuição binomial e superdispersão dos dados desse processo. Para tanto foi construído um teste baseado na estatística

qui quadrado, o qual revelou ausência de superdispersão (p -valor = 0,2850). A aderência à distribuição binomial foi verificada também a partir do teste qui quadrado (p -valor = 0,757).

A independência pode ser avaliada pelo processo de amostragem, uma vez que as coletas foram realizadas em locais, dias e horários distintos, portanto, independentes. Também foram construídos os gráficos da função de autocorrelação e autocorrelação parcial amostral. Os gráficos da Figura 2 mostram que os dados não possuem autocorrelação; portanto, os gráficos Shewhart e CUSUM podem ser aplicados.

Figura 2 – Gráficos da função de autocorrelação e função de autocorrelação parcial amostral



Para planejamento dos gráficos, de início foi cogitado considerar um $ARL_0 = 370$ ($\alpha = 0,27\%$). Neste caso α , significa probabilidade de alarme falso. Considerando $ARL_0 = 370$ ($\alpha = 0,27\%$) existe a probabilidade de 0,27% do gráfico de controle acusar incorretamente uma mudança na média do processo a cada 370 subgrupos. Entretanto, para este estudo o limite convencional de Shewhart implicava um valor de ARL_0 menor e, conseqüentemente, uma taxa de alarmes

falsos maior. A opção, então, foi calcular o limite exato (ou probabilístico) a partir da distribuição binomial. Assim, o limite superior de controle é $LSC = 11,0$ com $ARL_0 = 465,5$ ($\alpha = 0,21\%$).

Para fins de comparação, foi desenvolvido um gráfico do tipo Shewhart individual com o mesmo $ARL_0 = 465,5$ e $LSC = 11,0$. O gráfico do tipo Shewhart, Figura 3, sinaliza, antes do combinado (28ª amostra – abril de 2009), uma mudança na média (np) sob controle.

Um gráfico CUSUM planejado para detectar $p_1 = 0,05$ ($s = 127$), o limite estabelecido pela resolução, com $k = 6,9$; $h = 5,21$ e $ARL_0 = 465,5$; está na Figura 4. Este gráfico tem o mesmo desempenho que o Shewhart, sinalizando uma situação fora de controle estatístico, na 28ª amostra (abril de 2009). Ademais, ainda indica a partir da 30ª amostra, haver tendência de aumento do número de análises não conformes, o que não aparece no gráfico do tipo Shewhart.

Figura 3 – Gráfico Shewhart para os dados de coliformes totais.

