

# Estudos dos padrões de não aleatoriedade dos gráficos de controle de *Shewhart*: um enfoque probabilístico

Quintiliano Siqueira Schroden Nomelini

Eric Batista Ferreira

Marcelo Silva de Oliveira



## Resumo

*Além de investimento em novas tecnologias e modernização do parque industrial, a competitividade dos produtos brasileiros está ligada ao aumento de produtividade. O Controle Estatístico de Processo (CEP), junto com os gráficos de controle, inseriu-se neste contexto no que diz respeito ao controle das características significativas do produto e do processo, em tempo real, garantindo níveis de qualidade, a um custo exigido pelo mercado. Um procedimento importante é a interpretação dos gráficos de controle. Um estudo foi feito para ilustrar e discutir padrões de não aleatoriedade, que auxiliam na interpretação de tais gráficos. Alguns padrões de não aleatoriedade foram ilustrados e, por meio de cálculo de probabilidade, mostrou-se quais deles rejeitaram a hipótese de controle estatístico, a um nível de significância pré-especificado. Concluiu-se que os padrões mencionados na literatura não eram tão informativos. No que diz respeito ao nível de significância, observou-se então que em alguns padrões a um certo nível de significância retornaram resultados diferentes dos mencionados na literatura.*

**Palavras-chave:** Controle Estatístico de Processo – CEP. Padrões de não aleatoriedade. Processo estável. Causas especiais. Gráfico de controle.

## 1 Introdução

Nas últimas duas décadas, o cenário competitivo brasileiro foi alterado em função da exposição dos consumidores a produtos de melhor qualidade e menor custo. Além de investimento em novas tecnologias e modernização do parque industrial, a competitividade dos produtos brasileiros está ligada ao aumento de produtividade, com o uso de ferramentas estatísticas de controle de processo (TORMINATO, 2004).

Segundo Montgomery (2004), o processo de fabricação deve ser estável e todos os indivíduos envolvidos no processo (incluindo operadores, engenheiros, pessoal da garantia da qualidade e gerência) devem procurar, continuamente, melhorar o desempenho do processo e reduzir a variabilidade nos parâmetros-chave. O CEP é uma ferramenta primordial para a obtenção desse objetivo e considerada uma das mais importantes na metodologia Seis Sigma.

O CEP é uma ferramenta simples e sua efetividade é testemunhada por uma repetição fisicamente estabelecida nas indústrias por todo o mundo. Por meio dela, consegue-se controlar características significativas do produto e do processo, em tempo real, garantindo níveis de qualidade,

a um custo suportado pelo mercado. O Controle Estatístico de Processo é, sem dúvida, uma das mais poderosas metodologias desenvolvidas, visando auxiliar no controle eficaz da qualidade do produto e seus processos produtivos, usando a estatística como metodologia para analisar as limitações do processo.

Entre as principais ferramentas do CEP, os gráficos de controle se destacam. Eles estão entre os mais utilizados no que diz respeito ao monitoramento de processos. Segundo Werkema (1995), a interpretação de gráficos de controle é feita por meio de padrões de não aleatoriedade, por isso, este trabalho se propõe a ilustrar e discutir tais padrões, pautando-se em conceitos estatísticos e probabilísticos.

## 2 Material e métodos

Montgomery (2004) afirma que, em qualquer processo de produção, independentemente de quão bem planejado ou cuidadosamente mantido ele seja, certa quantidade de variabilidade inerente ou natural sempre existirá.

A contribuição de Deming (1990) neste campo foi muito importante; ele sempre destacou a importância de se saber distinguir as causas comuns das causas especiais. A causa aleatória (ou comum) de variação é uma fonte de variação imprevisível, inerente ao processo, que afeta todos os valores individuais de uma determinada característica deste, sob medição. Causas atribuíveis (especiais, assinaláveis ou ainda identificáveis) de variação constituem um fator identificável oriundo de eventos passageiros, que gera variações não aleatórias (padrões anormais dos dados) que afetam o processo de maneira previsível. A remoção completa de causas especiais de variação deve ser feita para que se tenha um processo estável, condição essencial para a implementação de melhoria neste processo. A identificação das causas especiais e o monitoramento para prevenção é feito por meio do CEP e sua principal ferramenta, os gráficos de controle, com a identificação dos padrões de não aleatoriedade.

Segundo Montgomery (2004), o controle estatístico de processo (CEP) é uma poderosa coleção de ferramentas de resolução de problemas, sendo útil na obtenção da estabilidade do processo e na melhoria da capacidade por meio da redução da variabilidade.

O gráfico de controle (GC), também conhecido como gráfico de controle de Shewhart, é uma técnica estatística usada para medição e análise do comportamento de um processo. O GC é um gráfico que consiste numa linha central (LC), um limite inferior de controle (LIC) e um limite superior de controle (LSC), e valores do parâmetro de interesse (uma característica do processo) grafados sequencialmente ao longo do tempo, que representam o estado atual de um processo.

