

EDUARDO QUADROS DA SILVA

**NOVA FUNÇÃO DENSIDADE DE PROBABILIDADE
APLICÁVEL À CIÊNCIA FLORESTAL**

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

Orientador: Prof. Dr. Sylvio Péllico Netto

CURITIBA

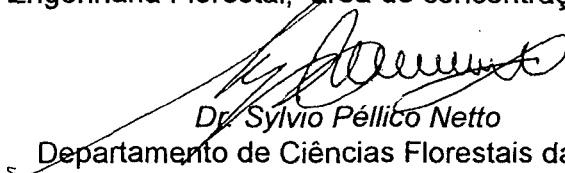
2003




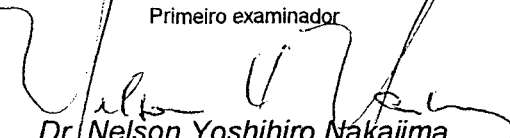
Universidade Federal do Paraná
Setor de Ciências Agrárias – Centro de Ciências Florestais e da Madeira
Programa de Pós-Graduação em Engenharia Florestal
Av. Lothário Meissner, 3400 – Jardim Botânico – CAMPUS III
80210-170 - CURITIBA - Paraná
Tel. (41) 360.4212 - Fax. (41) 360.4211 – <http://www.floresta.ufpr.br/pos-graduacao>
e-mail: pinheiro@floresta.ufpr.br

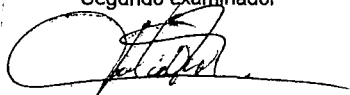
PARECER
Defesa nº 506

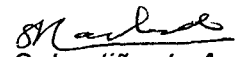
A banca examinadora, instituída pelo colegiado do Curso de Pós-Graduação em Engenharia Florestal, do Setor de Ciências Agrárias, da Universidade Federal do Paraná, após argüir o doutorando *EDUARDO QUADROS DA SILVA* em relação ao seu trabalho de tese intitulado "NOVA FUNÇÃO DENSIDADE DE PROBABILIDADE APLICÁVEL À CIÊNCIA FLORESTAL", é de parecer favorável à **APROVAÇÃO** do acadêmico, habilitando-o ao título de *Doutor* no Curso de Pós-Graduação em Engenharia Florestal, área de concentração em *Manejo Florestal*.


Dr. Sylvio Péllico Netto
Departamento de Ciências Florestais da UFPR
Orientador e presidente da banca examinadora

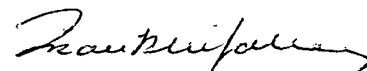

Dr. Marco André Mazzarotto
Pontifícia Universidade Católica do Paraná - PUC
Primeiro examinador


Dr. Nelson Yoshihiro Nakajima
Fundação Universidade Regional de Blumenau - FURB
Segundo examinador


Dr. Julio Eduardo Arce
Departamento de Ciências Florestais da UFPR
Terceiro examinador


Dr. Sebastião do Amaral Machado
Departamento de Ciências Florestais da UFPR
Quarto examinador

Curitiba, 13 de março de 2003.


Nivaldo Eduardo Rizzi
Coordenador do Curso de Pós-Graduação em Engenharia Florestal
Franklin Galvão
Vice-coordenador

*A Deus,
luz que ilumina o meu caminho.
A meus pais Manoel (in memoriam) e Romilda,
pela maneira com que me prepararam para a vida.
À Terezinha, minha esposa,
pelo apoio nos momentos em que tive dificuldades.
A meu filho Guilherme Eduardo,
pelos momentos felizes que tem me proporcionado,
A meu irmão Fernando,
companheiro na minha caminhada.*

DEDICO.

AGRADECIMENTOS

À Universidade Federal do Paraná e ao Centro de Ciências Florestais e da Madeira pela oportunidade de cursar o doutorado.

À Pontifícia Universidade Católica do Paraná pela concessão de bolsa-auxílio.

Aos professores do comitê de orientação Carlos Roberto Sanquetta e Sebastião do Amaral Machado pelas aulas, conselhos e informações importantes para a realização do meu trabalho.

Ao professor orientador Sylvio Péllico Netto, por ter conduzido com sabedoria os meus estudos e por estar sempre presente com suas valiosas contribuições, a minha incomensurável gratidão pelo apoio incessante.

Aos professores responsáveis pelas disciplinas que cursei por terem transmitido não apenas conhecimentos mas também exemplos e lições de vida.

Ao estatístico e professor Sérgio Aparecido Ignácio pelas contribuições que fez para o meu trabalho e pelo apoio nos momentos em que tive dificuldade.

Ao secretário Reinaldo pela competência com que desempenha o seu trabalho.

Ao amigos Luciano Farinha e Fabio Minoru Yamaji, sempre dispostos a tirar minhas dúvidas a respeito da Área de Manejo Florestal.

Aos professores Júlio Eduardo Arce, Nelson Yoshihiro Nakajima e Marco André Mazarotto, integrantes da banca examinadora, pelas contribuições para o aperfeiçoamento do meu trabalho.

Aos amigos Engenheiros Florestais que cursaram disciplinas comigo pela amizade e pelos conhecimentos da área florestal que dividiram comigo.

A todos que contribuíram para que este trabalho pudesse ser realizado.

BIOGRAFIA

Eduardo Quadros da Silva, filho de Manoel Quadros da Silva e Romilda Quadros da Silva, nasceu em 13 de outubro de 1967, na cidade de União da Vitória, PR.

Concluiu o 1.º grau no Colégio Estadual Padre Arnaldo Jansen e o 2.º grau no Colégio Ideal, ambos em São José dos Pinhais, PR.

Em 1985, ingressou no curso de Licenciatura em Ciências-Habilitação Matemática e Bacharelado em Matemática da Universidade Católica do Paraná, concluindo-o em dezembro de 1988.

Em 1986 e 1987 foi professor de Matemática nos colégios estaduais Padre Arnaldo Jansen e Afonso Pena em São José dos Pinhais e em 1988 no Colégio Positivo Júnior em Curitiba.

De 1989 a 1991 trabalhou como professor de Matemática, Física e Desenho Geométrico no primeiro e no segundo graus do Colégio Positivo de Jaguariaíva, escola conveniada à fábrica de papel PISA situada em Jaguariaíva, PR.

Em 1992 trabalhou como professor de Matemática no Colégio Positivo 2.º Grau em Curitiba e iniciou suas atividades como autor de material didático de Matemática.

Em 1993 trabalhou como Coordenador dos Autores de Material Didático das Áreas de Matemática, Física e Química para 1.º e 2.º graus do Sistema Educacional Positivo.

Também em 1993, cursou especialização em Didática do Ensino Superior na Pontifícia Universidade Católica do Paraná.

De 1994 a 1995 foi Coordenador de Matemática do 2.º Grau do Sistema Educacional Positivo.

De 1994 até 1997 cursou mestrado em Educação com dissertação referente ao ensino de logaritmos inserido no Cálculo Diferencial e Integral. Obteve o grau de Mestre em 8 de Julho de 1997.

Em 1994 ingressou na Pontifícia Universidade Católica do Paraná, como professor de Cálculo Diferencial e Integral nos cursos de Matemática, Arquitetura, Química Industrial e Engenharias.

Em 1995 iniciou suas atividades como professor de Matemática para os cursos de Agronomia e Zootecnia do Centro de Ciências Agrárias e Ambientais da PUCPR onde permanece até hoje.

De 1997 a 1998 foi Coordenador da Área de Exatas do Colégio Novo Ateneu em Curitiba.

Também em 1998 prestou serviço como autor de material didático para o Colégio Terceiro Milênio.

Em abril de 1999 iniciou o Doutorado em Engenharia Florestal, com concentração em Manejo Florestal.

Em 2001 iniciou suas atividades como professor de Cálculo Diferencial, Geometria Analítica e Álgebra Linear no curso de Engenharia Ambiental da PUCPR onde permanece até hoje.

Em 6 de Janeiro de 2003 assumiu o cargo de Diretor do Curso de Matemática da PUCPR.

SUMÁRIO

LISTA DE FIGURAS	viii
LISTA DE FÓRMULAS E DEFINIÇÕES	ix
LISTA DE TABELAS	xi
RESUMO	xii
ABSTRACT	xiii
1 INTRODUÇÃO	1
1.1 JUSTIFICATIVA	2
1.2 OBJETIVOS	5
1.2.1 Objetivo Geral	5
1.2.2 Objetivos Específicos	5
2 REVISÃO DE LITERATURA	6
2.1 TÓPICOS DE MATEMÁTICA	6
2.1.1 Classificação das Funções	6
2.1.2 Funções Crescentes e Decrescentes	7
2.1.3 Definição de Limite	8
2.1.4 Continuidade	8
2.1.5 Derivabilidade e Continuidade	9
2.1.6 Teorema Fundamental do Cálculo	11
2.1.7 Integrais Impróprias	11
2.1.8 Função Gama	13
2.1.9 Ajuste de Curvas por Quadrados Mínimos Não Lineares no Matlab	13
2.2 TÓPICOS DE ESTATÍSTICA	15
2.2.1 Variáveis Aleatórias Discretas e Contínuas	15
2.2.2 Função Densidade de Probabilidade	16
2.2.3 Função de Distribuição Acumulada	17
2.2.4 Média e Variância de uma Função Densidade de Probabilidade	18
2.3 PRINCIPAIS DISTRIBUIÇÕES CONTÍNUAS	18
2.3.1 Distribuição Retangular Uniforme	19
2.3.2 Distribuição Exponencial	20
2.3.3 Distribuição Gama	21
2.3.4 Distribuição Beta	23
2.3.5 Distribuição Weibull	25
2.3.6 Distribuição χ^2	27

2.3.7	Distribuição Normal ou de Gauss	28
3	DESENVOLVIMENTO DA FUNÇÃO DENSIDADE DE PROBABILIDADE	30
3.1	DIAGNÓSTICO DE AJUSTES FEITOS COM DISTRIBUIÇÕES EXISTENTES.....	30
3.2	ESTUDO DE MODELO PARA DADOS DO JEQUITIBÁ-ROSA.....	33
3.3	ESTUDO DE FUNÇÕES COM CURVAS ASSIMÉTRICAS	37
3.4	ELABORAÇÃO DA DISTRIBUIÇÃO POLINOMIAL	41
3.4.1	Caracterização e Processo de Elaboração	41
3.4.2	Aplicação para o Caso do Jequitibá-Rosa	43
4	RESULTADOS.....	54
4.1	CÁLCULOS COM A FDP ELABORADA PARA O JEQUITIBÁ-ROSA.....	54
4.1.2	Cálculo da Média pela Distribuição Polinomial.....	55
4.1.3	Cálculo da Variância pela Distribuição Polinomial	55
4.2	MODELO ELABORADO PARA O JEQUITIBÁ-BRANCO	56
4.2.1	Cálculo da Média para o Jequitibá-Branco.....	61
4.2.2	Cálculo da Variância para o Jequitibá-Branco	61
4.3	TESTES COMPARATIVOS DO JEQUITIBÁ-BRANCO E DO JEQUITIBÁ-ROSA..	62
4.3.1	Teste dos Dados do Jequitibá-Rosa com a Distribuição Polinomial.....	64
4.3.2	Teste dos Dados do Jequitibá-Rosa com a Distribuição Gama	65
4.3.3	Teste dos Dados do Jequitibá-Rosa com a Distribuição Beta	66
4.3.4	Teste dos Dados do Jequitibá-Rosa com a Distribuição Weibull Dois Parâmetros	67
4.3.5	Teste dos Dados do Jequitibá-Branco com a Distribuição Polinomial.....	68
4.3.6	Teste dos Dados do Jequitibá-Branco com a Distribuição Gama	70
4.3.7	Teste dos Dados do Jequitibá-Branco com a Distribuição Beta.....	70
4.3.8	Teste dos Dados do Jequitibá-Branco com a Distribuição Weibull	71
4.3.9	Análise Exploratória dos Dados do Jequitibá-Rosa	72
4.3.10	Análise Exploratória dos Dados do Jequitibá-Branco	72
4.4	AJUSTE DO MODELO POLINOMIAL A DADOS DE CIRCUNFERÊNCIA A 60 cm DE ALTURA DE PLANTAS DE PALMITO (<i>Euterpe edulis</i>).....	73
4.4.1	Caracterização do Conjunto de Dados.....	73
4.4.2	Elaboração do Modelo Polinomial para os Dados de Circunferências a 60 cm de Altura de Plantas de Palmito.....	75
4.4.3	Cálculo da Média e da Variância pelo Modelo Polinomial	81
4.4.4	Testes de Aderência do Modelo Polinomial aos Dados de Circunferência a 60cm de Altura de Plantas de Palmito.....	82
4.4.5	Ajuste de Outras Distribuições Contínuas aos Dados de Circunferência a 60 cm de Altura de Plantas de Palmito.....	83

