
CEPpt: Um pacote R para o Controle Estatístico de Processo

Rafael L. Bastos^{1†}, Eric B. Ferreira²

¹ Mestrando em Estatística e Experimentação Agropecuária, Universidade Federal de Lavras.

² Professor Adjunto III, Instituto de Ciências Exatas, Universidade Federal de Alfenas.

Resumo: *A utilização de pacotes computacionais tem sido cada vez mais frequente no meio acadêmico devido, principalmente, aos avanços tecnológicos. Na campo de ensino do Controle Estatístico de Processo não poderia ser diferente, pois esses pacotes permitem aos acadêmicos fixarem-se mais na análise e interpretação de resultados. Para auxiliá-los, a representação gráfica computacional é essencial por economizar tempo e esforço nos procedimentos. No Controle Estatístico de Processo, uma das principais ferramentas de monitoramento é o gráfico de controle. Com ele é possível detectar se um processo está sob controle estatístico e possíveis padrões de não-aleatoriedade. O presente trabalho apresenta a criação de um pacote estatístico chamado CEPpt (Controle Estatístico de Processo - Português), desenvolvido em linguagem R, um ambiente computacional livre, dando opções aos usuários de construir os mais importantes gráficos de controle. Para construir os gráficos de controle, nesse trabalho, foram utilizados dados fictícios. Pode-se concluir que o pacote CEPpt é uma ferramenta alternativa que alcançou os objetivos de ser versátil, dinâmico e promissor.*

Palavras-chave: Software R, Controle Estatístico de Processo, Gráficos de Controle, Pacote CEPpt.

Abstract: *The use of computer packages have been increasingly common in academia, mainly due to technological advances. In the field of teaching Statistical Process Control could not be different, because these packages allow students to stick to the analysis and interpretation of results. To assist them, the graphical representation is essential for saving time and effort in the proceedings. In Statistical Process Control, one of the main monitoring tools is the quality control chart. It's possible to detect whether a process is under statistical control and possible patterns of non-randomness. This work presents the creation of a statistical package called CEPpt (Statistical Process Control - Portuguese), developed using R, an open computing environment, giving users the options to build the most important control charts. To construct the control charts, in this work, we used dummy data. It can be concluded that the CEPpt package is an alternative tool and achieved the objectives to be versatile, dynamic and promising.*

Keywords: R Software. Statistical Process Control. Control Charts. CEPpt Package.

[†]Autor correspondente: ie12702@yahoo.com.br.

Introdução

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. Hoje em dia, muitas fábricas no mundo aplicam pelo menos algumas ferramentas simples do *CEP* para melhoria dos processos industriais (MONTGOMERY, 2009).

Uma das principais e mais sofisticadas ferramentas do *CEP* é o Gráfico de Controle. Com ele é possível monitorar e rastrear toda variabilidade do processo, sendo possível detectar se o mesmo está sob controle estatístico - por meio dos padrões de não-aleatoriedade. Um processo é considerado fora de controle - ou instável - quando são detectados padrões de não-aleatoriedade. De acordo com Nomelini, Ferreira e Oliveira (2009), no processo sob controle estatístico, a característica de qualidade do conjunto dos itens produzidos possui distribuição normal e quando a variabilidade se torna não-normal, as amostras indicarão que o processo de fabricação se modificou e ficou fora de controle.

Os gráficos que estão entre as mais importantes e úteis técnicas de controle e monitoramento de processos são: média amostral, medida individual, média móvel, amplitude móvel, amplitude e desvio padrão. Eles são geralmente utilizados em pares, sendo um gráfico para uma medida de posição e outro para uma medida de variabilidade. Por exemplo: média amostral e amplitude, média amostral e desvio padrão, medida individual e amplitude móvel etc. De acordo com Borges (2009), o uso de pares se justifica no fato de que pode haver presença de causas especiais em relação ao parâmetro de posição e não em relação ao parâmetro de dispersão, ou vice-versa.

Os limites externos do gráfico de controle para a medida de posição escolhida são funções da medida de variabilidade que compõe o par. Por exemplo, nos gráficos da *média amostral utilizando a amplitude*, calculam-se o Limite Inferior de Controle (*LIC*) e o Limite Superior de Controle (*LSC*) do gráfico da média amostral utilizando a amplitude.

Segundo Montgomery (2009), o uso dos gráficos de controle apresenta várias vantagens, tais como redução de custos, aumento da produção, prevenção de defeitos, fornece informações de diagnóstico, estima a capacidade do processo em atender às especificações do projeto, entre outras. Vale notar que o uso incorreto dos gráficos de controle poderá conduzir estimativas errôneas e interferências desnecessárias no processo. Segundo Borges (2009), o fator principal para usar incorretamente os gráficos de controle é o desconhecimento das técnicas e ferramentas estatísticas, não compreendendo a informação contida no padrão de pontos apresentado por eles.

A utilização de pacotes computacionais tem sido cada vez mais frequente no meio acadêmico (HENNING; ALVES; VIEIRA, 2007). Não poderia ser diferente no campo do ensino de Controle Estatístico de Processo que acompanha os avanços da tecnologia para atender à complexidade dos atuais processos. Houve grande evolução nesta área, na medida em que os pacotes permitiram aos acadêmicos fixarem-se mais na análise e interpretação de resultados.

A função `qcc()` (SCRUCCA, 2004), componente do pacote de mesmo nome, desenvolvida no software R (R CORE TEAM, 2012) por Lucca Scrucca do Departamento de Ciências Estatísticas da “Università degli Studi di Perugia”, Itália. O autor desenvolveu este pacote computacional para proporcionar aos seus alunos de graduação uma ferramenta para a aprendizagem dos conceitos básicos de *CEP* permitindo a construção dos gráficos de controle, exibindo os pontos médios observados e as linhas de controle.

O pacote *CEPpt* (BASTOS, 2011) foi desenvolvido no software R por Rafael Lemos Bastos em seu Trabalho de Conclusão de Curso na Universidade Federal de Alfenas, Brasil. O presente autor desenvolveu este pacote computacional para proporcionar aos usuários uma ferramenta para aprendizagem dos conceitos básicos de *CEP* que permite a construção dos mais importantes gráficos de controle, exhibe os pontos médios observados ao longo dos subgrupos racionais, as linhas de controle e linhas que delimitam as zonas para melhor compreensão dos padrões de não aleatoriedade, com opção de gerar uma lista com todas informações necessárias e detalhadas, caso o usuário deseje fazer uma análise.

