

## 2 Fundamentos Conceituais e Trabalhos Precedentes

### 2.1. Fundamentos Conceituais

Para controlar estatisticamente um processo produtivo é prática comum utilizar um par de gráficos de controle, um para monitoramento da média e outro para monitoramento da dispersão do processo. Dentre os gráficos existentes para o monitoramento da dispersão de um processo, objeto de estudo neste trabalho, destacam-se os gráficos da amplitude amostral  $R$  e do desvio-padrão amostral  $S$  como os mais utilizados.

O gráfico de  $R$  é largamente utilizado devido à facilidade com que as amplitudes são calculadas, mas deixa a desejar quanto à estimação da dispersão do processo, já que para o cálculo da amplitude são utilizados apenas os valores extremos da amostra (valor máximo e mínimo da amostra). Para Montgomery (2001) “o gráfico do desvio-padrão  $S$  oferece maior eficiência na estimação da dispersão do processo e é mais flexível para aplicações que envolvam tamanhos de subgrupos maiores e/ou diferentes que o gráfico de  $R$ ”. Além disso, com a difusão do uso de planilhas eletrônicas, os cálculos tornam-se automatizados e o gráfico de  $S$  torna-se tão simples de utilizar quanto o de  $R$ , cuja sobrevivência só se explica por razões históricas e pelo desconhecimento da vantagem do gráfico de  $S$  sobre este.

Os limites de controle para os gráficos de  $S$  devem ser baseados na distribuição da estatística  $S$ . Supondo que a variável a ser monitorada siga uma distribuição normal com média  $\mu$  e variância  $\sigma^2$ , a variância amostral  $S^2$  é um estimador não-viesado para  $\sigma^2$ . Entretanto, o mesmo não pode ser dito a respeito do desvio-padrão amostral  $S$ . De fato, é sabido que:

$$\mu_S = E(S) = c_4\sigma \quad (2.1.1)$$

onde  $c_4$  é determinado a partir do tamanho de amostra  $n$  e dado por:

$$c_4 = \frac{\Gamma\left(\frac{n}{2}\right)}{\Gamma\left(\frac{n-1}{2}\right)} \sqrt{\frac{2}{n-1}} \quad (2.1.2)$$

O desvio-padrão do estimador  $S$  é dado por:

$$\sigma_S = \sqrt{\text{Var}(S)} = \sqrt{\sigma^2 - c_4^2 \sigma^2} = \sigma \sqrt{1 - c_4^2} \quad (2.1.3)$$

Para a dedução da forma da distribuição de  $S$ , de seus parâmetros e da expressão para  $c_4$ , ver, por exemplo, Hald (1952). Tabelas dos valores de  $c_4$  costumam ser encontradas em livros de Controle Estatístico da Qualidade, como Montgomery (2001) ou Costa et al. (2005).

A partir das equações (2.1.1) e (2.1.3) pode-se calcular a linha central ( $LC$ ) e os limites de controle de três-sigma ( $3\sigma$ ) para o gráfico de  $S$  ( $LSC_S$  e  $LIC_S$ ), de acordo com as seguintes equações:

$$LSC_S = c_4 \sigma_0 + 3 \sigma_0 \sqrt{1 - c_4^2} \quad (2.1.4)$$

$$LC_S = c_4 \sigma_0 \quad (2.1.5)$$

$$LIC_S = \max \left\{ 0, c_4 \sigma_0 - 3 \sigma_0 \sqrt{1 - c_4^2} \right\} \quad (2.1.6)$$

onde  $\sigma_0$  é o desvio-padrão do processo em controle. Os limites de três-sigma apresentados nas fórmulas acima são usuais devido à sua simplicidade. Neles, a distância entre cada um dos limites e a linha central do gráfico de controle equivale a três desvios-padrão da estatística em questão, no caso,  $S$ .

Se  $\sigma_0$  não for conhecido, então ele deve ser estimado. O procedimento recomendado para a sua estimação consiste na retirada de  $m$  amostras do processo, cada uma com tamanho  $n$ , durante um período em que se tenha confiança de que o

processo está em controle, isento de causas especiais de variabilidade; seguida do cálculo da média  $\bar{S}$  dos  $m$  desvios-padrão  $S_j$ ,  $j = 1, 2, \dots, m$ , obtidos para cada amostra, dada por:

$$\bar{S} = \frac{1}{m} \sum_{j=1}^m S_j \quad (2.1.7)$$

No caso de controle de processos, trabalha-se com essa média dos  $S_j$  por ser um estimador mais robusto (menos sensível a observações anômalas) que o estimador  $S$  calculado com todos os  $(m, n)$  pontos considerados como uma amostra única [ver, por exemplo, Wheeler (1995) ou Costa et al. (2005)].

A partir das equações (2.1.4), (2.1.5), (2.1.6) e da constatação de que  $\bar{S}/c_4$  é um estimador não-viesado de  $\sigma$ , através de (2.1.1) e (2.1.7), os limites de controle de  $3\sigma$  e a linha central do gráfico de  $S$  são:

$$LSC_S = \bar{S} + 3 \frac{\bar{S}}{c_4} \sqrt{1 - c_4^2} \quad (2.1.8)$$

$$LC_S = \bar{S} \quad (2.1.9)$$

$$LIC_S = \max \left\{ 0, \bar{S} - 3 \frac{\bar{S}}{c_4} \sqrt{1 - c_4^2} \right\} \quad (2.1.10)$$

De maneira geral, quando se monitora a dispersão do processo deseja-se detectar rapidamente aumentos na dispersão do mesmo. Nessa situação, utiliza-se um gráfico de controle unilateral (para detectar mudanças somente em uma direção) para a dispersão do processo, em que se observa se as estatísticas plotadas não excedem um limite superior de controle calculado.

É conhecido que  $(n-1)S^2/\sigma^2$  segue uma distribuição qui-quadrado com  $n-1$  graus de liberdade [ver, por exemplo, Costa et al. (2005)]. Portanto, o valor

de  $LSC$  que, se o processo tiver desvio-padrão igual a  $\sigma_0$ , é ultrapassado com probabilidade  $\alpha$  é o valor que satisfaz à seguinte relação:

$$\chi_{n-1, \alpha}^2 = \frac{(n-1)LSC_S^2}{\sigma_0^2} \quad (2.1.11)$$

onde  $\chi_{n-1, \alpha}^2$  é quantil de  $(1-\alpha)$  da variável qui-quadrado com  $n-1$  graus de liberdade. Substituindo, em (2.1.11),  $\sigma_0$  desconhecido por sua estimativa  $\hat{\sigma}_0$ , obtém-se:

$$LSC_S = \sqrt{\frac{\chi_{n-1, \alpha}^2 \hat{\sigma}_0^2}{(n-1)}} \quad (2.1.12)$$

Note-se que a expressão (2.1.12) acima se aplica ao gráfico de controle unilateral de  $S$ , sem  $LIC$ . Este é o tipo de gráfico que será considerado neste trabalho, partindo da suposição que o interesse é detectar eventuais aumentos na dispersão do processo.

