

SAULO HENRIQUE WEBER

**Desenvolvimento de uma Função Densidade de Probabilidade
Multimodal Aplicável à Ciência Florestal**

Tese apresentada como requisito parcial para a obtenção do título de Doutor em Engenharia Florestal no Programa de Pós-Graduação em Engenharia Florestal da Universidade Federal do Paraná.

Orientador: Prof. Dr. Sylvio Péllico Netto
Co-orientador: Prof. Dr. Julio Eduardo Arce

CURITIBA
2011

DEDICATÓRIA

Dedico esse trabalho à Deus,
à minha esposa Ariadne
e à minha família.

AGRADECIMENTOS

Agradeço à Deus, por ter me dado forças e coragem para enfrentar esse caminho que por muitas vezes foi árduo. Ele me ajudou a permanecer perseverando, mesmo nos momentos em que o corpo fraquejava.

Agradeço à Universidade Federal do Paraná - UFPR, em especial à Pós-Graduação em Engenharia Florestal, pela estrutura física como computadores, biblioteca, auditório e laboratórios. Mas agradeço a essa instituição de ensino superior principalmente por ter disponibilizado um corpo docente tão preparado e comprometido com o processo de ensino/aprendizagem. As disciplinas cursadas foram essenciais para o embasamento e desenvolvimento do presente trabalho.

Sou grato ao Conselho Nacional de Desenvolvimento Científico e Tecnológico - CNPq, pela bolsa de estudos a mim concedida. Sem ela não seria possível a realização do presente trabalho. Saliento a importância dessa modalidade de auxílio para o desenvolvimento tecnológico no Brasil.

Ao Professor Doutor Julio Eduardo Arce, meus agradecimentos. Foi de valia incomensurável todo apoio e orientação desde o mestrado até o doutorado. É um professor dinâmico e objetivo, que me ajudou decisivamente para o desenvolvimento e conclusão dessa pesquisa.

Ao Professor Doutor Sylvio Péllico Netto meus sinceros agradecimentos. Trabalhar com ele durante minha graduação foi decisivo para a mudança no rumo de minha carreira. Sua orientação nos projetos de pesquisa, as idéias para elaboração de artigos, o trabalho na Revista Acadêmica - Ciências Agrárias e Ambientais, tudo isso ajudou na minha formação profissional e pessoal. Contribuiu para a melhoria da minha escrita. Auxiliou com seus conhecimentos nas áreas de estatística e engenharia florestal, mas principalmente pela sua vasta experiência.

À minha amada esposa Ariadne Josiane Castoldi Silva Weber, meu obrigado. Obrigado por ter me aguentado nos momentos de stress, de tristeza, de insegurança, enfim, de dificuldades. Mas agradeço principalmente por ter estado ao meu lado quando os resultados surgiam, nos momentos de alegria, de felicidade, por me apoiar incondicionalmente, por estar sempre contente, motivada, confiante e otimista. Ela me dá sorte. Foi durante conversas com ela pela internet que alguns dos principais resultados positivos foram alcançados.

Agradeço ao Reinaldo Pinheiro, secretário do curso de Pós-Graduação em Engenharia Florestal da Universidade Federal do Paraná, por todas as orientações sobre os trâmites legais

e processuais dentro da instituição. Agradeço também pelo bom humor e pelo pronto atendimento às minhas dúvidas.

Manifesto aqui minha gratidão ao Professor Pós-Doutor Carlos Roberto Sanquetta, pelos dados sobre a Araucária, coletados no município de São João do Triunfo, bem como ao Professor Pós-Doutor Afonso Figueiredo Filho, pelos dados da Araucária coletados na Flona de Irati. Agradeço ainda ao Professor Doutor Sylvio Péllico Netto por disponibilizar a distribuição diamétrica da castanheira. As informações foram essenciais para o desenvolvimento do modelo aqui apresentado.

Quero agradecer à minha família, por ter me apoiado nessa caminhada. Um agradecimento especial faço à minha mãe, sempre presente e compreensiva. Agradeço à minha avó, meu pai, sua esposa, meus irmãos, minha oma, meu opa, meus tios e tias, meus primos e primas.

Agradeço ainda ao meu avô, Carlos Osmar Lenz (*in memoriam*), pela minha educação na adolescência, pelos inúmeros momentos felizes contigo compartilhados, pelo amor que sentias, pela retidão nas atitudes e por tantos outros motivos que tenho para agradecer-te.

BIOGRAFIA

Nascido em Curitiba, no estado do Paraná, no sétimo dia do mês de janeiro, no ano de 1982, Saulo Henrique Weber é filho de Luis Carlos Behlau Weber e Lílian Regina Lenz. Logo cedo, ainda nas primeiras semanas de vida, mudou-se para Itapoá, Santa Catarina, junto com seus pais, onde morou até o final de seus 12 anos. Lá cursou os seis primeiros anos do ensino fundamental, no Colégio Estadual Nereu Ramos.

Retornou à Curitiba em 1994. No ano seguinte continuou seus estudos no Colégio Estadual Santo Agostinho, até o início de 1997, quando se mudou para o Capão Raso, bairro da mesma cidade. Lá cursou os últimos anos do ensino médio, no Colégio Estadual Doutor Francisco de Azevedo Macedo.

Prestou vestibular de inverno para o curso de Licenciatura em Matemática na Pontifícia Universidade Católica do Paraná no ano 2000, sendo aprovado. Graduou-se quatro anos mais tarde, em 2004. Durante esse período estagiou na área de informática na Procuradoria do Trabalho. Posteriormente foi estagiário na Qualogic, empresa fornecedora de material pedagógico para a rede municipal de ensino de Curitiba, bem como realizadora de treinamento para seu uso.

Do quarto ao oitavo semestre de graduação foi monitor de cálculo diferencial e integral no curso de Engenharia Agrônômica da Pontifícia Universidade Católica do Paraná, no município de São José dos Pinhais, sob a orientação do Professor Eduardo Quadros da Silva. Nesse período foi bolsista do Projeto Sassafrás – Bioecologia e Uso Sustentável durante dois anos. Trabalhou como bolsista também no Programa Ecológico de Longa Duração – PELD, financiado pelo CNPq.

Depois de formado atuou como bolsista no Projeto Robalo, desenvolvido no município de Guaratuba, Paraná, e coordenado pelo Professor Doutor Sylvio Péllico Netto. Esse projeto visava o repovoamento da baía de Guaratuba com alevinos da espécie.

No mesmo ano de sua formatura foi aceito no Programa de Pós-Graduação em Engenharia Florestal da Universidade Federal do Paraná. Realizou seu mestrado no período de um ano e oito meses, tendo defendido sua dissertação sobre funções densidade de probabilidade em novembro de 2006, quando ingressou no curso de doutorado da mesma instituição.

Atualmente é professor da Universidade do Centro-Oeste – Unicentro, desde 2009. É também professor da Pontifícia Universidade Católica do Paraná desde o ano de 2010.

SUMÁRIO

1. INTRODUÇÃO.....	1
1.1 JUSTIFICATIVA.....	3
1.2 OBJETIVOS.....	5
1.2.1 Objetivo Geral.....	5
1.2.2 Objetivos Específicos.....	5
2. REVISÃO DE LITERATURA.....	6
2.1 MATEMÁTICA.....	6
2.1.1 Funções.....	7
2.1.2 Limites e continuidade.....	7
2.1.3 Derivadas.....	8
2.1.4 Integrais.....	8
2.1.5 Teorema Fundamental do Cálculo.....	9
2.2 MODELAGEM.....	9
2.3 ESTATÍSTICA.....	11
2.4 PRINCIPAIS DISTRIBUIÇÕES DISCRETAS.....	13
2.4.1 Bernoulli.....	13
2.4.2 Binomial.....	14
2.4.3 Poisson.....	15
2.4.4 Multinomial.....	15
2.4.5 Hipergeométrica.....	16
2.4.6 Distribuição geométrica.....	16
2.5 PRINCIPAIS DISTRIBUIÇÕES CONTÍNUAS.....	16
2.5.1 Distribuição Exponencial.....	20
2.5.2 Distribuição Gama.....	22
2.5.3 Distribuição Beta.....	23
2.5.4 Distribuição Weibull.....	26
2.5.5 Distribuição Normal ou de Gauss.....	28
2.5.6 Distribuição de Quadros.....	30
2.5.7 Funções Spline.....	31
2.6 BIOLOGIA.....	31
2.7 <i>Bertholletia excelsa</i>	31
2.8 <i>Araucaria angustifolia</i>	33
3. MATERIAIS E MÉTODOS.....	35

3.1 MÉDIA E VARIÂNCIA DE UMA FUNÇÃO DENSIDADE DE PROBABILIDADE	35
3.2 EXPECTATIVA MATEMÁTICA	35
3.3 DISTRIBUIÇÃO EXPONENCIAL	35
3.4 DISTRIBUIÇÃO GAMA	36
3.5 DISTRIBUIÇÃO BETA	38
3.6 DISTRIBUIÇÃO WEIBULL	39
3.7 DISTRIBUIÇÃO NORMAL	39
3.8 DISTRIBUIÇÃO DE QUADROS	40
3.9 FUNÇÕES SPLINE	41
3.10 MODELO MULTIMODAL PROPOSTO	41
3.11 AVALIAÇÃO DO AJUSTE	42
3.11.1 Erro padrão da estimativa recalculado	42
3.11.2 Kolmogorov e Smirnov	42
4. DESENVOLVIMENTO DO MODELO	44
4.1 INTRODUÇÃO	44
4.2 ETAPAS DO DESENVOLVIMENTO	46
4.3 CONJUNTO DE DADOS	46
4.4 A FUNÇÃO ORIGINAL	46
4.4.1 Teste 1	46
4.4.2 Teste 2	47
4.4.3 Teste 3	48
4.4.4 Teste 4	49
4.4.5 Teste 5	50
4.4.6 Teste 6	51
4.4.7 Teste 7	51
4.4.8 Teste 8	52
4.4.9 Teste 9	53
4.4.10 Teste 10	54
5. RESULTADOS E DISCUSSÕES	57
5.1 INTRODUÇÃO	57
5.2 <i>Bertholletia excelsa</i> Bonpl	58
5.2.1 Estatística descritiva	60
5.2.2 Normal	60

5.2.3 Exponencial	62
5.2.4 Beta.....	64
5.2.5 Gama.....	66
5.2.6 Weibull	68
5.2.7 Quadros.....	70
5.2.8 Spline Cúbica.....	73
5.2.9 Modelo Proposto	75
5.2.10 Resumo dos resultados dos modelos	78
5.3 <i>Araucaria angustifolia</i>	79
5.3.1 Estatística descritiva	81
5.3.2 Normal	81
5.3.3 Exponencial	83
5.3.4 Beta.....	85
5.3.5 Gama.....	87
5.3.6 Weibull	89
5.3.7 Quadros.....	91
5.3.8 Spline Cúbica.....	94
5.3.9 Modelo Proposto	96
5.3.10 Resumo dos resultados dos modelos	99
5.4 <i>Araucaria angustifolia</i>	101
5.4.1 Estatística descritiva	102
5.4.2 Normal	103
5.4.3 Exponencial	104
5.4.4 Beta.....	106
5.4.5 Gama.....	108
5.4.6 Weibull	110
5.4.7 Quadros.....	112
5.4.8 Spline Cúbica.....	115
5.4.9 Modelo Proposto	117
5.5 COMPARAÇÃO ENTRE OS MODELOS (GERAL).....	121
6. CONCLUSÕES	124
7. REFERÊNCIAS	126

LISTA DE FIGURAS

FIGURA 01 – Representação da área sob a curva	13
FIGURA 02 – DISTRIBUIÇÃO EXPONENCIAL, QUANDO $\beta = 1$	21
FIGURA 03 - REPRESENTAÇÃO GRÁFICA DA DISTRIBUIÇÃO GAMMA, PARA $\alpha=1$ (SÓLIDO), $\alpha=1,5$ (PONTOS), $\alpha=2$ (TRAÇOS) E $\alpha=3$ (PONTOS E TRAÇOS) E $\beta=0,5$	23
FIGURA 04 – REPRESENTAÇÃO GRÁFICA DA DISTRIBUIÇÃO BETA, PARA $\alpha=1$ E $\beta = 2$ (SÓLIDO), $\alpha=2$ E $\beta = 1$ (PONTOS) E $\alpha=1$ E $\beta =$ 1 (TRAÇOS).....	24
FIGURA 05 – REPRESENTAÇÃO GRÁFICA DA DISTRIBUIÇÃO BETA, PARA $\alpha =0,3$ E $\beta = 0,8$ (SÓLIDO) E $\alpha =0,3$ E $\beta = 0,3$ (PONTOS).....	25
FIGURA 06 – REPRESENTAÇÃO GRÁFICA DA DISTRIBUIÇÃO BETA, PARA $\alpha =2$ E $\beta = 3$ (SÓLIDO), $\alpha =5$ E $\beta = 3$ (PONTOS) E $\alpha =2$ E $\beta =$ 2 (TRAÇOS).....	25
FIGURA 07 – REPRESENTAÇÃO GRÁFICA DA DISTRIBUIÇÃO BETA, PARA $\alpha = 0,5$ E $\beta = 1,5$ (SÓLIDO) E $\alpha = 3$ E $\beta = 0,5$ (PONTOS).....	26
FIGURA 08 - REPRESENTAÇÃO GRÁFICA DA DISTRIBUIÇÃO WEIBULL, COM $a = 1$, $b = 10$ e $c = 0,7$ (SÓLIDO), COM $a = 1$, $b = 10$ e $c = 1$ (PONTOS), COM $a = 1$, $b = 10$ e $c = 2$ (TRAÇOS), COM $a = 1$, $b = 10$ e c $= 3,6$ (PONTOS E TRAÇOS)	27
FIGURA 09 – DISTRIBUIÇÃO NORMAL, $\sigma = 1$ E $\mu = 4$	28
FIGURA 10 – Resultado do primeiro teste	47
FIGURA 11 – Resultado do segundo teste.....	48
FIGURA 12 - Resultado do terceiro teste	49
FIGURA 13 – Resultado do quarto teste.....	49
FIGURA 14 – Resultado do quinto teste.....	50
FIGURA 15 – Resultado do sexto teste.....	51
FIGURA 16 – Resultado do sétimo teste	52
FIGURA 17 – Resultado do oitavo teste	52
FIGURA 18 – Resultado do nono teste	53
FIGURA 19 – Resultado do décimo teste	55
FIGURA 20 – Estrutura populacional da espécie <i>Bertholletia excelsa</i> Bonpl.....	60
FIGURA 21 – Distribuição diamétrica da castanheira e curva Normal a ela ajustada.	62
FIGURA 22 – Resíduos em percentagem da distribuição Normal	62
FIGURA 23 – Distribuição diamétrica da castanheira e distribuição Exponencial a ela ajustada.	64

FIGURA 24 – Resíduos em percentagem da distribuição Normal	64
FIGURA 25 – Distribuição diamétrica da castanheira e distribuição Beta a ela ajustada.	66
FIGURA 26 – Resíduos em percentagem da distribuição Normal	66
FIGURA 27 – Distribuição diamétrica da castanheira e distribuição Gama a ela ajustada.	68
FIGURA 28 – Resíduos em percentagem da distribuição Normal	68
FIGURA 29 – Distribuição diamétrica da castanheira e distribuição Weibull a ela ajustada.	70
FIGURA 30 – Resíduos em percentagem da distribuição Normal	70
FIGURA 31 – Distribuição diamétrica da castanheira e distribuição de Quadros a ela ajustada.	72
FIGURA 32 – Resíduos em percentagem da distribuição Normal	72
FIGURA 33 – Distribuição diamétrica da castanheira e função Spline Cúbica a ela ajustada.	75
FIGURA 34 – Resíduos em percentagem da distribuição Normal	75
FIGURA 35 – Distribuição diamétrica da Castanheira e modelo proposto a ela ajustada.	77
FIGURA 36 – Resíduos em percentagem modelo proposto	77
FIGURA 37 – Estrutura populacional da Araucária em uma área de 3,5 ha em São João do triunfo	81
FIGURA 38 – Distribuição diamétrica da Araucária e curva Normal a ela ajustada.	83
FIGURA 39 – Resíduos em percentagem da distribuição Normal	83
FIGURA 40 – Distribuição diamétrica da Araucária e distribuição Exponencial a ela ajustada.	85
FIGURA 41 – Resíduos em percentagem da distribuição Exponencial.....	85
FIGURA 42 – Distribuição diamétrica da Araucária e distribuição Beta a ela ajustada.	87
FIGURA 43 – Resíduos em percentagem da distribuição Beta	87
FIGURA 44 – Distribuição diamétrica da Araucária e distribuição Gama a ela ajustada.	89
FIGURA 45 – Resíduos em percentagem da distribuição Gama	89
FIGURA 46 – Distribuição diamétrica da Araucária e distribuição Weibull a ela ajustada.	91
FIGURA 47 – Resíduos em percentagem da distribuição Weibull.....	91
FIGURA 48 – Distribuição diamétrica da Araucária e distribuição de Quadros a ela ajustada.	93
FIGURA 49 – Resíduos em percentagem da distribuição de Quadros	93

FIGURA 50 – Distribuição diamétrica da Araucária e função Spline Cúbica a ela ajustada.	96
FIGURA 51 – Resíduos em percentagem da função Spline Cúbica	96
FIGURA 52 – Distribuição diamétrica da Araucária e modelo proposto a ela ajustada.	98
FIGURA 53 – Resíduos em percentagem modelo proposto	98
FIGURA 54 – Estrutura populacional da Araucária em uma área de 25 ha em um Fragmento de Floresta Ombrófila Mista em Irati	102
FIGURA 55 – Distribuição diamétrica da Araucária e curva Normal a ela ajustada.	104
FIGURA 56 – Resíduos em percentagem da distribuição Normal	104
FIGURA 57 – Distribuição diamétrica da Araucária e distribuição Exponencial a ela ajustada.	106
FIGURA 58 – Resíduos em percentagem da distribuição Exponencial.....	106
FIGURA 59 – Distribuição diamétrica da Araucária e distribuição Beta a ela ajustada.	108
FIGURA 60 – Resíduos em percentagem da distribuição Beta	108
FIGURA 61 – Distribuição diamétrica da Araucária e distribuição Gama a ela ajustada.	110
FIGURA 62 – Resíduos em percentagem da distribuição Gama	110
FIGURA 63 – Distribuição diamétrica da Araucária e distribuição Weibull a ela ajustada.	112
FIGURA 64 – Resíduos em percentagem da distribuição Weibull.....	112
FIGURA 65 – Distribuição diamétrica da Araucária e distribuição de Quadros a ela ajustada.	114
FIGURA 66 – Resíduos em percentagem da distribuição de Quadros	114
FIGURA 67 – Distribuição diamétrica da Araucária e função Spline Cúbica a ela ajustada.	116
FIGURA 68 – Resíduos em percentagem da função Spline Cúbica	117
FIGURA 69 – Distribuição diamétrica da Araucária e modelo proposto a ela ajustada.	119
FIGURA 70 – Resíduos em percentagem modelo proposto	119

LISTA DE TABELAS

TABELA 01 – Dados simulados utilizados para o desenvolvimento do modelo.	46
TABELA 02 – Frequências absolutas (f_{obs}) e acumuladas (F_{obs}) observadas por classe diamétrica da espécie <i>Bertholletia excelsa</i> Bonpl.	59
TABELA 03 – Distribuição diamétrica, frequências acumuladas esperadas pela distribuição Normal ajustada e diferença absoluta entre as distribuições acumuladas esperada e observada.	61
TABELA 04 – Distribuição diamétrica, frequências acumuladas esperadas pela distribuição Exponencial ajustada e diferença absoluta entre as distribuições acumuladas esperada e observada.	63
TABELA 05 – Distribuição diamétrica, frequências acumuladas esperadas pela distribuição Beta ajustada e diferença absoluta entre as distribuições acumuladas esperada e observada.	65
TABELA 06 – Distribuição diamétrica, frequências acumuladas esperadas pela distribuição Gama ajustada e diferença absoluta entre as distribuições acumuladas esperada e observada.	67
TABELA 07 – Distribuição diamétrica, frequências acumuladas esperadas pela distribuição Weibull ajustada e diferença absoluta entre as distribuições acumuladas esperada e observada.	69
TABELA 08 – Distribuição diamétrica, frequências acumuladas esperadas pela distribuição de Quadros ajustada e diferença absoluta entre as distribuições acumuladas esperada e observada.	71
TABELA 09 – Distribuição diamétrica, frequências acumuladas esperadas pela distribuição de Quadros ajustada e diferença absoluta entre as distribuições acumuladas esperada e observada.	74
TABELA 10 – Distribuição diamétrica, frequências acumuladas esperadas pelo modelo proposto ajustado e diferença absoluta entre as distribuições acumuladas esperada e observada.	78
TABELA 11 – Comparação entre os modelos para a castanheira.	79
TABELA 12 – Frequências absolutas (f_{obs}) e acumuladas (F_{obs}) observadas por classe diamétrica da Araucária em uma área de 3,5 ha em São João do triunfo	80
TABELA 13 – Distribuição diamétrica, frequências acumuladas esperadas pela distribuição Normal ajustada e diferença absoluta entre as distribuições acumuladas esperada e observada.	82
TABELA 14 – Distribuição diamétrica, frequências acumuladas esperadas pela distribuição Exponencial ajustada e diferença absoluta entre as distribuições acumuladas esperada e observada.	84
TABELA 15 – Distribuição diamétrica, frequências acumuladas esperadas pela distribuição Beta ajustada e diferença absoluta entre as distribuições acumuladas esperada e observada.	86

TABELA 16 – Distribuição diamétrica, frequências acumuladas esperadas pela distribuição Gama ajustada e diferença absoluta entre as distribuições acumuladas esperada e observada.	88
TABELA 17 – Distribuição diamétrica, frequências acumuladas esperadas pela distribuição Weibull ajustada e diferença absoluta entre as distribuições acumuladas esperada e observada.	90
TABELA 18 – Distribuição diamétrica, frequências acumuladas esperadas pela distribuição de Quadros ajustada e diferença absoluta entre as distribuições acumuladas esperada e observada.	92
TABELA 19 – Distribuição diamétrica, frequências acumuladas esperadas pela distribuição de Quadros ajustada e diferença absoluta entre as distribuições acumuladas esperada e observada.	95
TABELA 20 – Distribuição diamétrica, frequências acumuladas esperadas pelo modelo proposto ajustado e diferença absoluta entre as distribuições acumuladas esperada e observada.	99
TABELA 21 – Comparação entre os modelos para a Araucária.....	100
TABELA 22 – Frequências absolutas (f_{obs}) e acumuladas (F_{obs}) observadas por classe diamétrica da Araucária em uma área de 3,5 ha em São João do triunfo	102
TABELA 23 – Distribuição diamétrica, frequências acumuladas esperadas pela distribuição Normal ajustada e diferença absoluta entre as distribuições acumuladas esperada e observada.	103
TABELA 24 – Distribuição diamétrica, frequências acumuladas esperadas pela distribuição Exponencial ajustada e diferença absoluta entre as distribuições acumuladas esperada e observada.	105
TABELA 25 – Distribuição diamétrica, frequências acumuladas esperadas pela distribuição Beta ajustada e diferença absoluta entre as distribuições acumuladas esperada e observada.	107
TABELA 26 – Distribuição diamétrica, frequências acumuladas esperadas pela distribuição Gama ajustada e diferença absoluta entre as distribuições acumuladas esperada e observada.	109
TABELA 27 – Distribuição diamétrica, frequências acumuladas esperadas pela distribuição Weibull ajustada e diferença absoluta entre as distribuições acumuladas esperada e observada.	111
TABELA 28 – Distribuição diamétrica, frequências acumuladas esperadas pela distribuição de Quadros ajustada e diferença absoluta entre as distribuições acumuladas esperada e observada.	113
TABELA 29 – Distribuição diamétrica, frequências acumuladas esperadas pela distribuição de Quadros ajustada e diferença absoluta entre as distribuições acumuladas esperada e observada.	116
TABELA 30 – Distribuição diamétrica, frequências acumuladas esperadas pela distribuição de Quadros ajustada e diferença absoluta entre as distribuições acumuladas esperada e observada.	120
TABELA 31 – Comparação entre os modelos para a Araucária.....	121

RESUMO

A tendência da distribuição diamétrica é um excelente referencial para o estudo de florestas, sejam elas nativas ou plantadas. Uma maneira de se compreender esse comportamento é por meio de funções densidade de probabilidade. O objetivo da presente pesquisa foi propor um novo modelo matemático que fosse suficientemente flexível para representar mais de uma moda. A fim de avaliar a eficácia do modelo proposto, o ajuste desse foi comparado com o das distribuições probabilísticas contínuas clássicas (Gauss, Exponencial, Gama, Beta e Weibull), bem como a de Quadros e a função Spline. Três distribuições diamétricas distintas foram usadas: uma da Castanheira, em que foram observadas três modas e a distribuição passa pela origem; e duas da Araucária, as quais apresentavam duas modas, mas uma passava pela origem e a outra não. A Distribuição Exponencial assume apenas forma decrescente e é potencialmente viável quando se deseja representar conjuntos de dados diamétricos de regeneração natural, logo foi impróprio para a realidade da presente pesquisa. A distribuição Normal, frequentemente utilizada para representar distribuições diamétricas, não teve um resultado satisfatório, em função de sua baixa flexibilidade, não representando mais que uma moda. As distribuições de Weibull, Beta e Gama são mais flexíveis, mas não se adequam a conjuntos de dados com mais de uma moda. O modelo de Quadros é mais flexível, mas não se aderiu a todos os conjuntos de dados observados no presente trabalho. A função Spline apresenta uma alta flexibilidade, mas os resultados não foram satisfatórios, tanto na questão do ajuste, quanto na questão da sua interpretação biológica, uma vez que gera probabilidades negativas e não é capaz de prever a ocorrência de indivíduos nas classes diamétricas superiores. O modelo aqui proposto foi aderente às três distribuições diamétricas, sendo capaz de apresentar mais de uma moda (as quais são fáceis de serem implementadas), é flexível, pode ou não passar pela origem, cumpre todos os requisitos de função densidade de probabilidade (é positiva, contínua e integrável em todo o seu domínio), a média e variância podem ser obtidas por meio de integrais, de acordo com a teoria de distribuições probabilísticas.

Palavras-chave: Distribuição de Probabilidade, Moda, Floresta, *Araucaria angustifolia*, *Bertholletia excelsa*.

ABSTRACT

Development of a multimodal probability density function

The trend of DBH distribution is an excellent referential to the forest study, native or manmade forests. Density Probability Functions can be satisfactorily used for that purpose. Therefore, the main objective of this research was to present a new mathematical model, sufficiently flexible to represent two or more modes. As objective of evaluating the model effectiveness, its adjustment was compared to the classic continuous probabilistic distributions (Gauss, Exponential, Gamma, Beta and Weibull), as well as Quadros and Spline Function. Three distinct diametric distributions were used: one for Brazilian Nut tree, with three modes and passing through the origin, and two for Araucaria, with two modes, one passing through the origin and the other no. The Exponential Distribution assumes just a decrescent form and is potentially viable when represent natural regeneration diametric data is desirable; therefore it is not good to this research. The Gaussian Distribution, commonly used to represent diametric distributions, provided a non satisfactory result, because of its low flexibility, in which occur one mode. The Weibull, Gamma and Beta Distributions were more flexible, but can not represent two or more modes. The model developed by Quadros is more flexible, but it was not adherent to all data distributions observed in this work. The Spline Function presented a high flexibility, but the results were not satisfactory, as considering the adjustment, as considering the biological interpretation, since it generates negative probabilities and can not predict the occurrence of individuals in the higher diameter classes. The proposed model was adherent to all diametric distributions, being capable to present two or more modes (easily implemented); it is flexible; it can or can not pass through the origin; it has all characteristics of a density probabilistic function (it is always positive, continuous and can be integrated in its domain), the average and variance can be obtained by means of calculus, according to the probabilistic distribution theory.

Keywords: Probability distribution, Mode, Forest, *Araucaria angustifolia*, *Bertholletia excelsa*.

1. INTRODUÇÃO

A palavra, hoje amplamente conhecida, “estatística”, segundo Boyer (1996), surgiu da expressão em Latim *statisticum collegium* (palestra sobre os assuntos do Estado). Esse termo originou o verbete em língua italiana *statista*, que significa "homem de estado", ou político. Em alemão, *Statistik* designa a análise de dados sobre o Estado. O sentido adotado atualmente surgiu no início do século 19.

Com o passar dos anos, novos métodos estatísticos foram necessários para a realização de determinados experimentos biológicos. Essa parte inicial e fundamental da experimentação foi então desenvolvida por biólogos, muitas vezes estudiosos de conhecimentos matemáticos (MOOD; GRAYBILL, 1978). Surgiram, ao longo do tempo, diversas teorias nas mais variadas áreas do conhecimento e é importante lembrar que a ciência, como ela é conhecida atualmente, deve-se substancialmente à interação entre os pesquisadores dessas áreas.

A definição de estatística mais habitual atualmente diz que essa é uma área do conhecimento que utiliza teorias probabilísticas para explicação de eventos, estudos e experimentos. Essa teoria tem por objetivo obter, organizar e analisar dados, determinar suas correlações, tirar delas suas consequências para descrição, explicar o que passou e fazer previsões (WEBER, 2006).

Segundo Johnson e Leone (1964), pode-se dividir a estatística em três partes, a saber:

- Estatística descritiva: em que o objetivo é, em geral, a descrição dos dados coletados por meio de sua sumarização, transformando-os em informação;
- Estudo das probabilidades: situações em que envolvam acaso. Em outras palavras, são casos em que se deseja determinar a chance ou a possibilidade de ocorrência de um determinado evento; e
- Inferência estatística: é a análise e interpretação de dados amostrais.

De acordo com Eues (1995), o passo decisivo para fundamentação teórica da inferência estatística, associa-se ao desenvolvimento do cálculo das probabilidades. Nessa publicação ele cita que as probabilidades são utilizadas para exprimir a chance de ocorrência de determinado evento.

Pode-se definir amostragem como ato de obter uma amostra de uma população da qual se deseja fazer inferência. O levantamento por amostragem, quando comparado com o censo, apresenta custos menores, resultados em menor tempo e objetivos mais amplos. Uma amostragem probabilística ocorre quando cada elemento da população, da qual é escolhida a amostra, tem uma probabilidade conhecida e diferente de zero de pertencer a ela.

Com objetivo de ampliar os estudos e aprofundar os resultados, foram desenvolvidas distribuições de frequência, que são agrupamentos de dados em classes. Uma distribuição de frequências relativas, para os resultados de um espaço amostral, constitui uma distribuição de probabilidade. Essas distribuições podem ser contínuas ou descontínuas, de acordo com os tipos de variáveis em questão, discretas ou contínuas. Enquanto uma variável aleatória discreta pode ser escolhida de possíveis valores finitos ou infinitos numeráveis, as variáveis aleatórias contínuas apresentam um conjunto de possibilidades incontáveis ou não numeráveis.

Existe um interesse especial nas distribuições contínuas, chamada Funções Densidade de Probabilidade (FDP). Elas surgiram em virtude da ocorrência de variáveis contínuas como tempo e distância, entre outras. Salienta-se que seu desenvolvimento e aplicação estão intimamente ligados a duas teorias: ajuste de funções e cálculo infinitesimal. É fácil compreender que, dado um conjunto de dados e um modelo probabilístico contínuo, há a necessidade de se conhecer um método que estime os seus coeficientes, o qual é chamado de regressão. O outro ponto, desenvolvimento do cálculo infinitesimal está vinculado ao cálculo das probabilidades, uma vez que a área entre o eixo das abscissas e o gráfico da distribuição probabilística contínua representa a probabilidade de ocorrência de um determinado evento. Essa área pode ser calculada por meio do cálculo integral.

Algumas FDPs surgiram ao longo dos tempos, dentre as quais cabe ressaltar a Distribuição Normal que, segundo Ross (2000), foi a primeira e mais utilizada função desenvolvida. Criada pelo matemático francês Abraham De Moivre, em 1733, com a finalidade de aproximar probabilidades referentes a lançamentos de moedas, foi apenas em 1809 que o conceituado matemático Johann Carl Friedrich Gauss a usou como uma parte integral da sua aproximação para predizer o local de entidades astronômicas, publicada na terceira seção do livro *Theoria Motus*. Nesse modelo a área sob a curva entre a média e um ponto arbitrário é função do número de desvios padrão entre a média e aquele ponto. Essa distribuição desempenha um papel essencial na história do estudo das populações e é amplamente aplicada na área biológica, resultando em uma excelente aderência.

Apesar das inúmeras qualidades inerentes à curva normal, pesquisas revelaram que em alguns casos os dados distribuem-se de forma assimétrica, caso não contemplado por essa notável distribuição. Essas novas conformações fomentaram uma busca por novas soluções, dentre as quais se destacaram aquelas de fácil ajuste e relativa flexibilidade. Vale lembrar que a implementação computacional é muito recente e que há apenas algumas décadas todos os

cálculos eram realizados sem o auxílio de computadores. Funções complexas e com cálculos complicados para ajuste e estimativa de probabilidade eram imediatamente descartadas.

Dentre as distribuições que evoluíram, resalta-se um dos modelos mais utilizados atualmente, a função Weibull. Foi desenvolvida em Stockholm, na Suécia, por Ernst Hjalmar Waloddi Weibull em 1939 com a finalidade de realizar um planejamento estatístico sobre fadiga de materiais. A distribuição Beta é utilizada principalmente para análises Bayesianas, PERT (Program Evaluation and Review Technique) e CPM (Critical Path Method), por ser binomial. A função Gama é chamada de distribuição χ_n^2 (lê-se qui-quadrado) com n graus de liberdade quando $\lambda = \frac{1}{2}$ e $t = \frac{n}{2}$, sendo n um inteiro positivo. A Exponencial pode ser considerada como caso particular da Gama e tem sido utilizada para prever o período de tempo necessário até a ocorrência de um evento. No entanto, até mesmo essas distribuições vêm tendo desempenho abaixo das expectativas quando a variabilidade dos dados é alta.

As distribuições diamétricas e hipsométricas das espécies arbóreas podem apresentar uma tendência simétrica ou assimétrica, mas com comportamento amplamente estudado. Muitos são os estudos sobre espécies do gênero Pinus e Eucalyptus. Todavia, as florestas naturais do Brasil possuem características únicas em exuberância e heterogeneidade, bem como abrangem extensas áreas. Dessa maneira, a única alternativa para abordá-las é a amostragem, uma vez que nessas circunstâncias torna-se inviável o censo, em função do tempo e, principalmente, do custo requerido para essa operação. Muitos estudos têm sido realizados no Brasil a fim de entender a dinâmica de crescimento de espécies nativas.

Outros pontos importantes surgem ao se avaliar mais aprofundadamente o comportamento do crescimento de florestas nativas brasileiras, como a floresta amazônica, e florestas plantadas com espécies exóticas. Entre os casos de relevância cabe salientar a ocorrência de duas ou mais modas.

1.1 JUSTIFICATIVA

Estudar a distribuição diamétrica de uma floresta é uma maneira eficiente de se entender seu comportamento, por fornecer estimativas de parâmetros imprescindíveis para seu manejo, como crescimento e produção. Uma forma de se realizar esse estudo é por meio do ajuste de modelos, dos quais se destacam os probabilísticos. Embora existam muitos modelos descritos na literatura, algumas características têm sido ignoradas ao se utilizar curvas que não descrevem com precisão a tendência biológica ocorrente.

A busca pela precisão tem sido cada vez mais almejada com o desenvolvimento das ciências. Isso se deve principalmente pela crescente interdisciplinaridade que vem ocorrendo, em que se podem observar estudos integrados entre as diferentes áreas do conhecimento (biológicas, exatas, médicas, humanas, sociais e outras). Aliado a essa nova realidade está o uso da informática como facilitadora dos processos de cálculo. Esse é sem dúvida um dos motivos da disseminação do uso de técnicas estatísticas cada vez mais avançadas.

Em florestas em que há ação antrópica, ou onde ocorreram distúrbios significativos, a distribuição diamétrica sofre uma mudança na sua tendência natural. Nesses casos, se não houver necessidade de precisão, um modelo generalista será de grande valia, atendendo satisfatoriamente ao estudo em questão. Contudo, quando a pesquisa requer o entendimento das sutilezas da natureza, surge a necessidade de se criar um novo modelo.

A distribuição diamétrica da castanheira, conforme especificada no relatório do inventário florestal realizado na região de Altamira, apresentou a ocorrência de mais de uma moda. Nesse caso, isso deve-se nitidamente a distúrbios naturais que ocorrem na floresta. Posteriormente, estudando a distribuição da Araucária em um remanescente no estado do Paraná, foi verificado um evento similar ao da castanheira, porém o motivo é claramente outro: a ação antrópica, mais precisamente, o seu corte para fins comerciais.

Fundamenta-se esta pesquisa na necessidade de desenvolver uma função densidade de probabilidade capaz de representar, de forma satisfatória, a distribuição diamétrica, seja de espécies florestais nativas ou plantadas, manejadas ou não, em que se observe a ocorrência de duas ou mais modas. Esse modelo deve ser suficientemente flexível para representar cada uma das modas e ter algumas características biológicas.

1.2 OBJETIVOS

1.2.1 Objetivo Geral

O objetivo principal da presente pesquisa foi propor uma nova função densidade de probabilidade com ocorrência de dois ou mais pontos de máximo local.

1.2.2 Objetivos Específicos

- Elaborar e propor um modelo capaz de:
 - Interceptar ou não o eixo das ordenadas na origem, de acordo com a premissa admitida;
 - Ser integrável em seu domínio;
 - Ser contínuo em todo o seu domínio;
 - Representar satisfatoriamente a tendência observada empiricamente;
 - Ter média e variância calculadas por meio de integrais;
 - Ser definido por apenas uma sentença;
 - Tender para zero quando X tender para infinito, ou seja, deve estimar que a quantidade de indivíduos nas classes diamétricas superiores tenda a zero, mas não deve desconsiderar a possibilidade da sua existência;
 - Apresentar tantas modas quantas as observadas a campo; e
 - Ser de fácil interpretação e ajuste.
- Comparar a eficiência do modelo proposto com a das distribuições clássicas (Normal, Exponencial, Weibull, Beta e Gama), de Quadros e Spline.

2. REVISÃO DE LITERATURA

2.1 MATEMÁTICA

Courant (1965) afirmou que variável real é um símbolo que representa qualquer um dos números de um conjunto real de números. Este conjunto é o campo ou domínio da variável. Constante é uma quantidade cujo valor permanece inalterado num dado problema. O campo ou domínio de uma constante é, simplesmente, um número.

Entende-se por uma função f uma terna $(A, B, a \mapsto b)$ onde A e B são dois conjuntos e $a \mapsto b$, uma regra que permite associar a cada elemento a de A um único elemento b de B . O conjunto A é o *domínio* de f e indica-se por D_f , assim $A=D_f$. O conjunto B é o *contradomínio* de f . O único b de B associado ao elemento a de A é indicado por $f(a)$ (GUIDORIZZI, 1994).

Seja S um espaço em que se defina uma função de probabilidade. Seja x uma função de valores reais definida em S (a função x transforma pontos de S em pontos de x). Diz-se que x é uma variável aleatória (variável aleatória unidimensional) (MOOD; GRAYBILL, 1978). As variáveis aleatórias podem dividir-se em duas classes: 1) uma variável aleatória discreta se toma somente um número finito ou uma infinidade numerável de valores; 2) Uma variável aleatória contínua pode assumir qualquer valor de certo intervalo ou coleção de intervalos sobre esse x , sem a restrição de que aqueles sejam números isolados.

Uma variável discreta pode assumir um número finito ou infinito numerável de valores, enquanto que uma variável contínua pode assumir um número infinito de valores não numeráveis.

Diz-se que X é uma variável aleatória discreta unidimensional se é uma variável aleatória que toma somente um número finito ou infinito numerável de valores desse x . Suponha-se que x toma unicamente valores $X_1, X_2, \dots, X_n, \dots$ com probabilidades $f(X_1), f(X_2), \dots, f(X_n), \dots$ e imagine-se que A é qualquer subconjunto dos pontos $X_1, X_2, \dots, X_n, \dots$, a probabilidade $P(A)$ do sucesso A (probabilidade de que X esteja em A) se define como $P(A) = \sum_A f(X)$, em que $\sum_A f(X)$ representa a soma de $f(X)$ para aqueles valores X_i que pertencem a A (MOOD; GRAYBILL, 1978).

Mood e Graybill (1978) definem que x é uma variável aleatória contínua unidimensional se existe uma função f tal que $f(x) \geq 0$ para todo x do intervalo $-\infty < x < \infty$

e tal que para qualquer sucesso de A $P(A) = P(x \text{ esteja em } A) = \int_A f(x)dx$, em que $f(x)$ se denomina função densidade de x .

2.1.1 Funções

De acordo com Stewart (2011) uma função é uma lei a qual para cada elemento x em um conjunto A faz corresponder um elemento chamado $f(x)$, em um conjunto B . O conjunto A é chamado domínio da função f , enquanto B , conjunto de todos os objetos que correspondem a elementos de A , é chamado de contradomínio da função f (MOISE, 1972). As mesmas operações válidas para números são válidas para funções, ou seja, pode-se produzir novas funções procedendo a operações como adicionar, subtrair, multiplicar e dividir (SWOKOWSKI, 1994). Se uma função for definida por $f(x) = a_n x^n + a_{n-1} x^{n-1} + \dots + a_1 x + a_0$, em que a_0, a_1, \dots, a_n são números reais $a_n \neq 0$ e n é um número inteiro não negativo, então $f(x)$ será chamada de função polinomial de grau n (STEWART, 2011).

De acordo com Anton (2007) uma função que pode ser expressa como uma razão de dois polinômios $P(x)$ e $Q(x)$ é chamada racional (EQUAÇÃO 01). Seu domínio será o conjunto de todos os valores de x tais que $Q(x) \neq 0$.

$$f(x) = \frac{P(x)}{Q(x)} \quad \text{EQ. 01}$$

Para qualquer x real, define-se e^x como aquele número e cujo logaritmo é x , ou seja, $f(x) = e^x$ (APOSTOL, 1977). Pode-se, assim como as demais funções, utilizar as famílias oriundas dessa função.

2.1.2 Limites e continuidade

Sejam f uma função e p um ponto do domínio de f ou extremidade de um dos intervalos que compõem o domínio de f . Diz-se que f tem limite L em p , se, para todo $\varepsilon > 0$ dado, existir um $\delta > 0$ tal que, para todo $x \in D_f$, $0 < |x - p| < \delta \Rightarrow |f(x) - L| < \varepsilon$. Tal número L , que quando existe e é único, é indicado por $\lim_{x \rightarrow p} f(x)$. Assim,

$$\lim_{x \rightarrow p} f(x) = L \Leftrightarrow \begin{cases} \forall \varepsilon > 0, \exists \delta > 0 \text{ tal que, para todo } x \in D_f \\ 0 < |x - p| < \delta \Rightarrow |f(x) - L| < \varepsilon \end{cases} \quad (\text{GUIDORIZZI, 1994}).$$

De acordo com Flemming e Gonçalves (1992), uma função f é contínua no ponto a se as seguintes condições forem satisfeitas:

f é definida no ponto a ;

$\lim_{x \rightarrow a} f(x)$ existe; e

$\lim_{x \rightarrow a} f(x) = f(a)$.

2.1.3 Derivadas

A derivada de uma função $y = f(x)$ em relação a x é definida por $\frac{dy}{dx} = f'(x) = y' = \lim_{\Delta x \rightarrow 0} \frac{\Delta y}{\Delta x} = \lim_{\Delta x \rightarrow 0} \frac{f(x + \Delta x) - f(x)}{\Delta x}$ desde que o limite exista (PISKOUNOV, 1993). Toda função derivável num ponto x_1 é contínua nesse ponto (PISKOUNOV, 1993).

Dado que $f(x)$ existe para todos os valores de $x \in (a, b)$ e f tem um extremo relativo em c , onde $a < c < b$, se $f'(c)$ existe, então $f'(c) = 0$. Se f for derivável em todos os pontos de seu domínio, então f será contínua em todos os pontos de seu domínio (GRANVILLE; LONGLEY, 1966).

A teoria de derivadas tem aplicações amplas, desde estudo de comportamento de funções (crescentes, decrescentes, pontos críticos: máximos, mínimos e de inflexão), taxas de variação, problemas de otimização, movimento retilíneo uniformemente variado, entre outros. Na estatística desempenha um papel fundamental, pois permite, a partir da função acumulada, encontrar a função densidade de probabilidade.

2.1.4 Integrais

De acordo com Swokowski (1994), supondo-se f contínua em um intervalo fechado $[a, b]$:

- Se a função G é definida por $G(x) = \int_a^x f(t) dt$ para todo x em $[a, b]$, então G é uma antiderivada de f em $[a, b]$.
- Se F é qualquer antiderivada de f em $[a, b]$, então $\int_a^b f(x) dx = F(b) - F(a)$.

Swokowski (1994) define as integrais impróprias como:

- Se f é contínua em $[a, \infty)$ então $\int_a^\infty f(x) dx = \lim_{t \rightarrow \infty} \int_a^t f(x) dx$, desde que o limite exista; e

- Se f é contínua em $(-\infty, a]$, então $\int_{-\infty}^a f(x)dx = \lim_{t \rightarrow \infty} \int_t^a f(x)dx$, desde que o limite exista.

2.1.5 Teorema Fundamental do Cálculo

De acordo com Swokowski (1994), supondo-se f contínua em um intervalo fechado $[a, b]$:

- Se a função G é definida por $G(x) = \int_a^x f(t)dt$ para todo x em $[a, b]$, então G é uma antiderivada de f em $[a, b]$.
- Se F é qualquer antiderivada de f em $[a, b]$, então $\int_a^b f(x)dx = F(b) - F(a)$.

A teoria de integrais, conforme Stewart (2011), possui uma ampla área de utilização, dentre as quais se destacam: o cálculo de áreas entre curvas, volumes e áreas de sólidos de revolução, volume de sólidos quaisquer, comprimento de arco, na física (pressão hidrostática e força, momentos e centro de massa), na economia (excedente do consumidor), na biologia (circulação sanguínea, capacidade cardíaca), na estatística (em funções densidade de probabilidade: cálculos de probabilidade, valores médios, variância, assimetria e curtose).

2.2 MODELAGEM

De acordo com Stevenson (1981), um dos principais instrumentos extensamente usados na estatística é o modelo. Os modelos são versões simplificadas de algum problema ou situação da vida real. São usados para ilustrar certos aspetos da situação, evitando grande número de detalhes que talvez sejam irrelevantes para o problema; podem, assim, ajudar a reduzir o grau de complexidade de um problema.

Segundo Meyer (1969), todas as vezes que se emprega matemática, a fim de estudar alguns fenômenos de observação, deve-se essencialmente começar por construir um modelo matemático (determinístico ou probabilístico) para eles. Inevitavelmente, o modelo deve simplificar as coisas e certos pormenores devem ser desprezados. O bom resultado do modelo depende de que os pormenores desprezados sejam ou não realmente sem importância na elucidação do fenômeno estudado. A resolução do problema matemático pode estar correta e, não obstante, estar em grande discordância com os dados observados, simplesmente porque as hipóteses básicas feitas não forem confirmadas. Geralmente é bastante difícil afirmar com certeza se um modelo matemático especificado é ou não adequado, antes que alguns dados de

observação sejam obtidos. A fim de verificar a validade de um modelo deve-se deduzir certo número de consequências do modelo e, a seguir, comparar esses resultados previstos com observações.

Um modelo determinístico refere-se àquele em que as condições sob as quais um experimento é executado determinam o resultado do experimento. Nesses modelos empregam-se “considerações físicas” para prever o resultado, enquanto que em um modelo não-determinístico (também chamado de probabilístico ou estocástico) emprega-se a mesma espécie de considerações para especificar uma distribuição de probabilidade (MEYER, 1969).

Para Stevenson (1981), a “estimação” é o processo que consiste em utilizar dados amostrais para estimar os valores de parâmetros populacionais desconhecidos. Essencialmente, qualquer característica de uma população pode ser estimada a partir de uma amostra aleatória. Entre os mais comuns estão a média e o desvio padrão de uma população e a proporção populacional. A avaliação de inventários, a estimação do custo de projetos, a avaliação de novas fontes de energia, predições sobre a realização de empreendimentos, estimativas de tempo médio são todas situações que envolvem estimação. A estimativa pontual é a estimativa única de um parâmetro populacional, enquanto que a estimativa intervalar dá um intervalo de valores possíveis, no qual se admite que esteja o parâmetro populacional.

A modelagem para a Ciência Florestal teve seu início no inventário florestal, uma vez que essa técnica possibilita a descrição quantitativa e qualitativa das florestas naturais ou plantadas. A descrição quantitativa pode ser feita a partir do volume de madeira e do número de árvores disponíveis, enquanto que a qualitativa versa sobre as espécies contidas na floresta, suas características e utilizações.