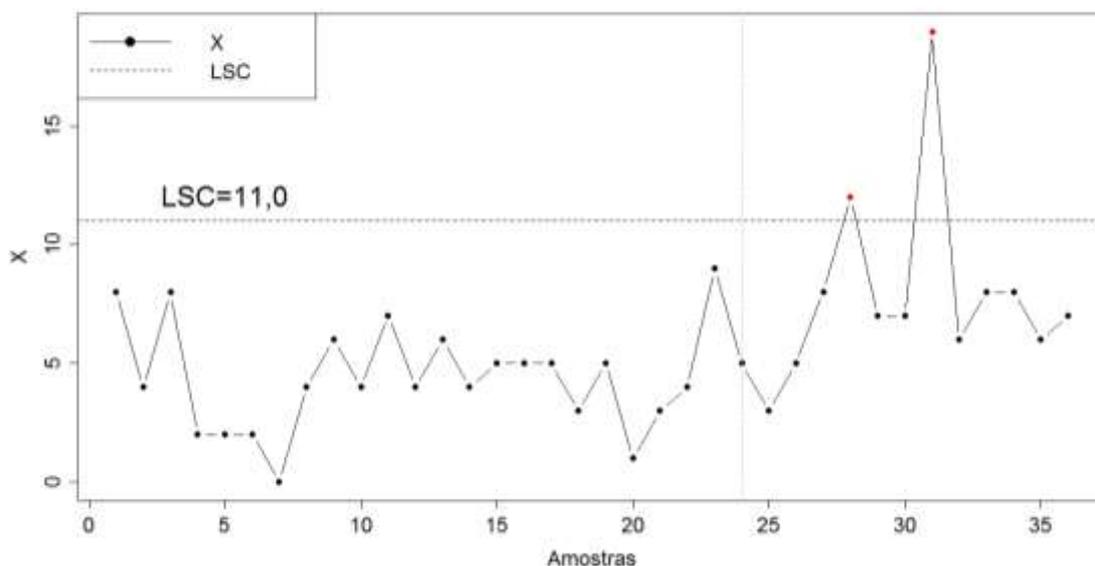
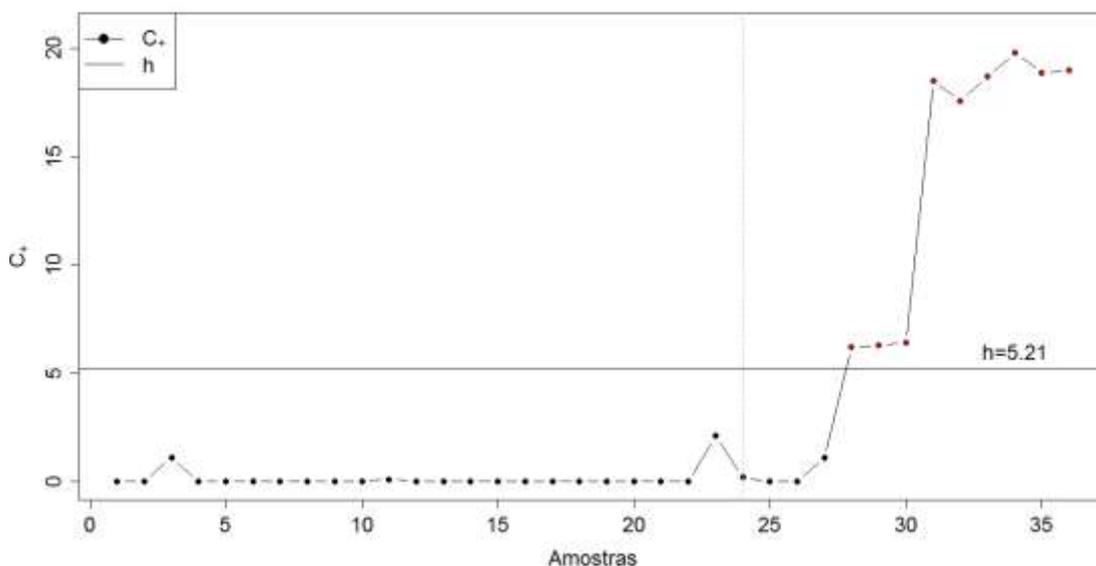


