

Controle Estatístico da Dispersão

Os trabalhos existentes na área de Controle Estatístico de Processos multicanal voltados para a detecção das causas especiais que afetam individualmente os canais do processo são dedicados exclusivamente à sinalização de descontrole na média dessas componentes individuais. Ainda não há trabalhos que proponham novos esquemas ou estudem a sensibilidade dos esquemas já existentes contra alterações na dispersão das componentes individuais de um processo multicanal.

Uma das contribuições desta tese é o estudo da sensibilidade dos esquemas de controle já existentes (voltados para o monitoramento das médias das componentes individuais) a aumentos na *dispersão* do processo (especificamente, são considerados os esquemas de controle da média de Mortell e Runger, 1995, Runger et al., 1996 e Barbosa, 2008). Tal proposta tem o intuito de verificar se os esquemas de controle existentes possuem desempenho bom o suficiente para considerar apenas um gráfico de controle para o monitoramento da *média* e da *dispersão* das componentes individuais, sem a necessidade de utilizar um esquema de controle dedicado exclusivamente ao controle da dispersão.

Cabe ressaltar que a análise que será realizada possui caráter inédito, pois não há nenhum outro trabalho na literatura que mostra tais resultados para o controle estatístico da dispersão em processos multicanal.

Para tal estudo, considera-se nesta tese que as alterações na dispersão ocorrem somente em um canal do processo. Em relação ao GCG de diferenças em relação ao nível-base, neste capítulo são desenvolvidas analiticamente as probabilidades, no caso de alteração na dispersão, dos seguintes eventos: sinalização associada ao canal afetado pela causa especial e sinalização por um canal diferente do afetado pela causa especial. Tais eventos são considerados para dar maior entendimento sobre os tipos de sinais que podem ocorrer no monitoramento da dispersão, dada a suposição de ocorrência de descontrole em apenas um dos c canais do processo.

A distribuição dos valores de RL (*Run Lengths* – comprimento de corrida) dos gráficos de controle de Shewhart é geométrica, então, como a média da distribuição geométrica é o inverso da probabilidade do evento ocorrer, o NMA (número médio de amostras até um sinal; valor médio de *RL*) pode ser obtido facilmente pela inversa da probabilidade de sinal em uma amostra qualquer.

Para alguns dos eventos que serão considerados na análise, duas medidas de desempenho serão utilizadas: a probabilidade de sinalização pelo gráfico de controle e o número médio de amostras até um sinal, considerando a ocorrência de alguma causa especial de variação que tenha afetado a *dispersão* do processo.

As probabilidades acima descritas serão obtidas somente para os esquemas de controle do tipo de Shewhart, pois são os gráficos de controle cuja probabilidade de sinalização (supondo os parâmetros do processo estáticos, seja em controle e fora de controle) é constante nas diversas amostras. Nos esquemas EWMA, as probabilidades de sinalização de descontrole não são constantes porque as observações não são independentes. Então, não se pode obter o NMA pela inversa da probabilidade de sinal em uma amostra qualquer. No entanto, pode-se obter uma frequência de sinalização em uma amostra, calculando a inversa do NMA_1 que deve ser obtido por outro método (como por exemplo, por um modelo de Cadeias de Markov).

Outra contribuição desta tese é o detalhamento do projeto estatístico de dois gráficos de controle de grupos (GCG) do tipo de Shewhart dedicados ao monitoramento da *dispersão* de um canal em particular do processo. Foram desenvolvidos dois desses gráficos: um, para dados subgrupados (amostras maior que uma unidade em cada canal) e outro, para observações individuais. Além dos gráficos de controle de Shewhart citados, foram também desenvolvidas versões EWMA correspondentes, para garantir detecções mais rápidas de aumentos de magnitude pequena a moderada na *dispersão* do canal afetado. Tais esquemas de controle serão detalhados neste capítulo. Os projetos ótimos dos gráficos de controle de EWMA aqui desenvolvidos serão mostrados nos próximos capítulos.

3.1

Estudo analítico das medidas de desempenho do GCG das diferenças no caso de aumentos da dispersão de um canal

Na ocorrência de algum aumento por um fator γ na dispersão da parcela individual e_i de um canal qualquer (mas de apenas um dos canais), passando este canal afetado a ter dispersão $\sigma_1^2 = \gamma \sigma_0^2$, a diferença média \hat{e}_{ii} terá a sua variância afetada. Tais alterações sofridas afetarão as probabilidades de sinalização do GCG das diferenças de Barbosa (2008). Especificamente será afetada tanto a probabilidade de sinalização associada ao canal afetado, quanto à probabilidade de sinalização associada aos demais canais do processo.

Nesta seção são formuladas analiticamente as equações que fornecem as probabilidades acima mencionadas, permitindo assim o cálculo das medidas de desempenho do GCG de Barbosa (2008) para diversos tipos de sinais na presença de causas especiais de variação que afetam a dispersão da parcela individual de algum canal em particular do processo. Vale salientar que Barbosa (2008) não avaliou o GCG de diferenças para tais casos, somente analisou os casos de alterações na média.

É conveniente lembrar o modelo Mortell e Runger (1995) reescrito por Barbosa (2008) para processos multicanal é o seguinte:

$$x_{ij} = b_t + e_{ij} \quad (2.26)$$

a) Probabilidade de Sinalização associada ao Canal Afetado pela Causa Especial:

Relembrando, a *diferença média de um canal i* qualquer obtida em (2.40) é escrita como:

$$\hat{e}_{ii} = e_{ii} - e_{t..} \quad (2.40)$$

onde:

$e_{t..} \sim N(0, \sigma^2/nc)$ considerada a independência dos e_{ij} para diversos valores de t, i e j .