Segundo Lourenço Filho (1976), (apud ARIENTE et al. 2005), a verificação de que o processo está ou não sob controle é feita pelo exame de amostras extraídas periodicamente. No processo sob controle, a característica de qualidade do conjunto dos itens produzidos possui distribuição normal. Quando a variabilidade se torna anormal, as amostras indicarão que o processo de fabricação se modificou e ficou fora de controle.

Segundo Costa et al. (2004), nos estudos de gráficos de controle de Shewhart é necessário que os valores observados da variável monitorada sejam independentes. Em consequência, espera-se que as variáveis sejam estacionárias e não correlacionadas, para se ter eficiência em detectar condições de fora de controle de modo confiável (MONTGOMERY, 2004).

De certa maneira, o gráfico de controle é o teste da hipótese de que o processo está em estado de controle estatístico, hipótese nula, ou está fora de controle estatístico, hipótese alternativa. A interpretação dos gráficos de controle e a definição do momento em que o processo se encontra fora de controle são feitas por meio do exame da ocorrência (ou não) dos padrões de não aleatoriedade.

Alguns métodos para procurar padrões de não aleatoriedade, que indicam falta de controle, segundo Werkema (1995), são:

- **Pontos fora dos limites de controle:** esta é a indicação mais evidente de falta de controle de um processo, exigindo investigação imediata da causa de variação assinalável responsável pela sua ocorrência. Estes podem vir de resultados de erros de registro dos dados, de cálculos ou de medição ou, ainda, de algum instrumento descalibrado, de um erro do operador ou de defeitos nos equipamentos.
- **Padrões cíclicos ou de periodicidade:** acontecem quando os pontos, repetidamente, apresentam uma tendência para cima e para baixo, em intervalos de tempo que têm, aproximadamente, a mesma amplitude. Algumas causas especiais que podem acarretar em periodicidade são: mudanças sistemáticas nas condições ambientais, cansaço do operador, entre outros.
- **Sequência ou deslocamento de nível do processo:** é uma configuração em que vários pontos consecutivos do gráfico de controle aparecem em apenas um dos lados da linha média. As sequências consideradas anormais são: sete ou mais pontos consecutivos; uma sequência com menos de sete pontos consecutivos, em que pelo menos dez de onze pontos consecutivos aparecem do mesmo lado da linha média; pelo menos doze de quatorze pontos consecutivos aparecem em um mesmo lado da linha média; e pelo menos dezesseis de vinte pontos consecutivos aparecem em um mesmo lado da linha média.
- **Tendência:** é constituído por um movimento contínuo dos pontos do gráfico de controle em uma direção ascendente ou descendente.
- **Mistura ou aproximação dos limites de controle:** é quando os pontos tendem a cair próximo ou levemente fora dos limites de controle, com relativamente poucos pontos próximos da linha média. Neste caso, podem existir duas distribuições sobrepostas, por exemplo, duas máquinas trabalhando de maneira diferente.
- **Estratificação ou aproximação da linha média:** nesse caso, a maioria dos pontos está próximo da linha média, apresentando uma variabilidade menor do que a esperada. Pode ter ocorrido erro nos cálculos dos limites de controle ou que os subgrupos racionais (amostras) foram formados de maneira inadequada. Portanto, a aproximação da linha média não significa estar sob controle, mas, sim, a mistura de dados provenientes de populações distintas.

### 3 Resultados e discussão

Segue abaixo uma forma de mostrar os cálculos probabilísticos para cada um dos padrões de não aleatoriedade comentados anteriormente.

### 3.1 Pontos fora dos limites de controle

O processo de exame é feito da seguinte forma: o operador retira uma amostra de tamanho  $n$ , a uma determinada frequência. Por exemplo, para o gráfico de controle de médias, são plotadas as médias das  $n$  amostras. Logo após, verifica-se se o ponto plotado ou se a média está fora dos limites de controle; caso esteja, então, o processo está fora de controle e medidas de ações corretivas devem ser tomadas.

Esta situação pode ser enunciada estatisticamente com o espaço amostral sendo igual a  $\Omega = \{D, F\}$ , em que  $D$  é o ponto “estar dentro” dos limites de controle e  $F$  é “estar fora”. Dessa forma, pode-se pensar em modelar uma distribuição de Bernoulli, em que a variável aleatória e a função densidade de probabilidade são (Equações 1 e 2):

$$X = \begin{cases} 0, & \text{se dentro;} \\ 1, & \text{se fora;} \end{cases} \quad (1)$$

$$P(X = x) = \begin{cases} (1-p), & \text{para } x=0; \\ p, & \text{para } x=1; \\ 0 & \text{para outros valores de } x. \end{cases} \quad (2)$$

As hipóteses a serem testadas são as seguintes (Equação 3):

$$\begin{cases} H_0 : \text{Processo sob controle;} \\ H_1 : \text{Processo fora de controle;} \end{cases} \quad (3)$$

