



Controle estatístico de qualidade usando o método bootstrap

Statistical quality control using the bootstrap method

C. M. Silva^{1,2*}; J. W. Vieira^{3,4}; N. A. A. Neto⁵; G. L. Gonçalves⁵

¹Universidade de Pernambuco, Campus Mata Norte, Departamento de Biologia, Rua Amaro Maltês de Farias, S/N, CEP- 55800-000, Nazaré da Mata, Pernambuco, Brasil

²Mestrado em Gestão do Desenvolvimento Local Sustentável (GDLS)-FCAP-UPE, Av. Sport Club do Recife, 252, bloco C, 4º andar – Madalena – CEP 50740-500 – Recife, Pernambuco, Brasil

³Departamento de Energia Nuclear da Universidade Federal de Pernambuco, Avenida Professor Luiz Freire, 1000, Cidade Universitária, CEP-50740-540, Recife - PE - Brasil. Grupo de Estudos em Radioecologia (RAE)

⁴Instituto Federal de Educação, Ciência e Tecnologia-Campus Recife, CEP-50740-540, Recife-PE, Brasil.

⁵Universidade de Pernambuco, Campus Benfica, Departamento Básico, Rua Benfica, 455, Madalena, CEP-50720-001, Recife, Pernambuco, Brasil

*cleomacio@hotmail.com

(Recebido em 09 de janeiro de 2017; aceito em 09 de março de 2018)

O controle estatístico de qualidade é um procedimento de monitoração e manutenção que utiliza ferramentas estatísticas para avaliar a qualidade de produtos e serviços. Quando uma decisão precisa ser tomada para aceitar ou rejeitar produtos numa linha de produção, verifica-se o funcionamento do processo através da análise estatística do gráfico de controle. Somente assim, será possível verificar se um processo deve ser continuado ou deve ser ajustado para atingir a qualidade total. A ferramenta estatística conhecida como método bootstrap pode ser testada com a finalidade de verificar sua operacionalidade no controle estatístico de qualidade. O método bootstrap tem sido bastante utilizado em diferentes áreas do conhecimento. Este método consiste de procedimentos computacionais intensivos usados para estudar estatísticas paramétrica e não paramétrica. Assim, o objetivo da presente pesquisa foi avaliar o uso do método bootstrap no controle estatístico de qualidade. Para isto, construiu-se um programa computacional em C++. Concluiu-se então, que o método bootstrap mostrou-se bastante eficiente na análise estatística de controle de qualidade, evitando a utilização de um modelo padrão de distribuição de probabilidade.

Palavras-chave: Modelagem estatística, Controle, Processo, Análise estatística

Statistical quality control is a monitoring and maintenance procedure that uses statistical tools to evaluate the quality of products and services. When a decision must be made to accept or reject products on a production line, the process is verified through the statistical analysis of the control chart. Only then will it be possible to verify whether a process must be continued or must be adjusted to achieve full quality. The statistical tool known as bootstrap method can be tested with the purpose of verifying its operability in statistical quality control. The bootstrap method has been widely used in different areas of knowledge. This method consists of intensive computational procedures used to study parametric and nonparametric statistics. Thus, the objective of the present study was to evaluate the use of the bootstrap method in statistical quality control. For this, a computational program was constructed in C ++. It was concluded that the bootstrap method proved to be quite efficient in the statistical analysis of quality control, avoiding the use of a standard model of probability distribution.

Keywords: Statistical modeling, Control, Process, Statistical analysis

1. INTRODUÇÃO

Muitas pessoas já passaram pela experiência de comprar um produto e descobrir que estava com defeito. Isto poderia ser uma mochila nova com um zíper quebrado ou uma impressora de computador com defeito. Como consumidores, esperamos que os produtos que compramos funcionem como pretendido. No entanto, os produtores sabem que nem sempre é possível inspecionar todos os produtos de uma linha de produção ao mesmo tempo. O desafio é projetar formas de maximizar a capacidade de monitorar a qualidade dos produtos que estão sendo

produzidos e eliminar os defeitos. Uma maneira de garantir um produto de qualidade é criar qualidade no processo. Para tanto, é necessário à utilização de ferramentas específicas que possam ajudar na tomada de decisão no controle de qualidade. Essas ferramentas provêm da área de estatísticas e são usadas para ajudar a identificar problemas de qualidade no processo de produção, bem como no próprio produto [1]. O controle estatístico de qualidade (CEQ) é um procedimento padrão que utiliza ferramentas estatísticas, tais como a estatística descritiva, estatística do processo de controle de qualidade e da amostragem por aceitação, no gerenciamento da qualidade total. As ferramentas em cada uma dessas categorias fornecem diferentes tipos de informações para uso em análise estatística de qualidade [1,2].