Figura 4 – Gráfico CUSUM para os dados de coliformes totais

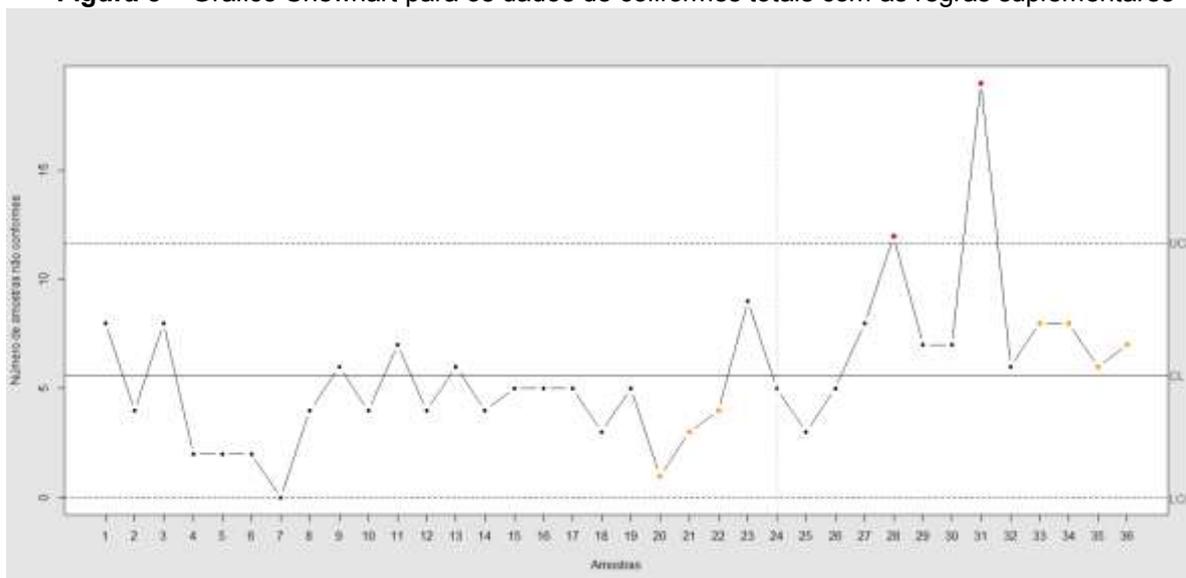


Ao serem aplicadas as regras suplementares – além dos limites de controle, ao procedimento de Shewhart – mencionadas na seção 2.2, elas também indicam que o processo está fora de controle estatístico, nas amostras 33 a 36 (ver Figura 5). O uso destas regras, no entanto, implica aumento do risco α , ou seja, probabilidade maior de ocorrência de alarmes falsos. Além disso, a interpretação do gráfico se torna mais complexa. A intenção é chamar atenção para a tendência de aumento de amostras não conformes, o que é possível verificar a partir do gráfico CUSUM (Figura 4) e do uso das regras suplementares no gráfico de Shewhart (Figura 5).

A primeira situação (pontos 20, 21 e 22) refere-se à violação da regra 2, onde dois ou três pontos sucessivos se encontram fora dos limites de alerta de dois σ . Esta situação não caracteriza problema pois, pelo fato de se aproximar do limite inferior, indica melhoria no processo: menos amostras com a presença de coliformes. A segunda situação, a partir da amostra 33, remete à existência de sete ou mais pontos acima da linha central, evidenciando problemas com o aumento de amostras não conformes.

Pode ser verificado que o gráfico CUSUM permite sinalizar esta tendência de aumento, sem implicar no aumento da probabilidade de ocorrência de alarmes falsos.

Figura 5 – Gráfico Shewhart para os dados de coliformes totais com as regras suplementares



Os dois procedimentos, isto é, CUSUM e Shewhart, indicaram que a média (np) do processo mudou, o que não é desejável. Implica que coliformes totais estão presentes em número maior de análises em 2009 e, por isso, seriam recomendáveis investigações na linha de distribuição. Vale ressaltar que no ano anterior (2008), o estado de Santa Catarina, incluindo alguns bairros de Joinville, foi acometido por uma grande enchente. Além disso, em 2009 houve troca de governo municipal, o que pode ter resultado alterações dos profissionais que fazem o serviço de manutenção da rede. Este exemplo de aplicação permitiu avaliar a aplicabilidade de um gráfico CUSUM binomial, além de possibilitar observar outros aspectos interessantes:

- a) Há dificuldade em estabelecer os limites do gráfico Shewhart para alcançar o ARL_0 de projeto;
- b) O desempenho do CUSUM individual foi similar ao gráfico Shewhart, sinalizando a mesma amostra, com a vantagem de detectar a tendência de aumento no processo. No gráfico Shewhart, isto somente é possível com auxílio das regras suplementares, pois estas proporcionam ao gráfico de controle maior sensibilidade para detectar mudanças no processo.

Assim, um gráfico CUSUM planejado para detecção de p_1 correspondente a 0,05 (5%) pode ser recomendado para monitoramento deste particular processo. Nas análises realizadas, o gráfico CUSUM foi tão efetivo quanto o gráfico do tipo Shewhart, porém com a vantagem de trazer informações sobre a tendência do processo.

5 CONCLUSÕES E CONSIDERAÇÕES FINAIS

Os gráficos de controle estatístico, e em particular os gráficos CUSUM binomiais, são técnicas estatísticas utilizadas para monitorar a qualidade de um processo. Os referidos gráficos podem ser utilizados com o mesmo propósito, mas com vantagens sobre os gráficos de Shewhart quando o objetivo for detectar pequenas variações, considerando aspectos estatísticos e econômicos. Do ponto de vista estatístico, pode haver uma redução no número médio de amostras até que o gráfico sinalize a ocorrência de uma causa especial. Sob o ponto de vista econômico, pode haver uma redução nos custos relacionados ao controle do processo, que é causada pela diminuição na taxa de amostragem quando o gráfico indicar um processo sob controle.

Neste contexto, este trabalho teve como proposta analisar a aplicação de um gráfico CUSUM binomial como ferramenta para monitorar um indicador de qualidade da água de abastecimento no que respeita à presença de coliformes totais.

Os aspectos práticos decorrentes da referida aplicação desse gráfico no monitoramento da presença de coliformes totais no processo avaliado de tratamento de água, conforme resultados obtidos, indicam haver diferença entre a utilização dos gráficos CUSUM e de Shewhart no que tange ao desempenho estatístico.

