**Universidade de São Paulo
Escola Superior de Agricultura “Luiz de Queiroz”**

**Índices de capacidade do processo para distribuições não normais:
uma aplicação na indústria metalúrgica**

Patricia Shizue Matsumura Ueda Gonzalez

Dissertação apresentada para obtenção do título de
Mestra em Ciências. Área de concentração:
Estatística e Experimentação Agronômica

**Piracicaba
2013**

Patricia Shizue Matsumura Ueda Gonzalez
Bacharel em Estatística

**Índices de capacidade do processo para distribuições não normais: uma
aplicação na indústria metalúrgica**

Versão revisada de acordo com a resolução CoPGr 6018 de 2011

Orientadora:
Prof.^a Dr.^a **SONIA MARIA DE STEFANO PIEDADE**

Dissertação apresentada para obtenção do título de
Mestra em Ciências. Área de concentração:
Estatística e Experimentação Agronômica

Piracicaba
2013

Dados Internacionais de Catalogação na Publicação
DIVISÃO DE BIBLIOTECA - ESALQ/ USP

Ueda Gonzalez, Patricia Shizue Matsumura

Índices de capacidade do processo para distribuições não normais: uma aplicação na indústria metalúrgica / Patricia Shizue Matsumura Ueda Gonzalez.- - versão revisada de acordo com a resolução CoPGr 6018 de 2011. - - Piracicaba, 2013.
68 p: il.

Dissertação (Mestrado) - - Escola Superior de Agricultura "Luiz de Queiroz", 2013.

1. Controle estatístico do processo 2. Distribuição não normal 3. Índices de capacidade 4. Indústria metalúrgica 5. R - Software estatístico I. Título

CDD 519.5
U22i

"Permitida a cópia total ou parcial deste documento, desde que citada a fonte - O autor"

DEDICATÓRIA

*Ao meu marido Roger
e ao meu filho Vitor*

AGRADECIMENTOS

À minha orientadora, Profa. Dra. Sonia Maria De Stefano Piedade, pelo conhecimento transmitido, pela compreensão demonstrada nos momentos difíceis ocorridos durante o meu caminho até a conclusão desta dissertação, pela paciência das revisões e pelo carinho demonstrado.

Aos colegas de mestrado com os quais compartilhei meus anseios, dificuldades e vitórias.

Ao Prof. Dr. Carlos Tadeu e à secretária Solange Sabadin pela atenção com que trataram minha necessidade de extensão de prazo do curso.

À secretária Luciane Brajão, por toda sua atenção e disposição em ajudar. Mais do que uma profissional do programa de pós-graduação, ela se tornou uma grande amiga durante o mestrado.

Ao Eduardo Martins, gestor e amigo, pelo apoio e flexibilização do horário de trabalho para que eu pudesse frequentar as aulas.

Por fim, agradeço à minha família que também teve que abdicar de alguns momentos juntos para que fosse possível concluir este trabalho e o curso.

SUMÁRIO

RESUMO.....	9
ABSTRACT	11
LISTA DE FIGURAS.....	13
LISTA DE TABELAS	15
LISTA DE ABREVIATURAS E SIGLAS.....	17
1 INTRODUÇÃO	19
1.1 Justificativa.....	20
1.2 Objetivos	21
1.3 Metodologia.....	21
1.4 Estrutura.....	22
1.5 Limitações	23
2 REVISÃO BIBLIOGRÁFICA.....	25
2.1 Capacidade do processo.....	25
2.1.1 Índice C_p	28
2.1.2 Índice C_{pk}	30
2.1.3 Índices C_{pm} e C_{pmk}	32
2.1.4 Formulação única dos índices.....	33
2.2 Capacidade do processo não normal.....	34
2.2.1 Método 1	40
2.2.2 Método 2	44
2.2.3 Método 3	46
3 ESTUDOS DE CASO	49
3.1 Estudo comparativo entre os métodos de análise de capacidade do processo não normal	49
3.2 Aplicação dos métodos de análise de capacidade do processo não normal.....	53
4 CONSIDERAÇÕES FINAIS	57
REFERÊNCIAS.....	59
BIBLIOGRAFIA CONSULTADA.....	61
ANEXOS	63

RESUMO

Índices de capacidade do processo para distribuições não normais: uma aplicação na indústria metalúrgica

Recentemente a análise de capacidade do processo para dados com distribuição não normal começou a ser explorada, porém são raras as referências em língua portuguesa que abordem este assunto. Este trabalho apresenta os principais índices para análise da capacidade do processo com dados não normais encontrados na literatura, sendo o primeiro deles o método proposto por Clements em 1989, o segundo, proposto por Pearn e Chen (1997) e o terceiro método apresentado é o proposto por Chen e Ding em 2001. Os métodos de Clements (1989) e Pearn e Chen (1997) são parecidos em muitos aspectos, mas cada um deles apresenta uma novidade em relação ao cálculo dos índices de capacidade. Em comum, ambos os métodos apresentam índices semelhantes aos índices tradicionais, que supõem normalidade dos dados. O método de Chen e Ding (2001) traz como novidade o uso do índice proposto por eles para a estimação do número de itens não conformes apresentados pelo processo em estudo. Por fim, dois estudos de caso são apresentados neste trabalho. No primeiro deles tem-se a comparação entre os métodos de Clements (1989) e de Pearn e Chen (1997) para que se conclua sobre qual dos métodos é o melhor para o cálculo da capacidade do processo não normal. No segundo estudo de caso, todos os três métodos apresentados neste trabalho são aplicados em um conjunto de dados reais, obtidos de uma indústria metalúrgica. Para execução das análises nos estudos de caso, foi aplicado o software livre R versão 2.10. Os resultados apresentados mostram que o método proposto por Pearn e Chen (1997) é mais eficiente do que o proposto por Clements (1989) por considerar a assimetria do processo. Quanto ao processo da indústria metalúrgica observou-se que a etapa analisada é relativamente capaz.

Palavras-chave: Controle estatístico do processo; Índices de capacidade; Distribuição não normal; Itens não conformes

ABSTRACT

Process capability indices for non-normal distributions: an application in the metallurgical industry

The analysis of process capability for non-normal distribution data started to be explored recently, but is rare references that address this issue in Portuguese. This dissertation presents the main indices for process capability analysis for non-normal data. The first method was proposed in 1989 by Clements, the second was proposed by Pearn e Chen (1997) and the third was proposed by Chen e Ding in 2001. The methods of Clements (1989) and Pearn e Chen (1997) are similar in many aspects, but each presents something new in the calculation of capability indices. In common, both methods have indices similar to the traditional indices, for normal data. The method proposed by Chen e Ding (2001) innovates using the index in the estimative of number of non-conforming in the process. Finally, two case studies are presented. The first compares the methods of Clement (1989) and Pearn e Chen (1997) to show the best method of calculation the capability of non-normal process. In the second, the three methods showing in this dissertation are applied in a metallurgical industry data. The R software version 2.10 was used in the analysis. The results show that the method proposed by Pearn and Chen (1997) is more efficient than that proposed by Clements (1989) for considering the asymmetry of the process. And the metallurgical process stage analyzed is relatively capable.

Keywords: Statistical process control; Capability indices; Non-normal distribution; Non-conforming items

LISTA DE FIGURAS

Figura 1(a) – Processo sob controle estatístico	20
Figura 1(b) – Processo fora de controle estatístico	20
Figura 2 – Padrão de variação de uma carta de controle.....	26
Figura 3 – Limites naturais de variabilidade do processo.....	26
Figura 4 – Relações entre os limites naturais do processo e os limites de especificação	27
Figura 5 – Comparação dos histogramas com as especificações.....	30
Figura 6 – Interpretação dos índices C_p e C_{pk}	31
Figura 7 – Distribuições dos processos A, B e C	35
Figura 8 – Distribuições de probabilidade (a) t-Student, (b) Gama, (c) Lognormal e (d) Weibull.....	37
Figura 9 – Histogramas das distribuições assimétricas.....	50
Figura 10(a) – Histograma dos valores de dureza	55
Figura 10(b) – Normal-plot dos valores de dureza	55

LISTA DE TABELAS

Tabela 1 – Índices de capacidade, sob suposição de normalidade, para os processos A, B e C	35
Tabela 2 – Erros absolutos em PPM obtidos nas 4 distribuições estudadas no caso unilateral	38
Tabela 3 – C_{pk} da distribuição normal equivalente ao número de não conformes estimados nas 4 distribuições	38
Tabela 4 – Índices de capacidade do processo obtidos pelo método de Clements ..	43
Tabela 5 – Índices de capacidade do processo propostos por Pearn e Chen.....	45
Tabela 6 – Características de distribuição de 6 processos não normais.....	47
Tabela 7 – Comparação entre S_{pmk} e C_{Npmk}	48
Tabela 8 – Estatísticas descritivas das distribuições assimétricas	50
Tabela 9 – “Distâncias relativas” da mediana em relação ao valor nominal.....	51
Tabela 10 – Índices de capacidade para as distribuições geradas	51
Tabela 11 – Comparação entre os índices de Clements e os índices de Pearn e Chen	53
Tabela 12 – Estatísticas descritivas da medição de dureza	55
Tabela 13 – Índices de capacidade do processo produtivo de barras de aço	56

LISTA DE ABREVIATURAS E SIGLAS

LIC	limite inferior de controle
LIE	limite inferior de especificação
LNI	limite natural inferior
LNS	limite natural superior
LSC	limite superior de controle
LSE	limite superior de especificação
PPM	partes por milhão

1 INTRODUÇÃO

A questão da qualidade nunca esteve em tanta evidência quanto nos tempos atuais. O que as empresas buscam hoje é atender as necessidades e exigências dos seus clientes por meio de processos cada vez mais eficientes. O objetivo das empresas é produzir com a menor perda possível, de tempo, de recursos e de custos, sem deixar de atender aos desejos e necessidades dos clientes.

Muitas empresas utilizam o controle de qualidade para atingir seus objetivos. Um ramo do controle da qualidade é o controle estatístico do processo que coleta, analisa e interpreta dados para sua utilização nas atividades de melhoria e controle da qualidade de produtos e serviços.