As estatísticas descritivas são usadas para avaliar certas características de qualidade, como a tendência central e a variabilidade dos dados observados. Embora descrições de certas características sejam úteis, não são suficientes para nos ajudar a avaliar se há um problema com a qualidade. A amostragem de aceitação pode nos ajudar a fazer isso. A amostragem de aceitação nos ajuda a decidir se a qualidade desejável foi alcançada para um lote de produtos e se aceita ou rejeita os itens produzidos. Embora esta informação seja útil na tomada da decisão da aceitação da qualidade, após o produto ter sido produzido, ela não nos ajuda a identificar um problema de qualidade durante o processo da produção. Para tanto, precisamos de ferramentas na categoria do controle estatístico de processo. As três dessas categorias estatísticas de controle de qualidade são úteis para medir e avaliar a qualidade dos produtos ou serviços [1]. No entanto, o controle estatístico de processo é usado com mais frequência, pois identifica problemas de qualidade durante o processo de produção. As ferramentas de controle de qualidade não são utilizadas apenas para medir o valor de uma característica de qualidade, elas também ajudam a identificar uma mudança ou variação em algumas características de qualidade do produto ou processo [2].

Os fatores particulares de variação são detectados por pontos fora dos limites de controle ou pela presença de um comportamento sistemático, não aleatório, no gráfico de controle. Um dos principais objetivos no controle de qualidade é a redução da variabilidade no produto. Os fatores particulares são o motivo principal da falta de uniformidade dos itens produzidos. Só existe um modo seguro e com base científica de se descobrir as causas da variabilidade, ou seja, de identificar fatores particulares de variação, é a aplicação de técnicas estatísticas.

A estatística baseia-se em técnicas úteis para tomar decisões sobre um processo ou uma população com base em uma análise das informações contidas em uma amostra dessa população. Os métodos estatísticos desempenham um papel vital na melhoria e qualidade do controle e melhoria. Eles fornecem o principal meio pelo qual um produto é amostrado, testado e avaliado, e a informação nesses dados é usado para controlar e melhorar o processo e o produto. Além disso, a estatística é o idioma em que profissionais que lidam com o controle de qualidade usam na tomada de decisão sobre um determinado processo [2]. As causas aleatórias podem provocar alterações no desempenho de uma linha de produção. Assim, qualquer intervenção para melhorar a funcionalidade do sistema, deve ser executada somente quando ele estiver sob controle. Caso, isto não seja realizado, fica difícil constatar a causa da variabilidade do processo. A distribuição de probabilidade normal é a mais utilizada no controle estatístico de qualidade. Entretanto, em alguns casos, a distribuição normal apresenta limitações de uso. Assim, o método bootstrap constitui-se uma alternativa viável na inferência do controle estatístico de processo. Como o bootstrap utiliza-se de procedimentos computacionais intensivos com dados gerados pelo método Monte Carlo, é possível fazer diversas análises estatísticas robustas usando pequenas amostras. Isto é uma grande vantagem do uso do método bootstrap na análise estatística de controle de qualidade, principalmente, por questões econômicas. Fazendo uso de um programa construído em C++, o objetivo da presente pesquisa foi estudar a viabilidade do método bootstrap no controle estatístico de qualidade.

2. MATERIAL E MÉTODOS

2.1. CRITÉRIOS DE SELEÇÃO DAS AMOSTRAS

Para testar o programa computacional desenvolvido foram obtidos valores reais de linha de produção industrial de uma empresa localizada na região Metropolitana do Recife. A empresa

colaboradora é do ramo da fabricação de tubos PVC rígido com padrões técnicos para uso em transporte de água fria residencial. Foram obtidos dados de medição do diâmetro externo de seis amostras de tubos de transporte de água residencial, conforme procedimentos técnicos padrões estabelecidos pela Associação Brasileira de Normas Técnicas (ABNT). As amostras foram coletadas diretamente da linha de produção da empresa pelo auditor do processo de controle de qualidade da empresa.

