

Análise crítica da aplicação de métodos estatísticos em processos definidos por dados que não apresentam distribuição normal

Celso Francisco de Moraes (UNIFEI) celso1moraes@terra.com.br

João Roberto Ferreira (UNIFEI) jorofe@unifei.edu.br

Pedro Paulo Balestrassi (UNIFEI) pedro@unifei.edu.br

Resumo

Este artigo apresenta uma análise crítica da utilização de determinadas ferramentas da Qualidade e de alguns procedimentos em Engenharia de Produção nos casos em que os dados estatísticos coletados não se apresentam normalmente distribuídos. O problema geral aqui identificado diz respeito à ocorrência de decisões questionáveis tomadas a partir da análise de dados considerados normais. A justificativa primordial para este estudo reside na relevância da análise dos dados durante sua coleta e tratamento em trabalhos científicos nos mais variados níveis. Os principais objetivos deste artigo são: apresentar alguns exemplos relacionados ao Controle Estatístico de Processo e à Metodologia Seis Sigma que possam induzir a conclusões duvidosas devido à adoção incorreta de “normalidade”; propor revisão crítica efetuando as transformações de Box-Cox e/ou Yeo-Johnson; e, finalmente, discutir as conclusões e decisões estabelecidas através da comparação entre os dados originalmente coletados e os dados transformados.

Palavras chave: Métodos estatísticos; Dados não normais, Transformação de dados.

1. Introdução

A proposta deste artigo em analisar a aplicação de alguns métodos estatísticos está inserida em um contexto mais amplo que se refere ao Controle Estatístico da Qualidade e aos aspectos básicos que devem ser observados pelas organizações que buscam a plena satisfação de seus clientes. Segundo Feigenbaum (1994) a Estatística desempenha papel importante nos programas de Controle da Qualidade, pois ao longo dos anos suas técnicas e metodologias tornaram-se cada vez mais amplamente utilizadas e aceitas nas organizações.

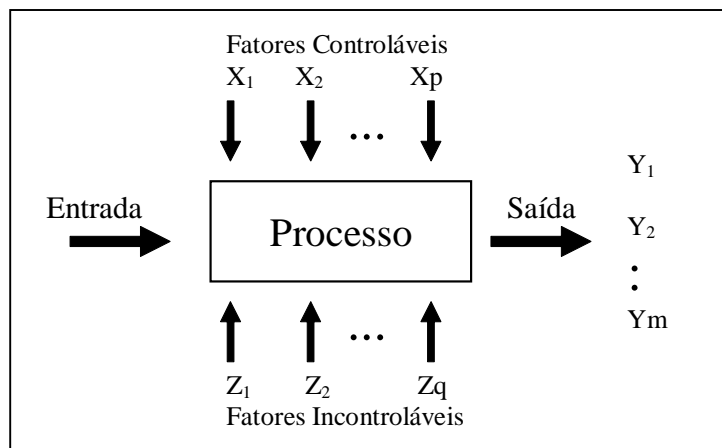
Na visão de Montgomery (1985) o objetivo primário do Controle Estatístico da Qualidade é a redução sistemática da variabilidade nas características-chave do produto, fornecendo as ferramentas necessárias para avaliação e melhoria de processos, produtos e serviços de forma robusta e abrangente. Entretanto, tais ferramentas também dependem da qualidade da informação e da forma como os dados são coletados, processados e interpretados. Por este motivo é de fundamental importância o rigor metodológico na coleta e tratamento dos dados.

No desenvolvimento deste artigo pretende-se analisar a utilização de alguns métodos estatísticos, tais como os gráficos de controle em Controle Estatístico de Processo e avaliação de desempenho na Metodologia Seis Sigma com enfoque na correta interpretação dos dados coletados, especificamente no que se refere à suposição de sua distribuição normal, com revisão crítica através das transformações propostas por Box & Cox (1964) e Yeo & Johnson (2000). As conclusões e decisões estabelecidas serão discutidas por meio de comparação entre os dados originalmente coletados e os dados transformados em distribuição normal.

2. Resumo dos métodos em análise

2.1. Gráficos de controle no Controle Estatístico de Processo (CEP)

Com base na norma NBR ISO 9000:2000, define-se processo como um conjunto de atividades inter-relacionadas ou interativas que transformam insumos (entradas) em produtos (saídas). A Figura 1 exemplifica a situação esquemática de um processo com suas entradas, seus fatores controláveis, seus fatores incontroláveis (ruídos) e suas saídas.



Fonte: Adaptado de Montgomery & Runger (2003).

Figura 1 – Visão esquemática de um processo

Controle Estatístico de Processo (CEP) pode ser definido como um conjunto de ferramentas que tem o propósito de indicar se um processo está funcionando de forma ideal, quando apenas causas comuns de variação estão presentes, ou se este processo está desordenado, e necessita de algum tipo de ação corretiva, ou seja, quando existem causas especiais de variação. De acordo com Montgomery & Runger (2003) o gráfico de controle é a mais poderosa ferramenta do CEP, pois evidencia a variação do processo em relação aos limites de controle estipulados.