Antes da revolução industrial a qualidade dos produtos era controlada pela experiência dos artesãos, que dominavam todo o processo produtivo. Com o surgimento das indústrias, a produção foi dividida em operações que eram realizadas por diferentes funcionários, pois com isso conseguia-se produzir em massa e com baixo custo. Para garantir que os produtos finais tivessem a qualidade esperada pelos consumidores era realizada a inspeção 100%, ou seja, todos os produtos finais eram inspecionados. Com o crescimento da demanda e a intensificação da produção em massa, tornou-se necessário o uso de métodos estatísticos para a garantia da qualidade. A primeira técnica usada então foi a de inspeção por amostragem. Porém a inspeção final não melhorava a qualidade do produto, apenas fornecia o percentual de itens conformes e não conformes. Para melhorar a qualidade do produto final foi preciso reconhecer que a variabilidade era um fator presente nos processos produtivos e que ela deveria ser estudada.

Foi partindo desse estudo da variabilidade que se chegou ao conceito de processo estatisticamente sob controle. Um processo é dito estatisticamente sob controle quando somente causas comuns de variabilidade estiverem presentes (Figura 1(a)). Ao contrário, se um processo apresentar, além das causas comuns de

variabilidade, causas especiais, ele será dito fora de controle estatístico (Figura 1(b)).

Figura 1(a) - Processo sob controle estatístico

Figura 1(b) - Processo fora de controle estatístico

Para verificar se o processo está sob controle estatístico pode-se fazer uso de gráficos de controle. Quando se verifica que o processo encontra-se sob controle estatístico, pode-se medir o quanto esse processo consegue gerar produtos que atendam as especificações de projeto. Essas especificações refletem os desejos e exigências de seus clientes. Para isso, faz-se uso dos índices de capacidade do processo, que quantificam a capacidade do processo de uma empresa em gerar produtos que atendam às especificações.

1.1. Justificativa

A literatura sobre capacidade do processo tem abordado o tema dando maior enfoque a análise de dados normalmente distribuídos. Para este tipo de distribuição de dados, tem-se uma grande variedade de bibliografias que tratam dos índices de capacidade, mas é rara a bibliografia que aborde os índices de capacidade para dados que não atendam a suposição de normalidade.

Como na prática existem situações em que a normalidade dos dados não é atendida, surge a necessidade e a curiosidade de conhecer quais as alternativas para contornar essa limitação. Mais recentemente a análise de dados que não apresentam distribuição normal, denominados na literatura como dados não

normais, começaram a ser explorados, porém em português ainda são raras as referências que tratem do assunto.

Então, este trabalho tem como objetivo fazer um levantamento dos principais métodos de análise de capacidade do processo para dados que tenham distribuição não normal e compara-los por meio de um estudo de caso.

1.2. Objetivos

Esta dissertação tem como objetivo fazer um levantamento dos principais métodos de análise de capacidade do processo para dados que não sejam normalmente distribuídos e aplicá-los em um conjunto de dados reais, fazendo uso do software livre R para ilustrar como os índices podem ser utilizados em uma situação prática.

Outro objetivo é apresentar quais as distorções que ocorrem na interpretação dos índices quando não se tem a suposição de normalidade dos dados atendida, por meio de um estudo que aplica tanto os índices de capacidade para dados normais, já bem descritos na literatura, quanto os índices para dados não normais, que são abordados por este trabalho, num mesmo conjunto de dados que sejam não normalmente distribuídos.

Tem-se também por objetivo fazer um estudo comparativo entre os métodos de análise de capacidade para dados não normais para que se conclua sobre qual dos métodos apresentados é melhor.

1.3. Metodologia

Este trabalho é constituído basicamente de uma pesquisa exploratória, já que seu objetivo maior é proporcionar uma maior familiaridade com um problema específico, neste caso o uso de índices de capacidade quando não se tem normalidade dos dados de interesse. Conforme Silva e Menezes (2001), este tipo de pesquisa envolve, em geral, um levantamento bibliográfico, entrevistas com pessoas

que tiveram experiências práticas com o problema pesquisado e análise de exemplos que estimulem a compreensão.

A pesquisa exploratória pode assumir, em geral, as formas de pesquisas bibliográficas e estudos de caso. Neste trabalho foram feitas as duas formas de pesquisa. Inicialmente foi realizada a pesquisa bibliográfica, com a coleta de material em diversas fontes da literatura. E posteriormente foi descrito um estudo de caso da literatura e também foi realizado um estudo de caso prático, a fim de exemplificar os tipos de índices obtidos na pesquisa bibliográfica visando facilitar o conhecimento mais detalhado destes índices.

Para a realização da pesquisa bibliográfica foi feita uma revisão em livros e periódicos científicos, além de busca em sites da Internet sobre os métodos de análise da capacidade do processo, tanto para dados que apresentam distribuição normal como para dados que não tenham distribuição normal.

Para descrever o estudo de caso da literatura parte dos resultados da pesquisa de Pearn e Chen (1997) foi reproduzida neste trabalho. E para o estudo de caso prático, foram coletados dados reais, referentes à medição de dureza de barras de aço, de uma indústria metalúrgica da região de Campinas. Para a comparação dos métodos de análise propostos nos periódicos foi utilizado o software livre R, que se encontra disponível na internet para qualquer usuário que tenha um mínimo de familiaridade com linguagem de programação e conhecimento em estatística.

1.4. Estrutura

O trabalho está estruturado da seguinte maneira.

No capítulo 1 foi feita uma breve introdução sobre a utilização de índices de capacidade no controle estatístico do processo, ramo do controle de qualidade. Além disso, neste primeiro capítulo são apresentados os objetivos deste trabalho, suas justificativa, metodologia e limitações.

No capítulo 2 é feita uma revisão bibliográfica sobre capacidade do processo para dados normalmente distribuídos, explorando os quatro índices tradicionais que são amplamente abordados pela literatura. Em seguida é feita a revisão sobre a capacidade do processo para dados com distribuição não normal e são apresentados alguns dos principais métodos de análise de capacidade para este tipo de processo.

No capítulo 3 são apresentados estudos de comparação e de aplicação dos métodos apresentados no item 2.2, para distribuições não normais.

No capítulo 4 são apresentadas as considerações finais deste trabalho.

1.5. Limitações

Esta dissertação não tem a pretensão de reunir todos os métodos de análise de capacidade de processos não normalmente distribuídos. Dos diversos artigos pesquisados, apenas 3 métodos principais foram reunidos neste trabalho. Um método que seria interessante, mas que não abordado, é a transformação da variável não normal em normalmente distribuída. Chou *et. al.* (1998) sugere uma transformação de dados que pode ser aplicada em controle estatístico do processo.

Uma segunda limitação deste trabalho se refere ao item 3.1, onde são feitas as comparações entre os métodos de análise de capacidade para distribuições não normais. Neste item apenas dois dos três métodos abordados neste trabalho são comparados. O terceiro método não foi considerado no estudo comparativo do item 3.1, pois o autor deste método não deixou claro em sua publicação qual a escala que se utiliza para a interpretação do índice.

Por fim, outra limitação deste trabalho se refere à aplicação dos métodos de análise de capacidade de processos não normais. Como o banco de dados real mostrado no item 3.2 não apresentava itens fora das especificações, não se pôde estimar a diferença entre os índices de capacidade normais e não normais quanto ao erro na estimativa do número de itens não conformes. Porém, apesar de não ter sido possível aplicar esta comparação no banco de dados, no início do item

2.2 é apresentado um estudo feito por Somerville e Montgomery (1996-97) que calcula o erro cometido na estimação de número de itens não conformes para 4 diferentes distribuições simuladas com diferentes parâmetros, para o caso unilateral.

2 REVISÃO BIBLIOGRÁFICA

Um conjunto de procedimentos que visam produzir um serviço ou um produto pode ser definido como processo. Todos os processos possuem variabilidade, decorrentes das condições do ambiente ou do sistema de medição (WERKEMA, 1995).

A análise da capacidade do processo é uma técnica que quantifica a variabilidade do processo e permite uma comparação com as especificações do produto. O índice de capacidade do processo compara, geralmente por meio de uma razão, a amplitude da especificação de uma característica com a amplitude do processo de produção dessa característica.

2.1. Capacidade do Processo

Quando as causas especiais de variação são eliminadas de um processo com distribuição normal quanto à característica de interesse, diz-se que o processo está sob controle estatístico ou que se trata de um processo estável. Quando isso ocorre, é estabelecido um padrão de variação (Figura 2) e pode-se verificar em uma carta de controle que cerca de dois terços dos pontos situam-se próximos à linha central, que poucos pontos ficam juntos aos limites de controle e que ocorre uma distribuição balanceada dos pontos abaixo e acima do valor central. Porém, mesmo em processos estáveis, itens defeituosos são produzidos. Logo, não é suficiente colocar e manter um processo sob controle; é fundamental avaliar se o processo é capaz de atender às especificações estabelecidas a partir das necessidades dos clientes. É esta avaliação que constitui a análise da capacidade do processo.

Figura 2 - Padrão de variação de uma carta de controle

A capacidade do processo deve ser obtida para processos estáveis, sendo definida a partir dos limites naturais de sua variabilidade. Esses limites são dados por:

$$\text{Limite natural superior: } \mu + 3\sigma \quad (1)$$

$$\text{Limite natural inferior: } \mu - 3\sigma \quad (2)$$

onde μ é a média do processo e σ é o desvio-padrão do processo, sendo que a figura 3 ilustra estes limites.

Figura 3 - Limites naturais de variabilidade do processo

Para processos em que os dados tenham distribuição normal e que estejam centrados na média do processo, essa utilização de 3σ resultará em 99,73% da distribuição dentro dos limites estabelecidos. Neste ponto vale ressaltar que apesar de parecer pouco o percentual fora dos limites de especificação (0,27%) isto representa 2700 itens não conformes em um milhão de itens produzidos. Além disso,

se a distribuição do processo não for normal, então o percentual fora dos limites naturais do processo poderá diferir consideravelmente de 0,27%.

Para avaliar a capacidade do processo devem-se comparar os limites naturais do processo com os limites de especificação. A figura 4 mostra quatro situações possíveis de relacionamento entre esses limites. Em (a) o processo centrado atende às especificações, apresentando uma variabilidade menor que a exigida pelas especificações. Em (b) o processo, apesar de centrado, não atende às especificações, tendo em seus limites naturais valores maiores que os especificados. Em (c) o processo, apesar de descentrado, tem seus limites naturais de variabilidade menores que os de especificação, conseguindo atender às especificações. Em (d), além de descentrado, o processo apresenta um percentual de itens fora dos limites de especificação devido à variabilidade excessiva do processo (WERNER, 1996).