A análise de desempenho dos gráficos de controle consiste em medir a sensibilidade dos mesmos a perturbações, sensibilidade essa que pode ser medida pela sua probabilidade (ou por sua rapidez, ou ainda por sua eficiência – razão entre medidas de “rapidez” e de “custo”) de sinalizar um deslocamento dos parâmetros em relação ao valor sob controle (valor-alvo). Na literatura de CEP, tal análise de desempenho pode ser desenvolvida segundo a lógica dos testes de hipóteses, em que:

$$H_0 = \text{processo livre de causas especiais } (\sigma = \sigma_0) \quad (2.1.13)$$

$$H_1 = \text{processo sob a influência de causas especiais } (\sigma = \sigma_1) \quad (2.1.14)$$

onde  $\sigma_0$  representa, como já visto, o valor-alvo para a dispersão do processo e  $\sigma_1$  representa a dispersão do processo quando este está sob a influência de causas

especiais. Se o processo estiver sob controle estatístico (a hipótese  $H_0$  é verdadeira),  $\alpha$  representa o risco (probabilidade) de erroneamente se considerar o processo fora de controle – probabilidade de alarme falso ou erro do tipo I. Para o gráfico de  $S$ , o risco  $\alpha$  corresponde a:

$$\alpha = P [ S > LSC_S \text{ ou } S < LIC_S \mid \sigma = \sigma_0 ] \quad (2.1.15)$$

Se o processo estiver fora de controle (a hipótese  $H_0$  é falsa),  $\beta$  representa o risco (probabilidade) de erroneamente se considerar o processo em controle – probabilidade de não-detecção ou erro do tipo II. Portanto corresponde a:

$$\beta = P [ LIC_S \leq S \leq LSC_S \mid \sigma = \sigma_1 ] \quad (2.1.16)$$

Freqüentemente deseja-se determinar a probabilidade de sinalização de uma causa especial pelo gráfico, calculada como o complemento da probabilidade do erro do tipo II. Essa probabilidade, portanto, que será simbolizada pela sigla  $Pd$ , representa o poder do gráfico de controle, e corresponde a:

$$Pd = 1 - \beta = P [ S > LSC_S \text{ ou } S < LIC_S \mid \sigma = \sigma_1 ] \quad (2.1.17)$$

Como é possível observar, as expressões (2.1.15) e (2.1.17) para  $\alpha$  e para  $Pd$  seriam equivalentes se não fosse pela condição em que cada probabilidade é calculada:  $\alpha$  representa a probabilidade de sinal (ponto fora dos limites) quando o processo está em controle ( $\sigma = \sigma_0$ ); e  $Pd$  representa a probabilidade de sinal quando o processo está fora de controle ( $\sigma = \sigma_1$ ).

O cálculo dessas probabilidades requereria, em princípio, o conhecimento da distribuição de probabilidade da variável aleatória  $S$ . Segundo Costa et al. (2005), a distribuição do desvio-padrão  $S$  não tem sido tabelada. Entretanto, como para qualquer  $a > 0$ ,  $Pr [S^2 > a^2] = Pr [S > a]$ , é possível obter os valores de  $\alpha$ ,  $\beta$  ou  $Pd$  do gráfico de  $S$  pela distribuição de qui-quadrado, já que  $(n-1)S^2/\sigma^2$  segue uma distribuição de qui-quadrado ( $\chi^2$ ) com  $n-1$  graus de liberdade.

Para fixar a probabilidade de alarme falso em um nível especificado, pode-se utilizar, alternativamente aos limites de três-sigma, limites de probabilidade, que são obtidos escolhendo-se diretamente a probabilidade do erro do tipo I e determinando-se então, os limites de controle correspondentes a este risco  $\alpha$ , em função da distribuição de probabilidade da estatística usada no gráfico (no caso,  $S$ ).

Segundo Montgomery (2001), “é prática padrão, nos Estados Unidos, determinar os limites de controle como um múltiplo do desvio-padrão da estatística mostrada no gráfico”. Em geral utiliza-se o múltiplo três, pelo fato de darem bons resultados na prática e dificilmente se conhecer a distribuição de probabilidade da característica de qualidade, impossibilitando o cálculo de limites de probabilidade exatos. Wheeler (1995) diz que limites de três-sigma levam a uma baixa probabilidade de alarme falso com qualquer distribuição e Montgomery (2001) diz: “Se a distribuição da característica de qualidade pode ser razoavelmente aproximada pela distribuição normal, então haverá pouca diferença entre os limites três-sigma e de probabilidade 0,001”.

Outras medidas de desempenho além de  $\alpha$ ,  $\beta$  e  $Pd$  podem ser utilizadas para medir a eficiência dos gráficos de controle. Dentre as mais freqüentemente utilizadas destacam-se o número médio de amostras até o sinal ( $NMA$ , em inglês  $ARL$ ), o número médio de amostras até um alarme falso ( $NMA_0$ , em inglês  $ARL_0$ ), o tempo esperado até o sinal ( $TES$ ) e o tempo médio até um alarme falso ( $TMAF$ ).

Para o gráfico de controle de  $S$ , assim como para os demais gráficos de Shewhart, em que as amostras são independentes, o número de pontos que serão plotados (portanto, de amostras que serão retiradas) até que o primeiro ponto exceda um limite de controle (sinal), segue uma distribuição geométrica com parâmetro  $p$ , média  $1/p$  e desvio-padrão  $\sqrt{1-p}/p$ . Quando a hipótese  $H_0$  for verdadeira ( $\sigma = \sigma_0$ )  $p$  será igual a  $\alpha$  e  $NMA_0$ , o número médio de amostras até o sinal (alarme falso), será calculado como a média dessa distribuição geométrica de parâmetro  $\alpha$ , como segue:

$$NMA_0 = \frac{1}{\alpha} \quad (2.1.18)$$

Quando a hipótese  $H_0$  for falsa ( $\sigma = \sigma_1$ ),  $p$  será igual ao  $Pd$  do gráfico de controle e  $NMA$ , o número médio de amostras até o sinal verdadeiro, será calculado como a média dessa distribuição geométrica de parâmetro  $Pd$ , como segue:

$$NMA = \frac{1}{Pd} = \frac{1}{(1 - \beta)} \quad (2.1.19)$$

O tempo esperado até o sinal ( $TES$ ) e o tempo médio até um alarme falso ( $TMAF$ ) são medidas de desempenho que derivam das medidas  $NMA_0$  e  $NMA$ . O  $TES$  representa o tempo médio transcorrido desde a ocorrência da causa especial (que alterou os parâmetros do processo) até que seja gerado um sinal no gráfico de controle, na situação em que  $\sigma = \sigma_1$ . Se as amostras são retiradas a intervalos regulares de tempo  $h$ , então o  $TES$  é dado por [ver Costa et al. (2005)]:

$$TES = (NMA - 0,5)h \quad (2.1.20)$$

O tempo médio até um alarme falso ( $TMAF$ ) indica o tempo esperado até a ocorrência de um alarme falso (com  $\sigma = \sigma_0$ ). Se as amostras são retiradas a intervalos regulares de tempo  $h$ , então o  $TMAF$  é obtido a partir da equação:

$$TMAF = NMA_0 h \quad (2.1.21)$$

Essas medidas de desempenho são úteis para o *projeto dos gráficos de controle*, que consiste na escolha dos parâmetros de implementação ( $n$ ,  $h$ , limites de controle, etc.) que garantam uma frequência de alarmes falsos suficientemente baixa (representada por valores baixos para  $\alpha$  e valores altos para  $NMA_0$  e  $TMAF$ ) e uma rápida detecção de descontroles no processo (representada por valores baixos de  $NMA$ ,  $TES$  e  $\beta$  e altos valores de  $Pd$ ).

Há diversas classes de modelos para este fim: modelos estatísticos, que não consideram o tempo entre amostras  $h$  (por exemplo, fixando  $NMA_0$  e buscando o valor de  $n$  que garante certo poder especificado de detectar uma alteração nos parâmetros do processo de magnitude considerada relevante, por levar a um

aumento significativo da probabilidade de produzir itens defeituosos e que, portanto, se quer detectar rapidamente) ou até consideram o tempo entre amostras, mas sem entrar em considerações de custos; modelos de *design* econômico, que consideram diferentes componentes de custo (em geral, no mínimo os custos de investigar alarmes falsos e verdadeiros, de amostragem e de se produzir com o processo fora de controle) e procuram determinar os parâmetros do processo que minimizam uma função “custo total por unidade de tempo”; modelos semi-econômicos, que consideram alguns componentes de custo, mas não todos, ou os consideram indiretamente; modelos econômicos com o acréscimo de restrições estatísticas (como restrições ao  $NMA_0$  e ao  $NMA$  para diversas magnitudes de alteração).