Atualmente os modelos são amplamente utilizados nas mais diversas áreas do conhecimento, das quais vale ressaltar as Ciências Exatas e da Terra (astronomia, geografia, engenharias), Ciências Agrárias e Ambientais (agronomia, medicina veterinária, engenharia florestal, zootecnia), Ciências Biológicas e Biomédicas (biologia marinha e molecular, medicina, enfermagem) e Ciências Humanas (estudos comportamentais). O inventário florestal fornece a situação da floresta no momento em que essa foi medida. (PENG, 2000).

A modelagem pode ser feita a partir de árvores individuais, que são obtidas pelo inventário de pontos permanentes e podem indicar a mortalidade do indivíduo em estudo, além de outros aspectos. A modelagem empírica é a base da modelagem utilizada em manejos florestais e a mecanicista é mais direcionada para pesquisas. A modelagem florestal está intimamente ligada ao manejo de florestas, pois, com a condição de se predizer o futuro de

uma floresta, é mais fácil explorá-la de forma racional, sem que haja agressões ao meio ambiente (VANCLAY et al., 1997).

Para que um modelo matemático possa ser considerado uma função densidade de probabilidade, é necessário que as condições descritas no próximo item sejam atendidas.

2.3 ESTATÍSTICA

Segundo Johnson e Leone (1964), a estatística pode ser dividida em três partes: 1) Estatística descritiva: Compreende a organização, o resumo e, em geral, a simplificação de informações, que podem ser muito complexas; 2) Estudo sobre probabilidade: Útil para analisar situações que envolvam acaso, como os jogos de dados, cartas e lançamentos de moedas; e 3) Inferência estatística: Diz respeito à análise e interpretação de dados amostrais.

Ainda, de acordo com Stevenson (1981), as probabilidades são utilizadas para exprimir a chance de ocorrência de determinado evento. A probabilidade de ocorrência de uma evento é dada por um número e que pode variar de 0 a 1. A definição clássica de probabilidade, segundo Mood e Graybill (1978), diz que se um sucesso pode ocorrer de n maneiras mutuamente excludentes e igualmente verossímeis e se n_A dessas possuem um atributo A , a probabilidade de A é a fração $\frac{n_A}{n}$.

A finalidade da amostragem, conforme descrito por Johnson e Leone (1964), é obter uma indicação do valor de um ou mais parâmetros de uma população, tais como a média, o desvio padrão, ou a proporção de itens que possuem determinada característica. As estatísticas amostrais que correspondem a esses parâmetros populacionais são usadas para aproximar os valores desconhecidos daqueles parâmetros. Assim é que a média amostral é usada para estimar a média da população, o desvio padrão amostral é usado para estimar o desvio padrão populacional, e a proporção amostral serve para estimar a proporção na população. Uma distribuição amostral é uma distribuição de probabilidades que indica até que ponto uma estatística amostral tende a variar devido a variações causais na amostragem aleatória.

Uma distribuição de probabilidade é uma distribuição de frequências relativas para os resultados de um espaço amostral; mostra a proporção das vezes em que a variável aleatória tende a assumir cada um dos diversos valores (PAGANO; GAUVREAU, 2003). As distribuições descontínuas de probabilidades envolvem variáveis aleatórias relativas a dados que podem ser contados, como o número de ocorrências por amostra, ou o número de ocorrências por unidade num intervalo de tempo, de área, ou de distância (MILONE, 2009).

De acordo com Meyer (1969), seja X uma variável discreta. Portanto, R_X , o contradomínio de X , será formado no máximo por um número infinito numerável de valores X_1, X_2, \dots . A cada possível resultado X_i associa-se um número $p(X_i) = P(X = X_i)$, denominado probabilidade de x_i . Os números $p(X_i), i = 1, 2, \dots$ devem satisfazer as seguintes condições, em que a função p é denominada função de probabilidade da variável aleatória X :

a) $p(X_i) \geq 0$, para todo i ,

b) $\sum_{i=1}^{\infty} p(X_i) = 1$.

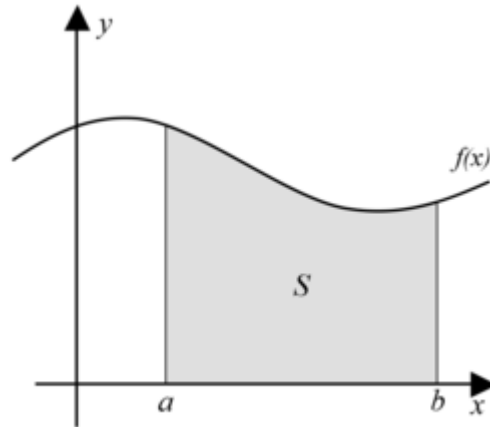
Função densidade de probabilidade contínuas, denotada por fdp, são limites de uma infinita sequência de distribuições discretas, cujas variáveis formam um conjunto enumerável de valores (ANDERSON; SWEENEY; WILLIAMS, 2008). Tais distribuições devem atender às seguintes condições, segundo Silva (2003):

- $f(x) \geq 0$, para todos os valores de X , pois não existe probabilidade negativa;
- $\int_{-\infty}^{+\infty} f(x)dx = 1$, o que determina que a área entre a curva representativa da função $f(x)$ em todo o intervalo seja igual a 1; e
- $\int_a^b f(x)dx = P(a < X \leq b), b > a$, ou seja, a probabilidade da variável aleatória X assumir valor em um intervalo será dada pela integral da função nesse intervalo. Geometricamente fica estabelecido que a probabilidade em um determinado intervalo corresponde à área determinada sob a região plana delimitada pela função nesse intervalo.

Propriedades decorrentes da definição de função densidade de probabilidade (MEYER, 1969):

a) $P(a < X < b)$ representa a área sob a curva no gráfico da figura 01, entre $X = a$ e $X = b$.

FIGURA 01 – Representação da área sob a curva



b) Uma consequência decorrente da descrição probabilística de X acima é que, para qualquer

valor especificado de X , X_0 , tem-se $P(X = X_0) = 0$, porque $P(X = X_0) = \int_{x_0}^{x_0} f(x)dx = 0$;

c) Se uma função f satisfizer as condições $f(x) \geq 0$ para todo x , e $\int_{-\infty}^{+\infty} f(x)dx = k$, em que k é

um número real positivo (não necessariamente igual a 1), então f não satisfaz todas as condições para ser uma fdp. No entanto, pode-se facilmente definir uma nova função g , em

termos de f , assim, $g(x) = \frac{f(x)}{k}$ para todo x . Dessa maneira, g satisfará todas as condições de

uma fdp.

d) Se X tomar valores somente em algum intervalo finito $[a, b]$, pode-se simplesmente atribuir $f(x) = 0$ para todo $x \notin [a, b]$. Em consequência, a fdp ficará definida para todos os valores

reais de x , e é possível arbitrar que $\int_{-\infty}^{+\infty} f(x)dx = 1$.

2.4 PRINCIPAIS DISTRIBUIÇÕES DISCRETAS

2.4.1 Bernoulli

Segundo Ross (2000), a Distribuição de Bernoulli aplica-se a experimentos em que o resultado possa ser classificado como sucesso ou falha. Assumindo $X = 1$ para sucesso e $X = 0$ para falha, então a função de probabilidade de X é dada por:

$$p(0) = P\{X = 0\} = 1 - p$$

$$p(1) = P\{X = 1\} = p$$

Em que $0 \leq p \leq 1$ é a probabilidade de sucesso.

Esta variável aleatória X será chamada de variável aleatória de Bernoulli, para algum $p \in (0, 1)$.

O matemático suíço Jacques Bernoulli (1654 – 1705) estudou experimentos independentes tendo uma probabilidade de sucesso comum p . Em seu livro *Ars Conjectandi* (a arte de conjecturar), publicado pelo seu sobrinho Nicolas oito anos após sua morte em 1713, Bernoulli mostrou que se um número de tais experimentos for grande, então a proporção deles que terá sucesso se aproximará de p com uma probabilidade próxima de 1. Jacques Bernoulli é da primeira geração da família dos mais famosos matemáticos de todos os tempos. Sua família fez contribuições fundamentais para a probabilidade, estatística e matemática. Uma dificuldade em saber o número exato de contribuições de cada membro da família consiste no fato de que alguns deles possuíam o mesmo nome. Outra dificuldade é que eles utilizavam nomes diferentes em diferentes localidades, por exemplo, Jacques, aqui citado, assinava também como Jaques. Entretanto, seus números, influência e resultados foram prodígios. Assim como os Bach para a música, os Bernoullis para a matemática constituíram uma família que entrou para a história (ROSS, 2000).

2.4.2 Binomial

Usa-se o termo binomial para designar situações em que os resultados de uma variável aleatória podem ser agrupados em duas classes ou categorias. Os dados são, pois, nominais. As categorias devem ser mutuamente excludentes, de modo a deixar perfeitamente claro a qual categoria pertence determinada observação; e as classes devem ser coletivamente exaustivas, de forma que nenhum resultado fora delas é possível (STEVENSON, 1981). A distribuição binomial é, provavelmente, a de uso mais frequente entre as distribuições discretas, nas aplicações da teoria estatística. Essa distribuição está associada com as provas repetidas de um mesmo sucesso (MOOD; GRAYBILL, 1978). Essa distribuição tem ampla aplicabilidade no campo da confiabilidade, principalmente quando se está interessado no tempo de duração de um determinado evento, duração média de um evento, entre outros (TRIOLA, 2008).

Ela é descrita conforme expressão 02 (MONTGOMERY; RANGER, 2003).

$$f(X) = \binom{n}{p} p^x (1-p)^{n-x}, X = 0, 1, 2, \dots, n \quad \text{EQ. 02}$$

A distribuição probabilística binomial descreve a distribuição da variável X , o número de sucessos em n repetições, se o experimento satisfaz as seguintes condições (FREUND; WILSON, 1996):

- O experimento consiste em n tentativas idênticas;
- Cada tentativa resulta em uma de duas possibilidades mutuamente exclusivas, uma chamada sucesso e outra falha;
- A probabilidade de sucesso em uma repetição é igual a p e permanece constante durante todo o experimento; e
- Os eventos são independentes.

2.4.3 Poisson

De acordo com Larson e Farber (2010), a distribuição de Poisson é útil para descrever as probabilidades do número de ocorrências num campo ou intervalo contínuo (em geral tempo ou espaço). Ela é utilizada em uma ampla gama de situações físicas, como modelo matemático para descrever, em dado intervalo de tempo, eventos tais como emissão de partículas de uma substância radiotiva, chegadas de passageiros a um terminal aéreo, distribuição de partículas de poeira que atingem certo local, chegadas de carros a um cruzamento (SOONG, 1986). Diz-se que a variável X se distribui segundo uma distribuição de

Poisson se a função de quantidade é $f(X) = \frac{e^{-m} m^X}{X!}$, com $X = 0, 1, 2, \dots$ em que m é

qualquer número positivo (MOOD; GRAYBILL, 1978). Essa distribuição baseia-se nas seguintes hipóteses:

- 1) A probabilidade de uma ocorrência é a mesma em todo o campo de observação;
- 2) A probabilidade de mais de uma ocorrência num único ponto é aproximadamente zero; e
- 3) O número de ocorrências em qualquer intervalo é independente do número de ocorrências em outros intervalos.

2.4.4 Multinomial

A distribuição multinomial, segundo Stevenson (1981), é usada em situações onde há mais de dois resultados mutuamente excludentes. Tal como na binomial, exige-se que as provas sejam independentes, com probabilidade constante. Um exemplo de distribuição

multinomial é a jogada de um dado (SPIEGEL, 2009). Essa distribuição admite que hajam mais de dois eventos possíveis $E_1, E_2, E_3, \dots, E_n$.

2.4.5 Hipergeométrica

A distribuição hipergeométrica refere-se a situações com dois ou mais resultados, em que a probabilidade de sucesso varia de uma prova para outra (LEVINE et al., 2008). Nesse caso as tentativas não são independentes e as amostras são frequentemente selecionadas sem reposição (MONTGOMERY, RUNGER, 2003). A função de quantidade hipergeométrica está apresentada na equação 03, em que se tem N objetos, entre K classificados como sucessos, $N - K$ como falhas, em que $K \leq N$, $n \leq N$ e x é a variável aleatória que representa o número de sucessos na amostra (MOOD; GRAYBILL, 1978).

$$f(X) = \frac{\binom{K}{X} \binom{N-K}{n-X}}{\binom{N}{n}}, \text{ com } X = 0, 1, 2, \dots, r \quad \text{EQ. 03}$$

2.4.6 Distribuição geométrica

Em situações em que se deseja saber o número de provas necessárias para a ocorrência do primeiro sucesso utiliza-se a distribuição geométrica (SOONG, 1986). Ela é denotada como uma série de tentativas independentes de Bernoulli, com probabilidade constante p de um sucesso, X denota o número de tentativas até que o primeiro sucesso ocorra, assim a distribuição será $f(X) = (1-p)^{X-1} p$, em que $X = 1, 2, \dots$ (MONTGOMERY, RUNGER, 2003)

2.5 PRINCIPAIS DISTRIBUIÇÕES CONTÍNUAS

Dentre as funções densidade de probabilidade clássicas destacam-se a Exponencial, Gama, Beta, Weibull e Normal (Gauss), que serão apresentadas a seguir.

Alguns trabalhos importantes em que foram utilizadas funções densidade de probabilidade no setor florestal estão apresentados a seguir. Neles se observa a grande versatilidade dos modelos probabilísticos.

Eisfeld et al. (2005) trabalharam com modelagem do crescimento e da produção de *Pinus taeda* L. por meio das distribuições Gamma, Lognormal, Beta, SB Johnson, SBB de Johnson, Weibull, Exponencial e a Normal. Nesse trabalho os autores utilizaram dados de 325 parcelas permanentes de *Pinus taeda* sem desbaste, da empresa International Paper do Brasil.

O melhor ajuste foi obtido com a distribuição Weibull. Alguns autores citam a distribuição Gama com dois “m” (Gamma). Isso é decorrente da expressão na língua inglesa. Em português o termo existe e escreve-se com apenas um “m”.

Maestri (1992) estudou a contribuição de variáveis ambientais na redução de erro das estimativas de crescimento e produção de madeira de povoamentos clonais de híbridos naturais de *Eucalyptus grandis* implantados no norte do Espírito Santo e Sul da Bahia. O sistema completo incluiu a modelagem da sobrevivência de árvores, área basal, variância dos diâmetros, ajuste da função Weibull à distribuição diamétrica, relação hipsométrica genérica e função de afilamento do tronco.

Abreu (2002), com dados precoces de *Eucalyptus grandis*, modelou os atributos da floresta (sítio, sobrevivência, relação hipsométrica genérica, diâmetro mínimo, diâmetro máximo, variância dos diâmetros, área basal e média aritmética dos diâmetros), por meio do ajuste das funções probabilísticas como Beta, Weibull e S_B de Johnson por diferentes métodos e avaliou a veracidade das prognoses e a eficiência do modelo para realizá-las precocemente. O estudo foi desenvolvido com dados de *Eucalyptus grandis* provenientes da empresa Votorantim Celulose e Papel S/A – VCP, situada no município de Luiz Antônio, no Estado de São Paulo. Nesse estudo, foi verificado que a função de distribuição Weibull, ajustada pelo método dos momentos, mostrou-se a mais apropriada.

As funções Beta, Gamma, Normal, S_B de Johnson e Weibull com três parâmetros foram ajustadas e testadas por Valejos (2003) no desenvolvimento de um sistema de simulação de crescimento e produção de *Populus spp* no Chile. O sistema foi denominado SALICA e utilizou a plataforma do software Excel 2000.

Meira-Neto, Martins e Souza (2005) ajustaram a curva Normal aos dados de frequência do número de parcelas pelas classes de valores de cobertura média do solo, utilizando-se fotogrametria para os meses de agosto de 1993 e de janeiro de 1994, na mata da Silvicultura, Viçosa, MG. O objetivo do trabalho foi estabelecer a relação existente entre espécies do estrato herbáceo-arbustivo e a cobertura do dossel.

Lima et al. (2003) concluíram que a distribuição normal ajustou-se satisfatoriamente para as concentrações de pH, Al, Ca, Mg, H, P, K, C, areia, silte, argila, soma de bases ($S = Ca + Mg + Na + K$), capacidade de troca catiônica ($CTC = S + H + Al$), saturação por bases ($V = 100.S/T$) e saturação por alumínio. O objetivo do trabalho foi agrupar parte das espécies arbóreas de uma floresta primária do Estado do Amapá, Amazônia Oriental, por meio das características do solo.

Bartoszeck et al. (2004) mostraram que a distribuição diamétrica da bracatinga (*Mimosa scabrella* Benth), em que a densidade inicial foi de 2000 árvores/ha, seguiu a mesma tendência apresentada pela distribuição normal. O estudo teve como objetivo estudar o comportamento de funções densidade de probabilidade ajustadas a distribuições diamétricas da bracatinga nos municípios da Região Metropolitana de Curitiba, Paraná.

Silva et al. (2008) compararam as distribuições Log-normal, Gama, Weibull e Exponencial para determinar a que melhor caracteriza a distribuição dos valores mensais de deficiência hídrica no solo da região de abrangência da estação meteorológica de Santa Maria e demonstraram que as distribuições lognormal e exponencial foram as mais eficientes para tal propósito. Para os dados decendiais, os melhores resultados foram obtidos pelos modelos lognormal, exponencial e gama.

Cunha Neto et al. (1994) desenvolveram um modelo de predição presente e futura para *Eucalyptus grandis*, utilizando a distribuição de Weibull 3 parâmetros para representar a distribuição diamétrica em diferentes idades, densidades e classes de produtividade dos sítios. Os dados para a pesquisa são provenientes da CELPAV (Cia. Votorantim Celulose e Papel), no município de Luiz Antônio, São Paulo.

Santana et al. (2008) aplicaram o modelo de Weibull para desenvolver um estudo preliminar da velocidade e direção dos ventos na região norte da cidade Cuiabá, Mato Grosso. Por meio dos coeficientes obtidos no ajuste desse modelo foi possível concluir que a primavera é a estação com ventos mais intensos e que não é viável a exploração eólica para geração de energia elétrica na região estudada.

Silvino et al. (2007) usaram as distribuições de probabilidade Log-Normal, Normal, Exponencial, Gama e Weibull para estimar as vazões máximas e mínimas anuais para o rio Paraguai, utilizando dados de uma estação fluviométrica localizada na cidade de Cáceres, MT. Foi observado que as distribuições Gama e Weibull foram as mais aderentes, respectivamente, para se modela tais eventos aleatórios

Mourão Júnior e Barbosa (2005) utilizaram as distribuições Normal, Log-Normal, Log-Normal 3p, Exponencial, Exponencial modificada, Gama, Gama 3 P, Log Gama, Gama invertida, Gumbel mínima, Gumbel máxima, Frechet mínimo e máximo, 2 e 3 P, Weibull máximo, mínimo, 2 e 3 P para estudo de parâmetros dendrométricos do caimbé (*Curatella americana* L.) nos ambientes de savana aberta do extremo norte amazônico. Apesar do baixo número de indivíduos amostrados no estudo, a distribuição que mais aderiu aos dados de diâmetro de copa e da base foi a lognormal, enquanto que para a altura foi a normal.

Mesquita et al. (2007) testaram uma alternativa para a caracterização da condutividade hidráulica saturada no solo da Região Centro-Oeste do Estado de São Paulo por meio das distribuições Normal e Log-Normal. Os autores concluíram que a probabilidade de ocorrência associada ao modelo lognormal ajustado é uma alternativa para a representação de tal evento aleatório.

Catalunha et al. (2002) avaliaram a precipitação fluvial no estado de Minas Gerais por meio das distribuições de probabilidade Exponencial, Gama, Log-normal (a dois e três parâmetros), Normal e Weibull. As distribuições Weibull e Exponencial foram superiores às demais para valores diários de precipitação (exceto para decêncios do período seco), enquanto que para valores totais o melhor resultado foi obtido por meio da Exponencial.

Furlan (2009) ajustou a distribuição Gama para avaliação de variabilidade temporal e espacial das chuvas em Rondônia, analisando se o desmatamento ocorrido nos últimos 30 anos foi suficiente para contribuir para a modificação do regime hídrico e pluviométrico na região. Ela concluiu que a ação antrópica não modificou significativamente o regime pluviométrico e que esse decorre da variabilidade climática natural.

Santos (2010) determinou o volume de madeira de um híbrido de *Eucalyptus tereticornis* por meio das distribuições Normal, Student, Exponencial, Cauchy e Logísticas I e II. O experimento foi realizado na região cujas coordenadas geográficas de posição são 07°29' 00S e 40°36' 00W e altitude de 816 metros.

Rossato (2006) aplicou os modelos Normal, Log-Normal, Beta, Gama e Weibull para explicar o comportamento dos teores nutricionais para classificação nos diversos tecidos vegetais de *Acacia mearnsii*, *Eucalyptus* sp. e *Pinus* sp.. As funções aplicadas, exceto a beta, podem explicar a distribuição dos resultados das análises químicas de tecidos vegetais.

Arce (2004) avaliou a estrutura de florestas clonais de *Populus deltoides* por meio do ajuste das distribuições Normal, Gama, Beta, SB de Johnson e Weibull com dois e três parâmetros. Nesse estudo a função Weibull apresentou os menores valores de desvios relativos elevados à quarta potência.

Machado, Mello e Barros (2000) compararam os modelos Normal, Log-Normal, Gama, Weibull e Beta para analisar seus ajustes aos dados de volume total de pinheiro (*Araucaria angustifolia* (Bert.) O. Ktze) por unidade de área.

Machado et al. (2009) consideraram dados dendrométricos de um fragmento de Floresta Ombrófila Mista, localizado em Curitiba, Paraná, em que foram ajustados os modelos Gama, Gama adaptada, Weibull com 2 e 3 parâmetros, SB de Johnson, Exponencial e Weber. O objetivo foi caracterizar a estrutura diamétrica das espécies *Cedrela fissilis*

(Cedro), *Luehea divaricata* (Açoita Cavallo), *Gochnatia polymorpha* (Cambará), *Sebastiania commersoniana* (Branquilha) e *Casearia sylvestris* (Cafezeiro). A distribuição de Weber foi a melhor para toda a floresta, cedro, cambará e branquilha, enquanto que a Gama foi mais eficiente para o cafezeiro e nenhuma ajustou-se aos dados do açoita cavalo.

Binoti et al. (2010) ajustaram o modelo de Weibull com 3 parâmetros aos dados de distribuição diamétrica para plantios de eucalipto submetido a desbaste, localizado na região Nordeste do Estado da Bahia.

Em um povoamento de *Pinus elliottii* no município de Itapeva, São Paulo, Jorge e Veiga (1990) ajustaram a função Weibull.

Higuchi (2007) ajustou a função probabilística de Weibull aos dados da floresta amazônica, tendo obtido resultados satisfatórios. Ele concluiu, ainda, que esse modelo pode auxiliar na fiscalização do desmatamento da região.

Weber (2006) testou uma nova função densidade de probabilidade para avaliação de regeneração natural do Sassafrás (*Ocotea odorifera*), considerando sua distribuição diamétrica. Nesse estudo foram aplicadas as distribuições Normal, Beta, Gama, Weibull e Exponencial. Foi possível observar que apesar da grande flexibilidade de alguns dos modelos ajustados, a distribuição diamétrica do sassafrás é muito variável e somente o novo modelo apresentado foi aderente.

Weber, Arce e Péllico Netto (2009) aplicaram a distribuição de Weber, bem como as distribuições clássicas (Normal, Beta, Gama, Weibull e Exponencial) para determinar as probabilidades de ocorrência de pinhas verdes, segundo seus pesos. De acordo com o teste de Kolmogorov-Smirnov, o modelo de Weber ajustou-se satisfatoriamente aos dados coletados, fornecendo uma alternativa para estudos dessa natureza.

Observa-se que as funções densidade de probabilidade podem ser aplicadas em diversas áreas do conhecimento com ajustes satisfatórios. O uso desses modelos permite avaliações e interpretações durante todo o período de observação ou em parte dele. A forma da distribuição de frequências observada é fundamental para a escolha correta do modelo probabilístico, em que devem ser considerados os pressupostos teóricos de cada caso.

2.5.1 Distribuição Exponencial

A distribuição exponencial desempenha um papel importante na descrição de uma grande classe de fenômenos, particularmente nos assuntos da teoria da confiabilidade (MEYER, 1974). Para Johnson e Leone (1964), esse modelo pode ser considerado como um caso particular da distribuição gama, que será apresentada no tópico a seguinte. A distribuição

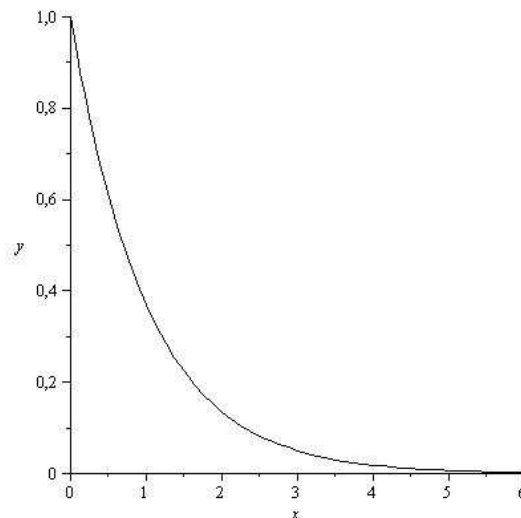
exponencial tem sido utilizada para prever o período de tempo necessário até um evento ocorrer, por exemplo, o tempo até acontecer um terremoto, ou até começar uma nova guerra ou até o telefone tocar, você atender e perceber que discaram errado (ROSS, 2000). A distribuição exponencial, de acordo com Stevenson (1981), envolve probabilidade ao longo do tempo ou da distância entre ocorrências num intervalo contínuo. Por exemplo, a exponencial é usada como modelo do tempo entre falhas de equipamento elétrico, tempo de chegada de clientes em um supermercado, tempo entre chamadas telefônicas, entre outros. Soong (1986) afirma que há uma ligação muito estrita entre as distribuições de Poisson e exponencial.

Ainda segundo Meyer (1974), a função densidade de probabilidade exponencial, com variável aleatória contínua x , em que x representa o número de chegadas no intervalo de tempo $[0, t]$, distribuído sendo a lei de Poisson, é dada por (EQUAÇÃO 04):

$$f(x) = \begin{cases} \frac{1}{\beta} \cdot e^{-\frac{x}{\beta}} & \text{para } x \geq 0, \beta > 0 \\ 0 & \text{e.o.c.} \end{cases} \quad \text{EQ. 04}$$

Diz-se que a distribuição exponencial, também chamada de exponencial negativa, em função de o expoente ser negativo, é o produto entre o inverso do parâmetro β e a exponencial do quociente entre o oposto da variável aleatória x e o parâmetro β . O gráfico do modelo é mostrado na Figura 02.

FIGURA 02 – DISTRIBUIÇÃO EXPONENCIAL, QUANDO $\beta = 1$



A média e a variância, através da expectativa matemática, são calculadas através das seguintes fórmulas, respectivamente, (EQUAÇÕES 05 e 06):

$$E(X) = \beta \quad \text{EQ. 05}$$

$$\sigma_x^2 = \beta^2 \quad \text{EQ. 06}$$

2.5.2 Distribuição Gama

Roussas (1997) apresenta a distribuição Gama de acordo com a equação 07.

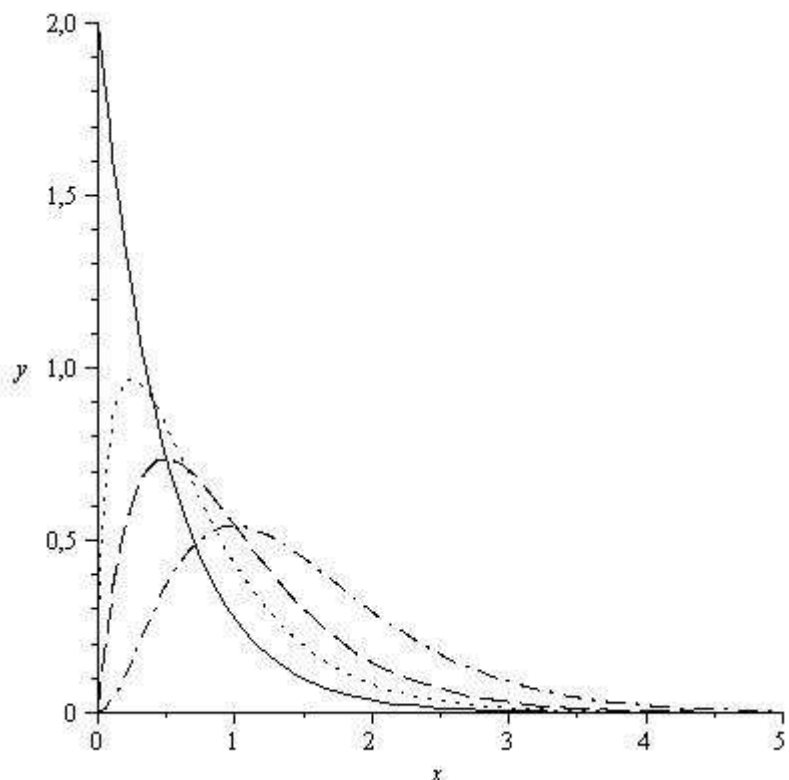
$$f(x) = \begin{cases} \frac{x^{\alpha-1} \cdot e^{-x/\beta}}{\beta^\alpha \cdot \Gamma(\alpha)} & \text{para } x \geq 0 \\ 0 & \text{e.o.c,} \end{cases} \quad \text{EQ. 07}$$

Em que α e β são valores positivos e afetam a forma e a escala, respectivamente (ROUSSAS, 1997). Diz-se que a distribuição Gama é denotada como o quociente entre: 1) o produto entre: a) a variável aleatória x elevada ao parâmetro α subtraído de uma unidade e; b) exponencial do quociente entre o oposto da variável aleatória x e o parâmetro β ; e 2) o produto entre a) o parâmetro β elevado à α e b) a função gama de α .

A Figura 03 representa a curva gerada a partir da função gama, em que β é um valor constante (0,5) e α varia entre 1 e 3.

A distribuição Gama com $\lambda = \frac{1}{2}$ e $t = \frac{n}{2}$ (n sendo um inteiro positivo) é chamada de distribuição χ_n^2 (lê-se qui-quadrado) com n graus de liberdade.

FIGURA 03 - REPRESENTAÇÃO GRÁFICA DA DISTRIBUIÇÃO GAMMA, PARA $\alpha=1$ (SÓLIDO), $\alpha=1,5$ (PONTOS), $\alpha=2$ (TRAÇOS) E $\alpha=3$ (PONTOS E TRAÇOS) E $\beta=0,5$



A média e a variância são dadas pelas equações 08 e 09, respectivamente.

$$E(X) = \alpha \cdot \beta \quad \text{EQ. 08}$$

$$\sigma^2 = \alpha \cdot \beta^2 \quad \text{EQ. 09}$$

Como a distribuição Gama é unilateral, serve frequentemente de modelo para quantidades físicas que só tomam valores positivos (SOONG, 1986). Ela apresenta uma grande variedade de formas, sendo, dessa maneira, amplamente utilizada.

2.5.3 Distribuição Beta

Pode-se dizer que a distribuição Beta é o quociente entre: 1) o produto de: a) função gama da soma de alfa e beta; b) a diferença da variável aleatória x e a , elevada a alfa subtraído de uma unidade; c) a diferença entre b e x , elevado a beta menos 1; 2) o produto entre: a) gama de alfa; b) gama de beta; e c) a diferença entre b e a elevada a soma de alfa e beta subtraída de uma unidade. Considera-se a distribuição de probabilidades beta como sendo (EQUAÇÃO 10).

$$f(x) = \begin{cases} \frac{\Gamma(\alpha + \beta)}{\Gamma(\alpha) \cdot \Gamma(\beta) \cdot (b-a)^{\alpha+\beta-1}} \cdot (x-a)^{\alpha-1} \cdot (b-x)^{\beta-1} & \text{se } a < x < b \\ 0 & \text{e.o.c.} \end{cases} \quad \text{EQ. 10}$$

A função beta pode assumir diversas formas, de acordo com os valores de α e β , como observa-se nas Figuras 04 a 07. Essa distribuição pode ser usada como modelo de fenômenos aleatórios cujo conjunto de valores possíveis é um intervalo finito $[a, b]$, podendo ser transformada para um intervalo $[0, 1]$.

FIGURA 04 – REPRESENTAÇÃO GRÁFICA DA DISTRIBUIÇÃO BETA, PARA $\alpha=1$ E $\beta=2$ (SÓLIDO), $\alpha=2$ E $\beta=1$ (PONTOS) E $\alpha=1$ E $\beta=1$ (TRAÇOS)

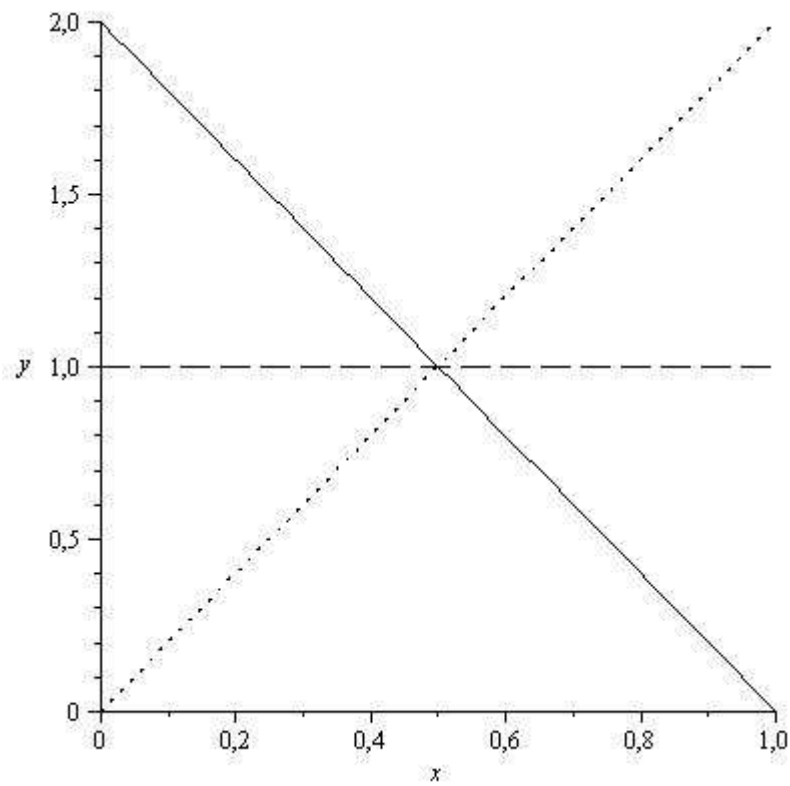


FIGURA 05 – REPRESENTAÇÃO GRÁFICA DA DISTRIBUIÇÃO BETA, PARA $\alpha=0,3$ E $\beta=0,8$ (SÓLIDO) E $\alpha=0,3$ E $\beta=0,3$ (PONTOS)

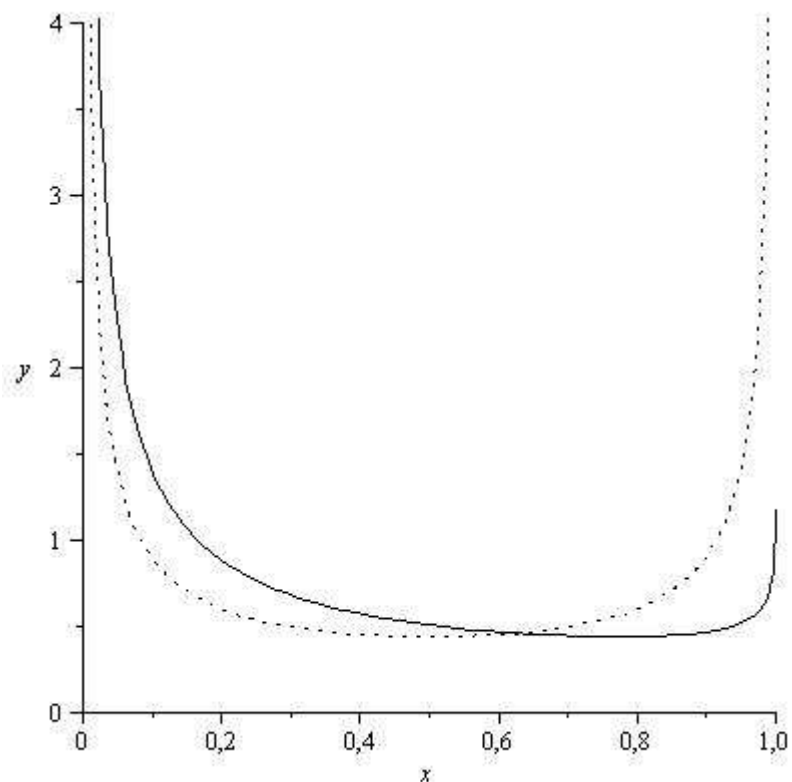


FIGURA 06 – REPRESENTAÇÃO GRÁFICA DA DISTRIBUIÇÃO BETA, PARA $\alpha=2$ E $\beta=3$ (SÓLIDO), $\alpha=5$ E $\beta=3$ (PONTOS) E $\alpha=2$ E $\beta=2$ (TRACÇOS)

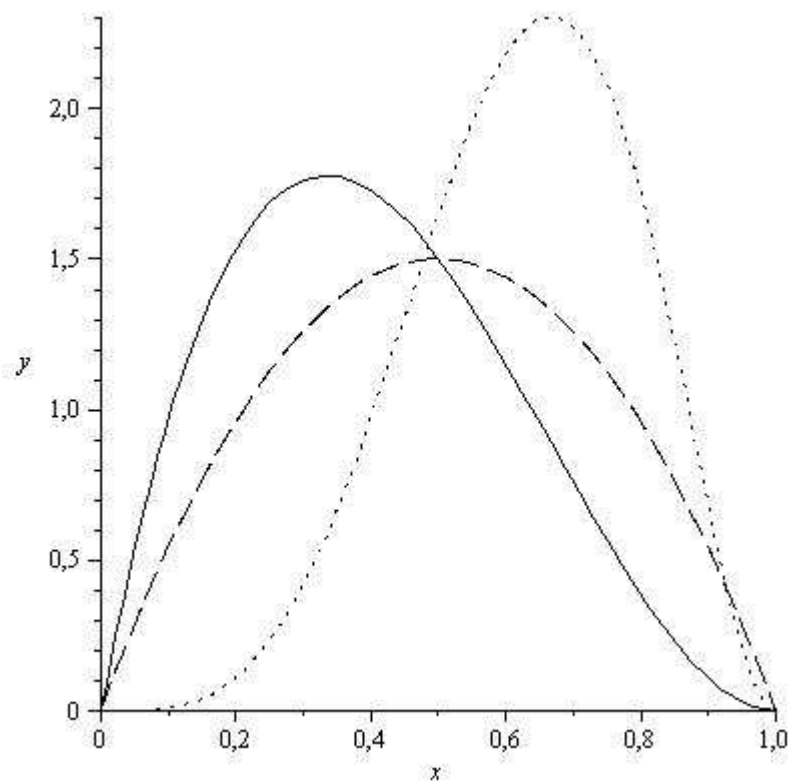
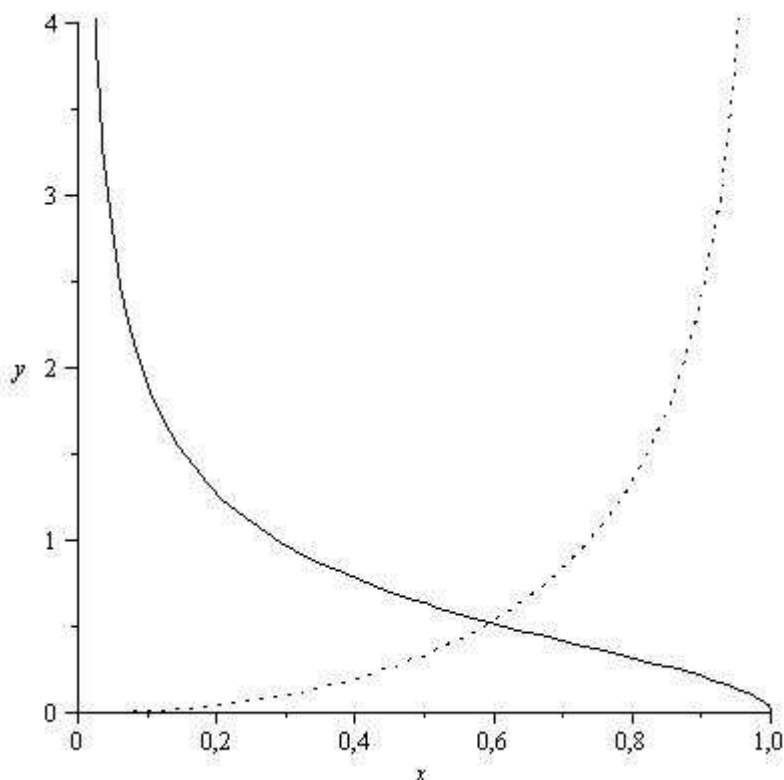


FIGURA 07 – REPRESENTAÇÃO GRÁFICA DA DISTRIBUIÇÃO BETA, PARA $\alpha = 0,5$ E $\beta = 1,5$ (SÓLIDO) E $\alpha = 3$ E $\beta = 0,5$ (PONTOS)



A média é dada pela Equação 11 e a variância pela Equação 12.

$$E(X) = \frac{\alpha}{\alpha + \beta} \quad \text{EQ. 11}$$

$$\sigma^2 = \frac{\alpha \cdot \beta}{(\alpha + \beta)^2 \cdot (\alpha + \beta + 1)} \quad \text{EQ. 12}$$

2.5.4 Distribuição Weibull

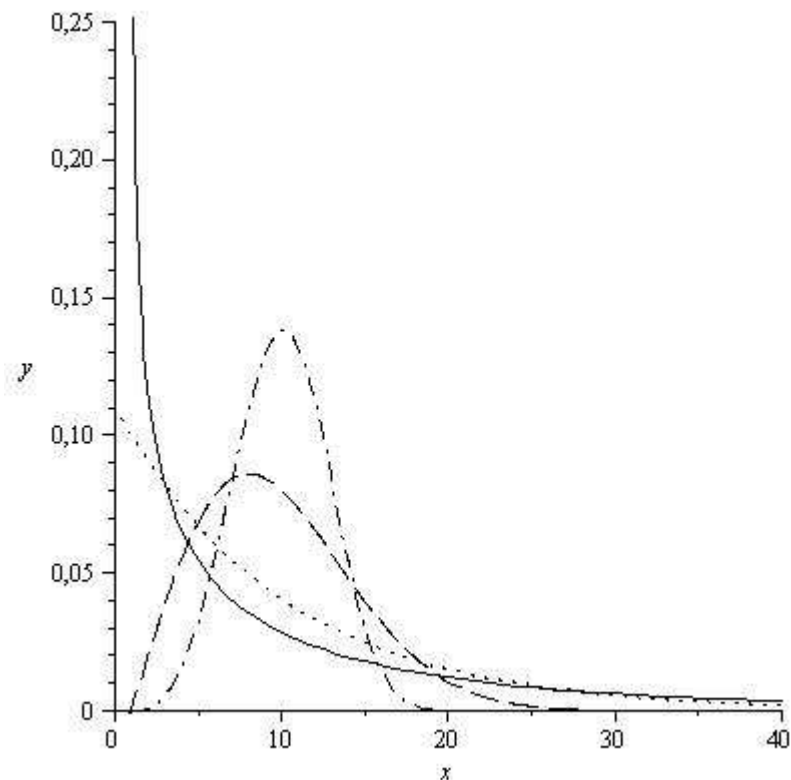
Esta distribuição é amplamente utilizada para testes de confiabilidade e vida útil (JOHNSON; LEONE, 1964). A função densidade de probabilidade da distribuição Weibull apresenta-se conforme transcrita a seguir (EQUAÇÃO 13). Pode-se afirmar que essa é mais uma das distribuições desenvolvidas a partir do conceito de exponenciais, logo, da mesma família. Certamente mais flexível que as demais, mas ainda limitada por seu fator de variação (exponencial). Pode ser descrita como sendo o produto entre: 1) o quociente entre c e b ; 2) o quociente entre a diferença entre a variável aleatória x e a , e b , elevado a c subtraído de uma unidade; e 3) a exponencial do quociente entre a diferença de x e a , e b , elevado a c , multiplicado por menos 1.

$$f(x) = \begin{cases} \frac{c}{b} \cdot \left(\frac{x-a}{b}\right)^{c-1} \cdot e^{-\left(\frac{x-a}{b}\right)^c} & \text{se } x \geq a, b > 0 \text{ e } c > 0 \\ 0 & \text{e.o.c.} \end{cases} \quad \text{EQ. 13}$$

Assim como a função beta, a Weibull também pode assumir várias formas (FIGURA 08).

A distribuição Weibull é usualmente utilizada na engenharia devido à sua versatilidade. Foi originalmente proposta para interpretação de fadiga de materiais em Stockholm por Ernst Hjalmar Waloddi Weibull, 1939, mas agora seu uso estendeu-se a muitos outros problemas de engenharia. Em particular, é comumente usada no campo dos fenômenos da vida, como uma distribuição de tempo de vida de algum objeto, particularmente quando o modelo de ligação mais fraca é apropriado para o objeto, isto é, considerar um objeto com muitas partes e supor que existirá uma falha nele. Sob essas condições, foi mostrado que a distribuição Weibull fornece uma boa aproximação com a distribuição do tempo de vida do item (ROSS, 2000).

FIGURA 08 - REPRESENTAÇÃO GRÁFICA DA DISTRIBUIÇÃO WEIBULL, COM $a = 1$, $b = 10$ e $c = 0,7$ (SÓLIDO), COM $a = 1$, $b = 10$ e $c = 1$ (PONTOS), COM $a = 1$, $b = 10$ e $c = 2$ (TRAÇOS), COM $a = 1$, $b = 10$ e $c = 3,6$ (PONTOS E TRAÇOS)



A média e a variância são dadas pelas Equações 14 e 15 respectivamente.

$$E(X) = a + b \cdot \Gamma\left(1 + \frac{1}{c}\right) \quad \text{EQ. 14}$$

$$\sigma_X^2 = b^2 \cdot \sigma_Y^2 \quad \text{EQ. 15}$$

em que σ_Y^2 é calculado pela Equação 16.

$$\sigma_Y^2 = \Gamma\left(1 + \frac{2}{c}\right) - \Gamma^2\left(1 + \frac{1}{c}\right) \quad \text{EQ. 16}$$

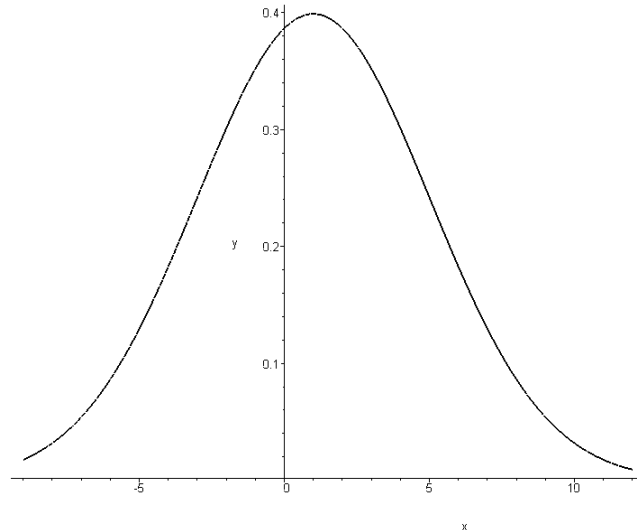
2.5.5 Distribuição Normal ou de Gauss

De acordo com Meyer (1974), a distribuição Normal, desenvolvida por Gauss, é de extrema importância no campo das probabilidades, com aplicações em diversas áreas do conhecimento. Esse modelo é definido como se segue (EQUAÇÃO 17):

$$f(x) = \begin{cases} \frac{1}{\sqrt{2\pi} \cdot \sigma} \cdot e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} & -\infty < x < +\infty \end{cases} \quad \text{EQ. 17}$$

A distribuição normal pode ser representada graficamente pela Figura 09:

FIGURA 09 – DISTRIBUIÇÃO NORMAL, $\sigma = 1$ E $\mu = 4$



As Equações 18 e 19 representam a média e a variância, respectivamente.

$$E(X) = \text{média aritmética} = \mu \quad \text{EQ. 18}$$

$$\text{var}(X) = \sigma^2 \quad \text{EQ. 19}$$

Esta distribuição apresenta as seguintes propriedades:

1. A curva normal tem forma de sino;
2. É simétrica em relação à média;
3. É definida para qualquer valor pertencente ao conjunto dos números reais, variando de infinito negativo a infinito positivo;
4. Cada distribuição normal fica completamente especificada por sua média e seu desvio padrão; há uma distribuição normal distinta para cada combinação de média e desvio padrão;
5. A área total sob a curva é considerada como 100%;
6. A área sob a curva entre dois pontos é a probabilidade de uma variável normalmente distribuída entre esses pontos;
7. Como há um número ilimitado de valores no intervalo $-\infty$ a $+\infty$, a probabilidade de uma variável aleatória distribuída normalmente assumir exatamente determinado valor é zero. Assim, as probabilidades se referem a intervalos de valores;
8. A área sob curva entre a média e um ponto arbitrário é função do número de desvios padrões entre a média e aquele ponto; e
9. Uma vez que não existe forma analítica para se calcular a probabilidade em um intervalo, na distribuição normal, tabelas são utilizadas (FREUND, WILSON, 1996). As probabilidades foram obtidas por meio de cálculo numérico.