Figura 4 - Relações entre os limites naturais do processo e os limites de especificação

Segundo Montgomery (1997), a análise da capacidade do processo é uma parte importante de um programa de melhoria da qualidade, tendo aplicação em muitos segmentos de um ciclo produtivo, podendo ser citados alguns dos seus mais importantes usos:

- i) Predição de como um processo irá conservar-se dentro das especificações;
- ii) Seleção e/ou modificação de um processo;
- iii) Estabelecimento de um intervalo entre amostragens num processo de monitoramento;
- iv) Especificação de desempenho necessária para um novo equipamento;
- v) Seleção entre fornecedores concorrentes;
- vi) Estabelecimento de uma sequencia no processo de produção quando há efeito interativo do processo sobre as especificações;
- vii) Redução na variabilidade de um processo de produção.

Como técnicas usadas na análise de capacidade do processo podem ser citadas: a análise gráfica, feita por histograma ou gráfico de probabilidade, cartas de controle e os índices de capacidade do processo. Como o interesse desta dissertação é estudar os índices de capacidade, apenas eles serão abordados.

Os índices de capacidade são números adimensionais que permitem uma quantificação do desempenho de processos. Quando o processo está sob controle estatístico e a variável de interesse tem distribuição normal, quatro índices são citados na literatura, sendo eles: C_p , C_{pk} , C_{pm} e C_{pmk} .

2.1.1. Índice C_p

O índice C_p , chamado de índice de capacidade potencial do processo mede quanto o processo pode produzir produtos satisfazendo as especificações. Este índice considera que o processo está centrado no valor nominal da

especificação. Caso a variável de interesse tenha distribuição bilateral, o índice C_p é definido por:

$$C_p = \frac{LSE - LIE}{6\sigma} \quad (3)$$

em que LSE é o limite superior de especificação, LIE é o limite inferior de especificação e σ é o desvio-padrão do processo.

Como este índice relaciona a variabilidade permitida ao processo (especificada no projeto) com a variabilidade natural do processo, tem-se que quanto maior for o valor de C_p , maior será a capacidade do processo em satisfazer às especificações, desde que a média esteja centrada no valor nominal. Existe uma regra prática para a análise deste índice:

- i) Quando $C_p < 1$, diz-se que o processo é incapaz de atender às especificações (Figura 5(a)). Neste caso seria necessário tentar diminuir a variabilidade deste processo ou usar outro processo que fosse mais adequado às especificações, pois a produção de itens não conformes está acima de 2700 partes por milhão (PPM) da produção.
- ii) Quando $1 \leq C_p \leq 1,33$, diz-se que o processo é aceitável ou relativamente capaz de produzir dentro das especificações (Figura 5(b)). O uso de cartas de controle neste tipo de processo é indicado para manter o processo sob controle e evitar a produção de itens fora da especificação. A produção de itens não conformes fica entre 64 PPM e 2700 PPM da produção.
- iii) E quando $C_p \geq 1,33$, diz-se que o processo é capaz de atender às especificações (Figura 5(c)). Quando se tem este tipo de processo, podem-se coletar amostras periódicas para o acompanhamento do processo, já que o número de itens não conformes produzidos está abaixo de 64 PPM.

Figura 5 - Comparação dos histogramas com as especificações

O índice C_p , por não levar em consideração a localização do processo, só depende da amplitude do intervalo de especificação e da variabilidade natural do processo para o seu cálculo. Como consequência disto, para um determinado valor de C_p , pode-se ter qualquer percentual de itens fora das especificações, este percentual vai depender apenas de onde esteja localizada a média do processo. Então, o índice C_p dá uma ideia de quanto o processo é potencialmente capaz de produzir dentro do intervalo especificado em projeto. Caso a média coincida com o valor nominal da especificação facilmente se observa esta ideia.

2.1.2. Índice C_{pk}

Como na prática nem sempre o processo está centrado no valor nominal da especificação, e então o uso do índice C_p pode levar a conclusões erradas, Kane (1986) propôs o índice de desempenho C_{pk} , que leva em consideração a distância da média do processo em relação aos limites de especificação. Este índice é dado por:

$$C_{pk} = \text{mínimo} \left(\frac{LSE - \mu}{3\sigma}, \frac{\mu - LIE}{3\sigma} \right) \quad (4)$$

onde LSE é o limite superior de especificação, LIE é o limite inferior de especificação, μ é a média do processo e σ é o desvio-padrão do processo.

Se o processo estiver centrado no valor nominal de especificação, $C_p = C_{pk}$. Então, caso C_p seja diferente de C_{pk} , sabe-se que o processo está descentrado, isto é, que a média não coincide com o valor nominal das especificações. As interpretações do índice C_{pk} podem ser feitas pela regra mostrada para o índice C_p , já que a análise da capacidade do processo é feita usando estes dois índices em conjunto.

Já sendo compreendidos os dois índices, C_p e C_{pk} , pode-se visualizar o risco de interpretações erradas do índice C_p quando o processo não está centrado no valor nominal de especificação pela figura 6. Enquanto os gráficos (a), (b) e (c) representam processos centrados, os gráficos (d), (e) e (f) representam processos descentrados. Como C_{pk} considera a distância entre a média do processo e o valor nominal de especificação, este índice fornece uma interpretação mais precisa da capacidade do processo que C_p . Note que em (d) $C_{pk} < 1$ enquanto $C_p = 1$. Em (b) e (e) tem-se processos com menor variabilidade. Como C_p relaciona apenas a variabilidade do processo com a variabilidade permitida (especificação), apresenta valor maior que 1, indicando que o processo é potencialmente capaz. Mas em (e), apesar da menor variabilidade, o processo está descentrado, e verifica-se novamente a diferença de interpretação entre os dois índices. Já em (c) e (f), como os processos apresentam variabilidade natural maior que a permitida, ambos os índices indicam que o processo não é capaz.

Figura 6 - Interpretação dos índices C_p e C_{pk} .

2.1.3. Índices C_{pm} e C_{pmk}

Esses índices são alternativos aos índices anteriores e consideram, além da variância do processo, a distância de sua média em relação ao valor nominal da especificação.

O índice C_{pm} é dado por:

$$C_{pm} = \frac{LSE - LIE}{6\sqrt{\sigma^2 + (\mu - T)^2}} \quad (5)$$

onde LSE é o limite superior de especificação, LIE é o limite inferior de especificação, μ é a média do processo, σ é o desvio-padrão do processo e T é o valor nominal da especificação.

Pela definição do índice C_{pm} , um aumento na variabilidade do processo faz com que o denominador do índice aumente e consequentemente o valor do índice diminuirá. Além disso, um distanciamento maior do processo em relação ao valor nominal também poderá provocar um aumento no denominador do índice, tornando-o menor.

Uma vantagem do índice C_{pm} em relação ao índice C_p é que ele fornece uma boa ideia da capacidade do processo tanto para os processos que se apresentam próximos ao valor nominal quanto para os que se apresentam mais afastados dele. Além disso, segundo Chan *et. al.* (1988), se o processo segue uma distribuição normal e a média do processo está centrada no valor nominal de especificação, o índice C_{pm} apresenta distribuição semelhante à do índice C_p , mas seu estimador $\hat{C}_{pm}$ é mais eficiente e apresenta viés menor que $\hat{C}_p$, estimador de C_p .

Porém este índice considera apenas a variabilidade permitida ao processo no numerador. Para refinar ainda mais a análise da capacidade do processo, outro índice é proposto: o C_{pmk} , que considera a menor distância entre a

média do processo em relação aos limites de especificação em seu numerador, e é dado pela expressão abaixo:

$$C_{pmk} = \text{mínimo} \left(\frac{LSE - \mu}{3\sqrt{\sigma^2 + (\mu - T)^2}}, \frac{\mu - LIE}{3\sqrt{\sigma^2 + (\mu - T)^2}} \right) \quad (6)$$

onde LSE é o limite superior de especificação, LIE é o limite inferior de especificação, μ é a média do processo, σ é o desvio-padrão do processo e T é o valor nominal da especificação.

2.1.4. Formulação única dos índices

Vännman (1995) construiu uma fórmula única da qual podem ser derivados esses índices de capacidade, isto é, os quatro índices básicos (C_p , C_{pk} , C_{pm} , C_{pmk}) poderiam ser vistos como casos especiais. Essa fórmula é definida da seguinte maneira:

$$C_p(u, v) = \frac{d - u|\mu - m|}{3\sqrt{\sigma^2 + v(u - T)^2}} \quad (7)$$

em que: μ é a média e σ é o desvio-padrão do processo, $d = \frac{(LSE - LIE)}{2}$ é a metade do comprimento do intervalo de especificação, $m = \frac{(LSE + LIE)}{2}$ é o ponto médio entre os limites de especificação, T é o valor nominal e $u, v \geq 0$.

Para se chegar aos índices básicos, basta substituir u e v por combinações de 0 e 1 na fórmula inicial. Assim,

- $C_p(0,0) = C_p$;
- $C_p(1,0) = C_{pk}$;
- $C_p(0,1) = C_{pm}$;
- $C_p(1,1) = C_{pmk}$.

A relação entre os quatro índices pode ser estabelecida da seguinte forma:

$$C_{pm} = C_p \left[1 + \left(\frac{\mu - T}{\sigma} \right)^2 \right]^{-\frac{1}{2}} \quad (8) \text{ e}$$

$$C_{pmk} = C_{pk} \left[1 + \left(\frac{\mu - T}{\sigma} \right)^2 \right]^{-\frac{1}{2}}. \quad (9)$$

Além disso, segundo Pearn e Chen (1997) pode-se obter um *ranking* dos quatro índices em termos da sensibilidade de captar a distância entre a média do processo e o valor nominal de especificação, sendo eles, do mais sensível ao menos sensível: C_{pmk} , C_{pm} , C_{pk} e C_p . E, finalizando, tem-se que para um processo com limites de especificação simétricos em relação ao valor nominal e que apresenta a média sobre esse valor nominal, os índices C_{pk} , C_{pm} e C_{pmk} tornam-se iguais ao índice C_p .

2.2. Capacidade do processo não normal

Muitas indústrias, pelo desconhecimento dos métodos de análise da capacidade de processos não normais, assumem que seu processo fornece dados que sejam normalmente distribuídos e calculam os índices citados no capítulo anterior para analisar a capacidade de produção de itens conformes. Porém essa prática pode gerar interpretações errôneas sobre a capacidade do processo, já que os índices C_p , C_{pk} , C_{pm} e C_{pmk} supõem normalidade da variável de interesse.

