

UNIVERSIDADE FEDERAL DE ITAJUBÁ

Celso Francisco de Moraes

**ESTUDO DA UTILIZAÇÃO DO GRÁFICO DE
CONTROLE INDIVIDUAL E DO ÍNDICE DE
CAPABILIDADE SIGMA PARA DADOS NÃO
NORMAIS**

Dissertação submetida ao Programa de Pós-Graduação em
Engenharia de Produção como requisito parcial à obtenção
do título de *Mestre em Engenharia de Produção*

Orientador: Prof. João Roberto Ferreira, Dr.

Co-Orientador: Anderson Paulo de Paiva, MSc.

Itajubá

Abril de 2006

UNIVERSIDADE FEDERAL DE ITAJUBÁ

Celso Francisco de Moraes

**ESTUDO DA UTILIZAÇÃO DO GRÁFICO DE
CONTROLE INDIVIDUAL E DO ÍNDICE DE
CAPABILIDADE SIGMA PARA DADOS NÃO
NORMAIS**

Dissertação aprovada por banca examinadora em 26 de abril de 2006,
conferindo ao autor o título de *Mestre em Engenharia de Produção*

Banca Examinadora:

Prof. Messias Borges Silva, Dr.

Prof. João Batista Turrioni, Dr.

Anderson Paulo de Paiva, MSc. (Co-Orientador)

Prof. João Roberto Ferreira, Dr. (Orientador)

Itajubá

Abril de 2006

Dedico este trabalho às minhas filhas Amanda e Renata.

AGRADECIMENTOS

Primeiramente agradeço a Deus pelo privilégio de possuir a saúde, a motivação e os recursos materiais necessários para a continuidade desta importante caminhada educativa. À minha família, Patrícia, Amanda e Renata, pela base emocional, convívio afetivo e compreensão, fundamentais para consecução de qualquer desafio.

Particular agradecimento aos orientadores acadêmicos, Prof. Dr. João Roberto Ferreira, pela supervisão das atividades relativas à elaboração da dissertação; ao MSc. Anderson Paulo de Paiva pelas críticas e sugestões apresentadas, determinantes para o direcionamento do trabalho; e ao Prof. Dr. Pedro Paulo Balestrassi, pela assistência na definição do projeto de pesquisa.

Ao diretor do Instituto de Engenharia de Produção e Gestão da UNIFEI Prof. PhD. Luiz Gonzaga Mariano de Souza, ao coordenador do Programa de Pós-Graduação em Engenharia de Produção Prof. Dr. Carlos Eduardo Sanches da Silva e aos demais docentes do curso, em especial ao Prof. Dr. Carlos Henrique Pereira Mello, ao Prof. Dr. João Batista Turrioni e ao Prof. Dr. Renato da Silva Lima, que também contribuíram através de orientação e avaliação dos artigos produzidos a partir do desenvolvimento do tema em questão durante os seminários de dissertação.

Também devo considerações aos funcionários da secretaria, biblioteca, administração, manutenção e demais setores do campus pela importante colaboração direta ou indireta. Aos mestrandos do Programa de Pós-Graduação em Engenharia de Produção da UNIFEI, em especial aos matriculados a partir de 2004, que dividiram comigo suas experiências pessoais, profissionais e estudantis. Ao colega Juliano Dias Calderaro pela troca de informações e fornecimento dos dados reais utilizados na pesquisa.

Sou grato aos meus amigos e aos colegas de trabalho pelo apoio e incentivo demonstrados, destacadamente Celso Pereira Cobra e Ediraldo Bernardi Carvalho, que aprovaram e avalizaram esta empreitada acadêmica.

Especial e sincero agradecimento à Laura Maria Anselmo Rodrigues e Raquel Aparecida Anselmo pelo maior presente que se pode receber em toda uma vida: o acesso à educação.

“Suponho que todas as coisas que podem ser abrangidas pelo conhecimento humano estão interligadas, como na Matemática”.

René Descartes (1596-1650)

SUMÁRIO

Dedicatória.....	iii
Agradecimentos	iv
Epígrafe	v
Sumário.....	vi
Lista de Figuras	viii
Lista de Tabelas	xii
Lista de Quadros	xiii
Lista de Abreviaturas e Siglas	xiv
Lista de Símbolos	xv
Resumo	xvii
<i>Abstract</i>	xviii
1. INTRODUÇÃO.....	1
1.1. Considerações Iniciais	1
1.2. Descrição do Problema.....	3
1.3. Objetivos.....	4
1.4. Metodologia Adotada	5
1.5. Estrutura do Trabalho	7
2. FUNDAMENTAÇÃO TEÓRICA	9
2.1. Visão Conceitual da Qualidade	9
2.2. Métodos Estatísticos Aplicados à Qualidade	12
2.3. Análise e Tratamento de Dados.....	14
2.3.1. Aspectos Gerais	14
2.3.2. Distribuição de Frequências	16
2.3.3. A Distribuição Normal	19
2.4. Controle Estatístico de Processo	22
2.4.1. Definição e Conceitos Básicos	22
2.4.2. Características Gerais	24
2.4.3. Gráficos de Controle.....	26
2.4.4. Índices de Capabilidade.....	33
2.5. Metodologia Seis Sigma.....	38
2.5.1. Definição e Conceitos Básicos	38
2.5.2. Origens da Metodologia Seis Sigma	39
2.5.3. Características Gerais	40
2.5.4. Determinação do Nível Sigma.....	43
3. ESTUDO DA NÃO NORMALIDADE	48
3.1. Análise de Normalidade	48
3.2. Efeitos da Não Normalidade	50
3.3. Não Normalidade, Gráficos de Controle e Capabilidade	53
3.4. Transformação de Dados	55
3.4.1. Aspectos Gerais	55
3.4.2. Transformação de Box-Cox.....	56
3.4.3. Transformação de Johnson	62

4.	APLICAÇÃO DA METODOLOGIA	64
4.1.	Considerações Gerais	64
4.1.1.	Aspectos Metodológicos	64
4.1.2.	Caracterização dos Cenários da Pesquisa.....	65
4.2.	Investigação Experimental Através de Simulação de Dados	68
4.2.1.	Análise de Dados Modelados pela Distribuição Gamma	68
4.2.2.	Transformação de Dados Modelados pela Distribuição Gamma.....	70
4.2.3.	Gráficos de Controle Individuais – Distribuição Gamma	75
4.2.4.	Avaliação de Capabilidade – Distribuição Gamma.....	79
4.2.5.	Análise de Dados Modelados pela Distribuição Beta	84
4.2.6.	Transformação de Dados Modelados pela Distribuição Beta	86
4.2.7.	Gráficos de Controle Individuais – Distribuição Beta	88
4.2.8.	Avaliação de Capabilidade – Distribuição Beta	91
4.3.	Pesquisa Exploratória Através de Estudo de Caso	94
4.3.1.	Análise de Dados Reais – Furos 3/4”	94
4.3.2.	Transformação de Dados Reais – Furos 3/4”	99
4.3.3.	Cálculo de Capabilidade de Dados Reais – Furos 3/4”	103
4.3.4.	Análise de Dados Reais – Furos 5/8”	108
4.3.5.	Transformação de Dados Reais – Furos 5/8”	111
4.3.6.	Cálculo de Capabilidade de Dados Reais – Furos 5/8”	116
5.	DISCUSSÃO DOS RESULTADOS	122
5.1.	Análise de Resultados na Investigação Experimental	122
5.1.1.	Análise dos Gráficos de Controle – Distribuição Gamma	122
5.1.2.	Análise dos Gráficos de Controle – Distribuição Beta.....	123
5.1.3.	Considerações sobre Normalidade e Gráficos de Controle na Simulação	124
5.1.4.	Análise do Cálculo de Capabilidade – Distribuição Gamma	125
5.1.5.	Análise do Cálculo de Capabilidade – Distribuição Beta.....	126
5.1.6.	Considerações sobre Normalidade e Capabilidade na Simulação.....	127
5.2.	Análise de Resultados na Pesquisa Exploratória.....	128
5.2.1.	Análise do Cálculo de Capabilidade – Dados Reais “3/4AE”	128
5.2.2.	Análise do Cálculo de Capabilidade – Dados Reais “3/4AD”	129
5.2.3.	Análise do Cálculo de Capabilidade – Dados Reais “5/8TE”	130
5.2.4.	Análise do Cálculo de Capabilidade – Dados Reais “5/8TD”	130
5.2.5.	Considerações sobre Normalidade e Capabilidade no Estudo de Caso.....	131
6.	CONCLUSÃO.....	133
6.1.	Conclusão do Trabalho	133
6.2.	Considerações Finais e Recomendações	134
	APÊNDICE A – Dados Simulados – Distribuição Gamma.....	137
	APÊNDICE B – Dados Simulados – Distribuição Beta.....	138
	APÊNDICE C – Dados Reais – Furos 3/4”	139
	APÊNDICE D – Dados Reais – Furos 5/8”	140
	REFERÊNCIAS BIBLIOGRÁFICAS	141
	BIBLIOGRAFIA COMPLEMENTAR CONSULTADA.....	148
	ANEXO A – Tabela das Áreas sob a Curva Normal Padronizada.....	151
	ANEXO B – Fatores para Construção de Gráficos de Controle	153
	ANEXO C – Tabela de Referência entre Índice Sigma e PPM.....	154
	ANEXO D – Publicação XII SIMPEP – Bauru, SP, Brasil, Novembro de 2005	155

LISTA DE FIGURAS

Figura 1.1 – Distribuições diferentes com mesma localização e dispersão.....	3
Figura 2.1 – Representação de um processo produtivo	12
Figura 2.2 – Diagrama de utilização dos métodos para garantia da qualidade	14
Figura 2.3 – Exemplo de um diagrama de caixa (Box Plot)	15
Figura 2.4 – Exemplo de uma distribuição de frequências com histograma.....	17
Figura 2.5 – Principais aspectos da curva correspondente à distribuição Normal	19
Figura 2.6 – Curva normal reduzida ou padronizada	20
Figura 2.7 – Estratégia para melhorias em Controle Estatístico de Processo.....	26
Figura 2.8 – Exemplo típico de um gráfico de controle	27
Figura 2.9 – Fluxograma para escolha do tipo de gráfico de controle	33
Figura 2.10 – Exemplo de avaliação de capacidade de um processo com histograma.....	34
Figura 2.11 – A curva normal e os limites de especificação	40
Figura 2.12 – Deslocamento da média do processo em 1,5 sigma	41
Figura 2.13 – Esquema simplificado para cálculo do nível sigma	45
Figura 2.14 – Fluxograma geral para determinação do índice de capacidade seis sigma	47
Figura 3.1 – Limites de controle em termos de percentis para distribuições não normais.....	53
Figura 3.2 – Exemplo de análise de dados não normais.....	59
Figura 3.3 – Exemplo de um gráfico relativo à transformação de Box-Cox.....	60
Figura 3.4 – Exemplo de análise dos dados após transformação de Box-Cox.....	61
Figura 3.5 – Exemplo de um gráfico relativo à transformação de Johnson	63
Figura 4.1 – Fluxograma das atividades da pesquisa experimental	66
Figura 4.2 – Fluxograma das atividades da pesquisa exploratória	67
Figura 4.3 – Análise dos dados para “Gamma 220”.....	69
Figura 4.4 – Análise dos dados para “Gamma 220” sem <i>outliers</i>	69
Figura 4.5 – Avaliação melhor distribuição aplicável aos dados “Gamma 220” sem <i>outliers</i>	70
Figura 4.6 – Transformação de Box-Cox para os dados em “Gamma 220”	71
Figura 4.7 – Análise dos dados transformados por Box-Cox em “Gamma 220”.....	71
Figura 4.8 – Transformação de Box-Cox para os dados em “Gamma 220” sem <i>outliers</i>	72
Figura 4.9 – Análise dos dados transformados por Box-Cox em “Gamma 220” sem <i>outliers</i>	72
Figura 4.10 – Transformação de Johnson para os dados em “Gamma 220”.....	73
Figura 4.11 – Análise dos dados transformados por Johnson em “Gamma 220”	73
Figura 4.12 – Transformação de Johnson para os dados em “Gamma 220” sem <i>outliers</i>	74

Figura 4.13 – Análise dos dados transformados (Johnson) em “Gamma 220” sem <i>outliers</i> ...	74
Figura 4.14 – Gráfico de controle para “Gamma 220” considerando distribuição Normal	75
Figura 4.15 – Gráfico de controle “Gamma 220” sem <i>outliers</i> assumindo normalidade.....	76
Figura 4.16 – Gráfico de controle “Gamma 220” com limites de controle em percentis	76
Figura 4.17 – Gráfico de controle “Gamma 220” sem <i>outliers</i> com limites em percentis.....	77
Figura 4.18 – Gráfico de controle para “Gamma 220” após transformação de Box-Cox.....	77
Figura 4.19 – Gráfico de controle para “Gamma 220” sem <i>outliers</i> após Box-Cox.....	78
Figura 4.20 – Gráfico de controle para “Gamma 220” após transformação de Johnson	78
Figura 4.21 – Gráfico de controle para “Gamma 220” sem <i>outliers</i> após Johnson	79
Figura 4.22 – Índices de capacidade para “Gamma 220” considerando distribuição Normal	80
Figura 4.23 – Índices de capacidade “Gamma 220” sem <i>outliers</i> assumindo normalidade ..	80
Figura 4.24 – Índices de capacidade de “Gamma 220” considerando distribuição Gamma .	81
Figura 4.25 – Índices de capacidade de “Gamma 220” sem <i>outliers</i> (distribuição Weibull)	81
Figura 4.26 – Índices de capacidade de “Gamma 220” após transformação de Box-Cox.....	82
Figura 4.27 – Índices de capacidade de “Gamma 220” sem <i>outliers</i> após Box-Cox.....	82
Figura 4.28 – Índices de capacidade de “Gamma 220” após transformação de Johnson	83
Figura 4.29 – Índices de capacidade de “Gamma 220” sem <i>outliers</i> após Johnson	83
Figura 4.30 – Análise dos dados para “Beta 2550”	85
Figura 4.31 – Análise dos dados para “Beta 2550” sem <i>outliers</i>	85
Figura 4.32 – Transformação de Box-Cox para os dados em “Beta 2550”	86
Figura 4.33 – Análise dos dados transformados por Box-Cox em “Beta 2550”	86
Figura 4.34 – Transformação de Johnson para os dados em “Beta 2550”	87
Figura 4.35 – Análise dos dados transformados por Johnson em “Beta 2550”	87
Figura 4.36 – Gráfico de controle para “Beta 2550” considerando distribuição Normal	88
Figura 4.37 – Gráfico de controle “Beta 2550” sem <i>outliers</i> assumindo distribuição Normal	89
Figura 4.38 – Gráfico de controle “Beta 2550” com limites de controle em percentis.....	89
Figura 4.39 – Gráfico de controle para “Beta 2550” após transformação de Box-Cox	90
Figura 4.40 – Gráfico de controle para “Beta 2550” após transformação de Johnson.....	90
Figura 4.41 – Índices de capacidade de “Beta 2550” considerando distribuição Normal	91
Figura 4.42 – Índices de capacidade de “Beta 2550” sem <i>outliers</i> assumindo normalidade.	92
Figura 4.43 – Índices de capacidade de “Beta 2550” considerando distribuição Weibull	92
Figura 4.44 – Índices de capacidade de “Beta 2550” após transformação de Box-Cox.....	93
Figura 4.45 – Índices de capacidade de “Beta 2550” após transformação de Johnson	93
Figura 4.46 – Análise do conjunto de dados originais “3/4AE”	95
Figura 4.47 – Análise do conjunto de dados “3/4AE” sem <i>outliers</i>	95

Figura 4.48 – Avaliação da melhor distribuição aplicável ao conjunto de dados “3/4AE”	96
Figura 4.49 – Análise do conjunto de dados originais “3/4AD”	97
Figura 4.50 – Análise do conjunto de dados “3/4AD” sem <i>outliers</i>	97
Figura 4.51 – Avaliação da melhor distribuição aplicável ao conjunto de dados “3/4AD”	98
Figura 4.52 – Avaliação distribuição aplicável ao conjunto de dados “3/4AD” sem <i>outliers</i> ..	98
Figura 4.53 – Transformação de Box-Cox para o conjunto de dados “3/4AE”	99
Figura 4.54 – Análise do conjunto de dados “3/4AE” transformados por Box-Cox	99
Figura 4.55 – Transformação de Johnson para o conjunto de dados “3/4AE”	100
Figura 4.56 – Transformação de Box-Cox para o conjunto de dados “3/4AD”	100
Figura 4.57 – Análise do conjunto de dados “3/4AD” transformados por Box-Cox	101
Figura 4.58 – Transformação de Johnson para o conjunto de dados “3/4AD”	101
Figura 4.59 – Transformação de Box-Cox para o conjunto de dados “3/4AD” sem <i>outliers</i> ..	102
Figura 4.60 – Transformação de Johnson para o conjunto de dados “3/4AD” sem <i>outliers</i> ..	102
Figura 4.61 – Índices de capacidade do conjunto de dados originais “3/4AE”	103
Figura 4.62 – Índices de capacidade dados “3/4AE” baseados na distribuição Loglogistic ..	104
Figura 4.63 – Índices de capacidade dos dados “3/4AE” transformados por Johnson	104
Figura 4.64 – Índices de capacidade do conjunto de dados “3/4AE” sem <i>outliers</i>	105
Figura 4.65 – Índices de capacidade do conjunto de dados originais “3/4AD”	105
Figura 4.66 – Índices de capacidade dados “3/4AD” baseados na distribuição Logistic	106
Figura 4.67 – Índices de capacidade do conjunto de dados “3/4AD” sem <i>outliers</i>	106
Figura 4.68 – Índices capacidade dados “3/4AD” sem <i>outliers</i> (distribuição Lognormal) ..	107
Figura 4.69 – Análise do conjunto de dados originais “5/8TE”	108
Figura 4.70 – Avaliação da melhor distribuição aplicável ao conjunto de dados “5/8TE”	109
Figura 4.71 – Análise do conjunto de dados originais “5/8TD”	109
Figura 4.72 – Análise do conjunto de dados “5/8TD” sem <i>outliers</i>	110
Figura 4.73 – Avaliação da melhor distribuição aplicável ao conjunto de dados “5/8TD” ...	110
Figura 4.74 – Avaliação de distribuição aplicável ao conjunto dados “5/8TD” sem <i>outliers</i> ..	111
Figura 4.75 – Transformação de Box-Cox para o conjunto de dados “5/8TE”	112
Figura 4.76 – Análise do conjunto de dados “5/8TE” transformados por Box-Cox	112
Figura 4.77 – Transformação de Johnson para o conjunto de dados “5/8TE”	113
Figura 4.78 – Transformação de Box-Cox para o conjunto de dados “5/8TD”	113
Figura 4.79 – Análise do conjunto de dados “5/8TD” transformados por Box-Cox	114
Figura 4.80 – Transformação de Johnson para o conjunto de dados “5/8TD”	114
Figura 4.81 – Transformação de Box-Cox para o conjunto de dados “5/8TD” sem <i>outliers</i> ..	115
Figura 4.82 – Transformação de Johnson para o conjunto de dados “5/8TD” sem <i>outliers</i> ..	115

Figura 4.83 – Índices de capacidade do conjunto de dados originais “5/8TE”.....	116
Figura 4.84 – Índices de capacidade dados “5/8TE” baseados na distribuição Lognormal	117
Figura 4.85 – Índices de capacidade dos dados “5/8TE” transformados por Johnson	117
Figura 4.86 – Índices de capacidade dos dados originais “5/8TD”	118
Figura 4.87 – Índices de capacidade dados “5/8TD” baseados na distribuição Lognormal	118
Figura 4.88 – Índices de capacidade dos dados “5/8TD” transformados por Johnson.....	119
Figura 4.89 – Índices de capacidade do conjunto de dados “5/8TD” sem <i>outliers</i>	119
Figura 4.90 – Índices de capacidade dados “5/8TD” sem <i>outliers</i> (distribuição Lognormal)..	120
Figura 4.91 – Índices de capacidade dos dados “5/8TD” sem <i>outliers</i> (transf. Johnson)....	120

LISTA DE TABELAS

Tabela 2.1 – Referências para o nível sigma	38
Tabela 2.2 – Referências entre nível sigma e PPM (com e sem desvio).....	41
Tabela 3.1 – Comparação da proporção de itens não conformes em distribuições diferentes .	52
Tabela 3.2 – Exemplos típicos da transformação de Box-Cox.....	58
Tabela 3.3 – Exemplo de um conjunto de dados não normais	59
Tabela 3.4 – Fórmulas associadas com as famílias da transformação de Johnson.....	62
Tabela 4.1 – Cálculos comparativos do nível sigma para “Gamma 220”	84
Tabela 4.2 – Cálculos comparativos do nível sigma para “Beta 2550”	94
Tabela 4.3 – Cálculos comparativos do nível sigma para “3/4AE” e “3/4AD”	107
Tabela 4.4 – Cálculos comparativos do nível sigma para “5/8TE” e “5/8TD”	121
Tabela 5.1 – Interpretação dos gráficos de controle individuais - dados simulados	124
Tabela 5.2 – Determinação do nível sigma - dados simulados	127
Tabela 5.3 – Determinação do nível sigma - dados reais	132

LISTA DE QUADROS

Quadro 1.1 – Conseqüências da suposição incorreta de normalidade.....	4
Quadro 1.2 – Resumo da metodologia adotada.....	6
Quadro 2.1 – Evolução da Inspeção à Garantia da Qualidade	11
Quadro 2.2 – Campos de aplicação de algumas distribuições estatísticas	18
Quadro 2.3 – Fórmulas dos tipos mais comuns de gráficos de controle para variáveis.....	30
Quadro 2.4 – Fórmulas dos tipos de gráficos de controle para atributos	31
Quadro 2.5 – Classificação de processos a partir do índice C_p	37
Quadro 2.6 – Visão geral do DMAIC.....	42
Quadro 2.7 – Papéis desempenhados na Metodologia Seis Sigma	43
Quadro 2.8 – Definição de conceitos para atributos na Metodologia Seis Sigma.....	46
Quadro 4.1 – Resumo dos aspectos gerais da pesquisa.....	64

LISTA DE ABREVIATURAS E SIGLAS

ANOVA – *Analysis of Variance*

ASQC – *American Society for Quality Control*

BB – *Black Belt*

CEP – Controle Estatístico de Processo

CI – Confidence Interval

CTQ – *Critical to Quality*

DMAIC – *Define, Measure, Analyze, Improve, Control*

DOE – *Design of Experiment*

DPMO – Defeitos por Milhão de Oportunidades

DPO – Defeitos por Oportunidade

DPU – Defeitos por Unidade

GB – *Green Belt*

IQR – intervalo interquartil

ISO – *International Organization for Standardization*

LB – Lower Bound

LCL – Lower Control Limit

LIC – Limite Inferior de Controle

LIE – Limite Inferior de Especificação

LM – Linha Média

Lower CL – Lower Confidence Limit

LSC – Limite Superior de Controle

LSE – Limite Superior de Especificação

LSXY – Least Squares Method

MBB – *Master Black Belt*

PDCA – *Plan, Do, Check, Act*

PPM – Partes por Milhão

psi – *pound square inch*

StDev – Standard Deviation

UB – Upper Bound

UCL – Upper Control Limit

Upper CL – Upper Confidence Limit

LISTA DE SÍMBOLOS

- A_2 – constante para definição dos limites de controle das médias no gráfico $\bar{X}-R$
- A_3 – constante para definição dos limites de controle das médias no gráficos $\bar{X}-S$
- B_3 – constante para definição do limite de controle inferior do desvio-padrão no gráfico $\bar{X}-S$
- B_4 – constante para definição do limite de controle superior do desvio-padrão no gráfico $\bar{X}-S$
- c – número total de defeitos em todas as unidades da amostra em estudo
- c_4 – constante para definição da linha média no gráfico de controle para desvios-padrão
- C_P – índice de capacidade potencial do processo
- C_{PK} – índice de capacidade real do processo
- d – número de peças defeituosas da amostra em estudo
- d_2 – constante para definição da linha média no gráfico de controle para amplitudes
- D_3 – constante para definição do limite de controle inferior da amplitude no gráfico $\bar{X}-R$
- D_4 – constante para definição do limite de controle superior da amplitude no gráfico $\bar{X}-R$
- e – constante equivalente ao número irracional 2,7182...
- E_2 – constante para definição dos limites de controle no gráfico de valores individuais
- $E(y)$ – valor esperado de uma variável dependente
- $f(x, \theta)$ – função resposta da variável dependente
- H_0 – hipótese nula
- H_1 – hipótese alternativa
- m – número de amostras em estudo
- n – tamanho da amostra em estudo
- P_P – índice de desempenho potencial do processo
- P_{PK} – índice de desempenho real do processo
- Q_1 – 1º quartil do conjunto de dados em estudo
- Q_3 – 3º quartil do conjunto de dados em estudo
- R – amplitude da amostra em estudo
- \bar{R} – média das amplitudes das amostras em estudo
- S – desvio-padrão da amostra em estudo
- \bar{S} – média dos desvios-padrão das amostras em estudo
- S_B – família *bounded* (limitada) da transformação de Johnson
- S_L – família Lognormal da transformação de Johnson
- S_U – família *unbounded* (ilimitada) da transformação de Johnson

x – variável aleatória em estudo

\bar{X} – média da amostra em estudo

y_u – variável dependente em estudo

$y^{(\lambda)}$ – variável dependente após a transformação de Box-Cox

z – variável aleatória normal padrão ou escore padronizado

z_{bench} – escore padronizado (referência no cálculo do índice de capacidade seis sigma)

z_{LT} – escore padronizado de longo prazo (*long term*)

z_{shift} – desvio da média do processo ao longo do tempo

z_{ST} – escore padronizado de curto prazo (*short term*)

α – probabilidade de cometer erro tipo I em um teste de hipóteses

γ – parâmetro que define a transformação de Johnson

ϵ_u – erros observados na função resposta da variável dependente

η – parâmetro que define a transformação de Johnson

θ – conjunto de parâmetros que afetam a função resposta da variável dependente

λ – parâmetro que define a transformação de Box-Cox

μ – média da população em estudo

$\hat{\mu}$ – estimativa da média do processo em estudo

π – constante equivalente ao número irracional 3,1415...

σ – desvio-padrão da população em estudo

$\hat{\sigma}$ – estimativa do desvio-padrão do processo em estudo

σ^2 – variância da população em estudo

$\Phi(z)$ – função de distribuição cumulativa de uma variável aleatória normal padrão

$\Psi^{(\lambda, x)}$ – função de transformação para a condição de normalidade

RESUMO

Esta dissertação apresenta uma análise crítica da utilização de alguns métodos estatísticos aplicados à Qualidade nos casos em que os dados coletados não se apresentam normalmente distribuídos. O problema geral aqui identificado diz respeito à ocorrência de decisões questionáveis tomadas a partir da interpretação de dados considerados normais. A justificativa primordial para este estudo reside na relevância da análise dos dados durante sua coleta e tratamento na prática industrial e em trabalhos científicos nos mais variados níveis. Os principais objetivos deste trabalho são: apresentar algumas situações de utilização de gráficos de controle e índices de capacidade seis sigma que possam induzir a conclusões duvidosas devido à adoção incorreta de normalidade; propor revisão crítica efetuando as transformações de Box-Cox e/ou Johnson; e, finalmente, discutir as conclusões e decisões estabelecidas através da comparação entre os dados brutos originalmente coletados, os dados previamente analisados e processados e os dados transformados. A abordagem metodológica combina pesquisa experimental através da análise de dados simulados gerados a partir de um *software* estatístico bem como pesquisa exploratória na forma de estudo de caso em um processo industrial de medição de furos de precisão.

Palavras chave: Dados não normais, Gráfico de controle individual, Índice de capacidade sigma, Transformação de dados.

ABSTRACT

This dissertation presents a critical analysis regarding the use of some statistical methods applied to Quality in cases in which the collected data are not normally distributed. The general problem refers to the occurrence of questionable decision takings starting from the interpretation of data considered normally distributed. The primordial justification for this study is the relevance of the data analysis during its collection and handling in the industrial practice and in scientific works in the most varied levels. The main objectives of this work are: to present some examples related to the usage of control charts and six sigma capability index that may induce to doubtful conclusions due to the incorrect assumption of normality; to propose critical review proceeding to the Box-Cox and/or Johnson transformations; and, finally, to discuss the conclusions and decisions established through the comparison among the gross data originally collected, the data previously analyzed and processed and the transformed data. The methodological approach combines experimental research through the analysis of simulate data generated starting from a statistical software as well as exploratory research with case study in an industrial process of precision hole measurement.

Key Words: Individual control chart, Data transformation, Non-normal data, Sigma capability index.

1. INTRODUÇÃO

1.1 Considerações Iniciais

Em muitas oportunidades a realização de uma pesquisa de cunho científico contribui para a descoberta de novos conhecimentos, para o desenvolvimento de novas teorias ou para a confirmação experimental da aplicação de tecnologias inovadoras; em outras situações pode proporcionar o estabelecimento de fronteiras entre áreas de conhecimento afins, motivando debates que podem introduzir desafios ou questões para as quais nem sempre há respostas imediatas; também é possível que determinadas investigações científicas complementem ou até mesmo superem teorias e doutrinas anteriormente consagradas.

A proposta desta dissertação em analisar a aplicação de alguns métodos estatísticos está inserida em um contexto amplo que se refere ao Controle Estatístico da Qualidade e aos aspectos básicos que devem ser observados pelas organizações que buscam a plena satisfação de seus clientes. O tema da pesquisa diz respeito à análise crítica da utilização de conceitos e técnicas relativas ao Controle Estatístico de Processo (CEP) e à Metodologia Seis Sigma, especificamente nos casos em que os dados originalmente coletados não apresentam uma forma de distribuição que seja modelada pela distribuição Normal. Esta investigação tem importância no contexto acadêmico e principalmente na prática dos processos operacionais dos mais variados tipos de organizações, e se justifica, pois é de extrema relevância a correta análise dos dados durante sua coleta e tratamento.

A aplicação adequada de técnicas estatísticas em atividades industriais e empresariais contribui de maneira decisiva para a otimização dos processos produtivos, com conseqüente redução de custos e melhoria da qualidade. Segundo Liberatore (2001), no âmbito industrial a pesquisa estatística visa fornecer respostas para algumas questões importantes, como as que se seguem:

- a) Existe relação entre a variável “A” e a variável “B” em processos de fabricação ou de prestação de serviços?
- b) Qual é o melhor ajuste das máquinas e dispositivos empregados no processo?
- c) Qual dentre os diversos métodos de processamento disponíveis é superior?
- d) Que parte da variação observada nos processos pode ser atribuída à variação inerente ao sistema de medição?
- e) Os programas de treinamento e desenvolvimento atingem o resultado esperado?

Diversos autores destacam a crescente utilização de técnicas estatísticas para melhoria da qualidade em produtos e serviços, entretanto nem sempre os cuidados necessários para a implementação de tais técnicas são adequadamente observados. Como exemplo, Balestrassi (2000) cita um trabalho de Alwan & Roberts no qual foi avaliada a utilização dos gráficos de controle em uma amostragem de 235 (duzentas e trinta e cinco) situações reais e, surpreendentemente, identificou-se violação a conceitos estatísticos básicos em 86% dos casos. Desconhecimento de conceitos estatísticos básicos pode conduzir a suposições incorretas, como, por exemplo, acreditar que os dados coletados em qualquer processo observado na prática sempre sejam perfeitamente representados pela distribuição Normal.

O cálculo de qualquer medida estatística de desempenho requer uma suposição acerca da forma da distribuição de probabilidades da característica da qualidade analisada. Woodall (2000) descreve que muitos estudos teóricos, bem como simulações em Controle Estatístico de Processo, baseiam-se na premissa de normalidade da distribuição e independência das amostras ao longo do tempo. Considerando tal afirmação, é recomendável cautela na aplicação e na análise de métodos estatísticos, pois em muitas situações reais a suposição de normalidade não se confirma na prática, o que implica em especial atenção em relação aos dados coletados.

Recentemente as aplicações estatísticas com base nos programas de melhoria Seis Sigma têm sido muito difundidas nos mais variados tipos de organizações e, de acordo com Hoerl (2001), embora esta metodologia demande a utilização de ferramentas estatísticas, a mesma não deve ser encarada simplesmente como uma coletânea de dispositivos e ferramentas. Esta distinção advém da constatação de que um dos papéis fundamentais desempenhados pelos principais personagens envolvidos na implementação da metodologia é, justamente, a garantia da correta utilização das técnicas estatísticas que, por não serem novas, merecem uma abordagem diferenciada.

No desenvolvimento deste trabalho pretende-se analisar a utilização de gráficos de controle individuais em CEP e a avaliação de capacidade na Metodologia Seis Sigma, com enfoque na correta interpretação dos dados coletados, especificamente no que se refere à suposição de normalidade, com revisão crítica através de transformações matemáticas de variáveis. As conclusões e decisões estabelecidas devem ser discutidas por meio de comparação entre os dados brutos originalmente coletados, os dados previamente tratados e os dados transformados para uma nova condição modelada pela distribuição Normal.

1.2 Descrição do Problema

Segundo Bayeux (2001) uma condição relevante na análise estatística de processos é que esses tenham comportamento modelado pela distribuição Normal, ou pelo menos que a distribuição, embora não normal, seja conhecida. Na verdade, em boa parte dos casos práticos, a condição de normalidade absoluta e estável é um evento raro sendo comum algumas situações de não-normalidade, tais como assimetria unilateral, distribuição com ramos mais densos, uniformidade na frequência dos valores observados, entre outras.

A Figura 1.1 apresenta uma comparação entre a distribuição Normal e outros tipos de distribuições observadas em situações reais. Nota-se que, embora as distribuições mostradas possuam a mesma localização e mesma dispersão dos dados em relação à distribuição Normal, existe uma sensível diferença entre as áreas delimitadas pelos limites de especificação inferior e superior. Tal fato altera, por exemplo, os cálculos relativos à taxa esperada de produtos não-conformes na saída do processo e conseqüentemente os índices de capacidade ou capacidade, além de outros indicadores úteis em análise estatística.

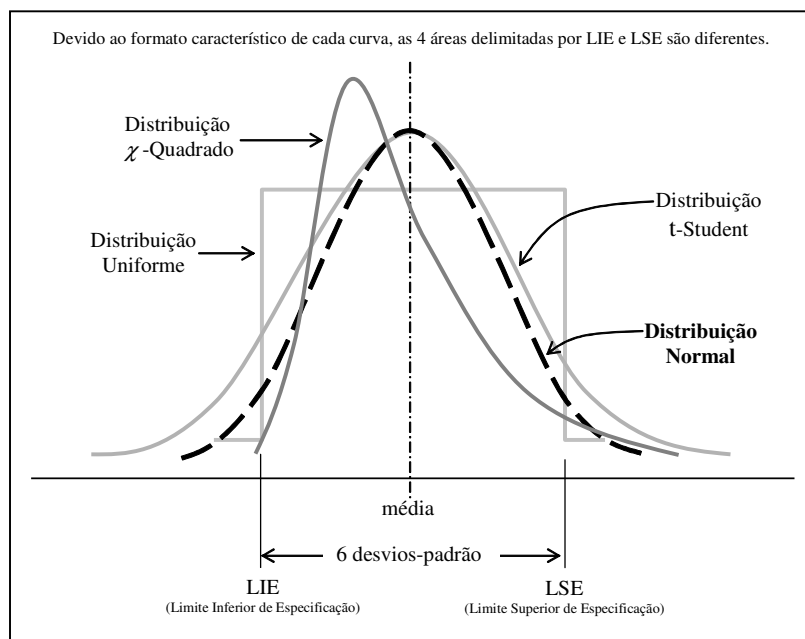


Figura 1.1 – Distribuições diferentes com mesma localização e dispersão

Fonte: Adaptado de Bayeux (2001)

O problema que motivou a pesquisa referente a esta dissertação é a ocorrência de decisões questionáveis tomadas a partir da análise de dados considerados normalmente distribuídos sem a devida confirmação de tal suposição. O Quadro 1.1 mostra um resumo das

conseqüências na aplicação de alguns métodos estatísticos quando se assume normalidade para dados que na realidade não apresentam distribuição Normal.

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

Quadro 1.1 – Conseqüências da suposição incorreta de normalidade
Fonte: Adaptado de Rath & Strong Management Consultants (2000)

As principais questões a serem investigadas, com base na definição do problema e compreensão de suas conseqüências, são as seguintes:

- a) Identificação de falsas causas especiais de variação na análise dos gráficos de controle individuais em CEP;
- b) Cálculo incorreto do nível sigma na avaliação de capacidade na Metodologia Seis Sigma.

1.3 Objetivos

O objetivo geral desta pesquisa é efetuar um aprofundamento dos aspectos relativos aos efeitos da não normalidade de dados na aplicação de métodos estatísticos, bem como das técnicas de transformação matemática de variáveis e sua comparação com outros métodos, de modo a destacar a relevância da correta análise de dados.

De forma complementar e garantindo que objetivo proposto seja alcançado, os seguintes objetivos específicos foram estabelecidos:

- a) Discutir a utilização de gráficos de controle individuais em alguns exemplos simulados em situações que possam induzir a conclusões duvidosas devido à adoção incorreta de normalidade dos dados;
- b) Avaliar a determinação do índice de capacidade sigma em alguns exemplos reais e simulados em situações que possam induzir a conclusões duvidosas devido à adoção incorreta de normalidade dos dados.

1.4 Metodologia Adotada

A classificação dos tipos de pesquisa científica apresenta diversos critérios que variam de acordo com o contexto geral do estudo e com o enfoque adotado. Esta diversidade está relacionada, por exemplo, aos objetivos, objetos de estudo, procedimentos e condições de pesquisa. Por este motivo, antes de caracterizar a metodologia adotada, é interessante apresentar algumas definições em relação aos tipos e métodos de pesquisa para um melhor entendimento do cenário que originou o planejamento desta dissertação.

Severino (2002) destaca que os métodos e as técnicas empíricas de pesquisa possibilitam a existência de diversas formas de investigação científica, tais como, pesquisa experimental, pesquisa bibliográfica, pesquisa de campo, pesquisa documental, pesquisa histórica, pesquisa fenomenológica, pesquisa clínica, pesquisa lingüística, entre outras.

No que tange aos objetivos, Gil (1991) classifica os tipos de pesquisa científica da seguinte maneira:

- a) Pesquisas exploratórias – proporcionam maior familiaridade com determinado tema e, quanto aos procedimentos técnicos utilizados, podem ser divididas em pesquisa bibliográfica e estudo de caso;
- b) Pesquisas descritivas – possibilitam descrição detalhada de características de grupos e, em termos de procedimentos técnicos, assumem a forma de levantamento através de censo ou amostragem;
- c) Pesquisas explicativas – identificam os fatores que causam ou contribuem para ocorrência de determinados fenômenos e, quanto aos procedimentos técnicos utilizados, caracterizam-se como pesquisa experimental.

Em termos de metodologias que utilizam simulação e modelagem, Bertrand & Fransoo (2002) classificam as pesquisas de Gestão de Operações em duas categorias:

- a) Pesquisa axiomática – tipicamente normativa, a investigação é orientada por modelos idealizados, onde o objetivo primário do pesquisador é obter soluções no modelo definido e garantir que tais soluções forneçam as descobertas na estrutura do problema;
- b) Pesquisa empírica – podendo ser normativa ou descritiva, este tipo de pesquisa é orientada por resultados e medições empíricas e está interessada na criação de modelos que adequadamente descrevam as relações causais existentes em processos reais e que conduzam a um melhor entendimento destes processos.

Segundo Severino (2002), diretamente relacionado com o tipo de pesquisa estão os métodos e as técnicas a serem adotados. Entende-se por métodos os procedimentos mais amplos de raciocínio, enquanto técnicas são procedimentos mais restritos que operacionalizam os métodos, mediante emprego de instrumentos adequados.

Com base nas referências citadas, a presente pesquisa pode ser caracterizada como uma combinação de dois métodos distintos, definidos como pesquisa experimental e pesquisa exploratória, associados a duas técnicas diferentes, respectivamente simulação e estudo de caso. O Quadro 1.2 apresenta um resumo da categorização e da proposta deste trabalho.

MÉTODO	TÉCNICA	PROPOSTA	COLETA DE DADOS
Pesquisa Experimental	Simulação	Investigar interpretação de dados não normais simulados em gráficos de controle individuais e na determinação do nível sigma	Dados gerados através do <i>software</i> de aplicações estatísticas Minitab®
Pesquisa Exploratória	Estudo de Caso	Investigar interpretação de dados reais não normais na determinação do índice de capacidade e nível sigma do processo	Dados obtidos em um processo industrial de furação de precisão

Quadro 1.2 – Resumo da metodologia adotada

Neste trabalho a investigação experimental consiste em efetuar análise crítica da interpretação de gráficos de controle individuais e da avaliação de capacidade através da determinação do nível sigma a partir de dados não normais gerados com auxílio do *software* estatístico Minitab®. Dois conjuntos de dados simulados, propositalmente não normais, foram analisados quanto ao seu comportamento em relação à suposição de normalidade em aplicações de interpretação de gráficos de controle individuais e cálculo do índice de capacidade seis sigma.

Através da pesquisa exploratória foi efetuado um aprofundamento dos aspectos relativos aos efeitos da não normalidade de dados no cálculo de capacidade e determinação do nível sigma de qualidade, além da discussão acerca da aplicabilidade da estratégia de transformação matemática de variáveis pelo método de Box-Cox e de Johnson. Para consecução desta tarefa foi adotada a técnica de estudo de caso em um processo industrial de furação de precisão em uma empresa do interior do estado de São Paulo.

1.5 Estrutura do Trabalho

Esta dissertação é o resultado final de um projeto de pesquisa relativo à investigação da aplicação de métodos estatísticos em processos ou estudos científicos nos quais não existe a confirmação da premissa de normalidade dos dados. Em uma abordagem inicial, pretendia-se realizar uma pesquisa que possibilitasse tal investigação em diversos métodos estatísticos sensíveis à suposição de normalidade. Posteriormente, quando da delimitação do tema, optou-se pela definição de uma análise restrita a apenas dois métodos bem conhecidos e difundidos nas esferas acadêmica e industrial, os gráficos de controle e a avaliação de capacidade.

Neste capítulo introdutório da dissertação já foram apresentados, pela ordem, a relevância do tema no contexto da Engenharia de Produção, a descrição do problema e as questões a serem investigadas, a determinação dos objetivos e a definição da metodologia de pesquisa adotada. Os comentários a seguir visam promover um entendimento mais abrangente da estrutura pela qual o trabalho foi organizado e orientar a sua leitura, interpretação e avaliação crítica.

O capítulo 2 trata da fundamentação teórica, que reflete a revisão da literatura iniciada logo após a definição do tema e que se estendeu por quase todo período de pesquisa. Nos tópicos iniciais são abordados vários aspectos conceituais, tais como a importância da qualidade nos contextos industrial e empresarial, além do valor da utilização de métodos e técnicas estatísticas para melhoria da qualidade de produtos e serviços. Também são apresentados resumos sobre a Análise de Dados, com ênfase nos principais aspectos da distribuição Normal, o Controle Estatístico de Processo, com destaque para os gráficos de controle e índices de capacidade, e a Metodologia Seis Sigma, ressaltando as fórmulas para o cálculo do índice de capacidade através do nível sigma de qualidade.

Para a revisão bibliográfica relativa ao tema central da pesquisa, que aborda as considerações gerais acerca de não normalidade e aplicação de métodos estatísticos, foi reservada uma parte específica na dissertação: o capítulo 3. Primeiramente são apresentadas algumas definições necessárias ao entendimento da análise de normalidade das amostras em estudo, tais como, amostragem, intervalos de confiança e teste de hipóteses. Em seguida, são comentados alguns efeitos da condição de não normalidade em estudos estatísticos e algumas técnicas utilizadas para contornar esta situação, como, por exemplo, as transformações de variáveis, destacadamente as transformações de Box-Cox e Johnson.

No capítulo 4, que discorre sobre o desenvolvimento da pesquisa, é apresentada inicialmente uma visão geral sobre os aspectos metodológicos e os dois tipos de pesquisa

utilizados de acordo com os cenários definidos. Em um primeiro cenário, os dados simulados utilizados na investigação experimental são analisados, tratados e discutidos em relação à interpretação dos gráficos de controle e ao cálculo do nível sigma. No segundo cenário referente à pesquisa exploratória, os dados reais obtidos em um processo industrial também são analisados, tratados e discutidos em relação ao cálculo do índice sigma de capacidade. Os cálculos e análises gráficas foram efetuados por meio do *software* estatístico Minitab®.

A discussão detalhada dos resultados obtidos no desenvolvimento deste estudo é apresentada no capítulo 5, que foi dividido em dois tópicos principais: análise dos resultados decorrentes da investigação experimental e análise dos resultados obtidos por meio da pesquisa exploratória. O capítulo 6, que é de conclusão, encerra a argumentação apresentada neste trabalho através das considerações finais advindas da análise dos resultados obtidos nos cenários estabelecidos em comparação com as proposições teóricas referenciadas. Também são indicadas algumas sugestões para pesquisas futuras, provenientes, tanto das limitações encontradas durante o desenvolvimento do trabalho, quanto das possibilidades, inicialmente não vislumbradas, que emergiram pela revisão da literatura e aplicação da metodologia.

Nos Apêndices A e B estão disponíveis para consulta, os conjuntos de dados simulados utilizados durante a pesquisa experimental e nos Apêndices C e D podem ser visualizados os dados reais da pesquisa exploratória, complementando assim a argumentação apresentada ao longo dos capítulos. Logo após as Referências Bibliográficas há uma lista com a Bibliografia Complementar Consultada que, mesmo sem citação direta, influenciou o entendimento do tema, a condução da pesquisa e a realização deste trabalho.

E por último, ilustrando os diversos aspectos abordados, os Anexos que incluem, na seqüência, a tabela com as diversas áreas sob a curva correspondente à distribuição normal padronizada (Anexo A), os fatores utilizados para cálculo dos limites nos gráficos de controle (Anexo B) e uma tabela com alguns valores de equivalência entre nível sigma e proporção de defeitos expressa em partes por milhão (Anexo C). O Anexo D é um artigo elaborado durante o desenvolvimento desta pesquisa, que foi submetido, aprovado e apresentado no XII Simpósio de Engenharia de Produção (SIMPEP) realizado em novembro de 2005 na Universidade Estadual Paulista (UNESP) em Bauru. Este mesmo artigo foi selecionado para compor o próximo número da revista GEPROS (Gestão da Produção, Operações e Sistemas) editada pelo Departamento de Engenharia de Produção da Faculdade de Engenharia de Bauru.

2. FUNDAMENTAÇÃO TEÓRICA

2.1 Visão Conceitual da Qualidade

Constantemente utilizado de maneira informal, o termo qualidade provoca contradições em relação a sua correta conceituação, pois os aspectos técnicos associados podem ser os mais diversos possíveis, desde características esperadas em processos de produção ou de serviços até situações que dizem respeito às mais complexas estratégias empresariais. Dos conceitos propostos por diversos especialistas para o termo qualidade, Prazeres (1997) destaca os seguintes:

- “Qualidade é a satisfação do cliente” e “melhoria contínua” (William E. Deming).
- “O nível de satisfação alcançado por um determinado produto no atendimento aos objetivos do usuário, durante seu uso, é chamado de adequação ao uso. Este conceito de adequação ao uso, popularmente conhecido por alguns nomes, tal como qualidade, é um conceito universal aplicável a qualquer tipo de bem ou serviço” (Joseph M. Juran).
- “Qualidade é a composição total das características de marketing, engenharia, fabricação e manutenção de um produto ou serviço, através das quais o mesmo produto ou serviço, em uso, atenderá as expectativas do cliente” (Armand Feigenbaum).
- “Qualidade é conformidade com os requisitos” (Philip B. Crosby).
- “Qualidade é a totalidade de requisitos e características de um produto ou serviço que estabelece a sua capacidade de satisfazer determinadas necessidades” (ASQC – *American Society for Quality Control*).
- “Qualidade é a totalidade de características de uma entidade que lhe confere a capacidade de satisfazer as necessidades explícitas e implícitas” (ISO – *International Organization for Standardization*).

De acordo com Garvin (1984) o termo qualidade possui um aspecto dinâmico e seu conceito pode sofrer alterações que conduzem a cinco abordagens gerais de definição:

- a) Abordagem transcendental – considera a qualidade como sinônimo de excelência inata, uma característica, propriedade ou estado que torna um produto ou serviço aceitável plenamente, embora tal aceitação seja derivada não de análise e estudos efetuados, mas da constatação prática, proveniente, em muitas das vezes, da experiência;
- b) Abordagem baseada no produto – entende a qualidade como um conjunto mensurável e preciso de características requeridas para satisfazer o consumidor;

- c) Abordagem baseada em valor – define a qualidade em termos de custo e preço, considerando que produto ou serviço é de boa qualidade quando apresenta o melhor desempenho a um preço aceitável;
- d) Abordagem baseada em manufatura – considera produtos ou serviços que estejam livres de erros, relacionando-se ao esforço produtivo no sentido de produzir o item em conformidade com as suas especificações básicas.
- e) Abordagem baseada no usuário – assegura que o produto ou serviço está adequado a seu propósito, sendo o usuário a fonte da avaliação sobre a qualidade, ou seja, não se pode pensar em qualidade sem antes identificar o desejo do consumidor.

Na visão de Campos (1992) um produto ou serviço de qualidade é aquele que atende perfeitamente, de forma confiável, de forma acessível, de forma segura e no tempo certo às necessidades do cliente. Segundo Montgomery (1985) há dois aspectos gerais de qualidade: qualidade de projeto e qualidade de conformidade. Todos os bens e serviços são produzidos em vários graus ou níveis de qualidade e quando estas variações são intencionais o termo técnico apropriado é qualidade de projeto. A qualidade de conformidade é o nível de adequação do produto em relação às especificações e tolerâncias requeridas pelo projeto.

Slack *et al.* (1999) preconizam que o sucesso das organizações depende de sua vantagem competitiva baseada em produção, que está associada a cinco objetivos de desempenho: qualidade, rapidez, confiabilidade, flexibilidade e custo. A qualidade é um objetivo de desempenho particularmente importante, pois afeta diretamente consumidores internos e externos, além de possibilitar redução de custos, aumento de confiabilidade e conseqüentemente a satisfação dos clientes. De acordo com Paladini (1990) esta importância atribuída à qualidade é decorrente de sua profunda ligação com o objetivo básico de qualquer empresa que, em um contexto mais amplo, é sobreviver de forma a manter sua faixa de atuação no mercado.

Os diversos aspectos conceituais e as várias abordagens do termo qualidade, bem como a enorme quantidade de estudos científicos associados ao tema, evidenciam a sua importância e relevância nos mais variados cenários, sendo classificada como uma disciplina associada à Engenharia de Produção, merecendo assim uma denominação específica de campo de estudo, neste caso grafada pelo substantivo próprio Qualidade.

Um breve histórico da evolução da Qualidade nas organizações, baseado em Slack *et al.* (1999), destaca que originalmente a qualidade dos produtos era atingida por inspeção, com separação dos defeitos antes de serem percebidos pelos consumidores. Em seguida o conceito

de Controle da Qualidade desenvolveu uma abordagem mais sistemática, não apenas para detectar, mas também para tratar dos problemas de qualidade. A Garantia da Qualidade ampliou a responsabilidade da qualidade ao incluir outras funções, além das operações diretas e introduzindo o uso de técnicas estatísticas mais sofisticadas. O Quadro 2.1 sintetiza alguns aspectos da evolução da Inspeção até a Garantia da Qualidade.

IDENTIFICAÇÃO DE CARACTERÍSTICAS	ETAPA DO MOVIMENTO DA QUALIDADE		
	Inspeção	Controle Estatístico da Qualidade	Garantia da Qualidade
Preocupação básica	Verificação	Controle	Coordenação
Visão da Qualidade	Um problema a ser resolvido	Um problema a ser resolvido	Um problema a ser resolvido, mas enfrentado proativamente
Ênfase	Uniformidade dos produtos	Uniformidade dos produtos com menos inspeção	Contribuição de todos os grupos funcionais para prevenir falhas da qualidade
Métodos	Aparelhos de medida e mensuração	Instrumentos e técnicas estatísticas	Programas e sistemas
Papel dos profissionais da Qualidade	Inspeção, classificação, contagem e avaliação	Solução de problemas e aplicação de métodos estatísticos	Mensuração da qualidade, planejamento da qualidade e projeto de programas
Responsável pela Qualidade	Departamento de Inspeção	Departamentos de Fabricação e Engenharia	Todos os departamentos
Orientação e abordagem	“inspeção” da qualidade	“controle” da qualidade	“construção” da qualidade

Quadro 2.1 – Evolução da Inspeção à Garantia da Qualidade

Fonte: Garvin (2002)

Para Slack *et al.* (1999) a Administração da Qualidade pode ser vista como uma evolução natural das abordagens anteriores e, em resumo, envolve os seguintes aspectos:

- a) Atendimento das necessidades e expectativas dos consumidores;
- b) Inclusão de todas as partes da organização;
- c) Inclusão de todas as pessoas da organização;
- d) Exame de todos os custos relacionados com qualidade;
- e) Construção da qualidade desde o projeto ao invés de apenas inspecionar;
- f) Desenvolvimento de sistemas e procedimentos que apoiem qualidade e melhoria;
- g) Desenvolvimento de um processo de melhoria contínua.

2.2 Métodos Estatísticos Aplicados à Qualidade

Estatística pode ser definida como uma ciência que trata da organização, descrição, análise e interpretação de dados experimentais. De acordo com Feigenbaum (1994) a Estatística desempenha papel importante nos programas de Controle da Qualidade Total, pois ao longo dos anos suas técnicas e metodologias tornaram-se cada vez mais amplamente utilizadas e aceitas nas organizações. A Figura 2.1 representa o papel que as técnicas estatísticas desempenham na avaliação e controle de um processo produtivo esquematizado em um conjunto de entradas, fatores controláveis e incontroláveis e saídas.

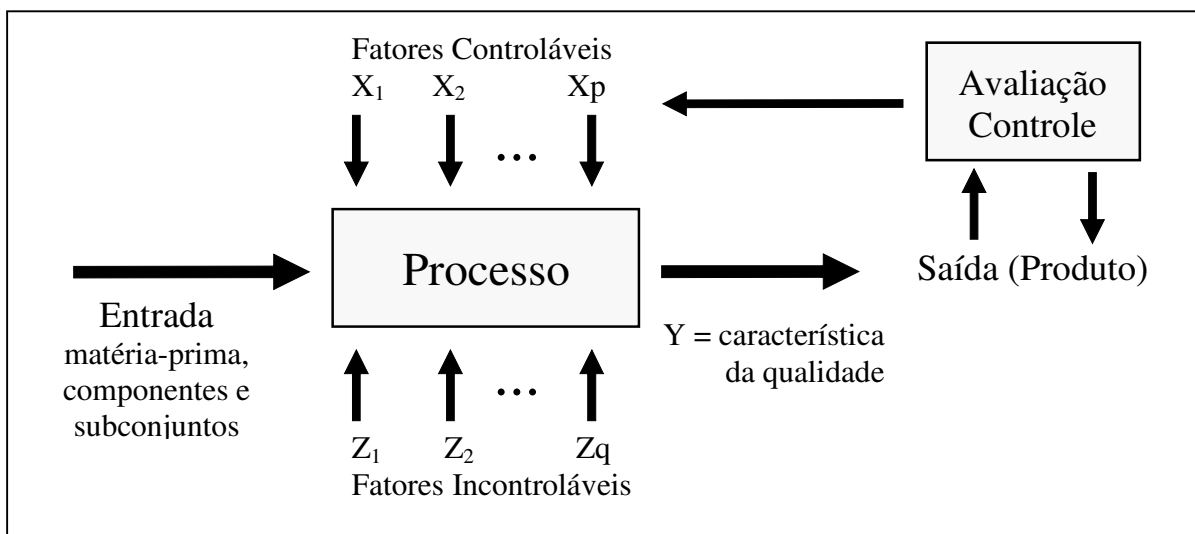


Figura 2.1 – Representação de um processo produtivo

Fonte: Montgomery (1985)

Conforme resumo apresentado por Montgomery (1985), Juran & Godfrey (1999) e Garvin (2002) a história de utilização de técnicas estatísticas em fabricação e garantia da qualidade teve início em 1924 com Walter A. Shewhart da Bell Telephone Laboratories com o desenvolvimento do conceito de gráficos de controle estatístico. Ainda na década de 1920 Harold F. Dodge e Harold G. Romig, também da Bell Telephone Laboratories, desenvolveram a aceitação por amostragem estatisticamente baseada como alternativa à inspeção 100%. Em meados da década de 1930 métodos de controle estatístico da qualidade foram amplamente utilizados na Western Electric. Durante a Segunda Guerra Mundial o uso de conceitos estatísticos de aceitação e controle de qualidade nas indústrias tornaram-se mais difundidos e posteriormente as demais organizações perceberam que as técnicas de controle da qualidade poderiam ser úteis em todos os tipos de produtos e serviços. Nas décadas de 1950 e 1960 houve desenvolvimento tanto em garantia da qualidade como em custos da

qualidade, confiabilidade e o surgimento de uma forma de gerenciamento das organizações sob o ponto de vista da qualidade. Mais recentemente o desenvolvimento em garantia da qualidade também se relaciona às questões comportamentais, tais como motivação de funcionários e responsabilidade do produto.

Atualmente é inquestionável o papel fundamental que os métodos estatísticos desempenham na melhoria da qualidade. Montgomery & Runger (2003) destacam algumas de suas aplicações:

- a) Planejamento e desenvolvimento de produtos – podem ser utilizados experimentos planejados para comparar, por exemplo, diferentes materiais, componentes ou ingredientes e auxiliar na determinação das tolerâncias com redução significativa de custo e tempo de desenvolvimento;
- b) Determinação da capacidade de um processo de fabricação – o Controle Estatístico de Processo pode ser usado para melhorar sistematicamente um processo pela redução da variabilidade;
- c) Investigação de melhorias no processo – técnicas de Planejamento de Experimentos podem promover melhorias que conduzem a rendimentos maiores e menores custos de fabricação;
- d) Fornecimento de dados de desempenho e confiabilidade do produto – testes de vida podem conduzir a novos e melhores projetos e produtos que possuam vidas úteis mais longas e menores custos operacionais e de manutenção.

De acordo com Juran & Godfrey (1999) as ferramentas e os métodos estatísticos têm contribuído de modo determinante para o planejamento e melhoria da qualidade e, em alguns casos específicos tais ferramentas são mais do que úteis, pois os problemas de qualidade simplesmente não podem ser solucionados por completo sem a sua adequada aplicação.

Estas ferramentas estatísticas, que são essenciais nas atividades de melhoria da qualidade em produtos e serviços, quando agrupadas de modo organizado, constituem o Controle Estatístico da Qualidade. Segundo Sarkadi & Vincze (1974) há duas formas principais de Controle Estatístico da Qualidade: aceitação por amostragem e técnicas de controle estatístico de processo.

Na visão de Montgomery (1985) o objetivo primário do Controle Estatístico da Qualidade é a redução sistemática da variabilidade nas características-chave do produto. A Figura 2.2 ilustra a forma como isso ocorre ao longo do tempo:

- a) Nos estágios iniciais, quando a aceitação por amostragem é a técnica mais utilizada, as falhas no processo ou os itens em desacordo com as especificações, constituem alta percentagem das saídas do processo;
- b) A introdução de técnicas de controle estatístico de processo permite a estabilização do processo e redução da variabilidade;
- c) Finalmente, a adoção de experimentos planejados, empregados em conjunto com o controle de processo, pode minimizar ainda mais a variabilidade, resultando em um processo produtivo virtualmente livre de defeitos.

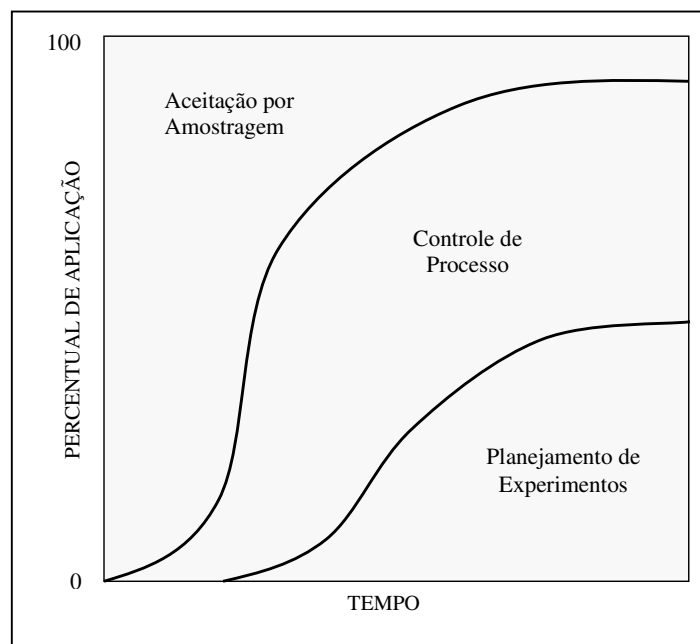


Figura 2.2 – Diagrama de utilização dos métodos para garantia da qualidade

Fonte: Montgomery (1985)

A argumentação até aqui exposta, confirma o entendimento de que os métodos estatísticos aplicados à qualidade fornecem as ferramentas necessárias para avaliação e melhoria de projetos, processos, produtos e serviços de forma robusta e abrangente.

2.3 Análise e Tratamento de Dados

2.3.1 Aspectos Gerais

A análise de dados é parte fundamental em aplicações que envolvem o uso da Estatística. Na visão de Onishi (2002), o papel da análise de dados é possibilitar que os dados coletados revelem a sua verdade, mas infelizmente isso não é tão simples, pois para tanto é

necessário estabelecer e manter condições adequadas. Para Kume (1993) os dados são os guias das ações, pois a partir deles apreendem-se os fatos pertinentes e tomam-se as providências de maneira embasada.

A atividade que precede a etapa de análise é a coleta de dados e, para que sua execução ocorra de maneira adequada, são necessários alguns cuidados, dentre os quais segundo Soares (2003), destacam-se os seguintes:

- a) A precisão dos instrumentos;
- b) As normas de medição observadas e padronizadas;
- c) O treinamento do pessoal responsável pela coleta

Segundo Montgomery & Runger (2003) sumários e apresentações de dados bem constituídos são essenciais para o bom julgamento estatístico, pois permitem a visualização de características importantes e sua correta interpretação. Um exemplo de organização de dados através de um diagrama de caixa (*Box Plot*) é mostrado através da Figura 2.3.