A altura de um homem, a velocidade em qualquer direção e o erro na medida de uma quantidade física apresenta um comportamento que aproxima-se da distribuição de probabilidades normal (ROSS, 2000).

A distribuição normal foi introduzida pelo matemático francês Abraham De Moivre, em 1733. De Moivre, que usou essa distribuição para aproximar probabilidades referentes a lançamentos de moedas, a chamou de curva exponencial em forma de sino (exponential bell-shaped curve). Entretanto, foi apresentada realmente apenas em 1809, quando o famoso matemático alemão K. F. Gauss a usou como uma parte integral da sua aproximação para prever o local de entidades astronômicas. Como resultado tornou-se comum após essa data chamá-la de distribuição Gaussiana (ROSS, 2000).

Do meio para o final do século XIX, entretanto, muitos estatísticos passaram a acreditar que a maioria dos intervalos de dados teria histogramas de acordo com a forma de sino Gaussiano. Além disso, tornou-se aceitável que fosse “normal” para qualquer conjunto de dados bem comportados seguir esta curva. Como resultado, seguindo o estatístico britânico Karl Pearson, as pessoas começaram a se referir à curva de Gauss chamando-a de curva normal (ROSS, 2000).

Karl Friederich Gauss (1777 – 1855), um dos primeiros usuários da curva normal, foi um dos maiores matemáticos de todos os tempos. O conhecido matemático historiador Eric Temple Bell expressou em seu livro “Men of mathematics” (Homens da matemática), em um capítulo intitulado “The prince of mathematicians” (O príncipe dos matemáticos): “Arquimedes, Newton e Gauss; esses três estão em uma classe por si mesmos entre grandes matemáticos e não seriam pobres mortais capazes de classificá-los em ordem de mérito. Os três iniciaram tendências em matemática pura e aplicada. Arquimedes estimou sua matemática pura mais que suas aplicações; Newton pareceu ter encontrado a justificativa chefe para suas invenções matemáticas nos usos científicos nos quais ele colocou; enquanto Gauss declarou que para tudo trabalha-se no lado puro e aplicado” (ROSS, 2000).

2.5.6 Distribuição de Quadros

Eduardo Quadros da Silva, em 2003, defendeu sua tese intitulada nova função densidade de probabilidade aplicável à ciência florestal, na qual foram sugeridas novas opções para o ajuste de distribuições de probabilidades. O modelo da distribuição polinomial (assim chamada pelo autor na época) foi desenvolvido para aplicação a variáveis tomadas em árvores de floresta natural, podendo ser estendido também a povoamentos e situações em que os modelos existentes não demonstram aderência (SILVA, 2003). A função é definida genericamente como consta na equação 20:

$$f(x) = \frac{1}{k} \begin{cases} c_1 x^d & \text{se } 0 < x < l_1 \\ a_1 x^n + a_2 x^{n-1} + \dots + a_m & \text{se } l_1 \leq x \leq l_2 \\ \frac{c_2}{x^h} & \text{se } x > l_2 \\ 0 & \text{e.o.c} \end{cases} \quad \text{EQ. 20}$$

Em que n , d e h são inteiros positivos, $a_1, a_2, \dots, a_m, c_1, c_2$ são números reais,

$k = \int_0^{\infty} f(x) dx$, l_1 é o limite superior da classe onde será ajustada a função $c_1 x^d$ e l_2 é o limite

superior da última classe onde o polinômio produz um bom ajuste. Tal modelo foi inicialmente aplicado à distribuição de alturas do Jequitibá-Rosa, em 12 parcelas permanentes localizadas no município de Cássia, MG, sendo considerado aderente aos dados.

2.5.7 Funções Spline

Em uma Função Spline cada ponto amostrado é chamado de nó. De acordo com Ruggiero (1998) uma spline cúbica, $S_3(x)$ é uma função polinomial truncada, contínua, onde cada parte, $s_k(x)$ é um polinômio de terceiro grau no intervalo $[x_{k-1}, x_k]$, $k = 1, 2, \dots, n$. $S_3(x)$ tem a primeira e segunda derivadas contínuas, o que a torna sem bicos e nem troque abruptamente de curvatura nos nós.

Supondo que $f(x)$ esteja tabelada nos pontos x_i , $i = 0, 1, 2, \dots, n$, a função $S_3(x)$ é chamada spline cúbica interpolante de $f(x)$ nos nós x_i , $i = 0, \dots, n$ se existem n polinômios de grau 3, $s_k(x)$, $k = 1, \dots, n$ tais que:

i) $S_3(x) = s_k(x)$ para $x \in [x_{k-1}, x_k]$, $k = 1, \dots, n$

ii) $S_3(x_i) = f(x_i)$, $i = 0, 1, \dots, n$

iii) $s_k(x_k) = s_{k+1}(x_k)$, $k = 1, 2, \dots, (n-1)$

iv) $s'_k(x_k) = s'_{k+1}(x_k)$, $k = 1, 2, \dots, (n-1)$

v) $s''_k(x_k) = s''_{k+1}(x_k)$, $k = 1, 2, \dots, (n-1)$

Para simplicidade de notação, pode ser escrito $s_k(x) = a_k(x - x_k)^3 + b_k(x - x_k)^2 + c_k(x - x_k) + d_k$, $k = 1, 2, \dots, n$.

2.6 BIOLOGIA

Em uma floresta, o aumento em tamanho ocorrido em um intervalo de tempo é definido como crescimento e é dado pela atividade das árvores vivas, porém sua somatória não representa o crescimento da floresta como um todo, pois árvores morrem, são cortadas ou recrutadas. O crescimento das árvores, mais convenientemente medido pelo incremento da circunferência ou diâmetro à altura do peito é de grande interesse da silvicultura e do manejo florestal (GOMIDE, 1997).

2.7 *Bertholletia excelsa*

Bertholletia excelsa Bonpl, popularmente conhecida como Castanheira, pertence à família Lecythidaceae, é uma das espécies arbóreas de maior valor ecológico e econômico na floresta amazônica (PERES, BAIDER, 1997).

Segundo o relatório da FAO (1986) *Bertholletia excelsa* é geralmente encontrada em florestas tropicais quentes e úmidas. É comum também em regiões quentes de florestas transicionais que recebem precipitação anual de apenas 1500 mm, período seco bem definido, de 4 a 5 meses. Sua área de ocorrência vai desde o extremo sul das guianas até o Alto Beni - 14° de latitude Sul no sul do Mato Grosso (MULLER et al., 1980). Seu limite leste encontra-

se a noroeste do Maranhão, bacia do Rio Gurupi, estando ausente em toda a parte ocidental da hiléia em ocorrência espontânea (DUCKE; BLACK, 1954). É extremamente bem adaptada à solos muito argilosos e latossolos de baixa fertilidade da região, mas não é muito tolerante a áreas alagadas. Naturalmente ocorre em agrupamentos relativamente extensos (popularmente conhecida como “bolas” ou “reboleiras”) denominados castanhais, grupos dispersos de 50 a 100 árvores, separados por um quilômetro ou mais de distância, cuja densidade aproximada de 0,1 a 2,5 árvores por hectare foi observada em inventários. Em um castanhal pode ser encontrada de 1 a 15 castanheiras por hectare, sendo que a média geral para a Amazônia é de 1,5 árvore/ha (SALOMÃO, 1991; SALOMÃO et al., 2002). Em função dessa sua dinâmica de estabelecimento, surgem pontos de máximo local, ou modas. O banco de mudas do sub-bosque é impedido de se desenvolver em função do sombreamento ocasionado pelo dossel superior. É necessário um distúrbio (como dinâmica de clareira) para que essas plântulas invistam em altura.

A castanheira é uma árvore de grande porte, que pode atingir até 50 metros de altura (sendo reto por aproximadamente 20 m ou mais) e 2 metros de diâmetro na base (LORENZI, 1992). Segundo Lorenzi (1992), a madeira da Castanheira, potencialmente, pode ser usada na construção civil e naval, bem como para esteios e obras externas. De acordo com o relatório da FAO (1986), a madeira é de boa qualidade, com nervura retilínea, fácil de trabalhar, com densidade de 0,6 a 0,75 g/cm³, textura média e brilhosa, com cerne marrom avermelhado claro, cuja seiva é acinzentada. É excelente para acabamentos finos. O corte das árvores tem sido intensificado em muitas áreas, apesar das leis que o proíbem, em função do alto valor do seu fruto. O pericarpo do fruto é usado como combustível e algumas vezes esculpido para venda em lojas de artesanato.

As sementes levam entre 6 meses até 2 ou 3 anos para germinar, a menos que seja procedida uma boa escarificação ou remoção das conchas, podendo o tempo de germinação ser reduzido para 20 dias. Mudas levemente sombreadas crescem cerca de 50 a 80 cm no primeiro ano, podendo atingir até 1 metro na perpendicular. Essa taxa de crescimento permanece dessa forma nos primeiros 5 anos, tendo, após esse período, uma diminuição em sua velocidade de ascensão. Árvores de 40 anos têm cerca de 20m e 50 cm de DAP, enquanto que aquelas com 60 anos têm 60 m e 80 cm de altura e DAP, respectivamente. A frutificação inicia entre os 12 e 16 anos nos exemplares presentes na floresta nativa, enquanto que em povoamentos bem manejados podem chegar aos 8 anos de idade, entretanto já foi apresentado um caso de frutificação aos 4 anos e meio. Essa variação na produção anual de sementes pode acarretar dois ou mais pontos de máximo local.

Outro fator que contribui para a ocorrência de mais modas é a exploração seletiva para determinados fins. Em alguns casos, classes diamétricas inteiras são cortadas.

2.8 *Araucaria angustifolia*

Araucaria angustifolia Bertol. (Kuntze) é classificada como gimnosperma arbórea perenifólia, podendo atingir uma altura de 40 m e até 120 cm de DAP. Pode ser observado um tronco reto, quase cilíndrico, principal motivo da viabilidade econômica para sua exploração (VELOSO; RANGEL-FILHO; LIMA, 1991). Apesar dos valores descritos por esses autores, cabe salientar que já foram encontrados indivíduos maiores, com DAP de 170 cm (no município de General Carneiro) e 197 cm (em Castro). Já na serra catarinense foi encontrado um pinheiro com aproximadamente 311 cm de DAP.

Segundo Machado e Siqueira (1980) a área original de ocorrência da espécie no Brasil era de 182.295 km², ou 22,50% da superfície total dos estados de São Paulo, Paraná, Santa Catarina e Rio Grande do Sul. A Araucária ocorria naturalmente no sul do Brasil, grande parte dos estados do Paraná, Santa Catarina e Rio Grande do Sul (HUECK, 1972). Atualmente, estima-se que os remanescentes ocupem entre 1% a 4% da área original (LIMA; CAPOBIANCO 1997, GUERRA *et al.* 2002).

Hertel (1980) classifica em quatro as fases distintas de desenvolvimento após a germinação da *Araucaria angustifolia*, conforme descrito a seguir.

- Puladendro, enquanto plântula.
- Tirodendro, a planta atingindo a maturidade. Nessa fase sua copa assume uma forma típica, cônica, como o modo dilucional baseado em “verticilos” (pseudoverticilos) ou “coroas”, cujo raio diminui harmônica e proporcionalmente à medida que a altura de inserção aumenta, resultando em uma perfeita forma cônica ou, em silhueta, a figura de um triângulo isósceles com o vértice voltado para cima. Esta forma, piramidada, permanece até o início do processo de encopamento, quando retarda-se o crescimento em altura do eixo vertical e modifica-se a silhueta da árvore.
- Ferodendro, período em que a árvore é produtiva e que ocorre a completa abscisão dos galhos mais baixos e a copa assume a forma de uma umbela.
- Senadendro, em que o indivíduo protela sua morte e sua forma assemelha-se a de uma taça ou cálice.

De acordo com Nogueira (1989), quando a árvore atinge esta fase é um sinal de declínio de sua vitalidade. Esta conformação arquitetural é abreviada quando a planta cresce

em sítios pobres, podendo ocorrer mesmo antes dos vinte anos de idade, ainda que a altura da árvore seja relativamente baixa.

O crescimento da araucária está diretamente correlacionado com a qualidade do sítio (DE HOOG, 1981). Entretanto, a velocidade de crescimento varia com o decorrer dos anos. No início o crescimento é lento e aumenta a partir do terceiro ano. À medida que a competição entre as árvores se faz presente, o ritmo de crescimento a aumenta, desacelerando na fase adulta e permanecendo constante até sua fase senil, quando a dinâmica para e o indivíduo sobrevive até o fim de sua vida (NUTTO, 2001). Esse período completo pode durar mais de 300 anos (CARVALHO, 1994). Carvalho (1994) afirma que existem plantios que alcançam 26 m³/ha/ano, enquanto outros mostram uma produção vinte vezes menor.

A araucária, em condições naturais, na fase jovem, se comporta como uma espécie pioneira (KLEIN, 1960). Na fase adulta se comporta como secundária longeva, formando o estrato emergente da Floresta Ombrófila Mista (KLEIN, 1960; CARVALHO, 1994). De acordo com Carvalho (1994) a araucária tolera sombreamento quando jovem, porém não tolera sombreamento lateral quando plantada na capoeira alta e quando adulta é heliófila. Segundo Inoue, Galvão e Torres (1979) a araucária é tolerante a níveis moderados de sombra na fase juvenil e quando a espécie é exposta em intensidade luminosas na ordem de 25% da luminosidade plena há uma maior atividade fotossintética, o que indica, aparentemente, melhores ritmos de crescimento. A formação do lenho inicial pode estar relacionada com a maior incidência luminosa sobre a araucária, comparando árvores de borda e no interior da floresta (SANTAROSA et al., 2007).

3. MATERIAIS E MÉTODOS

Para o desenvolvimento do modelo foram utilizadas funções compostas, conforme apresentado na revisão bibliográfica. As características desejadas foram testadas por meio de seu ajuste a um conjunto de dados simulados, apresentados no tópico desenvolvimento do modelo. O método utilizado para ajustar o modelo está descrito no presente capítulo, bem como as Funções Densidade de Probabilidade com as quais a função ajustada foi comparada. Informações mais detalhadas sobre essas estão também apresentadas no presente tópico. Os procedimentos de comparação entre as funções ajustadas estão apresentados a seguir, a saber: erro padrão da estimativa recalculado, teste de Kolmogorov-Smirnov e análise de resíduos.

3.1 MÉDIA E VARIÂNCIA DE UMA FUNÇÃO DENSIDADE DE PROBABILIDADE

A média e a variância podem ser determinadas por meio da Função Geratriz dos Momentos ou da Expectativa Matemática. Essa última será a utilizada no presente estudo.

3.2 EXPECTATIVA MATEMÁTICA

Seja X uma variável aleatória contínua com função densidade de probabilidade f . Define-se o valor esperado de X como $E(X)$, em que $E(X)$ é a média, conforme Equação 21 a seguir (GUIDORIZZI, 1994):

$$\mu = E(X) = \int_a^b x \cdot f(x) dx \quad \text{EQ. 21}$$

A variância pode ser escrita da seguinte forma (EQUAÇÃO 22), após estabelecer um passo intermediário (EQUAÇÃO 23):

$$E(X^2) = \int_a^b x^2 \cdot f(x) dx \quad \text{EQ. 22}$$

$$\sigma_x^2 = E(X^2) - [E(X)]^2 \quad \text{EQ. 23}$$

3.3 DISTRIBUIÇÃO EXPONENCIAL

Segundo Meyer (1974), a distribuição exponencial pode ser definida conforme equação 04.

$$f(x) = \begin{cases} \frac{1}{\beta} \cdot e^{-\frac{x}{\beta}} & \text{para } x \geq 0, \beta > 0 \\ 0 & \text{e.o.c.} \end{cases} \quad \text{EQ. 04}$$

Em cada distribuição deve ser assegurado que esta é probabilística, ou seja, que a probabilidade total no seu domínio deve ser igual a 1. Como a distribuição exponencial é contínua, a probabilidade de ocorrência de um determinado evento descrito por ela pode ser calculado por meio da integral da função no intervalo de seu domínio. Para se obter a probabilidade total basta integrar a função de 0 até infinito (domínio) (EQUAÇÃO 24).

$$\int_0^{\infty} f(x)dx = \int_0^{\infty} \frac{1}{\beta} e^{-\frac{x}{\beta}} dx = 1 \quad \text{EQ. 24}$$

Conforme enunciado anteriormente, as expressões da média e da variância podem ser obtidas por meio de expectativa matemática, integrando a função multiplicada por x e x^2 . Dessa forma obtém-se a média (EQUAÇÃO 25) e a variância (EQUAÇÃO 26).

$$E(X) = \int_0^{\infty} x \cdot f(x)dx = \int_0^{\infty} x \cdot \frac{1}{\beta} e^{-\frac{x}{\beta}} dx = \beta \quad \text{EQ. 25}$$

$$\begin{aligned} \sigma_x^2 &= E(X^2) - [E(X)]^2 = \int_0^{\infty} x^2 \cdot f(x)dx - [E(X)]^2 = \\ &= \int_0^{\infty} x^2 \cdot \frac{1}{\beta} e^{-\frac{x}{\beta}} dx - \beta^2 = 2\beta^2 - \beta^2 = \beta^2 \end{aligned} \quad \text{EQ. 26}$$

O ajuste é realizado utilizando-se o método dos momentos, ou seja, é utilizado como coeficiente β a média dos dados observados. Uma vez que a exponencial fica totalmente determinada pelo parâmetro β , esse é o único cálculo necessário para o ajuste.

3.4 DISTRIBUIÇÃO GAMA

A distribuição Gama é definida a seguir (EQUAÇÃO 07):

$$f(x) = \begin{cases} \frac{x^{\alpha-1} \cdot e^{-x/\beta}}{\beta^{\alpha} \cdot \Gamma(\alpha)} & \text{para } x \geq 0 \\ 0 & \text{e.o.c.} \end{cases} \quad \text{EQ. 07}$$

A distribuição Gama possui alguns termos que não são habituais à maioria das pessoas. Por esse motivo serão apresentadas as definições da função gama $\Gamma(\alpha)$ para valores de α inteiro e não inteiro. Para valores de α inteiro, a função gama é igual fatorial de α subtraído de uma unidade (EQUAÇÃO 27). Quando o valor de α não pertencer ao conjunto dos números inteiros positivos, o valor da função gama para um valor z adicionado de uma unidade será o resultado da integral imprópria no intervalo de 0 a mais infinito, do produto de t elevado a z e exponencial do oposto de t em função da variável t (EQUAÇÃO 28).

$$\Gamma(\alpha) = (\alpha - 1)! \quad \text{EQ. 27}$$

$$\Gamma(z + 1) = \int_0^{\infty} t^z \cdot e^{-t} dt \quad \text{EQ. 28}$$

Salienta-se, neste momento, a diferença da função gama e a função densidade de probabilidade Gama. A primeira é descrita pelas equações 27 e 28, enquanto que a segunda pela 07. A primeira expressão é um fatorial e a segunda a probabilidade de ocorrência de um determinado evento que tenha a tendência da distribuição Gama. Para que seja considerada uma função densidade de probabilidade é necessário que a área entre o gráfico da função e o eixo x seja igual a 1, determinando, assim, que sua probabilidade total não seja maior que 100%. Essa verificação é realizada por meio da EQUAÇÃO 29, na qual se observa que a premissa é verificada.

$$\int_0^{\infty} \frac{x^{\alpha-1} \cdot e^{-x/\beta}}{\beta^{\alpha} \cdot \Gamma(\alpha)} dx = 1 \quad \text{EQ. 29}$$

Outros dois fatores importantes nas distribuições probabilísticas, conforme mencionado anteriormente, são as expressões analíticas para se determinar a média e a variância. Utilizando-se a expectativa matemática, pode-se facilmente chegar às expressões da média (EQUAÇÃO 30) e variância (EQUAÇÃO 31).

$$\mu = \int_0^{\infty} x \frac{x^{\alpha-1} \cdot e^{-x/\beta}}{\beta^{\alpha} \cdot \Gamma(\alpha)} dx = \alpha\beta \quad \text{EQ. 30}$$

$$\sigma_x^2 = E(X^2) - [E(X)]^2 = \int_0^{\infty} x^2 \frac{x^{\alpha-1} \cdot e^{-x/\beta}}{\beta^{\alpha} \cdot \Gamma(\alpha)} dx - [E(X)]^2 = \alpha\beta^2 \quad \text{EQ. 31}$$

Os ajustes da distribuição Gama serão realizados pelo método dos momentos, utilizando-se as expressões 32 e 33, apresentadas anteriormente. Como a distribuição Gama possui dois parâmetros (α e β), que estão relacionados à média e à variância, que podem ser calculadas empiricamente, basta estabelecer as duas igualdades (EQUAÇÕES 32 e 33). Como se tem duas incógnitas e duas equações, o sistema é possível e determinado.

$$\bar{X} = \alpha\beta \quad \text{EQ. 32}$$

$$\sigma_x^2 = \alpha \cdot \beta^2 \quad \text{EQ. 33}$$

3.5 DISTRIBUIÇÃO BETA

A FDP Beta pode ser apresentada de acordo com a equação 10. Essa é a maneira expandida da distribuição, a qual pode ser aplicada a qualquer conjunto de dados, independente do intervalo. Ressalta-se que a distribuição, como foi concebida, ou seja, sem os valores de a e b , estendia-se de 0 a 1.

$$f(x) = \begin{cases} \frac{\Gamma(\alpha + \beta)}{\Gamma(\alpha) \cdot \Gamma(\beta) \cdot (b-a)^{\alpha+\beta-1}} \cdot (x-a)^{\alpha-1} \cdot (b-x)^{\beta-1} & \text{se } a < x < b \\ 0 & \text{e.o.c.} \end{cases} \quad \text{EQ. 10}$$

Deve-se, como nas demais distribuições, verificar se a soma das probabilidades no intervalo de seu domínio é igual a 1, conforme EQUAÇÃO 34.

$$\int_a^b \frac{\Gamma(\alpha + \beta)}{\Gamma(\alpha) \cdot \Gamma(\beta) \cdot (b-a)^{\alpha+\beta-1}} \cdot (x-a)^{\alpha-1} \cdot (b-x)^{\beta-1} dx = 1 \quad \text{EQ. 34}$$

É possível, pela manipulação da expectativa matemática, resolver a integral que gera o primeiro e segundo momento, obtendo-se dessa forma as equações 35 e 36.

$$\mu = \int_a^b x \frac{\Gamma(\alpha + \beta)}{\Gamma(\alpha) \cdot \Gamma(\beta) \cdot (b-a)^{\alpha+\beta-1}} \cdot (x-a)^{\alpha-1} \cdot (b-x)^{\beta-1} dx = \frac{\alpha}{\alpha + \beta} \quad \text{EQ. 35}$$

$$\begin{aligned} \sigma_x^2 &= E(X^2) - [E(X)]^2 = \\ &\int_a^b x^2 \frac{\Gamma(\alpha + \beta)}{\Gamma(\alpha) \cdot \Gamma(\beta) \cdot (b-a)^{\alpha+\beta-1}} \cdot (x-a)^{\alpha-1} \cdot (b-x)^{\beta-1} dx - [E(X)]^2 = \\ &\frac{\alpha\beta}{(\alpha + \beta)^2 (\alpha + \beta + 1)} \end{aligned} \quad \text{EQ. 36}$$

Como a média e a variância empírica aproximam-se dos parâmetros populacionais, podem ser utilizadas as estimativas da média e da variância para ajustar o modelo. Da mesma forma que na Gama, a distribuição beta possui dois coeficientes a serem ajustados, alfa e beta, os outros dois (a e b) são os limites inferior e superior, respectivamente (EQUAÇÕES 37 e 38). Esse método é também conhecido como método dos momentos.

$$\bar{X} = \frac{\alpha}{\alpha + \beta} \quad \text{EQ. 37}$$

$$\sigma_x^2 = \frac{\alpha\beta}{(\alpha + \beta)^2 (\alpha + \beta + 1)} \quad \text{EQ. 38}$$

3.6 DISTRIBUIÇÃO WEIBULL

A distribuição de Weibull pode ser definida de acordo com a EQUAÇÃO 14.

$$f(x) = \begin{cases} \frac{c}{b} \cdot \left(\frac{x-a}{b}\right)^{c-1} \cdot e^{-\left(\frac{x-a}{b}\right)^c} & \text{se } x \geq a, b > 0 \text{ e } c > 0 \\ 0 & \text{e.o.c.} \end{cases} \quad \text{EQ. 14}$$

Há necessidade de se averiguar se essa função é probabilística ou não, de acordo com as premissas já citadas. Estabelecendo a integral da função no intervalo de seu domínio, ou seja, de 0 a infinito positivo, obtém-se a EQUAÇÃO 39.

$$\int_0^{\infty} \frac{c}{b} \cdot \left(\frac{x-a}{b}\right)^{c-1} \cdot e^{-\left(\frac{x-a}{b}\right)^c} dx = 1 \quad \text{EQ. 39}$$

Os cálculos da média e variância são obtidos por meio das EQUAÇÕES 40, 41 e 42.

$$\mu = \int_0^{\infty} x \frac{c}{b} \cdot \left(\frac{x-a}{b}\right)^{c-1} \cdot e^{-\left(\frac{x-a}{b}\right)^c} dx = a + b \cdot \Gamma\left(1 + \frac{1}{c}\right) \quad \text{EQ. 40}$$

$$\sigma_x^2 = b^2 \cdot \sigma_y^2 \quad \text{EQ. 41}$$

em que σ_y^2 é calculado pela Equação 42.

$$\sigma_y^2 = \Gamma\left(1 + \frac{2}{c}\right) - \Gamma^2\left(1 + \frac{1}{c}\right) \quad \text{EQ. 42}$$

A FDP de Weibull será ajustada pelo método de regressão não-linear, procedimento de Marquardt, conforme será apresentado mais adiante.

3.7 DISTRIBUIÇÃO NORMAL

A distribuição de Gauss é comumente apresentada de acordo com a EQUAÇÃO 17. Verifica-se em todas as distribuições que há uma partícula nas definições, em que se atribui o valor zero e.o.c. (em outros casos), exceto nessa. Isso deve-se ao fato de que a distribuição normal tem como domínio o conjunto dos números reais, não havendo possibilidades de outros casos.

$$f(x) = \begin{cases} \frac{1}{\sqrt{2\pi} \cdot \sigma} \cdot e^{-\frac{1}{2} \left(\frac{x-\mu}{\sigma}\right)^2} & -\infty < x < +\infty \end{cases} \quad \text{EQ. 17}$$

As probabilidades da distribuição normal ou Gaussiana, como é amplamente conhecida, apenas podem ser calculadas por meio de cálculo numérico, donde se conclui a respeito de sua probabilidade total (EQUAÇÃO 43).

$$\int_{-\infty}^{+\infty} \frac{1}{\sqrt{2\pi} \cdot \sigma} \cdot e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} dx = 1 \quad \text{EQ. 43}$$

A FDP normal inaugurou uma nova era na construção de funções probabilísticas, em que os seus coeficientes assumem valores iguais à média e à variância (EQUAÇÕES 44 e 45). Esses mesmos valores são usados no ajuste da função aos dados coletados, admitindo-se que a média e variância observadas aproximam-se da real.

$$\bar{x} \cong \mu \quad \text{EQ. 44}$$

$$s_x^2 \cong \sigma_x^2 \quad \text{EQ. 45}$$

3.8 DISTRIBUIÇÃO DE QUADROS

A distribuição de Quadros é apresentada de acordo com a EQUAÇÃO 20.

$$f(x) = \frac{1}{k} \begin{cases} c_1 x^d & \text{se } 0 < x < l_1 \\ a_1 x^n + a_2 x^{n-1} + \dots + a_m & \text{se } l_1 \leq x \leq l_2 \\ \frac{c_2}{x^h} & \text{se } x > l_2 \\ 0 & \text{e.o.c} \end{cases} \quad \text{EQ. 20}$$

Em que n , d e h são inteiros positivos, $a_1, a_2, \dots, a_m, c_1, c_2$ são números reais,

$k = \int_0^{\infty} f(x) dx$, l_1 é o limite superior da classe na qual será ajustada a função $c_1 x^d$ e l_2 é o

limite superior da última classe na qual o polinômio produz um bom ajuste.

Não há para essa distribuição fórmulas analíticas prontas. Logo, é preciso proceder o ajuste da função aos dados empíricos para, posteriormente, calcular a média e a variância estimadas pelo modelo.

Etapas para ajuste

1. Ajuste de um polinômio que mais se aproxime dos dados;
2. Esboçar o gráfico do polinômio produzido;
3. Desconsiderar as classes em que o polinômio assume valores negativos ou contraria a tendência dos dados observados;
4. Elaborar funções para as partes inicial e final;
5. Formar a função definida pelas 3 funções anteriormente mencionadas;
6. Calcular k, que é a área entre a função e o eixo X no intervalo de 0 a ∞ ; e

7. Multiplicar a função por $\frac{1}{k}$, a fim de se obter a função probabilística.

Pela forma como foi concebida, conforme item 7 anterior, essa função é sempre probabilística, respeitando as premissas de FDPs.

3.9 FUNÇÕES SPLINE

É importante salientar que as funções Spline não foram originalmente concebidas para serem probabilísticas, dessa forma, serão apresentados os processos de transformação em FDP. A função Spline é definida conforme equação 46.

$$s_k(x) = a_k(x - x_k)^3 + b_k(x - x_k)^2 + c_k(x - x_k) + d_k, k = 1, 2, \dots, n. \quad \text{EQ. 46}$$

Calcula-se a constante L , resultado da integral da Função Spline no seu domínio e multiplica-se seu inverso pela função ajustada. Com isso obtém-se a FDP a partir da Spline Cúbica.

A média é obtida integrando-se o produto da FDP por x no seu intervalo. A diferença entre a integral do produto entre a FDP e x^2 , e a média, representa a variância.

3.10 MODELO MULTIMODAL PROPOSTO

O modelo multimodal proposto foi ajustado por meio de regressão não linear, segundo o procedimento de Marquardt, pelo método de mínimos quadrados, com múltiplas iterações, de acordo com Marquardt (1963). Esse método é realizado computacionalmente por meio do software Table curve 2d.

A maioria dos algoritmos para estimar parâmetros não lineares pelo método de mínimos quadrados centraram-se sobre qualquer uma de duas aproximações. Por um lado, o modelo pode ser expandido como uma série de Taylor e as correções dos parâmetros calculadas a cada iteração na suposição de linearidades locais. Por outro lado, várias modificações do método gradiente estão sendo usadas. O método de Marquardt utiliza o método de máxima verossimilhança, que executa uma ótima interpolação entre o método da série de Taylor e o método gradiente. A interpolação é baseada em máxima verossimilhança, na qual a série de Taylor truncada fornece uma representação adequada de um modelo não-linear (MARQUARDT, 1963).

Seja $E(y) = f(x_1, x_2, \dots, x_m; \beta_1, \beta_2, \dots, \beta_k) = f(x, \beta)$ o modelo a ser ajustado aos dados, em que x_1, x_2, \dots, x_m são as variáveis independentes, $\beta_1, \beta_2, \dots, \beta_k$ são os k parâmetros e $E(y)$ o valor esperado da variável y . Denotando-se os pontos amostrados como

$(Y_i, X_{1i}, X_{2i}, \dots, X_{ni})$, $i = 1, 2, \dots, n$, o problema consiste em computar essas estimativas dos parâmetros que minimizem $\phi = \sum_{i=1}^n [Y_i - \hat{Y}_i]^2 = \|Y - \hat{Y}\|^2$, em que \hat{Y}_i é o valor de y predito por $E(y)$ no i -ésimo ponto amostrado (MARQUARDT, 1963).

3.11 AVALIAÇÃO DO AJUSTE

A aderência dos modelos ajustados será comparada por meio do erro padrão da estimativa recalculado em percentagem; pelo teste de Kolmogorov-Smirnov e pela análise dos resíduos.

3.11.1 Erro padrão da estimativa recalculado

O erro padrão da estimativa recalculado pode ser calculado utilizando a Equação 47:

$$S_{yx \text{ recalculado}} = \sqrt{\frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{n - p}} \quad \text{EQ. 47}$$

O erro padrão da estimativa em percentagem é dado pela razão entre o erro padrão da estimativa e a média (EQUAÇÃO 48):

$$S_{yx} \% = \frac{S_{yx}}{\bar{x}} \cdot 100 \quad \text{EQ. 48}$$

Em que:

y_i é frequência observada;

\hat{y}_i é a frequência estimada pelos modelos;

\bar{x} é a média aritmética;

n é o número de dados; e

p o número de parâmetros estimados pelo modelo em questão.

3.11.2 Kolmogorov e Smirnov

O teste de aderência de Kolmogorov e Smirnov consiste em comparar a máxima diferença entre a frequência observada e a frequência estimada, dividida pelo número de observações, e pode ser obtido pela Equação 49:

$$d_{calc} = \frac{(\max|F_o(X) - F_e(X)|)}{n} \quad \text{EQ. 49}$$

Em que $F_o(X)$ é a frequência observada acumulada, $F_e(X)$ é a frequência esperada acumulada estimada pelo modelo e n o número de observações.

Se d_{calc} for menor que o valor limite tabelado, aceita-se o ajuste. Se n for menor que 50, para $\alpha = 5\%$, e $\alpha = 1\%$, determina-se o valor crítico por meio das Equações 50 e 51.

$$\alpha = 5\% \rightarrow 1,36 \cdot \sqrt{\frac{n_1 + n_2}{n_1 \cdot n_2}} \quad \text{EQ. 50}$$

$$\alpha = 1\% \rightarrow 1,63 \cdot \sqrt{\frac{n_1 + n_2}{n_1 \cdot n_2}} \quad \text{EQ. 51}$$

4. DESENVOLVIMENTO DO MODELO

O modelo foi desenvolvido visando atender uma necessidade específica: flexibilidade capaz de representar duas ou mais modas. A seguir será apresentada uma breve introdução sobre as distribuições atualmente utilizadas e posteriormente as fases que culminaram no modelo proposto.

4.1 INTRODUÇÃO

O surgimento de funções densidade de probabilidade trouxe um novo horizonte no campo dos estudos de eventos aleatórios para variáveis contínuas, tendo sido desenvolvido de forma brilhante. Esse conhecimento viabilizou o desenvolvimento de diversas áreas do conhecimento, principalmente nas áreas concernentes à biologia. Tendo sido originalmente desenvolvida para estudos de eventos determinísticos em jogos de azar, rapidamente se verificou a grande aplicabilidade desse conceito de grande valia.

As distribuições clássicas têm inestimável valor nesse âmbito, dentre eles se destacam a Curva de Gauss, Exponencial, Gama, Beta e Weibull. A história mostra que a evolução no processo de construção de funções probabilísticas é oriunda das necessidades que surgem hodiernamente. Nesse aspecto cabe destacar acerca de três pontos, a saber: flexibilidade, integrabilidade e interpretação dos coeficientes.

A principal distribuição contínua é notadamente a Normal. Ela é definida brilhantemente, de tal sorte que fica totalmente determinada e ajustada conhecendo-se apenas a média e o desvio padrão. Salienta-se ainda que seus coeficientes são denotados exatamente como μ e σ , média e desvio padrão. Entretanto, é importante ressaltar que esses são parâmetros populacionais. Nesse momento se faz prudente lembrar a diferença primordial entre coeficiente e parâmetro. Coeficientes são as constantes dos modelos que devem ser ajustados por meio de métodos estatísticos. Eles determinam as diferentes conformações que as distribuições podem assumir. Parâmetros são valores indicativos de uma determinada população, entre eles se destacam a média, o desvio padrão, a mediana, a moda, entre outros.

Retomando à questão da Distribuição Normal, sua moda, mediana e média são coincidentes, ou seja, estão localizadas no mesmo ponto em X. Entretanto, essa característica, que tão brilhantemente foi descoberta por Abraham De Moivre, encerra um dos pontos dignos de nota. O primeiro é que a função se adequa à maioria dos conjuntos de dados, quando considera-se a população como um todo. O segundo ponto importante é que quando se está lidando com amostras pequenas, em que a distribuição amostral dos dados é assimétrica, a

Curva Normal tem um desempenho aquém do esperado, não sendo, muitas vezes, aplicável. Com relação à sua integrabilidade, ela é complexa, entretanto viável e cujas probabilidades estão totalmente tabeladas, sendo amplamente divulgada e constante em livros de estatística (tabela z), não constituindo, atualmente, um problema.

Existe um conjunto amplo de situações em que os dados distribuem-se de forma decrescente, tais como os concernentes à radioatividade, intervalos de chegada ou saída de veículos, questões sobre resistência de materiais, regeneração natural de espécies arbóreas, entre outros. Verificou-se que uma distribuição contínua muito simples resolveria essa situação: a exponencial. Essa é uma função cujo ajuste é muito fácil, em função de ter apenas um coeficiente a ser determinado. O mesmo problema enfrentado pela Curva Normal está implícito nessa distribuição: a flexibilidade.

Em consequência da situação supracitada, novas distribuições foram desenvolvidas visando equacionar este problema: a flexibilidade. Dentre as mais utilizadas estão a Gama, Beta e Weibull. Outras distribuições foram desenvolvidas mais recentemente, entre elas a distribuição de Quadros. Essa tem especial interesse, uma vez que foi pensada em função de um problema florestal, como a que está sendo discutida na presente pesquisa. Apesar de todas as vantagens apresentadas por Quadros (2003), aquela função é truncada, ou seja, composta por mais de uma sentença. Cabe ressaltar a capacidade do modelo de Quadros ter mais de uma moda.

Posteriormente, Weber (2006) pesquisou acerca de uma distribuição que tivesse uma flexibilidade capaz de se ajustar aos mais variados conjuntos de dados. Durante sua pesquisa verificou-se que havia um termo no denominador da função que “gerava” uma moda. Sobre esse denominador foi despendido um tempo longo, a fim de se resolver o problema da divisão por zero e de aumentar substancialmente a flexibilidade do modelo. A título de inovação, foi repetido o termo que “gerava” uma nova moda, tendo surgido, por consequência, outra moda. No momento daquela pesquisa sabia-se da importância dessa descoberta para o campo dos estudos de distribuições probabilísticas. Contudo, naquele momento, o conjunto de dados com o qual se estava trabalhando apresentava distribuição unimodal, ou seja, com apenas uma moda.

O interesse em estudar com maior rigor o termo gerador de modas se deu em função do conhecimento por parte do Professor Doutor Sylvio Péllico Netto, de espécies Amazônicas com distribuições diamétricas multimodais. Serão apresentados alguns dos testes realizados e a sequência de pensamentos, desde a concepção até a conclusão.

4.2 ETAPAS DO DESENVOLVIMENTO

4.3 CONJUNTO DE DADOS

O conjunto de dados utilizados para o desenvolvimento do modelo foi simulado, conforme tabela a seguir. É possível observar que existem duas modas nessa distribuição de dados: em $X = 11,75$ e $X = 19,25$. Entende-se por moda o valor X no qual ocorre a maior frequência. O termo multimodal significa ocorrência de dois ou mais pontos de máximos locais.

TABELA 01 – Dados simulados utilizados para o desenvolvimento do modelo.

X	Y
0,00	0
4,25	12
6,75	26
9,25	48
11,75	52
14,25	33
16,75	49
19,25	100
21,75	67
24,25	20
26,75	10

4.4 A FUNÇÃO ORIGINAL

A função original utilizada foi aquela desenvolvida por Weber (2006), conforme EQUAÇÃO 52 a seguir.

$$f(x) = \frac{x^a}{(b + c \cdot x)^d} \quad \text{EQ. 52}$$

Nota-se que o termo $(b + c \cdot x)^d$ é o termo chamado “gerador de modas”. Essa é uma função unimodal. É de interesse na presente pesquisa encontrar distribuições com mais de uma moda.

4.4.1 Teste 1

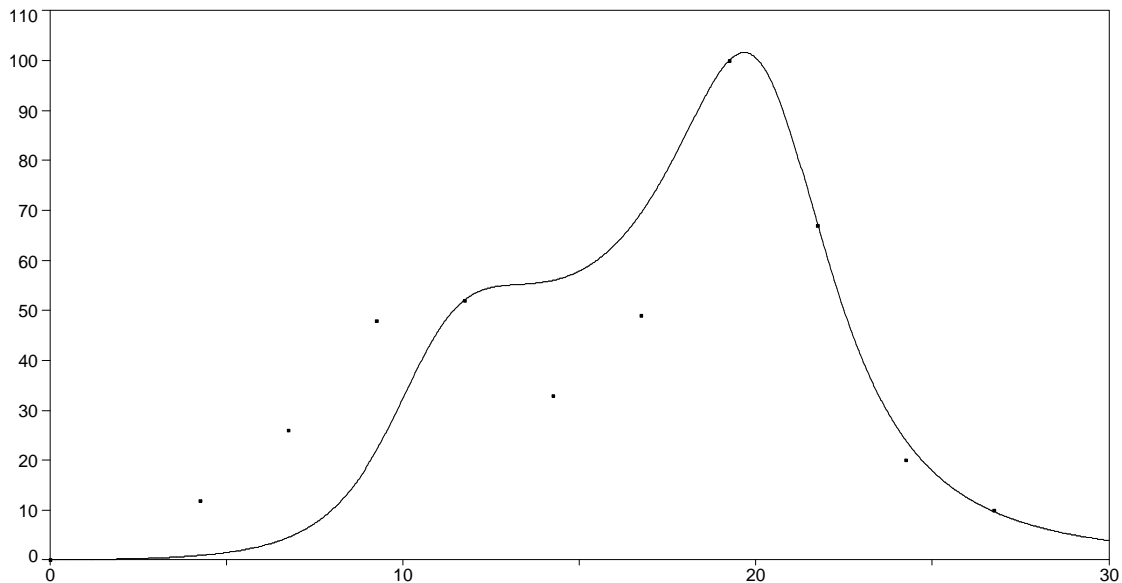
O primeiro teste realizado foi com a função denotada pela equação 53. Nela se flexibilizou o numerador, utilizando-se uma função do segundo grau sem o termo independente. Com isso se determinou que a distribuição passaria pela origem, pois quando $X = 0$ a função assume valor igual a zero. Outro ponto importante refere-se ao termo d , pois ele garante que o denominador jamais se anulará, resultando em um ponto de descontinuidade.

Além disso, verifica-se a existência dos valores M1 e M2, que são as duas modas assumidas pelo modelo.

$$f(x) = \frac{ax^2 + bx}{c(x - M1)^2(x - M2)^2 + d} \quad \text{EQ. 53}$$

O resultado do ajuste pode ser observado pela figura 10. Por meio dos valores do coeficiente de determinação múltiplo ajustado, R^2 adj, (0,52) verifica-se que o ajuste foi apenas razoável. Pode-se observar que surgiram, como era desejado, duas modas, mas que não foi satisfatório o ajuste, pois a distribuição não conseguiu moldar-se aos dados nas classes iniciais.

FIGURA 10 – Resultado do primeiro teste



Os testes que se seguiram visavam exclusivamente aumentar a flexibilidade. Cabe ressaltar que para que esse modelo fosse escolhido, seria necessário testar sua integrabilidade.

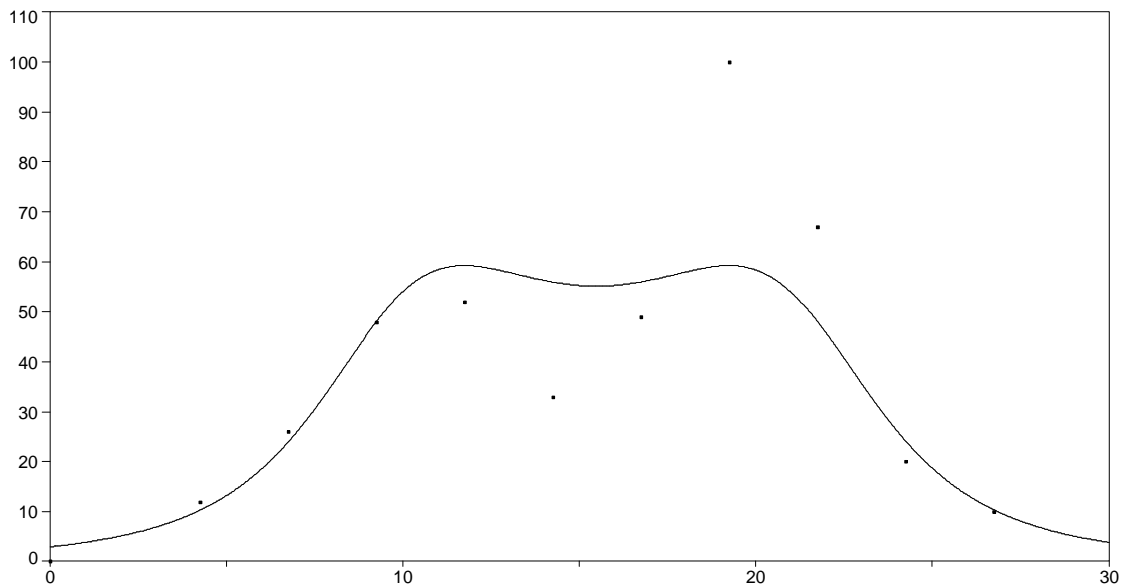
4.4.2 Teste 2

Conforme pode-se observar no segundo teste (EQUAÇÃO 54), a tentativa realizada consistiu em mudar a característica do numerador, uma vez que aparentemente o problema a ser resolvido consistia em aumentar a flexibilidade desse componente. É importante salientar que esse modelo não passa pela origem, o que não é desejável.

$$f(x) = \frac{(x^2 - a)^2}{b(x - M1)^2(x - M2)^2 + c} \quad \text{EQ. 54}$$

Por meio do figura 11 observa-se algo interessante, uma vez que a amplitude das modas foi próxima, entretanto isso não era de interesse. Salienta-se que as classes iniciais e finais do conjunto de dados foram muito bem explicadas, mas o modelo não teve capacidade de ajustar-se às classes intermediárias. Nesse caso, apesar de o software *Table curve* apresentar os valores de R^2 adj, não será feita uma discussão acerca dele, afinal a tendência do modelo, bem como sua flexibilidade, são mais importantes que estatísticas de ajuste nesse momento. Esse modelo, em relação ao número de coeficientes, é melhor que o primeiro, mas ainda não cabe essa discussão, pois o resultado do ajuste não foi satisfatório.