Segundo Alwan & Roberts *apud* Toledo (2005) algumas suposições precisam ser satisfeitas para que, independentemente do tipo do gráfico de controle utilizado, os resultados sejam válidos:

- a) É necessário que as observações sejam independentes e identicamente distribuídas, ou seja, que as amostras sejam retiradas de forma aleatória e que o processo que as gerou esteja sob controle estatístico;
- b) As observações devem seguir alguma distribuição de probabilidade específica, tais como Normal, Binomial ou Poisson.

2.2. Avaliação de desempenho na Metodologia Seis Sigma

De acordo com Harry (1998) Seis Sigma é uma metodologia de melhoria em processos com o objetivo de reduzir os defeitos a uma taxa de 3,4 partes por milhão nas características críticas de qualidade para os clientes. Linderman *et al.* (2003) definem Seis Sigma como um método sistemático para melhoria de processos estratégicos, desenvolvimento de novos produtos e

serviços que se baseia em métodos estatísticos e científicos para obtenção de uma dramática redução nas taxas de defeitos.

A Metodologia Seis Sigma é baseada em um sistema de acompanhamento conhecido como DMAIC, sigla que denota as seguintes etapas: Definir (*Define*), Medir (*Measure*), Analisar (*Analyze*), Melhorar (*Improve*) e Controlar (*Control*). Breyfogle (2003) destaca que empresas que seguiram disciplinadamente estas etapas obtiveram significativa vantagem competitiva.

A letra grega minúscula sigma (σ) é o símbolo estatístico para desvio padrão, que é uma medida de variabilidade de um processo. O símbolo σ pode ser usado como um nível de indicação de desempenho, sendo que quanto maior o nível sigma, melhor é o processo. A Tabela 1 apresenta a relação do nível sigma com a quantidade de defeitos por milhão de oportunidades, bem como suas implicações em termos de custos e competitividade.

NÍVEL SIGMA	DPMO (defeitos por milhão de oportunidades)	CUSTO DA NÃO QUALIDADE	CATEGORIA
6	3,4	< 10 % das vendas	Empresa de classe mundial
5	233	10-15 % das vendas	Empresa comum
4	6.210	15-20 % das vendas	
3	66.807	20-30 % das vendas	
2	308.537	30-40 % das vendas	Empresa não competitiva
1	690.000	-	

Fonte: Adaptado de Harry (1998)

Tabela 1 – Referências para o nível sigma

Conforme norma NBR ISO 9000:2000 capacidade é definida como a aptidão de uma organização, sistema ou processo de realizar um produto que irá atender aos requisitos especificados para este produto. Em termos estatísticos, a capacidade e o desempenho de um processo são representados por meio das equações (1), (2), (3) e (4) (BALESTRASSI, 2004):

$$C_P = \frac{LSE - LIE}{6\sigma} \quad (1)$$

$$C_{PK} = \min. \left[\frac{LSE - \mu}{3\sigma}, \frac{\mu - LIE}{3\sigma} \right] \quad (2)$$

$$P_P = \frac{LSE - LIE}{6S} \quad (3)$$

$$P_{PK} = \min. \left[\frac{LSE - \mu}{3S}, \frac{\mu - LIE}{3S} \right] \quad (4)$$

onde:

C_P = índice de capacidade potencial (processo centrado na média das especificações)

C_{PK} = índice de capacidade real (processo não centrado na média das especificações)

P_P = índice de desempenho potencial (processo centrado na média dos dados)

P_{PK} = índice de desempenho real (processo não centrado na média dos dados)

LSE = limite superior de especificação

LIE = limite inferior de especificação

μ = média do processo

S = desvio padrão (amostra)

σ = desvio-padrão do processo (população)

Para Montgomery & Runger (2003) um processo com $C_{PK} = 2$ é referido como um processo seis sigma, pois a distância a partir da média do processo até a especificação mais próxima é de seis desvios-padrão. Se a média do processo se deslocar do centro por 1,5 desvio-padrão, o C_{PK} diminuirá para $4,5\sigma/3\sigma = 1,5$. Considerando um processo distribuído normalmente, a fração não conforme do processo deslocado é de 3,4 partes por milhão. Assim sendo, a média de um processo seis sigma pode se deslocar de 1,5 desvio-padrão do centro das especificações e ainda manter uma fração não conforme de apenas 3,4 partes por milhão.

Boas aproximações podem ser obtidas para determinação do nível sigma através da equação (5) (BALESTRASSI, 2004) e da equação (6) (BREYFOGLE *apud* BALESTRASSI, 2004)

$$\text{Nível Sigma} \approx 3.P_{PK} + 1,5 \quad (5)$$

$$\text{Nível Sigma} \approx 0,8406 + \sqrt{29,37 - 2,221 \cdot \ln(dpmo)} \quad (6)$$

3. Análise da distribuição e transformação de dados

3.1. Considerações gerais

Os métodos estatísticos discutidos neste artigo pressupõem que os dados em estudo sigam uma distribuição de probabilidade conhecida. A análise e as conclusões que resultam da aplicação da metodologia são válidas apenas nos casos onde a suposição da distribuição se confirme correta. Segundo Feigenbaum (1994), independentemente de quão precisos sejam os métodos estatísticos utilizados, na realidade os dados é que são os elementos importantes.