A equação (2.40) pode ser reescrita sob a seguinte forma analítica:

$$\hat{e}_{ii.} = e_{ii.} - \frac{1}{c} \sum_{k=1}^c e_{tk.} \quad (3.1)$$

ou, de forma mais simplificada:

$$\hat{e}_{ii.} = \left(\frac{c-1}{c} \right) e_{ii.} - \frac{1}{c} \sum_{\substack{k=1 \\ k \neq i}}^c e_{tk.} \quad (3.2)$$

onde:

$$e_{ii.} \sim N(0, \sigma^2/n)$$

$$\hat{e}_{ii.} \sim N\left(0, \left(\frac{c-1}{c}\right) \frac{\sigma^2}{n}\right)$$

Partindo de (3.1), o valor esperado é obtido por:

$$\begin{aligned} E(\hat{e}_{ii.}) &= E(e_{ii.}) - \frac{1}{c} \sum_{k=1}^c E(e_{tk.}) \\ &= 0 \end{aligned} \quad (3.3)$$

A ocorrência de uma causa especial de variação que altera a dispersão de um canal i qualquer de σ_0^2 para $\gamma^2 \sigma_0^2$ não afeta o valor esperado do ruído $e_{ii.}$. Já a variância do ruído $e_{ii.}$ se altera para $\gamma^2 \sigma^2/n$ e a variância dos ruídos dos demais canais permanece a mesma, i.e., σ^2/n . A variância de $\hat{e}_{ii.}$ pode ser determinada facilmente por:

$$Var(\hat{e}_{ii.}) = Var\left(\left(\frac{c-1}{c}\right) e_{ii.} - \frac{1}{c} \sum_{\substack{k=1 \\ k \neq i}}^c e_{tk.}\right) \quad (3.4)$$

Por hipótese, as diferenças médias teóricas são independentes uma das outras ($e_{ii.}$ independente de $e_{tk.}$ para $i \neq k$). Então a equação (3.4) se reduz simplesmente a:

$$\begin{aligned} Var(\hat{e}_{ii.}) &= \left(\frac{c-1}{c}\right)^2 Var(e_{ii.}) + \left(\frac{1}{c}\right)^2 \sum_{\substack{k=1 \\ k \neq i}}^c Var(e_{tk.}) \\ &= \left(\frac{c-1}{c}\right)^2 Var(e_{ii.}) + \left(\frac{1}{c}\right)^2 (c-1) Var(e_{tk.}) \end{aligned} \quad (3.5)$$

Ou seja, a equação (3.5) tem o seu valor expresso por:

$$\begin{aligned} \text{Var}(\hat{e}_{ii.}) &= \left(\frac{c-1}{c}\right)^2 \frac{\sigma^2 \gamma^2}{n} + \left(\frac{1}{c}\right)^2 (c-1) \frac{\sigma^2}{n} \\ &= \frac{\sigma^2}{n} \left[\frac{\gamma^2 (c-1)^2 + (c-1)}{c^2} \right] \end{aligned} \quad (3.6)$$

Sintetizando, a distribuição amostral da *diferença média do canal i* ($\hat{e}_{ii.}$) na presença de alguma causa especial de variação que afete a dispersão do canal i é determinada por:

$$\hat{e}_{ii.} \sim N\left(0, \frac{\sigma^2}{n} \left[\frac{\gamma^2 (c-1)^2 + (c-1)}{c^2} \right]\right)$$

Ao considerar $\gamma^2 = 1$, a variância obtida em (3.6) será exatamente igual à variância da diferença média do canal i ($\hat{e}_{ii.}$) quando o processo encontra-se em estado de controle.

$$\begin{aligned} \text{Var}(\hat{e}_{ii.}) &= \frac{\sigma^2}{n} \left[\frac{(c-1)^2 + (c-1)}{c^2} \right] \\ &= \frac{\sigma^2}{n} \left(\frac{c-1}{c} \right) \end{aligned} \quad (3.7)$$

Após ter obtido a distribuição amostral da *diferença média do canal i* ($\hat{e}_{ii.}$) fora de controle, torna-se fácil o cálculo da probabilidade de sinalização associada ao canal afetado pela causa especial. Tal probabilidade é determinada analiticamente por:

$$\begin{aligned}
P[(\hat{e}_{i.} > LSC) \cup (\hat{e}_{i.} < LIC)] &= 1 - [P(\hat{e}_{i.} < LSC) - P(\hat{e}_{i.} < LIC)] \\
&= 1 - \left[\Phi \left(\frac{K_{dnb} \sigma \sqrt{\frac{c-1}{nc}}}{\sqrt{\frac{\sigma^2}{n} \left[\frac{\gamma^2 (c-1)^2 + (c-1)}{c^2} \right]}} \right) - \Phi \left(\frac{-K_{dnb} \sigma \sqrt{\frac{c-1}{nc}}}{\sqrt{\frac{\sigma^2}{n} \left[\frac{\gamma^2 (c-1)^2 + (c-1)}{c^2} \right]}} \right) \right] \\
&= 1 - \left[\Phi \left(\frac{K_{dnb} \sqrt{c}}{\sqrt{\gamma^2 (c-1) + 1}} \right) - \Phi \left(\frac{-K_{dnb} \sqrt{c}}{\sqrt{\gamma^2 (c-1) + 1}} \right) \right] \tag{3.8}
\end{aligned}$$

b) Probabilidade de Sinalização por um Canal Diferente do Afetado pela Causa Especial:

A diferença média de qualquer outro canal k que não foi afetado pela causa especial é diretamente obtida por:

$$\hat{e}_{tk.} = x_{tk.} - \hat{b}_t \tag{3.9}$$

Para a obtenção da distribuição amostral da diferença média de qualquer outro canal k não afetado é necessário um cálculo prévio do valor esperado do nível-base estimado em função desses outros canais não afetados pela causa especial. Relembrando o resultado obtido em (2.38), o nível-base estimado pode ser obtido sob a forma de:

$$\hat{b}_t = \frac{1}{c} \sum_{k=1}^c x_{tk.} \tag{3.10}$$

onde:

$$x_{tk.} = b_t + e_{tk.} \tag{3.11}$$

ou, de forma mais direta, (3.10) pode ser ainda reescrita em função das diferenças médias de outros canais k .

$$\begin{aligned}
\hat{b}_t &= \frac{1}{c} \sum_{k=1}^c b_t + e_{tk}. \\
&= b_t + \frac{1}{c} \sum_{k=1}^c e_{tk}. \\
&= b_t + e_{t..}
\end{aligned} \tag{3.12}$$

A equação (3.12) pode ser ainda reescrita mais uma vez, decompondo o somatório $\sum_{k=1}^c e_{tk}$ em duas parcelas, tal como segue:

$$\hat{b}_t = b_t + \frac{1}{c} \left(e_{ti.} + \sum_{\substack{k=1 \\ k \neq i}}^c e_{tk} \right) \tag{3.13}$$

Então, o valor esperado do nível-base estimado é determinado por:

$$\begin{aligned}
E(\hat{b}_t) &= E(b_t) + \frac{1}{c} E \left(e_{ti.} + \sum_{\substack{k=1 \\ k \neq i}}^c e_{tk} \right) \\
&= b_t
\end{aligned} \tag{3.14}$$

Obtido o valor esperado do nível-base estimado torna-se fácil o cálculo analítico do valor esperado das diferenças médias de qualquer outro canal k não afetado. Logo, partindo de (3.9), tem-se que:

$$\begin{aligned}
E(\hat{e}_{tk.}) &= E(x_{tk.} - \hat{b}_t) \\
&= E(x_{tk.}) - E(\hat{b}_t) \\
&= b_t + E(e_{tk.}) - b_t \\
&= 0
\end{aligned} \tag{3.15}$$

Tal como o resultado obtido em (3.3), pode-se comprovar que a ocorrência de causas especiais que afetam a dispersão de um canal i qualquer não interfere no

valor esperado das diferenças médias dos outros k canais não afetados. Contudo, a variância de tais diferenças não tem o mesmo comportamento, pois é afetada pela alteração ocorrida na dispersão das diferenças médias do canal que sofreu alteração no processo.

Então, partindo da equação (3.9) e utilizando os resultados de (3.11) e (3.12), a variância das diferenças médias dos outros k canais não afetados é obtida analiticamente na forma de:

$$\begin{aligned}
 Var(\hat{e}_{tk.}) &= Var(x_{tk.} - \hat{b}_t) \\
 &= Var(b_t + e_{tk.} - b_t - e_{t..}) \\
 &= Var(e_{tk.} - e_{t..}) \\
 &= Var\left(e_{tk.} - \frac{1}{c} \sum_{i=1}^c e_{ti.}\right)
 \end{aligned} \tag{3.16}$$

Decompondo o somatório $\sum_{i=1}^c e_{ti.}$, (3.16) pode ser reescrita na seguinte forma:

$$\begin{aligned}
 Var(\hat{e}_{tk.}) &= Var\left(e_{tk.} - \frac{1}{c} e_{tk.} - \frac{1}{c} \sum_{\substack{i=1 \\ i \neq k}}^c e_{ti.}\right) \\
 &= Var\left(\left(\frac{c-1}{c}\right) e_{tk.} - \frac{1}{c} \sum_{\substack{i=1 \\ i \neq k}}^c e_{ti.}\right)
 \end{aligned} \tag{3.17}$$

Analisando o somatório $\sum_{\substack{i=1 \\ i \neq k}}^c e_{ti.}$ é observado que, dentre os $(c-1)$ termos gerados por ele, apenas um deles tem variância alterada para $\gamma^2 \sigma^2$, pertinente ao único canal que sofreu alteração, enquanto os outros $(c-2)$ termos têm variância inalterada igual a σ^2 .

Logo, considerando a hipótese de que as diferenças médias teóricas são independentes uma das outras ($e_{i.}$ independente de $e_{i.k.}$ para $i \neq k$), (3.17) pode ser reescrita sob a seguinte forma analítica:

$$\begin{aligned} \text{Var}(\hat{e}_{i.k.}) &= \left(\frac{c-1}{c}\right)^2 \frac{\sigma^2}{n} + \left(\frac{1}{c}\right)^2 \left[(c-2) \frac{\sigma^2}{n} + \gamma^2 \frac{\sigma^2}{n} \right] \\ &= \frac{\sigma^2}{nc^2} \left[(c-1)^2 + (c-2 + \gamma^2) \right] \\ &= \frac{\sigma^2}{nc^2} \left[c^2 - c - 1 + \gamma^2 \right] \end{aligned} \quad (3.18)$$

Sintetizando, a distribuição amostral da *diferença média dos outros k canais* não afetados ($\hat{e}_{i.k.}$) na presença de alguma causa especial de variação que afete a dispersão do canal i é determinada por:

$$\hat{e}_{i.k.} \sim N\left(0, \frac{\sigma^2}{nc^2} \left[c^2 - c - 1 + \gamma^2 \right] \right)$$

Ao considerar $\gamma^2 = 1$, a variância obtida em (3.18) será exatamente igual a variância da *diferença média dos outros k canais* ($\hat{e}_{i.k.}$) quando o processo encontra-se em estado de controle.