O número de trabalhos sobre o desempenho de gráficos de controle e sobre o seu projeto (escolha dos valores “ótimos” para seus parâmetros  $n$ ,  $h$ , limites de controle, segundo variados critérios - medida de desempenho ou de custo) é enorme. Quaisquer que sejam, porém, a classe de modelos ou o modelo específico em questão, são raríssimos, excepcionais, os casos em que não sejam usados como argumentos as probabilidades  $\alpha$ ,  $\beta$  ou  $Pd$ , que a grande maioria dos modelos considera conhecidos com precisão absoluta.

Assim, as probabilidades  $\alpha$  e  $\beta$  (ou  $Pd$ ) podem ser consideradas as medidas de desempenho básicas de um gráfico de controle, porque as demais ( $NMA$ ,  $NMA_0$ ,  $TES$  e  $TMAF$ ), como visto, são função delas. Elas dependem do valor dos limites de controle (ver equações 2.1.15 a 2.1.17), que por sua vez, são determinados usando estimativas dos parâmetros do processo (no caso em estudo – do gráfico de  $S$  – a estimativa do desvio-padrão do processo). Assim, erros na estimação do desvio-padrão do processo resultam em erros nos valores calculados para as medidas de desempenho do gráfico de controle.

Outra ferramenta do controle estatístico de qualidade com grande utilização são os índices de capacidade do processo. Como mencionado anteriormente, assim como ocorre com os limites dos gráficos de controle, a análise de capacidade de processos também é baseada em estimativas dos parâmetros dos mesmos. A precisão dos índices de capacidade, que verificam a capacidade do processo de atender às especificações de projeto (limites de especificação), é determinada pela precisão das estimativas da dispersão do processo.

A seguir é apresentado um breve resumo de índices de capacidade, para o leitor menos familiarizado.

Segundo Montgomery (2001), “a capacidade do processo refere-se à sua uniformidade, que é medida através da variabilidade no processo.” Essa capacidade depende então, das especificações de projeto e da variabilidade do processo, de forma que alterações provocadas por causas especiais reduzem a sua capacidade e aumentam o número de itens não conformes produzidos.

Uma forma de expressar quantitativamente a capacidade de um processo consiste em utilizar índices ou razões de capacidade para medir o quanto o processo consegue atender às especificações. Existem vários índices de capacidade do processo; dentre eles, destacam-se os índices  $C_p$ ,  $C_{pk}$  e  $C_{pm}$  como os mais utilizados. Quanto maior o valor do índice, maior a capacidade do processo em atender às especificações. Os índices  $C_p$ ,  $C_{pk}$  e  $C_{pm}$  são obtidos através das seguintes equações:

$$C_p = \frac{LSE - LIE}{6\sigma} \quad (2.1.22)$$

$$C_{pk} = \text{Min} \left\{ \frac{LSE - \mu}{3\sigma}, \frac{\mu - LIE}{3\sigma} \right\} \quad (2.1.23)$$

$$C_{pm} = \frac{LSE - LIE}{6\sqrt{\sigma^2 + (d - \mu)^2}} \quad (2.1.24)$$

onde  $d = (LSE + LIE)/2$ , e  $LSE$  e  $LIE$  representam, respectivamente, os limites superior e inferior de especificação.

Ramos (2006) comenta que o numerador do  $C_p$  dá o tamanho da amplitude sobre a qual as observações do processo podem variar enquanto que o denominador, o tamanho da amplitude sobre a qual o processo está variando realmente. Como o valor do índice não é sensível ao valor da média do processo, ele é útil como uma medida de capacidade do processo quando centrado. Se o processo não estiver centrado no ponto médio das especificações, o índice  $C_p$  só

tem sentido como um indicador da capacidade potencial do processo, quando sua média for reajustada.

Embora não seja sensível à média do processo, o índice  $C_p$  é afetado diretamente pelo seu desvio-padrão. Da mesma forma, em consequência, sua precisão é afetada pela precisão da estimativa do desvio-padrão do processo; isso porque, na maioria das aplicações práticas, o desvio-padrão  $\sigma$  é quase sempre desconhecido e deve ser substituído por uma estimativa. English & Taylor (1993) chamam a atenção para a possibilidade de ocorrência de grandes erros ao se estimar os índices de capacidade com base em dados amostrais e questionam a efetividade desses índices em medir a capacidade do processo quando isso ocorre.

Para um processo ser considerado capaz pelo índice  $C_p$ , deve apresentar um  $C_p$  igual ou superior a 1,00. Segundo Kotz & Johnson (1993), supondo que os dados sejam normalmente distribuídos e  $\sigma$  seja corretamente estimado,  $C_p = 1$  garante que a proporção de itens defeituosos esperada nunca será menor do que 0,27%. Entretanto, valores superiores a  $C_p = 1,00$  têm sido considerados como valores mínimos recomendáveis para este índice. O valor de  $C_p = 1,33$  é um mínimo recomendável muito utilizado e garante que os dados do processo utilizam no máximo 75% da amplitude de especificação. Em outras palavras, garante uma taxa de itens defeituosos muito baixa, aproximadamente 0,007%. Montgomery (2001) recomenda que seja utilizado um  $C_p = 1,33$  para processos existentes, o que corresponde a uma diferença entre os limites superior e inferior de especificação de  $8\sigma$ ; e um  $C_p = 1,55$  para processos novos, correspondente a uma diferença entre os limites superior e inferior de especificação de  $9\sigma$ .

Na verdade, o valor de  $C_p$  considerado desejável (alvo) depende do processo particular em questão, do tipo de produto, da gravidade das consequências de se produzir uma unidade não conforme. Em qualquer caso, o índice é uma medida de capacidade do processo e precisa ser calculado e confrontado com o valor de referência, qualquer que seja ele.

O índice  $C_{pk}$ , diferentemente do índice  $C_p$ , leva em consideração a centralidade do processo e é definido como um  $C_p$  unilateral para o limite de especificação mais próximo da média do processo (pior situação). Isto é, o índice  $C_{pk}$ , além de avaliar a variabilidade natural do processo em relação à variabilidade permitida, verifica também a posição do processo em relação aos limites (superior e inferior) de especificação. Entretanto, para que o  $C_{pk}$  seja uma

medida adequada de centralidade de um processo, ele deve ser comparado com o  $C_p$ . Quando se verificar  $C_{pk} = C_p$ , o processo estará centrado e quando  $C_p > C_{pk}$ , o processo estará descentrado. Para que um processo seja considerado capaz, deve possuir  $C_{pk} \geq 1$ .

Em relação à variabilidade do processo, para um valor fixo de  $\mu$ ,  $C_{pk}$  varia inversamente em relação a  $\sigma$  (da mesma forma que  $C_p$ ) – diminui quando  $\sigma$  aumenta e aumenta quando  $\sigma$  tende para zero. Este fato pode desqualificar  $C_{pk}$  como um indicador de centralidade do processo. Dessa limitação do índice  $C_{pk}$  originou-se uma razão da capacidade para melhor avaliar um processo descentrado, o índice  $C_{pm}$ . Quando  $d = \mu$ , os três índices se igualam ( $C_{pm} = C_{pk} = C_p$ ) e à medida que  $\mu$  se afasta de  $d$ ,  $C_{pm}$  e  $C_{pk}$  diminuem em relação à  $C_p$ .