2.2. ANÁLISE ESTATÍSTICAS DOS DADOS

Os dados experimentais foram analisados pelo método bootstrap desenvolvido por Efron (1982) [3]. Este método consiste na reamostragem dos dados originais com repetição. Em cada reamostragem cria-se um conjunto de dados, obtendo a média em cada procedimento. Depois de n interações bootstrap, calcula-se a média das médias. Assim, obtém-se a distribuição bootstrap, de onde são realizadas todas as análises estatísticas. Neste caso, seja o conjunto $X = (x_1, x_2, \dots, x_n)$ os dados obtidos da amostra experimental [4]. Depois dos procedimentos de reamostragens, obtém-se o conjunto de dados da amostra bootstrap $X^* = (x_1^*, x_2^*, \dots, x_n^*)$ com reposição dos dados experimentais. Do conjunto experimental $X = (x_1, x_2, \dots, x_n)$ correspondem estimadores, dados por $\bar{x}_1, \bar{x}_2, \dots, \bar{x}_n$ que formam a distribuição bootstrap [4]. O método bootstrap não é limitado apenas ao cálculo da média, mas a outros parâmetros estatísticos [5]. Computacionalmente, a distribuição bootstrap F é construída pelo método Monte Carlo, com n interações, com repetições. A lei dos Grandes Números garante que o estimador bootstrap é o representante robusto do conjunto de dados experimentais [3]. A Figura 1 mostra a representação do algoritmo construído na linguagem C++ para calcular as estatísticas bootstrap. Neste caso, foram consideradas as seguintes etapas:

- (1) Dos dados experimentais foram realizados sorteios usando gerador de números aleatórios [6] com reposição, obtendo as amostras bootstrap com o mesmo tamanho da amostra experimental.
- (2) Calculou-se a média aritmética em cada etapa de reamostragem.
- (3) O item (2) foi repetido n vezes, até obter a estatística desejada.
- (4) As médias obtidas formam a distribuição bootstrap.
- (5) Determinou-se o estimador bootstrap.

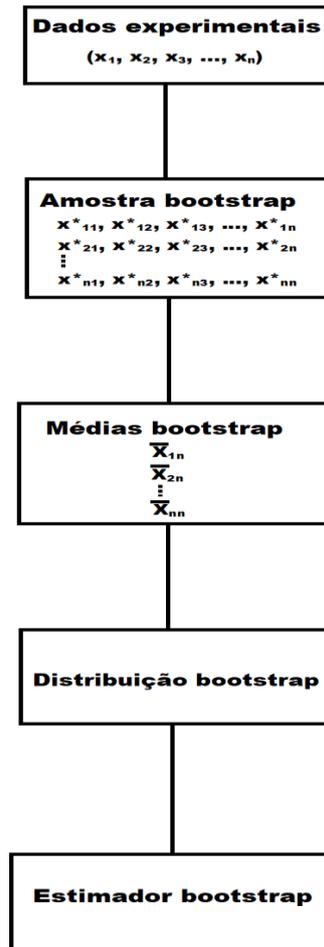


Figura 1: Esquema do algoritmo construído em C++.

3. RESULTADOS E DISCUSSÃO

Na Tabela 1 encontram-se apresentados os valores dos diâmetros externos obtidos da linha de produção de seis amostras de tubos de transporte de água fria residencial coletadas na linha de produção da empresa pesquisada. Por questão de anonimato, o nome da empresa não foi divulgado.

Tabela 1: Valores dos diâmetros externos (mm)

Amostra	Visual	Diâmetro Padrão	Resultado
01	Conforme	274,10	Aprovado
02	Conforme	274,10	Aprovado
03	Conforme	274,10	Aprovado
04	Conforme	274,00	Aprovado
05	Conforme	274,20	Aprovado
06	Conforme	274,20	Aprovado

Fonte: Empresa pesquisada

No relatório do auditor de controle de qualidade da empresa consta que as amostras analisadas atenderam as especificações técnicas da norma da ABNT. O programa computacional de controle

