

2 Revisão Conceitual e Trabalhos Precedentes

Neste capítulo, é feita inicialmente uma revisão conceitual das principais medidas de desempenho de gráficos de controle, e é realizada uma revisão bibliográfica dos trabalhos precedentes envolvendo outros tipos de gráficos e procedimentos de controle por atributos que buscam uma maior eficiência que a dos gráficos tradicionais de Shewhart na detecção de alterações de pequena magnitude na característica de qualidade do processo. Uma maior ênfase é dada aos trabalhos acerca de gráficos adaptativos e gráficos EWMA. Além disso, é feita uma revisão dos esquemas EWMA que utilizam a estratégia de gráficos adaptativos, no caso, para controle por variáveis, pois não encontramos na literatura trabalho envolvendo esquemas EWMA adaptativos por atributos.

2.1. Revisão Conceitual: Principais Medidas de Desempenho de Gráficos de Controle

As medidas de desempenho de gráficos de controle podem ser usadas para análise da capacidade dos mesmos de detectar alterações no processo. A seguir, são definidas as principais medidas de desempenho de gráficos de controle.

A probabilidade de alarme falso é a probabilidade de erroneamente considerar-se o processo fora de controle, se o processo estiver em controle, ou seja, ocorre um alarme falso quando há um sinal indevido de que o processo está sob a influência de alguma causa especial, demandando, portanto, ajustes. A consequência de ordem prática associada ao alarme falso é intervir no processo na hora errada, quando ele encontra-se no mais perfeito estado de controle estatístico, o que implica em custo de mão de obra, atraso da produção, caso o processo seja interrompido, e possível desajuste do mesmo na tentativa de corrigir um problema inexistente.

Chamamos de NAS o número de amostras até o gráfico produzir um sinal de existência de causa especial, quando o processo está fora de controle. O valor

esperado de NAS será denotado por NMA (número médio de amostras até o sinal), ou seja, o NMA é o número esperado de amostras retiradas até o gráfico fornecer um alarme verdadeiro.

Quando o processo encontra-se em controle, e o gráfico sinaliza indevidamente um descontrole (alarme falso), o número médio de amostras até um alarme falso é denotado por NMAF.

O TMAF representa o tempo médio até a ocorrência de um alarme falso. Este tempo é função do NMAF e do intervalo de tempo adotado entre amostras. Um alto valor de TMAF é desejável, de modo a se garantir, assim, uma baixa frequência de alarmes falsos.

Quando o processo se inicia fora de controle (ou pelo menos o seu monitoramento se inicia com ele fora de controle), chamaremos de TES_f o tempo esperado até a sinalização do descontrole pelo gráfico. O índice f evidencia justamente o fato de o processo já estar fora de controle. Se h for o intervalo de tempo entre amostras, podemos escrever $TES_f = E(h \times NAS) = h \times E(NAS) = h \times NMA$.

Como já foi dito no *Capítulo 1*, estamos supondo neste trabalho que ao iniciar-se o monitoramento, o processo está em controle, e só sai deste estado em algum momento aleatório no futuro, devido à ocorrência de alguma causa especial. Supõe-se também que a causa especial ocorra instantaneamente, fazendo a característica de qualidade de interesse ser aumentada de seu valor-alvo para outro valor, e que o processo se mantenha nesse estado enquanto nenhuma ação corretiva for empreendida.

O TES é uma medida de desempenho que representa o valor esperado do tempo até o sinal ($E[TS]$, onde TS representa o tempo até o sinal), quando o monitoramento é iniciado com o processo em controle e a causa especial só ocorre após o processo ter permanecido em controle por um certo período de tempo, ou seja, o TES mede o intervalo de tempo esperado entre a alteração no processo e o alarme verdadeiro. Supondo que a alteração ocorra bruscamente, então se Q for o intervalo de tempo entre o momento da retirada da última amostra antes da alteração no processo e o instante desta alteração, podemos escrever $TES = E(TS) = E(h \times NAS - Q) = h \times NMA - E(Q)$. Como a alteração no processo pode ocorrer a qualquer instante dentro de um intervalo entre amostras

(de duração h), então, conforme Reynolds *et al.* (1988), pode-se supor para Q uma distribuição uniforme em $[0, h]$, e assim $E(Q) = h/2$. Essa suposição é bastante razoável na grande maioria dos casos, levando a uma aproximação bastante precisa do real valor de $E(Q)$. Na verdade, para se obter a expressão exata para $E(Q)$, seria necessário conhecer a distribuição do tempo em que o processo está em controle; no entanto, de acordo com Epprecht, Costa & Mendes (2003), essa é uma excelente aproximação para $E(Q)$ se o intervalo de tempo entre amostras for menor que 10% do tempo médio em controle. Adota-se, aqui, então, esta aproximação por simplicidade e para generalidade dos resultados. Portanto, $TES = h \times NMA - h/2 = TES_f - h/2$.

A *Figura 1*, retirada de Costa, Epprecht & Carpinetti (2004), ilustra a definição apresentada do tempo até o sinal, TS, quando uma característica de qualidade tem seu valor médio alterado de μ_0 para μ_1 por uma causa especial.

Para as análises de desempenho realizadas nesta pesquisa, serão usados o TMAF e o TES como medidas de desempenho.

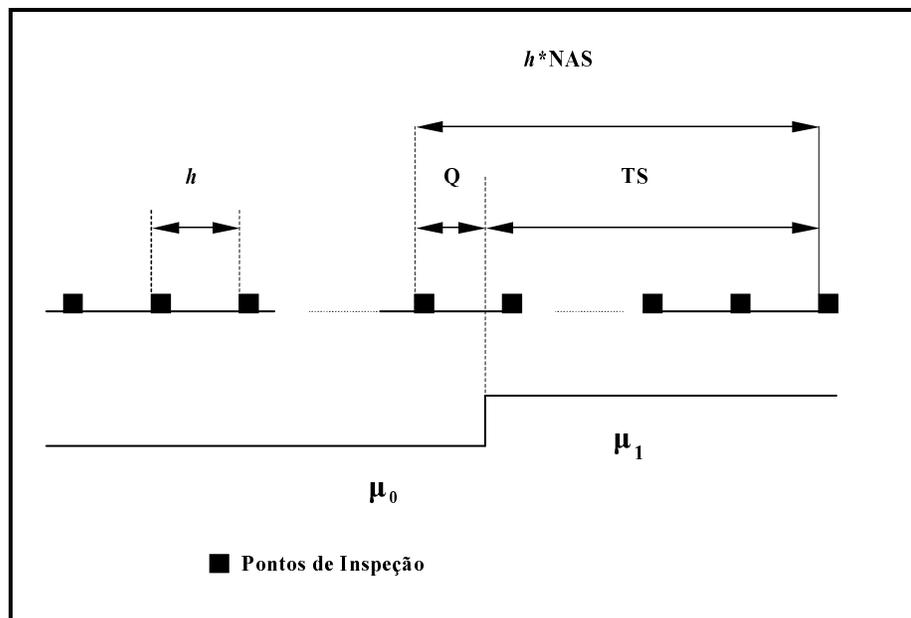


Figura 1 - Tempo até o Sinal – TS

2.2. Trabalhos Precedentes

Um primeiro procedimento visando aumento da eficiência dos gráficos de controle envolve o uso de **regras suplementares** de decisão (*runs rules*). Para gráficos por atributos, destacam-se os estudos de Nelson (1997) e de Acosta-Mejia (1999).