Segundo Costa et al. (2005), uma desvantagem do  $C_{pm}$  se deve ao fato deste índice penalizar os processos mais por falta de centralidade do que pela proporção de itens não conformes. Sendo assim, processos que produzam iguais proporções de itens não conformes podem ter valores de  $C_{pm}$  muito diferentes e processos que produzam proporções de itens não conformes muito diferentes podem ter valores de  $C_{pm}$  próximos.

## 2.2. Trabalhos Precedentes

A influência da estimação dos parâmetros do processo no desempenho dos gráficos de controle é reconhecida há bastante tempo. As recomendações clássicas encontradas nos livros de CEP, desde Shewhart, acerca do número ( $m$ ) e tamanho ( $n$ ) de amostras iniciais necessárias para estabelecimento dos limites para os gráficos de controle é de 20 a 30 amostras de tamanho 4 ou 5, tomadas quando o processo está estável (“em controle”), ou seja, com a variável de interesse apresentando uma distribuição constante, com média e variância constantes. Por exemplo, Montgomery (2001) recomenda usar 20 a 25 amostras (ou “subgrupos” de tamanho 3, 4 ou 5) para calcular limites de controle provisórios que sejam confiáveis. No entanto, há vários artigos que apresentam resultados de pesquisa mostrando que o número de amostras necessário para se ter confiança no desempenho dos gráficos (que, por exemplo, a taxa de alarmes falsos não será muito maior que a especificada) deve ser bem maior que tais 20, 25 ou 30 amostras.

Nesta Seção, são apresentados alguns trabalhos já existentes sobre o desempenho dos gráficos de controle de Shewhart quando os parâmetros são estimados, bem como as recomendações de alguns autores a respeito do tamanho e número de amostras necessário para se atingir determinado grau de eficiência sob a ótica de determinada medida de desempenho.

Jensen *et al.* (2006) apresentam uma revisão bibliográfica sobre o efeito da estimação dos parâmetros nas propriedades dos gráficos de controle. Trata-se de um trabalho abrangente, pois, além de reunir os principais feitos nesta área em relação a diversos tipos de gráficos, desde a década de 50 até o presente, traz recomendações e sugestões para trabalhos futuros que possam responder questões até o momento em aberto.

Os autores reportam que, na prática, os parâmetros dos gráficos de controle raramente são conhecidos e, por isso, são frequentemente baseados em estimativas. Ao se utilizar estimativas no lugar de parâmetros conhecidos, a variabilidade dos estimadores pode resultar em um desempenho do gráfico diferente daquele construído com parâmetros conhecidos. Das recentes pesquisas

realizadas sobre o efeito da estimação dos parâmetros nas propriedades dos gráficos de controle, os autores destacam três questões que surgem com frequência nos trabalhos sobre o tema:

- a) “Quanto se perde em termos de desempenho quando são utilizados parâmetros estimados ao invés de conhecidos?”
- b) “Que tamanho de amostra é necessário na *Fase I* para garantir desempenho adequado na *Fase II*?”, e
- c) “Como os limites da *Fase II* poderiam ser ajustados para compensar o tamanho da amostra na *Fase I*?”.

Além disso, os autores argumentam que uma importante questão consiste em selecionar o melhor estimador para cada parâmetro e apresentam os principais trabalhos que analisam o desempenho de diferentes estimadores para fornecer recomendações sobre qual usar.

Jensen *et al.* (2006) também discutem a questão das medidas de desempenho utilizadas em duas situações: quando os parâmetros são conhecidos e quando necessitam ser estimados. De maneira geral, quando os parâmetros são conhecidos e durante o período em que o processo permanece estável e ajustado (sob controle), em gráficos de Shewhart (nos quais os valores sucessivos são independentes e identicamente distribuídos), o *RL* (*Run Length*, em inglês), número de amostras até que um sinal ocorra, é uma variável aleatória que segue uma distribuição geométrica, cujo parâmetro  $p$  é a probabilidade de uma estatística cair fora dos limites de controle (que corresponde à probabilidade de alarme falso,  $\alpha$ , quando o processo está sob controle, e a  $1 - \beta$ , o poder do gráfico, quando ele está fora de controle).

Entretanto, quando os parâmetros são estimados, o *RL* não segue uma distribuição geométrica e, portanto, a probabilidade de um sinal não possui a mesma interpretação. Esta afirmação requer um maior esclarecimento, e só vale num certo contexto.

O *RL* do gráfico de controle é dependente dos valores das estimativas dos parâmetros, que, antes dos dados serem coletados na *Fase I*, são variáveis aleatórias. Devido a essa dependência, o *RL* também se torna uma variável aleatória (denote-se esta variável por  $T$ ), que possui uma distribuição diferente para cada valor possível das estimativas dos parâmetros ( $v, w$ ); de modo que cabe

estudar a função densidade de probabilidade condicional do  $RL$ ,  $f_{TW,W}(t|v,w)$ . Infelizmente, esta densidade não depende apenas das estimativas  $(v, w)$ , mas ainda dos próprios valores reais dos parâmetros, que são desconhecidos (ou teríamos o caso de gráficos com parâmetros conhecidos, em que não são necessárias estimativas dos parâmetros e as medidas de desempenho dos gráficos são determinadas com exatidão, pelos modelos tradicionais). Em consequência, no caso de parâmetros estimados, nunca é possível obter valores exatos para as medidas de desempenho de um determinado gráfico de controle. Entretanto, a distribuição condicional do  $RL$  pode ser utilizada para determinar um melhor e um pior cenários para gráficos com parâmetros estimados.

Uma alternativa (para quando não são conhecidos os parâmetros) para a distribuição de probabilidade condicional do  $RL$  é a utilização da distribuição de probabilidade marginal do  $RL$ ,  $f_T(t)$ , que permite avaliar a variabilidade aleatória introduzida no gráfico de controle pela estimação dos parâmetros, sem requerer o conhecimento dos parâmetros. A distribuição marginal do  $RL$  pode ser obtida ponderando a distribuição condicional do  $RL$  por todos os valores possíveis de estimadores dos parâmetros, e pode ser usada antes da coleta dos dados na *Fase I*, para tomar decisões a respeito do tamanho de amostra necessário para se atingir determinado desempenho. Nesse caso, Jensen et al. (*op. cit.*) recomendam complementar a avaliação do gráfico com outras medidas de desempenho, além do *NMA*, tais como o *SDRL* (desvio-padrão do  $RL$ ) e outros percentis. Até porque, evidentemente, a distribuição marginal do  $RL$  é mais dispersa que a distribuição condicional. Como será visto, o uso da distribuição marginal do  $RL$  é a abordagem adotada pela grande maioria dos trabalhos na literatura pesquisada. Não é a abordagem adotada nesta dissertação.

Os autores encerram com a afirmação de que muitas pesquisas têm sido realizadas para se determinar as propriedades dos gráficos de controle quando os parâmetros são estimados, mas que trabalhos adicionais são necessários. Na maioria dos casos, quando os parâmetros são estimados, é necessário coletar mais dados na *Fase I* do que o usualmente recomendado; e o impacto da estimação depende da direção do erro de estimação e do tipo de gráfico em questão.

Hillier (1969) é um dos trabalhos pioneiros mais relevantes. O autor avalia o desempenho dos gráficos de  $\bar{X}$  e  $R$  quando seus limites são calculados com base em um pequeno número inicial de subgrupos, e propõe um método de determinação dos limites de controle para esses gráficos de forma a garantir para eles desempenho idêntico ao dos gráficos com limites baseados em parâmetros conhecidos, independentemente da quantidade de subgrupos utilizada.