Lipson & Sheth (1973) indicam que na falta de alguma evidência em contrário, pode-se assumir, em uma primeira abordagem, que os dados utilizados em métodos estatísticos sejam normalmente distribuídos. Contudo, Montgomery (1985) alerta que em muitas situações práticas existem razões para se duvidar da validade da suposição de “normalidade”, o que implica em especial atenção na análise dos dados.

Visto que existem situações onde a distribuição de dados não corresponde a uma distribuição normal, então para estes casos tanto a suposição de “normalidade” quanto utilização da curva normal como referência certamente se revelarão inadequadas. Na realidade a existência de “não normalidade” em distribuição de dados é bastante comum, principalmente quando o número de observações não é muito grande.

O Quadro 1 resume o que ocorre com a aplicação de alguns métodos estatísticos quando se assume incorretamente “normalidade” para dados que não apresentam distribuição normal.

MÉTODO ESTATÍSTICO	CONSEQÜÊNCIA DA “NÃO NORMALIDADE”
Controle Estatístico de Processo	Falsas causas especiais detectadas nos gráficos de controle individuais
Seis Sigma	Cálculo incorreto do nível sigma
Teste de Hipóteses	Conclusões incorretas sobre diferenças entre grupos
Análise de Regressão	Identificação equivocada de fatores e erros em predições
Planejamento de Experimentos	Conclusões incorretas sobre importância e efeito dos fatores

Quadro 1 – Conseqüências da suposição incorreta de “normalidade”

Tais fatos demonstram a importância do rigor metodológico na coleta e no tratamento dos dados, pois conclusões e definições questionáveis podem ocorrer devido à incorreta suposição de aderência à distribuição normal. Para evitar este problema é necessário efetuar o teste de “normalidade” dos dados antes de se aplicar os procedimentos metodológicos descritos.

Um exemplo de análise gráfica utilizando o *software* estatístico MiniTab® é mostrado através da Figura 2. No teste de “normalidade” de Anderson-Darling considera-se normal a distribuição que apresentar *p-value* maior que 0,05, o que significaria uma probabilidade maior que 5% em cometer erro ao rejeitar a hipótese de “normalidade” da distribuição em análise. Os dados referem-se ao tempo médio de atendimento (em minutos) extraído diariamente ao longo de um ano em uma empresa fictícia de prestação de serviços denominada “Gamma 220 Ltda”. Neste caso, como *p-value* é menor que 0,05, a distribuição não pode ser considerada normal.

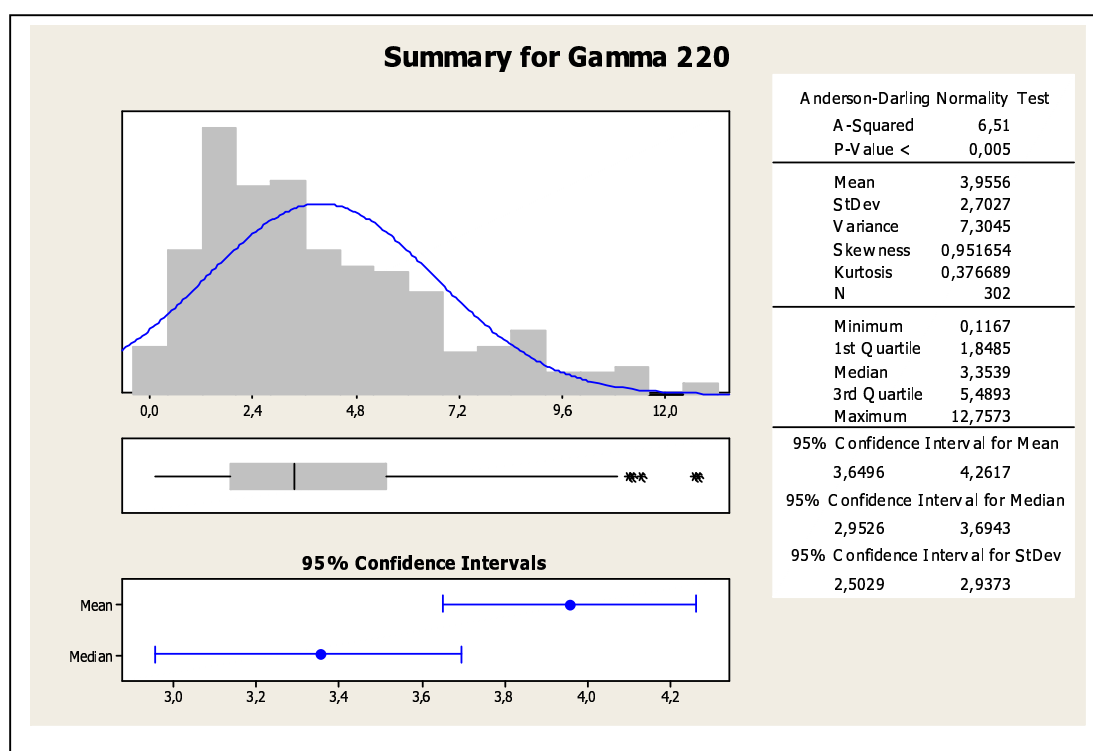


Figura 2 – Exemplo de análise gráfica de uma distribuição

Juran & Gryna (1992) estabeleceram alguns critérios de análise para abordagem dos dados em caso de não aderência à distribuição normal:

- a) Examinar os dados para verificar se há alguma explicação não estatística para o padrão distribucional não convencional;
- b) Analisar os dados em termos de médias ao invés de valores individuais, pois médias de amostra seguem de perto uma distribuição de probabilidade normal, mesmo se a população de valores individuais não é distribuída normalmente (teorema central do limite);
- c) Utilizar como referência outro tipo de distribuição que se enquadre mais adequadamente ao conjunto de dados coletados;

- d) Efetuar transformação matemática da característica original para uma nova característica que se aproxime de uma distribuição normal.