Considerando-se a hipótese de normalidade e os limites de 3-sigma, a probabilidade no gráfico de média de que pontos estejam aleatoriamente fora dos limites de controle é de, aproximadamente,  $P(X = 1) = p = 0.27\%$ , pois (Equações 4 e 5):

$$Z_{\mu \pm 3\sigma} = \frac{(\mu \pm 3\sigma) - \mu}{\sigma} = \pm 3; \quad (4)$$

$$P(\text{fora dos limites } 3\sigma) = P(Z < -3) + P(Z > 3) = 2 \cdot 0,00135 = 0,0027; \quad (5)$$

Então, a probabilidade de se cometer o erro tipo I, rejeitar  $H_0$ , verdadeira, é muito pequena. Logo, se pode rejeitar  $H_0$ , pois o  $p$ -valor é menor que 5%.

Estatisticamente, pode-se enunciar esta situação com espaço amostral que leve em consideração todas as amostras já observadas, até que uma observação esteja fora dos limites de controle. Este espaço pode ser expresso como (Equação 6):

$$\Omega = \left\{ \begin{array}{c} DF \\ DDF \\ DDDF \\ \dots \end{array} \right\} \quad (6)$$

Associando-se a uma variável aleatória  $X$ , que é o número de pontos até que ocorra um ponto fora dos limites, com distribuição geométrica, a função densidade de probabilidade é (Equação 7):

$$P(X = k) = p(1-p)^{k-1}, \quad (7)$$

em que  $p$  é a probabilidade do ponto estar fora dos limites de controle.

Então, para  $k = 2, 3, 4, 5$ , tem-se (Equação 8):

$$\begin{aligned} P(X = 2) &= 0,0027 (0,9973)^1 = 0,0027; \\ P(X = 3) &= 0,0027 (0,9973)^2 = 0,0027; \\ P(X = 4) &= 0,0027 (0,9973)^3 = 0,0027; \\ P(X = 5) &= 0,0027 (0,9973)^4 = 0,0027. \end{aligned} \quad (8)$$

Logo, a probabilidade de rejeitar a hipótese  $H_0$ , verdadeira, é igual à do caso anterior.

### 3.2 Padrões cíclicos ou periodicidade

Esta situação não é tão simples quanto a anterior, podendo-se pensar na utilização da ferramenta séries temporais para estudar um gráfico de controle e verificar se há ocorrência de sazonalidade, que é uma ciclicidade. Mas, para se ter bons resultados com esta ferramenta, deve-se utilizar sequências relativamente grandes que, em termos práticos, não é boa ideia, pois, levará muito tempo para detectar tal padrão. Então, para ilustrar essa situação, simulou-se um caso de ciclicidade absoluta (sem erros aleatórios), em amostras crescentes, para se verificar o comportamento das decisões do teste de normalidade de Shapiro-Wilk, a 5% de significância.

No software R, versão 2.4.0, simularam-se pontos vindos de uma função seno, que é um exemplo de padrão cíclico, cujos argumentos variavam como valores discretos de 1 a 100. Noventa e oito amostras foram compostas, seus tamanhos variando de 3 a 100. A menor amostra analisada tinha tamanho três, devido a uma exigência do teste. O teste de normalidade foi aplicado a todas as amostras para se verificar em que ponto detectava-se a falta de normalidade. Observou-se, então, que até dezesseis pontos a normalidade não era detectada (Figura 1). A linha pontilhada é o limiar de 5%, ou seja, para o teste utilizado a 5%, são necessários dezesseis pontos cíclicos para se poder afirmar que o processo está fora de controle. Na prática, esperar por essa quantidade de amostras para descobrir que o processo está fora de controle não é o ideal: imagine-se um processo em que o intervalo de amostragem é de meia hora; então, somente após oito horas e meia verificar-se-ia que o processo não estava sob controle. Além disso, este teste não é muito robusto para pequenos tamanhos de amostras, como pode ser visto nos  $p$ -valores. Sugere-se, então, uma discussão melhor sobre o tema, como simular pontos da função seno com diferentes períodos e também comparar resultados de diferentes testes de normalidade, como, por exemplo, o

teste de Shapiro-Wilk multivariado com as aproximações de Royston (1992 e 1993), que tornou o teste univariado mais robusto, segundo Ferreira e Dias (2006).