Nelson (1997) verifica que o uso de regras suplementares para um gráfico de np é particularmente vantajoso quando não existe um limite inferior de controle, ou seja, quando $np - 3\sqrt{np(1-p)} \leq 0$, onde n é o tamanho de amostra e p é a fração média de defeituosos, pois um teste suplementar é capaz de sinalizar a ocorrência de uma causa especial indicando que a fração de defeituosos diminuiu, e não há problema se o valor de p for pequeno, uma vez que um número suficiente de subgrupos sucessivos contendo zero itens defeituosos pode sinalizar uma diminuição em p . O autor mostra que as quantidades de subgrupos sucessivos e de itens defeituosos para os testes suplementares dependem dos valores de n e p , fornece um programa computacional que calcula os valores apropriados do número de subgrupos sucessivos e do número de itens defeituosos para toda combinação de n e p em um gráfico de np , e apresenta, também, um teste suplementar para detectar um aumento em p .

Acosta-Mejia (1999) analisa o desempenho, baseado na distribuição do número de amostras até um sinal (*run length*), de alguns gráficos alternativos para monitorar aumentos e diminuições na fração média de defeituosos p . Os gráficos alternativos estudados são: gráficos que usam os limites de controle modificados de Ryan & Schwertman (1997), gráficos que substituem o limite inferior de controle por regras suplementares, e gráficos que usam uma transformação arco-seno. É mostrado que a substituição do limite inferior de controle por regras suplementares pode resultar em um aumento geral no desempenho do gráfico.

Há, ainda, estudos acerca de gráficos de controle baseados no número de itens entre ocorrências, ou seja, no número de itens conformes encontrados entre dois itens não-conformes produzidos. Estes gráficos são também chamados de **gráficos de contagem inter-eventos**. Destinam-se a processos com fração defeituosa ou frequência de ocorrência de defeitos em controle muito baixas.

Podemos citar os estudos de Bourke (1991), Nelson (1994), McCool & Joyner-Motley (1998), Wu, Yeo & Spedding (2001) e Wu & Yeo (2001).

Bourke (1991) apresenta dois gráficos de Shewhart, denominados gráficos RL_1 e RL_2 , e um gráfico de Somas Acumuladas (CUSUM, de *Cumulative Sum*; o procedimento CUSUM será descrito mais adiante) denominado RL CUSUM, também conhecidos como gráficos CRL (de *Conforming Run Length*), baseados no número de itens conformes entre dois itens não-conformes produzidos, e mostra o desempenho desses gráficos na detecção de alterações na fração de defeituosos quando é utilizada uma inspeção a 100% na ordem de produção.

Nelson (1994) sugere também a contagem inter-eventos como uma alternativa para monitorar a proporção de itens não-conformes quando a frequência de itens defeituosos é muito pequena. O autor assume que a taxa de ocorrência de itens não-conformes é adequadamente aproximada por uma distribuição de Poisson, e, assim, os tempos entre ocorrências destes itens não-conformes são independentes e exponencialmente distribuídos. No entanto, a distribuição exponencial é muito assimétrica, de modo que o autor sugere o uso de uma transformação simples para permitir que testes suplementares possam ser aplicados: a distribuição exponencial é transformada em uma distribuição de Weibull com parâmetros tais que esta última é adequadamente aproximada por uma distribuição normal, ou seja, é feita a transformação da variável aleatória X (número de itens amostrados até um item não-conforme ser encontrado) em $Y = X^{0,2777}$, onde Y é uma variável aleatória com distribuição aproximadamente normal.

McCool & Joyner-Motley (1998) propõem uma alternativa para o caso de fração não-conforme extremamente pequena, usando uma transformação logarítmica para o número de itens amostrados até um item não-conforme ser encontrado; com a constatação de que o uso da transformação logarítmica gera um NMAF (número médio de amostras até um alarme falso) inaceitavelmente pequeno, os autores propuseram aplicar à estatística transformada um esquema da Média Móvel Ponderada Exponencialmente (EWMA, de *Exponentially Weighted Moving Average*, esquema que será descrito mais adiante), verificando que há melhorias em seu desempenho.

Wu, Yeo & Spedding (2001) propõem um novo gráfico de controle, o gráfico de controle sintético, para a detecção de aumentos na fração média de não-

conformes p . Os autores descrevem a operação, o projeto (escolha dos valores para os parâmetros) e o desempenho deste gráfico, que é uma combinação das operações de um gráfico de np com um gráfico de contagem inter-eventos (chamado pelos autores de CRL, de *Conforming Run Length*), onde o primeiro é o gráfico do número de defeituosos, e o segundo é o gráfico do número de unidades inspecionadas entre duas unidades não-conformes consecutivas. O gráfico sintético funciona da mesma maneira que um gráfico CRL, com exceção do fato de cada ponto no gráfico CRL ser substituído por uma amostra de n unidades no gráfico sintético; em outras palavras, em um gráfico CRL, a fração de não-conformes p é a probabilidade de ocorrência de uma unidade não-conforme, mas, em um gráfico sintético, p corresponde à probabilidade de ocorrência de uma amostra não-conforme. Resultados revelam que o gráfico sintético possui um maior poder na detecção de aumentos na fração de não-conformes do processo que ambos os gráficos de np e CRL, quando é usada a inspeção a 100%. A implementação direta deste esquema sintético para atributos pode ser vista em Wu & Yeo (2001), onde os autores apresentam um programa em C que pode ser usado para fazer o projeto e calcular a medida de desempenho TES_f (tempo esperado até a sinalização pelo gráfico do descontrole, quando o processo já se inicia fora de controle) de gráficos de controle sintéticos por atributos.