O objetivo do presente trabalho é apresentar a criação do pacote CEPpt, desenvolvido no software R, mais dinâmico e completo e compará-lo ao pacote desse mesmo software que constrói gráficos de controle.

Controle Estatístico de Processo

O Controle Estatístico de Processo é um conjunto de sete ferramentas de monitoramento que auxilia no controle de qualidade com a finalidade de reduzir a variabilidade do processo, propiciando níveis melhores de qualidade nos resultados de produção, diminuindo também os custos. Suas sete ferramentas são: Gráfico de ramo e folhas, Folhas de verificação, Gráfico de Pareto, Diagrama de causa e efeito, Diagrama de concentração de defeito, Diagrama de dispersão, Gráfico de Controle.

Uma das principais e mais sofisticadas ferramentas é a Carta de Shewhart (MONTGOMERY, 2009), também conhecida como Gráfico de Controle, que é definida por Werkema (1995) como monitoramento da variabilidade e avaliação da estabilidade de um processo, por meio dos padrões de não-aleatoriedade.

Gráficos de Controle

O gráfico de controle sinaliza a presença de causas especiais quando o processo se encontra fora de controle. Sendo que quando um processo está somente sujeito a ações das causas aleatórias, ou seja, variabilidade natural do processo, diz-se que ele está sob controle, quando, além das causas aleatórias, causas especiais estiverem presentes, diz-se que ele está fora de controle, pois desloca a distribuição da variável aleatória X , tornando a operação anormal. A causa especial pode ser corrigida ou eliminada.

De acordo com Borges (2009), para utilizar o gráfico de controle os dados que serão estudados devem seguir uma distribuição Normal, caso não sigam, deve-se utilizar outras ferramentas estatísticas para monitorar a estabilidade do processo. De acordo com Carvalho e Paladini (2005), os pontos ao longo do tempo dos gráficos de controle são chamados de subgrupos racionais, eles são médias de pequenas amostras representando o estado atual do processo. Estes são geralmente unidos por segmentos de reta para facilitar a visualização da evolução da sequência de pontos ao longo do tempo.

Os gráficos de controle possuem três linhas paralelas: uma linha central (LC) e duas linhas externas, denominadas limite inferior de controle (LIC) e limite superior de controle (LSC).

Para Montgomery (2009), a especificação dos limites é uma decisão crítica, pois:

- Se afastar os limites de controle da linha central, diminui o risco de um erro tipo I (o risco de um ponto cair fora dos limites de controle, indicando uma condição de fora de controle, quando nenhuma causa atribuível está presente) porém, estará aumentando o erro tipo II (o risco de um ponto cair entre os limites de controle quando o processo está, na verdade, fora de controle).
- Se os limites de controle estiverem próximos da linha central, obtém-se o efeito oposto, aumenta-se o risco de um erro tipo I, enquanto o erro tipo II é diminuído.

Por esses motivos, os limites de controle são determinados por meio de um múltiplo do desvio-padrão da estatística que segundo Montgomery (2009), em geral, é 3, ou seja, os limites normalmente empregados são de três-sigma, independentemente do tipo de gráfico usado e a justificativa do uso deste limite é o fato de dar bom resultado na prática definindo entre os dois limites um nível de confiança de 99,73%, esse número significa que um ponto fora de controle pode ocorrer uma vez em 370 subgrupos.

De acordo com Montgomery (2009), os gráficos de controle são geralmente utilizados em pares pois quando se lida com uma variável é necessário um para monitorar o valor médio e outro para monitorar a sua variabilidade, sendo importante manter sob controle ambos.

Os mais importantes gráficos (média amostral, medida individual, média móvel, amplitude móvel, amplitude e desvio padrão) serão descritos a seguir:

Gráfico de controle \bar{X} (Média amostral)

De acordo com Ferreira e Oliveira (2008a) e Montgomery (2009), o gráfico \bar{X} (média amostral) exibe a média de cada subgrupo racional, refletindo a variabilidade no processo ao longo do tempo, sendo que a variabilidade é representada pela amplitude ou pelo o desvio padrão.

Gráfico de controle X (Medida Individual)

De acordo com Borges (2009), o gráfico de controle X é quando o característico a ser medido pode ser expresso por um valor em uma escala contínua de medidas e os subgrupos racionais tem tamanho 1, ou quando as medições estão largamente espaçadas no decorrer do tempo.

Gráfico de controle \bar{X}_m (Média Móvel)

De acordo com Ferreira e Oliveira (2008a), esses gráficos são úteis quando há interesse na tendência do desempenho do processo no decorrer do tempo. Em cada tempo é calculada a média e a trama que liga todas as médias é a média móvel. Esse gráfico é uma alternativa mais apropriada quando a magnitude do deslocamento da média do processo de interesse for pequena.

Gráfico de controle R (Amplitude)

O gráfico R indica a dispersão existente dentro dos subgrupos racionais, ou seja, a variabilidade em um dado instante de tempo, sendo por definição, a diferença entre a maior e a menor observação.

Na Tabela 1 são apresentadas as vantagens e desvantagens do uso do gráfico R :

Tabela 1: Vantagens e desvantagens do uso do gráfico R

Gráfico de controle	Vantagens	Desvantagens
R	- Fácil cálculo; - O estimador R não é viesado, portanto, pode ser adotado em amostras pequenas	- Sensível a mudanças que afetam somente parte do processo, sendo assim menos informativo; - Quanto maior for a amostra, menos eficiente será.

Gráfico de controle S (Desvio Padrão)

De acordo com Ferreira e Oliveira (2008a), quando são adotadas amostras (subgrupos) de tamanhos maiores que 10, o estimador R da variabilidade do processo perde muito em eficiência em relação a S , o desvio padrão do processo, e, mesmo sendo mais simples, não deve ser mais empregado para avaliar a variabilidade do processo. Em seu lugar, utiliza-se o gráfico S .

Na Tabela 2 são apresentadas as vantagens e desvantagens do uso do gráfico S :

Tabela 2: Vantagens e desvantagens do uso do gráfico S

Gráfico de controle	Vantagens	Desvantagens
S	- Não é sensível a mudanças, sendo assim mais informativo; - Quanto maior for a amostra, mais eficiente será.	- Difícil de calcular no “chão de fábrica”; - O estimador S é viesado, sendo necessário, que a amostra seja grande.

Gráfico de controle Rm (Amplitude Móvel)

De acordo com Montgomery (2009), o gráfico de controle para medir a variabilidade quando a medida de posição for o gráfico X , será o gráfico Rm . A amplitude móvel é definida como $Rm_i = |x_i - x_{i-1}|$, sendo $i = 2, 3, \dots, m$ e m o número de subgrupos racionais.

