



UM MÉTODO SIMPLES PARA O PROJETO ÓTIMO DE GRÁFICOS DE \bar{X}

Eugenio K. Epprecht

Alessandra Bastos dos Santos

Departamento de Engenharia Industrial – PUC-Rio

Rua Marquês de São Vicente, 225 – Gávea

22453-900 – Rio de Janeiro – RJ

email: eke@rdc.puc-rio.br

Resumo

É desenvolvido um método simples e eficiente para a seleção de parâmetros (tamanho de amostra, intervalo de tempo entre amostras, e largura dos limites de controle) para gráficos de \bar{X} para controle estatístico de processos, que fornece soluções ótimas em termos da relação entre a rapidez de detecção e os recursos alocados à inspeção. O método é superior a procedimentos baseados em propriedades estatísticas do gráfico (tais como o ARL) e à adoção de valores empíricos gerais, e não apresenta as dificuldades de implementação dos métodos econômicos (economic design): não requer dados difíceis de obter ou estimar nem programas de computador para solução. Os resultados obtidos diferem de alguns resultados bem difundidos na literatura; mostra-se que os últimos são subótimos, por serem baseados em modelos que desconsideram aspectos importantes do problema. Uma contribuição adicional do método é o tratamento explícito e detalhado de restrições opcionais ao tamanho de amostra e/ou ao intervalo de tempo entre amostras. A simplicidade do método o torna viável para implementação em qualquer ambiente, e também apropriado para a sala de aula.

Palavras-chave: gráfico de \bar{X} , gráficos de Shewhart, controle estatístico de processo, controle estatístico de qualidade, projeto de gráficos de controle.

1. Introdução

A implementação de gráficos de Shewhart para controle estatístico de qualidade de um processo traz inevitavelmente consigo a questão do projeto (*design*) do gráfico, i.e., da seleção dos valores para os parâmetros: n (tamanho de amostra), h (intervalo de tempo entre amostras consecutivas), e L (distância dos limites de controle à linha central do gráfico, em número de desvios padrão da estatística considerada: \bar{X} , R , S , p , c , ou outra).

Como em todo teste estatístico, o problema é de um compromisso entre: a) o poder do gráfico em detectar descontroles no processo (mudanças no valor médio da característica de qualidade de interesse), b) a frequência de ocorrência de alarmes falsos (sinalização de um descontrole pelo gráfico quando na realidade o processo está em controle), e c) o gasto com inspeção.

Quanto menor L , maior o poder do gráfico (probabilidade de detectar um descontrole), mas maior também a probabilidade de alarme falso; para melhorar o poder do gráfico sem aumentar a frequência de ocorrência de alarmes falsos, é preciso aumentar n — o que, porém, aumenta o gasto com inspeção. Quanto ao valor de h , este não afeta nem o poder do gráfico nem a probabilidade de alarmes falsos, mas altera os *tempos* até a detecção de um descontrole, e até a ocorrência de um alarme falso. Evidentemente, há também restrições de ordem prática e custos associados ao valor de h . Estes múltiplos objetivos e restrições, de minimizar (ou limitar) o tempo até a detecção, a incidência de alarmes falsos e os gastos com inspeção, tornam o projeto de gráficos de controle um problema não trivial de otimização.

As abordagens adotadas para solução deste problema podem ser classificadas em 3 grandes categorias:

- **empíricas:** adotando para os parâmetros valores que, em grande parte dos casos, fornecem bons resultados. Escolhas tradicionais para L e n são $L=3$ e $n = 4$ ou 5 (MONTGOMERY, 1991, 1997). Quanto ao

intervalo de tempo entre duas amostras, não existe um valor estabelecido: “*The heuristic usually attributable to Shewhart (...) samples once every hour for high-volume processes*” (SANIGA, 1989); “*samples (...) taken fairly frequently*” (DUNCAN, 1986, *apud* SANIGA, 1989); “*frequent enough to detect the effect of special causes while the special cause itself can still be identified. (...) as a rule of thumb you should average about 1 out of control point per (...) 25 groups*” (PYZDEK, 1990). Os inconvenientes desta abordagem são óbvios: valores gerais não consideram a especificidade de cada situação; nem sempre proporcionam um bom desempenho, e só por acaso otimizam o uso dos recursos.

- **estatísticas:** n e L são selecionados segundo medidas de desempenho estatístico do gráfico, tais como a probabilidade de alarme falso, a probabilidade de não-detecção, ou o *ARL* (*average run length*) — valor esperado do número de amostras extraídas desde a ocorrência de um descontrole até a sua sinalização pelo gráfico. As principais limitações desta abordagem são: o fato de ela não produzir indicações para a escolha do valor de h ; e a conseqüente subotimização dos resultados obtidos, devido à não consideração da possibilidade de se “jogar” com o valor de h para, por exemplo, acelerar a detecção de um descontrole sem aumentar o tamanho n da amostra. Em outras palavras, a abordagem estatística não considera todas as alternativas, por não considerar todas as dimensões do problema.
- **econômicas:** este tipo de abordagem procura levar em conta todos os elementos do problema que resultem em alguma forma de custo, incluindo vários elementos que não são considerados explicitamente nas outras abordagens (tais como o custo de investigar a(s) causa(s) de um descontrole, o custo de um alarme falso, e a distribuição de probabilidades do tempo que o processo permanece em controle), e combiná-los em uma única

função-objetivo, a ser minimizada: o custo total por unidade de tempo (que seria, em última instância, o critério do cliente). Em tese, portanto, esta abordagem é superior: a mais completa e rigorosa. Na prática, porém, raramente é utilizada, devido não só à complexidade dos modelos resultantes e à necessidade de programas de computador para sua solução, mas, principalmente, devido à dificuldade de se obterem todos os dados necessários (como quantificar, por exemplo, o custo de um alarme falso?). Além disso, os modelos, ainda que matematicamente sofisticados, simplificam certos aspectos do problema real, produzindo por vezes soluções “ótimas” inaceitáveis na prática (ver discussões em: WOODALL, 1986, que inclusive fornece um exemplo cuja solução ótima produziria um alarme falso a cada 3 amostras; e em KEATS *et al.*, 1997). SANIGA (1989), entre outros, propõe soluções para este tipo de problema, mas não para as demais dificuldades de implementação destes modelos.

Há ainda abordagens híbridas ou intermediárias, como veremos na Seção 7.

Este artigo propõe uma abordagem que procura reduzir os inconvenientes das acima citadas, e que define objetivos e restrições em termos dos seguintes critérios: a) o tempo médio até a detecção de um descontrole, b) o número médio de amostras até um alarme falso e c) a intensidade de inspeção (número de medidas por unidade de tempo) alocada ao esquema de controle.