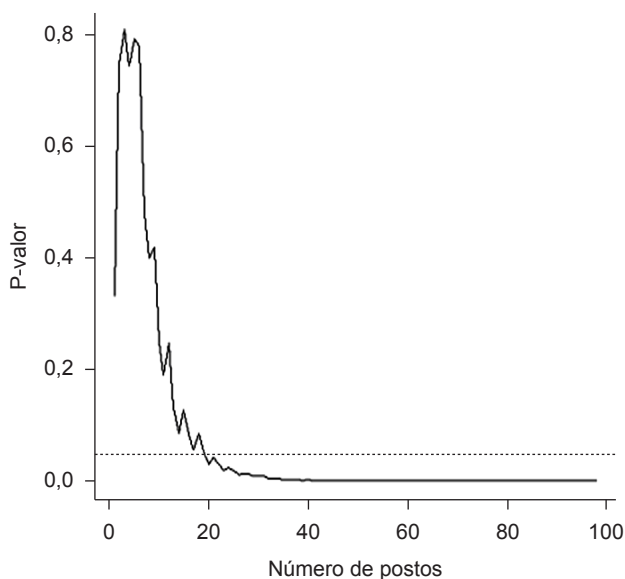
### 3.3 Sequência ou deslocamento de nível do processo

Pode-se observar neste padrão de não aleatoriedade uma familiaridade com algumas estatísticas não paramétricas, como por exemplo, testes dos sinais e teste de aleatoriedade. Este último verifica se o conjunto de números tem uniformidade e independência, e se aplicado no conjunto de dados da Figura 2, verificaríamos a não aleatoriedade dos dados.

Para ilustrar esta situação, modelou-se o problema da seguinte forma: definiu-se o espaço amostral e os quatro eventos contidos nele, mutuamente excludentes, os quais são as regiões A, B, C e D do gráfico de controle (Figura 2). Para o caso de pontos consecutivos abaixo ou acima da linha média, calculou-se a probabilidade da união dos eventos B e C, os quais podem seguir uma distribuição binomial cujo sucesso é estar abaixo ou acima da linha média e o fracasso é, respectivamente, o complementar. Definiram-se também as hipóteses a serem testadas, estas baseadas na normalidade. Logo, a probabilidade encontrada é a de rejeitar  $H_0$  sendo esta verdadeira, ou seja, encontra-se o p-valor. Então, com um nível de significância ( $\alpha$ ), rejeita-se  $H_0$ , caso o p-valor seja menor que  $\alpha$ . Ilustra-se da seguinte forma:

- hipóteses a serem testadas (Equação 9):

$$\begin{cases} H_0: \text{Processo sob controle;} \\ H_1: \text{Processo fora de controle;} \end{cases} \quad (9)$$



**Figura 1.** Gráfico do p-valor em relação ao número de pontos no gráfico de controle que estejam seguindo um padrão cíclico.

- para o caso de pontos consecutivos, tem-se (Equação 10):

$$P(B \cup C) = P(B) + P(C) = 2P(B)$$

$$P(B) = P(X = m) = C_{m,m} p^m (1-p)^{m-m}$$

$$P(B \cup C) = 2p^m = 2 \cdot 0,4987^m < 0,01 \quad (10)$$

$$\ln(2) + m \ln(0,4987) < \ln(0,01)$$

$$0,6931 + m(-0,6957) < -4,6052$$

$$m > 7$$

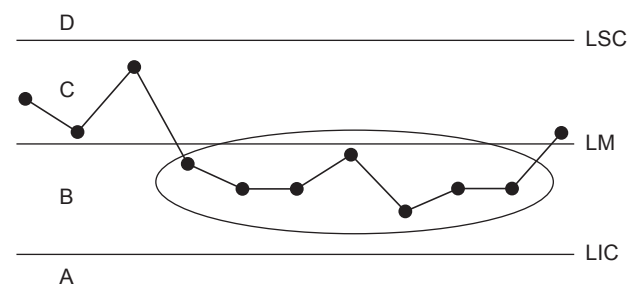
- Portanto, se a probabilidade de se rejeitar  $H_0$ ,  $H_0$  sendo verdadeiro, for muito pequena, pode-se concluir que o processo está fora de controle. Logo, para uma sequência com mais de seis pontos consecutivos, pode-se afirmar que o processo está fora de controle a 5% de significância. Já para um  $\alpha$  de 1%, o processo só estará fora de controle para oito ou mais pontos consecutivos abaixo ou acima da linha média (Tabela 1).

Os cálculos de probabilidades podem ser feitos fazendo-se uma aproximação da distribuição binomial à distribuição normal. A seguir, um exemplo para o caso de sete pontos consecutivos.

Segundo o teorema de DeMoivre e Laplace para se ter uma boa aproximação Normal o número de tentativas(n)

**Tabela 1.** Possibilidades de ocorrência da união dos eventos B e C, caso consecutivo.

m	P(B ∪ C)
4	12,36%
5	6,16%
6	3,08%
7	1,54%
8	0,76%



**Figura 2.** Exemplo de sequência de sete pontos consecutivos abaixo da linha média (LM), mas entre os limites, inferior (LIC) e superior de controle (LSC).

deve ser  $n > 10$  e o  $p$  deve estar próximo de 0,5, para outros valores de  $p$  é necessário aumentar o valor de  $n$ . Este teorema é um caso particular do teorema central do limite.