FIGURA 11 – Resultado do segundo teste

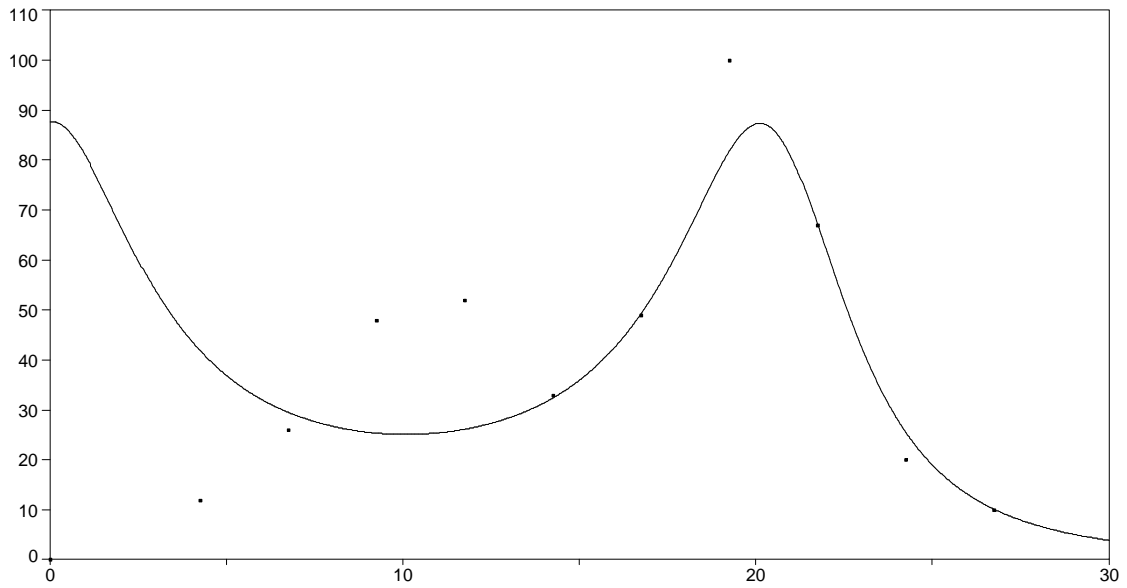


4.4.3 Teste 3

No teste 3 (EQUAÇÃO 55) os valores de M1 e M2 não são mais as modas, mas são valores que facilitam para o ajuste por parte do programa. Além disso, esperava-se melhorar o ajuste, entretanto o resultado foi muito pior, conforme se observa no figura 12. Isso é decorrente do método de ajuste, que priorizou as classes finais e deslocou a primeira moda para o eixo Y, o que não era desejado. Este modelo também não passa pela origem.

$$f(x) = \frac{(x^2 - a)^2}{(bx - M1)^2 (cx - M2)^2 + d} \quad \text{EQ. 55}$$

FIGURA 12 - Resultado do terceiro teste

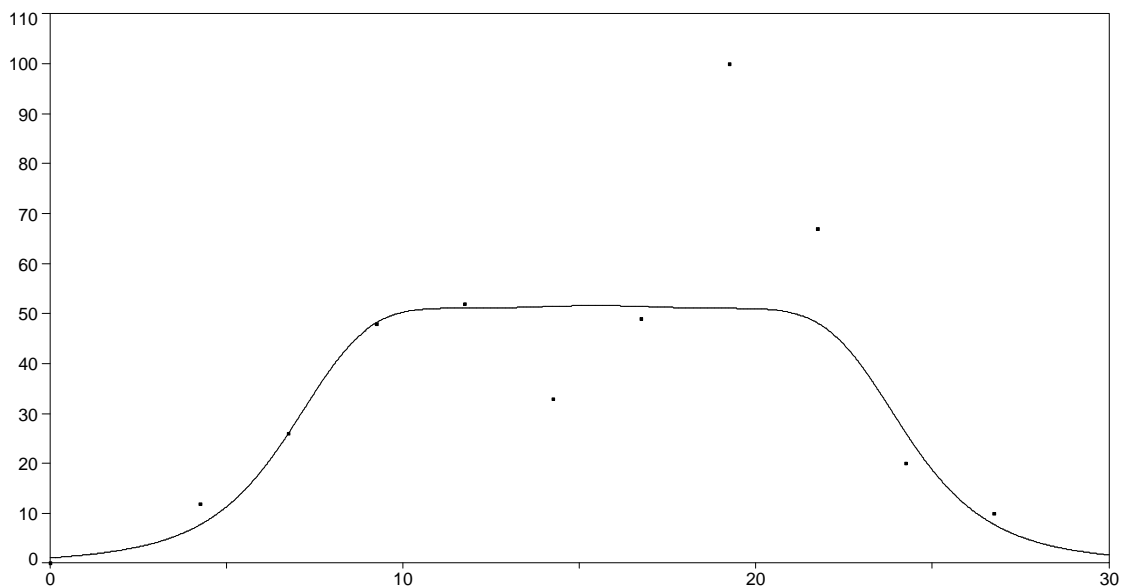


4.4.4 Teste 4

Como introduzir um novo coeficiente em cada gerador de modas não surtiu o efeito desejado, esses foram retirados (EQUAÇÃO 56). Ainda buscando a flexibilização, os termos do denominador foram elevados ao cubo, portanto aumentados, mas o efeito foi uma piora significativa no ajuste (FIGURA 13). Mesmo o fato de não passar pela origem não ajudou a tendência.

$$f(x) = \frac{(x^2 - a)^2}{b(x - M1)^3(x - M2)^3 + c} \quad \text{EQ. 56}$$

FIGURA 13 – Resultado do quarto teste



Estão sendo apresentados aqui alguns dos testes realizados. Cabe frisar que outros ajustes intermediários foram feitos, mas apenas aqueles dignos de nota estão sendo apresentados, uma vez que agregaram algum conhecimento e entendimento para a elaboração de um modelo com ajuste suficientemente satisfatório, sendo cabível propô-lo a nível científico.

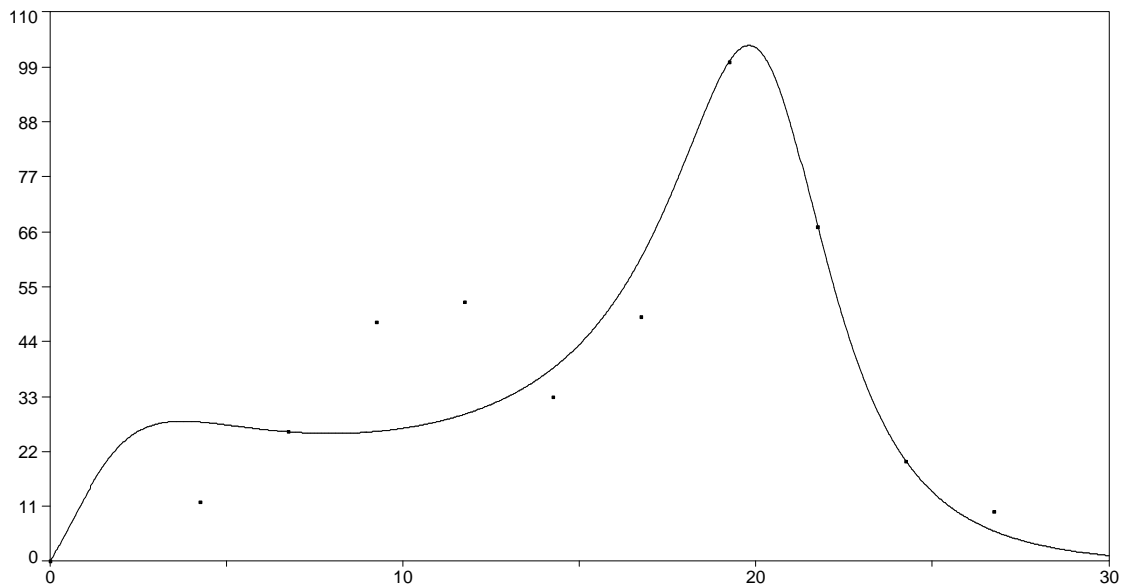
4.4.5 Teste 5

Esse é mais um modelo da família de funções que não passa pela origem (EQUAÇÃO 57). Tentou-se, novamente, flexibilizar o denominar e o numerador, mas o resultado não foi tão bom quanto se imaginava, apesar de ter sido muito melhor que os anteriores. Deve-se notar que há sempre, em todos os modelos, um termo independente no denominador. Ele é necessário para evitar pontos de descontinuidade. Isso está sendo enfatizado, pois é uma premissa do modelo que já está definida até o momento.

$$f(x) = \frac{ax^2 + bx + c}{(dx - M1)^2 (ex - M2)^2 + f} \quad \text{EQ. 57}$$

Observa-se no figura 14 que as duas modas foram mais evidentes e que as classes finais foram bem representadas, mas as inferiores e intermediárias não. Nota-se que a função passa muito próxima da origem.

FIGURA 14 – Resultado do quinto teste

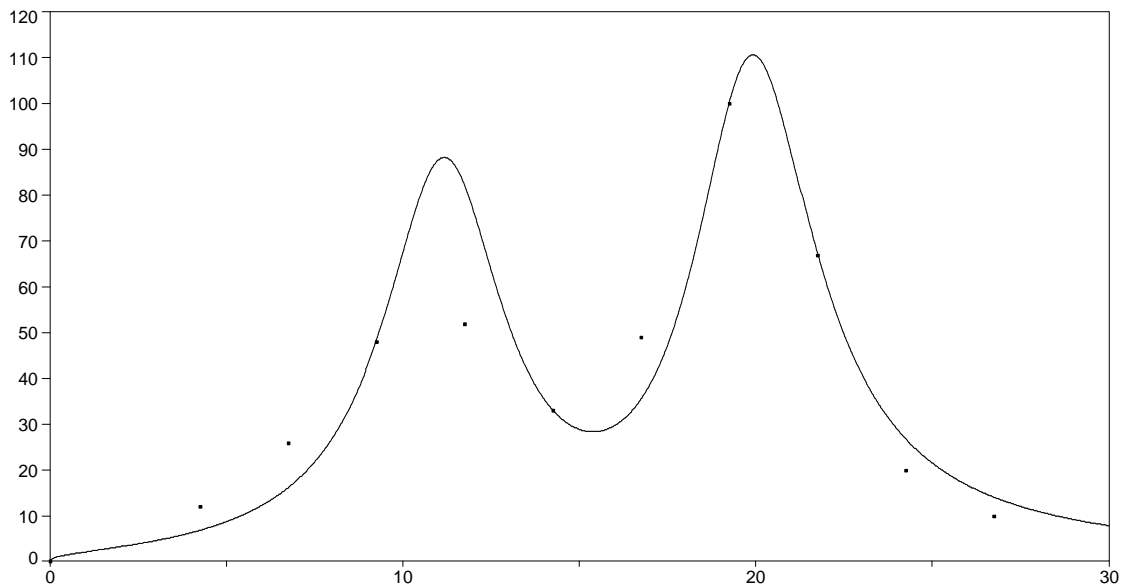


4.4.6 Teste 6

Ao se adicionar o produto da soma das partes geradoras de moda no numerador com X elevado a um coeficiente e retirando-se os coeficientes de dentro dos parênteses (EQUAÇÃO 58) o ajuste melhora significativamente (FIGURA 15). Esse modelo passa obrigatoriamente pela origem.

$$f(x) = \frac{x^a \left((x - M1)^2 + (x - M2)^2 \right)}{b(x - M1)^2 (x - M2)^2 + c} \quad \text{EQ. 58}$$

FIGURA 15 – Resultado do sexto teste



Nesse caso observa-se que a primeira moda superestima a frequência. Assim, o objetivo é reduzir a amplitude da primeira moda.

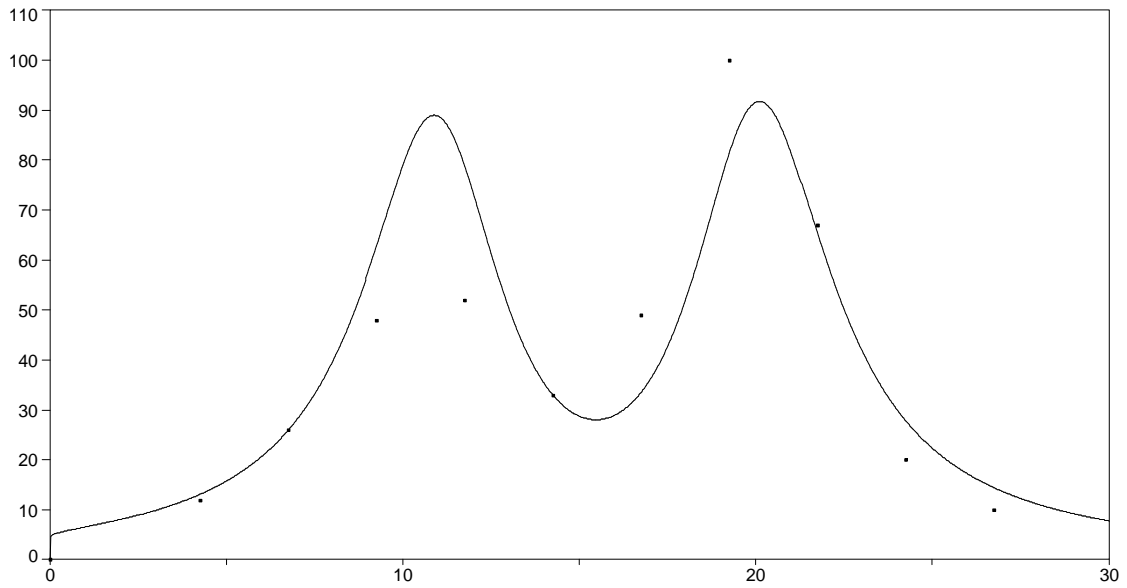
4.4.7 Teste 7

Fixando-se o valor de a no modelo anterior (EQUAÇÃO 59) tentou-se reduzir a primeira moda, entretanto o que se conseguiu foi obter duas modas de mesma amplitude, o que não é desejável, logo o modelo foi rejeitado.

$$f(x) = \frac{x^{0,05} \left((x - M1)^2 + (x - M2)^2 \right)}{a(x - M1)^2 (x - M2)^2 + b} \quad \text{EQ. 59}$$

O resultado do ajuste da Equação 59 está apresentado no figura 16.

FIGURA 16 – Resultado do sétimo teste

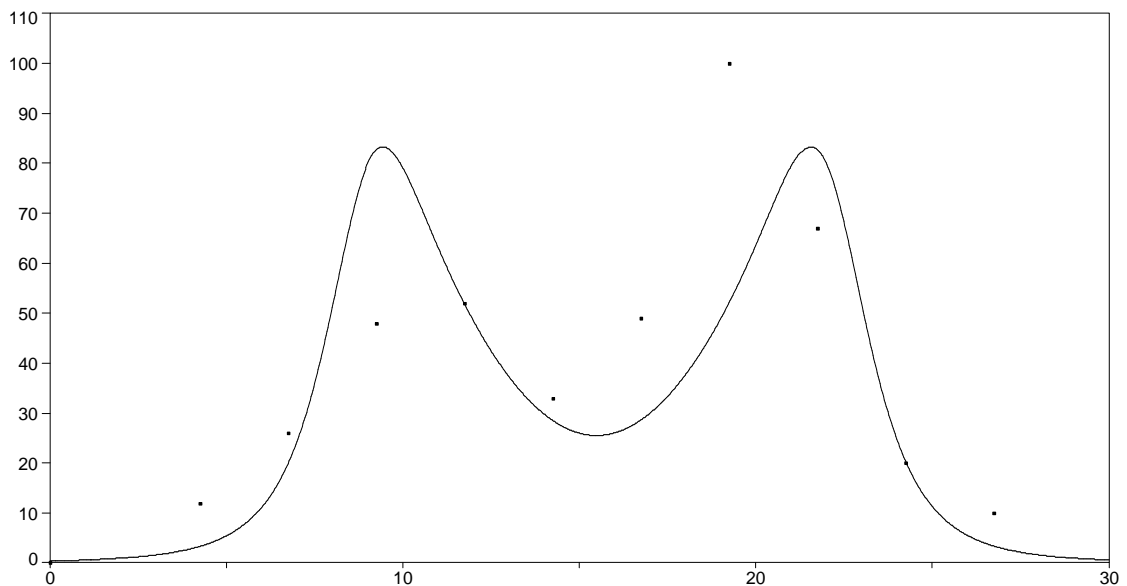


4.4.8 Teste 8

Retirar o termo X elevado a um valor constante e elevar os termos geradores de moda a quarta potência (EQUAÇÃO 60) não foi eficiente (FIGURA 17).

$$f(x) = \frac{a((x - M1)^2 + (x - M2)^2)}{(x - M1)^4(x - M2)^4 + b} \quad \text{EQ. 60}$$

FIGURA 17 – Resultado do oitavo teste

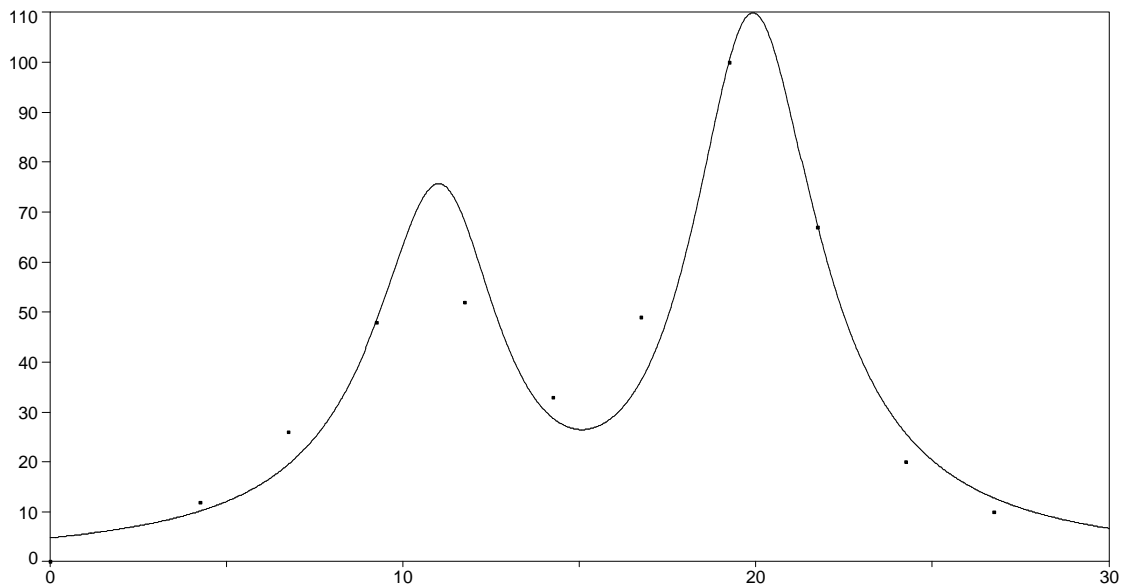


4.4.9 Teste 9

Um resultado satisfatório (FIGURA 18) foi obtido quando se multiplicou a soma dos geradores de moda no numerador por coeficientes (EQUAÇÃO 61). Observa-se que esse modelo não passa pela origem, mas é interessante proceder os testes adicionais.

$$f(x) = \frac{a(x - M1)^2 + b(x - M2)^2}{(x - M1)^2(x - M2)^2 + c} \quad \text{EQ. 61}$$

FIGURA 18 – Resultado do nono teste



É importante observar as estatísticas de ajuste desse modelo, em que o R^2 adj foi igual a 0,91, considerado bom. Pelo gráfico é possível verificar que o modelo teve flexibilidade suficiente para ajustar-se aos dados ao longo de toda a distribuição.

A partir do resultado obtido, procedeu-se o teste de integrabilidade. Aqui cabe lembrar que não há necessidade do modelo ser derivável, mas sim integrável. Isso deve-se a três motivos:

1. A integral da função no seu domínio representa a área entre a curva e o eixo x . O inverso desse resultado deve ser multiplicado pela função a fim de se obter uma nova função densidade de probabilidade, uma vez que a integral em seu domínio valerá, agora, 1;
2. A integral do produto da FDP por x no intervalo representará a média da distribuição;
3. A integral do produto da FDP por x^2 no intervalo representará o segundo momento. A diferença entre o resultado do segundo momento e o quadrado do primeiro (média) será igual à variância; e

Essa é a razão da necessidade de que as funções sejam integráveis e que a diferença entre o grau do numerador e denominador seja maior que 2. Cabe ressaltar que o domínio

dessas distribuições varia de zero até mais infinito. Logo, a função deve tender a zero quando x tende a infinito.

Integrando a função em relação a x em seu intervalo (EQUAÇÃO 61) verifica-se que essa não é integrável.

A integral da função ajustada em relação a x em seu intervalo (EQUAÇÃO 62) é igual a 1.156, o que é um bom resultado. Já a integral da função multiplicada por x (EQUAÇÃO 63) tende para infinito, logo esse modelo deve ser descartado.

$$\int_0^{\infty} \frac{628(x-11,75)^2 + 430(x-19,25)^2}{(x-11,75)^2(x-19,25)^2 + 353} dx \quad \text{EQ. 62}$$

$$\int_0^{\infty} \frac{x(628(x-11,75)^2 + 430(x-19,25)^2)}{(x-11,75)^2(x-19,25)^2 + 353} dx \quad \text{EQ. 63}$$

4.4.10 Teste 10

O raciocínio utilizado nesse teste foi o seguinte: no denominador, utilizar o gerador de modas, mas usando-se x elevado ao quadrado. Permanecer com o termo independente no denominador para, como dito anteriormente, não proporcionar um ponto de descontinuidade na função. No numerador foi utilizada uma parte da diferencial do denominador, excluindo-se algumas constantes.

$$f(x) = \frac{x \cdot (a \cdot x^2 - b) + x \cdot (c \cdot x^2 - d)}{e \cdot (ax^2 - b)^2 \cdot (cx^2 - d)^2 + f} \quad \text{EQ. 64}$$

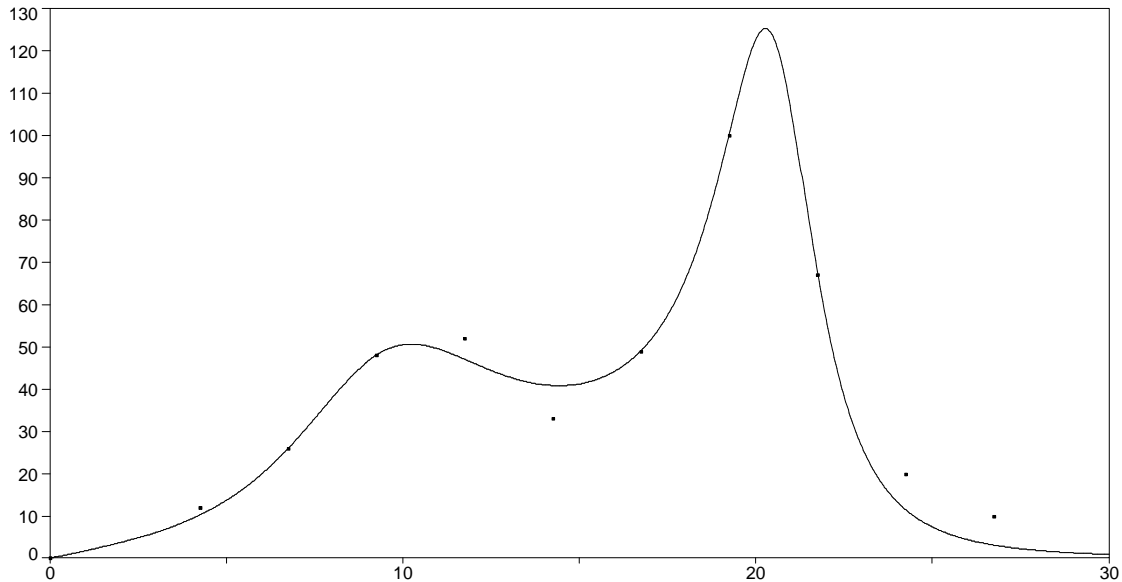
Ao se fazer algumas análises matemáticas sobre o modelo observa-se que:

1. Passa pela origem. Condição garantida pelo x multiplicando ambos os termos no numerador;
2. Não possui ponto de descontinuidade. Fato determinado pelo acréscimo da constante f no denominador;
3. Integrabilidade. Com relação a esse ponto, três informações devem ser lembradas:
 - a. A diferença entre o grau do numerador e do denominador é 5;
 - b. A função é assintótica no eixo x ; e
 - c. A função possui integral definida no intervalo de seu domínio, mesmo quando multiplicada por x ou por x^2 , a fim de se obter o primeiro e segundo momento.

Executando-se o ajuste dessa função aos dados simulados encontrou-se a equação 65, que gerou o figura 19.

$$f(x) = \frac{x(0.5339x^2 - 45.0117) + (-0.4476x^2 + 181.9487)x}{27.7982 - 8,1 \cdot 10^{-7}(0.5339x^2 - 45.0117)^2(-0.4476x^2 + 181.9487)^2} \quad \text{EQ. 65}$$

FIGURA 19 – Resultado do décimo teste



Quanto às estatísticas, verifica-se que o ajuste foi satisfatório, R^2_{adj} igual a 0,9378. O software utilizou apenas 7 iterações. A integral da função em seu domínio ($0 \leq x < \infty$) foi igual a 1037,16. Multiplicando-se a função pelo inverso dessa constante obtém-se a FDP. A integral da FDP em seu intervalo é igual a 1, como se desejava, uma vez que a área sob a curva de uma FDP seja igual à probabilidade total de ocorrência de um evento, ou seja, 100%. A média empírica estimada a partir dos dados resultou em 16,08, enquanto que a obtida pelo modelo foi de 18,86 e, logo, muito próximas. A média da FDP foi obtida pela integral da FDP por x em seu domínio. A variância empírica resultou em 31,06, enquanto que a obtida pelo modelo foi de 30,77, também muito próximas. Esses resultados comprovam a eficiência do modelo, tendo sido, dessa maneira, escolhido para ser proposto como nova função densidade de probabilidade bimodal. Foi realizado o teste não paramétrico de Kolmogorov-Smirnov para avaliar a aderência do modelo aos dados. O valor calculado foi de 0,03, enquanto que o tabelado foi de 0,09, logo, o modelo foi considerado aderente e, conseqüentemente, ajustou-se bem aos dados amostrais.

Após alguns testes foi verificado que o numerador é um fator limitante do modelo, ou seja, a estrutura, como apresentada, a torna menos flexível, embora em muitos casos não comprometa significativamente a qualidade do ajuste. Durante a aplicação do modelo a dados reais com três modas observou-se um decréscimo abrupto nas classes diamétricas superiores.

Por esse motivo o modelo foi alterado, conforme equação 66, retirando-se a relação existente entre o numerador e o denominador. No numerador foi utilizado um polinômio de grau quatro, enquanto que no denominador foi introduzida uma função de grau 8. Com isso foi estabelecida uma diferença de 4 unidades entre o grau do numerador e do denominador. Essa flexibilização do numerador gera um problema de integração, que é resolvida pelo cálculo numérico.

$$f(x) = \frac{g \cdot x^4 + h \cdot x^3 + i \cdot x^2 + j \cdot x}{e \cdot (a \cdot x^2 - b)^2 \cdot (c \cdot x^2 - d)^2 + f} \quad \text{EQ. 66}$$

Os coeficientes do numerador podem ser omitidos, dependendo do caso em questão, desde que não reduza significativamente a qualidade do ajuste.

Caso se deseje obter distribuições multimodais, basta proceder da mesma forma que na bimodal. Para uma distribuição de dados com M modas devem ser acrescentados M geradores de modas no denominador. Cada gerador de modas tem grau 4, logo o denominador terá grau $4 \cdot M$. Dessa maneira, no numerador haverá um polinômio de grau menor ou igual a $4 \cdot M - 4$. Por exemplo, o grau do polinômio no numerador estará entre 0 e 4 para uma distribuição unimodal, menor que 8 para uma bimodal, menor que 12 para trimodal, e assim sucessivamente.

5. RESULTADOS E DISCUSSÕES

5.1 INTRODUÇÃO

O modelo multimodal, como está sendo chamado na presente pesquisa, foi inicialmente concebido no campo teórico, em que foram consideradas duas etapas fundamentais. A primeira parte consistiu em determinar famílias de funções que tivessem algumas propriedades primordiais, como serão brevemente apresentadas. Essas características estão intimamente relacionadas com as premissas de distribuições probabilísticas, bem como vinculadas ao propósito maior do presente trabalho: determinar um modelo multimodal, que atenda às necessidades da Ciência Florestal.

As propriedades do modelo proposto são:

1. Ter domínio no campo dos números reais positivos, ou seja, ser válida para qualquer valor real entre 0 (inclusive) e infinito positivo. Isso se deve ao fato de não haver probabilidade de ocorrência negativa.
2. Interceptar o eixo das abscissas na origem. Essa condição está relacionada com as premissas de distribuições probabilísticas;
3. Ser assintótica ao eixo das abscissas. Ou seja, as frequências absolutas das classes diamétricas devem tender para zero quando X tender a infinito, em que X é o DAP. Essa característica está associada a duas pressuposições. A primeira refere-se às premissas de distribuições de probabilidade, uma vez que a área entre a curva que representa o modelo e o eixo das abscissas, no seu domínio, deve ser igual a 1 (100%). A segunda diz respeito à Ciência Florestal, ao confirmar que sempre haverá chance de se encontrar indivíduos em qualquer classe diamétrica, inclusive com DAP muito grande. O conceito de muito grande pode variar de espécie para espécie;
4. Ser integrável. Essa é uma propriedade vinculada a uma questão matemática. Por meio de integrais será calculada a área entre a curva e o eixo das abscissas. Isso garante que será possível determinar probabilidades em qualquer intervalo do domínio da função ajustada;
5. Ter a mesma quantidade de modas quanto às observadas empiricamente. Esse é um fator que vem sendo cada vez mais importante na ciência florestal, tanto para povoamentos florestais com espécies plantadas, quanto para florestas naturais. No primeiro caso, a necessidade encontra-se em aumentar a precisão dos resultados, gerando valores de produção mais acurados e permitindo um planejamento mais dinâmico e acurado. Já no caso de espécies nativas o interesse é outro e está vinculado à preservação e conservação das espécies; e

6. Ter flexibilidade. O modelo proposto não só deve ter uma conformação variável, com tantas modas quanto se queira, como deve ser maleável, a fim de representar de forma satisfatória a grande variabilidade de comportamentos de variáveis quantitativas encontradas nos meios biológicos.

Com essas características em mente, o modelo foi desenvolvido no campo teórico, inicialmente, em que foram determinadas famílias de funções que tivessem as propriedades 1, 2, 3 e 5. A partir de um conjunto teórico de dados, buscou-se um modelo flexível. Tendo encontrado algumas famílias de funções, retornou-se à questão da integrabilidade. Nota-se que, apesar de haver uma sequência lógica das propriedades no processo de elaboração, essa não foi seguida. Isso porque se imaginava que a integrabilidade seria uma característica “fácil” de ser obtida, o que não se comprovou durante a pesquisa, como pode ser evidenciado no capítulo anterior.

No presente capítulo serão aplicados alguns dos principais modelos probabilísticos utilizados na Ciência Florestal: Normal, Beta, Gama, Exponencial, Weibull, Quadros e Spline. Cabe ressaltar que a função spline não é comumente utilizada como função densidade de probabilidade, pois necessita de algumas modificações em sua estrutura para atender aos requisitos desse fim. Essa função foi utilizada devido à grande variabilidade de formas que pode assumir. Para comprovar a eficiência do modelo proposto, esse e os modelos apresentados anteriormente serão ajustados a três distribuições diamétricas: *Bertholletia excelsa* Bonpl e dois de araucária.

5.2 *Bertholletia excelsa* Bonpl

As informações referentes à distribuição diamétrica da *Bertholletia excelsa* Bonpl foram extraídas do relatório do inventário florestal realizado no Plano de Colonização Agrícola de Altamira, ao longo da Rodovia Transamazônica, no ano 1976, contratado pela SUDAM (INCRA, 1976).

Segundo Hamelin (1991), a área do Projeto Integrado de colonização (PIC) Altamira, parte integrante do PIN (Plano de Integração Nacional), era uma faixa de terra com largura de 20 a 24 km, atravessada no meio pela Transamazônica, com extensão de 480 km (240 km para cada lado a partir da cidade de Altamira). Entretanto, o projeto só foi executado no interior de um segmento situado a Oeste de Altamira, do km 20 ao km 120; chamado de “colonização dirigida”, quanto ao resto, designou-se como “colonização espontânea”. A área que se estende do Km 20 ao Km 120 foi a única onde foi colocada em prática uma parte das diretrizes descritas em Altamira I: construção de agrovilas, postos de saúde, algumas escolas,

etc. A única também onde foram colocados em prática sistemas planejados de produção (WOOD; SCHMINK, 1982).

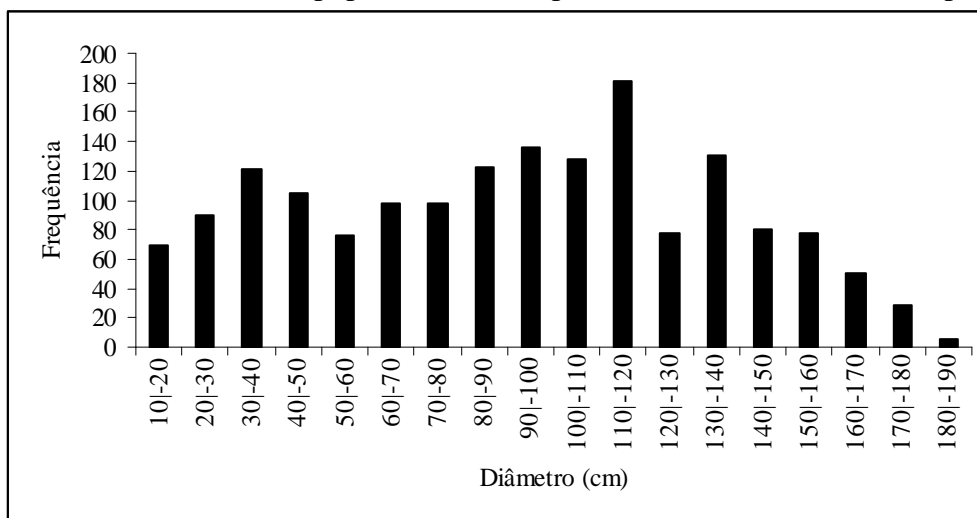
O processo de amostragem utilizado foi o de conglomerados, conforme foi apresentado em Péllico Netto e Brena (1997). A unidade amostral utilizada no conglomerado é composta por quatro sub-unidades, em cruz de malta, de acordo com o que foi proposto por Péllico Netto (1981), onde a área de cada sub-unidade tinha ¼ ha (10 x 250m), totalizando 1 ha para cada conglomerado. A espécie em estudo foi amostrada em todos os conglomerados e a totalidade das frequências foi distribuída em classes diamétricas, com intervalos de 10 cm, visando facilitar a visualização da sua gama de variação em toda área inventariada.

Na tabela 02 estão apresentadas as frequências em todas as classes diamétricas. Foi elaborado um histograma (FIGURA 20) para uma melhor visualização da distribuição empírica dos dados.

TABELA 02 – Frequências absolutas (f_{obs}) e acumuladas (F_{obs}) observadas por classe diamétrica da espécie *Bertholletia excelsa* Bonpl.

Li-ls	Xi	f_{obs}	F_{obs}
10 -20	15	70	70
20 -30	25	90	160
30 -40	35	121	281
40 -50	45	105	386
50 -60	55	76	462
60 -70	65	98	560
70 -80	75	98	658
80 -90	85	123	781
90 -100	95	136	917
100 -110	105	128	1045
110 -120	115	181	1226
120 -130	125	78	1304
130 -140	135	130	1434
140 -150	145	80	1514
150 -160	155	78	1592
160 -170	165	50	1642
170 -180	175	28	1670
180 -190	185	5	1675
Total		1675	

FIGURA 20 – Estrutura populacional da espécie *Bertholletia excelsa* Bonpl.



5.2.1 Estatística descritiva

A média aritmética calculada empiricamente para dados agrupados foi de 91,26 cm de DAP e variância de 1872,06 cm², logo o desvio padrão foi de 43,27 cm. Sabe-se que na natureza pode haver uma variação aleatória na quantidade de árvores, explicada por fatores ambientais, tais como características edafoclimáticas, ou fatores antrópicos. A presente série de dados foi considerada trimodal, ou seja, com três modas, nas classes diamétricas 30|-40, 110|-120 e 130|-140 cm.

Serão apresentados a seguir os resultados das distribuições Normal, Exponencial, Beta, Gama, Weibull, Quadros, Spline e do modelo proposto na presente pesquisa. A seguir será feita a comparação entre os modelos e a discussão sobre suas aderências ou não aos dados observados.

5.2.2 Normal

A distribuição Normal ajustada encontra-se na equação 67, enquanto que a média e a variância são apresentadas nas equações 68 e 69, respectivamente. Na tabela 03 encontram-se as frequências observadas a campo e esperadas pela função ajustada. Os valores da média e desvio padrão estimados pela distribuição são iguais aos amostrados.

$$f(x) = \left\{ \frac{1}{\sqrt{2\pi} \cdot 43,27} \cdot e^{-\frac{1}{2} \left(\frac{x-91,26}{43,27} \right)^2} \right. \quad -\infty < x < +\infty \quad \text{EQ. 67}$$

$$\mu = 91,26 \text{ cm} \quad \text{EQ. 68}$$

$$\sigma_x^2 = 1872,06 \text{ cm}^2 \quad \text{EQ. 69}$$

TABELA 03 – Distribuição diamétrica, frequências acumuladas esperadas pela distribuição Normal ajustada e diferença absoluta entre as distribuições acumuladas esperada e observada.

$Li-ls$	X_i	f_{obs}	F_{obs}	F_{esp}	$ F_{esp}-F_{obs} $
10 -20	15	70	70	65,32	4,68
20 -30	25	90	160	105,26	54,74
30 -40	35	121	281	162,08	118,92
40 -50	45	105	386	238,71	147,29
50 -60	55	76	462	336,71	125,29
60 -70	65	98	560	455,56	104,44
70 -80	75	98	658	592,21	65,79
80 -90	85	123	781	741,21	39,79
90 -100	95	136	917	895,24	21,76
100 -110	105	128	1045	1046,24	1,24
110 -120	115	181	1226	1186,59	39,41
120 -130	125	78	1304	1310,30	6,30
130 -140	135	130	1434	1413,69	20,31
140 -150	145	80	1514	1495,61	18,39
150 -160	155	78	1592	1557,18	34,82
160 -170	165	50	1642	1601,04	40,96
170 -180	175	28	1670	1630,67	39,33
180 -190	185	5	1675	1649,65	25,35
Total		1675			

As diferenças absolutas entre a frequência acumulada e esperada pela curva de Gauss ajustada servem para realizar o teste de Kolmogorov-Smirnov (KS). De acordo com a equação 49, os valores de KS tabelados são 0,0470 e 0,0563, a 95% e 99%, respectivamente. O valor de KS calculado foi de 0,0879, mostrando que a distribuição não foi aderente.

Encontra-se na figura 21 o gráfico dos valores observados, bem como a curva ajustada. Com o auxílio dessa imagem é possível ratificar que a distribuição não é capaz de representar de forma satisfatória a distribuição diamétrica da castanheira. Outro bom indicador encontra-se na figura 22, onde estão apresentados os resíduos percentuais entre valores estimados pelo modelo e observados a campo.

FIGURA 21 – Distribuição diamétrica da castanheira e curva Normal a ela ajustada.

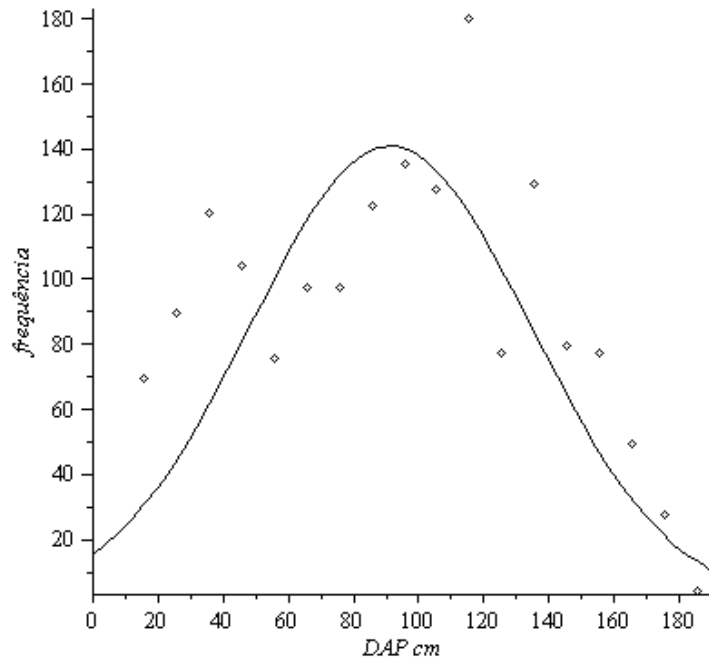
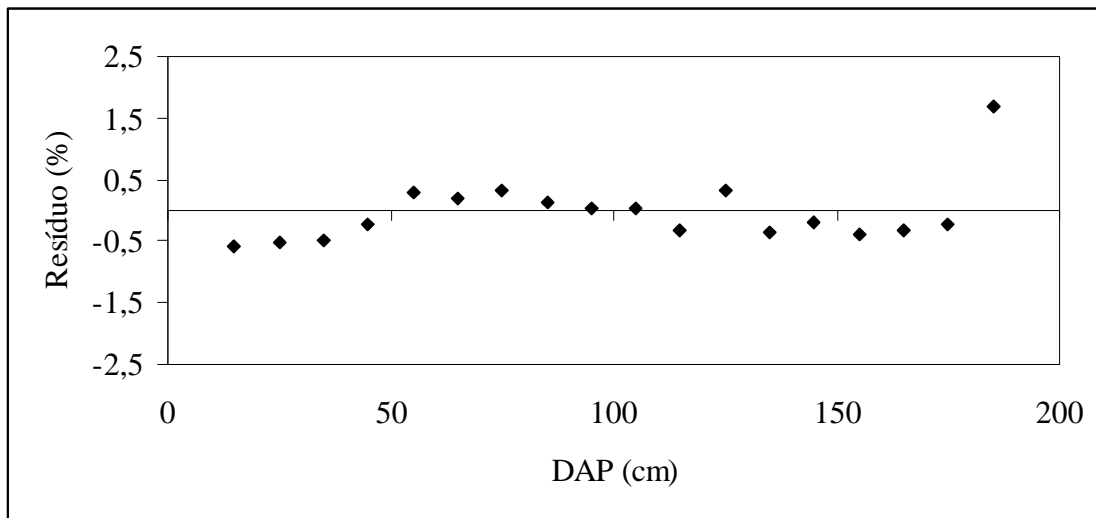


FIGURA 22 – Resíduos em porcentagem da distribuição Normal



5.2.3 Exponencial

A distribuição Exponencial ajustada encontra-se na equação 70, enquanto que a média e a variância são apresentadas nas equações 71 e 72, respectivamente. Na tabela 04 encontram-se as frequências observadas a campo e esperadas pela função ajustada. O valor da média estimada pela distribuição é igual ao amostrado, enquanto que o desvio padrão estimado é igual à média amostral, logo, muito diferente do observado.

$$f(x) = \begin{cases} \frac{1}{91,26} \cdot e^{\frac{-x}{91,26}} & \text{para } x \geq 0, \beta > 0 \\ 0 & \text{e.o.c.} \end{cases} \quad \text{EQ. 70}$$

$$E(X) = 91,26 \text{ cm} \quad \text{EQ. 71}$$

$$\sigma_x^2 = 8327,79 \text{ cm}^2 \quad \text{EQ. 72}$$

TABELA 04 – Distribuição diamétrica, frequências acumuladas esperadas pela distribuição Exponencial ajustada e diferença absoluta entre as distribuições acumuladas esperada e observada.

Li-ls	Xi	f _{obs}	F _{obs}	F _{esp}	F _{esp} -F _{obs}
10 -20	15	70	70	253,89	183,89
20 -30	25	90	160	401,38	241,38
30 -40	35	121	281	533,57	252,57
40 -50	45	105	386	652,04	266,04
50 -60	55	76	462	758,21	296,21
60 -70	65	98	560	853,37	293,37
70 -80	75	98	658	938,65	280,65
80 -90	85	123	781	1015,07	234,07
90 -100	95	136	917	1083,57	166,57
100 -110	105	128	1045	1144,95	99,95
110 -120	115	181	1226	1199,97	26,03
120 -130	125	78	1304	1249,27	54,73
130 -140	135	130	1434	1293,46	140,54
140 -150	145	80	1514	1333,06	180,94
150 -160	155	78	1592	1368,55	223,45
160 -170	165	50	1642	1400,35	241,65
170 -180	175	28	1670	1428,86	241,14
180 -190	185	5	1675	1454,41	220,59
Total		1675			

Comparando a razão da máxima diferença entre as frequências acumuladas observada e estimada pela distribuição Exponencial (KS = 0,1768) com os valores tabelados, observa-se que ela não foi aderente, como já se esperava. Sabe-se que a distribuição Exponencial é muito útil para estudos de regeneração natural, pois tem a forma de um “J” invertido, ou seja, é decrescente, mas não tem flexibilidade para ser unimodal, quanto menos trimodal, como é o caso do presente estudo.

Encontra-se na figura 23 o gráfico dos valores observados, bem como a curva ajustada. Com o auxílio dessa imagem é possível ratificar que a distribuição não é capaz de representar de forma satisfatória a distribuição diamétrica da castanheira. Outro bom indicador encontra-se na figura 24, onde estão apresentados os resíduos percentuais entre valores estimados pelo modelo e observados a campo.

FIGURA 23 – Distribuição diamétrica da castanheira e distribuição Exponencial a ela ajustada.

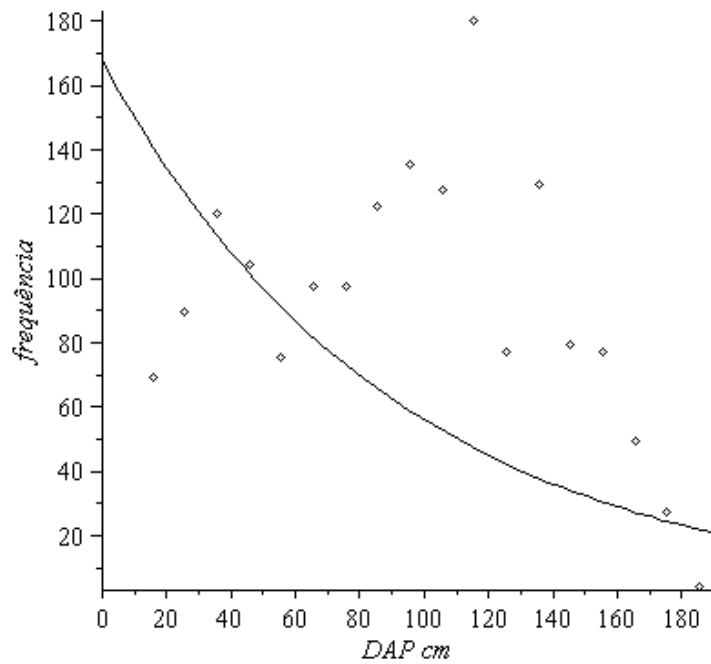
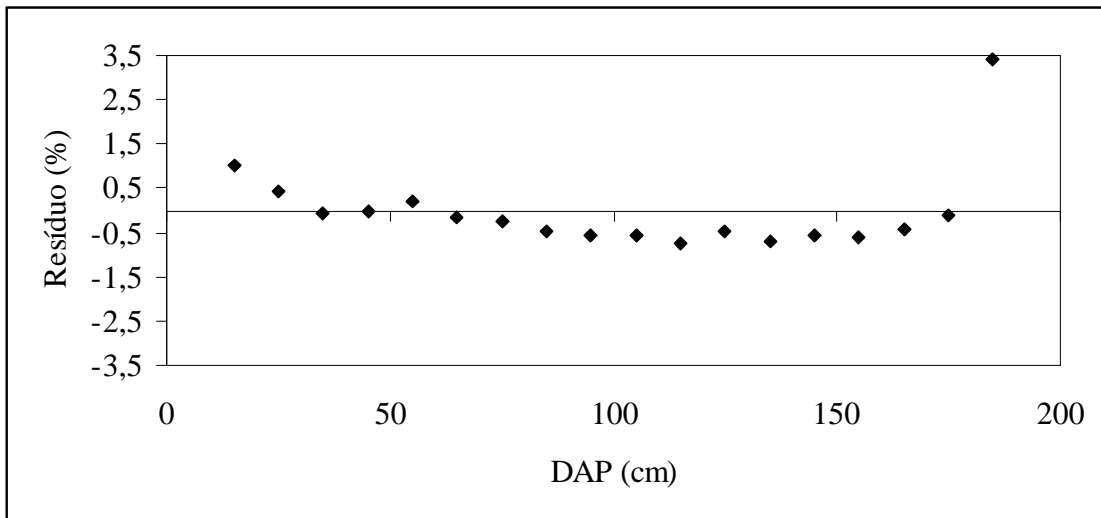


FIGURA 24 – Resíduos em percentagem da distribuição Exponencial



5.2.4 Beta

A distribuição Beta ajustada encontra-se na equação 73, enquanto que a média e a variância são apresentadas nas equações 74 e 75, respectivamente. Na tabela 05 encontram-se as frequências observadas a campo e esperadas pela função ajustada. De acordo com as fórmulas analíticas pode-se verificar que a distribuição subestimou um pouco a média, apesar de ter sido muito próximo, entretanto superestimou o desvio padrão em 2,5 cm.

$$f(x) = \begin{cases} 0.0001655 \cdot (x-10)^{0,27} \cdot (190-x)^{0,56} & \text{se } 10 < x < 190 \\ 0 & \text{e.o.c.} \end{cases} \quad \text{EQ. 73}$$

$$\mu = 90,74 \text{ cm} \quad \text{EQ. 74}$$

$$\sigma_x^2 = 2097,54 \text{ cm}^2 \quad \text{EQ. 75}$$

TABELA 05 – Distribuição diamétrica, frequências acumuladas esperadas pela distribuição Beta ajustada e diferença absoluta entre as distribuições acumuladas esperada e observada.

Li-ls	Xi	f _{obs}	F _{obs}	F _{esp}	F _{esp} -F _{obs}
10 -20	15	70	70	8,64	61,36
20 -30	25	90	160	47,27	112,73
30 -40	35	121	281	103,62	177,38
40 -50	45	105	386	173,12	212,88
50 -60	55	76	462	253,22	208,78
60 -70	65	98	560	342,15	217,85
70 -80	75	98	658	438,51	219,49
80 -90	85	123	781	541,10	239,90
90 -100	95	136	917	648,83	268,17
100 -110	105	128	1045	760,67	284,33
110 -120	115	181	1226	875,61	350,39
120 -130	125	78	1304	992,57	311,43
130 -140	135	130	1434	1110,41	323,59
140 -150	145	80	1514	1227,77	286,23
150 -160	155	78	1592	1343,01	248,99
160 -170	165	50	1642	1453,90	188,10
170 -180	175	28	1670	1556,99	113,01
180 -190	185	5	1675	1645,08	29,92
Total		1675			

O valor de KS (0,2092) calculado para a distribuição Beta foi maior que o tabelado, logo a distribuição não pode ser considerada aderente aos dados observados da castanheira.