Outro exemplo de novos tipos de gráficos de controle que vêm sendo desenvolvidos são os chamados **gráficos adaptativos**. Nesse tipo de gráfico, alguns ou todos os parâmetros de projeto (tamanho de amostra, intervalo de tempo entre amostras e limites de controle) passam a variar durante a operação do processo. O procedimento para implementação dos gráficos adaptativos consiste, basicamente, em uma regra simples para variar esses parâmetros. Por exemplo, pode-se usar dois tamanhos de amostra, que chamaremos, por simplicidade, de “menor” e “maior”, dependendo da posição da estatística registrada no gráfico: se o ponto se encontra dentro dos limites de controle, mas próximo a um deles (indicando que o processo provavelmente está necessitando de ajustes), então o tamanho da próxima amostra deve ser “maior”, para aumentar o poder de detecção do gráfico; ao contrário, se o valor da estatística registrada se encontra perto da linha central do gráfico, então o tamanho da próxima amostra deve ser “menor”. Alternativamente, ao invés de se variar o tamanho de amostra, pode-se variar o tempo de espera até a retirada da próxima amostra, em função da posição do ponto

registrado no gráfico: se este se encontrar dentro dos limites de controle, mas próximo a um deles, espera-se um intervalo “mais curto” de tempo para retirar a próxima amostra; se o ponto se encontrar perto da linha central, espera-se um intervalo “mais longo” de tempo para retirar a próxima amostra. Pode-se também variar o coeficiente de abertura dos limites de controle, e ainda qualquer combinação desses 3 parâmetros de projeto do gráfico: tamanho de amostra, intervalo de tempo entre amostras e limites de controle.

Se os parâmetros do gráfico adaptativo forem bem escolhidos, ele apresentará uma maior rapidez média para sinalizar alterações no processo que o gráfico tradicional (parâmetros fixos), sem requerer maior taxa média de amostragem (número de unidades inspecionadas por unidade de tempo) e sem reduzir a proteção contra alarmes falsos. Os gráficos adaptativos só não são mais eficientes que os gráficos tradicionais quando as alterações no processo são muito grandes, de modo que o poder dos gráficos tradicionais (probabilidade de sinalização em uma amostra) já é muito alto.

Os parâmetros podem variar entre mais de dois valores cada um; porém Reynolds *et al.* (1988), Reynolds (1989, 1995), Runger & Pignatiello (1991), Runger & Montgomery (1993), Prabhu *et al.* (1994), e Zimmer *et al.* (1998) mostraram que o ganho de desempenho marginal por admitir mais de dois valores para os parâmetros é muito pequeno e não compensa, em geral, a complexidade operacional que acarreta.

A revisão de Tagaras (1998) apresenta as publicações até 1997 focadas no projeto estatístico e econômico, nas propriedades, e na operação, de esquemas adaptativos. Em seu artigo, inicialmente está explicada a distinção entre estes gráficos (também chamados de dinâmicos) e gráficos estáticos com parâmetros variáveis: nos gráficos adaptativos, pelo menos um de seus parâmetros de projeto varia em tempo real, em função da informação da amostra corrente, e nos gráficos estáticos com parâmetros variáveis, apesar de seus parâmetros não serem constantes, estes não variam em função da informação da amostra corrente. Também nesse artigo, são discutidas várias medidas de desempenho (estatísticas e econômicas) relevantes para o projeto (escolha dos valores para os parâmetros) de gráficos adaptativos. A revisão prossegue com a apresentação dos principais trabalhos, artigos e resultados sobre esquemas de controle que variam dinamicamente os parâmetros: tamanho de amostra, intervalo de amostragem e

limites de controle, e fornece comparações entre os gráficos adaptativos e gráficos estáticos. Além disso, identifica lacunas na literatura existente que podem constituir áreas para futuras pesquisas em gráficos adaptativos. De acordo com esse autor, a flexibilidade dos gráficos adaptativos irá aparentemente resultar em um monitoramento mais eficiente, uma vez que estes gráficos fazem melhor uso das informações do que os gráficos tradicionais (de parâmetros fixos).

O esquema adaptativo que varia somente o intervalo de tempo entre amostras é denominado VSI (de *Variable Sampling Interval*). No caso de gráficos para atributos, Woodall (1997) e Tagaras (1998) apontavam que o único trabalho existente até a ocasião era o de Vaughan (1993), que propõe um esquema de controle adaptativo para gráficos de np usando uma abordagem bayesiana, e levando em conta as mudanças na probabilidade de o processo estar em controle. O autor desenvolve um esquema VSI para gráficos de np onde o intervalo de amostragem é determinado pelo número de não-conformes observados na amostra anterior, resultando, no entanto, em um uso de mais de 2 intervalos de amostragem. Seu modelo bayesiano é baseado nas probabilidades a priori e a posteriori de o processo estar fora de controle, requer o conhecimento da distribuição do tempo em controle, e sua política é conduzida da seguinte maneira: a cada amostra, o número observado de unidades defeituosas estabelece o tempo até a retirada da próxima amostra (que deve ser retirada quando a probabilidade a posteriori de o processo estar fora de controle atingir um valor limite predefinido).

O esquema adaptativo que varia somente o tamanho de amostra é conhecido como VSS (de *Variable Sample Size*). Para gráficos VSS para atributos, pode-se citar o trabalho de Epprecht & Costa (2001), que investigam as propriedades dos gráficos de np e de c quando o tamanho das amostras pode assumir dois valores. As propriedades dos gráficos de np VSS e de c VSS são obtidas usando cadeias de Markov, mostrando o ganho na velocidade com a qual estes gráficos adaptativos detectam deteriorações no processo (deteriorações que aumentam o número médio de não-conformes e de não-conformidades durante a produção). Os gráficos adaptativos de np e de c com tamanho de amostra variável são também comparados com gráficos de np e de c de parâmetros fixos, em termos de suas velocidades de detecção de mudanças no processo.

O esquema adaptativo que varia simultaneamente o tamanho de amostra e o intervalo de tempo entre amostras é denominado VSSI (de *Variable Sample Size*

and Sampling Interval). Este método, aplicado a gráficos por atributos, foi analisado por Cesar (2000). Os resultados indicam que o desempenho do gráfico VSSI é em geral superior ao dos gráficos VSS ou VSI correspondentes, considerados como casos particulares do esquema VSSI. O modelo VSI de Cesar (2000) vem, assim, se juntar ao esquema proposto por Vaughan (1993); a diferença é que, em Cesar (2000), só há 2 valores considerados para o intervalo de tempo entre amostras, e o critério de escolha do intervalo não é bayesiano.