3.2. Transformação de Box-Cox

Uma estratégia eficiente para normalizar os dados não normais é através de transformação das variáveis em estudo. Entretanto a escolha do tipo adequado de transformação não parece ser uma tarefa óbvia, pois matematicamente existem inúmeras possibilidades e apenas o método de “tentativa e erro” nem sempre é o mais recomendado. Além disso, nem sempre uma transformação matemática produz os resultados esperados. A transformação linear, por exemplo, altera a escala da distribuição, mas não altera sua forma; já a transformação exponencial é mais eficiente para este propósito.

Box & Cox (1964) efetuaram um estudo detalhado na análise de dados representados pelas observações x_1, x_2, \dots, x_n consideradas normalmente distribuídas, com variância constante e valores esperados especificados por modelos lineares. O mais importante e significativo resultado deste estudo foi a definição da seguinte família de transformação exponencial da variável x para $x^{(\lambda)}$:

$$x^{(\lambda)} = \begin{cases} \frac{x^\lambda - 1}{\lambda} & (\lambda \neq 0) \\ \log x & (\lambda = 0) \end{cases} \quad (7)$$

Segundo Box & Cox (1964) esta transformação é definida somente para variáveis com valores positivos ($x > 0$) e o parâmetro λ , possivelmente um vetor, é o elemento que define a transformação específica e que, com frequência, resulta em “normalidade”. Como a análise de variância não é afetada por uma transformação linear, na prática a equação (7) pode ser simplificada para a seguinte forma:

$$x^{(\lambda)} = \begin{cases} x^\lambda & (\lambda \neq 0) \\ \log x & (\lambda = 0) \end{cases} \quad (8)$$

Para o conjunto de dados apresentado na Figura 2 foi efetuada aplicação da transformação de Box-Cox utilizando o *software* estatístico MiniTab®. As características da transformação, com destaque para determinação do parâmetro λ , estimado com 95% de confiança, são mostradas na Figura 3.

Após transformação de Box-Cox os dados obtidos podem ser bem representados por uma distribuição normal, visto que $p\text{-value} = 0,476$, ou seja, existe grande probabilidade de cometer erro ao rejeitar a hipótese de “normalidade” da distribuição obtida. A Figura 4 apresenta a análise gráfica da distribuição resultante da transformação.

