

GRÁFICOS DE CONTROLE ROBUSTO COM R

Nathalia Miranda Correa

Departamento de Engenharia de Produção
Universidade de Estado de Santa Catarina
Joinville, Brasil
mirandacnathalia@gmail.com

Elisa Henning

Departamento de Matemática
Universidade de Estado de Santa Catarina
Joinville, Brasil
elisa.henning@udesc.br

Andrea Cristina Konrath

Departamento de Informática e Estatística
Universidade Federal de Santa Catarina
Florianópolis, Brasil
andrea.ck@ufsc.br

Custodio da Cunha Alves

ENGETEC
Universidade da Região de Joinville
Joinville, Brasil
custodio.alves@gmail.com

Resumo— O Controle Estatístico de Processos (CEP) é um conjunto de ferramentas que se utiliza da aplicação de métodos estatísticos com a finalidade de monitorar processos e fornecer embasamento para estratégias e prevenção de falhas. Se um determinado conjunto possui muitas observações discrepantes, é provável que a acuracidade do gráfico de controle diminua. Estudos sobre os impactos dos *outliers* levaram ao desenvolvimento de novas estatísticas, os estimadores robustos. Com o intuito de aprimorar e expandir o conhecimento relacionado à estatística robusta aplicada à qualidade, este artigo propôs a criação de rotinas no software R com o uso de estimadores robustos em gráficos de controle de Shewhart. Para mostrar os efeitos de tais estatísticas, é usado um conjunto de dados simulado.

Palavras-chave— Gráficos de Controle Robustos. *Outliers*. *M-Estimadores*.

I. INTRODUÇÃO

O Controle Estatístico de Processo (CEP) trata da utilização de métodos estatísticos em todos os estágios considerados críticos ou de muita importância no desenvolvimento de um produto ou processo. O gráfico de controle, ferramenta pertencente ao CEP, foi desenvolvido em 1920 por Walter A. Shewhart, ao aliar os conceitos de estatística à realidade industrial. Hoje em dia, as ferramentas do CEP difundiram-se por diversos ramos, não mais sendo restrito apenas aos processos industriais [1]. Gráficos de controle são aplicados à saúde pública, monitoramento ambiental, detecção de fraudes, entre muitas outras aplicações. O principal objetivo dos gráficos de controle está na detecção dos níveis de qualidade de um processo por meio da análise e monitoramento de sua variabilidade [2].

As observações usadas para a construção dos gráficos são de máximo interesse, pois seu funcionamento depende das características do conjunto de dados. Se um determinado conjunto de dados possui muitas observações discrepantes, é provável que a acuracidade da carta de controle diminua. A

presença de valores discrepantes (*outliers*) diminui a sensibilidade dos gráficos de controle, pois limites de controle ficam mais alargados, o que torna mais difícil a detecção de *outliers* [3]. Estudos vêm sendo desenvolvidos sobre como tratar os *outliers* para que o desempenho dos gráficos de controle estatístico não seja comprometido, por meio de estimadores robustos [3][4][5]. Nesse sentido, este artigo apresenta resultados parciais de uma pesquisa nos quais gráficos de controle robustos foram investigados para aplicação na sinalização de vazamentos em redes de abastecimento e alterações no consumo de água. Nesta tipologia de aplicação, valores discrepantes devem ser adequadamente tratados por representarem elementos de vital interesse para a pesquisa. Apresentam-se neste documento propostas de rotinas em R para tais gráficos.

A partir da presente introdução este trabalho está assim estruturado: inicialmente descrevem-se brevemente os gráficos de controle; na seção seguinte definem-se os estimadores robustos; em sequência estão alguns resultados da proposta e discussão e por fim, as considerações finais.

II. GRÁFICOS DE CONTROLE

Um gráfico de controle é uma ferramenta visual que permite a análise da variabilidade de uma série de observações, geralmente ao longo do tempo. O gráfico de controle mais popular é o de Shewhart e basicamente é composto por uma linha central (LC), que representa a média das amostras; os pontos, que denota as médias dos subgrupos racionais ou valores individuais; um limite superior de controle (LSC) e um limite inferior de controle (LIC). Quando um processo está sob controle estatístico, as observações estarão plotadas entre os limites inferior e superior. Caso algum ponto fique além dos limites ou possua algum padrão (não aleatoriedade), este é interpretado como uma evidência de que o processo pode se encontrar fora de controle estatístico de processo. Isto implica numa investigação mais aprofundada para que possam ser aplicadas medidas corretivas e as causas desta variabilidade sejam corrigidas, se for o caso [1][2].

A. Estimadores Robustos

Para [4], a presença de *outliers* pode impactar em limites de controle mais largos, podendo aumentar a ocorrência de erros tipo I, alarmes falsos, e tipo II, perturbações não detectadas. Os estimadores robustos são estimadores estatísticos insensíveis a pequenos desvios causados por valores discrepantes presentes num conjunto de dados. O que se espera de resultado dessa aplicação é que as estatísticas que são plotadas no gráfico e seus limites de controle sejam pouco sensíveis a *outliers*.

A utilização de estimadores robustos aplicados aos gráficos de controle vem aumentando nas últimas décadas. Exemplificando: foram propostos gráficos de controle para média e amplitude com limites determinados a partir da média aparada [4]; limites determinados a partir da média dos intervalos interquantis [3]. Um modelo de gráfico com uso de M-estimadores foi proposto por [5] e comparados os desempenhos de alguns gráficos com estimadores robustos, entre eles a média aparada e os M-estimadores de Huber e *bisquare*. Ao final, foi observado que o M-estimador obtido a partir da função ρ *bisquare* possuía o melhor desempenho frente aos outros estimadores indicados[5].