$$\begin{aligned} \text{Var}(\hat{e}_{i.k.}) &= \frac{\sigma^2}{nc^2} \left[c^2 - c - 1 + \gamma^2 \right] \\ &= \frac{\sigma^2}{n} \left(\frac{c-1}{c} \right) \end{aligned} \quad (3.19)$$

A probabilidade do GCG de diferenças, sob a presença da causa especial considerada, *signalizar em qualquer outro canal k* que não tenha sido afetado pela causa especial é calculada pela seguinte forma:

$$P[(\hat{e}_{i.k.} > LSC) \cup (\hat{e}_{i.k.} < LIC)] = 1 - [P(\hat{e}_{i.k.} < LSC) - P(\hat{e}_{i.k.} < LIC)] \quad (3.20)$$

Após ter obtido a *distribuição amostral da diferença média dos outros k canais não afetados* ($\hat{e}_{i.k.}$), quando o processo está fora de controle, torna-se fácil

calcular analiticamente a probabilidade dada pela equação (3.20). Desenvolvendo, tem-se que:

$$\begin{aligned}
 P[(\hat{e}_{ik.} > LSC) \cup (\hat{e}_{ik.} < LIC)] &= \\
 &= 1 - \left[\Phi \left(\frac{K_{dnb} \sigma \sqrt{\frac{c-1}{nc}}}{\sqrt{\frac{\sigma^2}{nc^2} [c^2 - c - 1 + \gamma^2]}} \right) - \Phi \left(\frac{-K_{dnb} \sigma \sqrt{\frac{c-1}{nc}}}{\sqrt{\frac{\sigma^2}{nc^2} [c^2 - c - 1 + \gamma^2]}} \right) \right] \\
 &= 1 - \left[\Phi \left(\frac{K_{dnb} \sqrt{(c-1)c}}{\sqrt{(c^2 - c - 1 + \gamma^2)}} \right) - \Phi \left(\frac{-K_{dnb} \sqrt{(c-1)c}}{\sqrt{(c^2 - c - 1 + \gamma^2)}} \right) \right] \quad (3.21)
 \end{aligned}$$

Esta é a probabilidade de sinalização de um canal k não afetado pela causa especial de variação, individualmente. Caso haja independência entre os valores $\hat{e}_{ik.}$ para os diversos canais do processo (onde $k = 1, 2, \dots, c; k \neq i$), o número de sinais simultâneos por canais não afetados seria uma variável aleatória binomial, e o cálculo da probabilidade de sinal por algum (pelo menos um) canal k não afetado seria dado por:

$$1 - (1 - P_{\hat{e}_{ik.}})^{c-1} \quad (3.22)$$

onde $P_{\hat{e}_{ik.}}$ é definido pela equação (3.21).

Entretanto, como os valores de $\hat{e}_{ik.}$ são correlacionados, ao utilizar a equação (3.22), a probabilidade de sinal por um ou mais canais não afetados seria superestimada.

Dependendo da conduta do pessoal responsável pelo monitoramento do processo e do tipo de teste utilizado para o diagnóstico do estado do processo na presença de algum sinal, o sinal em um canal não afetado pode ser considerado alarme falso ou alarme verdadeiro. O alarme pode ser considerado verdadeiro nos casos em que a causa especial é descoberta no canal afetado, mesmo que não seja diagnosticada nos canais não afetados, devido à maior investigação realizada por causa do sinal de descontrole.

Na análise dos eventos deve-se ter certa cautela ao definir o que é um sinal de descontrole, pois há diversos tipos de sinais que podem ocorrer ao controlar estatisticamente o processo.

A seguir são desenvolvidos novos gráficos de controle de grupos para o controle da dispersão de um canal específico no processo.

3.2

Gráficos específicos para o controle da dispersão

Nesta seção são detalhadas propostas de gráficos de controle de grupos voltados especificamente para o monitoramento da dispersão em um determinado canal do processo. As propostas diferem conforme a quantidade de amostras extraídas de cada canal. Além disso, para cada caso, há versões Shewhart e EWMA dos gráficos de controle.

Considere inicialmente os casos onde é possível a retirada de mais de uma amostra de cada canal do processo. Embora seja uma situação pouco freqüente na prática, devido a certas restrições operacionais e financeiras dos processos produtivos, são casos que devem ser analisados, pois permitem uma melhora de desempenho substancial.

Contudo, existem casos mais específicos e mais freqüentes na prática, onde existem fortes restrições operacionais e financeiras à extração de amostras dos canais do processo. Em algumas situações específicas, quando a inspeção do sistema produtivo é automatizada, ou quando a taxa de produção é muito lenta, deve ser extraída apenas uma observação de cada canal (observações individuais) para não interromper o processo, ou gerar custos maiores de produção. Nesses casos mais específicos, para os gráficos de controle de Shewhart, usualmente a estatística de controle considerada é o próprio valor amostrado do processo.

3.2.1

GCG versão Shewhart

Quando é possível a extração de mais de uma amostra ($n > 1$) de cada um dos canais do processo, as duas estatísticas de controle mais utilizadas para o monitoramento da dispersão são a amplitude amostral (R) e o desvio-padrão (S).

A estatística da variância amostral (S^2) é a menos utilizada, devido à dificuldade de interpretação pelo pessoal responsável pelo controle do processo. Na prática, o gráfico de controle da amplitude amostral é mais utilizado que o gráfico de controle do desvio-padrão e da variância amostral, devido à facilidade de implementação e interpretação no ambiente produtivo.

Embora o uso da estatística de amplitude amostral seja mais difundido que o uso da estatística de desvio-padrão e da estatística de variância amostral para o controle estatístico de processos, esse comportamento não se justifica mais, dada a maior eficiência do gráfico de controle de S (e de S^2) contra alterações na dispersão do processo e o fato de que não há mais dificuldade operacional (esforço computacional empregado) em se calcular S ou S^2 .