Na análise realizada, o gráfico CUSUM binomial revelou ser tão efetivo quanto o gráfico do tipo Shewhart, porém com a vantagem de trazer informações sobre a tendência ao longo do período analisado. No gráfico do tipo Shewhart, como as observações são vistas separadamente, é difícil detectar, nos dados, padrões

cíclicos ou tendenciosos, o que é revelado na sua fraca sensibilidade na detecção de pequenas mudanças do referido processo.

O gráfico CUSUM binomial, planejado para detecção de p_1 correspondente a 0,05 (5%), pode ser recomendado para monitoramento deste particular processo. Como sugestão para continuidade dos trabalhos, alguns pesquisadores citam a aplicação, com posterior análise, de um gráfico CUSUM planejado para detectar alterações maiores (WU; JIAO; LIU, 2008), como fazem referência, também, ao uso de gráficos combinados Shewhart-CUSUM.

Outra técnica que pode ser incorporada é a análise da curva *Receiver Operating Characteristic* (ROC) – método gráfico para avaliação, organização e seleção de sistemas de diagnósticos e predição. O gráfico ROC está baseado na probabilidade de detecção – ou taxa de verdadeiros positivos – e na probabilidade de falsos alarmes, ou taxa de falsos positivos.

REFERÊNCIAS

AIAG - AUTOMOTIVE INDUSTRY ACTION GROUP. **Análise dos sistemas de medição**. 3 ed. São Paulo: 2004.

AIZENCHTADT, E.; INGMAN, D.; FRIEDLER, E. Quality control of wastewater treatment: a new approach. **European Journal of Operations Research**, v. 189, n. 2, p. 445-458, 2008.

ALBERS, W. Improved binomial charts for high-quality processes. **Produção**, v. 21, n. 2, p. 209-216, 2011.

ALVES, C. C. **Gráficos de controle CUSUM: um enfoque dinâmico para a análise estatística de processos**. 2003. 119f. Dissertação (Mestrado em Engenharia de Produção), Centro Tecnológico, Universidade Federal de Santa Catarina, Florianópolis.

ALVES, C. C. **Método da equação integral com quadratura gaussiana para otimizar os parâmetros do gráfico de controle multivariado de somas acumuladas**. 2009. 196f. Tese (Doutorado em Engenharia de Produção), Centro Tecnológico, Universidade Federal de Santa Catarina, Florianópolis.

ALVES, C. C.; SAMOHYL, R. W. A utilização dos gráficos de controle CUSUM para o monitoramento de processos industriais. In: ENCONTRO NACIONAL DE

ENGENHARIA DE PRODUÇÃO, 24., 2004, Florianópolis. **Anais....** Florianópolis, UFSC, 2004.

APPOLINÁRIO, F. **Dicionário de metodologia científica**: um guia para a produção do conhecimento científico. São Paulo: Atlas, 2004.

BARROS, A. J. S.; LEHFELD, N. A. S. **Fundamentos de metodologia**: um guia para a Iniciação Científica. 2.ed. São Paulo: Makron Books, 2000.

BOURKE, P. D. Sample size and the Binomial CUSUM Control Chart: the case of 100% inspection. **Metrika**, v. 53, n. 1, p. 51-70, 2001.

BRASIL. Ministério da Saúde. **Secretaria de vigilância em saúde. coordenação-geral de vigilância em saúde ambiental**, Portaria no. 518/2004 do Ministério da Saúde. Disponível em:
http://portal.saude.gov.br/portal/arquivos/pdf/portaria_518_2004.pdf. Acesso em: 11 Jul. 2012.

BRASIL. Ministério da Saúde. **Vigilância e controle da qualidade da água para consumo humano**. Secretaria de Vigilância em Saúde. Brasília, 2006. 212 p. Disponível em:
http://bvsms.saude.gov.br/bvs/publicacoes/vigilancia_controle_qualidade_agua.pdf. Acesso em: 19 set. 2012.

BROOK, D.; EVANS, D. A., Na Approach to the Probability Distribution of Cusum Run Length, **Biometrika**, v. 59, n. 3, p. 539-549, 1972.

COSTA, A. F. B.; EPPRECHT, E. K.; CARPINETTI, L. C. R. **Controle estatístico de qualidade**. 2.ed. Atlas: São Paulo, 2005.

