

## **Índices de Capacidade do Processo para Dados não normais: Um Estudo do Consumo de Água em uma Indústria do Setor Calçadista**

Heloisa Helena Weber (UFRGS) [heloisa@producao.ufrgs.br](mailto:heloisa@producao.ufrgs.br)  
Liane Werner (UFRGS) [liane@producao.ufrg.br](mailto:liane@producao.ufrg.br)

### **Resumo:**

A análise dos índices de capacidade de processo é uma ferramenta que quando não aplicada corretamente pode trazer informações inexatas e interpretações errôneas sobre o processo. Os primeiros índices foram desenvolvidos a partir de 1974, por Juran, e desde então melhorias foram propostas a esta metodologia. Este trabalho apresenta o exemplo de uma aplicação de como analisar a capacidade do processo quando os dados coletados não apresentam distribuição normal, para tanto foi utilizado a transformação Box-Cox. Após isso, a interpretação dos resultados segue os mesmos princípios dos índices de capacidade para distribuição normal. Para o estudo realizado verificou-se que o processo não é capaz.

**Palavras chave:** Capacidade do processo, dados não normais, transformação de Box-Cox.

## **Process Capability Indices for Non-Normal Data: a Study of Water Use in the Shoe Manufacturing Industry**

### **Abstract**

The analysis of process capability index is a tool that when not applied correctly can bring inaccuracies and misinterpretations about the process. The first indices have been developed since 1974 by Juran, and since then improvements have been proposed to this methodology. This paper presents an application example of how to analyze the ability of the process when the data collected do not have normal distribution. For this was used the Box-Cox transformation after that the interpretation of results follows the same principles of capability indices for normal distribution. For the study it was found that the process is not capable.

**Key-words:** Process Capability, no normal data, Box-Cox transformation

### **1. Introdução**

Um processo pode ser definido como um conjunto de procedimentos com o objetivo de produzir um efeito, que é um serviço ou um produto do processo. Todos os processos possuem variabilidade, decorrentes das condições do ambiente ou do sistema de medição (WERKEMA, 2006). Segundo Silva, Ribeiro e Melo (2008), essa variabilidade pertence ao processo, e sob o ponto de vista econômico e físico, é inviável eliminá-la.

Analisando-se a variabilidade do processo, têm-se duas classificações possíveis para ele: quando somente estas causas inviáveis de eliminação estão presentes (causa comuns), sendo o processo é dito estatisticamente sob controle e quando, o processo apresenta, além das causas comuns de variabilidade, variabilidade anômala (causas especiais), sendo conhecido como processo fora de controle estatístico.

O estudo de capacidade de processos é uma ferramenta existente no controle estatístico de processo que desde o início da década de 80 vem sendo utilizado nas indústrias do Japão, país referência quando se fala em qualidade (MIRANDA, 2005). Segundo Ramos (2005), esta ferramenta tem como objetivo avaliar se um determinado processo consegue gerar produtos que atendam às especificações de desejadas.

Para Montgomery (2004), a qualidade dos produtos produzidos, conhecida como conformidade, pode ser atingida quando os processos que a produzem são capazes, isto é, produzem as peças de acordo com o que foi definido no projeto. Quando a produção de itens defeituosos ocorre acima do admitido, o processo é considerado incapaz, basicamente por dois motivos: a variabilidade do processo é muito grande em relação aos limites de especificação e/ou a média do processo não está centrada no alvo. Tendo conhecimento destas informações, pode-se então tomar ações de melhorias no processo (MONTGOMERY, 2004).

Existem ainda muitas críticas e descrença quanto ao uso da estatística e de índices de capacidade para avaliar a capacidade de processos produtivos. Estas críticas surgem muitas vezes devido a má utilização dos índices, que necessitam do conhecimento de suas propriedades estatísticas para avaliação mais correta da real capacidade (MIRANDA, 2005). Por exemplo, os dados podem apresentar distribuição normal ou não normal. Se uma distribuição for não normal e receber tratamento de normalidade, a análise apresentará estimativa errada do número de itens não conformes e conseqüentemente uma interpretação equivocada sobre este processo (GONZALES, WERNER, 2009).