estatísticos dos processos utiliza a distribuição normal no sistema de padronização interna, como valores padrões de diâmetros externos variando de 274,00 a 274,60 mm. Assim, os valores apresentados na Tabela 1 encontram-se no intervalo dos valores padrões. Quando as causas aleatórias não interferem significativamente no processo, pode-se afirmar que ele se encontra estatisticamente controlado, sendo representado pela distribuição de probabilidade normal. Assim, é possível observar os produtos que se encontram fora dos padrões estabelecidos. Neste caso, o gráfico de controle é de fundamental importância na análise estatística do sistema. Com ele é possível estimar o grau de variabilidade. No caso do processo não está sob controle, podemos afirmar que ele é um processo instável, cuja capacidade diverge da padronização do sistema. Todos os parâmetros de padronização usados num controle estatístico de processo são obtidos da distribuição de probabilidade normal. Através da análise do histograma é possível verificar a forma da característica da qualidade em estudo, como também estimar a capacidade independentemente de termos a informação sobre as especificações do produto. Apesar da distribuição de probabilidade ser bastante utilizada em controle estatístico de qualidade, pode ocorrer flutuações causadas por eventos aleatórios. O desempenho de qualquer processo deve ser analisado do resultado de saída. O desempenho de qualquer processo de produção é caracterizado por sua variabilidade interna. Isto faz com que tenhamos informações do próprio sistema em si.

As variabilidades do sistema deveriam ser o alvo supremo dos esforços dos auditores do controle de qualidade. Em muitas situações é possível diminuir as flutuações causadas no sistema, com a finalidade de se obter um produto que atenda todas as especificações técnicas estabelecidas. Para tanto, a estatística desempenha função preponderante na tomada de decisões. Em muitas situações, por questões econômicas, a preparação de grande quantidade de amostras padrões inviabilizam o processo de controle de qualidade. Assim, muitas empresas utilizam apenas procedimentos de controle de qualidade que consideram apenas que os dados padronizados tenham distribuição normal, sem o objetivo de realizar um estudo acurado de todos os fatores que interferem nos procedimentos de preparação dos padrões. Assim, não cara as vezes, o processo de controle de qualidade está sujeito às limitações da quantidade de padrões, como também das causas aleatórias inerentes à preparação. Portanto, é necessário um procedimento estatístico que compense a restrição no número de preparação de amostras padrões, e também reduza significativamente as flutuações estatísticas nas medidas realizadas.

Quando um conjunto de dados é formado por pequenas amostras, a média aritmética obtida é bastante sensível às flutuações estatísticas. Neste caso, o tratamento estatístico fica comprometido. Porém, o método bootstrap pode ser usado de maneira eficiente quando o conjunto de dados é formado por pequenas amostras. Para um número de n iterações bootstrap, a lei dos Grandes Números garante que, se n for suficientemente grande, os estimadores bootstrap possuem distribuição normal [7]. Porém, as medidas de tendências centrais, tais como as médias aritmética, geométrica e a mediana são praticamente iguais na distribuição bootstrap [8]. Tais medidas podem ser convenientemente usadas no controle estatístico de qualidade. A Figura 1 mostra a distribuição dos dados apresentados na Tabela 1. Observa-se na Figura 2 que os dados, realmente tenderam a uma distribuição normal, conforme citado por Costa et al. (2005) [1]. Segundo estes autores, dados coletados de linha de produção tendem à distribuição normal. Daí este tipo de distribuição ser bastante utilizada em controle estatístico de qualidade. Por outro lado, a Figura 3 mostra elevada dispersão no histograma dos dados da Tabela 1, confirmando assim os estudos realizados por Helene e Vanin (2002) [7]. A Figura 4 mostra a distribuição bootstrap das reamostragens dos dados apresentados na Tabela 1, para 50 iterações, conforme estabelecido por Efron (1982) [3]. Observa-se que os dados reamostrados tenderam para a distribuição de probabilidade normal, pois segundo Helene e Vanin (2002) [7], para valores de n suficientemente grande, os estimadores bootstrap convergem para os seus valores verdadeiros. A Figura 5 mostra o histograma da distribuição bootstrap, confirmando assim, a diminuição significativa da dispersão.