Encontra-se na figura 25 o gráfico dos valores observados, bem como a curva ajustada. Com o auxílio dessa imagem é possível ratificar que a distribuição não é capaz de representar de forma satisfatória a distribuição diamétrica da castanheira. Sabe-se que a distribuição Beta é uma das mais flexíveis que existem, mas ainda assim não é capaz de ajustar-se a esses dados trimodais. Outro bom indicador encontra-se na figura 26, onde estão apresentados os resíduos percentuais entre valores estimados pelo modelo e observados a campo.

FIGURA 25 – Distribuição diamétrica da castanheira e distribuição Beta a ela ajustada.

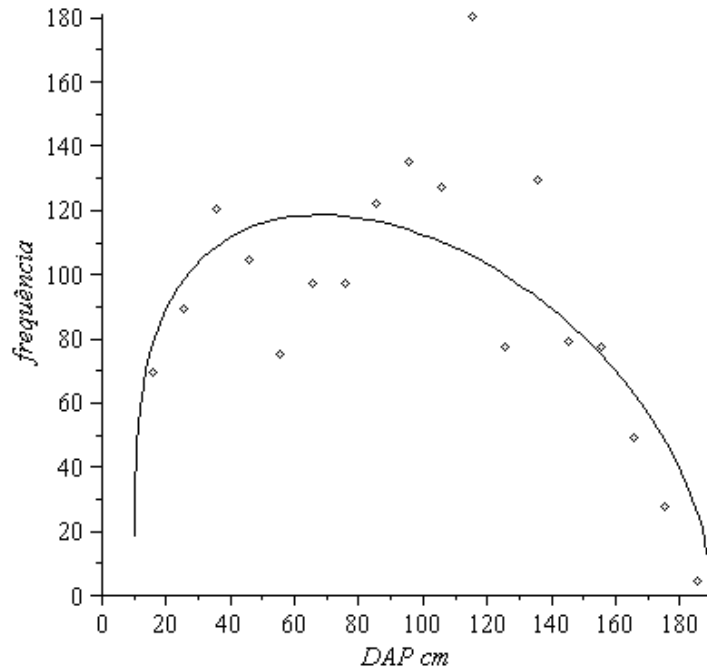
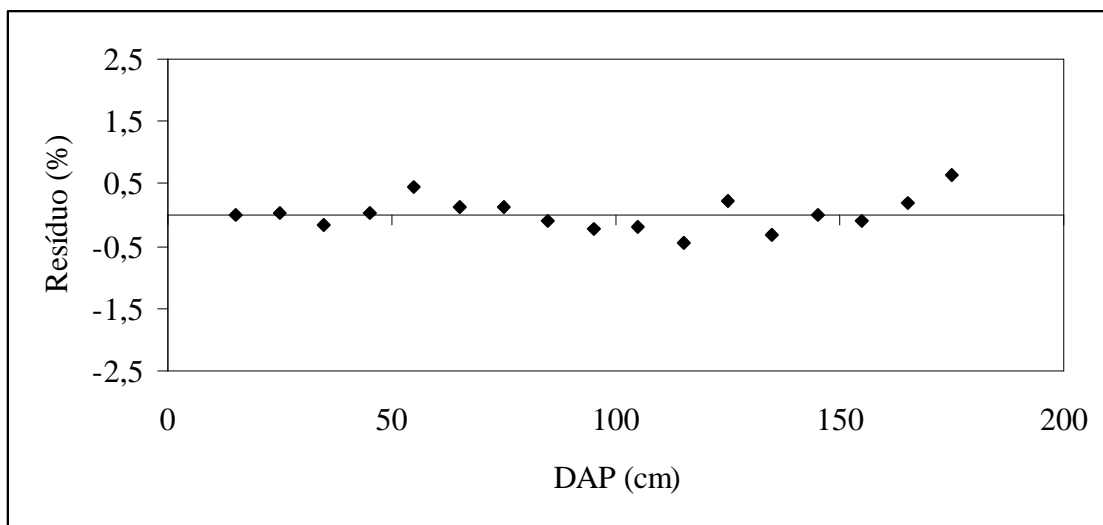


FIGURA 26 – Resíduos em porcentagem da distribuição Beta



5.2.5 Gama

A distribuição Gama ajustada encontra-se na equação 76, enquanto que a média e a variância são apresentadas nas equações 77 e 78, respectivamente. Na tabela 06 encontram-se as frequências observadas a campo e esperadas pela função ajustada. A média e o desvio padrão são iguais aos observados em campo, em função do método de ajuste.

$$f(x) = \begin{cases} \frac{x^{3,45} \cdot e^{-x/20,51}}{20,51^{4,45} \cdot \Gamma(4,45)} & \text{para } x \geq 0 \\ 0 & \text{e.o.c,} \end{cases} \quad \text{EQ. 76}$$

$$\mu = 91,26 \text{ cm} \quad \text{EQ. 77}$$

$$\sigma_x^2 = 1872,06 \text{ cm}^2 \quad \text{EQ. 78}$$

TABELA 06 – Distribuição diamétrica, frequências acumuladas esperadas pela distribuição Gama ajustada e diferença absoluta entre as distribuições acumuladas esperada e observada.

Li-ls	Xi	f _{obs}	F _{obs}	F _{esp}	F _{esp} -F _{obs}
10 -20	15	70	70	4,78	65,22
20 -30	25	90	160	31,56	128,44
30 -40	35	121	281	96,47	184,53
40 -50	45	105	386	203,42	182,58
50 -60	55	76	462	345,08	116,92
60 -70	65	98	560	508,29	51,71
70 -80	75	98	658	678,99	20,99
80 -90	85	123	781	845,29	64,29
90 -100	95	136	917	998,71	81,71
100 -110	105	128	1045	1134,33	89,33
110 -120	115	181	1226	1250,11	24,11
120 -130	125	78	1304	1346,16	42,16
130 -140	135	130	1434	1423,95	10,05
140 -150	145	80	1514	1485,67	28,33
150 -160	155	78	1592	1533,77	58,23
160 -170	165	50	1642	1570,70	71,30
170 -180	175	28	1670	1598,66	71,34
180 -190	185	5	1675	1619,59	55,41
Total		1675			

A distribuição Gama não foi aderente aos dados (KS = 0,1102). Apesar da flexibilidade desse modelo, não há combinação valores para seus coeficientes que a torne bimodal. Dessa forma já se esperava que o ajuste não seria satisfatório, como pode ser observado na figura 27, onde está apresentado o gráfico dos valores observados, bem como a curva ajustada. Outro bom indicador encontra-se na figura 28, onde estão apresentados os resíduos percentuais entre valores estimados pelo modelo e observados a campo.

FIGURA 27 – Distribuição diamétrica da castanheira e distribuição Gama a ela ajustada.

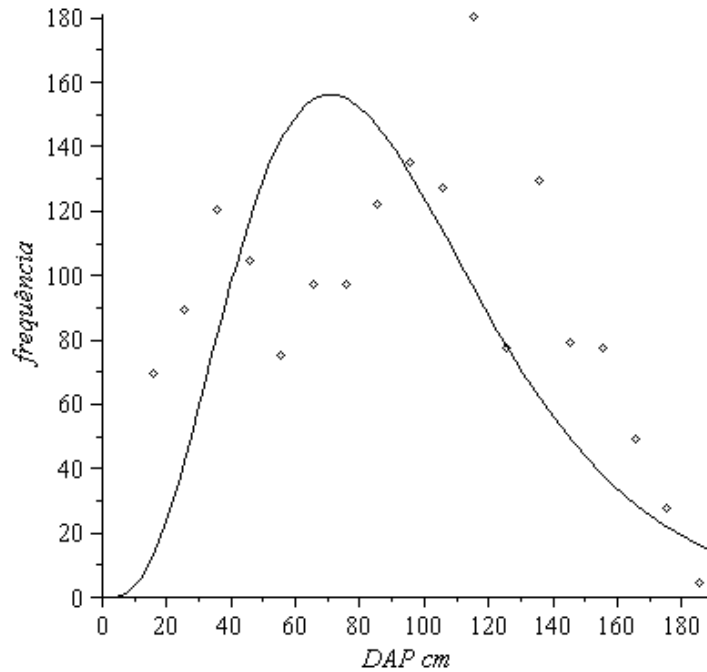
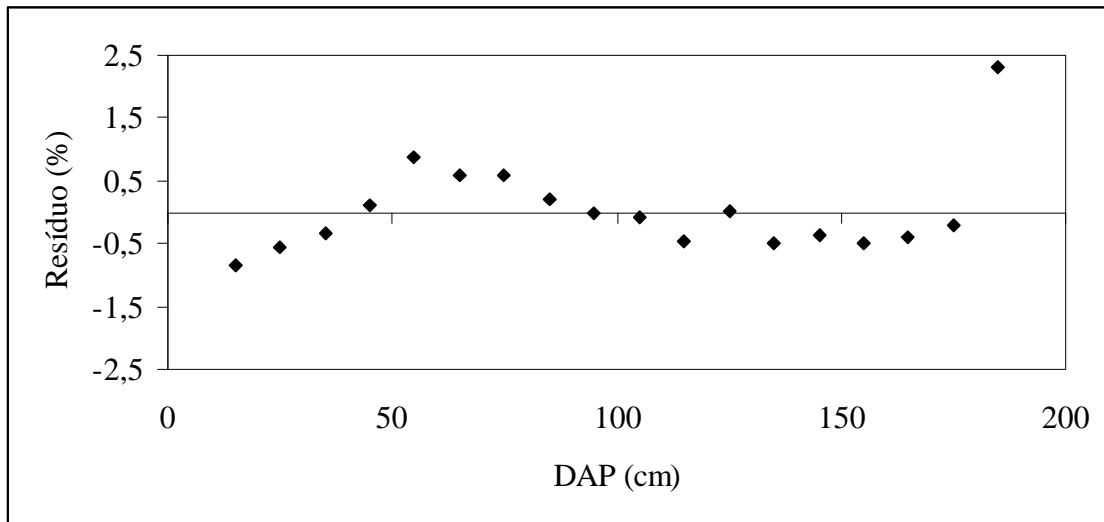


FIGURA 28 – Resíduos em porcentagem da distribuição Gama



5.2.6 Weibull

A distribuição Weibull ajustada encontra-se na equação 49. Na tabela 07 encontram-se as frequências observadas a campo e esperadas pela função ajustada. A média e o desvio padrão estimados pela distribuição de Weibull foram muito próximos dos observados a campo. Essa é uma distribuição muito utilizada na Ciência Florestal, em função de sua grande versatilidade.

$$f(x) = \begin{cases} 0,02 \left(\frac{x-14,95}{119,66} \right)^{1,711} \cdot e^{-\left(\frac{x-14,95}{119,66} \right)^{2,711}} & \text{se } x \geq a, b > 0 \text{ e } c > 0 \\ 0 & \text{e.o.c.} \end{cases} \quad \text{EQ. 79}$$

A média e a variância estão apresentadas nas Equações 80 e 81, respectivamente.

$$\mu = 91,48 \text{ cm} \quad \text{EQ. 80}$$

$$\sigma_x^2 = 1790,48 \text{ cm}^2 \quad \text{EQ. 81}$$

TABELA 07 – Distribuição diamétrica, frequências acumuladas esperadas pela distribuição Weibull ajustada e diferença absoluta entre as distribuições acumuladas esperada e observada.

Li-ls	Xi	f _{obs}	F _{obs}	F _{esp}	F _{esp} -F _{obs}
10 -20	15	70	70	38,59	31,41
20 -30	25	90	160	83,18	76,82
30 -40	35	121	281	149,35	131,65
40 -50	45	105	386	237,99	148,01
50 -60	55	76	462	348,06	113,94
60 -70	65	98	560	476,51	83,49
70 -80	75	98	658	618,58	39,42
80 -90	85	123	781	768,15	12,85
90 -100	95	136	917	918,47	1,47
100 -110	105	128	1045	1062,88	17,88
110 -120	115	181	1226	1195,58	30,42
120 -130	125	78	1304	1312,23	8,23
130 -140	135	130	1434	1410,29	23,71
140 -150	145	80	1514	1489,08	24,92
150 -160	155	78	1592	1549,53	42,47
160 -170	165	50	1642	1593,79	48,21
170 -180	175	28	1670	1624,67	45,33
180 -190	185	5	1675	1645,19	29,81
Total		1675			

A distribuição foi originalmente desenvolvida para testar resistência de materiais, mas logo foi aplicada às demais áreas do conhecimento, em função de sua flexibilidade. Apesar disso, o valor de KS (0,0884) mostra que esse modelo não se aderiu à distribuição diamétrica da castanheira. Esse era um fato esperado, pois, assim como as demais FDPs clássicas, essa não assume uma forma trimodal. Pode ser observado na figura 29 o ajuste da distribuição Weibull aos dados observados dos diâmetros da castanheira. Outro bom indicador encontra-se na figura 30, onde estão apresentados os resíduos percentuais entre valores estimados pelo modelo e observados a campo.

FIGURA 29 – Distribuição diamétrica da castanheira e distribuição Weibull a ela ajustada.

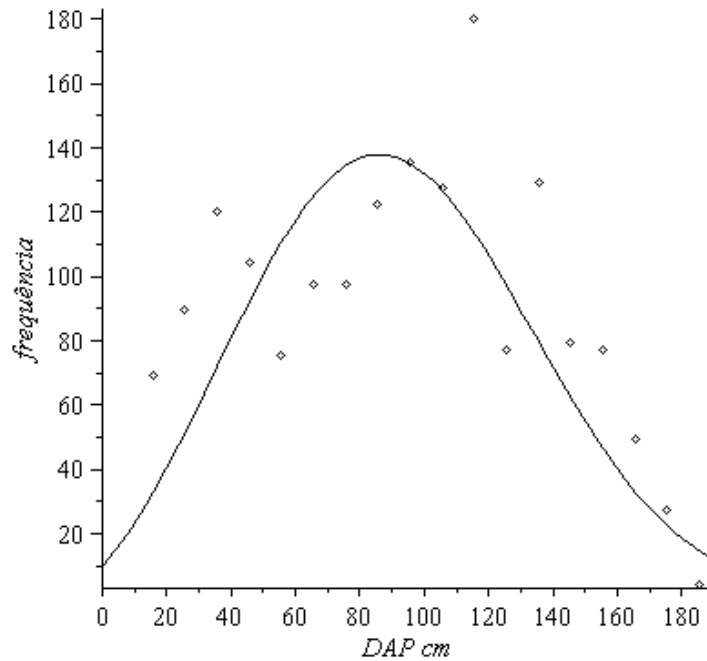
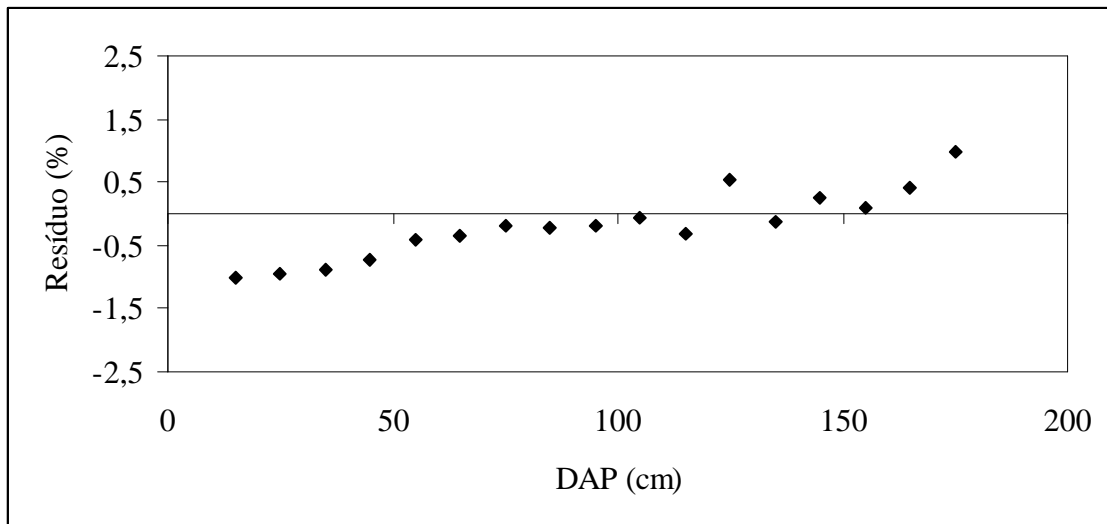


FIGURA 30 – Resíduos em porcentagem da distribuição Weibull



5.2.7 Quadros

A distribuição de Quadros ajustada encontra-se na equação 82. Na tabela 08 encontram-se as frequências observadas a campo e esperadas pela função ajustada. Os valores da média e do desvio padrão foram muito próximos dos observados, mostrando a alta qualidade do modelo. Ao se realizar o ajuste da função observou-se os valores dos coeficientes, muito próximos de zero. Normalmente estes valores podem ser retirados da função sem perda significativa de flexibilidade, o que não aconteceu na presente pesquisa. Foi

verificado que retirando-se apenas um coeficiente, a distribuição perdia sua característica mais notável sob o aspecto da tendência biológica observada: ser bimodal.

$$f(x) = \frac{1}{16890,01} \begin{cases} 7,49x^{0,8} & \text{se } 0 < x < 15 \\ -8,746 \cdot 10^{-10} x^6 + 5,656 \cdot 10^{-7} x^5 - 1,417 \cdot 10^{-4} x^4 + 1,717 \cdot 10^{-2} x^3 \\ -1,033x^2 + 28,68x - 183,6 & \text{se } 15 \leq x \leq 185 \\ \frac{1,0462 \cdot 10^{114}}{x^{50}} & \text{se } x > 185 \\ 0 & \text{e.o.c} \end{cases} \quad \text{EQ. 82}$$

A média e a variância estão apresentadas nas Equações 83 e 84, respectivamente.

$$\mu = 89,34 \text{ cm} \quad \text{EQ. 83}$$

$$\sigma_x^2 = 1981,35 \text{ cm}^2 \quad \text{EQ. 84}$$

TABELA 08 – Distribuição diamétrica, frequências acumuladas esperadas pela distribuição de Quadros ajustada e diferença absoluta entre as distribuições acumuladas esperada e observada.

Li-ls	Xi	f _{obs}	F _{obs}	F _{esp}	F _{esp} -F _{obs}
10 -20	15	70	70	87,27	17,27
20 -30	25	90	160	190,80	30,80
30 -40	35	121	281	295,91	14,91
40 -50	45	105	386	391,15	5,15
50 -60	55	76	462	480,25	18,25
60 -70	65	98	560	572,18	12,18
70 -80	75	98	658	674,91	16,91
80 -90	85	123	781	792,07	11,07
90 -100	95	136	917	922,07	5,07
100 -110	105	128	1045	1058,82	13,82
110 -120	115	181	1226	1193,73	32,27
120 -130	125	78	1304	1318,01	14,01
130 -140	135	130	1434	1425,01	8,99
140 -150	145	80	1514	1511,64	2,36
150 -160	155	78	1592	1578,51	13,49
160 -170	165	50	1642	1659,88	17,88
170 -180	175	28	1670	1663,12	6,88
180 -190	185	5	1675	1663,58	11,42
Total		1675			

Pelo teste de Kolmogorov-Smirnov pode-se concluir que o modelo foi aderente aos valores observados (KS = 0,0193). Esse modelo foi originalmente desenvolvido para aplicação na Ciência Florestal, mas, em função da forma como foi concebido, tem sido aplicado em outras áreas do conhecimento. Essa é uma distribuição recente e infelizmente

diversos pesquisadores ainda não têm contato nem muitas informações sobre ela. Pode ser observado na figura 31 o ajuste da distribuição de Quadros aos dados observados dos diâmetros da castanheira. Essa função teve a habilidade singular, até agora, de apresentar uma conformação bimodal. Outro bom indicador encontra-se na figura 32, onde estão apresentados os resíduos percentuais entre valores estimados pelo modelo e observados a campo.

FIGURA 31 – Distribuição diamétrica da castanheira e distribuição de Quadros a ela ajustada.

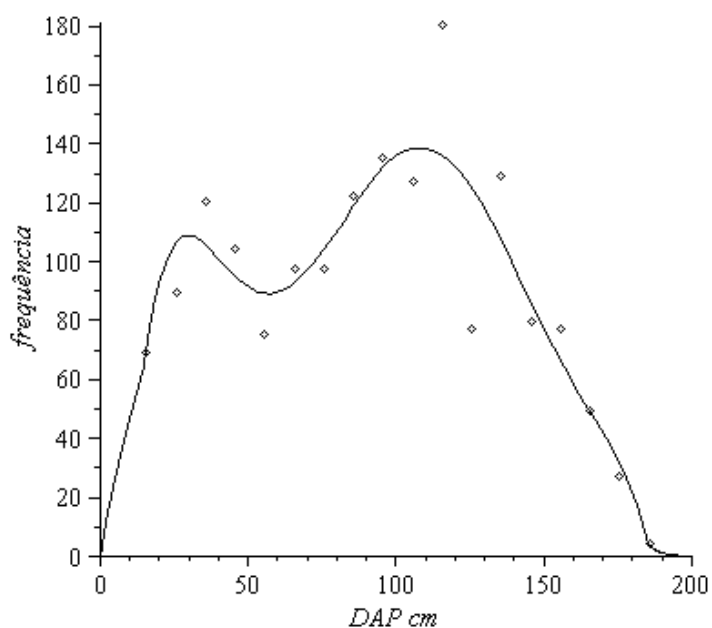
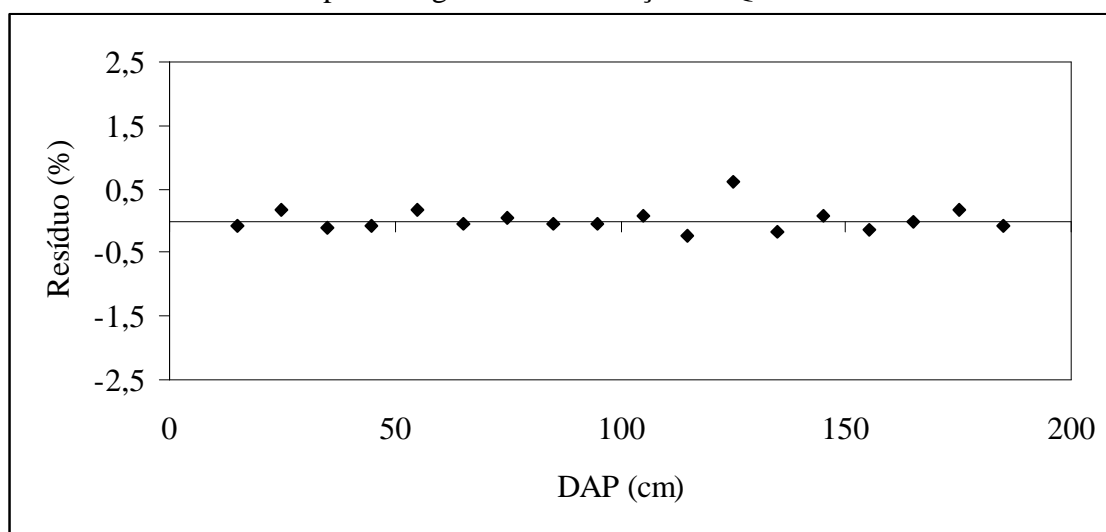


FIGURA 32 – Resíduos em porcentagem da distribuição de Quadros



5.2.8 Spline Cúbica

A Função Spline Cúbica ajustada encontra-se na equação 85. Na tabela 09 encontram-se as frequências observadas a campo e esperadas pela função ajustada. Apesar do tamanho da função truncada, ela é simples, quando analisada em partes, uma vez que é composta por funções de terceiro grau. Por isso leva o nome de Spline Cúbica.

$$f(x) = \frac{1}{24076,86} \left\{ \begin{array}{ll} 5,6948x - 0,0046x^3 & 0 \leq x < 15 \\ -64,2278 + 18,5404x - 0,8564x^2 - 0,0145x^3 & 15 \leq x < 25 \\ 476,976 - 46,4041x + 1,7414x^2 - 0,0202x^3 & 25 \leq x < 35 \\ -741,5667 + 58,0425x - 1,2428x^2 + 0,0082x^3 & 35 \leq x < 45 \\ -1921,9769 + 136,7365x - 2,9915x^2 + 0,0212x^3 & 45 \leq x < 55 \\ 6436,0064 - 319,1535x + 5,2974x^2 - 0,0290x^3 & 55 \leq x < 65 \\ -7566,9860 + 327,1384x - 4,6456x^2 + 0,0219x^3 & 65 \leq x < 75 \\ 6659,5682 - 241,9237x + 2,9419x^2 - 0,0118x^3 & 75 \leq x < 85 \\ 6721,4062 - 244,1062x + 2,9676x^2 - 0,0119x^3 & 85 \leq x < 95 \\ -46542,9935 + 1437,9274x - 14,7380x^2 + 0,0502x^3 & 95 \leq x < 105 \\ 135666,2598 - 3768,0512x + 34,8427x^2 - 0,1071x^3 & 105 \leq x < 115 \\ -272645,8516 + 6883,5691x - 57,7801x^2 + 0,1613x^3 & 115 \leq x < 125 \\ 486120,0875 - 11326,8135x + 87,9030x^2 - 0,2272x^3 & 125 \leq x < 135 \\ -1279151,7409 + 27901,4494x - 202,6767x^2 + 0,4903x^3 & 135 \leq x < 145 \\ 2904888,0556 - 58664,8912x + 394,3325x^2 - 0,8821x^3 & 145 \leq x < 155 \\ -3575620,3718 + 66764,3041x - 414,8881x^2 + 0,8581x^3 & 155 \leq x < 165 \\ 2132084,8988 - 37012,1553x + 214,0601x^2 - 0,4125x^3 & 165 \leq x < 175 \\ -521612,9866 + 8479,8084x - 45,8939x^2 + 0,0827x^3 & 175 \leq x < 187 \end{array} \right.$$

EQ. 85

A média e a variância estão apresentadas nas Equações 86 e 87 respectivamente. Observa-se que a função Spline Cúbica superestimou a média e o desvio padrão.

$$\mu = 109,14 \text{ cm} \quad \text{EQ. 86}$$

$$\sigma_x^2 = 2263,00 \text{ cm}^2 \quad \text{EQ. 87}$$

TABELA 09 – Distribuição diamétrica, frequências acumuladas esperadas pela função Spline ajustada e diferença absoluta entre as distribuições acumuladas esperada e observada.

$Li-ls$	Xi	f_{obs}	F_{obs}	F_{esp}	$ F_{esp}-F_{obs} $
10 -20	15	70	70	66,73	3,27
20 -30	25	90	160	130,29	29,71
30 -40	35	121	281	212,59	68,41
40 -50	45	105	386	285,02	100,98
50 -60	55	76	462	340,28	121,72
60 -70	65	98	560	406,90	153,10
70 -80	75	98	658	476,41	181,59
80 -90	85	123	781	561,64	219,36
90 -100	95	136	917	654,52	262,48
100 -110	105	128	1045	748,18	296,82
110 -120	115	181	1226	864,72	361,28
120 -130	125	78	1304	930,51	373,49
130 -140	135	130	1434	1005,00	429,00
140 -150	145	80	1514	1107,26	406,74
150 -160	155	78	1592	1576,91	15,09
160 -170	165	50	1642	1655,24	13,24
170 -180	175	28	1670	1655,72	14,28
180 -190	185	5	1675	1675,00	0,00
Total		1675			

Ao observar a figura 33, na qual se encontra a representação gráfica da distribuição diamétrica da castanheira e a função Spline Cúbica a ela ajustada, observa-se que ela passa sobre todos os pontos observados. Era de se esperar que fosse, dessa maneira, aderente à distribuição, o que não é confirmado pelo teste de Kolmogorov-Smirnov ($KS = 0,2661$). Esse fato traz consigo duas conclusões muito importantes:

- 1) Ter frequência absoluta estimada pelo modelo igual às observadas empiricamente não garante que a distribuição seja satisfatória para explicar esse conjunto de dados.
- 2) A função Spline Cúbica ignora por completo quaisquer comportamentos ou tendências biológicas, uma vez que não foi criada para isso. Logo, deve ser utilizada com bastante cuidado. Outro bom indicador encontra-se na figura 34, onde estão apresentados os resíduos percentuais entre valores estimados pelo modelo e observados a campo.

FIGURA 33 – Distribuição diamétrica da castanheira e função Spline Cúbica a ela ajustada.

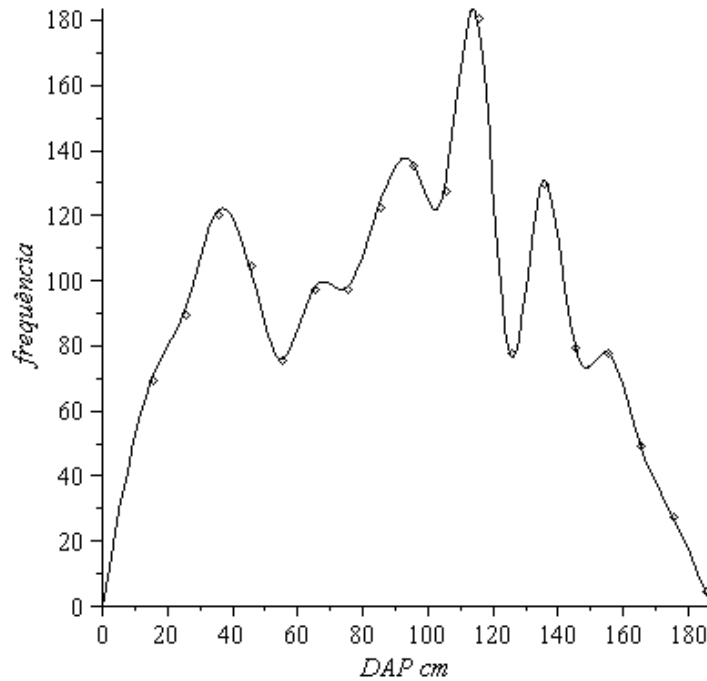
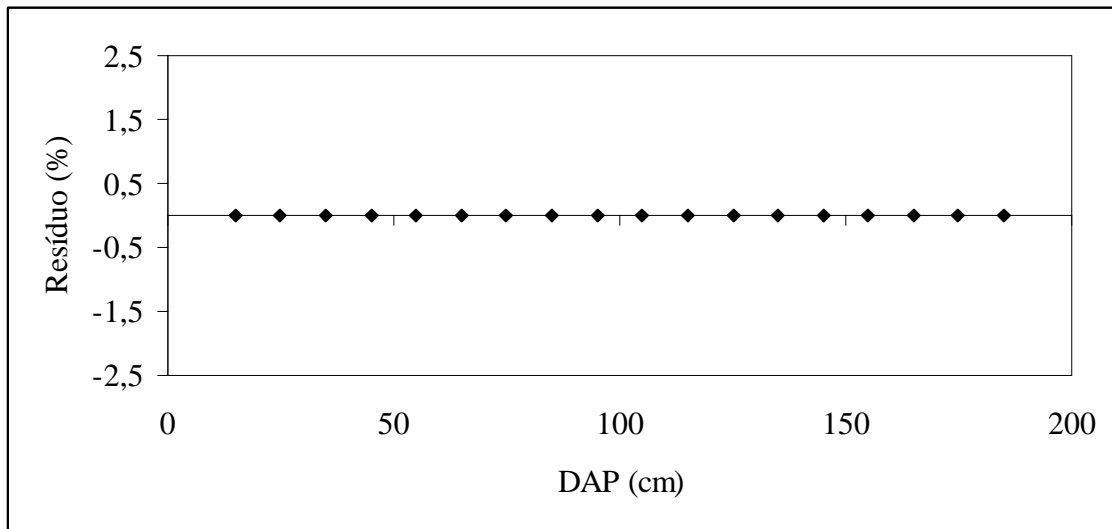


FIGURA 34 – Resíduos em percentagem da função Spline Cúbica



5.2.9 Modelo Proposto

Na proposta do presente estudo foi apresentada uma distribuição multimodal. Como citado anteriormente, os dados apresentam 3 modas bem definidas (DAP = 35, 115 e 135 cm), característica que se almeja do modelo (EQUAÇÃO 88).

$$f(x) = \frac{ax^8 + bx^6 + cx}{d + e \cdot (fx^2 - g)^2 \cdot (hx^2 - i)^2 \cdot (jx^2 - k)^2} \quad \text{EQ. 88}$$

Os coeficientes g , i e k são dependentes de f , h e j , respectivamente, e são obtidos por meio das modas, logo não são ajustados por meio de regressão linear. Os cálculos e resultados estão apresentados nas Equações 89, 90 e 91.

$f \cdot x^2 - g$ Parte da função que gera a primeira moda. Será definida com PGM1 (Parte Geradora da Moda 1).

$f \cdot x^2 - g = 0$ Quando o valor de x for igual a 35 a PGM1 deverá ser igual a zero.

$$g = f \cdot 35^2 \quad \text{EQ. 89}$$

$h \cdot x^2 - i$ PGM2

$h \cdot x^2 - i = 0$ Quando o valor de x for igual a 115 a PGM2 deverá ser igual a zero.

$$i = h \cdot 115^2 = \quad \text{EQ. 90}$$

$j \cdot x^2 - k$ PGM3

$j \cdot x^2 - k = 0$ Quando o valor de x for igual a 135 a PGM3 deverá ser igual a zero.

$$k = j \cdot 135^2 = \quad \text{EQ. 91}$$

Os valores dos coeficientes calculados por meio de regressão não linear foram $a = 0,106786 \times 10^{-3}$; $b = -2,54215328$; $c = 0,370413 \times 10^{11}$; $d = 0,117937 \times 10^{11}$; $e = 0,985599 \times 10^{-6}$; $f = 0,175603222$; $h = 0,87781519 \times 10^{-1}$ e $j = 0,17560322 \times 10^{-1}$. Substituindo-se esses valores no modelo e realizando as simplificações pertinentes, tem-se a equação ajustada (EQUAÇÃO 92).

$$f(x) = \frac{1,07 \cdot 10^{-4} x^8 - 2,54x^6 + 3,70 \cdot 10^{-10} x}{1,18 \cdot 10^{10} + 9,86 \cdot 10^{-7} \cdot (0,18x^2 - 215,11)^2 \cdot (0,09x^2 - 1160,91)^2 \cdot (0,02x^2 - 320,04)^2}$$

EQ. 92

Integrando essa função no intervalo de zero até mais infinito obtém-se uma constante $k = 16.055,87$, que representa a área entre a curva ajustada e o eixo x (DAP). O inverso de k deve ser multiplicado pela função ajustada, a fim de se encontrar a FDP (EQUAÇÃO 93).

$$FDP = \frac{1}{16.055,87} \cdot \frac{1,07 \cdot 10^{-4} x^8 - 2,54x^6 + 3,70 \cdot 10^{-10} x}{1,18 \cdot 10^{10} + 9,86 \cdot 10^{-7} \cdot (0,18x^2 - 215,11)^2 \cdot (0,09x^2 - 1160,91)^2 \cdot (0,02x^2 - 320,04)^2}$$

EQ. 93

A representação gráfica da Equação 92, bem como os pontos referentes aos dados observados, encontram-se na figura 35. As frequências esperadas acumuladas estão apresentadas na tabela 10. Os resíduos encontram-se na figura 36.

FIGURA 35 – Distribuição diamétrica da Castanheira e modelo proposto a ela ajustada.

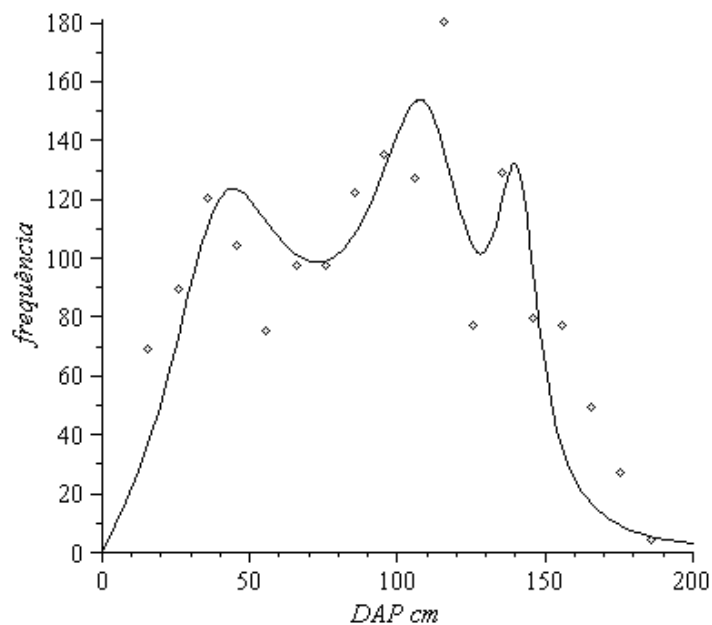
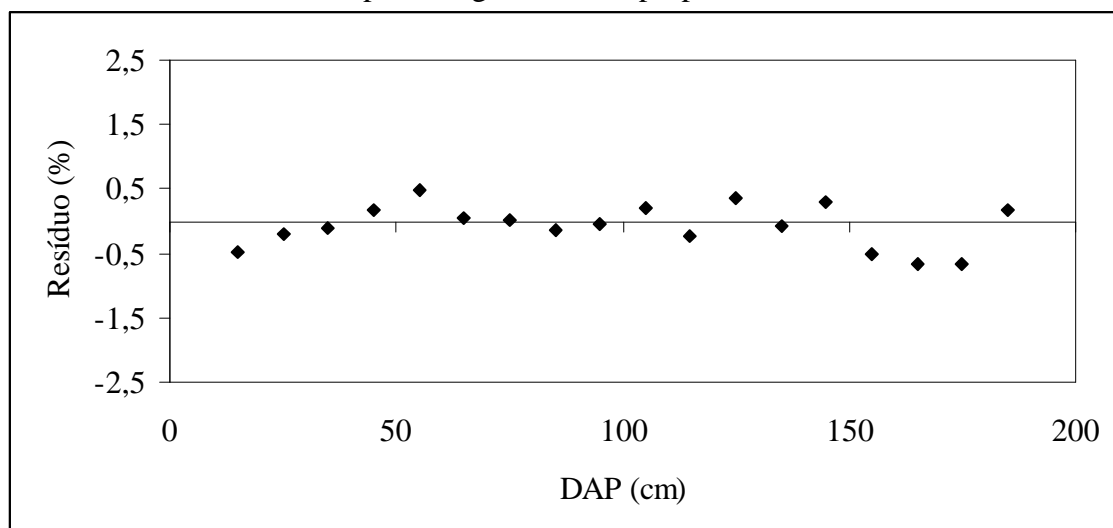


FIGURA 36 – Resíduos em porcentagem modelo proposto



A integral do produto entre a FDP e x , no intervalo de zero a mais infinito, ou seja, no seu domínio, é igual à média aritmética estimada pelo modelo ($\mu = 90,51$ cm). A variância é obtida fazendo-se a diferença entre a integral do produto da FDP por x^2 , e a média elevada ao quadrado. O resultado obtido foi de 2015,55 cm².

TABELA 10 – Distribuição diamétrica, frequências acumuladas esperadas pelo modelo proposto ajustado e diferença absoluta entre as distribuições acumuladas esperada e observada.

Li-ls	Xi	f _{obs}	F _{obs}	F _{esp}	F _{esp} -F _{obs}
10 -20	15	70	70	48,34	21,66
20 -30	25	90	160	122,44	37,56
30 -40	35	121	281	235,53	45,47
40 -50	45	105	386	363,47	22,53
50 -60	55	76	462	481,82	19,82
60 -70	65	98	560	588,37	28,37
70 -80	75	98	658	691,96	33,96
80 -90	85	123	781	804,48	23,48
90 -100	95	136	917	938,80	21,80
100 -110	105	128	1045	1095,85	50,85
110 -120	115	181	1226	1240,28	14,28
120 -130	125	78	1304	1351,95	47,95
130 -140	135	130	1434	1475,25	41,25
140 -150	145	80	1514	1581,13	67,13
150 -160	155	78	1592	1622,85	30,85
160 -170	165	50	1642	1641,23	0,77
170 -180	175	28	1670	1651,19	18,81
180 -190	185	5	1675	1653,57	21,43
Total		1675			

O valor calculado para o modelo proposto no teste de aderência de Kolmogorov-Smirnov foi de 0,0401. Comparando-se com os valores críticos, pode-se dizer que o modelo ajustou-se satisfatoriamente aos dados. Adicionalmente, pode-se observar que esse foi o único modelo que apresentou três modas, conforme observado a campo.

5.2.10 Resumo dos resultados dos modelos

Na tabela 11 estão apresentados de maneira resumida os coeficientes dos modelos ajustados e os valores de KS e Syx, a fim de se realizar uma análise comparativa. Os coeficientes da função Spline Cúbica e a distribuição de Quadros não serão inseridos nessa tabela, uma vez que são em grande número (Spline = 74, Quadros = 11). Os valores de referência do teste de Kolmogorov-Smirnov são 0,0470 e 0,0563, para 95 e 99% de probabilidade, respectivamente.

TABELA 11 – Comparação entre os modelos para a castanheira.

Modelo	Coeficientes				KS	Syx%
Normal	$\mu = 91,26$	$\sigma = 43,27$			0,0879**	3,43
Exponencial	$\beta = 91,26$				0,1768**	7,48
Beta	a=10	b=190	$\alpha = 1,27$	$\beta = 1,56$	0,2092**	5,26
Gama	$\alpha = 4,45$	$\beta = 20,51$			0,1102**	5,11
Weibull	a=-14,95	b=119,66	c=2,71		0,0884**	3,49
Quadros					0,0193ns	2,37
Spline					0,2561**	11,39
Modelo proposto					0,0401ns	2,70

Valores seguidos de * e ** indicam significância a 95% e 99% de probabilidade, respectivamente, e ns indica não significância.

5.3 *Araucaria angustifolia*

A distribuição diamétrica utilizada para a presente aplicação do modelo proposto é oriunda de um remanescente florestal inserido na Fazenda Experimental da Universidade Federal do Paraná, localizada no município de São João do Triunfo, estado do Paraná, a uma distância de aproximadamente 125 km de Curitiba. Esta unidade experimental possui área de cerca de 32 hectares e os experimentos foram lá implantados em julho de 1995, sendo 3 parcelas de 1 hectare (100 x 100m) e 1 parcela de 0,5 hectare (50 x 100 m). Estas parcelas foram subdivididas em subparcelas de 10 x 10 m.

O clima da região, segundo a classificação de Köppen, é classificado como do tipo Cfb – clima Temperado Sempre Úmido com temperatura média inferior a 22°C durante o mês mais quente, sem estação seca, com verões frescos e mais de cinco geadas noturnas durante o ano (PIZATTO, 1999).

De acordo com Pizatto (1999), o solo da floresta foi classificado como tipo ALISSOLOS com uma pequena porção do solo tipo CAMBISSOLOS. A ordem dos ALISSOLOS foi extinta em função do processo de formação (que implica na concentração de alumínio no solo), não ser considerado como principal, tendo em vista que o referido processo ocorre em diversas classes de solos, independente do processo de formação principal. Os solos contidos nesta classe foram redistribuídos nas ordens dos Argissolos e Nitossolos. Os CAMBISSOLOS são solos constituídos por material mineral com horizonte B incipiente subjacente a qualquer tipo de horizonte superficial, exceto hístico com 40cm ou mais de espessura, ou horizonte A chernozêmico, quando o B incipiente apresentar argila de atividade alta e saturação por bases alta. Plintita e petroplintita, horizonte glei e horizonte vértico, se presentes, não satisfazem os requisitos para Plintossolos, Gleissolos e Vertissolos, respectivamente (EMBRAPA, 2006).

A parcela das Araucárias possui o dossel, quase que totalmente, dominado pela *Araucaria angustifolia* acompanhada no estrato inferior por espécies arbóreas como: *Nectandra grandiflora*, *Ilex paraguariensis* e outros gêneros. Segundo Pizatto (1999) e Longhi (1980), uma distância moderada existe entre o dossel e o primeiro estrato das espécies de folhosas. Adicionalmente, Pizatto (1999) afirma que é provável que nesta área tenha existido uma intervenção antrópica e, dessa forma, criou-se condições para a reconstituição da floresta, favorecendo de certa forma a regeneração natural.

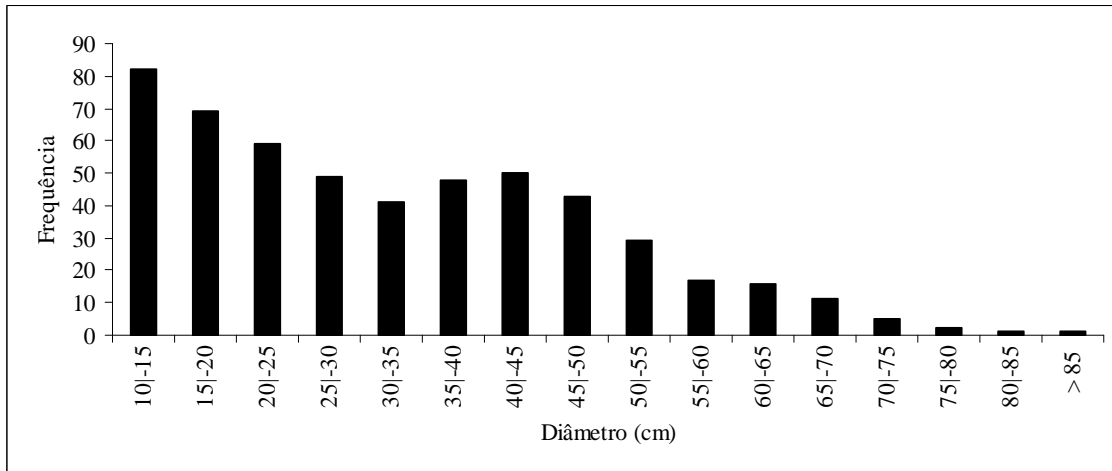
Todas as árvores com DAP (diâmetro à altura do peito - 1,30 m) maior ou igual a 10 cm foram identificadas, pintadas na altura de medição e numeradas com etiquetas de plástico. Exsicatas de todas as espécies foram preparadas e armazenadas em herbário da Universidade Federal do Paraná.

Na tabela 12, estão apresentados os dados sumariados, bem como as frequências em todas as classes diamétricas. Foi elaborado um histograma (FIGURA 37), para uma melhor visualização da distribuição empírica dos dados.

TABELA 12 – Frequências absolutas (f_{obs}) e acumuladas (F_{obs}) observadas por classe diamétrica da Araucária em uma área de 3,5 ha em São João do triunfo

L_i-L_s	X_i	f_{obs}	F_{obs}
10 -15	12,5	82	82
15 -20	17,5	69	151
20 -25	22,5	59	210
25 -30	27,5	49	259
30 -35	32,5	41	300
35 -40	37,5	48	348
40 -45	42,5	50	398
45 -50	47,5	43	441
50 -55	52,5	29	470
55 -60	57,5	17	487
60 -65	62,5	16	503
65 -70	67,5	11	514
70 -75	72,5	5	519
75 -80	77,5	2	521
80 -85	82,5	1	522
> 85	87,5	1	523
Total		523	

FIGURA 37 – Estrutura populacional da Araucária em uma área de 3,5 ha em São João do triunfo



5.3.1 Estatística descritiva

A média aritmética calculada empiricamente para dados agrupados foi de 32,77 cm de DAP e variância de 266,21 cm², logo o desvio padrão foi de 16,32 cm. Pode-se observar um padrão semelhante ao da castanheira, quando considerado o número de modas: duas, entretanto observa-se que a distribuição diamétrica da Araucária não apresentou outlier. As duas modas observadas encontram-se nas classes diamétricas de 10 a 15 cm e 40 a 45 cm.

Serão apresentados a seguir os resultados das distribuições Normal, Exponencial, Beta, Gama, Weibull, Quadros, Spline e do modelo proposto na presente pesquisa. A seguir será feita a comparação entre os modelos e a discussão sobre suas aderências ou não aos dados observados.

5.3.2 Normal

A distribuição Normal ajustada encontra-se na equação 94, enquanto que a média e a variância são apresentadas nas equações 95 e 96, respectivamente. Na tabela 13 encontram-se as frequências observadas a campo e esperadas pela função ajustada. Os valores da média e desvio padrão estimados pela distribuição são iguais aos amostrados.

$$f(x) = \left\{ \frac{1}{\sqrt{2\pi} \cdot 16,32} \cdot e^{-\frac{1}{2} \left(\frac{x-32,77}{16,32} \right)^2} \right. \quad -\infty < x < +\infty \quad \text{EQ. 94}$$

$$\mu = 32,77 \text{ cm} \quad \text{EQ. 95}$$

$$\sigma_x^2 = 266,21 \text{ cm}^2 \quad \text{EQ. 96}$$

TABELA 13 – Distribuição diamétrica, frequências acumuladas esperadas pela distribuição Normal ajustada e diferença absoluta entre as distribuições acumuladas esperada e observada.