Nesta tese é proposta a adoção da estatística S^2 em um GCG para o controle de processos, onde é possível a retirada de mais de uma amostra por canal. Embora esta estatística não seja a mais difundida (a usualmente adotada no controle de processos), o GCG de S^2 é análogo ao GCG de S , ao trabalhar com limites de probabilidade em ambos os gráficos. A estatística de S não é aqui utilizada porque, embora com limites de controle de 3σ os gráficos de S e de S^2 tenham desempenho diferente (até pelo fato de apresentarem diferentes curtoses), no caso de gráficos de controle de grupos é preciso trabalhar com limites de probabilidade para controlar o valor de α_{global} e, neste caso, os gráficos de S^2 e S possuem igual desempenho, sendo o LSC de um igual ao quadrado do LSC do outro.

Quando não é possível a retirada de mais de uma amostra por canal, outra estatística de controle deve ser utilizada para o controle da dispersão das componentes individuais de variação, pois não há como obter a estatística S^2 . Woodall (2000, pág. 347) afirma que a estatística de amplitude móvel (MR) é usualmente adotada na literatura para o controle da dispersão de processos univariados.

Sullivan e Woodall (1996) e Rigdon et al. (2004), entre outros autores, mostram que a amplitude móvel é pouco efetiva para o controle da dispersão em processos univariados. Contudo, Roes et al. (1993), Adke and Hong (1997) chegam a conclusão oposta, afirmando que o gráfico de controle baseado na amplitude móvel contém informação útil sobre os aumentos na dispersão do

processo. Amin e Ethridge (1998) mostram que, quando associado a um gráfico de controle de observações individuais, o gráfico de controle de MR pode ajudar na detecção de aumentos na dispersão.

Esta tese propõe um GCG da amplitude móvel das diferenças em relação ao nível-base (MR DNB) na versão Shewhart para o controle da dispersão da componente individual dos canais de um PMC (no caso de monitoramento do processo por observações individuais).

A seguir o GCG de S^2 e o GCG de MR DNB são detalhados.

3.2.1.1

Dados subgrupados: GCG de S^2

Esta subsecção descreve o gráfico de controle de grupos de S^2 na versão Shewhart, que será aplicado diretamente às observações de cada canal do processo, uma vez que a variância amostral das observações de um canal é igual à variância das diferenças destas observações em relação ao nível-base do processo.

Considere a extração de mais de uma medida (i.e., subgrupo de tamanho $n > 1$) por canal em cada instante de amostragem t . A *variância amostral no i -ésimo canal no instante t* é dada por:

$$S_{ii}^2 = \frac{1}{n-1} \sum_{j=1}^n (x_{tij} - x_{ti.})^2 \quad (3.23)$$

onde

x_{tij} é a j -ésima medida obtida no canal i no instante t ;

$x_{ti.}$ é o valor médio das medidas obtidas no canal i no instante t .

Como o interesse é detectar aumentos na dispersão de $\hat{\sigma}_{ii}$, o esquema proposto só possui um limite (superior) de controle. Este é dado por:

$$LSC_{GCGS^2} = \frac{\sigma^2}{n-1} \chi_{1-\alpha_{individual}, n-1}^2 \quad (3.24)$$

onde $\alpha_{individual}$ é determinado através da correção de Dunn-Sidak (Dunn, 1958; Sidak, 1967) de maneira que a probabilidade de alarme falso global do GCG seja igual a um valor de α_{global} especificado. Relembrando, o cálculo do $\alpha_{individual}$ obtido na equação (2.37) mostra que

$$\alpha_{individual} = 1 - \sqrt[c]{1 - \alpha_{global}} \quad (2.37)$$

A correção de Dunn-Sidak garante a probabilidade de alarme falso α_{global} especificado, pois as variâncias amostrais dos diversos canais são independentes umas das outras.

Se a variância amostral máxima obtida dentre todos os canais for maior que o limite superior de controle definido em (3.24), há forte evidência de descontrole na dispersão de alguma das componentes individuais dos canais do processo. Ou seja, o processo é considerado fora de controle se

$$\max_{i=2,3,\dots,c} (S_{ii}^2) > LSC_{GCGS^2}$$

Tal como foi mencionado anteriormente, é indiferente, para o desempenho do GCG, usar como estatística de controle S^2 ou S (os valores de NMA_1 são idênticos). A única transformação necessária para utilizar o GCG de S (que pode ser mais aceito no meio industrial) está nas equações (3.23) e (3.24), onde deverá ser calculada a raiz quadrada da estatística e do limite superior de controle de S^2 . E, de forma análoga, o máximo dos desvios-padrão ($\max(S_{ii})$, para $i = 2,3,\dots,c$) é comparado ao limite superior de controle, para verificação do estado do processo (se está em controle ou fora de controle).

3.2.1.2

Observações individuais: GCG de MR

Uma componente de variação do processo é a variação do nível-base. Trabalhar diretamente com um GCG da amplitude móvel das observações individuais de cada canal teria o inconveniente de que essa amplitude incorporaria

também a variação do nível-base, reduzindo a sensibilidade do gráfico de controle a aumentos na variação das componentes individuais.