Para exemplificar que tipo de interpretação errônea pode ser provocado pelo uso desses índices básicos em distribuições não normais, Pearn e Chen (1997) consideram três processos A, B e C, todos com distribuição qui-quadrado com 2 graus de liberdade. Na figura 7 são mostradas as distribuições dos três processos e suas relações com os limites de especificação, sendo que o processo A apresenta média igual a 30, o processo B, média igual a 37, e o processo C apresenta média igual a 44. Ambos os processos apresentam desvio-padrão igual a 2.

Figura 7 - Distribuições dos processos A, B e C

Observando a figura 7, nota-se que o processo A apresenta média igual ao LIE, a média do processo B é igual ao valor nominal de especificação e que o processo C tem média igual ao LSE. A partir destas informações, tem-se que o processo B está centrado no valor nominal e que os processos A e C apresentam-se igualmente, porém opostamente, distantes do valor nominal. Calculando-se os quatro índices de capacidade, supondo normalidade destes processos, verificam-se, conforme a tabela 1, que eles não permitem estimar um percentual correto de itens não conformes, já que esse percentual para o processo A é de 63% e para C é de 37%, mas os índices dos processos A e C não apresentaram diferença entre si.

Tabela 1 - Índices de capacidade, sob suposição normalidade, para os processos A, B e C

Processo	C_p	C_{pk}	C_{pm}	C_{pmk}
A	1,17	0,00	0,32	0,00
B	1,17	1,17	1,17	1,17
C	1,17	0,00	0,32	0,00

Somerville e Montgomery (1996-1997) fizeram um estudo para verificar os efeitos da falta de normalidade dos dados no cálculo do número de itens não conformes produzidos por um processo em quatro famílias de distribuições não normais:

(a) t-Student com função densidade de probabilidade (fdp)

$$f(t) = \frac{\Gamma\left(\frac{k+1}{2}\right)}{\sqrt{k\pi}(k/2)} \left(\frac{t^2}{k} + 1\right)^{-\frac{k+1}{2}}, \quad -\infty < t < \infty. \quad \text{A distribuição}$$

apresenta k graus de liberdade, média $\mu=0$ e variância

$$\sigma^2 = \frac{k}{k-2}.$$

(b) Gama com fdp $f(t) = \frac{1}{\beta^\alpha \Gamma(\alpha)} x^{\alpha-1} e^{-\frac{x}{\beta}}$, onde $\alpha \geq 1$ e $\beta > 0$

sendo que α é o parâmetro de forma e β é o parâmetro de escala. A distribuição tem média $\mu=\beta\alpha$ e variância $\sigma^2=\alpha\beta^2$.

(c) Lognormal com fdp $f(x) = \frac{1}{\sqrt{2\pi}\beta x} \exp\left(-\frac{1}{2\beta^2}(\ln x - \alpha)^2\right)$, $x \geq 0$.

A distribuição tem média $\mu = e^{\alpha+\beta^2/2}$ e variância $\sigma^2 = e^{2\alpha+\beta}(e^2 - 1)$;

(d) Weibull com fdp $f(x) = \frac{\alpha}{\beta} \left(\frac{x}{\beta}\right)^{\alpha-1} \exp\left[-\left(\frac{x}{\beta}\right)^\alpha\right]$, onde $x > 0$, $\alpha > 0$

e $\beta > 0$, sendo que α é o parâmetro de forma e β é o parâmetro de escala. A distribuição tem média

$$\mu = \beta \Gamma\left(1 + \frac{1}{\alpha}\right) \text{ e variância } \sigma^2 = \beta^2 \left\{ \Gamma\left(1 + \frac{2}{\alpha}\right) - \left[\Gamma\left(1 + \frac{1}{\alpha}\right) \right]^2 \right\}.$$

Na figura 8 pode-se observar o comportamento destas famílias de distribuições para diferentes parâmetros.

Neste estudo diversos valores de C_{pk} foram usados para estimar o número de itens não conformes, em partes por milhão (PPM), considerando que o processo seguia uma distribuição normal. Com isso os autores verificaram que pequenos afastamentos de uma curva em relação a normal produziam sérios erros de estimação do número de não conformes. As tabelas 2 e 3 resumem alguns resultados obtidos pelo estudo.

Figura 8 - Distribuições de probabilidade

(a) t-Student, (b) Gama, (c) Lognormal e (d) Weibull

Tabela 2 - Erros absolutos em PPM obtidos nas 4 distribuições estudadas no caso unilateral

Valor C_{pk}	t-Student		Gama ($\beta = 1$)		Lognormal		Weibull ($\beta = 1$)	
	5 gl	20 gl	$\alpha = 1,5$	$\alpha = 3$	$\beta = 0,25$	$\beta = 1$	$\alpha = 1$	$\alpha = 3$
0,67	1906	1164	25396	21440	15788	14224	27037	2801
1,00	4512	1100	14475	10446	7063	16698	16966	148
1,33	1755	180	5063	2905	1668	9831	6706	-5
1,67	664	18	1617	696	332	5833	2478	-0,17
2,00	287	2	508	159	65	3658	912	-0,0009

Tabela 3 - C_{pk} da distribuição normal equivalente ao número de não conformes estimados nas 4 distribuições

Valor C_{pk}	t-Student		Gama ($\beta = 1$)		Lognormal		Weibull ($\beta = 1$)	
	5 gl	20 gl	$\alpha = 1,5$	$\alpha = 3$	$\beta = 0,25$	$\beta = 1$	$\alpha = 1$	$\alpha = 3$
0,67	-	-	0,55	0,57	0,59	0,60	0,55	0,65
1,00	-	-	0,72	0,75	0,80	0,70	0,70	0,99
1,33	-	-	0,86	0,92	0,98	0,78	0,82	1,35
1,67	0,97	1,17	0,98	1,07	1,13	0,84	0,94	1,72
2,00	-	-	1,10	1,20	1,28	0,89	1,04	2,12

Para a distribuição t-Student, Somerville e Montgomery (1996-97) obtiveram os erros para diversos graus de liberdade tanto para o caso unilateral quanto para o caso bilateral, porém as tabelas 3 e 4 apresentam apenas alguns valores obtidos para 5 e 20 graus de liberdades. Pelos valores de erros obtidos verificou-se que quanto maior o número de graus de liberdade, menor era o erro de estimação do número de itens não conformes, já que quanto maior os graus de liberdade, mais a distribuição t-Student se aproxima da distribuição normal. Além disso, verificou-se que quanto maior era a capacidade do processo, menor era o erro de estimação dos itens não conformes. Para $C_{pk}=1,33$ e 5 graus de liberdade, o erro unilateral é de 1755 PPM; e para $C_{pk}=2,00$ e 20 graus de liberdade, o erro unilateral é de apenas 2 PPM.

Para a distribuição gama, os autores calcularam os erros em relação ao número de não conformes produzidos pela suposição de normalidade de dados que seguissem a distribuição gama com especificação unilateral. O parâmetro β foi mantido constante ($\beta=1$) e como resultado, tinha-se que quanto maior o valor de α , menor era o erro obtido. E assim como para a distribuição t-Student, quanto maior a capacidade do processo, menor era o erro na estimação de itens não conformes. Para $\alpha=1,5$ e $C_{pk}=0,67$ o erro obtido foi de 25396 PPM. Para se ter esse número de não conformes na distribuição normal, o C_{pk} deveria ser igual a 0,55. Já para valores maiores α e C_{pk} , $\alpha=3,00$ e $C_{pk}=2,00$ o erro obtido foi de 159 PPM. Para esse número de não conformes na distribuição normal, o C_{pk} seria igual a 1,20.

A distribuição lognormal, segundo Somerville e Montgomery (1996-97), é a mais sensível às variações dos parâmetros quando se considera a capacidade do processo, isto é, o número de não conformes varia muito com pequenas mudanças nos parâmetros α e β . Para uma distribuição lognormal com $\beta=1$, ao supor normalidade dos dados e calcular-se o C_{pk} , apesar de obter-se um valor de 1,33 para o índice, o número de não conformes estimados é de 9863 PPM. Se os dados realmente seguissem uma distribuição normal, o número de itens não conformes estimados para este valor de C_{pk} seria de 32 PPM. Com isso tem-se que o uso do índice inadequado gerou um erro de 9831 PPM na estimação do número de itens não conformes.

A distribuição Weibull apresenta diversos formatos, dependendo dos valores de seus parâmetros. Na figura 8(d) tem-se $\beta=1$ e alguns valores de α . Quanto maior o valor de α , mais próxima da distribuição normal a distribuição Weibull se torna, sendo que para $\alpha>3$, ela apresenta uma área sob sua cauda menor que a da distribuição normal. Com isso, Somerville e Montgomery (1996-97) encontraram valores negativos para a estimativa do número de não conformes quando se supunha normalidade para dados com distribuição Weibull. Na tabela 3 observa-se que para $\alpha=3$ e valores de C_{pk} maiores que 1,33, têm-se erros negativos na estimação do número de itens não conformes.

Com essas quatro distribuições não normais Somerville e Montgomery (1996-1997) exemplificaram como a suposição de normalidade, quando não atendida, pode produzir estimativas erradas do número de itens não conformes. Então, quando a suposição de normalidade dos dados não é atendida, existem métodos alternativos para a análise da capacidade do processo. Nesta dissertação serão abordados 3 desses métodos.

O primeiro deles, proposto por Clements (1989), se baseia em medidas de percentil para obter os índices de capacidade C'_p e C'_{pk} . A partir desse método de Clements; Pearn e Kotz; Pearn e Chen *apud* Pearn e Chen (1997) propuseram uma generalização da formulação única apresentada na equação (7) para que se pudessem obter quatro índices com interpretações semelhantes aos dos índices para dados normais.

O segundo método, proposto por Pearn e Chen (1997) também faz uso dos percentis para o cálculo dos índices, diferindo apenas quanto aos índices C'_{pk} e C'_{pmk} , pois no cálculo do denominador destes índices, é considerada a metade da distância entre os percentis da distribuição, enquanto no primeiro método se tem a distância entre a mediana e o percentil, superior ou inferior.

O terceiro método foi proposto por Chen e Ding (2001) e considera em seu cálculo a variabilidade do processo, a distância entre a média do processo e o valor nominal, e a proporção de itens não conformes, obtendo o índice S_{pmk} . Este índice proposto por neste terceiro método, consegue refletir exatamente o número de não conformes do processo, sendo superior aos outros dois métodos anteriores (CHEN; DING, 2001).