DUNCAN, A. J. **Quality control and industrial statistics**. 5.ed. Irwin Homewood IL, 1986. 1296 p.

GAN, F. F. An optimal design of CUSUM control charts for binomial counts. **Journal of Applied Statistics**, v. 20, n. 4, p. 445-460, 1993.

GIL, A. C. **Como elaborar projetos de pesquisa**. 5.ed. São Paulo: Atlas, 2010.

HAWKINS, D.; OLWELL, D. **Cumulative sum charts and charting for quality improvement**. New York: Springer Verlag Inc, 1998.

HAWKINS, D. M. Evaluation on Average Run Lengths of Cumulative Sum Charts for an Arbitrary Data Distribution. **Communication in Statistics, Simulation and Computation**, v. 21, n. 4, p. 1001-1020, 1992.

HENNING, E. **Aperfeiçoamento e desenvolvimento dos gráficos combinados Shewhart-Cusum binomiais**. 250f. Tese (Doutorado em Engenharia de Produção), Centro Tecnológico, Universidade Federal de Santa Catarina, Florianópolis, 2010.

HENNING, E. ; SAMOHYL, R.W.; WALTER, O.M.F.C.; LINDOSO, R.L. Gráfico combinado Shewhart-CUSUM aplicado a um processo de rotulagem de garrafas. **Gestão Industrial**, v. 08, n. 4: p. 204-221, 2012.

HENNING, E.; SAMOHYL, R. W.; ALVES, C. C. Aproximação do limite superior de um gráfico CUSUM binomial. In: ENCONTRO NACIONAL DE ENGENHARIA DE PRODUÇÃO, 30., 2010, São Carlos. **Anais...** São Carlos, UFSCar, 2010.

HÖHLE, M. Surveillance: an R package for the monitoring of infectious diseases. **Computational Statistics**, v. 22, n. 4, p. 571-572, 2007.

JONES, G.; GOVINDARAJU, K. A Graphical Method for Checking Attribute Control Charts Assumptions. **Quality Engineering**, v. 13, n. 1, p. 19-26, 2000.

JUN, C.; CHOI, M. S.; Simulating the Average Run Length for CUSUM Schemes using Variance Reduction Techniques. **Communication in Statistics - Simulation and Computation**, v. 22, n. 3, p. 877-887, 1993.

KEMP, K. W. The Use of Cumulative Sums for Sampling Inspection Schemes. **Applied Statistics**, v. 11, n. 1, p. 16-31, 1962.

KHOO, M. B. C. A moving average control chart for monitoring the fraction non-conforming. **Quality Reliability Engineering International**, v. 20, n. 6, p. 617-635, 2004.

LANEY, D. B. Improved Control Charts for Attributes. **Quality Engineering**, v. 14, n. 4, p. 531-537, 2002.

LUCAS, J. M. Counted Data CUSUM's. **Technometrics**, v. 27, n. 2, p. 129-144, 1985.

LUCAS, J. Control schemes for low count level. **Journal of Quality Technology**, v. 21, n. 3, p. 109-201, 1989.

LUCAS, J. M.; DAVIS, D. J.; SANIGA, E. M. Detecting improvement using Shewhart attribute control charts when the lower control limit is zero. **IIE Transactions**, v. 38, n. 8, p. 659-669, 2006.

MESNIL, B.; PETITGAS, P. Detection of changes in time-series of indication using CUSUM control charts. **Aquatic Living Resources**, v. 22, n. 2, p. 187-192, 2009.

MONTGOMERY, D. C. **Introduction to Statistical Quality Control**. 6th ed. John Wiley & Sons, Inc., Hoboken: New Jersey, 2008.

MORAIS, M. C. **Ordenação estocástica na análise de desempenho de esquemas de controle de qualidade**. 194f. Tese (Doutorado em Matemática), Instituto Superior Técnico, Universidade Técnica de Lisboa, 2002. Disponível em: <http://www.math.ist.utl.pt/~mjmorais/thesis.pdf>. Acesso em: 19 jan. 2013.

NOMELINI, Q.S.S.; FERREIRA, E.B.; OLIVEIRA, M.S. Estudos dos padrões de não aleatoriedade dos gráficos de controle de Shewhart: um enfoque probabilístico. **Gestão & Produção**, v. 16, n. 3, 2009.

PAGE, E. S. Continuous Inspection Schemes. **Biometrika**, v. 41, n. 1/2, p. 100-114, 1954.

