

## WHITE PAPER SOBRE O ASSISTENTE DO MINITAB

Este artigo é parte de uma série de artigos que explicam a pesquisa conduzida pelos estatísticos do Minitab para desenvolver os métodos e verificações de dados usados no Assistente no Minitab Statistical Software.

# Cartas de controle para variáveis

## Visão geral

As cartas de controle são usadas para monitorar regularmente um processo, a fim de determinar se ele está sob controle. O Assistente do Minitab inclui duas cartas de controle amplamente utilizadas para monitorar cartas com dados contínuos:

- Cartas Xbar-R ou Xbar-S. Estas cartas são usadas quando os dados são coletados em subgrupos. O Minitab usa o desvio padrão combinado para estimar o desvio padrão dentro do subgrupo. A carta R fornece uma estimativa efetiva da variação para subgrupos de tamanho até aproximadamente 10 (AIAG, 1995; Montgomery, 2001). Para tamanhos de subgrupos maiores, a carta S oferece uma estimativa melhor dos desvios padrão dentro de subgrupo individuais. Para acompanhar as convenções passadas do Minitab e sermos conservadores, recomendamos uma carta S quando o tamanho do subgrupo for maior do que 8. Para tamanhos de subgrupo menores ou iguais a 8, as cartas R e S oferecem resultados semelhantes.
- Carta I-MR. Esta carta é usada quando não existem subgrupos. O Minitab usa um método de amplitude móvel média de comprimento 2 para estimar o desvio padrão.

Os limites de controle para uma carta de controle normalmente são estabelecidos na fase de controle de um projeto Seis Sigma. Uma boa carta de controle deve ser sensível o suficiente para sinalizar rapidamente quando uma causa especial existir. Esta sensibilidade pode ser avaliada calculando-se o número médio de subgrupos necessário para sinalizar uma causa especial. Da mesma forma, uma boa carta de controle raramente indica um "falso positivo" quando o processo está sob controle. A taxa de falsos positivos pode ser avaliada calculando-se o percentual de subgrupos considerados "fora de controle" quando o processo está sob controle.

Em geral, as cartas de controle são otimizadas quando cada observação vem de uma distribuição normal, cada observação é independente e apenas a variabilidade de causas comuns existe dentro dos subgrupos. Portanto, a Carta de Relatório do assistente executa automaticamente as verificações de dados a seguir para avaliar estas condições:

- Normalidade
- Estabilidade
- Quantidade de dados
- Dados correlacionados

Neste artigo, nós investigamos como uma carta de controle de variáveis se comporta quando estas condições variam e quando nós descrevemos como estabelecemos um conjunto de orientações para avaliar os requisitos para estas condições.

# Verificações dos dados

## Normalidade

As cartas de controle não são baseadas na suposição de que os dados do processo são distribuídos normalmente, mas os critérios usados nos testes para causas especiais são baseados nesta suposição. Se os dados forem rigorosamente assimétricos ou se os dados ficarem concentrados nas extremidades da distribuição (“com caudas pesadas”), os resultados do teste podem não ser precisos. Por exemplo, o gráfico pode indicar falsos positivos a uma taxa mais alta do que o esperado.

## Objetivo

Nós investigamos o efeito de dados não normais nas cartas Xbar e I. Nosso objetivo era determinar como a não normalidade afeta a taxa de falsos positivos. Especificamente, queríamos determinar se dados não normais aumentam significativamente a taxa na qual uma carta indica que pontos estão fora de controle quando o processo está realmente no controle (falsos positivos).

## Método

Realizamos simulações com 10.000 subgrupos e diferentes níveis de não normalidade e registramos o percentual de falsos positivos. As simulações permitem o teste de várias condições para determinar os efeitos da não normalidade. Nós escolhemos a distribuição assimétrica à direita e as distribuições simétricas com caudas pesadas porque estes são casos comuns de distribuições não normais na prática. Consulte o Anexo D para obter mais detalhes.

## Resultados

### **CARTA XBAR (TAMANHO DO SUBGRUPO $\geq 2$ )**

Nossa simulação mostrou que a taxa de falsos positivos não aumenta significativamente quando os dados são não normais no caso de o tamanho do subgrupo ser 2 ou mais. Com base neste resultado, nós não verificamos a normalidade para as cartas Xbar-R ou Xbar-S. Mesmo quando os dados são altamente assimétricos ou extremamente pesados na cauda, a taxa de falsos positivos para o teste 1 e o teste 2 é menor do que 2%, o que não é notavelmente mais alto do que a taxa de falsos positivos de 0,7% para a distribuição normal.