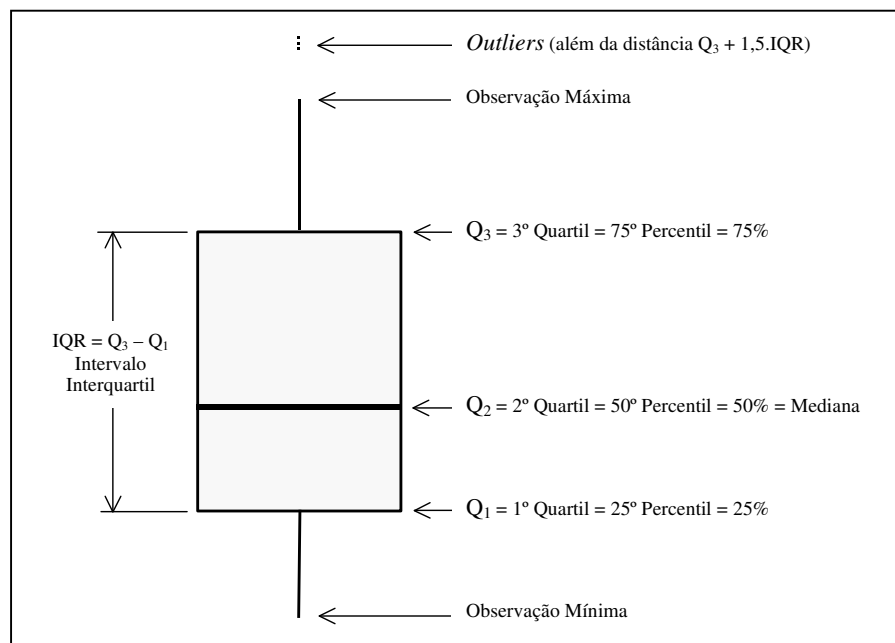


Figura 2.3 – Exemplo de um diagrama de caixa (*Box Plot*)

Fonte: Adaptado de Montgomery & Runger (2003)

O diagrama de caixa (*Box Plot*), que pode ter alinhamento vertical ou horizontal, é uma apresentação gráfica que descreve simultaneamente várias características importantes de um conjunto de dados, tais como quartis, mediana, dispersão, desvio de simetria e identificação das observações que estão muito distantes do restante dos dados. Estas observações que não seguem o padrão da distribuição são também chamadas de *outliers*.

Os *outliers* podem refletir uma propriedade típica de um fenômeno subjacente ou podem ser originados por erros de medição ou outras anomalias resultantes da coleta de dados. No diagrama de caixa os *outliers* são representados por pontos individuais além da faixa de 1,5 interquartil em relação ao 1° quartil ou em relação ao 3° quartil, podendo ser calculado conforme indicado nas Equações 2.1 e 2.2 (MONTGOMERY & RUNGER, 2003):

$$Outlier = \{x / x < Q_1 - 1,5 \cdot IQR\} \quad (2.1)$$

ou

$$Outlier = \{x / x > Q_3 + 1,5 \cdot IQR\} \quad (2.2)$$

Onde:

x = valor observado

Q_1 = 1° Quartil

Q_3 = 3° Quartil

IQR = Intervalo Interquartil = $Q_3 - Q_1$

2.3.2 Distribuição de Frequências

De acordo com Feigenbaum (1994) a distribuição de frequências pode ser definida como tabulação ou registro do número de vezes que uma medição de uma dada característica da qualidade ocorre na amostra do produto sob verificação. As distribuições estatísticas representam uma população que é a coleção de todas as observações potenciais acerca de um determinado fenômeno. O conjunto de dados efetivamente observado ou extraído constitui a amostra desta população, sendo que um parâmetro está para a população assim como uma estatística está para a amostra.

Para construir uma distribuição de frequências deve-se dividir a faixa de dados em intervalos de classe ou células. A quantidade de intervalos depende da quantidade total de observações e da dispersão dos dados. Como regra prática pode-se utilizar um número de intervalos de classe aproximadamente igual à raiz quadrada do número de observações. A distribuição de frequências pode ser apresentada graficamente através de histogramas, que são de fácil interpretação, principalmente quando os intervalos de classe possuem a mesma largura.

A Figura 2.4 apresenta um exemplo de distribuição de frequências com histograma relacionado à medição da resistência à compressão de um determinado corpo-de-prova.

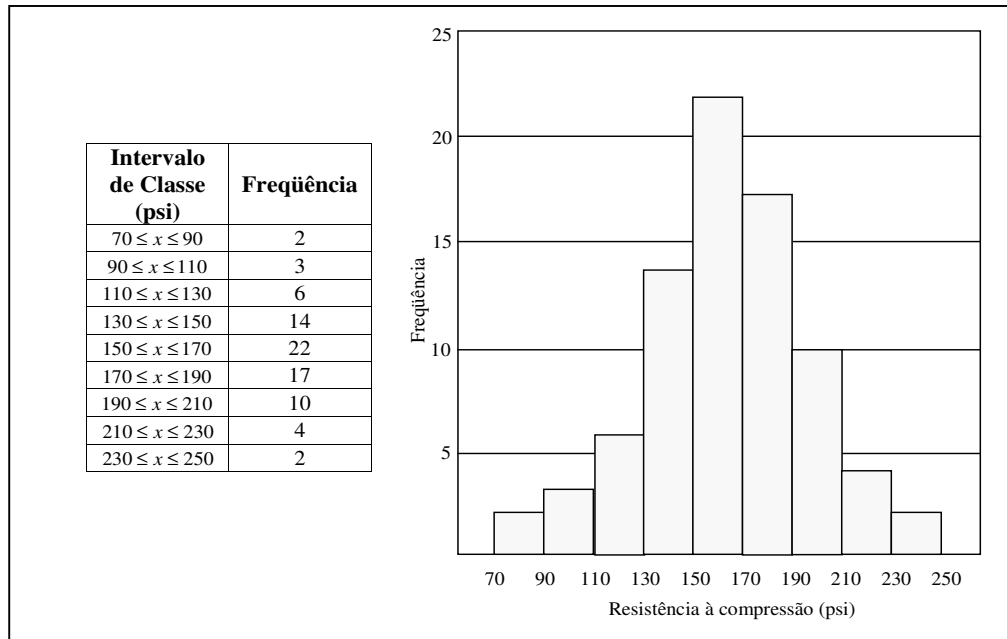


Figura 2.4 – Exemplo de uma distribuição de frequências com histograma

Fonte: Montgomery & Runger (2003)

As distribuições podem estar associadas a variáveis aleatórias discretas ou contínuas. Uma variável aleatória é dita discreta quando possui uma faixa finita ou uma faixa infinita contável, tais como, contagem de peças defeituosas em um determinado lote, proporção de defeituosos, número de produtos vendidos, etc. São exemplos de distribuições discretas:

- a) Distribuição Binomial;
- b) Distribuição de Poisson;
- c) Distribuição Geométrica;
- d) Distribuição de Pascal;
- e) Distribuição Multinomial;
- f) Distribuição Hipergeométrica.

Uma variável aleatória contínua é uma variável aleatória com um intervalo finito ou infinito de números reais para sua faixa, tais como corrente elétrica, comprimento, pressão, temperatura, peso, etc. São exemplos de distribuições contínuas:

- a) Distribuição Normal;
- b) Distribuição Uniforme;
- c) Distribuição Chi-Quadrado;
- d) Distribuição F de Fisher;
- e) Distribuição t de Student;
- f) Distribuição Beta;

- g) Distribuição Cauchy;
- h) Distribuição Exponencial;
- i) Distribuição Gamma;
- j) Distribuição Laplace,
- k) Distribuição Lognormal;
- l) Distribuição Weibull.

O Quadro 2.2 apresenta exemplos de aplicação de algumas distribuições estatísticas.

DISTRIBUIÇÃO ESTATÍSTICA	CAMPOS DE APLICAÇÃO
Normal	Diversas propriedades físicas, mecânicas, elétricas, químicas, etc.
Lognormal	Fenômenos de ciclo de vida; situações assimétricas onde as ocorrências são concentradas na cauda da distribuição, nas quais diferenças em observações são de grande ordem de magnitude.
Weibull (2 parâmetros)	Mesmos casos da Lognormal e também em situações onde as ocorrências (por exemplo, taxa de falhas) podem diminuir, aumentar ou permanecer constantes com o aumento da característica medida.
Weibull (3 parâmetros)	Mesmos casos da Weibull (2 parâmetros) e, em adição, diversas propriedades físicas, mecânicas, elétricas, químicas, exceto aquelas caracterizadas pela distribuição Normal.
Exponencial	Ciclo de vida de sistemas, conjuntos, etc. Para componentes e situações onde as falhas ocorrem ao acaso isoladamente e não dependem do tempo em serviço.
Binomial	Número de defeitos em um tamanho de amostra “n” tomada de um grande lote tendo a fração “p” de defeituosos; situações envolvendo tipos de observações “passa/não passa” e “bom/ruim”. Proporção do lote não se altera apreciavelmente como resultado da escolha da amostra.
Hipergeometrica	Inspeção de peças mecânicas, elétricas de um pequeno lote tendo percentual conhecido de defeituosos. Mesmos casos da Binomial, com a diferença que a proporção do lote pode se alterar como resultado da escolha da amostra.
Poisson	Situação onde o número de vezes que um evento ocorre pode ser observado, mas não o número de vezes que o evento não ocorre. Aplicável aos eventos aleatórios distribuídos no tempo.

Quadro 2.2 – Campos de aplicação de algumas distribuições estatísticas

Fonte: Adaptado de Lipson & Sheth (1973)

2.3.3 A Distribuição Normal

Em diversos fenômenos observados na natureza, bem como em processos experimentais, de produção e de prestação de serviços existem repetidas observações que diferem entre si e que variam em torno de um valor central em uma distribuição aproximadamente simétrica na qual os pequenos desvios ocorrem com muito mais freqüência do que os grandes desvios. Para Box et al. (1978) uma distribuição contínua que se destaca por representar tal situação e ocupa uma importante posição no estudo da Estatística é a distribuição Normal ou Gaussiana. A função densidade de probabilidade para uma distribuição Normal é dada através da Equação 2.3 (MONTGOMERY, 1984).

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}} \text{ para } -\infty < x < \infty \quad (2.3)$$

Onde:

x = variável aleatória em estudo

μ = média da distribuição ($-\infty < \mu < \infty$)

σ = desvio-padrão da distribuição ($\sigma > 0$)

A curva correspondente à distribuição Normal, também conhecida como curva normal, curva do sino (*Bell Curve*) ou curva de Gauss é representada através da Figura 2.5.

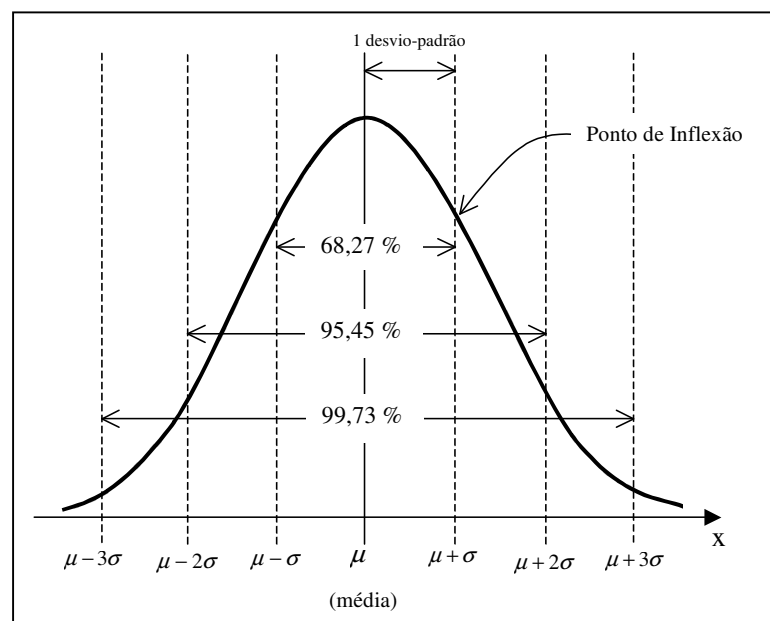


Figura 2.5 – Principais aspectos da curva correspondente à distribuição Normal

Fonte: Adaptado de Montgomery (1985) e Kume (1993)

Segundo Lipson & Sheth (1973), a distribuição Normal que é a mais conhecida e a mais amplamente utilizada das distribuições estatísticas, teve seus conceitos desenvolvidos no século XVIII, de forma independente, através dos trabalhos de Gauss, Laplace e De Moivre. Esta distribuição apresenta algumas características marcantes, dentre as quais destacam-se:

- A curva é simétrica em relação à média e esta coincide com a moda e a mediana;
- A notação $N(\mu; \sigma^2)$ indica uma distribuição Normal com média μ e variância σ^2 ;
- Apresenta coeficiente de assimetria igual a zero e de curtose igual a três;
- O desvio padrão σ indica a distância da média μ ao ponto de inflexão da curva;
- 68,27% dos dados encontram-se na faixa entre $\pm 1 \sigma$ (desvio padrão) da média;
- 95,45% dos dados encontram-se na faixa entre $\pm 2 \sigma$ (desvios padrão) da média;
- 99,73% dos dados encontram-se na faixa entre $\pm 3 \sigma$ (desvios padrão) da média.

Cálculos de probabilidade em relação à distribuição Normal podem ser efetuados a partir da curva normal reduzida ou padronizada $N(0;1)$, que é uma distribuição Normal com média $\mu = 0$ e variância $\sigma^2 = 1$ (Anexo A). A variável pertencente a esta distribuição é denominada variável aleatória normal padrão e é denotada por z . A variável z representa a distância de x a partir de sua média em termos de desvios padrão e é, portanto, fundamental para cálculo de probabilidades para uma variável aleatória normal arbitrária x . A Figura 2.6 ilustra uma curva normal reduzida ou padronizada.

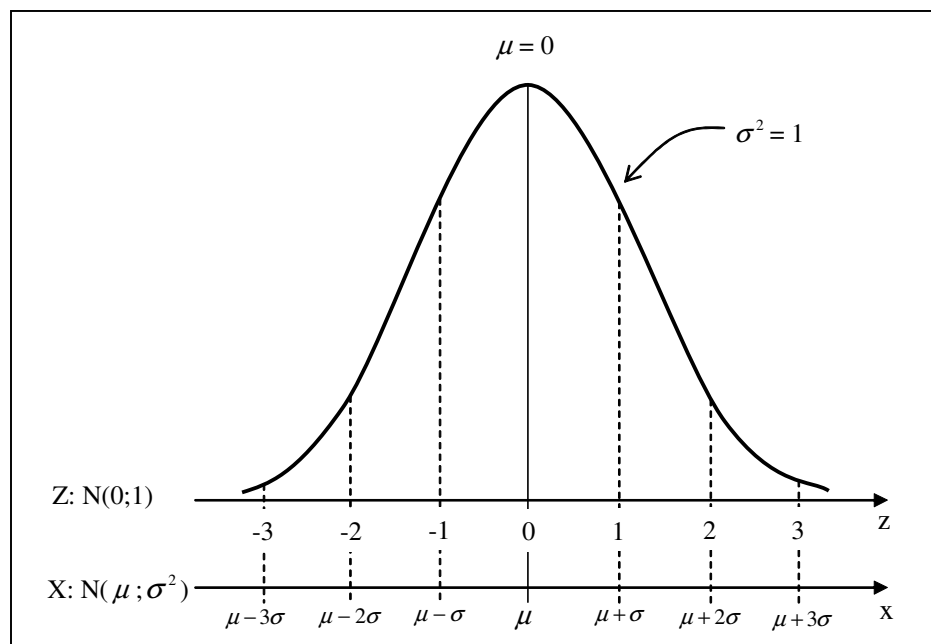


Figura 2.6 – Curva normal padronizada

Fonte: Adaptado de Montgomery (1985)

O cálculo da variável z , também chamada de escore padronizado, é dado pela fórmula indicada através da Equação 2.4 (MONTGOMERY, 1985).

$$z = \frac{x - \mu}{\sigma} \quad (2.4)$$

Onde:

z = variável aleatória normal padrão ou escore padronizado

x = variável aleatória em estudo

σ = desvio-padrão da distribuição ($\sigma > 0$)

✓ TEOREMA CENTRAL DO LIMITE:

O Teorema Central do Limite (Equação 2.5), considerado um dos mais úteis teoremas da Estatística, estabelece que a soma e, conseqüentemente, a média de um grande número de variáveis aleatórias independentes apresenta uma distribuição aproximadamente normal, indiferentemente do tipo de distribuição de probabilidade dos valores individuais. Esta aproximação da condição de normalidade melhora à medida que a quantidade de variáveis independentes aumenta (MONTGOMERY, 1985).

$$z_n = \frac{y - \sum_{i=1}^n \mu_i}{\sqrt{\sum_{i=1}^n \sigma_i^2}} \sim N(0;1) \quad (2.5)$$

Onde:

x_i = variáveis aleatórias independentes

$y = x_1 + x_2 + x_3 + \dots + x_n$

z_n = escore padronizado

μ_i = média de x_i

σ_i^2 = variância de x_i

As condições acima mencionadas são freqüentemente observadas em diversos métodos estatísticos aplicados à Qualidade. De acordo com Montgomery & Runger (2003), como conseqüência do Teorema Central do Limite, sempre que ocorrerem replicações de um experimento aleatório, a variável aleatória que for igual ao resultado médio das réplicas tenderá a ter uma distribuição Normal, à medida que o número de réplicas se torne grande.

2.4 Controle Estatístico de Processo

2.4.1 Definição e Conceitos Básicos

De acordo com Woodall (2000) o Controle Estatístico de Processo (CEP) é uma subárea do Controle Estatístico da Qualidade e consiste em métodos para o entendimento, o monitoramento e a melhoria do desempenho de processos ao longo do tempo, priorizando a compreensão da variação das características da qualidade como sendo de fundamental importância para o controle do processo. Segundo Silva (1999) o Controle Estatístico de Processo é o ramo do Controle de Qualidade que consiste na coleta, análise e interpretação de dados, estabelecimento de padrões, comparação de desempenhos e verificação de desvios para utilização em atividades de melhoria e controle da qualidade de produtos e serviços. Na visão de Campos (1992) o CEP é parte integrante dos programas de Qualidade Total adotados por inúmeras empresas como estratégia habilitadora de sua permanência e expansão nos mercados globalizados.

Galuch (2002) afirma que o movimento originado pelo CEP resultou no treinamento de muitos supervisores e operadores em ferramentas estatísticas básicas, capacitando-os a entender melhor o comportamento de processos e produtos e muitos deles aprenderam que decisões baseadas na coleta e análise de dados superam as decisões baseadas no empirismo. Entretanto, Soares (2001) salienta que o CEP é um sistema de decisão e não um substituto da experiência, ou seja, os métodos estatísticos auxiliam a detecção e isolamento de desarranjos em processos, indicando suas causas; com isso a gerência e a equipe técnica podem indicar e aplicar as ações necessárias.

Em linhas gerais o Controle Estatístico de Processo pode ser definido como um conjunto de ferramentas que tem o propósito de indicar se um processo está funcionando de maneira ideal, quando apenas causas comuns de variação estão presentes, ou se este processo está fora de controle e necessita de algum tipo de ação corretiva, ou seja, quando existem causas especiais de variação. Para um completo entendimento deste conceito é necessário estabelecer algumas definições que sustentam o Controle Estatístico de Processo, tais como, processo, variabilidade, causas de variação e capacidade.

✓ PROCESSO:

Com base na norma NBR ISO 9000:2000, define-se processo como um conjunto de atividades inter-relacionadas ou interativas que transformam insumos (entradas) em produtos (saídas). Toledo (2005) também define processo como um conjunto de transformações que

determinado produto sofre em função de suas entradas, fatores controláveis e fatores que não podem ser controlados (ruídos), resultando em uma determinada saída.

Segundo Harrington (1993) existem dois tipos de processo: o empresarial e o produtivo. Um processo empresarial consiste em um grupo de tarefas interligadas logicamente que fazem uso de recursos da organização para gerar resultados definidos em apoio aos objetivos da organização. Para Soares (2001) o processo produtivo é o que melhor se identifica nas indústrias, pois envolve a manufatura, isto é, envolve contato direto com o produto ou serviço que será fornecido ao cliente externo.

✓ VARIABILIDADE:

O estudo da variabilidade é de grande importância na aplicação de métodos estatísticos, pois é algo inerente a todo processo produtivo. Para Montgomery (1985) a variabilidade devido à aleatoriedade inerente à natureza, torna praticamente impossível a produção de dois produtos ou serviços idênticos. Se essa variabilidade for pequena, ou seja, se não causar impacto perceptível para o consumidor, é tolerável, caso contrário será indesejável ou mesmo inaceitável.

Na visão de Reis (2001) o conceito de variabilidade é absolutamente crucial para a compreensão de todas as técnicas não somente de Controle Estatístico de Processo, mas de qualquer método estatístico e, não são raros os casos em que os envolvidos na aplicação das técnicas de CEP não dispõem de uma clara idéia a respeito.

✓ CAUSAS DE VARIAÇÃO:

Conforme definido por diversos autores as causas de variação em um processo podem ser classificadas como causas especiais e causas comuns, sendo que, de acordo com Montgomery & Runger (2003), os termos inicialmente utilizados por Walter A. Shewhart eram causas atribuídas e causas casuais, respectivamente. Deming (1986) afirma que as causas especiais correspondem a 6% dos problemas e as causas comuns relacionam-se a 94% das situações observadas nos processos.

As causas especiais também chamadas de causas assinaláveis indicam que o processo não está sob controle estatístico. Normalmente, causas especiais são as que realmente afetam a qualidade do produto e na maioria das oportunidades faz-se necessário o seu entendimento e eliminação, para que o processo se mantenha estável e sob controle. As causas especiais podem ser causas esporádicas, como por exemplo, condições ambientais passageiras, avarias ou problemas no equipamento de produção, ferramenta inadequada, lote isolado de matéria-prima com defeito e mão-de-obra. Segundo Pires (2000) o monitoramento e a ação corretiva

sobre as causas especiais é de responsabilidade da área operacional da empresa e, uma vez identificadas, pode-se atuar sobre as mesmas, buscando a estabilização do processo.

As causas comuns também denominadas causas sistêmicas, são causas inerentes ao processo produtivo, pertencendo assim ao próprio sistema, não afetando de maneira incisiva a qualidade do produto, caracterizando um processo sob controle estatístico. A influência de algumas destas causas só pode ser removida ou detectada com iniciativas gerenciais. De acordo com Souza (2003) quando apenas as causas comuns estão atuando no processo, a variabilidade se mantém em uma faixa estável, conhecida como faixa característica do processo. Alguns exemplos de causas comuns são treinamento inadequado dos operadores, utilização de equipamentos cronicamente imprecisos, manutenção deficiente, falta de normalização, falta de documentação, aquisição sistemática de matéria-prima de baixa qualidade, entre outros. Para Grant & Leavenworth (1972) a atuação sobre as causas comuns normalmente requer investimentos na melhoria de equipamentos e troca de matérias-primas e, uma vez mantidas em níveis razoáveis, as mesmas não afetam de maneira nociva a qualidade dos itens manufaturados.

✓ CAPACIDADE:

Conforme norma NBR ISO 9000:2000 capacidade é definida como a aptidão de uma organização, sistema ou processo de realizar um produto que irá atender aos requisitos especificados para este produto. Para Prazeres (1997) a capacidade de um processo pode ser definida como a habilidade intrínseca de um processo em desempenhar suas funções nas condições de trabalho, satisfazendo certas especificações e tolerâncias, sendo uma medida de uniformidade inerente ao processo. Aspectos gerais em relação aos índices de capacidade ou capacidade serão discutidos posteriormente.

2.4.2 Características Gerais

De acordo com Montgomery & Runger (2003) o Controle Estatístico de Processo (CEP) é um poderoso recurso a ser utilizado em tempo real (*on line*) para avaliar a condição de estabilidade de um processo e para melhorar sua capacidade através da redução da variabilidade. Costuma-se entender o CEP como um conjunto de ferramentas para resolver problemas, aplicável a qualquer processo. Segundo Montgomery & Runger (2003) as ferramentas mais importantes do CEP são as seguintes:

- a) Histograma – apresentação gráfica de uma distribuição de frequências ou série de distribuições quantitativas por meio de retângulos justapostos de modo a revelar a quantidade de variação existente dentro de um processo;
- b) Gráfico de Pareto – representação gráfica do princípio enunciado pelo economista italiano Vilfredo Pareto em 1897, sobre desigualdade na distribuição de renda, adaptado aos problemas de qualidade, ou seja, as melhorias podem ser obtidas se a atenção estiver concentrada primeiramente na direção dos poucos problemas vitais e em seguida na direção de suas poucas causas vitais;
- c) Diagrama de causa e efeito – conhecido como Diagrama de Ishikawa ou Diagrama de Espinha de Peixe, é uma técnica visual que interliga os efeitos com as respectivas causas, possibilitando uma visualização das variáveis do processo;
- d) Diagrama de defeito-concentração – esquema ou desenho de produto com identificação e localização dos defeitos de maior ocorrência;
- e) Gráfico de controle – a mais poderosa ferramenta do CEP, é utilizada para mostrar a variação do processo em relação aos limites de controle estipulados;
- f) Diagrama de dispersão – representação gráfica para investigar possível correlação entre duas variáveis e para provar possível relação entre causa e efeito;
- g) Folha de verificação – formulário no qual os dados podem ser coletados e registrados de maneira ordenada, permitindo rápida interpretação de resultados.

Na opinião de Reis (2001) o CEP permite o monitoramento contínuo do processo, possibilitando uma ação imediata no momento em que um problema é detectado, sustentando a filosofia que preconiza a construção da Qualidade dentro do processo e a prevenção de problemas. Segundo Silva (1999) os gráficos de controle ou cartas de controle, utilizados no CEP, são normalmente aplicados em processos contínuos e semi-contínuos onde uma dimensão ou característica do produto é medida em uma amostra aleatória tirada sob um determinado período de tempo com o objetivo de avaliar a estabilidade ou estado de controle estatístico de um processo.

Para Montgomery & Runger (2003) além das ferramentas citadas também existem os aspectos comportamentais envolvidos na implementação do CEP, isto é, a atitude voltada para a melhoria contínua na qualidade e na produtividade através da redução sistemática da variabilidade, que afinal de contas é o objetivo principal do Controle Estatístico de Processo.

A Figura 2.7 apresenta a identificação das etapas envolvidas na operacionalização do CEP e a utilização de informações provenientes dos gráficos de controle para auxiliar a tomada de decisões sobre ações de melhoria.

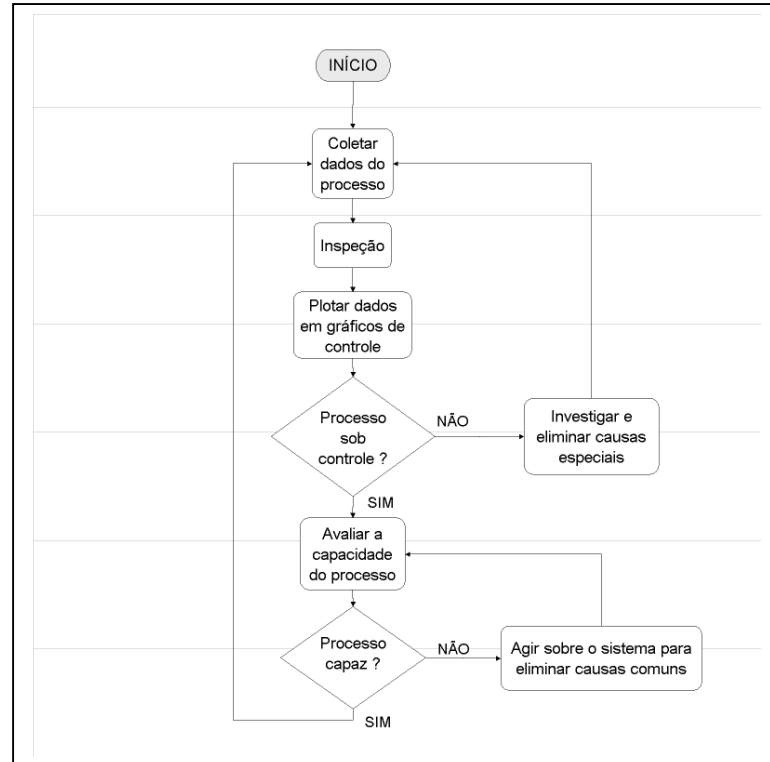


Figura 2.7 – Estratégia para melhorias em Controle Estatístico de Processo

Fonte: Adaptado de Schissatti (1998)

2.4.3 Gráficos de Controle

Conforme mencionado anteriormente a principal e mais poderosa ferramenta do CEP é o gráfico de controle que foi introduzido por Walter A. Shewhart da Bell Telephone Laboratories em 1924 como mecanismo de distinção entre as chamadas causas comuns (casuais) e causas especiais (atribuídas) de variação de um processo. Segundo Almeida (2003), devido ao fato de os gráficos de controle serem ferramentas extremamente úteis para identificação de variações observadas em um processo em decorrência de causas comuns ou especiais, sua utilização é amplamente difundida em Controle Estatístico de Processo. Pode-se dizer que tais gráficos são comparações do desempenho do processo, que pode ser avaliado por meio de qualquer característica da qualidade que seja de interesse, como por exemplo, a

média de determinado comprimento, os valores de diâmetros de furos, o número de itens faltantes, dentre outros.

De acordo com Montgomery & Runger (2003) existe forte conexão entre gráficos de controle e teste de hipóteses. Essencialmente o gráfico de controle é um teste da hipótese nula de que o processo está em um estado de controle estatístico e pode ser definido da seguinte maneira:

H_0 : O processo está sob controle estatístico (hipótese nula)

H_1 : O processo não está sob controle estatístico (hipótese alternativa)

Através da determinação dos limites de controle pode-se avaliar se o processo encontra-se sob controle estatístico, verificando as características dos pontos registrados no gráfico. Se os mesmos distribuem-se de acordo com padrões aleatórios, indicam processo sob controle estatístico, pois somente causas comuns estão presentes no mesmo; porém se o gráfico apresenta padrões não aleatórios, tais como pontos além dos limites de controle, significa que o processo pode estar sendo influenciado por causas especiais de variação. Neste caso deve ser efetuada uma análise cuidadosa do processo para possibilitar o diagnóstico do problema e a tomada de ações para que o processo retorne à condição anterior. Um gráfico de controle típico é mostrado através da Figura 2.8:

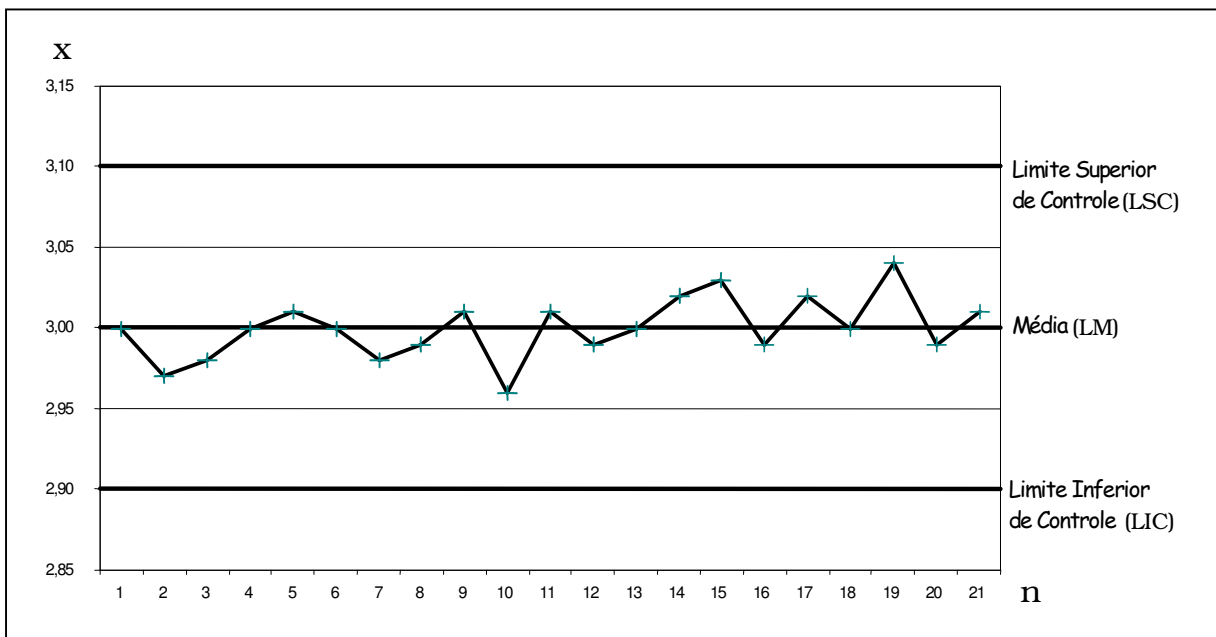


Figura 2.8 – Exemplo de um gráfico de controle

Fonte: Adaptado de Montgomery & Runger (2003)

Segundo Reis (2001), de maneira geral, o procedimento utilizado para a construção de gráficos de controle baseia-se na coleta de amostras de tamanhos fixos em intervalos amostrais, sendo que seu registro deve respeitar a ordem cronológica para que seja possível identificar quais são as causas de variação não aleatórias. Através dos valores amostrais coletados são obtidas estimativas para a média e dispersão do processo e, a partir de então, são determinados os limites de controle inferior e superior e a linha central do gráfico.

Para Montgomery & Runger (2003) os gráficos de controle têm uma longa história de utilização industrial e há, no mínimo, cinco razões para sua popularidade:

- a) Gráficos de controle são comprovadamente úteis para melhoria da qualidade;
- b) Gráficos de controle são efetivos na prevenção de defeitos;
- c) Gráficos de controle podem prever ajustes desnecessários no processo;
- d) Gráficos de controle fornecem informação sobre diagnóstico;
- e) Gráficos de controle fornecem informação sobre a capacidade do processo.

Além disso, nos últimos anos a moderna tecnologia computacional simplificou a coleta de dados e a implementação de gráficos de controle em qualquer tipo de processo, possibilitando que a análise possa ser efetuada em um microcomputador ou em um terminal de rede em tempo real no próprio local de trabalho.

Na visão de Reis (2001) o conceito de processo sob controle estatístico é extremamente importante, pois é através deste que se evidencia a presença de padrões não aleatórios. Há inúmeros padrões não aleatórios que podem ser encontrados nos gráficos de controle e para permitir sua correta identificação os seguintes critérios foram estabelecidos (ASQC/AIAG, 1992):

- a) Existência de pontos além dos limites superior ou inferior de controle;
- b) Ocorrência de um padrão, uma tendência ou um ciclo que obviamente não se mostre aleatório;
- c) Regra da seqüência de oito: existência de oito pontos consecutivos, todos acima ou todos abaixo da linha central;
- d) Seis pontos consecutivos, todos crescentes ou todos decrescentes;
- e) Catorze pontos consecutivos se alternado para cima e para baixo;
- f) Dois ou três pontos consecutivos além dos limites de controle fixados à distância de dois desvios-padrão a contar da linha central;
- g) Quatro dentre cinco pontos consecutivos fora dos limites de controle fixados à distância de um desvio padrão a contar da linha central.

Para utilização dos gráficos de controle alguns aspectos importantes merecem atenção especial para que não ocorram análises incorretas em função da falta de entendimento de suas peculiaridades e características. De acordo com Alwan & Roberts (1995) algumas suposições precisam ser satisfeitas para que, independentemente do tipo do gráfico de controle empregado, os resultados sejam válidos:

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

As suposições apresentadas acima são extremamente importantes, pois segundo Reis (2001) os limites de controle calculados e as regras para identificação da presença de padrões não aleatórios pressupõem que estas suposições sejam satisfeitas, caso contrário o valor encontrado nestes gráficos seria no mínimo questionável.

Existem diversos critérios para classificação dos tipos de gráficos de controle conforme argumentação de Ramos (2003). Quanto à característica de controle, por exemplo, os gráficos de controle distinguem-se em entre gráficos de controle para variáveis, isto é medidas em geral e gráficos de controle para atributos, ou seja, dados contáveis.

De acordo com Montgomery (1985) os principais gráficos de controle para variáveis são os seguintes:

- a) Gráfico \bar{X} ou X-bar – É conhecido como gráfico das médias e apresenta o registro das médias da característica da qualidade em estudo, de modo a controlar o valor médio da característica de interesse. Na construção deste gráfico supõe-se a condição de normalidade das médias amostrais.
- b) Gráfico S – É o gráfico de desvio padrão, onde os desvios-padrão da característica da qualidade em estudo são registrados de modo a controlar a variabilidade da característica de interesse. Na construção deste gráfico também se supõe a normalidade dos dados amostrais e, geralmente é preferido em relação ao gráfico de intervalos ou das amplitudes.
- c) Gráfico R – É o gráfico de intervalo ou das amplitudes, no qual as amplitudes da característica da qualidade em estudo são registradas de forma a controlar a variabilidade da característica de interesse.

Na prática a utilização conjunta dos gráficos X-bar-R ou X-bar-S é muito comum, pois ambos permitem um melhor acompanhamento e controle da característica da qualidade em estudo, sendo que este último é preferível para amostras de maior tamanho.

As fórmulas para definição dos limites de controle dos tipos mais comuns de gráficos de controle para variáveis são apresentadas no Quadro 2.3. As constantes $A_2, A_3, B_3, B_4, D_3, D_4$ e E_2 , que permitem a estimativa de $\mu \pm 3\sigma$, variam em função do tamanho das amostras (Anexo B) e encontram-se disponíveis em referências tais como Montgomery (1985), Feigenbaum (1994), Juran & Godfrey (1999), Montgomery & Runger (2003), entre outros.

TIPO DE GRÁFICO	LIMITES DE CONTROLE	
	Gráfico	Fórmulas
\bar{X} e R	Média	$LSC_{\bar{x}} = \bar{\bar{X}} + A_2 \cdot \bar{R}$ $LM_{\bar{x}} = \bar{\bar{X}}$ $LIC_{\bar{x}} = \bar{\bar{X}} - A_2 \cdot \bar{R}$
	Amplitude	$LSC_R = D_4 \cdot \bar{R}$ $LM_R = \bar{R}$ $LIC_R = D_3 \cdot \bar{R}$
\bar{X} e S	Média	$LSC_{\bar{x}} = \bar{\bar{X}} + A_3 \cdot \bar{S}$ $LM_{\bar{x}} = \bar{\bar{X}}$ $LIC_{\bar{x}} = \bar{\bar{X}} - A_3 \cdot \bar{S}$
	Desvio-Padrão	$LSC_S = B_4 \cdot \bar{R}$ $LM_S = \bar{S}$ $LIC_S = B_3 \cdot \bar{S}$
\tilde{X} e R	Mediana	$LSC_{\tilde{x}} = \bar{\tilde{X}} + A_2 \cdot \bar{R}$ $LM_{\tilde{x}} = \bar{\tilde{X}}$ $LIC_{\tilde{x}} = \bar{\tilde{X}} - A_2 \cdot \bar{R}$
	Amplitude	$LSC_R = D_4 \cdot \bar{R}$ $LM_R = \bar{R}$ $LIC_R = D_3 \cdot \bar{R}$
X_i e R	Individuais (X)	$LSC_{X_i} = \bar{X} + E_2 \cdot \bar{Rm}$ $LM_{X_i} = \bar{X}$ $LIC_{X_i} = \bar{X} - E_2 \cdot \bar{Rm}$
	Amplitude	$LSC_R = D_4 \cdot \bar{Rm}$ $LM_R = \bar{Rm}$ $LIC_R = D_3 \cdot \bar{Rm}$

Quadro 2.3 – Fórmulas dos tipos mais comuns de gráficos de controle para variáveis

Fonte: Galuch (2002)

Segundo Montgomery (1985) os principais gráficos de controle para atributos são:

- Gráfico p – Registro das frações de itens defeituosos da amostra, isto é, do número de itens defeituosos dividido pelo número total de itens da amostra. A distribuição de probabilidades da fração defeituosa é binomial. É um gráfico utilizado para atributos do tipo classificação
- Gráfico np – Similar ao anterior, este gráfico apresenta o registro do número de itens da amostra que foram classificados como defeituosos.
- Gráfico c – Atributos do tipo contagem devem ser controlados por este tipo de gráfico, que registra o número total de defeitos ou não-conformidades de cada amostra inspecionada. A distribuição de probabilidade é a de Poisson e é necessário que as amostras tenham sempre o mesmo tamanho.
- Gráfico u – Este gráfico registra o número de defeitos dividido pelo número de unidades inspecionadas, ou seja, a taxa de defeitos (atributos do tipo contagem). A distribuição de probabilidade também é de Poisson, mas não é necessário que as amostras tenham sempre o mesmo tamanho.

O Quadro 2.4 apresenta as fórmulas para definição dos limites de controle dos gráficos de controle para atributos. Os termos indicados nas fórmulas possuem o seguinte significado:

n = tamanho da amostra;

m = número de amostras;

c = número total de defeitos em todas as unidades da amostra;

d = número de peças defeituosas.

TIPOS DE GRÁFICOS	FÓRMULAS	
	Linha Média	Limites de Controle
p (proporção de defeituosos)	$\bar{p} = \frac{\sum d}{\sum n}$	$\bar{p} \pm 3 \cdot \sqrt{\frac{\bar{p}(1-\bar{p})}{n}}$
np (número total de defeituosos)	$n \bar{p}$	$n \bar{p} \pm 3 \cdot \sqrt{n \bar{p}(1-\bar{p})}$
c (número de defeitos da amostra)	$\bar{c} = \frac{\sum c}{m}$	$\bar{c} \pm 3 \cdot \sqrt{\bar{c}}$
u (defeitos por unidade)	$\bar{u} = \frac{\sum c}{\sum n}$	$\bar{u} \pm 3 \cdot \sqrt{\frac{\bar{u}}{n}}$

Quadro 2.4 – Fórmulas dos tipos de gráficos de controle para atributos

Fonte: Galuch (2002)

No projeto de um gráfico de controle é necessário especificar o tamanho da amostra a ser utilizada, bem como a frequência de amostragem. Montgomery (1985) destaca que a largura dos limites de controle é inversamente proporcional ao tamanho n da amostra para um determinado múltiplo do valor do desvio-padrão. Uma idéia fundamental no uso de gráficos de controle é coletar dados de acordo com o que Walter A. Shewhart chamou de subgrupo racional. Na visão de Montgomery & Runger (2003) este conceito significa que subgrupos ou amostras devem ser selecionados de modo que a variabilidade das observações, dentro de um subgrupo, inclua toda variabilidade casual ou natural. Desta forma, os limites de controle representarão fronteiras para toda variabilidade natural e as causas especiais tenderão a gerar pontos além dos limites de controle estabelecidos.

As formas mais utilizadas de gráficos de controle são os gráficos \bar{X} -R e os gráficos individuais, sendo que uma importante desvantagem destes gráficos é a sua relativa insensibilidade em detectar pequenas mudanças no processo, como algo da ordem de $1,5\sigma$ ou menos. Segundo Bower (2000) dois tipos de gráficos são primariamente utilizados para detectar pequenos desvios: CUSUM e EWMA.

O gráfico CUSUM (*Cumulative Sum*) ou gráfico da soma cumulativa foi desenvolvido em 1954 por E.S. Page e é construído através do registro da soma cumulativa dos desvios dos valores amostrais em relação a um valor alvo. De acordo com Montgomery & Runger (2003) os gráficos de soma cumulativa são indicados para detectar pequenos desvios, pois combinam informação proveniente de várias amostras. Por este motivo, o gráfico de controle da soma cumulativa é um bom candidato para uso em indústrias químicas e de processos, onde subgrupos racionais são frequentemente de tamanho igual a 1.

Uma técnica alternativa para detectar pequenas mudanças no processo é o uso do gráfico EWMA (*Exponentially Weighted Moving Average*) ou gráfico com média móvel exponencialmente ponderada, desenvolvido por S.W. Roberts em 1959. Bower (2000) afirma que este tipo de gráfico possui algumas propriedades particularmente atrativas, tais como:

- a) Diferentemente dos gráficos de Shewhart, todos os dados coletados ao longo do tempo podem ser usados para determinar a situação de controle do processo;
- b) É frequentemente superior ao gráfico CUSUM para detectar grandes desvios;
- c) Pode ser aplicado para monitorar desvios-padrão em adição à média do processo;
- d) Possibilidade de uso de EWMA na previsão de valores de média do processo;
- e) A metodologia EWMA é insensível à suposição de normalidade.

Escolher a ferramenta adequada para execução das atividades de monitoramento de um processo pode parecer uma tarefa complexa devido à existência de inúmeros gráficos de controle. Contudo, a escolha do tipo de gráfico mais adequado em função dos diferentes tipos de dados pode ser auxiliada por meio do fluxograma apresentado na Figura 2.9.

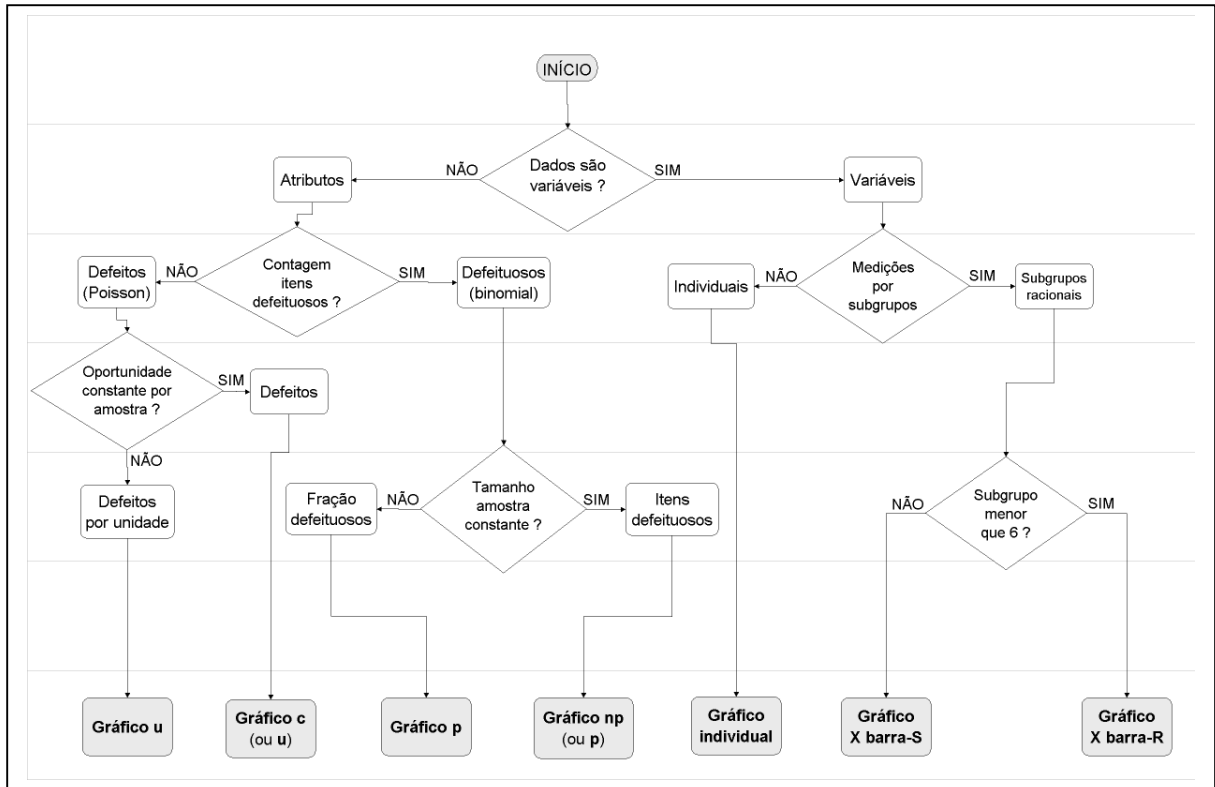


Figura 2.9 – Fluxograma para escolha do tipo de gráfico de controle

Fonte: ASQC/AIAG (1992)

2.4.4 Índices de Capacidade

Os gráficos de controle têm por objetivo avaliar a estabilidade do processo estudado, monitorando seus parâmetros ao longo do tempo; ao passo que o estudo de capacidade visa definir se um processo, cujo comportamento seja conhecido, é capaz de produzir itens ou prestar serviços conforme as especificações estabelecidas pelo cliente. Almas (2003) ressalta que o fato de um processo estar dentro dos limites de controle estatístico não significa que os produtos resultantes atendem às especificações de qualidade exigidas.

Segundo Pires (2000), o estudo da capacidade dos processos é um procedimento que evolui a partir do estudo dos gráficos de controle e, somente após a eliminação das causas

especiais, pode-se avaliar se um processo é capaz de atender as especificações de uma determinada característica da qualidade. Para Kotz & Johnson (2002) os índices de capacidade ou capacidade fornecem uma estimativa sobre o potencial do processo ao invés de expressar a real situação na qual o mesmo se encontra.

Os termos capacidade, capacidade, aptidão e desempenho de processo possuem o mesmo significado segundo Toledo (2005), pois todos representam medidas estatísticas que indicam a taxa de variação existente no processo em relação às especificações dos clientes. Um exemplo de avaliação da capacidade em um processo de fabricação de rotores é representado através da Figura 2.10. Neste exemplo a avaliação é efetuada por meio de comparação do histograma dos valores das medidas de interesse com os seus limites de especificação.

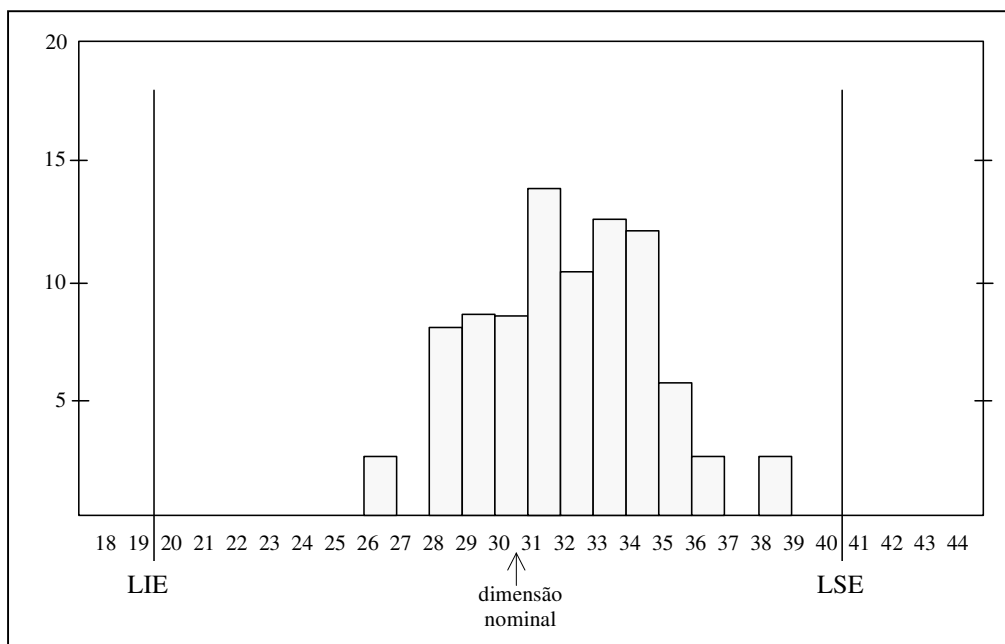


Figura 2.10 – Exemplo de avaliação de capacidade de um processo com histograma

Fonte:Montgomery & Runger (2003)

Montgomery (1985) estabelece que as principais aplicações da análise de capacidade de um processo são as seguintes:

- a) Prever como o processo se manterá dentro das tolerâncias;
- b) Auxiliar na seleção ou modificação do processo;
- c) Auxiliar no estabelecimento de um intervalo entre as amostragens no processo;
- d) Especificar requisitos de desempenho para novos equipamentos;
- e) Auxiliar na escolha entre fornecedores;

- f) Planejar a seqüência dos processos produtivos quando há interação entre estes processos e as tolerâncias;
- g) Reduzir a variabilidade em um processo produtivo;
- h) Aferir o sistema de obtenção de medidas.

De acordo com Roth (2005) os índices de capacidade são calculados a partir de dados do processo em estudo, mas os valores são afetados pelo fato de o processo estar sob controle estatístico e ter um desvio-padrão aumentado. Portanto, os seguintes cuidados devem ser observados para que o cálculo dos índices seja adequado:

- a) O processo deve produzir uma distribuição Normal;
- b) O processo deve estar sob controle estatístico;
- c) O tamanho da amostra deve permitir que o cálculo do desvio-padrão seja racional.

Para avaliar a capacidade do processo é necessário conhecer a distribuição de probabilidade da variável de interesse e estimar a média e a variabilidade dos valores individuais, para então determinar os limites naturais do processo. Se a variável de interesse seguir uma distribuição Normal seus limites naturais podem ser calculados como seis desvios-padrão. Para Roth (2005) os índices de capacidade possibilitam, através de rápida observação, determinar se o processo é capaz de atender as especificações, sendo definidos como uma taxa da amplitude das tolerâncias em relação à dispersão do processo, da seguinte forma (Equação 2.6):

$$\text{índice} = \frac{\text{esperado}}{\text{real}} \quad (2.6)$$

A dispersão do processo pode ser calculada a partir da estimativa do desvio-padrão, que no caso dos gráficos de controle para \bar{X} -R e \bar{X} -S é obtida da seguinte maneira:

$$\hat{\sigma} = \frac{\bar{R}}{d_2} \quad (2.7) \quad \text{ou} \quad \hat{\sigma} = \frac{\bar{S}}{C_4} \quad (2.8)$$

Onde:

$\hat{\sigma}$ = estimativa do desvio-padrão do processo em estudo

\bar{R} = média das amplitudes das amostras em estudo

\bar{S} = média dos desvios-padrão das amostras em estudo

C_4, d_2 = constantes tabeladas que variam em função do tamanho das amostras

Existem diversos índices de capacidade e cada um deles apresenta diferentes informações, por este motivo é de fundamental importância o perfeito entendimento de seu significado de modo a possibilitar correta interpretação e adequada tomada de ações. Os índices de capacidade, fundamentados nas suposições de normalidade dos dados e de controle estatístico do processo são comumente conhecidos como C_P e C_{PK} . Em processos onde as amostras são relativamente pequenas ou para os quais não se tenha certeza de sua estabilidade são utilizados os índices P_P e P_{PK} . A seguir são apresentadas as fórmulas de alguns índices de capacidade de acordo com Roth (2005):

- a) Índice C_P (Equação 2.9) – Fornece uma medida indireta da habilidade potencial do processo em atender às especificações da característica da qualidade de interesse, considerando o processo centrado na média das especificações.

$$C_P = \frac{LSE - LIE}{6\hat{\sigma}} \quad (2.9)$$

Onde:

C_P = índice de capacidade potencial do processo

LSE = limite superior de especificação

LIE = limite inferior de especificação

$\hat{\sigma}$ = estimativa do desvio-padrão do processo

- b) Índice C_{PK} (Equação 2.10) – Fornece uma medida da habilidade real do processo em atender às especificações, pois quantifica a capacidade em função da pior metade dos dados; não considera o processo centrado na média das especificações.

$$C_{PK} = \min. \left[\frac{LSE - \hat{\mu}}{3\hat{\sigma}}, \frac{\hat{\mu} - LIE}{3\hat{\sigma}} \right] \quad (2.10)$$

Onde:

C_{PK} = índice de capacidade real do processo

$\hat{\mu}$ = estimativa da média do processo

- c) Índice P_P (Equação 2.11) – Fornece uma medida do desempenho potencial, mas ao invés da estimativa $\hat{\sigma}$ para o desvio-padrão do processo, utiliza o desvio-padrão S da amostra em estudo, considerando o processo centrado na média das especificações.

$$P_P = \frac{LSE - LIE}{6S} \quad (2.11)$$

Onde:

P_P = índice de desempenho potencial do processo

S = desvio padrão da amostra em estudo

- d) Índice P_{PK} (Equação 2.12) – Fornece uma medida do desempenho real através do desvio-padrão S da amostra em estudo e não considera o processo centrado na média das especificações, pois o índice é calculado em função da pior metade dos dados.

$$P_{PK} = \min. \left[\frac{LSE - \mu}{3S}, \frac{\mu - LIE}{3S} \right] \quad (2.12)$$

Onde:

P_{PK} = índice de desempenho real do processo

O Quadro 2.5 mostra uma regra geral baseada em Werkema (1995) e Ramos (2003), que pode ser utilizada para análise, classificação e interpretação do índice C_P de capacidade ou capacidade do processo a partir de um código de identificação por cores.

ÍNDICE C_P	CLASSIFICAÇÃO	INTERPRETAÇÃO
$C_P < 1,00$	Processo “vermelho” (incapaz)	A capacidade do processo é inadequada à especificação exigida. Nesta situação o responsável pela produção deverá tentar diminuir a variabilidade do processo ou realizar o trabalho em outro processo que atenda as especificações.
$1,00 \leq C_P \leq 1,33$	Processo “amarelo” (razoável)	A capacidade do processo está dentro da especificação exigida. Nesta situação o responsável pela produção deverá tentar diminuir a variabilidade do processo. Gráficos de controle são úteis para manter o processo sob controle estatístico, evitando a produção de unidades não conformes.
$C_P > 1,33$	Processo “verde” (capaz)	A capacidade do processo é adequada à especificação exigida. Nesta situação o responsável pela produção não necessita de maiores preocupações com o processo, a menos que se queira reduzir a variabilidade para aumentar a qualidade dos produtos.

Quadro 2.5 – Classificação de processos a partir do índice C_P

Fonte: Adaptado de Werkema (1995) e Ramos (2003)

2.5 Metodologia Seis Sigma

2.5.1 Definição e Conceitos Básicos

Seis Sigma é uma metodologia de melhoria em processos fundamentada em métodos estatísticos com o objetivo de reduzir os erros através da identificação e eliminação das causas de variação nos diversos processos empresariais. Segundo Linderman *et al.* (2003) Seis Sigma é um método sistemático e organizado para melhoria de processos, novos produtos e desenvolvimento de serviços que se baseia em métodos estatísticos e científicos para efetuar uma drástica redução nas taxas de defeitos definidas pelos clientes. Harry (1998) afirma que a aplicação da metodologia permite às organizações incrementar seus lucros por meio da otimização das operações, melhoria da qualidade e eliminação de defeitos, erros e falhas e, por este motivo, tem sido adotada por diversas empresas ao redor do mundo.

Para Hoerl (2001) a iniciativa de melhoria através da metodologia Seis Sigma tem se tornado extremamente popular nos últimos anos e está conduzindo a uma série de discussões e aplicações relacionadas aos diversos métodos estatísticos. Segundo Rotondaro (2002), as ferramentas utilizadas e os métodos aplicados são bem conhecidos dos profissionais da Qualidade, mas a estruturação do programa e o foco na redução da variabilidade do processo tornam a Metodologia Seis Sigma extremamente eficaz.

De acordo com Harry & Schroeder (2000) o conceito sigma foi criado na década de 1980 como uma forma de desenvolver uma métrica universal de qualidade para mensuração de processos, independentemente de sua complexidade. A metodologia Seis Sigma está associada à redução da variação e o desvio-padrão, representado pela letra σ do alfabeto grego, quantifica a variabilidade em um processo. A Tabela 2.1 mostra uma relação do nível sigma com a proporção de defeitos definida em partes por milhão (PPM).

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

Tabela 2.1 – Referências para o nível sigma
Fonte: Harry (1998)

Quanto maior o valor na escala sigma, menor é a probabilidade de um processo produzir defeitos e, como resultado, há a redução da necessidade de testes e inspeções, os ciclos e os custos diminuem e a satisfação dos clientes aumenta, atingindo assim elevados níveis de excelência operacional e de competitividade.

Segundo Harry (2003) a Metodologia Seis Sigma pode apresentar diversas aplicações práticas e benefícios econômicos em uma empresa, dentre os quais destacam-se:

- a) *Benchmark* e metas de qualidade;
- b) Medida de desempenho;
- c) Filosofia de melhoria;
- d) Caracterização e otimização de parâmetros;
- e) Projeto de novos sistemas;
- f) Simulação de processos e de produção;
- g) Análise de confiabilidade.

2.5.2 Origens da Metodologia Seis Sigma

Conforme relatado por Perez-Wilson (1998), Harry & Schroeder (2000), Harry (2003) e sintetizado por Souza (2002), Reis (2003) e Toledo (2005), a Metodologia Seis Sigma nasceu e se desenvolveu em meados da década de 1980 na empresa Motorola, com o objetivo de torná-la capaz de enfrentar os concorrentes estrangeiros, que estavam fabricando produtos de melhor qualidade com menores custos. A empresa passou a direcionar esforços no sentido de reduzir a variação dos processos de manufatura, administrativos e demais atividades na organização, tanto que, após quatro anos de sua implantação, já havia economizado aproximadamente US\$ 2,2 bilhões (dois bilhões e duzentos milhões de dólares).

Após a divulgação destes primeiros resultados, empresas como Asea Brown Boveri, Allied Signal, General Electric e Sony, também passaram a utilizar a nova metodologia. A Asea Brown Boveri implantou Seis Sigma em sua unidade de negócios de transformadores, conseguindo uma redução de 30% nos custos dos produtos e 68% nos níveis de defeitos, o que resultou em uma economia de US\$ 898 milhões (oitocentos e noventa e oito milhões de dólares) em um período de dois anos. Em seguida, Jack Welch, então CEO (*Chief Executive Office*) da General Electric, aplicou o programa em sua organização com grande sucesso, pois evoluiu de uma abordagem de solução de problemas para uma visão mais ampla de estratégia de negócio, fazendo com que a Metodologia Seis Sigma obtivesse uma repercussão notável em termos mundiais.

2.5.3 Características Gerais

Em um processo que apresenta uma variação de 3σ entre a média e cada um dos limites de especificação, a área sob a curva normal indica uma proporção de itens não conformes de 1350 PPM em cada uma das duas caudas da distribuição. Considerando os dois lados, a fração de itens não conformes é de 2700 PPM. Por outro lado, um processo com 6σ entre a média e cada um dos limites de especificação apresenta 0,001 PPM em cada lado da distribuição, ou seja, a fração total de itens não conformes é da ordem de 0,002 PPM. A Figura 2.11 ilustra esta situação.

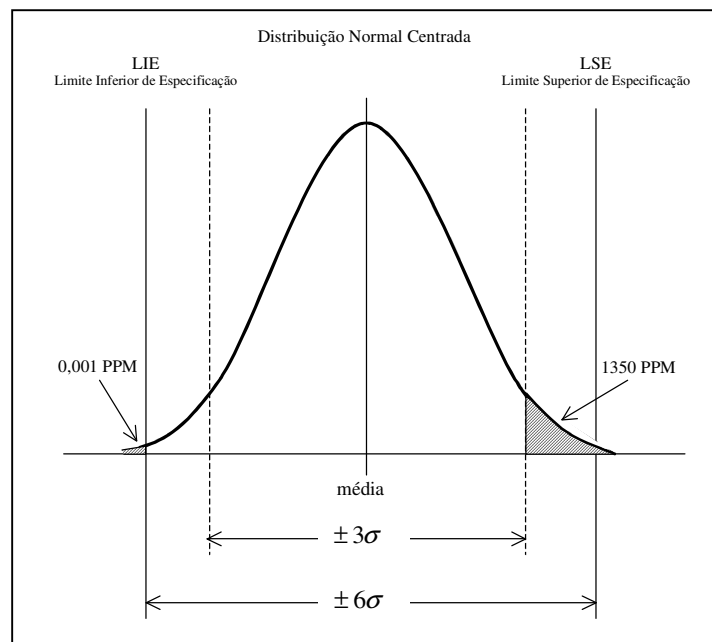


Figura 2.11 – A curva normal e os limites de especificação

Fonte: Adaptado de Reis (2003)

Entretanto, dificilmente um processo permanece centralizado ao longo do tempo e é esperado que apresente um desvio (*shift*) em relação ao valor nominal. A empresa Motorola adotou mais ou menos $1,5\sigma$ como valor representativo para este desvio. Com isso, um processo operando com a Metodologia Seis Sigma apresenta 3,4 DPMO (defeitos por milhão de oportunidades). Segundo Reis (2003) a maior parte das empresas adota o critério estabelecido pela Motorola, embora existam questionamentos em relação ao valor de $1,5\sigma$. De acordo com Harry (2003) o desvio de 1,5 sigma está associado a uma correção estatisticamente embasada para compensar a degradação do desempenho a longo prazo, atribuída unicamente à influência de erros aleatórios inerentes ao processo. A Figura 2.12 apresenta o deslocamento da média do processo em relação à curva normal.