Para solucionar esta situação têm-se duas alternativas, a primeira é por meio de uso de índices específicos para dados não normais, a outra consiste na transformação dos dados em dados normalmente distribuídos.

Este artigo tem por objetivo verificar o que ocorre com a capacidade do processo por meio de um exemplo de aplicação da transformação de Box-Cox, para quando os dados do processo apresentam distribuição não normal. O trabalho não contempla a melhoria deste processo, mas pretende contribuir para futuros trabalhos relacionados a análise e transformação de dados não normais.

## **2. Revisão Teórica**

### **2.1 Análise do Processo**

Os estudos de capacidade avaliam se um determinado processo é capaz de gerar produtos que atendam aos limites especificados. Especificação de projeto é o valor alvo característica de produtos que se querem produzir mais as tolerâncias admitidas em torno destes alvos. Estudo de capacidade de processo é uma técnica que tem utilidade em muitos pares do ciclo do produto e é também fundamental para programas de melhoria da qualidade (MONTGOMERY, 2004).

A metodologia orienta a obter uma amostra de itens produzidos em condições normais de operação e que a partir de cálculos estatísticos se determina a capacidade do processo em fabricar itens conformes (RAMOS, 2005).

Montgomery (2004) explica que o estudo da capacidade do processo pode ser útil para quantificar a variabilidade do processo, analisar essa variabilidade em relação às exigências do produto e auxilia na redução ou eliminação dessa variabilidade.

Segundo Miranda (2005), os índices de capacidade de processo podem ser classificados em três gerações. Os primeiros estudos sobre capacidade de processo foram desenvolvidos em 1974, por Juran, que analisou a relação entre a variabilidade do processo e as especificações

do cliente. Mas foi a partir dos anos 80 que esses índices, considerados da primeira geração dos índices de capacidade, passaram a ser utilizados na indústria japonesa.

Em 1985 Taguchi criou o conceito de função perda, que apresenta uma penalidade quadrática por um produto não estar no alvo de especificação. Essa função foi incorporada aos índices existentes anteriormente e considera-se a segunda geração de índices de capacidade. Segundo Gentilini, Stroeike e Werner (2010), um produto fora do alvo pode gerar sucatas, retrabalhos, insatisfação do cliente, gastos adicionais para reposição ou reparação do produto.

A partir da década de 90 surgiram então os índices que compõe a terceira geração de índices de capacidade. Nesse período surgiram diversos modelos, cada um deles criados para lidar com determinadas características de dados. Os índices dessa terceira geração são considerados de difícil interpretação e por esse motivo a alternativa é tratar os dados para que se possível aplicar os cálculos dos índices de primeira geração.

A figura 1 reflete como proceder a análise de capacidade do processo. Inicia-se com a coleta de dados, proveniente das cartas de controle ou de um levantamento específico para a análise de capacidade do processo. Após realiza-se uma verificação da normalidade dos dados, caso sim, obtem-se os índices já conhecidos e faz-se uma avaliação da capacidade do processo. No caso dos dados não apresentarem distribuição normal, avalia-se a possibilidade de transformá-los em dados normais. Caso sim, obtem-se os índices como no passo anterior, caso não calcular a capacidade do processo por meio de índices específicos para a situação de não normalidade.