Acredita-se que seja fácil para o usuário dos gráficos exprimir seus objetivos e restrições por meio destes critérios. O modelo leva (como será mostrado) a procedimentos simples e diretos para o projeto do gráfico. Assim, a abordagem proposta:

- é obviamente superior à adoção cega de valores empíricos gerais;
- é mais vantajosa que a utilização de critérios meramente estatísticos (baseados por exemplo no *ARL*), pois, ao contrário destes, contempla o intervalo de tempo entre amostras e

as alternativas de combinação de valores de todos os parâmetros; e

- elimina os problemas de implementação dos métodos econômicos, ao mesmo tempo que considera, implicitamente, os elementos de custo usualmente tidos como os mais relevantes: o custo de inspeção (ao utilizar a intensidade de inspeção como critério) e o custo de produzir em condições degradadas (por intermédio do critério do tempo até a detecção). E o terceiro critério levado em conta pelo modelo — o número médio de amostras até um alarme falso — provavelmente é mais adequado para o usuário do que o *custo* de um alarme falso, que, como visto acima, pode levar a soluções inaceitáveis em termos de frequência de alarmes falsos.

O presente artigo se concentra no projeto de gráficos de \bar{X} , isoladamente. No entanto, a análise aqui apresentada pode se aplicar a outros gráficos, ou conjuntos de gráficos. A Seção 8 menciona extensões pretendidas nesta direção.

2. Modelo Matemático

Suponha-se um processo de produção discreto, com uma característica de qualidade numérica e contínua X . A intervalos regularmente espaçados de h unidades de tempo, é extraída uma amostra de n itens, e as sucessivas médias amostrais \bar{X}_i são plotadas em um gráfico de Shewhart para médias (“gráfico de \bar{X} ”). Supõe-se ainda que as medidas de X em itens consecutivos sejam i.i.d., com média μ_0 e desvio padrão σ_0 . Suponha que o processo está sujeito a deslocamentos bruscos (*shifts*) em sua média. Um ponto \bar{X}_i fora do intervalo definido pelos limites $\mu_0 + L\sigma_0/\sqrt{n}$ e $\mu_0 - L\sigma_0/\sqrt{n}$ é interpretado como sinal de que o processo está fora de controle. Chamando de μ_1 o valor da média após o *shift*, defina-se $d = |\mu_1 - \mu_0|/\sigma_0$. Obviamente o poder do gráfico (probabilidade de um descontrole ser sinalizado em uma amostra) é função do valor de d . Assim, para o projeto do gráfico, requer-se que o usuário seja capaz de indicar qual o valor

d de interesse: a magnitude do afastamento da média em relação ao valor-alvo que se torna significativo para a qualidade do processo. Este valor d pode ser definido com base em considerações de capacidade de processo. Por exemplo, d pode ser definido como o afastamento da média que resulta em uma proporção P_{int} , considerada intolerável, de itens não conformes produzidos. Se a característica X possuir distribuição normal, então d pode ser calculado por

$$d = z_{esp} \cdot \sigma_0 - \Phi^{-1}(1 - P_{int}) \quad (1)$$

onde $z_{esp} \cdot \sigma_0$ é a distância do(s) limite(s) de especificação ao valor-alvo, e Φ é a função de distribuição normal padrão acumulada.

Supõe-se que há um limite máximo de recursos de inspeção que se está disposto a investir, na forma de um número máximo de itens a serem inspecionados por dia (KEATS *et al.*, 1995, citam como um exemplo de razão para este limite a baixa disponibilidade de inspetores ou aparelhos de medida, que são compartilhados por outros processos).

Definam-se as seguintes variáveis:

$AATS$: tempo médio entre a ocorrência de um descontrole e a sua sinalização pelo gráfico (*detecção* do descontrole),
 r : intensidade de inspeção (número de itens inspecionados por unidade de tempo), dada por

$$r = n/h \quad (2)$$

ARL_0 : número médio de amostras observadas até a ocorrência de um alarme falso.

O problema de projeto ótimo do gráfico pode então ser formulado como:

$$\min_{d, L, n, h} AATS(d, L, n, h) \quad (3)$$

sujeito a

$$r \leq r_{max} \quad (4)$$

$$ARL_0 \geq ARL_{0min} \quad (5)$$

Os dados são: d , r_{max} e ARL_{0min} ; as variáveis de decisão são L , n e h .

Para não sobrecarregar a notação, daqui em diante os argumentos d , L , n e h só serão

indicados quando se quiser chamar a atenção para as dependências funcionais.

Assume-se ainda que: o tempo para tirar uma amostra é muito inferior a h , e portanto, desprezível; o número de unidades de cada amostra, n , é muito menor que o número de unidades produzidas no intervalo de tempo h ; e a probabilidade de ocorrer um descontrole durante a retirada da amostra é desprezível.

Duas condições adicionais são requeridas para a relevância prática da solução:

$$AATS \ll MTIC \quad (6)$$

onde $MTIC$, que depende do processo, é o tempo médio até a ocorrência de um descontrole (tempo médio entre 2 descontroles consecutivos); e

$$AATS \ll ARL_0 \cdot h \quad (7)$$

Esta segunda condição equivale a requerer que o poder do gráfico — probabilidade de detectar um descontrole — seja muito maior que a probabilidade de alarme falso.

Estas duas condições, por dependerem de $AATS$, só podem ser verificadas após a solução, mas na prática é razoável assumir que a solução ótima do problema as satisfaz. Com efeito, é fácil verificar, examinando as soluções obtidas, na Seção 4, que elas satisfazem a (7). Quanto à condição (6), como ela depende de $MTIC$, dependerá de cada processo, mas não só na prática é razoável — como já foi dito — esperar que seja satisfeita, como também veremos adiante uma formulação alternativa (“dual”) do problema, em que esta condição pode ser forçada.

Assume-se ainda que \bar{X} tem distribuição de probabilidades normal (portanto, se X não tiver distribuição normal, os resultados obtidos conterão alguma imprecisão para pequenas amostras).

Definam-se as seguintes variáveis aleatórias (ver Figura 1): S , o tempo decorrido entre a extração da última amostra anterior ao descontrole e o descontrole ($0 \leq S \leq h$); RL , o número de amostras extraídas entre a ocorrência de um descontrole e o sinal emitido pelo gráfico; e TS , o tempo entre o descontrole e o sinal. Então,

$$TS = (RL)h - S \quad (8)$$

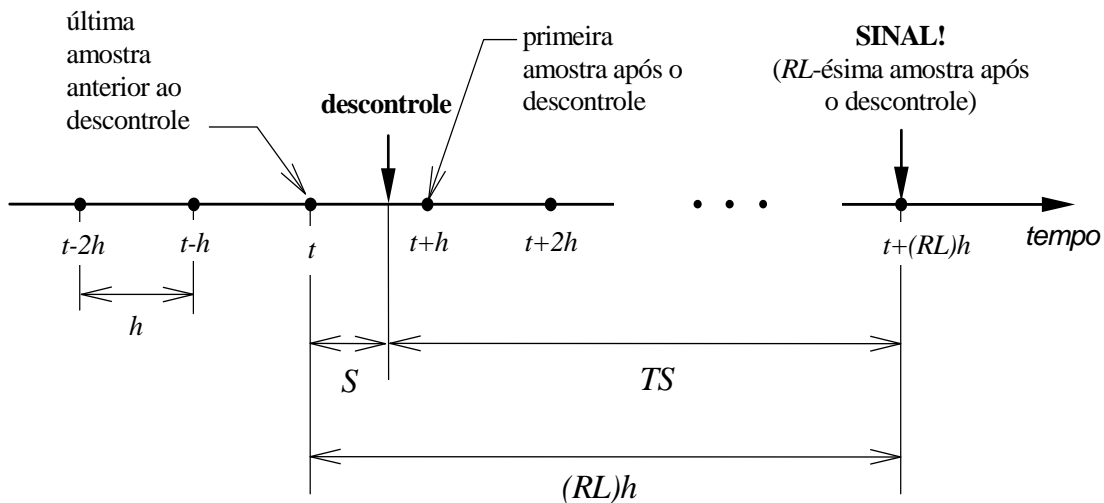


Figura 1 – Eventos no controle de processo e ilustração da fórmula (8)