$Li-ls$	Xi	f_{obs}	F_{obs}	F_{esp}	$ F_{esp}-F_{obs} $
10 -15	12,5	82	82	56,00	26,00
15 -20	17,5	69	151	91,37	59,63
20 -25	22,5	59	210	138,37	71,63
25 -30	27,5	49	259	195,29	63,71
30 -35	32,5	41	300	258,08	41,92
35 -40	37,5	48	348	321,18	26,82
40 -45	42,5	50	398	378,95	19,05
45 -50	47,5	43	441	427,14	13,86
50 -55	52,5	29	470	463,77	6,23
55 -60	57,5	17	487	489,12	2,12
60 -65	62,5	16	503	505,11	2,11
65 -70	67,5	11	514	514,30	0,30
70 -75	72,5	5	519	519,11	0,11
75 -80	77,5	2	521	521,40	0,40
80 -85	82,5	1	522	522,40	0,40
> 85	87,5	1	523	522,79	0,21
Total		523			

As diferenças absolutas entre a frequência acumulada e esperada pela curva de Gauss ajustada servem para realizar o teste de Kolmogorov-Smirnov (KS). De acordo com a equação 49, os valores de KS tabelados são 0,0841 e 0,1008, a 95% e 99%, respectivamente. O valor de KS calculado foi de 0,1370, mostrando que a distribuição não foi aderente.

Encontra-se na figura 38 o gráfico dos valores observados, bem como a curva ajustada. Com o auxílio dessa imagem é possível ratificar que a distribuição não é capaz de representar de forma satisfatória a distribuição diamétrica da Araucária. Outro bom indicador encontra-se na figura 39, onde estão apresentados os resíduos percentuais entre valores estimados pelo modelo e observados a campo.

FIGURA 38 – Distribuição diamétrica da Araucária e curva Normal a ela ajustada.

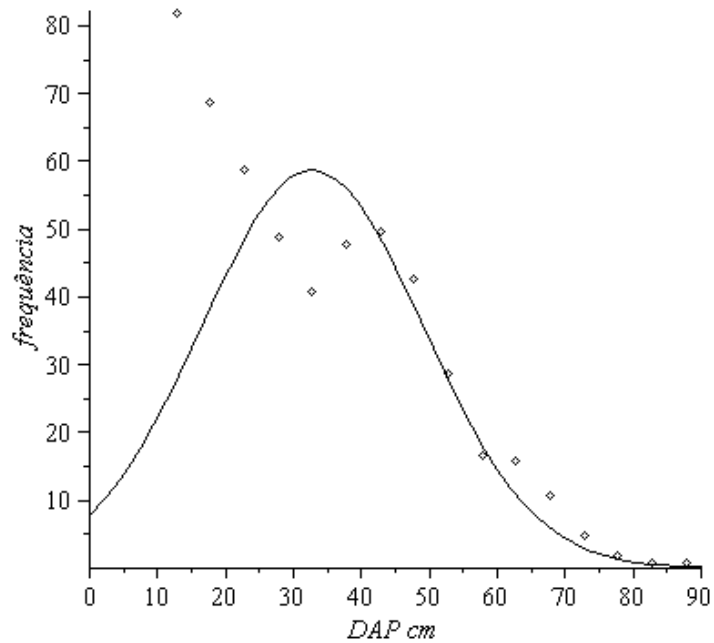
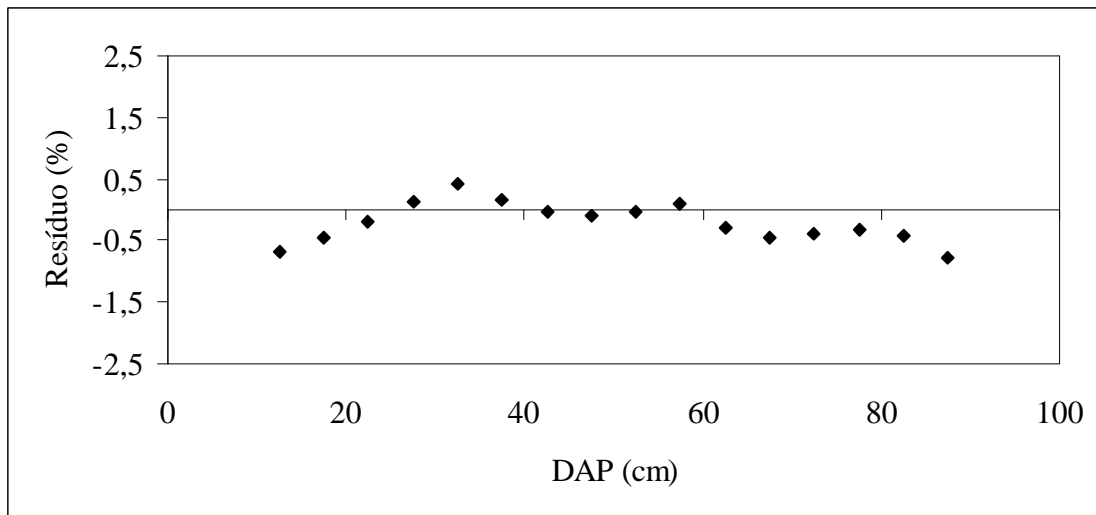


FIGURA 39 – Resíduos em porcentagem da distribuição Normal



5.3.3 Exponencial

A distribuição Exponencial ajustada encontra-se na equação 97, enquanto que a média e a variância são apresentadas nas equações 98 e 99, respectivamente. Na tabela 14 encontram-se as frequências observadas a campo e esperadas pela função ajustada. O valor da média estimada pela distribuição é igual ao amostrado, enquanto que o desvio padrão é igual à média, logo, muito diferente do observado.

$$f(x) = \begin{cases} \frac{1}{32,77} \cdot e^{\frac{-x}{32,77}} & \text{para } x \geq 0, \beta > 0 \\ 0 & \text{e.o.c.} \end{cases} \quad \text{EQ. 97}$$

$$\mu = 32,77 \text{ cm} \quad \text{EQ. 98}$$

$$\sigma_x^2 = 1073,72 \text{ cm}^2 \quad \text{EQ. 99}$$

TABELA 14 – Distribuição diamétrica, frequências acumuladas esperadas pela distribuição Exponencial ajustada e diferença absoluta entre as distribuições acumuladas esperada e observada.

Li-ls	Xi	f _{obs}	F _{obs}	F _{esp}	F _{esp} -F _{obs}
10 -15	12,5	82	82	165,87	83,87
15 -20	17,5	69	151	216,41	65,41
20 -25	22,5	59	210	259,80	49,80
25 -30	27,5	49	259	297,04	38,04
30 -35	32,5	41	300	329,02	29,02
35 -40	37,5	48	348	356,47	8,47
40 -45	42,5	50	398	380,04	17,96
45 -50	47,5	43	441	400,27	40,73
50 -55	52,5	29	470	417,64	52,36
55 -60	57,5	17	487	432,55	54,45
60 -65	62,5	16	503	445,35	57,65
65 -70	67,5	11	514	456,34	57,66
70 -75	72,5	5	519	465,77	53,23
75 -80	77,5	2	521	473,87	47,13
80 -85	82,5	1	522	480,82	41,18
> 85	87,5	1	523	486,79	36,21
Total		523			

Comparando a razão da máxima diferença entre as frequências acumuladas observada e estimada pela distribuição Exponencial (KS = 0,1604) com os valores tabelados, observa-se que ela não foi aderente, como já se esperava, uma vez que essa distribuição não apresenta modas. Este modelo tem sido testado e utilizado amplamente em função de sua simplicidade de ajuste e bons resultados em muitos casos de florestas naturais, contudo, na distribuição que está sendo utilizada, não gerou um resultado satisfatório.

Encontra-se na figura 40 o gráfico dos valores observados, bem como a curva ajustada. Com o auxílio dessa imagem é possível ratificar que a distribuição não é capaz de representar de forma satisfatória a distribuição diamétrica da Araucária. Outro bom indicador encontra-se na figura 41, onde estão apresentados os resíduos percentuais entre valores estimados pelo modelo e observados a campo.

FIGURA 40 – Distribuição diamétrica da Araucária e distribuição Exponencial a ela ajustada.

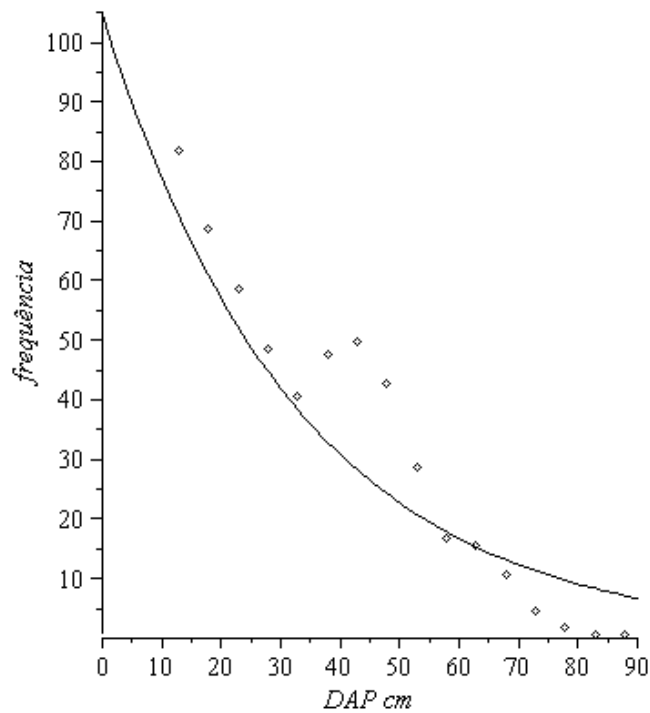
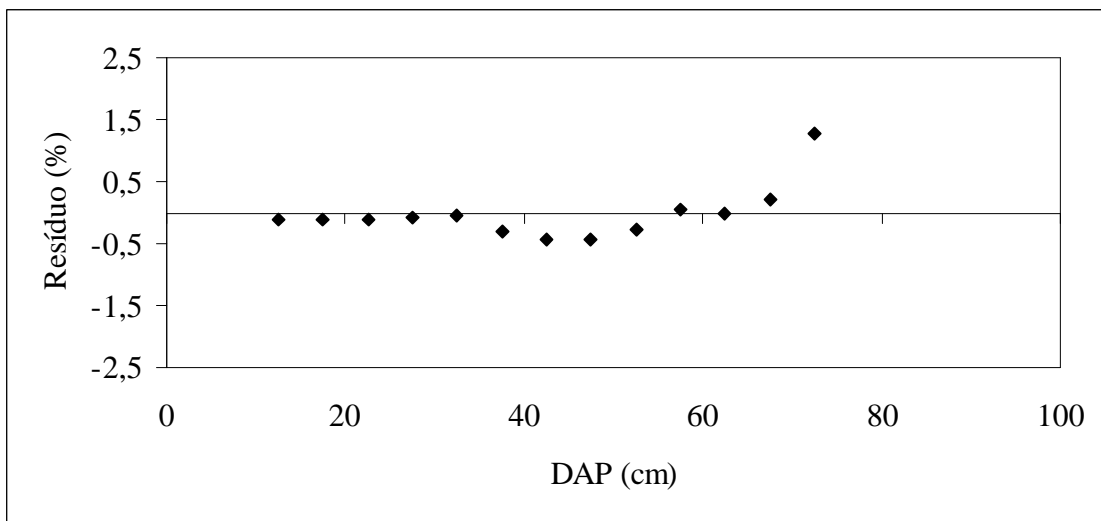


FIGURA 41 – Resíduos em percentagem da distribuição Exponencial



5.3.4 Beta

A distribuição Beta ajustada encontra-se na equação 100, enquanto que a média e a variância são apresentadas nas equações 101 e 102, respectivamente. Na tabela 15 encontram-se as frequências observadas a campo e esperadas pela função ajustada. De acordo com as fórmulas analíticas pode-se verificar que as estimativas da média e do desvio padrão da distribuição foram muito próximos dos obtidos a partir dos dados amostrais.

$$f(x) = \begin{cases} 0,0000749 \cdot (x-10)^{-1420} \cdot (90-x)^{1,4577} & \text{se } 10 < x < 90 \\ 0 & \text{e.o.c.} \end{cases} \quad \text{EQ. 100}$$

$$\mu = 30,70 \text{ cm} \quad \text{EQ. 101}$$

$$\sigma_x^2 = 284,44 \text{ cm}^2 \quad \text{EQ. 102}$$

TABELA 15 – Distribuição diamétrica, frequências acumuladas esperadas pela distribuição Beta ajustada e diferença absoluta entre as distribuições acumuladas esperada e observada.

Li-ls	Xi	f _{obs}	F _{obs}	F _{esp}	F _{esp} -F _{obs}
10 -15	12,5	82	82	0,08	81,92
15 -20	17,5	69	151	1,22	149,78
20 -25	22,5	59	210	4,32	205,68
25 -30	27,5	49	259	9,95	249,05
30 -35	32,5	41	300	18,60	281,40
35 -40	37,5	48	348	30,71	317,29
40 -45	42,5	50	398	46,72	351,28
45 -50	47,5	43	441	67,07	373,93
50 -55	52,5	29	470	92,19	377,81
55 -60	57,5	17	487	122,60	364,40
60 -65	62,5	16	503	158,84	344,16
65 -70	67,5	11	514	201,57	312,43
70 -75	72,5	5	519	251,65	267,35
75 -80	77,5	2	521	310,30	210,70
80 -85	82,5	1	522	379,59	142,41
> 85	87,5	1	523	464,67	58,33
Total		523			

O valor de KS (0,7224) calculado para a distribuição Beta foi maior que o tabelado, logo a distribuição não pode ser considerada aderente aos dados observados da Araucária.

Encontra-se na figura 42 o gráfico dos valores observados, bem como a curva ajustada. Com o auxílio dessa imagem é possível ratificar que a distribuição não é capaz de representar de forma satisfatória a distribuição diamétrica da Araucária. Sabe-se que a distribuição Beta é uma das mais flexíveis que existem, mas ainda assim não é capaz ajustar-se a esses dados bimodais. Outro bom indicador encontra-se na figura 43, onde estão apresentados os resíduos percentuais entre valores estimados pelo modelo e observados a campo.

FIGURA 42 – Distribuição diamétrica da Araucária e distribuição Beta a ela ajustada.

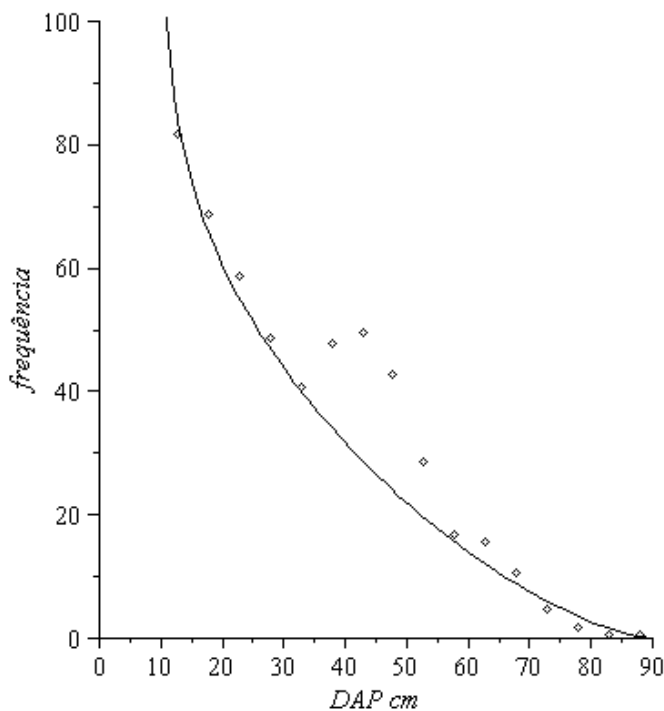
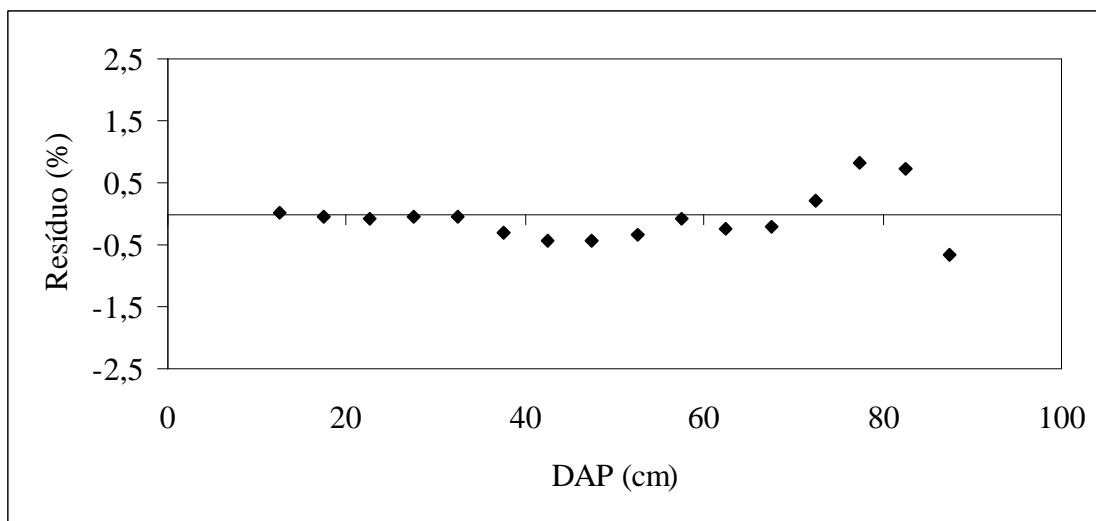


FIGURA 43 – Resíduos em porcentagem da distribuição Beta



O gráfico dos resíduos em porcentagem do ajuste da distribuição Beta aos dados da Araucária foi limitado entre -150 e 150%, com o objetivo de padronização. Os resíduos da distribuição Beta até 47,5 cm extrapolaram esse limite, logo não aparecem no gráfico.

5.3.5 Gama

A distribuição Gama ajustada encontra-se na equação 103, enquanto que a média e a variância são apresentadas nas equações 104 e 105, respectivamente. Na tabela 16 encontram-

se as frequências observadas a campo e esperadas pela função ajustada. A média e o desvio padrão são iguais aos observados em campo, em função do método de ajuste.

$$f(x) = \begin{cases} \frac{X^{3,03} \cdot e^{-x/8,12}}{8,12^{4,03} \cdot \Gamma(4,03)} & \text{para } x \geq 0 \\ 0 & \text{e.o.c,} \end{cases} \quad \text{EQ. 103}$$

$$\mu = 32,77 \text{ cm} \quad \text{EQ. 104}$$

$$\sigma_x^2 = 266,21 \text{ cm}^2 \quad \text{EQ. 105}$$

TABELA 16 – Distribuição diamétrica, frequências acumuladas esperadas pela distribuição Gama ajustada e diferença absoluta entre as distribuições acumuladas esperada e observada.

Li-ls	Xi	f _{obs}	F _{obs}	F _{esp}	F _{esp} -F _{obs}
10 -15	12,5	82	82	35,52	46,48
15 -20	17,5	69	151	87,29	63,71
20 -25	22,5	59	210	154,26	55,74
25 -30	27,5	49	259	225,61	33,39
30 -35	32,5	41	300	292,78	7,22
35 -40	37,5	48	348	350,83	2,83
40 -45	42,5	50	398	397,94	0,06
45 -50	47,5	43	441	434,37	6,63
50 -55	52,5	29	470	461,51	8,49
55 -60	57,5	17	487	481,11	5,89
60 -65	62,5	16	503	494,91	8,09
65 -70	67,5	11	514	504,42	9,58
70 -75	72,5	5	519	510,86	8,14
75 -80	77,5	2	521	515,16	5,84
80 -85	82,5	1	522	517,98	4,02
> 85	87,5	1	523	519,82	3,18
Total		523			

A distribuição Gama não foi aderente aos dados (KS = 0,1218). Apesar da flexibilidade desse modelo, não há combinação de valores para seus coeficientes que a torne bimodal. Dessa forma já se esperava que o ajuste não seria satisfatório, como pode ser observado na figura 44, onde está apresentado o gráfico dos valores observados, bem como a curva ajustada. Outro bom indicador encontra-se na figura 45, onde estão apresentados os resíduos percentuais entre valores estimados pelo modelo e observados a campo.

FIGURA 44 – Distribuição diamétrica da Araucária e distribuição Gama a ela ajustada.

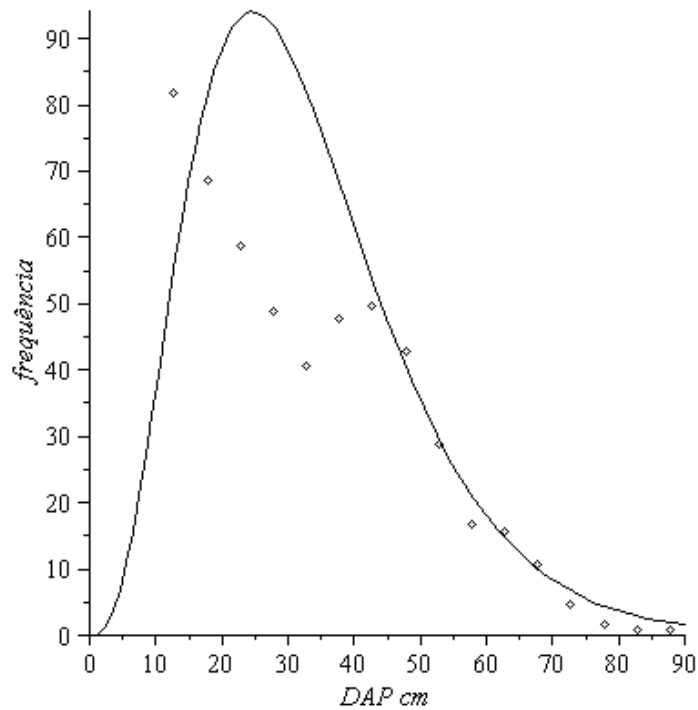
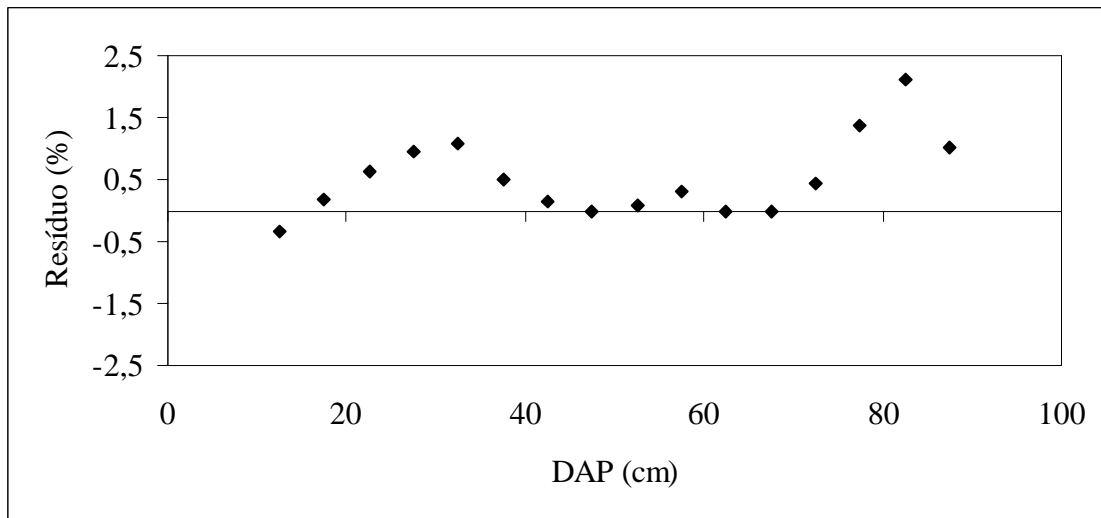


FIGURA 45 – Resíduos em percentagem da distribuição Gama



5.3.6 Weibull

A distribuição Weibull ajustada encontra-se na equação 106. Na tabela 17 encontram-se as frequências observadas a campo e esperadas pela função ajustada. A média estimada pela distribuição de Weibull foi muito próxima da observada a campo. Porém, a distribuição superestimou o desvio padrão. Essa é uma distribuição muito utilizada nas Ciências Florestais, em função de sua grande versatilidade.

$$f(x) = \begin{cases} 0,04 \left(\frac{x-10}{19,61} \right)^{-0,1664} \cdot e^{-\left(\frac{x-10}{19,61} \right)^{0,8336}} & \text{se } x \geq a, b > 0 \text{ e } c > 0 \\ 0 & \text{e.o.c.} \end{cases} \quad \text{EQ. 106}$$

A média e a variância estão apresentadas nas Equações 107 e 108 respectivamente.

$$\mu = 31,60 \text{ cm} \quad \text{EQ. 107}$$

$$\sigma_x^2 = 678,96 \text{ cm}^2 \quad \text{EQ. 108}$$

TABELA 17 – Distribuição diamétrica, frequências acumuladas esperadas pela distribuição Weibull ajustada e diferença absoluta entre as distribuições acumuladas esperada e observada.

Li-ls	Xi	f _{obs}	F _{obs}	F _{esp}	F _{esp} -F _{obs}
10 -15	12,5	82	82	85,97	3,97
15 -20	17,5	69	151	189,10	38,10
20 -25	22,5	59	210	259,88	49,88
25 -30	27,5	49	259	312,35	53,35
30 -35	32,5	41	300	352,58	52,58
35 -40	37,5	48	348	384,06	36,06
40 -45	42,5	50	398	409,02	11,02
45 -50	47,5	43	441	429,03	11,97
50 -55	52,5	29	470	445,20	24,80
55 -60	57,5	17	487	458,34	28,66
60 -65	62,5	16	503	469,10	33,90
65 -70	67,5	11	514	477,93	36,07
70 -75	72,5	5	519	485,22	33,78
75 -80	77,5	2	521	491,26	29,74
80 -85	82,5	1	522	496,27	25,73
> 85	87,5	1	523	500,45	22,55
Total		523			

O valor de KS (0,1020) mostra que esse modelo não se aderiu à distribuição diamétrica da Araucária por pouco, uma vez que o valor de referência a 99% é igual a 0,1008. Isso mostra a alta versatilidade desse modelo, embora a tendência desejada (bimodal) não tenha sido obtida. Pode ser observado na figura 46 o ajuste da distribuição Weibull aos dados observados dos diâmetros da Araucária. Outro bom indicador encontra-se na figura 47, onde estão apresentados os resíduos percentuais entre valores estimados pelo modelo e observados a campo.

FIGURA 46 – Distribuição diamétrica da Araucária e distribuição Weibull a ela ajustada.

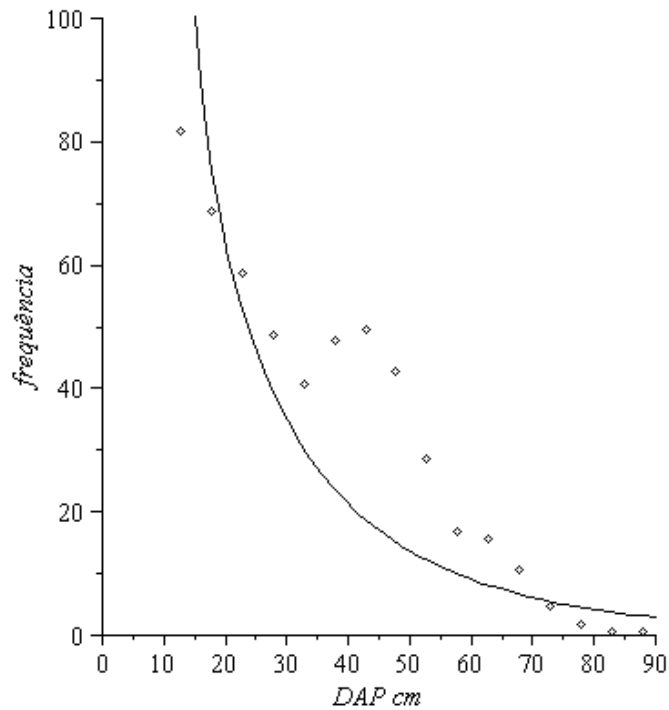
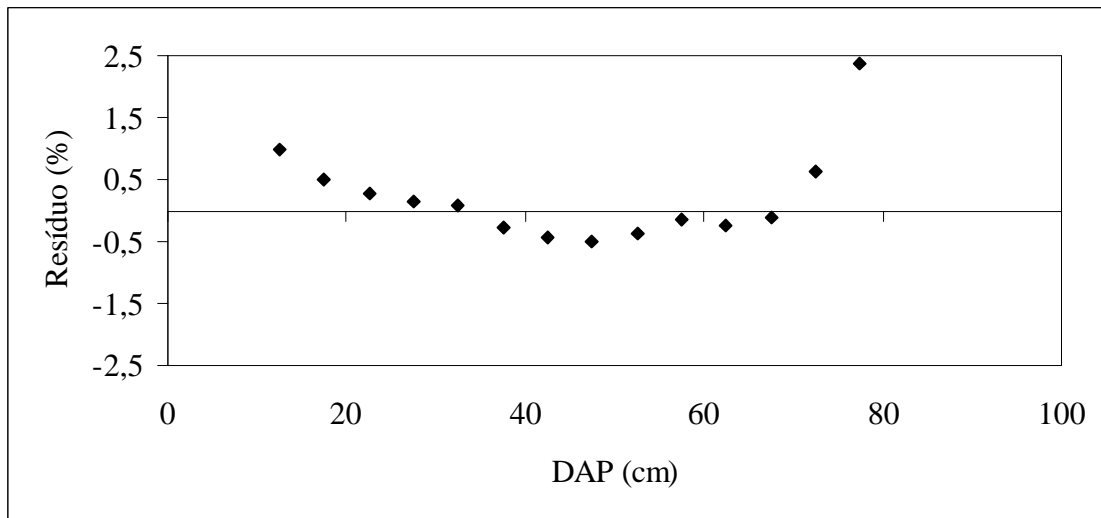


FIGURA 47 – Resíduos em percentagem da distribuição Weibull



5.3.7 Quadros

A distribuição de Quadros ajustada encontra-se na equação 107. Na tabela 18 encontram-se as frequências observadas a campo e esperadas pela função ajustada. O valor da média estimado pela função foi abaixo do observado, enquanto que a variância foi superestimada. Isso ocorre em decorrência das funções obtidas na primeira e na terceira parte. A distribuição de Quadros será eficiente, na forma como foi concebida, quando as modas estiverem nas classes intermediárias, o que não foi observado na presente pesquisa. Já foi mencionado, em tópicos anteriores, que os coeficientes muito próximos de zero não podem

ser retirados, pois acarretam perda significativa de flexibilidade, tornando o modelo inapto. Apesar de toda a flexibilidade apresentada pelo modelo polinomial, ele não foi capaz de representar a moda intermediária.

$$f(x) = \frac{1}{3411,97} \begin{cases} 0,019x^{0,1} & \text{se } 0 < X < 10 \\ -5,0536 \cdot 10^{-7} x^5 + 1,4703 \cdot 10^{-4} x^4 - 1,5867 \cdot 10^{-2} x^3 + 0,78309 \cdot x^2 \\ -18,404x + 219,51 & \text{se } 10 \leq X \leq 90 \\ \frac{8383431436}{x^5} & \text{se } X > 90 \\ 0 & \text{e.o.c} \end{cases} \quad \text{EQ. 109}$$

A média e a variância estão apresentadas nas Equações 110 e 111, respectivamente.

$$\mu = 27,56 \text{ cm} \quad \text{EQ. 110}$$

$$\sigma_x^2 = 439,85 \text{ cm}^2 \quad \text{EQ. 111}$$

TABELA 18 – Distribuição diamétrica, frequências acumuladas esperadas pela distribuição de Quadros ajustada e diferença absoluta entre as distribuições acumuladas esperada e observada.

Li-ls	Xi	f _{obs}	F _{obs}	F _{esp}	F _{esp} -F _{obs}
10 -15	12,5	82	82	179,90	97,90
15 -20	17,5	69	151	230,17	79,17
20 -25	22,5	59	210	273,17	63,17
25 -30	27,5	49	259	312,91	53,91
30 -35	32,5	41	300	350,87	50,87
35 -40	37,5	48	348	386,94	38,94
40 -45	42,5	50	398	420,13	22,13
45 -50	47,5	43	441	449,17	8,17
50 -55	52,5	29	470	472,98	2,98
55 -60	57,5	17	487	490,99	3,99
60 -65	62,5	16	503	503,27	0,27
65 -70	67,5	11	514	510,58	3,42
70 -75	72,5	5	519	514,24	4,76
75 -80	77,5	2	521	515,84	5,16
80 -85	82,5	1	522	516,87	5,13
> 85	87,5	1	523	523,00	0,00
Total		523			

Pelo teste de Kolmogorov-Smirnov pode-se concluir que o modelo não foi aderente aos valores observados (KS = 0,1872). Pode ser observado na figura 48 o ajuste da distribuição de Quadros aos dados observados dos diâmetros da Araucária. Na figura 49 estão

apresentados os resíduos percentuais entre valores estimados pelo modelo e observados a campo.

FIGURA 48 – Distribuição diamétrica da Araucária e distribuição de Quadros a ela ajustada.

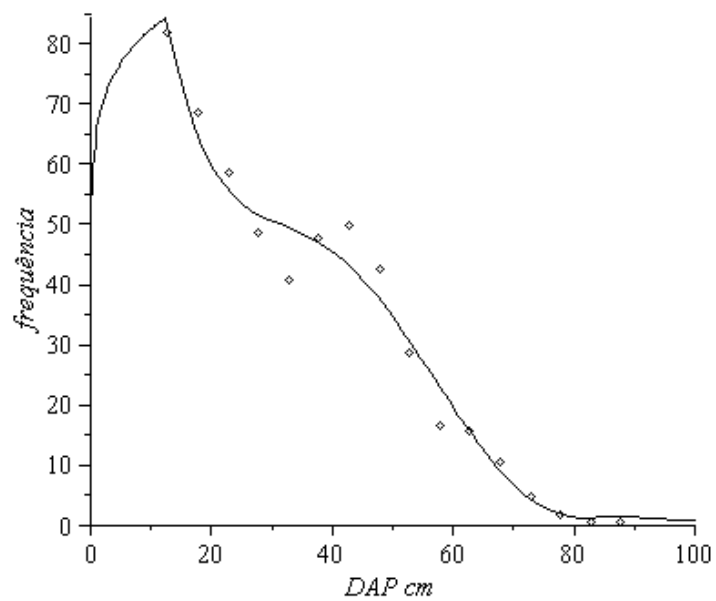
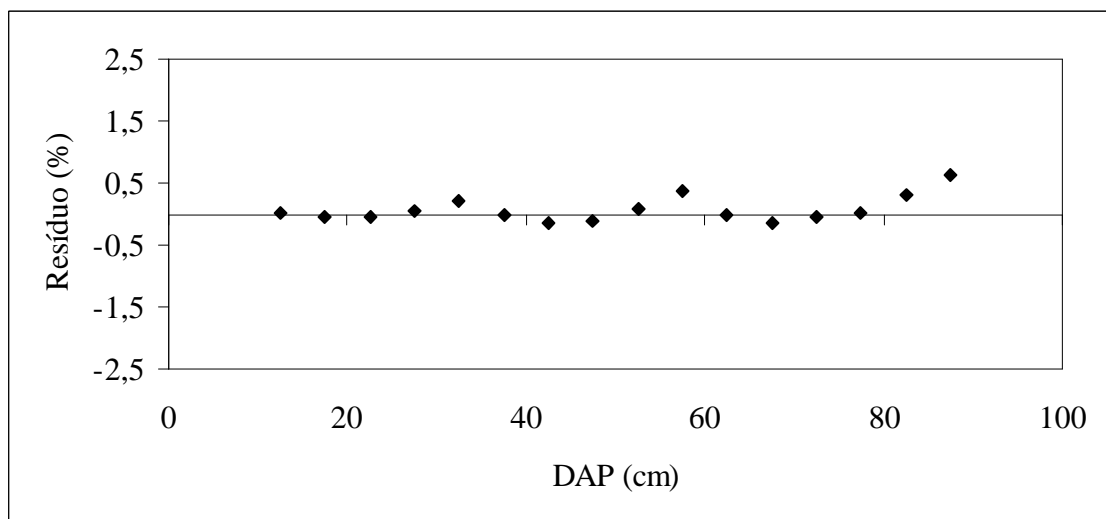


FIGURA 49 – Resíduos em percentagem da distribuição de Quadros



5.3.8 Spline Cúbica

A Função Spline Cúbica ajustada encontra-se na equação 112. Na tabela 19 encontram-se as frequências observadas a campo e esperadas pela função ajustada.

$$f(x) = \frac{1}{356,86} \begin{cases} 10,0181x - 0,0221x^3 & 0,00 \leq x < 12,5 \\ -192,1371 + 56,1310x - 3,6890x^2 + 0,0762x^3 & 12,5 \leq x < 17,5 \\ 351,6834 - 37,0953x + 1,6382x^2 - 0,0252x^3 & 17,5 \leq x < 22,5 \\ 56,6441 + 2,2432x - 0,1102x^2 + 0,0007x^3 & 22,5 \leq x < 27,5 \\ -730,7902 + 88,1452x - 3,2339x^2 + 0,0385x^3 & 27,5 \leq x < 32,5 \\ 2336,5778 - 194,9965x + 5,4781x^2 - 0,0508x^3 & 32,5 \leq x < 37,5 \\ -593,0473 + 39,3735x - 0,7717x^2 + 0,0047x^3 & 37,5 \leq x < 42,5 \\ -219,7268 + 13,0215x - 0,1517x^2 - 0,0001x^3 & 42,5 \leq x < 47,5 \\ -1494,8698 + 93,5568x - 1,8471x^2 + 0,0118x^3 & 47,5 \leq x < 52,5 \\ -3415,1777 + 203,2887x - 3,9373x^2 + 0,0250x^3 & 52,5 \leq x < 57,5 \\ 8938,2826 - 441,2397x + 7,2719x^2 - 0,0399x^3 & 57,5 \leq x < 62,5 \\ -4404,6249 + 199,2199x - 2,9754x^2 + 0,0147x^3 & 62,5 \leq x < 67,5 \\ -1443,1357 + 67,5982x - 1,0255x^2 + 0,0051x^3 & 67,5 \leq x < 72,5 \\ 1656,9289 - 60,6804x + 0,7439x^2 - 0,0030x^3 & 72,5 \leq x < 77,5 \\ 647,1401 - 21,5918x + 0,2395x^2 - 0,0009x^3 & 77,5 \leq x < 82,5 \\ 951,6681 - 32,6655x + 3,737x^2 - 0,0014x^3 & 82,5 \leq x < 90,0 \end{cases} \quad \text{EQ. 112}$$

A média e a variância estão apresentadas nas Equações 113 e 114 respectivamente. Observa-se que a função Spline Cúbica subestimou a média e superestimou o desvio padrão.

$$\mu = 28,94 \text{ cm} \quad \text{EQ. 113}$$

$$\sigma_x^2 = 314,72 \text{ cm}^2 \quad \text{EQ. 114}$$

TABELA 19 – Distribuição diamétrica, frequências acumuladas esperadas pela função Spline Cúbica ajustada e diferença absoluta entre as distribuições acumuladas esperada e observada.

$Li-ls$	Xi	f_{obs}	F_{obs}	F_{esp}	$ F_{esp}-F_{obs} $
10 -15	12,5	82	82	145,07	63,07
15 -20	17,5	69	151	204,48	53,48
20 -25	22,5	59	210	254,88	44,88
25 -30	27,5	49	259	296,76	37,76
30 -35	32,5	41	300	332,62	32,62
35 -40	37,5	48	348	373,35	25,35
40 -45	42,5	50	398	415,81	17,81
45 -50	47,5	43	441	452,32	11,32
50 -55	52,5	29	470	477,16	7,16
55 -60	57,5	17	487	492,26	5,26
60 -65	62,5	16	503	505,67	2,67
65 -70	67,5	11	514	515,07	1,07
70 -75	72,5	5	519	519,48	0,48
75 -80	77,5	2	521	521,25	0,25
80 -85	82,5	1	522	522,14	0,14
> 85	87,5	1	523	523,00	0,00
Total		523			

Ao observar a figura 50, na qual se encontra a representação gráfica da distribuição diamétrica da Araucária e a função Spline Cúbica a ela ajustada, nota-se que ela passa sobre todos os pontos observados. Por esse motivo, esperava-se que ela fosse aderente à distribuição de dados amostrais, o que não é confirmado pelo teste de Kolmogorov-Smirnov ($KS = 0,1206$). É interessante notar que nesse caso a função Spline Cúbica representou muito bem a moda amostrada, como pode ser visto na figura 51. Nota-se que, diferentemente do caso da castanheira, aqui a função Spline teve uma tendência muito mais “comportada”. Com isso pode-se anexar à Spline mais uma característica: pode ser considerada uma boa alternativa quando os dados não geram pontos de inflexão seguidos.

FIGURA 50 – Distribuição diamétrica da Araucária e função Spline Cúbica a ela ajustada.

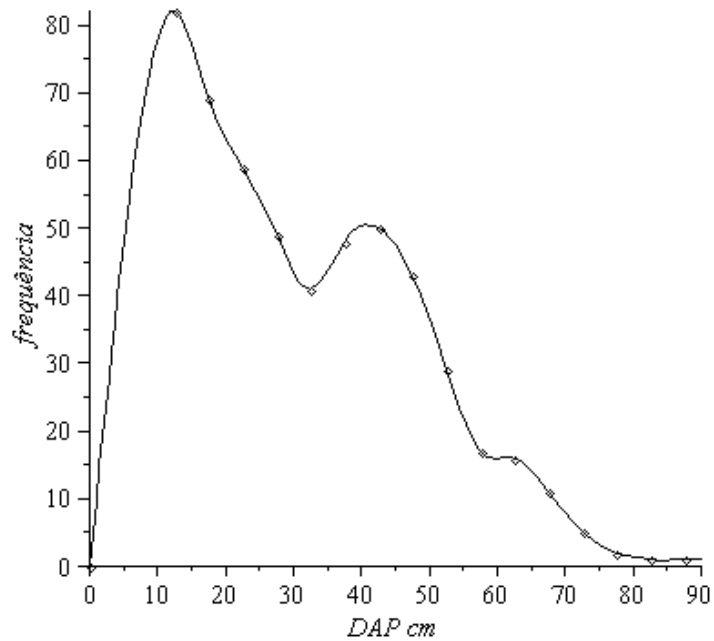
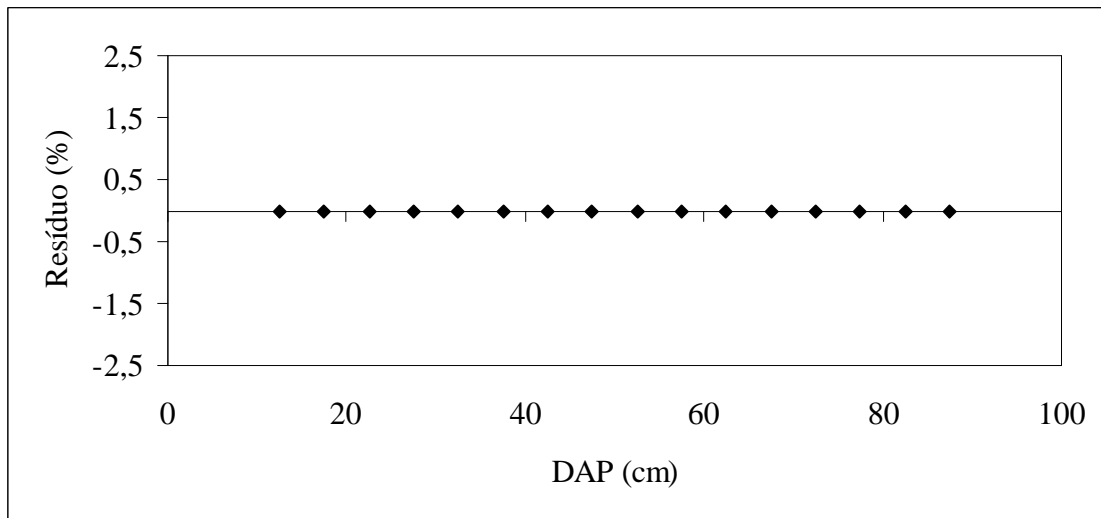


FIGURA 51 – Resíduos em percentagem da função Spline Cúbica



5.3.9 Modelo Proposto

Na proposta do presente estudo foi apresentada uma distribuição multimodal. Uma vez que o conjunto de dados observados por Pizatto (1999) apresentou duas modas ($X = 12,5$ e $42,5$ cm), será proposto um modelo com essa característica, bimodal (EQUAÇÃO 115).

$$f(x) = \frac{x \cdot (a \cdot x^2 - b)^2 + x \cdot (c \cdot x^2 - d)^2}{e + f \cdot (a \cdot x^2 - b)^2 \cdot (c \cdot x^2 - d)^2} \quad \text{EQ. 115}$$

Os coeficientes b e d são dependentes de a e c , respectivamente, e são obtidos por meio das modas. Os cálculos e resultados estão apresentados nas Equações 116 e 117.

$a \cdot x^2 - b$ Parte da função que gera a segunda moda. Será definida com PGM1 (Parte Geradora da Moda 1).

$a \cdot x^2 - b = 0$ Quando o valor de x for igual a 42,5 a PGM1 deverá ser igual a zero.

$$b = a \cdot 42,5^2 = 1806,25 \cdot a \quad \text{EQ. 116}$$

$c \cdot x^2 - d$ PGM2

$c \cdot x^2 - d = 0$ Quando o valor de x for igual a 12,5 a PGM1 deverá ser igual a zero.

$$d = c \cdot 12,5^2 = 156,25 \quad \text{EQ. 117}$$

Os valores dos coeficientes calculados por meio de regressão não linear foram $a = 2,0518$, $b = 320,5903$, $c = -4,3133$, $d = -7790,8129$, $e = 9741790$ e $f = 1,5182 \cdot 10^{-7}$. Substituindo-se esses valores no modelo e realizando as simplificações pertinentes, tem-se a equação ajustada (EQUAÇÃO 118).

$$f(x) = \frac{x \cdot (2,0518 \cdot x^2 - 320,5903)^2 + x \cdot (-4,3133 \cdot x^2 - 7790,8129)^2}{9741790 + 1,5182 \cdot 10^{-7} \cdot (2,0518 \cdot x^2 - 320,5903)^2 \cdot (-4,3133 \cdot x^2 - 7790,8129)^2}$$

EQ. 118

Integrando essa função no intervalo de zero até mais infinito obtém-se uma constante $k = 2817,7912$, que representa a área entre a curva ajustada e o eixo x (DAP). O inverso de k deve ser multiplicado pela função ajustada, a fim de se encontrar a FDP (EQUAÇÃO 119).

$$FDP(x) = \frac{0,0003548879026 \cdot x \cdot (2,0518 \cdot x^2 - 320,5903)^2 + x \cdot (-4,3133 \cdot x^2 - 7790,8129)^2}{9741790 + 1,5182 \cdot 10^{-7} \cdot (2,0518 \cdot x^2 - 320,5903)^2 \cdot (-4,3133 \cdot x^2 - 7790,8129)^2} \quad \text{EQ. 119}$$

A representação gráfica da Equação 118, bem como os pontos referentes aos dados observados, encontram-se na figura 52. Os resíduos encontram-se na figura 53.

FIGURA 52 – Distribuição diamétrica da Araucária e modelo proposto a ela ajustada.