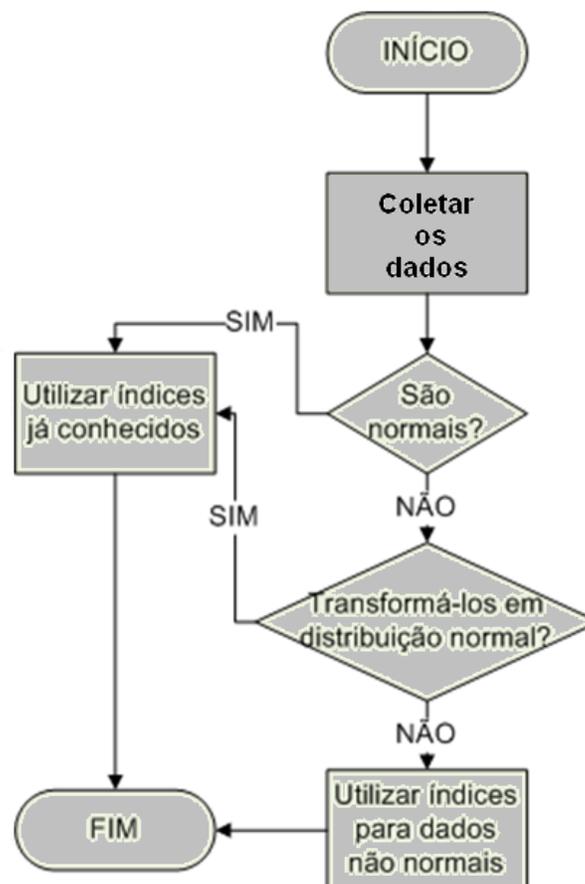
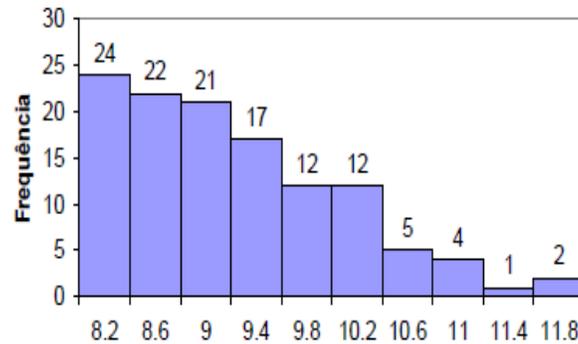


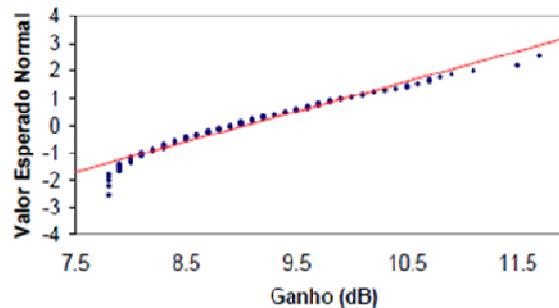
FIGURA 1 – Fluxograma proposto para análise de capacidade do processo

## 2.2 Capacidade do Processo

Antes de utilizar algum dos índices de capacidade é necessário verificar se os dados coletados seguem uma distribuição normal ou não normal. Miranda (2005) afirma que este procedimento pode ser feito utilizando ferramentas gráficas, como um histograma ou papel de probabilidade, apresentados na figura 2 ou por meio da realização de testes de normalidade, como Anderson-Darling ou Komogorov-Smirnoff.



(a)



(b)

FIGURA 2: Exemplo de histograma para dados não normais (a) e papel de probabilidade (b).

Quando os dados seguem uma distribuição normal eles podem ser analisados diretamente com a utilização dos índices de capacidade da primeira geração, como o  $C_p$ , ou o  $C_{pk}$  (MONTGOMERY, 2004).

O índice  $C_p$  mede a capacidade potencial do processo e segundo Montgomey (2004) um  $C_p$  menor que 1,33 significa que o processo não é capaz de produzir somente unidades conformes. Ele é obtido aplicando a fórmula (1).

$$C_p = \frac{LSE - LSI}{6\sigma} \quad (1)$$

onde: LSE e LSI são, respectivamente, os limites superior e inferior de especificação e  $\sigma$  o desvio padrão do processo.

Montgomery (2004) desenvolveu uma tabela de valores mínimos recomendados da razão da capacidade de processo, que leva em consideração se essa análise se refere a processo existente ou novo, processos que envolvem parâmetros críticos.

Este índice não leva em consideração a média do processo em relação aos limites especificados, ou seja, mede somente a dispersão das especificações em relação a seis desvios padrões do processo.

Desta deficiência do Cp, surgiu o índice Cpk, que além dos limites superior e inferior de especificação também leva em conta a centralização do processo. Esse índice é obtido quando aplicada a equação (2).

$$Cpk = \min \left\{ \frac{LSE - \mu}{3\sigma} ; \frac{\mu - LIE}{3\sigma} \right\} \quad (2)$$

onde: LSE e LSI são, respectivamente, os limites superior e inferior de especificação,  $\mu$  é a média do processo e  $\sigma$  o desvio padrão do processo.