Cabe comentar aqui que, embora o esquema de Cesar (2000) tenha sido chamado de VSSI, a variação dos limites de controle também é efetuada no esquema. Há uma diferença entre os gráficos por atributos e os gráficos por variáveis no que diz respeito aos limites de controle. Nos gráficos por variáveis, como a densidade de probabilidade da estatística registrada (por exemplo, \bar{X}) é contínua, pode-se estabelecer os limites de controle em função de uma probabilidade de alarme falso. Assim, o limite de controle é dado por um número, K , de desvios-padrão da estatística, e pode-se variar ou não o valor de K (obtendo assim um esquema de controle com “limites variáveis” ou “limites constantes” (significando na verdade uma probabilidade de alarme falso variável ou constante), independentemente de o tamanho de amostra ser constante ou variável). Já no caso de gráficos por atributos, como a probabilidade de alarme falso (ou verdadeiro) é uma função descontínua (“em escada”) dos limites de controle, não é possível obter uma mesma probabilidade de alarme falso com dois tamanhos quaisquer de amostra, de modo que não há sentido em falar em “limites constantes” ou “limites variáveis” da mesma maneira que para gráficos adaptativos por variáveis. Desta forma, os gráficos VSS e VSSI por atributos são sempre “de limites variáveis”. E, como nos casos considerados em Cesar (2000) escolheram-se sempre os limites de controle procurando o melhor desempenho, este esquema VSSI poderia ser chamado de V_p (de *Variable Parameters*), um esquema adaptativo que varia o tamanho de amostra, o intervalo de tempo entre amostras, e o coeficiente de abertura dos limites de controle, ou seja, todos os parâmetros de projeto.

Estendendo esse trabalho, Epprecht, Costa & Mendes (2003) propõem e analisam o desempenho de um esquema adaptativo para c em que todos os parâmetros do gráfico variam entre dois valores, de acordo com o que se observa do processo, e fornecem uma tabela com projetos ótimos para uma série de

valores de u_0 e u_1 (frequências de não-conformidades em controle e fora de controle). Os resultados estendem-se a um gráfico adaptativo de np e também podem ser usados para gráficos de u e de p . Os autores também estudam casos degenerados, nos quais um, dois ou todos os parâmetros são mantidos constantes, e uma comparação de desempenho é feita entre os esquemas. Trata-se assim de um modelo genérico para os esquemas adaptativos duais para gráficos de Shewhart por atributos. Devido às semelhanças entre o esquema EWMA adaptativo proposto nesta tese e o esquema de Epprecht, Costa & Mendes (2003), vale a pena deter-nos um pouco numa descrição mais detalhada deste último. Os autores consideram um processo de produção, cuja estatística de interesse x_k é o número de não-conformidades da k -ésima amostra, composta de m unidades de inspeção. Assume-se que x_k segue uma distribuição de Poisson com média $c = mu$, onde u é o número médio de não-conformidades por unidade de inspeção. Supõe-se, também, que as amostras sejam independentes. Considera-se que o processo inicia sua operação em controle com $u = u_0$, mas após a ocorrência de uma causa especial, ele muda para um estado fora de controle, que faz o valor de u aumentar de u_0 para u_1 . O processo permanece nesse estado fora de controle até que uma intervenção seja feita para trazê-lo de volta ao estado em controle. O gráfico adaptativo de c é definido através de dois conjuntos de parâmetros, Θ_1 e Θ_2 , definidos como $\Theta_1 = \{m_S, h_L, LSC_1, LSA_1\}$ e $\Theta_2 = \{m_L, h_S, LSC_2, LSA_2\}$, onde m_S é o tamanho menor de amostra e m_L é o tamanho maior de amostra (em unidades de inspeção), h_L é o intervalo mais longo de tempo entre amostras e h_S é o intervalo mais curto de tempo entre amostras, LSC_i é o limite superior de controle e LSA_i é o limite superior de advertência ($i = 1, 2$). O gráfico é construído sem limite inferior de controle, pois o interesse dos autores é apenas detectar a eventual ocorrência de deteriorações no processo que venham a aumentar o número esperado de não-conformidades durante a produção. Assume-se, por definição, que $0 < LSA_i < LSC_i$ ($i = 1, 2$), $m_S < m_L$ e $h_S < h_L$; na verdade, $m_S \leq m_L$ e $h_S \leq h_L$, considerando os casos degenerados. A operação do gráfico adaptativo é relativamente simples; cada amostra é retirada seguindo-se o conjunto de parâmetros Θ_1 ou Θ_2 , de acordo com a seguinte regra: se $x_k \leq LSA_i$, usa-se o conjunto de parâmetros Θ_1 para a retirada da $(k+1)$ -ésima amostra; se $LSA_i < x_k \leq LSC_i$, usa-se o conjunto Θ_2 para a retirada da $(k+1)$ -ésima amostra; se

$LSC_i < x_k$ (alarme), inicia-se a procura pela causa especial ocorrida no processo. O princípio básico das regras de operação do gráfico totalmente adaptativo é simples: se x_k encontra-se abaixo do limite de advertência, o controle é relaxado porque espera-se um tempo mais longo que o usual para retirada da $(k+1)$ ésima amostra, esta $(k+1)$ ésima amostra é menor que o usual, e o limite superior de controle é maior que o usual; inversamente, se x_k encontra-se entre o limite de controle e o limite de aviso, o controle é mais rigoroso porque espera-se menos tempo que o usual para retirada da $(k+1)$ ésima amostra, esta $(k+1)$ ésima amostra é maior que o usual, e o limite superior de controle é menor que o usual para este tamanho de amostra. O método de cadeias de Markov é utilizado para obtenção das propriedades do gráfico adaptativo. Os estados da cadeia relacionam-se à estatística x_k da seguinte maneira: $x_k \leq LSA_i$ está relacionado ao estado 1; $LSA_i < x_k \leq LSC_i$, ao estado 2; e $LSC_i < x_k$, ao estado 3 (absorvente); o índice i em LSA_i e LSC_i refere-se ao conjunto de parâmetros usado para a k -ésima amostra. A medida de desempenho usada para analisar a rapidez com que o gráfico adaptativo detecta um aumento na incidência de não-conformidades é o TES: o tempo médio entre a ocorrência da causa especial e sua sinalização. Para comparar os desempenhos do gráfico V_p com os outros esquemas adaptativos degenerados e com o esquema tradicional (de parâmetros fixos), os autores fornecem vários projetos de gráficos para diferentes valores de u_0 ($u_0 = 0,5, 1,0, 1,5, 2,0, 3,0$ e $4,0$), os valores de TES desses gráficos para vários aumentos $\gamma = u_1/u_0$ na frequência de ocorrência de não-conformidades ($\gamma = 1,5; 2,0; 2,5; 3,0; 3,5; 4,0; 4,5$ e $5,0$) e os desvios-padrão do tempo até o sinal para cada valor de γ . O esquema tradicional (gráfico de c de parâmetros fixos) é usado como referência para comparação. Sem perda de generalidade, adota-se, como a unidade de tempo, o intervalo de tempo entre amostras h do esquema tradicional, e, como a unidade de inspeção, o tamanho de amostra m do esquema tradicional; então, $h = 1$ e $m = 1$, fazendo, portanto, $u_0 = c_0$. Os autores assumem que o processo permanece em controle durante a maior parte do tempo de sua operação. Os gráficos de controle adaptativos e o gráfico tradicional devem ser comparados sob iguais condições, oferecendo a mesma proteção contra alarmes falsos e requerendo a mesma quantidade de itens inspecionados. Para tal, seus TMAF's devem ser os mesmos, assim como, na média, seus comprimentos de intervalos de tempo e tamanhos de