### **CARTA I (TAMANHO DE SUBGRUPO = 1)**

Nossa simulação mostrou que a carta I é sensível a dados não normais. Quando os dados são não normais, a carta I produz uma taxa de falsos positivos que é 4 a 5 vezes mais alta do que quando os dados são normais. Para tratar a sensibilidade da carta I para dados não normais, o Assistente faz o seguinte:

- Executa um teste de Anderson-Darling se os dados puderem ser altamente não-normais, como indicado por um número de pontos maior do que o esperado fora

dos limites de controle (ou seja, 2 ou mais pontos e 2% ou mais dos pontos estão fora dos limites de controle).

- Se o teste de Anderson-Darling sugerir que os dados são não normais, o Assistente transforma os dados utilizando o lambda ideal de Box-Cox. O teste de Anderson-Darling é realizado nos dados transformados. Se o teste não rejeitar a hipótese nula (de que os dados são normais), o Assistente sugere a utilização dos dados transformados se o processo produzir naturalmente dados não normais.

A transformação de Box-Cox é eficiente somente para dados não normais que forem assimétricos à direita. Se a transformação não for eficiente para seus dados não normais, talvez seja necessário considerar outras opções. Além disso, como o teste de Anderson-Darling e a transformação de Box-Cox são afetadas pelas observações extremas, você deve omitir os pontos com causas especiais conhecidas antes de transformar seus dados.

Quando procurar pela normalidade, o Cartão de Relatório do Assistente para a carta I exibe os indicadores de status a seguir:

Status	Condição
	Se < 2 pontos ou < 2% dos pontos estiverem fora dos limites de controle na carta I ou O teste de normalidade de Anderson-Darling para os dados ou para os dados transformados de Box-Cox falha em rejeitar a hipótese nula (de que os dados são normais) em $\alpha = 0,01$ .
	Se as condições acima não forem sustentadas.

## Estabilidade

Para cartas de controle de variáveis, podem ser executados oito testes para avaliar a estabilidade do processo. O uso simultâneo destes testes aumenta a sensibilidade da carta de controle. Entretanto, é importante determinar a finalidade e o valor agregado de cada teste porque a taxa de falsos positivos aumenta quanto mais testes forem adicionados à carta de controle.

### Objetivo

Nós desejávamos determinar qual dos oito testes para estabilidade deveria ser incluído com as cartas de controle de variáveis no Assistente. Nosso primeiro objetivo era identificar os testes que aumentam significativamente a sensibilidade para as condições fora de controle sem elevar significativamente a taxa de falsos positivos. Nosso segundo objetivo era garantir a simplicidade e praticidade da carta. Nossa pesquisa se concentrou nos testes para a carta Xbar e carta I. Para as cartas R, S e MR, nós usamos somente o Teste 1, que sinaliza quando um ponto cai fora dos limites de controle

### Método

Nós realizamos simulações e uma revisão da literatura para avaliar como a utilização de uma combinação de testes para estabilidade afeta a sensibilidade e a taxa de falsos positivos das

cartas de controle. Além disso, nós avaliamos a prevalência das causas especiais associadas ao teste. Para obter mais detalhes sobre os métodos usados para cada teste, consulte a seção Resultados abaixo e o Anexo B.

## Resultados

Descobrimos que os testes 1, 2 e 7 foram os mais úteis para avaliação da estabilidade da carta Xbar e da carta I:

### TESTE 1: IDENTIFICA PONTOS FORA DOS LIMITES DE CONTROLE

O teste 1 identifica pontos > 3 desvios padrão da linha central. O teste 1 é reconhecido universalmente como necessário para a detecção de situações fora de controle. Ele tem uma taxa de falsos positivos de somente 0,27%.

### TESTE 2: IDENTIFICA OS DESLOCAMENTOS NAS MÉDIAS

O teste 2 sinaliza quando 9 pontos em sequência caem no mesmo lado da linha central. Nós realizamos uma simulação usando 4 médias diferentes, definidas para múltiplos do desvio padrão, e determinamos o número de subgrupos necessários para detectar um sinal. Nós definimos os limites de controle com base na distribuição normal. Descobrimos que adicionar o Teste 2 aumenta significativamente a sensibilidade da carta para detectar deslocamentos pequenos na média. Quando o teste 1 e o teste 2 são usados juntos, é necessário um número significativamente menor de subgrupos para detectar um deslocamento pequeno na média do que seria necessário se o teste 1 fosse usado sozinho. Portanto, adicionar o teste 2 ajuda a detectar situações fora de controle e aumenta a sensibilidade o suficiente para garantir um pequeno aumento na taxa de falsos positivos.