Da mesma forma que Cp e o Cpk para indicar um processo capaz deve ser maior que 1,33 e quando Cp = Cpk então o processo está centrado no ponto médio das especificações.

Gonzáles e Werner (2009) explicam que ao se utilizar índices convencionais em dados não normais, pode-se concluir que o processo é capaz quando na realidade não é. Esta situação é agravada ainda mais, quando se tem distribuições muito diferentes da distribuição normal, como pode ser observado na figura 3.

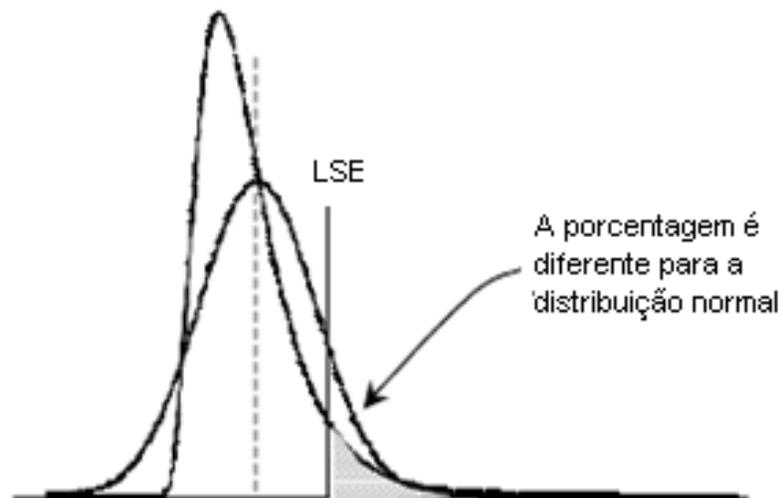


FIGURA 3: Comparação entre uma distribuição normal e uma distribuição não normal

Fonte: Adaptado de Oliveira (2005)

Se a condição de normalidade não for satisfeita então eles deverão ser analisados com índices elaborados para distribuições não normais de dados ou podem ser transformados e posteriormente ser analisados com os índices de primeira geração.

Na literatura são abordadas algumas ferramentas para realizar a transformação de dados, como as de Box-Cox e Johnson, além do ajuste de distribuição (OLIVEIRA, 2005). Por sua simplicidade e praticidade, este estudo utilizará a transformação de Box-Cox.

### 2.3 Transformação de Box-Cox

Box e Cox (1964) propuseram um conjunto de transformações sobre uma variável positiva Y. A tabela 1 apresenta algumas transformações mais utilizadas na prática, de como proceder

para transformar uma variável  $Y$ , de uma distribuição não-normal, em uma variável de distribuição normal.

O método orienta a realizar a transformação com vários valores de  $\lambda$ . A alternativa onde os dados após a transformação apresentar uma distribuição normal e apresentar o menor desvio padrão, deve ser a opção utilizada.

Valor de $\lambda$	$Y' = Y$
-2	$\frac{1}{y^2}$
-1	$\frac{1}{y}$
-0,5	$\frac{1}{\sqrt{y}}$
0	$\ln y$
0,5	$\sqrt{y}$
1	Sem transformação
2	$y^2$

TABELA 1 - Transformação para alguns valores de  $\lambda$

Fonte: Oliveira (2005)

### 3. Estudo Aplicado

Este estudo foi aplicado visando avaliar a capacidade do processo de consumo de água em uma empresa do setor calçadista. Tal estudo irá permitir a empresa verificar se o consumo de água está dentro dos padrões estabelecidos.

A coleta de dados foi realizada através da leitura de um hidrômetro instalado na saída da caixa d'água, de modo a controlar a quantidade de água utilizada durante o dia. A leitura do hidrômetro foi realizada todos os dias às seis horas da manhã, uma hora antes do início das atividades na empresa. Deste modo se garantiu que o consumo de todo um dia de trabalho, incluindo horas extras, seria registrado nessa leitura.