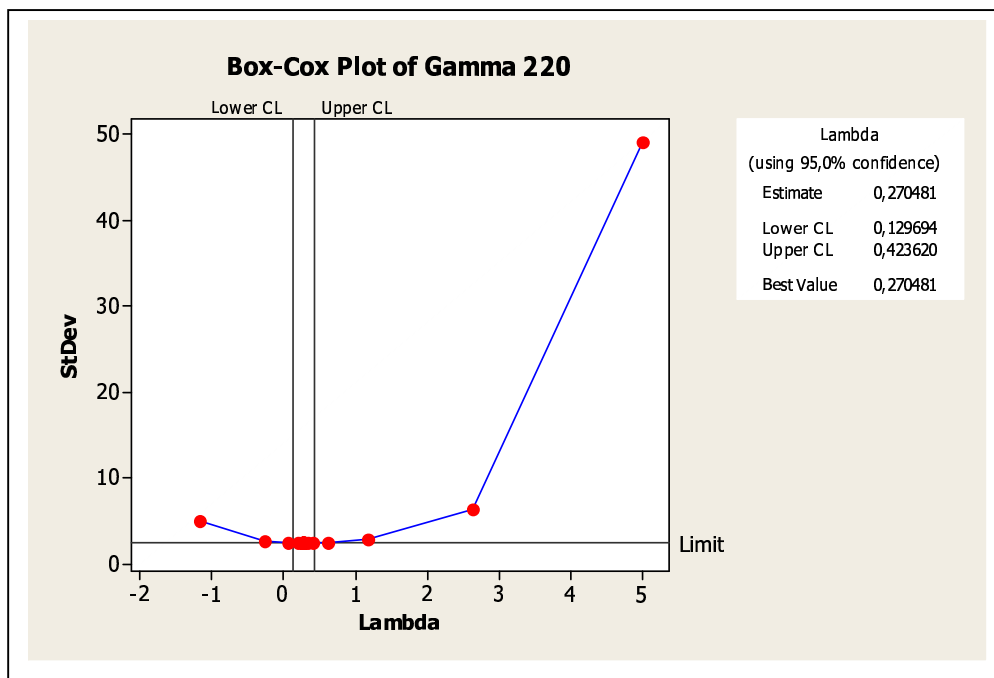


Figura 3 – Transformação de Box-Cox

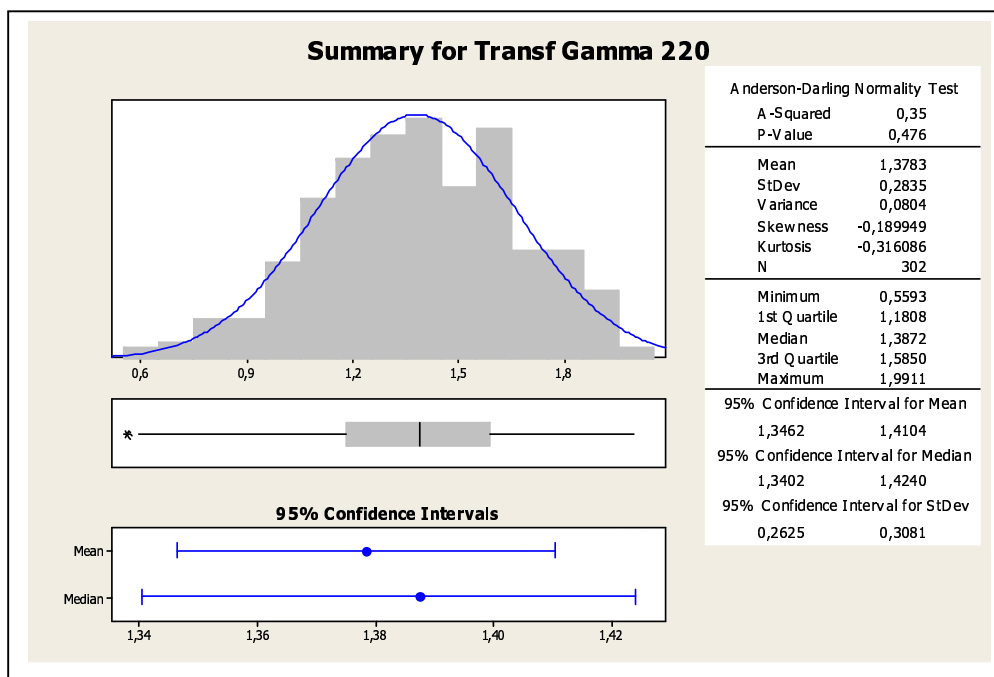


Figura 4 – Análise dos dados transformados (Box-Cox)

Para Box & Cox (1964) após a adequada transformação da variável x para $x^{(\lambda)}$, pode-se presumir que os valores esperados das observações transformadas apresentem as seguintes características:

- a) Sejam descritos por um modelo de estrutura simples;
- b) A variância seja constante;
- c) As observações estejam normalmente distribuídas.

3.3. Transformação de Yeo-Johnson

De acordo com Yeo & Johnson (2000) a contribuição de Box & Cox (1964) foi o maior passo na determinação de uma maneira objetiva de se efetuar transformação de dados. Entretanto, como a transformação de Box-Cox é válida apenas para valores positivos de x , havia espaço para algum tipo de melhoria. Embora seja possível efetuar uma troca de parâmetros em caso de valores negativos para utilização da transformação de Box-Cox, existe o inconveniente de tal ação afetar a teoria que suporta a definição do intervalo de confiança de λ .

Yeo & Johnson (2000) propuseram uma nova família de transformação de dados, válida tanto para valores positivos como para valores negativos da variável x . Sua fórmula, definida como uma função $\psi: \mathbb{R} \times \mathbb{R} \rightarrow \mathbb{R}$, é apresentada a seguir:

$$\Psi^{(\lambda, x)} = \begin{cases} \{(x+1)^\lambda - 1\} / \lambda & (x \geq 0, \lambda \neq 0) \\ \log(x+1) & (x \geq 0, \lambda = 0) \\ -\{(-x+1)^{2-\lambda} - 1\} / 2 - \lambda & (x < 0, \lambda \neq 2) \\ -\log(-x+1) & (x < 0, \lambda = 2) \end{cases} \quad (9)$$

Para o conjunto de dados apresentados na Figura 2 também foi efetuada aplicação da transformação de Yeo-Johnson utilizando o *software* estatístico MiniTab®. As características da transformação, com destaque para determinação da equação de transformação das variáveis, são apresentadas na Figura 5. Após transformação de Yeo-Johnson os dados podem ser melhor representados por uma distribuição normal, visto que $p\text{-value} = 0,950$, ou seja, existe elevada probabilidade de cometer erro ao rejeitar a hipótese de “normalidade” da distribuição obtida.