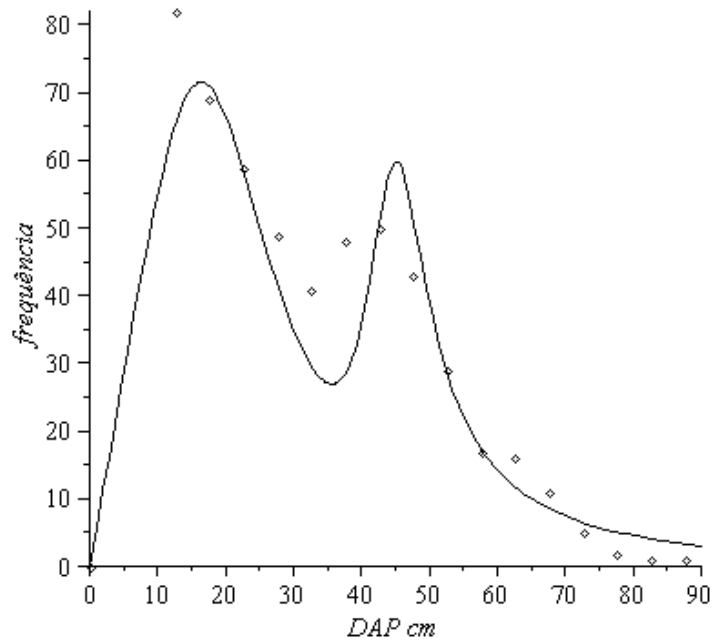
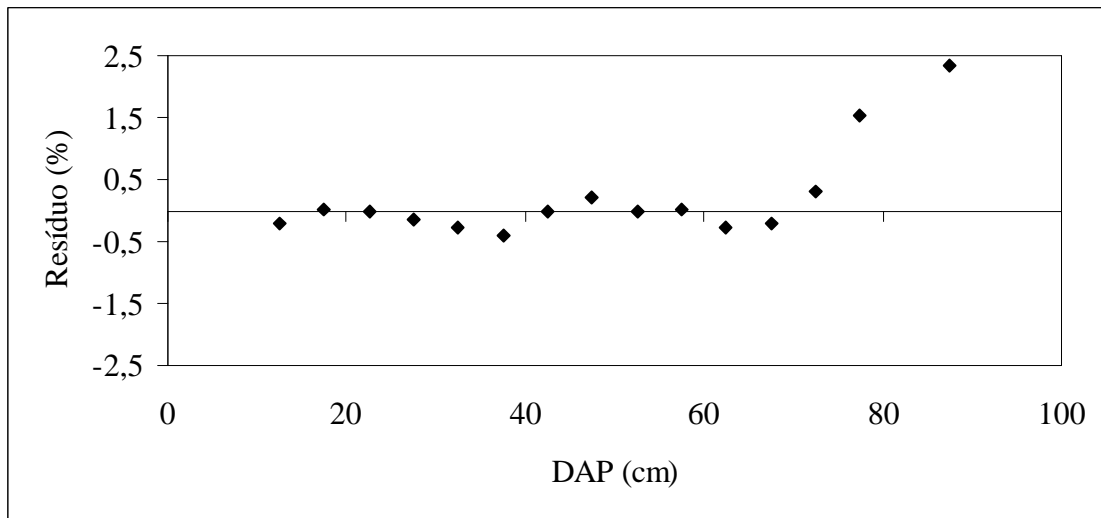


FIGURA 53 – Resíduos em percentagem modelo proposto



A integral do produto entre a FDP e X , no intervalo de zero a mais infinito, ou seja, no seu domínio, é igual à média aritmética estimada pelo modelo ($\mu = 37,82$ cm) (EQUAÇÃO 120).

$$\int_0^{+\infty} FDP(X) \cdot X \cdot dX = \frac{0,0003548879026 \cdot X^2 \cdot (2,0518 \cdot X^2 - 320,5903)^2 + X \cdot (-4,3133 \cdot X^2 - 77908129)^2}{9741790 + 1,51818 \cdot 10^{-7} \cdot (2,0518 \cdot X^2 - 320,5903)^2 \cdot (-4,3133 \cdot X^2 - 77908129)^2} \text{ EQ. 120}$$

$$= 37,82$$

A variância é obtida fazendo-se a diferença entre a integral do produto da FDP por X^2 , e a média elevada ao quadrado. No ajuste em questão, a variância estimada pelo modelo foi

de 56.863,56 cm², ou seja, desvio padrão de 238,46 cm. Esse valor elevado deve-se ao fato de o modelo proposto superestimar as classes diamétricas superiores.

O modelo proposto foi o único considerado aderente aos dados, por meio do teste de Kolmogorov-Smirnov (0,0616), uma vez que o valor calculado foi menor que o crítico, conforme tabela 20.

TABELA 20 – Distribuição diamétrica, frequências acumuladas esperadas pelo modelo proposto ajustado e diferença absoluta entre as distribuições acumuladas esperada e observada.

Li-ls	Xi	f _{obs}	F _{obs}	F _{esp}	F _{esp} -F _{obs}
10 -15	12,5	82	82	111,93	29,93
15 -20	17,5	69	151	177,05	26,05
20 -25	22,5	59	210	231,55	21,55
25 -30	27,5	49	259	270,63	11,63
30 -35	32,5	41	300	298,62	1,38
35 -40	37,5	48	348	325,72	22,28
40 -45	42,5	50	398	371,44	26,56
45 -50	47,5	43	441	419,02	21,98
50 -55	52,5	29	470	446,33	23,67
55 -60	57,5	17	487	462,75	24,25
60 -65	62,5	16	503	473,80	29,20
65 -70	67,5	11	514	481,80	32,20
70 -75	72,5	5	519	487,86	31,14
75 -80	77,5	2	521	492,60	28,40
80 -85	82,5	1	522	496,40	25,60
> 85	87,5	1	523	522,40	0,60
Total		523			

5.3.10 Resumo dos resultados dos modelos

Na tabela 21 estão apresentados de maneira resumida os coeficientes dos modelos ajustados e os valores de KS e Syx, a fim de se realizar uma análise comparativa. Os coeficientes da função Spline Cúbica e a distribuição de Quadros não serão inseridos nessa tabela, uma vez que são em grande número. Os valores de referência do teste de Kolmogorov-Smirnov são 0,0841 e 0,1008, para 95 e 99% de significância, respectivamente. Considerando esse teste não paramétrico pode-se afirmar que apenas o modelo proposto foi aderente à distribuição diamétrica da Araucária. No cálculo do erro padrão da estimativa recalculado em porcentagem é considerado o número de coeficientes ajustados em cada modelo (Normal = 2, Exponencial = 1, Beta = 4, Gama = 2, Weibull = 3, Quadros = 10, Spline = 62 e modelo proposto = 6), em que o modelo proposto apresentou o menor valor (será mais aderente aquele modelo que apresentar valor mais baixo).

TABELA 21 – Comparação entre os modelos para a Araucária.

Modelo	Coeficientes						KS	Syx%
Normal	$\mu = 32,77$	$\sigma = 16,32$					0,1370**	7,24
Exponencial	$\beta = 32,77$						0,1604**	13,19
Beta	$a=10$	$b=90$	$\alpha = 0,86$	$\beta = 2,46$			0,7224**	26,42
Gama	$\alpha = 4,03$	$\beta = 8,12$					0,1218**	8,34
Weibull	$a=10,00$	$b=19,61$	$c=0,83$				0,1020**	7,35
Quadros							0,1872**	14,12
Spline							0,1206**	9,43
Modelo proposto	$a=2,0518$	$b=320,5903$	$c=-4,33133$	$d=-7790,8129$	$e=9.741.790$	$f=1,5182.10^{-7}$	0,0616 ^{ns}	5,64

Valores seguidos de * e ** indicam significância a 95% e 99% de probabilidade, respectivamente, e ns indica não significância.

5.4 *Araucaria angustifolia*

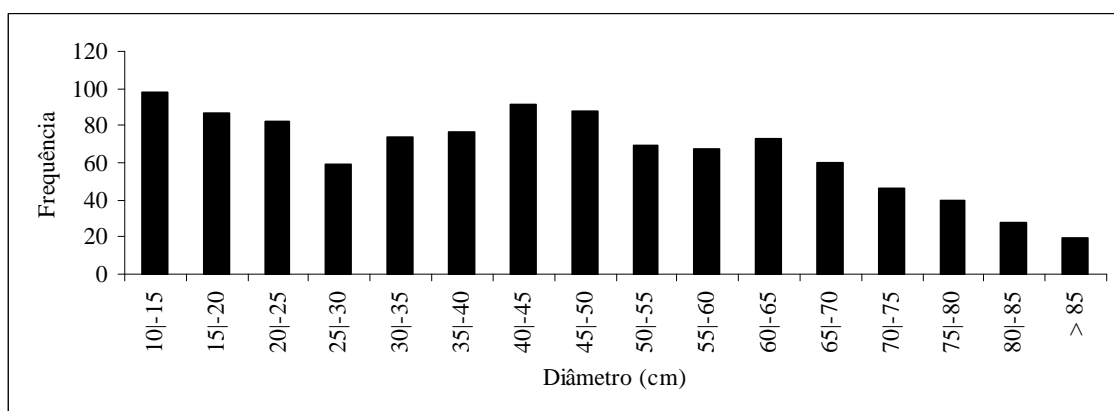
Criada em 1942, a Floresta Nacional (FLONA) de Irati é uma unidade de conservação que está localizada no Centro Fernandes Pinheiro, nas coordenadas geográficas 25° 25' de latitude sul, 50° 36' de longitude oeste e 25°17' de latitude sul, 50°30' de longitude oeste (IBAMA, 2010)

De acordo com Köppen o clima microrregional é do tipo Cfb Mesotérmico caracterizado pelos verões frescos, geadas severas e seca (SPVS, 1996). A altitude média do local é de 885 metros, o solo é classificado como podzólico vermelho-amarelo e apresenta acidez média. O relevo varia de irregular a plano e a vegetação predominante na área de estudos é a Floresta Ombrófila Mista, de acordo com a classificação do IBGE (IBAMA, 2010). As parcelas permanentes foram instaladas entre os anos de 2000 e 2001, em um fragmento de Ombrófila Mista existente naquela FLONA, explorada até a sua criação. Em virtude disso, pode ser classificada como uma vegetação primária alterada ou uma secundária em estágio avançado (RODE, 2008). As parcelas permanentes foram instaladas por professores do Departamento de Engenharia Florestal da UNICENTRO e engloba uma área de 25 hectares, formada por 25 blocos de 1 hectare cada um. Cada bloco é dividido em 4 parcelas quadradas de 50 m x 50 m (2.500 m²) e cada parcela é subdividida em faixas de 10 m x 50 m (500 m²). Em 2002 todas as árvores com diâmetro a altura do peito (DAP) igual ou superior a 10 cm foram medidas. Na tabela 22, estão apresentados os dados sumariados, bem como as frequências em todas as classes diamétricas. Foi elaborado um histograma (FIGURA 54), para uma melhor visualização da distribuição empírica dos dados.

TABELA 22 – Frequências absolutas (f_{obs}) e acumuladas (F_{obs}) observadas por classe diamétrica da Araucária em uma área de 25 ha em Irati, PR

Li ls	X_i	f_{obs}	F_{obs}
10 15	12,5	98	98
15 20	17,5	87	185
20 25	22,5	82	267
25 30	27,5	59	326
30 35	32,5	74	400
35 40	37,5	77	477
40 45	42,5	91	568
45 50	47,5	88	656
50 55	52,5	69	725
55 60	57,5	67	792
60 65	62,5	73	865
65 70	67,5	60	925
70 75	72,5	46	971
75 80	77,5	40	1011
80 85	82,5	28	1039
> 85	87,5	19	1058
Total		1058	

FIGURA 54 – Estrutura populacional da Araucária em uma área de 25 ha em um Fragmento de Floresta Ombrófila Mista em Irati



5.4.1 Estatística descritiva

A média aritmética calculada empiricamente para dados agrupados foi de 43,53 cm de DAP e variância de 436,53 cm², logo o desvio padrão foi de 20,89 cm. Pode-se observar um padrão diferente ao da Araucária de São João do Triunfo, com duas modas bem definidas, mas com uma tendência diferente no início. Em comparação com o fragmento de São João do Triunfo, a de Irati resultou em uma média aritmética maior.

Serão apresentados a seguir os resultados das distribuições Normal, Exponencial, Beta, Gama, Weibull, Quadros, Spline e do modelo proposto na presente pesquisa. A seguir

será feita a comparação entre os modelos e a discussão sobre suas aderências ou não aos dados observados.

5.4.2 Normal

A distribuição Normal ajustada encontra-se na equação 121, enquanto que a média e a variância são apresentadas nas equações 122 e 123, respectivamente. Na tabela 23 encontram-se as frequências observadas a campo e esperadas pela função ajustada. Os valores da média e desvio padrão estimados pela distribuição são iguais aos amostrados.

$$f(x) = \begin{cases} \frac{1}{\sqrt{2\pi} \cdot 20,89} \cdot e^{-\frac{1}{2} \left(\frac{x-43,53}{20,89} \right)^2} & -\infty < x < +\infty \end{cases} \quad \text{EQ. 121}$$

$$\mu = 43,53 \text{ cm} \quad \text{EQ. 122}$$

$$\sigma_x^2 = 436,53 \text{ cm}^2 \quad \text{EQ. 123}$$

TABELA 23 – Distribuição diamétrica, frequências acumuladas esperadas pela distribuição Normal ajustada e diferença absoluta entre as distribuições acumuladas esperada e observada.

Li -ls	Xi	f _{obs}	F _{obs}	F _{esp}	F _{esp} -F _{obs}
10 -15	12,5	98	98	72,77	25,23
15 -20	17,5	87	185	112,62	72,38
20 -25	22,5	82	267	166,24	100,76
25 -30	27,5	59	326	234,38	91,62
30 -35	32,5	74	400	316,19	83,81
35 -40	37,5	77	477	408,94	68,06
40 -45	42,5	91	568	508,29	59,71
45 -50	47,5	88	656	608,81	47,19
50 -55	52,5	69	725	704,88	20,12
55 -60	57,5	67	792	791,60	0,40
60 -65	62,5	73	865	865,55	0,55
65 -70	67,5	60	925	925,12	0,12
70 -75	72,5	46	971	970,45	0,55
75 -80	77,5	40	1011	1003,02	7,98
80 -85	82,5	28	1039	1025,14	13,86
> 85	87,5	19	1058	1039,32	18,68
Total		1058			

As diferenças absolutas entre a frequência acumulada e esperada pela curva de Gauss ajustada servem para realizar o teste de Kolmogorov-Smirnov (KS). De acordo com a equação 49, os valores de KS tabelados são 0,0591 e 0,0709, a 95% e 99%, respectivamente. O valor de KS calculado foi de 0,0952, mostrando que a distribuição não foi aderente.

Encontra-se na figura 55 o gráfico dos valores observados, bem como a curva ajustada. Pode-se notar que a curva ajusta-se bem às frequências de algumas classes, mas não

tem flexibilidade para explicar as classes iniciais. Outro bom indicador encontra-se na figura 56, onde estão apresentados os resíduos percentuais entre valores estimados pelo modelo e observados a campo.

FIGURA 55 – Distribuição diamétrica da Araucária e curva Normal a ela ajustada.

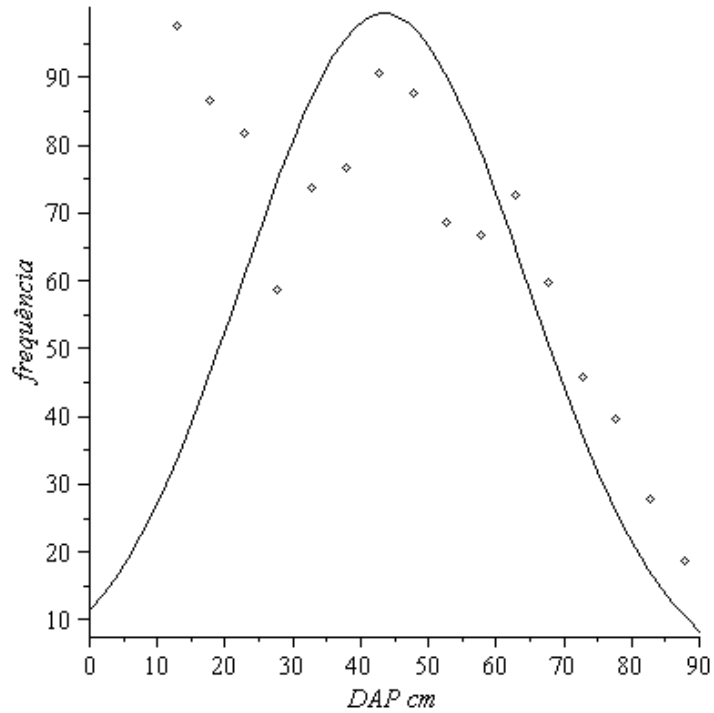
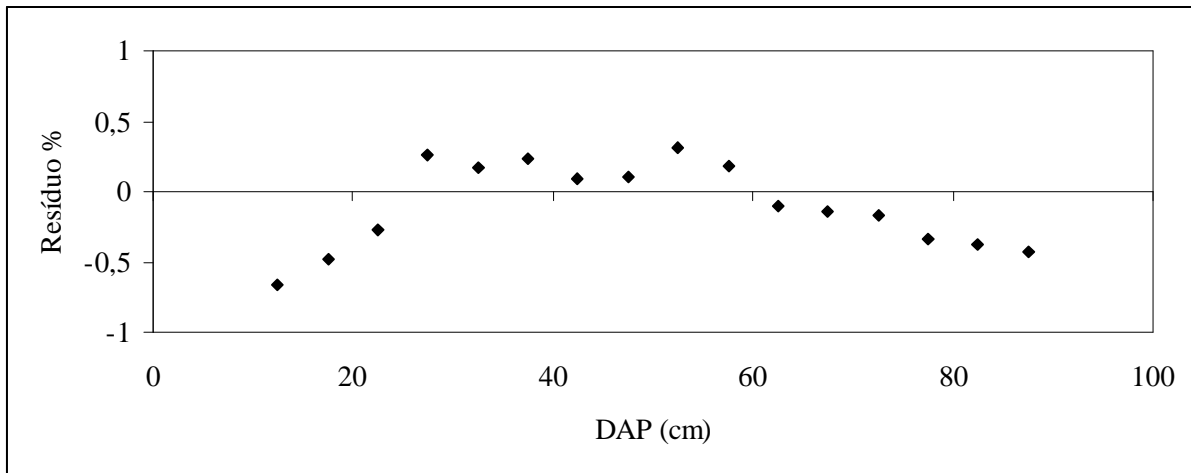


FIGURA 56 – Resíduos em percentagem da distribuição Normal



5.4.3 Exponencial

A distribuição Exponencial ajustada encontra-se na equação 124, enquanto que a média e a variância são apresentadas nas equações 125 e 126, respectivamente. Na tabela 24 encontram-se as frequências observadas a campo e esperadas pela função ajustada. O valor da

média estimada pela distribuição é igual ao amostrado, enquanto que o desvio padrão é igual à média, logo, muito diferente do observado.

$$f(x) = \begin{cases} \frac{1}{43,53} \cdot e^{\frac{-x}{43,53}} & \text{para } x \geq 0, \beta > 0 \\ 0 & \text{e.o.c.} \end{cases} \quad \text{EQ. 124}$$

$$\mu = 43,53 \text{ cm} \quad \text{EQ. 125}$$

$$\sigma_x^2 = 1894,47 \text{ cm}^2 \quad \text{EQ. 126}$$

TABELA 24 – Distribuição diamétrica, frequências acumuladas esperadas pela distribuição Exponencial ajustada e diferença absoluta entre as distribuições acumuladas esperada e observada.

Li-ls	Xi	f _{obs}	F _{obs}	F _{esp}	F _{esp} -F _{obs}
10 -15	12,5	98	98	264,11	166,11
15 -20	17,5	87	185	350,26	165,26
20 -25	22,5	82	267	427,07	160,07
25 -30	27,5	59	326	495,54	169,54
30 -35	32,5	74	400	556,58	156,58
35 -40	37,5	77	477	610,99	133,99
40 -45	42,5	91	568	659,50	91,50
45 -50	47,5	88	656	702,75	46,75
50 -55	52,5	69	725	741,30	16,30
55 -60	57,5	67	792	775,67	16,33
60 -65	62,5	73	865	806,31	58,69
65 -70	67,5	60	925	833,62	91,38
70 -75	72,5	46	971	857,97	113,03
75 -80	77,5	40	1011	879,68	131,32
80 -85	82,5	28	1039	899,03	139,97
> 85	87,5	19	1058	916,28	141,72
Total		1058			

Comparando a razão da máxima diferença entre as frequências acumuladas observada e estimada pela distribuição Exponencial (KS = 0,1602) com os valores tabelados, observa-se que ela não foi aderente, como já se esperava, uma vez que essa distribuição não apresenta modas.

Encontra-se na figura 57 o gráfico dos valores observados, bem como a curva ajustada. Com o auxílio dessa imagem é possível ratificar que a distribuição não é capaz de representar de forma satisfatória a distribuição diamétrica da Araucária. Outro bom indicador encontra-se na figura 58, onde estão apresentados os resíduos percentuais entre valores estimados pelo modelo e observados a campo.

FIGURA 57 – Distribuição diamétrica da Araucária e distribuição Exponencial a ela ajustada.

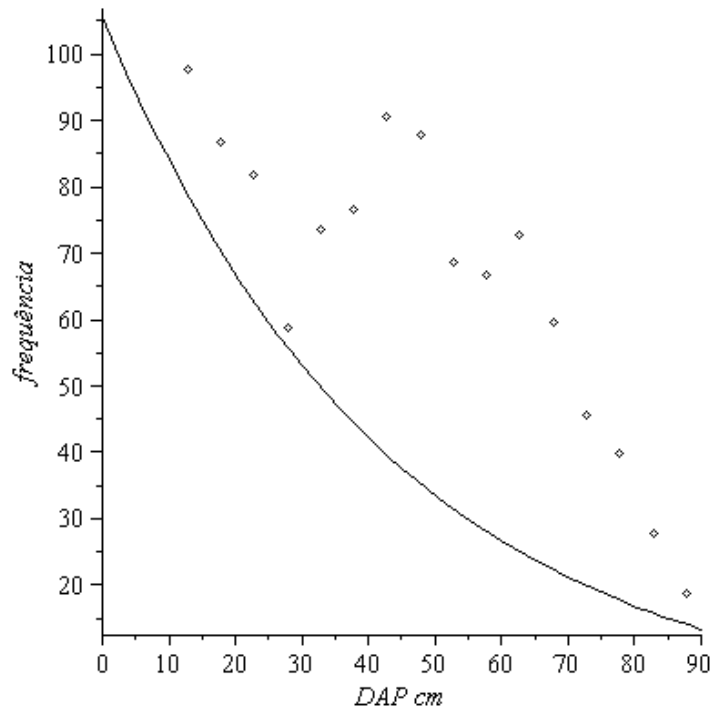
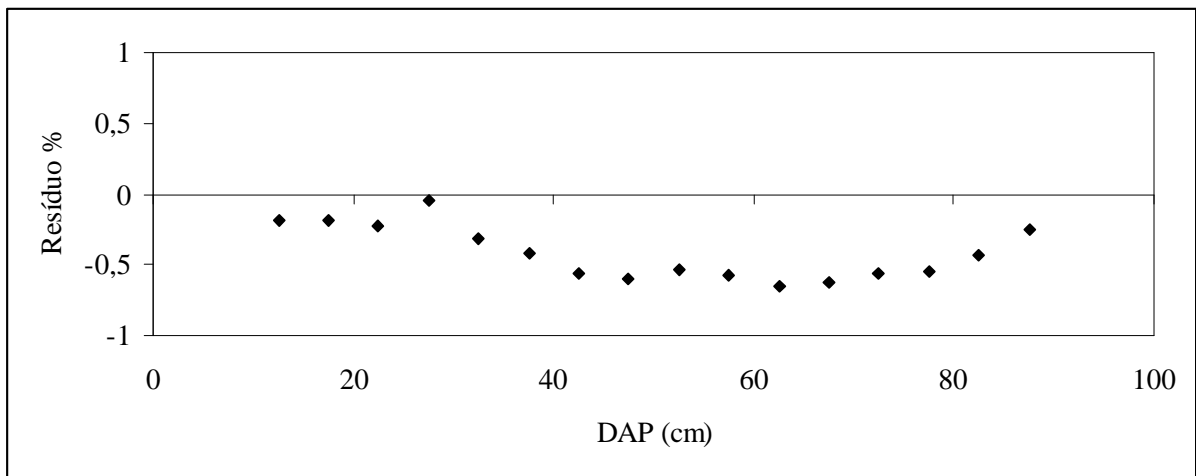


FIGURA 58 – Resíduos em porcentagem da distribuição Exponencial



5.4.4 Beta

A distribuição Beta ajustada encontra-se na equação 127, enquanto que a média e a variância são apresentadas nas equações 128 e 129, respectivamente. Na tabela 25 encontram-se as frequências observadas a campo e esperadas pela função ajustada. De acordo com as fórmulas analíticas pode-se verificar que estimativa da média e do desvio padrão da distribuição foi muito próxima da observada.

$$f(x) = \begin{cases} 0,007683 \cdot (x-10)^{-0,1196} \cdot (90-x)^{0,2479} & \text{se } 10 < x < 90 \\ 0 & \text{e.o.c.} \end{cases} \quad \text{EQ. 127}$$

$$\mu = 43,09 \text{ cm} \quad \text{EQ. 128}$$

$$\sigma_x^2 = 496,21 \text{ cm}^2 \quad \text{EQ. 129}$$

TABELA 25 – Distribuição diamétrica, frequências acumuladas esperadas pela distribuição Beta ajustada e diferença absoluta entre as distribuições acumuladas esperada e observada.

Li-ls	Xi	f _{obs}	F _{obs}	F _{esp}	F _{esp} -F _{obs}
10 -15	12,5	98	98	12,13	85,87
15 -20	17,5	87	185	47,99	137,01
20 -25	22,5	82	267	91,19	175,81
25 -30	27,5	59	326	139,43	186,57
30 -35	32,5	74	400	191,76	208,24
35 -40	37,5	77	477	247,66	229,34
40 -45	42,5	91	568	306,82	261,18
45 -50	47,5	88	656	369,06	286,94
50 -55	52,5	69	725	434,32	290,68
55 -60	57,5	67	792	502,59	289,41
60 -65	62,5	73	865	573,98	291,02
65 -70	67,5	60	925	648,69	276,31
70 -75	72,5	46	971	727,11	243,89
75 -80	77,5	40	1011	809,92	201,08
80 -85	82,5	28	1039	898,53	140,47
> 85	87,5	19	1058	996,92	61,08
Total		1058			

O valor de KS (0,2751) calculado para a distribuição Beta foi maior que o tabelado, logo a distribuição não pode ser considerada aderente aos dados observados da Araucária.

Encontra-se na figura 59 o gráfico dos valores observados, bem como a curva ajustada. Apesar de a distribuição Beta ser altamente flexível, não há combinação de coeficientes que a faça assumir uma forma bimodal, como é o caso da presente pesquisa. Outro bom indicador encontra-se na figura 60, onde estão apresentados os resíduos percentuais entre valores estimados pelo modelo e observados a campo.

FIGURA 59 – Distribuição diamétrica da Araucária e distribuição Beta a ela ajustada.

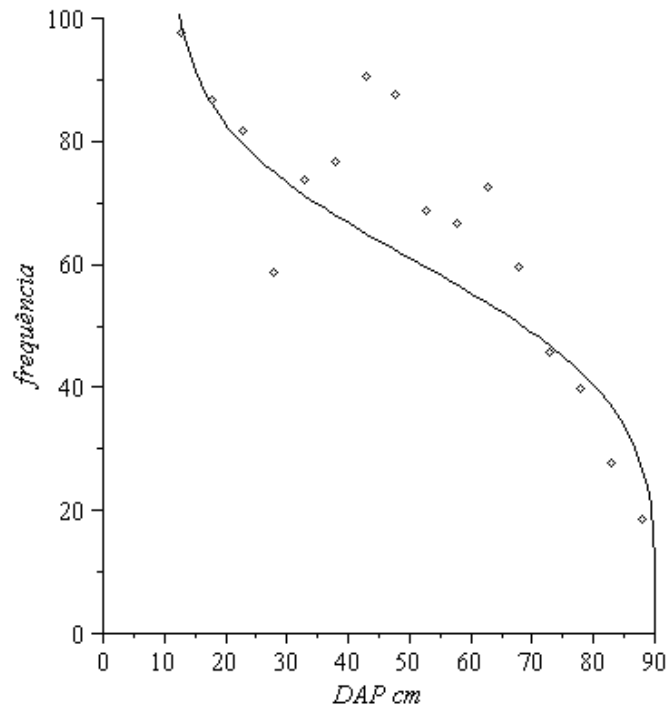
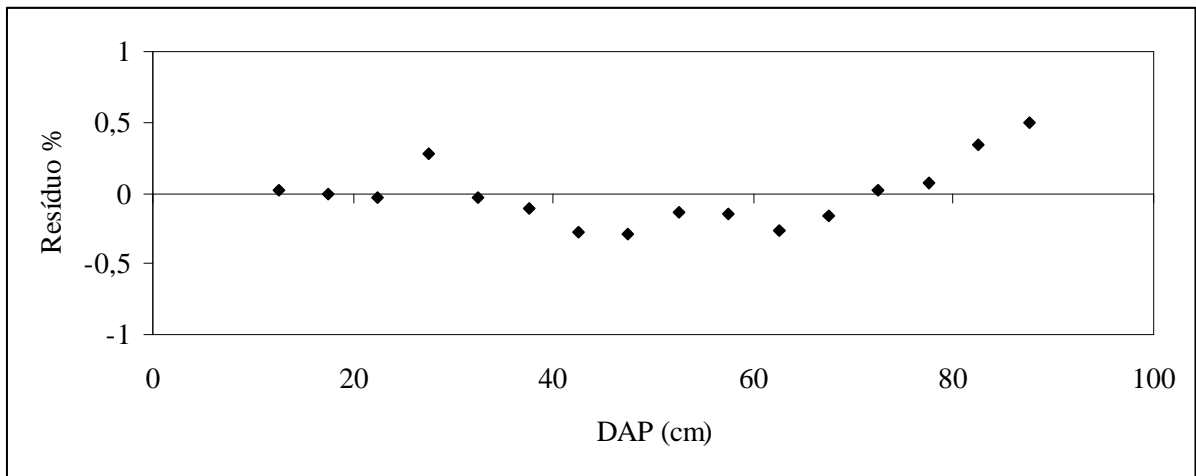


FIGURA 60 – Resíduos em percentagem da distribuição Beta



5.4.5 Gama

A distribuição Gama ajustada encontra-se na equação 130, enquanto que a média e a variância são apresentadas nas equações 131 e 132, respectivamente. Na tabela 26 encontram-se as frequências observadas a campo e esperadas pela função ajustada. A média e o desvio padrão são iguais aos observados em campo, em função do método de ajuste.

$$f(x) = \begin{cases} \frac{x^{3,34} \cdot e^{-x/10,03}}{10,03^{4,33} \cdot \Gamma(4,33)} & \text{para } x \geq 0 \\ 0 & \text{e.o.c,} \end{cases} \quad \text{EQ. 130}$$

$$\mu = 43,53 \text{ cm} \quad \text{EQ. 131}$$

$$\sigma_x^2 = 496,21 \text{ cm}^2 \quad \text{EQ. 132}$$

TABELA 26 – Distribuição diamétrica, frequências acumuladas esperadas pela distribuição Gama ajustada e diferença absoluta entre as distribuições acumuladas esperada e observada.

Li ls	Xi	f _{obs}	F _{obs}	F _{esp}	F _{esp} -F _{obs}
10 15	12,5	98	98	25,14	72,86
15 20	17,5	87	185	73,70	111,30
20 25	22,5	82	267	150,45	116,55
25 30	27,5	59	326	248,65	77,35
30 35	32,5	74	400	358,38	41,62
35 40	37,5	77	477	470,01	6,99
40 45	42,5	91	568	576,02	8,02
45 50	47,5	88	656	671,53	15,53
50 55	52,5	69	725	754,08	29,08
55 60	57,5	67	792	823,06	31,06
60 65	62,5	73	865	879,12	14,12
65 70	67,5	60	925	923,64	1,36
70 75	72,5	46	971	958,29	12,71
75 80	77,5	40	1011	984,81	26,19
80 85	82,5	28	1039	1004,79	34,21
> 85	87,5	19	1058	1019,66	38,34
Total		1058			

A distribuição Gama não foi aderente aos dados (KS = 0,1102). Apesar da flexibilidade desse modelo, não há combinação de valores para seus coeficientes que a torne bimodal. Dessa maneira já era esperado que o ajuste não seria satisfatório, como pode ser observado na figura 61, onde está apresentado o gráfico dos valores observados, bem como a curva ajustada. Outro bom indicador encontra-se na figura 62, onde estão apresentados os resíduos percentuais entre valores estimados pelo modelo e observados a campo.

FIGURA 61 – Distribuição diamétrica da Araucária e distribuição Gama a ela ajustada.

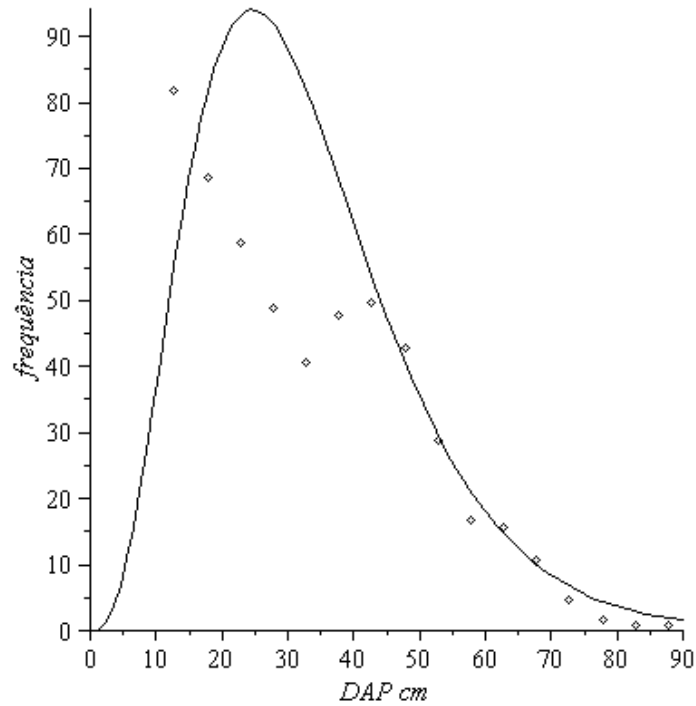
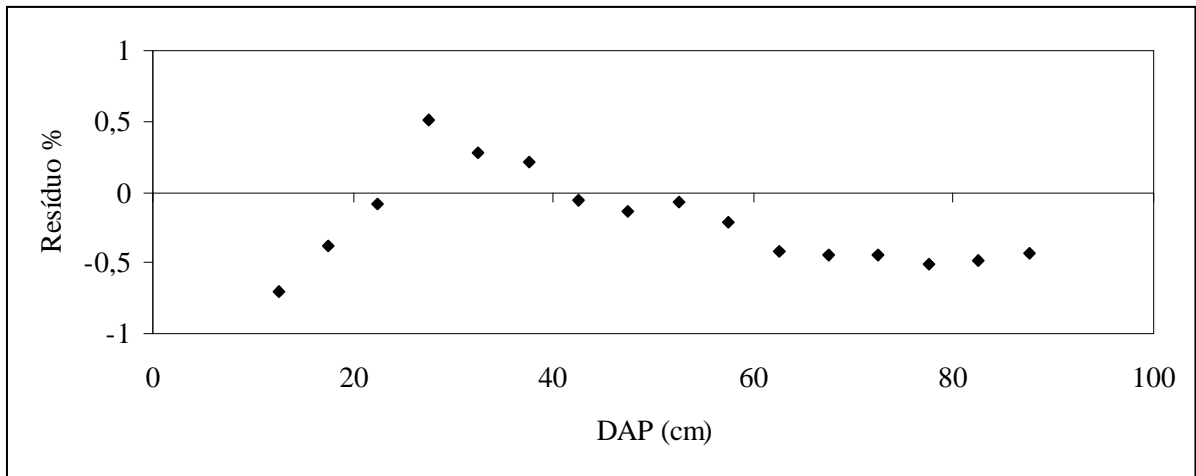


FIGURA 62 – Resíduos em percentagem da distribuição Gama



5.4.6 Weibull

A distribuição Weibull ajustada encontra-se na equação 133. Na tabela 27 encontram-se as frequências observadas a campo e esperadas pela função ajustada. A média e o desvio padrão estimados pela distribuição de Weibull foram muito próximos dos observados a campo.

$$f(x) = \begin{cases} 0,04 \left(\frac{x-8,56}{38,92} \right)^{0,6453} \cdot e^{-\left(\frac{x-8,56}{38,92} \right)^{1,6453}} & \text{se } x \geq a, b > 0 \text{ e } c > 0 \\ 0 & \text{e.o.c.} \end{cases} \quad \text{EQ. 133}$$

A média e a variância estão apresentadas nas Equações 134 e 135 respectivamente.

$$\mu = 43,37 \text{ cm} \quad \text{EQ. 134}$$

$$\sigma_x^2 = 471,52 \text{ cm}^2 \quad \text{EQ. 135}$$

TABELA 27 – Distribuição diamétrica, frequências acumuladas esperadas pela distribuição Weibull ajustada e diferença absoluta entre as distribuições acumuladas esperada e observada.

Li-ls	Xi	f _{obs}	F _{obs}	F _{esp}	F _{esp} -F _{obs}
10 -15	12,5	98	98	24,18	73,82
15 -20	17,5	87	185	90,04	94,96
20 -25	22,5	82	267	178,43	88,57
25 -30	27,5	59	326	278,76	47,24
30 -35	32,5	74	400	383,12	16,88
35 -40	37,5	77	477	485,58	8,58
40 -45	42,5	91	568	581,85	13,85
45 -50	47,5	88	656	669,15	13,15
50 -55	52,5	69	725	745,97	20,97
55 -60	57,5	67	792	811,78	19,78
60 -65	62,5	73	865	866,82	1,82
65 -70	67,5	60	925	911,86	13,14
70 -75	72,5	46	971	947,97	23,03
75 -80	77,5	40	1011	976,36	34,64
80 -85	82,5	28	1039	998,29	40,71
> 85	87,5	19	1058	1014,93	43,07
Total		1058			

O valor de KS (0,0898) mostra que esse modelo não se aderiu à distribuição diamétrica da Araucária por pouco, uma vez que o valor de referência a 99% é igual a 0,1008. Isso mostra a alta versatilidade desse modelo, embora a tendência desejada (bimodal) não tenha sido obtida. Pode ser observado na figura 63 o ajuste da distribuição Weibull aos dados observados dos diâmetros da Araucária. Outro bom indicador encontra-se na figura 64, onde estão apresentados os resíduos percentuais entre valores estimados pelo modelo e observados a campo.

FIGURA 63 – Distribuição diamétrica da Araucária e distribuição Weibull a ela ajustada.

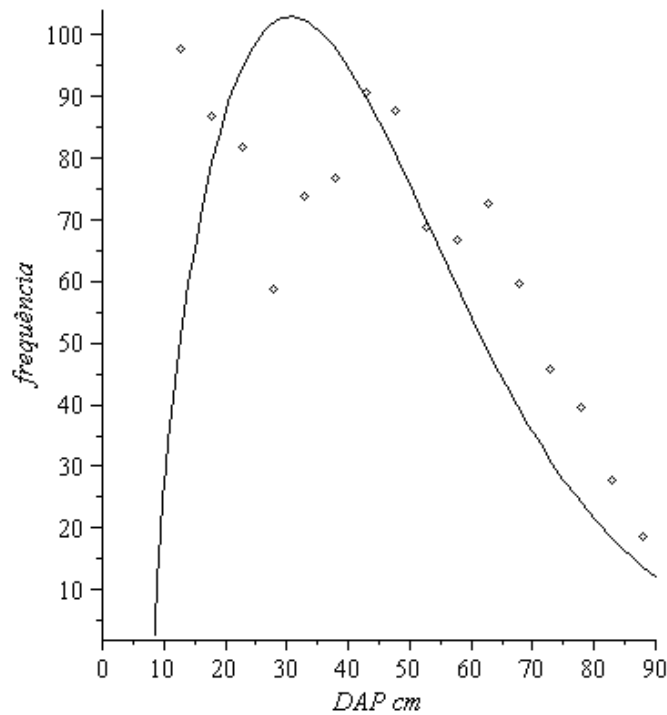
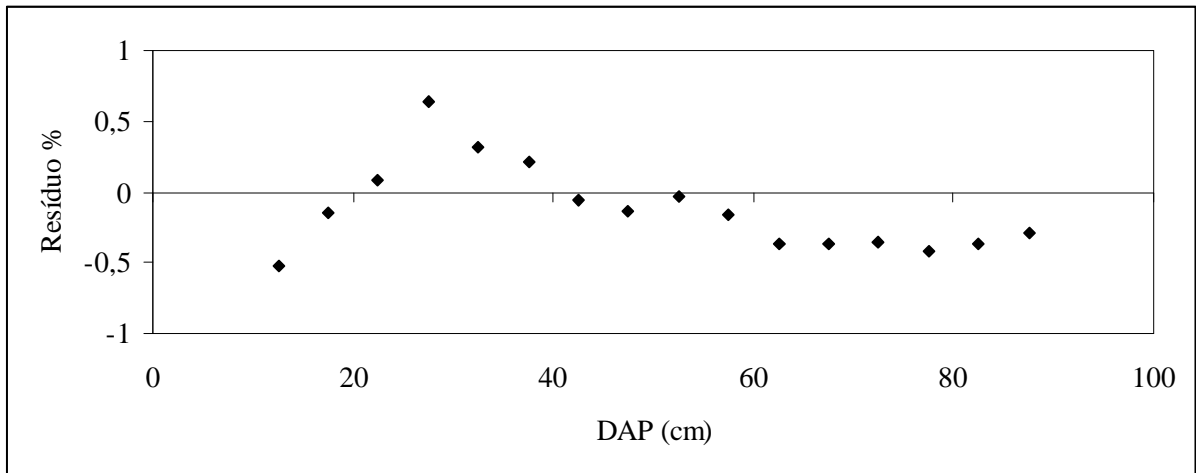


FIGURA 64 – Resíduos em percentagem da distribuição Weibull



5.4.7 Quadros

A distribuição de Quadros ajustada encontra-se na equação 136. Na tabela 28 encontram-se as frequências observadas a campo e esperadas pela função ajustada. O valor da média estimado pela função foi abaixo do observado, enquanto que a variância foi superestimada. Isso ocorre em decorrência das funções obtidas na primeira e na terceira parte. A distribuição de Quadros será eficiente, na forma como foi concebida, as modas estiverem nas classes intermediárias, o que não foi observado na presente pesquisa. Já foi mencionado, em tópicos anteriores, que os coeficientes muito próximos de zero não podem ser retirados, pois acarretam perda significativa de flexibilidade, tornando o modelo inapto. Apesar de toda

a flexibilidade apresentada pelo modelo polinomial, ele não foi capaz de representar a moda intermediária.

$$f(x) = \frac{1}{5724,498} \begin{cases} 0,000331x^5 & \text{se } 0 < x < 10 \\ -2,6070 \cdot 10^{-7} x^5 + 9,1728 \cdot 10^5 x^4 - 0,011935x^3 & \\ +0,68862x^2 - 17,492x + 233,2 & \text{se } 10 \leq x \leq 90 \text{ EQ. 136} \\ \frac{9,82 \cdot 10^{10}}{x^5} & \text{se } x > 90 \\ 0 & \text{e.o.c} \end{cases}$$

A média e a variância estão apresentadas nas Equações 137 e 138 respectivamente.

$$\mu = 48,11 \text{ cm} \quad \text{EQ. 137}$$

$$\sigma_x^2 = 896,02 \text{ cm}^2 \quad \text{EQ. 138}$$

TABELA 28 – Distribuição diamétrica, frequências acumuladas esperadas pela distribuição de Quadros ajustada e diferença absoluta entre as distribuições acumuladas esperada e observada.

Li -ls	Xi	f _{obs}	F _{obs}	F _{esp}	F _{esp} -F _{obs}
10 -15	12,5	98	98	104,15	6,15
15 -20	17,5	87	185	180,52	4,48
20 -25	22,5	82	267	249,43	17,57
25 -30	27,5	59	326	317,10	8,90
30 -35	32,5	74	400	386,67	13,33
35 -40	37,5	77	477	458,99	18,01
40 -45	42,5	91	568	533,33	34,67
45 -50	47,5	88	656	607,95	48,05
50 -55	52,5	69	725	680,60	44,40
55 -60	57,5	67	792	748,93	43,07
60 -65	62,5	73	865	810,86	54,14
65 -70	67,5	60	925	864,74	60,26
70 -75	72,5	46	971	909,55	61,45
75 -80	77,5	40	1011	944,92	66,08
80 -85	82,5	28	1039	971,10	67,90
> 85	87,5	19	1058	1056,00	2,00
Total		1058			

Pelo teste de Kolmogorov-Smirnov pode-se concluir que o modelo foi aderente a 95% de probabilidade aos valores observados (KS = 0,0642). Pode ser observado na figura 65 o ajuste da distribuição de Quadros aos dados observados dos diâmetros da Araucária. Na figura 66 estão apresentados os resíduos percentuais entre valores estimados pelo modelo e

observados a campo. Observa-se um ponto em que a curva faz um “bico”, como se diz habitualmente em cálculo diferencial. Esse não é um problema, como já foi discutido anteriormente. A questão primordial, para fins estatísticos, é a integrabilidade, que nesse caso é garantida.

FIGURA 65 – Distribuição diamétrica da Araucária e distribuição de Quadros a ela ajustada.

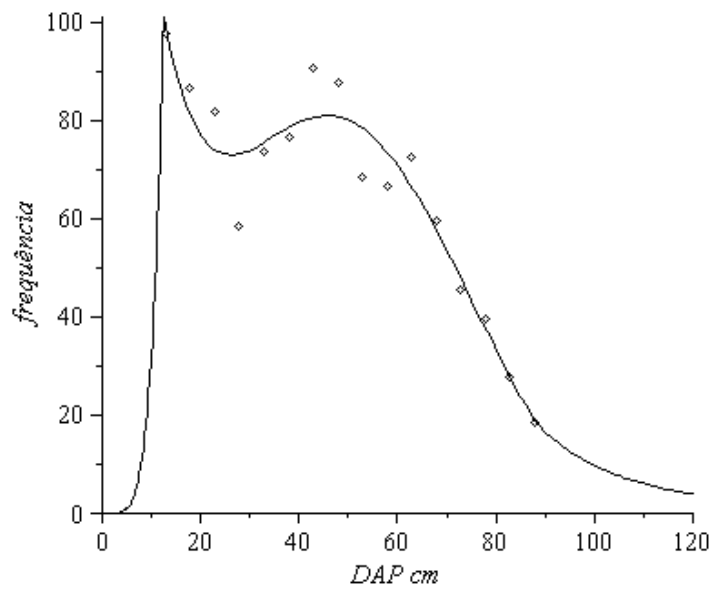
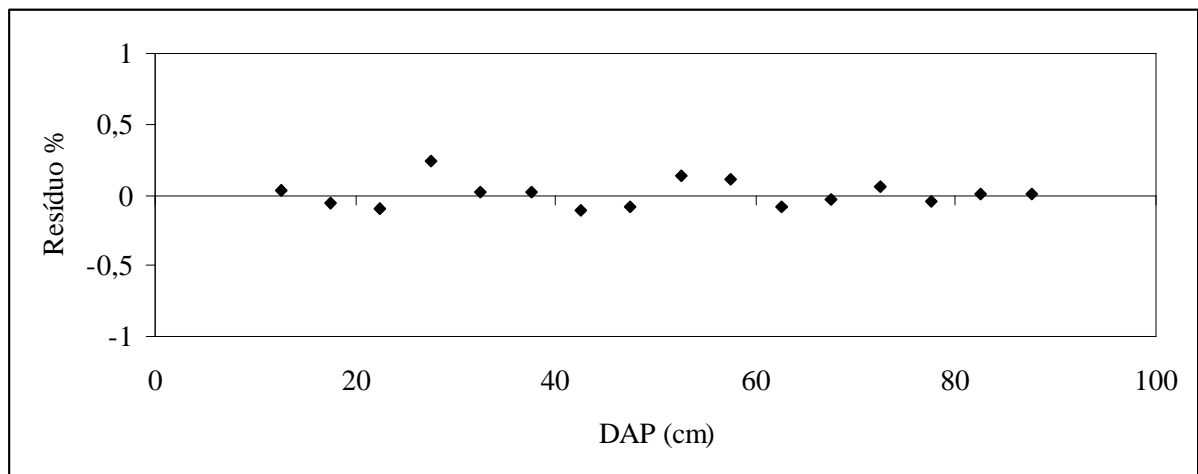


FIGURA 66 – Resíduos em porcentagem da distribuição de Quadros



5.4.8 Spline Cúbica

A Função Spline Cúbica ajustada encontra-se na equação 139. Na tabela 29 encontram-se as frequências observadas a campo e esperadas pela função ajustada.

$$f(x) = \frac{1}{5811,86} \begin{cases} 11,8442x + 0,0256x^3 & 0,00 \leq x < 12,5 \\ -266,7181 + 75,8565x - 5,1210x^2 + 1,109x^3 & 12,5 \leq x < 17,5 \\ 983,0788 - 138,3944x + 7,1219x^2 - 1,223x^3 & 17,5 \leq x < 22,5 \\ -2529,7134 + 329,9779x - 13,6946x^2 + 1,861x^3 & 22,5 \leq x < 27,5 \\ 4964,9580 - 487,6226x + 16,0363x^2 - 1,742x^3 & 27,5 \leq x < 32,5 \\ -4822,2050 + 415,8078x - 11,7616x^2 + 1,109x^3 & 32,5 \leq x < 37,5 \\ 5516,3753 - 411,2786x + 10,2941x^2 - 0,852x^3 & 37,5 \leq x < 42,5 \\ -1476,5289 + 82,3382x - 1,3204x^2 + 0,0059x^3 & 42,5 \leq x < 47,5 \\ -8300,5119 + 513,3266x - 10,3939x^2 + 0,0696x^3 & 47,5 \leq x < 52,5 \\ 4691,9862 - 229,1019x + 3,7476x^2 - 0,0202x^3 & 52,5 \leq x < 57,5 \\ 12396,0156 - 631,0513x + 10,7380x^2 - 0,0607x^3 & 57,5 \leq x < 62,5 \\ -13942,8631 + 633,2149x - 9,4902x^2 + 0,0471x^3 & 62,5 \leq x < 67,5 \\ -4405,2214 + 209,3197x - 3,2103x^2 + 0,0161x^3 & 67,5 \leq x < 72,5 \\ 16872,6847 - 671,1453x + 8,9341x^2 - 0,0397x^3 & 72,5 \leq x < 77,5 \\ -15877,0672 + 596,5870x - 7,4238x^2 + 0,0307x^3 & 77,5 \leq x < 82,5 \\ 7475,8463 - 252,6099x + 2,8695x^2 - 0,0109x^3 & 82,5 \leq x < 90,0 \end{cases} \text{ EQ. 139}$$

A média e a variância estão apresentadas nas Equações 140 e 141 respectivamente. Observa-se que a função Spline Cúbica subestimou a média e superestimou o desvio padrão.

$$\mu = 40,16 \text{ cm} \quad \text{EQ. 140}$$

$$\sigma_x^2 = 510,57 \text{ cm}^2 \quad \text{EQ. 141}$$

TABELA 29 – Distribuição diamétrica, frequências acumuladas esperadas pela função Spline Cúbica ajustada e diferença absoluta entre as distribuições acumuladas esperada e observada.