4.5 ANÁLISE DE RESULTADOS.....	84
4.5.1 Análise dos Resultados da Variável Altura Obtidos para o Jequitibá-Rosa	84
4.5.2 Análise dos Resultados Obtidos para o Jequitibá-Branco	86
4.5.3 Análise dos Resultados dos Testes de Aderência do Modelo Polinomial e de outros Modelos aos Dados de Circunferência à 60 cm de Altura de Plantas de Palmito	86
CONCLUSÕES.....	88
REFERÊNCIAS	91
GLOSSÁRIO	93
ANEXO 1 - AJUSTE DOS DADOS DE ALTURAS DE <i>Eucalyptus saligna</i> PELA DISTRIBUIÇÃO WEIBULL COM DOIS PARÂMETROS.....	94
ANEXO 2 - AJUSTE DOS DADOS DE DAP COM CASCA DE <i>pinus sp</i> PELA DISTRIBUIÇÃO BETA.....	95
ANEXO 3 - AJUSTE DOS DADOS DE DAP COM CASCA DE <i>pinus sp</i> PELA DISTRIBUIÇÃO WEIBULL COM DOIS PARÂMETROS	96
ANEXO 4 - GRÁFICO DE ALTURA TOTAL DO JEQUITIBÁ BRANCO JUNTAMENTE COM PONTOS OBSERVADOS E AJUSTES FEITOS PELAS DISTRIBUIÇÕES BETA, WEIBULL DOIS PARÂMETROS, GAMA E POLINOMIAL	97
ANEXO 5 - GRÁFICO DE ALTURA TOTAL DO JEQUITIBÁ ROSA JUNTAMENTE COM PONTOS OBSERVADOS E AJUSTES FEITOS PELAS DISTRIBUIÇÕES POLINOMIAL, GAMA, BETA E WEIBULL DE DOIS PARÂMETROS	98

LISTA DE FIGURAS

1	DERIVABILIDADE E CONTINUIDADE	10
2	INTEGRAL IMPRÓPRIA $[a, \infty)$	12
3	INTEGRAL IMPRÓPRIA $(-\infty, a]$	12
4	DISTRIBUIÇÃO RETANGULAR UNIFORME.....	19
5	DISTRIBUIÇÃO EXPONENCIAL.....	20
6	DISTRIBUIÇÃO GAMA.....	22
7	DISTRIBUIÇÃO BETA PARA ALGUNS VALORES DOS PARÂMETROS.....	25
8	DISTRIBUIÇÃO WEIBULL PARA ALGUNS VALORES DOS PARÂMETROS.....	26
9	DISTRIBUIÇÃO NORMAL.....	28
10	FREQÜÊNCIAS DO JEQUITIBÁ-ROSA.....	36
11	GRÁFICO DA FUNÇÃO $f(x) = 1 / (x^2 - 2x + 2)$	37
12	GRÁFICO DE $f(x) = 1 / (x^2 + 6x + 12)$	38
13	GRÁFICO DE $f(x) = 1 / (x^2 - 6x + 12)$	39
14	GRÁFICO DA FUNÇÃO $f(x) = -1 / (x^2 - 2x + 2)$	39
15	POLINÔMIO AJUSTADO PARA OS DADOS DO JEQUITIBÁ-ROSA JUNTAMENTE COM VALORES OBSERVADOS.....	44
16	GRÁFICO DO POLINÔMIO AJUSTADO MOSTRANDO AS RAÍZES E O INTERVALO ONDE É NEGATIVO.....	45
17	GRÁFICO DE $g(x)$, FUNÇÃO DEFINIDA POR TRÊS SENTENÇAS.....	49
18	GRÁFICO DE $g(x)$ COM OS VALORES OBSERVADOS DO JEQUITIBÁ-ROSA.....	50
19	GRÁFICO DO POLINÔMIO GERADO PARA O JEQUITIBÁ-BRANCO JUNTAMENTE COM OS VALORES OBSERVADOS	57
20	GRÁFICO DO POLINÔMIO GERADO PARA O JEQUITIBÁ-BRANCO COM OS INTERVALOS ONDE DEVERÃO SER INSERIDAS NOVAS FUNÇÕES.....	58
21	GRÁFICO DA FUNÇÃO $g(x)$ DEFINIDA PARA O JEQUITIBÁ-BRANCO	60
22	FREQÜÊNCIA DOS DADOS DE CIRCUNFERÊNCIA À 60 cm DE ALTURA DE PLANTAS DE PALMITO.....	74
23	GRÁFICO DO POLINÔMIO AJUSTADO PARA OS DADOS DE CIRCUNFERÊNCIAS DE PLANTAS DE PALMITO COM DE 60cm DE ALTURA	76
24	GRÁFICO DA FUNÇÃO $g(x)$ ELABORADA PARA OS DADOS DE CIRCUNFERÊNCIA DE PLANTAS DE PALMITO COM 60 cm DE ALTURA.....	79

LISTA DE FÓRMULAS E DEFINIÇÕES

1	FUNÇÃO GAMA	13
2	FUNÇÃO DE DISTRIBUIÇÃO ACUMULADA	17
3	MÉDIA DE UMA FUNÇÃO DENSIDADE DE PROBABILIDADE PELA EXPECTATIVA MATEMÁTICA.....	18
4	VARIÂNCIA DE UMA FUNÇÃO DENSIDADE DE PROBABILIDADE PELA EXPECTATIVA MATEMÁTICA.....	18
5	FUNÇÃO DENSIDADE DE PROBABILIDADE RETANGULAR UNIFORME - MODELO CONTÍNUO $f(x)$	19
6	FUNÇÃO DE DISTRIBUIÇÃO PARA O MODELO RETANGULAR UNIFORME $F(x)$	20
7	MÉDIA DA DISTRIBUIÇÃO RETANGULAR UNIFORME.....	20
8	VARIÂNCIA DA DISTRIBUIÇÃO RETANGULAR UNIFORME.....	20
9	FUNÇÃO DENSIDADE DE PROBABILIDADE EXPONENCIAL $f(x)$	20
10	FUNÇÃO DE DISTRIBUIÇÃO PARA O MODELO EXPONENCIAL $F(x)$	21
11	MÉDIA DA FDP EXPONENCIAL	21
12	VARIÂNCIA DA FDP EXPONENCIAL	21
13	FUNÇÃO DENSIDADE DE PROBABILIDADE GAMA	21
14	FUNÇÃO DE DISTRIBUIÇÃO PARA O MODELO GAMA	23
15	MÉDIA DA FDP GAMA.....	23
16	VARIÂNCIA DA FDP GAMA.....	23
17	FUNÇÃO DENSIDADE DE PROBABILIDADE BETA - MODELO GENERALIZADO.....	23
18	FUNÇÃO DENSIDADE DE PROBABILIDADE BETA - CASO PARTICULAR.....	23
19	FUNÇÃO DE DISTRIBUIÇÃO PARA O MODELO BETA	24
20	MÉDIA DA FDP BETA.....	24
21	VARIÂNCIA DA FDP BETA.....	24
22	FUNÇÃO DENSIDADE DE PROBABILIDADE WEIBULL.....	26
23	MÉDIA DA FDP WEIBULL.....	26
24	VARIÂNCIA DA FDP WEIBULL.....	26
25	FUNÇÃO DENSIDADE DE PROBABILIDADE DO QUI-QUADRADO.....	27
26	MÉDIA DA FDP DO QUI-QUADRADO	27
27	VARIÂNCIA DA FDP DO QUI-QUADRADO	27
28	FUNÇÃO DENSIDADE DE PROBABILIDADE NORMAL OU DE GAUSS.....	28
29	FUNÇÃO $f(x) = \frac{a}{bx^2 + cx + d}$	38

30	FUNÇÃO DENSIDADE DE PROBABILIDADE POLINOMIAL OU FUNÇÃO DE CINCO ESTÁGIOS	41
31	TESTE DO QUI-QUADRADO χ^2	62
32	TESTE DE KOLMOGOROV-SMIRNOV	63
33	FÓRMULAS QUE PRODUZEM O VALOR d_{tab} NO TESTE DE KOLMOGOROV SMIRNOV	63
34	FÓRMULA PARA TRANSFORMAR UM CONJUNTO DE VALORES EM UM INTERVALO ENTRE 0 E 1	66
35	FÓRMULA PARA DETERMINAR O PARÂMETRO α NA DISTRIBUIÇÃO BETA	67
36	FÓRMULA PARA DETERMINAR O PARÂMETRO β NA DISTRIBUIÇÃO BETA	67
37	FÓRMULA PARA DETERMINAR O PARÂMETRO c NA DISTRIBUIÇÃO WEIBULL	68
38	FÓRMULA PARA DETERMINAR O PARÂMETRO b NA DISTRIBUIÇÃO WEIBULL	68

LISTA DE TABELAS

1	FREQÜÊNCIAS OBSERVADAS POR CLASSE DE ALTURA DE <i>Eucalyptus saligna</i> - ANÁLISE EXPLORATÓRIA	31
2	DAP COM CASCA DE <i>Pinus</i> sp - ANÁLISE EXPLORATÓRIA.....	32
3	FREQÜÊNCIA DO JEQUITIBÁ-ROSA	36
4	AJUSTE DOS DADOS DO JEQUITIBÁ-ROSA À DISTRIBUIÇÃO POLINOMIAL	65
5	AJUSTE DOS DADOS DO JEQUITIBÁ-ROSA AO MODELO GAMA.....	66
6	AJUSTE DOS DADOS DO JEQUITIBÁ-ROSA AO MODELO BETA.....	66
7	AJUSTE DOS DADOS DO JEQUITIBÁ ROSA AO MODELO WEIBULL DOIS PARÂMETROS	67
8	AJUSTE DO JEQUITIBÁ-BRANCO À DISTRIBUIÇÃO POLINOMIAL.....	69
9	AJUSTE DO JEQUITIBÁ-BRANCO PELA DISTRIBUIÇÃO GAMA	70
10	AJUSTE DO JEQUITIBÁ-BRANCO PELA DISTRIBUIÇÃO BETA	71
11	AJUSTE DO JEQUITIBÁ-BRANCO PELA DISTRIBUIÇÃO WEIBULL.....	71
12	JEQUITIBÁ-ROSA – ANÁLISE EXPLORATÓRIA.....	72
13	JEQUITIBÁ-BRANCO- ANÁLISE EXPLORATÓRIA.....	73
14	MEDIDAS EMPÍRICAS DOS DADOS DE CIRCUNFERÊNCIAS A 60 CM DE ALTURA DE PLANTAS DE PALMITO	74
15	TESTES DE ADERÊNCIA DOS DADOS DO PALMITO AO MODELO POLINOMIAL	82
16	AJUSTE DOS DADOS DO PALMITO PELA DISTRIBUIÇÃO GAMA.....	83
17	AJUSTE DOS DADOS DO PALMITO PELA DISTRIBUIÇÃO BETA	83
18	AJUSTE DOS DADOS DO PALMITO PELA DISTRIBUIÇÃO WEIBULL.....	84
19	AJUSTE DOS DADOS DO PALMITO PELA DISTRIBUIÇÃO NORMAL	84

RESUMO

Este trabalho tem como objetivo principal apresentar novas opções para o ajuste de distribuições de probabilidades que são utilizadas na Ciência Florestal. Alguns modelos contínuos apresentam certas distorções ao serem implementados no cálculo com dados oriundos de florestas naturais devido à grande variabilidade que se encontra nessas situações. Esse fato foi constatado principalmente quando se estudou a variável altura. Embora o modelo seja construído com dados de alturas de florestas naturais, as fórmulas desenvolvidas poderão ser aplicadas para outras grandezas, principalmente se o gráfico apresentar assimetria, o que afasta a possibilidade de estudo por meio da distribuição normal. Para a realização do trabalho inicialmente foram estudadas maneiras de modelar uma função matemática que pudesse ser transformada em função densidade de probabilidade. A função deveria assumir somente valores positivos, ser contínua e sua integral, considerando todo o intervalo real, deveria convergir para um. Foram feitas várias tentativas com funções matemáticas que, apesar de atenderem às condições de uma função densidade de probabilidade, não eram suficientemente flexíveis para se adaptar às características dos dados de uma floresta natural. Finalmente chegou-se a uma função que é definida por três sentenças, formada por um polinômio de grau n , uma curva crescente de 5.º grau preferencialmente e uma curva decrescente positiva tendendo a zero com integral convergente no infinito. O polinômio explicou a maior parte dos dados e, para as classes onde este não produziu bom ajuste, foram elaboradas outras duas funções. Para os testes iniciais foram utilizados dados de alturas de Jequitibá-Rosa (*Cariniana legalis*), provenientes da Fazenda Reata, situada no município de Cássia, Minas Gerais. Para a referida espécie, com frequência total observada de 493, o valor obtido da frequência esperada foi 492,9. A média obtida no modelo desenvolvido foi 17,8 metros, valor considerado excelente comparado a 18 metros, obtido de forma empírica por meio dos dados observados. Um ótimo ajuste também foi obtido para a variância, o que levou a aplicar a pesquisa a dados de outras espécies e comparar os resultados obtidos com distribuições já conhecidas.

Palavras-chave: probabilidade; ajuste de polinômio; florestas naturais; distribuições assimétricas.

ABSTRACT

The main objective of this research was to introduce new options for the fitting of probability distributions used in Forestry. Some continuous probability distributions present certain distortion when used in calculus with data from natural forests due to the high variability in such situations. This fact was especially noticed in studies of the variable height. Although the object of the study was natural forest's height data, the developed formulas may be applied for other variables as well, especially if the graph is asymmetric which prevents the study to be made by the normal distribution curve. Before the study could be carried out, we did some work to model a mathematical function that could be changed into probability density function, that is, the functional values had to be positive, the function should be continuous, and its integral – considering the whole real interval – had to converge to one. Several attempts were made with mathematical functions that fulfilled the requirements of probability density function, but none was flexible enough to suit the data of a natural forest. Finally, a function was obtained which was defined by several sentences, formed by a n -polynomial, preferably a 5th degree increasing curve, a positive decreasing curve tending to zero, and the integral converging to the infinity. The polynomial explains most of the data; for the cases in which it fails to produce a good fitting, two other functions were created. The species used at first was Jequitibá-Rosa (*Cariniana legalis*), whose data came from a farm located in the municipality of Cássia, in Minas Gerais, Brazil. The total frequency observed for Jequitibá-Rosa, which was 493, was explained by the model with 492.9. The average achieved in the developed model 17.8 m is very close to the one that was calculated directly through observed data whose value is 18 m. An optimal adjustment was also achieved for the variance, leading to extending the research to other species and comparing the data obtained with existing distributions.