Segundo Triola (1999), em uma boa aproximação, tem-se que verificar (Equação 11):

$$\begin{cases} mp \geq 5 \rightarrow 7 \cdot 0,4987 = 3,49 < 5 \\ mq \geq 5 \rightarrow 7 \cdot 0,5013 = 3,51 < 5 \end{cases} \quad (11)$$

Observa-se, então, que a aproximação não é boa para sete pontos, mas para  $m > 10$ , sim. Mas, como já vimos acima, para oito pontos ou mais ( $p - \text{valor} < 5\%$ ), o processo se encontra fora de controle, então, não se usa esperar  $m$  ser maior que vinte para verificar este padrão.

Pode-se, ainda, ilustrar tal padrão com o espaço amostral composto por observações que estejam dentro dos limites, mas estão caindo apenas de um lado da linha média. Então, o sucesso é  $k$  observações de um lado da linha média e o fracasso é  $k - 1$  do outro lado da linha média. Como o objetivo é analisar a união dos dois eventos abaixo ou acima da linha média, tem-se (Equação 12):

$$\begin{aligned} P(B \cup C) &= 2P(X = k) = \\ &= 2 \cdot 0,4987 \cdot (0,5013)^{k-1} < 0,01 \quad k > 7 \end{aligned} \quad (12)$$

Logo, ilustrar no contexto de distribuição geométrica ou binomial, não faz nenhuma diferença, chegando-se assim, ao mesmo resultado. Já para os casos não consecutivos, não é possível o cálculo dessas probabilidades pelo contexto da geométrica, então segue em binomial.

Já para o caso em que nem todos os pontos são consecutivos e  $p = 0,4987$ , tem-se (Equação 13):

$$\begin{aligned} P(B \cup C) &= P(B) + P(C) = 2P(B) \\ P(B) &= P(X = v) + C_{m,v} p^v (1-p)^{m-v} \end{aligned} \quad (13)$$

- Dez de onze são consecutivos abaixo ou acima da linha média (Equação 14):

$$\begin{aligned} P(B) &= P(X = 10) = C_{11,10} p^{10} (1-p)^1 = \\ &= C_{11,10} 0,4987^{10} (0,5013)^1 = \\ &= 0,52\%; \end{aligned} \quad (14)$$

Logo, a probabilidade de se rejeitar  $H_0$ , verdadeiro, é muito pequena, ou seja, a possibilidade de se cometer este erro é baixa. Então, se pode rejeitar  $H_0$  a 5% de significância, quando se tem dez de onze pontos consecutivos abaixo ou acima da linha média. Já a 1% não se pode constatar o mesmo, mas, pelos dados da Tabela 2, pode-se perceber que o  $p$ -valor ficou muito próximo de ser menor que 1%.

- Doze de quatorze são consecutivos (Equação 15):

$$\begin{aligned} P(X = 12) &= C_{14,12} p^{12} (1-p)^2 = \\ &= C_{14,12} 0,4987^{12} (0,5013)^2 = \\ &= 0,54\%; \end{aligned} \quad (15)$$

Da mesma forma, pode-se rejeitar  $H_0$ , a 5% de significância, quando se tem doze de quatorze pontos consecutivos abaixo ou acima da linha média. E, para 1%, acontece o mesmo do caso acima.

- Dezesseis de vinte são consecutivos (Equação 16):

$$\begin{aligned} P(X = 16) &= C_{20,16} p^{16} (1-p)^4 = \\ &= C_{20,16} 0,4987^{16} (0,5013)^4 = 0,45\% \end{aligned} \quad (16)$$

Pode-se, então, rejeitar  $H_0$ , a 5% de significância, quando se tem dezesseis de vinte pontos consecutivos abaixo ou acima da linha média. E, para 1%, pode-se afirmar a mesma coisa, rejeitando-se, assim,  $H_0$ .

### 3.4 Tendência

Segundo Montgomery (2004), um processo com comportamento estacionário é visto por Shewhart como um processo sob controle. O que leva uma série (ou um conjunto de observações) a ser não estacionária é a existência de sazonalidade e/ou tendência. Logo, um processo que apresente tendência não é um processo com comportamento estacionário, conseqüentemente, não está sob controle. Werkema (1995) comenta também que quando se observam sete ou mais pontos consecutivos (ascendentes ou descendentes), o processo encontra-se fora de controle.

Um teste baseado em regressão linear simples poderia ser proposto se fosse identificado um valor  $\tau$ , para o  $R^2$  (coeficiente de determinação desta reta), tal que, acima deste valor ( $R^2 \geq \tau$ ), os pontos observados seriam claramente não normais e abaixo do qual haveria normalidade. Espera-se que  $\tau$  seja dependente do teste de normalidade usado, ou seja, cada teste de normalidade deve ter um  $\tau$  acima do qual os pontos são considerados não normais. Espera-se que o  $\tau$  seja tão menor quanto mais rigoroso for o teste de normalidade.