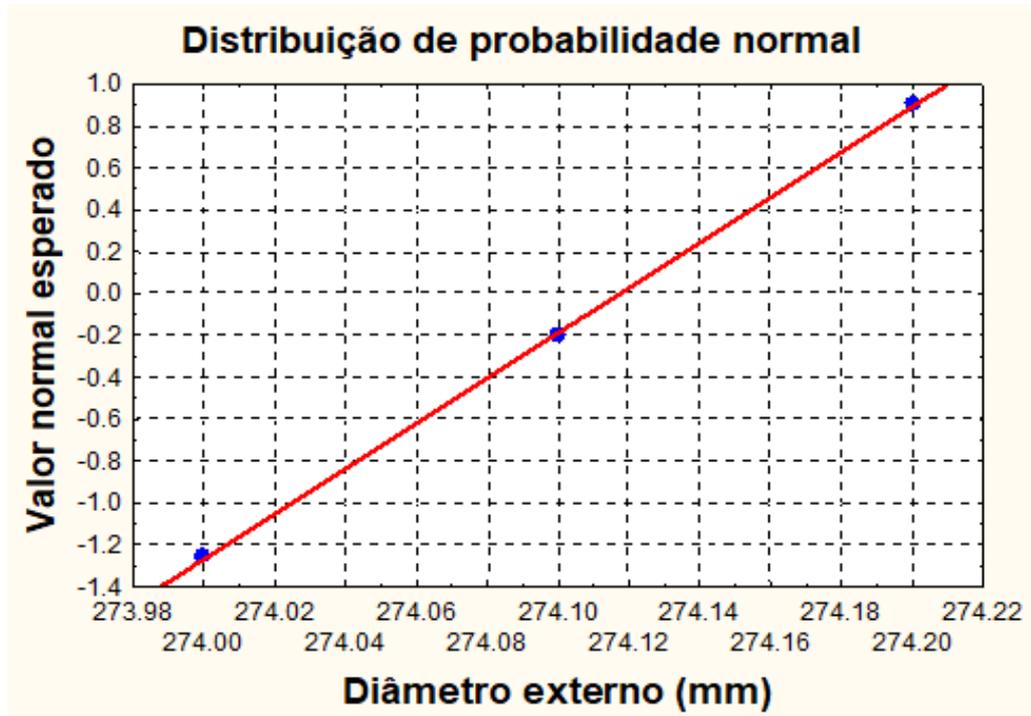


Figura 2: Ajuste dos dados experimentais à distribuição normal.

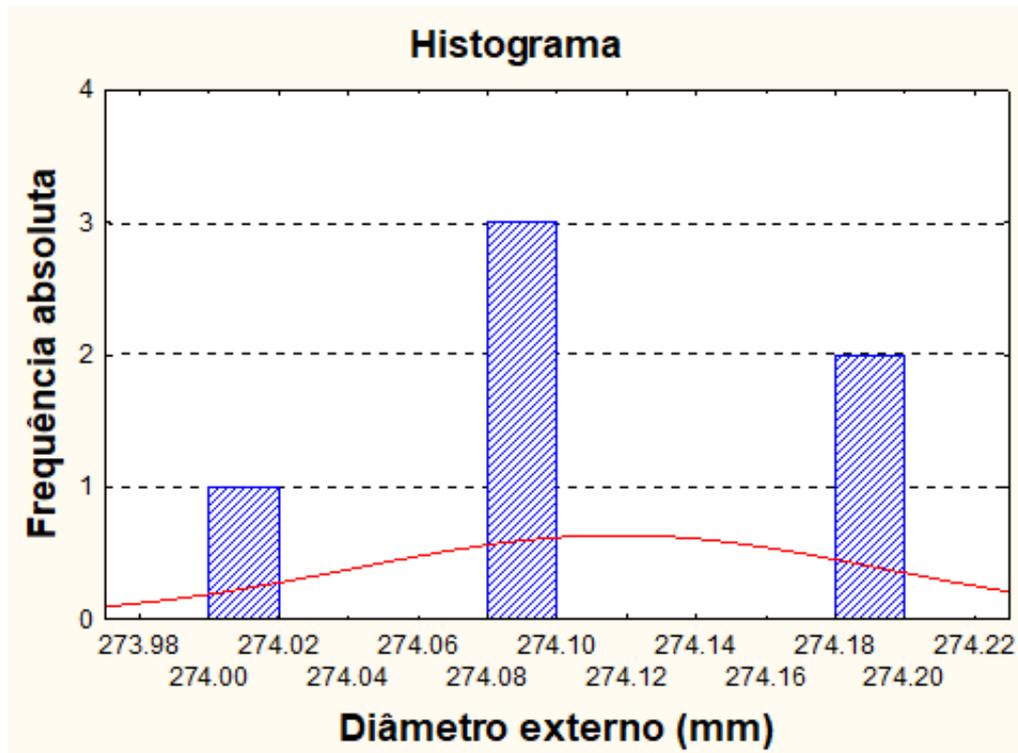


Figura 3: Distribuição dos dados experimentais.