amostras, com o processo em controle, devem ser equivalentes, isto é, o intervalo médio de tempo entre amostras (\bar{h}) dos esquemas adaptativos deve ser $\bar{h} = h = 1$, e o tamanho médio de amostra (\bar{m}) dos esquemas adaptativos deve ser $\bar{m} = m = 1$. O tamanho médio de amostra (\bar{m}) e o intervalo médio de tempo entre amostras (\bar{h}) podem ser calculados da seguinte maneira: $\bar{m} = p_1 m_S + p_2 m_L$ e $\bar{h} = p_1 h_L + p_2 h_S$, onde p_1 e p_2 representam, respectivamente, as probabilidades das estatísticas caírem nas regiões relacionadas ao estado 1 e ao estado 2 da cadeia (até um alarme falso). No artigo, os autores garantiram um menor (ou igual) custo com amostragem e uma maior (ou igual) proteção contra alarmes falsos dos gráficos adaptativos, usando $TMAF(\text{adaptativo}) \geq TMAF(\text{tradicional})$, $\bar{h} \geq 1$ e $\bar{m} \leq 1$. Respeitando essas restrições, os autores fizeram a escolha dos parâmetros para os gráficos adaptativos em função da magnitude γ^* da alteração contra a qual se desejava obter maior proteção. Os projetos fornecidos minimizam o TES para três valores de γ^* ($\gamma^* = 1,5; 2,0$ e $3,0$). O intervalo mais curto de tempo entre amostras (h_S) foi colocado como parâmetro de entrada, e, para mostrar como o valor de TES varia em função de h_S , foram usados dois diferentes valores para este parâmetro ($h_S = 0,10$ e $0,50$). Em geral, o desempenho do gráfico mostrou-se melhor quando se utilizou o menor valor possível para este parâmetro. Resultados mostraram a maior eficiência do esquema Vp na detecção de aumentos de pequena magnitude na incidência de não-conformidades.

A *Tabela 1* mostra um quadro sintético dos artigos aqui mencionados envolvendo regras suplementares, gráficos de contagem inter-eventos e gráficos adaptativos para atributos.

Controle por atributos		
Regras suplementares	Gráficos de contagem inter-eventos	Esquemas adaptativos
Nelson (1997);	Bourke (1991);	Vaughan (1993);
Acosta-Mejia (1999);	Nelson (1994);	Tagaras (1998) (revisão bibliográfica);
	McCool & Joyner-Motley (1998);	Cesar (2000);
	Wu, Yeo & Spedding (2001);	Epprecht & Costa (2001);
	Wu & Yeo (2001);	Epprecht, Costa & Mendes (2003);

Tabela 1 - Artigos sobre regras suplementares, gráficos de contagem inter-eventos e gráficos adaptativos para atributos.

Outros exemplos de gráficos indicados para o monitoramento de processos sujeitos a pequenas perturbações são o **gráfico das Somas Acumuladas (CUSUM, de *Cumulative Sum*)** e o **gráfico da Média Móvel Ponderada Exponencialmente (EWMA, de *Exponentially Weighted Moving Average*)**. O gráfico de controle CUSUM utiliza informações de diversas amostras para decidir sobre o estado do processo: à medida que as amostras são retiradas, são acumuladas as diferenças entre os valores observados da característica de qualidade de interesse e o parâmetro de referência k . Um sinal de processo fora de controle é produzido se estas somas acumuladas excederem ou se igualarem ao valor do intervalo de decisão d . O gráfico de controle EWMA leva em consideração as médias ponderadas exponencialmente das observações da característica de qualidade de interesse, através da constante de amortecimento λ . Por média ponderada exponencialmente entende-se uma média ponderada das observações, com pesos que decaem em progressão geométrica decrescente com a idade das observações (a última tem o maior peso); a razão da progressão geométrica é $(1-\lambda)$. Desse modo, quando um desses mecanismos de controle (CUSUM ou EWMA) está em uso, a decisão sobre o estado do processo é baseada na informação acumulada de diversas amostras, e não apenas na última delas, de modo que, acumulando-se a “pequena evidência” que cada amostra fornece do estado do processo, consegue-se maior rapidez na sinalização de pequenas

alterações. Uma descrição mais detalhada do esquema EWMA será realizada no *Capítulo 3*.

Descreveremos, inicialmente, os principais artigos acerca dos esquemas CUSUM aplicados ao controle de processos por atributos. Podemos destacar os estudos de Gan (1993), para o controle da fração média de não-conformes; de Lucas (1985, 1989), White & Keats (1996), e White, Keats & Stanley (1997), para o controle do número médio de não-conformidades; de Bourke (1991) e Reynolds & Stoumbos (1999), para a contagem inter-eventos no monitoramento de unidades não-conformes; e de Lucas (1985), Vardeman & Ray (1985), Gan (1994), e Gan & Choi (1994), para a contagem inter-eventos no monitoramento de não-conformidades.

Gan (1993) propõe uma estratégia de projeto ótimo de gráficos de controle CUSUM para fração de defeituosos que permite que os parâmetros do gráfico possam ser calculados facilmente. Para uma alteração específica, e dados o tamanho de amostra n e o NMAF (número médio de amostras até um alarme falso), um gráfico CUSUM com um valor do parâmetro de referência k tal que o gráfico apresente o menor NMA (número médio de amostras até o sinal) entre todos os outros gráficos CUSUM com mesmos n e NMAF é conhecido como um gráfico CUSUM ótimo. O artigo fornece figuras com curvas de NMA que permitem que o limite do gráfico CUSUM ótimo seja determinado facilmente. A escolha ótima do tamanho de amostra e a relação entre o gráfico CUSUM e o teste sequencial da razão de probabilidades (SPRT) são também investigados. Resultados mostram que um menor tamanho de amostra é preferível a um maior tamanho de amostra, pois um gráfico CUSUM com um menor tamanho de amostra é capaz de detectar uma fração de defeituosos fora de controle um pouco mais rapidamente; além disso, verificou-se que o valor de referência de um SPRT é de fato a escolha ótima para o parâmetro de referência k de um gráfico CUSUM.