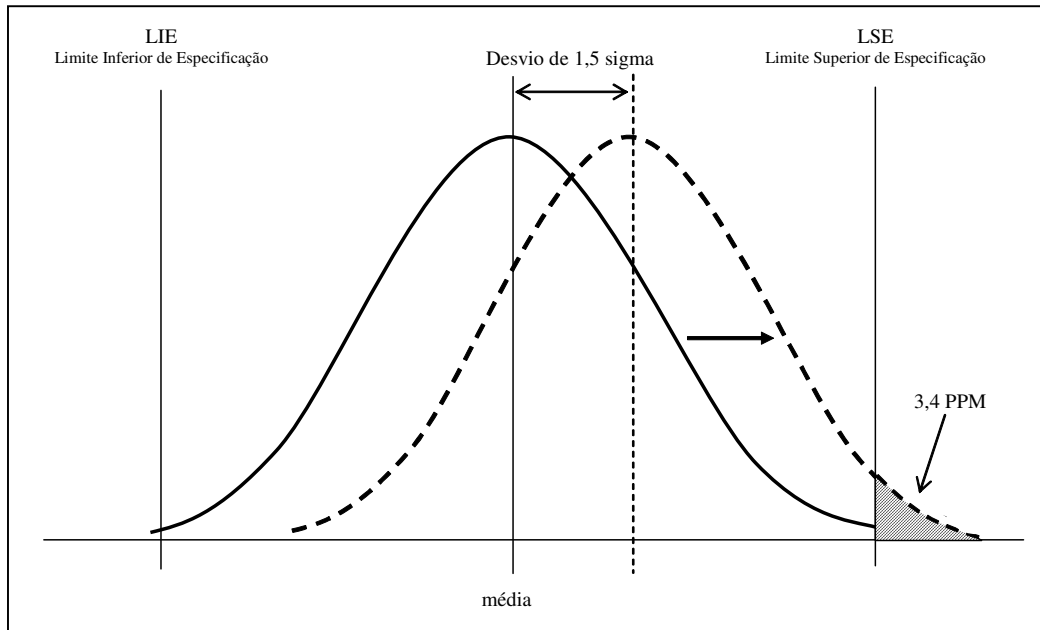


Figura 2.12 – Deslocamento da média do processo em 1,5 sigma

Fonte: Adaptado de Montgomery & Runger (2003)

A quantidade de itens além dos limites de especificação para alguns valores de desvios-padrão em um processo centralizado, assim como em um processo não centralizado, com desvio de 1,5 sigma, é mostrada na Tabela 2.2.

NÍVEL SIGMA	PPM (partes por milhão)	
	Sem Shift	Com Shift
± 1 sigma	317300	697700
± 2 sigma	45500	308700
± 3 sigma	2700	66810
± 4 sigma	63	6210
± 5 sigma	0,57	233
± 6 sigma	0,002	3,4

Tabela 2.2 – Referências entre nível sigma e PPM (com e sem desvio)

Fonte: Adaptado de Reis (2003)

A Metodologia Seis Sigma é baseada em um sistema de acompanhamento conhecido como DMAIC, sigla que denota as seguintes etapas: Definir (*Define*), Medir (*Measure*), Analisar (*Analyze*), Melhorar (*Improve*) e Controlar (*Control*). O Quadro 2.6 apresenta uma visão geral dos passos associados ao modelo DMAIC.

ETAPA	DESCRIÇÃO
DEFINIÇÃO	Identificar CTQs (itens críticos para qualidade na visão do cliente)
	Desenvolver escopo de atuação da equipe de trabalho
	Definir mapa do processo
MEDIÇÃO	Selecionar característica do CTQ
	Definir padrão de desempenho
	Analisar o sistema de medição e coleta de dados
ANÁLISE	Estabelecer a capacidade do processo
	Definir objetivos de desempenho
	Identificar origens de variação
MELHORIA	Filtrar causas potenciais de variação
	Descobrir relações entre as variáveis e propor soluções
	Estabelecer tolerâncias operacionais e solução piloto
CONTROLE	Validar o sistema de medição
	Determinar a capacidade do processo
	Implementar sistema de controle do processo

Quadro 2.6 – Visão geral do DMAIC

Fonte: Adaptado de Harry (1998)

Para Rotondaro (2002) a abordagem DMAIC é na realidade uma evolução do método introduzido por W.E. Deming, o ciclo PDCA (*Plan, Do, Check, Act*), que é um dos exemplos mais populares de metodologia que tem sido utilizada para a melhoria dos processos.

A Metodologia Seis Sigma necessita de pessoal especializado para sua aplicação e, segundo Harry (1998), os principais papéis são desempenhados pelos seguintes elementos:

- a) Comitê de Liderança;
- b) Patrocinadores (*Champions*);
- c) Especialista Máster (*Master Black Belt*);
- d) Especialista em Seis Sigma (*Black Belt*);
- e) Participantes das Equipes (*Green Belts*).

Os termos *Black Belt* e *Green Belt* são uma analogia aos especialistas em artes marciais. As principais atribuições deste grupo incluem:

- a) Estabelecimento das funções nos programas;
- b) Definição e fornecimento de infra-estrutura e recursos;
- c) Seleção de projetos específicos;
- d) Revisão regular do andamento dos projetos;
- e) Atuação como patrocinadores dos projetos;
- f) Remoção de obstáculos ao trabalho dos grupos.

O Quadro 2.7 apresenta a comparação dos papéis na aplicação da Metodologia Seis Sigma, assim como uma apresentação sucinta das necessidades de treinamento.

	Champion	Master Black Belt	Black Belt	Green Belt
Qualificações	Familiaridade com ferramentas estatísticas básicas e avançadas. Executivos seniores e gerentes.	Recomendável formação técnica. Domínio de ferramentas estatísticas básicas e avançadas. Um MBB poderia ser, por exemplo, um gerente ou engenheiro chefe.	Recomendável formação ou orientação técnica. Domínio de ferramentas estatísticas básicas. Um BB poderia ser engenheiro ou profissional com 5 ou mais anos de experiência.	Base e suporte técnico. Familiaridade com as ferramentas estatísticas básicas. Sua posição deve estar associada com o problema que está sendo resolvido.
Treinamento	3 a 5 dias de treinamento específico	Aproximadamente 200 horas de treinamento e desenvolvimento de projetos	Aproximadamente 160 horas de treinamento e desenvolvimento de projetos	Aproximadamente 80 horas de treinamento e desenvolvimento de projetos
Número de Pessoas Treinadas	1 Champion por unidade de negócio	1 MBB para cada 20-30 BB	1 BB para cada 50-100 pessoas	1 GB para cada 10-20 pessoas

Quadro 2.7 – Papéis desempenhados na Metodologia Seis Sigma

Fonte: Adaptado de Harry & Schroeder (2000)

2.5.4 Determinação do Nível Sigma

Como forma de estimar o desempenho das atividades desenvolvidas, as organizações estabelecem medidores de desempenho claros, tais como partes por milhão (PPM), índices de capacidade de processo, custos da qualidade, entre outros. De acordo com Carvalho (2002) o índice de capacidade seis sigma apresenta algumas diferenças em relação às métricas tradicionais e, dentre as métricas mais utilizadas, as que mais se assemelham ao referido índice são C_{PK} e P_{PK} (Equações 2.10 e 2.12).

Em um processo cuja média esteja equidistante dos limites de especificação e o índice de capacidade seja igual a 1 ($C_{PK} = 1$), os limites de especificação distam 3 desvios-padrão da média. Caso a média não esteja centralizada em relação aos limites de especificação, adota-se a pior capacidade, ou seja, utiliza-se o limite de especificação mais próximo. Para um processo com padrão seis sigma o índice C_{PK} deve ser igual a 2, isto é, o processo é considerado capaz se a sua média estiver a 6 desvios-padrão dos limites de especificação.

O índice utilizado para determinar a capacidade seis sigma mede a distância da média à especificação mais próxima (LIE ou LSE) em quantidades de desvios-padrão (sigmas), utilizando como referência a curva normal padronizada $N(0;1)$. Através de uma tabela de distribuição Normal Padrão e conforme mostrado na Figura 2.11, pode-se verificar que o nível seis sigma corresponde a aproximadamente 0,002 PPM, ou seja, 2 (dois) defeitos por bilhão de oportunidades. A quantidade de 3,4 partes por milhão, definida como padrão seis sigma, relaciona-se ao valor z igual a 4,5. Esta diferença de 1,5 sigma, segundo estudos da Motorola relatados por Harry (2003), advém da variação e mudanças do processo ao longo do tempo conforme anteriormente representado através da Figura 2.12.

De acordo com resumo efetuado por Almas (2003), os dados coletados em um intervalo de tempo pequeno e, portanto, sem causas especiais de variação, representam o índice z_{ST} de curto prazo (*short term*). Os dados obtidos em intervalos de tempo maiores, com causas comuns e especiais de variação, equivalem ao índice z_{LT} de longo prazo (*long term*). Como no índice de curto prazo os dados utilizados não contêm causas especiais de variação, normalmente seu valor é maior que os índices utilizados com os dados de longo prazo.

A referência para a determinação do nível sigma é o índice conhecido como z_{bench} que é obtido através da verificação da proporção de defeitos ou partes por milhão de acordo com os limites de especificação inferior e superior da distribuição ou do processo em estudo (Equações 2.13 e 2.14), que somadas resultam na proporção total de defeitos (Equação 2.15).

$$z_{LIE} = \frac{\bar{X} - LIE}{S} \Rightarrow PPM_{LIE} \quad (2.13)$$

$$z_{LSE} = \frac{LSE - \bar{X}}{S} \Rightarrow PPM_{LSE} \quad (2.14)$$

$$PPM_{TOTAL} = PPM_{LIE} + PPM_{LSE} \quad (2.15)$$

Onde:

z_{LIE} = coordenada correspondente ao limite inferior de especificação da distribuição em estudo

z_{LSE} = coordenada correspondente ao limite superior de especificação da distribuição em estudo

\bar{X} = média da amostra do processo em estudo

PPM_{LIE} = fração não conforme abaixo do limite inferior de especificação (partes por milhão)

PPM_{LSE} = fração não conforme acima do limite superior de especificação (partes por milhão)

PPM_{TOTAL} = fração não conforme total em partes por milhão

Com a proporção total de defeitos do processo em estudo obtém-se, através da tabela de distribuição Normal Padrão, a estatística z_{bench} , que corresponde ao escore padronizado da curva normal reduzida $N(0;1)$. Esta estatística equivale ao índice z_{LT} , pois refere-se ao z_{bench} de longo prazo conforme mencionado anteriormente. Em seguida, obtém-se o valor do índice de curto prazo z_{ST} , que é a capacidade potencial do processo conforme Equação 2.16.

$$Z_{bench(ST)} = Z_{bench(LT)} + Z_{shift} \tag{2.16}$$

Onde:

$Z_{bench(ST)}$ = índice de curto prazo equivalente ao nível sigma

$Z_{bench(LT)}$ = índice de longo prazo

Z_{shift} = desvio da média do processo ao longo do tempo

Adotando-se o valor de 1,5 sigma para Z_{shift} , a Equação 2.16 também pode ser expressa de modo simplificado de acordo com Equação 2.17:

$$\text{Nível Sigma} = Z_{bench(LT)} + 1,5 \tag{2.17}$$

Através da Figura 2.13 é apresentado um esquema extremamente simplificado que possibilita a visualização e compreensão da lógica utilizada na determinação do índice ou nível sigma de capacidade.

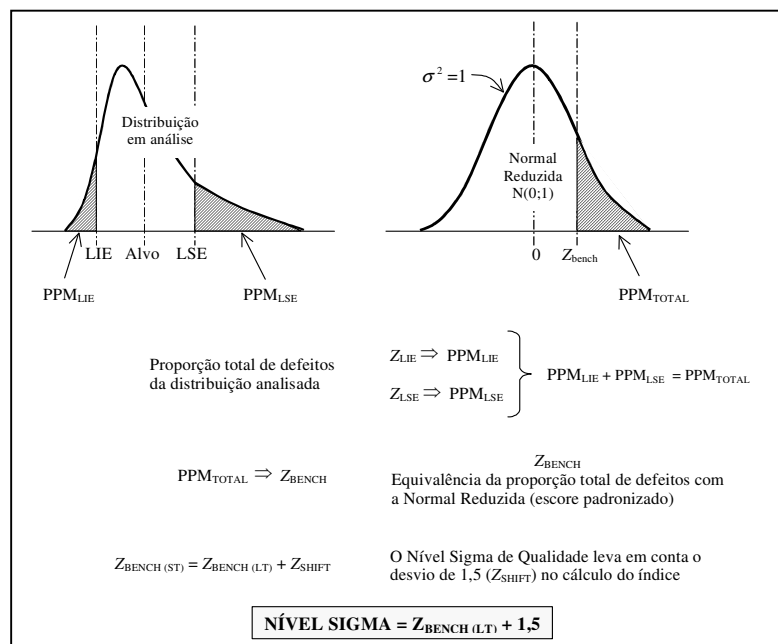


Figura 2.13 – Esquema simplificado para cálculo do nível sigma

Como método alternativo as seguintes aproximações, indicadas através das Equações 2.18 e 2.19, ainda podem ser utilizadas no cálculo do nível sigma (BREYFOGLE, 1999):

$$\text{Nível Sigma} \approx 3.P_{PK} + 1,5 \quad (2.18)$$

$$\text{Nível Sigma} \approx 0,8406 + \sqrt{29,37 - 2,221 \cdot \ln(ppm)} \quad (2.19)$$

A relação entre nível sigma e a proporção de defeitos em partes por milhão (PPM) é definida em função da curva normal reduzida $N(0;1)$, podendo ser obtida através das tabelas da distribuição Normal Padrão (Anexo A), disponíveis em livros de Estatística e Controle Estatístico de Processo ou ainda por meio de *softwares* estatísticos. Uma tabela de conversão em PPM para alguns valores de z_{bench} está disponível no Anexo C desta dissertação.

Para o cálculo do nível sigma nos casos onde a variável em estudo é do tipo atributo, percorre-se o caminho inverso, isto é, a partir da proporção de defeitos determinam-se os índices de longo e curto prazo respectivamente. Segundo Carvalho (2002) a empresa Motorola introduziu uma forma de ajustar as medidas em função da complexidade do que está sendo avaliado, ou seja, o número de defeitos por milhão de oportunidades (DPMO). Essa abordagem permite que sejam comparados os desempenhos de diferentes produtos e serviços, fato que amplia o leque de possibilidades de utilização da Metodologia Seis Sigma. Uma síntese dos conceitos básicos para atributos é mostrada no Quadro 2.8.

CONCEITO	DEFINIÇÃO
Defeito	Qualquer não conformidade às especificações
Defeituoso	Unidade que apresenta um ou mais defeitos
Unidade	Saída do processo que será avaliada segundo a presença de defeitos
Oportunidade	Formas nas quais o processo pode se desviar do que é especificado para cada unidade, gerando não conformidade
Defeitos por unidade	$DPU = \frac{\text{número de defeitos}}{\text{número de unidades}}$
Defeitos por oportunidades	$DPO = \frac{\text{número de defeitos}}{\text{número de oportunidades} \times \text{número de unidades}}$
Defeitos por milhão de oportunidades	$DPMO = \frac{\text{número de defeitos}}{\text{número de oportunidades} \times \text{número de unidades}} \times 10^6$

Quadro 2.8 – Definição de conceitos para atributos na Metodologia Seis Sigma

Fonte: Adaptado de Carvalho (2002)

Um fluxograma geral resumido para determinação do índice de capacidade seis sigma é apresentado através da Figura 2.14.

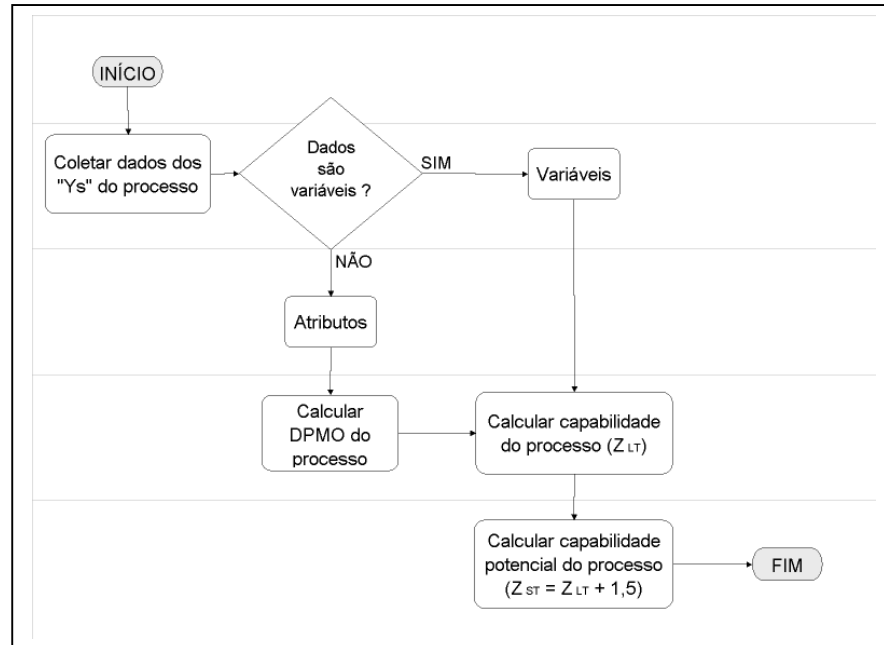


Figura 2.14 – Fluxograma geral para determinação do índice de capacidade seis sigma

Fonte: Adaptado de Carvalho (2002)

3. ESTUDO DA NÃO NORMALIDADE

3.1 Análise de Normalidade

A distribuição Normal ou Gaussiana é importante em diversos estudos, pois este tipo de distribuição representa de maneira bem aproximada o comportamento de várias situações reais. Além disso, a distribuição de muitos dos testes estatísticos é normal ou apresenta uma forma derivada da distribuição Normal. Soares (2003) relembra que há alguns anos atrás existiam estatísticos que afirmavam com convicção que se um processo não fosse distribuído normalmente, poderia haver algo de errado ou o mesmo estaria fora de controle. Sob este ponto de vista, o gráfico de controle seria apenas uma ferramenta para determinar a não normalidade de um processo com vistas a possibilitar sua correção. Atualmente reconhece-se que o uso dos gráficos de controle provém, principalmente, de sua simplicidade.

A normalidade pode ser verificada na forma gráfica ou por meio de técnicas estatísticas como teste de hipóteses. Segundo Miranda (2005) a forma gráfica pode ser, por exemplo, através de histograma pela observação da existência de assimetria dos dados, ou gráfico normal de probabilidade pela comparação do valor esperado nominal com o valor ordenado das observações, esperando-se uma relação linear quando existir normalidade.

Para que haja um completo entendimento da verificação de normalidade efetuada por meio de testes estatísticos é necessário estabelecer alguns conceitos que os sustentam, tais como, amostragem, intervalo de confiança, teste de hipóteses, nível de significância estatística e valor P (*p-value*).

✓ AMOSTRAGEM:

De acordo com Prazeres (1997) uma amostra é uma parte ou porção representativa de um conjunto ou população a ser medida, analisada ou ensaiada. Um dos principais objetivos da amostragem é a redução de custos na determinação de algum parâmetro da população que se deseja analisar. Em processos produtivos é possível determinar tais parâmetros por meio de algumas amostras e o risco de errar, que depende do tamanho da amostra, também tem um custo associado e pode ser calculado.

✓ INTERVALO DE CONFIANÇA:

Quando se deseja estimar parâmetros de uma população a partir de amostras, não se tem certeza sobre o verdadeiro valor do parâmetro em questão. O intervalo de confiança de uma determinada estatística da amostra indica a amplitude ou faixa de valores na qual o verdadeiro valor do parâmetro está contido com certa probabilidade.

✓ TESTE DE HIPÓTESES:

Conforme definição de Prazeres (1997) o teste de hipóteses é um procedimento estatístico para testar hipóteses sobre um parâmetro populacional a partir de uma distribuição amostral conhecida. Consideram-se duas hipóteses: a hipótese nula (H_0), que é a suposição que se pretende testar, e a hipótese alternativa (H_1). Os testes podem ser unilaterais ou bilaterais de acordo com a formulação da hipótese alternativa. Existem dois erros que podem ser cometidos quando se testa uma hipótese: erro tipo I, que ocorre quando se rejeita uma hipótese nula sendo a mesma verdadeira, e o erro tipo II quando se aceita a hipótese nula sendo ela falsa.

✓ NÍVEL DE SIGNIFICÂNCIA ESTATÍSTICA:

O nível de significância estatística de um resultado é uma medida estimada do grau no qual este resultado é verdadeiro, no sentido de ser representativo da população. Este parâmetro representa o índice decrescente de confiabilidade de um resultado, isto é, quanto maior o nível de significância, menor a possibilidade de que a relação observada entre variáveis da amostra seja um indicador confiável da relação entre as respectivas variáveis da população. Segundo Montgomery & Runger (2003) a probabilidade α de cometer o erro tipo I também é chamada de nível de significância ou tamanho do teste. Um nível de significância igual a 0,05 indica que há uma probabilidade de 5% de que a relação entre variáveis encontrada na amostra seja apenas coincidência ou acidental. Em muitas áreas de pesquisa utiliza-se o nível de significância estatística de 0,05 como um limite de erro aceitável.

✓ VALOR P OU *P-VALUE*:

De acordo com Montgomery & Runger (2003) o valor P ou *p-value* é o menor nível de significância que conduz à rejeição da hipótese nula com os dados fornecidos. Em outras palavras, *p-value* é a probabilidade de que a estatística de teste assuma um valor, no mínimo, tão extremo quanto o valor observado da estatística para que a hipótese nula seja considerada verdadeira. Se z_0 for o valor calculado da estatística de teste, o valor P pode ser definido conforme Equação 3.1:

$$P = \begin{cases} 2 \cdot [1 - \Phi(|z_0|)] & (\text{teste bilateral}) \\ 1 - \Phi(z_0) & (\text{teste unilateral superior}) \\ \Phi(z_0) & (\text{teste unilateral inferior}) \end{cases} \quad (3.1)$$

Onde:

$\Phi(z)$ = função de distribuição cumulativa de uma variável aleatória normal padrão

Existem diversos testes de normalidade disponíveis, que são conduzidos por meio da avaliação de determinados parâmetros específicos. Os testes de Anderson-Darling, Kolmogorov-Smirnov, Shapiro-Wilk e Ryan-Joiner são exemplos de testes que podem ser utilizados para confirmação de normalidade. No desenvolvimento da pesquisa que originou a presente dissertação foi utilizada a opção de teste de normalidade Anderson-Darling através do *software* estatístico Minitab®.

Pelo procedimento de teste de Anderson-Darling efetua-se a comparação de uma função de distribuição cumulativa observada com uma função de distribuição cumulativa esperada. Rejeita-se a hipótese nula de que a distribuição seja de uma forma específica se a estatística de teste A^2 for maior do que um valor crítico; sendo que estes valores críticos são tabelados para determinados tamanhos de amostra.

As análises de normalidade discutidas durante a pesquisa em questão foram efetuadas considerando o conceito de significância estatística no teste de Anderson-Darling da seguinte maneira:

H_0 (hipótese nula):	Os dados seguem a distribuição normal.
H_1 (hipótese alternativa):	Os dados não seguem a distribuição normal.

A abordagem utilizada nesta pesquisa compara o valor P (*p-value*) associado à estatística de teste A^2 com um nível de significância α arbitrado para aceitação da hipótese nula de normalidade. Em todas as análises efetuadas foi adotado $\alpha = 0,05$, ou seja, ao longo desta dissertação, sempre que *p-value* apresentar valor maior que 0,05 a distribuição será considerada normal.

3.2 Efeitos da Não Normalidade

Os métodos estatísticos discutidos nesta dissertação pressupõem que os dados em estudo sigam uma distribuição de probabilidade conhecida. A análise e as conclusões que resultam da aplicação da metodologia são válidas apenas nos casos onde a suposição da distribuição se confirme verdadeira.

Castagliola & Tsung (2005) ressaltam que devido às modernas tecnologias de medição e inspeção é bastante comum uma coleta rotineira de grande quantidade de dados de unidades individuais em um intervalo de tempo muito curto; e tais medições contínuas podem resultar em dados que tendem a apresentar auto-correlação e distribuição não normal. Lipson & Sheth (1973) indicam que na falta de alguma evidência em contrário, pode-se assumir, em uma primeira abordagem, que os dados utilizados em métodos estatísticos sejam normalmente

distribuídos. Contudo, Montgomery (1985) alerta que em muitas situações práticas existem razões para se duvidar da validade da suposição de normalidade, o que implica em especial atenção na análise dos dados. De acordo com Spedding & Rawlings (1994), em produção contínua a suposição de normalidade é freqüentemente justificada, mas a distribuição da variação do processo é mais crítica quando são considerados pequenos lotes de fabricação.

Segundo Alwan (2001) muitos dos procedimentos estatísticos são derivados de um conjunto de suposições específicas acerca da origem dos dados e, por este motivo, diversas investigações são conduzidas para determinar quão sensíveis são as conclusões tiradas em situações de desvio de tais premissas, pois as conseqüências da violação destas premissas nas inferências efetuadas podem ser sérias.

Em sentido oposto, Sall (2005) recomenda bom senso, pois não há motivo para o que ele chama de medo irracional da não normalidade, visto que em grandes amostras a sua detecção é simples, mas neste caso sua influência não é tão danosa e em pequenas amostras a não normalidade pode ser significativa, entretanto não se pode detectá-la facilmente.

Visto que existem situações onde a distribuição de dados não corresponde a uma distribuição Normal, então para estes casos tanto a suposição de normalidade quanto utilização da curva normal como referência certamente se revelarão inadequadas. Na realidade a existência de não normalidade em distribuição de dados é bastante comum, principalmente quando o número de observações não é muito grande. Conforme apresentado no tópico 1.2, a suposição incorreta de normalidade dos dados pode acarretar algumas das seguintes conseqüências:

- a) Falsos alarmes na análise de gráficos de controle individuais;
- b) Cálculo incorreto de índices de capacidade;
- c) Conclusões incorretas sobre diferenças entre grupos em Teste de Hipóteses;
- d) Erros em predições na Análise de Regressão;
- e) Conclusões incorretas sobre importância e efeito de fatores em Planejamento de Experimentos (DOE).

Para Castagliola & Tsung (2005), muitas das técnicas existentes em Controle Estatístico de Processo não foram concebidas para atender tais situações, visto que as técnicas de CEP são afetadas por dados auto-correlacionados e assimétricos. Sob o ponto de vista destes autores, muitos estudos foram efetuados para investigar os efeitos da não normalidade dos dados e, mesmo que as conclusões apresentem algumas contradições de acordo com o

ponto de vista adotado, parece muito claro que os impactos de assimetria no desempenho dos gráficos de controle podem ser substanciais.

Por outro lado, Borrer *et al.* (1999) alegam que diversos autores defendem o argumento de que a não normalidade na aplicação de gráficos de controle não é tão relevante, quando se trabalha com amostras utilizando subgrupos de tamanho $n > 1$. Uma observação importante de um estudo efetuado por Alwan (2001) indica que o efeito da não normalidade diminui à medida que aumenta o tamanho da amostra por grupo e este efeito aumenta com o aumento do número de grupos ou subgrupos.

Com relação à avaliação de capacidade, Kotz & Johnson (2002) alertam que a ausência de normalidade indica necessidade de muita atenção no uso dos índices para julgamento adequado da qualidade do processo em estudo. Como exemplo, Miranda (2005) cita um estudo efetuado por M. Deleryd no qual são apontados alguns efeitos de distribuições não normais na estimação do número de defeituosos, considerando três processos com distribuições diferentes em comparação com a distribuição normal. A Tabela 3.1 indica que as quantidades de itens não conformes nas três distribuições são bem diferentes entre si e em relação à distribuição normal. Portanto, caso a normalidade fosse considerada nos três primeiros processos, ocorreria um erro grosseiro na determinação da quantidade de itens defeituosos e conseqüentemente dos respectivos índices de capacidade.

TIPO DE DISTRIBUIÇÃO	FRAÇÃO NÃO CONFORME (considerando limites de $\pm 3\sigma$)
Chi-Quadrado com 4,5 graus de liberdade	14000 PPM
t-Student com 8 graus de liberdade	4000 PPM
Uniforme	0
Normal	2700 PPM

Tabela 3.1 – Comparação da proporção de itens não conformes em distribuições diferentes

Fonte: Baseado em Miranda (2005)

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

3.3 Não Normalidade, Gráficos de Controle e Capabilidade

Conforme argumentação enfatizada no tópico 2.4.3, a aplicação de CEP pressupõe que as observações do processo em estudo sejam independentes e identicamente distribuídas, além de seguirem uma distribuição de probabilidade específica. Juran & Gryna (1992) estabeleceram alguns critérios para abordagem dos dados em caso de não normalidade:

- Examinar os dados para verificar se há alguma explicação não estatística para o padrão distribucional não convencional;
- Analisar os dados em termos de médias ao invés de valores individuais, pois médias de amostra seguem de perto uma distribuição de probabilidade normal, mesmo quando a população de valores individuais não é distribuída normalmente (Teorema Central do Limite);
- Utilizar como referência outro tipo de distribuição que se enquadre mais adequadamente ao conjunto de dados coletados;
- Efetuar transformação matemática da característica original para uma nova característica que se aproxime de uma distribuição Normal.

Quando os dados seguem uma distribuição não normal conhecida, as análises devem ser efetuadas considerando as propriedades características deste conjunto de dados. Uma alternativa para a construção de gráficos de controle nestes casos é demonstrada através da Figura 3.1.

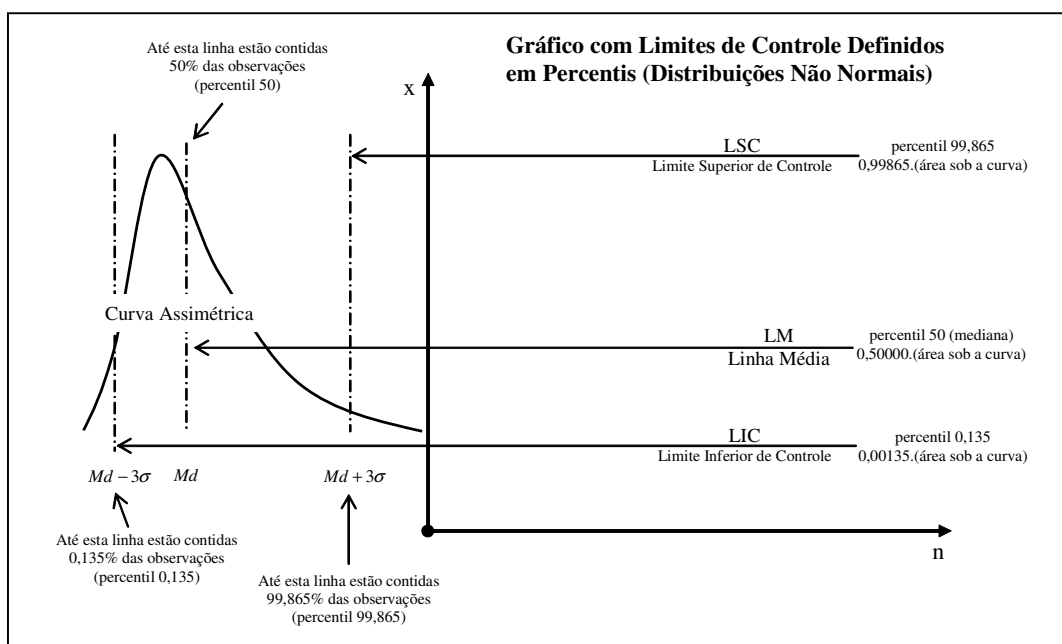


Figura 3.1 – Limites de controle em termos de percentis para distribuições não normais

Com base em Levinson (2000), conhecendo-se a função densidade de probabilidade da distribuição não normal em questão, é possível determinar o limite inferior de controle, a linha média e o limite superior de controle através dos valores coletados correspondentes aos percentis 0,135, 50 e 99,865, respectivamente, da seguinte forma:

$$LIC = \mu - 3\sigma \text{ (corresponde a 0,135\% da área sob a curva da distribuição)}$$

$$LM = \mu \text{ (corresponde a 50\% da área sob a curva da distribuição)}$$

$$LSC = \mu + 3\sigma \text{ (corresponde a 99,865\% da área sob a curva da distribuição)}$$

Com relação ao cálculo de capacidade a mesma abordagem pode ser adotada quando a distribuição não normal em estudo é bem conhecida, determinando-se a taxa de defeituosos por meio dos parâmetros da distribuição em questão. De acordo com Miranda (2005), o seguinte método de cálculo do índice de capacidade através de percentis desenvolvido por J.A. Clements em 1989 pode ser convenientemente aplicado (Equações 3.2 e 3.3):

$$C_{P(q)} = \frac{LSE - LIE}{x_{0,99865} - x_{0,00135}} \quad (3.2)$$

$$C_{PK(q)} = \min. \left[\frac{LSE - x_{0,50000}}{x_{0,99865} - x_{0,50000}}, \frac{x_{0,50000} - LIE}{x_{0,50000} - x_{0,00135}} \right] \quad (3.3)$$

Onde:

$C_{P(q)}$ = índice de capacidade potencial do processo com percentis

$C_{PK(q)}$ = índice de capacidade real do processo com percentis

LSE = limite superior de especificação

LIE = limite inferior de especificação

$x_{0,00135}$ = percentil 0,135 correspondente ao valor $\mu - 3\sigma$ da distribuição normal

$x_{0,50000}$ = percentil 50 correspondente à média μ da distribuição normal

$x_{0,99865}$ = percentil 99,865 correspondente ao valor $\mu + 3\sigma$ da distribuição normal

$x_{0,99865} - x_{0,00135}$ = amplitude correspondente ao valor 6σ da distribuição normal

Uma outra possibilidade de avaliação em processos definidos por dados modelados por distribuições não normais, a transformação matemática das variáveis, é apresentada com maiores detalhes no próximo tópico.

3.4 Transformação de Dados

3.4.1 Aspectos Gerais

As técnicas de transformação de variáveis dependentes e independentes em modelos estatísticos podem ser úteis para melhorar a interpretação dos dados e a sua aderência em modelos hipotéticos. Box & Tidwell (1962) afirmam que para representar as relações entre uma resposta e um número de variáveis independentes, é preferível sempre que possível, trabalhar com um modelo simples com variáveis transformadas do que com modelos mais complexos na forma original das variáveis. Segundo estes autores muita pesquisa experimental diz respeito ao estudo de uma função resposta do tipo $E(y) = f(x, \theta)$, onde x representa o nível das variáveis e θ , o conjunto de parâmetros que afetam a resposta. Nestes casos o objetivo usualmente é:

- a) Verificar a adequação da forma funcional assumida;
- b) Estimar os valores dos parâmetros θ e, conseqüentemente a resposta $E(y)$;
- c) Obter medidas de precisão das estimativas.

Supondo que n observações sejam efetuadas; então a u -ésima observação y_u em um nível conhecido das variáveis $x_u = x_{u1}, x_{u2}, \dots, x_{uk}$ é dado por meio da Equação 3.4:

$$y_u = f(x_u, \theta) + \varepsilon_u \quad (3.4)$$

Onde:

θ = conjunto de p parâmetros $\theta_1, \theta_2, \dots, \theta_p$ que afetam a função resposta

ε_u = erros observados na função resposta da variável dependente

A premissa básica de uma transformação de variáveis independentes é que os erros ε_u na variável dependente sejam normalmente e independentemente distribuídos com variância σ^2 constante, de modo a converter a função destas variáveis transformadas em uma forma mais simples quanto possível. A transformação de variáveis dependentes é freqüentemente aplicada para encontrar as suposições necessárias para certos modelos nos quais os resíduos sejam normalmente distribuídos com variância constante. Para ilustrar, se em uma distribuição a dispersão dos resíduos aumenta com valores maiores da variável dependente, uma transformação com raiz-quadrada da variável dependente estabilizaria a variância e, além disso, se necessário, poderia ao mesmo tempo normalizar os resíduos.

Embora modelos de regressão usualmente não necessitem de suposições acerca da distribuição de variáveis independentes, uma variável independente pode ser transformada para melhor adequação ou para tornar esta distribuição mais simétrica e aliviar a excessiva influência de um valor fora do padrão (*outlier*). As transformações podem ser aplicadas às variáveis dependentes e/ou independentes para converter um modelo de regressão não linear em linear. Por exemplo, uma transformação logarítmica da variável dependente mudaria um modelo de crescimento exponencial em uma regressão linear, simplificando, deste modo, os procedimentos de estimação.

Como as transformações têm múltiplos efeitos, seu uso requer alguns cuidados. Por exemplo, quando se transforma uma variável para encontrar a suposição de normalidade e variância constante dos resíduos, pode-se omitir o fato de que esta transformação poderia alterar as relações existentes entre as variáveis originais. Uma análise de resíduos poderia ser usada extensivamente para avaliar a adequação da transformação específica. As transformações também afetam os coeficientes dos modelos obtidos. Se desejado, para algumas transformações, uma função inversa pode ser aplicada para obter coeficientes que permitam interpretação nas unidades originais.

De acordo com Chen *et al.* (2003) muitos testes estatísticos são baseados na suposição de normalidade; quando os dados se afastam desta condição, uma transformação adequada pode freqüentemente produzir um conjunto de dados que sigam de forma aproximada a distribuição Normal. Para Box & Tidwell (1962) um dos mais simples tipos de transformação que pode ser empregado em diversas aplicações é a transformação de potência (*power transformation*), que inclui muitas das formas comumente utilizadas, tais como, transformação inversa e transformação com raiz quadrada. No tópico seguinte é abordado um tipo muito útil de transformação de potência, a transformação de Box-Cox.

3.4.2 Transformação de Box-Cox

As técnicas para análise de modelos lineares, exemplificados pela Análise de Variância e pela Análise de Regressão, são baseadas nas seguintes premissas de acordo com Box & Cox (1964):

- a) Simplicidade da estrutura para $E(y)$;
- b) Variância constante;
- c) Normalidade da distribuição;
- d) Independência das observações.

Conforme previamente comentado, a estratégia para fazer com que dados não normais se aproximem de uma distribuição Normal através de transformação das variáveis em estudo também deve ser considerada. Entretanto a escolha do tipo adequado de transformação não parece ser uma tarefa óbvia, pois em termos matemáticos existem inúmeras possibilidades e apenas o método de “tentativa e erro” nem sempre é o mais recomendado. A transformação linear, por exemplo, altera a escala da distribuição, mas em alguns casos não altera sua forma; já a transformação de potência é mais eficiente para este propósito.

Box & Cox (1964) efetuaram um estudo detalhado em análise de dados, com ênfase na transformação de variáveis dependentes, onde a idéia principal era dedicar atenção em transformações indexadas por um parâmetro λ desconhecido e a partir de então estimar os outros parâmetros do modelo através de métodos convencionais de inferência. O resultado mais importante deste estudo foi a definição da seguinte família de transformação de potência da variável y para $y^{(\lambda)}$ (Equação 3.5):

$$y^{(\lambda)} = \begin{cases} \frac{y^\lambda - 1}{\lambda} & (\lambda \neq 0) \\ \ln y & (\lambda = 0) \end{cases} \quad (3.5)$$

Onde:

$y^{(\lambda)}$ = variável dependente após a transformação

y = variável dependente antes da transformação

λ = parâmetro que define a transformação

Segundo Box & Cox (1964) esta transformação é definida somente para variáveis com valores positivos ($x > 0$) e o parâmetro λ é o elemento que define a transformação específica e que, com freqüência, resulta em normalidade. Como existem transformações lineares que não afetam a análise de variância, na prática a Equação 3.5 pode ser simplificada para a seguinte forma (Equação 3.6):

$$y^{(\lambda)} = \begin{cases} y^\lambda & (\lambda \neq 0) \\ \ln y & (\lambda = 0) \end{cases} \quad (3.6)$$

Em outras palavras, a transformação de Box-Cox indica o valor do parâmetro λ que minimiza o erro quadrático (e conseqüentemente o desvio-padrão) da variável transformada, sendo a transformação resultante igual a y^λ quando λ é diferente de zero e $\ln y$ quando λ é igual a zero.

A aplicação da transformação de Box-Cox através do *software* estatístico Minitab® está estruturada na pesquisa dos valores de λ no intervalo e -5 até 5, de modo que quando a transformação é solicitada, o *software* apresenta um gráfico com as seguintes características:

- A melhor estimativa de λ para a transformação pretendida;
- Dois valores concorrentes para o verdadeiro valor de λ ;
- Um intervalo com 95% de confiança para o verdadeiro valor de λ .

Alguns exemplos típicos de transformação e o seu correspondente significado para valores de λ variando entre -2 e 2 são apresentados na Tabela 3.2.

VALOR DE λ	TRANSFORMAÇÃO $y^{(\lambda)} = y'$	DENOMINAÇÃO COMUM
2	$y' = y^2$	Quadrado
1	$y' = y$	Não Há Transformação
0,5	$y' = \sqrt{y}$	Raiz Quadrada
0	$y' = \ln y$	Logaritmo Natural
-0,5	$y' = \frac{1}{\sqrt{y}}$	Inverso da Raiz Quadrada
-1	$y' = \frac{1}{y}$	Inverso
-2	$y' = \frac{1}{y^2}$	Inverso do Quadrado

Tabela 3.2 – Exemplos típicos da transformação de Box-Cox

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

- Sejam descritos por um modelo de estrutura simples;
- A variância seja constante;
- As observações sejam normalmente distribuídas.

Para exemplificar as técnicas de transformação de variáveis, um conjunto de dados não normais contendo 48 (quarenta e oito) observações é apresentado através da Tabela 3.3.

EXEMPLO COM 48 OBSERVAÇÕES NÃO NORMAIS			
0,31	0,82	0,43	0,45
0,45	1,10	0,45	0,71
0,46	0,88	0,63	0,66
0,43	0,72	0,76	0,62
0,36	0,92	0,44	0,56
0,29	0,61	0,35	1,02
0,40	0,49	0,31	0,71
0,23	1,24	0,40	0,38
0,22	0,30	0,23	0,30
0,21	0,37	0,25	0,36
0,18	0,38	0,24	0,31
0,23	0,29	0,22	0,33

Tabela 3.3 – Exemplo de um conjunto de dados não normais

A Figura 3.2 apresenta o resultado da análise da distribuição referente ao conjunto de dados da Tabela 3.3. A condição de não normalidade é confirmada, pois o valor P (*p-value*) mostrou-se inferior a 0,05 no teste de normalidade de Anderson-Darling.

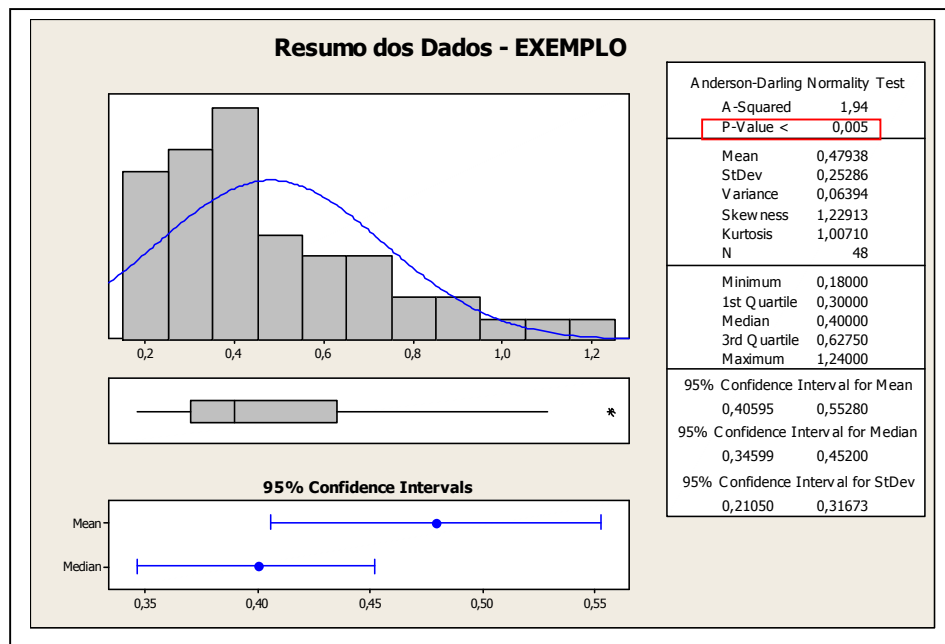


Figura 3.2 – Exemplo de análise de dados não normais

Para este mesmo exemplo, após a aplicação da transformação de Box-Cox pode-se avaliar a validade da transformação através do diagrama representado pela Figura 3.3.

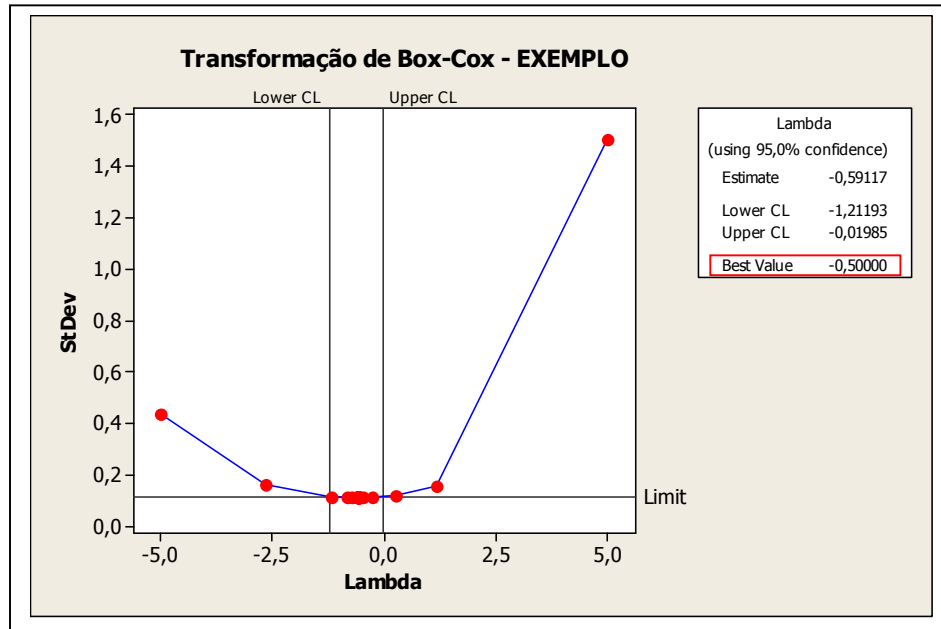


Figura 3.3 – Exemplo de um gráfico relativo à transformação de Box-Cox

Através deste diagrama, podem ser visualizados na parte superior direita, o valor estimado (-0,59117) e o melhor valor (-0,50000) para λ , sendo que o melhor valor é aquele efetivamente utilizado na transformação. Além disso, são indicados os valores limítrofes inferior (-1,21193) e superior (-0,01985), destacados pelas linhas verticais. O intervalo com 95% de confiança inclui todos os valores de λ que produzem dados transformados com desvio-padrão menor ou igual à linha horizontal, sendo que um valor de λ muito próximo a 1 não resultaria em transformação útil.

Em situações de aplicação prática é preferível que este valor corresponda a transformações que possam ser facilmente entendidas, tais como raiz quadrada ($\lambda = 0,5$) ou logaritmo natural ($\lambda = 0$). Neste exemplo a escolha razoável é -0,5 (inverso da raiz quadrada), pois além de representar uma transformação explicável, a mesma está contida no intervalo de 95% de confiança.

Após a transformação de Box-Cox com utilização do parâmetro λ igual a -0,5 é possível efetuar uma nova verificação em relação à condição de normalidade dos dados. A Figura 3.4 mostra a análise dos dados transformados com destaque para o teste de normalidade de Anderson-Darling, que apresenta *p-value* igual a 0,519; logo a distribuição transformada deve ser considerada normal.

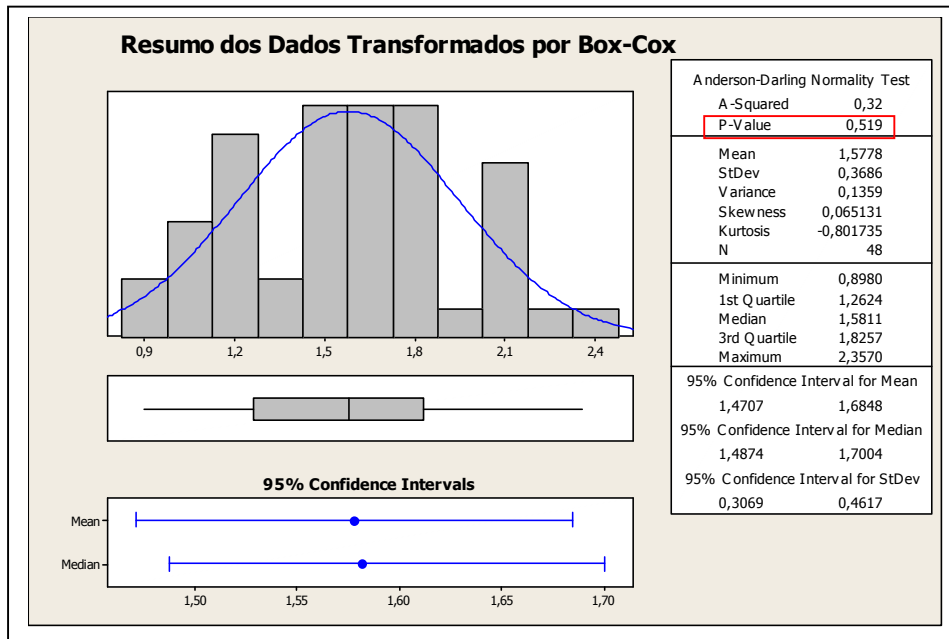


Figura 3.4 – Exemplo de análise dos dados após transformação de Box-Cox

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

Yeo & Johnson (2000) propuseram uma nova família de transformação de dados, válida tanto para valores positivos como para valores negativos. Sua fórmula, definida como uma função $\psi: \mathbb{R} \times \mathbb{R} \rightarrow \mathbb{R}$, é apresentada através da Equação 3.7.

$$\psi^{(\lambda, x)} = \begin{cases} \frac{\{(x+1)^\lambda - 1\}}{\lambda} & (x \geq 0, \lambda \neq 0) \\ \log(x+1) & (x \geq 0, \lambda = 0) \\ -\frac{\{(-x+1)^{2-\lambda} - 1\}}{2-\lambda} & (x < 0, \lambda \neq 2) \\ -\log(-x+1) & (x < 0, \lambda = 2) \end{cases} \quad (3.7)$$

3.4.3 Transformação de Johnson

Um outro método que pode ser empregado, conhecido como transformação de Johnson desenvolvido em 1949 por Norman L. Johnson, consiste em converter uma variável x para a distribuição normal padronizada. A transformação inclui três famílias de distribuição:

- Família S_B (onde o índice B provém de *bounded* ou limitado);
- Família S_L (onde o índice L significa Lognormal);
- Família S_U (onde o índice U provém de *unbounded* ou ilimitado).

Segundo Johnson (1949) a transformação é definida por uma das três fórmulas a partir da determinação dos parâmetros γ , η , ε e λ . As três fórmulas de transformação para a normal reduzida são mostradas na Tabela 3.4.

FAMÍLIA	TRANSFORMAÇÃO	PARÂMETROS
S_B	$\gamma + \eta \cdot \ln \left[\frac{x - \varepsilon}{\lambda + \varepsilon - x} \right]$	$\eta, \lambda > 0,$ $-\infty < \gamma < \infty$ $-\infty < \varepsilon < \infty$ $\varepsilon < x < \varepsilon + \lambda$
S_L	$\gamma + \eta \cdot \ln(x - \varepsilon)$	$\eta > 0,$ $-\infty < \gamma < \infty$ $-\infty < \varepsilon < \infty$ $\varepsilon < x$
S_U	$\gamma + \eta \cdot \text{Sinh}^{-1} \left[\frac{x - \varepsilon}{\lambda} \right]$	$\eta, \lambda > 0,$ $-\infty < \gamma < \infty$ $-\infty < \varepsilon < \infty$ $\varepsilon < x < \varepsilon + \lambda$

Tabela 3.4 – Fórmulas associadas com as famílias da transformação de Johnson

Fonte: Adaptado de Johnson (1949)

A aplicação da transformação de Johnson através do *software* estatístico Minitab® está estruturada de modo que o algoritmo execute as seguintes etapas:

- Considera as funções potenciais do sistema de Johnson;
- Estabelece a estimativa dos parâmetros γ , η , ε e λ ;
- Executa a transformação dos dados através da função de transformação;
- Efetua o teste de normalidade de Anderson-Darling para os dados transformados;
- Seleciona a função de transformação que apresente o maior *p-value*.

Caso o valor P (*p-value*) seja inferior ao valor α pré-estabelecido não se obtém transformação que resulte em normalidade.

Para o conjunto de dados não normais da Tabela 3.3 também foi efetuada aplicação da transformação de Johnson com as suas características destacadas na Figura 3.5.

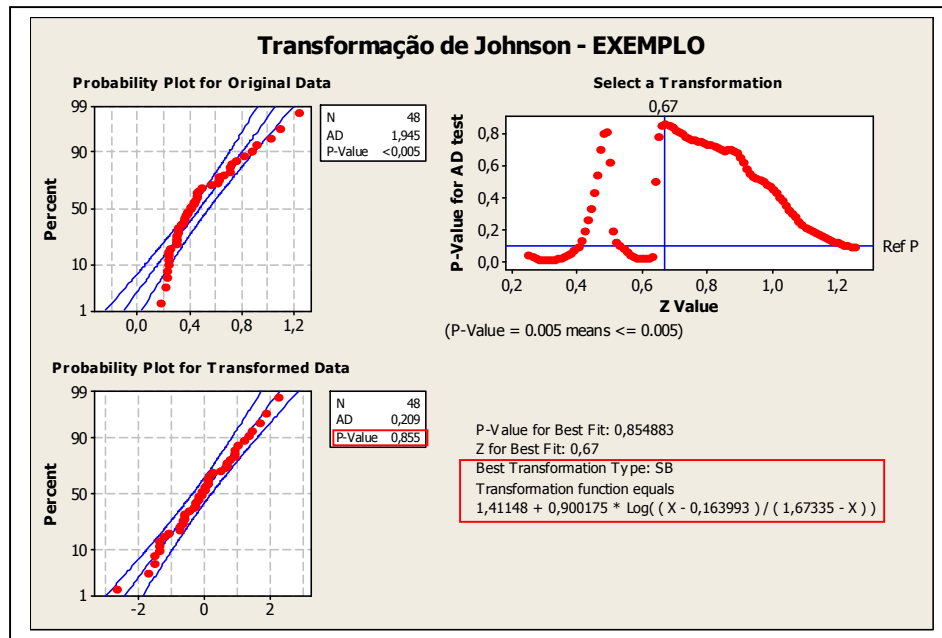


Figura 3.5 – Exemplo de um gráfico relativo à transformação de Johnson

A análise do diagrama acima, correspondente à transformação de Johnson, permite a discussão dos seguintes aspectos:

- Comparação entre os dados antes e depois da transformação através do gráfico normal de probabilidade. No exemplo, como já se sabia, os dados originais apresentam *p-value* menor que 0,05 (não normais), ao passo que os dados transformados possuem *p-value* igual a 0,855, caracterizando a adequada normalização dos dados;
- O diagrama de dispersão de *p-value versus z-value* em uma faixa de 0,25 a 1,25 indica a melhor função de transformação a ser selecionada para *z* igual a 0,67;
- O gráfico também mostra a função de transformação selecionada (família S_B) com os seus parâmetros correspondentes.

Observa-se através deste exemplo que, a probabilidade de 85,5% de que a distribuição transformada por Johnson seja Normal é bem maior do que aquela referente à distribuição resultante da transformação de Box-Cox (51,9%), ou seja, a transformação através do método de Johnson mostrou-se mais efetiva.

4. APLICAÇÃO DA METODOLOGIA

4.1 Considerações Gerais

4.1.1 Aspectos Metodológicos

A presente dissertação apresenta a questão referente à ocorrência de decisões questionáveis na aplicação de métodos estatísticos devido à interpretação inadequada dos dados coletados. Este é um problema para o qual já se obteve solução a partir de outras pesquisas, algumas delas mencionadas neste trabalho; entretanto, o que realmente motiva a abordagem deste tema é a necessidade de se destacar a relevância do rigor metodológico durante coleta e tratamento de dados estatísticos e promover uma maior divulgação das técnicas de transformação de variáveis.

Conforme definido no capítulo 1, este trabalho combina investigação experimental e pesquisa exploratória. Para permitir uma visualização completa dos aspectos gerais da pesquisa um resumo é apresentado através do Quadro 4.1. Na investigação experimental foram analisados os dados modelados pelas distribuições não normais Gamma (Apêndice A) e Beta (Apêndice B) em relação à interpretação dos gráficos de controle individuais e cálculo do nível sigma através da comparação das técnicas de transformação com a determinação dos percentis a partir da função densidade de probabilidade das distribuições. Na pesquisa exploratória foram investigados quatro conjuntos de dados reais não normais (Apêndices C e D) em relação ao cálculo de capacidade com utilização das técnicas de transformação de variáveis em comparação ao ajuste da melhor distribuição aplicável aos dados (*Best Fitting*).

TIPO DE PESQUISA		DADOS NÃO NORMAIS		QUESTÕES A INVESTIGAR	LIMITES		OBJETIVOS ESPECÍFICOS
MÉTODO	TÉCNICA	CONJUNTO DE DADOS	ORIGEM DOS DADOS		LIE	LSE	
Pesquisa Experimental	Simulação	Gamma 220 (Apêndice A)	Software Minitab®	Gráfico de Controle Individual e Nível Sigma	-	12	Comparar dados originais e análise de percentis c/ dados transformados
		Beta 2550 (Apêndice B)			8,20	8,50	
Pesquisa Exploratória	Estudo de Caso	3/4AE e 3/4AD (Apêndice C)	Processo Industrial	Cálculo de Capacidade e Nível Sigma	19,124	19,151	Comparar dados originais e distribuição adotada c/ dados transformados
		5/8TE e 5/8TD (Apêndice D)			15,950	15,977	

Quadro 4.1 – Resumo dos aspectos gerais da pesquisa

4.1.2 Caracterização dos Cenários da Pesquisa

Segundo Kotz & Johnson (2002) o rápido desenvolvimento científico no final do século XX, incluindo pesquisa em operações e estatística, foi verdadeiramente assombroso, de tal forma que muitos trabalhos efetuados há pouco mais de uma década hoje são classificados como obsoletos. Entretanto, devido à lacuna percebida entre teoria e prática, é pertinente a realização de um estudo exploratório para o caso relativo à aplicação de alguns métodos estatísticos de interesse acadêmico e industrial.

Para compor o contexto ilustrativo dos métodos estatísticos analisados no desenvolvimento da pesquisa foi elaborada uma investigação experimental baseada na utilização de construtos relacionados ao estudo de dados não normais simulados, modelados pelas distribuições Gamma e Beta, respectivamente. Para facilitar sua identificação, deste ponto em diante, os dados modelados pela distribuição Gamma serão chamados de “Gamma 220” e, sob a óptica de um contexto fictício, poderiam estar relacionados, por exemplo, à coleta do tempo médio diário de permanência em fila em uma empresa de prestação de serviços nos dias trabalhados ao longo de um ano. Neste contexto o tempo médio não deve exceder 12 minutos (limite superior de especificação). Estes dados, que podem ser consultados no Apêndice A, foram gerados através do *software* estatístico Minitab® da seguinte maneira:

✓ Função: *Calc / Set Base = 0*

✓ Função: *Calc / Random Data/ Gamma / Generate = 302 / Shape = 2 / Scale =2*

Ou seja, foram gerados 302 valores seguindo a distribuição Gamma com parâmetros *shape* igual a 2, *scale* igual a 2 e base para geração de dados igual a 0.

A mesma abordagem vale para os dados modelados pela distribuição Beta que serão denominados simplesmente como “Beta 2550”. No mesmo contexto fictício estes dados poderiam estar relacionados, por exemplo, à medição do diâmetro de peças usinadas. A medida do diâmetro desta peça usinada, neste cenário, deve ser $8,35 \pm 0,15$ mm, ou seja, deve estar situada entre 8,20 mm (limite inferior de especificação) e 8,50 mm (limite superior de especificação). Os dados brutos, detalhados no Apêndice B, foram gerados através do *software* estatístico Minitab® da seguinte maneira:

✓ Função: *Calc / Set Base = 0*

✓ Função: *Calc / Random Data / Beta / Generate = 200 / First Shape Parameter = 25 / Second Shape Parameter =5*

✓ Adição da constante 7,5 a cada um dos valores obtidos

Ou seja, foram gerados 200 valores seguindo a distribuição Beta com primeiro parâmetro *shape* igual a 25, segundo parâmetro *shape* igual a 5 e base para geração de dados igual a 0, com posterior transformação linear através do acréscimo da constante 7,5.

Os passos estabelecidos para a condução desta investigação experimental e os correspondentes tópicos dos capítulos da dissertação são mostrados na Figura 4.1.