Key-words: probability; polynomial fitting; natural forests; asymmetric distributions.

1 INTRODUÇÃO

A estatística tem contribuído muito para o desenvolvimento da Ciência Florestal. Seus conceitos, medidas e testes têm sido fundamentais, tanto para o diagnóstico de determinadas situações como para previsões e tomada de decisões.

Segundo PÉLLICO NETTO e BRENA (1997), atualmente a maioria das decisões humanas é tomada a partir da análise ou estudo de apenas uma parte, do todo que envolve o problema.

Uma teoria que tem sido utilizada de forma significativa na ciência é a das probabilidades. Pode-se diagnosticar uma doença a partir de uma pequena análise de tecido. A partir da análise de um lote significativo de placas de compensado é possível tirar conclusões a respeito de sua qualidade e, com isso, tomar decisões econômicas ou de interesse científico.

As circunstâncias levam a crer que estudar a complexidade e a precisão de fenômenos contribui de forma significativa para a Ciência e para a Engenharia mais especificamente. É indiscutível a importância do estudo mais preciso de fenômenos científicos. Existem muitas relações que, se estudadas, podem produzir informações que permitirão um diagnóstico mais preciso das situações.

O planejamento de um inventário florestal exige análise de fatores que, muitas vezes, precisam ser estimados, pois nem sempre é possível fazer uma completa enumeração ou censo por razões de ordem econômica e ambiental.

Segundo SOONG (1986), as idéias probabilísticas estão presentes em ampla diversidade de investigações científicas que envolvem aleatoriedade. E complementa:

... são inúmeros os fenômenos aleatórios na área científica: ruídos em sinais de rádio, intensidade de vendavais, vibrações mecânicas devidas a perturbações atmosféricas, movimento browniano de partículas em um líquido, números de chamadas telefônicas em certa localidade, comprimento de filas em um guichê, escolha do meio de transporte por um grupo de indivíduos, etc.

Dentro do contexto de aleatoriedade, são formuladas distribuições probabilísticas que permitem prever fenômenos a partir de modelos matemáticos, possibilitando o estudo do comportamento de variáveis do meio físico ou biológico denominadas "funções densidade de probabilidade". Neste trabalho foi efetuada uma análise destas funções, visando gerar contribuições para a Ciência Florestal.

Em um primeiro momento foi feita revisão de literatura com o objetivo de estruturar os principais conceitos matemáticos que seriam utilizados no desenvolvimento do trabalho e apresentar aspectos relevantes das principais distribuições probabilísticas contínuas e suas aplicações na área florestal.

Na seqüência, foi apresentada uma nova metodologia para elaboração de uma função densidade de probabilidade. Também foi elaborado um exemplo com dados de alturas do Jequitibá Rosa, com o objetivo de elucidar as etapas da elaboração do modelo.

Na etapa seguinte foram realizados testes para verificar a aderência do modelo a outros dados. Comparações com outras distribuições também estiveram presentes nessa etapa, a fim de que o leitor pudesse analisar as possíveis aplicações no cotidiano do engenheiro florestal. Uma das maneiras de analisar os resultados foi a aplicação do teste de Kolmogorov-Smirnov.

Finalmente foram feitas conclusões a respeito do trabalho. Nessa oportunidade foi possível verificar que a idéia apresentada aqui é um embrião para o desenvolvimento de muitos estudos. Exemplo disso seria compatibilizar com aplicativos computacionais as idéias que serão mencionadas, de modo que possam atender e tornar mais prático o trabalho do pesquisador em Engenharia Florestal.

1.1 JUSTIFICATIVA

Várias distribuições têm sido utilizadas na área florestal. Se o problema é determinar a probabilidade de árvores que serão obtidas a partir de sementes plantadas, estabelecer o número médio de árvores de uma determinada espécie ou outro objetivo é proposto, são utilizadas distribuições como binomial, Poisson ou outras conhecidas da literatura.

Muitas distribuições discretas são utilizadas. A distribuição binomial negativa pode ser utilizada, por exemplo, para verificar a probabilidade de um povoamento de Eucalipto ser atacada pelo inseto *Eusalasia eucerus*. Para outras situações são utilizadas distribuições contínuas, como Exponencial, Gama, Beta, Weibull, Normal, entre outras.

Em muitos casos o trabalho envolve variáveis que se configuram como contínuas, o que afasta a possibilidade de utilização das distribuições discretas. Já em outros, a situação é contínua, mas existe uma assimetria com relação à media, o que impossibilita a utilização da distribuição normal.

Para situações em que há assimetria são utilizadas principalmente as distribuições Gama, Beta e Weibull. Todavia, em alguns trabalhos, tem ocorrido a dificuldade em ajustar os dados aos modelos, por vários motivos.

Em muitos casos, os dados têm características que fogem aos modelos idealizados nas distribuições e acabam não se ajustando. Às vezes há necessidade de que o modelo seja mais flexível, devido a situações de grande variabilidade dos dados. Este fator faz com que as situações previstas nos modelos não se adaptem aos dados.

Depois de observado que dados de altura de árvores não se ajustavam aos modelos existentes, surgiu a idéia de criar uma distribuição que fosse modelada a partir dos dados, de suas características e peculiaridades. Com esse propósito surgiu a Distribuição Polinomial, que consiste em ajustar o polinômio que mais se aproxima dos dados observados transformando-o em função densidade de probabilidade.

O modelo Polinomial, idealizado para variáveis contínuas, foi concebido mediante a análise de fatores como flexibilidade, praticidade e aplicabilidade em florestas naturais onde o problema do ajuste com os modelos existentes é mais freqüente.

O principal componente da citada distribuição é um polinômio de grau n , ajustado a partir dos dados de cada situação pelo método dos Mínimos Quadrados, com o auxílio de aplicativo computacional. Este método consta de muitas literaturas

relacionadas à Estatística e ao Cálculo Numérico e foi escolhido após várias tentativas de interpolação com diversos métodos já existentes.

Segundo RUGGIERO e LOPES (1996), o método permite o ajuste do polinômio mais próximo a um conjunto de dados e permite extrapolar, o que é bastante oportuno para uma distribuição de probabilidade contínua.

Em muitos casos os polinômios ajustados apresentam raízes reais. Isto indica que, anteriormente ou posteriormente aos zeros, o polinômio apresenta valores negativos que devem ser eliminados, considerando-se que uma função densidade de probabilidade pode ter somente valores positivos.

Com a eliminação dos intervalos em que o polinômio ajustado apresenta valores negativos, surgiu a necessidade de elaborar funções cujo gráfico tivesse forma próxima ao conjunto de dados.

Quando o intervalo eliminado está contido nas primeiras classes do conjunto de dados, ajusta-se uma função polinomial de grau n , na maioria dos casos crescente, que se aproxima dos valores observados.

Quando a parte negativa do polinômio apresenta valores contidos nas últimas classes, o procedimento é semelhante. Entretanto deve-se tomar o cuidado de elaborar uma função cuja integral seja convergente no infinito na sua forma original e após ser transformada, quando se deseja obter a média e a variância do conjunto de dados.

Nos casos estudados neste trabalho, o modelo polinomial ficou constituído por três fórmulas o que, na estatística, se denomina "distribuição truncada". A primeira fórmula é uma função polinomial crescente, a segunda é o polinômio ajustado pelo método dos Mínimos Quadrados e a terceira é uma função decrescente e que tem a reta $y = 0$ como assíntota.

Este trabalho foi elaborado com a finalidade de oferecer mais uma opção em termos de distribuições contínuas. Embora o modelo tenha sua concepção voltada para dados de florestas naturais, ele pode ser utilizado em situações de povoamentos, e a forma como foi elaborado pode possibilitar o surgimento de aplicativos computacionais específicos para a Engenharia Florestal.

1.2 OBJETIVOS

1.2.1 Objetivo Geral

Contribuir com a Ciência Florestal utilizando conceitos matemáticos para a formulação de uma nova função densidade de probabilidade.

1.2.2 Objetivos Específicos

Os objetivos específicos foram:

- a) encontrar um polinômio de grau n que pudesse ser transformado para atender às condições de uma função probabilística;
- b) comparar os resultados da função Polinomial com outras funções densidade de probabilidade existentes;
- c) calcular a média e a variância de um conjunto de dados por meio do modelo desenvolvido;
- d) propor uma função que fosse formulada a partir da característica dos dados, atendendo aos requisitos de uma função densidade de probabilidade.

2 REVISÃO DE LITERATURA

Em GIOTTO (1986) tem-se o primeiro trabalho na área florestal no Brasil, que menciona a possibilidade da aplicação de polinômios em funções densidade de probabilidade.

A estrutura deste trabalho tem suas bases em muitos conceitos matemáticos. Tanto nas explicações como nos exemplos foram citadas definições e tiradas conclusões que envolvem tópicos da Matemática, seja ela elementar ou avançada. Desta forma optou-se por, à luz das devidas referências bibliográficas, fazer uma revisão nas definições que foram utilizadas no texto, para que as conclusões extraídas da pesquisa fossem fundamentadas em premissas comprovadamente verdadeiras.

2.1 TÓPICOS DE MATEMÁTICA

No desenvolvimento do modelo, muitos conceitos matemáticos foram mencionados. Tópicos inseridos nas definições estatísticas foram nesta etapa elucidados de forma a tornar mais clara a compreensão do modelo apresentado.

A seguir são feitas considerações sobre conceitos importantes para a estrutura do trabalho.

2.1.1 Classificação das Funções

Em LEITHOLD (1994) encontram-se as seguintes afirmações relacionadas com a classificação de funções. Se uma função $f(x)$ é dada por $f(x) = c$, sendo c um número real, então diz-se que $f(x)$ é função constante. O gráfico dessa função é uma reta horizontal. A função do primeiro grau definida por $f(x) = x$ é chamada função identidade. É um caso particular da função linear. Seu gráfico é a bissetriz dos quadrantes ímpares.

Se uma função for definida por $f(x) = a_n x^n + a_{n-1} x^{n-1} + \dots + a_1 x + a_0$, onde 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 .

A função linear é um caso particular da função polinomial em que o grau é 1. Se o grau de uma função polinomial for 2, ela será chamada de função quadrática, e se o grau for 3, será chamada de função cúbica. Se uma função puder ser expressa como o quociente de duas funções polinomiais, será chamada de função racional.

Uma classificação mais abrangente define a função algébrica como aquela formada por um número finito de operações algébricas sobre as funções identidade e constante. Essas operações algébricas incluem adição, subtração, multiplicação, divisão, potenciação e radiciação. As funções polinomiais e racionais são tipos especiais de funções algébricas.

Se uma função não é algébrica, ela pode ser transcendente. São funções transcendentess por exemplo, as funções trigonométricas e as logarítmicas.

2.1.2 Funções Crescentes e Decrescentes

A função $f: A \rightarrow B$ definida por $y = f(x)$ é crescente no conjunto $A_1 \subset A$ se, para dois valores quaisquer x_1 e x_2 pertencentes a A_1 , com $x_1 < x_2$, $f(x_1) < f(x_2)$, $A, B \subset \mathfrak{R}$ e A, B intervalos.

Em símbolos: f é crescente quando $(\forall x_1, x_2)(x_1 < x_2) \Rightarrow f(x_1) < f(x_2)$.

A função $f: A \rightarrow B$ definida por $y = f(x)$ é decrescente no conjunto $A_1 \subset A$ se, para dois valores quaisquer x_1 e x_2 pertencentes a A_1 , com $x_1 < x_2$, $f(x_1) > f(x_2)$.

Em símbolos: f é decrescente quando $(\forall x_1, x_2)(x_1 < x_2) \Rightarrow f(x_1) > f(x_2)$.

(IEZZI e MURAKAMI, 1993).

2.1.3 Definição de Limite

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 é único, é indicado por $\lim_{x \rightarrow p} f(x)$. Assim,

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

Em GUIDORIZZI (1987) pode ser encontrada esta definição.

É importante salientar que, para existir o limite, os seus limites laterais devem ser iguais. O limite expressa uma tendência da variável para um determinado valor.

2.1.4 Continuidade

Em PISKOUNOV (1993), são feitas as seguintes considerações a respeito de continuidade de funções:

Seja $y = f(x)$ uma função definida para o valor $x = x_0$ e numa certa vizinhança de centro x_0 . Seja $y_0 = f(x_0)$. Se for dado à variável x um acréscimo Δx positivo ou negativo, ela fica $x_0 + \Delta x$ e a função y sofre igualmente um acréscimo Δy . O novo valor da função é $y_0 + \Delta y = f(x_0 + \Delta x)$. O acréscimo da função é dado pela fórmula $\Delta y = f(x_0 + \Delta x) - f(x_0)$.