2.2.1. Método 1

Clements (1989) propôs um método simples de cálculo dos índices C_p e C_{pk} para qualquer tipo de distribuição dos dados de interesse, usando a família de curvas de Pearson. Essa família de curvas foi publicada em 1893, pelo matemático

Karl Pearson, e inclui diversas distribuições, sendo elas a normal, a lognormal, a t-student, a F, a beta e a gama.

Segundo o autor, seu método de cálculo dos índices apresenta algumas vantagens:

- i) Quando a distribuição dos dados é normal, os índices calculados por seu método são exatamente iguais àqueles calculados pelo método tradicional.
- ii) A única diferença em relação ao método tradicional de cálculo dos índices é a forma de calcular o comprimento e a posição das metades superior e inferior da distribuição em relação aos limites de especificação.
- iii) O método não necessita de transformações matemáticas dos dados.
- iv) O método é de fácil cálculo manual.
- v) O método pode ser facilmente usado como uma sub-rotina na maior parte das sub-rotinas de controle estatístico do processo existente.
- vi) A família de curvas de Pearson, em que está baseado o método de Clements, pode ser usada para fornecer estimativas de percentagens de itens não conformes para uma grande variedade de distribuições.
- vii) O método pode ser aplicado em qualquer família de curvas de probabilidade.

O índice C_p de Clements é definido por:

$$C'_p = \frac{LSE - LIE}{F_{99,865} - F_{0,135}} \quad (10)$$

em que LSE é o limite superior de especificação, LIE é o limite inferior de especificação e F_α é o α -ésimo percentil do processo.

Mas como o índice C'_p considera apenas a dispersão do processo e não leva em conta a sua posição em relação aos limites de especificação, ele pode levar a uma interpretação errada da capacidade do processo. Considerando também a posição do processo, tem-se o índice C'_{pk} de Clements é definido como o mínimo de dois índices, C'_{pI} e C'_{pS} , definidos abaixo:

$$C'_{pI} = \frac{M - LIE}{M - F_{0,135}} \quad (11)$$

$$C'_{pS} = \frac{LSE - M}{F_{99,865} - M} \quad (12)$$

onde M é a mediana do processo, LSE é o limite superior de especificação, LIE é o limite inferior de especificação e F_α é o α -ésimo percentil do processo.

A mediana é usada por esse método como medida de valor central para que se garanta que C_{pI} e C_{pS} meçam a relação entre as metades superior e inferior do processo com os valores superior e inferior de especificação.

Quando a distribuição dos dados é normal, a distância entre a mediana e cada um dos limites de especificação é igual a 3σ . Com isso o índice C'_p de Clements se reduz ao índice C_p tradicional, mostrado na equação (3). Além disso, o índice C'_{pk} de Clements também se reduz ao índice C_{pk} tradicional da equação (4).

Posteriormente Pearn e Kotz; Pearn e Chen *apud* Pearn e Chen (1997), baseados neste método de Clements, propuseram uma generalização da formulação única mostrada na equação (7), a partir da qual se obtém quatro índices para dados com distribuição não normal. Essa generalização é dada por:

$$C'_p(u,v) = (1-u) \times \frac{LSE - LIE}{6\sqrt{\left[\frac{F_{99,865} - F_{0,135}}{6}\right]^2 + v(M-T)^2}} + u \times \text{mínimo} \left\{ \frac{LSE - M}{3\sqrt{\left[\frac{F_{99,865} - M}{3}\right]^2 + v(M-T)^2}}, \frac{M - LIE}{3\sqrt{\left[\frac{M - F_{0,135}}{3}\right]^2 + v(M-T)^2}} \right\} \quad (13)$$

em que LSE e LIE são os limites superior e inferior de especificação, respectivamente; F_α é o α -ésimo percentil do processo, M é a mediana, T é o valor nominal do processo e $u,v \geq 0$.

Substituindo em (u,v) a combinação de 0 e 1, obtêm-se os quatro índices equivalentes aos índices tradicionais, citados no item 2.1, que podem ser interpretados da mesma forma, porém atendendo a qualquer distribuição, não se restringindo apenas à normal. Esses índices são mostrados na tabela 4.

Tabela 4 - Índices de capacidade do processo obtidos pelo método de Clements

(u,v)	$C'_p(u,v)$
(0,0)	$C'_p = \frac{LSE - LIE}{F_{99,865} - F_{0,135}} \quad (10)$
(0,1)	$C'_{pk} = \text{mínimo} \left(\frac{LSE - M}{F_{99,865} - M}, \frac{M - LIE}{M - F_{0,135}} \right) \quad (14)$
(1,0)	$C'_{pm} = \frac{LSE - LIE}{6\sqrt{\left[\frac{F_{99,865} - F_{0,135}}{6}\right]^2 + (M - T)^2}} \quad (15)$
(1,1)	$C'_{pmk} = \text{mínimo} \left(\frac{LSE - M}{3\sqrt{\left[\frac{F_{99,865} - M}{3}\right]^2 + (M - T)^2}}, \frac{M - LIE}{3\sqrt{\left[\frac{M - F_{0,135}}{3}\right]^2 + (M - T)^2}} \right) \quad (16)$

2.2.2. Método 2

Fazendo uso da formulação única apresentada na equação (7), Pearn e Chen (1997) propuseram uma generalização, chamada de $C_{Np}(u,v)$, que atende casos em que a suposição de normalidade não seja verificada. Essa generalização é dada por:

$$C_{Np}(u,v) = \frac{d - u|M - m|}{3\sqrt{\left[\frac{F_{99,865} - F_{0,135}}{6}\right]^2 + v(M - T)^2}} \quad (17)$$

onde F_α é o α -ésimo percentil, M é a mediana da distribuição, $d = \frac{(LSE - LIE)}{2}$ é a metade do comprimento do intervalo de especificação, $m = \frac{(LSE + LIE)}{2}$ é o ponto médio entre os limites de especificação, T é o valor nominal e $u, v \geq 0$.

Para desenvolver essa generalização, Pearn e Chen (1997) substituíram o desvio padrão σ de $C_p(u,v)$ por $\frac{F_{99,865} - F_{0,135}}{6}$ em $C_{Np}(u,v)$, pois assim é mantida a ideia de medir a variabilidade inerente ao processo, mas não se restringindo à distribuição normal. Outra substituição feita foi da média μ em $C_p(u,v)$ pela mediana M em $C_{Np}(u,v)$, já que a mediana é uma medida mais robusta para tendência central que a média quando existe uma distribuição assimétrica, especialmente se essa distribuição tiver caudas longas.

Da mesma forma que no primeiro método, substituindo em (u,v) a combinação de 0 e 1, obtêm-se os quatro índices equivalentes aos citados no item 2.1, que também podem ser interpretados da mesma forma, porém atendendo a qualquer distribuição, não se restringindo apenas à normal. Note que $C_{Np} = C'_p$ e que $C_{Npk} = C'_{pk}$. Os índices $C_{Np}(u,v)$ são mostrados na tabela 5.

Tabela 5 - Índices de capacidade do processo propostos por Pearn e Chen (1997)

(u,v)	$C_{Np}(u,v)$
(0,0)	$C_{Np} = \frac{LSE - LIE}{F_{99,865} - F_{0,135}} \quad (10)$
(0,1)	$C_{Npk} = \min \left(\left[\frac{LSE - M}{\frac{F_{99,865} - F_{0,135}}{2}} \right], \left[\frac{M - LIE}{\frac{F_{99,865} - F_{0,135}}{2}} \right] \right) \quad (18)$
(1,0)	$C_{Npm} = \frac{LSE - LIE}{6 \sqrt{\left[\frac{F_{99,865} - F_{0,135}}{6} \right]^2 + (M - T)^2}} \quad (15)$
(1,1)	$C_{Npmk} = \min \left(\frac{LSE - M}{3 \sqrt{\left[\frac{F_{99,865} - F_{0,135}}{6} \right]^2 + (M - T)^2}}, \frac{M - LIE}{3 \sqrt{\left[\frac{F_{99,865} - F_{0,135}}{6} \right]^2 + (M - T)^2}} \right) \quad (19)$

Assim como ocorre com os índices tradicionais do item 2.1, pode ser estabelecida uma relação entre os índices: $C_{Npm} = C_{Np} \{1 + [(\mu - T)/\sigma'']^2\}^{-1/2}$ e $C_{Npmk} = C_{Npk} \{1 + [(\mu - T)/\sigma'']^2\}^{-1/2}$, onde $\sigma'' = (F_{99,865} - F_{0,135})/6$. Também existe um *ranking* (PEARN e CHEN, 1997), para os quatro índices propostos, segundo sua sensibilidade em captar a distância entre a mediana do processo e o valor nominal; do mais sensível ao menos, tem-se: C_{Npmk} , C_{Npm} , C_{Npk} e C_{Np} . Quando as especificações são simétricas em relação ao valor nominal, $C_{Npk} = (1-k)C_{Np}$ e $C_{Npmk} = (1-k)C_{Npm}$, onde $k = |M - T|/d$ e $d = \frac{(LSE - LIE)}{2}$. Se o processo com distribuição não normal está centrado no valor nominal, isto é, $M=T$, $C_{Np} = C_{Npk} = C_{Npm} = C_{Npmk} = d/3\sigma''$. Mais ainda, se a distribuição dos dados for normal, tem-se $M = m$ e $s'' = s$, e então vale ressaltar que as generalizações $C_{Np}(u,v)$ se reduzirão aos quatro índices básicos $C_p(u,v)$, obtendo-se $C_{Np} = C_p$, $C_{Npk} = C_{pk}$, $C_{Npm} = C_{pm}$ e $C_{Npmk} = C_{pmk}$.

2.2.3. Método 3

O terceiro método se refere à proposta de Chen e Ding (2001). Os autores propuseram um índice denominado S_{pmk} , que atende a qualquer distribuição de dados, e considera em seu cálculo a variabilidade do processo, a distância da média do processo em relação ao valor nominal e a proporção de não conformes. Esse índice é dado por:

$$S_{pmk} = \frac{\Phi^{-1}\left(\frac{1 + F(LSE) - F(LIE)}{2}\right)}{3\sqrt{1 + \left(\frac{\mu - T}{\sigma}\right)^2}} \quad (20)$$

em que $F(x)$ é a função de distribuição acumulada da variável aleatória inerente ao processo, μ é a média do processo, σ é o desvio padrão do processo, T é o valor nominal de especificação, LIE é o limite inferior de especificação e LSE é o limite superior de especificação.