Lucas (1985) descreve o projeto e a implementação de gráficos CUSUM aplicados ao controle do número de defeitos, para dois modelos diferentes: o denominado Poisson CUSUM, que usa a distribuição de Poisson para modelar o número de defeitos observados por amostra, ou seja, um esquema CUSUM para não-conformidades; e o denominado CUSUM para tempo entre eventos, que usa a distribuição exponencial para modelar o tempo entre defeitos. O autor também avalia os benefícios do uso do procedimento de resposta inicial rápida (FIR, de

Fast Initial Response), que torna a detecção de descontroles mais rápida, quando o processo já está fora de controle ao iniciar o monitoramento ou ao reiniciá-lo após um alarme verdadeiro, e também verifica as propriedades de um esquema CUSUM robusto, para analisar os efeitos de observações atípicas ou de *outliers*. O esquema de controle Poisson CUSUM acumula a soma das diferenças entre o valor observado Y_i e um valor de referência k , onde Y_i é o número de não-conformidades em uma amostra i de tamanho n . Para um esquema Poisson CUSUM usado para detectar um aumento no número médio de não-conformidades, a estatística CUSUM é $S_i = \max(0, Y_i - k + S_{i-1})$, enquanto que, para um esquema Poisson CUSUM usado para detectar uma diminuição no número médio de não-conformidades, a estatística CUSUM é calculada por $S_i = \max(0, k - Y_i + S_{i-1})$. Uma vez que a estatística CUSUM não assume valores negativos, o processo é considerado em controle se as amostras estiverem entre zero e o valor do intervalo de decisão d . Um esquema CUSUM padrão usa o valor inicial $S_0 = 0$. Se a estratégia de resposta inicial rápida é utilizada, atribui-se a esse valor inicial S_0 um valor positivo. Um valor recomendado, e usado pelo autor no artigo em questão, é a metade do intervalo de decisão, ou seja, $S_0 = d/2$. Com tal valor positivo para S_0 , o esquema CUSUM sinaliza mais rapidamente o descontrole se o processo já se encontra fora de controle no início do monitoramento, mas, se o processo estiver em controle, a estatística CUSUM decai rapidamente, passando em seguida a oscilar em torno de zero, e o desempenho deste esquema CUSUM (com $S_0 = d/2$) torna-se praticamente igual ao desempenho do esquema CUSUM padrão (com $S_0 = 0$). Para um esquema Poisson CUSUM, o valor de referência k é escolhido de modo a estar entre o número médio aceitável de não-conformidades por amostra (ou seja, quando o processo está em controle), denominado μ_a , e o número médio de não-conformidades por amostra que se deseja que o esquema CUSUM detecte rapidamente, denominado μ_d . Para um gráfico Poisson CUSUM, k é calculado por $k = (\mu_d - \mu_a) / (\ln(\mu_d) - \ln(\mu_a))$, ou seja, é o mesmo valor de referência de um SPRT testando a hipótese nula de que a média é igual a μ_a e a hipótese alternativa de que a média é igual a μ_d . Quando $k \geq 1$, seu valor é arredondado para o mais próximo inteiro. O intervalo de decisão d é escolhido de modo a fornecer um valor aceitável de NMAF, e um valor pequeno de NMA para detectar μ_d . O autor

fornece tabelas de NMA's dos esquemas Poisson CUSUM (e também para os esquemas CUSUM para tempo entre eventos) para uma gama de valores de μ_a , μ_d , k , d , usando-se ou não a estratégia de resposta inicial rápida, e para os casos de um aumento ou de uma diminuição na incidência de não-conformidades. Os NMA's foram calculados modelando a evolução de S_i por cadeias de Markov.

Em outro artigo, Lucas (1989) discute o controle do número de defeitos que ocorrem a níveis muito baixos, descrevendo dois procedimentos práticos baseados em esquemas CUSUM e suas propriedades: o primeiro esquema tem parâmetros $k = 0$ e $d = 1$, portanto sinaliza a cada ocorrência de um defeito; já o segundo esquema CUSUM fornece um sinal de descontrole se existirem dois defeitos dentro de um intervalo de t amostras, e apresenta parâmetros $k = 1/t$ e $d = 1$, podendo se usar FIR com o valor $S_0 = 1 - k$. São fornecidos os valores de NMA, calculados a partir da teoria de cadeias de Markov, para os dois esquemas.

White & Keats (1996) apresentam um programa em QuickBASIC que usa a abordagem das cadeias de Markov para calcular valores de NMA e de momentos de ordens maiores do número de amostras até o sinal para o esquema de controle Poisson CUSUM. O programa pode ser usado para obter o NMA para qualquer esquema Poisson CUSUM, tanto para o caso de um aumento quanto para o caso de uma diminuição do número médio de defeitos, para qualquer magnitude de alteração, usando-se ou não a estratégia de resposta inicial rápida.

White, Keats & Stanley (1997) apresentam uma comparação entre dois procedimentos para controle do número médio de defeitos na amostra: o gráfico Poisson CUSUM e o gráfico de c tradicional (gráfico de Shewhart). Os esquemas são comparados em termos dos NMA's obtidos por cada esquema, tendo os NMA's do esquema Poisson CUSUM sido calculados usando a teoria das cadeias de Markov. Os resultados evidenciam que o gráfico Poisson CUSUM é uma alternativa superior ao gráfico de c , pois fornece uma detecção mais rápida de condições de descontrole que o gráfico de c , gerando ainda um maior NMAF.

Como já citado anteriormente, Bourke (1991) apresenta um gráfico CUSUM, denominado RL CUSUM para fração defeituosa, baseado no número de itens conformes entre dois itens não-conformes produzidos, ou seja, para variáveis com distribuição geométrica, analisando seu desempenho quando é utilizada uma inspeção a 100% na ordem de produção.

Reynolds & Stoumbos (1999) consideram um gráfico CUSUM, denominado Bernoulli CUSUM, baseado nas observações da variável de Bernoulli correspondente à inspeção de itens individuais classificados como defeituosos ou não-defeituosos. Os autores mostram que o gráfico Bernoulli CUSUM é equivalente a um gráfico CUSUM para variáveis aleatórias com distribuição geométrica, baseado na contagem do número de itens não-defeituosos entre dois itens defeituosos. Resultados mostram que o gráfico Bernoulli CUSUM detecta mudanças na fração defeituosa p de maneira substancialmente mais rápida que a abordagem tradicional de agrupar itens em amostras e aplicar um gráfico de p de Shewhart. Além disso, o gráfico Bernoulli CUSUM também se mostra mais eficiente do que agrupar itens em amostras e aplicar um gráfico CUSUM às estatísticas das amostras (gráfico CUSUM para variáveis com distribuição binomial).

Vardeman & Ray (1985) apresentam a resolução de equações integrais para cálculo de NMA's de gráficos CUSUM quando as observações são exponencialmente distribuídas, e é realizada uma aplicação destes gráficos para controlar a intensidade de um processo de Poisson.