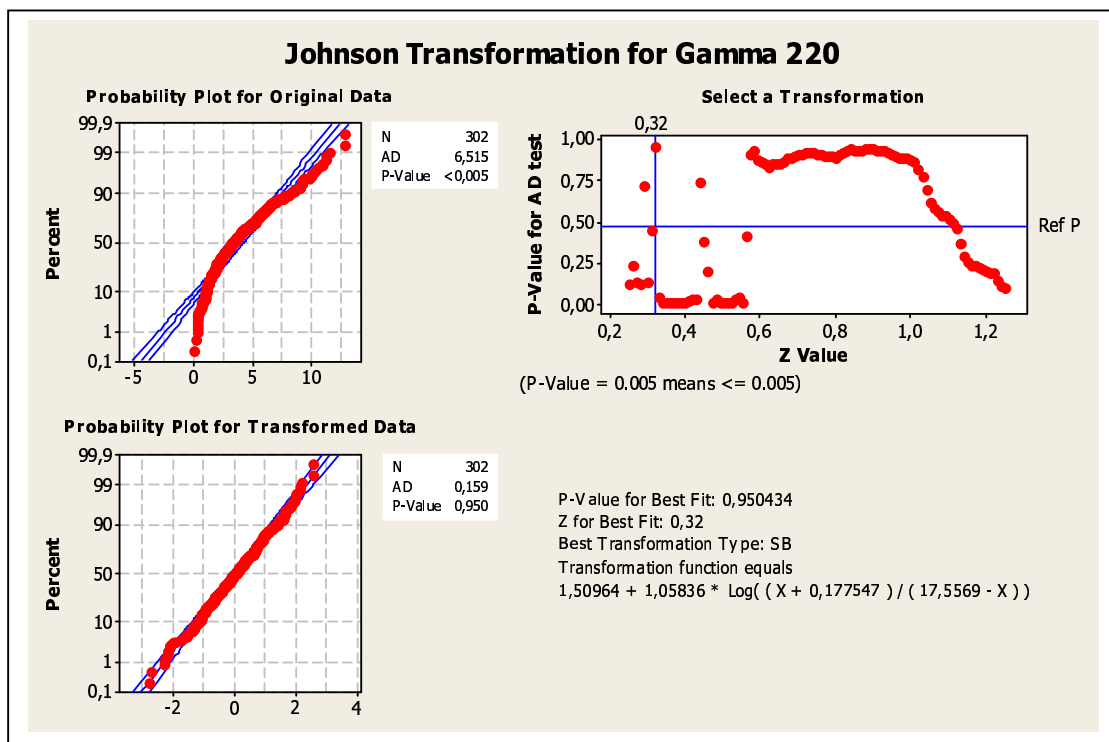


Figura 5 – Análise dos dados transformados (Yeo-Johnson)

4. Exemplos de análise dos métodos estatísticos em questão

4.1. Aspectos metodológicos

Segundo Asti Vera (1973) o ponto de partida de uma pesquisa é a existência de um problema que se deverá definir, examinar, avaliar e analisar criticamente para, em seguida, ser proposta sua solução. No presente artigo é apresentado o problema de ocorrência de decisões questionáveis na aplicação de métodos estatísticos devido à interpretação inadequada dos dados coletados. Este é um problema para o qual já se obteve solução a partir de outras pesquisas, algumas delas mencionadas neste trabalho; entretanto, o que realmente motiva o tema é a necessidade de promover uma maior divulgação das técnicas de transformação de variáveis para destacar a relevância do rigor metodológico durante coleta e tratamento de dados estatísticos.

Utilizando como referência os critérios de classificação de Gil (1991), quanto aos objetivos este trabalho caracteriza-se como uma pesquisa exploratória, pois deverá proporcionar maior familiaridade com as técnicas de transformação de variáveis durante coleta e tratamento de dados em trabalhos científicos, além de possibilitar análise crítica através da comparação entre os dados originalmente coletados e os dados transformados.

Köche (1984) destaca que a ciência proporciona a conceptualização da realidade e os conceitos com que ela opera são chamados de construtos. Os construtos são adotados ou inventados conscientemente com um significado específico. Na visão de Kelinger *apud* Köche (1984) conceito e construtos são semelhantes, sendo que a diferença reside no fato de que o construto possui um significado construído intencionalmente a partir de um marco teórico, devendo ser delimitado, traduzido em proposições particulares observáveis e mensuráveis. O objetivo do construto é fazer com que não haja ambigüidade no referencial empírico dos conceitos utilizados pelos pesquisadores.

A análise crítica da utilização de métodos estatísticos no presente artigo está fundamentada na utilização de um construto relacionado à análise de dados em uma empresa fictícia de prestação de serviços. O foco do negócio desta empresa fictícia, denominada “Gamma 220 Ltda”, é o contato direto com seus clientes através de rotinas específicas, cujo detalhamento é desnecessário para o propósito deste estudo. Os dados existentes referem-se ao tempo médio de atendimento aos clientes da empresa coletados diariamente ao longo dos 302 dias trabalhados no ano de 2004. Conforme análise gráfica, já demonstrada por meio da Figura 2, a distribuição dos dados referentes ao tempo médio de atendimento na empresa “Gamma 220 Ltda” não segue uma distribuição normal.

A seguir serão apresentadas, analisadas e discutidas as formas de utilização de gráficos de controle e avaliação de desempenho através da determinação do nível sigma a partir dos dados da empresa “Gamma 220 Ltda” (dados gerados pelo *software* estatístico MiniTab®: 302 valores seguindo a distribuição Gamma com parâmetros $shape = 2$ e $scale = 2$).

4.2. Identificação de falsas causas especiais nos gráficos de controle individuais

Três formas de utilização de gráfico de controle para avaliação do processo em termos de tempo médio de atendimento na empresa “Gamma 220 Ltda” são discutidas a seguir. A Figura 6 representa os dados plotados no *software* estatístico MiniTab® sem a preocupação de testar a aderência dos mesmos à distribuição normal. Observa-se através deste gráfico de controle individual que, de acordo com os limites estabelecidos, existem duas causas especiais de variação.