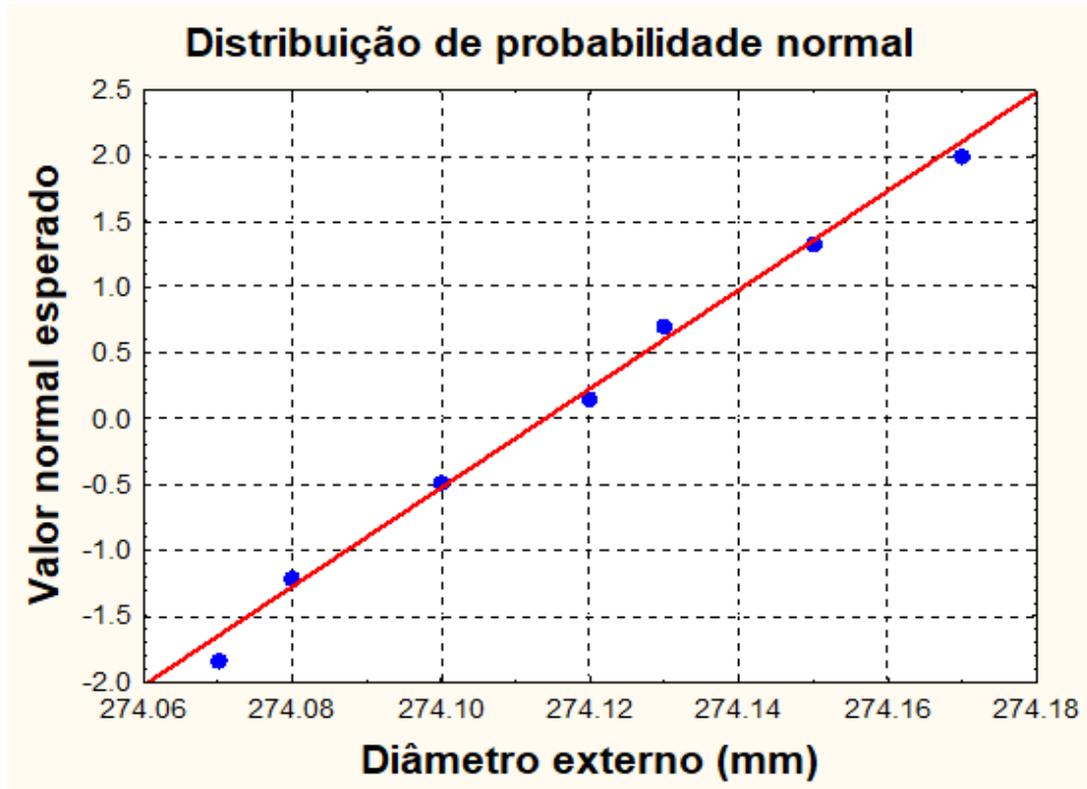


Figura 4: Ajuste dos dados bootstrap à distribuição normal.