A função $y = f(x)$ diz-se contínua para o valor $x = x_0$ se ela está definida numa certa vizinhança do ponto x_0 (e igualmente no ponto x_0) e se $\lim_{\Delta y \rightarrow 0} \Delta y = 0$, o que equivale a $\lim_{\Delta x \rightarrow 0} [f(x_0 + \Delta x) - f(x_0)] = 0$.

A condição de continuidade pode ser escrita da seguinte forma:

$$\lim_{x \rightarrow x_0} f(x) = f(x_0).$$

Geometricamente, a continuidade de uma função num dado ponto significa que a diferença das ordenadas do gráfico da função $y = f(x)$ nos pontos $x_0 + \Delta x$ e x_0 é arbitrariamente pequena em valor absoluto desde que $|\Delta x|$ seja suficientemente pequeno.

Uma função $y = f(x)$ contínua em todo o ponto do intervalo (a, b) , onde $a < b$, diz-se contínua neste intervalo.

Se a função é definida para $x = a$ e se $\lim_{x \rightarrow a^+} f(x) = f(a)$, diz-se que a função $f(x)$ é contínua à direita no ponto $x = a$. Se $\lim_{x \rightarrow b^-} f(x) = f(b)$, diz-se que ela é contínua à esquerda no ponto $x = b$.

Se a função $f(x)$ é contínua em cada ponto do intervalo (a, b) , bem como nas extremidades desse intervalo, diz-se que a função $f(x)$ é contínua no intervalo fechado ou no segmento $[a, b]$.

2.1.5 Derivabilidade e Continuidade

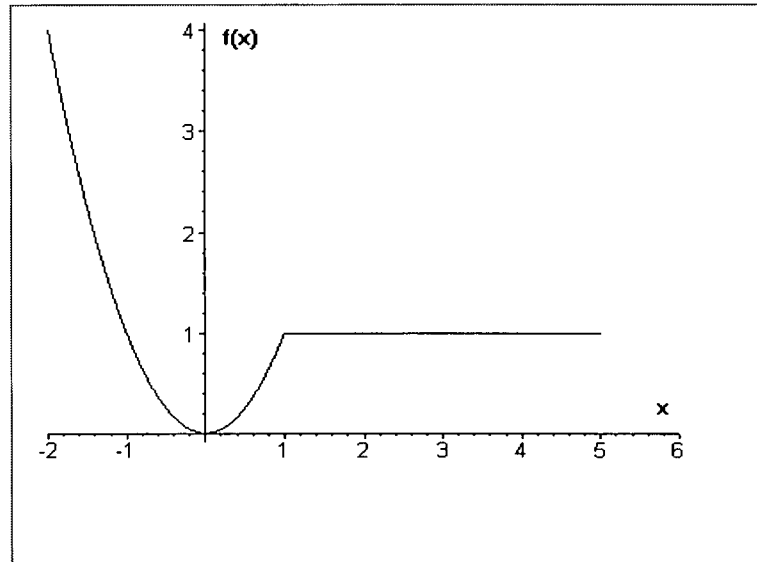
Muitas funções como, por exemplo, $f(x) = |x|$ são contínuas em determinado ponto mas não são deriváveis neste ponto. Deste modo, pode-se concluir que continuidade não implica derivabilidade. Entretanto derivabilidade implica continuidade.

Se uma função $y = f(x)$ tem uma derivada no ponto $x = x_0$, isto é, se o limite $\lim_{\Delta x \rightarrow 0} \frac{\Delta y}{\Delta x} = \lim_{\Delta x \rightarrow 0} \frac{f(x_0 + \Delta x) - f(x_0)}{\Delta x}$ existir, dir-se-á que a função é derivável para o valor $x = x_0$. Se a função tem uma derivada em cada ponto de um intervalo (a, b) , diz-se que ela é derivável sobre este intervalo.

GUIDORIZZI (1987) esclarece a situação apresentando o enunciado de um teorema e um exemplo explicativo. "Se f for derivável em p , então f será contínua em p ". Desta maneira, nos pontos de descontinuidade uma função não pode ter derivada. Como já foi citado, a proposição inversa não é verdadeira, isto é, se uma função $y = f(x)$ é contínua no ponto $x = x_0$, não se pode afirmar que ela seja derivável no ponto x_0 .

A figura 1 mostra um caso em que a função é contínua mas não é derivável.

FIGURA 1 - DERIVABILIDADE E CONTINUIDADE



Neste exemplo percebe-se que f é contínua em $x=1$ mas não é derivável nesse ponto; o gráfico de f apresenta um "bico" no ponto $(1, f(1))$.

O fato de f não ser derivável indica que o limite não existe no valor considerado. De fato a fórmula que define a função é

$$f(x) = \begin{cases} x^2 & \text{se } x \leq 1 \\ 1 & \text{se } x > 1 \end{cases}$$

Para verificar se f é contínua em $x=1$, deve-se verificar se o limite existe.

De fato $\lim_{x \rightarrow 1^+} f(x) = \lim_{x \rightarrow 1^-} f(x) = 1 = f(1)$, então f é contínua em 1.

Para verificar se f é derivável em 1, faz-se

$$\frac{f(x) - f(1)}{x - 1} = \begin{cases} \frac{x^2 - 1}{x - 1} & \text{se } x < 1 \\ 0 & \text{se } x > 1 \end{cases}$$

e verifica-se que o limite pela esquerda vale 2 e pela direita vale zero. Portanto o limite não existe e a função não é derivável em $x=1$ (GUIDORIZZI, 1987).

2.1.6 Teorema Fundamental do Cálculo

Segundo SWOKOWSKI (1994) este teorema foi desenvolvido independentemente por Sir Isaac Newton na Inglaterra e por Gottfried Wilhelm Leibniz na Alemanha. É em razão dessa descoberta simultânea que se atribui a ambos a invenção do cálculo. O autor trata de um conceito importante na referida obra, antes de enunciar o teorema: "Uma função F é uma antiderivada de f em um intervalo I se $F'(x) = f(x)$ para todo x em I ".

Suponha-se f contínua em um intervalo fechado $[a, b]$.

- Parte I: 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 $[a, b]$.
- Parte II: Se F é qualquer antiderivada de f em $[a, b]$, então

$$\int_a^b f(x)dx = F(b) - F(a).$$

Este teorema, além de estabelecer uma relação entre derivadas e integrais, fundamenta o conceito de integral definida e suas aplicações em áreas, volumes e situações em que ele pode ser utilizado.

2.1.7 Integrais Impróprias

Segundo SWOKOWSKI (1994) as integrais impróprias são definidas da seguinte maneira:

- (i) Se f é contínua em $[a, \infty)$, então

$$\int_a^{\infty} f(x)dx = \lim_{t \rightarrow \infty} \int_a^t f(x)dx \text{ e}$$

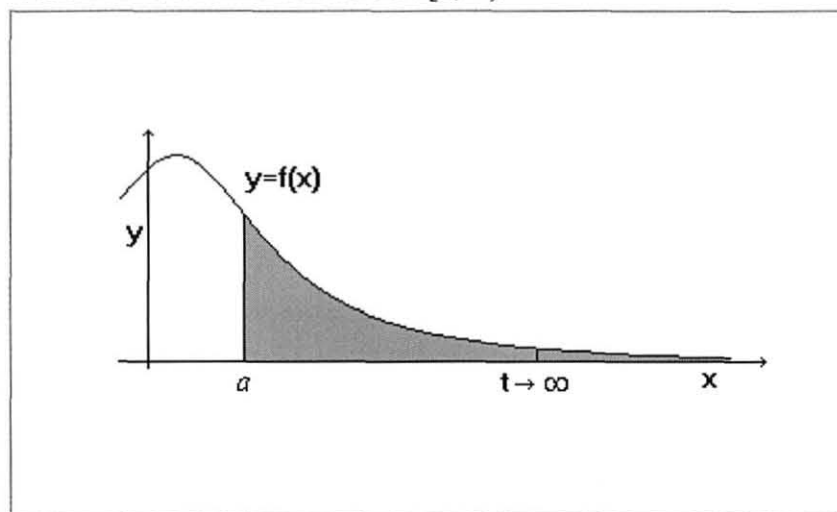
- (ii) Se f é contínua em $[-\infty, a)$, então $\int_{-\infty}^a f(x)dx = \lim_{t \rightarrow -\infty} \int_t^a f(x)dx$, desde que o

limite exista.

As expressões das definições (i) e (ii) são denominadas integrais impróprias. Elas diferem das integrais definidas, pelo fato de um dos limites de integração não ser número real. Diz-se que uma integral imprópria converge se o limite existe; o limite é então o valor da integral imprópria. Se o limite não existe, a integral imprópria diverge.

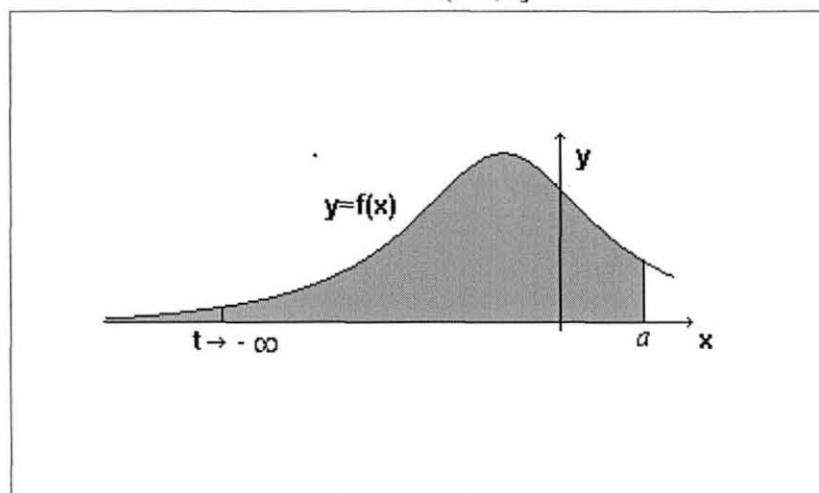
A figura 2 contribui para elucidar o enunciado da definição (i).

FIGURA 2 - INTEGRAL IMPRÓPRIA $[a, \infty)$



A figura 3 ilustra o que é definido em (ii).

FIGURA 3 - INTEGRAL IMPRÓPRIA $(-\infty, a]$



Seja f contínua para todo x . Se a é um número real arbitrário, então;

$$\int_{-\infty}^{\infty} f(x) dx = \int_{-\infty}^a f(x) dx + \int_a^{\infty} f(x) dx$$

desde que as integrais à esquerda e à direita sejam convergentes.

2.1.8 Função Gama

Segundo MEYER (1984) a função gama, denotada por $\Gamma(n)$, é definida por

$$\Gamma(n) = \int_0^{\infty} x^{n-1} e^{-x} dx \quad (1)$$

que é convergente para $n > 0$. A fórmula de recorrência para a função gama é $\Gamma(n+1) = n\Gamma(n)$.

2.1.9 Ajuste de Curvas por Quadrados Mínimos Não Lineares no Matlab

No software Matlab é possível obter o ajuste de um polinômio a um conjunto de dados pelo método dos Quadrados Mínimos Não Lineares.

Em MATSUMOTO (2001) encontra-se o comando "polyfit" na lista de funções do Matlab, o que ele executa e qual método matemático é utilizado: "polyfit: ajuste por polinômio. Dados uma série de valores x , y e o grau do polinômio n , retorna coeficientes do polinômio de grau n que melhor aproxima a curva dada pela série em x, y (método dos mínimos quadrados)".

Em PENNY e LINDFIELD (2000) também podem ser encontrados exemplos de ajuste de polinômios pela função polyfit. Nessa obra o assunto é tratado no capítulo referente à regressão polinomial.

Muitas vezes, em cálculo numérico trabalha-se com o conceito de interpolação. Tendo um conjunto de dados e suas respectivas ordenadas, na interpolação dispõe-se de vários métodos em que o polinômio interpolador "passa" obrigatoriamente por todos os pontos. Entretanto em algumas situações a interpolação não é o melhor método.

Segundo RUGGIERO e LOPES (1996), a interpolação não é aconselhável quando:

- a) é preciso obter um valor aproximado da função em algum ponto fora do intervalo de tabelamento, ou seja, quando se quer extrapolar;
- b) os valores tabelados são resultantes de algum experimento físico ou de alguma pesquisa, porque, nesses casos, estes valores poderão conter erros inerentes que, em geral, não são previsíveis.

E complementam: "Surge então a necessidade de ajustar a estas funções tabeladas uma função que seja uma "boa aproximação" para os valores estabelecidos e que permita "extrapolar" com certa margem de segurança".

Para o ajuste de polinômios o método de mínimos quadrados é bastante utilizado. É o que afirma CUNHA (2000):

O método dos quadrados mínimos é provavelmente a técnica de aproximação mais usada na análise numérica ou em problemas práticos. Isto se deve tanto à sua simplicidade quanto ao fato de que, em geral, buscam-se aproximações para valores que são medidas obtidas experimentalmente com um certo grau de incerteza. Este método foi publicado, pela primeira vez, por Adrien-Marie Legendre (1752-1833) em 1805, mas ele já era usado por Carl Friedrich Gauss (1777-1855), em seus cálculos na Astronomia, desde 1795.

O método dos quadrados mínimos, como o próprio nome sugere, consiste em aproximar os dados minimizando os resíduos. Portanto é necessário minimizar a distância entre cada valor estimado e o correspondente observado.