Gan (1994) usa o monitoramento da média β de uma distribuição exponencial como um método alternativo para o controle estatístico da taxa de ocorrência de eventos. Em seu artigo, são fornecidas curvas para se obter facilmente os parâmetros de um gráfico CUSUM ótimo para variáveis com distribuição exponencial: as curvas mostram valores de d versus k para gráficos CUSUM unilaterais (ou seja, para aumento ou diminuição da média exponencial β), com valores de NMAF iguais a 100, 150, 200, ..., 500, 600, 700, ..., 1000. Um gráfico CUSUM ótimo é definido aqui como aquele que produz, para um valor de NMAF específico, o menor valor de NMA para detectar uma determinada alteração. O autor também compara os desempenhos do esquema CUSUM proposto e do esquema Poisson CUSUM, verificando que: para pequenas diminuições na média β da distribuição exponencial, ou seja, para pequenos aumentos na taxa de ocorrência de eventos, os esquemas apresentam desempenhos similares; para grandes diminuições em β , o esquema CUSUM proposto é um pouco mais sensível que o esquema Poisson CUSUM quando são usados intervalos pequenos de tempo entre amostras, e, quando são usados intervalos

grandes de tempo entre amostras, o gráfico Poisson CUSUM não é sensível a grandes diminuições em β ; já para pequenos aumentos em β , os esquemas apresentam desempenhos similares; e, para grandes aumentos em β , o esquema Poisson CUSUM mostra-se mais sensível.

Gan & Choi (1994) fornecem um programa para calcular NMA's de gráficos CUSUM quando as observações são exponencialmente distribuídas. Os cálculos são baseados nos resultados de Vardeman & Ray (1985), que, como já foi visto, fornecem expressões exatas para os NMA's desses gráficos através da solução de equações diferenciais simples de primeira ordem.

Alguns estudos incorporaram a estratégia de esquemas adaptativos a métodos CUSUM. No entanto, a grande maioria destes estudos está voltada para o controle de processos por variáveis. Para atributos, o único trabalho existente envolvendo o método CUSUM adaptativo é o de Rendtel (1990), que introduz e investiga esquemas CUSUM com intervalos de amostragem e tamanhos de amostra variáveis, para controlar a fração defeituosa. Seu procedimento utiliza vários tamanhos de amostra e intervalos de amostragem, e o desempenho do procedimento é analisado aplicando a teoria de cadeias de Markov, com um grande número de estados transientes.

A *Tabela 2* lista os principais artigos envolvendo o método CUSUM para o controle por atributos.

Estatística amostral	Esquema CUSUM para atributos	
	Com parâmetros fixos	Adaptativo
não-conformes	Gan (1993);	Rendtel (1990);
não-conformidades	Lucas (1985, 1989); White & Keats (1996); White, Keats & Stanley (1997);	-----
contagem inter-eventos (não-conformes)	Bourke (1991); Reynolds & Stoumbos (1999);	-----
contagem inter-eventos (não-conformidades)	Lucas (1985); Vardeman & Ray (1985); Gan (1994); Gan & Choi (1994);	-----

Tabela 2 - Principais artigos sobre o esquema CUSUM para atributos.

No caso do método EWMA, a grande maioria dos estudos e a totalidade dos que envolvem a estratégia de gráficos adaptativos também está voltada para gráficos de controle por variáveis. Alguns estudos estão voltados para o controle de processo por atributos, mas não há na literatura trabalho envolvendo o esquema EWMA adaptativo para atributos.

A *Tabela 3* mostra esquematicamente os principais artigos que envolvem o esquema EWMA relacionado ao controle de processos por variáveis (monitorando a estatística \bar{X}) e por atributos (para np , para c , e para contagem inter-eventos para o monitoramento de não-conformes e para o monitoramento de não-conformidades). A seguir, descreveremos os artigos citados nesta tabela.

Estatística amostral	Esquema EWMA	
	Com parâmetros fixos	Adaptativo
média amostral	Crowder (1987a, 1987b, 1989); Lucas & Saccucci (1990); Saccucci & Lucas (1990);	Sacucci, Amin & Lucas (1992); Reynolds & Arnold (2001);
não-conformes	Gan (1990a);	-----
não-conformidades	Gan (1990b); Borrer, Champ & Rigdon (1998);	-----
contagem inter-eventos (não-conformes)	McCool & Joyner-Motley (1998);	-----
contagem inter-eventos (não-conformidades)	Gan (1998); Gan & Chang (2000);	-----

Tabela 3 - Principais artigos sobre o esquema EWMA.

Crowder (1987a) fornece um programa que calcula NMA's de gráficos EWMA para variáveis, usando a solução de sistemas de equações lineares algébricas, através de quadratura gaussiana.

O mesmo autor (Crowder (1987b)) apresenta um procedimento numérico, usando equações integrais, para o cálculo de momentos do número de amostras até um sinal de gráficos EWMA. NMA's e desvios-padrão do número de amostras até um sinal são calculados para o caso de gráficos EWMA bilaterais, assumindo que as variáveis são independentes e normalmente distribuídas. O procedimento fornecido, usando equações integrais, pode ser estendido a casos onde a

distribuição das observações não é normal e a versões unilaterais de gráficos EWMA.

Ainda Crowder (1989) mostra um procedimento simples para o projeto de um esquema EWMA para controle de processo por variáveis. A aplicação do procedimento é facilitada com o uso de gráficos, que são fornecidos, para identificação dos parâmetros ótimos do esquema EWMA. Os parâmetros escolhidos são considerados ótimos no sentido em que eles produzem, para um valor de NMAF específico, o menor valor de NMA para detectar um determinado desvio da média. O autor realiza, também, uma análise de sensibilidade destes parâmetros.

Lucas & Saccucci (1990) avaliam, através de cálculos usando cadeia de Markov, as propriedades de um esquema de controle EWMA para monitoramento da média do processo por observações individuais usando limites de controle constantes (baseados na variância assintótica). São fornecidos vários projetos alternativos que produzem um NMAF = 500 (vários pares da constante de amortecimento do esquema EWMA, λ , e larguras dos limites de controle, $L's$), e, para cada projeto, os NMA's para diversos desvios da média. São considerados os casos *zero-state* e *steady-state*. *Zero-state* corresponde ao caso em que o processo começa em controle, e *steady-state*, ao caso em que o desvio ocorre quando o efeito do valor inicial já se tornou desprezível. Os autores possuem (embora não forneçam no artigo) mais tabelas para outros valores de NMAF. É fornecido um procedimento para o projeto do gráfico, com o qual se obtém uma tabela com parâmetros ótimos de esquemas EWMA: para um NMAF específico e um dado desvio da média, são fornecidos os parâmetros do esquema EWMA que apresenta o menor valor de NMA para esse dado desvio. Depois, são analisadas e discutidas as propriedades de alguns aprimoramentos para esquemas EWMA. Um primeiro aprimoramento é o procedimento de resposta inicial rápida (FIR, de *Fast Initial Response*), que torna o esquema mais sensível a causas especiais quando do início do monitoramento (*start-up*), aumentando a rapidez de sinalização das mesmas. Outro aprimoramento é o uso um esquema Shewhart-EWMA combinado, que fornece proteção tanto para desvios de pequena magnitude quanto para desvios de grande magnitude no processo. Por último, é apresentado o esquema EWMA “robusto”, que fornece proteção contra *outliers* ocasionais nos dados. É feita

também uma extensa comparação entre os esquemas EWMA e CUSUM, verificando-se que seus NMA's são similares.