PEDRINI, D. C.; SANT'ANA, A. M. O.; TEN CATEN, C. S. Gráfico EWMAREG aplicado no monitoramento de processos industriais. **Produção Online**, v.11, n. 4, p. 1141-1159, out./dez. 2011. Disponível em: <<http://producaoonline.org.br/rpo/article/view/801>>. Acesso em: 17 jan. 2013.

R DEVELOPMENT CORE TEAM. **R: A language and environment for statistical computing**. R Foundation for Statistical Computing, Vienna, Austria, 2011. Disponível em: <http://www.R-project.org>.

REYNOLDS, M. R.; STOUMBOS, Z. G. The SPRT chart for monitoring a proportion. **IIE Transactions**, v. 30, n. 6, p. 545-561, 1998.

REYNOLDS, M. R.; STOUMBOS, Z. G. A CUSUM Chart for Monitoring a Proportion When Inspecting Continuously. **Journal of Quality Technology**, v. 31, n.1, p. 87-108, 1999.

REYNOLDS, M. R.; STOUMBOS, Z. G. A general approach to modeling CUSUM charts for a proportion. **IIE Transactions**, v. 32, n. 6, p. 515-535, 2000.

REYNOLDS, M. R.; STOUMBOS, Z. G. **Monitoring a Proportion using CUSUM and SPRT charts**. In: LENZ, H. J.; WILRICH, P. T. *Frontiers in Statistical Quality Control*, Physika-Verlag Heidelberg: 2001.

ROGERSON, P. A. Formulas for the Design of CUSUM Quality Control Charts. **Communications in Statistics – Theory and Methods**, v. 35, p. 373-383, 2006.

ROSS, S. M. Variance reduction in simulation via random hazards. **Probability in the Engineering and Informational Sciences**, v.4, n. 3, p. 299-309, 1990.

RUSSO, S. L.; CAMARGO, M. E.; SAMOBYL, R. W. Gráficos de Controle baseado nos resíduos do modelo de regressão de Poisson. **Produção Online**, v. 8, n. 4, Dez.

2008. Disponível em: <<http://producaoonline.org.br/index.php/rpo/article/viewFile/138/212>>. Acesso em: 17 jan. 2013.

SAMOHYL, R. W. **Controle estatístico de qualidade**. Rio de Janeiro: Elsevier, 2009.

SANIGA, E. M.; DAVIS, D. J.; MCWILLIAMS, T.P. Economic, Statistical, and economic statistical design of attribute charts. **Journal of Quality Technology**, v. 27, n. 1, p. 56-61, 1995.

SCRUCCA, L. qcc: An R package for quality control charting and statistical process control. **R News**, v. 4, n.1, p. 11-17, 2004.

SIEGMUND, D. **Sequential Analysis: test and Confidence Intervals**. New-York: Springer Verlag, 1985.

SONESSON, C.; BOCK, D. A review and discussion of prospective statistical surveillance in public health. **Journal of the Royal Statistical Society A**. v. 166, n. 1, p. 5-21, 2003.

SOUZA, G. P; DOMINGOS FILHO, M. D.; SAMOHYL, R. W. Aplicação dos Conceitos de Controle Estatístico de Processo (CEP) em uma Indústria de Fundição do Norte Catarinense. **Produção Online**, v. 7, n. 2, p. 64-84, Ago. 2007. Disponível em: <<http://producaoonline.org.br/rpo/article/view/73>>. Acesso em: 17 jan. 2013.

SMETI, E. M.; THANASOULIAS, N. C.; KOUSOURIS, L. P.; TZOUMERKARS, P. C. An approach for the application of statistical process control techniques for quality improvement of treated water. **Desalination**, v. 213, n.1/3, p. 273-281, 2007.

UFV. Qualidade da Água. Disponível em: <http://www.ufv.br/dea/lqa/qualidade.htm>. Acesso em: 19 set. 2012.

SMETI, E. M.; KORONAKIS, D. E.; GOLFINOPOULOS, S. K. Control charts for the toxicity of finished water -Modeling the structure of toxicity. **Water Research**, v. 41, n. 12, p. 2679-2689, 2007.

THOR, J. et al. Application of statistical process control in healthcare improvement: systematic review. **Quality Management in Healthcare**, v. 16, p. 387-399, 2007.
WALD, A. **Sequential Analysis**. New York: Dover Publications, 1947.