A minimização da distância é feita com base nos conceitos estudados no Cálculo Diferencial e Integral. Em Cálculo estuda-se que, nos pontos críticos, a derivada de primeira ordem deve ser nula. A partir dessas idéias são deduzidas as equações normais que fazem parte do método citado.

Como o método dos mínimos quadrados é bastante conhecido, bem como suas equações normais, preferiu-se omitir a sua estruturação matemática, uma vez que está disponível em vários textos tanto de cálculo numérico como de estatística.

Vale ressaltar que, por uma questão de praticidade, o método será aplicado com o auxílio do computador, o que já foi fundamentado na seção 2.1.9.

2.2 TÓPICOS DE ESTATÍSTICA

2.2.1 Variáveis Aleatórias Discretas e Contínuas

Existem funções que associam números reais aos eventos de um espaço amostral. Essas funções são chamadas de "variáveis aleatórias". Isso quer dizer que o resultado de um experimento pode ser descrito por meio de números, o que torna mais fácil o tratamento matemático do fenômeno.

Considere-se o experimento jogar um dado e a seguinte função $X =$ "o triplo do produto obtido mais dois"; X será uma variável aleatória que poderá assumir os valores decorrentes dos resultados do dado. Esses valores são 5, 8, 11, 14, 17, 20 e têm todas probabilidades iguais a $1/6$. Assim tem-se $(5, 1/6)$, $(8, 1/6)$, $(11, 1/6)$, $(14, 1/6)$, $(17, 1/6)$ e $(20, 1/6)$, o conjunto de pares de cada variável aleatória.

Para a variável aleatória em questão, estão associados os pares possíveis resultantes de sua definição. Esse é o exemplo de uma variável aleatória discreta, pois os possíveis resultados da variável formam um conjunto enumerável de valores. Nesses casos a distribuição de probabilidade é caracterizada por uma função de probabilidade que associa probabilidades não nulas aos possíveis valores da variável aleatória, e zero aos demais valores. A função probabilidade pode ser dada por uma tabela, um gráfico, uma expressão analítica ou outra forma.

Considere-se agora que os possíveis resultados de um experimento são representados pelos infinitos valores de um intervalo contínuo. Assim, a variável aleatória será dita contínua e pode, em determinado intervalo, assumir qualquer valor. Por exemplo: ao estudar a vida de uma árvore, a sua idade pode ser considerada uma variável aleatória contínua, pois pode assumir qualquer valor maior do que 0 até um valor determinado ou indeterminado, de acordo com a espécie.

Diante do que foi exposto, as variáveis contínuas podem assumir uma infinidade de valores em determinados intervalos. Além disso, a probabilidade correspondente a cada possível valor individualmente considerado passa a ser zero, o que não significa que seja um resultado impossível, uma vez que a probabilidade deve ser considerada como uma de

área. Portanto, nas variáveis aleatórias contínuas interessam os valores de probabilidades que a variável assume em determinados intervalos. Esses valores podem ser determinados por meio das distribuições de probabilidade.

2.2.2 Função Densidade de Probabilidade

Muitas situações na prática envolvem variáveis contínuas. Por exemplo: ao se medir a altura de uma árvore como o Jequitibá-Rosa, pode ser encontrado o valor 52,3 metros. Como a altura pode assumir infinitos resultados, é considerada uma variável contínua e nessas situações costuma-se estudar a variável inserida em intervalos, por exemplo em (50, 55].

Em grande parte dos estudos na área florestal as variáveis são consideradas contínuas. Este fato ocorre porque medidas como DAP (diâmetro à altura do peito), CAP (circunferência à altura do peito), áreas e volumes, podem ser expressos por uma infinidade de valores dentro de um intervalo considerado.

Segundo PÉLLICO NETTO (1993), entendem-se as distribuições contínuas como limites de uma infinita seqüência de distribuições discretas.

No caso de variável aleatória contínua, a função de probabilidade recebe a denominação de *função densidade de probabilidade*, que deve verificar as seguintes condições:

$$1^a) f(x) \geq 0, \text{ para todos os valores de } x;$$

$$2^a) \int_{-\infty}^{+\infty} f(x) dx = 1;$$

$$3^a) \int_a^b f(x) dx = P(a < X \leq b), \quad b > a;$$

- a primeira condição decorre do fato de não haver probabilidade negativa;
- a segunda condição afirma que a área sob o gráfico da função é igual a 1.

- a última condição indica que 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.

Os resultados impossíveis, no caso de variáveis aleatórias contínuas, serão caracterizados por uma função densidade de probabilidade nula no intervalo considerado.

2.2.3 Função de Distribuição Acumulada

Pode-se caracterizar a distribuição de probabilidade de uma variável aleatória por meio de sua função de distribuição acumulada, que será representada pela letra maiúscula correspondente à minúscula utilizada para a função densidade.

Esta função, definida por $F(x) = P(X \leq x)$, fornece para qualquer ponto considerado a probabilidade de que a variável aleatória assuma um valor menor ou igual ao correspondente a esse ponto. Para as variáveis aleatórias contínuas a função é dada por $F(x) = \int_{-\infty}^x f(t) dt$. (2)

Portanto, conhecida a função densidade de probabilidade, pode-se determinar a sua função de distribuição.

De forma contrária, sendo conhecida a função de distribuição, pode-se determinar a função densidade de probabilidade. A função densidade de probabilidade será obtida por derivação da função de distribuição, pois $F(x)$ resulta da integração de $f(x)$. Segundo GMURMAN (1983) A função de distribuição $F(x)$ tem as seguintes propriedades.

- $F(a) < F(b)$ para qualquer par de valores reais a e b tal que $b > a$;
- $0 \leq F(x) \leq 1$;
- $F(-\infty) = 0$ e $F(+\infty) = 1$;
- $F(x)$ é sempre crescente;
- $F(b) - F(a) = P(a < X \leq b)$, $b > a$.

2.2.4 Média e Variância de uma Função Densidade de Probabilidade

Pode-se determinar essas medidas por meio da Expectativa Matemática ou pela Função Geratriz de Momentos.

Em GUIDORIZZI (1994) encontram-se as seguintes definições para média e variância:

Seja X uma variável aleatória contínua com função densidade de probabilidade f . Define-se o valor esperado $E(X)$ de X por

$$E(X) = \int_{-\infty}^{+\infty} x \cdot f(x) dx \quad (3)$$

e a variância σ_x^2 de X por $\sigma_x^2 = \int_{-\infty}^{+\infty} [x - E(x)]^2 \cdot f(x) dx$ desde que as integrais impróprias sejam convergentes.

Também se pode escrever a expressão da variância da seguinte forma:

$$\sigma_x^2 = \int_{-\infty}^{+\infty} x^2 \cdot f(x) dx - [E(X)]^2 \quad (4)$$

Estas expressões serão utilizadas para calcular a média e a variância dos resultados obtidos no modelo proposto neste trabalho. Pretende-se comparar os valores obtidos no modelo desenvolvido com os obtidos em outras distribuições já conhecidas e com a média e a variância calculadas de maneira empírica utilizando os dados observados.

2.3 PRINCIPAIS DISTRIBUIÇÕES CONTÍNUAS

Vários pesquisadores desenvolveram fórmulas para expressar a probabilidade em situações em que a variável se configura como contínua.

Muitas distribuições resultam de casos particulares de outras, como é o da distribuição exponencial, obtida quando se atribui o valor 1 para o parâmetro α na distribuição gama. Outros modelos contínuos como, por exemplo, o normal, são amplamente utilizados quando se tem simetria no conjunto de dados. Em casos onde há assimetria no conjunto de dados recorre-se também a distribuições como Beta e Weibull.

Neste trabalho foi desenvolvido o modelo Polinomial, cuja formulação segue uma lógica diferente da utilizada em outras distribuições. Foram feitas comparações com resultados obtidos em outros modelos contínuos tanto na média e na variância como em testes de aderência. Com a finalidade de tornar mais prática a comparação feita no capítulo 4, é apresentado a seguir um resumo das principais distribuições contínuas.

2.3.1 Distribuição Retangular Uniforme

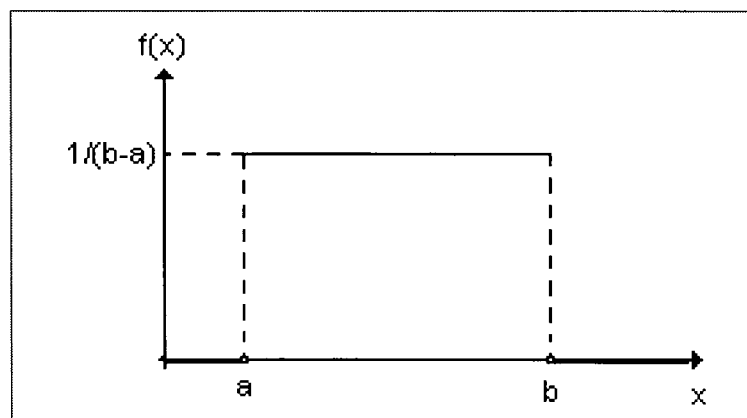
Considere-se um jogo de roleta em que qualquer valor pode ocorrer com a mesma probabilidade. Neste caso, pode-se dizer que a distribuição é uniforme e passa a ser definida assim:

Segundo MAGALHÃES e LIMA (2002) uma variável aleatória X tem distribuição *retangular uniforme* no Intervalo $[a, b]$, $a < b$, se sua função densidade de probabilidade é dada por:

$$f(x) = \begin{cases} \frac{1}{b-a}, & a \leq x \leq b \\ 0, & \text{e.o.c.} \end{cases} \quad (5)$$

A figura 4 mostra a o gráfico da distribuição.

FIGURA 4 - DISTRIBUIÇÃO RETANGULAR UNIFORME



A função de distribuição é dada pela seguinte expressão:

$$F(x) = \begin{cases} 0 & \text{se } x < a \\ \frac{x-a}{b-a} & \text{se } a < x \leq b \\ 1 & \text{se } x > b \end{cases} \quad (6)$$

A média e a variância através da expectativa matemática são:

- média da distribuição:

$$\mu_x = \frac{a+b}{2} \quad (7)$$

- variância:

$$\sigma^2 = \frac{(b-a)^2}{12} \quad (8)$$

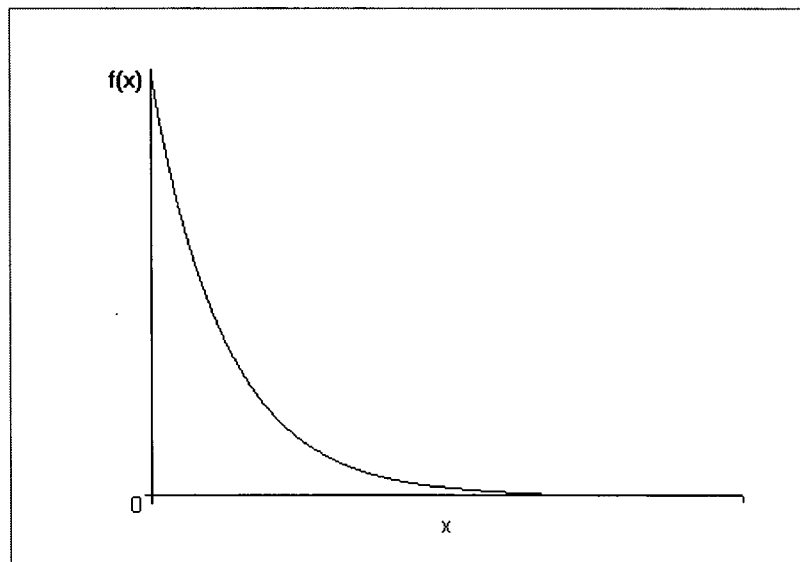
2.3.2 Distribuição Exponencial

A função densidade de probabilidade exponencial é dada por

$$f(x) = \begin{cases} \frac{1}{\beta} \cdot e^{-\frac{x}{\beta}} & \text{para } x \geq 0, \beta > 0 \\ 0 & \text{e.o.c} \end{cases} \quad (9)$$

Um gráfico do modelo é mostrado na figura 5.

FIGURA 5 - DISTRIBUIÇÃO EXPONENCIAL



A função de distribuição para a função densidade de probabilidade exponencial é dada por

$$F(x) = \begin{cases} 0 & \text{se } x < 0 \\ 1 - e^{-x/\beta} & \text{se } x \geq 0 \end{cases} \quad (10)$$

A média e a variância através da expectativa matemática são:

- média da distribuição:

$$E(x) = \beta \quad (11)$$

- variância:

$$\sigma_x^2 = \beta^2 \quad (12)$$

Nessa distribuição, à medida que x aumenta, a função $f(x)$ resulta em probabilidades menores, pelo fato de se tratar do inverso de $e^{x/\beta}$. Desta forma pode-se concluir que a função é sempre decrescente. Este fato traz algumas limitações para sua aplicação.

Há uma série de situações em que os dados se distribuem pelo modelo exponencial. Especialmente no campo biológico e físico há inúmeros casos, como a vida de colônias de bactérias, a vida útil de lâmpadas etc.

A referida distribuição também pode ser utilizada para descrever o comportamento do diâmetro de uma floresta, em que $f(x)$ é a frequência e x representa a variável contínua dos diâmetros.

2.3.3 Distribuição Gama

Segundo ROUSSAS (1997), uma distribuição probabilística em que a variável X se caracteriza na função de frequência como segue,

$$f(x) = \begin{cases} \frac{x^{\alpha-1} \cdot e^{-x/\beta}}{\beta^\alpha \cdot \Gamma(\alpha)} & \text{para } x \geq 0 \\ 0 & \text{e.o.c.} \end{cases} \quad (13)$$

onde α e β são valores positivos, é conhecida como função densidade de probabilidade gama que dá nome à distribuição.