### 3.5 Mistura ou aproximação dos limites de controle

Para ilustrar esta situação, supuseram-se limites bem largos próximos dos limites de controle, como os limites pontilhados (Figura 3), seguindo-se com estudos de

**Tabela 2.** Possibilidades de ocorrência da união dos eventos B e C, casos não consecutivos.

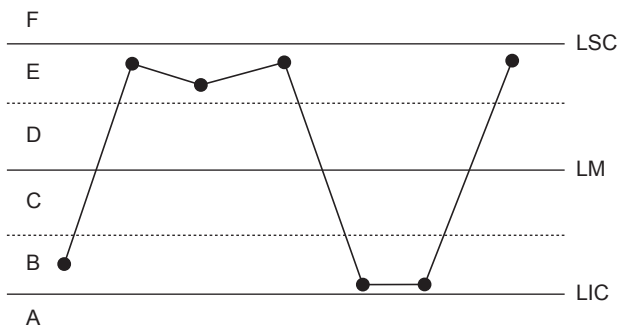
x	P(B ∪ C)
10 de 11 pontos	1,05%
12 de 14 pontos	1,08%
16 de 20 pontos	0,89%

probabilidades semelhantes aos feitos no item 3.3 acima. Considera-se que existem no espaço amostral seis eventos, A, B, C, D, E, F, e calculam-se as probabilidades da união entre B e E. Construíram-se limites que variam em torno de 2,0-sigma a 2,4-sigma. O objetivo é obter a probabilidade de ter pontos consecutivos ou não nos eventos B e E, e verificar a possibilidade de rejeição da hipótese de que o processo está sob controle ou normalidade. Supondo-se normalidade, tem-se a probabilidade ( $p$ ) de cada ponto cair nos referidos limites, sendo esta probabilidade constante e os  $m$  pontos independentes. Então, pode-se modelar uma distribuição binomial ( $m; p$ ). O sucesso consiste em caírem pontos dentro dos eventos B e E, nos quais estes são dados pelos limites construídos, e o fracasso é cair pontos fora deles. Sejam os casos:

- Para o caso de pontos consecutivos e pela a Equação (10), tem-se (Equação 18):

$$\begin{aligned} P(B \cup E) &= 2p^m < 0,01 \\ \ln(2) + m\ln(p) &< \ln(0,01) \\ 0,6931 + m\ln(p) &< -4,6052 \quad (18) \\ m &> \frac{5,2983}{-\ln(p)} \end{aligned}$$

Então, para dois ou mais pontos consecutivos dentro dos intervalos, o processo está fora de controle, isso com nível de significância a 1% (Tabela 3).



**Figura 3.** Exemplo de mistura em um processo.

**Tabela 3.** Número de pontos nos intervalos, em que pontos consecutivos iguais ou acima de  $m$  indicam que o processo está fora de controle com  $\alpha = 0,01$ .

k-sigma	p	m
2,0	0,0228	2
2,1	0,0179	2
2,2	0,0139	2
2,3	0,0107	2
2,4	0,0082	2

- Para pontos não consecutivos, pela Equação (14), tem-se o resultado da Tabela 4.

Então, para quatro ou mais pontos dentro do primeiro intervalo, pode-se dizer que o processo está fora de controle. E, em intervalos de 2,2-sigma, estará fora de controle com três ou mais pontos, isso com nível de significância a 1%.

### 3.6 Estratificação ou aproximação da linha média

Para ilustrar esta situação, supuseram-se limites bem estreitos em torno da linha média, como os limites pontilhados (Figura 4), seguindo-se com estudos de probabilidades semelhantes aos feitos no item 5. Construíram-se limites que variam em torno de 0,1-sigma a 0,6-sigma, supondo-se pontos consecutivos dentro dos limites e também não consecutivos, em que estes, por exemplo, viriam de uma amostra de tamanho  $m = 24$ . Agora, o objetivo é obter a probabilidade de se ter vários pontos dentro dos eventos C e D e verificar se rejeitaríamos a hipótese de que o processo está sob controle.

Observe-se que foi construído o espaço amostral semelhante ao do item 3.5. Supondo-se normalidade, tem-se que a probabilidade ( $p$ ) de cada ponto cair nos referidos limites é constante, os eventos mutuamente excludentes e os  $m$  pontos independentes. Então, pode-se modelar uma distribuição binomial ( $m; p$ ), em que o sucesso é caírem pontos dentro dos limites, nos quais estes formarão os eventos C e D, e o fracasso é caírem pontos fora deles. Sejam os casos:

- Para o caso de pontos consecutivos e pela Equação (18), tem-se (Equação 19):

$$\begin{aligned} P(C \cup D) &= 2p^m = 2p^m < 0,01 \\ m &> \frac{5,2983}{-\ln(p)} \quad (19) \end{aligned}$$

Portanto, a quantidade de pontos consecutivos  $m$  encontrados dentro destes intervalos mostra um padrão de não aleatoriedade do processo para todo  $m$  maior ou igual ao encontrado na Tabela 5. Por exemplo, para o intervalo de 0,6-sigma, tem-se que, para  $m \geq 4$ , o processo se encontra fora de controle estatístico, ou seja, com quatro

**Tabela 4.** Número de pontos nos intervalos, supondo  $m = 24$ , em que pontos iguais ou acima de  $x$  indicam que o processo está fora de controle com  $\alpha = 0,01$ .