$Li-ls$	X_i	f_{obs}	F_{obs}	F_{esp}	$ F_{esp}-F_{obs} $
10 -15	12,5	98	98	183,76	85,76
15 -20	17,5	87	185	263,87	78,87
20 -25	22,5	82	267	336,91	69,91
25 -30	27,5	59	326	393,12	67,12
30 -35	32,5	74	400	459,18	59,18
35 -40	37,5	77	477	430,26	46,74
40 -45	42,5	91	568	612,18	44,18
45 -50	47,5	88	656	691,48	35,48
50 -55	52,5	69	725	755,19	30,19
55 -60	57,5	67	792	816,60	24,60
60 -65	62,5	73	865	882,01	17,01
65 -70	67,5	60	925	936,67	11,67
70 -75	72,5	46	971	979,01	8,01
75 -80	77,5	40	1011	1014,98	3,98
80 -85	82,5	28	1039	1040,71	1,71
> 85	87,5	19	1058	1058,00	0,00
Total		1058			

Ao observar a figura 67, na qual se encontra a representação gráfica da distribuição diamétrica da Araucária e a função Spline Cúbica a ela ajustada, nota-se que ela passa sobre todos os pontos observados. De acordo com o teste de Kolmogorov-Smirnov ($KS = 0,0811$), a função não pode ser considerada aderente. É interessante notar que a função Spline Cúbica representou três modas amostradas, como pode ser visto na figura 68, mas não foi aderente.

FIGURA 67 – Distribuição diamétrica da Araucária e função Spline Cúbica a ela ajustada.

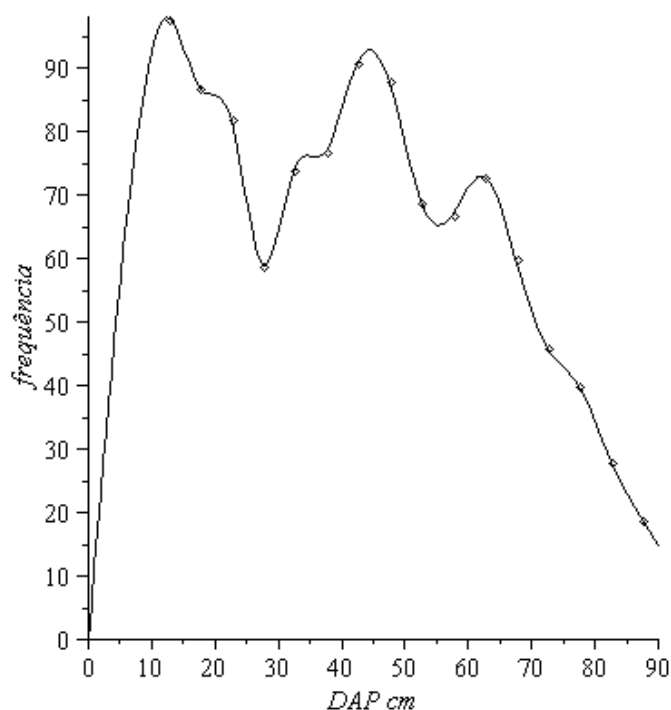
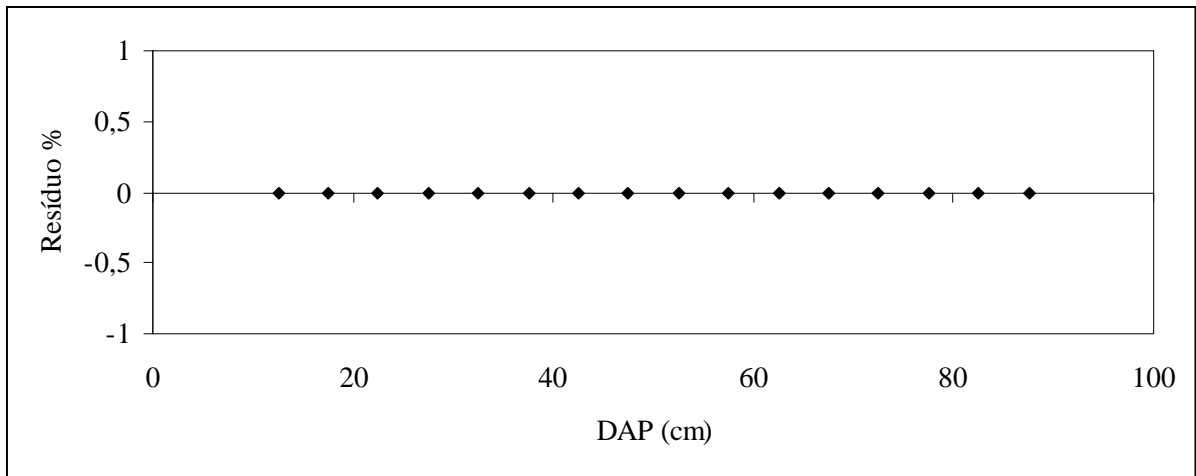


FIGURA 68 – Resíduos em percentagem da função Spline Cúbica



5.4.9 Modelo Proposto

A distribuição diamétrica apresentada pela Araucária, conforme observado por Orellana (2009) seguiu uma conformação diferente dos outros dois apresentados anteriormente. Por esse motivo foi ajustado um modelo que intercepta o eixo das ordenadas em um ponto diferente de zero (EQUAÇÃO 142). Por meio da estatística descritiva pode-se observar três modas, entretanto foram consideradas duas para o ajuste.

$$f(x) = \frac{a \cdot x^4 + b \cdot x + c}{d + e(f \cdot x^2 - g)^2 \cdot (h \cdot x^2 - i)^2} \quad \text{EQ. 142}$$

Os coeficientes g e i são dependentes de f e h , respectivamente, e são obtidos por meio das modas. Os cálculos estão apresentados a seguir, cujos resultados estão apresentados nas Equações 143 e 144.

$$f \cdot x^2 - g$$

Parte da função que gera a primeira moda. Será definida com PGM1 (Parte Geradora da Moda 1). Foi considerada uma moda hipotética negativa em -20 a fim de contemplar a tendência descrita anteriormente.

$$f \cdot x^2 - g = 0$$

Quando o valor de x for igual a -20 a PGM1 deverá ser igual a zero.

$$g = -f \cdot 20^2 = -400 \cdot f$$

EQ. 143

$$h \cdot x^2 - i$$

PGM2

$$h \cdot x^2 - i = 0$$

Embora tenha se admitido que quando o valor de x fosse igual a 42,5 a PGM2 deverá ser igual a zero para formar uma moda, o ajuste melhorou significativamente quando utilizou-se 20.

$$i = h \cdot 20^2 = 400 \cdot h \quad \text{EQ. 144}$$

Os valores dos coeficientes calculados por meio de regressão não linear foram $a = 55.799,48$, $b = -7,14 \cdot 10^9$, $c = 4,37 \cdot 10^{11}$, $d = 3,66 \cdot 10^9$, $e = 3,01 \cdot 10^{-6}$, $f = 1,90$ e $h = 1,90$. Substituindo-se esses valores no modelo e realizando as simplificações pertinentes, tem-se a equação ajustada (EQUAÇÃO 145).

$$f(x) = \frac{55.799,48 \cdot x^4 - 7.14 \cdot 10^9 \cdot x + 4,37 \cdot 10^{11}}{3,66 \cdot 10^9 + 3,01 \cdot 10^{-6} (1,90 \cdot x^2 + 760,16)^2 \cdot (1,90 \cdot x^2 - 760,16)^2} \quad \text{EQ. 145}$$

Integrando essa função no intervalo de 10 até mais 90 obtém-se uma constante $k = 5274,39$, que representa a área entre a curva ajustada e o eixo x (DAP) em seu domínio. O inverso de k deve ser multiplicado pela função ajustada, a fim de se encontrar a FDP (EQUAÇÃO 146).

$$FDP(x) = \frac{1}{5274,39} \cdot \frac{55.799,48 \cdot x^4 - 7.14 \cdot 10^9 \cdot x + 4,37 \cdot 10^{11}}{3,66 \cdot 10^9 + 3,01 \cdot 10^{-6} (1,90 \cdot x^2 + 760,16)^2 \cdot (1,90 \cdot x^2 - 760,16)^2} \quad \text{EQ. 146}$$

A representação gráfica da Equação 145, bem como os pontos referentes aos dados observados, encontram-se na figura 69. Os resíduos podem ser observados na figura 70.

FIGURA 69 – Distribuição diamétrica da Araucária e modelo proposto a ela ajustada.

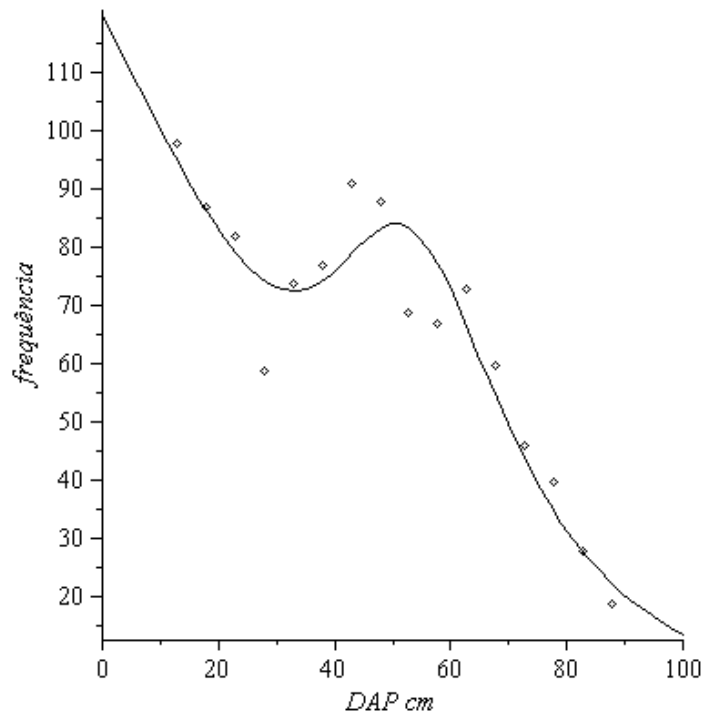
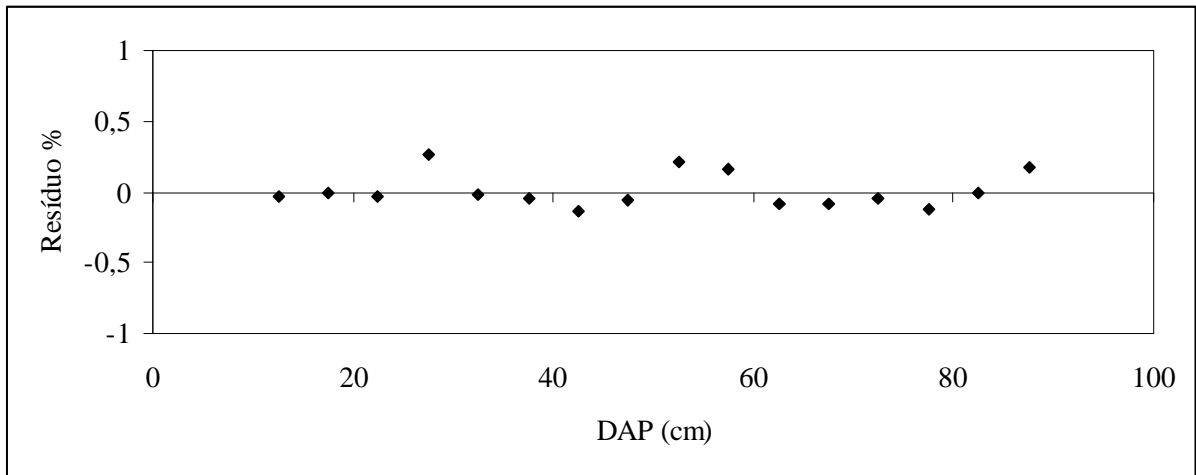


FIGURA 70 – Resíduos em percentagem modelo proposto



A integral do produto entre a FDP e x no intervalo de zero a mais infinito é igual à média aritmética estimada pelo modelo ($\mu = 60,46$ cm) (EQUAÇÃO 147).

$$\begin{aligned}
 & \int_0^{+\infty} FDP(x) \cdot x \cdot dx = \\
 & = \int_0^{+\infty} \frac{x}{5274,39} \cdot \frac{55.799,48 \cdot x^4 - 7.14 \cdot 10^9 \cdot x + 4,37 \cdot 10^{11}}{3,66 \cdot 10^9 + 3,01 \cdot 10^{-6} (1,90 \cdot x^2 + 760,16)^2 \cdot (1,90 \cdot x^2 - 760,16)^2} dx \quad \text{EQ. 147} \\
 & = 60,46 \text{ cm}
 \end{aligned}$$

A variância é obtida fazendo-se a diferença entre a integral do produto da FDP por x^2 ,

e a média elevada ao quadrado. No ajuste em questão, a variância estimada pelo modelo foi de 1597,88 cm², ou seja, desvio padrão de 238,46 cm.

Segundo o teste de aderência de Kolmogorov-Smirnov (0,0135) o modelo ajustou-se aos dados observados, representando as modas satisfatoriamente, conforme apresentado na tabela 30.

TABELA 30 – Distribuição diamétrica, frequências acumuladas esperadas pelo modelo proposto ajustado e diferença absoluta entre as distribuições acumuladas esperada e observada.

Li-ls	Xi	f _{obs}	F _{obs}	F _{esp}	F _{esp} -F _{obs}
10 -15	12,5	98	98	95,75	2,25
15 -20	17,5	87	185	182,82	2,18
20 -25	22,5	82	267	262,61	4,39
25 -30	27,5	59	326	337,34	11,34
30 -35	32,5	74	400	410,05	10,05
35 -40	37,5	77	477	484,21	7,21
40 -45	42,5	91	568	562,64	5,36
45 -50	47,5	88	656	645,49	10,51
50 -55	52,5	69	725	728,82	3,82
55 -60	57,5	67	792	806,32	14,32
60 -65	62,5	73	865	873,26	8,26
65 -70	67,5	60	925	928,27	3,27
70 -75	72,5	46	971	972,38	1,38
75 -80	77,5	40	1011	1007,49	3,51
80 -85	82,5	28	1039	1035,50	3,50
> 85	87,5	19	1058	1057,99	0,01
Total		1058			

5.4.10 Resumo dos resultados dos modelos

Na tabela 31 estão apresentados de maneira resumida os coeficientes dos modelos ajustados e os valores de KS e Syx, a fim de se realizar uma análise comparativa. Os coeficientes da função Spline Cúbica e a distribuição de Quadros não serão inseridos nessa tabela, uma vez que são em grande número. Os valores de referência do teste de Kolmogorov-Smirnov são 0,0841 e 0,1008, para 95 e 99% de probabilidade, respectivamente.

TABELA 31 – Comparação entre os modelos para a Araucária.

Modelo	Coeficientes				KS	Syx%
Normal	$\mu = 43,53$	$\sigma = 20,89$			0,0953**	2,38
Exponencial	$\beta = 43,53$				0,1602**	13,19
Beta	a=10	b=90	$\alpha = 0,88$	$\beta = 1,25$	0,2751**	26,42
Gama	$\alpha = 4,34$	$\beta = 10,03$			0,1102**	8,34
Weibull	a=8,56	b=38,92	c=1,65		0,0897**	7,36
Quadros					0,0642*	14,04
Spline					0,0811**	8,89
Modelo proposto					0,0135ns	1,50

Valores seguidos de * e ** indicam significância a 95% e 99% de probabilidade, respectivamente, e ns indica não significância.

5.5 COMPARAÇÃO ENTRE OS MODELOS (GERAL)

Em florestas plantadas são realizados desbastes, nas quais são retiradas algumas árvores para algum uso específico, como, por exemplo, postes. Quando isso é feito, pode surgir mais de uma moda. Logo, quando se deseja estimar com precisão a quantidade de madeira nesse plantio, é imprescindível que se utilize um modelo multimodal flexível.

Em florestas nativas quatro são os fatores determinantes para a formação de modas:

- Distúrbios naturais: como dinâmica de clareira, eventos climáticos como ventos, enchentes ou secas;
- Intervenção antrópica: com o corte seletivo da floresta para extração de madeira, seja por manejo sustentado ou não;
- Características inerentes à espécie: algumas espécies têm uma dinâmica de reprodução e estabelecimento próprios, como o caso apresentado na presente pesquisa, em que a Castanheira ocorre em “reboleiras” ou agrupamentos; e
- Aleatório: algumas modas podem ocorrer randomicamente, em casos em que não se consegue determinar com precisão o motivo de suas ocorrências.

A forma que a curva do modelo toma após o ajuste aos dados observados é inerente a cada espécie, área de amostragem e fatores como os descritos anteriormente. A primeira aplicação do modelo foi à distribuição diamétrica da castanheira, na qual foram observadas três modas. A ocorrência dessas modas pode estar relacionada principalmente com a dinâmica de reprodução e formação de agrupamentos de indivíduos dessa espécie clímax. Assim, a baixa disponibilidade de energia solar inibe o crescimento das plantas mais baixas, que podem permanecer no sub-bosque por um longo intervalo de tempo. Com a queda ou quebra de uma ou mais árvores, espaços são abertos no dossel, facilitando o crescimento de algumas plantas do sub-bosque simultaneamente, o que pode gerar uma moda em alguma classe diamétrica. Esse evento natural, conhecido como dinâmica de clareira, repete-se diversas vezes e pode

gerar duas ou mais modas. Considerou-se nesse conjunto de dados que a distribuição interceptaria a origem, ou seja, que não haveriam indivíduos com DAP igual a 0 cm.

A segunda aplicação do modelo foi à distribuição diamétrica da araucária, na qual foram observadas apenas duas modas. A explicação mais provável para o surgimento dessas modas está relacionada com a intervenção antrópica, uma vez que essa espécie foi amplamente explorada para fins comerciais, onde foram abatidas árvores selecionando-se os diâmetros desejados. Como pressuposto teórico, foi utilizado um modelo que intercepta a origem.

A diferença da segunda para a terceira aplicação está apenas na premissa teórica. Nesse último caso, foi utilizado um modelo que não passa pela origem, ou seja, que considera que há indivíduos com DAP igual a 0 cm. Nos três casos selecionados considerou-se que a função seria assintótica no eixo das abscissas (DAP). Biologicamente isso corresponde a supor que árvores com DAPs maiores do que os observados em campo possam ocorrer.

As diferentes formas que as funções densidade de probabilidade podem assumir, de acordo com a variação de seus coeficientes foram apresentadas no tópico referente aos materiais utilizados e métodos aplicados. Pode-se observar que a distribuição Exponencial possui uma forma definida, a saber, decrescente. Essa forma é comumente descrita na literatura como “J invertido”. Ela é muito utilizada e potencialmente viável quando se deseja representar conjuntos de dados diamétricos de regeneração natural de espécies arbóreas. Entretanto, uma vez que no presente trabalho estão sendo utilizadas distribuições multimodais, esse modelo torna-se impróprio para representar tais realidades biológicas.

A distribuição Normal, frequentemente utilizada para representar distribuições diamétricas, não resultou satisfatória em função de sua baixa flexibilidade. Sabe-se que esse modelo é amplamente aplicado em conjuntos de dados biológicos, mas que a sua forma pouco variável (unimodal e simétrica em relação à média) a torna excelente nos casos em que se está trabalhando com censo (quando todos os indivíduos da população são avaliados) ou quando se tem um padrão “normal”.

As distribuições de Weibull, Beta e Gama são mais flexíveis, mas não se adequam a conjuntos de dados com mais de uma moda. Cabe salientar que a função de Weibull tem sido amplamente aplicada no setor florestal desde 1973, quando Bailey e Dell a utilizaram para modelagem em florestas, considerando-a matematicamente simples e flexível e mais fácil de ser utilizada do que a distribuição Beta, pelo fato de não requerer integração numérica para o cálculo do número de indivíduos nas classes.

O modelo apresentado por Silva et al. (2003), em que foi ajustada uma função truncada, definida por três sentenças, à distribuição de alturas de jequitibá-rosa, representou um ganho substancial no estudo de dados com formas mais anômalas, como no caso da ocorrência de multimodalidade. Mesmo sendo mais flexível que os demais modelos, houve deficiência na qualidade do modelo, quando ajustado aos dados utilizados no presente estudo.

A função Spline é definida por n funções, portanto é um modelo truncado. Ela não foi originalmente desenvolvida para cálculos de probabilidade, mas, devido à sua alta flexibilidade, surgiu o interesse de testá-la uma vez mais. Os resultados não foram satisfatórios, tanto na questão do ajuste, quanto na questão da sua interpretação biológica. Nela não se pode realizar uma extrapolação para diâmetros maiores do que os observados, uma vez que o comportamento do polinômio final pode gerar tendência para o infinito positivo ou negativo, o que não tem sentido em florestas.

Em todos os casos apresentados na literatura pertinente foi observado um comportamento unimodal ou decrescente. Nesses casos os modelos clássicos e o de Quadros funcionam bem, aderindo aos dados de forma satisfatória. Mas quando observam-se modas bem definidas, esses modelos falham, uma vez que não foram desenvolvidos para esse fim. O único modelo teoricamente capaz de representar satisfatoriamente essa tendência é o de Quadros. Entretanto, ele se aderiu satisfatoriamente apenas aos dados no caso da Castanheira, segundo o teste de Kolmogorov-Smirnov.

A função Spline não foi aderente aos dados nos três casos testados, bem como não representa a tendência biológica da distribuição diamétrica: interceptar ou não a origem, ser assintótica ao eixo das abscissas quando o DAP tender para infinito positivo e ter tantas modas quantas forem observadas. Cabe salientar que tão importante quanto ser aderente, o modelo deve representar a tendência natural de ocorrência das plantas, mesmo quando a floresta for manejada. Logo, funções altamente flexíveis, mas que ignoram o comportamento biológico, devem ser desconsideradas.

O modelo aqui apresentado foi aderente às três distribuições diamétricas, tendo representado os conjuntos variáveis de dados, dois de Araucária e um de Castanheira. Isso mostra que o modelo foi eficiente.

Ao aumentar a flexibilidade e melhorar a aderência, algumas classes diamétricas superiores foram superestimadas. Dessa forma a variância aumenta significativamente.

6. CONCLUSÕES

A distribuição Exponencial não tem aderência com os dados diamétricos da castanheira e os dois casos de araucária observados. Isso se deve à sua forma pouco flexível, uma vez que ela pode assumir apenas a tendência decrescente ou “J” invertido. Esse modelo tem sua importância no estudo de regeneração natural, mas jamais poderá ser empregado em pesquisas nas quais ocorram modas.

Os modelos Normal, Beta, Gama e Weibull não são aderentes às distribuições diamétricas observadas no presente estudo. Essas funções não devem ser utilizadas quando os dados apresentam duas ou mais modas. Nesses casos eles poderão apresentar estimativas errôneas e tendências que não condizem com a realidade biológica.

Apesar de o modelo de Quadros ser mais flexível do que os citados acima, ele não é capaz de representar satisfatoriamente a tendência observada na presente pesquisa, exceto no caso da Castanheira. As sentenças que definem a função nas porções inicial e final são elaboradas de acordo com o comportamento observado. As classes centrais estão sujeitas ao ajuste de um polinômio de grau n , que pode ou não seguir a tendência biológica observada. Quando o grau do polinômio é aumentado, a fim de se obter flexibilização, o modelo deixa de seguir forma bi ou trimodal observada empiricamente.

A função Spline não se adere aos dados do presente estudo, apesar de passar sobre todos os pontos observados, pois as probabilidades são calculadas por meio das áreas e não dos valores esperados pela função. Essa função desconsidera completamente a tendência biológica. A função pode passar pela origem, se esse ponto pertencer ao conjunto de dados. A função gera muitas modas, as quais, na maioria dos casos, não existem. Ela não tende a zero quando x tende a infinito, devendo, desta maneira, ser utilizada apenas para interpolação e nunca para extrapolação. Ela tenderá para infinito positivo ou negativo quando o diâmetro tender a infinito, o que inviabiliza seu uso em extrapolações de predição de ocorrência de indivíduos nas classes diamétricas superiores.

O modelo proposto pertence a uma família de funções que atendem à necessidade básica requerida, ou seja, são capazes de apresentar mais de uma moda. A classe diamétrica onde ela ou elas ocorrerão, bem como quantas existirão, são facilmente implementadas no modelo.

Nos casos da Araucária, considerou-se que quando o DAP for igual a zero, não há indivíduos. Já no caso da Castanheira determinou-se que haveria indivíduos nesse ponto. O modelo proposto pode ou não interceptar o eixo das ordenadas no ponto zero, ou seja, na

origem. O que muda de um caso para o outro é apenas um coeficiente independente (não multiplicado por x) no denominador: com ele a função não passará pela origem e sem ele, passará.

A função ajustada é sempre positiva. Isso é fundamental, uma vez que não existe probabilidade de ocorrência negativa.

O modelo proposto ajustado é integrável em todo domínio da função, ou seja, os reais positivos, por cálculo numérico, o que é necessário em função da avaliação de uma eventual ocorrência de árvores com diâmetros maiores do que aqueles amostrados. Logo, o cálculo das probabilidades é também obtido por cálculo numérico. Isso não torna o modelo melhor nem pior, uma vez que são inúmeros os programas que realizam integração numérica atualmente.

A adição ao denominador de uma constante independente garante a continuidade do modelo proposto frente a eventuais pontos de indeterminação ($0/0$, zero sobre zero) ou inexistência (constante dividido por zero).

Apesar de não ser uma premissa originalmente requerida, ao se desenvolver o modelo, a função encontrada resultou em diferenciabilidade em todo o seu domínio. Operacionalmente isso não trás grandes ganhos, tanto do ponto de vista matemático, quanto com relação ao ajuste.

A média e a variância podem ser obtidas por meio de integrais, de acordo com a teoria de distribuições probabilísticas. Salienta-se que tal condição é conseguida utilizando-se uma diferença de pelo menos quatro unidades entre o numerador e o denominador, ou seja, o polinômio do numerador deve ter grau menor ou igual a $n - 4$ e a função do denominador n .

Uma única sentença define o modelo. Pode-se dizer que o modelo proposto é a razão entre um polinômio de grau n e a adição de uma constante com o produto de PGMs, sendo que a quantidade de PGMs é igual ao número de modas.

O modelo é assintótico ao eixo x , ou seja, tende para zero quando x tende para infinito. Isso garante que exista a possibilidade de ocorrência de indivíduos nas classes diamétricas superiores, mas que esta probabilidade tende a zero.

Sugere-se que testes adicionais com conjuntos multimodais sejam conduzidos para a avaliação de sua eficácia.

Para avaliar a eficácia do modelo proposto, sugere-se que esse seja aplicado a outros conjuntos de dados, como a distribuição de alturas do Jequitibá branco e de outras espécies nativas do Brasil, bem como a distribuições diamétricas de povoamentos florestais desbastados.

7. REFERÊNCIAS

- ABREU, E. C. R. **Modelagem para prognose precoce do volume por classe diamétrica para *Eucalyptus grandis***. Lavras, 2000. 69f. Dissertação (Mestrado em Ciências Florestais), Universidade Federal de Lavras.
- ANDERSON, D. R.; SWEENEY, D. J.; WILLIAMS, T. A. **Estatística aplicada à Administração e Economia**. Trad. SANTOS, J. C. B. 2 ed. São Paulo: Editora CENGAGE Learning. 597 p. 2008.
- ANTON, H. **Cálculo**. 8 ed. São Paulo: Bookman companhia editora. Volume 1, p. 650. 2007.
- APOSTOL, T. M. **Calculus: Cálculo con funciones de una variable, con una introducción al álgebra lineal**. 2 ed. Barcelona: Editorial Reverté, S.A. Volume 1, p. 813. 1977.
- ARCE, J. E. Modelagem da estrutura de florestas clonais de *Populus deltoides* Marsh. Através de distribuições diamétricas probabilísticas. **Ciência Florestal**, Santa Maria, v. 14, n. 1, p. 149-164, 2004.
- BARTOSZECK, A. C. P. S.; MACHADO, S. A.; FIGUEIREDO FILHO, A.; OLIVEIRA, E. B. Distribuição diamétrica para bracatingais em diferentes idades, sítios e densidades na região metropolitana de Curitiba. **Floresta**, Curitiba-PR, V. 34, N. 3, p. 305-323, set/Dez 2004.
- BINOTI, D. H. B.; LEITE, H. G.; NOGUEIRA, G. S.; SILVA, M. L. M.; GARCIA, S. L. R.; CRUZ, J. P. Uso da função Weibull de três parâmetros em um modelo de distribuição diamétrica para plantios de eucalipto submetidos a desbaste. **Árvore**, Viçosa-MG, v.34, n.1, p.147-156, 2010.
- BOYER, C. B. **História da matemática**. Ed. 2. Edgard Blucher. 1996.
- CARVALHO, P. E. R. **Espécies florestais brasileiras: recomendações silviculturais, potencialidades e uso da madeira**. Empresa Brasileira de Pesquisa Agropecuária; Centro Nacional de Pesquisa de Florestas – Colombo: EMBRAPA – CNPF; Brasília: EMBRAPA – SPI, 1994. 640p.
- CATALUNHAA, M. J.; SEDIYAMA, G. C.; LEAL, B. G.; SOARES, C. P. B.; RIBEIRO, A.. Avaliação de cinco funções densidade de probabilidade a séries de precipitação pluvial no Estado de Minas Gerais. **Revista Brasileira de Agrometeorologia**, v.10, n.1, p.153-162, 2002.
- COURANT, R. **Cálculo diferencial e integral**. 1 ed. Quarta impressão. Rio de Janeiro: Editora Globo. Volume 1, p. 616. 1965.
- CUNHA NETO, F. R. S.; SCOLFORO, J. R. S.; CALEGARIO, N.; OLIVEIRA, A. D.; KANEGAE JUNIOR, H. Modelo de predição da produção por classe de diâmetro para *Eucalyptus grandis*. **Cerne**, V. 1, N. 1, p. 108-122, 1994.
- DE HOOG, R.J. **Site-Nutrition-Growth relationships of *Araucaria angustifolia* (Bert.) O. Kuntze in southern Brazil**. Tese (Doutorado em Ciência Florestal) – Albert-Ludwigs-Universität. 1981.

- DUCKE, A.; G. A.; BLACK. G. A. Notas sobre a fitogeografia da Amazônia brasileira. **Boletim Técnico do Instituto Agrônomo do Norte** **29**, 1954. p. 1-48.
- EISFELD, R. L.; SANQUETTA, C. R.; ARCE, J. E.; MAESTRI, R.; WEBER, K. S. Modelagem do crescimento e da produção de *Pinus taeda* L. por meio de função probabilística. **Floresta**, Curitiba, PR, v. 35, n. 2, mai./ago., p. 317 – 328, 2005.
- EMBRAPA - CNPS. **Sistema Brasileiro de Classificação de Solos**. Brasília: Embrapa-SPI; Rio de Janeiro: Embrapa-Solos, 2006. 306 p.
- EUES, H. **Introdução à história da matemática**. UNICAMP. 1995.
- FAO. **Food and Fruit-bearing forest species**. 3: examples from Latin America. Roma. 1986.
- FLEMMING, D. M.; GONÇALVES, M. B. **Cálculo A**: funções, limite, derivação, integração. 5ª ed, São Paulo, SP: Makron. P. 617, 1992.
- FREUND, R. J.; WILSON, W. J. **Statistical methods**. Academic Press. Revised edition. P. 684. 1996.
- FURLAN, D. N. **Variabilidade temporal e espacial das chuvas e do balanço hídrico no estado de Rondônia**: caracterização e análise de tendência. 129 f. Dissertação (Mestrado em Agronomia). Esalq, Piracicaba. 2009.
- GOMIDE, G. L. A. **Estrutura e dinâmica de crescimento de florestas tropicais primária e secundária no Estado do Paraná**. Curitiba: Dissertação (Mestrado em Engenharia Florestal) - UFPR, 1997.
- GRANVILLE, W. A.; LONGLEY, P. F. S. **elementos de cálculo diferencial e integral**. 5 ed. Rio de Janeiro: Editora Científica. P. 703. 1966.
- GUERRA, M. P.; SILVEIRA, V.; REIS, M. S.; SCHNEIDER, L. Exploração, manejo e conservação da araucária (*Araucaria angustifolia*). In: SIMÕES, L. L.; LINO, C. F., (orgs) **Sustentável Mata Atlântica**: a exploração de seus recursos florestais. Senac, São Paulo, p.85-102. 2002.
- GUIDORIZZI, H. L. **Um curso de cálculo**. 2.ed. São Paulo: LTC, p. 579, 1994.
- HAMELIN, P, O fracasso anunciado. **Museu Paraense Emilio Goeldi**: Colepio Eduardo Galvão, p. 161-176. 1991.
- HERTEL, R. J. G. **Interpretação morfológica da *Araucaria angustifolia***. Curitiba. 1980. 143 p. Tese (Professor titular). Universidade Federal do Paraná. 1980.
- HIGUCHI, F. G. **A influência do tamanho da parcela na precisão da função de distribuição de diâmetro de Weibull da floresta primária da Amazônia central**. 59 f. Dissertação (Mestrado em Ciências Florestais). UFPR, Curitiba, 2007.
- HUECK. K. **As florestas da América do Sul**. Brasília; Editora da Universidade de Brasília. 1972. 466 p.

- IBAMA – Instituto Brasileiro do Meio Ambiente e dos Recursos Naturais Renováveis. Disponível em <www.ibama.gov.br> Acesso em 10/05/2010.
- INOUE, M. T.; GALVAO, F.; TORRES, D. V. Estudo ecofisiológico sobre *Araucaria angustifolia* (Bert.) O. Ktze.: fotossíntese em dependência a luz no estágio juvenil. **Floresta**, Curitiba, v. 10, n. 1, p. 5-9, 1979.
- INSTITUTO NACIONAL DE COLONIZAÇÃO E REFORMA AGRÁRIA (Brasília, DF). **Inventário florestal do projeto integrado de colonização de Altamira-PA**. Curitiba: INCRA / Universidade Federal do Paraná - Centro de Pesquisas Florestais, 1976. 129p.
- JOHNSON, N. L.; LEONE, F. C. **Statistics and experimental design in the engineering and the physical sciences**. V. 1, p. 523, 1964.
- JORGE, L. A. B.; VEIGA, R. A. A. A função Weibull no estudo de distribuições diamétricas em povoamento de *Pinus elliottii* na estação experimental de Itapeva. **IPEF**, n.43/44, p.54-60, jan./dez.1990.
- KLEIN, R.M. O aspecto dinâmico do pinheiro brasileiro. **Sellowia**. 12. p. 17-44. 1960.
- LARSON, R.; FARBER, B. **Estatística aplicada**. 4 ed. São Paulo: Pearson Prentice Hall. 637 p. 2010.
- LEVINE, D. M.; STEPHAN, D.; KREHBIEL, T. C.; BERENSON, M. L. **Estatística: teoria e aplicações**. Trad. SOUZA, T. C. P. 5 ed. Rio de Janeiro: LTC. 752 p.. 2008.
- LIMA, A. R.; CAPOBIANCO, J. P. R. (coords.). **Mata Atlântica: avanços legais e institucionais para sua conservação**. Documentos do ISA n. 004. Instituto Sócio Ambiental, Brasília. 1997.
- LIMA, J. A. S.; MENEGUELLI, N. A.; GAZEL FILHO, A. B.; PÉREZ, D. V. Agrupamento de espécies arbóreas de uma floresta tropical por características de solo. **Pesquisa Agropecuária Brasileira**, Brasília, v. 38, n. 1, p. 109-116, jan. 2003
- LONGHI, S. J. **A estrutura de uma floresta natural de *Araucaria angustifolia* (Bert.) O. Ktze, no Sul do Brasil**. Curitiba: UFPR, 1980, 198 p. Dissertação (Mestrado em Ciência Florestal) – Universidade Federal do Paraná, 1980.
- LORENZI, H. **Arvores Brasileiras: manual de identificação e cultivo de arb[oreas nativas do Brasil**. Nova Odessa, SP: Editora Plantarum, 1992. 368p.
- MACHADO, S. A.; AUGUSTYNCZIK, A. L. D.; NASCIMENTO, R. G. M.; TÊO, S. J.; MIGUEL, E. P.; FIGURA, M. A.; SILVA, L. C. R. Funções de distribuição diamétrica em um fragmento de Floresta Ombrófila Mista. **Ciência Rural**, v. 39, n. 8, p. 2428-2434, nov, 2009.
- MACHADO, S. A.; MELLO, J. M.; BARROS, D. A. Comparação entre métodos para avaliação de volume total de madeira por unidade de área, para o pinheiro do Paraná, na região Sul do Brasil. **Cerne**, Lavras, v. 6, n. 2, p. 55-66, 2000.

- MACHADO, S. A.; SIQUEIRA, J. D. P. Distribuição natural da *Araucaria angustifolia* (Bert.) O. Ktze. In: Encontro da IUFRO. Problemas florestais do gênero *Araucaria*. **Anais...** Curitiba, FUPEF, 1980. p. 4-9.
- MAESTRI R. **Estimativa de produção presente e futura de volume de madeira e peso de casca para povoamentos de Acácia - Negra (*Acacia mearnsii* de Wild)**. Curitiba, 1992. 102f. Dissertação (Mestrado em Ciências Florestais) Setor de Ciências Agrárias, Universidade Federal do Paraná.
- MARQUADT, D. W. An algorithm for least squares estimation of nonlinear parameters, **SIAM Journal**, V. 11, p. 431- 441, 1963.
- MEIRA-NETO, J. A. A.; MARTINS, F. R.; SOUZA, A. L. Influência da cobertura e do solo na composição florística do sub-bosque em uma floresta estacional semidecidual em Viçosa, MG, Brasil. **Acta Botânica Brasílica**. V.19, N. 3, p. 473-486, 2005.
- MESQUITA, M. G. B. F.; MORAES, S. O.; PERUCHI, F.; TEREZA, M. C. Alternativa para caracterização da condutividade hidráulica saturada do solo utilizando probabilidade de ocorrência. **Ciência agrotécnica**, Lavras, v. 31, n. 6, p. 1605-1609, nov./dez., 2007.
- MEYER, P. L. **Probabilidade: aplicações à estatística**. Rio de Janeiro, p. 391, 1974.
- MILONE, G. **Estatística geral e aplicada**. São Paulo: Editora CENGAGE Learning. 479 p. 2009.
- MOISE, E. E. **Cálculo: um curso universitário**. São Paulo, Edgard Blucher. Volume 1, p. 487. 1972.
- MONTGOMERY, D. C.; RUNGER, G. C. **Estatística aplicada e probabilidade para engenheiros**. 2 ed. Editora LTC. P. 463 2003.
- MOOD, A. M.; GRAYBILL, F. A. **Introducción a la teoría de la estadística**. Madrid, p. 536, 1978.
- MOURÃO JÚNIOR, M.; BARBOSA, R. I. Bioecologia do caimbé [*Curatella americana* L. (Dilleniaceae)] (i): Distribuições de probabilidade de parâmetros dendrométricos. **Embrapa**. Comunicado Técnico 18. Dezembro de 2005.
- MULLER, C. H.; RODRIGUES, L. A.; MULLER, A. A.; MULLER; N. R. M. Castanha-do-Brasil - Resultados de Pesquisas. **Miscelânea**. V.2, p. 1-25. 1980.
- NOGUEIRA. A. C. **Reação do crescimento radial da *Araucaria angustifolia* (Bert.) O. Kuntze em florestas naturais que sofreram corte seletivo**. Curitiba. PR. 1989. 107 f. Tese (Doutorado). Universidade Federal do Paraná. 1989.
- NUTTO, L. Manejo do crescimento diamétrico de *Araucaria angustifolia* (Bert.) O. Ktze. baseado na árvore individual. **Ciência Florestal**, Santa Maria, v. 11, n. 2, p. 9-25, 2001.
- PAGANO, M.; GAUVREAU, K. **Princípios de Bioestatística**. Trad. PAIVA, L. S. C. 2 ed. São Paulo: Editora CENGAGE Learning. 506 p. 2003.
- PÉLLICO NETTO, S. Estimativas volumétricas de árvores individuais. Síntese teórica. **Floresta** (UFPR), Curitiba - PR, v. 9, n. 2, p. 63-73, 1981.

- PÉLLICO NETTO, S.; BRENA, D. A. **Inventário Florestal**. V.1, Curitiba, 1997. 316 p.
- PENG, C. Growth and yield models for uneven-aged stands: past, present and future. **For. Ecol. Manage.**, v. 132, p. 259-279, 2000.
- PERES, C. A.; BAIDER, C Seed dispersal, spatial distribution and population structure of Brazilnut trees *Bertholletia excelsa* in southeastern Amazonia. **Journal of Tropical Ecology**, 13, pp 595-616. 1997.
- PISKOUNOV, N. **Cálculo diferencial e integral**. 16^a ed. Porto, Portugal: Lopes da Silva, V1, 1993.
- PIZATTO, W. **Avaliação biométrica da estrutura e da dinâmica de uma Floresta Ombrófila Mista em São João do Triunfo – PR: 1995 a 1998**. 172 f. Dissertação (Mestrado em Ciências Florestais), UFPR, 1999.
- RODE, R. **Avaliação florística e estrutural de uma floresta ombrófila mista e de uma vegetação arbórea estabelecida sob um povoamento de *Araucaria angustifolia* de 60 anos**. 158f. Dissertação (Mestrado em Ciências Florestais), UFPR, 2008.
- ROSS, S. M. **Introduction to probability models**. 7 ed. Academic Press. P. 693. 2000.
- ROSSATO, R. A. R. **Limites de precaução e controle em análises de tecidos vegetais de espécies florestais**. 54 f. Dissertação (mestrado em Engenharia Florestal). UFSM, Santa Maria, 2006.
- ROUSSAS, G. G. **A course in mathematical statistics**. 2^a ed. Londres: Adisson Wesley Publishing Company, p. 536, 1997.
- RUGGIERO, M. A. G.; LOPES, V. L. R. **Cálculo Numérico: aspectos teóricos e computacionais**, segunda edição, Makron Books, 1998.
- SALOMÃO, R. P., Estrutura e densidade de *Bertholletia excelsa* H. & B. (“castanheira”) nas regiões de Carajás e Marabá, Estado do Pará. **Boletim do Museu Paraense Emílio Goeldi, série Botânica**, V. 7, N. 1, p. 47-68, 1991.
- SALOMÃO, R. P.; MATOS, A. H.; CORRÊA, M. S.; ROSÁRIO, C. S.; SILVA, C. A. S.; SANTOS, M. R.; NEGREIROS, C. O.; ROSÁRIO, D. O.; PINHO, J. A. G. B.; SILVA, D.F.; PANTOJA, P.E. Inventário florestal em 772,25 ha de castanheiras (*Bertholletia excelsa* H. & B. - Lecythidaceae) no Platô Almeidas, Floresta Nacional Sarcá-Taquera/Ibama, Porto Trombetas, município de Oriximiná, Pará. **Relatório Técnico**. Mineração Rio do Norte: Cooperativa de Tecnologia Organizacional. Porto Trombetas - Pará. 100 p. 2002.
- SANTANA, F. B.; ARRUDA, P. H. Z.; FARIA, J. L. B.; NOGUEIRA, J. S. Estudo preliminar da velocidade e direção dos ventos, em Cuiabá, MT, Brasil. **Revista Brasileira de Agrometeorologia**, v.16, n. 2, p.175-180, agosto/2008.
- SANTAROSA, E.; OLIVEIRA, J. M.; ROIG, F. A.; DEPATTA PILLAR, V. Crescimento Sazonal em *Araucaria angustifolia*: Evidências Anatômicas. **Revista Brasileira de Biociências**, Porto Alegre, v. 5, supl 1, p. 618-620, jul. 2007.

- SANTOS, A. S. A. **Modelos simétricos transformados não lineares com aplicação na estimativa volumétrica em híbrido de *Eucalyptus tereticornis* no pólo gesseiro de Araripe – PE.** 94 f. Dissertação (mestrado em Biometria e Estatística). UFRPE, Recife, 2010.
- SILVA, E. Q. **Nova função densidade de probabilidade aplicável à Ciência Florestal.** 2003. 98f. Tese (Doutorado em Engenharia Florestal) – Universidade Federal do Paraná, UFPR, Curitiba, PR, 2003.
- SILVA, E. Q.; PÉLLICO NETTO, S.; MACHADO, S. A.; SANQUETTA, C. R. Função densidade de probabilidade aplicável à Ciência Florestal. **Floresta**, Curitiba-PR, V. 33, N. 3, 2003.
- SILVA, J. C.; HELDWEIN, A. B.; MARTINS, F. B.; GRIMM, E. L. Funções de distribuição de probabilidade decendial e mensal para a deficiência hídrica no solo. **Ciência Rural**, v. 38, n. 7, out, 2008.
- SILVINO, A. N. O.; SILVEIRA, A.; MUSIS, C. R.; WYREPKOWSKI, C. C.; CONCEIÇÃO, F. T. Determinação de vazões extremas para diversos períodos de retorno para o rio Paraguai utilizando métodos estatísticos. **Geociências**. São Paulo, V. 26, N. 4, p. 369-378, 2007.
- SOONG, T.T. **Modelos Probabilísticos em Engenharia e Ciências.** Universidade Estadual de New York, em Buffalo, New York. ICEX-UFMG, 1986.
- SPIEGEL, M. R.; STEPHENS, L. J. **Estatística.** Trad. NASCIMENTO, J. L. 4 ed. Porto Alegre: Bookman. 597 p. 2009.
- SPVS - Sociedade de Pesquisa em Vida Selvagem e Educação Ambiental. **Manual para recuperação da reserva florestal legal.** Curitiba: FNMA, 84p., 1996.
- STEVENSON, W. J. **Estatística aplicada à administração.** P. 495, 1981.
- STEWART, J. **Cálculo.** 6. ed. – São Paulo: Pioneria Thomson Learning. Volume 1, p. 535. 2011.
- SWOKOWSKI, E. W. **Cálculo com geometria analítica.** 2. ed. São Paulo: Makron Books. Volume 1, p. 744. 1994.
- TRIOLA M. F. **Introdução à estatística.** Trad. FARIAS E FLORES, V. R. L. 10 ed. Rio de Janeiro: LTC. 696 p. 2008.
- VALEJOS, O. S. B. **Sistema de simulação de crescimento e produção de *Populus spp.*, no Chile (Sistema SALICA).** Curitiba, 2003. 110f. Tese (Doutorado em Ciências Florestais) Setor de Ciências Agrárias, Universidade Federal do Paraná.
- VANCLAY, J. K.; GILLICON, A. N.; KEENAN, R. J. Using functional attributes to quantify site productivity and growth patterns in mixed forests. **Forest Management**, v. 94, p. 149-163,1997.
- VELOSO. H. P.; RANGEL-FILHO. A .L.; LIMA. J. C. A. **Classificação da vegetação brasileira adaptada a um sistema universal.** IBGE. Rio de Janeiro. 1991. 123 p.

- WEBER, S. H. **Desenvolvimento de nova função densidade de probabilidade para avaliação de regeneração natural**. 74 f. Dissertação (Mestrado em Ciências Florestais). UFPR, Curitiba-PR. 2006.
- WEBER, S. H.; ARCE, J. E.; PELLICO NETTO, S. Aplicação da Distribuição de Weber ao peso de pinhas verdes. *Árvore*, Viçosa, v. 33, n. 5, Oct. 2009.
- WOOD, C. H.; SCHMINK, M. Blaming the victim: small Farmer production in an Amazon colonization project. *Studies in Third World Societies*, v. 7, p. 77-93. 1979.