WERKEMA, M. C. C. **Ferramentas estatísticas básicas para o gerenciamento de processos**. Belo Horizonte: QFCO, 1995.

WOODALL, W. H. Control Charts based in Attribute Data: Bibliography and Review. **Journal of Quality Technology**, v. 29, n. 2, p. 172-183, 1997.

WOODALL, W. H. The Use of Control Charts in Health-Care and Public-Health Surveillance. **Journal of Quality Technology**, v. 38, n. 2, p. 89-104, 2006.

WOODALL, W. H.; ADAMS, B. M.; BENNEYAN, J. C. **Statistical Process Control in Healthcare**. In: FALTIN, F.; KENNET, R.; RUGGIERI, F. Statistical Methods in Healthcare. West Sussex: John Wiley & Sons, 2012.

WU, Z.; YEO, S. H.; SPEDING, T. A. A synthetic control chart for detecting fraction nonconforming increases. **Journal of Quality Technology**, v. 33, n. 1, p. 104-111, 2001.

WU, Z.; LUO, H. Three-triplet np control charts. **European Journal of Operational Research**, v.149, n. 3, p. 614-624 , 2003.

WU, Z.; LUO, H.; ZHANG, X. L. Optimal np chart with curtailment. **European Journal of Operational Research**, v. 174, n. 3, p. 1723-1741, 2006.

WU, Z.; JIAO, J.; LIU, Y. A binomial CUSUM chart for detecting large shifts in fraction non conforming. **Journal of Applied Statistics**, v. 35, n.11, p. 1267-1276, 2008.

ZHOU, W.; BECK, B. F.; PETTIT, A. K.; WANG. J. Application of water quality control charts to spring monitoring in karst terranes. **Environmental Geology**, v. 53, n. 6, p. 1311-1321, 2008.



Artigo recebido em 19/09/2012 e aceito para publicação em 22/11/2013
DOI: <http://dx.doi.org/10.14488/1676-1901.v14.i1.1452>

ANEXO A

Hawkins (1992) e Hawkins e Olwell (1998) reforçaram a ligação entre o k do CUSUM e do SPRT, ao desenvolverem uma expressão aplicável a distribuições contínuas e discretas pertencentes à família exponencial. Se uma sequência de variáveis aleatórias y seguindo uma distribuição membro da família exponencial na forma

$$f(y, \theta) = e^{a(y)b(\theta)+c(y)+d(\theta)}, \quad (\text{A.1})$$

com função densidade de probabilidades f_{dp} para o caso contínuo e função de probabilidades f_p para o caso discreto, e fazendo Y_i igual a $Y_i = a(Y_i)$. Suponha o teste de hipóteses $H_0: \theta = \theta_0$ contra a hipótese $H_1: \theta = \theta_1$, com $\theta_1 > \theta_0$, onde θ é o parâmetro de interesse (média ou proporção, por exemplo). O SPRT é baseado na soma acumulada da quantidade $(Y_i - k)$ onde o valor de referência k é dado por

$$k = -\frac{d(\theta_1) - d(\theta_0)}{b(\theta_1) - b(\theta_0)}, \quad (\text{A.2})$$

e segue a forma de f_{dp} e f_p e a escolha de θ_0 e θ_1 . Deste modo é possível definir facilmente o valor de referência k , para variáveis com distribuições pertencentes a membros da família exponencial. Então, para a distribuição binomial tem-se

$$k = -\frac{d(\theta_1) - d(\theta_0)}{b(\theta_1) - b(\theta_0)} = -n \left(\frac{\ln\left(\frac{1-p_1}{1-p_0}\right)}{\ln\left(\frac{p_1(1-p_0)}{p_0(1-p_1)}\right)} \right), \quad (\text{A.3})$$

com $f(y | m, p) = e^{y \ln(p/(1-p)) + n \ln(1-p) + \ln n! - \ln(y!(n-y)!)}$. Neste caso, $d(\theta) = d(p) = n \ln(1-p)$ e

$$b(\theta) = b(p) = \ln\left(\frac{p}{1-p}\right).$$

A expressão (A.3) é a mesma da Equação (2). Até então foi abordado apenas um CUSUM com início em zero (0). Outro valor inicial, que corresponde ao que recebe o nome de Resposta Inicial Rápida (RIR) não afeta a escolha de k . O valor de referência k do SPRT, utilizado neste trabalho como valor de referência k do CUSUM binomial, é determinado pela escolha dos dois valores de parâmetros p_0 e p_1 associados às hipóteses nula e alternativa.