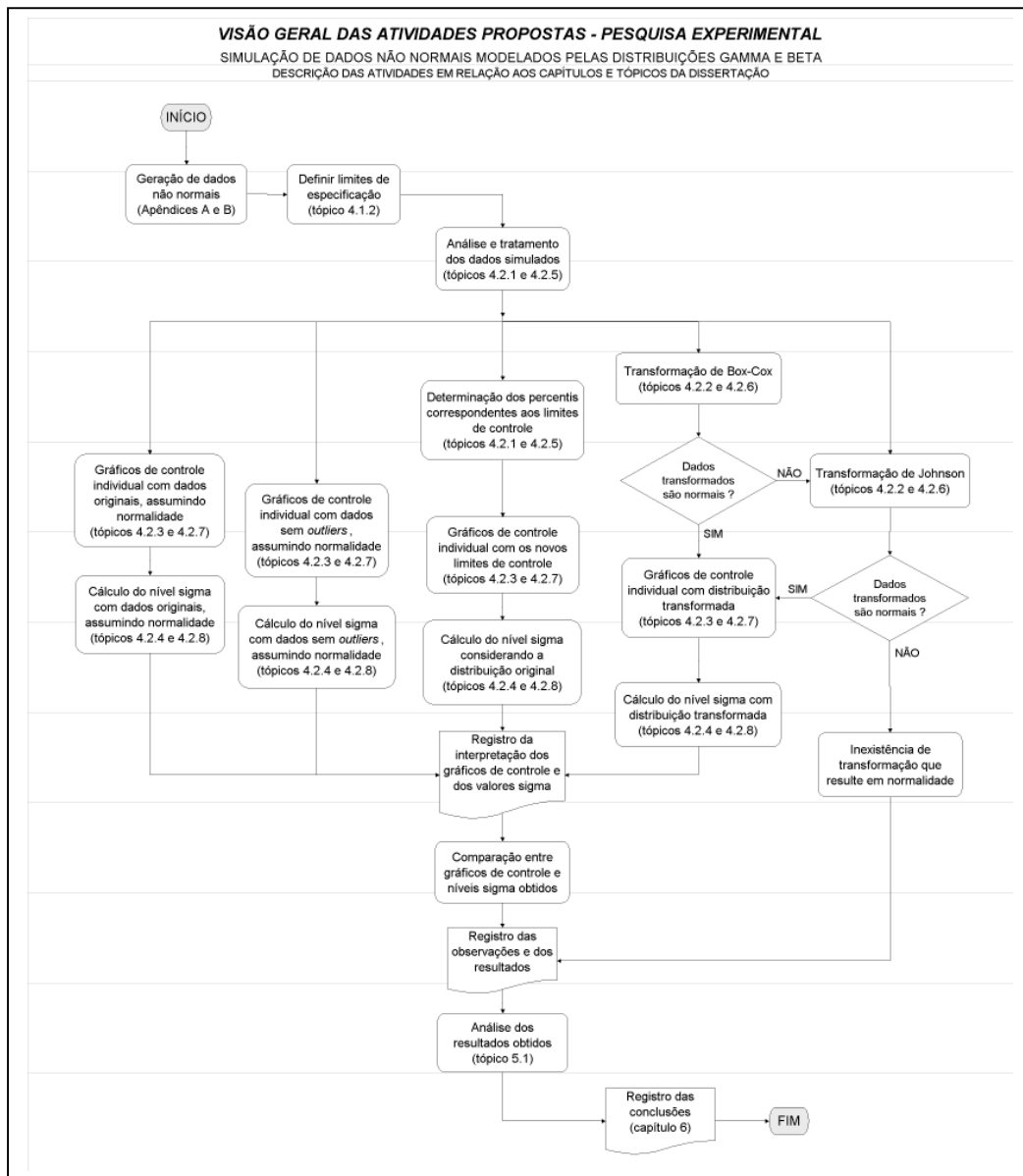


Figura 4.1 – Fluxograma das atividades da pesquisa experimental

A pesquisa exploratória, sob a forma de estudo de caso, teve o propósito de efetuar um exame detalhado em um processo específico de furação de precisão em uma empresa localizada no interior do estado de São Paulo. Os passos estabelecidos para a condução desta pesquisa exploratória e os correspondentes tópicos dos capítulos são mostrados na Figura 4.2.

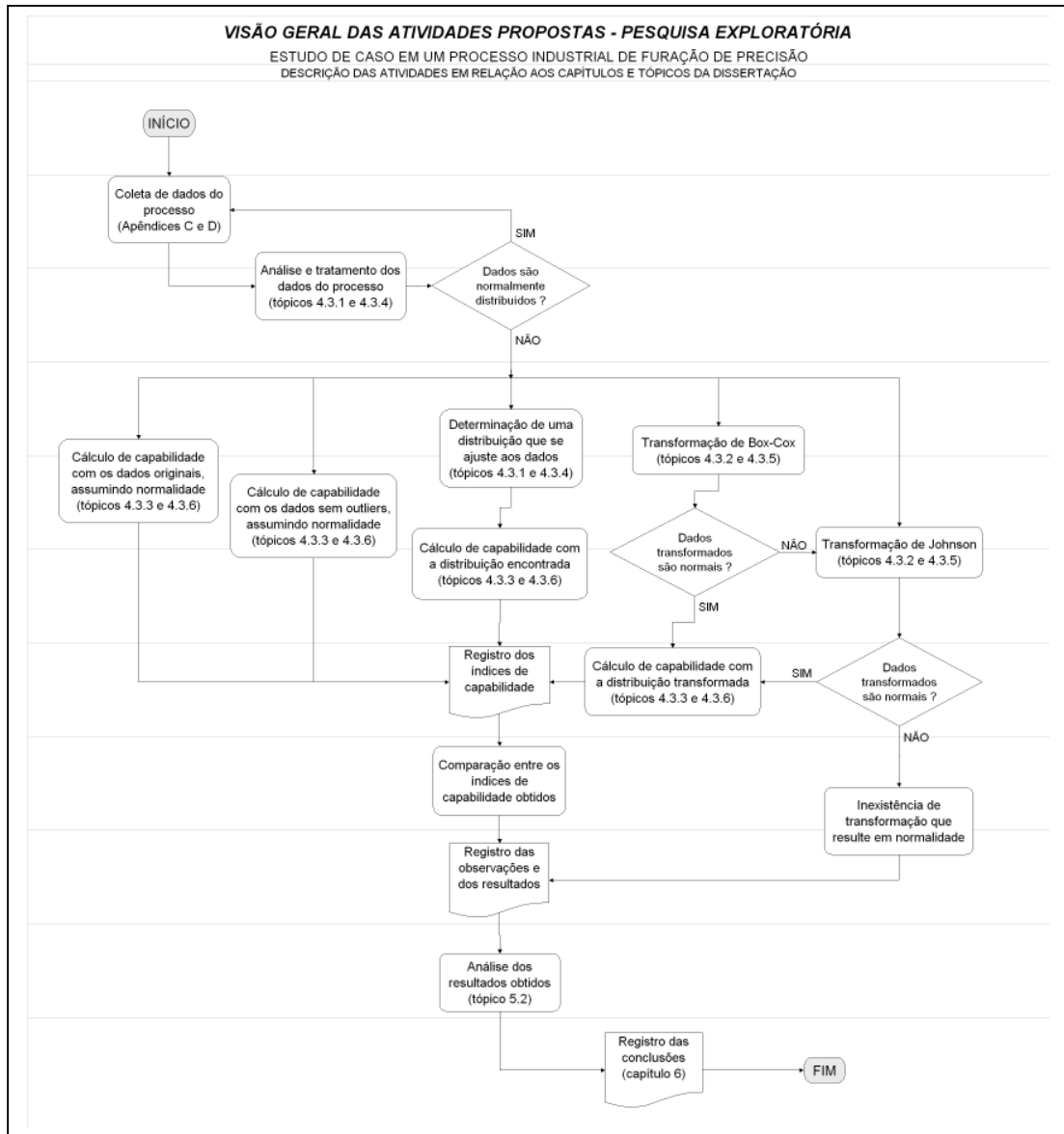


Figura 4.2 – Fluxograma das atividades da pesquisa exploratória

Devido a uma demanda específica do setor de Engenharia e Projeto da empresa, houve a necessidade de se determinar a capacidade do referido processo. Após coleta dos dados obtidos em um determinado período, os mesmos foram divididos em dois conjuntos de acordo com a medida nominal do diâmetro do furo e o tipo de material do componente a ser furado.

Os dados brutos do primeiro conjunto, que podem ser consultados através do Apêndice C, são relativos a dois grupos de medições de diâmetro de furos em um componente estrutural de alumínio para fixação de parafusos com diâmetro de 3/4 de polegada. No desenvolvimento do trabalho estes grupos de medições serão denominados simplesmente por:

- ✓ “3/4AE” – medições de furos na parte esquerda do componente de alumínio; e
- ✓ “3/4AD” – medições de furos na parte direita do componente de alumínio.

As tolerâncias de projeto especificadas para estes furos são as seguintes:

- LIE (Limite Inferior de Especificação) = 19,124 mm
- LSE (Limite Superior de Especificação) = 19,151 mm

Os dados do segundo conjunto estão detalhados no Apêndice D e referem-se aos valores de dois grupos de medições de diâmetro de furos em um componente estrutural de titânio para fixação de parafusos com diâmetro de 5/8 de polegada. São eles:

- ✓ “5/8TE” – medições de furos na parte esquerda do componente de titânio; e
- ✓ “5/8TD” – medições de furos na parte direita do componente de titânio.

As tolerâncias de projeto especificadas para estes furos são as seguintes:

- LIE (Limite Inferior de Especificação) = 15,950 mm
- LSE (Limite Superior de Especificação) = 15,977 mm

4.2 Investigação Experimental Através de Simulação de Dados

4.2.1 Análise de Dados Modelados pela Distribuição Gamma

De acordo com as considerações apresentadas no capítulo 1 as principais questões a serem investigadas através da pesquisa experimental são:

- a) Identificação de falsas causas especiais de variação nos gráficos de controle individuais em Controle Estatístico de Processo;
- b) Cálculo incorreto do nível sigma na avaliação de capacidade utilizando a Metodologia Seis Sigma.

Antes da aplicação dos métodos estatísticos citados foi necessário efetuar uma análise dos dados disponíveis em “Gamma 220” quanto a sua aderência à distribuição Normal. A partir dos dados simulados relativos ao tempo médio de fila em “Gamma 220” (Apêndice A) foram efetuados a análise da distribuição e o teste de normalidade de Anderson-Darling utilizando o programa estatístico.

Conforme argumentação indicada no capítulo 3, a verificação de normalidade das distribuições estudadas nesta pesquisa baseou-se na análise do valor P (*p-value*) associado à estatística A^2 do teste de normalidade de Anderson-Darling. Visto que para “Gamma 220” o valor P de 0,005 mostrou-se menor que o nível de significância arbitrado ($\alpha = 0,05$), a hipótese nula de normalidade foi rejeitada e, obviamente a distribuição em questão não pôde ser considerada normalmente distribuída. A Figura 4.3 apresenta o resultado da análise dos dados em “Gamma 220”.

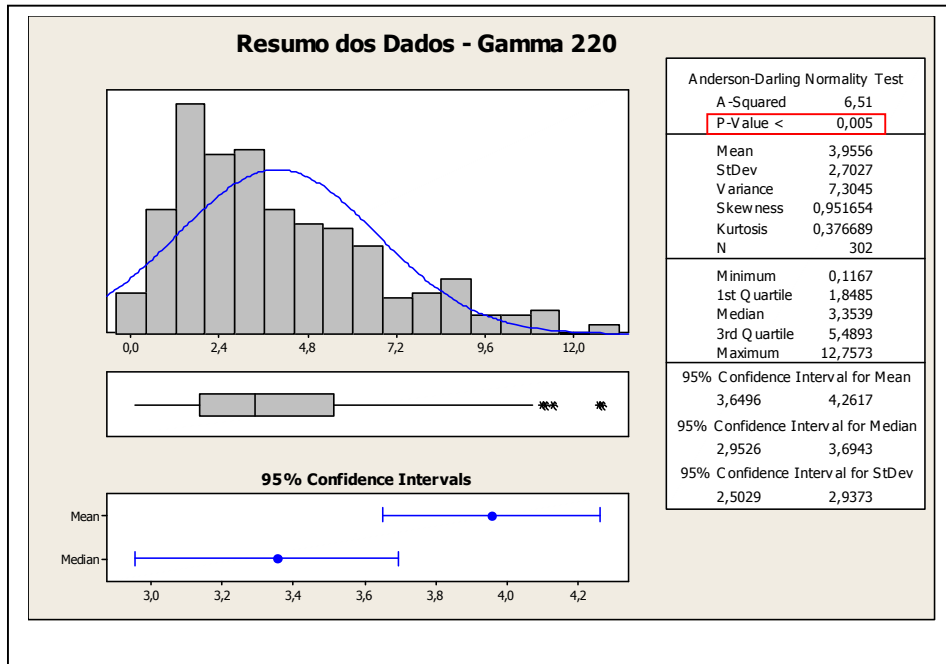


Figura 4.3 – Análise dos dados para “Gamma 220”

A análise do conjunto de dados “Gamma 220” também indicou a existência de alguns pontos extremos fora do padrão seguido pelo restante dos dados, isto é, de acordo com os conceitos apresentados no tópico 2.3.1, tais pontos foram caracterizados como *outliers*. Através da análise do conjunto de dados “Gamma 220” sem a presença dos *outliers*, apresentada na Figura 4.4 como “Gamma 220*”, concluiu-se que a distribuição resultante não era normal, pois o valor de *p-value* permaneceu inferior a 0,05.

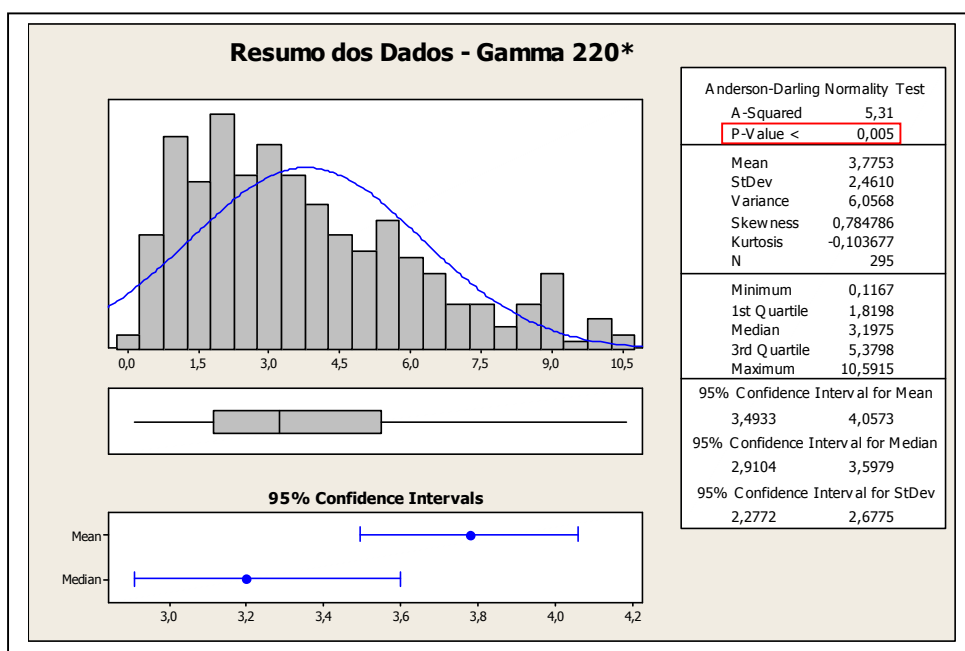


Figura 4.4 – Análise dos dados para “Gamma 220” sem *outliers*

Como o conjunto de dados “Gamma 220” sem *outliers* também não seguia a distribuição Normal, foi efetuada a verificação da distribuição que melhor representasse estes dados por meio das técnicas de análise de correlação. Conforme mostrado na Figura 4.5, a distribuição Weibull com 3 parâmetros foi a melhor opção encontrada, pois indicou o maior coeficiente de correlação (0,998) entre as distribuições analisadas para um intervalo de 95% de confiança.

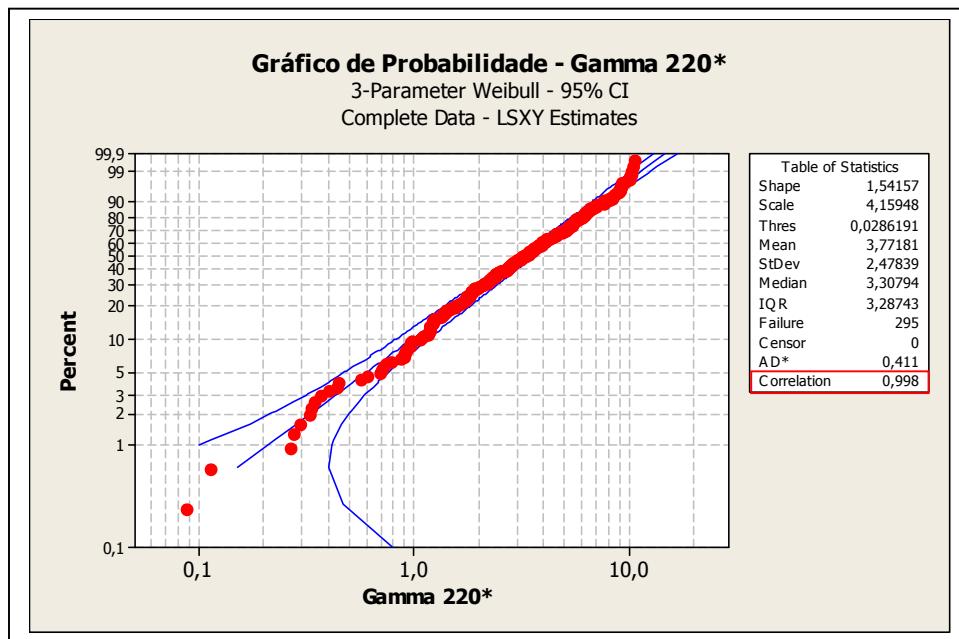


Figura 4.5 – Avaliação da melhor distribuição aplicável aos dados “Gamma 220” sem *outliers*

4.2.2 Transformação de Dados Modelados pela Distribuição Gamma

Para o conjunto de dados “Gamma 220” (Apêndice A) foi efetuada aplicação da transformação de Box-Cox. As características da transformação, com destaque para determinação do parâmetro $\lambda = 0,27$, estimado com 95% de confiança, são mostradas na Figura 4.6.

Após transformação de Box-Cox, admitindo $\alpha = 0,05$, os dados obtidos puderam ser representados por uma distribuição Normal, visto que *p-value* (0,476) resultou maior que 0,05. Em outras palavras, havia uma probabilidade de 47,6% de que a distribuição fosse normalmente distribuída. A Figura 4.7 apresenta a análise gráfica da distribuição resultante da transformação de Box-Cox.

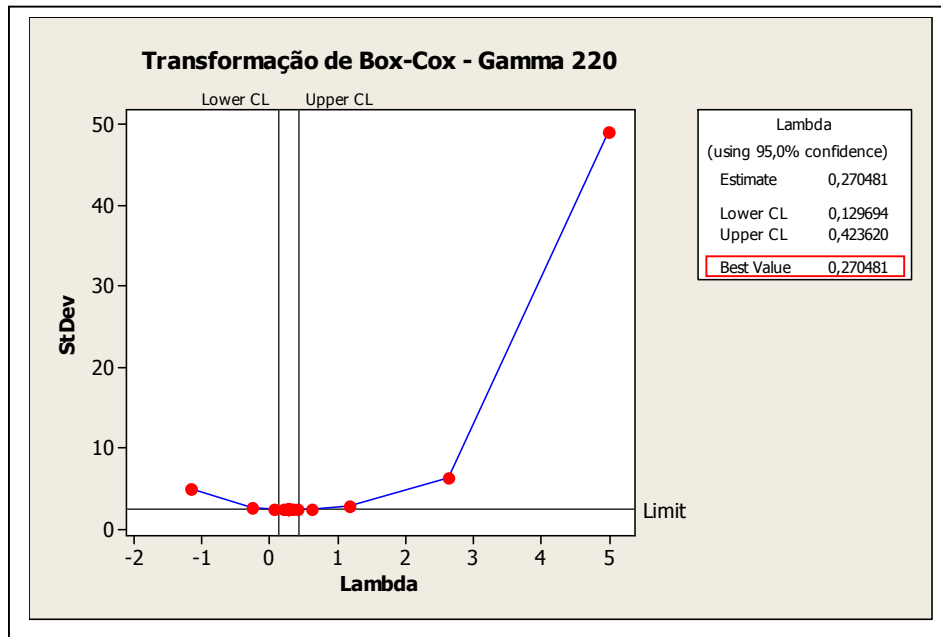


Figura 4.6 – Transformação de Box-Cox para os dados em “Gamma 220”

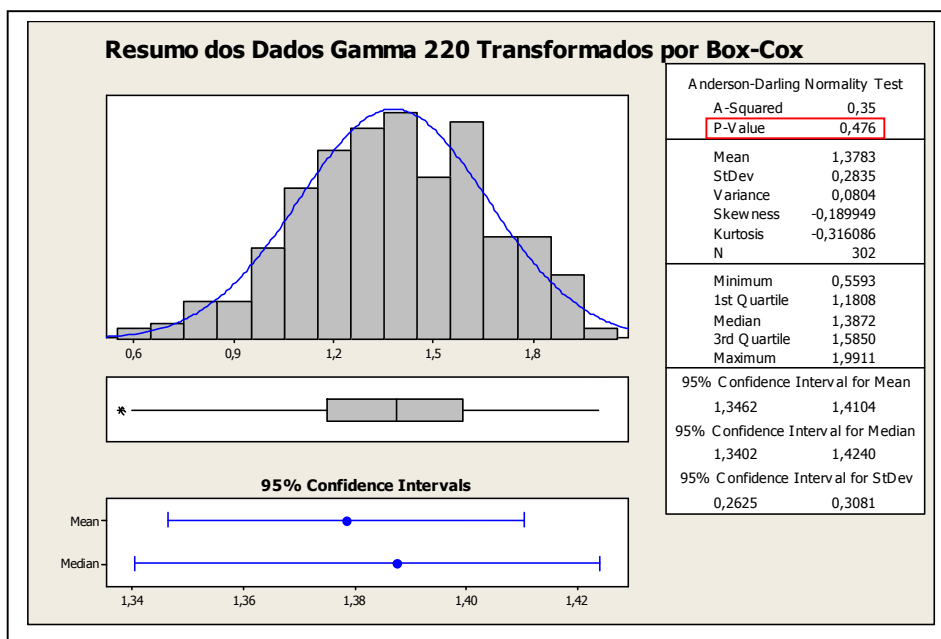


Figura 4.7 – Análise dos dados transformados por Box-Cox em “Gamma 220”

Para complementar a análise, foi efetuada aplicação da transformação de Box-Cox para o conjunto de dados “Gamma 220” sem outliers (ou “Gamma 220*”). As características da transformação, com destaque para determinação do parâmetro $\lambda = 0,31$, estimado com 95% de confiança, são mostradas na Figura 4.8.

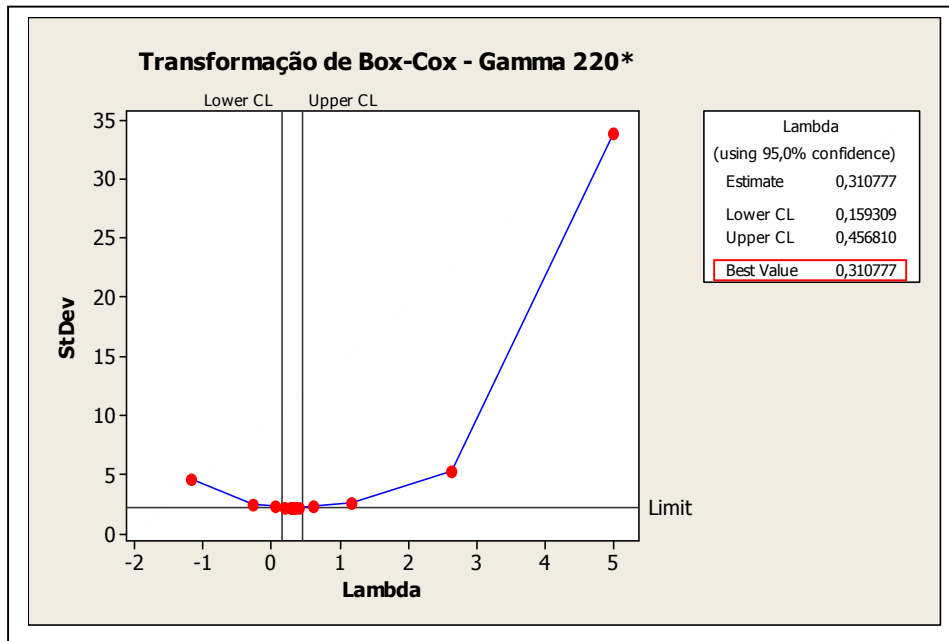


Figura 4.8 – Transformação de Box-Cox para os dados em “Gamma 220” sem outliers

Após transformação de Box-Cox, admitindo $\alpha = 0,05$, os dados obtidos puderam ser representados por uma distribuição Normal, visto que *p-value* (0,233) resultou maior que 0,05. A Figura 4.9 apresenta a análise gráfica da distribuição resultante da transformação de Box-Cox.

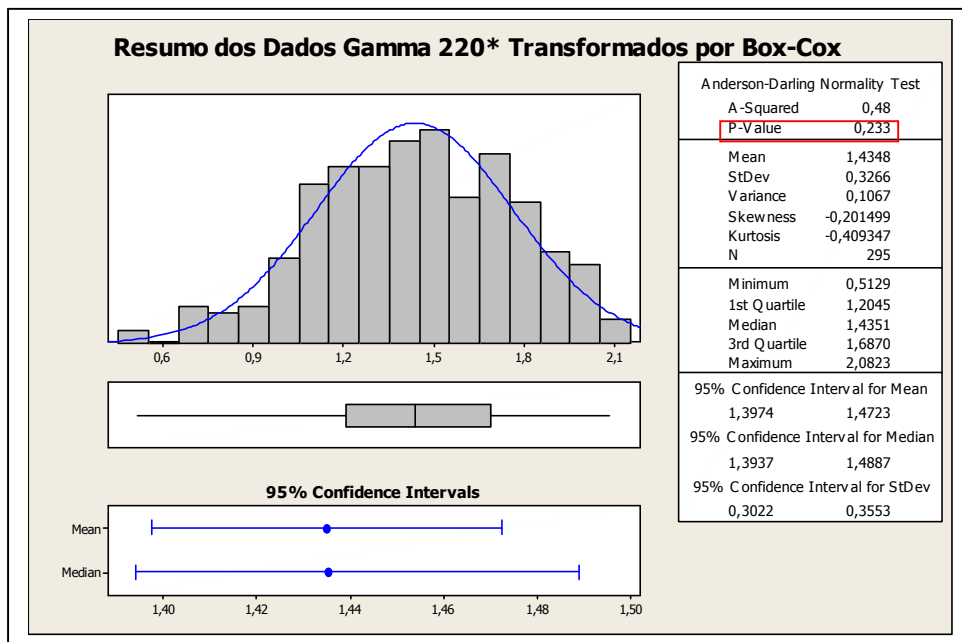


Figura 4.9 – Análise dos dados transformados por Box-Cox em “Gamma 220” sem outliers

Para o conjunto de dados “Gamma 220” também foi efetuada a transformação de Johnson. As características desta transformação, com destaque para determinação da família de transformação e a correspondente equação, são apresentadas na Figura 4.10.

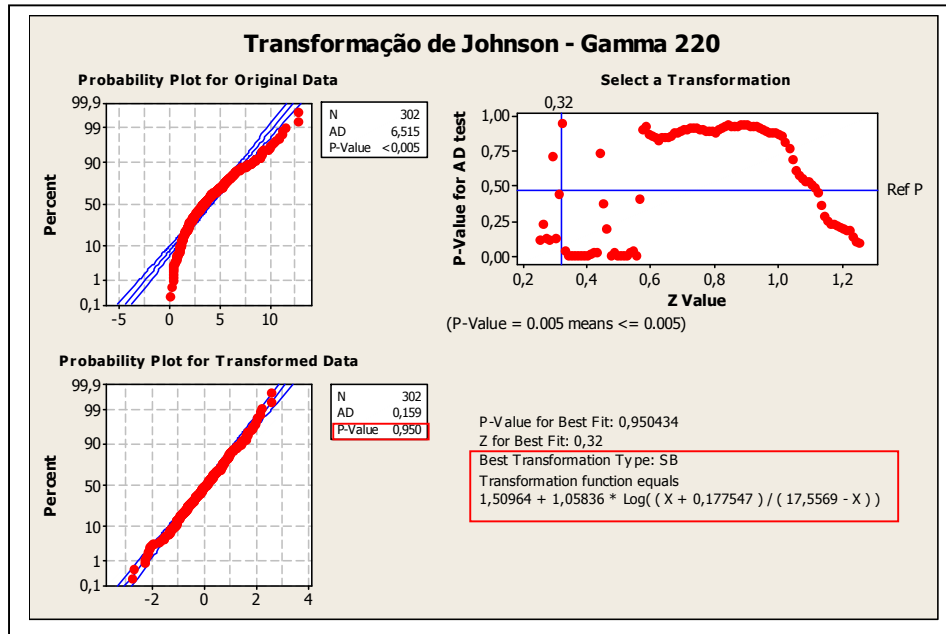


Figura 4.10 – Transformação de Johnson para os dados em “Gamma 220”

Após transformação de Johnson, admitindo $\alpha = 0,05$, os dados foram considerados como sendo normais (*p-value* igual a 0,950) conforme mostrado na Figura 4.11.

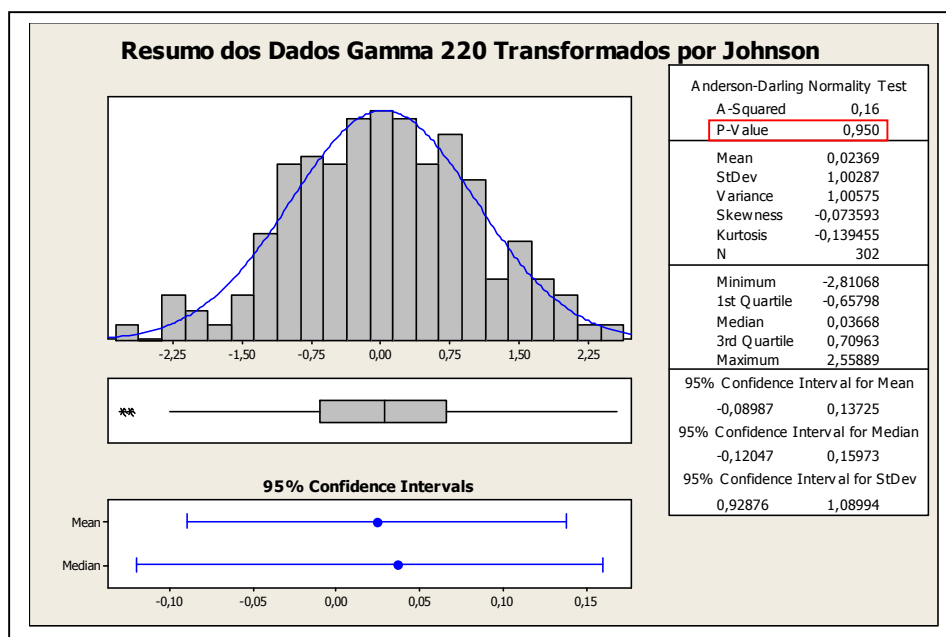


Figura 4.11 – Análise dos dados transformados por Johnson em “Gamma 220”

A análise foi complementada através da transformação de Johnson para o conjunto de dados “Gamma 220” sem *outliers* (ou “Gamma 220*”). As características da transformação são apresentadas na Figura 4.12.

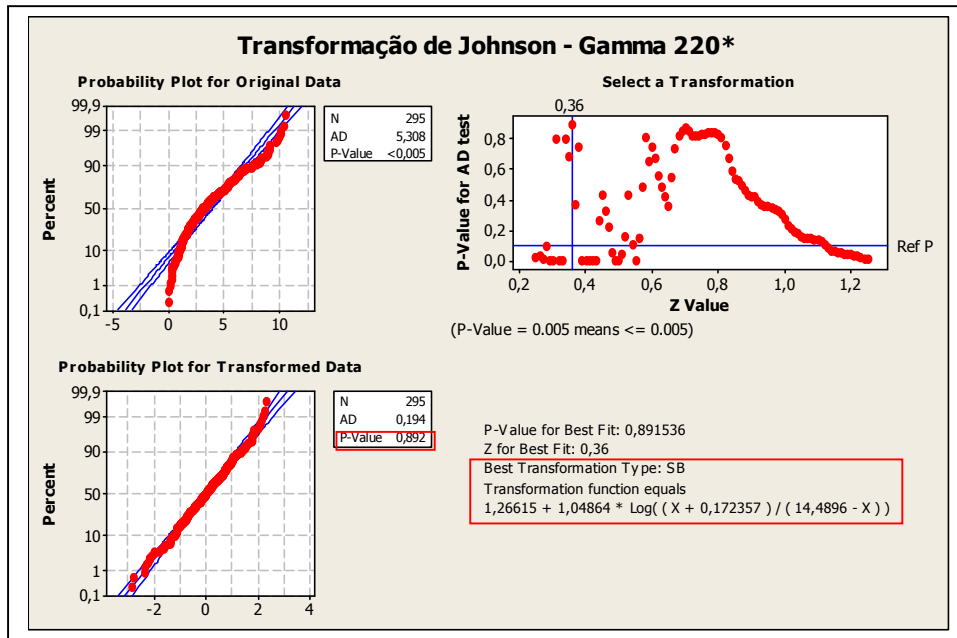


Figura 4.12 – Transformação de Johnson para os dados em “Gamma 220” sem *outliers*

Após transformação de Johnson, admitindo $\alpha = 0,05$, os dados foram considerados como sendo normais (*p-value* igual a 0,892) conforme mostrado na Figura 4.13.

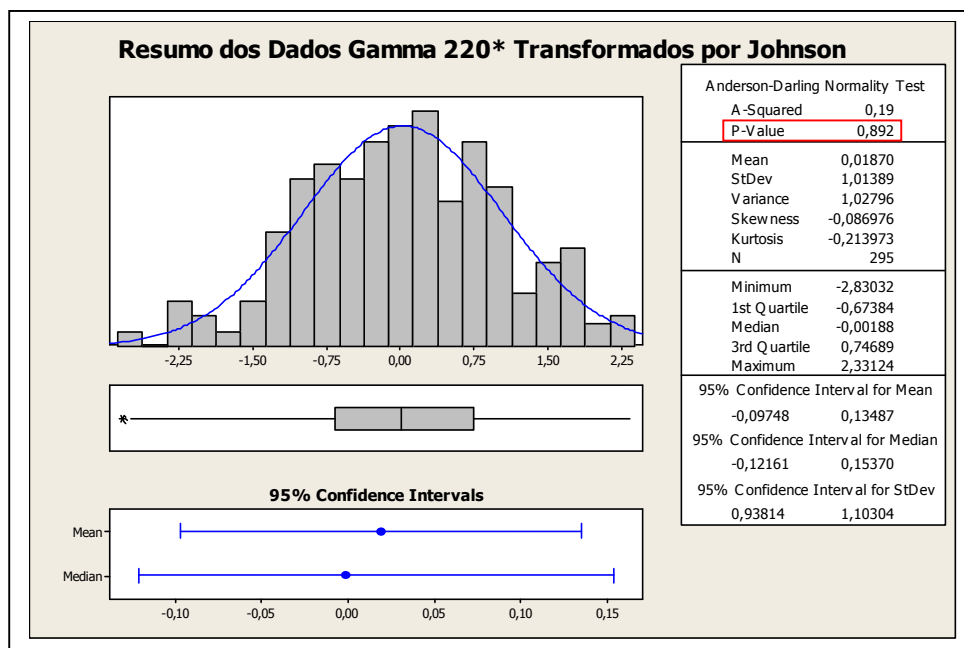


Figura 4.13 – Análise dos dados transformados (Johnson) em “Gamma 220” sem *outliers*

4.2.3 Gráficos de Controle Individuais – Distribuição Gamma

Algumas formas distintas de utilização de gráfico de controle para avaliação do processo em termos de tempo médio de fila em “Gamma 220” são discutidas a seguir. Em uma primeira abordagem, representada pela Figura 4.14, os dados originais foram inseridos para análise no programa estatístico sem a preocupação de testar a aderência dos mesmos à distribuição Normal, ou seja, admitindo incorretamente uma condição de normalidade dos dados.

Analisando o gráfico da Figura 4.14 em relação ao primeiro critério de identificação de padrões não aleatórios, estabelecido anteriormente no tópico 2.4.3, também conhecido como “teste 1”, que avalia a ocorrência de pontos além dos limites de controle estipulados, observou-se a existência de 2 (dois) pontos localizados acima do limite superior de controle. Este fato denotaria a presença de duas causas especiais de variação, caso a suposição de normalidade estivesse correta.

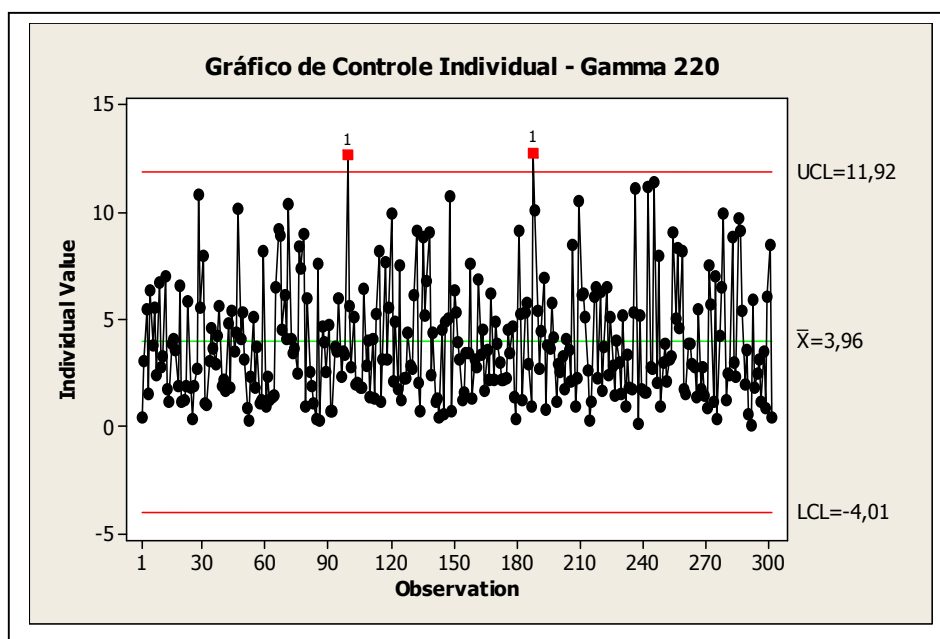


Figura 4.14 – Gráfico de controle para “Gamma 220” considerando distribuição Normal

No entanto, conforme mencionado anteriormente, o conjunto de dados “Gamma 220” apresentava algumas observações consideradas fora do padrão distribucional, isto é, alguns pontos caracterizavam-se como *outliers*. Por este motivo, também foi construído o gráfico de controle individual do conjunto de dados “Gamma 220” sem a presença dos *outliers* (ou “Gamma 220*”) e assumindo normalidade dos dados. Conforme demonstrado através da Figura 4.15 não mais se visualizaram pontos além dos limites de controle.

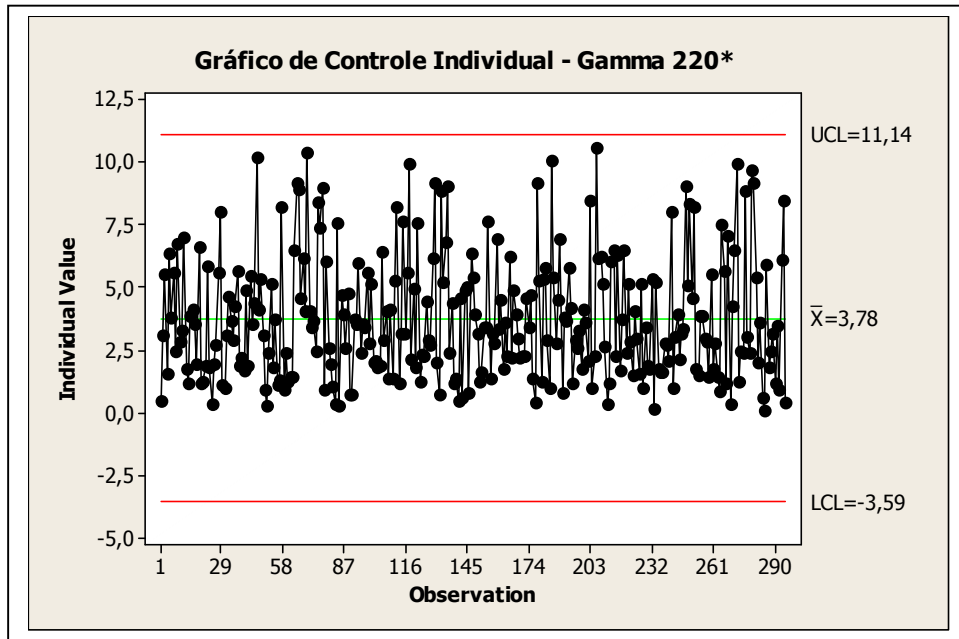


Figura 4.15 – Gráfico de controle “Gamma 220” sem outliers assumindo distribuição Normal

Uma segunda abordagem quanto à avaliação da estabilidade dos dados em “Gamma 220” foi adotada conforme critérios apresentados no tópico 3.3, ou seja, foram estabelecidos limites de controle por meio do percentil 0,135 (0,105767) para LIC, percentil 50 ou mediana (3,356690) para a LM e percentil 99,865 (17,8004) para LSC. A Figura 4.16 indica que para esta situação não foram encontrados pontos fora dos limites de controle.

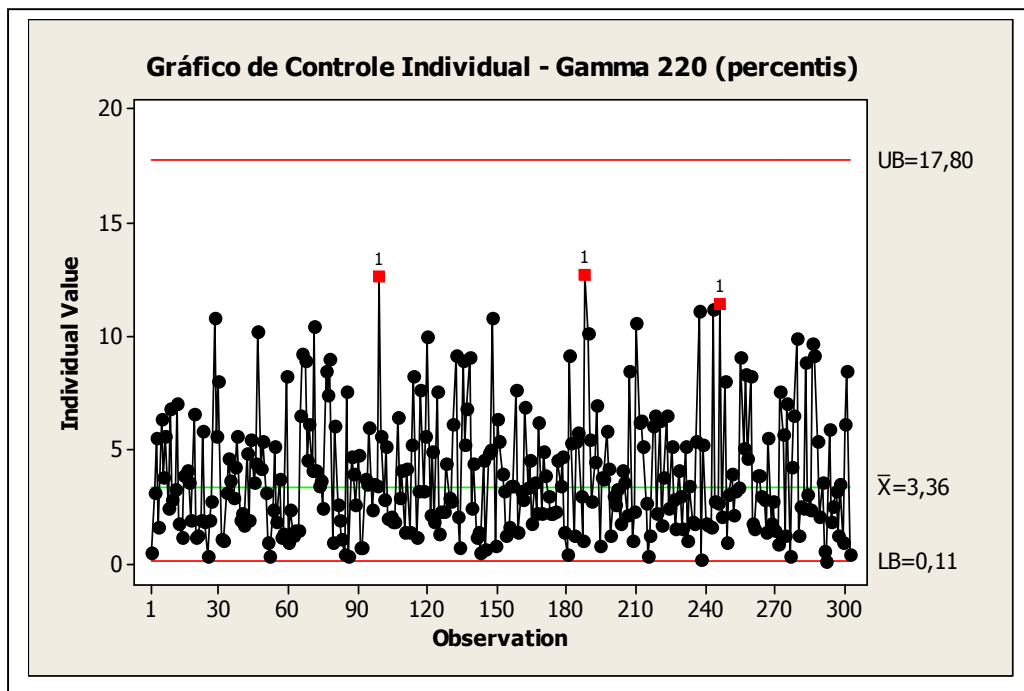


Figura 4.16 – Gráfico de controle “Gamma 220” com limites de controle em percentis

A Figura 4.17 apresenta a mesma análise para “Gamma 220” já sem os *outliers* e, nesta situação, também não se verificaram pontos além dos limites de controle.

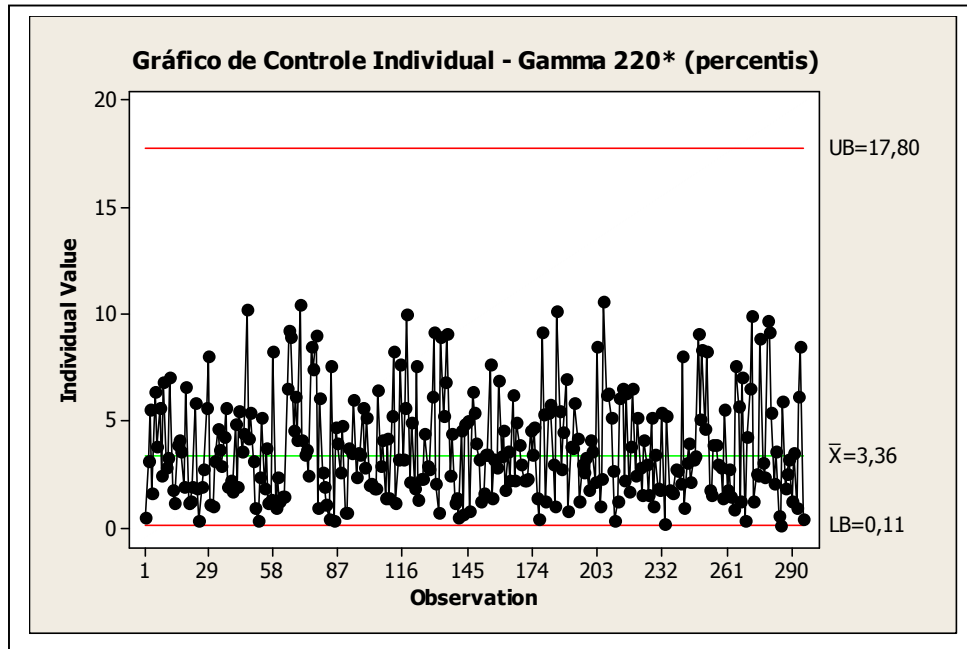


Figura 4.17 – Gráfico de controle “Gamma 220” sem *outliers* com limites em percentis

Com relação à terceira abordagem, mostrada na Figura 4.18, o gráfico de controle individual, que foi obtido após prévia transformação dos dados originais para a condição de normalidade via transformação de Box-Cox, não indicou pontos além dos limites de controle.

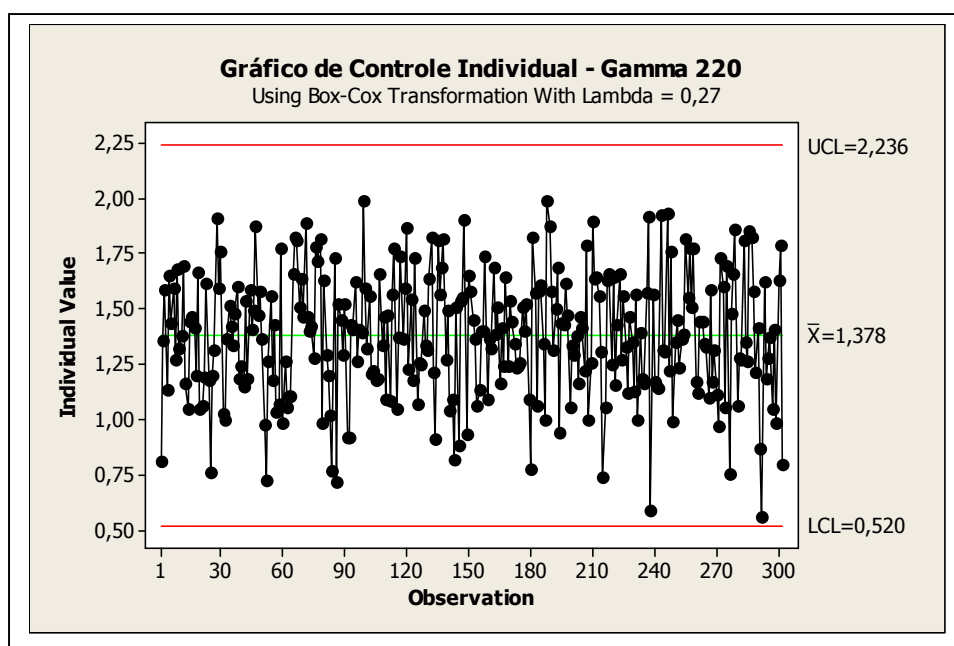


Figura 4.18 – Gráfico de controle para “Gamma 220” após transformação de Box-Cox

A Figura 4.19 mostra que os dados de “Gamma 220” transformados por Box-Cox sem os outliers (“Gamma 220*”) também não apresentaram pontos além dos limites de controle.

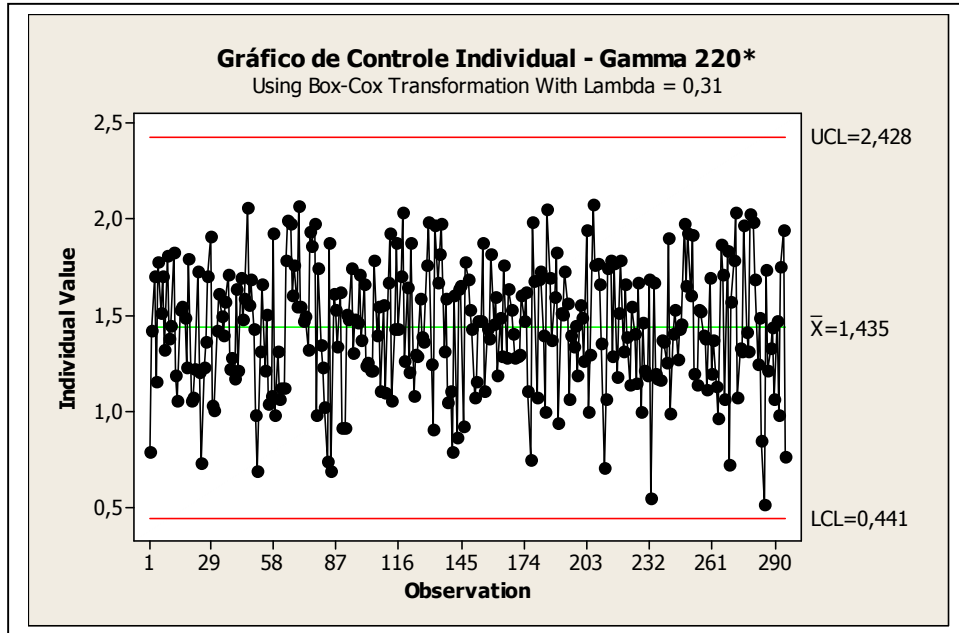


Figura 4.19 – Gráfico de controle para “Gamma 220” sem outliers após Box-Cox

Na quarta abordagem, a análise do gráfico de controle individual após a transformação de Johnson, mostrada na Figura 4.20, não indicou pontos fora dos limites de controle, assim como nos dados transformados por Box-Cox.

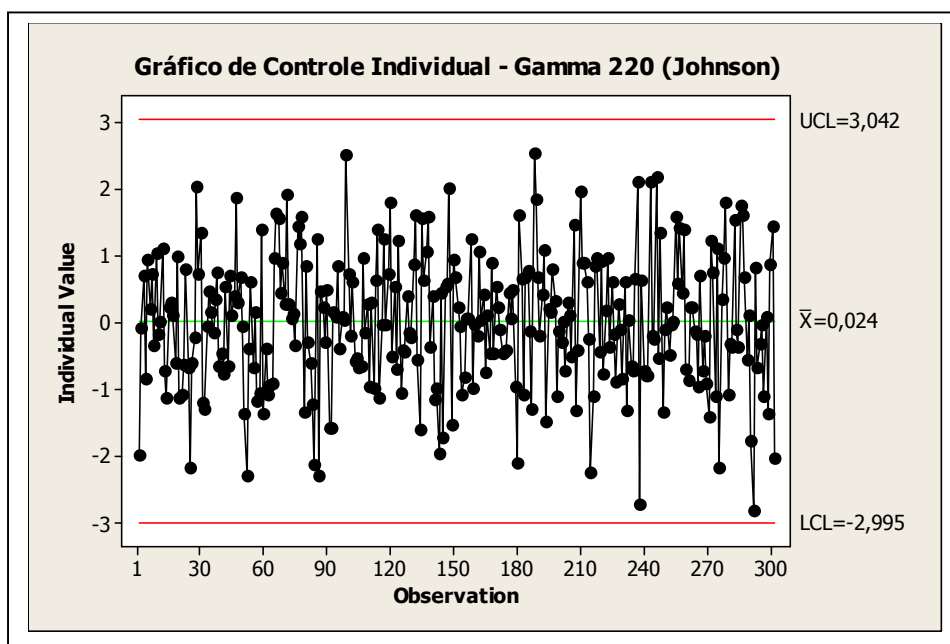


Figura 4.20 – Gráfico de controle para “Gamma 220” após transformação de Johnson

A Figura 4.21 indica que os dados de “Gamma 220” transformados por Johnson sem os *outliers* (“Gamma 220*”) também não apresentaram pontos além dos limites de controle.

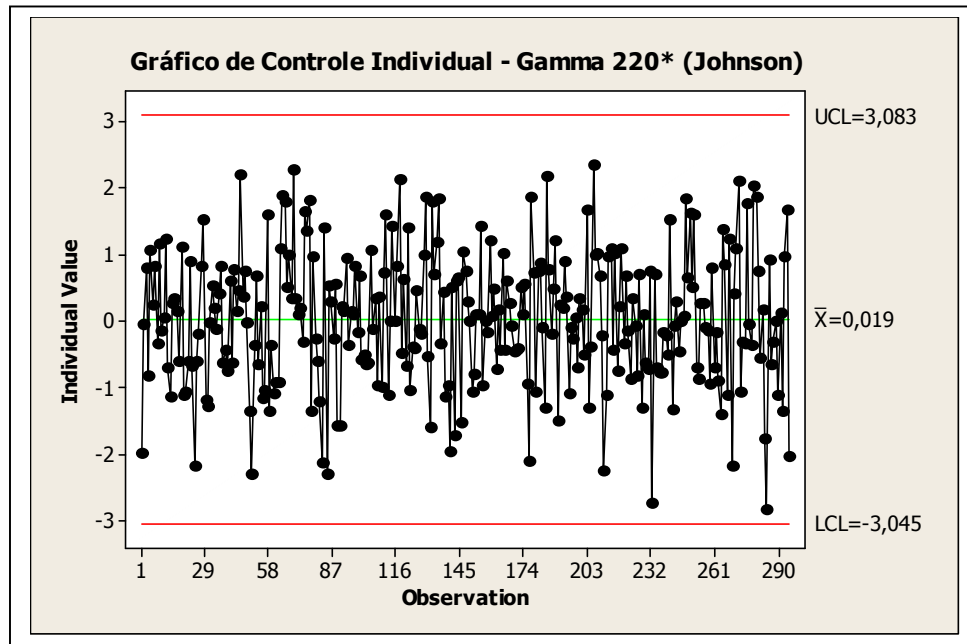


Figura 4.21 – Gráfico de controle para “Gamma 220” sem *outliers* após Johnson

4.2.4 Avaliação de Capabilidade – Distribuição Gamma

A avaliação de capacidade para o conjunto de dados “Gamma 220” teve como requisito um limite superior de especificação de 12 minutos. Foram calculados diversos índices de capacidade Z_{bench} (Z_{LT} e Z_{ST}) diferentes:

- Em uma primeira abordagem o cálculo foi efetuado com os dados originais e os dados tratados (sem *outliers*) como se fossem normalmente distribuídos;
- No segundo caso a condição assimétrica da distribuição foi considerada no cálculo, tanto para os dados originais quanto para os dados sem *outliers*.
- Na terceira abordagem os dados originais e sem *outliers* foram previamente transformados por meio da transformação de Box-Cox;
- Na quarta situação foi utilizada a transformação de Johnson.

A Figura 4.22 apresenta os resultados do cálculo de capacidade considerando os dados originais “Gamma 220” como se fossem normalmente distribuídos. A Figura 4.23 apresenta os resultados considerando os dados originais “Gamma 220” sem a presença de *outliers* (“Gamma 220*”) como se fossem normalmente distribuídos.

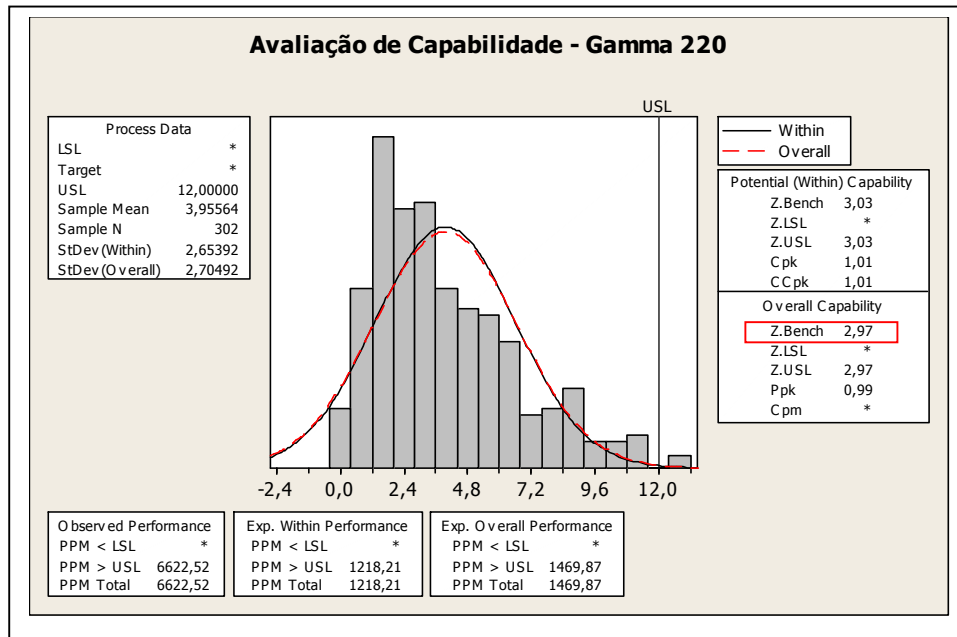


Figura 4.22 – Índices de capacidade de “Gamma 220” considerando distribuição Normal

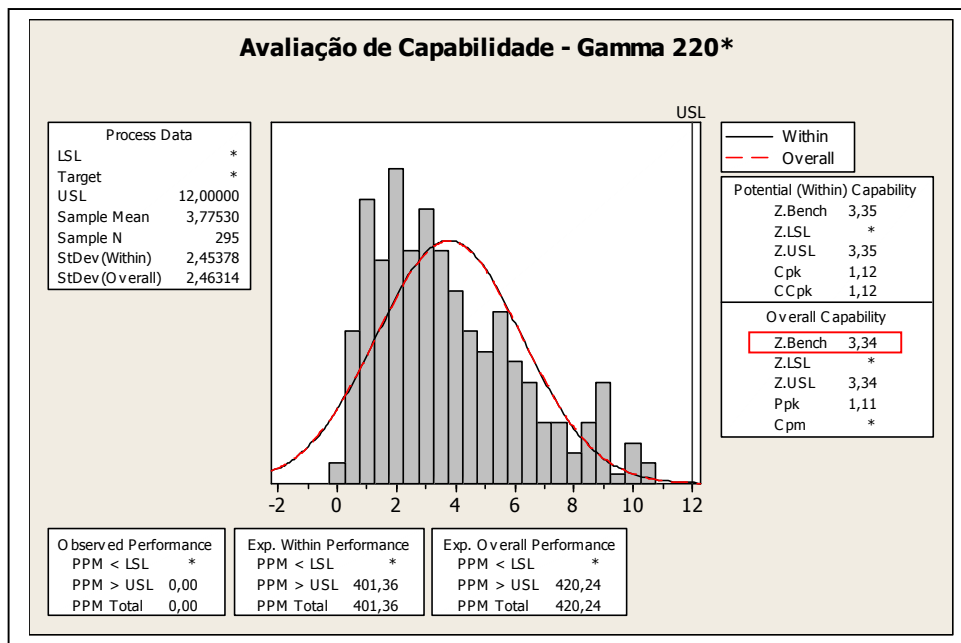


Figura 4.23 – Índices de capacidade de “Gamma 220” sem outliers assumindo normalidade

Como o conjunto de dados “Gamma 220”, de fato, não era modelado pela distribuição Normal e sim pela distribuição Gamma, o cálculo de capacidade foi feito levando em conta os parâmetros da função densidade de probabilidade da distribuição Gamma, conforme demonstrado através da Figura 4.24.

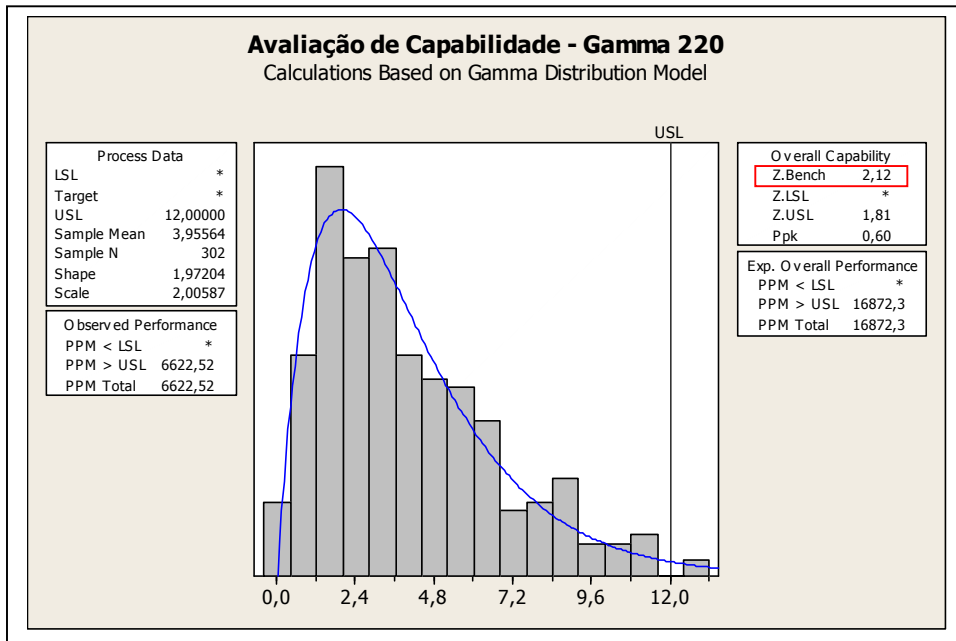


Figura 4.24 – Índices de capabilidade de “Gamma 220” considerando distribuição Gamma

A eliminação dos *outliers* do conjunto de dados “Gamma 220” resultou em um padrão diferente da distribuição Gamma. Portanto, antes do cálculo de capabilidade foi verificado que a distribuição Weibull com 3 parâmetros, representaria o melhor ajuste para os dados de “Gamma 220” sem a presença de *outliers* (“Gamma 220*”), para um intervalo de 95% de confiança. Os resultados nesta situação são apresentados através da Figura 4.25.

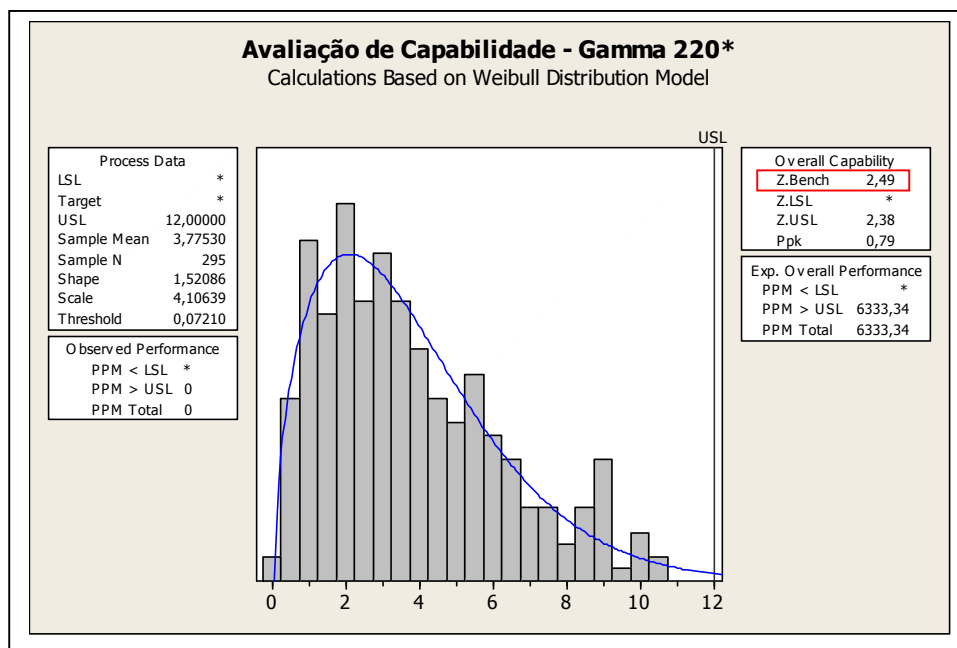


Figura 4.25 – Índices de capabilidade de “Gamma 220” sem *outliers* (distribuição Weibull)

A Figura 4.26 apresenta os resultados do cálculo de capacidade do processo após aplicação da transformação de Box-Cox no conjunto de dados originais “Gamma 220”.