Diz-se que a função gama é completa quando os limites de sua integração são tomados entre 0 e ∞ .

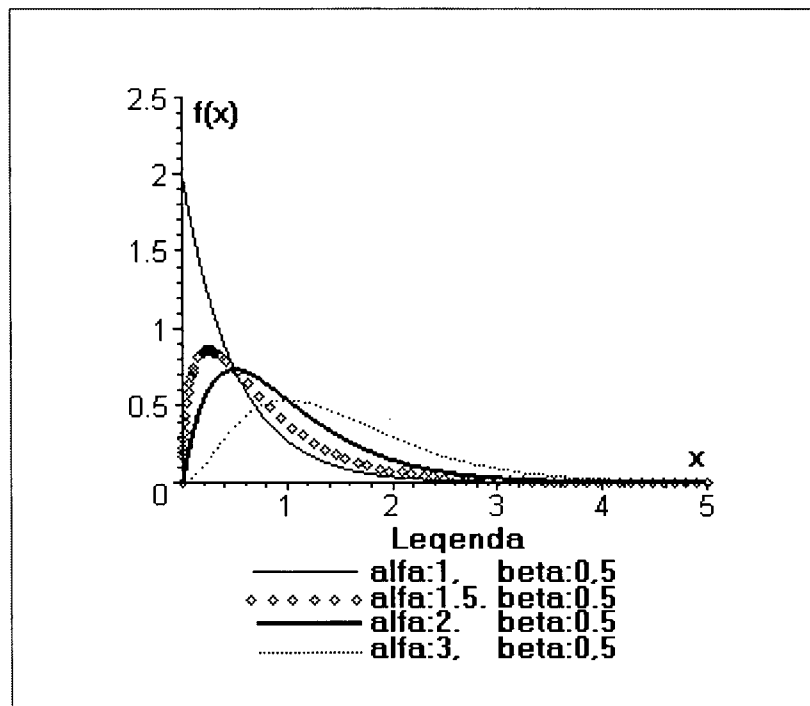
Esta função somente poderá ser avaliada pelas técnicas analíticas se α for um valor inteiro. Quando α não for um valor inteiro, esta integral somente poderá ser avaliada por integração numérica ou por expansão de séries. Por estas razões existem tabelas para a função gama.

A expressão da função gama, fórmula (1), também recebe o nome de função fatorial.

Para se ter uma idéia melhor de operacionalização da distribuição gama, diversas formas foram tabuladas a partir da fórmula 13.

Conforme são atribuídos valores aos parâmetros, surgem curvas que expressam diferentes formas que a distribuição gama pode assumir. A figura 6 mostra algumas destas curvas e suas mudanças no gráfico, quando se mantém β constante e se aumenta o valor de α .

FIGURA 6 - DISTRIBUIÇÃO GAMA



É possível demonstrar que β funciona como fator de escala, e que α é o parâmetro que afeta a forma da função de frequência. Esta função é geralmente assimétrica, mas sua assimetria diminui à medida que α aumenta.

A função de distribuição acumulativa para o modelo gama é:

$$F(x) = \begin{cases} \frac{1}{\beta^\alpha \cdot \Gamma(\alpha)} \int_0^x u^{\alpha-1} e^{-u/\beta} du & \text{se } x \geq 0 \\ 0 & \text{e.o.c} \end{cases} \quad (14)$$

$$\text{A média da distribuição é dada por } E(x) = \alpha \cdot \beta \quad (15)$$

$$\text{e sua variância é dada por } \sigma^2 = \alpha \cdot \beta^2 \quad (16)$$

Esta distribuição tem sido aplicada a funções que representem distribuição diamétrica em diversos tipos de florestas.

2.3.4 Distribuição Beta

Considere-se uma variável x que se distribui segundo a distribuição beta com parâmetros α e β . Sua função de frequência é conhecida como

$$f(x) = \frac{\Gamma(\alpha + \beta)}{\Gamma(\alpha) \cdot \Gamma(\beta) \cdot (b - a)^{\alpha + \beta - 1}} (x - a)^{\alpha - 1} (b - x)^{\beta - 1}$$

se $a < x < b$, ou 0 em outros casos. (17)

De acordo com ROSS(1998), distribuição é mais conhecida na sua forma mais simplificada quando $a=0$ e $b=1$. A função de frequência se reduz para a forma

$$f(x) = \frac{\Gamma(\alpha + \beta)}{\Gamma(\alpha) \cdot \Gamma(\beta)} x^{\alpha - 1} (1 - x)^{\beta - 1} \quad \text{com } 0 < x < 1 \quad (18)$$

A função de distribuição para o modelo beta é dada pela seguinte expressão:

$$f(x) = \begin{cases} 0 & \text{para } x \leq 0 \\ \frac{\Gamma(\alpha + \beta)}{\Gamma(\alpha)\Gamma(\beta)} \int_0^x u^{\alpha-1}(1-u)^{\beta-1} du & \text{se } 0 \leq x \leq 1 \\ 1 & \text{se } x > 1 \end{cases} \quad (19)$$

$$\text{A média da distribuição é dada por } E(x) = \frac{\alpha}{\alpha + \beta} \quad (20)$$

$$\text{e sua variância é dada por } \sigma^2 = \frac{\alpha \cdot \beta}{(\alpha + \beta)^2 (\alpha + \beta + 1)} \quad (21)$$

ambas obtidas pela expectativa matemática.

Esta distribuição tem uma aplicação ampla no estudo de vários fenômenos biológicos em que a variável x é aleatória e contínua. Dada sua grande flexibilidade de adaptação, atende a quase todos os casos de distribuições diamétricas em diversas formações florestais.

Entretanto existem os casos em que há dificuldade para ajustar os parâmetros α e β , como por exemplo, quando eles são expressos em números não inteiros. Além disso, é importante salientar que em alguns casos, por mais que se faça uma variação nos parâmetros, estes podem não produzir a curva que se deseja.

Se $\alpha = 1$ e $\beta = 2$ ou vice-versa o gráfico é uma reta inclinada. E se os valores de α e β forem iguais a 1, o gráfico é uma reta horizontal.

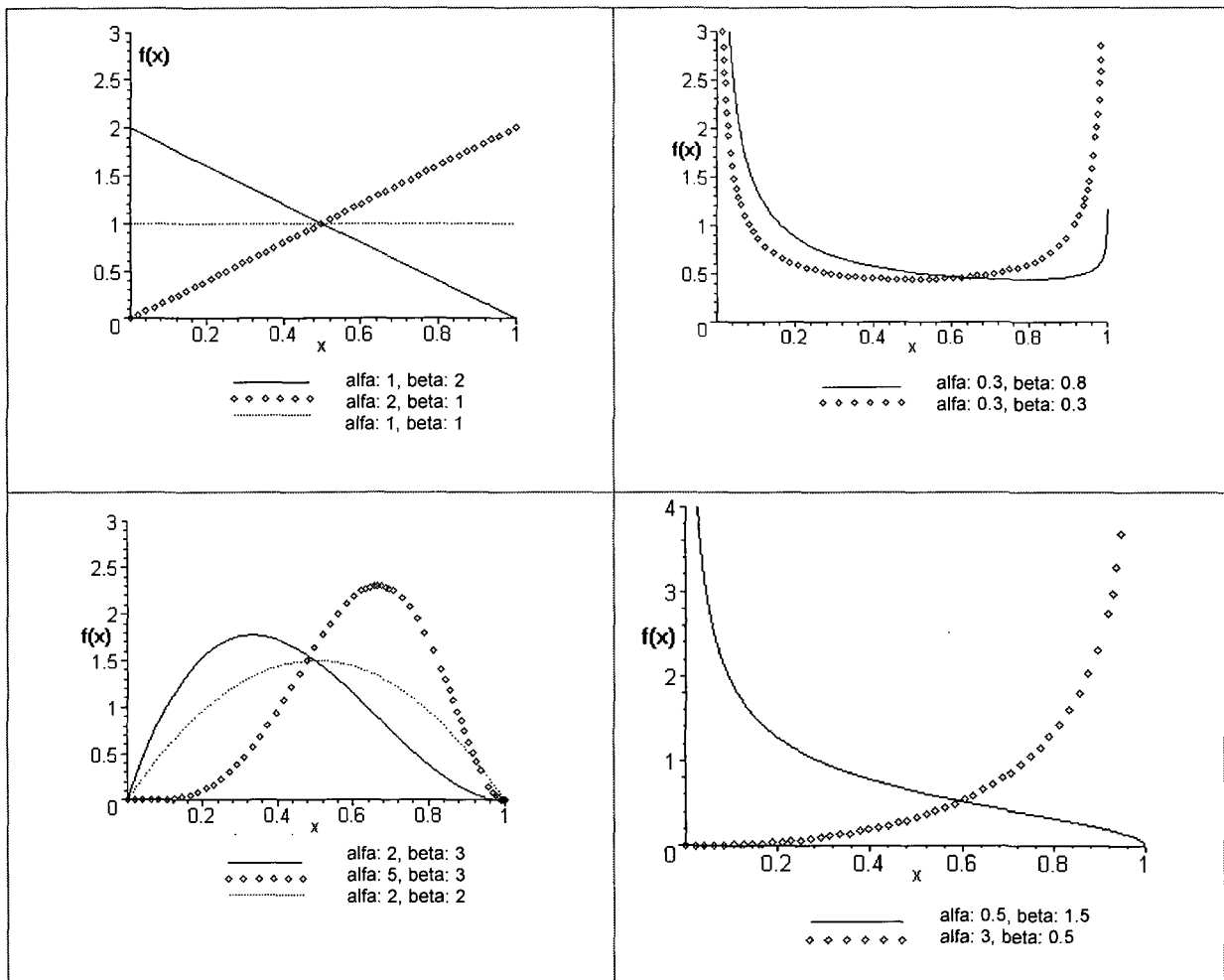
Se os valores de α e β forem 0,3 e 0,8 combinados, o gráfico tem a forma de um **U**.

A curva assume um aspecto de simetria quando α e β tem o valor 2. A assimetria é positiva quando $\alpha = 2$ e $\beta = 3$ e negativa quando $\alpha = 5$ e $\beta = 3$.

Quando $\alpha = 0,5$ e $\beta = 1,5$ tem-se o formato de **J invertido** e quando $\alpha = 3$ e $\beta = 0,5$ o formato de **J**.

A figura 7 mostra as diferentes formas que a função beta pode assumir conforme variam os seus parâmetros.

FIGURA 7 – DISTRIBUIÇÃO BETA PARA ALGUNS VALORES DOS PARÂMETROS



2.3.5 Distribuição Weibull

A distribuição de Weibull usada crescentemente no setor florestal, vem demonstrando ser muito flexível e apropriada para ajustamento das distribuições diamétricas em que há assimetria.

Esta distribuição pode ser considerada como uma forma mais genérica da distribuição exponencial por ter três parâmetros que determinam a amplitude, a forma e também os momentos.

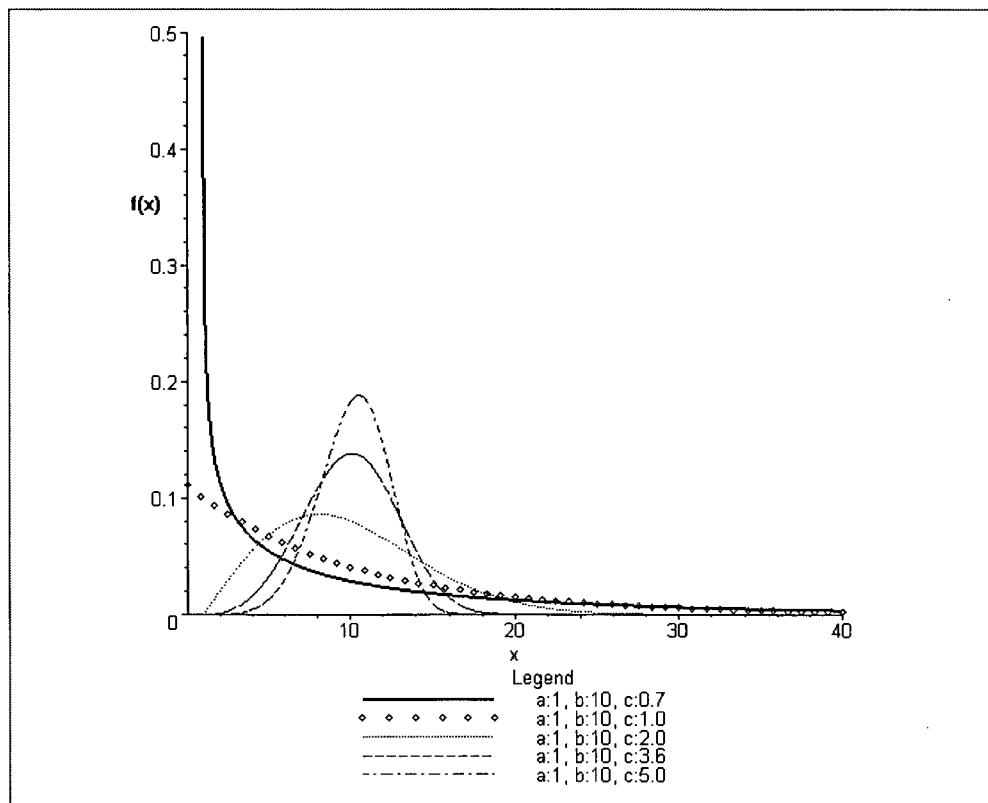
Sua expressão é a seguinte

$$f(x) = \begin{cases} \frac{c}{b} \left(\frac{x-a}{b} \right)^{c-1} \cdot e^{-\left[\left(\frac{x-a}{b} \right)^c \right]}, & x \geq a, b > 0, c > 0 \\ 0 & \text{e.o.c} \end{cases} \quad (22)$$

A função não é linear e apresenta a dificuldade de não ser linearizável para permitir a obtenção dos estimadores dos parâmetros pelo método dos mínimos quadrados. Para sua utilização em distribuições diamétricas, o parâmetro "a" será tomado como diâmetro mínimo dado que $x \geq a$. O parâmetro b caracteriza a escala da curva e o c é o que caracteriza a forma, ou seja, define o tipo de curva.