III. RESULTADOS

As rotinas no software R [6] são elaboradas para o uso de estimadores robustos em gráficos de controle. Com o objetivo de mostrar os efeitos de tais estatísticas, é utilizado um conjunto de dados simulado e feita a comparação entre o gráfico de controle de Shewhart clássico e aplicando um estimador robusto. Neste documento apresenta-se o gráfico de controle aplicando o M-estimador com a função ρ *bisquare*. É simulado um conjunto de dados com vinte subgrupos de tamanho $n = 5$, contemplando a presença de *outliers*. A proposta de rotina utiliza funções do pacote *qcc* [7] e *smoothest* [8]. Os gráficos resultantes podem ser visualizados nas Figuras 1 e 2. No gráfico de Shewhart clássico (Figura 1) nenhum ponto fora de controle estatístico é observado, apesar do conjunto de dados gerado incluir alguns *outliers*.

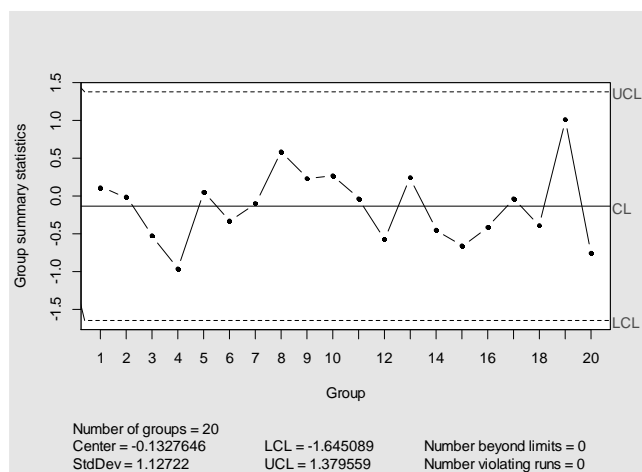


Fig. 1. Gráfico Shewhart clássico para os dados simulados.

Na Figura 2, com um gráfico de Shewhart com M-estimador pode-se visualizar que os *outliers* gerados artificialmente são sinalizados pelo gráfico.

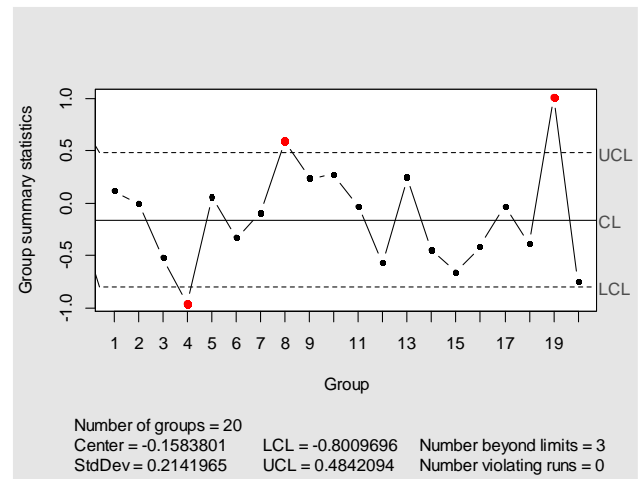


Fig. 2. Gráfico com M-estimador (*bisquare*) para os dados simulados.

As rotinas elaboradas são de fácil implementação e aplicação, portanto, apresentam-se como opções para o monitoramento de processos.

IV. CONSIDERAÇÕES FINAIS

Neste trabalho são propostas rotinas no ambiente R para a construção de gráficos de controle robustos. Apresenta-se uma aplicação a partir de dados simulados de um gráfico no qual é aplicado um M-estimador com a função ρ *bisquare*. Na aplicação é possível visualizar o desempenho superior do gráfico robusto na presença de *outliers*. A continuidade dos trabalhos prevê a aplicação em dados reais.

AGRADECIMENTOS

Os autores agradecem ao CNPq Proc. No. 460310/2014-3, pela concessão do auxílio financeiro e ao programa PIBIC CNPq/UEDESC pela bolsa de iniciação científica.

REFERÊNCIAS

- [1] E. Henning et al, "Aplicação de gráficos de controle estatístico de processos para o monitoramento dos casos de meningite no município de Joinville", *Produção em Foco*, Joinville, v. 2, n. 1, pp.01-26, Nov. 2012.
- [2] M.M. Carvalho et al, *Gestão da qualidade: Teoria e casos*. 10. ed. Rio de Janeiro: Elsevier, 2005.
- [3] D.M. Roche, "Robust Control Charts", *Technometrics*, v. 31, n. 2, p.173-184, Mai. 1989.
- [4] Langenberg, P. & Iglewicz, B, "Trimmed mean X and R charts", *Journal of Quality Technology*, 18, 152-161, 1986.
- [5] A. Maddahi, S. Hamid & A.H. Shokouhi, "A robust X-bar control chart based on M-estimators in presence of outliers", *The International Journal of Advanced Manufacturing Technology*, v. 56, n. 5, p.711-719, Set. 2011.
- [6] R Core Team (2015). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. <https://www.R-project.org/>.
- [7] L. Scrucca, "qcc: an R package for quality control charting and statistical process control", *R News* 4/1, 11-17, 2004.
- [8] C. Hennig, *smoothest: Smoothed M-estimators for 1-dimensional location*. R package version 0.1-2. <https://CRAN.R-project.org/package=smoothest>, 2012.