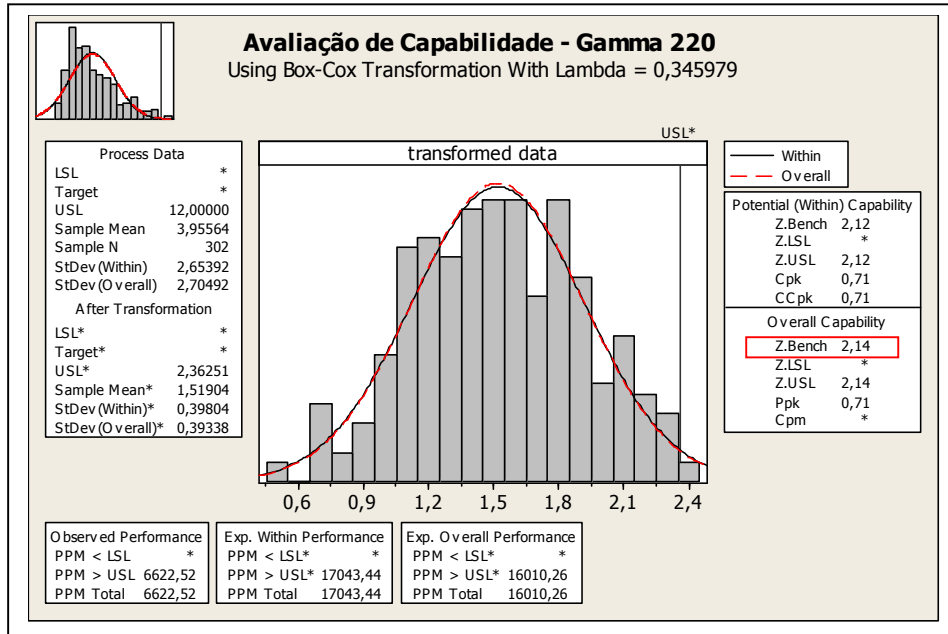


Figura 4.26 – Índices de capacidade de “Gamma 220” após transformação de Box-Cox

O cálculo de capacidade após a transformação de Box-Cox também foi efetuado para o conjunto de dados “Gamma 220” sem a presença de outliers (“Gamma 220*”) e os resultados deste cálculo estão evidenciados na Figura 4.27.

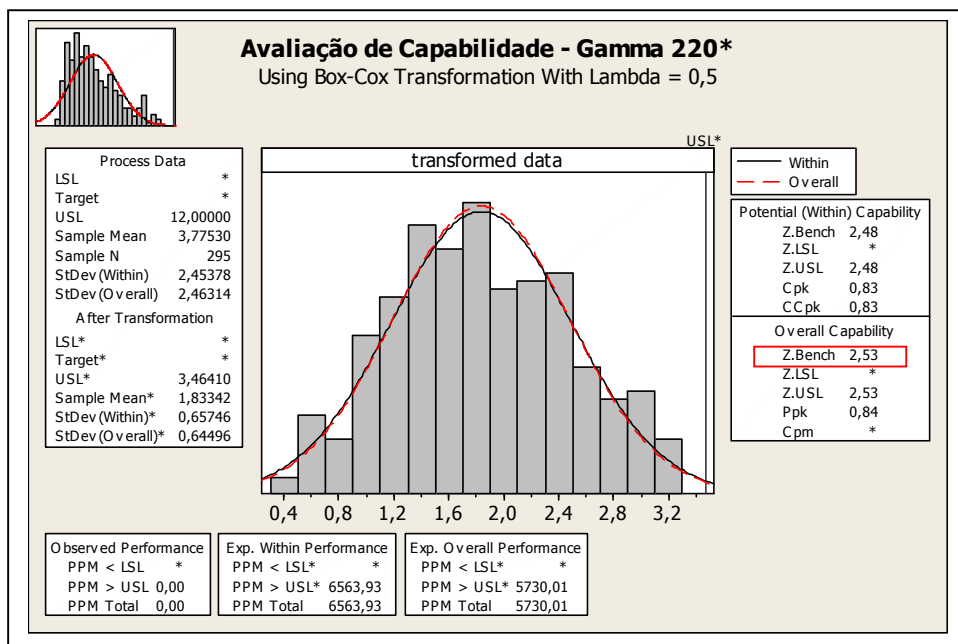


Figura 4.27 – Índices de capacidade de “Gamma 220” sem outliers após Box-Cox

A Figura 4.28 apresenta os resultados do cálculo de capacidade do processo após utilização da transformação de Johnson no conjunto de dados originais “Gamma 220”.

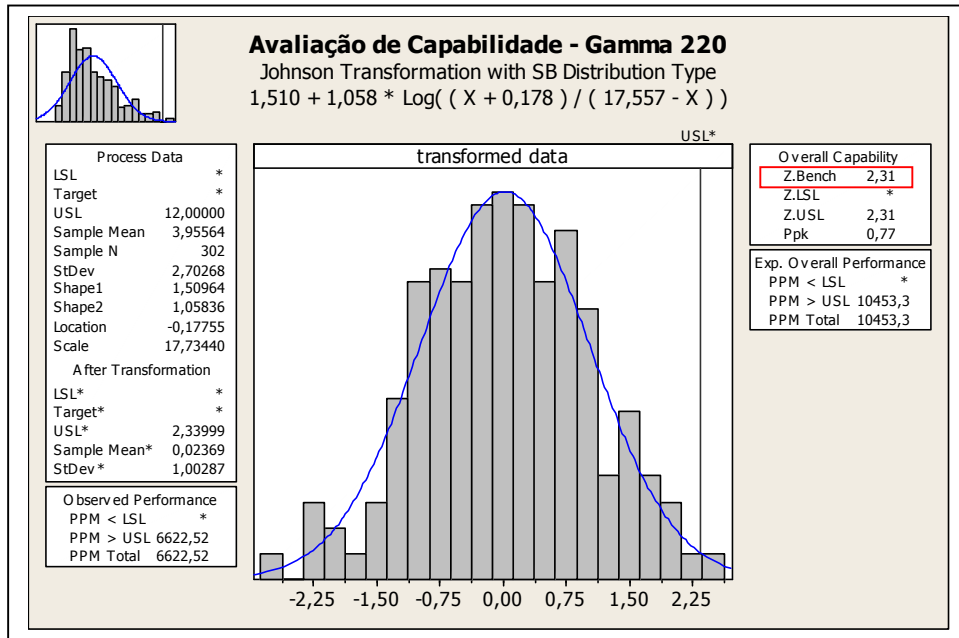


Figura 4.28 – Índices de capacidade de “Gamma 220” após transformação de Johnson

O cálculo de capacidade após a transformação de Johnson também foi efetuado para o conjunto de dados “Gamma 220” sem a presença de outliers (“Gamma 220*”) e os resultados deste cálculo são mostrados na Figura 4.29.

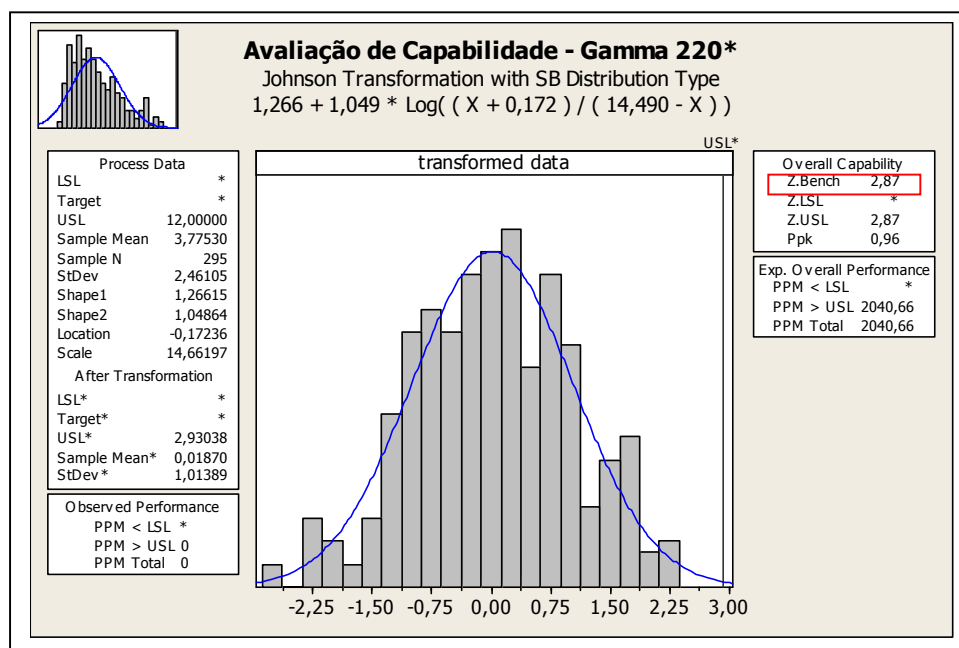


Figura 4.29 – Índices de capacidade de “Gamma 220” sem outliers após Johnson

A Tabela 4.1 apresenta a comparação dos valores obtidos nas diversas situações apresentadas, com o correspondente cálculo do nível sigma e a fração não conforme em partes por milhão para o conjunto de dados “Gamma 220”, considerando Z_{shift} igual a 1,5 sigma.

PREMISSA ADOTADA PARA CÁLCULO	LSE (min)	ÍNDICE Z_{LT}	NÍVEL SIGMA ($Z_{LT}+1,5$)	PPM
Dados originais considerados normalmente distribuídos	12	2,97	4,47	1470
Dados sem <i>outliers</i> considerados normalmente distribuídos	12	3,34	4,84	420
Dados originais considerando distribuição Gamma	12	2,12	3,62	16872
Dados sem <i>outliers</i> considerando distribuição Weibull	12	2,49	3,99	6333
Dados originais não normais transformados por Box-Cox	12	2,14	3,64	16010
Dados sem <i>outliers</i> transformados por Box-Cox	12	2,53	4,03	5730
Dados originais transformados por Johnson	12	2,31	3,81	10453
Dados sem <i>outliers</i> transformados por Johnson	12	2,87	4,37	2041

Tabela 4.1 – Cálculos comparativos do nível sigma para “Gamma 220”

4.2.5 Análise de Dados Modelados pela Distribuição Beta

Antes da aplicação dos métodos estatísticos citados foi necessário efetuar uma análise dos dados disponíveis em “Beta 2550” quanto a sua aderência à distribuição Normal. A partir dos dados simulados relativos aos valores de diâmetro de pino especial em “Beta 2550” (Apêndice B) foram efetuados a análise da distribuição e o teste de normalidade utilizando o programa estatístico, sendo que a verificação de normalidade das distribuições estudadas nesta pesquisa baseou-se na análise do valor P (*p-value*) associado à estatística A^2 do teste de normalidade de Anderson-Darling.

A Figura 4.30 apresenta o resultado da análise dos dados em “Beta 2550”. Admitindo $\alpha = 0,05$, a distribuição em questão foi considerada não normal visto que o *p-value* obtido (0,023) resultou menor que 0,05. Assim sendo, como a hipótese nula de normalidade foi rejeitada, obviamente a distribuição em questão não pôde ser considerada normalmente distribuída.

A análise do conjunto de dados “Beta 2550” também indicou a existência de alguns pontos extremos fora do padrão seguido pelo restante dos dados, isto é, de acordo com os conceitos apresentados no tópico 2.3.1, tais pontos foram caracterizados como *outliers*.

Através da análise do conjunto de dados “Beta 2550” sem a presença dos *outliers*, apresentada na Figura 4.31 como “Beta 2550*”, concluiu-se que a distribuição resultante poderia ser bem representada pela distribuição normal, pois o *p-value* encontrado (0,582) resultou superior a 0,05. Como a hipótese nula de normalidade foi aceita, obviamente o conjunto de dados “Beta 2550” pôde ser considerado normalmente distribuído.

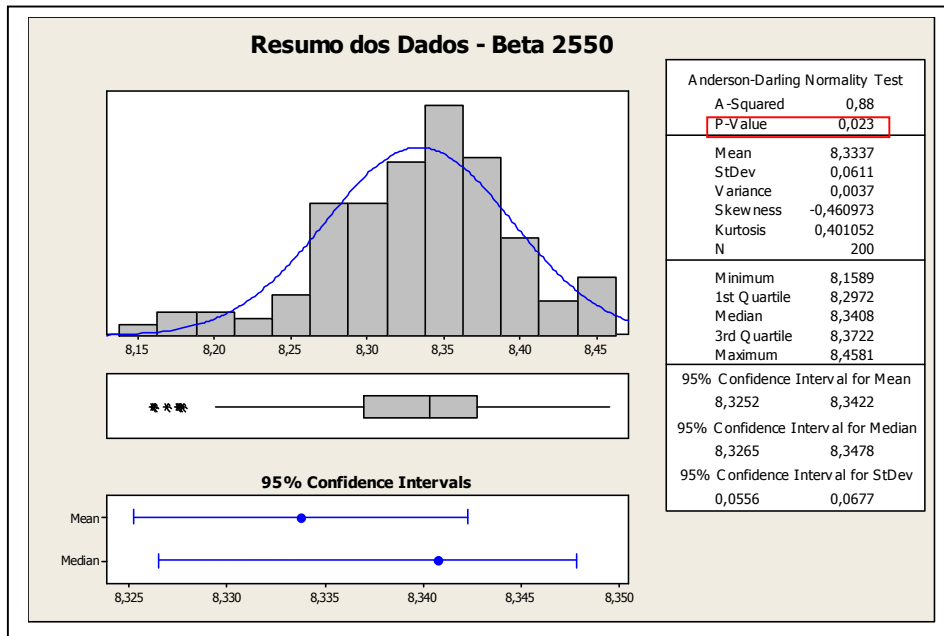


Figura 4.30 – Análise dos dados para “Beta 2550”

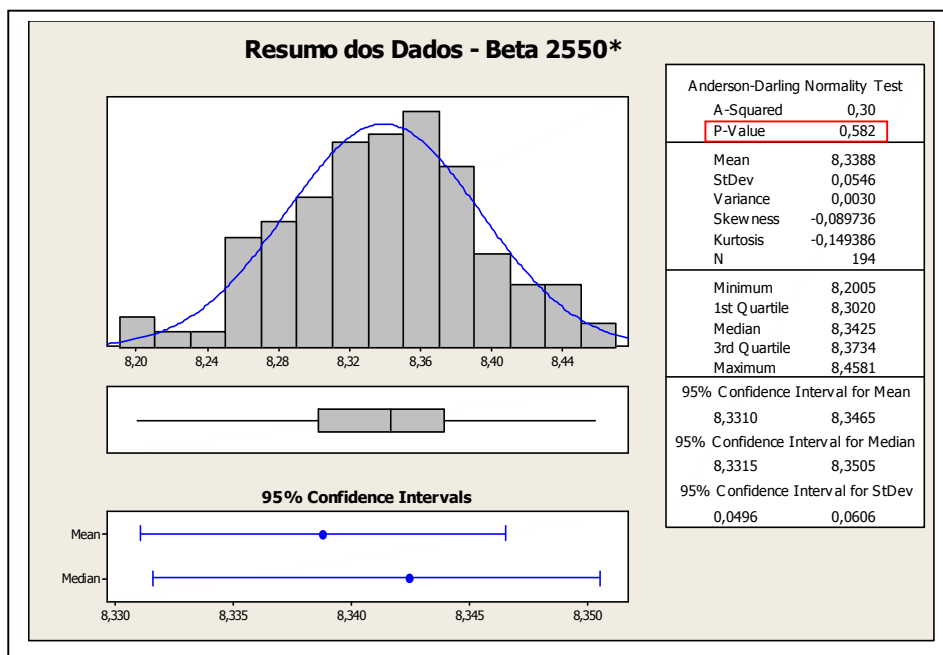


Figura 4.31 – Análise dos dados para “Beta 2550” sem *outliers*

4.2.6 Transformação de Dados Modelados pela Distribuição Beta

Para o conjunto de dados “Beta 2550” (Apêndice B) foi efetuada a transformação de Box-Cox. As características desta transformação são mostradas através da Figura 4.32 e a análise gráfica da distribuição Normal resultante, com *p-value* igual a 0,065, é apresentada na Figura 4.33.

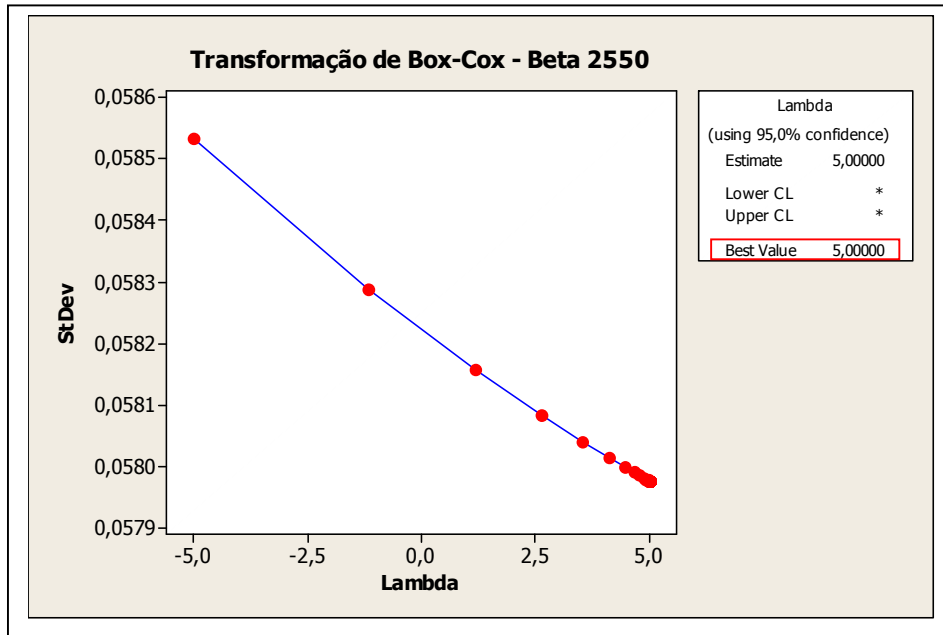


Figura 4.32 – Transformação de Box-Cox para os dados em “Beta 2550”

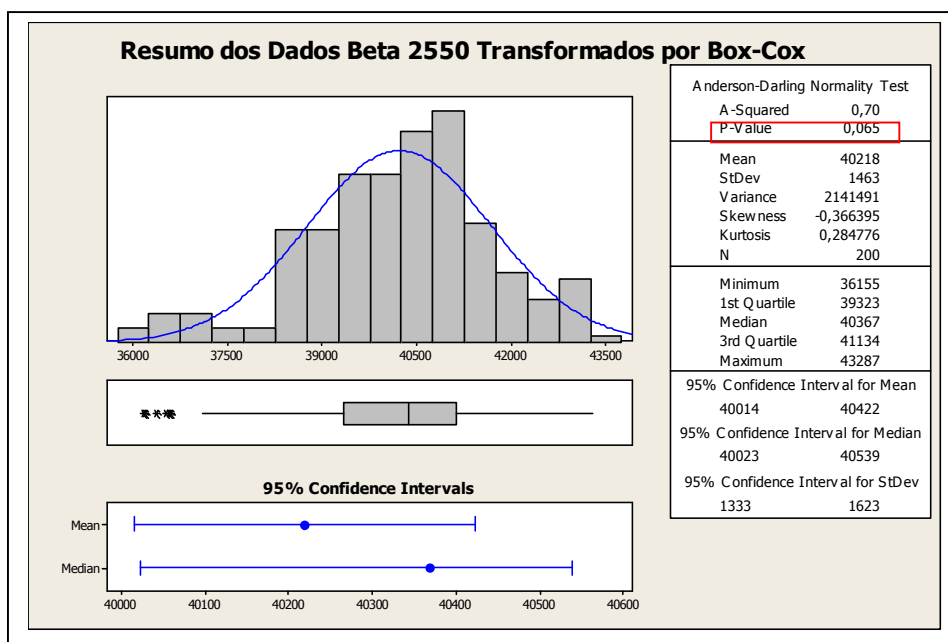


Figura 4.33 – Análise dos dados transformados por Box-Cox em “Beta 2550”

Para o conjunto de dados “Beta 2550” também foi efetuada aplicação da transformação de Johnson. As características desta transformação, a determinação da família de transformação e a correspondente equação são apresentadas na Figura 4.34.

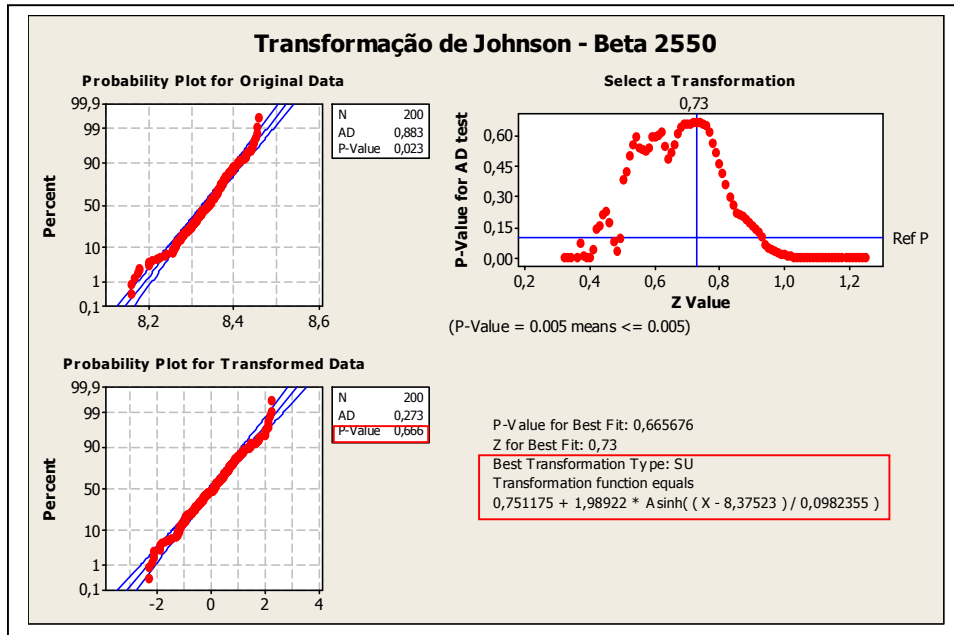


Figura 4.34 – Transformação de Johnson para os dados em “Beta 2550”

Após transformação de Johnson os dados passaram a ser descritos adequadamente pela distribuição Normal (*p-value* igual a 0,666) conforme mostra a Figura 4.35.

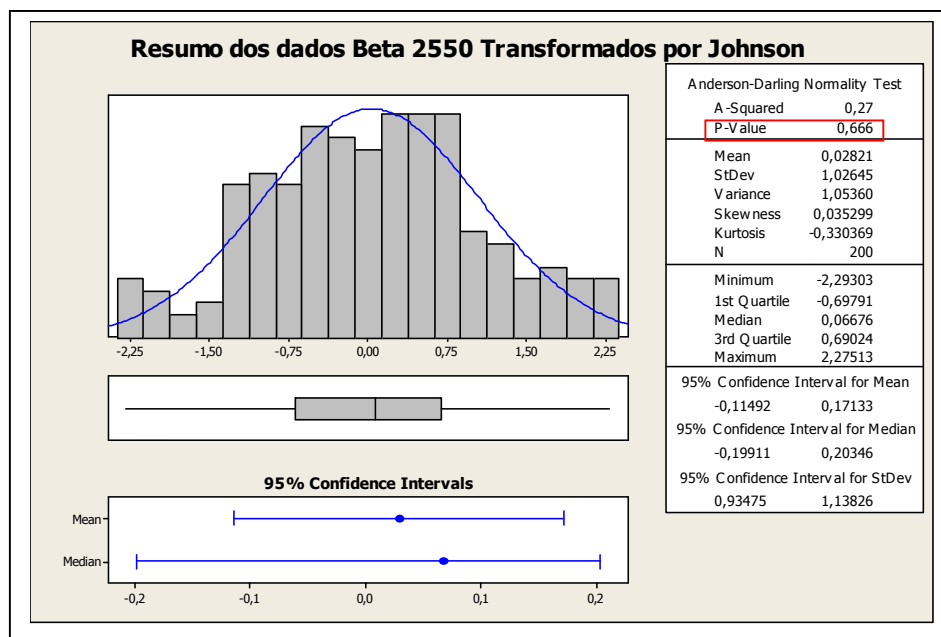


Figura 4.35 – Análise dos dados transformados por Johnson em “Beta 2550”

Como a análise do conjunto de dados “Beta 2550” sem a presença de *outliers* indicou que o mesmo passou a ser bem representado pela distribuição Normal, não foram necessárias as transformações de Box-Cox ou Johnson para esta condição.

4.2.7 Gráficos de Controle Individuais – Distribuição Beta

Durante avaliação dos dados em “Beta 2550” algumas formas de utilização de gráfico de controle para avaliação do processo em termos de medida do diâmetro do pino especial foram discutidas. Na primeira situação, representada pela Figura 4.36, os dados foram inseridos para análise no programa estatístico sem a preocupação de testar a aderência dos mesmos à distribuição Normal.

Analisando o gráfico da Figura 4.36 em relação ao primeiro critério de identificação de padrões não aleatórios, estabelecido no tópico 2.4.3, também conhecido como “teste 1”, que avalia a ocorrência de pontos além dos limites de controle estipulados, observou-se a existência de 1 (um) ponto localizado abaixo do limite inferior de controle. Este fato denotaria a presença de uma causa especial de variação, caso a suposição de normalidade estivesse correta.

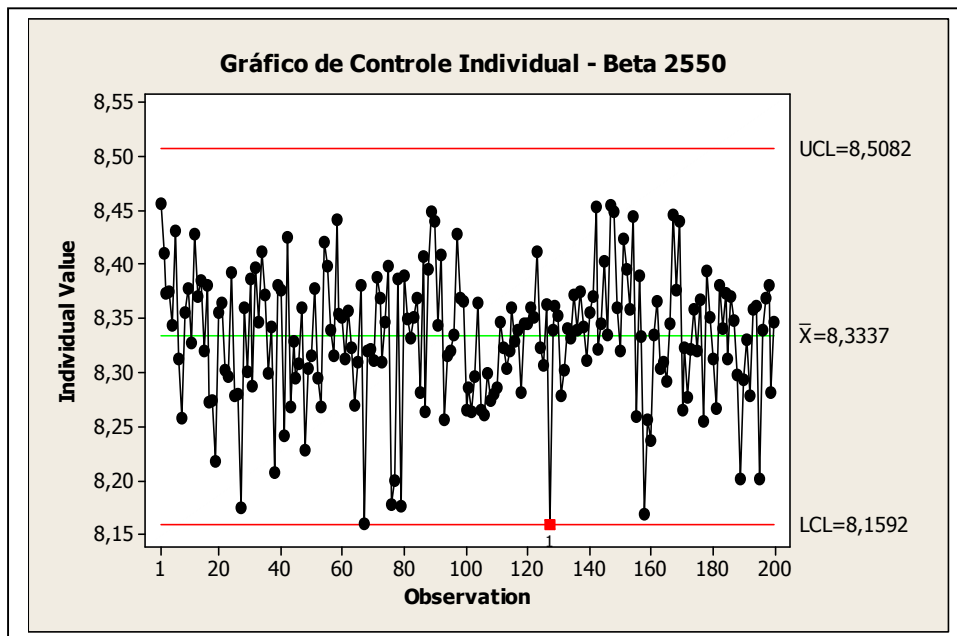


Figura 4.36 – Gráfico de controle para “Beta 2550” considerando distribuição Normal

Por outro lado, a análise do gráfico de controle individual para os dados em “Beta 2550” sem *outliers* (ou “Beta 2550*”) assumindo distribuição Normal, conforme demonstrado

na Figura 4.37, não indicou a existência de pontos além dos limites de controle. Importante ressaltar que neste caso a suposição de normalidade correspondia à realidade dos dados.

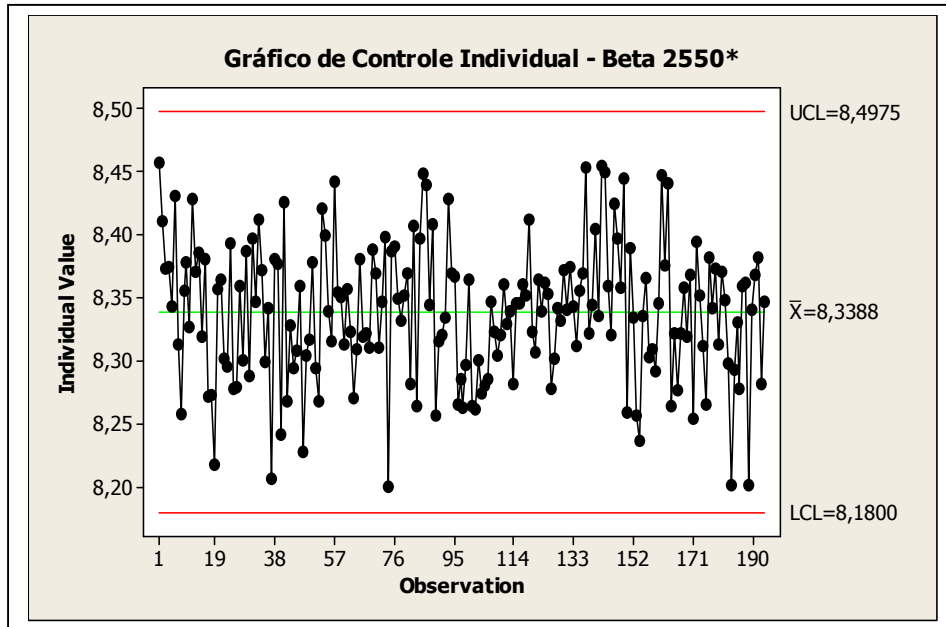


Figura 4.37 – Gráfico de controle “Beta 2550” sem outliers assumindo distribuição Normal

Na segunda forma de avaliar a condição de estabilidade dos dados em “Beta 2550” foram considerados percentil 0,135 (8,085346) para LIC, percentil 50 ou mediana (8,340784) para a LM e percentil 99,865 (8,471053) para LSC conforme indicado na Figura 4.38.

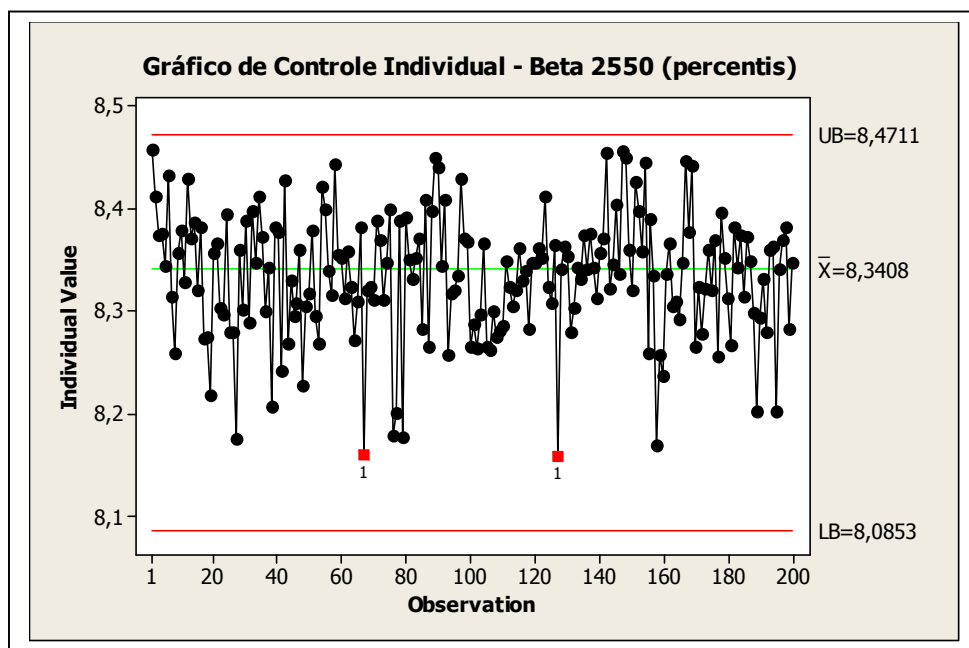


Figura 4.38 – Gráfico de controle “Beta 2550” com limites de controle em percentis

Na terceira abordagem o gráfico de controle individual foi obtido após prévia transformação dos dados utilizando a transformação de Box-Cox conforme mostrado na Figura 4.39. Neste caso não foram observados pontos além dos limites de controle.

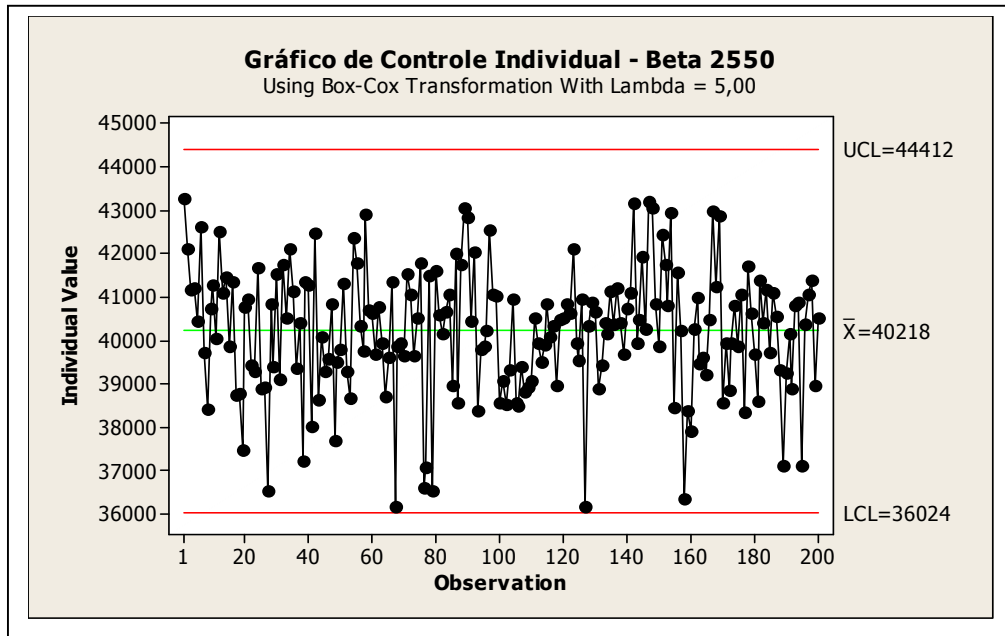


Figura 4.39 – Gráfico de controle para “Beta 2550” após transformação de Box-Cox

A quarta situação de análise com o gráfico de controle individual após a transformação de Johnson, também não apresentou pontos além dos limites estipulados (Figura 4.40).

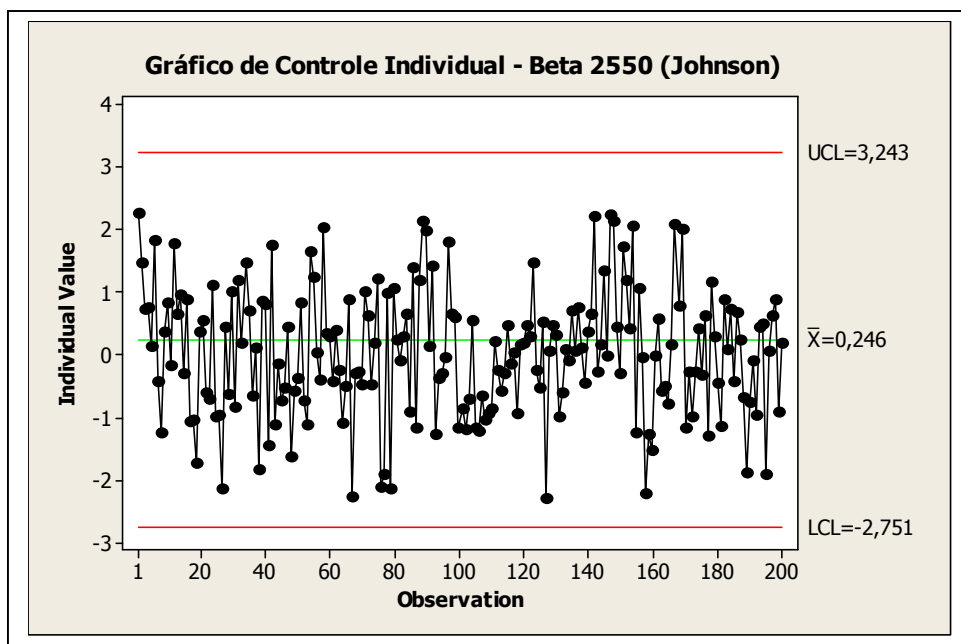


Figura 4.40 – Gráfico de controle para “Beta 2550” após transformação de Johnson

Como a análise do conjunto de dados “Beta 2550” sem a presença de *outliers* (“Beta 2550*”) indicou que o mesmo passou a ser bem representado pela distribuição Normal, não foram necessárias as verificações dos gráficos de controle individuais após as transformações de Box-Cox ou Johnson para esta condição.

4.2.8 Avaliação de Capabilidade – Distribuição Beta

A avaliação de capacidade em “Beta 2550” teve como requisito um limite inferior de especificação de 8,20 mm e um limite superior de especificação de 8,50 mm para a medida do diâmetro do pino especial. Foram calculados diversos índices de capacidade Z_{bench} (Z_{LT} e Z_{ST}) diferentes:

- a) Em uma primeira abordagem o cálculo foi efetuado com os dados originais e os dados tratados (sem *outliers*) como se fossem normalmente distribuídos;
- b) No segundo caso a condição assimétrica da distribuição foi considerada no cálculo dos dados originais.
- c) Na terceira abordagem os dados originais foram previamente transformados por meio da transformação de Box-Cox;
- d) Na quarta situação foi utilizada a transformação de Johnson.

A Figura 4.41 apresenta os resultados do cálculo de capacidade para “Beta 2550” considerando os dados normalmente distribuídos.

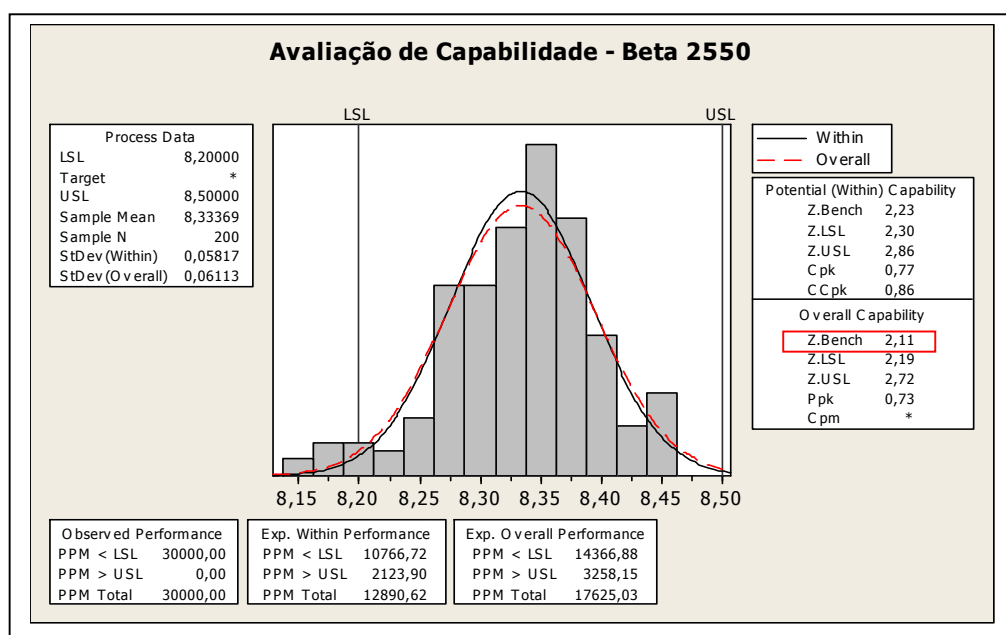


Figura 4.41 – Índices de capacidade de “Beta 2550” considerando distribuição Normal

A Figura 4.42 apresenta os resultados com os dados originais “Beta 2550” sem a presença de *outliers*, considerando distribuição Normal, suposição que se mostrou verdadeira.

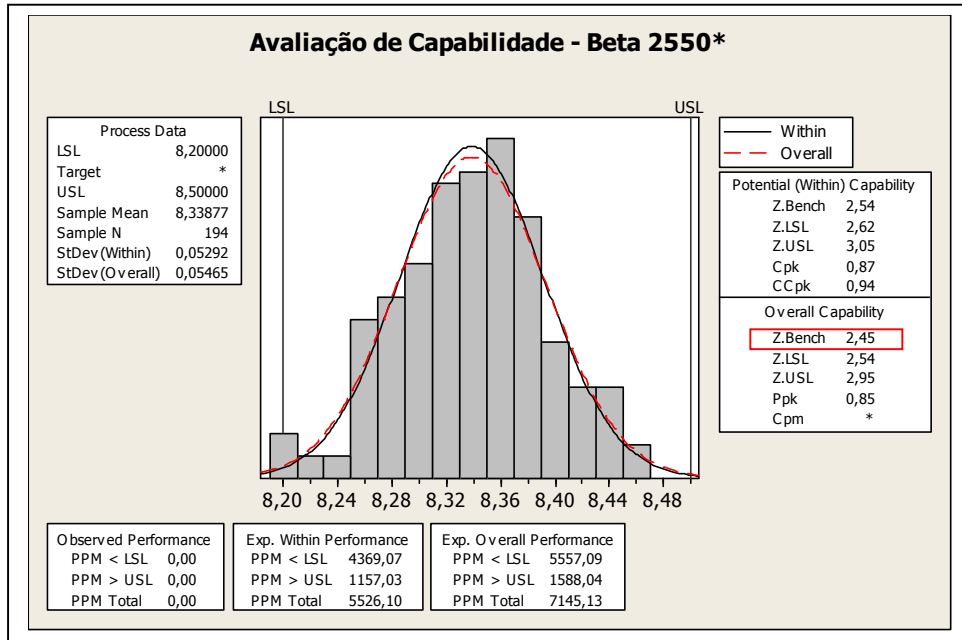


Figura 4.42 – Índices de capacidade de “Beta 2550” sem *outliers* assumindo normalidade

Por não seguir a distribuição Normal, o cálculo de capacidade dos dados em “Beta 2550” foi efetuado considerando a distribuição Weibull com 3 parâmetros (Figura 4.43), pois não havia a opção de cálculo com distribuição Beta no programa estatístico utilizado.

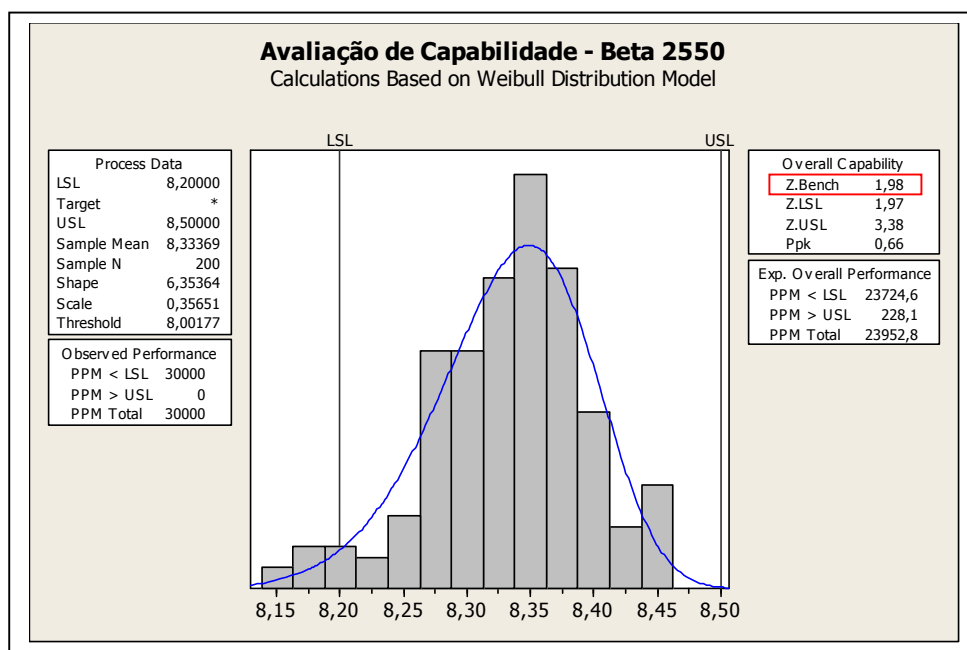


Figura 4.43 – Índices de capacidade de “Beta 2550” considerando distribuição Weibull

Não foi necessária a verificação de capacidade após adequação de outra distribuição aos dados “Beta 2550” sem outliers, pois os mesmos já eram normalmente distribuídos.

A Figura 4.44 apresenta os resultados do cálculo de capacidade do processo após utilização da transformação de Box-Cox e a Figura 4.45 apresenta os mesmos resultados após utilização da transformação de Johnson.

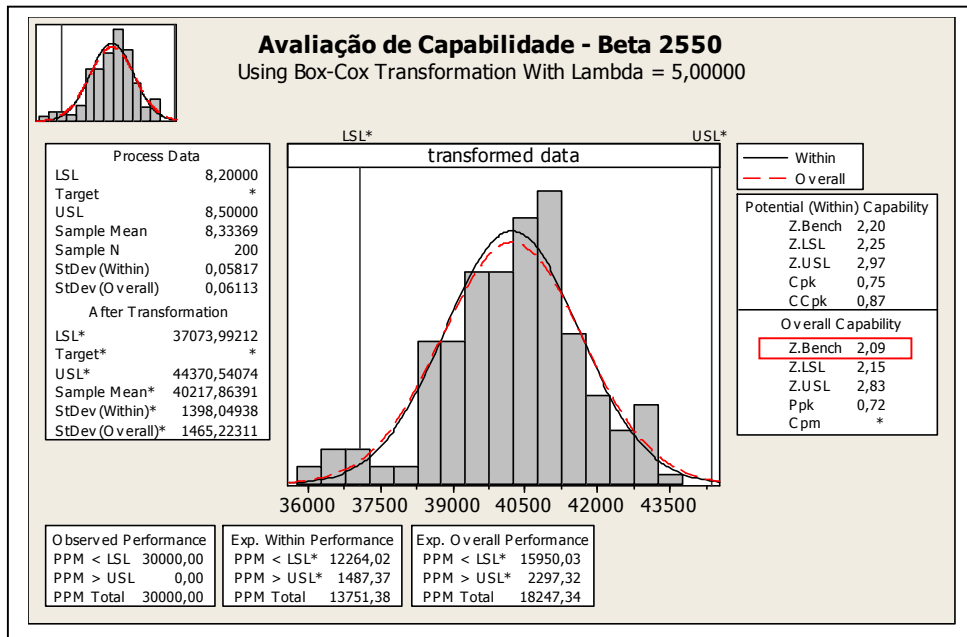


Figura 4.44 – Índices de capacidade de “Beta 2550” após transformação de Box-Cox

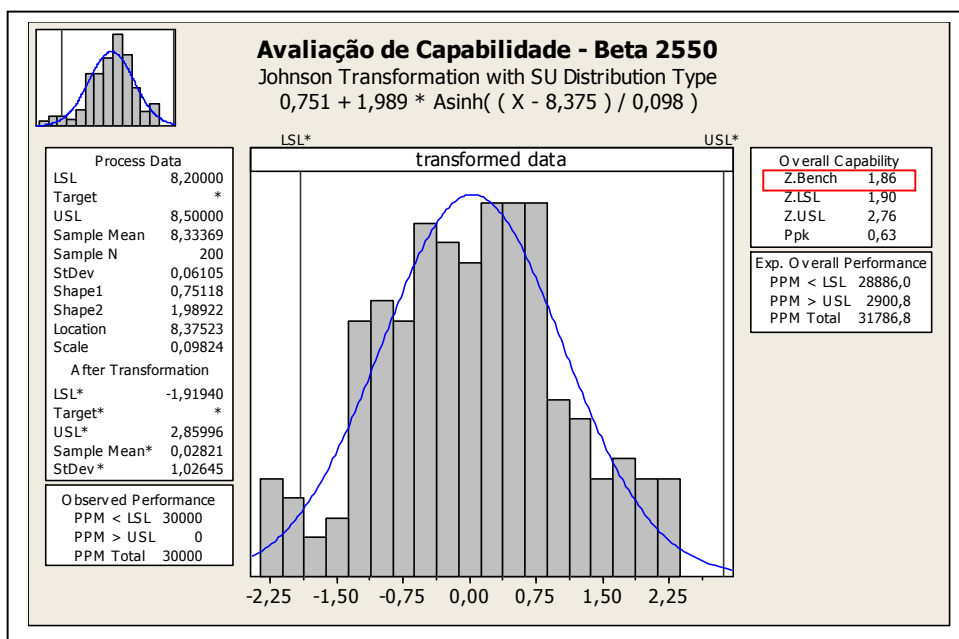


Figura 4.45 – Índices de capacidade de “Beta 2550” após transformação de Johnson

Não foi necessária a verificação de capacidade após transformação de Box-Cox ou Johnson para o conjunto de dados “Beta 2550” sem *outliers* (“Beta 2550*”), pois os mesmos já eram normalmente distribuídos.

A Tabela 4.2 apresenta a comparação dos valores obtidos nas várias situações apresentadas, com o correspondente cálculo do nível sigma e a fração não conforme em partes por milhão para o conjunto de dados “Beta 2550”, considerando Z_{shift} igual a 1,5 sigma.

PREMISSA ADOTADA PARA CÁLCULO	LIE (mm)	LSE (mm)	ÍNDICE Z_{LT}	NÍVEL SIGMA ($Z_{LT}+1,5$)	PPM
Dados originais considerados normais	8,20	8,50	2,11	3,61	17625
Dados sem <i>outliers</i> considerados normais	8,20	8,50	2,45	3,95	7145
Dados originais considerando distribuição Weibull	8,20	8,50	1,98	3,48	23953
Dados originais transformados por Box-Cox	8,20	8,50	2,09	3,59	18247
Dados originais transformados por Johnson	8,20	8,50	1,86	3,36	31787

Tabela 4.2 – Cálculos comparativos do nível sigma para “Beta 2550”

4.3 Pesquisa Exploratória Através de Estudo de Caso

4.3.1 Análise de Dados Reais – Furos 3/4”

Antes da aplicação dos métodos estatísticos desta pesquisa exploratória foi necessário efetuar uma análise dos dados disponíveis em “3/4AE” e “3/4AD” quanto a sua aderência à distribuição Normal.

A partir dos dados reais relativos aos valores de diâmetro dos furos do componente estrutural em “3/4AE” (Apêndice C) foram efetuados a análise da distribuição e o teste de normalidade utilizando os recursos disponíveis do programa estatístico, sendo que a verificação de normalidade das distribuições estudadas nesta pesquisa baseou-se na análise do valor P (*p-value*) associado à estatística A^2 do teste de normalidade de Anderson-Darling.

A Figura 4.46 apresenta o resultado da análise dos dados em “3/4AE”. Admitindo $\alpha = 0,05$, a distribuição em questão foi considerada não normal visto que o *p-value* obtido (0,005) resultou menor que 0,05. Assim sendo, como a hipótese nula de normalidade foi rejeitada, obviamente a distribuição em questão não pôde ser considerada normalmente distribuída.

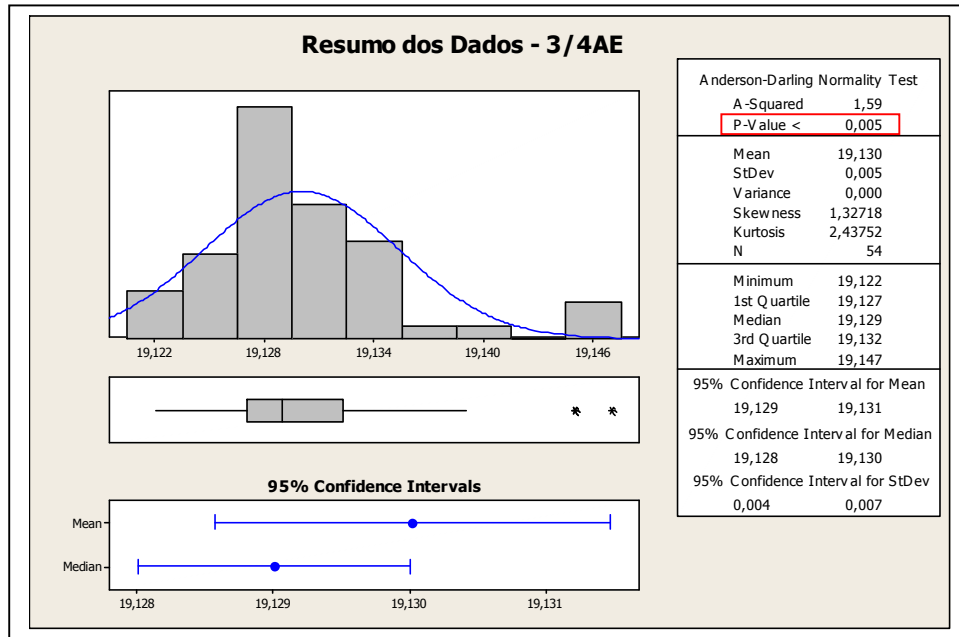


Figura 4.46 – Análise do conjunto de dados originais “3/4AE”

A análise do conjunto de dados “3/4AE” também indicou a existência de pontos extremos fora do padrão seguido pelo restante dos dados (*outliers*). Através da análise dos dados “3/4AE” sem a presença dos *outliers*, apresentada na Figura 4.47 como “3/4AE*”, concluiu-se que a distribuição resultante poderia ser bem representada pela distribuição Normal, pois o *p-value* encontrado (0,301) resultou superior a 0,05.

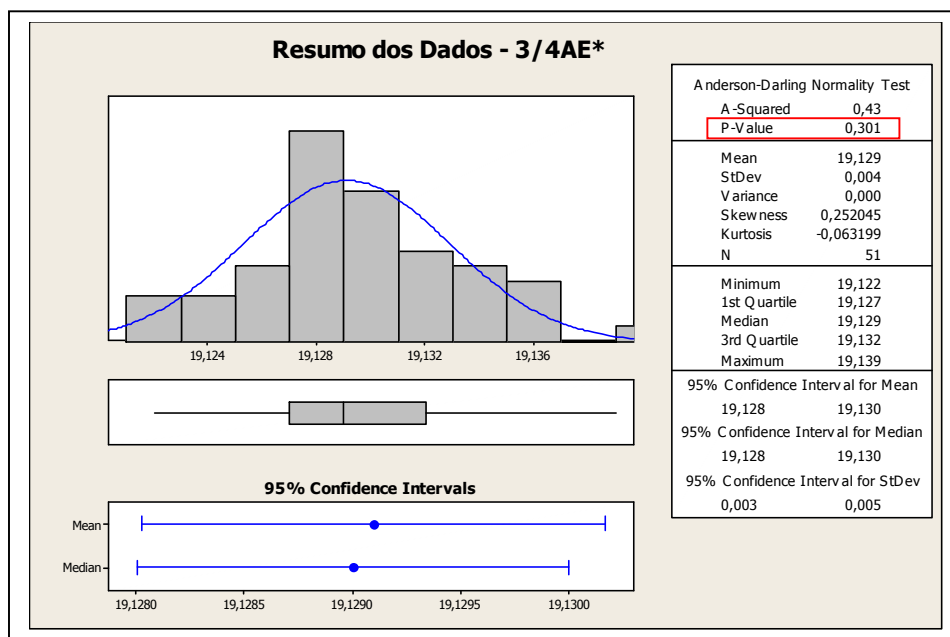


Figura 4.47 – Análise do conjunto de dados “3/4AE” sem *outliers*

Como o conjunto de dados “3/4AE” não seguia a distribuição Normal, foi efetuada a verificação da distribuição que melhor representasse estes dados. Conforme mostrado na Figura 4.48, a distribuição Loglogistic com 3 parâmetros foi a melhor opção encontrada, pois indicou o maior coeficiente de correlação (0,951) entre as distribuições analisadas para um intervalo de 95% de confiança.

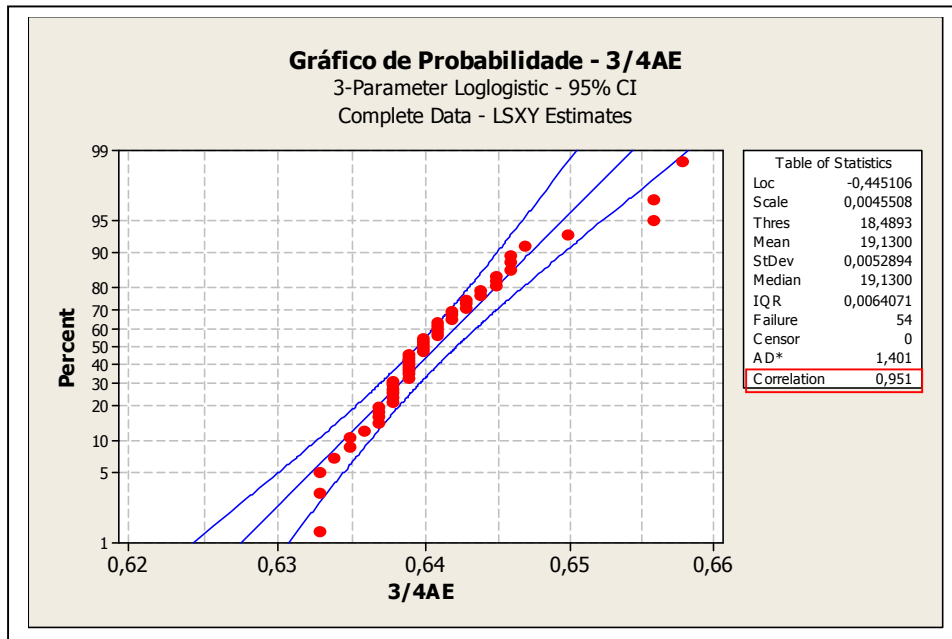


Figura 4.48 – Avaliação da melhor distribuição aplicável ao conjunto de dados “3/4AE”

A partir dos dados reais relativos aos valores de diâmetro dos furos do componente estrutural em “3/4AD” (Apêndice C) foram efetuados a análise da distribuição e o teste de normalidade de Anderson-Darling. Conforme argumentação indicada no capítulo 3, a verificação de normalidade das distribuições estudadas nesta pesquisa baseou-se na análise do valor P (*p-value*) associado à estatística A^2 do teste de normalidade de Anderson-Darling.

A Figura 4.49 apresenta o resultado da análise dos dados em “3/4AD”. Admitindo $\alpha = 0,05$, a distribuição em questão foi considerada não normal visto que o *p-value* obtido (0,034) resultou menor que 0,05. Portanto, como a hipótese nula de normalidade foi rejeitada, obviamente a distribuição em questão não pôde ser considerada normalmente distribuída.

A análise do conjunto de dados “3/4AD” também indicou a existência de *outliers*. Através da análise do conjunto de dados “3/4AD” sem a presença dos *outliers*, apresentada na Figura 4.50 como “3/4AD*”, concluiu-se que a distribuição resultante não era normal, pois o valor de *p-value* permaneceu inferior a 0,05.

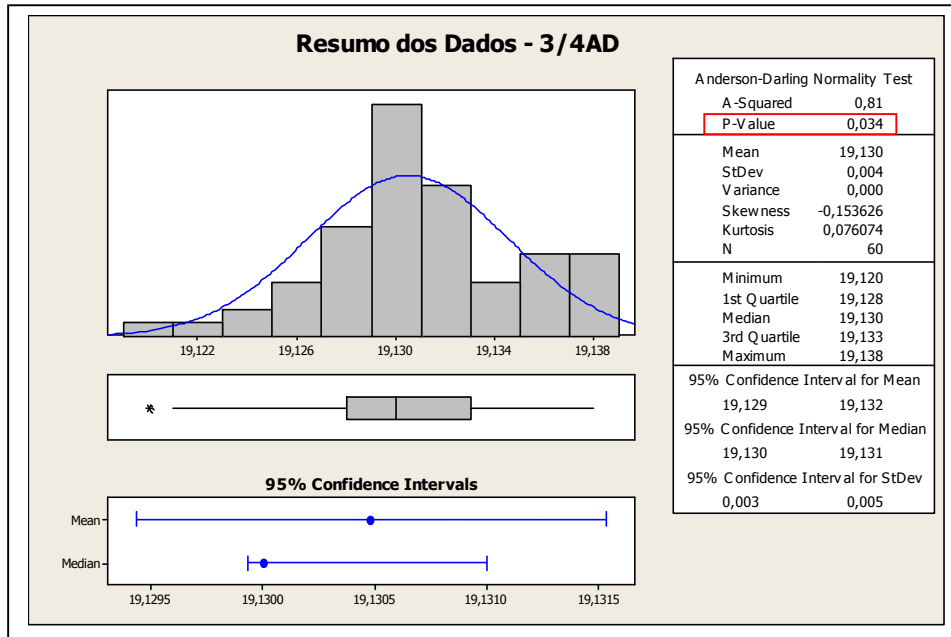


Figura 4.49 – Análise do conjunto de dados originais “3/4AD”

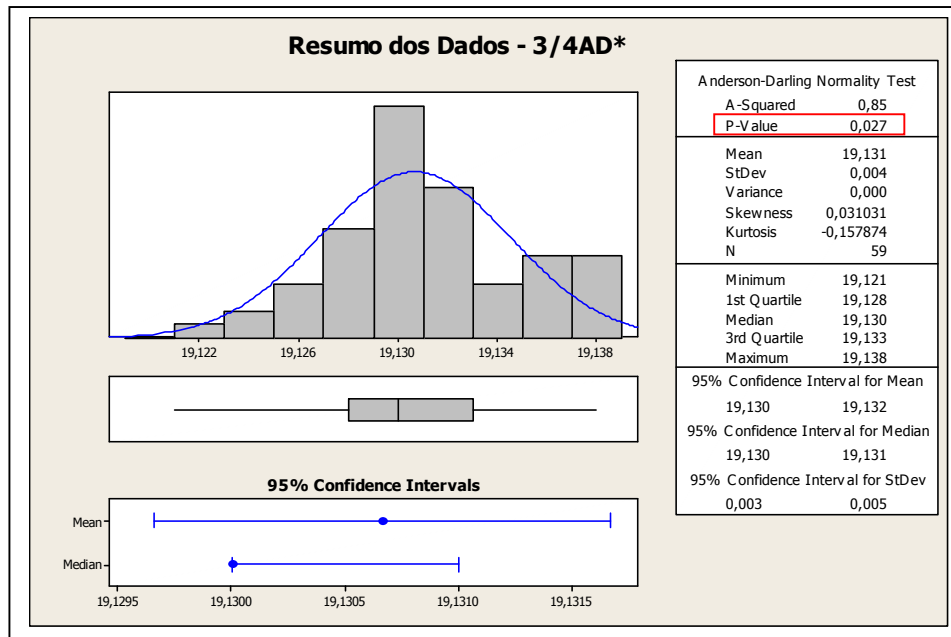


Figura 4.50 – Análise do conjunto de dados “3/4AD” sem outliers

Como o conjunto de dados “3/4AD” não seguia a distribuição Normal, foi efetuada a verificação da distribuição que melhor representasse estes dados. Conforme indicado através da Figura 4.51, a distribuição Logistic foi a melhor opção encontrada, pois indicou o maior coeficiente de correlação (0,983) entre as distribuições analisadas para um intervalo de 95% de confiança.