Os valores eram inseridos todos os dias em uma planilha e a diferença entre a leitura do dia  $x$  e o dia  $(x-1)$  representa o consumo de água do dia  $(x-1)$ . Na Tabela 2 são apresentados os valores do consumo de água, em metros cúbicos, durante os 15 dias da coleta de dados.

Realizou-se o teste estatístico denominado Kolmogorov-Smimoff, utilizando o software estatístico SPSS, sendo que o teste rejeitou a hipótese de normalidade, indicando a não normalidade dos dados ao nível de 5% de significância.

Consumo de água		
65,7	65,5	11,9
63,9	80	8,5
65,4	75,5	64
15,6	69,31	70,1
12	69	76,9

TABELA 2- Consumo de água

Realizou-se o teste estatístico denominado Kolmogorov-Smimoff, utilizando o software estatístico SPSS, sendo que o teste rejeitou a hipótese de normalidade, indicando a não normalidade dos dados ao nível de 5% de significância.

A seguir, realizaram-se as transformações sugeridas no método Box-Cox apresentadas na Tabela 1. Os resultados obtidos são apresentados na Tabela 3. Após realizar estas transformações a metodologia orienta a aplicar testes de normalidade a estes novos dados. Os resultados obtidos ao usar o software SPSS são apresentados na Figura 4.

**Hypothesis Test Summary**

	Null Hypothesis	Test	Sig.	Decision
1	The distribution of Quadrado is normal with mean 3,610.472 and standard deviation 2,262.59.	One-Sample Kolmogorov-Smirnov Test	.100	Retain the null hypothesis.
2	The distribution of RaizQuadrada is normal with mean 7.031 and standard deviation 2.264.	One-Sample Kolmogorov-Smirnov Test	.017	Reject the null hypothesis.
3	The distribution of LogNatural is normal with mean 3.766 and standard deviation 0.824.	One-Sample Kolmogorov-Smirnov Test	.011	Reject the null hypothesis.
4	The distribution of InvRaizQuadrada is normal with mean 0.166 and standard deviation 0.081.	One-Sample Kolmogorov-Smirnov Test	.008	Reject the null hypothesis.
5	The distribution of Inverso is normal with mean 0.034 and standard deviation 0.035.	One-Sample Kolmogorov-Smirnov Test	.008	Reject the null hypothesis.
6	The distribution of InversoQuadrado is normal with mean 0.002 and standard deviation 0.004.	One-Sample Kolmogorov-Smirnov Test	.009	Reject the null hypothesis.

Asymptotic significances are displayed. The significance level is .05.

FIGURA 4: Teste de Normalidade dos dados após a transformação Box-Cox

Analisando os resultados obtidos na figura 4, observa-se que o teste estatístico Kolmogorov-Smimov rejeitou a hipótese de normalidade, indicando a não normalidade dos dados transformados em Raiz Quadrada, Logaritmo Natural, Inverso da Raiz Quadrada, Inverso e Inverso do Quadrado ao nível de 5% de significância. A única distribuição considerada normal por este teste foi a distribuição do Quadrado ( $Y^2$ ).

Original Y	Quadrado Y <sup>2</sup>	Raiz Quadrada $\sqrt{Y}$	Logaritmo Natural ln (Y)	Inverso da Raiz Quadrada $1/\sqrt{Y}$	Inverso 1/Y	Inverso ao Quadrado 1/Y <sup>2</sup>
65,7	4316,490	8,106	4,185	0,123	0,015	0,000232
63,9	4083,210	7,994	4,157	0,125	0,016	0,000245
65,4	4277,160	8,087	4,181	0,124	0,015	0,000234
15,6	243,360	3,950	2,747	0,253	0,064	0,004109
12	144,000	3,464	2,485	0,289	0,083	0,006944
65,5	4290,250	8,093	4,182	0,124	0,015	0,000233
80	6400,000	8,944	4,382	0,112	0,013	0,000156
75,5	5700,250	8,689	4,324	0,115	0,013	0,000175
69,31	4803,876	8,325	4,239	0,120	0,014	0,000208
69	4761,000	8,307	4,234	0,120	0,014	0,000210
11,9	141,610	3,450	2,477	0,290	0,084	0,007062
8,5	72,250	2,915	2,140	0,343	0,118	0,013841
64	4096,000	8,000	4,159	0,125	0,016	0,000244
70,1	4914,010	8,373	4,250	0,119	0,014	0,000203
76,9	5913,610	8,769	4,343	0,114	0,013	0,000169