O tempo médio entre o descontrolamento e o sinal, AATS, é o valor esperado de TS :

$$AATS = E(TS) = hE(RL) - E(S) \quad (9)$$

É usual aproximar a média de S por:

$$E(S) \cong h/2 \quad (10)$$

O argumento é que, se $h \ll MTIC$ (o que decorre de (6)), então o valor da densidade de probabilidade do tempo τ até o descontrolamento varia pouco entre t e $t+h$, para qualquer t . Em consequência, a densidade de probabilidade de τ condicionada a que o descontrolamento tenha ocorrido entre t e $t+h$ é aproximadamente uniforme; neste caso, $E(\tau)$ é aproximadamente igual a $t+h/2$. Fazendo t igual ao instante da última amostra anterior ao descontrolamento, por definição $S = \tau - t$, o que implica em (10).

Quanto a $E(RL)$, chamado comumente de ARL , é de conhecimento geral que, para o gráfico de \bar{X} , ele é dado por:

$$ARL(d, L, n) = \frac{1}{1 - \Phi(L - d\sqrt{n}) + \Phi(-L - d\sqrt{n})} \quad (11)$$

Substituindo então (10) em (9), e $E(RL)$ por ARL , chega-se a:

$$AATS(d, L, n, h) = [ARL(d, L, n) - 1/2].h \quad (12)$$

Para reduzir as dimensões do problema, um exame de (11) mostra que, para quaisquer valo-

res fixos de d e n , quanto menor L , menor o valor de $ARL(d, n, L)$ e, portanto, da função-objetivo AATS em (12). Por outro lado, a restrição (5), $ARL_0 \geq ARL_{0min}$, determina um limite inferior para o valor de L . Assim, a solução ótima possui L tão pequeno quanto admissível por esta restrição, e isto ocorre quando

$$ARL_0 = ARL_{0min} \quad (13)$$

De (11), fazendo $d=0$ e $ARL(0, L, n) = ARL_{0min}$, chega-se facilmente a

$$L = -\Phi^{-1}[1/(2 ARL_{0min})] \quad (14)$$

Assim, a restrição (5) do problema, $ARL_0 \geq ARL_{0min}$, fixa o valor de L .

Com L e d prefixados, o valor de AATS em (12) se torna função apenas de n e h . Para deixar (12) explicitamente em termos da intensidade de inspeção $r=n/h$ (que é um dos elementos do problema), pode-se multiplicar ambos os membros por r , obtendo, após substituir, no membro direito, rh por n :

$$AATS_{d,L}(n, r).r = [ARL(d, L, n) - 1/2].n \quad (15)$$

Esta equação (15) mostra três coisas importantes:

1) Supondo d e L fixos, para qualquer n , AATS é inversamente proporcional à intensidade de inspeção r e vice-versa.

2) Portanto, qualquer que seja o valor ótimo para n , n_{ot} , a solução ótima tem $r = r_{max}$ e (da definição de r)

$$h_{ot} = n_{ot}/r_{max} \quad (16)$$

3) Para r fixo (e igual a r_{max}), o mínimo de $AATS$ é obtido minimizando a função

$$g_{d,L}(n) = [ARL(d,L,n) - 1/2]n \quad (17)$$

que é o segundo membro de (15). O problema reduz-se então a um problema de minimização de uma função de uma variável (n), parametrizada em d e L .

Antes de prosseguir, observemos que o problema poderia ter recebido outra formulação, dual à descrita por (3), (4) e (5). A saber,

$$\min r = \frac{n}{h} \quad (18)$$

sujeito a

$$AATS(d, L, n, h) \leq AATS_{max} \quad (19)$$

$$ARL_0 \geq ARL_{0min} \quad (20) \text{ (5 repetida)}$$

Um exame de (15) mostra que este problema se reduz ao mesmo problema central que o problema “primal” representado por (3), (4) e (5): L é determinado da mesma maneira, por (14), e n é determinado pela minimização de $g(n)$, que levará, para um mesmo d , a um mesmo valor de n_{ot} que o problema “primal”. Apenas h , no problema “primal”, será determinado por (16), enquanto que, no “dual”, será determinado por

$$h_{ot} = n_{ot}AATS_{max} / g(n_{ot}) \quad (21)$$

(pois para minimizar $r=g(n)/AATS$ é preciso fazer $AATS=AATS_{max}$), ou por

$$h_{ot} = AATS_{max} / (ARL(n_{ot}) - 0,5) \quad (22)$$

(obtida de (12), fazendo $AATS = AATS_{max}$).

A função $g(n)$, portanto, é uma medida da razão custo/benefício de cada projeto (custo: de inspeção, medido em termos de $r=n/h$; benefício: em termos da “rapidez” de detecção, medida pelo inverso de $AATS$). Quanto menor $g(n)$, mais eficiente o projeto: menor o $AATS$ que se obtém

com uma dada intensidade de inspeção fixa, e menor a intensidade de inspeção necessária para obter um $AATS$ pré-especificado.

O modelo foi desenvolvido sob a hipótese de que o processo se inicia em controle, mas será válido também para o caso em que ele se inicie fora de controle, desde que se adote como instante da primeira amostra $t_1 = h/2$ — o que, aliás, é recomendável como proteção adicional contra possíveis problemas no início do processo. Isto deverá ser feito a cada reinício do processo e do esquema de controle.

3. Solução do Modelo

Fixados d e L , o fato de que n é inteiro e que razões de ordem prática limitam seu valor, eliminam a necessidade de qualquer método dinâmico de busca do ótimo: é mais simples e imediato calcular exaustivamente o valor da função $g(n)$ para $n=1$ até $n=30$, 50 ou 100, e simplesmente selecionar o menor valor obtido.