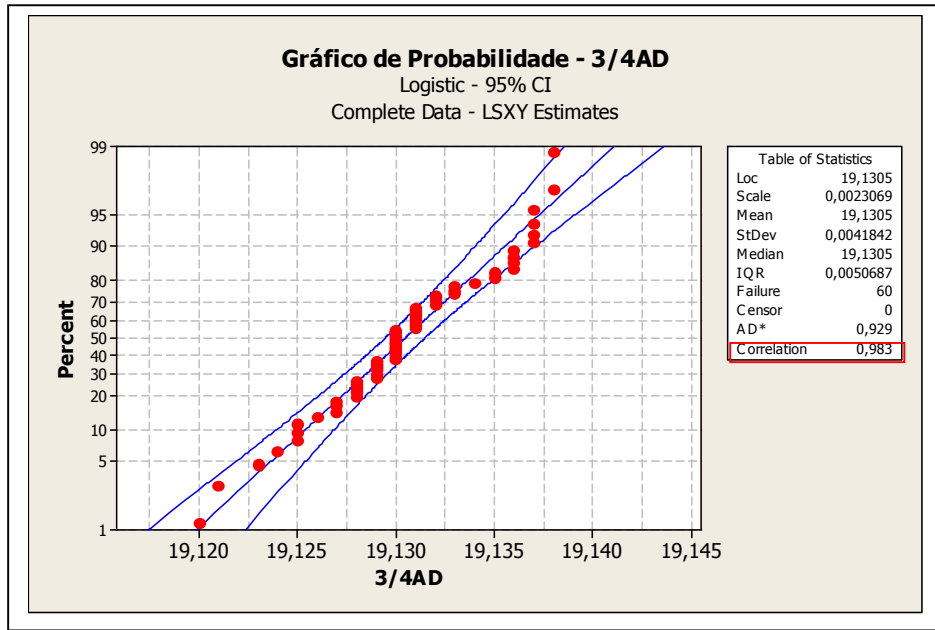


Figura 4.51 – Avaliação da melhor distribuição aplicável ao conjunto de dados “3/4AD”

Como o conjunto de dados “3/4AD” sem outliers (“3/4AD*”) também não seguia a distribuição Normal, foi efetuada a verificação da distribuição que melhor representasse estes dados. Conforme mostrado na Figura 4.52, a distribuição Lognormal com 3 parâmetros foi a melhor opção encontrada, pois indicou o maior coeficiente de correlação (0,984) entre as distribuições analisadas para um intervalo de 95% de confiança.

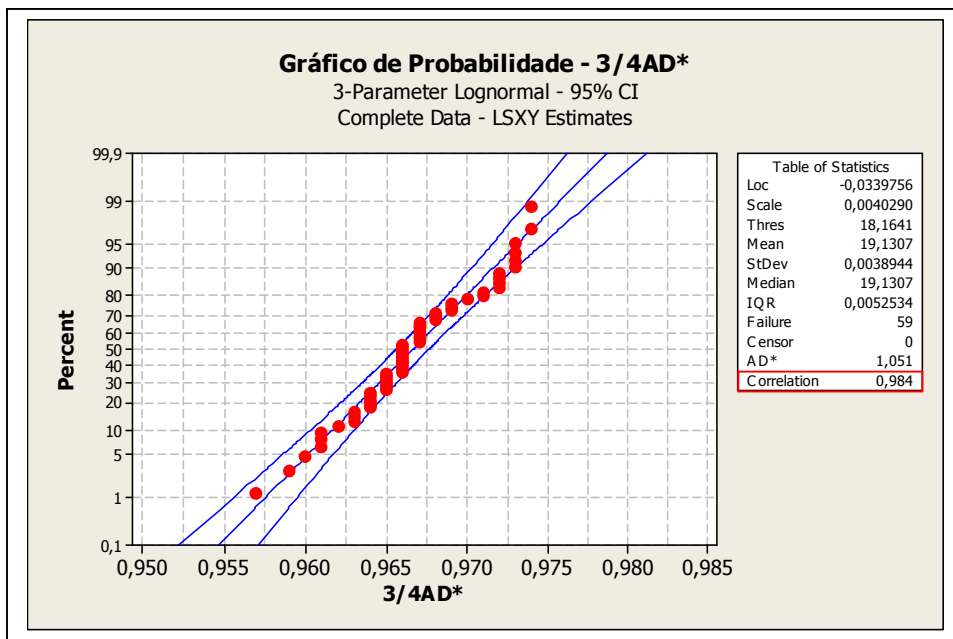


Figura 4.52 – Avaliação de distribuição aplicável ao conjunto de dados “3/4AD” sem outliers

4.3.2 Transformação de Dados Reais – Furos 3/4”

Para o conjunto de dados “3/4AE” foi efetuada aplicação da transformação de Box-Cox conforme indicado na Figura 4.53. Após transformação de Box-Cox, admitindo $\alpha = 0,05$, os dados obtidos não foram considerados normalmente distribuídos, visto que *p-value* (0,005) resultou menor que 0,05 (Figura 4.54).

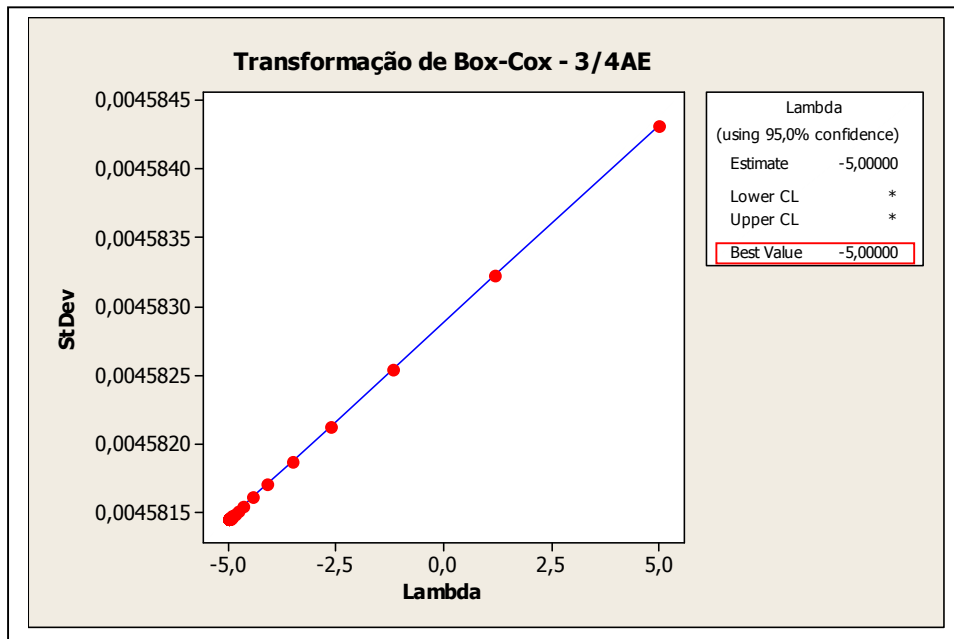


Figura 4.53 – Transformação de Box-Cox para o conjunto de dados “3/4AE”

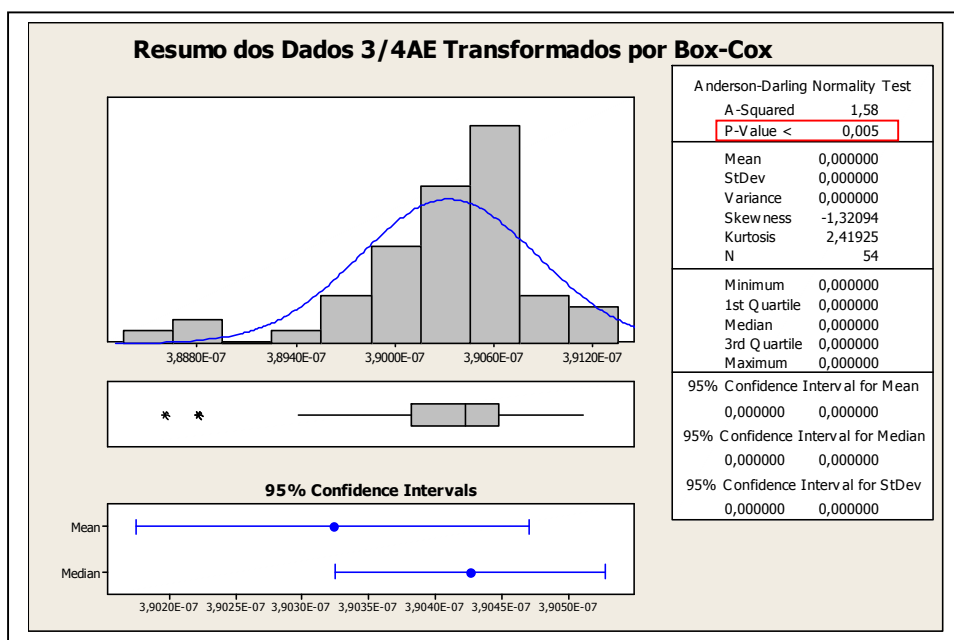


Figura 4.54 – Análise do conjunto de dados “3/4AE” transformados por Box-Cox

Para o conjunto de dados “3/4AE” também foi efetuada aplicação da transformação de Johnson conforme mostrado na Figura 4.55.

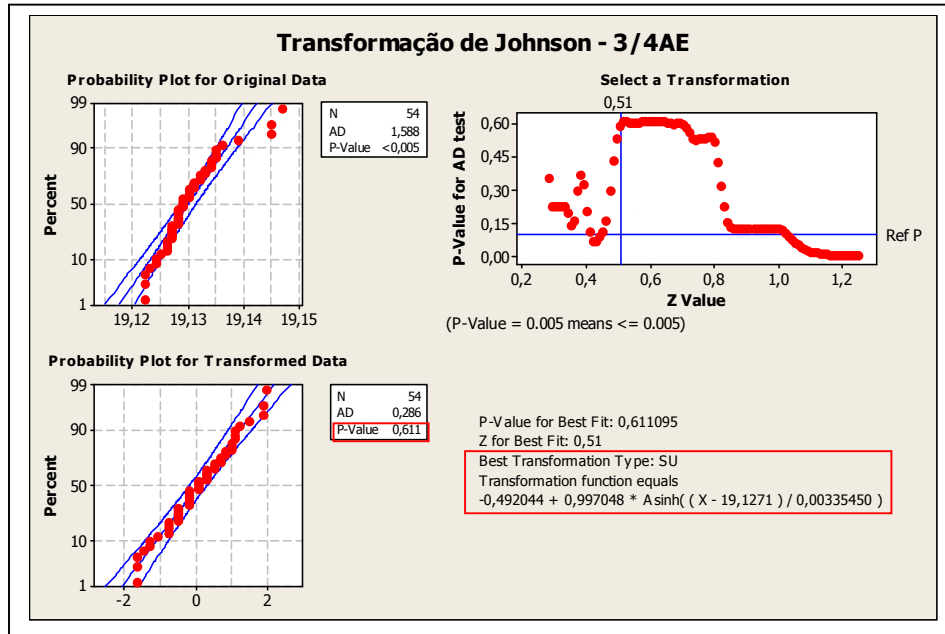


Figura 4.55 – Transformação de Johnson para o conjunto de dados “3/4AE”

Como a análise do conjunto de dados “3/4AE” sem a presença de outliers indicou que o mesmo era normalmente distribuído, não foram efetuadas as transformações nesta condição. A tentativa de transformação por Box-Cox para os dados “3/4AD” é mostrada na Figura 4.56.

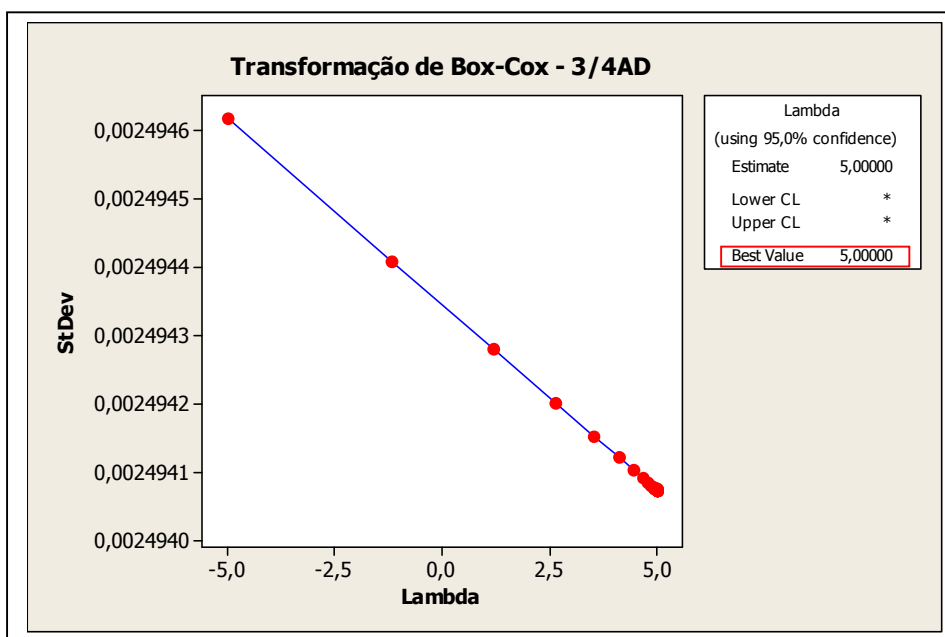


Figura 4.56 – Transformação de Box-Cox para o conjunto de dados “3/4AD”

Após transformação de Box-Cox, admitindo $\alpha = 0,05$, os dados obtidos não foram considerados normalmente distribuídos, visto que *p-value* (0,034) resultou menor que 0,05 conforme indicado através da Figura 4.57.

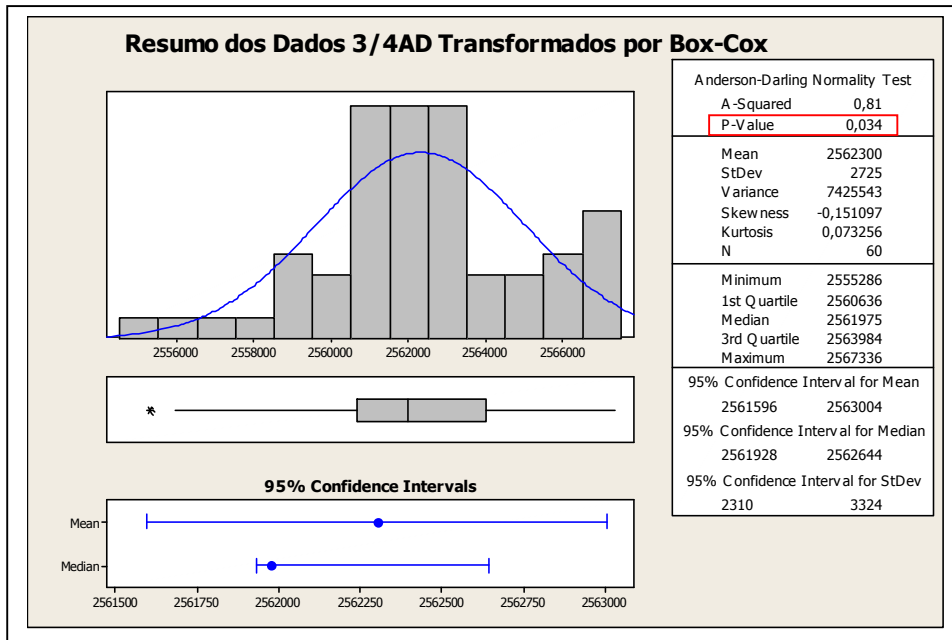


Figura 4.57 – Análise do conjunto de dados “3/4AD” transformados por Box-Cox

Para o conjunto de dados “3/4AD” também foi efetuada aplicação da transformação de Johnson (Figura 4.58), não sendo encontrada equação que resultasse em normalidade.

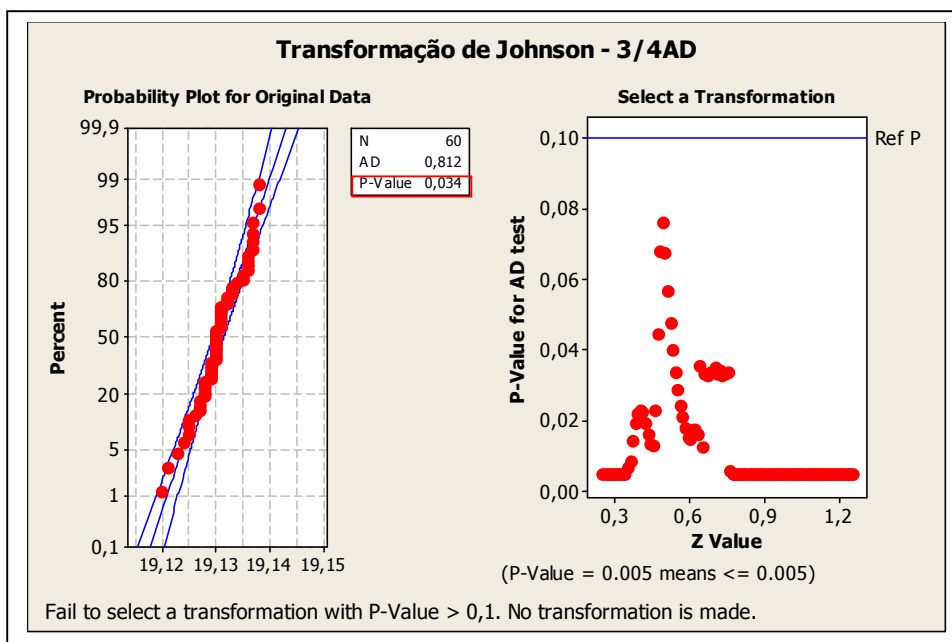


Figura 4.58 – Transformação de Johnson para o conjunto de dados “3/4AD”

Para o conjunto de dados “3/4AD” sem *outliers* (ou “3/4AD*”) foi tentada a transformação dos dados para condição de normalidade através da aplicação do método de Box-Cox conforme indicado através da Figura 4.59.

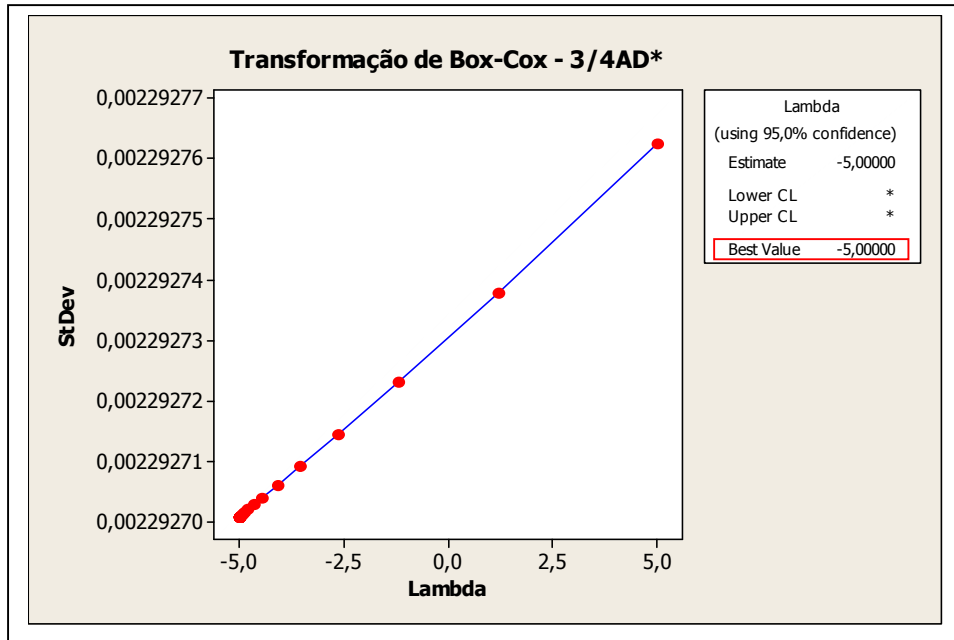


Figura 4.59 – Transformação de Box-Cox para o conjunto de dados “3/4AD” sem *outliers*

Para o conjunto de dados “3/4AD” sem *outliers* também foi efetuada a transformação de Johnson (Figura 4.60), não sendo encontrada equação que resultasse em normalidade.

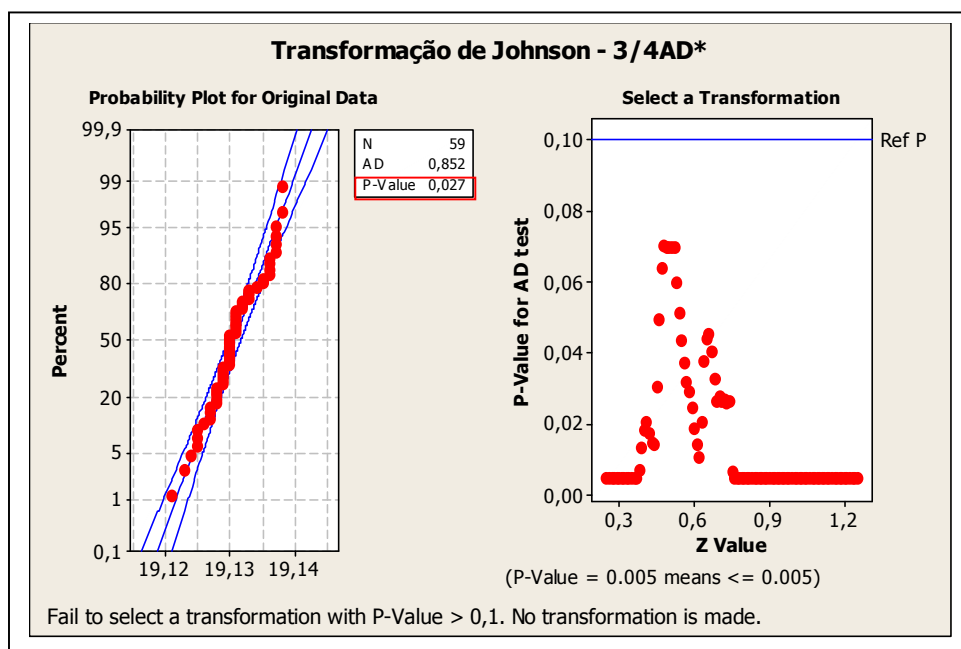


Figura 4.60 – Transformação de Johnson para o conjunto de dados “3/4AD” sem *outliers*

4.3.3 Cálculo de Capabilidade de Dados Reais – Furos 3/4”

A avaliação de capabilidade para os conjuntos de dados “3/4AE” e “3/4AD” (Apêndice C) teve como requisito um limite inferior de especificação de 19,124 mm e um limite superior de especificação de 19,151 mm para as medidas de diâmetro dos furos do componente estrutural. Foram calculados diversos índices de capabilidade z_{bench} (Z_{LT} e Z_{ST}) diferentes:

- Em uma primeira abordagem o cálculo foi efetuado com os dados originais e os dados tratados (sem *outliers*) como se fossem normalmente distribuídos;
- No segundo caso a condição não normal da distribuição foi considerada no cálculo dos dados originais e, quando aplicável, nos dados sem a presença de *outliers*.
- Na terceira abordagem os dados originais foram modificados por meio da transformação de Box-Cox;
- Na quarta situação foi utilizada a transformação de Johnson.

A Figura 4.61 apresenta os resultados do cálculo de capabilidade considerando os dados originais “3/4AE” como se fossem normalmente distribuídos.

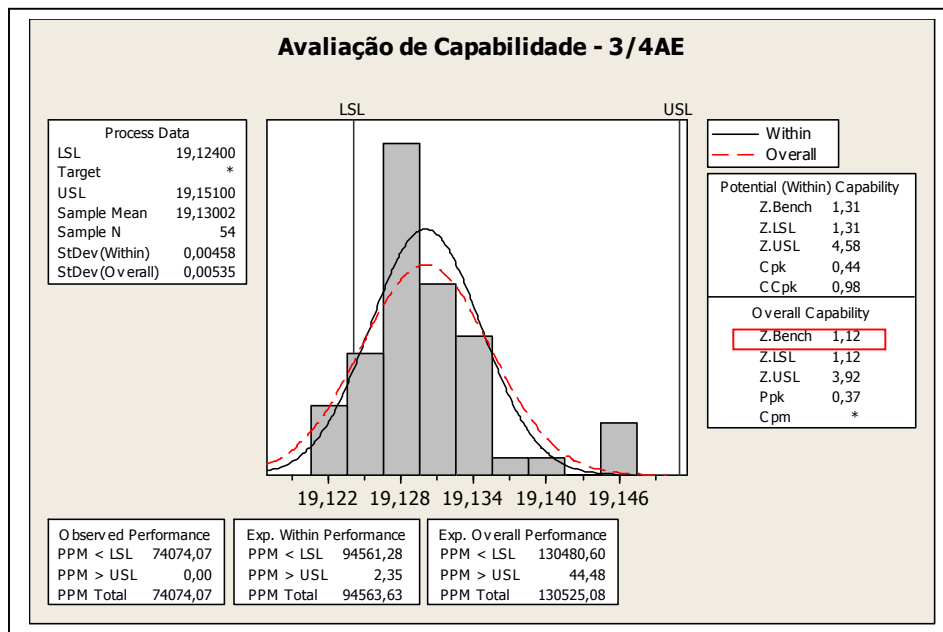


Figura 4.61 – Índices de capabilidade do conjunto de dados originais “3/4AE”

Como o conjunto de dados “3/4AE” não se mostrou normalmente distribuído, o cálculo de capabilidade foi efetuado a partir da distribuição Loglogistic com 3 parâmetros, correspondente ao melhor ajuste para os dados de “3/4AE” em um intervalo de 95% de confiança. Os resultados nesta situação são apresentados através da Figura 4.62.

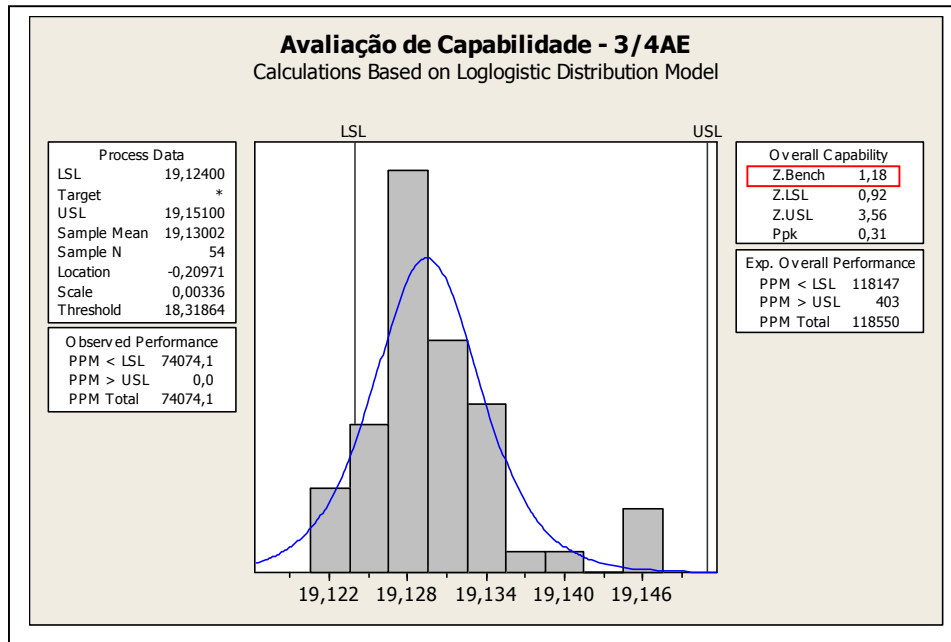


Figura 4.62 – Índices de capabilidade dos dados “3/4AE” baseados na distribuição Loglogistic

Não foi efetuado cálculo de capabilidade após transformação de Box-Cox dos dados “3/4AE”, pois a mesma não resultou em normalidade.

A Figura 4.63 apresenta os resultados do cálculo de capabilidade do processo após utilização da transformação de Johnson no conjunto de dados originais “3/4AE”.

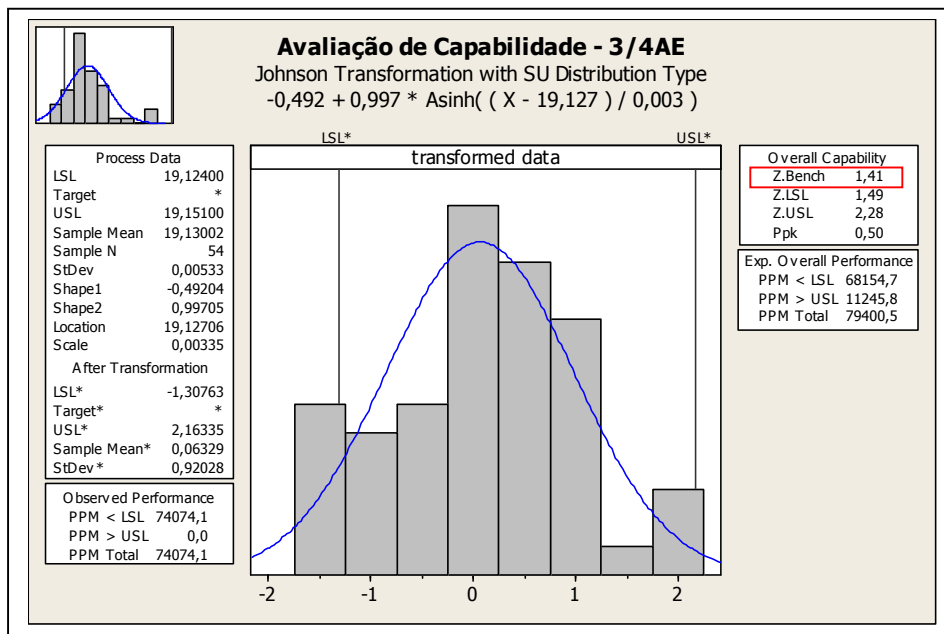


Figura 4.63 – Índices de capabilidade dos dados “3/4AE” transformados por Johnson

A Figura 4.64 apresenta os resultados considerando os dados originais “3/4AE” sem a presença de *outliers*, considerando distribuição Normal, suposição que se mostrou verdadeira.

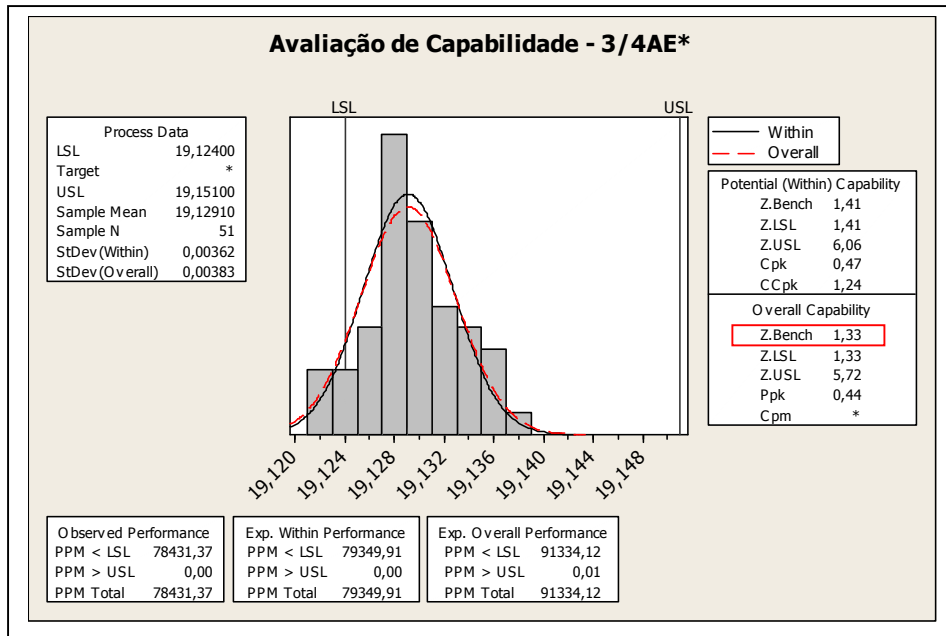


Figura 4.64 – Índices de capacidade do conjunto de dados “3/4AE” sem *outliers*

Não foram efetuados os cálculos de capacidade após ajuste da distribuição e transformações em “3/4AE” sem *outliers*, pois o mesmo já era normal. A Figura 4.65 mostra o cálculo de capacidade para os dados “3/4AD” como se fossem normalmente distribuídos.

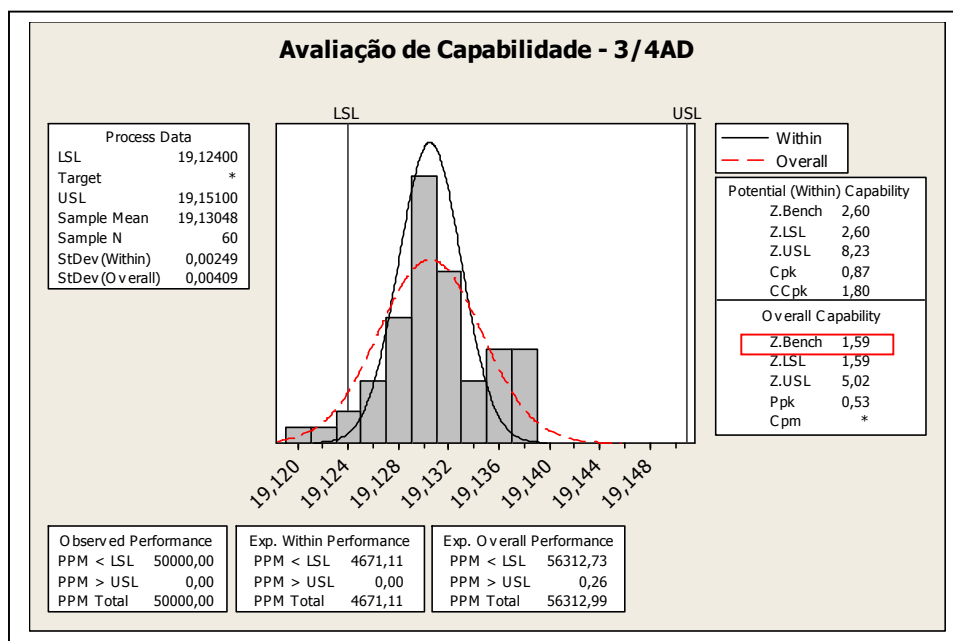


Figura 4.65 – Índices de capacidade do conjunto de dados originais “3/4AD”

Como o conjunto de dados “3/4AD” não se mostrou normalmente distribuído, o cálculo de capacidade foi efetuado a partir da distribuição Logistic, correspondente ao melhor ajuste para os dados de “3/4AD” em um intervalo de 95% de confiança (Figura 4.66).

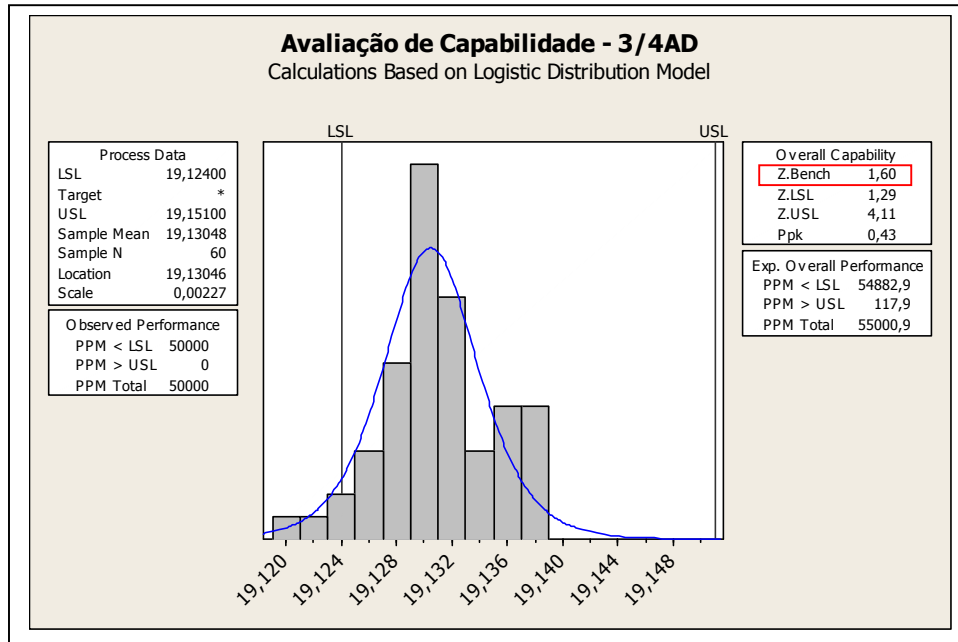


Figura 4.66 – Índices de capacidade dos dados “3/4AD” baseados na distribuição Logistic

A Figura 4.67 apresenta os resultados considerando os dados originais “3/4AD” sem a presença de outliers e considerando os dados como se fossem normalmente distribuídos.

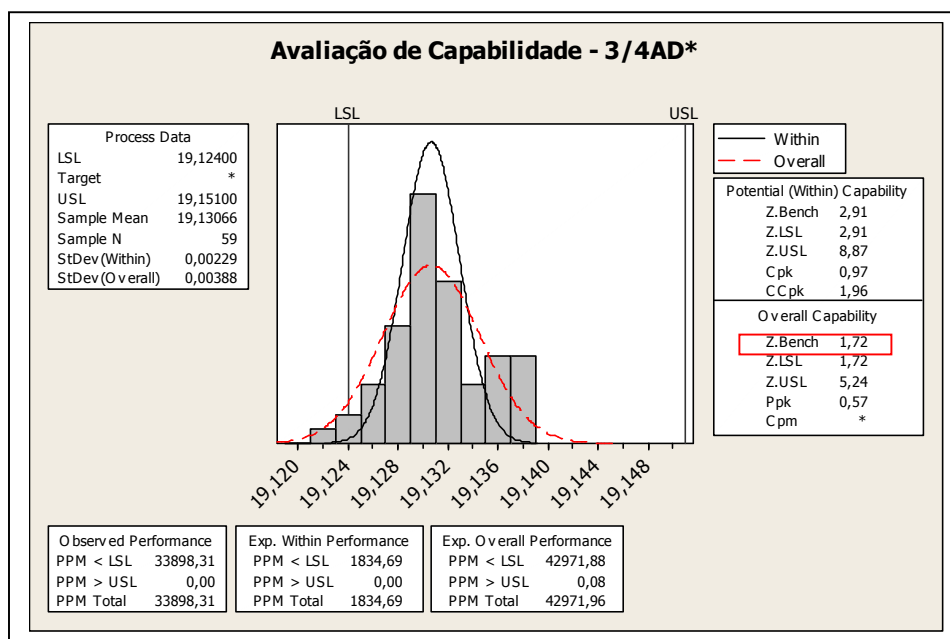


Figura 4.67 – Índices de capacidade do conjunto de dados “3/4AD” sem outliers

Como o conjunto de dados “3/4AD” sem *outliers* não se mostrou normalmente distribuído, o cálculo de capacidade foi efetuado a partir da distribuição Lognormal com 3 parâmetros, conforme indicado através da Figura 4.68.

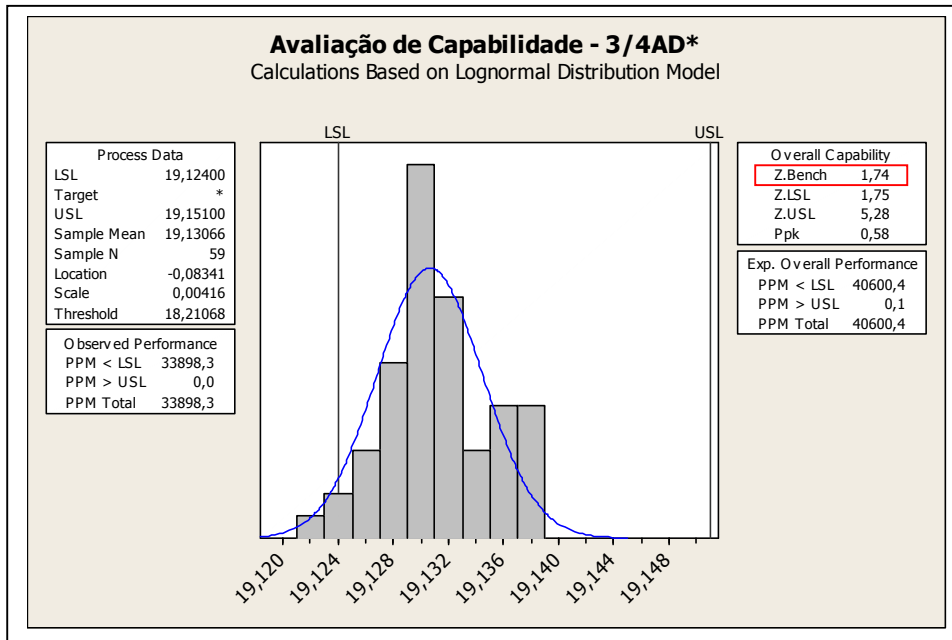


Figura 4.68 – Índices capacidade dos dados “3/4AD” sem *outliers* (distribuição Lognormal)

Não foram efetuados os cálculos de capacidade após as transformações de Box-Cox e Johnson para os dados “3/4AD” e para “3/4AD” sem *outliers*, pois as transformações citadas não resultaram em normalidade. Um resumo dos valores obtidos nas diversas situações descritas para os dados reais “3/4AE” e “3/4AD” é apresentado na Tabela 4.3.

PREMISSA ADOTADA PARA CÁLCULO	LIE (mm)	LSE (mm)	ÍNDICE Z_{LT}	NÍVEL SIGMA ($Z_{LT}+1,5$)	PPM
Dados “3/4AE” considerados normais	19,124	19,151	1,12	2,62	130525
Dados “3/4AE” sem <i>outliers</i> supondo normais	19,124	19,151	1,33	2,83	91334
Dados “3/4AE” (distribuição Loglogistic)	19,124	19,151	1,18	2,68	118550
Dados “3/4AE” transformados (Johnson)	19,124	19,151	1,41	2,91	79401
Dados “3/4AD” considerados normais	19,124	19,151	1,59	3,09	56313
Dados “3/4AD” sem <i>outliers</i> supondo normais	19,124	19,151	1,72	3,22	42972
Dados “3/4AD” (distribuição Logistic)	19,124	19,151	1,60	3,10	55001
Dados “3/4AD” sem <i>outliers</i> (Lognormal)	19,124	19,151	1,74	3,24	40600

Tabela 4.3 – Cálculos comparativos do nível sigma para “3/4AE” e “3/4AD”

4.3.4 Análise de Dados Reais – Furos 5/8”

Para dar continuidade à aplicação dos métodos estatísticos desta pesquisa exploratória também foi necessário efetuar uma análise dos dados disponíveis em “5/8TE” e “5/8TD” quanto a sua aderência à distribuição Normal.

A partir dos dados reais relativos aos valores de diâmetro dos furos do componente estrutural em “5/8TE” (Apêndice D) foram efetuados a análise da distribuição e o teste de normalidade utilizando os recursos disponíveis do programa estatístico, sendo que a verificação de normalidade baseou-se na análise do valor P (*p-value*) associado à estatística A^2 do teste de normalidade de Anderson-Darling.

A Figura 4.69 apresenta o resultado da análise dos dados em “5/8TE”. Admitindo $\alpha = 0,05$, a distribuição em questão foi considerada não normal visto que o *p-value* obtido (0,005) resultou menor que 0,05. Portanto, como a hipótese nula de normalidade foi rejeitada, obviamente a distribuição em questão não pôde ser considerada normalmente distribuída. Além disso, foi observado que o conjunto de dados “5/8TE” não possuía *outliers*.

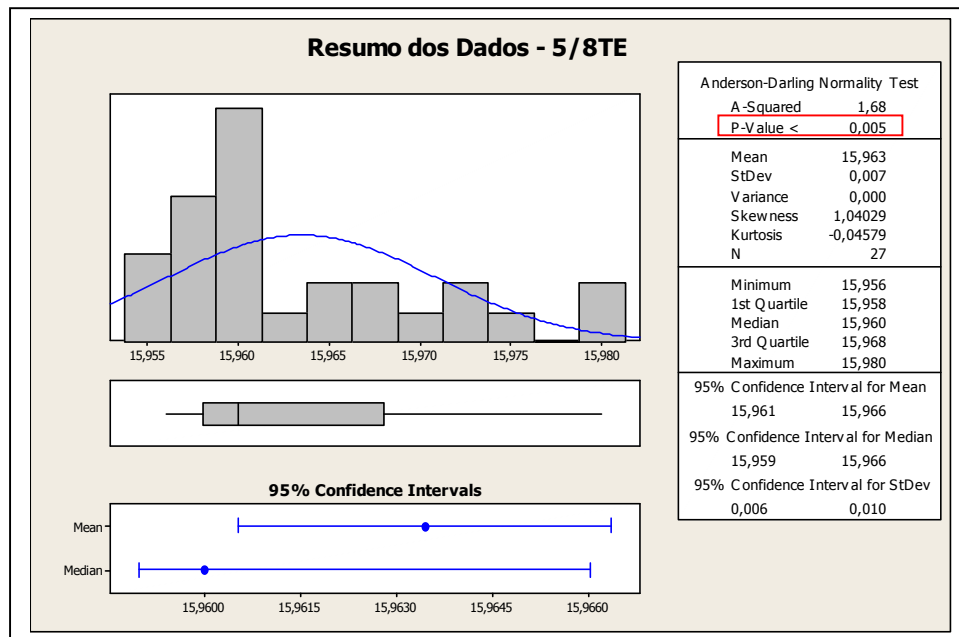


Figura 4.69 – Análise do conjunto de dados originais “5/8TE”

Como o conjunto de dados “5/8TE” não seguia a distribuição Normal, foi efetuada a verificação da distribuição que melhor representasse estes dados. Conforme mostrado na Figura 4.70, a distribuição Lognormal com 3 parâmetros foi a melhor opção encontrada, pois indicou o maior coeficiente de correlação (0,924) entre as distribuições analisadas para um intervalo de 95% de confiança.

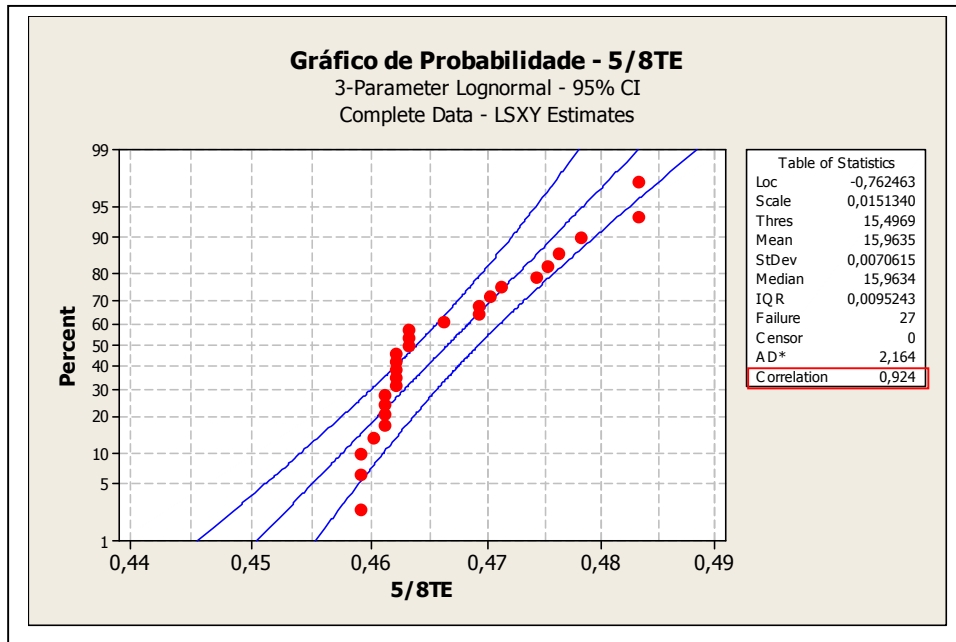


Figura 4.70 – Avaliação da melhor distribuição aplicável ao conjunto de dados “5/8TE”

A partir dos dados reais relativos aos valores de diâmetro dos furos do componente estrutural em “5/8TD” (Apêndice D) foram efetuados a análise da distribuição e o teste de normalidade de Anderson-Darling conforme indicado na Figura 4.71. Admitindo $\alpha = 0,05$, a distribuição em questão foi considerada não normal visto que o *p-value* obtido (0,015) resultou menor que 0,05.

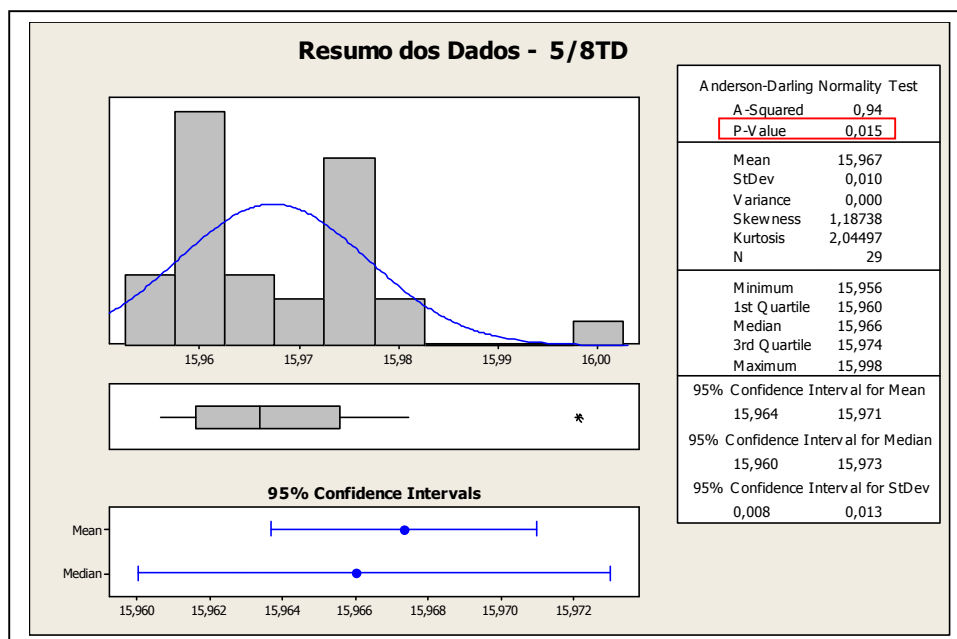


Figura 4.71 – Análise do conjunto de dados originais “5/8TD”

Pela análise dos dados “5/8TD” sem a presença dos *outliers*, apresentada na Figura 4.72 como “5/8TD*”, concluiu-se que a distribuição não era normal (*p-value* inferior a 0,05).

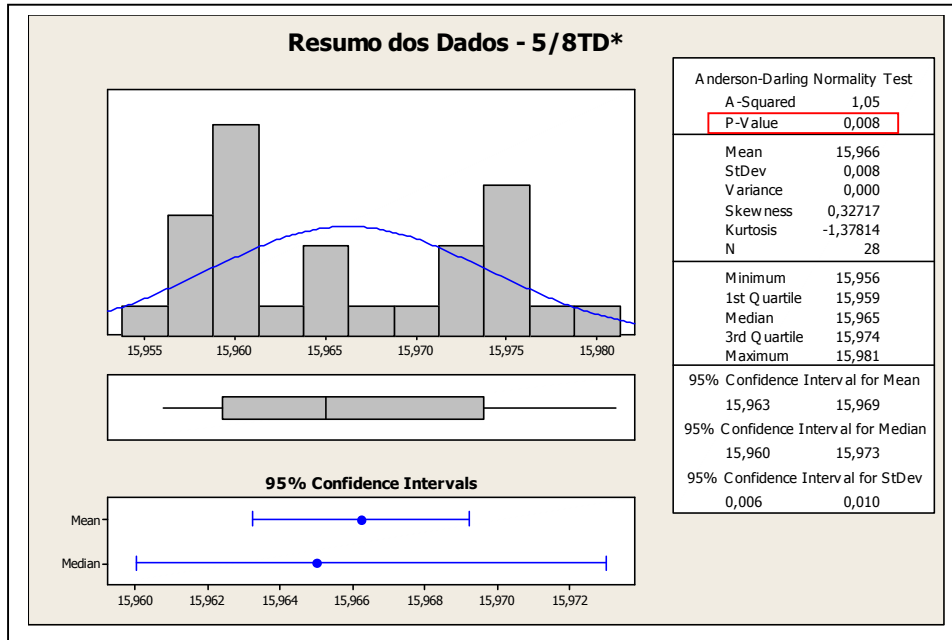


Figura 4.72 – Análise do conjunto de dados “5/8TD” sem *outliers*

Para o conjunto “5/8TD” foi efetuada a verificação de uma distribuição que melhor representasse os dados. Conforme mostrado na Figura 4.73, a distribuição Lognormal com 3 parâmetros foi a melhor opção encontrada.

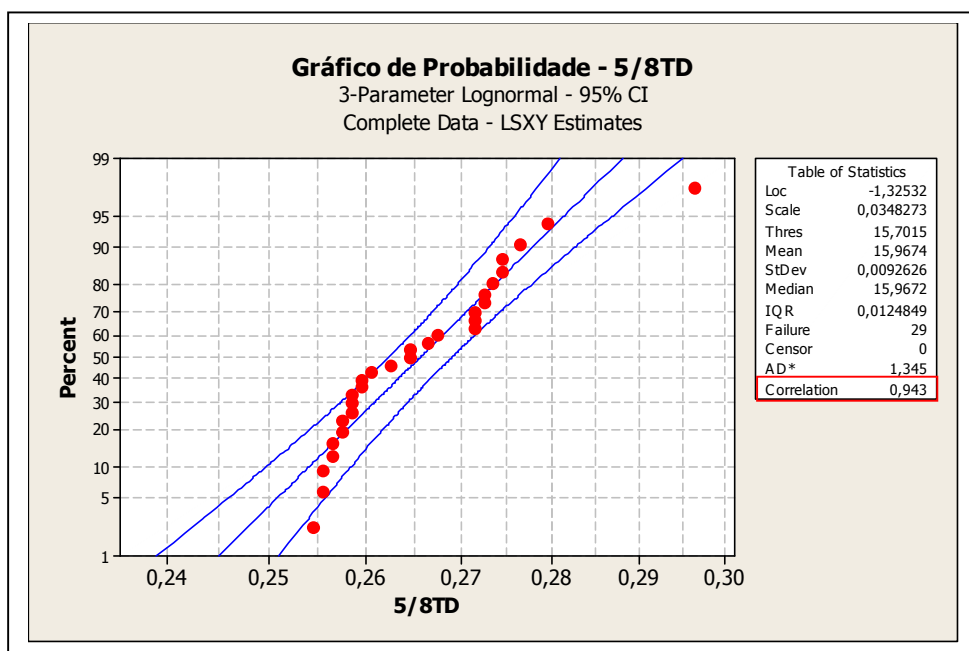


Figura 4.73 – Avaliação da melhor distribuição aplicável ao conjunto de dados “5/8TD”

Como o conjunto de dados “5/8TD” sem *outliers* (ou 5/8TD**) também não seguia a distribuição Normal, foi efetuada a verificação da distribuição que melhor representasse estes dados utilizando o programa estatístico. Conforme demonstrado através da Figura 4.74, a distribuição Lognormal com 3 parâmetros foi a melhor opção encontrada, pois indicou o maior coeficiente de correlação (0,960) entre as distribuições analisadas para um intervalo de 95% de confiança.

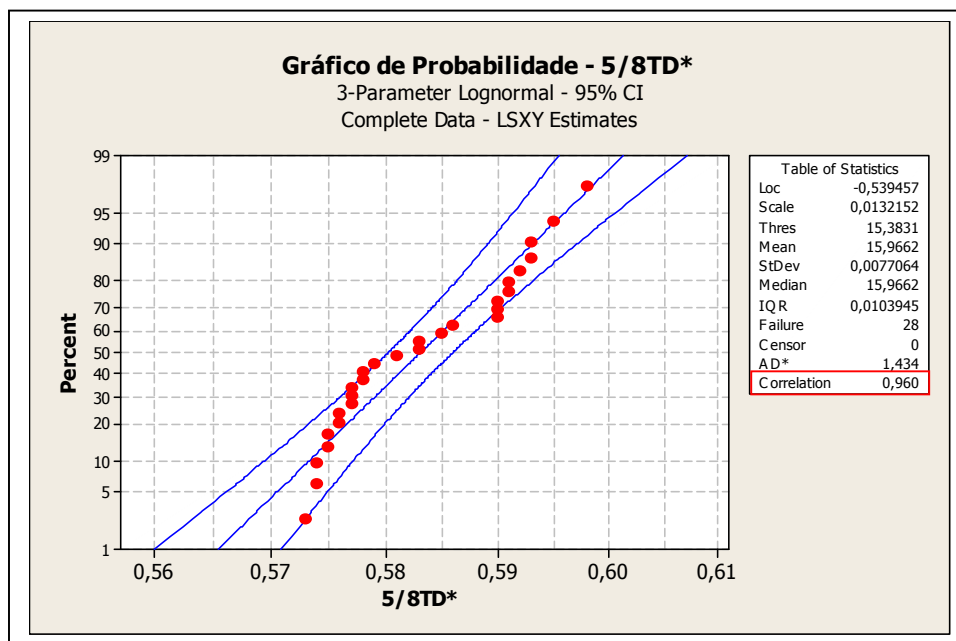


Figura 4.74 – Avaliação de distribuição aplicável ao conjunto de dados “5/8TD” sem *outliers*

4.3.5 Transformação de Dados Reais – Furos 5/8”

Para o conjunto de dados “5/8TE” foi efetuada uma tentativa de transformação por Box-Cox utilizando o programa estatístico. As características desta transformação, com destaque para o valor do parâmetro λ estimado com 95% de confiança, são apresentadas através da Figura 4.75.

Após transformação de Box-Cox, admitindo $\alpha = 0,05$, os dados obtidos não foram considerados normalmente distribuídos, visto que *p-value* (0,005) resultou menor que 0,05, ou seja, rejeitou-se a hipótese nula de normalidade da distribuição obtida. A Figura 4.76 apresenta a análise gráfica da distribuição resultante da transformação de Box-Cox.

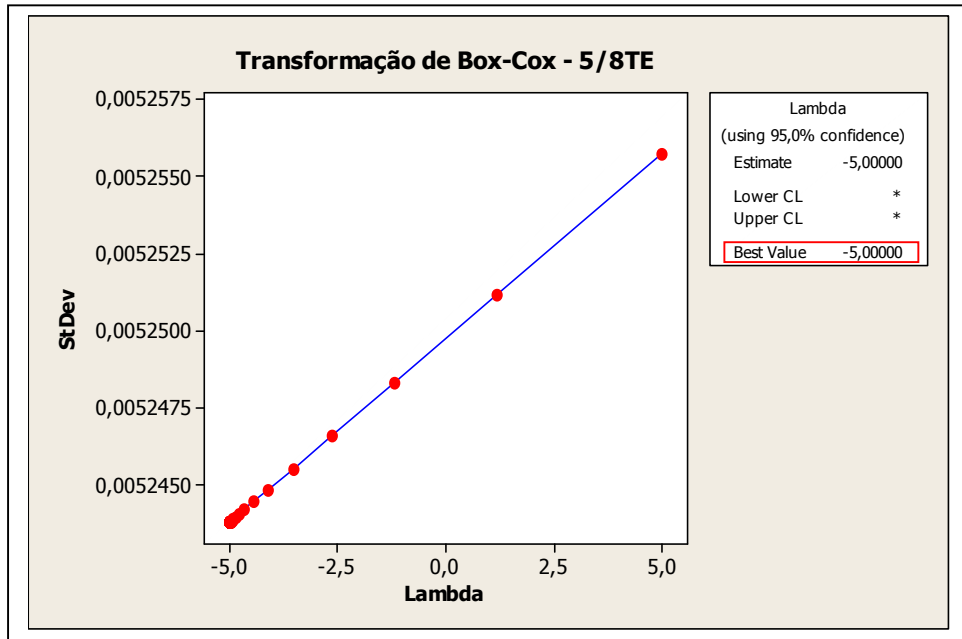


Figura 4.75 – Transformação de Box-Cox para o conjunto de dados “5/8TE”

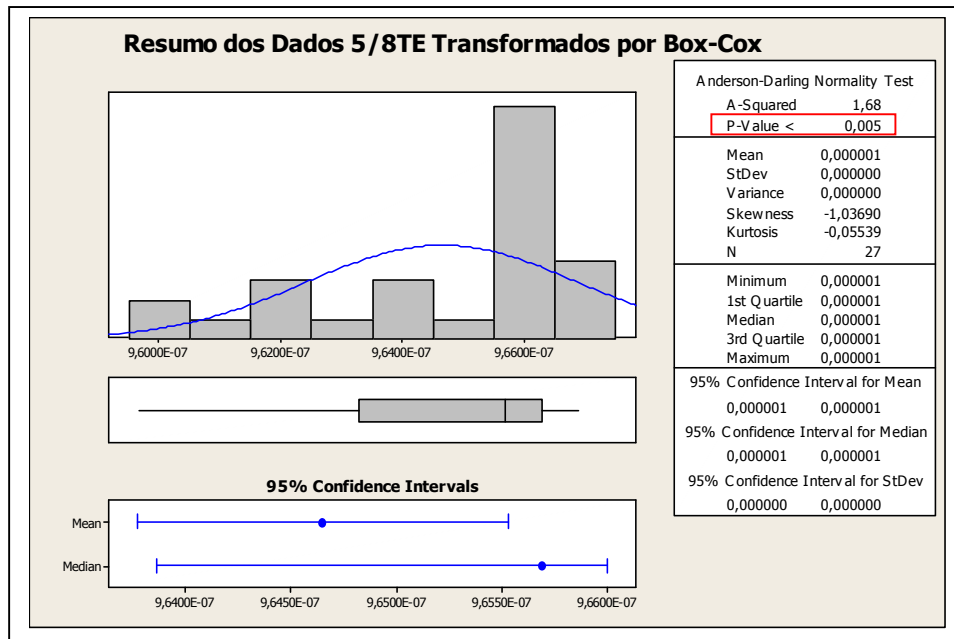


Figura 4.76 – Análise do conjunto de dados “5/8TE” transformados por Box-Cox

Para o conjunto de dados “5/8TE” também foi efetuada aplicação da transformação de Johnson por meio do programa estatístico. O diagrama da Figura 4.77 apresenta as características desta transformação, com destaque para determinação da família de transformação e a correspondente equação.