### TESTE 7: IDENTIFICA OS LIMITES DE CONTROLE QUE SÃO LARGOS DEMAIS

O teste 7 sinaliza quando 12-15 pontos em sequência caem dentro de 1 desvio padrão da linha central. O teste 7 é usado somente para a carta Xbar quando os limites de controle são estimados a partir dos dados. Quando este teste falha, a causa é geralmente uma fonte sistêmica de variação (estratificação) dentro de um subgrupo, que é muitas vezes o resultado da não formação de subgrupos racionais. Como a formação de subgrupos racionais é fundamental para garantir que a carta de controle possa detectar com precisão as situações fora do controle, o Minitab usa um teste 7 modificado ao estimar os limites de controle a partir dos dados. O teste 7 indica uma falha quando o número de pontos em sequência está 12 e 15, dependendo do número de subgrupos:

<b>k = (número de subgrupos) x 0,33</b>	<b>Pontos necessários</b>
k < 12	12
k ≥ 12 e k ≤ 15	Inteiro ≥ k
k > 15	15

## Testes não incluídos no Assistente

### TESTE 3: K PONTOS EM SEQUÊNCIA", TODOS CRESCENTES OU DECRESCENTES

O teste 3 é projetado para detectar deslocamentos na média do processo (Davis e Woodall, 1988). No entanto, quando o teste 3 é utilizado além dos testes 1 e 2, não há um aumento significativo da sensibilidade da carta para detectar deslocamentos na média do processo. Como nós já decidimos usar os testes 1 e 2 com base em nossos resultados de simulação, a inclusão do teste 3 não agregaria nenhum valor significativo à carta.

### TESTE 4: PONTOS K EM SEQUÊNCIA, ALTERNANDO ACIMA E ABAIXO

Embora este padrão possa ocorrer na prática, recomendamos que você procure por tendências ou padrões incomuns em vez do teste para um padrão específico.

### TESTE 5: K DE K=1 PONTOS > 2 DESVIOS PADRÃO DA LINHA CENTRAL

Para garantir a simplicidade da carta, excluímos esse teste porque ele não identifica de modo singular situações de causa especial.

### TESTE 6: K DE K+1 PONTOS > 1 DESVIO PADRÃO DA LINHA CENTRAL

Para garantir a simplicidade da carta, excluímos esse teste porque ele não identifica de modo singular situações de causa especial.

### TESTE 8: K PONTOS EM SEQUÊNCIA > DE 1 DESVIO PADRÃO DA LINHA CENTRAL

Para garantir a simplicidade da carta, excluímos esse teste porque ele não identifica de modo singular situações de causa especial.

Quando procura por estabilidade, o Cartão de Relatório do Assistente exibe os indicadores de status a seguir:

Status	Condição
	<p>Não há falhas nos testes na carta para a média (carta I ou carta Xbar) e o gráfico de variação (carta AM, R ou S).</p> <p>Os testes usados para cada carta são:</p> <ul style="list-style-type: none"><li>• Carta I: Teste 1 e Teste 2.</li><li>• Carta Xbar: Teste 1, Teste 2 e Teste 7. Teste 7 é realizado somente quando os limites de controle são estimados a partir dos dados.</li><li>• Cartas AM, R e S: Teste 1.</li></ul>
	<p>Se a condição acima não vigorar.</p>

## Quantidade de dados

Se não existem valores conhecidos para os limites de controle, eles podem ser estimados a partir dos dados. Para obter estimativas precisas dos limites, você deve ter dados suficientes. Se a quantidade de dados for suficiente, os limites de controle podem estar longe dos limites "verdadeiros" devido à variabilidade da amostra. Para aprimorar a precisão dos limites, é possível aumentar o número de observações.

### Objetivo

Investigou-se o número de observações que são necessárias para se obter limites de controle precisos. Nosso objetivo foi determinar a quantidade de dados necessários para assegurar que os falsos positivos devido ao teste 1 não fossem mais do que 1%, com 95% de confiança.

### Método

Quando os dados são normalmente distribuídos e não há nenhum erro devido à variabilidade de amostragem, o percentual de pontos acima do limite de controle superior é 0,135%. Para determinar se o número de observações é adequado, seguimos o método descrito por Bischak (2007) para garantir que a taxa de falsos positivos devido aos pontos acima do limite de controle superior não fosse superior a 0,5%, com 95% de confiança. Devido à simetria dos limites de controle, esse método resulta em uma taxa de falsos positivos total de 1%, devido ao teste 1. Consulte o Anexo C para obter mais detalhes.