Para eliminar tal componente de variação e assim isolar as componentes individuais, a proposta desta tese, para o caso de observações individuais ($n=1$), é a utilização de um GCG da *amplitude móvel das diferenças em relação ao nível-base*. A amplitude móvel das diferenças é estimada por:

$$MR_{ti} = |\hat{e}_{ti} - \hat{e}_{(t-1)i}| \quad (3.25)$$

As diferenças em relação ao nível-base são distribuídas normalmente com média igual a zero e variância igual a $(c-1/c)\sigma^2/n$. É bem conhecido na literatura que, de modo geral, $E(MR_t) = d_2\sigma$ e $\sigma(MR_t) = d_3\sigma$, onde σ é o desvio-padrão do processo em controle. Então, como $Var(\hat{e}_{ti}) = (c-1/c)\sigma^2/n$, tem-se que:

$$\begin{aligned} E(MR_t) &= d_2\sigma(\hat{e}_{ti}) \\ &= d_2\sqrt{\left(\frac{c-1}{c}\right)}\sigma_0 \end{aligned} \quad (3.26)$$

$$\begin{aligned} DP(MR_t) &= d_3\sigma(\hat{e}_{ti}) \\ &= d_3\sqrt{\left(\frac{c-1}{c}\right)}\sigma_0 \end{aligned} \quad (3.27)$$

onde as constantes d_2 e d_3 são definidas para $n=2$.

A distância do limite (superior) de controle à linha média (dada por $d_2\sigma(\hat{e}_{ti})$) pode ser expressa como um múltiplo de K_{mr} do desvio-padrão de MR_t , onde K_{mr} é o fator de abertura do limite superior de controle, obtido por simulação para fornecer o NMA_0 especificado.

Portanto, o limite de controle do GCG de MR das diferenças é:

$$\begin{aligned} LSC_{GCGMR} &= d_2\sigma(\hat{e}_{ti}) + K_{MR}d_3\sigma(\hat{e}_{ti}) \\ &= (d_2 + K_{MR}d_3)\sigma\sqrt{\left(\frac{c-1}{c}\right)} \end{aligned} \quad (3.28)$$

É importante lembrar que, os e_{ij} 's não são diretamente observáveis, então, para estimativa de σ , ao se calcular o desvio-padrão dos resíduos (diferenças em relação ao nível-base estimado) de dados históricos através de qualquer estimador que seja, estar-se-á estimando diretamente $\sigma\sqrt{(c-1)/c}$, e não σ . Portanto, o uso do fator $\sqrt{(c-1)/c}$ multiplicando σ para o cálculo dos limites de controle de qualquer gráfico que utilize as diferenças em relação ao nível-base se aplica ao caso aqui presente em que se geram valores simulados de e_{ij} com um desvio-padrão σ .

Considera-se que algum canal do processo está fora de controle se

$$\max_{i=2,3,\dots,c}(MR_{ii}) > LSC_{GCG MR}$$

3.2.2

GCG versão EWMA

Os gráficos de controle na versão Shewhart são lentos para sinalizar alterações de magnitude pequena a moderada nos parâmetros do processo. Esquemas EWMA são usualmente utilizados na literatura para o controle de processos, quando o objetivo é adquirir maior sensibilidade a tais alterações.

Como será visto, no caso de extração de mais de uma amostra de cada canal do processo, o GCG de S^2 é mais eficiente para detectar sinais de descontrole na dispersão das componentes individuais do que o GCG de diferenças em relação ao nível-base, e do que os esquemas de controle de R_t de Mortell e Runger (1995) e de MEWMA de S^2 de Runger et al. (1996). Contudo, ainda se pode esperar que, para o caso de alterações de pequena magnitude na dispersão de um canal em particular do processo, esse desempenho possa ser melhorado, em termos de reduções nos valores de NMA_1 , pela incorporação de um procedimento EWMA.

Devido a certas peculiaridades dos esquemas de EWMA para o controle da dispersão, em vez de se aplicar o amortecimento diretamente à estatística S^2 , é mais conveniente (mais eficiente aplicá-lo) à estatística $\ln(S^2)$. Portanto é desenvolvido o GCG de EWMA de $\ln(S^2)$, como será visto na subseção 3.2.2.1.

Para os casos de extrações de observações individuais ($n=1$) de cada fluxo do processo, também é proposto e analisado nesta tese um GCG de EWMA de MR das diferenças em relação ao nível-base, visando obter melhor desempenho para os casos de alterações de pequena magnitude na dispersão de um canal em particular do processo. Tal gráfico de controle é desenvolvido na subseção 3.2.2.2.

3.2.2.1

Dados subgrupados: GCG de EWMA de $\ln(S^2)$

Na teoria clássica de gráficos de controle, alguns trabalhos propuseram utilizar esquemas de controle de EWMA para o monitoramento univariado da dispersão. Entre os esquemas univariados desenvolvidos poderiam ser destacados os seguintes: MacGregor e Harris (1993) propõem os esquemas EWMV (*Exponentially Weighted Moving Variance*) e EWMS (*Exponentially Weighted Mean Square Deviation*), Amin e Li (2000) propõem o esquema MaxMin EWMA, Castagliola (2000, 2005a e 2005b) propõe um esquema S-EWMA (para o controle do desvio-padrão), um esquema S^2 -EWMA (para o controle da variância) e um esquema R-EWMA (para o controle da amplitude amostral). Além disso, encontram-se na literatura algumas formas adaptativas⁶ para os esquemas de R-EWMA (Castagliola et al., 2006) e de S^2 -EWMA (Castagliola et al., 2007 e 2008).

No entanto, ainda não há trabalhos na literatura que proponham tais esquemas estatísticos para o controle da dispersão em processos multicanal.

Como visto na subseção 3.2.1.1, um dos esquemas de controle propostos (e analisados) nesta tese é um GCG de S^2 , na versão Shewhart, para o controle da dispersão de um canal em particular do processo. Contudo, com a correção de Dunn-Sidak aplicada à probabilidade de alarmes falsos, o limite superior de controle calculado de S^2 torna-se maior que o usualmente adotado num gráfico de controle de Shewhart clássico, o que, por consequência, reduz a sensibilidade do esquema a alterações de pequena magnitude na dispersão das componentes individuais do processo.