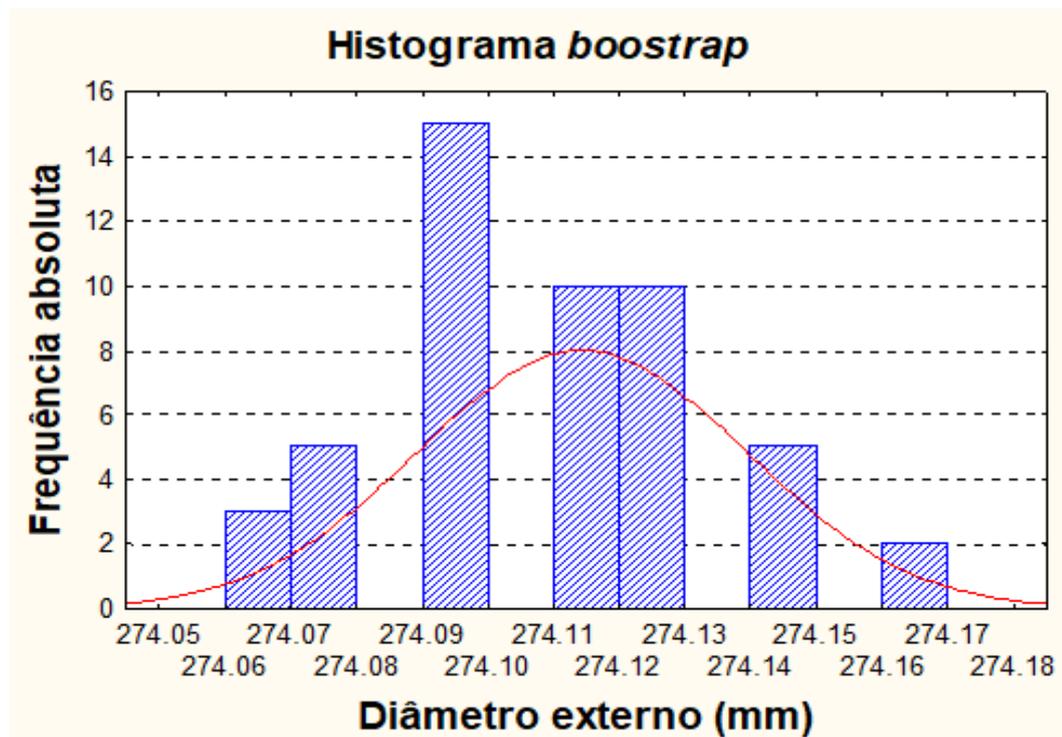


Figura 5: Histograma da distribuição bootstrap.

Na Tabela 2 encontram-se as medidas de tendência central para os valores experimentais e os valores *bootstrap*.

Tabela 2: Medidas de tendência central experimental e bootstrap

Estatística	Diâmetro externo (mm)
Média aritmética experimental \pm desvio padrão	274,12 \pm 0,07
Média geométrica experimental \pm desvio padrão	274,12 \pm 1,00
Mediana \pm desvio quartílico	274,10 \pm 0,08
Média aritmética <i>bootstrap</i> \pm desvio padrão	274,11 \pm 0,02
Média geométrica <i>bootstrap</i> \pm desvio padrão	274,11 \pm 1,00
Mediana <i>bootstrap</i> \pm desvio quartílico	274,10 \pm 0,05

Fonte: Empresa pesquisada

Observa-se na Tabela 2 que, tanto as grandezas de tendência central experimental, como as grandezas bootstrap pertencem ao intervalo de 274,00 a 274,60 mm, como estabelecido pelo sistema de controle de qualidade da empresa. Verifica-se que as médias aritméticas, geométricas e as medianas convergem para 274,12 mm, que é o valor da média aritmética experimental. Por outro lado, a média aritmética bootstrap mostrou-se mais eficiente do que as demais grandezas estatísticas analisadas, em relação à dispersão nos resultados obtidos. Assim, quando se aplica o método bootstrap na análise estatística do controle de qualidade, diminui-se significativamente a variabilidade no conjunto de dados da padronização. Com isto, é possível obter intervalos de confiança mais representativo na tomada de decisão no controle estatístico de qualidade. Na Tabela 3 encontram-se apresentados os valores dos limites inferiores e superiores para a média aritmética experimental e bootstrap.

Tabela 3: Intervalos de confiança para a média aritmética experimental e bootstrap

Estatística	Intervalo de confiança (mm)
Média aritmética experimental	[274,06; 274,18]
Média aritmética <i>bootstrap</i>	[274,07; 274,15]

Verifica-se na Tabela 3 que o intervalo de confiança da média aritmética experimental é mais amplo do que o da média aritmética bootstrap. Logo, o intervalo de confiança da média bootstrap é mais adequado para representar amplitude na sistematização da padronização de controle estatístico de qualidade. Assim, a inserção do método bootstrap no controle estatístico de qualidade, tornará o processo de padronização mais eficiente e menos dispendioso. Segundo Efron (1982) [3], quando um conjunto de dados é formado por pequenas amostras e possui flutuações sem a ocorrência de valores anômalos, como é o caso dos dados apresentados na Tabela 1, os valores de reamostragens bootstrap com 50 iterações são simétricos em relação à média da distribuição. Isto pode ser observado na Figura 6, que representa um sistema de controle de qualidade funcionando adequadamente dentro dos padrões de estabilidade operacional. Um sistema que funciona em condições de estabilidade como o representado pela Figura 6, é de fundamental importância para se ter um rígido controle estatístico de qualidade. Com isto, pode-se afirmar que o método bootstrap é uma ferramenta bastante útil na operacionalidade de sistemas de controle de qualidade.