### Resultados

Nós determinamos que, para quase todos os tamanhos de subgrupos, 100 observações totais são suficientes para se obter limites de controle precisos. Embora os tamanhos de subgrupo de 1 e 2 necessitassem de um pouco mais de observações, a taxa de falsos positivos ainda estava razoavelmente baixa (1,1%) com 100 observações. Portanto, para simplificar, usamos o corte de 100 observações totais para todos os tamanhos de subgrupos.

Com base nestes resultados, o Cartão do Relatório do Assistente mostra os indicadores de status a seguir na verificação da quantidade de dados:

Status	Condição
	Número de observações $\geq$ 100.
	Número de observações $<$ 100.

## Dados correlacionados

A autocorrelação é uma medida da dependência entre os pontos de dados que são coletados ao longo do tempo. A maioria dos dados de processo apresentam pelo menos um pequeno grau de autocorrelação. Se a autocorrelação for moderada ou alta, ela pode levar a resultados de testes incorretos. Normalmente, os dados autocorrelacionados exibem autocorrelação positiva, o que pode reduzir a variação dentro do subgrupo e conduzir a uma maior taxa de falsos positivos.

### Objetivo

Nós investigamos a relação entre a autocorrelação e a taxa de falsos positivos. Nosso objetivo foi determinar o nível de autocorrelação que gera uma taxa de falsos positivos inaceitável. Para simplificar, consideramos autocorrelação para lag 1 porque a autocorrelação para lag 1 é provavelmente maior que a autocorrelação para lags  $\geq 2$ .

### Método

Usando um modelo padrão para um processo autocorrelacionado, realizamos simulações com  $\phi = 0,2, 0,4, 0,5, 0,6$  e  $0,8$  ( $\phi$  é a autocorrelação do lag 1) para três tamanhos de subgrupo ( $n = 1, 3$  e  $5$ ). Usamos um conjunto inicial de 10.000 subgrupos para estabelecer os limites de controle. Em seguida, registramos o percentual de falsos positivos para um adicional de 2.500 subgrupos. Realizamos 10.000 iterações e registramos o percentual médio de falsos positivos. Consulte o Anexo D para obter mais detalhes.

### Resultados

Nossas simulações mostraram que mesmo níveis moderados de autocorrelação aumentam significativamente a taxa de falsos positivos. Quando a autocorrelação é  $\geq 0,4$ , a taxa de falsos positivos é muito alto e o gráfico de controle não tem significado. Para solucionar este problema, o Assistente executa um teste de autocorrelação se os dados puderem ser autocorrelacionados, como indicado por um número de pontos fora dos limites de controle maior do que o esperado (quando 2 ou mais pontos e 2% ou mais dos pontos estão do fora dos limites de controle). Nesse caso, os primeiros testes do Assistente, se for uma autocorrelação entre os pontos de dados sucessivos (lag = 1), é significativamente maior do que 0,2. Se a autocorrelação for significativamente maior do que 0,2, o Assistente testa se a autocorrelação entre os pontos de dados sucessivos (lag = 1) é significativamente maior do que 0,4.

Quando procura por dados correlacionados, o Cartão de Relatório do Assistente exibe os indicadores de status a seguir:

Status	Condição
	<p>O número de pontos fora dos limites de controle não é maior do que o esperado, ou seja, &lt; 2 pontos ou &lt; 2% dos pontos estão fora dos limites de controle.</p> <p>O número de pontos fora dos limites de controle é maior do que o esperado, mas um teste de autocorrelação = 0,2 contra a autocorrelação &gt; 0,2 não rejeita a hipótese nula no nível alfa = 0,01. Portanto, não há evidências suficientes para concluir que exista pelo menos um nível moderado de autocorrelação.</p>
	<p>Se as condições acima não forem sustentadas.</p> <p>Observação: se a hipótese nula de que a autocorrelação = 0,2 for rejeitada, realizamos um teste de acompanhamento de autocorrelação = 0,4 contra a autocorrelação &gt; 0,4. Se a autocorrelação = 0,4 teste for rejeitada, nós aumentamos a gravidade da mensagem de cautela.</p>

Para obter mais detalhes sobre o teste de hipótese para autocorrelação, consulte o Anexo D.