A principal função do gráfico de controle é detectar se o processo está sob controle estatístico por meio dos padrões de não-aleatoriedade.

Padrões de não-aleatoriedade

De acordo com Borges (2009), para verificar se um processo está sob controle estatístico (ausência de causas especiais de variação) utiliza-se os padrões de não aleatoriedade na configuração dos pontos do gráfico de controle. Em 1931, Shewhart propôs um critério para a análise da falta de estabilidade do processo. Posteriormente a Western Electric em 1958 e Nelson em 1985, propuseram critérios suplementares com objetivo de se aumentar a eficácia na detecção de padrões que indiquem a presença de causas especiais de variação.

O pacote CEPpt utiliza os padrões propostos por Nelson, pois é o mais completo contendo os de Shewhart e os da Western Electric. Esses padrões são apresentados por Borges (2009) em forma de oito critérios.

Para melhor compreensão destes critérios, a Figura será utilizada.

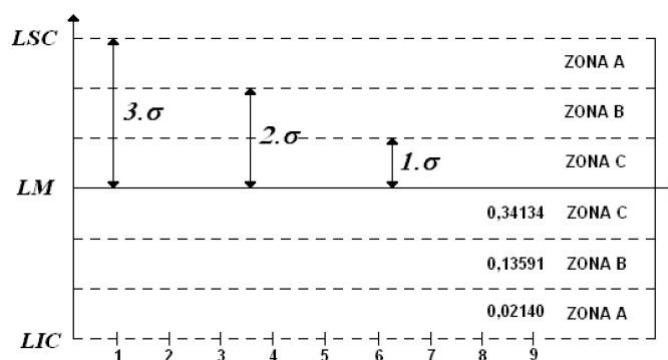


Figura 1: Zonas A, B e C.

Fonte: Adaptado de Borges (2009).

- **Critério 1:** Um ou mais pontos além da zona A, ou seja, acima do LSC ou abaixo do LIC .
- **Critério 2:** Nove pontos consecutivos na mesma metade do gráfico, ou seja, todos acima ou todos abaixo da LM .
- **Critério 3:** Seis pontos consecutivos aumentando ou diminuindo no gráfico.
- **Critério 4:** Quatorze pontos consecutivos alternando-se para cima e para baixo no gráfico.
- **Critério 5:** Dois em três pontos consecutivos na zona A.
- **Critério 6:** Quatro em cinco pontos consecutivos na zona A ou B.
- **Critério 7:** Quinze pontos consecutivos na zona C.
- **Critério 8:** Oito pontos consecutivos de ambos os lados da linha média, nenhum deles na zona C.

Caso pelo menos um critério for detectado no gráfico, o processo estará fora de controle.

De acordo com Montgomery (2009), o critério básico (critério 1) é o proposto por Shewhart, os outros são suplementares, usados para aumentar a sensibilidade dos gráficos de controle a uma pequena mudança no processo, de modo a responder mais rapidamente a uma causa especial.

Além da verificação dos padrões de não-aleatoriedade, os gráficos de controle também são utilizados para encontrar a capacidade do processo.

Capacidade do Processo

De acordo Ferreira e Oliveira (2008a), a capacidade do processo é definida a partir de uma comparação dos limites de especificação, os quais expressam as especificações esperadas do usuário para aquele processo, especificações que o processo pode alcançar ou não, com a realidade do processo, expressa pela média e variabilidade. Há o limite inferior de especificação (*LIE*), o limite superior de especificação (*LSE*) e a linha central de especificação (*LCE*).

C_p e C_{pk} são os índices específicos para esse fim. Quanto maior o valor de C_p , maior a capacidade do processo, sendo que um valor de $C_p > 1$ indica que o processo é capaz de atender às especificações do cliente. O cálculo do índice de capacidade de processo C_p é dado por:

$$C_p = \frac{LSE - LIE}{6\sigma} = \frac{LSE - LIE}{6 \frac{\bar{R}}{d_2}} = \frac{LSE - LIE}{6 \frac{\bar{S}}{c_4}}$$

em que d_2 e c_4 são fatores estatísticos para correção de viés.

O valor do denominador da função do índice C_p é 6σ , isso deve ao fato de ser o limite natural de tolerância que equivale a 99,73%.

A definição do índice C_p pressupõe que o processo está centrado no valor nominal da especificação; caso contrário, a capacidade real do processo é menor do que a indicada por C_p . Nesse caso, convém o uso do índice C_{pk} considerando um ajuste de C_p para o efeito de distribuição não centrada. O índice C_{pk} avalia a distância do valor central do processo aos limites da especificação, tomando aquela que é menor e, portanto, mais sujeita a propiciar resultados fora de especificação. O índice C_{pk} também é conhecido como a capacidade do processo unilateral para o limite de especificação mais próximo da média do processo, seu cálculo é dado por:

$C_{pi} = \frac{\bar{X} - LIE}{3 \frac{\bar{R}}{d_2}} = \frac{\bar{X} - LIE}{3 \frac{\bar{S}}{c_4}}$	$C_{ps} = \frac{LSE - \bar{X}}{3 \frac{\bar{R}}{d_2}} = \frac{LSE - \bar{X}}{3 \frac{\bar{S}}{c_4}}$	$C_{pk} = \min(C_{pi}, C_{ps})$
--	--	---------------------------------

Note que C_{pi} é a capacidade do processo inferior e C_{ps} é a capacidade do processo superior.

Na Figura 2 é apresentado um exemplo ilustrativo do gráfico da capacidade do processo utilizando a amplitude, construído pelo pacote CEPpt:

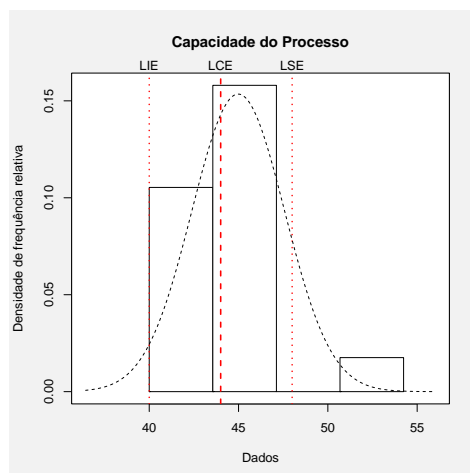


Figura 2: Exemplo de um Gráfico da capacidade do processo gerado pelo pacote CEPpt

Observa-se na Figura 2 que o gráfico da capacidade do processo é composto pela distribuição normal, isso ocorre pois sua interpretação usual se baseia nessa distribuição. Se a distribuição subjacente não é normal, então, o valor de C_p ou C_{pk} podem não ser verdadeiros. O uso do histograma juntamente com a média amostral \bar{X} e a amplitude A pode ajudar na estimação da capacidade de um processo, proporcionando informações imediatas pela impressão visual do desempenho do processo.