⁶ Que se permite variar os parâmetros de projeto do gráfico, tal como o tamanho da amostra ou o intervalo de amostragem, entre dois ou mais valores, em função da informação amostral mais recente.

Para garantir uma maior eficiência (valores menores de NMA_1) contra pequenos aumentos na dispersão de um canal, um GCG de EWMA de $\ln(S^2)$ é proposto e analisado.

O trabalho de Crowder e Hamilton (1992) fornece um esquema estatístico promissor para o controle da dispersão em processos univariados, ao utilizar um esquema EWMA unilateral superior (somente considera o limite superior de controle para sinalização de possíveis deteriorações na qualidade do processo) da variância amostral linearizada, $\ln(S^2)$. O esquema EWMA unilateral é um caso especial do gráfico de controle generalizado proposto por Champ et al. (1991).

Box et al. (1978) e outros autores recomendam fortemente o uso da variância amostral linearizada quando são realizadas inferências sobre variâncias de dados normalmente distribuídos. As razões para tal recomendação estão no fato de que tal transformação logarítmica na variância amostral, $\ln(S^2)$, obter estatísticas mais normalmente distribuídas do que as próprias variâncias amostrais (S^2). Crowder e Hamilton (1992) mostram que a variância de $\ln(S^2)$ é independente de σ^2 e depende somente do tamanho da amostra n . Mostram também que um aumento no parâmetro de escala da distribuição gamma (um aumento em σ^2) corresponde a um aumento no parâmetro de locação da distribuição log-gamma correspondente (a distribuição de $\ln(S^2)$).

Outras transformações poderiam ser aplicadas à variância amostral, tal como as transformações Box-Cox, para o controle da dispersão. Contudo, a transformação logarítmica é adotada por causa da simplicidade e das características acima listadas. A transformação logarítmica também foi utilizada na elaboração de esquemas para o controle de processos univariados por outros autores tais como Castagliola (2005a e 2005b) e Castagliola et al. (2008).

A estatística de controle EWMA da *variância linearizada em cada canal* é expressa pelo seguinte modelo:

$$Z_{\ln_i(t)} = \max\left[\left(\lambda \ln(S_{ii}^2) + (1 - \lambda)Z_{\ln_i(t-1)}\right), \ln(\sigma_0^2)\right] \quad (3.29)$$

onde:

S_{ii}^2 é a variância amostral do canal, dada pela equação (3.23);

A mesma constante de amortecimento (λ) é usada para todos os canais do processo.

Como o esquema de controle de EWMA de $\ln(S_{ii}^2)$ é projetado especialmente para detecção de aumentos na dispersão (não pretendendo ser um instrumento para a detecção de reduções da mesma), os autores propõem, para aumento da sensibilidade do gráfico, reinicializar o valor da estatística sempre que EWMA de $\ln(S_{ii}^2)$ for menor que $\ln(\sigma_0^2)$. Esta metodologia também é adotada no gráfico de controle de grupos de EWMA de $\ln(S^2)$ desenvolvido nesta tese.

A “barreira refletora” em $\ln(\sigma_0^2)$, como definida em (3.29) é utilizada para evitar o problema de inércia do gráfico de controle. Esta “barreira” serve apenas para os esquemas de controle de EWMA, pois com os gráficos de controle na versão de Shewhart ela não alteraria em nada o desempenho do gráfico, i.e., não reduzia os valores de NMA_1 , já que a estatística amostral não possui dependência serial e, portanto, não possui inércia.

Crowder e Hamilton (1992) demonstram, utilizando expansões de séries infinitas (veja Abramowitz e Stegun, 1965), que a variância de $\ln(S^2)$ é dada por:

$$Var[\ln(S_{ii}^2)] = \left[\frac{2}{n-1} + \frac{2}{(n-1)^2} + \frac{4}{3(n-1)^3} - \frac{16}{15(n-1)^5} \right] \quad (3.30)$$

Segundo Crowder e Hamilton (1992), a variância amostral da estatística EWMA de $\ln(S^2)$ não limitada pela “barreira refletora” pode ser utilizada para obter o limite superior de controle do gráfico. Sob a hipótese de independência de $\ln(S^2)$, a variância amostral de $Z_{\ln_i(t)}$ pode ser expressa como uma soma dos termos de uma progressão geométrica, utilizando a variância amostral definida em (3.30). Logo, esta variância é dada por:

$$\begin{aligned} Var[Z_{\ln_i(t)}] &= \left(\frac{\lambda}{2-\lambda} \right) [1 - (1-\lambda)^{2t}] Var[\ln(S_{ii}^2)] \\ &= \left(\frac{\lambda}{2-\lambda} \right) [1 - (1-\lambda)^{2t}] \left[\frac{2}{n-1} + \frac{2}{(n-1)^2} + \frac{4}{3(n-1)^3} - \frac{16}{15(n-1)^5} \right] \end{aligned} \quad (3.31)$$

Portanto, de forma análoga ao realizado em (2.6), o limite superior de controle do GCG de EWMA $\ln(S^2)$ representado em função da linha média (expressa por $\ln(\sigma_0^2)$) e da forma assintótica da variância obtida em (3.31) é o seguinte:

$$\begin{aligned} LSC_{GCGZ_{\ln}} &= \ln(\sigma_0^2) + K_{\ln,s} \sigma(Z_{\ln_i(t)}) \\ &= \ln(\sigma_0^2) + K_{\ln,s} \sqrt{\left(\frac{\lambda}{2-\lambda}\right) \left[\frac{2}{n-1} + \frac{2}{(n-1)^2} + \frac{4}{3(n-1)^3} - \frac{16}{15(n-1)^5} \right]} \end{aligned} \quad (3.32)$$

onde:

$K_{\ln,s}$ é o fator de abertura do limite superior de controle, obtido por simulação, que fornece o NMA_0 especificado.