# Referências

- AIAG (1995). *Statistical process control (SPC) reference manual*. Automotive Industry Action Group. AIAG (1995).
- Bischak, D.P., & Trietsch, D. (2007). The rate of false signals in  $\bar{X}$  control charts with estimated limits. *Journal of Quality Technology*, 39, 55–65.
- Bowerman, B.L., & O'Connell, R.T. (1979). *Forecasting and time series: An applied approach*. Belmont, CA: Duxbury Press.
- Chan, L. K., Hapuarachchi K. P., & Macpherson, B.D. (1988). Robustness of  $\bar{X}$  and  $R$  charts. *IEEE Transactions on Reliability*, 37, 117–123.
- Davis, R.B., & Woodall, W.H. (1988). Performance of the control chart trend rule under linear shift. *Journal of Quality Technology*, 20, 260–262.
- Montgomery, D. (2001). *Introduction to statistical quality control, 4<sup>th</sup> edition*. John Wiley & Sons.
- Schilling, E.G., & Nelson, P.R. (1976). The effect of non-normality on the control limits of  $\bar{X}$  charts. *Journal of Quality Technology*, 8, 183–188.
- Trietsch, D. (1999). *Statistical quality control: A loss minimization approach*. Singapore: World Scientific Publishing Company.
- Wheeler, D.J. (2004). *Advanced topics in statistical process control. The power of Shewhart's charts, 2<sup>nd</sup> edition*. Knoxville, TN: SPC Press.
- Yourstone, S.A., & Zimmer, W.J. (1992). Non-normality and the design of control charts for averages. *Decision Sciences*, 23, 1099–1113.

# Anexo A: Normalidade

## Simulação A1: Como a não normalidade afeta a taxa de falsos positivos

Para investigar como dados não normais afetam o desempenho da carta I e da carta Xbar, realizamos uma simulação para avaliar a taxa de falsos positivos associados a distribuições de dados não normais. Concentramos nossa atenção na distribuição assimétrica à direita e em distribuições simétricas com caudas pesadas, porque elas são as distribuições não normais comuns na prática. Em particular, examinamos 3 distribuições assimétricas (qui-quadrado com  $df = 3, 5$  e  $10$ ) e 2 distribuições com caudas pesadas (t com  $df = 3$  e  $5$ ).

Nós estabelecemos os limites de controle usando um conjunto inicial de 10.000 subgrupos. Registramos o percentual de falsos positivos por mais de 2.500 subgrupos. Em seguida, realizamos 10.000 iterações e calculamos o percentual médio de falsos positivos utilizando os testes 1 e 2 para causas especiais. Os resultados são apresentados na Tabela 1.

**Tabela 1** % de falsos positivos para o teste 1, teste 2 para a carta I ( $n=1$ ) e carta Xbar ( $n=2, 3, 4$  ou  $5$ )

Distribuição	Tamanho do subgrupo				
	n = 1	n = 2	n = 3	n = 4	n = 5
Normal (0, 1)	0,27, 0,39	0,27, 0,39	0,27, 0,39	0,27, 0,39	0,27, 0,39
Chi (3)	2,06, 1,17	1,18, 0,79	0,98, 0,62	0,86, 0,57	0,77, 0,53
Chi (5)	1,54, 0,83	0,93, 0,60	0,77, 0,53	0,67, 0,50	0,61, 0,47
Chi (10)	1,05, 0,60	0,67, 0,50	0,56, 0,46	0,50, 0,44	0,45, 0,43
t (3)	2,18, 0,39	1,30, 0,39	1,22, 0,39	1,16, 0,39	1,11, 0,39
t (5)	1,43, 0,39	0,93, 0,39	0,80, 0,39	0,71, 0,39	0,66, 0,39

Cada célula na Tabela 1 indica a taxa de falsos positivos calculados para os testes 1 e 2, respectivamente. Por exemplo, a célula associada com a distribuição normal padrão (Normal (0, 1)) e a carta I ( $n = 1$ ) indica que a carta I tem uma taxa de falsos positivos de 0,27% para o teste 1 e 0,39% para o teste 2, quando os dados são normalmente distribuídos.

Comparando as taxas de falsos positivos para a carta Xbar ( $n = 2, 3, 4, 5$ ) entre as distribuições normais e não normais, é possível observar que a taxa de falsos positivos não aumenta significativamente quando a suposição de normalidade é violada. Mesmo com distribuições altamente assimétricas (qui-quadrado, 3) e distribuições com caudas pesadas (t, 3), a taxa de falsos positivos para os testes 1 e 2 combinados é inferior a 2,0% (contra 0,7% para uma distribuição normal) para tamanhos de subgrupo pequenos, como 2, por exemplo. Assim, podemos concluir, em termos práticos, que a carta Xbar é resistente a violações da suposição de normalidade.