A Tabela 1 apresenta os valores de n ótimo em função de d para os limites de controle no valor consagrado de $L=3$. Para outros valores de L , tabelas e gráficos semelhantes podem ser levantados pelo procedimento acima descrito. Contudo, a escolha de $L=3$ não representa um mero empirismo, mas se justifica por:

- Imprecisão inerente na determinação da relação entre ARL_0 e L : quando não se conhece com certeza a distribuição de probabilidades da característica de qualidade X , mesmo que o Teorema do Limite Central garanta normalidade *aproximada* de \bar{X} , a região externa aos limites de controle corresponde à cauda da distribuição, onde, em razão da ordem de grandeza dos valores da densidade de probabilidade, os erros de aproximação se tornam proporcionalmente mais significativos. Neste caso, torna-se um preciosismo inútil tentar calcular por (14) um valor preciso para L com base num ARL_{0min} ideal.
- Imprecisão inerente na definição de ARL_{0min} pelo cliente: mesmo quando se pode estabelecer com precisão a relação entre L e ARL_0 ,

Tabela 1 – n ótimo, ARL e g , versus d , para $L=3$

| d | n_{ot} | $ARL(n_{ot})$ | $g(n_{ot})$ |
|-----|----------|---------------|-------------|
| 0,4 | 100 | 1,19 | 68,86 |
| 0,5 | 64 | 1,19 | 44,07 |
| 0,6 | 44 | 1,20 | 30,60 |
| 0,7 | 33 | 1,18 | 22,49 |
| 0,8 | 25 | 1,19 | 17,21 |
| 0,9 | 20 | 1,18 | 13,60 |
| 1,0 | 16 | 1,19 | 11,02 |
| 1,1 | 13 | 1,20 | 9,11 |
| 1,2 | 11 | 1,20 | 7,65 |
| 1,3 | 9 | 1,23 | 6,53 |
| 1,4 | 8 | 1,20 | 5,62 |
| 1,5 | 7 | 1,20 | 4,90 |
| 1,6 | 6 | 1,22 | 4,31 |
| 1,7 | 6 | 1,14 | 3,84 |
| 1,8 | 5 | 1,18 | 3,40 |
| 1,9 | 4 | 1,27 | 3,08 |
| 2,0 | 4 | 1,19 | 2,75 |
| 2,1 | 4 | 1,13 | 2,52 |
| 2,2 | 3 | 1,26 | 2,29 |

| d | n_{ot} | $ARL(n_{ot})$ | $g(n_{ot})$ |
|-----|----------|---------------|-------------|
| 2,3 | 3 | 1,19 | 2,08 |
| 2,4 | 3 | 1,14 | 1,92 |
| 2,5 | 3 | 1,10 | 1,80 |
| 2,6 | 2 | 1,33 | 1,66 |
| 2,7 | 2 | 1,26 | 1,52 |
| 2,8 | 2 | 1,20 | 1,41 |
| 2,9 | 2 | 1,16 | 1,31 |
| 3,0 | 2 | 1,12 | 1,24 |
| 3,1 | 2 | 1,09 | 1,18 |
| 3,2 | 2 | 1,07 | 1,14 |
| 3,3 | 2 | 1,05 | 1,10 |
| 3,4 | 1 | 1,53 | 1,03 |
| 3,5 | 1 | 1,45 | 0,95 |
| 3,6 | 1 | 1,38 | 0,88 |
| 3,7 | 1 | 1,32 | 0,82 |
| 3,8 | 1 | 1,27 | 0,77 |
| 3,9 | 1 | 1,23 | 0,73 |
| 4,0 | 1 | 1,19 | 0,69 |
| 5,0 | 1 | 1,02 | 0,52 |

é possível que a noção do limite tolerável para ARL_0 não seja precisa (principalmente se for difícil quantificar os custos de um alarme falso).

- Na prática, porém, qualquer que seja a distribuição de \bar{X} , a probabilidade de obter um ponto fora dos “limites de 3σ ” ($L=3$) com $d=0$ é suficientemente baixa para isto ser interpretado como um sinal de descontrole do processo, e para garantir um ARL_0 aceitável. Esta questão é bem discutida em WHEELER (1995).

Com a adoção de $L=3$, e definido o d relevante, o valor ótimo para n , n_{ot} pode ser lido diretamente da Tabela 1. Determinado n_{ot} , pode-se calcular o valor ótimo de h , h_{ot} : na forma “primal” do problema ($\min AATS$ s.a. $r \leq r_{max}$), h_{ot} é dado por (16); na forma “dual” do problema ($\min r$ s.a. $AATS \leq AATS_{max}$), h_{ot} pode ser calculado tanto por (22) como por (21).

Para eliminar a necessidade de calcular o valor de $ARL(n)$ para entrada em (22), ou de $g(n)$, para entrada em (21), estes valores são também fornecidos na Tabela 1, para cada par $(d, n_{ot}(d))$.

Para $d \leq 0,7$, os valores ótimos para n podem ser proibitivamente altos. Para estes valores de d , a Tabela 2 fornece $ARL(n)$ e $g(n)$ para $n=30$ e $n=25$, a fim de mostrar a sensibilidade da solução à substituição de n_{ot} por estes valores de n . Casos de restrições ao valor de n serão tratados adiante (Seção 5).

4. Exemplos Numéricos e Discussão

Suponha que o d de interesse é 1,2. Para $L=3$, a Tabela 1 fornece, para $d=1,2$, $n_{ot}=11$. Para ter um exemplo do ganho de eficiência que o método proposto proporciona em relação ao uso de valores empíricos, comparemos os desempe-

Tabela 2 – Projetos alternativos para d pequeno

| d | $n = 30$ | | $n = 25$ | |
|-----|-----------|---------|-----------|---------|
| | $ARL(30)$ | $g(30)$ | $ARL(25)$ | $g(25)$ |
| 0,4 | 4,78 | 128,4 | 6,30 | 145,1 |
| 0,5 | 2,52 | 60,59 | 3,24 | 68,53 |
| 0,6 | 1,63 | 33,69 | 2,00 | 37,50 |
| 0,7 | 1,25 | 22,60 | 1,45 | 23,66 |

nhos que o valor ótimo $n=11$ e o valor tradicional $n=4$ proporcionam para o gráfico. Já foi visto que $g(n)$ é uma medida da eficiência do gráfico. A tabela na Figura 2 mostra que $g(11) = 7,65$, enquanto $g(4) = 12,59$.

Suponha que se quer minimizar r (forma “dual” do problema, dada por (18), (19) e (20)), e que o limite de tolerância para AATS é $AATS_{max} = 2$ horas. Então, com $n=4$, $ARL = 3,65$ (pela tabela na Figura 2), e (de (12)) $AATS = (3,65 - 0,5).h$; isto leva a $h = AATS_{max}/3,15 = 0,64$ horas, ou 38 minutos. A intensidade de inspeção $r = n/h = 4/0,63 = 6,32$ itens/hora.

Com $n = 11$ (projeto ótimo), $ARL = 1,20$, $AATS = (1,20 - 0,5).h$, o que, para produzir o mesmo AATS de 2 horas, requer $h = AATS_{max}/0,70 = 2,86$ horas, ou 171 minutos; a intensidade de inspeção será $r = 11/2,86 = 3,85$ itens/hora. A economia de recursos terá sido de $100(1 - 3,85/6,32)\% = 39\%$ em relação ao *design* empírico.

O significado do tamanho ótimo de amostra n_{ot} pode ser ainda ilustrado pelo seguinte exemplo: suponha que se esteja operando com $h = 171$ minutos e $n = 4$ (mantenha-se $d=1,2$ e $L=3$): pela fórmula (12), AATS seria igual a 3,15 $h = 9$ horas. Caso se queira reduzir AATS para 2 horas mediante uma redução em h (mantendo n fixo no valor não ótimo $n=4$), será necessário reduzir h a $2/9$ do seu valor, o que faz r crescer 4,5 vezes. No entanto, chega-se ao mesmo AATS mantendo h constante e aumentando n para o valor ótimo $n=11$, o que equivale a multiplicar n (e portanto, r) por apenas 2,75. A razão entre as intensidades de inspeção resultantes $(= (2,75r_0/4,5r_0) = 0,61)$ é a mesma razão $g(11)/g(4)$: eficiência relativa entre os projetos.