A rigor, a barreira refletora altera a distribuição da estatística $Z_{\ln_i(t)}$, fazendo com que a sua variância difira da variância que ela teria sem a barreira, dada por (3.30). Isso, porém, não invalida o uso de (3.32) para o LSC, dado que o fator $K_{\ln,s}$ é obtido por simulação para fornecer o NMA_0 desejado.

Se o valor máximo da estatística EWMA dada por (3.29), dentre todos os canais, for maior que o limite superior de controle dado por (3.32), há forte evidência de descontrole na dispersão de alguma das componentes individuais dos canais do processo. Ou seja, algum canal do processo é considerado fora de controle se:

$$\max_{i=2,3,\dots,c} (Z_{\ln_i(t)}) > LSC_{GCGZ_{\ln}}$$

Os gráficos de controle de grupos de S^2 e de EWMA de $\ln(S^2)$ requerem a extração de amostras de tamanho maior que um ($n > 1$) por canal, para que se possa calcular S^2 . Contudo, existem casos específicos, e bem frequentes em situações práticas, em que fortes restrições operacionais e econômicas permitem somente a obtenção de observações individuais em cada canal. Para tais situações, nesta tese é proposto um GCG de EWMA da amplitude móvel das diferenças em relação ao nível-base, que é tratado na próxima subseção.

3.2.2.2

Observações individuais: GCG de EWMA de MR

Esta subsecção descreve o GCG de EWMA da amplitude móvel (MR) das diferenças em relação ao nível-base, visando garantir detecções mais rápidas contra alterações menores na dispersão ($\gamma \leq 2$) das componentes individuais (e_{ti}) dos canais do processo.

Como o objetivo aqui continua sendo a detecção de aumentos na dispersão, no esquema de controle proposto, tal como no caso do GCG de EWMA de $\ln(S^2)$, não há um limite inferior de controle e, além disso, uma barreira refletora é utilizada para reinicializar o valor de EWMA de MR sempre que o mesmo for menor que a linha média do gráfico, dada por $d_2\sigma(\hat{e}_{ti})$, onde $\sigma(\hat{e}_{ti})$ é o desvio-padrão das diferenças em relação ao nível-base.

A estatística de controle EWMA da amplitude móvel das diferenças é dada por:

$$\begin{aligned} Z_{MRi} &= \max\left[\left(\lambda MR_{ti} + (1-\lambda)Z_{MRi(t-1)}\right), d_2\sigma(\hat{e}_{ti})\right] \\ &= \max\left[\left(\lambda MR_{ti} + (1-\lambda)Z_{MRi(t-1)}\right), d_2\sqrt{\left(\frac{c-1}{c}\right)}\sigma_0\right] \end{aligned} \quad (3.33)$$

onde

MR_{ti} é a dada por (3.25);

d_2 é uma constante definida para $n=2$.

A função da “barreira refletora” em $d_2\sigma(\hat{e}_{ti})$, definida em (3.33) é evitar o problema de inércia do gráfico de controle. Se esta “barreira” fosse utilizada no GCG de MR na versão Shewhart, não alteraria os valores de NMA_1 , apenas poderia esconder a informação do verdadeiro valor de MR, quando este estivesse abaixo do limite superior de controle. O uso de tal “barreira” na versão Shewhart poderia truncar ou censurar os dados do processo, sem nenhum benefício adicional.

Portanto, de forma similar ao que foi feito em (2.6), o limite superior de controle do GCG de EWMA MR é dado por:

$$\begin{aligned} LSC_{GCG ZMR} &= d_2 \sigma_{\hat{e}i.} + K_{ZMR} d_3 \left(\frac{\lambda}{2-\lambda} \right) \sigma_{\hat{e}i.} \\ &= (d_2 + K_{ZMR} d_3) \sigma_0 \sqrt{\left(\frac{\lambda}{2-\lambda} \right) \left(\frac{c-1}{c} \right)} \end{aligned} \quad (3.34)$$

onde:

σ_0 é o desvio-padrão do processo em controle;

d_3 é uma constante definida para $n=2$.

K_{ZMR} é o fator de abertura do limite superior de controle, obtido por simulação, que fornece o NMA_0 especificado.

É verdade que o limite superior de controle para um valor de λ específico poderia ser obtido diretamente por uma busca por simulação, para fornecer o NMA_0 específico. Contudo, a constante $\lambda/(2-\lambda)$ é utilizada para a representação dos limites de controle por já ser utilizada na literatura por vários autores consagrados.

Na verdade, nem no caso de controle por variáveis independentes e normalmente distribuídas, o parâmetro K poderia ser determinado pelo quantil da distribuição normal, cuja probabilidade de a estatística ultrapassar o limite de controle corresponda ao inverso do NMA_0 desejado, devido à autocorrelação da série das estatísticas EWMA.

Se o valor máximo da estatística Z_{MRt} obtido em (3.33), dentre todos os canais, for maior que o limite superior de controle definido em (3.34), há forte evidência de descontrole na dispersão de alguma das componentes individuais dos canais do processo. Ou seja, algum canal do processo é considerado fora de controle se:

$$\max_{i=2,3,\dots,c} (Z_{MRt}) > LSC_{GCG ZMR}$$

Os resultados de desempenho para os GCG na versão Shewhart são mostrados no próximo capítulo e para os GCG na versão EWMA estão no Capítulo 7.