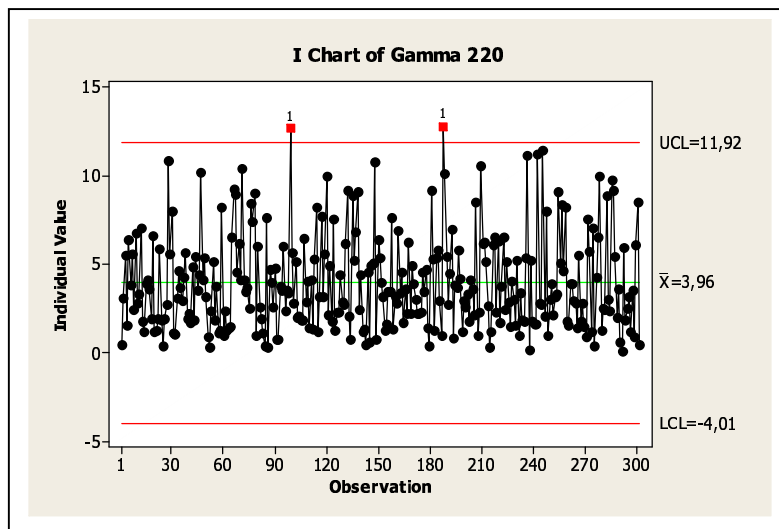


Figura 6 – Gráfico de controle para “Gamma 220” considerando os dados normalmente distribuídos

O gráfico de controle individual obtido após prévia transformação dos dados utilizando a transformação de Box-Cox é mostrado na Figura 7. Os dados transformados com $\lambda = 0,27$ assumem uma forma de distribuição que pode ser considerada normal, pois no teste de Anderson-Darling o parâmetro *p-value* é igual a 0,476. Neste caso o gráfico apresenta o mesmo comportamento do processo com os dados originais, porém, devido aos novos limites de controle, não há causas especiais de variação.

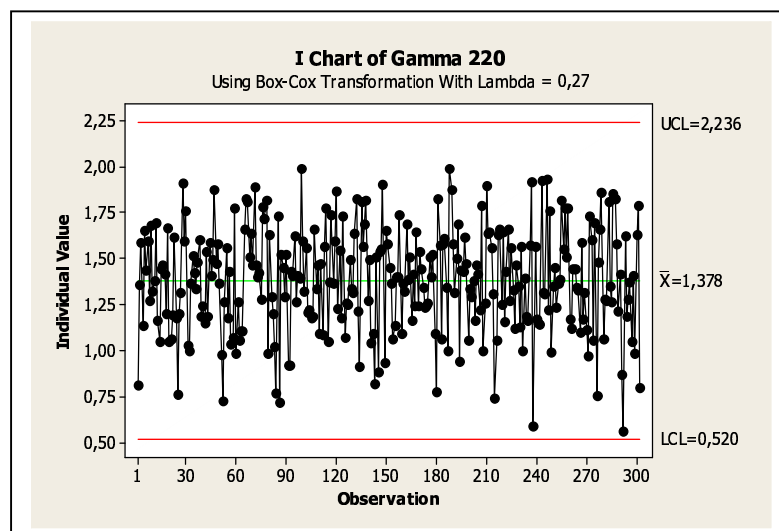


Figura 7 – Gráfico de controle para “Gamma 220” com transformação de Box-Cox

Utilizando-se a transformação de Yeo-Johnson os dados também assumem uma distribuição considerada normal, visto que no teste de Anderson-Darling o parâmetro *p-value* é igual a 0,950. A análise do gráfico de controle individual após a transformação de Yeo-Johnson indica o mesmo comportamento do processo em relação aos dados originais e, assim como nos dados transformados por Box-Cox, não se observam causas especiais de variação.

4.3. Cálculo incorreto do nível sigma

A avaliação de desempenho da empresa “Gamma 220 Ltda” tem como requisito um limite superior de especificação de 12 minutos, isto é, a média dos atendimentos diários não deve ultrapassar 12 minutos. Utilizando o *software* estatístico MiniTab® foram calculados três índices de desempenho P_{PK} diferentes: no primeiro caso o cálculo foi efetuado considerando os dados normalmente distribuídos; no segundo caso os dados foram previamente transformados por meio da transformação de Box-Cox; e no terceiro caso foi utilizada transformação de Yeo-Johnson. A Tabela 2 apresenta a comparação dos valores obtidos.

PREMISSA ADOTADA PARA CÁLCULO	LSE (min)	ÍNDICE PPK	DPMO	NÍVEL SIGMA
Dados considerados normalmente distribuídos	12	0,99	1470	4,47
Dados não normais transformados (Box-Cox)	12	0,71	16010	3,63
Dados não normais transformados (Yeo-Johnson)	12	0,77	10507	3,81

Tabela 2 – Cálculos comparativos do nível sigma para “Gamma 220”

A análise do nível sigma calculado nas três situações apresentadas indica grande disparidade entre os dados não transformados, isto é, considerados normais e os dados adequadamente transformados por Box-Cox ou Yeo-Johnson. O valor $4,47\sigma$ obtido com os dados originais denota um desempenho de 1470 defeitos por milhão de oportunidades (dpmo), ao passo que na realidade, devido à não aderência dos dados à distribuição normal, este desempenho estaria melhor representado pelos valores $3,63\sigma$ (16010 dpmo) ou $3,81\sigma$ (10507 dpmo) de acordo com o tipo de transformação adotado.