Outra maneira ainda de ver a questão é a seguinte: suponha a intensidade de inspeção fixa: $r=r_0$ (problema “primal”, definido por (3), (4) e (5)). Se a decisão for entre as alternativas: (a): $n=4$, $h=4/r_0$ e (b): $n=11$, $h=11/r_0$, a alternativa (a) tem $AATS_a = 3,15(4/r_0) = 12,6/r_0$, e a alternativa (b) tem $AATS_b = 0,7(11/r_0) = 7,7/r_0$. Portanto, $AATS_b = 0,61 AATS_a$. Para ter um exemplo numérico, suponha-se que se definiu $r_0 = 4$ itens por hora: a alternativa (a) seria $n=4$, $h=1$ hora, e teria $AATS_a = 3,15$ horas (189 minutos); a alternativa (b) seria $n=11$, $h=11/4$ horas (2 horas e 45 minutos), e teria $AATS_b = 1,93$ horas (115 minutos).

- Continuando no mesmo exemplo: caso este AATS de 115 minutos seja considerado inaceitável, e se esteja disposto a aumentar a intensidade de inspeção para reduzi-lo, então, o fato de o tamanho de amostra ótimo ser igual a 11 significa que é mais eficiente reduzir h que aumentar n . Por exemplo, suponhamos que se está disposto a dobrar a intensidade de inspeção: mantendo h inalterado e dobrando n (i.e., fazendo $n=22$), obtém-se $AATS = 5,54/r_0$ (83 minutos); porém mantendo $n=11$ e fazendo $h=11/8$ horas, o AATS se reduz a $3,83/r_0$ (58 minutos).

5. Problema com Restrições Adicionais a n e/ou h

Até aqui se tratou do problema sem restrições explícitas aos valores de n ou h . Contudo, é usual haver razões de ordem prática ou mesmo teórica impondo tais restrições. Por exemplo, o uso de um gráfico de R ou de S em

| n | ARL | g(n) |
|----|-------|-------|
| 1 | 27,82 | 27,32 |
| 2 | 10,38 | 19,77 |
| 3 | 5,61 | 15,32 |
| 4 | 3,65 | 12,59 |
| 5 | 2,66 | 10,81 |
| 6 | 2,10 | 9,61 |
| 7 | 1,76 | 8,79 |
| 8 | 1,53 | 8,25 |
| 9 | 1,38 | 7,90 |
| 10 | 1,27 | 7,71 |
| 11 | 1,20 | 7,65 |
| 12 | 1,14 | 7,69 |
| 13 | 1,10 | 7,82 |
| 14 | 1,07 | 8,02 |
| 15 | 1,05 | 8,28 |
| 16 | 1,04 | 8,60 |
| 17 | 1,03 | 8,95 |
| 18 | 1,02 | 9,33 |
| 19 | 1,01 | 9,75 |
| 20 | 1,01 | 10,18 |
| 21 | 1,01 | 10,63 |
| 22 | 1,00 | 11,09 |
| 23 | 1,00 | 11,57 |
| 24 | 1,00 | 12,05 |
| 25 | 1,00 | 12,53 |
| 26 | 1,00 | 13,02 |
| 27 | 1,00 | 13,52 |
| 28 | 1,00 | 14,01 |
| 29 | 1,00 | 14,51 |
| 30 | 1,00 | 15,01 |

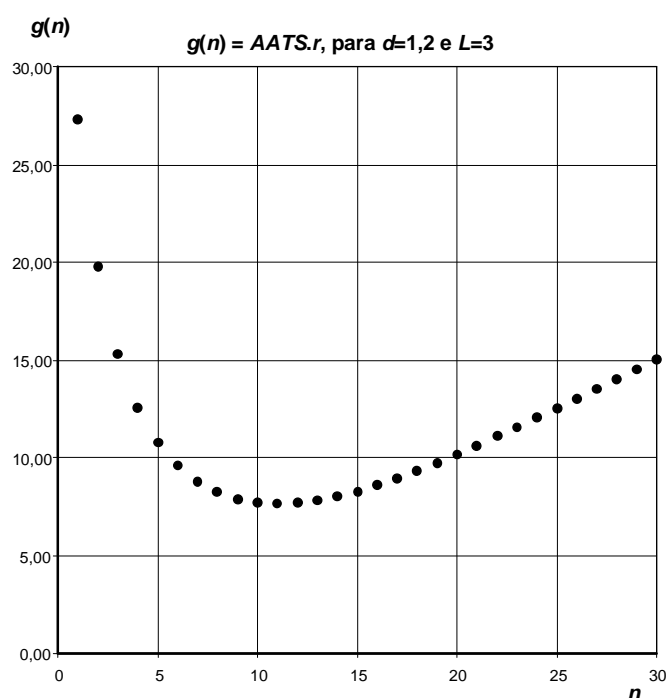


Figura 2 – $g(n)$ e $ARL(n)$ para $d=1,2$ e $L=3$

conjunto com o gráfico de \bar{X} , que é o procedimento usual, requer, para bom desempenho daquele, que n seja no mínimo igual a 4 unidades. Por outro lado, algumas situações típicas que limitam o valor máximo admissível para n são: limitações na taxa de produção, dificuldades operacionais de se separar uma amostra de mais de n_{max} itens de uma vez,

procedimentos sofisticados (demorados) de medição, ou que requeiram ensaios destrutivos. Finalmente, a interferência da extração de uma amostra sobre a operação do processo, ou ainda a existência de uma correlação não desprezível entre medidas de \bar{X} próximas no tempo (mesmo que medidas individuais, de X , possam apresentar variabilidade razoável) podem exigir um

espaçamento temporal mínimo entre duas amostras consecutivas. Estas situações resultam nas seguintes restrições, isoladamente ou em conjunto: $n \geq n_{min}$, $n \leq n_{max}$, e $h \geq h_{min}$.

A análise do problema com tal tipo de restrições requer o exame da forma da função $g(n)$ e o estudo detalhado de 7 casos diferentes, cada um dos quais se subdivide em duas ou mais situações. Esta análise pode ser encontrada em EPPRECHT & SANTOS (1998). Por falta de espaço, aqui são apresentados apenas os procedimentos para solução em cada caso (ver Quadro 1). Estes procedimentos fazem referência a duas fórmulas ainda não apresentadas, que são:

$$n'_{min} = r_{max}h_{min} \quad (23)$$

$$n'_{min} = \min\left(n \mid ARL(n) \leq \frac{AATS_{max}}{h_{min}} + \frac{1}{2}\right) \quad (24)$$

6. Uso do Método

O procedimento para uso do método na prática é sintetizado objetivamente abaixo. O usuário deve especificar: o d de interesse (que pode ser determinado pela fórmula (1) — ver discussão na Seção 2 deste artigo); o ARL_0 tolerado (ARL_{0min}); e um limite máximo para r (r_{max}) — ou, alternativamente, para $AATS_d$ ($AATS_{max}$). Opcionalmente, ele pode ainda especificar restrições adicionais na forma $n \geq n_{min}$ e/ou $n \leq n_{max}$ e/ou $h \geq h_{min}$. Requer-se que o usuário disponha de uma tabela de valores de n_{ot} em função de d , para o problema “irrestrito”, e para o ARL_0 (ou L) especificado. Aqui se recomenda (conforme argumentação na Seção 3) adotar $L=3$ (que fornece $ARL_0=370,4$). Neste caso, a tabela de n_{ot} é a Tabela 1 apresentada na Seção 3. Se, contudo, o usuário quiser especificar outro valor para ARL_0 , ele precisará, então:

- Calcular L pela fórmula (14), e
- Gerar a tabela para o L calculado. Isto pode ser feito em uma planilha eletrônica, em 2 etapas: primeiro, empregando as fórmulas (11) e (17), calcula-se $g(n)$ versus n para diversos valores de d , de modo a obter, para

cada valor de d , uma tabela e gráfico como na Figura 2; em seguida, em cada uma destas tabelas, determina-se $n_{ot}(d)$: o valor de n que produz o menor valor para $g(n)$ para aquele valor de d . Estes pares $(d, n_{ot}(d))$ constituem a tabela.