Software R

O software estatístico R é um projeto livre para qualquer pessoa ter acesso gratuitamente. Este software pode ser obtido gratuitamente em www.r-project.org, onde é apresentado em versões de acordo com o sistema operacional Windows, UNIX ou Macintosh. Além disso, encontram-se nesse site mais informações sobre a utilização e uma central de correspondências, onde profissionais de vários países podem contribuir para a implementação de novos recursos (PETERNELLI; MELLO, 2007).

Dentre 18 textos brasileiros pesquisados, incluindo artigos, resumos e dissertações que apresentaram ou não, sendo publicados de 2002 a 2009, os softwares utilizados na construção dos gráficos de controle, tem-se o seguinte resultado ilustrado na Tabela 3:

Tabela 3: Pesquisa sobre preferência de software para construção de gráficos de controle no Brasil

Software utilizado	$IC_{95\%}(p)$		
	LI	\hat{p}	LS
EXCEL	10,71%	27,78%	53,59%
STATISTICA	4,407%	16,67%	42,26%
R	1,95%	11,12%	36,07%
MINITAB	0,29%	5,55%	29,37%
CACs	0,29%	5,55%	29,37%
NWA Quality Analyst 5.1	0,29%	5,55%	29,37%
Não menciona	10,71%	27,78%	53,59%

Na Tabela 3, o software Excel é o preferido dos usuários. Considerado um dos mais populares aplicativos de computador até hoje, é um software proprietário. O segundo mais usado é o software STATISTICA, que também é proprietário. O software R é o terceiro mais usado, e como é um software livre, recente e tão completo como os outros, acredita-se que ao decorrer do tempo seja o mais utilizado. Com as vantagens do software R, o pacote CEPpt foi construído utilizando-o.

Pacote CEPpt

O pacote CEPpt (BASTOS, 2011) foi desenvolvido no software R (R CORE TEAM, 2012), sendo um ambiente computacional livre, dando opção ao usuário de construir os mais importantes gráficos do Controle Estatístico de Processo e calcula a capacidade do processo, sendo que seus gráficos e saídas são gerados no idioma Português.

Ele é composto de várias funções capazes de construir os gráficos de controle (Tabela 9) e está disponível no endereço eletrônico: <https://sites.google.com/site/ericbferreira>.

Nesse pacote também contém a função que constrói o gráfico para capacidade do processo, o argumento desta função é *CP*.

A seguir será listada a forma dos dados que cada função do pacote CEPpt recebe e o modelo da lista gerada por ela, caso o usuário queira. Serão apresentados também a linha central e os limites de controle usados na construção dos gráficos de controle deste pacote, essas informações foram retiradas do livro Introdução ao controle estatístico da qualidade (MONTGOMERY, 2009).

Função *Media.A*

Essa função recebe os dados na forma de uma matriz $n \times m$, em que n é o numero de tempos e m é o tamanho dos subgrupos racionais. Após acionar essa função, irá gerar o par de gráficos de controle *Media.A*. O usuário também tem a opção de receber uma lista contendo várias informações importantes, sendo essas o número de subgrupos racionais (n), o tamanho dos subgrupos racionais (m), o limite inferior de controle (*LIC*), a linha central (*LC*), o limite superior de controle (*LSC*) e todos padrões de não-aleatoriedade com suas respectivas frequências. Para receber essa lista, basta dar um nome a essa função e chamá-la com o novo nome, por exemplo:

```
a<-Media.A(dados, 'nome da variável') #gera o par de gráficos
```

Para calcular a linha central e os limites de controle do par de gráfico gerado por essa função, foram utilizados as fórmulas apresentadas na Tabela 4:

Tabela 4: Linha central e os limites de controle do par de gráficos *Media.A*

Gráfico de controle	Linha central	LIC	LSC
\bar{X}	$\bar{\bar{X}}$	$\bar{\bar{X}} - A_2\bar{R}$	$\bar{\bar{X}} + A_2\bar{R}$
R	\bar{R}	$D_3\bar{R}$	$D_4\bar{R}$

Note que $\bar{\bar{X}}$ é a média geral de todas médias dos subgrupos racionais, R é a amplitude, A_2 , D_3 e D_4 são fatores estatísticos para correção de viés.

Função *Media.DP*

Essa função recebe os dados na forma de uma matriz $n \times m$, em que n é o numero de tempos e m é o tamanho dos subgrupos racionais. Após acionar essa função, irá gerar o par de gráficos de controle *Media.DP*. O usuário também tem a opção de receber uma lista contendo várias informações importantes, sendo essas o número de subgrupos racionais (n), o tamanho dos subgrupos racionais (m), o limite inferior de controle (*LIC*), a linha central (*LC*), o limite superior de controle (*LSC*) e todos padrões de não-aleatoriedade com suas respectivas frequências. Para receber essa lista, basta dar um nome a essa função e chamá-la com o novo nome, por exemplo:

```
a<-Media.DP(dados, 'nome da variável') #gera o par de gráficos
```

Para calcular a linha central e os limites de controle do par de gráfico gerado por essa função, foram utilizados as fórmulas apresentadas na Tabela 5:

Tabela 5: Linha central e os limites de controle do par de gráficos `Media.DP`

Gráfico de controle	Linha central	LIC	LSC
\bar{X}	$\bar{\bar{X}}$	$\bar{\bar{X}} - A_3\bar{S}$	$\bar{\bar{X}} + A_3\bar{S}$
S	\bar{S}	$B_3\bar{S}$	$B_4\bar{S}$

Note que \bar{S} é o desvio padrão médio, S o desvio padrão e A_3 , B_3 e B_4 são fatores estatísticos para correção de viés.