Para a carta I, a Tabela 1 mostra uma taxa de falsos positivos de aproximadamente 3,2% para os testes 1 e 2 combinados quando a distribuição é fortemente assimétrica (qui-quadrado, 3); esta taxa de falsos positivos é quase 5 vezes maior do que a taxa de falsos positivos esperada quando os dados são normalmente distribuídos. A taxa de falsos positivos para os testes 1 e 2 combinados é de 2,6% para distribuições com caudas pesadas (t, 3), que é quase 4 vezes a taxa esperada quando os dados são normalmente distribuídos.

Portanto, os dados para a carta I-MR devem primeiramente ser testados quanto à normalidade. Se a hipótese nula do teste de qualidade de ajuste de Anderson-Darling para a normalidade for rejeitada (existe evidência estatisticamente significativa de que os dados não seguem uma distribuição normal), é possível usar uma transformação de Box-Cox para normalizar os dados. No entanto, a transformação Box-Cox é eficaz somente quando os dados são assimétricos à direita. Para outros casos de dados não normais são necessários mais estudos para determinar o melhor curso de ação. Além disso, como o teste de Anderson-Darling e a transformação Box-Cox são afetados por observações extremas, pontos com causas especiais conhecidas devem ser omitidos dos cálculos antes de realizar a análise.

# Anexo B: Estabilidade

## Simulação B1: Como a adição do teste 2 ao teste 1 afeta a sensibilidade

O teste 1 detecta pontos fora de controle sinalizando quando um ponto é maior do que 3 desvios padrão da linha central. O teste 2 detecta deslocamentos na média sinalizando quando 9 pontos em sequência caem no mesmo lado da linha central.

Para avaliar se o uso do teste 2 com o teste 1 melhora a sensibilidade das cartas da média (carta I e carta Xbar), estabelecemos limites de controle para uma distribuição normal (0, SD). Nós deslocamos a média da distribuição por um múltiplo do desvio padrão e, em seguida, registramos o número de subgrupos necessários para detectar um sinal para cada uma das 10.000 iterações. Os resultados são apresentados na Tabela 2.

**Tabela 2** Número médio de subgrupos até a falha de um teste 1 (Teste 1), falha de um teste 2 (Teste 2) ou falha do teste 1 ou teste 2 (Teste 1 ou 2). O deslocamento na média é igual ao múltiplo do desvio padrão (SD) e a simulação foi realizada para tamanhos de subgrupo (n) = 1, 3 e 5.

Turno	n=1			n=3			n=5		
	Teste 1	Teste 2	Teste 1 ou 2	Teste 1	Teste 2	Teste 1 ou 2	Teste 1	Teste 2	Teste 1 ou 2
0,5 SD	154	84	57	60	31	22	33	19	14
1 SD	44	24	17	10	11	7	4	10	4
1,5 SD	15	13	9	3	9	3	1,6	9	1,6
2 SD	6	10	5	1,5	9	1,5	1,1	9	1,1

Como visto nos resultados para a carta I (n = 1), quando ambos os testes são usados (coluna do Teste 1 ou 2) são necessários, em média, 57 subgrupos para detectar uma mudança de 0,5 desvio padrão da média, em comparação com uma média de 154 subgrupos necessários para detectar uma mudança de 0,5 desvio padrão quando o teste 1 é usado sozinho. Da mesma forma, a utilização de ambos os testes aumenta a sensibilidade para a carta Xbar (n = 3, n = 5). Por exemplo, para um subgrupo de tamanho 3, são necessários, em média, 22 subgrupos para a detecção de um deslocamento de 0,5 desvio padrão, quando ambos os testes 1 e 2 são utilizados, enquanto são necessários 60 subgrupos para detectar um deslocamento de 0,5 desvio padrão, quando o teste 1 é utilizado sozinho. Portanto, a utilização de ambos os testes aumenta significativamente a sensibilidade para detecção pequenos deslocamentos na média. À medida que o tamanho dos deslocamento aumenta, a adição do teste 2 não aumenta significativamente a sensibilidade.

## Simulação B2: Qual a eficiência do teste 7 na detecção da estratificação (várias fontes de variabilidade nos subgrupos)?

O teste 7 normalmente sinaliza uma falha quando um número de pontos entre 12 e 15 em sequência caem dentro de 1 desvio padrão da linha central. O Assistente utiliza uma regra modificada que se ajusta ao número de pontos necessários com base no número de subgrupos dos dados. Nós estabelecemos  $k = (\text{número de subgrupos} \times 0,33)$  e definimos os pontos em sequência necessários para uma falha do teste 7, como mostrado na Tabela 3.