O autor refere-se a resultados de trabalhos anteriores seus, que mostram que os limites convencionais “de três sigma” levam (com número de amostras pequeno) a uma probabilidade de alarme falso com valor maior que o nominal (o valor que ela teria se os parâmetros fossem conhecidos). Comenta ainda que essas probabilidades, na *Fase I*, são ligeiramente maiores que na *Fase II*, uma vez que, argumenta ele, um valor inusitadamente alto ou baixo de  $\bar{X}$  ou de  $R$  “puxaria”  $\bar{\bar{X}}$  ou  $\bar{\bar{R}}$  em sua direção, reduzindo ligeiramente a probabilidade de erro do tipo I em relação ao valor dessa probabilidade na *Fase II*.

Motivado por essas considerações, ele analisa as distribuições de  $\bar{\bar{X}}$  e de  $\bar{\bar{R}}$ , parametrizadas pelo número de amostras  $m$  e pelo tamanho de amostras  $n$ , e mostra, a partir delas, quais devem ser os valores “corrigidos” ou “exatos” para os fatores  $A_2$ ,  $D_3$  e  $D_4$ , usados tradicionalmente no cálculo dos limites de controle para os gráficos: constantes apropriadas (que denotou por  $A_2^{**}$ ,  $D_3^{**}$  e  $D_4^{**}$ ) que, se usadas no lugar dos valores convencionais, fornecem as probabilidades de alarme falso especificadas. Tais constantes são função não apenas do tamanho de amostra  $n$ , mas também do número de amostras  $m$ . Ele fornece tabelas de  $A_2^{**}$ ,  $D_3^{**}$  e  $D_4^{**}$  em função de  $m$ , para  $n=5$ , e para diversos valores da probabilidade de alarme falso especificada. Mais exatamente, na abordagem por ele proposta, o usuário dos gráficos deve especificar três probabilidades, correspondentes a três diferentes tipos de alarme falso:  $\alpha_2$ , probabilidade de que  $\bar{X}$  caia fora dos limites de controle do gráfico de  $\bar{X}$ ;  $\alpha_3$ , probabilidade de que  $R$  caia abaixo do limite inferior de controle do gráfico de  $R$ ; e  $\alpha_4$ , probabilidade de que  $R$  caia acima do limite superior de controle do gráfico de  $R$ .

Dada a diferença entre as distribuições da probabilidade de alarme falso na *Fase I* e na *Fase II*, os fatores para uma fase são diferentes dos fatores para a

outra; ele fornece, assim, tabelas de valores para cada fase (denotando os valores para cálculo dos limites para a *Fase II* por  $A_2^*$ ,  $D_3^*$  e  $D_4^*$ ).

Com tais fatores por ele propostos, o procedimento para estabelecimento dos limites na *Fase I* é idêntico ao tradicional: inspeciona-se um pequeno número de amostras ( $m$ ), calculando  $\bar{\bar{X}}$  e  $\bar{R}$ . Calculam-se em seguida os limites de controle provisórios; apenas, utilizando os fatores “corrigidos” ( $A_2^{**}$ ,  $D_3^{**}$  e  $D_4^{**}$ ) em vez dos usuais. Para cada subgrupo (amostra) inicial, verifica-se se a sua média  $\bar{X}$  e amplitude  $R$  estão dentro dos limites dos gráficos. Não sendo o caso, o subgrupo é descartado e os limites são recalculados. Uma vez que todos os subgrupos restantes estejam dentro dos limites, determinam-se os limites de controle para a *Fase II*, com o uso dos fatores  $A_2^*$ ,  $D_3^*$  e  $D_4^*$  (podendo-se especificar valores para  $\alpha_2$ ,  $\alpha_3$  e  $\alpha_4$  diferentes dos especificados para *Fase I*). Após estabelecer tais limites, subgrupos serão inspecionados periodicamente no futuro. Se um novo  $\bar{X}$  e/ou  $R$  cair fora dos limites de controle da *Fase II*, a interpretação será a mesma adotada para os limites de controle convencionais.

Dois considerações são importantes na escolha do número de subgrupos inicial  $m$ : o grau de necessidade por controle imediato e o poder do gráfico de controle em detectar que o processo está fora de controle. Embora os fatores fornecidos garantam que  $\alpha$  se mantenha no nível desejado com qualquer valor de  $m$  (os limites de controle devem ser confiáveis independentemente de quão pequeno seja  $m$ , e esse é o princípio da sua proposta de fatores corrigidos), valores muito pequenos de  $m$  reduzirão sensivelmente o poder dos gráficos, pelo “alargamento” dos limites de controle do gráfico de  $\bar{X}$ . Assim, caso se deseje iniciar cedo o controle do processo, é recomendável começar com um valor pequeno para  $m$ , calcular valores iniciais provisórios para os limites, e ir recalculando-os periodicamente, à medida que  $m$  aumenta. Hillier (*op. cit.*) chama ainda a atenção para a necessidade de expurgar pontos *passados* que caíam fora dos *novos* limites de controle, recalculados, pois eles são sinal de que o processo estava (e talvez ainda esteja) fora de controle; apenas, na ocasião da retirada das amostras a que correspondem tais pontos, não havia ainda evidência estatística suficiente para chegar a essa conclusão.

Considerações análogas cabem quanto à escolha dos valores de  $\alpha_2$ ,  $\alpha_3$  e  $\alpha_4$ . Há uma inclinação natural para especificar valores extremamente reduzidos (pequena probabilidade de erro do tipo I), mas isso resultaria em limites de controle muito “largos”, e eles necessitam ser relativamente “estreitos” para que os gráficos possam detectar mudanças no processo (pequena probabilidade de erro do tipo II). Em suma, ao se especificarem os valores de  $m$ ,  $\alpha_2$ ,  $\alpha_3$  e  $\alpha_4$ , deve-se buscar um equilíbrio, um compromisso, entre a probabilidade de erro do tipo I e a sensibilidade dos gráficos, levando-se ainda em consideração o custo envolvido.

Finalmente, embora os limites possam ser recalculados a cada nova amostra, isso pode não ser prático. Hillier (*op. cit.*) recomenda que eles sejam atualizados periodicamente, com frequência, mas não necessariamente a cada amostra.

Evidentemente, os  $NMA_0$ 's (ou seus inversos, os valores de  $\alpha$ ) obtidos por Hillier são valores médios (valores esperados de valores esperados, ponderados pela distribuição de probabilidades das estimativas de  $\bar{X}$  e  $\bar{R}$ ; ou ainda, os valores esperados das distribuições marginais dessas estimativas).

O trabalho apresentado por Yang & Hillier (1970) corresponde a uma versão do trabalho anterior de Hillier (1969) para os casos de gráficos de  $\bar{X}$  e  $S^2$  e de gráficos de  $\bar{X}$  e  $S$ . O conteúdo, a metodologia, os resultados, conclusões, recomendações, são análogos.

Quesenberry (1993) também estudou o efeito do tamanho de amostra e número de amostras sobre o  $NMA_0$  e o  $NMA_1$  de gráficos de  $\bar{X}$  e de  $X$ . Ele buscou a resposta para a seguinte questão: qual o número mínimo de amostras ( $m$ ) necessário (de um processo estável e normalmente distribuído) para garantir que o desempenho dos limites de controle estimados seja igual ao dos limites de controle “verdadeiros” (os limites que os gráficos teriam se os parâmetros do processo fossem conhecidos sem erro)?