A figura 8 mostra as formas que a distribuição Weibull pode assumir mantendo a e b constantes e variando o parâmetro c.

FIGURA 8 - DISTRIBUIÇÃO WEIBULL PARA ALGUNS VALORES DOS PARÂMETROS



A média da distribuição é $E(x) = a + b \cdot \Gamma\left(1 + \frac{1}{c}\right)$ (23)

e a sua variância é dada por $\sigma_x^2 = b^2 \cdot \sigma_y^2$ (24)

onde $\sigma_y^2 = \Gamma\left(1 + \frac{2}{c}\right) - \Gamma^2\left(1 + \frac{1}{c}\right)$.

A respeito desta distribuição observe-se que os parâmetros apesar de responsáveis pela forma ou pela escala, não conseguem muitas vezes fazer com que se produza uma curva que se ajuste bem a um determinado conjunto de dados, cujas características nem sempre seguem um modelo pré-fixado de curva. Seria então necessário saber qual a curva que atende aos dados, para posteriormente tentar ajustá-la pelos parâmetros da distribuição.

2.3.6 Distribuição χ^2

A distribuição do qui-quadrado é um caso particular da distribuição gama.

Para obter a distribuição faz-se na distribuição gama

$$\alpha = \frac{\nu}{2} \text{ e } \beta = 2$$

onde ν , número de graus de liberdade da amostragem, é um número inteiro positivo.

Obtém-se, então uma família de distribuições de um parâmetro com FDP

$$f(x) = \begin{cases} \frac{1}{2^{\nu/2} \Gamma(\nu/2)} x^{(\nu/2)-1} \cdot e^{-x/2}, & x \geq 0 \\ 0 & \text{e.o.c} \end{cases} \quad (25)$$

Uma variável aleatória x que tenha fdp dada pela equação anterior terá uma distribuição de qui-quadrado, com ν graus de liberdade.

$$\text{A média é dada por } E(x) = \nu \quad (26)$$

$$\text{e a variância é dada por } \sigma_x^2 = 2\nu \quad (27)$$

A distribuição do qui-quadrado tem várias aplicações importantes em inferência estatística, em virtude das quais a sua distribuição está tabulada para diferentes valores do parâmetro n . Deste modo, achando-se em tabela os valores, pode-se verificar se os resultados obtidos são aderentes ou não a modelos utilizados.

Outros exemplos de aplicação desta distribuição podem ser encontrados em testes de homogeneidade de variâncias, teste de ajuste de dados a qualquer

distribuição e em testes de hipóteses com dados discretos em geral. Ela é usada ainda para se estabelecer intervalo de confiança para a variância estimada na amostra (PÉLLICO NETTO, 1993).

Uma aplicação importante desta distribuição tem sido o teste de aderência, encontrado em quase todo o material bibliográfico produzido em estatística. Muitos livros contêm tabela com os valores do qui-quadrado. Em alguns casos surgem dúvidas de como estabelecer o número de graus de liberdade ou como saber se o teste pode ser aplicado para classes cuja frequência seja muito pequena. Por isso é interessante ser bastante criterioso, na aplicação do teste e verificar se o resultado está condizente com o que se quis testar.

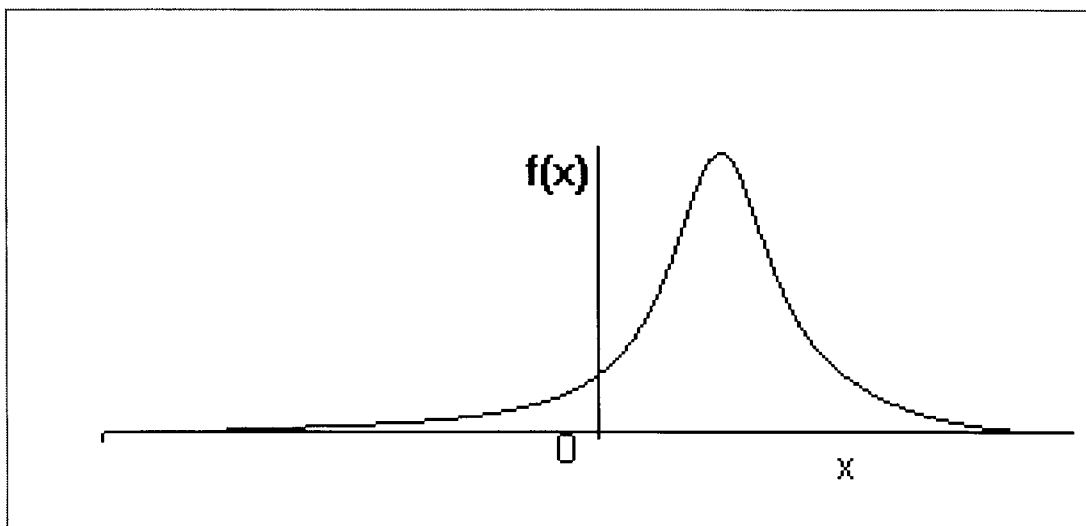
2.3.7 Distribuição Normal ou de Gauss

A expressão da distribuição normal é

$$f(x) = \frac{1}{\sqrt{2\pi} \sigma} e^{-\frac{1}{2} \left(\frac{x-\mu}{\sigma} \right)^2} \quad (28)$$

e sua representação gráfica está apresentada na figura 9. A média é igual a μ e a variância é igual a σ^2 .

FIGURA 9 - DISTRIBUIÇÃO NORMAL



A distribuição normal tem ampla aplicação no campo da estatística experimental, pois grande parte das variáveis contínuas se distribui segundo esta função probabilística.

Nos estudos na área florestal podem ocorrer casos de povoamentos ou outras situações que se distribuam de acordo com o modelo normal. Entretanto existem situações cujo conjunto de dados gera uma curva assimétrica. Nestas condições o modelo normal dificilmente terá um bom ajuste. Em florestas naturais, por exemplo, conforme a situação que se está estudando, dificilmente o modelo normal será aderente, pois, como não há controle das espécies pode se ter assimetria no conjunto de dados devido à sua grande variabilidade.

3 DESENVOLVIMENTO DA FUNÇÃO DENSIDADE DE PROBABILIDADE

Na elaboração do modelo proposto neste trabalho várias fórmulas foram analisadas até se chegar a uma expressão que poderia ser um caminho prático e útil para o trabalho do Engenheiro Florestal.

O objetivo era propor uma função densidade de probabilidade flexível e aplicável a dados de diâmetro, altura, circunferência e outras medidas de utilidade para as pesquisas na área florestal.

Pelo fato de a distribuição normal já ser uma excelente alternativa para os casos de curvas simétricas, optou-se por desenvolver fórmulas cujo gráfico fosse assimétrico ou que pudessem atender aos casos onde as distribuições existentes não produzem um bom ajuste.

3.1 DIAGNÓSTICO DE AJUSTES FEITOS COM DISTRIBUIÇÕES EXISTENTES

A pesquisa para desenvolver uma nova distribuição probabilística surgiu a da necessidade de ajustar funções quando as que existem na literatura não produziam bons resultados. Por exemplo, em determinado conjunto de dados muitas vezes torna-se difícil ajustar uma das funções já desenvolvidas porque os dados fazem com que o gráfico assumam formas que não estão previstas na combinação de parâmetros.

Para ilustrar o que foi citado foram selecionados dados de alturas de *Eucalyptus saligna*. Na tabela 1 é possível verificar a tendência dos dados.

TABELA 1 - FREQUÊNCIAS OBSERVADAS POR CLASSE DE ALTURA DE *Eucalyptus saligna* - ANÁLISE EXPLORATÓRIA

x_i	f_o	$x_i \cdot f_o$	$x_i^2 \cdot f_o$	$f_o (x-u)^3$	$f_o (x-u)^4$
3	1	3	9	-1821, 94	22252, 56
5	5	25	125	-5327, 35	10882, 32
7	7	49	343	-3878, 86	4551, 36
9	13	117	1053	-3118, 75	1490, 67
11	22	242	2662	-1645, 86	315, 23
13	62	806	10478	-672, 528	24, 01
15	122	1830	27450	-1, 18956	0, 01
17	168	2856	48552	957, 6766	10, 18
19	37	703	13357	2008, 485	205, 53
21	3	63	1323	581, 2171	1121, 04
	440	6694	105352	-12919, 1	40852, 93
Media =	15, 21			Assimetria =	-1, 29
Desvio padrão	2, 82			Curtose =	1, 45
				Variância	7,99

No anexo 1 foram colocados valores resultantes da aplicação do modelo Weibull de dois parâmetros aos dados. O teste de Kolmogorov-Smirnov indicou um valor calculado $d_{calc} = 0,28$, quando o tabelado foi $d_{tab} = 0,06$. Assim, pôde-se concluir que os dados da altura não foram aderentes ao modelo Weibull.

O resultado do teste do anexo 1 se reflete no somatório das frequências esperadas 403,60 sendo que o observado é 440. Apesar disso a média no modelo Weibull foi 13,89m, próxima de 15,21m, já a variância resultou em 30,81m² sendo a fornecida 7,99m².

Entretanto, analisando dados de diâmetro de pinus pôde-se constatar que as distribuições Beta e Weibull de dois parâmetros podem ser ajustadas. .

Os dados foram obtidos de uma parcela de 600 m² medidos em uma floresta de *Pinus* sp. com 12 anos de idade como mostra a tabela 2.

TABELA 2 - DAP COM CASCA DE *Pinus* sp - ANÁLISE EXPLORATÓRIA

x_i	f_o	$f_o \cdot x_i$	x_i^2	$x_i^2 \cdot f_o$
8,3	3	24,9	68,89	206,67
10,3	2	20,6	106,09	212,18
12,3	2	24,6	151,29	302,58
14,3	4	57,2	204,49	817,96
16,3	8	130,4	265,69	2125,52
18,3	9	164,7	334,89	3014,01
20,3	9	182,7	412,09	3708,81
22,3	12	267,6	497,29	5967,48
24,3	8	194,4	590,49	4723,92
26,3	7	184,1	691,69	4841,83
28,3	5	141,5	800,89	4004,45
30,3	2	60,6	918,09	1836,18
32,3	1	32,3	1043,29	1043,29
	72	1485,6		32804,88
MÉDIA	20,6333			
VARIÂNCIA	30,3099			
DESVIO PADRÃO	5,50544			

No anexo 2 tem-se os resultados do ajuste dos dados da tabela 2 ao modelo Beta. Obteve-se como freqüência esperada 71,6 sendo que a observada foi 72. Essa proximidade de valores se reflete no teste de Kolmogorov-Smirnov Onde o valor d_{calc} foi 0,0418 e o d_{tab} 0,1603 , o que indica um bom ajuste para 95% de confiança.

Para o modelo Weibull de dois parâmetros a freqüência esperada teve valor próximo de 71,6 . Na estatística de Kolmogorov-Smirnov $d_{calc}=0,0370$ e $d_{tab}=0,1603$ o que indica que os dados se ajustam bem ao modelo como mostra o anexo 3.

Em trabalho realizado com circunferências de palmito a 60 cm de altura do solo, verificou-se que houve um bom ajuste aos modelos Beta, Gama, Weibull de dois parâmetros e Normal.

Dos resultados citados e calculados nos anexos 1, 2 e 3 pôde-se concluir que o ajuste pode ser bom com as funções já existentes na literatura mas existem casos em que o ajuste não ocorre. Para os casos em que não há ajuste aos modelos já existentes estudou-se uma maneira de propor uma função cujo processo passa agora a ser descrito.

3.2 ESTUDO DE MODELO PARA DADOS DO JEQUITIBÁ-ROSA

Como o modelo Polinomial foi testado inicialmente com dados de Jequitibá-Rosa (*Cariniana legalis*), e Jequitibá-Branco (*Cariniana estrellensis*) foi feita uma pesquisa dos dados apresentados para estas espécies obtidos por meio de trabalhos realizados na Fazenda Reata, em Cássia, MG, por ocasião da realização do projeto PELD. Esses dados, complementados posteriormente por pesquisa em toda a área florestal, são resultantes de um acompanhamento contínuo dos Jequitibás.

Em inventário realizado na região para a dissertação de mestrado de CORAIOLA (1997), é possível verificar algumas características que podem auxiliar na caracterização do ambiente onde os Jequitibás estão inseridos.