Segundo os autores, a construção deste índice vem da seguinte propriedade dos índices:

$$C_{pmk} = C_{pm}(1 - K) = \frac{C_p}{\sqrt{1 + \left(\frac{\mu - T}{\sigma}\right)^2}} (1 - K) = \frac{C_{pk}}{\sqrt{1 + \left(\frac{\mu - T}{\sigma}\right)^2}} \quad (21)$$

e segundo Chen (2000), C_{pk} pode ser substituído por $\frac{1}{3}\Phi^{-1}\left(\frac{1 + F(LSE) - F(LIE)}{2}\right)$ em qualquer distribuição de probabilidade.

A proporção de itens não conformes pode ser estimada a partir de S_{pmk} por

$$P = 2 \times \left[1 - \Phi\left(3S_{pmk}\sqrt{1 + \left(\frac{\mu - T}{\sigma}\right)^2}\right) \right]. \quad (22)$$

Chen e Ding (2001) aplicaram o índice S_{pmk} em 6 processos não normais, cujas características de distribuição são mostradas na tabela 6. Para todos

estes processos a especificação nominal era de 17,8 e os limites inferior e superior tinham valores iguais a 10,0 e 25,6, respectivamente.

Tabela 6 - Características de distribuição de 6 processos não normais

Processo	Distribuição	μ	M	σ	$F_{0,135}$	$F_{99,865}$
A	$\chi_3^2 + 7$	10,00	9,366	2,45	7,030	22,630
B	$\chi_3^2 + 14,8$	17,80	17,166	2,45	14,830	30,430
C	$\chi_3^2 + 22,6$	25,60	24,966	2,45	22,630	38,230
D	Gama(6,3)	18,00	17,010	7,35	3,525	48,104
E	Gama(1,12)	12,00	8,318	12,00	0,016	79,292
F	Uniforme(17;25,8)	21,40	21,400	2,54	17,012	25,788

Chen e Ding (2001) afirmam em seu estudo que a proporção de itens não conformes obtidas pela equação (22) não difere da proporção real de não conformes, dada por $P(X < 10,0) + P(X > 25,6)$. Para mostrar o bom desempenho de seu índice, S_{pmk} , em relação à generalização realizada por Pearn e Chen (1997), descrita na equação (19), os autores compararam as proporções de itens não conformes estimados a partir dos dois índices. As estimativas dos índices e as respectivas proporções de não conformes são mostradas na tabela 7, sendo que

$$P'_{CNpmk} = 2 \times \left[1 - \Phi \left(3C_{Npmk} \sqrt{1 + \left(\frac{M - T}{(F_{99,865} - F_{0,135})/6} \right)^2} \right) \right] \quad (23)$$

$$e \quad ER = \frac{|P'_{CNpmk} - P|}{P}. \quad (24)$$

Tabela 7 - Comparação entre S_{pmk} e C_{Npmk}

Processo	P	S_{pmk}	C_{Npmk}	P'_{CNpmk}	ER (%)
A	0,6087	0,0511	0,0000	1,0000	64,3
B	0,0129	0,8292	0,8925	0,0059	54,3
C	0,3916	0,0856	0,0277	0,8074	106,2
D	0,2683	0,3689	0,3128	0,3454	28,7
E	0,7571	0,0928	0,0000	1,0000	32,1
F	0,0227	0,4378	0,3603	0,0041	87,9

A partir da tabela 7, tem-se que para todos os processos selecionados para o estudo de Chen e Ding (2001) existe uma diferença representativa entre as proporções de não conformes estimadas pelos índices S_{pmk} e C_{Npmk} . Segundo os autores, isto evidencia uma superioridade do índice proposto por eles em relação ao índice C_{Npmk} , apresentado no método 2 deste trabalho.

Quanto à interpretação do índice S_{pmk} , os autores mostram apenas um exemplo, onde o valor do índice obtido $S_{pmk}=0,4092$ é interpretado como reflexo de um processo não capaz.

3 ESTUDOS DE CASO

Uma vez apresentados alguns métodos de análise de capacidade para dados com distribuição não normal encontrados na literatura, é interessante exemplificar a aplicação destes métodos para que a compreensão destes seja facilitada. Este capítulo é composto por dois estudos de casos, que buscam mostrar um estudo comparativo e uma aplicação dos métodos apresentados no capítulo anterior.

O primeiro estudo de caso apresenta uma comparação entre os índices de capacidade para processos não normais, incluindo um estudo que foi realizado por Pearn e Chen (1997).

O segundo estudo de caso consiste da aplicação dos três métodos de análise de capacidade do processo não normal apresentados no item 2.2 em um banco de dados real, obtido de uma metalúrgica da região de Campinas.

Todos os resultados são obtidos a partir do software R versão 2.10. Os comandos usados para a execução do estudo comparativo encontram-se no anexo I e os comandos utilizados na aplicação dos métodos no banco de dados real encontram-se no anexo II.

3.1. Estudo comparativo entre os métodos de análise de capacidade do processo não normal

Neste estudo de caso é mostrado um estudo comparativo entre os dois primeiros métodos de análise de capacidade não normal citados no item 2.2 desta dissertação.

Relembrando, os dois métodos que serão comparados fazem uso dos percentis para o cálculo dos índices, porém enquanto os índices de Clements (1989)

consideram em seu denominador a distância entre a mediana e o percentil, superior ou inferior, para o cálculo dos índices C'_{pk} e C'_{pmk} , os índices de Pearn e Chen (1997) fazem uso da metade da distância entre os percentis da distribuição no cálculo do denominador dos índices C_{Npk} e C_{Npmk} .

Devido a essa diferença entre os métodos, quanto mais assimétrica for a distribuição dos dados, mais diferentes serão os denominadores desses índices e por consequência, mais diferentes serão os resultados obtidos da análise de capacidade do processo desses dois métodos. Para exemplificar essa diferença entre os dois métodos foram geradas três distribuições assimétricas: gama(2,100), gama(5,100) e gama(15,100), cada uma com 1000 repetições, cujos histogramas são mostrados na figura 9 e algumas estatísticas descritivas estão na tabela 8.

Figura 9 - Histogramas das distribuições assimétricas

Tabela 8 - Estatísticas descritivas das distribuições assimétricas

Distribuição	Mínimo	Máximo	Mediana	$d=(Max-Min)/2$
Gama(2,100)	0,0002	0,0881	0,0170	0,0440
Gama(5,100)	0,0087	0,1453	0,0470	0,0683
Gama(15,100)	0,0543	0,2895	0,1466	0,1176

Calculando a “distância relativa” da mediana em relação ao valor nominal para as distribuições simuladas, dada pela equação 25

$$dr = \frac{M - T}{\text{Max}(X) - \text{Min}(X)} \quad (25)$$

em que X é a variável de interesse que foi medida, T é o valor nominal de especificação e M é a mediana do processo, obtêm-se os resultados mostrados na tabela 9. Com esses resultados e observando os histogramas da figura 9, tem-se que o processo representado pela distribuição gama(2,100) é o pior deles, pois este processo é o que está centrado em um valor mais distante do seu valor nominal, e que o representado pela distribuição gama(15,100) é o melhor deles, já que a mediana e o valor nominal encontram-se mais próximos que nos outros dois processos simulados.

Tabela 9 - “Distâncias relativas” da mediana em relação ao valor nominal

Distribuições	Gama(2,100)	Gama(5,100)	Gama(15,100)
“Distância relativa” (dr)	-0,3068	-0,1560	0,1233

Para comparar os índices propostos por Clements (1989) e os índices propostos por Pearn e Chen (1997), foi admitido que cada um dos três processos possui limites de especificação superior e inferior coincidentes com seus valores de máximos e de mínimos, respectivamente, e valor nominal igual a $d = \frac{\text{max} - \text{min}}{2}$. Os índices obtidos para cada uma das três distribuições são mostrados na tabela 10.

Tabela 10 - Índices de capacidade para as distribuições geradas

Distribuição	Índices de Clements (1989)				Índices Pearn e Chen (1997)			
	C'_p	C'_{pk}	C'_{pm}	C'_{pmk}	C_{Np}	C_{Npk}	C_{Npm}	C_{Npmk}
Gama(2,100)	1,194	1,013	0,494	0,203	1,194	0,456	0,494	0,153
Gama(5,100)	1,050	1,012	0,750	0,517	1,050	0,590	0,750	0,264
Gama(15,100)	1,058	1,043	0,833	0,758	1,058	0,831	0,833	0,387

De um modo geral, observa-se que todos os índices acima, com exceção do índice C'_p e seu equivalente C_{Np} , calculados para distribuições não

normais, conseguem quantificar bem a relação de capacidade dos processos. O processo representado pela distribuição gama(15,100), que é o melhor dentre os três, possui capacidade maior em todos os índices, enquanto o processo representado pela distribuição gama(2,100) possui menor capacidade dentre os processos quanto aos índices calculados.

Comparando os índices C'_{pk} e seu equivalente C_{Npk} , e os índices C'_{pmk} e seu equivalente C_{Npmk} , que são os que diferem de um método para o outro, tem-se que para todas as três distribuições, os valores dos índices C'_{pk} e de seu equivalente C_{Npk} apresentam, além das diferenças numéricas, diferenças de interpretações quanto a capacidade do processo. Enquanto C'_{pk} avalia todos os três processos como potencialmente capazes, o índice C_{Npk} informa que os processos não são potencialmente capazes com valores de índices abaixo de 1. Quanto aos índices C'_{pmk} e C_{Npmk} , também se nota valores mais altos para o primeiro índice que para o segundo nos três processos simulados.

Como os processos são assimétricos e não estão centrados (segundo a mediana) no valor nominal, tem-se que os índices propostos por Pearn e Chen (1997) conseguem quantificar melhor a capacidade do processo em relação aos índices propostos por Clements (1989), que considerou, por exemplo, o processo representado pela distribuição gama(2,100) como potencialmente capaz segundo o índice C'_{pk} quando se nota, apenas olhando para o histograma do processo, que grande parte dos itens produzidos encontram-se abaixo do valor nominal.