Função `MediaM.AM`

Essa função recebe os dados na forma de um vetor. Após acionar essa função, irá gerar o par de gráficos de controle `MediaM.AM`. O usuário também tem a opção de receber uma lista contendo várias informações importantes, sendo essas o número de subgrupos racionais (n), o tamanho dos subgrupos racionais (m) que nesse caso será sempre $m = 1$, o limite inferior de controle (LIC), a linha central (LC), o limite superior de controle (LSC) e todos padrões de não-aleatoriedade com suas respectivas frequências. Para receber essa lista, basta dar um nome a essa função e chamá-la com o novo nome, por exemplo:

```
a<-MediaM.AM(dados, 'nome da variável') #gera o par de gráficos
```

Para construir o par de gráficos de controle `MediaM.AM`, a linha central e os limites de controle de ambos gráficos são apresentados na Tabela 6:

Tabela 6: Linha central e os limites de controle do par de gráficos `MediaM.AM`

Gráfico de controle	Linha central	LIC	LSC
$\bar{X}m$	$\bar{\bar{X}}m$	$\bar{\bar{X}}m + A_2\bar{R}m$	$\bar{\bar{X}}m + A_2\bar{R}m$
Rm	$\bar{R}m$	0	$D_4\bar{R}m$

Note que $\bar{\bar{X}}m$ é a média da média móvel, Rm é a amplitude móvel, $\bar{R}m = \frac{\sum_{j=1}^{m-1} Rm_j}{m-1}$ é a amplitude móvel média.

Função `Medidaind.AM`

Essa função recebe os dados na forma de um vetor. Após acionar essa função, irá gerar o par de gráficos de controle `Medidaind.AM`. O usuário também tem a opção de receber uma lista contendo várias informações importantes, sendo essas o número de subgrupos racionais (n), o tamanho dos subgrupos racionais (m) que nesse caso será sempre $m = 1$, o limite inferior de controle (LIC), a linha central (LC), o limite superior de controle (LSC) e todos padrões de não-aleatoriedade com suas respectivas frequências. Para receber essa lista, basta dar um nome a essa função e chamá-la com o novo nome, por exemplo:

```
a<-Medidaind.AM(dados, 'nome da variável') #gera o par de gráficos
```

Para construir o par de gráficos de controle *Medida.ind.AM*, a linha central e os limites de controle de ambos gráficos são apresentados na Tabela 7:

Tabela 7: Linha central e os limites de controle do par de gráficos *Medida.ind.AM*

Gráfico de controle	Linha central	LIC	LSC
X	\bar{X}	$\bar{X} - 3 \frac{\bar{R}m}{d_2(2)}$	$\bar{X} + 3 \frac{\bar{R}m}{d_2(2)}$
Rm	$\bar{R}m$	0	$D_4 \bar{R}m$

Note que $\bar{R}m = \frac{\sum_{j=1}^{m-1} Rm_j}{m-1}$ é a amplitude móvel média e $d_2(2) = 1,128$.

Função CP

Essa função recebe os dados na forma de uma matriz $n \times m$, em que n é o número de tempos e m é o tamanho dos subgrupos racionais. Após acionar essa função, irá gerar o gráfico da capacidade do processo, vale notar que o usuário tem a opção de usar a medida de variabilidade amplitude ou desvio padrão. O usuário também tem a opção de receber uma lista contendo várias informações, sendo essas o valor da capacidade do processo (Cp), o valor da capacidade do processo (Cpk), o limite inferior de especificação (LIE), a linha central de especificação (LCE) e o limite superior de especificação (LSE). Para receber essa lista, basta dar um nome a essa função e chamá-la com o novo nome, por exemplo:

```
a<-CP(dados, A=TRUE OU FALSE, LIE, LE, LSE, 'nome da variável')
```

em que $A = TRUE$ será quando desejar utilizar a medida de variabilidade amplitude e $A = FALSE$ quando desejar utilizar o desvio padrão.

Todas as funções incluídas no pacote *CEPpt* disponibilizam ao usuário a opção de colocar o nome da variável de estudo, sendo assim inserido no gráfico, lembrando que todos gráficos gerados nesse pacote são no idioma Português.

Na construção dos gráficos de controle do pacote *CEPpt*, foi necessário usar os fatores de correção para estimar os limites de controle.

Fatores de correção de viés

Em muitos trabalhos não são apresentados a expressão dos fatores de correção de viés, são apenas apresentados na tabela com seus valores variando em relação ao tamanho da amostra.

Nesse tópico será apresentado a expressão de d_2 , d_3 e c_4 de acordo com Costa, Epprecht e Carpinetti (2009).

A expressão do fator d_2 é:

$$d_2 = \int_0^{\infty} W f_n(W) dW \quad (1)$$

em que W é a amplitude relativa da amostra e $f_n(W)$ é a função densidade de probabilidade de W .

A expressão de W e $f_n(W)$ é dada a seguir:

$$W = \frac{R}{\sigma}$$

em que R é a amplitude de cada amostra e σ é o desvio padrão da variável X .

$$f_n(W) = n(n-1) \int_{-\infty}^{\infty} [F(X+W) - F(X)]^{n-2} f(X)f(X+W)dX$$

em que n é o tamanho da amostra e $F(X)$ é a distribuição acumulada de X .

A expressão do fator d_3 é:

$$d_3 = \sqrt{\int_0^{\infty} (W - d_2)^2 f_n(W)dW} \quad (2)$$

A expressão do fator c_4 é:

$$c_4 = \frac{\sqrt{2}}{\sqrt{n-1}} \frac{\Gamma(\frac{n}{2})}{\Gamma[\frac{(n-1)}{2}]} \quad (3)$$

em que Γ é a função Gama.

Com as equações (1), (2) e (3) é possível definir os fatores D_3 , D_4 , A_3 , A_4 , B_3 e B_4 , que são apresentados na Tabela 8:

Tabela 8: Fatores de correção de viés

$D_3 = 1 - \frac{3d_3}{d_2}$	$D_4 = 1 + \frac{3d_3}{d_2}$
$A_2 = \frac{3}{d_2\sqrt{n}}$	$A_3 = \frac{3}{c_4\sqrt{n}}$
$B_3 = 1 - \frac{3\sqrt{1-c_4^2}}{c_4}$	$B_4 = 1 + \frac{3\sqrt{1-c_4^2}}{c_4}$

Metodologia

Foi feita uma revisão de literatura sobre o Controle Estatístico de Processo por meio de livros clássicos e básicos e, eventualmente, artigos científicos que verssem de maneira condizente com os propósitos para o desenvolvimento desse pacote.

Todas as funções foram programadas em linguagem R, no software R. O pacote *qcc* e o pacote *IQCC* (RECCHIA; BARBOSA; GONÇALVES, 2010) foram importantes na construção das funções do pacote CEPpt. Por meio do pacote *IQCC* foram retiradas as funções programadas de d_2 , d_3 e c_4 .