Ele mostrou que a probabilidade de alarme falso,  $\alpha$ , desses gráficos quando  $n$  e  $m$  são pequenos é maior que a nominal (por exemplo, para o gráfico de  $\bar{X}$  com limites de três sigma, com  $m=30$  e  $n=5$ ,  $\alpha=0,00378$ , maior que o valor nominal 0,0027). Diz que, no entanto, os  $NMA_0$ 's são maiores, pois os eventos “sinal” não são independentes, e critica Hillier (1969), por ter desconsiderado esse

fato. Mostra que as diferenças  $(\bar{X}_i - U\hat{C}L)$  e  $(\bar{X}_j - U\hat{C}L)$  — onde  $\bar{X}_i$  e  $\bar{X}_j$  são valores de  $\bar{X}$  na *Fase II*, em diferentes amostras, e  $U\hat{C}L$  é o LSC “estimado” na *Fase I* (considerado, portanto, uma variável aleatória) — são correlacionadas positivamente (sua covariância é igual à variância de  $U\hat{C}L$ ), e portanto os eventos  $\bar{X}_i > U\hat{C}L$  e  $\bar{X}_j > U\hat{C}L$  não são independentes, e a seqüência de comparações de  $\bar{X}_k$  ( $k=1, 2, 3, \dots$ ) com os limites de controle não pode ser vista como uma seqüência de provas de Bernoulli, de modo que a “run length” não segue uma distribuição geométrica com média igual a  $1/\alpha$ .

O autor fornece uma tabela com essas correlações e, para obter os *NMA*'s ( $NMA_0$  e  $NMA_1$ ) na *Fase II*, para  $n=5$ , utilizou simulação de Monte-Carlo. Gerou aleatoriamente  $m$  amostras de tamanho  $n$  (no caso, fez isso apenas para  $n=5$ ), calculou os limites de controle “estimados” com base nesses dados e, em seguida, gerou observações com valores em controle ou fora de controle (média deslocada de  $\delta$  desvios-padrão), até obter um sinal. O número de amostras geradas até o sinal é uma valor (realização) da variável aleatória  $RL$ . O procedimento foi repetido um número suficiente de vezes para obter a média do  $RL$  — i.e., o *NMA* — com erro-padrão pequeno (o número de repetições esteve entre 6.000 e 40.000, dependendo do caso). Com  $\delta=0$  (valores em controle), tem-se o  $NMA_0$ ; com valores de  $\delta$  diferentes de 0, tem-se os valores de  $NMA_1$ .

O efeito da correlação positiva entre os valores de  $(\bar{X}_i - U\hat{C}L)$  e  $(\bar{X}_j - U\hat{C}L)$  (e entre os de  $(\bar{X}_i - L\hat{C}L)$  e  $(\bar{X}_j - L\hat{C}L)$ , onde  $U\hat{C}L$  e  $L\hat{C}L$  são os *LSC* e *LIC* “estimados” na *Fase I*, respectivamente) é um aumento dos *NMA*'s em relação aos valores “nominais” (calculados sob a suposição de independência dos eventos “sinal”), aumento este tanto mais pronunciado quanto menor for o valor de  $m$  (pois maior será a variância dos limites de controle estimados).

Como resultado desse trabalho, Quesenberry recomenda coletar  $m = 100$  amostras de tamanho  $n = 5$  cada uma, para o gráfico de  $\bar{X}$ . As simulações que ele realizou restringiram-se ao tamanho de amostra  $n = 5$ , porém, ele recomendou que, com outros tamanhos de amostra, o número de amostras  $m$  para o cálculo dos limites permanentes seja  $m = 400/(n - 1)$ .

Para o gráfico de  $X$ , o autor recomenda um número mínimo de 300 amostras para estabelecer os limites permanentes.

Durante a *Fase I*, porém, o autor recomenda, sempre que for possível, a substituição da prática tradicional de estabelecimento de limites de controle provisórios, a serem validados e revistos, pelo uso de gráficos  $Q$  (“*Q charts*”), cujos limites de controle garantem para o gráfico o mesmo desempenho que os limites baseados em valores conhecidos dos parâmetros.

Tais *Q-charts* pertencem a uma outra linha de trabalhos, que busca responder à pergunta (c) colocada por Jensen *et al.* (2006), citada mais acima: “Como os limites da *Fase II* poderiam ser ajustados para compensar o tamanho da amostra na *Fase I*?”. Em contraste com a prática usual de calcular os limites de controle por fórmulas que (implicitamente) consideram as estimativas dos parâmetros do processo como valores verdadeiros ao invés de variáveis aleatórias, Quesenberry (1991) analisa as distribuições em controle exatas das diferenças (ou razões, conforme o caso) entre estatísticas do  $(r+1)$ -ésimo subgrupo e os valores das mesmas estatísticas calculadas com base em  $r$  subgrupos precedentes, e estabelece limites de controle para essas diferenças ou razões — na realidade, para funções das mesmas, chamadas “estatísticas  $Q$ ”, com distribuição  $N(0, 1)$ . As estatísticas  $Q$  embutem a informação do número de amostras utilizado, evitando a necessidade de atualizar (recalcular) os limites de controle à medida que novas amostras vão sendo obtidas: os gráficos mantêm uma escala padrão, única, com limites por exemplo em  $+3$  e  $-3$ . A contrapartida é que os cálculos são mais trabalhosos, necessitando ser programados. Os cálculos são recursivos, baseados nos valores anteriores.

As *Q-charts* constituem o primeiro tipo de gráfico de Shewhart corretamente distribuído para um processo estável e normal. Então, como seus limites são baseados na distribuição exata das estatísticas plotadas, podem ser utilizadas com qualquer número de amostras, mesmo desde o início da amostragem do processo, desde que as primeiras amostras são retiradas, antes de que se disponha de estimativas dos parâmetros. Tornam-se, assim, muito úteis para processos com corridas curtas, mas também para o controle de qualquer processo, desde cedo e com probabilidade de alarme falso limitada. Constituem uma abordagem completamente diferente das demais, e que elimina em muitos aspectos a distinção entre a *Fase I* e a *Fase II* dos gráficos de controle.

O trabalho de Nedumaran e Pignatiello (2001) assim como o de Quesenberry (1993), apresentam o problema da construção dos gráficos de controle de Shewhart para a média (determinação dos limites de controle) quando os valores dos parâmetros do processo não são conhecidos. Apesar de o trabalho de Nedumaran e Pignatiello (2001) analisar o desempenho do mesmo tipo de gráfico, trata-se de um estudo sobre a construção dos gráficos de controle diferente daquele proposto por Quesenberry (1993), motivado pelos casos em que não é possível esperar pelo acúmulo do número de amostras recomendado (por exemplo, 100 amostras no caso de  $n=5$ ) para iniciar o controle do processo. Os autores propõem um procedimento baseado em um número inicial ( $m$ ) de subgrupos, com limites que são atualizados a cada  $k$  subgrupos. À medida que mais amostras são coletadas, o número  $k$  pode variar (quer dizer, podem-se atualizar os limites com menor frequência: por exemplo, pode-se iniciar com  $m=5$  ou 10 amostras iniciais, atualizando os limites de 5 em 5 amostras, até um total de 30 amostras, e então passar a atualizá-los a cada 10 novas amostras, até a 50ª amostra, quando se passa a atualizá-los de 25 em 25 amostras, fixando-os definitivamente com 75, ou 100, ou 150 amostras). Os autores levam em consideração a correlação existente entre o evento em que uma amostra ultrapassa os limites de controle ( $B_i$ ) e o evento em que outra amostra ultrapassa os limites de controle ( $B_j$ ). A abordagem proposta consiste em construir limites de controle tais que a probabilidade de que ocorra um alarme falso em algum dos próximos  $k$  subgrupos,  $\Pr[RL \leq k]$ , seja igual à de um gráfico com limites verdadeiros e risco  $\alpha$  especificado, i.e.,  $1 - (1 - \alpha)^k$ . Dada a correlação existente entre os eventos  $B_i$  e  $B_j$ , os limites de controle são estabelecidos com base em valores críticos,  $h'_{\gamma, m, k, \nu}$ , da distribuição normal  $k$ -variada das diferenças  $\bar{X}_i - \bar{X}$ . Eles fornecem tabelas de valores de  $h'_{\gamma, m, k, \nu}$  (onde  $\nu = m[n - 1]$  e  $\gamma = \Pr[RL \leq k] = 1 - (1 - \alpha)^k$ ) de acordo com  $\gamma$ ,  $m$ ,  $k$  e  $n$ .