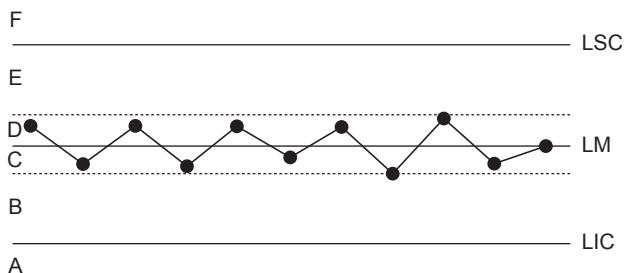
k-sigma	p	x
2,0	0,0228	4
2,1	0,0179	4
2,2	0,0139	3
2,3	0,0107	3
2,4	0,0082	3

**Tabela 5.** Número de pontos nos intervalos, em que pontos consecutivos acima de  $m$  indicam que o processo está fora de controle com  $\alpha = 0,01$ .

k-sigma	p	m
0,1	0,0398	2
0,2	0,0793	3
0,3	0,1179	3
0,4	0,1554	3
0,5	0,1915	4
0,6	0,2257	4

**Tabela 6.** Número de pontos nos intervalos, em que pontos iguais ou acima de  $x$  indicam que o processo está fora de controle, considerando-se  $m = 24$  e  $\alpha = 0,01$ .

k-sigma	p	x
0,1	0,0398	5
0,2	0,0793	7
0,3	0,1179	8
0,4	0,1554	10
0,5	0,1915	11
0,6	0,2257	12



**Figura 4.** Exemplo de estratificação ou aproximação da linha média.

ou mais pontos consecutivos próximos da linha média, à distância de 0,6-sigma, não há controle estatístico a 1% de significância.

Para o caso de pontos não consecutivos, pela Equação (13), tem-se o resultado da Tabela 6.

Logo, a quantidade de pontos  $x$  dentro dos intervalos construídos que indicarão um processo fora de controle será sempre maior ou igual aos valores de  $x$  encontrados na Tabela 6. Por exemplo, para o intervalo de 0,6-sigma, tem-se que, para  $x \geq 12$ , o processo se encontra fora de controle, ou seja, em 24 pontos, se 12 ou mais pontos se encontram próximos da linha média, à distância de 0,6-sigma, este processo está fora de controle, a 1% de significância.

## 4 Conclusões

Conclui-se que, para pontos fora dos limites de controle, as distribuições Bernoulli e Geométrica encontraram o mesmo resultado. As duas retornaram a mesma probabilidade de se rejeitar a hipótese de nulidade, sob normalidade, 0,27%. Já no caso da periodicidade, a literatura citada não comenta sobre o número de pontos que retorna um processo fora de controle. O estudo realizado estima este número, tendo o estudo sido bem específico. Utilizou-se a função seno do software R com parâmetro default. Logo, as conclusões que seguem devem ser vistas com ressalvas. Então, com o teste realizado, verificou-se que são precisos, no mínimo, dezesseis pontos cíclicos para se poder afirmar que o processo está fora de controle a 5% de significância.

A literatura cita que sete ou mais pontos consecutivos e certo número de pontos não consecutivos são padrões

de não aleatoriedade, chamados de sequência. Mas, isso não é tão informativo, como, por exemplo, qual seria a confiança desta informação? Assim, obtiveram-se os seguintes resultados: para os casos consecutivos, o processo se encontra fora de controle, com seis ou mais pontos a 5% de significância, ou ainda, com mais de sete pontos a 1% de significância. Nos casos em que 10 de 11 e 12 de 14 pontos são consecutivos, pode-se afirmar que são padrões de não aleatoriedade a 5% de significância, mas, o mesmo não pode ser dito a 1% de significância. Já no caso de 16 de 20 pontos consecutivos, a afirmação é verdadeira para os dois níveis de significância.