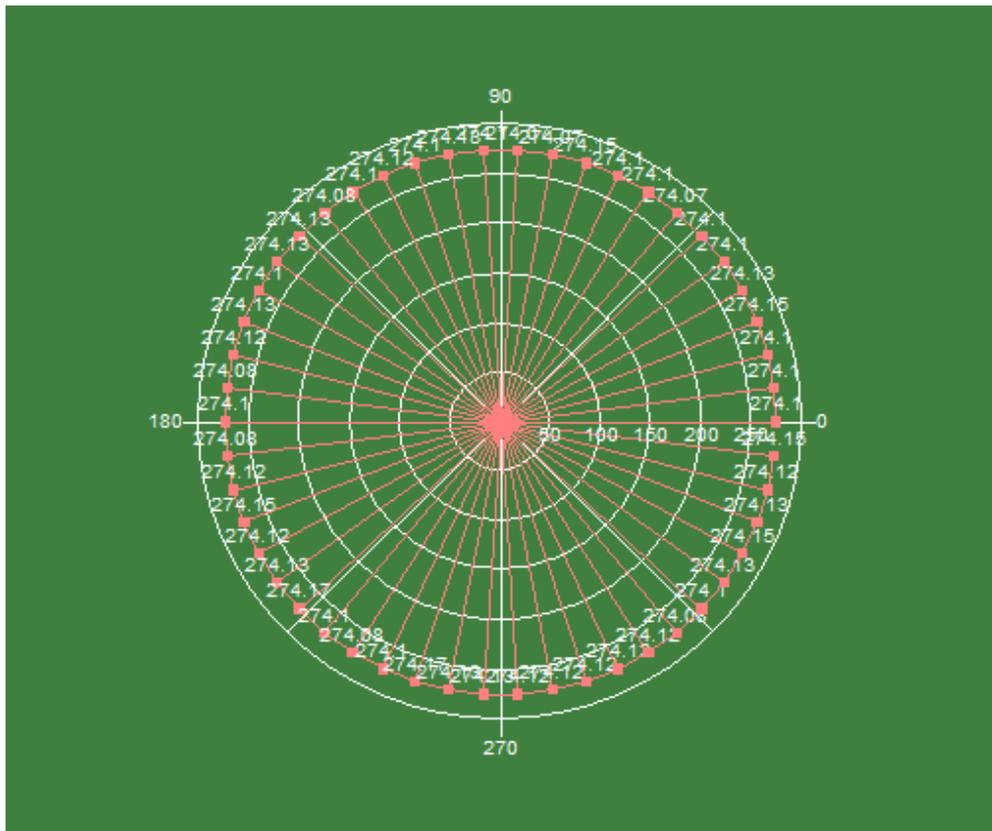


Figura 6: Distribuição de estabilidade das amostras bootstrap.

4. CONCLUSÃO

O método *bootstrap* mostrou-se bastante eficiente na sistematização de padronização de dados estatísticos de controle de qualidade, podendo ser utilizado para diminuir as flutuações estatísticas, como também, é ideal para construir intervalo de confiança.

5. REFERÊNCIAS BIBLIOGRÁFICAS

1. Costa AFB, Epprecht EK, Carpinetti C.R. Controle Estatístico de Qualidade. 2ª Ed. São Paulo: Atlas; 2005.
2. Montgomery DC. Introduction to statistical quality control. 6th edition. Arizona: John Wiley & Sons, Inc.; 2009.
3. Efron B. The jackknife, the bootstrap and other resampling plans. Bristol: J.W. Arrowsmith, Ltd.; 1982.
4. Efron B, Tibshirani RJ. An introduction to the bootstrap. New York: Chapman and Hall; 1993.
5. Efron B. Bootstrap methods: another look at the jackknife. *Annals of Statistics*. 1979;7:1–26.
6. Vieira JW. Construção de um modelo computacional de exposição para cálculos dosimétricos utilizando o código monte carlo egs4 e fantasmas de voxels. Tese de Doutorado do Programa de Pós-graduação em Tecnologias Energéticas e Nucleares da Universidade Federal de Pernambuco, 2004. 146p.
7. Helene O, Vanin VR. Analysis of discrepant data using a bootstrap procedure. *Nuclear Instruments and Methods in Physics Research A*. 2002;481:626-631, doi:10.1016/S0168-9002(01)01377-8.
8. Silva CM, Amaral RS, Vieira JW, Silva ANC, Júnior JAS, Alcoforado E S. Modelagem de tempo de sobrevida via método bootstrap. *Scientia Plena*. 2011;7(10):1-11.