TABELA 3 – Transformações de Box-Cox

A partir dos dados apresentados na figura 4 foi possível verificar que a única transformação pode ser utilizada como apresentando distribuição normal é o quadrado dos dados originais. Se mais de uma transformação apresentasse normalidade, utilizar-se-ia a de menor desvio padrão.

Para realizar a análise da capacidade do processo utilizou-se o índice Cpk, uma vez que se tem apenas o limite superior desejado pela empresa, que é de 65 metros cúbicos diários. O resultado segue:

$$C_{pk} = \min \left\{ \frac{LSE - \hat{\mu}}{3\hat{\sigma}} ; \frac{\hat{\mu} - LIE}{3\hat{\sigma}} \right\} = \min \left\{ \frac{65^2 - 3.610,47}{3 * 2.262,59} ; ** \right\} = 0,09$$

Analisando o índice obtido com os dados transformados verifica-se que o processo não é capaz.

#### 4. Discussão e Conclusões

Verificou-se que o processo é não capaz. Ao interpretar o resultado do índice Cpk fica claro que o consumo de água dessa empresa extrapolará o limite estabelecido de 65 metros cúbicos por dia em diversas vezes.

Os resultados desse estudo estão vinculados a uma amostra de que contém 15 dados. É possível que, caso a amostra utilizada para esta análise, os resultados fossem um pouco diferentes, desde os testes de normalidade. Porém a empresa estudada forneceu somente uma amostra com 15 dados.

Estudos de capacidade de processos para dados com distribuição normal estão consolidados na literatura (GONZALEZ e WERNER, 2009), mas referente a dados que não apresentam distribuição normal, poucos trabalhos são encontrados, o que oportunizou este estudo, que pretende colaborar para futuros trabalhos relacionados a análise e transformação de dados não normais.

## Referências

- BOX, G. E. P.; COX, D. R. An Analysis of transformations. *Journal of the Royal Statistical Society Series (Methodological)*, v.26, n. 2, p. 211-252. 1964.
- GENTILINI, M. M.; STROIEKE, R. E.; WERNER, L. *Aplicação do método de Taguchi na definição do Valor de tolerância para ensaio de Resistência ao Rasgo em Espumas de Poliuretano*. XVII SIMPEP – Simpósio de Engenharia de Produção. Bauru, 2010.
- GONZÁLES, P. U.; WERNER, L. *Comparação dos Índices de capacidade do processo para distribuições não normais*. Gestão e Produção. São Carlos, 2009.
- MIRANDA, R.G. *Um Modelo para análise de capacidade de processos com ênfase na transformação de dados* (Mestrado em Engenharia de Produção). Programa de Pós Graduação em Engenharia de Produção, Universidade Federal de Santa Catarina. Florianópolis, 2005.
- MONTGOMERY, D. C. *Introdução ao Controle Estatístico da Qualidade*. 4.ed. Rio de Janeiro: LTC, 2004.
- OLIVEIRA, E.S. *Análise de dados não normais no contexto da metodologia Sis Sigma* (Trabalho de Diplomação de curso de Engenharia de Produção e Gestão) Universidade Federal de Itajubá. 2005.
- RAMOS, A. W. *Estudos de capacidade para dados não normais*. Departamento de Engenharia de Produção da Universidade de São Paulo, 2005.
- SILVA, G. C. S.; RIBEIRO, F. C.; MÉLO, M. A. N. Aplicação do controle estatístico de processo para análise de sobrepeso de uma linha de desodorantes em uma indústria de higiene pessoal. *Anais: XXVIII ENEGEP – Encontro Nacional de Engenharia de Produção*. Rio de Janeiro, 2008.
- WERKEMA, Cristina. *Ferramentas estatísticas básicas para o gerenciamento de processos*. 1. Ed. Belo Horizonte, Editora Werkema, 2006.