Para ilustrar o procedimento de obtenção dos limites de controle, os autores apresentam um exemplo. Inicialmente, como não há dados históricos disponíveis, os parâmetros e os limites são obtidos a partir de  $m$  amostras de tamanho  $n$ . Entrando com os valores especificados de  $k$  e  $\alpha$  nas tabelas fornecidas, é possível

encontrar o valor correspondente para  $h'_{\gamma,m,k,v}$  e calcular os limites de controle que serão utilizados no monitoramento dos primeiros próximos  $k$  subgrupos. O processo de determinação dos limites é dinâmico, pois à medida que novos  $k$  subgrupos são coletados do processo, os parâmetros são atualizados (estimados com as  $m+k$  amostras anteriores) e os limites são construídos para os próximos  $k$  subgrupos.

Com exceção de Yang e Hillier (1970) e das *Q-charts* de Quesenberry (1991), os trabalhos relacionados até aqui aplicavam-se a gráficos de  $\bar{X}$  ou de  $\bar{X}$  e  $R$ . Chen (1998) analisa em seu trabalho as propriedades da distribuição do  $RL$  para os gráficos de controle de  $R$ ,  $S$  e  $S^2$ , para o caso em que o desvio-padrão do processo é estimado.

Ele obtém a função de distribuição incondicional (i.e., marginal) do  $RL$  integrando a densidade conjunta dos limites de controle e do  $RL$  ao longo dos valores dos limites de controle. A densidade conjunta, por sua vez, é obtida como o produto da densidade dos limites de controle pela densidade (discreta) condicional do  $RL$  (condicionada ao valor dos limites de controle). O autor obtém, a partir daí, os  $NMA$ 's e desvios-padrão dos  $RL$ 's, ou seja, obtém a média e o desvio-padrão da distribuição marginal dos  $RL$ 's. A análise foi feita para limites de probabilidade que levassem a um  $NMA_0$  marginal igual a 370.

Ele compara esses resultados com a distribuição do  $NMA$  quando o desvio-padrão é conhecido. É suposto que a variável de interesse segue uma distribuição normal e que as estimativas do desvio-padrão do processo utilizadas são  $\hat{\sigma}_{\bar{R}} = \bar{R}/d_2$  para o gráfico de  $R$ ;  $\hat{\sigma}_{\bar{S}} = \bar{S}/c_4$  para o gráfico de  $S$ ; e  $\hat{\sigma}_{S_p} = S_p^2$  (média ponderada dos desvios padrão amostrais) para o gráfico de  $S^2$ .

O autor define  $b$  como a razão entre o desvio-padrão para o processo fora de controle e o desvio-padrão-alvo;  $V$  como o número de amostras até que o primeiro sinal ocorra para o caso em que  $\sigma$  é conhecido e  $W$ , como o número de amostras até que o primeiro sinal ocorra para o caso em que  $\sigma$  é estimado. Em seu estudo, o autor compara a  $P(W \leq w)$  com  $P(V \leq v)$  para diferentes valores de  $m$ ,  $n$  e  $b$  e encontra resultados muito semelhantes para os gráficos de  $R$ ,  $S$  e  $S^2$ .

Para os três gráficos, quando  $b = 1$  (processo sob controle), o  $NMA$  é menor quando o desvio-padrão é estimado, ou seja, o processo tende a gerar, em média,

maior probabilidade de sinal ( $\alpha$ ) que na situação em que ele é conhecido. Quando a variabilidade do processo sofre uma alteração ( $b \neq 1$ ), o *NMA* é maior quando o desvio-padrão é estimado, sugerindo que mudanças na variabilidade não são detectadas rapidamente. O autor conclui, também, que o efeito da estimação de  $\sigma$  é maior quando a variabilidade aumenta ( $b > 1$ ) e que, se a variabilidade diminui ( $b < 1$ ), em média, o efeito da estimação parece ter um efeito muito pequeno na probabilidade de sinal. O autor também acrescenta que o efeito da estimação de  $\sigma$  depende de  $m$  e  $n$ , que em geral, diminui à medida que  $m$  e  $n$  aumentam: com  $4 \leq n \leq 10$ , o efeito é quase inexistente para  $m \geq 75$ .

Chen (1998) recomenda que, para que se tenha um bom desempenho em detectar mudanças no desvio-padrão do processo, com tamanhos de amostras entre 4 e 10 ( $4 \leq n \leq 10$ ), 75 amostras ( $m = 75$ ) devem ser coletadas na *Fase I*.

Maravelakis et al. (2002) estudaram os efeitos da estimação dos parâmetros sobre o desempenho dos gráficos de  $S$  com limites de três sigma e com limites de probabilidade, bem como sobre o desempenho do gráfico de observações individuais no monitoramento da dispersão do processo, estendendo os resultados de Chen (1998) para observações individuais e subgrupos racionais.

Eles obtiveram as médias (*NMA's*) e os desvios-padrão (*SDRL's*) da distribuição marginal da  $RL$ , por simulação, usando o mesmo procedimento que Quesenberry (1993), com 10.000 repetições.

De maneira geral, os principais resultados para o gráfico de  $S$  foram: os valores de *NMA* e *SDRL* variam na mesma direção, aumentando no caso de processo em controle e diminuindo no caso de processo fora de controle, com a ressalva que, quando  $m$  aumenta, o *NMA* se aproxima mais rapidamente do valor nominal que o *SDRL*. Quando o processo está sob controle são necessárias pelo menos  $m = 200$  amostras para garantir o bom desempenho do gráfico de controle; essa recomendação é praticamente independente do valor de  $n$ , ao qual o desempenho do gráfico é muito mais sensível do que a  $m$ . Contudo, com vistas ao desempenho do gráfico na situação do processo fora de controle, i.e., ao seu poder, especificar o valor de  $n$  se torna importante (há uma redução considerável nos valores do *NMA* à medida que  $n$  se torna maior). Recomenda-se utilizar  $n \geq 20$ . O efeito da estimação é severo para  $m \leq 20$ , especialmente quanto à detecção

de pequenas mudanças na dispersão do processo, é moderado para  $30 \leq m \leq 50$ , e satisfatoriamente pouco acentuado para  $m \geq 100$ .

Para o gráfico de observações individuais (gráfico de  $X$ ), para monitorar a dispersão do processo, os autores recomendam o valor mínimo de 300 observações para minimizar o efeito da estimação dos limites de controle do gráfico de  $X$ .

Em contraste com as recomendações dos autores precedentes, Chirico (1995) critica em seu trabalho as regras do tipo “receita de bolo” para o controle estatístico do processo (CEP) e a adoção dessas regras pelas empresas. O autor argumenta que a aplicação dessas regras pode não ser coerente com o processo em questão e que a utilização das mesmas inibe soluções criativas de engenharia para o problema da variabilidade dos processos. Para a maioria das empresas, tais regras seriam mais prejudiciais do que benéficas, pois são baseadas no estudo de processos ideais, na maioria das vezes com alto volume de produção e com um único canal.