As funções foram criadas no idioma português. Para apresentar os gráficos de controle, foram usados dados fictícios para melhor visualizar a presença de padrões de não-aleatoriedade.

Na construção da função CP, foi utilizado o livro Introdução à Estatística Básica com R (FERREIRA; OLIVEIRA, 2008b).

Resultados e discussão

No software R, o pacote que contém a função mais utilizada para construir gráficos de controle é o `qcc` (SCRUCCA, 2004). O pacote `IQCC` (RECCHIA; BARBOSA; GONÇALVES, 2010) e o pacote `qualityTools` (ROTH, 2011) também constroem estes gráficos utilizando dentro de si a função `qcc()`. A seguir, o pacote criado CEPpt será comparado a função `qcc()`.

Pacote CEPpt comparado a função `qcc()`

Os mais importantes tipos de pares de gráficos de controle do Controle Estatístico de Processo são construídos pelo pacote CEPpt, esses são apresentados na Tabela 9:

Tabela 9: Tipos de pares de gráficos de controle gerados pelo pacote CEPpt

Gráficos	Função	Descrição da função
$\bar{X} - R$	Media.A	Gráfico para média e amplitude amostrais
$\bar{X} - S$	Media.DP	Gráfico para média e desvio-padrão amostrais
$\bar{X}_m - R_m$	MediaM.AM	Gráfico para média móvel e amplitude móvel
$X - R_m$	Medidaind.AM	Gráfico para medidas individuais e amplitude móvel

Vale notar que o usuário tem a opção de gerar os gráficos de controle em pares ou separados, caso queira que eles estejam separados. A rotina utilizada para gerar, por exemplo, um par de gráficos de controle `Media.A` separados com sua respectiva lista será apresentada a seguir:

```
a<-Media.A(dados, 'nome da variável', junto=FALSE)
```

Os tipos de gráficos de controle que a função `qcc()` constrói serão apresentados na Tabela 10:

Tabela 10: Tipos de gráficos de controle gerados pela função `qcc()`

Tipo de gráfico	Argumento da função	Estatística
Média	"xbar"	Médias amostrais
Amplitude	"R"	Amplitudes amostrais
Desvio Padrão	"S"	Desvios padrões amostrais
Medida Individual	"xbar.one"	Medidas individuais amostrais
Proporção	"p"	Proporções amostrais
Contagem	"np"	Nº de unidades não-conformes
Contagem	"c"	Não-conformidades por unidade
Contagem	"u"	Não-conformidade média por unidade
Contagem	"g"	Nº de não-eventos entre eventos

A função `qcc()` dispõe somente do gráfico da medida individual, não havendo opção para construir os pares medida individual e amplitude móvel ou média móvel e amplitude móvel, pois quando os dados a serem estudados tiverem subgrupos racionais de tamanho 1, esses são os pares a serem utilizados.

No pacote CEPpt, é possível construir os gráficos para média móvel e amplitude móvel que possibilita os pares de gráficos para medida individual e amplitude móvel ou média móvel e amplitude móvel.

A função `qcc()` dispõe dos gráficos de controle `p`, `np`, `c`, `u` e `g` que o pacote `CEPpt` não dispõe, o motivo é que esses gráficos são poucos utilizados no controle de qualidade. Outra observação é que o gráfico de controle da média amostral (`xbar`) não deixa claro em sua documentação qual medida de variabilidade é utilizada no cálculo dos limites externos. Porém, analisando seu código, percebe-se que o gráfico da média amostral é construído utilizando o desvio padrão quando a amostra é maior que 25, e utilizando a amplitude em caso contrário. Não existe, para o usuário, a opção de construir o gráfico utilizando a medida de variabilidade que desejar.

Com a criação do pacote `CEPpt` (BASTOS, 2011), é possível dar opções ao usuário de construir o gráfico para média amostral utilizando o desvio padrão ou o gráfico para média amostral utilizando a amplitude, o que torna a função mais versátil ao se construir os pares média amostral e amplitude ou média amostral e desvio padrão.

Um exemplo dos gráficos de controle gerado pela função `qcc()`, utilizando medidas individuais de dados simulados, é apresentado na Figura 3:

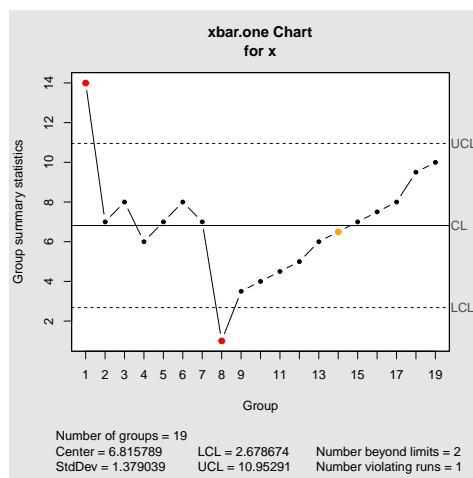


Figura 3: Exemplo do gráfico de controle da função `qcc`

No gráfico de controle da função `qcc()`, observando o modelo na Figura 3, o ponto na cor vermelha significa que o valor daquele subgrupo racional está extrapolando os limites de controle, o ponto na cor amarela significa que algum padrão de não-aleatoriedade foi detectado naquele instante de tempo. Os pontos na cor preta significam que não há irregularidade no processo.

Para visualizar um exemplo dos gráficos de controle do pacote `CEPpt` foi usado o par de gráficos de controle `Media.DP` apresentado na Figura 4, sendo os dados analisados fictícios.

No pacote `CEPpt`, os pontos que estiverem na cor preta significam que naquele tempo o processo está sob controle estatístico, caso o ponto esteja na cor azul indica presença de um ou mais critérios, ou seja, o processo está fora de controle naquele tempo. Os pontos na cor vermelha indicam a presença do critério 1, sendo que também poderá estar ocorrendo outros critérios, podendo ser detectado relacionando o gráfico com a lista gerada.

Na função `qcc()`, é plotado no gráfico o número de subgrupos, o valor da linha central, do limite inferior e superior de controle e o desvio padrão, as frequências dos padrões de não-aleatoriedade da seguinte forma: Number beyond limits (Número além dos limites) e Number violating runs (Número de violações), não apresentando qual tipo de padrão está sendo detectado, sendo mal informativo, pois a análise será visual podendo conduzir a conclusões errôneas.