ANEXO B

Tabela B1 - Conjunto de dados do exemplo de aplicação: Número de amostras não conformes (com presença de coliformes totais)

Ano	Mês	NC	Ano	Mês	NC	Ano	Mês	NC
2007	jan	8	2008	jan	6	2009	jan	3
2007	fev	4	2008	fev	4	2009	fev	5
2007	mar	8	2008	mar	5	2009	mar	8
2007	abr	2	2008	abr	5	2009	abr	12
2007	mai	2	2008	mai	5	2009	mai	7
2007	jun	2	2008	jun	3	2009	jun	7
2007	jul	0	2008	jul	5	2009	jul	19
2007	ago	4	2008	ago	1	2009	ago	6
2007	set	6	2008	set	3	2009	set	8
2007	out	4	2008	out	4	2009	out	8
2007	nov	7	2008	nov	9	2009	nov	6
2007	dez	4	2008	dez	5	2009	dez	7

Fonte: Relatórios anuais de qualidade dos anos de 2007, 2008 e 2009 da Companhia de Água e Saneamento Águas de Joinville.

OBS: NC corresponde a o número de amostras consideradas não conformes.

APÊNDICE A – ROTINA NO R PARA CONSTRUÇÃO DO GRÁFICO CUSUM BINOMIAL UNILATERAL SUPERIOR

```
f1=function (kmais,hmais,x) {
cmais=0
cmais[0]=0
for (i in 1:length(x)) {cmais[i]=max(0,cmais[i-1]+x[i]-kmais)}
plot(cmais,type='b',ylim=c(0,max(cmais+1,hmais)),main='
',ylab=expression(paste(C["+"])),xlab='Amostras',lwd=1,bg=c("white"))
points(cmais,pch=19)
abline(h=hmais,lty=1)
legend("topleft", c(expression(paste(C["+"])), "h"),
pch=c(19,NA),lty=c(1,1))
#text(2,hmais+0.5,paste("h=",hmais))
#grid(20)
return(cmais)
}
```

APÊNDICE B - PROGRAMAS PARA APROXIMAÇÃO DO LIMITE SUPERIOR DO CUSUM h

```

###função para aproximar h
###aproximação CD de Reynolds e Stoumbos
###com LambertW na resolução para b (h+correção)
###expandindo LambertW em serie com 2,3, 4 e 5 termos
###

hrs.6<-function(p0,p1,n,arl0) {
r1<- -log((1-p1)/(1-p0))
r2<- log((p1*(1-p0))/(p0*(1-p1)))
g <- r1/r2
k <- round(n*g,2)
l1<-log((exp(-((n *(r2*g + (-1 + exp(n* r2*g)) *(r1 - p0* r2)*
arl0))/(-1 + exp(n* r2 *g))))* n* r2 *g)/(-1 + exp(n* r2*g)))
l2<-log(-log((exp(-((n *r2* g)/(-1 + exp(n* r2* g))) + n *(-r1
+ p0* r2)* arl0)* n *r2 *g)/(-1 + exp(n* r2 *g))))
l3<-l2/l1
l4<-(l2*(l2-2))/(2*l1^2)
l5<-(l2*(6-9*l2+2*l2^2))/(6*l1^3)
b<-((-n*r2*g)/(1-exp(-n*r2*g))-(l1-l2)+n*arl0*(p0*r2-r1))/r2 #
2 termos
b1 <- ((-n*r2*g)/(1-exp(-n*r2*g))-(l1-l2+l3)+n*arl0*(p0*r2-
r1))/r2 # 3 termos
b2 <- ((-n*r2*g)/(1-exp(-n*r2*g))-(l1-l2+l3+l4)+n*arl0*(p0*r2-
r1))/r2 # 4 termos
b3 <- ((-n*r2*g)/(1-exp(-n*r2*g))-(l1-
l2+l3+l4+l5)+n*arl0*(p0*r2-r1))/r2 # 5 termos
h<- b-(1/3)*(1-2*p0)
h1<- b1-(1/3)*(1-2*p0)
h2<- b2-(1/3)*(1-2*p0)
h3<- b3-(1/3)*(1-2*p0)
structure(list(k=k,h=h,h1=h1,h2=h2,h3=h3))
}

```