De posse da tabela, o uso do método se torna extremamente simples:

- Entrando com d na tabela, obter o valor ótimo de n para o problema “irrestrito”; chamar este valor de n^* ;
- Se não houver restrições ao valor de n ou h :
 - $n_{ot} = n^*$;
 - calcular h_{ot} da seguinte maneira:
 - na forma “primal” do problema: $h_{ot} = n_{ot}/r_{max}$ (fórmula (16)).
 - na forma “dual”: h_{ot} pode ser calculado tanto por $h_{ot} = n_{ot}AATS_{max}/g(n_{ot})$ (fórmula (21)) como por $h_{ot} = AATS_{max}/(ARL(n_{ot})-0,5)$ (fórmula (22)).
- Se houver restrições a n e/ou a h , a solução é dada, conforme o caso, pelo Quadro 1.
- Em qualquer caso, uma vez determinados n e h , r é dado por $r=n/h$, e $AATS$ por $AATS=g(n)/r$, ou por $AATS=(ARL(n)-0,5)h$.

OBS.: $g(n)$ e $ARL(n)$, se não forem dados pela tabela, são calculados por (17) e (11).

Mesmo que o usuário não saiba especificar valores para $AATS_{max}$ ou r_{max} , ainda assim o modelo é aplicável, e útil: basta entrar com d na tabela para obter o valor n_{ot} que otimiza a relação entre $AATS$ e r . De posse deste valor, $g(n_{ot})$ mostra diretamente a relação entre o nível de recursos de inspeção alocados (representado pela intensidade de inspeção, r) e o desempenho obtido (representado pelo $AATS$). Esta relação é dada simplesmente por $AATS=g(n_{ot})/r$. Alternativamente, $AATS=[g(n_{ot})/n_{ot}]h$ fornece a relação entre h e $AATS$. Isto permite ao usuário examinar toda a gama de opções para decisão, ótimas em termos da relação entre $AATS$ e r .

Há ainda um procedimento alternativo para solução: caso se disponha, para o L prefixado, de um conjunto de tabelas e gráficos de $g(n)$ versus n , para diversos valores de d (tabelas e gráficos semelhantes aos da Figura 2 para $d=1,2$), pode-

Quadro 1 – Solução para o problema com restrições ao valor de n e/ou h (NOTA: n^* é o valor ótimo de n para o problema sem estas restrições)

| Caso | Restrições | Solução ótima |
|------|--|--|
| A | $n \geq n_{min}$ | 1. se $n_{min} \leq n^*$, $n_{ot} = n^*$; se $n^* \leq n_{min}$, $n_{ot} = n_{min}$; 2. h_{ot} é dado por: (16) (forma “primal”) ou por (21) ou (22) (forma “dual”). |
| B | $n \leq n_{max}$ | 1. se $n^* \leq n_{max}$, $n_{ot} = n^*$; se $n_{max} \leq n^*$, $n_{ot} = n_{max}$; 2. h_{ot} é dado como em A (passo 2). |
| C | $n_{min} \leq n \leq n_{max}$ | 1. se $n_{min} \leq n^* \leq n_{max}$, $n_{ot} = n^*$; se $n^* > n_{max}$, $n_{ot} = n_{max}$; se $n^* < n_{min}$, $n_{ot} = n_{min}$; 2. h_{ot} é dado como em A (passo 2). |
| D | $h \geq h_{min}$ | 1. Calcular n'_{min} : ⇒ problema “primal”: n'_{min} dado pela fórmula (23); ⇒ problema “dual”: n'_{min} dado pela fórmula (24); 2. Proceder como em A, com n'_{min} no lugar de n_{min} . |
| E | $n \geq n_{min}$ e $h \geq h_{min}$ | 1. Calcular n'_{min} como no caso D. 2. $n''_{min} = \max\{n'_{min}, n_{min}\}$; 3. Proceder como no Caso A, com n''_{min} no lugar de n_{min} . |
| F | $n \leq n_{max}$ e $h \geq h_{min}$ | 1. Calcular n'_{min} como no Caso D 2. se $n'_{min} \leq n_{max}$: proceder como no Caso C, com n'_{min} no lugar de n_{min} ; senão: ⇒ problema “primal”: $n_{ot} = n_{max}$ e $h_{ot} = h_{min}$; ⇒ problema “dual”: sem solução: é preciso relaxar alguma das restrições. |
| G | $n_{min} \leq n \leq n_{max}$ e $h \geq h_{min}$ | 1. Calcular n'_{min} como no Caso D 2. se $n'_{min} \leq n_{min}$: ignorar a restrição $h \geq h_{min}$ e proceder como no Caso C; senão: proceder como no passo 2 do Caso F. |

se, olhando a tabela/gráfico correspondente ao valor de d de interesse, visualizar imediatamente o valor ótimo para n , inclusive no caso com restrições ao valor de n . No caso de restrição ao valor de h , caso esta não seja satisfeita pelo valor h calculado para a solução ótima, basta determinar n'_{min} segundo a fórmula (23) ou (24), conforme a forma do problema, “primal” ou “dual”, e reinspecionar o gráfico e/ou tabela.

Inclusive, (24) é resolvida por inspeção à própria tabela, desde que ela contenha uma coluna de $ARL_d(n)$ (observe que $ARL_d(n)$ é monotonicamente decrescente). Uma vantagem adicional deste procedimento é permitir uma visualização imediata da sensibilidade da solução ao valor de n . EPPRECHT & SANTOS (1998) fornecem o conjunto de tabelas e curvas para $L=3$; em todo caso, lembremos que este conjunto pode ser

gerado sem dificuldade (para qualquer valor de L) em uma planilha eletrônica, usando as fórmulas (11) e (17).

O valor h_{ot} obtido para h pode ser pouco natural (por exemplo, 171 minutos). Caso se queira “arredondar” este valor para um valor operacionalmente mais cômodo, então, para não violar as restrições do problema, o arredondamento deve ser feito, na forma “primal”, “para cima” (aumentando h), e na forma “dual”, “para baixo”. A menos que se considere preferível, por razões operacionais, relaxar ligeiramente as restrições. Em qualquer caso, se o arredondamento for grande, deve-se reavaliar o valor ótimo de n . EPPRECHT & SANTOS (1998) fornecem uma orientação de como fazê-lo, que aqui é omitida por limitações de espaço.