**Tabela 3** Pontos em sequência necessários para uma falha no teste 7

<b>k = (número de subgrupos) x 0,33</b>	<b>Pontos necessários</b>
$k < 12$	12
$k \geq 12$ e $k \leq 15$	Inteiro $\geq k$
$k > 15$	15

Usando cenários comuns para o estabelecimento de limites de controle, foi realizada uma simulação para determinar a probabilidade de que o teste 7 indique uma falha utilizando os critérios acima. Especificamente, queríamos avaliar a regra para a detecção de estratificação durante a fase em que os limites de controle são calculados a partir dos dados.

Selecionamos aleatoriamente  $m$  subgrupos de tamanho  $n$  de uma distribuição normal com um desvio padrão (SD). Metade dos pontos em cada subgrupo teve uma média igual a 0 e a outra metade teve uma média igual ao deslocamento SD (0 SD, 1 SD ou 2 SD). Foram realizadas 10.000 iterações e registrada a porcentagem de cartas que mostraram pelo menos um sinal do teste de 7, como mostrado na Tabela 4.

**Table 4** Porcentagem de cartas que tiveram pelo menos um sinal do teste 7

<b>Número de subgrupos</b>		<b>m = 50</b>	<b>m = 75</b>	<b>m = 25</b>	<b>m = 38</b>	<b>m = 25</b>
<b>Tamanho do subgrupo</b>		<b>n = 2</b>	<b>n = 2</b>	<b>n = 4</b>	<b>n = 4</b>	<b>n = 6</b>
<b>Teste</b>		<b>15 em sequência</b>	<b>15 em sequência</b>	<b>12 em sequência</b>	<b>13 em sequência</b>	<b>12 em sequência</b>
<b>Turno</b>	<b>0 SD</b>	5%	8%	7%	8%	7%
	<b>1 SD</b>	23%	33%	17%	20%	15%
	<b>2 SD</b>	83%	94%	56%	66%	50%

Como pode ser visto na primeira linha Deslocamento da tabela (deslocamento = 0 SD), quando não há nenhuma estratificação, uma porcentagem relativamente pequena de cartas tem, pelo menos, um teste 7 com falha. No entanto, quando existe estratificação (deslocamento = 1 SD ou deslocamento = 2 SD), uma porcentagem muito maior de cartas

(na ordem de 94%) tem pelo menos uma falha no teste 7. Desta forma, o teste 7 pode identificar a estratificação na fase em que os limites de controle são estimados.

# Anexo C: Quantidade de dados

## Fórmula C1: Número de observações necessárias com base em um IC de 95% para o limite de controle superior

Quando os dados são normalmente distribuídos e não existe nenhum erro devido à variabilidade amostral, a porcentagem de observações acima do limite superior de controle é 0,135% para um processo sob controle. Para determinar se há observações suficientes para garantir que a taxa de falsos positivos fique razoavelmente baixa, seguimos Bischak (2007) e calculamos um limite inferior de 95% de confiança para o limite superior de controle da seguinte forma:

$$LSE_L = \hat{\mu} + \frac{h}{w_p}$$

em que

$\hat{\mu}$  = estimativa da linha central

$h = 3 \sigma$

$$w_p = 1 - \frac{z_p \sqrt{1 - c_4^2}}{c_4 \sqrt{\frac{n}{k}}}$$

$c_4$  = constante não viciada para uma carta Xbar

$n$  = número de observações

$k$  = tamanho do subgrupo

$z_p$  = fda inversa avaliada em  $p = 0,95$  para a distribuição normal com média = 0 e desvio padrão = 1

Nós definimos  $\frac{h}{w_p} = 2,65 \sigma$ , o que produz uma taxa de falsos positivos de 0,5% acima do limite de controle e resolve para  $n$ .

Para a carta I, estimamos  $\sigma$  usando uma amplitude móvel média com comprimento 2; portanto, bastou usar o número de observações da carta Xbar com tamanho do subgrupo igual a 2 para determinar a quantidade de dados necessária. Com base nestes cálculos, o número de observações necessárias para diferentes tamanhos do subgrupo são mostrados na Tabela 5.

Tabela 5 Número de observações necessárias para vários tamanhos de subgrupo

Tamanho do subgrupo	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
Número de observações	114	114	84	72	70	66	63	64	63	60	66	60	65	56	60

**Observação** O número de observações deve diminuir à medida que o tamanho do subgrupo aumenta. No entanto, as exceções a esta regra aparecem na Tabela 8. Estas exceções ocorrem porque o número de subgrupos foi arredondado para o número inteiro seguinte antes de ser multiplicado pelo número de observações em cada subgrupo, a fim de calcular o número total de observações necessárias.