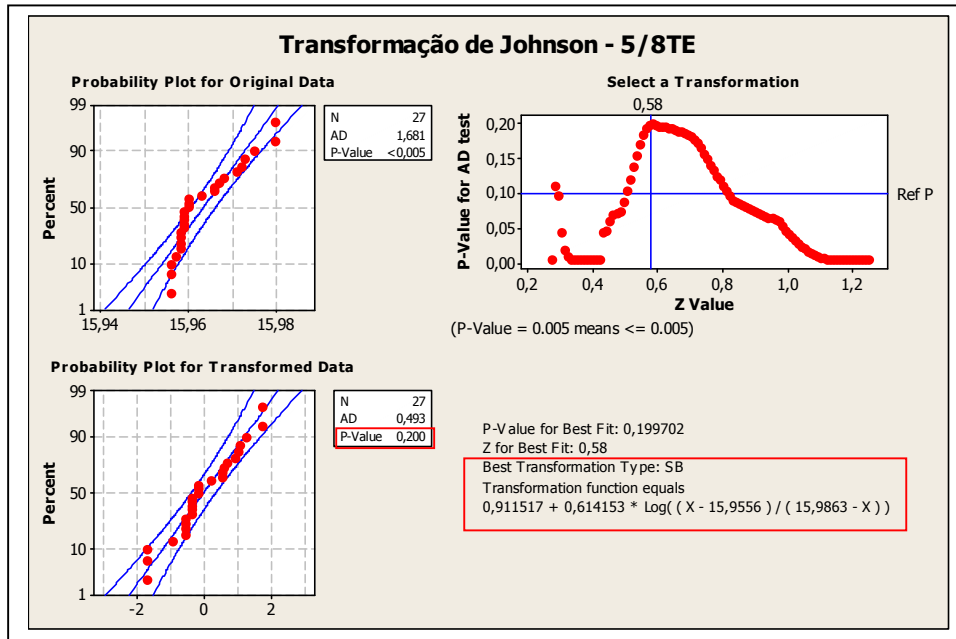


Figura 4.77 – Transformação de Johnson para o conjunto de dados “5/8TE”

Como a análise do conjunto de dados “5/8TE” não indicou a presença de *outliers*, não se fizeram necessárias aplicações das transformações de Box-Cox ou Johnson para esta condição.

Para o conjunto de dados “5/8TD” foi efetuada uma tentativa de transformação por Box-Cox conforme indicado através da Figura 4.78.

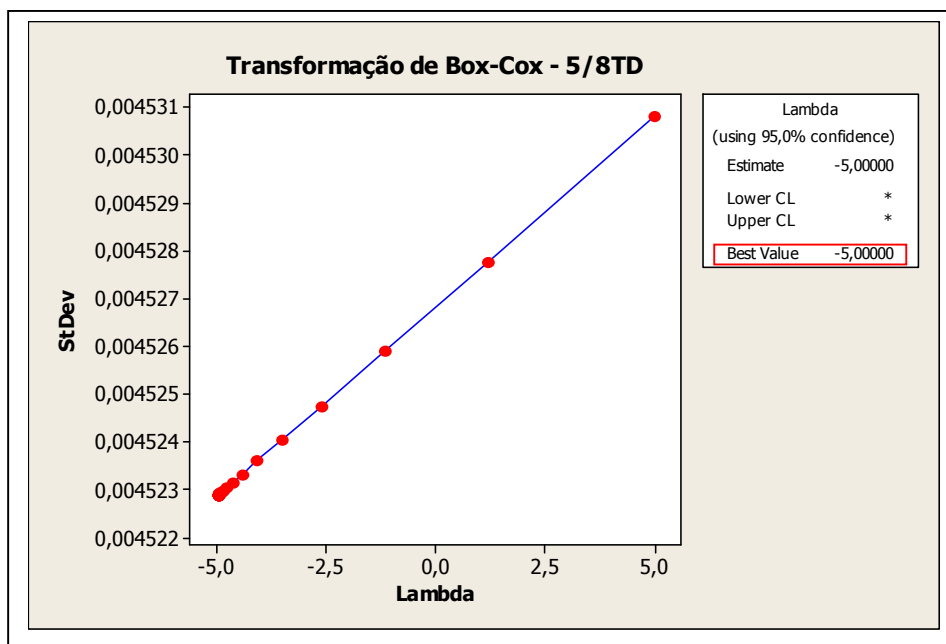


Figura 4.78 – Transformação de Box-Cox para o conjunto de dados “5/8TD”

Após transformação de Box-Cox, admitindo $\alpha = 0,05$, os dados obtidos não foram considerados normalmente distribuídos, visto que *p-value* (0,015) resultou menor que 0,05. A análise gráfica da distribuição transformada por Box-Cox é apresentada na Figura 4.79.

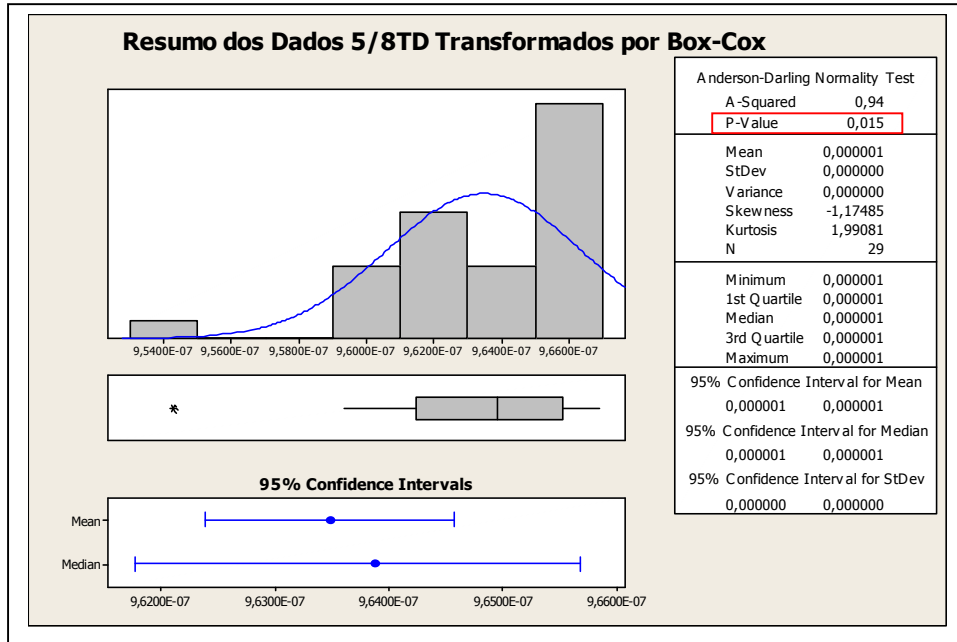


Figura 4.79 – Análise do conjunto de dados “5/8TD” transformados por Box-Cox

Para o conjunto de dados “5/8TD” também foi efetuada aplicação da transformação de Johnson conforme mostrado na Figura 4.80.

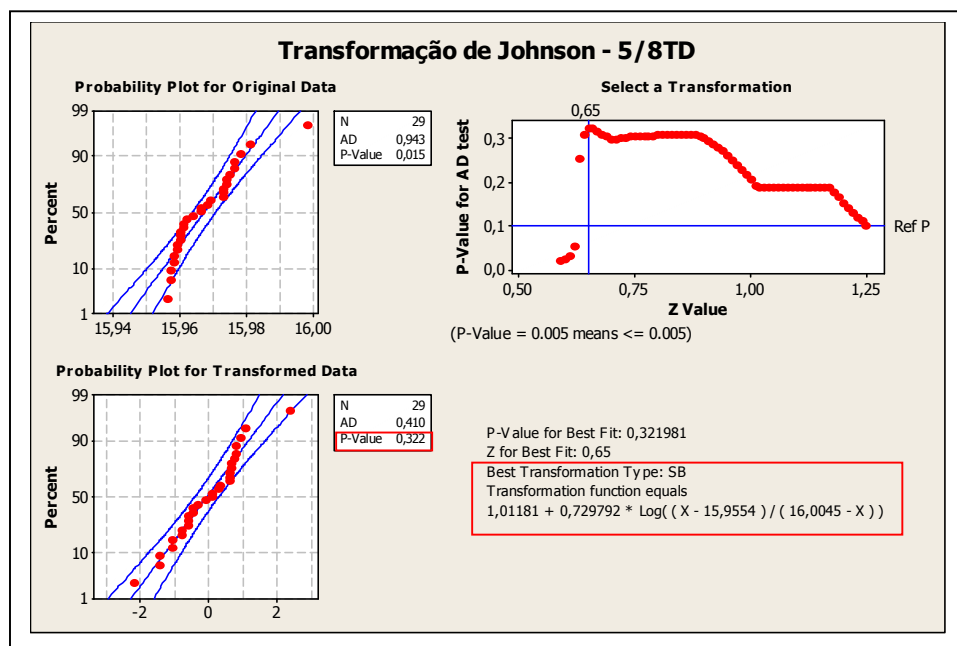


Figura 4.80 – Transformação de Johnson para o conjunto de dados “5/8TD”

Para o conjunto de dados “5/8TD” sem *outliers* (ou “5/8TD*”) foi tentada a transformação dos dados para condição de normalidade através da aplicação do método de Box-Cox conforme indicado através da Figura 4.81.

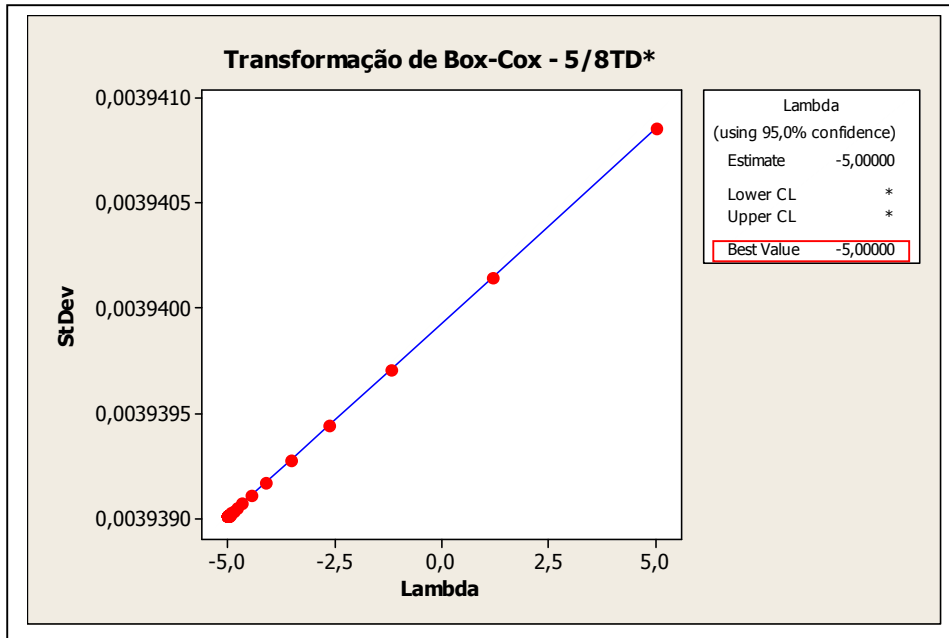


Figura 4.81 – Transformação de Box-Cox para o conjunto de dados “5/8TD” sem *outliers*

Para os dados “5/8TD” sem *outliers* (“5/8TD*”) também foi aplicado o método de transformação de Johnson de acordo com diagrama apresentado na Figura 4.82.

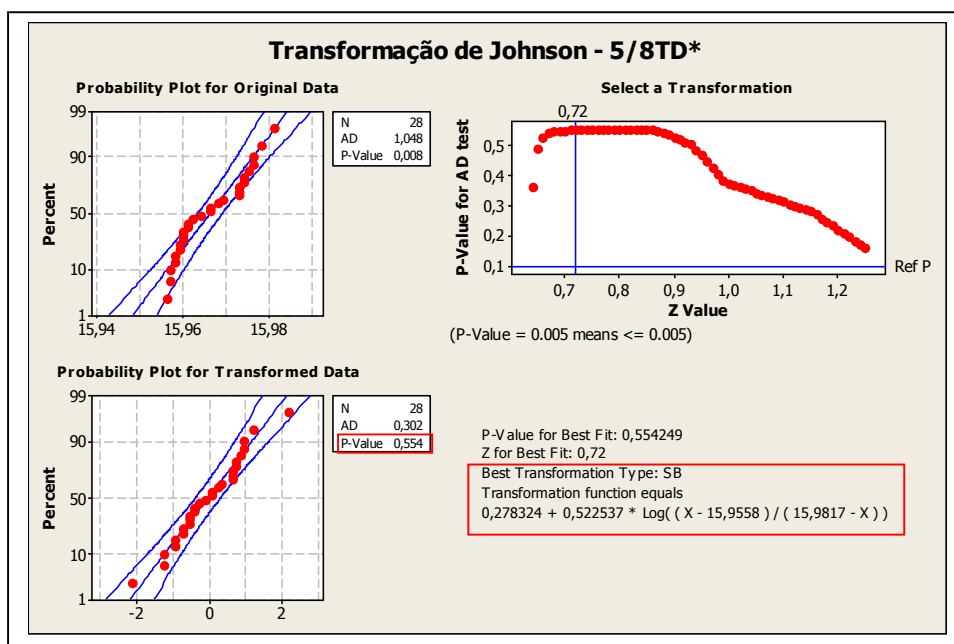


Figura 4.82 – Transformação de Johnson para o conjunto de dados “5/8TD” sem *outliers*

4.3.6 Cálculo de Capabilidade de Dados Reais – Furos 5/8”

A avaliação de capabilidade para os conjuntos de dados “5/8TE” e “5/8TD” (Apêndice D) teve como requisito um limite inferior de especificação de 15,950 mm e um limite superior de especificação de 15,977 mm para as medidas de diâmetro dos furos do componente estrutural. Foram calculados diversos índices de capabilidade Z_{bench} (Z_{LT} e Z_{ST}) diferentes:

- Em uma primeira abordagem o cálculo foi efetuado com os dados originais e os dados tratados (sem *outliers*) como se fossem normalmente distribuídos;
- No segundo caso a condição não normal da distribuição foi considerada no cálculo dos dados originais e, quando aplicável, nos dados sem a presença de *outliers*.
- Na terceira abordagem os dados originais foram modificados por meio da transformação de Box-Cox;
- Na quarta situação foi utilizada a transformação de Johnson.

A Figura 4.83 apresenta os resultados do cálculo de capabilidade considerando os dados originais “5/8TE” como se fossem normalmente distribuídos.

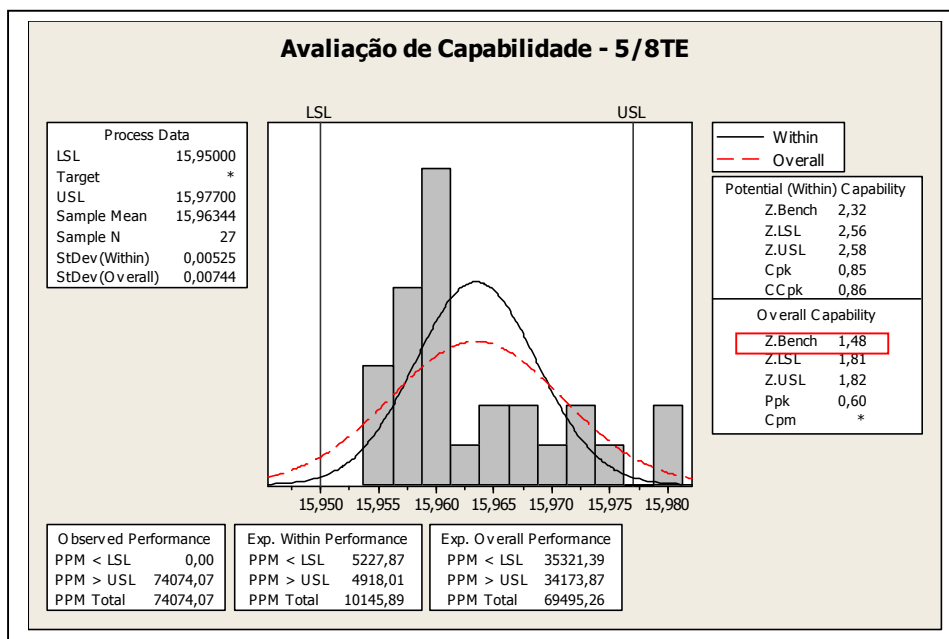


Figura 4.83 – Índices de capabilidade do conjunto de dados originais “5/8TE”

Como o conjunto de dados “5/8TE” não se mostrou normalmente distribuído, o cálculo de capabilidade foi efetuado a partir da distribuição Lognormal com 3 parâmetros, correspondente ao melhor ajuste para os dados de “5/8TE” em um intervalo de 95% de confiança. Os resultados nesta situação são apresentados através da Figura 4.84.

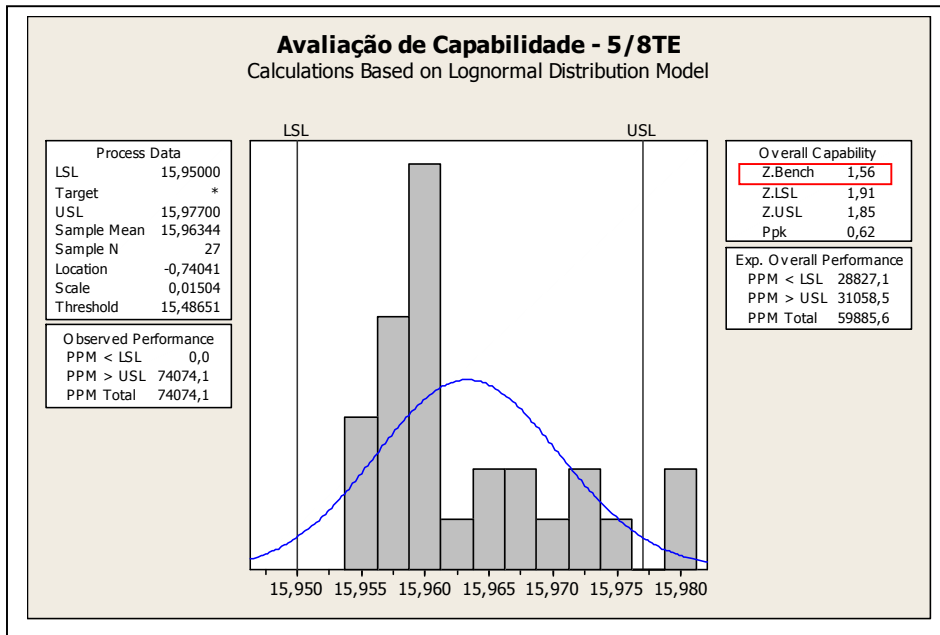


Figura 4.84 – Índices de capacidade dos dados “5/8TE” baseados na distribuição Lognormal

Não foi efetuado cálculo de capacidade após transformação de Box-Cox dos dados “5/8TE”, pois a mesma não resultou em normalidade.

A Figura 4.85 apresenta os resultados do cálculo de capacidade do processo após utilização da transformação de Johnson no conjunto de dados originais “5/8TE”.

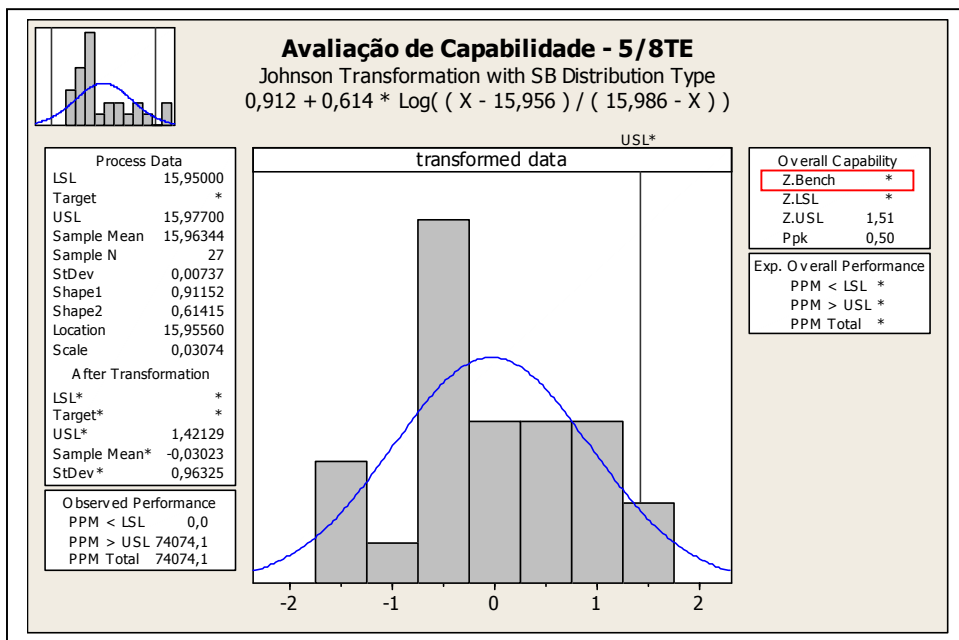


Figura 4.85 – Índices de capacidade dos dados “5/8TE” transformados por Johnson

A Figura 4.86 apresenta os resultados do cálculo de capacidade considerando os dados originais “5/8TD” como se fossem normalmente distribuídos.

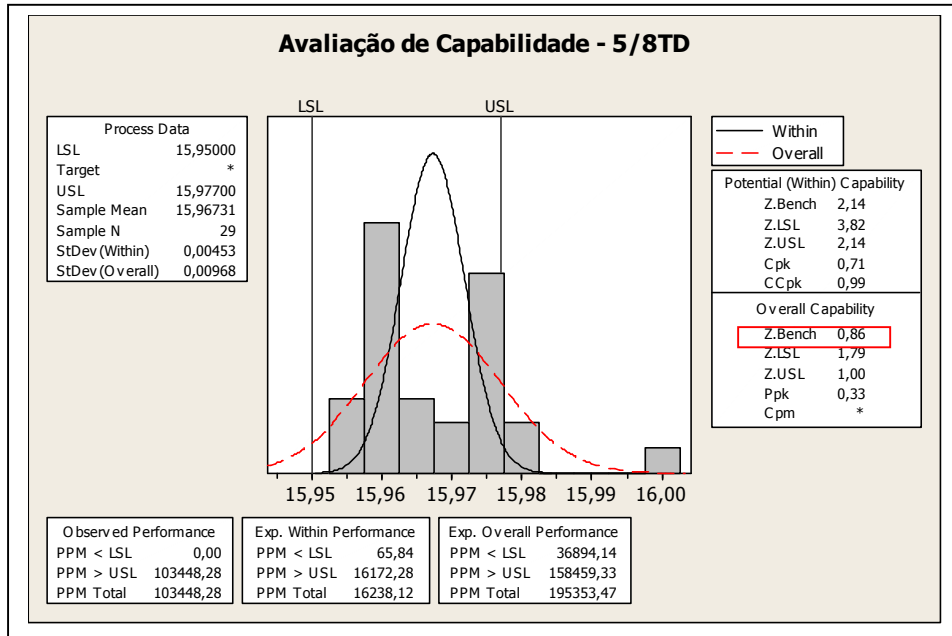


Figura 4.86 – Índices de capacidade dos dados originais “5/8TD”

Como o conjunto de dados “5/8TD” não se mostrou normalmente distribuído, o cálculo de capacidade foi efetuado a partir da distribuição Lognormal com 3 parâmetros. Os resultados nesta situação são apresentados através da Figura 4.87.

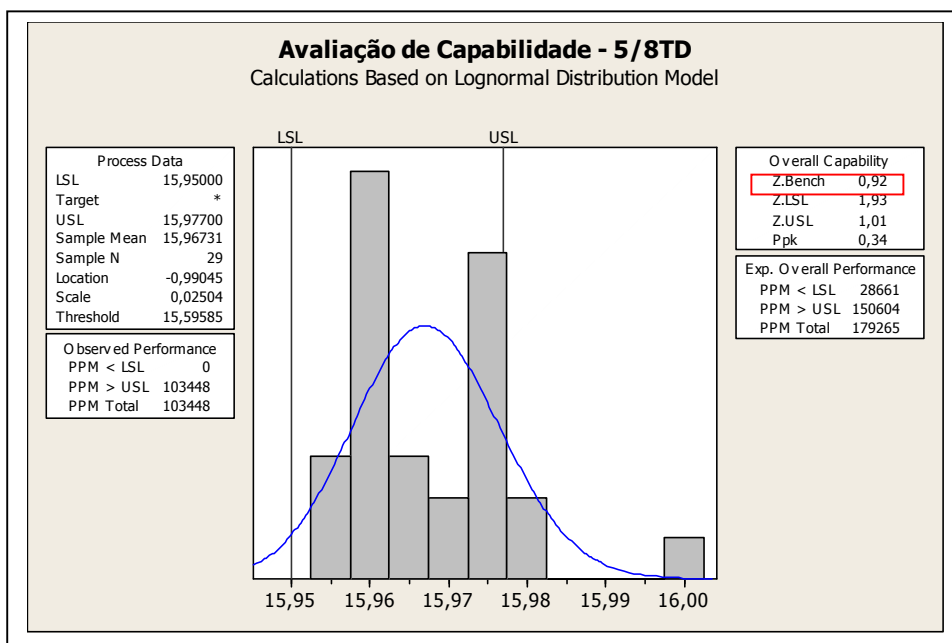


Figura 4.87 – Índices de capacidade dos dados “5/8TD” baseados na distribuição Lognormal

Não foi efetuado cálculo de capacidade após transformação de Box-Cox dos dados “5/8TD”, pois a mesma não resultou em normalidade. A Figura 4.88 apresenta os resultados do cálculo de capacidade do processo após transformação de Johnson em “5/8TD”.

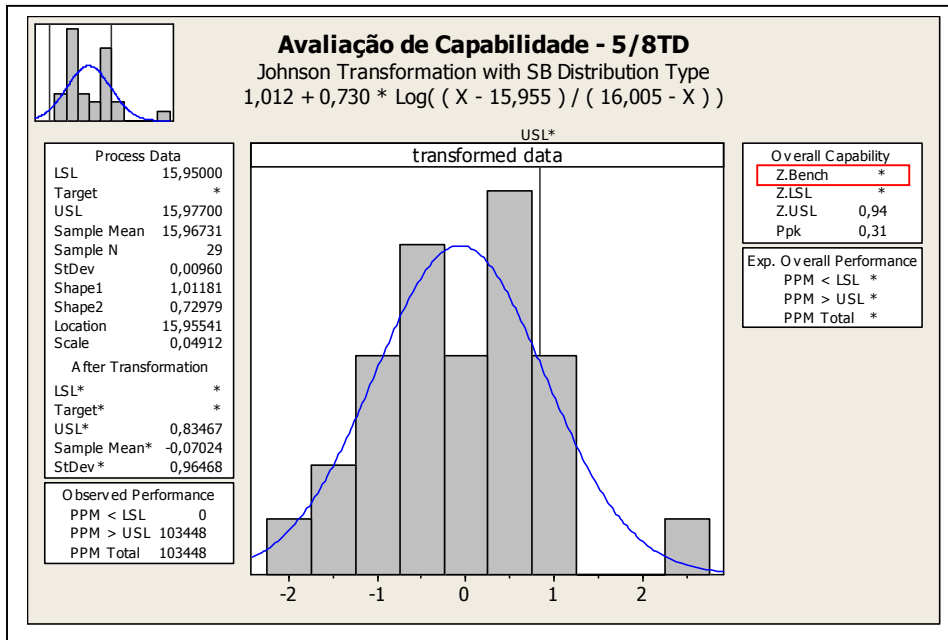


Figura 4.88 – Índices de capacidade dos dados “5/8TD” transformados por Johnson

A Figura 4.89 apresenta os resultados considerando os dados originais “5/8TD” sem a presença de outliers e considerando os dados como se fossem normalmente distribuídos.

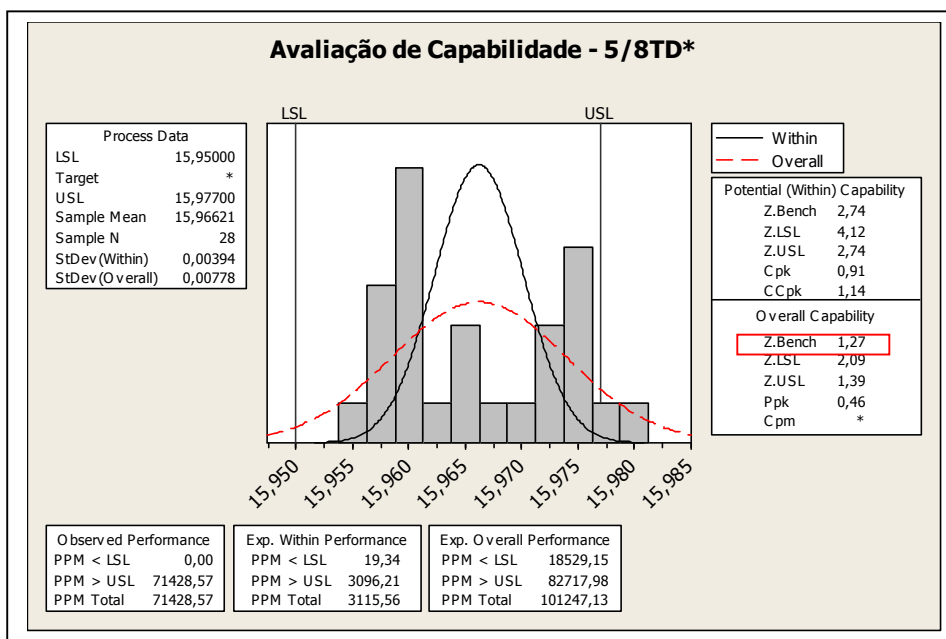


Figura 4.89 – Índices de capacidade do conjunto de dados “5/8TD” sem outliers

Como os dados “5/8TD” sem *outliers* não eram normalmente distribuído, o cálculo de capacidade foi efetuado a partir da distribuição Lognormal com 3 parâmetros (Figura 4.90).

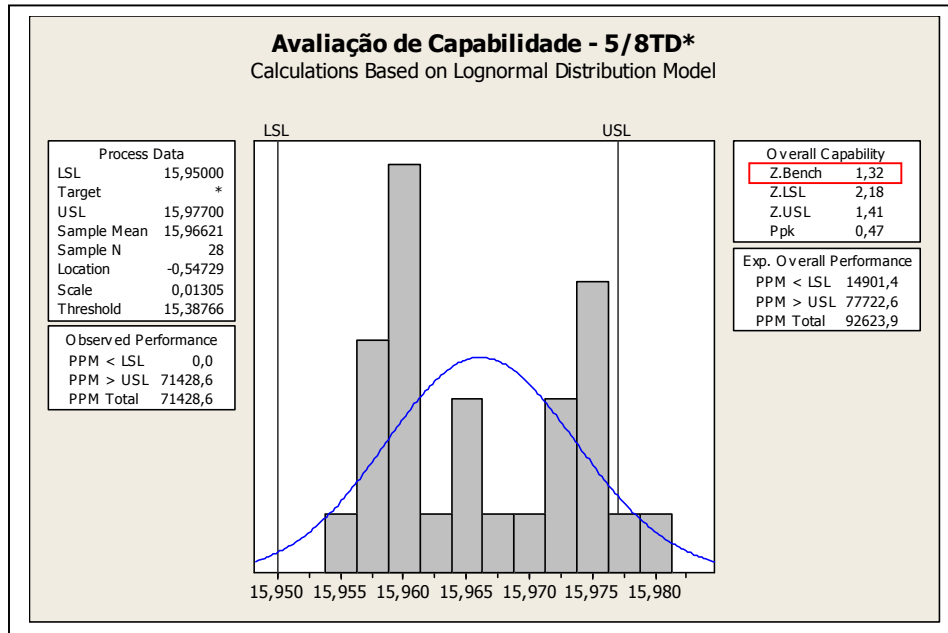


Figura 4.90 – Índices capacidade dos dados “5/8TD” sem *outliers* (distribuição Lognormal)

Não foi efetuado cálculo de capacidade após transformação de Box-Cox dos dados “5/8TD” sem *outliers*, pois a mesma não resultou em normalidade. A Figura 4.91 apresenta o cálculo de capacidade após transformação de Johnson dos dados “5/8TD” sem *outliers*.

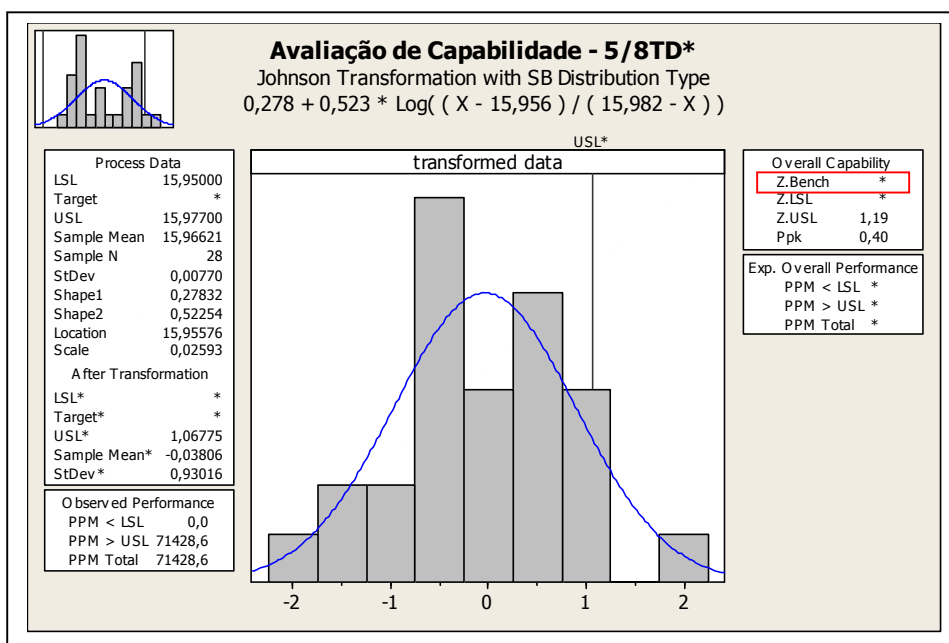


Figura 4.91 – Índices de capacidade dos dados “5/8TD” sem *outliers* (transf. Johnson)

Para os conjuntos de dados “5/8TE” e “5/8TD” não foram calculados os índices de capacidade após as transformações de Box-Cox, pois as mesmas não resultaram em normalidade. Também não foi possível obter índices de capacidade após as transformações de Johnson, uma vez que não foram encontrados valores de referência Z_{bench} a partir da aplicação da fórmula de transformação nos limites de especificação definidos. A Tabela 4.4 apresenta a comparação dos valores obtidos nas várias situações apresentadas, com o correspondente cálculo do nível sigma e a fração não conforme em partes por milhão para os conjuntos de dados “5/8TE” e “5/8TD”, considerando Z_{shift} igual a 1,5 sigma.

PREMISSA ADOTADA PARA CÁLCULO	LIE (mm)	LSE (mm)	ÍNDICE Z_{LT}	NÍVEL SIGMA ($Z_{LT}+1,5$)	PPM
Dados “5/8TE” considerados normais	15,950	15,977	1,48	2,98	69495
Dados “5/8TE” (distribuição Lognormal)	15,950	15,977	1,56	3,06	59886
Dados “5/8TD” considerados normais	15,950	15,977	0,86	2,36	195353
Dados “5/8TD” (distribuição Lognormal)	15,950	15,977	0,92	2,42	179265
Dados “5/8TD” sem <i>outliers</i> supondo normais	15,950	15,977	1,27	2,77	101247
Dados “5/8TD” sem <i>outliers</i> (Lognormal)	15,950	15,977	1,32	2,82	92624

Tabela 4.4 – Cálculos comparativos do nível sigma para “5/8TE” e “5/8TD”

5. DISCUSSÃO DOS RESULTADOS

5.1 Análise de Resultados na Investigação Experimental

A análise da simulação efetuada na pesquisa experimental com os dados não normais foi conduzida através da interpretação dos gráficos de controle individual e da avaliação de capacidade com os dados modelados pelas distribuições Gamma e Beta.

5.1.1 Análise dos Gráficos de Controle – Distribuição Gamma

Após análise dos gráficos de controle construídos durante a pesquisa experimental para os dados provenientes da distribuição Gamma, foram observadas quatro situações distintas, que estão abaixo detalhadas.

- a) Assumindo os dados modelados pela distribuição Gamma como se fossem normalmente distribuídos:
 - No chamado “teste 1”, que avalia a quantidade de pontos além dos limites de controle estipulados, foram identificados 2 (dois) pontos acima do limite superior de controle, que poderiam ser interpretados como causas especiais de variação.
 - Entretanto, o simples tratamento dos dados, por meio da eliminação dos pontos fora do padrão (*outliers*) foi suficiente para que o comportamento apresentado no gráfico de controle se enquadrasse em uma situação de controle estatístico, mesmo com a suposição incorreta de normalidade dos dados originais.
- b) Estabelecendo novos limites de controle inferior e superior, baseado nos percentis 0,135 e 99,865, respectivamente, com base na função densidade de probabilidade da distribuição Gamma, obviamente assimétrica:
 - Neste caso não foram observados pontos além dos limites de controle, tanto para os dados originais quanto para os dados tratados através da eliminação dos *outliers*.
- c) Efetuando transformação matemática dos dados modelados pela distribuição Gamma por meio do método de Box-Cox:
 - Após a aplicação da transformação de Box-Cox, não mais foram identificados pontos além dos limites de controle, tanto para os dados originais quanto para os dados tratados através da eliminação dos *outliers*.
- d) Efetuando transformação matemática dos dados modelados pela distribuição Gamma por meio do método de Johnson:

- De modo semelhante, após a transformação de Johnson, também não foram observados pontos além dos limites de controle, tanto para os dados originais quanto para os dados tratados através da eliminação dos *outliers*.

5.1.2 Análise dos Gráficos de Controle – Distribuição Beta

A análise dos gráficos de controle construídos durante a pesquisa experimental com dados modelados pela distribuição Beta, apresentou quatro situações diferentes para os dados originais brutos e tratados, conforme comentado a seguir.

- a) Assumindo os dados modelados pela distribuição Beta como se fossem normalmente distribuídos:
 - No “teste 1” de avaliação da quantidade de pontos além dos limites de controle, foi verificado 1 (um) ponto abaixo do limite inferior de controle, que poderia ser abordado como causa especial de variação.
 - É importante ressaltar que a análise de estabilidade do processo, obviamente não considerava os limites de especificação, pois além do ponto abaixo do limite de controle inferior, evidenciaram-se alguns pontos abaixo do limite inferior de especificação.
 - O simples tratamento dos dados, por meio da eliminação dos *outliers* foi suficiente para que os dados se ajustassem em uma distribuição considerada normal e sem pontos além dos limites de controle, isto é, os dados modelados pela distribuição Beta sem a presença de *outliers* mostraram-se normalmente distribuídos e o comportamento apresentado no gráfico de controle indicou uma situação de controle estatístico.
- b) Estabelecendo novos limites de controle inferior e superior, baseado nos percentis 0,135 e 99,865, respectivamente, com base na função densidade de probabilidade da distribuição assimétrica Beta:
 - Não foram observados pontos além dos limites de controle para os dados originais e, como a distribuição previamente tratada pode ser considerada normal, não foi necessária uma análise adicional dos novos limites de controle após a eliminação dos *outliers*.
- c) Efetuando transformação matemática dos dados modelados pela distribuição Beta por meio do método de Box-Cox:

- Após a aplicação da transformação de Box-Cox, não foram visualizados pontos além dos limites de controle e, como a distribuição previamente tratada pôde ser considerada normal, não foi efetuada transformação matemática após a eliminação dos *outliers*.

d) Efetuando transformação matemática dos dados modelados pela distribuição Gamma por meio do método de Johnson:

- Semelhantemente, após a transformação de Johnson, também não foram observados pontos além dos limites de controle e também não foi necessário efetuar transformação dos dados após a eliminação dos *outliers*.

5.1.3 Considerações sobre Normalidade e Gráficos de Controle na Simulação

A interpretação resumida da análise dos gráficos de controle individuais para os conjuntos de dados modelados pelas distribuições Gamma e Beta é mostrada na Tabela 5.1.

CONJUNTO DE DADOS SIMULADOS	TESTE 1 – QUANTIDADE DE PONTOS ALÉM DOS LIMITES DE CONTROLE (LC)			
	SUPONDO NORMAL LIC = $\bar{X} - E_2 \cdot \bar{Rm}$ LSC = $\bar{X} + E_2 \cdot \bar{Rm}$	NÃO NORMAL LC EM PERCENTIS LIC = percentil 0,135 LSC = percentil 99,865	TRANSF. BOX-COX LIC = $\bar{X}_{BC} - E_2 \cdot \bar{Rm}$ LSC = $\bar{X}_{BC} + E_2 \cdot \bar{Rm}$	TRANSF. JOHNSON LIC = $\bar{X}_J - E_2 \cdot \bar{Rm}$ LSC = $\bar{X}_J + E_2 \cdot \bar{Rm}$
Gamma 220	2 pontos acima do LSC	0	0	0
Gamma 220* (sem outliers)	0	0	0	0
Beta 2550	1 ponto abaixo do LIC	0	0	0
Beta 2550* (sem outliers) distribuição Normal	0	N/A (não aplicável)	N/A (não aplicável)	N/A (não aplicável)

Tabela 5.1 – Interpretação dos gráficos de controle individuais - dados simulados

Por meio da análise dos gráficos individuais de controle para os conjuntos de dados não normais simulados, ficou claro que a suposição incorreta de normalidade acarreta uma interpretação equivocada do comportamento do processo, pois conforme demonstrado através da avaliação dos dados modelados pelas distribuições Gamma e Beta, falsas causas especiais de variação foram evidenciadas. Por outro lado, o tratamento prévio dos dados por meio da eliminação dos *outliers*, a determinação de novos limites de controle baseados em percentis e

a transformação matemática das variáveis com as técnicas de Box-Cox e Johnson, mostraram-se procedimentos extremamente eficazes ao possibilitar uma interpretação mais coerente acerca da condição de estabilidade do processo em estudo.

5.1.4 Análise do Cálculo de Capabilidade – Distribuição Gamma

Após análise dos índices de capabilidade calculados durante a pesquisa experimental, utilizando o conceito seis sigma para os dados modelados pela distribuição Gamma em quatro situações distintas, foram observados os seguintes resultados.

- a) Assumindo os dados modelados pela distribuição Gamma como se fossem normalmente distribuídos:
 - Foi obtido um índice de $4,47\sigma$ com os dados originais, o que denotaria uma proporção de itens não conformes de 1470 PPM.
 - Após tratamento dos dados através da eliminação dos pontos fora do padrão (*outliers*), o desempenho, na verdade ilusório, apresentaria um índice de $4,84\sigma$, equivalente a 420 PPM.
- b) Adotando a característica assimétrica dos dados com base na função densidade de probabilidade da distribuição Gamma:
 - Neste caso, o índice obtido de $3,62\sigma$ apontou para um desempenho de 16872 PPM.
 - Através da eliminação dos *outliers*, obteve-se o índice de $3,99\sigma$ (6333 PPM), mais representativo da realidade dos dados assimétricos.
- c) Efetuando transformação matemática dos dados modelados pela distribuição Gamma por meio do método de Box-Cox:
 - Após a aplicação da transformação de Box-Cox, o novo nível sigma encontrado foi $3,64\sigma$ (16010 PPM), muito próximo ao valor de $3,62\sigma$ relativo ao cálculo com os dados considerados assimétricos.
 - O índice de $4,03\sigma$ (5730 PPM), obtido para os dados transformados após a remoção dos *outliers*, não é muito diferente do valor de $3,99\sigma$ referente ao cálculo dos dados assimétrico sem *outliers*.
- d) Efetuando transformação matemática dos dados modelados pela distribuição Gamma por meio do método de Johnson:

- Após a aplicação da transformação de Johnson, o nível sigma encontrado foi $3,81\sigma$ (10453 PPM), superior ao valor de $3,64\sigma$ obtido após transformação de Box-Cox e ao valor de $3,62\sigma$ calculado com os dados considerados assimétricos.
- O índice de $4,37\sigma$ (2041 PPM), obtido para os dados transformados após a remoção dos *outliers*, é consideravelmente superior ao valor de $4,03\sigma$ encontrado após transformação de Box-Cox e ao valor de $3,99\sigma$ referente ao cálculo dos dados assimétrico sem *outliers*.

5.1.5 Análise do Cálculo de Capabilidade – Distribuição Beta

A análise dos índices de capacidade calculados durante a pesquisa experimental, por meio da determinação do nível sigma para os dados modelados pela distribuição Beta em quatro situações diferentes, apresentou os resultados comentados a seguir.

- a) Assumindo os dados modelados pela distribuição Beta como se fossem normalmente distribuídos:
 - Foi obtido um índice de $3,61\sigma$ com os dados originais, o que denotaria uma proporção de itens não conformes de 17625 PPM.
 - Após tratamento dos dados através da eliminação dos *outliers*, os mesmos mostraram-se, de fato, normalmente distribuídos e o índice encontrado foi de $3,95\sigma$, equivalente a 7145 PPM.
- b) Adotando a característica assimétrica dos dados com base na função densidade de probabilidade da distribuição Beta:
 - O índice de $3,48\sigma$, obtido neste caso, indicou um desempenho de 23953 PPM.
 - Uma vez que a eliminação dos *outliers* resultou em normalidade, não foi necessária análise adicional com os dados tratados em sua característica assimétrica.
- c) Efetuando transformação matemática dos dados modelados pela distribuição Beta por meio do método de Box-Cox:
 - Após a aplicação da transformação de Box-Cox, o nível sigma encontrado foi $3,59\sigma$ (18247 PPM), não muito distante do valor de $3,48\sigma$ relativo ao cálculo com os dados considerados assimétricos.
 - Visto que a eliminação dos *outliers* resultou em normalidade, não foi necessária análise adicional com os dados tratados e transformados (Box-Cox).

- d) Efetuando transformação matemática dos dados modelados pela distribuição Gamma por meio do método de Johnson:
- Após a aplicação da transformação de Johnson, o nível sigma encontrado foi $3,36\sigma$ (31787 PPM), inferior ao valor de $3,59\sigma$ obtido após transformação de Box-Cox e ao valor de $3,48\sigma$ calculado com os dados considerados assimétricos.
 - Uma vez que a eliminação dos *outliers* resultou em normalidade, não foi necessária análise adicional com os dados tratados e transformados (Johnson).

5.1.6 Considerações sobre Normalidade e Capabilidade na Simulação

Um resumo da avaliação de capabilidade dos dados simulados modelados pelas distribuições Gamma e Beta é representado através da Tabela 5.2.

CONJUNTO DE DADOS SIMULADOS	AVALIAÇÃO DE CAPABILIDADE ATRAVÉS DO NÍVEL SIGMA											
	SUPONDO NORMAL			NÃO NORMAL (distribuição original)			TRANSF. BOX-COX			TRANSF. JOHNSON		
	Z _{LT}	Z _{ST} (σ)	PPM	Z _{LT}	Z _{ST} (σ)	PPM	Z _{LT}	Z _{ST} (σ)	PPM	Z _{LT}	Z _{ST} (σ)	PPM
Gamma 220	2,97	4,47	1470	2,12	3,62	16872	2,14	3,64	16010	2,31	3,81	10453
Gamma 220* (sl outliers)	3,34	4,84	420	2,49	3,99	6333	2,53	4,03	5730	2,87	4,37	2041
Beta 2550	2,11	3,61	17625	1,98	3,48	23953	2,09	3,59	18247	1,86	3,36	31787
Beta 2550* (sl outliers) distribuição Normal	2,45	3,95	7145	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A

Tabela 5.2 – Determinação do nível sigma - dados simulados

No cálculo do índice de capabilidade através do conceito seis sigma para o conjunto de dados não normais simulados, ficou demonstrado que a adoção inadequada de normalidade pode produzir valores que não representam o desempenho real do processo em estudo. Por exemplo, no caso dos dados gerados a partir da distribuição Gamma, o cálculo assumindo normalidade indicou índices de desempenho muito superiores aos valores obtidos através da consideração da característica assimétrica da distribuição e dos dados previamente transformados por Box-Cox e Johnson. Entretanto, em relação os dados gerados a partir da distribuição Beta, a adoção incorreta de normalidade não produziu valores tão discrepantes se comparados com aqueles obtidos por meio da característica não normal e mesmo em relação aos dados transformados.

5.2 Análise de Resultados na Pesquisa Exploratória

A análise do estudo de caso na pesquisa exploratória foi efetuada através da avaliação de capacidade com os conjuntos de dados reais “3/4AE”, “3/4AD”, “5/8TE” e “5/8TD”.

5.2.1 Análise do Cálculo de Capacibilidade – Dados “3/4AE”

A análise dos índices de capacidade calculados durante a pesquisa exploratória, por meio da determinação do nível sigma para o conjunto de dados “3/4AE” em quatro situações diferentes, apresentou os resultados comentados a seguir.

- a) Assumindo o conjunto de dados “3/4AE” como se fossem normalmente distribuídos:
 - Foi obtido um índice de $2,62\sigma$ com os dados originais, o que denotaria uma proporção de itens não conformes de 130525 PPM.
 - Após tratamento dos dados através da eliminação dos *outliers*, os mesmos mostraram-se, de fato, normalmente distribuídos e o índice encontrado foi de $2,83\sigma$, equivalente a 91334 PPM.
- b) Adotando a função densidade de probabilidade da distribuição que melhor se ajustou ao conjunto de dados (Loglogistic com 3 parâmetros):
 - O índice de $2,68\sigma$, obtido neste caso, indicou um desempenho de 118550 PPM.
 - Uma vez que a eliminação dos *outliers* resultou em normalidade, não foi necessária análise adicional com os dados tratados em sua característica assimétrica.
- c) Efetuando transformação matemática do conjunto de dados por meio do método de Box-Cox:
 - Não foi possível efetuar a transformação de Box-Cox através do *software* estatístico, isto é, não foi encontrado valor de λ que possibilitasse a normalização dos dados.
 - Visto que a eliminação dos *outliers* resultou em normalidade, não foi necessária análise adicional com os dados tratados e transformados (Box-Cox).
- d) Efetuando transformação matemática do conjunto de dados por meio do método de Johnson:
 - Após a aplicação da transformação de Johnson, o nível sigma encontrado foi $2,91\sigma$ (79401 PPM), um pouco superior ao valor $2,68\sigma$ calculado com os dados considerados assimétricos.

- Uma vez que a eliminação dos *outliers* resultou em normalidade, não foi necessária análise adicional com os dados tratados e transformados (Johnson).

5.2.2 Análise do Cálculo de Capabilidade – Dados “3/4AD”

A análise dos índices de capabilidade calculados durante a pesquisa exploratória, por meio da determinação do nível sigma para o conjunto de dados “3/4AD” em quatro situações diferentes, apresentou os resultados comentados a seguir.

- Assumindo o conjunto de dados “3/4AD” como se fossem normalmente distribuídos:
 - Foi obtido um índice de $3,09\sigma$ com os dados originais, o que denotaria uma proporção de itens não conformes de 56313 PPM.
 - Após tratamento dos dados através da eliminação dos *outliers*, o índice encontrado foi de $3,22\sigma$, o que representaria um desempenho equivalente a 42972 PPM.
- Adotando a função densidade de probabilidade da distribuição que melhor se ajustou ao conjunto de dados (Logistic):
 - O índice de $3,10\sigma$, obtido neste caso, indicou um desempenho de 55001 PPM, muito semelhante ao valor obtido com a suposição de normalidade.
 - Através da eliminação dos *outliers*, obteve-se o índice de $3,24\sigma$ (40600 PPM), também bastante próximo ao valor obtido com suposição de normalidade.
- Efetuada transformação matemática do conjunto de dados por meio do método de Box-Cox:
 - Não foi possível efetuar a transformação de Box-Cox através do *software* estatístico, isto é, não foi encontrado valor de λ que possibilitasse a normalização dos dados originais e dos dados sem *outliers*.
- Efetuada transformação matemática do conjunto de dados por meio do método de Johnson:
 - Não foi possível efetuar a transformação de Johnson através do *software* estatístico, isto é, não foi encontrada equação de transformação que possibilitasse a normalização dos dados originais e dos dados sem *outliers*.

5.2.3 Análise do Cálculo de Capabilidade – Dados “5/8TE”

A análise dos índices de capabilidade calculados durante a pesquisa exploratória, por meio da determinação do nível sigma para o conjunto de dados “5/8TE” em quatro situações diferentes, apresentou os resultados comentados a seguir.

a) Assumindo o conjunto de dados “5/8TE” como se fossem normalmente distribuídos:

- Foi obtido um índice de $2,98\sigma$ com os dados originais, o que denotaria uma proporção de itens não conformes de 69495 PPM.

- Como a análise do conjunto de dados “5/8TE” não indicou a presença de *outliers*, obviamente não se fez necessária uma segunda verificação após a sua eliminação.

b) Adotando a função densidade de probabilidade da distribuição que melhor se ajustou ao conjunto de dados (Lognormal com 3 parâmetros):

- O índice de $3,06\sigma$, obtido neste caso, indicou um desempenho de 59886 PPM, que tem a mesma ordem de grandeza do valor obtido com a suposição de normalidade.

- Como a análise do conjunto de dados “5/8TE” não indicou a presença de *outliers*, obviamente não se fez necessária uma segunda verificação após a sua eliminação.

c) Efetuando transformação matemática do conjunto de dados por Box-Cox:

- Não foi possível efetuar a transformação de Box-Cox através do *software* estatístico, isto é, não foi encontrado valor de λ que possibilitasse a normalização dos dados.

d) Efetuando transformação matemática do conjunto de dados por Johnson:

- Não foi possível efetuar a transformação de Johnson através do *software* estatístico, isto é, não foi encontrada equação de transformação que possibilitasse a normalização dos dados originais.

5.2.4 Análise do Cálculo de Capabilidade – Dados “5/8TD”

A análise dos índices de capabilidade calculados durante a pesquisa exploratória, por meio da determinação do nível sigma para o conjunto de dados “5/8TD” em quatro situações diferentes, apresentou os resultados comentados a seguir.

a) Assumindo o conjunto de dados “5/8TD” como se fossem normalmente distribuídos:

- Foi obtido um índice de $2,36\sigma$ com os dados originais, o que denotaria uma proporção de itens não conformes de 195353 PPM.

- Após tratamento dos dados através da eliminação dos *outliers*, o índice encontrado foi de $2,77\sigma$, o que representaria um desempenho equivalente a 101247 PPM.
- b) Adotando a função densidade de probabilidade da distribuição que melhor se ajustou ao conjunto de dados (Lognormal com 3 parâmetros):
 - O índice de $2,42\sigma$, obtido neste caso, indicou um desempenho de 179265 PPM, que tem a mesma ordem de grandeza do valor obtido com a suposição de normalidade.
 - Através da eliminação dos *outliers*, obteve-se o índice de $2,82\sigma$ (92624 PPM), também com a mesma ordem de grandeza do valor obtido com suposição de normalidade.
- c) Efetuando transformação matemática do conjunto de dados por meio do método de Box-Cox:
 - Não foi possível efetuar a transformação de Box-Cox através do *software* estatístico, isto é, não foi encontrado valor de λ que possibilitasse a normalização dos dados originais e dos dados sem *outliers*.
- d) Efetuando transformação matemática do conjunto de dados por meio do método de Johnson:
 - Não foi possível efetuar a transformação de Johnson através do *software* estatístico, isto é, não foi encontrada equação de transformação que possibilitasse a normalização dos dados originais e dos dados sem *outliers*.

5.2.5 Considerações sobre Normalidade e Capabilidade no Estudo de Caso

O cálculo do índice de capabilidade, através do conceito seis sigma com os dados reais não normais, indicou neste estudo de caso que a suposição incorreta de normalidade aparentemente não produziu índices de capabilidade extremamente discrepantes em relação aos índices obtidos a partir do ajuste prévio dos dados em uma distribuição conhecida (*best fitting*). Contudo, é importante deixar claro que este fato não elimina a necessidade de confirmação da premissa de normalidade, pois nem sempre a característica dos dados estudados pode garantir uma situação de robustez em relação ao tipo de distribuição.

Além disso, as técnicas de transformação de variáveis não se mostraram eficazes no cálculo do índice de capabilidade a partir dos dados reais selecionados, pois pelo método de Box-Cox não foi possível determinar transformação que resultasse em normalidade e pelo método de Johnson em apenas um dos quatro conjuntos estudados a transformação foi bem

sucedida e forneceu dados para comparação. Este fato pode ser atribuído muito mais à característica distribucional particular dos dados do que ao pequeno tamanho da amostra, pois em estudos preliminares com amostras maiores deste mesmo processo também não foi possível obter transformação que resultasse em normalidade.

Portanto, o ajuste dos dados em uma distribuição não normal adequadamente escolhida (*best fitting*) foi a melhor opção para tratamento prévio dos dados reais selecionados neste estudo de caso antes da determinação dos índices de capacidade. A Tabela 5.3 apresenta esta avaliação de capacidade de forma resumida.

CONJUNTO DE DADOS REAIS	AVALIAÇÃO DE CAPABILIDADE ATRAVÉS DO NÍVEL SIGMA											
	SUPONDO NORMAL			NÃO NORMAL (Best Fitting)			TRANSF. BOX-COX			TRANSF. JOHNSON		
	Z _{LT}	Z _{ST} (σ)	PPM	Z _{LT}	Z _{ST} (σ)	PPM	Z _{LT}	Z _{ST} (σ)	PPM	Z _{LT}	Z _{ST} (σ)	PPM
3/4AE	1,12	2,62	130525	1,18	2,68	118550	-	-	-	1,41	2,91	79401
3/4AE* (<i>sl outliers</i>) distribuição Normal	1,33	2,83	91334	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
3/4AD	1,59	3,09	56313	1,60	3,10	55001	-	-	-	-	-	-
3/4AD* (<i>sl outliers</i>)	1,72	3,22	42972	1,74	3,24	40600	-	-	-	-	-	-
5/8TE	1,48	2,98	69495	1,56	3,06	59886	-	-	-	-	-	-
5/8TE* (<i>sl outliers</i>)	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
5/8TD	0,86	2,36	195353	0,92	2,42	179265	-	-	-	-	-	-
5/8TD* (<i>sl outliers</i>)	1,27	2,77	101247	1,32	2,82	92624	-	-	-	-	-	-

Tabela 5.3 – Determinação do nível sigma - dados reais

6. CONCLUSÃO

6.1 Conclusão do Trabalho

Através deste trabalho de pesquisa foi efetuado um estudo sobre a utilização de alguns métodos estatísticos afetados pela suposição inadequada de normalidade dos dados, em especial os gráficos de controle individuais e a avaliação de capacidade através da métrica seis sigma. Os métodos estatísticos utilizados nesta pesquisa foram escolhidos por se tratarem de procedimentos que não são “livres da distribuição”, ou seja, os seus resultados dependem da suposição relativa ao tipo de distribuição envolvida.

De modo geral o principal objetivo da pesquisa, que previa um aprofundamento dos aspectos relativos aos efeitos da não normalidade na aplicação de métodos estatísticos com comparação de algumas técnicas para tratamento dos dados, foi devidamente atingido e as seguintes afirmações fundamentam esta conclusão:

- a) Métodos estatísticos baseados na premissa de normalidade dos dados coletados necessitam de confirmação prévia desta suposição. A análise e/ou tratamento adequado dos dados é condição *sine qua non* para continuidade das atividades planejadas.
- b) No caso dos gráficos de controle individuais, confirmou-se a condição de falha na interpretação da condição de estabilidade do processo em relação ao primeiro critério de identificação de padrões não aleatórios (teste 1), pois foi observada a ocorrência de pontos além dos limites de controle estipulados nos dois conjuntos de dados não normais simulados conforme distribuições Gamma e Beta, quando a condição de normalidade foi incorretamente assumida.
- c) Para os conjuntos de dados não normais simulados, a avaliação de capacidade indicou que índices incorretos podem ser obtidos a partir da suposição equivocada de normalidade e tal fato foi confirmado através da constatação de diferenças significativas em termos de fração não conforme em partes por milhão.
- d) Em relação aos conjuntos de dados reais não normais utilizados nesta pesquisa, as divergências observadas nos valores dos índices de capacidade não foram tão pronunciadas quanto para os dados simulados; entretanto, esta situação específica atribuída principalmente à característica particular dos dados escolhidos, não dispensa, em hipótese alguma, a confirmação prévia da premissa de normalidade.

6.2 Considerações Finais e Recomendações

Alguns dos procedimentos estatísticos mais conhecidos e utilizados fundamentam-se em métodos paramétricos, isto é, são baseados na suposição de que os processos analisados correspondam a amostras aleatórias provenientes de populações com distribuição de probabilidade conhecida, como por exemplo, a distribuição Normal. Entretanto, nem sempre a premissa de normalidade é plenamente satisfeita e, conforme comentado ao longo desta dissertação, algumas decisões equivocadas podem ser tomadas em decorrência deste erro de avaliação. Na prática, quando um fato desta natureza ocorre, a primeira reação dos indivíduos envolvidos nos procedimentos em questão é colocar em dúvida a eficácia dos métodos utilizados.

A revisão da literatura, que possibilitou o embasamento teórico da pesquisa, indicou que o rigor na coleta e no tratamento dos dados deve ser cuidadosamente considerado, pois em muitos casos a aplicação dos métodos de otimização e melhoria em processos não atinge os resultados esperados em função da falta de zelo no levantamento dos dados ou, até mesmo, pelo desconhecimento de conceitos estatísticos básicos. As seguintes considerações finais complementam a argumentação apresentada ao longo desta dissertação:

- a) Em qualquer processo ou sistema no qual se pretenda interagir para garantir algum tipo de controle ou promover melhorias, o cuidado na coleta e tratamento dos dados e das informações é tão importante quanto a metodologia escolhida. Em aplicações que fazem uso de métodos estatísticos este princípio deve ser encarado como regra essencial para o bom andamento das atividades.
- b) A escolha do método estatístico a ser empregado em qualquer tipo de processo, produtivo ou empresarial, deve levar em conta a natureza dos dados e das informações disponíveis com o objetivo de possibilitar que o procedimento escolhido seja aplicado em sua plenitude e não apenas de maneira incompleta ou mesmo equivocada. Eventuais adaptações ou aplicações parciais devem ser cuidadosamente avaliadas de forma a não transgredirem os conceitos básicos da metodologia em questão.
- c) Os indivíduos envolvidos na aplicação dos métodos estatísticos devem ser adequadamente treinados de acordo com o seu escopo de atuação e devem estar previamente informados sobre os objetivos primordiais e sobre eventuais limitações em termos de implementação.

- d) O tratamento dos dados que deve anteceder a aplicação do método estatístico pode envolver, por exemplo: a revisão no procedimento de coleta dos dados; a eliminação das observações defasadas em relação ao restante dos dados (*outliers*); ou a análise dos dados em termos de médias amostrais com base no Teorema Central do Limite, que estabelece que a média de um grande número de variáveis aleatórias independentes apresenta uma distribuição aproximadamente normal, indiferentemente do tipo de distribuição dos valores individuais.
- e) Caso a condição de desvio da normalidade permaneça, mesmo após o adequado tratamento de dados, podem ser utilizadas as alternativas de ajuste dos dados em uma distribuição conhecida ou técnicas de transformação de variáveis. A escolha entre ajuste dos dados e transformação deve ser analisada caso a caso, pois depende de vários fatores, tais como, natureza dos dados, métodos estatísticos envolvidos, disponibilidade de recursos computacionais, entre outros.
- f) O ajuste dos dados pode ser utilizado quando o método estatístico em questão também for válido para a distribuição específica que apresentar o maior coeficiente de correlação de acordo com o intervalo de confiança estipulado.
- g) Caso seja possível efetuar mais de um tipo de transformação para normalidade, deve ser escolhida aquela que mais se aproximar da distribuição normal, utilizando como referência, por exemplo, a transformação com maior valor P (*p-value*) conforme o intervalo de confiança determinado.