Para reforçar ainda mais que os índices propostos Pearn e Chen (1997) são mais eficientes na quantificação do que os índices propostos por Clements (1998), os autores Pearn e Chen (1997) fizeram um estudo comparativo entre seus índices e os propostos por Clements (1998), calculando-os para diversos valores de mediana, com $F_{99,865}$ e $F_{0,135}$ fixos e com $LIE = T - d$, onde

$$d = \frac{F_{99,865} - F_{0,135}}{2}.$$

Alguns desses valores foram reproduzidos na tabela 11, de onde

se notam que os índices de Pearn e Chen (1997) alcançam valor máximo quando a

mediana é igual ao valor nominal e que os índices de Clements (1989) alcançam valor máximo quando o processo é assimétrico, com mediana assumindo valor inferior ao valor nominal. Como o objetivo de um processo produtivo é produzir itens em torno do valor nominal, os índices de Pearn e Chen (1997) conseguem quantificar melhor a capacidade do processo.

Tabela 11 - Comparação entre os índices de Clements e os índices de Pearn e Chen

Mediana	Índices de Clements (1989)		Índices de Pearn e Chen (1997)	
	C'_{pk}	C'_{pmk}	C_{Npk}	C_{Npmk}
LIE	0,000	0,000	0,000	0,000
T-0,9d	0,200	0,036	0,100	0,035
T-0,8d	0,400	0,082	0,200	0,077
T-0,7d	0,600	0,139	0,300	0,129
T-0,6d	0,800	0,214	0,400	0,194
T-0,5d	1,000	0,316	0,500	0,277
T-0,4d	0,933	0,462	0,600	0,384
T-0,3d	0,867	0,680	0,700	0,520
T-0,2d	0,800	0,743	0,800	0,686
T-0,1d	0,733	0,719	0,900	0,862
T	0,667	0,667	1,000	1,000

3.2. Aplicação dos métodos de análise de capacidade do processo não normal

Neste estudo de caso serão aplicados os 3 métodos de análise de capacidade do processo não normal em dados coletados de uma metalúrgica da região de Campinas. Esses dados se referem à medição de dureza da etapa inicial de um processo produtivo de barras de aço fornecidas à indústria automobilística. A dureza pode ser definida como uma medida de resistência a deformações permanentes na superfície do material, servindo de controle de qualidade do tratamento térmico realizado.

Os valores são medidos por um equipamento chamado durômetro e são coletados após o processo de recozimento. O recozimento é um tipo de tratamento térmico que tem como objetivo mudar as propriedades do aço por meio de um conjunto de operações que incluem o aquecimento e o resfriamento, em condições controladas (SILVA; MEI, 2006).

A escala de medida é a brinell de dureza, com limite superior de especificação (LSE) de 248 brinell, limite inferior de especificação (LIE) de 202 brinell e valor nominal (T) de 230 brinell.

As peças com valores de dureza dentro dos limites de especificação são enviadas para a fase de acabamento do produto, de onde são liberados para o cliente; as peças abaixo do limite de especificação seguem para a etapa de acabamento, com restrições por afetar a eficiência desta etapa de acabamento; e as peças acima do limite de especificação voltam à etapa inicial do processo para que sejam retrabalhadas.

Após a etapa de acabamento é realizada uma nova medição para avaliar a conformidade das barras que devem seguir para o cliente externo. Somente a partir dos dados desta medição final que é possível concluir sobre a capacidade do processo, mas, por questão de sigilo empresarial, esses dados não foram fornecidos. Os dados trabalhados nesta dissertação se referem à medição da etapa inicial, que servem como ilustração da aplicação dos índices de capacidade, mas não sendo possível concluir sobre o processo da empresa estudada.

Para se conhecer melhor o processo produtivo das barras de aço utilizado neste estudo de caso, são mostradas na tabela 12 algumas estatísticas descritivas da medição de dureza. Os limites naturais de variação desta etapa do processo coincidem com os limites de especificação, como pode ser observado pelos valores de máximo e mínimo dos dados obtidos. Além disso, como a média é menor que a mediana, observa-se uma assimetria à esquerda (ou negativa).

Tabela 12 - Estatísticas descritivas da medição de dureza

Média	Desvio-padrão	Mínimo	Máximo	$F_{0,135}$	Mediana	$F_{99,865}$
228,71	8,30	202	248	220,27	230	245,56

Antes da aplicação dos três métodos de análise da capacidade do processo não normal apresentado no item 2.2, foram aplicados os testes de Shapiro-Wilk e Kolmogorov-Smirnov de normalidade nos dados. Por ambos os testes a suposição de normalidade dos dados é rejeitada. Para o teste de Shapiro-Wilk a estatística W obtida foi igual a 0,9564, com p-valor $=2,29 \times 10^{-12}$ e com o teste de Kolmogorov-Smirnov a rejeição da normalidade dos dados foi confirmada pelo valor da estatística $D=0,0907$, com p-valor $=9,85 \times 10^{-5}$. Visualmente pode-se observar a não normalidade dos dados por meio dos gráficos mostrados nas figuras 10(a) e 10(b).

Figura 10(a) - Histograma dos valores de dureza

Figura 10(b) - Normal-plot dos valores de dureza

Com esses resultados tornou-se necessária a aplicação dos métodos descritos no item 2.2 para a análise de capacidade deste processo não normal. Os resultados encontram-se na tabela 13, de onde se pode notar que, dependendo do índice escolhido, esta etapa do processo pode ser considerada relativamente capaz

ou incapaz, já que a maioria dos índices para dados não normais são maiores ou iguais a 1, mas alguns deles são menores que 1. Vale ressaltar em situações práticas geralmente apenas um índice é adotado para a avaliação da capacidade de um processo e dentre os métodos apresentados neste trabalho a melhor escolha seria do método 2 para a avaliação do processo, pois como foi mostrado no item 3.1 este é o melhor método dentre os métodos 1 e 2 apresentados.

Tabela 13 - Índices de capacidade do processo produtivo de barras de aço

Índices tradicionais: supondo normalidade		Método 1 Clements (1998)		Método 2 Pearn e Chen (1997)		Método 3 Chen e Ding (2001)	
C_p	0,923	C'_p	1,056	C_{Np}	1,056	S_{pmk}	0,894
C_{pk}	0,774	C'_{pk}	1,000	C_{Npk}	0,826		
C_{pm}	0,912	C'_{pm}	1,056	C_{Npm}	1,056		
C_{pmk}	0,765	C'_{pmk}	1,000	C_{Npmk}	0,643		

4 CONSIDERAÇÕES FINAIS

A análise de capacidade de processos sob suposição de normalidade foi apresentada no item 2.1 e sua extensão para processos onde a suposição de normalidade da variável de interesse não é atendida foi apresentada no item 2.2 por meio de um levantamento dos principais métodos existentes na literatura.

No início do item 2.2 foi apresentado um estudo realizado por Somerville e Montgomery (1996-1997) onde eram mostrados os erros obtidos com a suposição de normalidade para o uso dos índices de capacidade do processo quando os dados não eram normalmente distribuídos, inclusive quando a distribuição dos dados era muito próxima da normal, como era o caso da distribuição t-Student. Nas distribuições simuladas havia diferenças representativas entre o número de itens não conformes estimados pela suposição de normalidade e o número real de não conformes de cada uma das 4 distribuições para o caso unilateral.

Da comparação dos índices de capacidade não normais feitas no item 3.1 conclui-se que os índices propostos por Pearn e Chen (1997) são melhores que os propostos por Clements (1989), pois os índices C_{Npk} e C_{Npmk} consideram em seus cálculos a assimetria do processo e no estudo de Pearn e Chen (1997) alcançam valor máximo, para limites fixos, quando a mediana coincide com o valor nominal de especificação. Por outro lado, os índices de Clements (1989) apresentaram, no mesmo estudo, valor máximo quando a distribuição era nitidamente assimétrica, com mediana menor que o valor nominal.

Quanto ao processo de medição de dureza mostrado no item 3.2, cujas análises foram realizadas por meio do software R, pelos valores dos índices obtidos se concluiria que esta etapa do processo produtivo é relativamente capaz, mas somente pelos dados da segunda medição de dureza, que compõe a etapa final, é que se poderia concluir sobre a capacidade do processo produtivo estudado.

REFERÊNCIAS

- CHEN, J.P. Re-evaluating the process capability indices for non-normal distributions. **International Journal of Production Research**, London, v. 38, n. 6, p. 1311-1324, 2000.
- CHEN, J.P.; DING, C.G. A new process capability index for non-normal distributions. **The International Journal of Quality & Reliability Management**, Bradford, v. 18, n. 6/7, p. 762-770, 2001.
- CHOU, Y.M.; POLANSKY, A.M.; MASON, R.L. Transforming non-normal data to normality in statistical process control. **Journal of Quality Technology**, Milwaukee, v. 30, n. 2, p. 133-141, 1998.
- CLEMENTS, J.A. Process capability calculations for non-normal distributions. **Quality Progress**, Milwaukee, v. 22, n. 9, p. 95-100, 1989.
- MONTGOMERY, D.C. **Introduction to statistical quality control**. 3rd ed. New York: John Wiley, 1997. 734 p.
- PEARN, W.L.; CHEN, K.S. Capability indices for non-normal distributions with an application in electrolytic capacitor manufacturing. **Microelectronics and Reliability**, Kidlington, v. 37, n. 12, p. 1853-1858, 1997.
- THE R PROJECT FOR STATISTICAL COMPUTING. Disponível em: <<http://www.r-project.org>>. Acesso em: 01 nov. 2005.
- SILVA, A.L.V.C.; MEI, P.R. **Aços e ligas especiais**. 2. ed. São Paulo: Edgard Blusher, 2006. 646 p.
- SILVA, E.L.; MENEZES, E.M. **Metodologia da pesquisa e elaboração de dissertação**. 3. ed. Florianópolis: UFSC, Laboratório de Ensino a Distância, 2001. 114 p.
- SIQUEIRA, L.G.P. **Controle estatístico do processo**. São Paulo: Pioneira, 1997. 221 p.
- SOMERVILLE, S.E.; MONTGOMERY, D.C. Process capability indices and non-normal distributions. **Quality Engineering**, Monticello, v. 9, n. 2, p. 305-316, 1996-1997.
- VÄNNMAN, K. A unified approach to capability indices. **Statistica Sinica**, Beijing, v. 5, p. 805-820, 1995.
- WERKEMA, M.C.C. **Ferramentas estatísticas básicas para o gerenciamento de processos**. Belo Horizonte: Fundação Christiano Ottoni; Escola de Engenharia da UFMG, 1995. 404 p.