7. Comparação com Trabalhos Precedentes

O modelo proposto enquadra-se na categoria dos modelos *semi-econômicos* para projeto de gráficos de controle, por basear-se em critérios ligados implicitamente aos custos de inspeção e de produzir em condições degradadas, mas sem considerar *todos* os elementos de custo. Outros autores que propuseram modelos nesta categoria são: WEILER (1952), KEATS *et al.* (1995) e DEL CASTILLO *et al.* (1996).

WEILER (1952) desenvolveu um modelo para, com d e L fixos, minimizar o número esperado de itens inspecionados entre o descontrole e o sinal, $A(n)=ARL(n).n$. Chama a atenção a semelhança entre esta função e a função $g(n)$ aqui consideradas. Há duas diferenças, porém. A primeira é a não consideração por WEILER do tempo S entre o instante da amostra anterior ao descontrole e o instante do descontrole, o que leva a resultados subótimos em termos da razão entre os recursos empregados (r) e a rapidez de detecção ($1/AATS$). A divergência entre os tamanhos de amostra obtidos pelo seu modelo e os obtidos neste artigo é tão mais acentuada quanto menor o ARL_d . A segunda diferença é que, enquanto que neste artigo considera-se o número de itens inspecio-

nados *por unidade de tempo* (r), o objetivo de WEILER é minimizar o número médio de itens inspecionados *entre o instante do descontrole e a sua detecção*. A semelhança entre a função $A(n)$ e a função $g(n)$ decorre do fato que, na forma “dual” do problema aqui considerado, ao minimizar r se minimiza o número de itens inspecionados em *qualquer* intervalo de tempo e, em particular, ao se prefixar $AATS=AATS_{max}$, se minimiza o número médio de itens entre o instante do descontrole e a sua detecção. Porém os objetivos são diferentes (o que é mais evidente para a forma “primal” do problema, em que se trata de minimizar o *tempo* até a detecção).

Tais constatações são tanto mais importantes quanto o critério $A(n)$ de WEILER (1952) é difundido. PYZDEK (1990) reproduz, como recomendação de uso, uma tabela de tamanhos de amostra ótimos segundo aquele critério; WADSWORTH *et al.* (1986) apresentam o gráfico de uma família de curvas de $A(n)$ (que chamam de $I(n)$) *versus* d , parametrizadas para diversos valores de n (a idéia é adotar o n correspondente à curva de menor valor para o d de interesse); MONTGOMERY (1991, 1997) reproduz este gráfico, sem mostrar a relação entre $I(n)$ e $AATS$. Na última edição (1997), ele inclusive menciona o critério $ATS = ARL.h$ (que é usado por vários autores, entre os quais SANIGA, 1989), mas não o relaciona a $I(n)$.

Como medida de desempenho, o $AATS$ já foi utilizado por outros autores (por exemplo COSTA, 1997) com outros escopos que o projeto ótimo do gráfico (cabe notar que esta notação — “ $AATS$ ” — foi aqui mantida por uniformidade com a literatura).

Mais recentemente, KEATS *et al.* (1995) desenvolveram um modelo para o problema de minimizar o “comprimento médio de produção” APL (“*average production length*”), definido como o número médio de itens produzidos entre o descontrole e o sinal. Note que o tempo e o número de itens produzidos se relacionam pela taxa de produção, de modo que basta redefinir a unidade de tempo como “itens produzidos” para $AATS$ e APL representarem a mesma coisa. As

diferenças substanciais do modelo de KEATS *et al.* (1995) em relação ao nosso são: a) o cômputo do número de itens da amostra no cálculo do APL (o que não terá efeito significativo se — como foi assumido aqui — o número de itens produzidos entre duas amostras for suficientemente maior do que n); e b) o estabelecimento de um limite mínimo para o tempo médio até um alarme falso (na verdade, para o número médio de itens produzidos até um alarme falso, APL_0), em vez do número de amostras médio até um alarme falso, ARL_0 . Com isto, a “demora” até um alarme falso é representada na mesma unidade que a “demora” até a detecção. Pode-se considerar esta representação mais coerente que a utilização de $AATS_d$ e ARL_0 , como aqui é feito (observe, contudo, que estas demoras não são diretamente comparáveis: uma é o tempo médio entre alarmes falsos; mas a outra não é o tempo médio entre detecções; para obtê-lo seria preciso incluir no modelo o tempo médio entre descontroles, $MTIC$). O preço desta representação, porém, é que a solução do modelo de KEATS *et al.* (1995) requer a execução de um programa de computador. A vantagem do método aqui proposto é a simplicidade: a inspeção a uma tabela determina diretamente o projeto ótimo.

Na verdade, o método aqui proposto também permite representar uma restrição ao tempo médio até um alarme falso, ATS_0 , adicional à restrição $ARL_0 \geq ARL_{0min}$. Basta ver que tal restrição, na forma $ATS_0 \geq ATS_{0min}$, é equivalente, na forma “primal”, a uma restrição de h mínimo ($h \geq ATS_{0min}/ARL_{0min}$); e, na forma “dual”, a uma restrição de n mínimo (tal que $ARL_{d,L}(n) \leq AATS_{max}ARL_{0min}/AATS_{0min}$).

Outra vantagem do método aqui proposto em relação ao de KEATS *et al.* (1995) é que, neste último, restrições a n ou h não são contempladas; se necessário, estas teriam de ser embutidas no programa de computador que calcula a solução ótima.

DEL CASTILLO *et al.* (1996) desenvolveram um método interativo de apoio à decisão para o *design* de gráficos de controle, implemen-

tado mediante um programa de computador. A opinião dos presentes autores é que, por um lado, o método em questão não considera mais elementos do que a abordagem aqui proposta ou que a de KEATS *et al.* (1995) (à exceção do tempo médio que o processo permanece em controle, $MTIC$); por outro lado, ele traz mais dificuldades de utilização pelo menos que o método aqui proposto, por requerer uma estimativa do $MTIC$, por necessitar do programa de computador, e por exigir uma interação mais longa com o usuário do gráfico. Assim, não oferece vantagem significativa, enquanto apresenta limitações ao seu emprego.

O método aqui proposto não requer que se conheça o $MTIC$, mas, caso ele seja conhecido, pode orientar a escolha de r_{max} ou $AATS_{max}$ (conforme a versão do problema), pois quanto mais frequentes forem os descontroles (i.e., menor $MTIC$) mais se está disposto a investir em inspeção (maior r_{max}); e quanto mais raros os descontroles, mais raras também as perdas por produzir em condições degradadas; menor portanto o retorno do investimento em inspeção, e maior o $AATS_{max}$ tolerado, para reduzi-la.

8. Conclusões

O método proposto é bastante imediato para a seleção de parâmetros (tamanho de amostra n , intervalo de tempo entre amostras h , e abertura L dos limites de controle) para gráficos de \bar{X} . Para o caso sem restrições explícitas aos valores dos parâmetros, a Tabela 1 fornece diretamente os valores ótimos para n , e h é determinado seja pela fórmula (16), seja por (21) ou (22), conforme a versão do problema: minimizar o tempo médio até a detecção do descontrole ($AATS$) com recursos de inspeção limitados, ou minimizar gastos de inspeção com a restrição de um limite superior tolerado para o $AATS$. (Os resultados da Tabela 1 são para $L=3$, correspondendo a ARL_0 prefixado em 370; para outros valores de ARL_0 , tabela semelhante pode ser facilmente produzida pelo mesmo método, ilustrado claramente neste artigo e explicitado na Seção 6).