Existem algumas idéias e sugestões que refletem limitações encontradas durante o desenvolvimento da dissertação e até mesmo advindas de possibilidades, inicialmente não vislumbradas, que surgiram ao longo da revisão bibliográfica e da aplicação da metodologia. Tais oportunidades de continuidade em relação à presente pesquisa estão resumidas na forma de recomendações e sugestões para trabalhos futuros:

- a) Analisar criticamente outros casos reais onde efetivamente a premissa de normalidade dos dados seja inadequadamente assumida nos métodos estatísticos abordados por esta pesquisa.
- b) Efetuar análise semelhante em outros métodos estatísticos, tais como Planejamento de Experimentos, Teste de Hipóteses e Análise de Regressão.
- c) Avaliar os impactos da suposição incorreta de normalidade nos métodos estatísticos aplicados a processos industriais com enfoque em custos da qualidade.

- d) Estudar os aspectos relativos aos efeitos da não normalidade de dados na interpretação de gráficos de controle e no cálculo de capacidade, comparando as técnicas aplicadas em algumas publicações internacionais com as técnicas de transformação matemática propostas por Johnson (1949) e Box & Cox (1964).
- e) Discutir se existem distribuições que apresentam maior ou menor sensibilidade às transformações de Box-Cox e Johnson por meio de simulação.
- f) Investigar se os índices de capacidade calculados através das transformações variam significativamente de acordo com o tipo de transformação, utilizando Análise de Variância (ANOVA) ou Planejamento de Experimentos (DOE).
- g) Verificar se há diferenças significativas entre as estatísticas de teste para verificação de normalidade (por exemplo: Anderson-Darling, Ryan-Joiner e Kolmogorov-Smirnov) dependendo do tipo de transformação utilizado.
- h) Efetuar estudos similares às análises descritas anteriormente, comparando e discutindo os resultados de diversos *softwares* de aplicações estatísticas.

APÊNDICE B

Dados Simulados – Distribuição Beta

GERAÇÃO DE 200 VALORES MODELADOS PELA DISTRIBUIÇÃO BETA

MINITAB

FUNÇÃO: CALC / SET BASE = 0

FUNÇÃO: CALC / RANDOM DATA / GAMMA / GENERATE = 200 / 1ST SHAPE = 25 / 2ND SHAPE = 5

TRANSFORMAÇÃO LINEAR: VALOR OBTIDO + 7,5

SEQ.	VALOR	SEQ.	VALOR	SEQ.	VALOR	SEQ.	VALOR	SEQ.	VALOR
1	8,45807	41	8,24211	81	8,35032	121	8,36097	161	8,33628
2	8,41165	42	8,42695	82	8,33218	122	8,35226	162	8,36631
3	8,37419	43	8,26819	83	8,35262	123	8,41242	163	8,30414
4	8,37566	44	8,32911	84	8,37028	124	8,32334	164	8,30966
5	8,34414	45	8,2955	85	8,28237	125	8,30704	165	8,29207
6	8,43159	46	8,30838	86	8,40815	126	8,36446	166	8,34654
7	8,31396	47	8,36016	87	8,26452	127	8,15895	167	8,4471
8	8,25829	48	8,22799	88	8,39707	128	8,34006	168	8,37651
9	8,35646	49	8,30434	89	8,44931	129	8,36211	169	8,44165
10	8,37912	50	8,31687	90	8,44039	130	8,35366	170	8,26546
11	8,32773	51	8,37933	91	8,3445	131	8,27866	171	8,32307
12	8,42854	52	8,29512	92	8,40923	132	8,30233	172	8,2771
13	8,37079	53	8,26892	93	8,25715	133	8,34222	173	8,32263
14	8,38604	54	8,42167	94	8,31656	134	8,332	174	8,35912
15	8,32042	55	8,39958	95	8,32098	135	8,37317	175	8,32018
16	8,38134	56	8,33963	96	8,33512	136	8,34075	176	8,36889
17	8,27282	57	8,31593	97	8,42927	137	8,37552	177	8,25526
18	8,27416	58	8,44325	98	8,37	138	8,34316	178	8,39565
19	8,21798	59	8,35471	99	8,36738	139	8,31217	179	8,3524
20	8,35701	60	8,35188	100	8,26551	140	8,3567	180	8,31271
21	8,36542	61	8,31315	101	8,28658	141	8,37056	181	8,26656
22	8,30253	62	8,35793	102	8,26359	142	8,45418	182	8,38219
23	8,2963	63	8,32368	103	8,29687	143	8,32273	183	8,34218
24	8,39395	64	8,27083	104	8,3654	144	8,34553	184	8,37404
25	8,27878	65	8,30975	105	8,26508	145	8,40452	185	8,31349
26	8,28014	66	8,38163	106	8,26195	146	8,33615	186	8,37163
27	8,17534	67	8,1599	107	8,3006	147	8,45563	187	8,34937
28	8,36017	68	8,32049	108	8,27515	148	8,44982	188	8,29807
29	8,30082	69	8,3229	109	8,28111	149	8,36015	189	8,20219
30	8,38809	70	8,31127	110	8,28591	150	8,32073	190	8,29404
31	8,28877	71	8,38843	111	8,34802	151	8,42544	191	8,33183
32	8,39732	72	8,36971	112	8,32393	152	8,39719	192	8,27881
33	8,34748	73	8,31102	113	8,30508	153	8,35878	193	8,35994
34	8,41231	74	8,34722	114	8,32114	154	8,44541	194	8,36257
35	8,37241	75	8,39904	115	8,36113	155	8,25975	195	8,20187
36	8,2995	76	8,17852	116	8,32957	156	8,39035	196	8,34077
37	8,34269	77	8,20045	117	8,33991	157	8,33473	197	8,36922
38	8,20745	78	8,38738	118	8,2819	158	8,16859	198	8,38233
39	8,38114	79	8,17685	119	8,34647	159	8,25734	199	8,28242
40	8,37774	80	8,39094	120	8,34671	160	8,23742	200	8,34775

APÊNDICE C

Dados Reais – Furos 3/4”

COLETA DE 2 GRUPOS DE MEDIDAS DE DIÂMETRO DE FUROS 3/4”

ORIGEM DOS DADOS:

PROCESSO DE FURAÇÃO DE PRECISÃO EM 2 TIPOS DE COMPONENTE ESTRUTURAL DE ALUMÍNIO PARA FIXAÇÃO DE PARAFUSOS COM DIÂMETRO DE 3/4 POLEGADAS. EMPRESA DO INTERIOR DO ESTADO DE SÃO PAULO.

TOLERÂNCIAS DE PROJETO:

LIE = 19,124 MM

LSE = 19,151 MM

54 MEDICÕES (3/4AE)			
SEQ.	MEDIDA	SEQ.	MEDIDA
1	19,130	28	19,128
2	19,145	29	19,124
3	19,126	30	19,127
4	19,13	31	19,147
5	19,136	32	19,145
6	19,139	33	19,125
7	19,122	34	19,130
8	19,127	35	19,134
9	19,122	36	19,135
10	19,128	37	19,126
11	19,123	38	19,129
12	19,126	39	19,130
13	19,122	40	19,131
14	19,127	41	19,128
15	19,127	42	19,132
16	19,132	43	19,126
17	19,133	44	19,131
18	19,135	45	19,128
19	19,129	46	19,134
20	19,129	47	19,127
21	19,129	48	19,130
22	19,135	49	19,127
23	19,128	50	19,132
24	19,128	51	19,129
25	19,131	52	19,134
26	19,133	53	19,128
27	19,124	54	19,128

60 MEDICÕES (3/4AD)			
SEQ.	MEDIDA	SEQ.	MEDIDA
1	19,129	31	19,129
2	19,131	32	19,132
3	19,137	33	19,124
4	19,132	34	19,127
5	19,130	35	19,125
6	19,135	36	19,127
7	19,137	37	19,125
8	19,137	38	19,128
9	19,136	39	19,129
10	19,137	40	19,131
11	19,131	41	19,127
12	19,132	42	19,128
13	19,133	43	19,125
14	19,134	44	19,128
15	19,136	45	19,129
16	19,138	46	19,130
17	19,135	47	19,130
18	19,136	48	19,130
19	19,126	49	19,133
20	19,128	50	19,131
21	19,121	51	19,130
22	19,123	52	19,131
23	19,128	53	19,130
24	19,120	54	19,130
25	19,136	55	19,129
26	19,133	56	19,130
27	19,130	57	19,130
28	19,131	58	19,131
29	19,131	59	19,13
30	19,138	60	19,129

APÊNDICE D

Dados Reais – Furos 5/8”

COLETA DE 2 GRUPOS DE MEDIDAS DE DIÂMETRO DE FUROS 5/8”

ORIGEM DOS DADOS:

PROCESSO DE FURAÇÃO DE PRECISÃO EM 2 TIPOS DE COMPONENTE ESTRUTURAL DE TITÂNIO PARA FIXAÇÃO DE PARAFUSOS COM DIÂMETRO DE 5/8 POLEGADAS. EMPRESA DO INTERIOR DO ESTADO DE SÃO PAULO.

TOLERÂNCIAS DE PROJETO:

LIE = 15,950 MM

LSE = 15,977 MM

27 MEDIÇÕES (5/8TE)	
SEQ.	MEDIDA
1	15,956
2	15,980
3	15,956
4	15,980
5	15,973
6	15,967
7	15,963
8	15,968
9	15,972
10	15,975
11	15,966
12	15,971
13	15,959
14	15,959
15	15,959
16	15,966
17	15,959
18	15,959
19	15,958
20	15,957
21	15,956
22	15,960
23	15,958
24	15,958
25	15,960
26	15,96
27	15,958

29 MEDIÇÕES (5/8TD)	
SEQ.	MEDIDA
1	15,998
2	15,975
3	15,974
4	15,974
5	15,976
6	15,976
7	15,978
8	15,966
9	15,969
10	15,973
11	15,981
12	15,973
13	15,968
14	15,960
15	15,966
16	15,957
17	15,957
18	15,958
19	15,960
20	15,956
21	15,961
22	15,964
23	15,958
24	15,973
25	15,962
26	15,961
27	15,959
28	15,960
29	15,959

REFERÊNCIAS BIBLIOGRÁFICAS

ALMAS, F. – Implementação de Controle Estatístico de Processos em uma Empresa Têxtil. Dissertação de Mestrado do Programa de Pós-Graduação em Engenharia de Produção, Universidade Federal de Itajubá. Itajubá: UNIFEI, 2003.

ALMEIDA, S.S. – Desenvolvimento de Gráficos de Controle Aplicados ao Modelo Funcional de Regressão. Tese de Doutorado do Programa de Pós-Graduação em Engenharia de Produção, Universidade Federal de Santa Catarina. Florianópolis: UFSC, 2003.

ALWAN, L. and ROBERTS, H.V. – *The Problem of Misplaced Control Limits. Applied Statistics*, v. 44, n. 3, p. 269-278, 1995.

ASQC/AIAG – *American Society for Quality Control/Automotive Industry Action Group Task Force. Statistical Process Control: Reference Manual – Chrysler Corporation, Ford Motor Company and General Motor Corporation, 1992.*

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

AWAN, H.M. – *Effect of Departures from Standard Assumptions Used in Analysis of Variance. Journal of Research (Science), Bahauddin Zacariya University, Pakistan*, v. 12, n. 2, p. 180-188, December, 2001.

BALESTRASSI, P.P. – Identificação de Padrões em Gráficos de Controle Estatístico de Processos em Tempo Real Utilizando Séries Temporais e Redes Neurais Artificiais. Tese de Doutorado do Programa de Pós-Graduação em Engenharia de Produção, Universidade Federal de Santa Catarina. Florianópolis: UFSC, 2000.

BAYEUX, C. – Análise da Capabilidade e Processos (Parte III). *Banas Qualidade*, p. 75-78, Julho, 2001.

BERTRAND, J.W.M. and FRANSOO, J.C. – Modelling and Simulation: *Operations Management Research Methodologies using Quantitative Modeling. International Journal of Operations Production Management*, v. 22, n. 2, p. 241-264, 2002.

- BORROR, C.M., MONTGOMERY, D.C. and RUNGER, G.C. – *Robustness of EWMA Control Chart to Non-Normality*. *Journal of Quality Technology*, v. 31, n. 3, 1999.
- BOWER, K.M. – *Using Exponentially Weighted Moving Average (EWMA) Charts*. *Asia Pacific Process Engineer*, October, 2000.
- BOX, G.E.P. and COX, D.R. – *An Analysis of Transformations*. *Journal of Royal Statistical Society, B*, 39, p. 211-252, 1964.
- BOX, G.E.P. and TIDWELL, P.W. – *Transformation of Independent Variables*. *Technometrics*, v. 4, n. 4, p. 531-550, November, 1962.
- BOX, G.E.P., HUNTER, W.G. and HUNTER, J.S. – *Statistics for Experimenters*. New York: John Wiley & Sons, Inc., 1978.
- BREYFOGLE III, F.W. – *Implementing Six Sigma: Smarter Solutions Using Statistical Methods*. New York: John Wiley, 1999.
- CAMPOS, V.F. – *TQC: Controle da Qualidade Total*. 5ª ed. Belo Horizonte: Fundação Christiano Ottoni, 1992.
- CARVALHO, M.M. – *Medindo o Sigma do Processo*. In: ROTONDARO, R.G. (Org.) – *Seis Sigma: Estratégia Gerencial para Melhoria de Processos, Produtos e Serviços*. São Paulo: Editora Atlas, 2002. cap. 6, p. 165-175.
- CASTAGLIOLA, P. and TSUNG, F. – *Autocorrelated SPC for Non-Normal Situations*. *Quality and Reliability Engineering International*, 21, p. 131-161, January, 2005.
- CHEN, W., BAGHDASARYAN, L., BURANATHITI, T. and CAO, J. – *Model Validation via Uncertainty Propagation and Data Transformations*. Revision 2, submitted to the *AIAA Journal*, December, 2003.
- DEMING, W.E. – *Out of the Crisis*. Cambridge: Massachusetts Institute of Technology, Center for Advanced Engineering Study, 1986.
- FEIGENBAUM, A.V. – *Controle da Qualidade Total*, vol. 3, *Métodos Estatísticos Aplicados à Qualidade*. 3ª ed. São Paulo: Makron Books, 1994.

GALUCH, L. – Modelo para Implementação das Ferramentas Básicas do Controle Estatístico de Processo em Pequenas Empresas Manufatureiras. Dissertação de Mestrado do Programa de Pós-Graduação em Engenharia de Produção, Universidade Federal de Santa Catarina. Florianópolis: UFSC, 2002.

GARVIN, D. A. – Gerenciando a Qualidade: A Visão Estratégica e Competitiva. Rio de Janeiro: Qualitymark, 2002.

GARVIN, D.A. – *What Does “Product Quality” Really Mean? Sloan Management Review*, v. 26, n. 1, p. 25-43, Fall 1984.

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

GRANT, E.L. and LEAVENWORTH, R.S. – *Statistical Quality Control*. 4th ed. Tokyo: McGraw-Hill Kogakusha, Ltd, 1972.

HARRINGTON, J. – Aperfeiçoando Processos Empresariais. São Paulo: McGraw-Hill Ltda e Makron Books do Brasil Ltda, 1993.

HARRY, M.J. – *Resolving the Mysteries of Six Sigma: Statistical Constructs and Engineering Rationale*. Phoenix: Palladyne Publishing, 2003.

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

HARRY, M.J. and SCHROEDER, R. – *Six Sigma: The Breakthrough Management Strategy Revolutionizing the World's Top Corporation*. New York: Currency, 2000.

HOERL, R.W. – *Six Sigma Black Belt: What Do They Need to Know? Journal of Quality Technology*, v. 33, n. 4, p. 391-406, October 2001.

JOHNSON, N.L. – *Systems of Frequency Curves Generated by Methods of Translation*. *Biometrika*, v. 36, p. 149-176, 1949.

JURAN, J.M. and GODFREY, A.B. (Org.) – *Juran's Quality Handbook*. 5th ed. New York: McGraw-Hill, Inc., 1999.

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

KOTZ, S. and JOHNSON, N.L. – *Process Capability Indices - A Review, 1992-2000. Journal of Quality Technology*, v. 34, n. 1, p. 2-19, January, 2002.

KUME, H. – *Métodos Estatísticos para Melhoria da Qualidade*. 11^a ed. São Paulo: Editora Gente, 1993.

LEVINSON, W.A. – *SPC for Real-World Processes: What to Do When the Normality Assumption Doesn't Work. Annual Quality Congress, Indianapolis, May 2000*.

LIBERATORE, R.L. – *Teaching the Role of SPC in Industrial Statistics. Quality Progress*, p. 89-94, July, 2001.

LINDERMAN, K., SCHROEDER, R.G., ZAHEER, S. and CHOO, A.S. – *Six Sigma: A Goal-Theoretic Perspective. Journal of Operations Management*, v. 21, n. 2, p. 193-203, 2003.

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

MIRANDA, R.G. – *Um Modelo para Análise de Capacidade de Processos com Ênfase na Transformação de Dados*. Dissertação de Mestrado do Programa de Pós-Graduação em Engenharia de Produção, Universidade Federal de Santa Catarina. Florianópolis: UFSC, 2005.

MONTGOMERY, D.C. – *Design and Analysis of Experiments. 2nd ed. New York: John Wiley & Sons, Inc, 1984*.

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

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

ONISHI, H. – *A Knowledge-Based Variable Selection Method for Box-Cox Transformation. Journal Japan Statistical Society*, v. 32, n. 1, p. 15-42, 2002.

PALADINI, E.P. – *Controle de Qualidade: Uma Abordagem Abrangente*. São Paulo: Editora Atlas, 1990.

PEREZ-WILSON, M. – Seis Sigma: Compreendendo o Conceito, as Implicações e os Desafios. Rio de Janeiro: Qualitymark, 1998.

PIRES, V.T. – Implantação do Controle Estatístico de Processos em uma Empresa de Manufatura de Óleo de Arroz. Dissertação de Mestrado do Programa de Pós-Graduação em Engenharia de Produção, Universidade Federal do Rio Grande do Sul. Porto Alegre: UFRGS, 2000.

PRAZERES, P.M. – Mini-Dicionário de Termos da Qualidade. São Paulo: Editora Atlas, 1997.

RAMOS, E.M.L.S. – Aperfeiçoamento de Desenvolvimento de Ferramentas do Controle Estatístico da Qualidade Utilizando Quartis para Estimar o Desvio-Padrão. Tese de Doutorado do Programa de Pós-Graduação em Engenharia de Produção, Universidade Federal de Santa Catarina. Florianópolis: UFSC, 2003.

RATH & STRONG MANAGEMENT CONSULTANTS – *Black Belt Training, Normal Theory Module, 2000.*

REIS, D.A.F. – Seis Sigma: Um Estudo Aplicado ao Setor Eletrônico. Dissertação de Mestrado Profissionalizante em Engenharia, Universidade Federal do Rio Grande do Sul. Porto Alegre: UFRGS, 2003.

REIS, M.M. – Um Modelo para o Ensino do Controle Estatístico da Qualidade. Tese de Doutorado do Programa de Pós-Graduação em Engenharia de Produção, Universidade Federal de Santa Catarina. Florianópolis: UFSC, 2001.

ROTH II, G. – *Capability Indexes: Mytery Solved. Six Sigma Fórum Magazine, v. 4, n. 3, p. 17-21, May, 2005.*

ROTONDARO, R.G. (Org.) – Seis Sigma: Estratégia Gerencial para Melhoria de Processos, Produtos e Serviços. São Paulo: Editora Atlas, 2002.

SALL, J. – *Leptokurtosiphobia: Irrational Fear of Non-Normality. Six Sigma Forum Magazine, v. 4, n. 3, p. 11-15, May, 2005.*

SARKADI, K. and VINCZE, I. – *Mathematical Methods of Statistical Quality Control. New York and London: Academic Press, 1974.*

SCHISSATTI, M.L. – Uma Metodologia de Implantação de Cartas de Shewhart para o Controle de Processos. Dissertação de Mestrado do Programa de Pós-Graduação em Engenharia de Produção, Universidade Federal de Santa Catarina. Florianópolis: UFSC, 1998.

SEVERINO, A.J. – Metodologia do Trabalho Científico. 22ª ed. São Paulo: Cortez, 2002.

SILVA, L.S.C.V – Aplicação do Controle Estatístico de Processos na Indústria de Laticínios Lactoplasa: Um Estudo de Caso. Dissertação de Mestrado do Programa de Pós-Graduação em Engenharia de Produção, Universidade Federal de Santa Catarina. Florianópolis: UFSC, 1999.

SLACK, N., CHAMBERS, S., HARLAND, C., HARRISON, A. e JOHNSTON, R. – Administração da Produção. 1ª ed. São Paulo: Editora Atlas, 1999.

SOARES, G.M.V.P.P. - Aplicação do Controle Estatístico de Processos em Indústria de Bebidas: Um Estudo de Caso. Dissertação de Mestrado do Programa de Pós-Graduação em Engenharia de Produção, Universidade Federal de Santa Catarina. Florianópolis: UFSC, 2001.

SOARES, V.L.M.P. – Aplicação e Implantação do Controle Estatístico de Processo em Pintura Industrial. Dissertação de Mestrado do Programa de Pós-Graduação em Engenharia de Produção, Universidade Federal de Santa Catarina. Florianópolis: UFSC, 2003.

SOUZA, R.A. – Análise da Qualidade do Processo de Envase de Azeitonas Verdes Através de Algumas Ferramentas do Controle Estatístico de Processo. Dissertação de Mestrado do Programa de Pós-Graduação em Engenharia de Produção, Universidade Federal de Santa Catarina. Florianópolis: UFSC, 2003.

SOUZA, V.Z. – Otimização dos Parâmetros de Controle de um Processo de Fabricação de Latas: Uma Abordagem Seis Sigma. Dissertação de Mestrado do Programa de Pós-Graduação em Engenharia de Produção, Universidade Federal de Itajubá. Itajubá: UNIFEI, 2002.

SPEEDING, T.A. and RAWLINGS, P.L. – *Non-Normality in Statistical Process Control Measurements. The International Journal of Quality & Reliability Management*, v. 11, n. 6, p. 27-37, 1994.

TOLEDO, T.P.A. – Uma Investigação sobre Índices de Capabilidade com Ênfase na Metodologia Seis Sigma. Dissertação de Mestrado do Programa de Pós-Graduação em Engenharia de Produção, Universidade Federal de Itajubá. Itajubá: UNIFEI, 2005.

WERKEMA, M.C.C. – As Ferramentas da Qualidade no Gerenciamento de Processos. Belo Horizonte: Editora de Desenvolvimento Gerencial, 1995.

WOODALL, W.H. – *Controversies and Contradictions in Statistical Process Control. Journal of Quality Technology*, v. 32, n. 4, p. 341-350, October, 2000.

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

BIBLIOGRAFIA COMPLEMENTAR CONSULTADA

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

BALESTRASSI, P.P. – Curso de Estatística Aplicada. Apostila de Curso do Programa de Pós-Graduação em Engenharia de Produção, Universidade Federal de Itajubá. Itajubá: UNIFEI, 2004.

BALESTRASSI, P.P. – Metodologia Seis Sigma. Apostila de Curso do Programa de Pós-Graduação em Engenharia de Produção, Universidade Federal de Itajubá. Itajubá: UNIFEI, 2004.

BARROS, A.J.P. e LEHFELD, N.A.S. – Fundamentos de Metodologia: Um Guia para a Iniciação Científica. São Paulo: McGraw-Hill, 1986.

BASTOS, C.L. e KELLER, V. – Introdução à Metodologia Científica. 9ª ed. Petrópolis: 1997.

BREYFOGLE III, F.W. and CONNOLLY, M. – *Six Sigma Methods to Ensure Organizations' Health. R&D Magazine, v. 45, n. 4, p. 28-29, April, 2003.*

BRYMAN, A. – *Research Methods and Organization Studies. London: Unwin Hyman, 1989.*

COLLINS, M. (Org.) – Michaelis – Pequeno Dicionário Inglês-Português, Português-Inglês. 44ª ed. São Paulo: Melhoramentos, 1994.

DUCLOS, E. – *Contribution à la Maîtrise Statistique des Procédés cas des Procédés Non Normaux. These présentée à L'Ecole Supérieure D'Ingenieurs D'Annecy (ESIA), Université de Savoie, 1997.*

FERREIRA, A.B.H. – Mini-dicionário da Língua Portuguesa. 2ª ed. Rio de Janeiro: Editora Nova Fronteira, 1989.

JANISZEWSKI, S. and GEORGE, E. – *Integrating PSP, TSP and Six Sigma. American Society for Quality – ASQ/SQP, v. 6, n. 4, p. 4-13, September, 2004.*

KANHOLM, J. – ISO 9000 Explicada: Lista de Verificação com 65 Requisitos e Guia de Conformidade. São Paulo: Pioneira, 1995.

KEUZENKAMP, H.A. and MAGNUS, J.R. – *On Tests and Significance in Econometrics. Journal of Econometrics*, 67, p. 5-24, 1995.

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

LAKATOS, E.M. e MARCONI, M.A. – Fundamentos de Metodologia Científica. 3ª ed. São Paulo: Editora Atlas, 1991.

MARCONI, M.A. e LAKATOS, E.M. – Técnicas de Pesquisa: Planejamento e Execução de Pesquisas, Amostras e Técnicas de Pesquisas, Elaboração, Análise e Interpretação de Dados. 4ª ed. São Paulo: Editora Atlas, 1999.

MEMÓRIA, J.M.P. – Breve História da Estatística. Brasília: Embrapa Informação Tecnológica, 2004.

MINITAB INC. – *Minitab® Statistical Software, User's Guide, Release 14 for Windows, State College, Pennsylvania, 2003.*

MOURA, J.C.R. – Um Modelo de Melhoria Contínua de Processos para o Setor Público. Dissertação de Mestrado do Programa de Pós-Graduação em Engenharia de Produção, Universidade Federal de Santa Catarina. Florianópolis: UFSC, 2002.

NIST/SEMATECH. – *e-Handbook of Statistical Methods*. Disponível em: <http://www.itl.nist.gov/div898/handbook>. Acesso em 26 de janeiro de 2005.

SALIBY, E. – Repensando a Simulação: A Amostragem Descritiva. Rio de Janeiro: Editora da UFRJ, 1989.

SALVADOR, A.D. – Métodos e Técnicas de Pesquisa Bibliográfica. 9ª ed. Porto Alegre: Editora Sulina, 1981.

SEYMOUR-SMITH, M. – Os Cem Livros que mais Influenciaram a Humanidade: A História do Pensamento dos Tempos Antigos à Atualidade. 5ª ed. Rio de Janeiro: DIFEL, 2002.

STATSOFT, INC. – *Electronic Statistics Textbook. Tulsa, OK: StatSoft*. Disponível em: <http://www.statsoftinc.com/textbook/stathome.html>. Acesso em 01 de abril de 2005.

USEVICIUS, L.A. – Implantação da Metodologia Seis Sigma e Aplicação da Técnica Estatística Projeto de Experimentos na Resolução de Problemas e Otimização de Processos de Fabricação. Dissertação de Mestrado Profissionalizante em Engenharia, Universidade Federal do Rio Grande do Sul. Porto Alegre: UFRGS, 2004.

YI, J., PRYBUTOK, V.R. and CLAYTON, H.R. – *ARL Comparisons Between Neural Network Models and \bar{X} -Control Charts for Quality Characteristics that are Nonnormally Distributed. Economic Quality Control, v. 16, n. 1, p. 5-15, 2001.*

YIN, R.K. – *Case Study Research: Design and Methods. Sage: Newbury Park, 1989.*

ANEXO A

Tabela das Áreas sob a Curva Normal Padronizada

TABLE A Areas Under the Normal Curve Proportion of Total Area Under the Curve from $-\infty$ to $(x, - \mu)/\sigma$

$\frac{x - \mu}{\sigma}$	0.09	0.08	0.07	0.06	0.05	0.04	0.03	0.02	0.01	0.00
-3.5	0.00017	0.00017	0.00018	0.00019	0.00019	0.00020	0.00021	0.00022	0.00022	0.00023
-3.4	0.00024	0.00025	0.00026	0.00027	0.00028	0.00029	0.00030	0.00031	0.00033	0.00034
-3.3	0.00035	0.00036	0.00038	0.00039	0.00040	0.00042	0.00043	0.00045	0.00047	0.00048
-3.2	0.00050	0.00052	0.00054	0.00056	0.00058	0.00060	0.00062	0.00064	0.00066	0.00069
-3.1	0.00071	0.00074	0.00076	0.00079	0.00082	0.00085	0.00087	0.00090	0.00094	0.00097
-3.0	0.00100	0.00104	0.00107	0.00111	0.00114	0.00118	0.00122	0.00126	0.00131	0.00135
-2.9	0.0014	0.0014	0.0015	0.0015	0.0016	0.0016	0.0017	0.0017	0.0018	0.0019
-2.8	0.0019	0.0020	0.0021	0.0021	0.0022	0.0023	0.0023	0.0024	0.0025	0.0026
-2.7	0.0026	0.0027	0.0028	0.0029	0.0030	0.0031	0.0032	0.0033	0.0034	0.0035
-2.6	0.0036	0.0037	0.0038	0.0039	0.0040	0.0041	0.0043	0.0044	0.0045	0.0047
-2.5	0.0048	0.0049	0.0051	0.0052	0.0054	0.0055	0.0057	0.0059	0.0060	0.0062
-2.4	0.0064	0.0066	0.0068	0.0069	0.0071	0.0073	0.0075	0.0078	0.0080	0.0082
-2.3	0.0084	0.0087	0.0089	0.0091	0.0094	0.0096	0.0099	0.0102	0.0104	0.0107
-2.2	0.0110	0.0113	0.0116	0.0119	0.0122	0.0125	0.0129	0.0132	0.0136	0.0139
-2.1	0.0143	0.0146	0.0150	0.0154	0.0158	0.0162	0.0166	0.0170	0.0174	0.0179
-2.0	0.0183	0.0188	0.0192	0.0197	0.0202	0.0207	0.0212	0.0217	0.0222	0.0228
-1.9	0.0233	0.0239	0.0244	0.0250	0.0256	0.0262	0.0268	0.0274	0.0281	0.0287
-1.8	0.0294	0.0301	0.0307	0.0314	0.0322	0.0329	0.0336	0.0344	0.0351	0.0359
-1.7	0.0367	0.0375	0.0384	0.0392	0.0401	0.0409	0.0418	0.0427	0.0436	0.0446
-1.6	0.0455	0.0465	0.0475	0.0485	0.0495	0.0505	0.0516	0.0526	0.0537	0.0548
-1.5	0.0559	0.0571	0.0582	0.0594	0.0606	0.0618	0.0630	0.0643	0.0655	0.0668
-1.4	0.0681	0.0694	0.0708	0.0721	0.0735	0.0749	0.0764	0.0778	0.0793	0.0808
-1.3	0.0823	0.0838	0.0853	0.0869	0.0885	0.0901	0.0918	0.0934	0.0951	0.0968
-1.2	0.0985	0.1003	0.1020	0.1038	0.1057	0.1075	0.1093	0.1112	0.1131	0.1151
-1.1	0.1170	0.1190	0.1210	0.1230	0.1251	0.1271	0.1292	0.1314	0.1335	0.1357
-1.0	0.1379	0.1401	0.1423	0.1446	0.1469	0.1492	0.1515	0.1539	0.1562	0.1587
-0.9	0.1611	0.1635	0.1660	0.1685	0.1711	0.1736	0.1762	0.1788	0.1814	0.1841
-0.8	0.1867	0.1894	0.1922	0.1949	0.1977	0.2005	0.2033	0.2061	0.2090	0.2119
-0.7	0.2148	0.2177	0.2207	0.2236	0.2266	0.2297	0.2327	0.2358	0.2389	0.2420
-0.6	0.2451	0.2483	0.2514	0.2546	0.2578	0.2611	0.2643	0.2676	0.2709	0.2743
-0.5	0.2776	0.2810	0.2843	0.2877	0.2912	0.2946	0.2981	0.3015	0.3050	0.3085
-0.4	0.3121	0.3156	0.3192	0.3228	0.3264	0.3300	0.3336	0.3372	0.3409	0.3446
-0.3	0.3483	0.3520	0.3557	0.3594	0.3632	0.3669	0.3707	0.3745	0.3783	0.3821
-0.2	0.3859	0.3897	0.3936	0.3974	0.4013	0.4052	0.4090	0.4129	0.4168	0.4207
-0.1	0.4247	0.4286	0.4325	0.4364	0.4404	0.4443	0.4483	0.4522	0.4562	0.4602
-0.0	0.4641	0.4681	0.4721	0.4761	0.4801	0.4840	0.4880	0.4920	0.4960	0.5000

Fonte: Grant & Leavenworth (1972)

ANEXO A (continuação)

Tabela das Áreas sob a Curva Normal Padronizada (continuação)

TABLE A Areas Under the Normal Curve (continued)

$\frac{x_i - \mu}{\sigma}$	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
+0.0	0.5000	0.5040	0.5080	0.5120	0.5160	0.5199	0.5239	0.5279	0.5319	0.5359
+0.1	0.5398	0.5438	0.5478	0.5517	0.5557	0.5596	0.5636	0.5675	0.5714	0.5753
+0.2	0.5793	0.5832	0.5871	0.5910	0.5948	0.5987	0.6026	0.6064	0.6103	0.6141
+0.3	0.6179	0.6217	0.6255	0.6293	0.6331	0.6368	0.6406	0.6443	0.6480	0.6517
+0.4	0.6554	0.6591	0.6628	0.6664	0.6700	0.6736	0.6772	0.6808	0.6844	0.6879
+0.5	0.6915	0.6950	0.6985	0.7019	0.7054	0.7088	0.7123	0.7157	0.7190	0.7224
+0.6	0.7257	0.7291	0.7324	0.7357	0.7389	0.7422	0.7454	0.7486	0.7517	0.7549
+0.7	0.7580	0.7611	0.7642	0.7673	0.7704	0.7734	0.7764	0.7794	0.7823	0.7852
+0.8	0.7881	0.7910	0.7939	0.7967	0.7995	0.8023	0.8051	0.8079	0.8106	0.8133
+0.9	0.8159	0.8186	0.8212	0.8238	0.8264	0.8289	0.8315	0.8340	0.8365	0.8389
+1.0	0.8413	0.8438	0.8461	0.8485	0.8508	0.8531	0.8554	0.8577	0.8599	0.8621
+1.1	0.8643	0.8665	0.8686	0.8708	0.8729	0.8749	0.8770	0.8790	0.8810	0.8830
+1.2	0.8849	0.8869	0.8888	0.8907	0.8925	0.8944	0.8962	0.8980	0.8997	0.9015
+1.3	0.9032	0.9049	0.9066	0.9082	0.9099	0.9115	0.9131	0.9147	0.9162	0.9177
+1.4	0.9192	0.9207	0.9222	0.9236	0.9251	0.9265	0.9279	0.9292	0.9306	0.9319
+1.5	0.9332	0.9345	0.9357	0.9370	0.9382	0.9394	0.9406	0.9418	0.9429	0.9441
+1.6	0.9452	0.9463	0.9474	0.9484	0.9495	0.9505	0.9515	0.9525	0.9535	0.9545
+1.7	0.9554	0.9564	0.9573	0.9582	0.9591	0.9599	0.9608	0.9616	0.9625	0.9633
+1.8	0.9641	0.9649	0.9656	0.9664	0.9671	0.9678	0.9686	0.9693	0.9699	0.9706
+1.9	0.9713	0.9719	0.9726	0.9732	0.9738	0.9744	0.9750	0.9756	0.9761	0.9767
+2.0	0.9773	0.9778	0.9783	0.9788	0.9793	0.9798	0.9803	0.9808	0.9812	0.9817
+2.1	0.9821	0.9826	0.9830	0.9834	0.9838	0.9842	0.9846	0.9850	0.9854	0.9857
+2.2	0.9861	0.9864	0.9868	0.9871	0.9875	0.9878	0.9881	0.9884	0.9887	0.9890
+2.3	0.9893	0.9896	0.9898	0.9901	0.9904	0.9906	0.9909	0.9911	0.9913	0.9916
+2.4	0.9918	0.9920	0.9922	0.9925	0.9927	0.9929	0.9931	0.9932	0.9934	0.9936
+2.5	0.9938	0.9940	0.9941	0.9943	0.9945	0.9946	0.9948	0.9949	0.9951	0.9952
+2.6	0.9953	0.9955	0.9956	0.9957	0.9959	0.9960	0.9961	0.9962	0.9963	0.9964
+2.7	0.9965	0.9966	0.9967	0.9968	0.9969	0.9970	0.9971	0.9972	0.9973	0.9974
+2.8	0.9974	0.9975	0.9976	0.9977	0.9977	0.9978	0.9979	0.9979	0.9980	0.9981
+2.9	0.9981	0.9982	0.9983	0.9983	0.9984	0.9984	0.9985	0.9985	0.9986	0.9986
+3.0	0.99865	0.99869	0.99874	0.99878	0.99882	0.99886	0.99889	0.99893	0.99896	0.99900
+3.1	0.99903	0.99906	0.99910	0.99913	0.99915	0.99918	0.99921	0.99924	0.99926	0.99929
+3.2	0.99931	0.99934	0.99936	0.99938	0.99940	0.99942	0.99944	0.99946	0.99948	0.99950
+3.3	0.99952	0.99953	0.99955	0.99957	0.99958	0.99960	0.99961	0.99962	0.99964	0.99965
+3.4	0.99966	0.99967	0.99969	0.99970	0.99971	0.99972	0.99973	0.99974	0.99975	0.99976
+3.5	0.99977	0.99978	0.99978	0.99979	0.99980	0.99981	0.99981	0.99982	0.99983	0.99983

Fonte: Grant & Leavenworth (1972)

ANEXO B

Fatores para Construção de Gráficos de Controle

Tamanho da Amostra (n)	Fator para Limites de Controle				
	Gráfico \bar{X}			Gráfico R	
	A_1	A_2	d_2	D_3	D_4
2	3,760	1,880	1,128	0	3,267
3	2,394	1,023	1,693	0	2,575
4	1,880	0,729	2,059	0	2,282
5	1,596	0,577	2,326	0	2,115
6	1,410	0,483	2,534	0	2,004
7	1,277	0,419	2,704	0,076	1,924
8	1,175	0,373	2,847	0,136	1,864
9	1,094	0,337	2,970	0,184	1,816
10	1,028	0,308	3,078	0,223	1,777
11	0,973	0,285	3,173	0,256	1,744
12	0,925	0,266	3,258	0,284	1,716
13	0,884	0,249	3,336	0,308	1,692
14	0,848	0,235	3,407	0,329	1,671
15	0,816	0,223	3,472	0,348	1,652
16	0,788	0,212	3,532	0,364	1,636
17	0,762	0,203	3,588	0,379	1,621
18	0,738	0,194	3,640	0,392	1,608
19	0,717	0,187	3,689	0,404	1,596
20	0,697	0,180	3,735	0,414	1,586
21	0,679	0,173	3,778	0,425	1,575
22	0,662	0,167	3,819	0,434	1,566
23	0,647	0,162	3,858	0,443	1,557
24	0,632	0,157	3,895	0,452	1,548
25	0,619	0,153	3,931	0,459	1,541

Fonte: Montgomery & Runger (2003)

ANEXO C

Tabela de Referência entre Índice Sigma e PPM

Nível Sigma Z_{ST} (curto prazo)	$Z_{ST} = Z_{LT} + 1,5$ Z_{LT} (longo prazo)	PPM	Nível Sigma Z_{ST} (curto prazo)	$Z_{ST} = Z_{LT} + 1,5$ Z_{LT} (longo prazo)	PPM
1,0	-0,5	697.672,15	3,6	2,1	17.864,53
1,1	-0,4	660.082,92	3,7	2,2	13.903,50
1,2	-0,3	621.378,38	3,8	2,3	10.740,14
1,3	-0,2	581.814,88	3,9	2,4	8.197,56
1,4	-0,1	541.693,78	4,0	2,5	6.209,70
1,5	0	501.349,97	4,1	2,6	4.661,23
1,6	0,1	461.139,78	4,2	2,7	3.467,03
1,7	0,2	421.427,51	4,3	2,8	2.555,19
1,8	0,3	382.572,13	4,4	2,9	1.865,88
1,9	0,4	344.915,28	4,5	3,0	1.349,97
2,0	0,5	308.770,21	4,6	3,1	967,67
2,1	0,6	274.412,21	4,7	3,2	687,20
2,2	0,7	242.071,41	4,8	3,3	483,48
2,3	0,8	211.927,71	4,9	3,4	336,98
2,4	0,9	184.108,21	5,0	3,5	232,67
2,5	1,0	158.686,95	5,1	3,6	159,15
2,6	1,1	135.686,77	5,2	3,7	107,83
2,7	1,2	115.083,09	5,3	3,8	72,37
2,8	1,3	96.809,10	5,4	3,9	48,12
2,9	1,4	80.762,13	5,5	4,0	31,69
3,0	1,5	66.810,63	5,6	4,1	20,67
3,1	1,6	54.801,40	5,7	4,2	13,35
3,2	1,7	44.566,73	5,8	4,3	8,55
3,3	1,8	35.931,06	5,9	4,4	5,42
3,4	1,9	28.716,97	6,0	4,5	3,40
3,5	2,0	22.750,35			

Fonte: Adaptado de Breyfogle (1999)

ANEXO D

**Publicação XII SIMPEP – Bauru, SP, Brasil, Novembro de 2005
REVISTA GEPROS – Gestão da Produção, Operações e Sistemas**

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

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

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

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

Resumo

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

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

Abstract

This paper presents a critical analysis regarding the use of specific “quality tools” and some methods applied to Production Engineering in cases in which the collected data are not normally distributed. The general problem refers to the occurrence of questionable decision takings starting from the interpretation of data considered normally distributed. The primordial justification for this study is the relevance of the data analysis during its collection and handling in scientific works in the most varied levels. The main objectives of this work are: to present some examples related to the Statistical Process Control and the Six Sigma Methodology that may induce to doubtful conclusions due to the incorrect assumption of normality; to propose critical review proceeding to the Box-Cox and/or Johnson transformations; and, finally, to discuss the conclusions and decisions established through the comparison among the data originally collected and the transformed data.

Key Words: Statistical methods, Non-normal data, Data transformation.

1. Introdução

A proposta deste artigo em analisar a aplicação de alguns métodos estatísticos está inserida em um contexto mais amplo que se refere ao Controle Estatístico da Qualidade e aos aspectos básicos que devem ser observados pelas organizações que buscam a plena satisfação de seus clientes. A Estatística desempenha papel importante nos programas de Controle da Qualidade, pois ao longo dos anos suas técnicas e metodologias tornaram-se cada vez mais amplamente utilizadas e aceitas nas organizações (1).

O objetivo primário do Controle Estatístico da Qualidade é a redução sistemática da variabilidade nas características-chave do produto (2), fornecendo as ferramentas necessárias para avaliação e melhoria de processos, produtos e serviços de forma robusta e abrangente. Entretanto, tais ferramentas também dependem da qualidade da informação e da forma como os dados são coletados, processados e interpretados. Por este motivo é de fundamental importância o rigor metodológico na coleta e tratamento dos dados.

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

2. Resumo dos métodos em análise

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

Define-se processo como um conjunto de atividades inter-relacionadas ou interativas que transformam insumos (entradas) em produtos (saídas) (3). A Figura 1 exemplifica a situação esquemática de um processo com suas entradas, seus fatores controláveis, seus fatores incontroláveis (ruídos) e suas saídas.

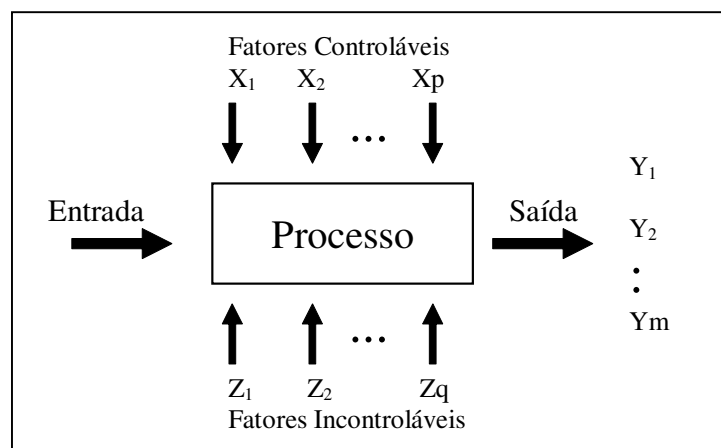


Figura 1 – Visão esquemática de um processo

Fonte: Adaptado de Montgomery & Runger (2003)

Controle Estatístico de Processo (CEP) pode ser definido como um conjunto de ferramentas que tem o propósito de indicar se um processo está funcionando de forma ideal, quando apenas causas comuns de variação estão presentes, ou se este processo está desordenado, e necessita de algum tipo de ação corretiva, ou seja, quando existem causas especiais de variação. O gráfico de controle é a mais poderosa ferramenta do CEP, pois evidencia a variação do processo em relação aos limites de controle estipulados (4).

Algumas suposições precisam ser satisfeitas para que, independentemente do tipo do gráfico de controle utilizado, os resultados sejam válidos (5):

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

2.2. Avaliação de desempenho na Metodologia Seis Sigma

Seis Sigma é uma metodologia de melhoria em processos com o objetivo de reduzir os defeitos a uma taxa de 3,4 partes por milhão nas características críticas de qualidade para os clientes (6). Também é definido como um método sistemático para melhoria de processos estratégicos, desenvolvimento de novos produtos e serviços que se baseia em métodos estatísticos e científicos para obtenção de uma drástica redução nas taxas de defeitos (7).

A Metodologia Seis Sigma é baseada em um sistema de acompanhamento conhecido como DMAIC, sigla que denota as seguintes etapas: Definir (*Define*), Medir (*Measure*), Analisar (*Analyze*), Melhorar (*Improve*) e Controlar (*Control*). As empresas que seguiram disciplinadamente estas etapas obtiveram significativa vantagem competitiva (8).

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

Tabela 1 – Referências para o nível sigma

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

Fonte: Harry (1998)

Capacidade é definida como a aptidão de uma organização, sistema ou processo de realizar um produto que irá atender aos requisitos especificados para este produto (3). Em termos estatísticos, a capacidade e o desempenho de um processo são representados por meio das equações 1, 2, 3 e 4 (9):

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

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

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

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

onde:

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

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

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

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

LSE = limite superior de especificação

LIE = limite inferior de especificação

μ = média do processo

S = desvio padrão (amostra)

σ = estimativa do desvio-padrão do processo (população)

Um processo com $C_{PK} = 2$ é referido como um processo seis sigma, pois a distância a partir da média do processo até a especificação mais próxima é de seis desvios-padrão. Se a média do processo se deslocar do centro por 1,5 desvio-padrão, o C_{PK} diminuirá para $4,5\sigma/3\sigma = 1,5$. Considerando um processo distribuído normalmente, a fração não conforme do processo deslocado é de 3,4 partes por milhão. Assim sendo, a média de um processo seis sigma pode se deslocar de 1,5 desvio-padrão do centro das especificações e ainda manter uma fração não conforme de apenas 3,4 partes por milhão (4). Boas aproximações podem ser obtidas para determinação do nível sigma através das equações 5 e 6 (9):

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

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

3. Análise da distribuição e transformação de dados

3.1. Considerações gerais

Os métodos estatísticos discutidos neste artigo pressupõem que os dados em estudo sigam uma distribuição de probabilidade conhecida. A análise e as conclusões que resultam da aplicação da metodologia são válidas apenas nos casos onde a suposição da distribuição se confirme correta. Independentemente de quão precisos sejam os métodos estatísticos utilizados, na realidade os dados é que são os elementos importantes (1).

Na falta de alguma evidência em contrário, pode-se assumir, em uma primeira abordagem, que os dados utilizados em métodos estatísticos sejam normalmente distribuídos (10). Contudo, em muitas situações práticas existem razões para se duvidar da validade da suposição de normalidade, o que implica em especial atenção na análise dos dados (2).

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

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

Quadro 1 – Conseqüências da suposição incorreta de normalidade

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

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

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

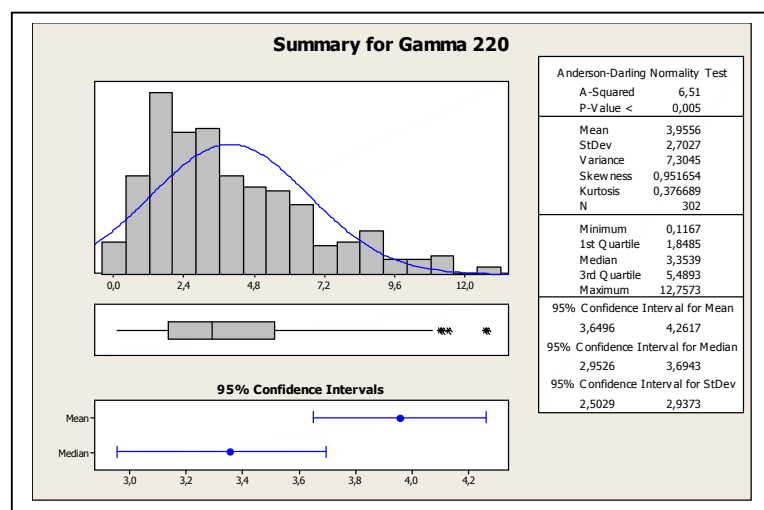


Figura 2 – Exemplo de análise gráfica de uma distribuição

A seguir são apresentados alguns critérios de análise para abordagem dos dados em caso de não aderência à distribuição normal (11):

- a) Examinar os dados para verificar se há alguma explicação não estatística para o padrão distribucional não convencional;
- b) Analisar os dados em termos de médias ao invés de valores individuais, pois médias de amostra seguem de perto uma distribuição de probabilidade normal, mesmo se a população de valores individuais não é distribuída normalmente (Teorema do Limite Central);
- c) Utilizar como referência outro tipo de distribuição que se enquadre mais adequadamente ao conjunto de dados coletados;
- d) Efetuar transformação matemática da característica original para uma nova característica que se aproxime de uma distribuição normal.

3.2. Transformação de Box-Cox

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

Um estudo detalhado na análise de dados representados pelas observações x_1, x_2, \dots, x_n consideradas normalmente distribuídas, com variância constante e valores esperados especificados por modelos lineares (12), levou à seguinte família de transformação exponencial da variável x para $x^{(\lambda)}$:

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

Esta transformação é definida somente para variáveis com valores positivos ($x > 0$) e o parâmetro λ , possivelmente um vetor, é o elemento que define a transformação específica e que, com frequência, resulta em normalidade. Como a análise de variância não é afetada por uma transformação linear, na prática a equação 7 pode ser simplificada para a seguinte forma (12):

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

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

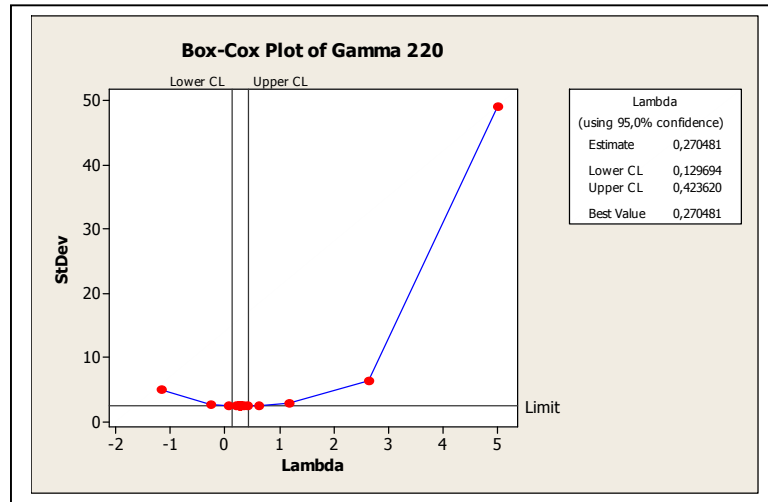


Figura 3 – Transformação de Box-Cox

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

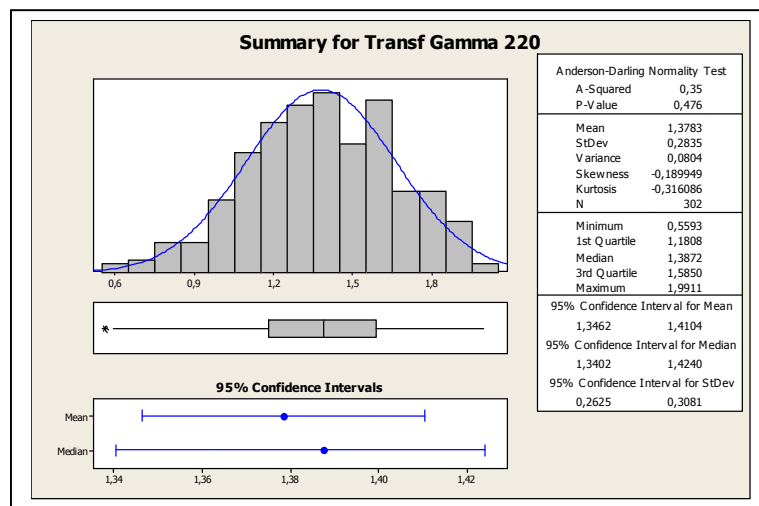


Figura 4 – Análise dos dados transformados pelo modelo Box-Cox

Após a adequada transformação da variável x para $x^{(\lambda)}$, pode-se presumir que os valores esperados das observações transformadas apresentem as seguintes características (12):

- a) Sejam descritos por um modelo de estrutura simples;

- b) A variância seja constante;
- c) As observações estejam normalmente distribuídas.

3.3. Transformação de Johnson

A transformação de Box-Cox foi um passo importante na determinação de uma maneira objetiva de se efetuar transformação de dados (13). Entretanto, como a transformação ela é válida apenas para valores positivos de x , havia espaço para algum tipo de melhoria. Embora seja possível efetuar uma troca de parâmetros em caso de valores negativos para utilização da transformação de Box-Cox, existe o inconveniente de tal ação afetar a teoria que suporta a definição do intervalo de confiança de λ .

Uma outra família de transformação de dados, válida tanto para valores positivos como para valores negativos da variável x foi desenvolvida posteriormente. Sua fórmula, definida como uma função $\psi: \mathbb{R} \times \mathbb{R} \rightarrow \mathbb{R}$, é apresentada a seguir (13):

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

A transformação de Johnson definida no *software* estatístico Minitab® é baseada na escolha de uma entre três famílias de distribuição (S_B , S_L ou S_U). Para o conjunto de dados apresentados na Figura 2 também foi efetuada aplicação da transformação de Johnson, resultando em uma distribuição normal, visto que $p\text{-value} = 0,950$ equivale a uma elevada probabilidade de cometer erro ao rejeitar a hipótese nula de normalidade da distribuição obtida. As características da transformação obtida no Minitab®, com destaque para determinação da equação de transformação das variáveis, são apresentadas na Figura 5.

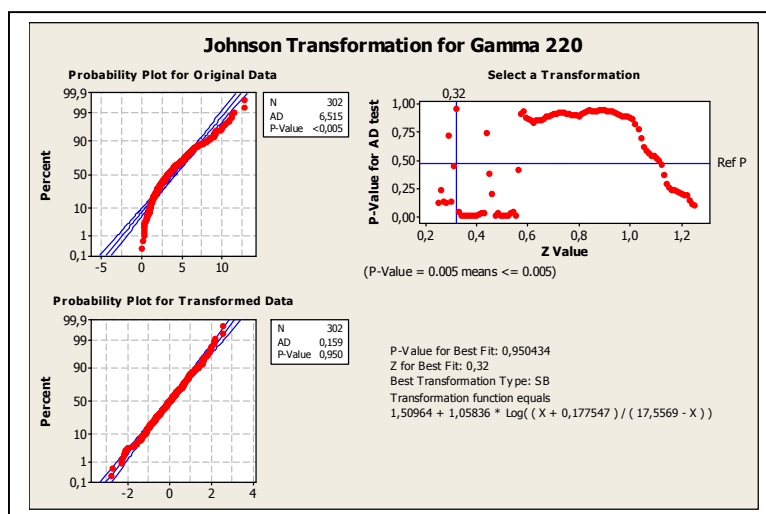


Figura 5 – Análise dos dados transformados pelo método de Johnson

4. Exemplos de análise dos métodos estatísticos em questão

4.1. Aspectos metodológicos

O ponto de partida de uma pesquisa é a existência de um problema que se deverá definir, examinar, avaliar e analisar criticamente para, em seguida, ser proposta sua solução (14). No presente artigo é apresentado o problema de ocorrência de decisões questionáveis na aplicação de métodos estatísticos devido à interpretação inadequada dos dados coletados. Este é um problema para o qual já se obteve solução a partir de outras pesquisas, algumas delas mencionadas neste trabalho; entretanto, o que realmente motiva o tema é a necessidade de promover uma maior divulgação das técnicas de transformação de variáveis para destacar a relevância do rigor metodológico durante coleta e tratamento de dados estatísticos.

Utilizando alguns critérios de classificação como referência (15), quanto aos objetivos este trabalho caracteriza-se como uma pesquisa exploratória, pois deverá proporcionar maior familiaridade com as técnicas de transformação de variáveis durante coleta e tratamento de dados em trabalhos científicos, além de possibilitar análise crítica através da comparação entre os dados originalmente coletados e os dados transformados.

A ciência proporciona a conceptualização da realidade e os conceitos com que ela opera são chamados de construtos (16). Os construtos são adotados ou inventados conscientemente com um significado específico. Conceito e construtos são semelhantes, sendo que a diferença reside no fato de que o construto possui um significado construído intencionalmente a partir de um marco teórico, devendo ser delimitado, traduzido em proposições particulares observáveis e mensuráveis (16). O objetivo do construto é fazer com que não haja ambigüidade no referencial empírico dos conceitos utilizados pelos pesquisadores.

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

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

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

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

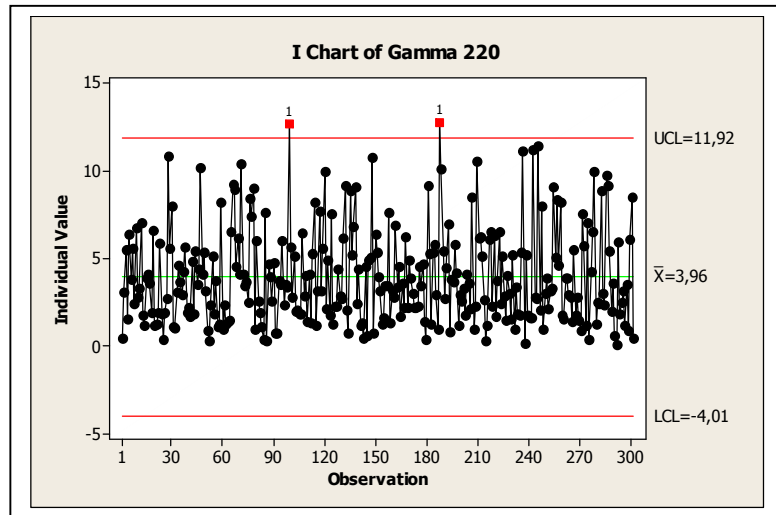


Figura 6 – Gráfico para “Gamma 220” considerando distribuição normal

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

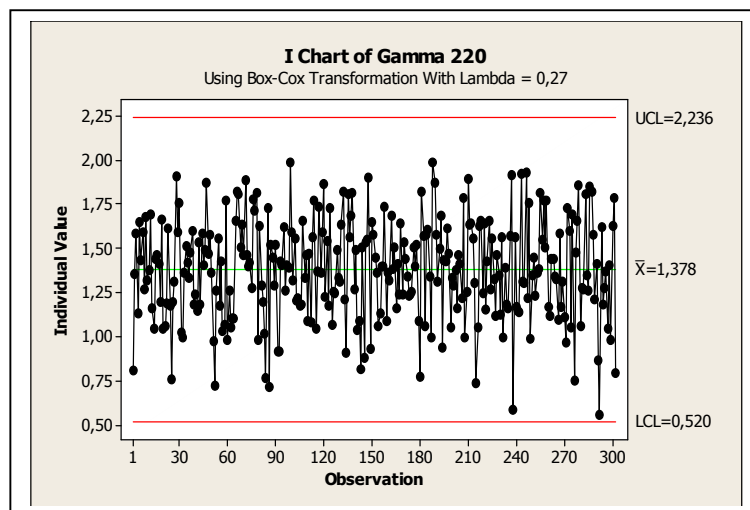


Figura 7 – Gráfico para “Gamma 220” com transformação de Box-Cox

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

4.3. Cálculo incorreto do nível sigma

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

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

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

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

5. Considerações finais

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

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

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

oportunidades (com a transformação de Johnson). Tais quantidades são muito significativas em termos de avaliação de capacidade e desempenho.

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

Referências

- (1) FEIGENBAUM, A.V. – Controle da Qualidade Total, vol. 3, Métodos Estatísticos Aplicados à Qualidade. 3ª ed. São Paulo: Makron Books, 1994.
- (2) MONTGOMERY, D.C. – Introduction to Statistical Quality Control. New York: John Wiley & Sons, Inc, 1985.
- (3) ASSOCIAÇÃO BRASILEIRA DE NORMAS TÉCNICAS – NBR ISO 9000: Sistema de Gestão da Qualidade – Fundamentos e Vocabulário, Rio de Janeiro, 2000.
- (4) MONTGOMERY, D.C. e RUNGER, G.C – Estatística Aplicada e Probabilidade para Engenheiros. 2ª ed. Rio de Janeiro: LTC – Livros Técnicos e Científicos Editora S.A., 2003.
- (5) TOLEDO, T.P.A. – Uma Investigação sobre Índices de Capacidade com Ênfase na Metodologia Seis Sigma. Dissertação de Mestrado. Universidade Federal de Itajubá, 2005.
- (6) HARRY, M.J. – Six Sigma: A Breakthrough Strategy for Profitability. New York: Quality Progress, v.31, n.5, p.60-64, May 1998.
- (7) LINDERMAN, K. et al. – Six Sigma: A Goal-Theoretic Perspective. Journal of Operations Management, v.21, n.2, p.193-203, 2003.
- (8) BREYFOGLE, F. – Six Sigma Methods to Ensure Organization's Health. . New York: Quality Progress, v.45, n.4, p.28-29, April 2003.
- (9) BALESTRASSI, P.P. – Metodologia Seis Sigma. Notas de Aula do Curso de Pós-Graduação em Engenharia de Produção. Universidade Federal de Itajubá, 2004.
- (10) LIPSON, C. and SHETH, N.J. – Statistical and Analysis of Engineering Experiments. New York: McGraw-Hill Book Company, 1973.
- (11) JURAN, J.M. e GRZYNA, F.M. – Controle da Qualidade Handbook, vol. 6, Métodos Estatísticos Clássicos Aplicados à Qualidade. 4ª ed. São Paulo: Makron Books, 1992.
- (12) BOX, G.E.P. and COX, D.R. – An Analysis of Transformations. Journal of Royal Statistical Society. B, 39, p.211-252, 1964.
- (13) YEO, I.K. and JOHNSON, R.A. – A New Family of Power Transformation to Improve Normality or Symmetry. Biometrika, 87, p.954-959, 2000.
- (14) ASTI VERA, A. – Metodologia da Pesquisa Científica. Porto Alegre: Globo, 1973.
- (15) GIL, A.C. Como Elaborar Projetos de Pesquisa. 3ª ed. São Paulo: Atlas, 1991.
- (16) KÖCHE, J.C. – Fundamentos de Metodologia Científica. 7ª ed. Caxias do Sul: Editora Vozes, 1984.