No pacote `CEPpt`, as informações não são plotadas no gráfico, deixando mais limpo, elas são apresentadas em forma de uma lista que é apresentada caso o usuário desejar, deixando mais informativo e eficiente a análise do processo, pois as frequências dos oito critérios dos padrões de não-aleatoriedade são geradas nessa lista. Para melhor compreensão dos padrões de não-aleatoriedade, nos gráficos são plotados limites que identificam as zonas A, B e C, com seus

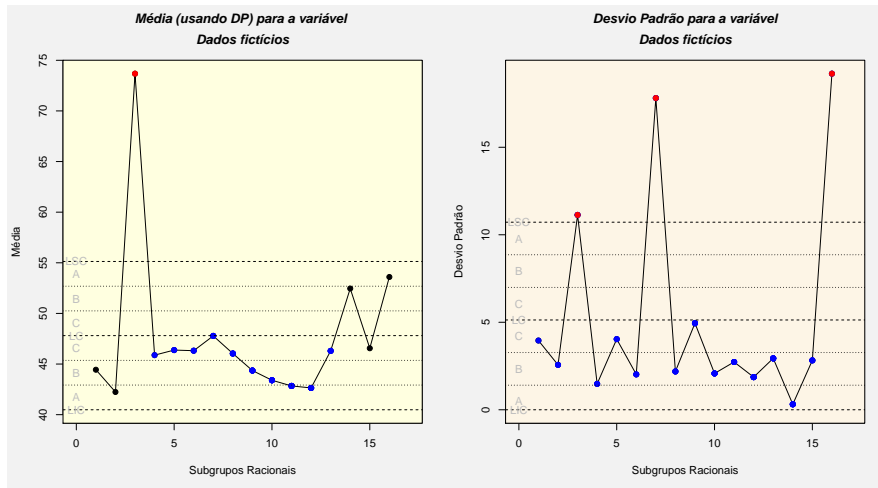


Figura 4: Par de gráficos de controle Média.DP gerado pelo pacote CEPpt

respectivos nomes. Os nomes dos limites de controle também são plotados.

Na Tabela 11 é apresentada a lista gerada do par de gráficos de controle da Figura 4.

Para fazer uma análise correta do par de gráficos de controle da Figura 4, basta relacionar com sua lista gerada e fazer a interpretação dos resultados. A interpretação dos resultados da Figura 4 com sua respectiva lista (Tabela 11) é apresentado a seguir:

No gráfico de controle da média utilizando o desvio padrão foram detectados um ponto acima do *LSC* (critério 1), duas sequências de nove pontos abaixo da *LC* (critério 2), uma sequência decrescente de seis pontos (critério 3), duas vezes dois em três pontos consecutivos na zona A (critério 5), duas vezes quatro em cinco pontos consecutivos na zona A ou B (critério 6), os outros critérios não citados não foram detectados.

No gráfico de controle do desvio padrão foram detectados três pontos acima do *LSC* (critério 1), duas sequências de 9 pontos abaixo da *LC* (critério 2), duas sequências de quatorze pontos alternando-se para cima e para baixo (critério 4) e cinco vezes quatro em cinco pontos consecutivos na zona A ou B (critério 6), os critérios não citados não foram detectados.

Um exemplo de gráfico da capacidade do processo construído pelo pacote *qcc*, especificando os limites da seguinte forma: $LIE = 25$, $LCE = 50$ e $LSE = 75$, é apresentado na Figura 5:

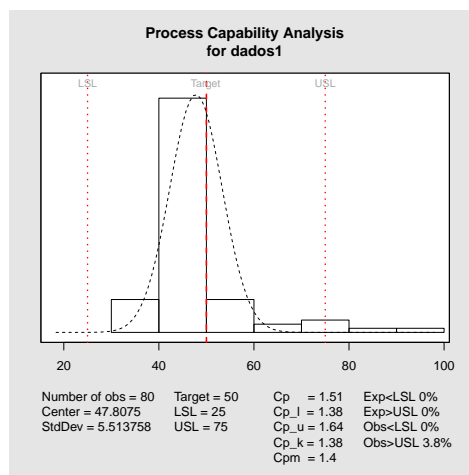


Figura 5: Gráfico da capacidade do processo usando a amplitude construído pelo pacote *qcc*

Tabela 11: Lista gerada pelo pacote CEPpt em relação ao par de gráficos da Figura 4

`Número de subgrupos racionais`	
Gráfico da Média usando Desvio Padrão	Gráfico do Desvio Padrão
16	16
`Tamanho dos subgrupos racionais`	
Gráfico da Média usando Desvio Padrão	Gráfico do Desvio Padrão
5	5
`Limite Inferior de Controle (LIC)`	
Gráfico da Média usando Desvio Padrão	Gráfico do Desvio Padrão
40.48395	0
`Linha Central (LC)`	
Gráfico da Média usando Desvio Padrão	Gráfico do Desvio Padrão
47.8075	5.131053
`Limite Superior de Controle (LSC)`	
Gráfico da Média usando Desvio Padrão	Gráfico do Desvio Padrão
55.13105	10.71876
`PADRÕES DE NÃO-ALEATORIEDADE`	
Nelson.1984	
`CRITÉRIO 1: Pontos acima do LSC ou abaixo do LIC`	
Gráfico da Média usando Desvio Padrão	Gráfico do Desvio Padrão
1	3
`CRITÉRIO 2: Nove pontos, em sequência, abaixo ou acima da LC`	
Gráfico da Média usando Desvio Padrão	Gráfico do Desvio Padrão
2	0
`CRITÉRIO 3: Seis pontos em sequência crescente ou decrescente`	
Gráfico da Média usando Desvio Padrão	Gráfico do Desvio Padrão
1	0
`CRITÉRIO 4: Quatorze pontos consecutivos alternando-se para cima e para baixo no gráfico`	
Gráfico da Média usando Desvio Padrão	Gráfico do Desvio Padrão
0	2
`CRITÉRIO 5: Dois em três pontos consecutivos na zona A`	
Gráfico da Média usando Desvio Padrão	Gráfico do Desvio Padrão
2	0
`CRITÉRIO 6: Quatro em cinco pontos consecutivos na zona A ou B`	
Gráfico da Média usando Desvio Padrão	Gráfico do Desvio Padrão
2	5
`CRITÉRIO 7: Quinze pontos consecutivos na zona C`	
Gráfico da Média usando Desvio Padrão	Gráfico do Desvio Padrão
0	0
`CRITÉRIO 8: Oito pontos consecutivos de ambos os lados da LC, nenhum deles na zona C`	
Gráfico da Média usando Desvio Padrão	Gráfico do Desvio Padrão
0	0

Na Figura 6, é apresentado o gráfico da capacidade do processo (C_p) dos “Dados fictícios” construído pelo pacote *CEPpt* usando o desvio padrão e as medidas especificadas na Figura 5:

Para receber a lista referente ao gráfico da Figura 6, utilizando como medida de variabilidade o desvio padrão, usou-se a seguinte função:

```
a<-CP(dados, A=FALSE, LIE=25, LCE=50, LSE=75, `Dados fictícios`)
```

Na Tabela 12, é apresentada a lista gerada com os seguintes resultados:

No pacote *gcc*, as informações do gráfico da capacidade de processo são plotadas no próprio gráfico (Figura 5), já no pacote *CEPpt*, as informações são apresentadas em forma de uma lista (Tabela 12), o que deixa o usuário com a opção de apresentar somente o gráfico ou ambos.