No caso do padrão mistura, não é comum encontrar qual o número de pontos, próximos dos limites de controle, que indiquem um padrão de não aleatoriedade. Então, em relação aos diferentes limites construídos, conclui-se que, no caso de pontos consecutivos em todos os limites, o número de pontos que confirma um processo fora de controle deve ser maior ou igual a dois, a 1% de significância. No caso não consecutivo, considerando-se que o gráfico tenha 24 pontos, tem-se que, para os limites de 2,0-sigma e 2,1-sigma são necessários quatro ou mais pontos dentro do intervalo entre os limites construídos e o limite de controle. Para os outros limites, 2,2-sigma, 2,3-sigma e 2,4-sigma, são necessários três ou mais pontos, todos a 1% de significância. Também não é comum encontrar, na literatura, números de pontos que indiquem estratificação. Logo, no caso consecutivo, tem-se que os números de pontos dentro do limite de 0,1-sigma, que confirmam um padrão de não aleatoriedade, são dois ou mais pontos. Nos limites 0,2-sigma, 0,3-sigma e 0,4-sigma são necessários três ou mais e, nos limites 0,5-sigma e 0,6-sigma, quatro ou mais pontos consecutivos. Agora, considerando-se também 24 pontos no gráfico, tem-se, que, no limite de 0,1-sigma, são necessários cinco ou mais pontos, para 0,2-sigma, sete ou mais, 0,3-sigma, oito ou mais, 0,4-sigma, dez ou mais, 0,5-sigma, onze ou mais e 0,6-sigma, doze ou mais pontos não consecutivos, para indicar um padrão de não aleatoriedade, a 1% de significância.

Para o padrão de tendência recomenda-se, sugere-se, utilizar a simulação Monte Carlo para a determinação de valores de  $\tau$ .

## Studies on non-random patterns in Shewhart control charts

### Abstract

*In the last two decades, the Brazilian consuming market was modified by the insertion of products of better quality and minor cost. Statistical Process Control (SPC) and control charts look after significant characteristics products and processes, in real time, guaranteeing quality levels, in a cost demanded by the market. In that context, an important procedure is the interpretation of the control charts; this is made through non-random patterns. A study was made illustrating and arguing such standards. Standards had been illustrated and one revealed which of them, rejected the nullity hypothesis, under a determined significance. One concluded that the standards mentioned in literature are not so informative and can be improved.*

**Keywords:** Statistical Process Control. Non-random patterns. Under control process. Special causes. Control charts.

### Referências bibliográficas

- COSTA, A. F. B.; EPPRECHT, E. K.; CARPINETTI, L. C. R. **Controle estatístico de qualidade**. São Paulo: Atlas, 2004.
- DEMING, W. E. **Qualidade: a revolução da administração**. Rio de Janeiro: Marques Saraiva, 1990. 1-367 p.
- FERREIRA, D. F.; DIAS, C. T. S. dos. **Apostila de estatística multivariada**. Lavras: Departamento de Ciências Exatas da UFLA, 2006. Cap. 3, p. 133-141. Disponível em: <<http://www.dex.ufla.br/~danielff/prefacioem.pdf>>. Acesso em: 22 Novembro 2006.
- LOURENÇO FILHO, R. C. B. **Controle estatístico de qualidade**. Rio de Janeiro: Livros Técnicos e Científicos, 1976.
- MONTGOMERY, D. C. **Introdução ao controle estatístico da qualidade**. Tradução Ana Maria Lima de Farias. 4 ed. Rio de Janeiro: LTC, 2004.
- R Development Core Team. **R: A language and environment for statistical computing**. Vienna, Austria: R Foundation for Statistical Computing, 2004. ISBN: 3-900051-07-0. Disponível em: <<http://www.R-project.org>>. Acesso em: 2007.
- ROYSTON, J. P. Approximating the Shapiro-Wilk's W-test for non-normality. **Statistics and Computing**, v. 2, n. 3, p. 117-119, 1992.
- ROYSTON, J. P. A toolkit for testing for testing for non-normality in complete and censored samples. **The Statistician**, v. 42, n. 1, p. 37-43, 1993.
- TORMINATO, S. M. **Análise da utilização da ferramenta CEP: um estudo de caso na manufatura de autopeças**. Campinas, 2004. Dissertação (Mestrado em Engenharia Mecânica) – Universidade Estadual de Campinas.
- TRIOLA, M. F. **Introdução à estatística**. Tradução Alfredo Alves de Farias. 7 ed. Rio de Janeiro: LTC, 1999. p. 133.
- WERKEMA, M. C. C. **Ferramentas estatísticas básicas para o gerenciamento de processos**. Belo Horizonte: QFCO, 1995.

### Sobre os autores

#### Quintiliano Siqueira Schrodin Nomelini

Faculdade de Ciências Integrada do Pontal, Universidade Federal de Uberlândia – FACIP/UFU  
Rua Antonio Augusto de Melo 529, Ituiutaba – MG  
e-mail: quintiliano@pontal.ufu.br

#### Eric Batista Ferreira

Departamento de Ciências Exatas (DEX/UFLA), Estatística e Experimentação Agropecuária – UFLA  
Caixa Postal 3037, CEP 37200-000, Lavras, MG  
e-mail: ericbferreira@netscape.net

#### Marcelo Silva de Oliveira

Departamento de Ciências Exatas, Universidade Federal de Lavras – DEX/UFLA  
Caixa Postal 3037, CEP 37200-000, Lavras, MG  
e-mail: marcelo.oliveira@ufla.br

Recebido: 7/3/2007  
Aceito: 31/8/2009