WERNER, L. **Controle estatístico de qualidade**. Porto Alegre: UFRGS, Instituto de Matemática, 1996. (Cadernos de Matemática e Estatística. Série B, 32).

BIBLIOGRAFIA CONSULTADA

AHMED, S.E. Assessing the process capability index for non-normal process. **Journal of Statistical Planning and Inference**, Amsterdam, v. 129, p. 195-206, 2005.

CHAN, L.K.; CHENG, S.W.; SPIRING, F.A. A new measure of process capability: C_{pm} . **Journal of Quality Technology**, Milwaukee, v. 20, n. 3, p. 162-175, 1988.

CORRÊA, J.M. **Estudo do controle e análise da capacidade de processo de produção de água potável**. 2007. 106 p. Dissertação (Mestrado em Programação Matemática) - Universidade Federal do Paraná, Curitiba, 2007.

DELERYD, M. The effect of skewness on estimates of some process capability indices. **International Journal of Applied Quality Management**, Bradford, v. 29, n. 20, p. 153-186, 1999.

FARNUM, N.R. Using Johnson curves to describe non-normal process data. **Quality Engineering**, Monticello v. 9, n. 2, p. 329-336, 1996-1997.

FEIGENBAUM, A.V. **Total quality control**. 3rd ed. New York: McGraw-Hill, 1983. 851 p.

FLAIG, J.J. A new approach to process capability analysis. **Quality Engineering**, Monticello, v. 9, n. 2, p. 205-211, 1996-1997.

GUNTER, B. The use and abuse of C_{pk} . **Quality Progress**, Milwaukee, v. 22, n. 1, p. 72-73, 1989a.

_____. The use and abuse of C_{pk} . Part 2. **Quality Progress**, Milwaukee, v. 22, n. 3, p. 108-109, 1989b.

_____. The use and abuse of C_{pk} . Part 3. **Quality Progress**, Milwaukee, v. 22, n. 5, p. 79-80, 1989c.

_____. The Use and Abuse of C_{pk} . Part 4. **Quality Progress**, Milwaukee, v. 22, n. 7, p. 86-87, 1989d.

JURAN, J.M. **Quality control handbook**. 3rd ed. New York: McGraw-Hill, 1979. 1 v.

PAN, J.N.; WU, S.L. Process capability analysis for non-normal relay test data. **Microelectronics and Reliability**, Kidlington, v. 37, n. 3, p. 421-429, 1997.

ANEXOS

ANEXO A

Comandos do software R utilizados no item 3.1

```

set.seed(20)
g1<-rgamma(1000,2,100)
g2<-rgamma(1000,5,100)
g3<-rgamma(1000,15,100)

par(mfrow=c(1,3))
hist(g1, main = paste("Gama(2,100)"))
hist(g2, main = paste("Gama(5,100)"))
hist(g3, main = paste("Gama(15,100)"))

max(g1)
max(g2)
max(g3)

min(g1)
min(g2)
min(g3)

#g1

f0135g1<-quantile(g1,0.00135)
f99865g1<-quantile(g1,0.99865)
mg1<-median(g1)

lseg1<-max(g1)
lieg1<-min(g1)
tg1<-(lseg1-lieg1)/2

cp1g1<-(lseg1-lieg1)/(f99865g1-f0135g1)
cpk1g1<-min(((lseg1-mg1)/(f99865g1-mg1)),((mg1-lieg1)/(mg1-f0135g1)))
cpm1g1<-(lseg1-lieg1)/(6*sqrt((((f99865g1-f0135g1)/6)^2)+((mg1-tg1)^2)))
cpmk1g1<-min(((lseg1-mg1)/(3*sqrt((((f99865g1-mg1)/3)^2)+((mg1-tg1)^2)))),((mg1-lieg1)/(3*sqrt((((mg1-f0135g1)/3)^2)+((mg1-tg1)^2))))))

cp2g1<-(lseg1-lieg1)/(f99865g1-f0135g1)
cpk2g1<-min(((lseg1-mg1)/((f99865g1-f0135g1)/2)),((mg1-lieg1)/((f99865g1-f0135g1)/2)))
cpm2g1<-(lseg1-lieg1)/(6*sqrt((((f99865g1-f0135g1)/6)^2)+((mg1-tg1)^2)))
cpmk2g1<-min(((lseg1-mg1)/(3*sqrt((((f99865g1-f0135g1)/6)^2)+((mg1-tg1)^2)))),((mg1-lieg1)/(3*sqrt((((f99865g1-f0135g1)/3)^2)+((mg1-tg1)^2))))))

```

#g2

```
f0135g2<-quantile(g2,0.00135)
f99865g2<-quantile(g2,0.99865)
mg2<-median(g2)
```

```
lseg2<-max(g2)
lieg2<-min(g2)
tg2<-(lseg2-lieg2)/2
```

```
cp1g2<-(lseg2-lieg2)/(f99865g2-f0135g2)
cpk1g2<-min(((lseg2-mg2)/(f99865g2-mg2)),((mg2-lieg2)/(mg2-f0135g2)))
cpm1g2<-(lseg2-lieg2)/(6*sqrt((((f99865g2-f0135g2)/6)^2)+((mg2-tg2)^2)))
cpmk1g2<-min(((lseg2-mg2)/(3*sqrt((((f99865g2-mg2)/3)^2)+((mg2-tg2)^2))))),((mg2-
lieg2)/(3*sqrt((((mg2-f0135g2)/3)^2)+((mg2-tg2)^2))))))
```

```
cp2g2<-(lseg2-lieg2)/(f99865g2-f0135g2)
cpk2g2<-min(((lseg2-mg2)/((f99865g2-f0135g2)/2)),((mg2-lieg2)/((f99865g2-
f0135g2)/2)))
cpm2g2<-(lseg2-lieg2)/(6*sqrt((((f99865g2-f0135g2)/6)^2)+((mg2-tg2)^2)))
cpmk2g2<-min(((lseg2-mg2)/(3*sqrt((((f99865g2-f0135g2)/6)^2)+((mg2-
tg2)^2))))),((mg2-lieg2)/(3*sqrt((((f99865g2-f0135g2)/3)^2)+((mg2-tg2)^2))))))
```

#g3

```
f0135g3<-quantile(g3,0.00135)
f99865g3<-quantile(g3,0.99865)
mg3<-median(g3)
```

```
lseg3<-max(g3)
lieg3<-min(g3)
tg3<-(lseg3-lieg3)/2
```

```
cp1g3<-(lseg3-lieg3)/(f99865g3-f0135g3)
cpk1g3<-min(((lseg3-mg3)/(f99865g3-mg3)),((mg3-lieg3)/(mg3-f0135g3)))
cpm1g3<-(lseg3-lieg3)/(6*sqrt((((f99865g3-f0135g3)/6)^2)+((mg3-tg3)^2)))
cpmk1g3<-min(((lseg3-mg3)/(3*sqrt((((f99865g3-mg3)/3)^2)+((mg3-tg3)^2))))),((mg3-
lieg3)/(3*sqrt((((mg3-f0135g3)/3)^2)+((mg3-tg3)^2))))))
```

```
cp2g3<-(lseg3-lieg3)/(f99865g3-f0135g3)
cpk2g3<-min(((lseg3-mg3)/((f99865g3-f0135g3)/2)),((mg3-lieg3)/((f99865g3-
f0135g3)/2)))
cpm2g3<-(lseg3-lieg3)/(6*sqrt((((f99865g3-f0135g3)/6)^2)+((mg3-tg3)^2)))
cpmk2g3<-min(((lseg3-mg3)/(3*sqrt((((f99865g3-f0135g3)/6)^2)+((mg3-
tg3)^2))))),((mg3-lieg3)/(3*sqrt((((f99865g3-f0135g3)/3)^2)+((mg3-tg3)^2))))))
```

ANEXO B

Comandos do software R utilizados no item 3.2

```
x<-scan("...endereço do banco de dados...")
```

```
f0135<-quantile(x,0.00135)
```

```
f99865<-quantile(x,0.99865)
```

```
m<-median(x)
```

```
mi<-mean(x)
```

```
s<-sd(x)
```

```
shapiro.test(x)
```

```
ks.test(x, "pnorm", mean = mean(x), sd = sd(x))
```

```
par(mfrow=c(2,2))
```

```
hist(x, xlim = range(200,250),nclass=20)
```

```
y<-rep(230,length(x))
```

```
lines(y,x)
```

```
qqnorm(x)
```

```
qqline(x)
```

```
lse<-248
```

```
lie<-202
```

```
t<-230
```

```
cp<-(lse-lie)/(6*s)
```

```
cpk<-min(((lse-mi)/(3*s)),((mi-lie)/(3*s)))
```

```
cpm<-(lse-lie)/(6*sqrt((s^2)+((mi-t)^2)))
```

```
cpmk<-min(((lse-mi)/(3*sqrt((s^2)+((mi-t)^2)))),((mi-lie)/(3*sqrt((s^2)+((mi-t)^2)))))
```

```
cp1<-(lse-lie)/(f99865-f0135)
```

```
cpk1<-min(((lse-m)/(f99865-m)),((m-lie)/(m-f0135)))
```

```
cpm1<-(lse-lie)/(6*sqrt((((f99865-f0135)/6)^2)+((m-t)^2)))
```

```
cpmk1<-min(((lse-m)/(3*sqrt((((f99865-m)/3)^2)+((m-t)^2))))),((m-lie)/(3*sqrt((((m-f0135)/3)^2)+((m-t)^2))))))
```

```
cp2<-(lse-lie)/(f99865-f0135)
```

```
cpk2<-min(((lse-m)/((f99865-f0135)/2)),((m-lie)/((f99865-f0135)/2)))
```

```
cpm2<-(lse-lie)/(6*sqrt((((f99865-f0135)/6)^2)+((m-t)^2)))
```

```
cpmk2<-min(((lse-m)/(3*sqrt((((f99865-f0135)/6)^2)+((m-t)^2))))),((m-lie)/(3*sqrt((((f99865-f0135)/3)^2)+((m-t)^2))))))
```

```
h<-hist(x, xlim = range(200,250),nclass=46)
```

```
spmk3<-(qnorm((1+(sum(h$counts)/length(x))-
```

```
(sum(h$counts[1])/length(x))/2))/(3*sqrt(1+(((mi-t)/s)^2)))
```