Para o caso de haver restrições adicionais aos valores de n e/ou h , o Quadro 1 resume os procedimentos para solução.

O método é superior ao projeto puramente estatístico de gráficos de controle, e tem, sobre os modelos econômicos, a vantagem de simplicidade; é quase tão imediato quanto utilizar valores empíricos gerais, levando porém a um uso mais eficiente dos recursos, como ilustrado no exemplo da Seção 4. Esta simplicidade torna viável sua implementação em qualquer ambiente, e mesmo sua inclusão em cursos introdutórios de controle de qualidade. Uma contribuição adicional do método é o tratamento detalhado dos casos de restrições a n e/ou h .

Os valores ótimos obtidos diferem dos propostos por WEILER (1952) (e recomendados por PYZDEK (1990), WADSWORTH *et al.* (1986) e MONTGOMERY (1991, 1997)); a diferença se deve a ter aquele autor desconsiderado a dimensão tempo e com isso a aleatoriedade do tempo entre a ocorrência do descontrole e o instante de extração da 1ª amostra subsequente, o que levou a resultados subótimos quanto ao uso dos recursos de inspeção.

Como o gráfico de \bar{X} quase sempre é utilizado em conjunto com um gráfico de R ou S , faz sentido estender a análise também para o problema de projeto conjunto de gráficos de \bar{X} e R (ou \bar{X} e S). As direções de prosseguimento planejadas para a pesquisa incluem esta análise, bem como a análise de outros esquemas de controle. Dependendo do esquema em questão, o cálculo do ARL pode se tornar bastante

complexo; a equação (15), porém, vale para todos os esquemas, e a função $g = AATS.r$ torna-se um critério universal e poderoso para estas análises.

No projeto de um esquema de controle, uma questão que precede a seleção dos parâmetros é a seleção do próprio esquema a adotar. Para d pequeno, por exemplo, cabe a pergunta se um esquema EWMA, CUSUM, misto (Shewhart-EWMA, Shewhart-CUSUM), ou outro, seria mais eficiente que o gráfico de \bar{X} . Ora, para responder a esta pergunta devem-se comparar as diversas alternativas *em seus respectivos projetos ótimos*. Os presentes resultados são portanto um subsídio indispensável para a análise comparativa de alternativas ao gráfico de \bar{X} .

Nesta direção de extensão da análise para outros esquemas de controle, e de comparação de alternativas, uma análise semelhante para o esquema EWMA também foi realizada (EPPRECHT *et al.*, 1998), e a análise para o gráfico de \bar{X} com n e h variáveis (ver COSTA, 1997, e PRABHU *et al.*, 1994) está em andamento.

Agradecimento

Esta pesquisa se insere no projeto de pesquisa do primeiro autor, e na pesquisa de mestrado da segunda autora, ambos apoiados pelo CNPq. A forma final deste artigo se beneficiou de valiosas sugestões de Antonio Fernando Branco Costa, FEG/UNESP, Guaratinguetá, e das críticas de um dos revisores.

Referências Bibliográficas

- COSTA, A.F.B.: " \bar{X} Charts with Variable Sample Sizes and Sampling Intervals". *Journal of Quality Technology*, 29: 197-204, 1997.
- DEL CASTILLO, E.; MACKIN, P. & MONTGOMERY, D.C.: "Multiple-criteria optimal design of \bar{X} control charts". *IIE Transactions*, 28: 467-474, 1996.

- EPPRECHT, E.K.; NINIO, A.L. & SOUZA, M.O.: "Projeto Ótimo de Gráficos de Médias Móveis Ponderadas Exponencialmente (EWMA) para Controle Estatístico de Processo". *Pesquisa Operacional* (a ser publicado). Também disponível (versão mais extensa) como Memorando Técnico n. 05/98, Departamento de Engenharia Industrial, PUC-Rio, Rio de Janeiro, 1998.

- EPPRECHT, E.K. & SANTOS, A.B.:** “Um Método Simples para o Projeto Ótimo de Gráficos de \bar{X} — com análise de restrições ao tamanho de amostra e frequência de amostragem”. Memorando Técnico n. 03/98, Departamento de Engenharia Industrial, PUC-Rio, Rio de Janeiro, 1998.
- KEATS, J.B.; DEL CASTILLO, E.; VON COLLANI, E. & SANIGA, E.M.:** “Economic modelling for statistical process control”. *Journal of Quality Technology*, 29(2): 144-147, 1997.
- KEATS, J.B.; MISKULIN, J.D. & RUNGER, G.C.:** “Statistical process control scheme design”. *Journal of Quality Technology*, 27(3): 214-225, 1995.
- MONTGOMERY, D.C.:** *Introduction to Statistical Quality Control*. New York, John Wiley & Sons, 1991 (2nd ed.), 1997 (3rd ed.).
- PRABHU, S.S.; MONTGOMERY, D.C. & RUNGER, G.C.:** “A Combined Adaptive Sample Size and Sampling Interval \bar{X} Control Scheme”. *Journal of Quality Technology*, 26: 164-176, 1994.
- PYZDEK, T.:** *Pyzdek's Guide to SPC: Vol. 1, Fundamentals*. Tucson, Quality Publishing, Inc., 1990.
- SANIGA, E.W.:** “Economic statistical control-chart designs with an application to \bar{X} and R charts”. *Technometrics*, 31(3), 313-320, 1989.
- WADSWORTH, W.H.; STEPHENS, K.S. & GODFREY, A.B.:** *Modern Methods for Quality Control and Improvement*. New York, John Wiley & Sons, 1986.
- WEILER, H.:** “On the most economical sample size for controlling the mean of a population”. *Annals of Mathematical Statistics*, 23: 247-254, 1952.
- WHEELER, D.J.:** *Advanced Topics in Statistical Process Control: The Power of Shewhart's Charts*. Knoxville, SPC Press, 1995.
- WOODALL, W.H.:** “Weaknesses of the economic design of control charts”. *Technometrics*, 28(4), 408-409, 1986.

A SIMPLE METHOD FOR OPTIMAL \bar{X} -CHART DESIGN

Abstract

A simple but efficient method has been developed for the design of \bar{X} -charts. The solutions provided are optimal in terms of the ratio between the number of items inspected per unit time and the detection speed (defined as the inverse of the average time to signal a shift in the mean). The method is superior to design procedures based exclusively on statistical properties of the chart (such as the ARL), and does not have the implementation drawbacks of economical design models. The results obtained differ from some results widely reported in the literature, which are shown to be sub-optimal as a consequence of their underlying model not considering some important aspects of the problem. An additional contribution of the method is the explicit treatment of optional constraints on the sample size and/or on the interval between samples. The ease of use of the method makes its implementation straightforward in any environment, and also makes it suitable for the classroom.

Key words: \bar{X} -chart, Shewhart charts, control chart design, semi-economic design, statistical process control, statistical quality control.