Em outro artigo (Saccucci & Lucas (1990)), os mesmos autores fornecem um programa em Fortran que calcula NMA's de *zero-state* e de *steady-state* usando a teoria de cadeias de Markov para esquemas EWMA e Shewhart-EWMA combinados, assumindo observações independentes e normalmente distribuídas.

Quanto a esquemas EWMA usando a estratégia de gráficos adaptativos, Saccucci, Amin & Lucas (1992) obtêm, através de cadeias de Markov, os valores de TES_f e de TES de esquemas de controle EWMA bilaterais usando VSI para monitoramento de observações normais, independentes e identicamente distribuídas. O artigo ainda fornece um procedimento útil para projeto do esquema e mostra comparações de desempenho com gráficos de \bar{X} e CUSUM que também usam intervalo de tempo entre amostras variável.

Reynolds & Arnold (2001) consideram gráficos de controle de \bar{X} VSS e de \bar{X} VSI, e desenvolvem métodos usando equações integrais e cadeias de Markov para calcular as propriedades destes gráficos. Gráficos EWMA com VSS e/ou VSI também são estudados em detalhe, e são investigadas diferentes maneiras de se definir a estatística de controle EWMA. São fornecidas orientações para ajudar na escolha dos tamanhos de amostra e dos intervalos de tempo entre amostras para estes gráficos. São feitas também comparações dos esquemas EWMA propostos com gráficos CUSUM e gráficos de \bar{X} que também usam esquemas VSS e/ou VSI. Os autores mostram que o uso de uma média amostral padronizada fornece bom desempenho e simplifica a especificação dos limites de controle requeridos. É mostrado também que o uso do esquema VSI em um gráfico EWMA fornece um aprimoramento substancial na rapidez de detecção de desvios pequenos e moderados na média do processo. O esquema VSI EWMA é recomendado para aplicações nas quais é factível variar-se o intervalo de tempo entre amostras. O uso do esquema VSS em gráficos EWMA, indicado para aplicações nas quais é factível variar-se o tamanho de amostra, mas não é factível variar-se o intervalo de tempo entre amostras, também fornece um aprimoramento significativo na capacidade de detecção de desvios pequenos e moderados na média do processo, mas o aprimoramento não é tão grande quando comparado com o uso do esquema VSI. Se é factível o uso tanto do esquema VSI quanto do esquema VSS, e se o

menor valor possível para o intervalo pequeno de tempo entre amostras é relativamente grande ou caso se esteja usando um esquema EWMA unilateral, algum aprimoramento adicional na capacidade de detecção em relação aos esquemas VSS e VSI pode ser obtido usando o esquema VSSI EWMA; já quando o menor valor possível do intervalo pequeno de tempo entre amostras é relativamente pequeno, não há muita vantagem em se usar o gráfico VSSI EWMA bilateral, pois o gráfico VSI EWMA bilateral já funciona suficientemente bem.

Para o controle de processo por atributos utilizando o esquema EWMA, o número de trabalhos existentes é bem menor, e, como já foi dito, os estudos ainda não envolvem esquemas adaptativos. Entretanto, por analogia com os resultados de esquemas EWMA adaptativos para controle por variáveis, é de se esperar que, também para controle por atributos, esquemas EWMA adaptativos proporcionem ganhos de desempenho em relação ao esquema EWMA de parâmetros fixos. Essa é a motivação desta pesquisa. Descreveremos, agora, então, os trabalhos existentes em esquemas EWMA (de parâmetros fixos) para atributos.

Gan (1990a) desenvolve um gráfico EWMA modificado (onde a estatística EWMA é arredondada para o inteiro mais próximo), para monitorar dados com distribuição binomial (número de defeituosos). Os NMA's e a função de probabilidade do número de amostras até um sinal são calculados usando cadeias de Markov. O gráfico EWMA modificado proposto demonstra ser, em geral, superior a um gráfico de Shewhart, em termos de NMA.

Em Gan (1990b), são desenvolvidos três gráficos EWMA modificados, para monitorar dados com distribuição de Poisson (número médio de defeitos na amostra). Em um primeiro caso, a estatística EWMA é arredondada para o inteiro mais próximo (gráfico denominado REWMA, onde o "R" vem de *rounded*); no segundo caso, a estatística EWMA denota o menor inteiro maior ou igual ao valor obtido (gráfico denominado CEWMA, onde o "C" vem de *ceiling*), e, no último caso, a estatística EWMA denota o maior inteiro menor ou igual ao valor obtido (gráfico denominado FEWMA, onde o "F" vem de *floor*). Os NMA's e a função de probabilidade do número de amostras até um sinal dos gráficos EWMA modificados são também calculados usando cadeias de Markov, e os gráficos propostos demonstram ser superiores a um gráfico de controle de Shewhart, em termos de NMA.

Borror, Champ & Rigdon (1998) introduzem um gráfico de controle EWMA para monitorar o número médio de defeitos na amostra. As medidas de desempenho do gráfico são obtidas usando uma cadeia de Markov, e seus NMA's são comparados com os de outros esquemas de controle para dados com distribuição de Poisson. São fornecidas, para alguns valores de número médio de defeitos, curvas de NMAF versus largura dos limites de controle, parametrizadas pelo valor da constante de amortecimento λ , de modo que, para um determinado valor de λ , o usuário possa escolher o valor da largura dos limites de controle que forneça um NMAF desejável.

Na contagem inter-eventos para o monitoramento de não-conformes, como já foi citado anteriormente, o gráfico proposto por McCool & Joyner-Motley (1998) tem seu desempenho melhorado quando é aplicado um esquema EWMA à estatística transformada.

Na contagem inter-eventos para o monitoramento de não-conformidades, destaca-se o artigo de Gan (1998), no qual os NMA's de gráficos EWMA aplicados a variáveis com distribuição exponencial são determinados por meio de equações diferenciais. É fornecido um procedimento simples para determinação dos parâmetros desses gráficos, unilaterais e bilaterais. O gráfico obtido usando o procedimento proposto é ótimo, no sentido que, para um mesmo valor de NMAF, é o que apresenta o menor valor de NMA para detectar uma determinada alteração no processo. São também comparados os NMA's de gráficos EWMA e CUSUM.

Gan & Chang (2000) apresentam um programa para determinar os perfis de NMA's (curvas NMA \times alteração no processo) de gráficos EWMA unilaterais e bilaterais com distribuição exponencial, que fornecem uma melhor compreensão do desempenho dos gráficos.