Em seu trabalho, o autor estabelece dois obstáculos no uso dos gráficos de controle em processos multi-canal ou de baixo volume: a determinação dos limites de controle e o desenvolvimento de estratégias de amostragem ótima. Ao mesmo tempo, o autor tenta desmistificar algumas questões: (1) não é necessário ter vinte e cinco subgrupos de dados para se calcular os limites de controle; (2) os limites são determinados por razões econômicas e não estatísticas; (3) as estratégias de formação de subgrupos são mais importantes que os limites de controle.

Para comprovar sua tese, o autor toma como base os trabalhos realizados por Shewhart e desenvolve um raciocínio que o leva a validar conceitualmente suas afirmativas. No estágio inicial de um gráfico de controle, os limites são definidos por alguma aproximação até que causas especiais sejam removidas do processo. Nessa fase, o autor afirma que é melhor estabelecer os limites com base no conhecimento do comportamento do processo do que com aplicação de regras dos gráficos de controle, permitindo distinguir mais efetivamente os tipos de variações (aleatórias ou especiais) nos estágios iniciais da análise. Adicionalmente, o autor considera que esses limites iniciais devem ser revisados periodicamente à medida que os dados são coletados, as causas especiais, removidas, e se adquire maior experiência. Outro questionamento realizado pelo

autor diz respeito à regra de limites de controle de “três-sigma”. O autor afirma que esta decisão constitui uma escolha econômica, pois permite balancear os benefícios e os custos decorrentes de se detectar causas especiais. Finalmente, o autor realiza recomendações sobre a estratégia de amostragem, advertindo que não é necessário utilizar uma quantidade de dados maior do que a necessária para se obter a sensibilidade e independência desejada, e que essa decisão deve considerar a experiência dos envolvidos no processo, e não somente seguir as “receitas” para gráficos de controle. Ele comenta, ainda, que o argumento de que com poucas amostras iniciais ainda não se tem precisão para estimação dos limites de controle tem o foco errado, e que, na *Fase Inicial*, em que se está conhecendo melhor e estabilizando o processo, manter a probabilidade de alarme falso em níveis muito baixos não é tão importante quanto a sensibilidade para detectar causas especiais, e que é muito mais indesejável perder (retardar) a oportunidade de detectá-las, seja porque a probabilidade de erro do tipo II é alta, seja porque ainda não se começou o controle *on line* do processo, enquanto se aguarda que o número supostamente suficiente de amostras iniciais seja atingido.

Os artigos aqui descritos constituem os principais trabalhos sobre a questão do efeito da estimação dos parâmetros do processo (e, portanto, do número e tamanho das amostras iniciais) sobre o desempenho dos gráficos de Shewhart para variáveis. Há muitos outros trabalhos, aplicados a outros tipos de gráficos, sem falar em CEP de processos autocorrelacionados ou CEP multivariado. A ênfase aqui foi nos gráficos de Shewhart para variáveis, por ser este o assunto desta dissertação. Ao leitor interessado em outros gráficos e no tema em geral, fica indicada a revisão, bastante abrangente, de Jensen *et al.* (2006), que, além disso, fornece uma extensa lista de referências.

O que se pode verificar através da revisão feita nesta Seção é que praticamente todos os estudos realizados foram voltados para (ou baseados na) distribuição *marginal* da “*run length*”, obtendo um risco  $\alpha$  médio ou (o que é equivalente) um *NMA* esperado (notar que o *NMA* já é um valor esperado; portanto trata-se do valor esperado de um valor esperado condicional).

Na nossa visão, há um problema com essa abordagem — ou pelo menos uma limitação. Os *NMA*'s assim obtidos são médias de “*run lengths*” de diferentes gráficos (i.e., de gráficos com diferentes valores para os limites). São médias

globais ou, como acabamos de dizer, valores esperados de valores esperados condicionais, i.e., valores esperados da distribuição marginal de “run lengths”. Porém, cada gráfico particular terá valores para os limites que, embora *a priori* sejam variáveis aleatórias, uma vez calculados tornam-se realizações, valores observados, dessas variáveis aleatórias e, portanto, constantes. Dado *um* gráfico com os seus limites, os eventos “sinal” passam a ter probabilidade constante e são independentes. A correlação que Quesenberry destaca entre  $(\bar{X}_i - U\hat{C}L)$  e  $(\bar{X}_j - U\hat{C}L)$  só existe quando se considera  $U\hat{C}L$  como variável aleatória, de maneira que o evento “sinal” passa a ser definido em função de duas variáveis aleatórias: não apenas do valor de  $\bar{X}$  da amostra, mas também da variável aleatória  $U\hat{C}L$ . Portanto, a dependência entre eventos “sinal” só existe no contexto da população de limites (ou de gráficos) possíveis de serem gerados; nesse contexto, ela pode ser entendida da seguinte forma: sem nenhum conhecimento do valor de  $U\hat{C}L$ , gráficos nos quais tenha ocorrido um sinal (ou um maior número de sinais) têm maior probabilidade de terem um menor valor de  $U\hat{C}L$  (e um maior valor de  $L\hat{C}L$  — em suma, limites mais estreitos) que gráficos em que não tenha ocorrido um sinal (ou em que tenha ocorrido um menor número de sinais); portanto, têm também maior probabilidade de registrarem novos sinais do que os últimos. Ora, esta visão considera o *universo de gráficos* (ou de valores possíveis para os limites de um gráfico, antes que eles se tornem constantes), e assim considera os sinais de *todos* os gráficos possíveis. Somente em tal universo é que os sinais de um mesmo gráfico não são independentes; em um único gráfico, eles *são* independentes. Em um único gráfico, a “run length” não segue a distribuição marginal, com uma média que é a média das médias condicionais, e uma variância que embute a própria variabilidade dessas médias; em um único gráfico, a probabilidade de alarme falso assume um único valor, existe um único poder para cada alteração no parâmetro do processo sendo monitorado (média, desvio-padrão), existe *um*  $NMA_0$ , e assim por diante. O  $NMA_0$  real será a média da distribuição do *RL condicionada ao valor dos limites de controle do gráfico*. Por isso, na nossa opinião, tem pouco sentido na prática considerar as médias das médias condicionais como valores únicos. Basear a decisão sobre  $m$  e  $n$  no valor médio da distribuição do  $NMA_0$  é arriscado, pois o  $NMA_0$  real do gráfico poderá

ser muito menor. Acreditamos que é mais útil, e faz mais sentido, para o praticante obter, para cada par  $(m, n)$ , a distribuição de probabilidade de  $\alpha$ , de  $NMA_0$  e de  $NMA_1$  ou o poder do gráfico. Para cada gráfico que se estabeleça, tais variáveis assumirão um único valor, e tal valor é uma variável aleatória cuja distribuição será assim conhecida. Poder-se-á, assim, conhecer o risco de  $\alpha$ , de  $NMA_0$  e de  $NMA_1$  ou do poder do gráfico atingirem valores considerados inaceitáveis, ou saber qual o número  $m$  de amostras iniciais necessário (dado o tamanho de amostra,  $n$ ) para que esse risco não ultrapasse um limite especificado. Esta é a motivação do presente trabalho, que tem como objetivo principal determinar a distribuição do risco  $\alpha$  e do poder para o gráfico de  $S$ , bem como o número necessário de amostras iniciais para limitar esse risco.

Quesenberry (1993) deixa bem claro este fato de que o  $NMA_0$  e o  $NMA_1$  de qualquer gráfico em particular não podem ser conhecidos com precisão, e que os resultados por ele apresentados se referem aos  $NMA_0$ 's e o  $NMA_1$ 's “médios”, “gerais”, de um programa de CEP de uma empresa, e que o  $NMA_0$  e o  $NMA_1$  de um gráfico qualquer em particular pode ser maior ou menor que os valores por ele apresentados. Outros autores não deixam este fato tão explícito, o que pode ser desorientador para um leitor menos atento.