Observando a lista gerada do gráfico da capacidade de processo utilizando o desvio padrão (Figura 6), verifica-se que a capacidade do processo é de 1,526629, como o valor de $C_p > 1$ indica que o processo é capaz de atender às especificações utilizadas.

Um destaque importante e diferenciado do pacote *CEPpt* é que os gráficos gerados para medidas de posição possuem cores da região interna diferentes dos gráficos gerados para medidas de variabilidade, sendo na cor amarelo luz e renda antiga, respectivamente.

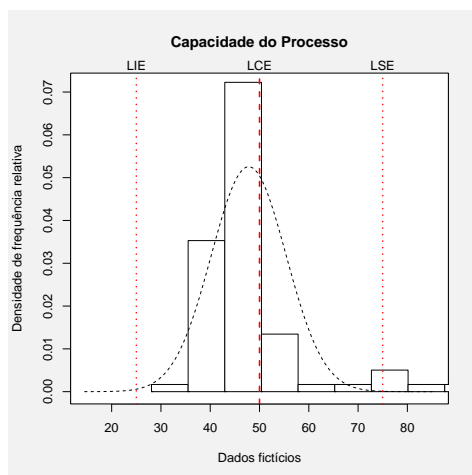


Figura 6: Gráfico da capacidade do processo dos dados fictícios utilizando o desvio padrão

Tabela 12: Lista gerada pelo pacote CEPpt do gráfico da capacidade do processo

```
'Capacidade do Processo (Cp)'  
1.526629  
'Capacidade do Processo (Cpk)'  
1.392743  
'Limite Inferior de Especificação'  
25  
'Linha Central de Especificação'  
50  
'Limite Superior de Especificação'  
75
```

Por fim, o fato de todos os gráficos gerados pela função `qcc()` estarem no idioma inglês limita seu uso entre estudantes brasileiros. Já gerados pelo pacote CEPpt estão no idioma português, (Brasil) sendo facilmente difundidos no ensino na graduação.

Conclusão

O pacote CEPpt mostrou-se promissor, completo e versátil nas análises de dados fictícios, tendo como suas principais vantagens: ter sido implementado em um ambiente computacional livre e gratuito, possibilitando a construção dos mais importantes gráficos de controle; calcula a capacidade do processo; apresenta suas saídas no idioma Português; verifica a ocorrência de padrões de não-aleatoriedade por meio dos oito critérios e gera essas informações na lista que está mais limpa e completa, dando ao usuário informações mais esclarecidas e corretas.

Agradecimentos

O presente artigo foi realizado com apoio da CAPES (Coordenação de Aperfeiçoamento de Pessoal de Nível Superior).

Referências

BASTOS, R. L. *CEPpt: Um pacote R para o Controle Estatístico de Processo*. 2011. 38p. Trabalho de Conclusão de Curso (Graduação em Matemática). Universidade Federal de

Alfenas, Alfenas, 2011.

BORGES, R. C. *Estudo de testes de estabilidade de processo em gráficos de controle de Shewhart*. 2009. 144p. Dissertação (Mestrado em Estatística e Experimentação Agropecuária) - UFLA/FAEPE, Lavras, 2009.

CARVALHO, M. M. DE.; PALADINI, E. P. *Gestão da qualidade: Teoria e casos*. Rio de Janeiro: Elsevier, 5ª reimpressão, 2005. 355p.

COSTA, A. F. B.; EPPRECHT, E. K.; CARPINETTI, L. C. R. *Controle Estatístico de Qualidade*. São Paulo: Atlas, 2.ed. 3ª reimpressão, 2009. 334p.

FERREIRA, E. B.; OLIVEIRA, M. S. DE. *Controle estatístico da qualidade*. Lavras: UFLA/FAEPE, 2008a. 87p.

FERREIRA, E. B.; OLIVEIRA, M. S. DE. *Introdução a estatística básica com R*. Lavras: UFLA/FAEPE, 2008b. 116p.

HENNING, E.; ALVES, C. DA C.; VIEIRA, V. O ambiente R como uma proposta de renovação para aprendizagem e monitoramento de processos em Controle Estatístico de Qualidade. In: XIV SIMPEP (Simpósio de Engenharia de Produção). Bauru-SP. *Anais...* 2007.

MONTGOMERY, D. C. *Introdução ao controle estatístico da qualidade*. 4ªed. LTC, 2009. 513p.

NOMELINI, Q. S. S.; FERREIRA, E. B.; OLIVEIRA, M. S. DE. Estudos dos padrões de não-aleatoriedade dos gráficos de controle de Shewart: um enfoque probabilístico. *Gestão & Produção*, São Carlos, v. 16, n.3, p. 414-421, jul.-set. 2009.

PETERNELLI, L. A.; MELLO, P. DE. *Conhecendo o R: Uma visão estatística*. Viçosa: Ed. UFV, 2007. 181p.

R CORE TEAM. *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria, 2012. ISBN 3-900051-07-0, URL <http://www.R-project.org/>.

RECCHIA, D. R.; BARBOSA, E. P.; GONÇALVES, E. DE J. *IQCC: Improved Quality Control Charts*. R package version 0.5. 2010. URL <http://CRAN.R-project.org/package=IQCC>.

ROTH, T. *qualityTools: Statistical Methods for Quality Science*. R package version 1.49. 2011. URL <http://CRAN.R-project.org/package=qualityTools>

SCRUCCA, L. *QCC: an R package for quality control charting and statistical process control*. R News 4/1, 11-17p. 2004.

WERKEMA, M. C. C. *As ferramentas da qualidade no gerenciamento de processos*. Belo Horizonte: UFMG, v.1. 1995. 108p.