5. Considerações finais

Através deste artigo foi efetuada uma breve revisão bibliográfica sobre alguns métodos estatísticos aplicados à gestão da qualidade que podem ser afetados pela suposição inadequada de “normalidade” dos dados, em especial o Controle Estatístico de Processo e a Metodologia Seis Sigma. Também foram apresentados alguns procedimentos de transformação utilizados para ajustar os dados coletados em uma distribuição normal de modo a evitar que conclusões duvidosas sejam assumidas como verdadeiras.

A análise dos dados da empresa fictícia “Gamma 220 Ltda”, adotada como exemplo neste trabalho, indicou que, quando os métodos estatísticos foram aplicados sem os cuidados necessários de confirmação de aderência à distribuição normal, ocorreram equívocos na determinação das causas especiais de variação do processo e na avaliação do desempenho da empresa ao longo do ano. Incorreções desta natureza podem prejudicar a compreensão dos cenários para a tomada de decisões e, até mesmo, comprometer o desdobramento de estratégias empresariais.

As duas falsas causas especiais de variação, identificadas no processo, se corretamente monitoradas, poderiam estimular uma ação gerencial desnecessária que, muito provavelmente, não agregaria benefícios aos resultados da empresa “Gamma 220 Ltda”. Com relação à definição incorreta do nível sigma, vale ressaltar que a defasagem entre dados originais e dados transformados corresponde a uma diferença de 14540 defeitos por milhão de oportunidades (com a transformação de Box-Cox) e 9037 defeitos por milhão de oportunidades (com a transformação de Yeo-Johnson). Tais quantidades são muito significativas em termos de avaliação de capacidade e desempenho.

A evolução desta pesquisa aponta para uma análise crítica de outros casos onde efetivamente a premissa de “normalidade” dos dados seja inadequadamente assumida nos métodos estatísticos abordados e, para trabalhos futuros, outros métodos, tais como Planejamento de Experimentos, Teste de Hipóteses e Análise de Regressão, também poderão ser analisados.

Referências

ASSOCIAÇÃO BRASILEIRA DE NORMAS TÉCNICAS – NBR ISO 9000: Sistema de Gestão da Qualidade – Fundamentos e Vocabulário, Rio de Janeiro, 2000.

ASTI VERA, A. – Metodologia da Pesquisa Científica. Porto Alegre: Globo, 1973.

BALESTRASSI, P.P. – Metodologia Seis Sigma. Notas de Aula do Curso de Pós-Graduação em Engenharia de Produção. Universidade Federal de Itajubá, 2004.

BOX, G.E.P. and COX, D.R. – *An Analysis of Transformations. Journal of Royal Statistical Society. B, 39, 211-252, 1964.*

BREYFOGLE, F. – *Six Sigma Methods to Ensure Organization's Health. . New York: Quality Progress, 45, 4, p.28-29, April 2003.*

FEIGENBAUM, A.V. – Controle da Qualidade Total, vol. 3, Métodos Estatísticos Aplicados à Qualidade. 3ª ed. São Paulo: Makron Books, 1994.

GIL, A.C. Como Elaborar Projetos de Pesquisa. 3ª ed. São Paulo: Atlas, 1991.

HARRY, M.J. – *Six Sigma: A Breakthrough Strategy for Profitability. New York: Quality Progress, 31, 5, p.60-64, May 1998.*

JURAN, J.M. e GRZYNA, F.M. – Controle da Qualidade Handbook, vol. 6, Métodos Estatísticos Clássicos Aplicados à Qualidade. 4ª ed. São Paulo: Makron Books, 1992.

KÖCHE, J.C. – Fundamentos de Metodologia Científica. 7ª ed. Caxias do Sul: Editora Vozes, 1984.

LINDERMAN, K. *et al.* – *Six Sigma: A Goal-Theoretic Perspective. Journal of Operations Management, v.21, n.2, p.193-203, 2003.*

LIPSON, C. and SHETH, N.J. – *Statistical and Analysis of Engineering Experiments. New York: McGraw-Hill Book Company, 1973.*

MONTGOMERY, D.C. – *Introduction to Statistical Quality Control. New York: John Wiley & Sons, Inc, 1985.*

MONTGOMERY, D.C. e RUNGER, G.C – Estatística Aplicada e Probabilidade para Engenheiros. 2ª ed. Rio de Janeiro: LTC – Livros Técnicos e Científicos Editora S.A., 2003.

TOLEDO, T.P.A. – Uma Investigação sobre Índices de Capabilidade com Ênfase na Metodologia Seis Sigma. Dissertação de Mestrado. Universidade Federal de Itajubá, 2005.

YEO, I.K. and JOHNSON, R.A. – *A New Family of Power Transformation to Improve Normality or Symmetry. Biometrika, 87, 954-959, 2000.*