Os resultados da Tabela 5 mostram que o número total de observações necessárias é menor ou igual a 100 para todos os tamanhos de subgrupos, exceto quando o tamanho do subgrupo é 1 ou 2. No entanto, mesmo quando o tamanho do subgrupo é 1 ou 2, a taxa de falsos positivos é de apenas cerca de 1,1%, com 100 observações. Portanto, 100 observações é um valor de corte eficaz para todos os tamanhos de subgrupos.

A análise acima assume que cada subgrupo terá a mesma quantidade de variação de causa comum. Na prática, os dados coletados em diferentes pontos no tempo podem ter diferentes quantidades de variabilidade de causa comum. Portanto, talvez você queira extrair amostras do processo em mais pontos no tempo do que o necessário para aumentar as chances de ter uma estimativa representativa da variação do processo.

# Anexo D: Autocorrelação

## Simulação D1: Como a autocorrelação afeta a taxa de falsos positivos

Definimos um processo autocorrelacionado usando o modelo a seguir:

$$x_t = \mu + \phi (x_{t-1} - \mu) + \varepsilon_t$$

em que

$x_t$  = Observação no tempo t

$\mu$  = média do processo

$\phi$  = correlação entre dois pontos consecutivos

$\varepsilon_t$  termo de erro no tempo t;  $\varepsilon \sim \text{Normal}(0, 1)$

Para simplificar, consideramos somente a autocorrelação para lag 1 porque é provável que a autocorrelação para lag 1 que seja maior do que autocorrelação para lags  $\geq 2$ .

Utilizando-se o modelo acima, foi realizada uma simulação com  $\phi = 0,2, 0,4, 0,5, 0,6$  e  $0,8$  e com três tamanhos de subgrupos ( $n = 1, 3$  e  $5$ ). Usamos um conjunto inicial de 10.000 subgrupos para estabelecer os limites de controle e, em seguida, registramos a porcentagem de falsos positivos por mais de 2.500 subgrupos. Realizamos 10.000 iterações e registramos o percentual médio de falsos positivos que utilizam os testes 1 e 2 para cada combinação de autocorrelação e tamanho de subgrupo. Os resultados são apresentados na Tabela 6 abaixo.

**Tabela 6** Porcentagem média de falsos positivos para o teste 1, teste 2 para um processo com autocorrelação ( $\phi$ ) e tamanho de subgrupo ( $n$ )

Tamanho do subgrupo	$\phi$				
	0,2	0,4	0,5	0,6	0,8
<b>n = 1</b>	0,73, 1,08	2,01, 2,99	3,38, 4,98	5,77, 8,34	17,94, 23,83
<b>n = 3</b>	1,48, 0,57	5,35, 0,98	9,07, 1,39	14,59, 2,19	33,91, 7,91
<b>n = 5</b>	1,49, 0,50	5,48, 0,68	9,40, 0,86	15,24, 1,18	35,52, 3,84

Observe que, conforme o valor de  $\phi$  aumenta, a taxa de falsos positivos aumenta. Na verdade, até mesmo níveis moderados de autocorrelação ( $\phi = 0,4$ ) levam a aumentos significativos na taxa de falsos positivos. Quando a autocorrelação é  $\geq 0,4$ , o gráfico de controle mostra uma grande porcentagem de falsos positivos, de forma que os resultados da carta são quase insignificantes.

## Fórmula D1: Teste para autocorrelação

Para determinar o nível de  $\phi$  de autocorrelação nos dados, seguimos Bowerman (1979) e realizamos um teste para avaliar a hipótese nula de  $\phi = r$  contra a hipótese alternativa de  $\phi > r$  usando a seguinte estatística de teste:

$$z = \frac{(\hat{\phi} - r)}{se_{\hat{\phi}}}$$

em que

$$\hat{\phi} = \frac{1}{m \hat{\sigma}} \sum_{i=1}^{m-1} (x_i - \mu)(x_{i+1} - \mu)$$

$m$  = número de observações

$\hat{\sigma}$  = variância da amostra

$$se_{\hat{\phi}} = \sqrt{m}$$

Ao calcular a autocorrelação dos dados em cartas Xbar, não incluímos a dependência entre o último ponto de um subgrupo e o primeiro ponto do subgrupo seguinte porque a inclusão dessas comparações levaria a uma subestimação da autocorrelação. Seguindo Bowerman (1979), calculamos o valor-p com base em uma distribuição normal padrão.