Na dissertação de Coraiola encontra-se o seguinte:

Os dados básicos utilizados neste trabalho foram obtidos em inventário florestal realizado numa área de floresta natural, localizada no município de Cássia, região sul do Estado de Minas Gerais, situada entre: Latitude 20°20' e 20°40' Sul e Longitude 46°40' e 47°00' Oeste. A área de estudo é formada por uma propriedade denominada como Fazenda Reata, possuindo cerca de 90 hectares de floresta natural.

A respeito das características geomorfológicas do local, obtiveram-se as seguintes informações no relatório RADAM-BRASIL (1978).

A região estudada é caracterizada pelo domínio morfoestrutural remanescente de Cadeias Dobradas, apresentando vestígios destas estruturas, com exposições eventuais de seus embasamentos. A área em questão está situada na Região dos Planaltos do Alto do Rio Grande, com altitudes médias em torno de 680m.

O solo da região onde os dados foram coletados tem as seguintes características:

Ocorrem solos classificados como Latossolo Vermelho-Escuro Distrófico, caracterizando solos minerais, não hidromórficos com horizontes B latossólico, diferenciado dos solos da classe Latossolo Vermelho-Amarelo por apresentarem teores mais elevados de Fe_2O_3 e, conseqüentemente, cores mais avermelhadas. Para os solos de textura mais argilosa e/ou muito argilosa, as percentagens de Fe_2O_3 no horizonte B variam de 8,8 a 16,6% e, para solos de textura média, estes valores variam de 3,9 a 8,5%. O horizonte B apresenta cores vermelho-escuras e vermelho-escuro-acinzentadas, com matiz 2,5 YR ou mais vermelho, valores iguais ou menores que 4 e cromas iguais ou inferiores a 6. Possui valor de V% variando de 3 a 28%, com horizonte A moderado e ocorrências de A proeminente, com teores de carbono variando de 0,5 a 2,4%. Os valores Ki, baixos no horizonte B, situam-se em torno de 1,20, e a relação molecular Kr, também baixa, varia de 0,44 a 1,57.

Estes solos geralmente ocorrem em relevos predominantemente ondulados, são profundos, acentuadamente drenados, com pouca diferenciação entre os horizontes. Possuem textura predominantemente argilosa (algumas médias), com relação silte/argila inferior a 0,6 em todos os perfis. Apresentam boas características físicas para o desenvolvimento das raízes e relevo satisfatório à mecanização.

No aspecto do solo, percebe-se a grande quantidade de informações técnicas que, mesmo não sendo necessárias ao estudo da probabilidade, podem evidenciar aspectos da região úteis a pesquisas de dinâmica florestal, silvicultura e outras áreas para os quais o tipo do solo possa ser mais determinante.

O tipo de relevo interfere no crescimento da árvore. Por esta razão acredita-se que é oportuno informar o tipo predominante na região. O relatório RADAM-BRASIL (1978) traz as informações a esse respeito:

Englobam relevos de dissecação estrutural moderada, orientados, constituindo cristas assimétricas e escarpas que coalescem com rampas de colúvio, "mares de morros" e colinas conexas". Possuem incisões de drenagem de 98 a 155m e declividades entre 11 e 24°, com formações superficiais espessas, de textura areno-argilosa, recobertas por pastagens e vegetação secundária.

É também importante definir o clima da região para que se possam fazer associações com outras produções científicas.

A região de Cássia-MG está situada a, aproximadamente, 680 metros de altitude, e apresenta temperaturas médias anuais de 26,5°C (máxima) e 19,5°C (mínima). O clima da região, segundo classificação de Köppen, é do tipo Cwa (Tropical de altitude), apresentando verões rigorosos e chuvosos.

A respeito da vegetação do local estudado tem-se:

A vegetação predominante na região estudada é a Floresta Estacional Semidecidual. O conceito ecológico da região da Floresta Estacional está preso ao clima de duas estações, uma chuvosa e outra seca, que condicionam uma estacionalidade foliar dos elementos arbóreos dominantes, os quais têm adaptação fisiológica à deficiência hídrica ou a baixa temperatura, durante certo tempo.

No caso das Florestas Semidecíduais, a porcentagem de árvores caducifólias no conjunto florestal, e não das espécies que perdem folhas individualmente, deve-se situar em torno de 20 a 50 % na época desfavorável.

Esta região apresenta uma grande distribuição espacial, abrangendo parte do Rio Grande do Sul e do Espírito Santo, os vales dos rios Paraíba e Sapucaí, os planaltos interioranos e parte da bacia do Paraná. Nestas áreas, exceto a bacia do Paraná, os ambientes da Floresta Estacional Semidecidual ocorrem sob um clima tropical estacional com mais de 60 dias secos por ano, e litologia Pré-Cambriana, nas mais variadas formas de relevo, até o Quaternário. A ocorrência predominante de arenitos na bacia do Paraná propiciou a

formação de solos com baixa capacidade de retenção de água e elevado déficit hídrico, que determinaram o caráter estacional do seu clima, sendo portanto, os ambientes desta região fitoecológica, nas áreas em questão, delimitados segundo este critério.

Nestes ambientes, a intensa ação do homem praticamente substituiu a cobertura original por pastagens, agricultura e vegetação secundária, principalmente. Atualmente, ainda se encontram locais de difícil acesso ou protegidos por lei, pequenos agrupamentos remanescentes das Formações Aluviais, das Terras Baixas, Submontana e Montana.

A formação encontrada na área de estudo compreende a Floresta Montana, que apresenta o maior número de agrupamentos remanescentes na área abrangida pelos ambientes da Floresta Estacional Semidecidual. Ela ocorre nas altitudes entre 500 e 1500 metros, revestindo os diques de diabásio da Formação Serra Geral, na bacia do Paraná, e sobre o relevo dissecado do embasamento de litologia variada. Apesar do número elevado de agrupamentos remanescentes, eles são na sua maioria pequenos, sendo os mais representativos encontrados próximos às cidades de Cordeiro-MG, Trajano de Moraes-RJ, Dolores de Turvo-MG, Alvinópolis-MG, Carmo do Meio-MG e Bocaína de Minas-MG. Nestas áreas são freqüentes espécies dos gêneros *Aspidosperma* (perobas), *Piptadenia* (angicos), *Cariniana* (jequitibás), *Ocotea* e *Nectandra* (canelas), e *Lecythis* (sapucaia), que ocupam o estrato dominante da floresta.

Após definidas as características da região, mas faltando ainda informações mais detalhadas sobre o inventário realizado, iniciou-se o desenvolvimento do modelo com uma tabela de dados do Jequitibá-Rosa. Segundo informações oriundas dos profissionais que coletaram os dados, foram feitas outras medições após o inventário que envolve os dados acima citados, com o objetivo de direcionar o estudo para as árvores que seriam de maior interesse para esta e outras pesquisas.

A composição das tabelas apresentadas para o Jequitibá-Rosa em um primeiro momento e para o Jequitibá-Branco posteriormente resulta de vários trabalhos realizados continuamente. Por este motivo, com o intuito de apenas ilustrar a região, foram apresentados dados que permitiram fazer associações com outros conceitos e, com base na literatura, entender o porquê dos dados se apresentarem desta maneira.

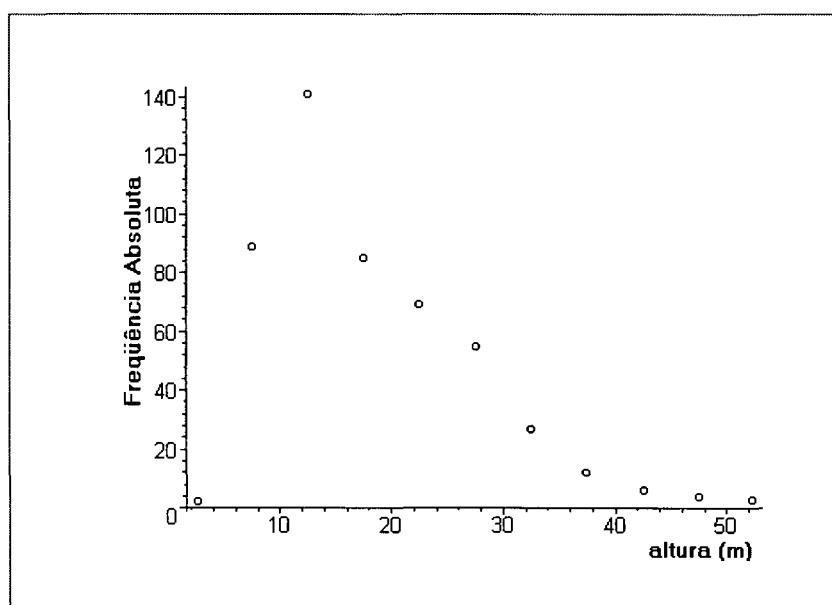
A tabela 3 e a figura 10 mostram os dados que serviram de base para a elaboração do modelo. No capítulo de resultados foram feitos testes com dados do Jequitibá-Branco.

TABELA 3 - FREQUÊNCIA DO JEQUITIBÁ-ROSA

INTERVALO DE CLASSE (m)	PONTO MÉDIO (m)	FREQUÊNCIA			
		Absoluta	Acumulada	Relativa	Relativa Acumulada
0 – 5	2,5	2	2	0,0041	0,0041
5 – 10	7,5	89	91	0,1805	0,1846
10 – 15	12,5	141	232	0,2860	0,4706
15 – 20	17,5	85	317	0,1724	0,6430
20 – 25	22,5	69	386	0,1400	0,7830
25 – 30	27,5	55	441	0,1116	0,8945
30 – 35	32,5	27	468	0,0548	0,9493
35 – 40	37,5	12	480	0,0243	0,9736
40 – 45	42,5	6	486	0,0122	0,9858
45 – 50	47,5	4	490	0,0081	0,9939
50 – 55	52,5	3	493	0,0061	1,0000
		493		1,0000	

Com auxílio do *software* Maple colocaram-se os pontos em um gráfico para que posteriores comparações e cálculos pudessem ser realizados. A figura 10 mostra a distribuição dos dados.

FIGURA 10 - FREQUÊNCIAS DO JEQUITIBÁ-ROSA



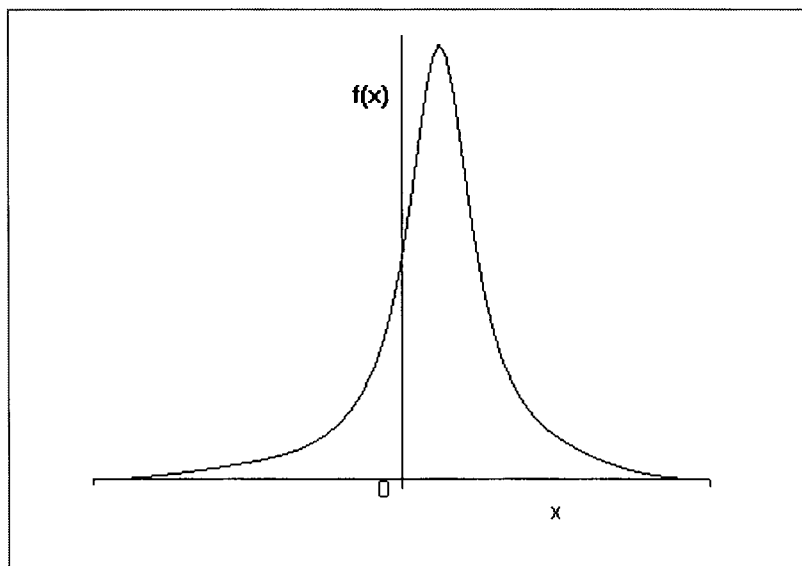
O gráfico de pontos permitiu que fossem feitas tentativas de funções por meio de análises de pontos. Com a função elaborada, foram calculadas as suas integrais em cada classe e medidas como média e variância.

3.3 ESTUDO DE FUNÇÕES COM CURVAS ASSIMÉTRICAS

No início do trabalho realizou-se um estudo de funções que produzissem curvas assimétricas. Em diversas obras de cálculo diferencial e integral observou-se que havia funções que se assemelhavam ao gráfico de pontos do Jequitibá-Rosa. Uma função bastante encontrada na literatura é $f(x) = \frac{1}{x^2 - 2x + 2}$ que começou a ser estudada por serem suas características aparentemente semelhantes ao que estava sendo procurado. A figura 11 mostra o gráfico, elaborado no *software* maple.

```
> plot(1/(x^2-2*x+2), x=-infinity..infinity, color=black).
```

FIGURA 11 - GRÁFICO DA FUNÇÃO $f(x) = 1 / (x^2 - 2x + 2)$



Para que uma expressão matemática atenda às exigências de uma função densidade de probabilidade (FDP), a sua integral no campo dos reais deve ser 1. Portanto o primeiro requisito exigido foi que a integral da função fosse convergente. Com o valor da integral imprópria os cálculos foram efetuados para que a função fosse transformada em outra cuja área fosse unitária.

Com o Maple , utilizando o comando “int” que calcula a integral, obteve-se:

```
> Int(1/(x^2-2*x+2), x=-infinity..infinity);
```

$$\int_{-\infty}^{\infty} \frac{1}{x^2 - 2x + 2} dx$$

> int(1/(x^2-2*x+2), x=-infinity..infinity);

= π

Multiplicando o valor da integral pelo seu inverso obteve-se o valor 1.

> (1/Pi)*int(1/(x^2-2*x+2), x=-infinity..infinity);

=1

$$\text{Assim, } \frac{1}{\pi} \cdot \int_{-\infty}^{\infty} \frac{1}{x^2 - 2x + 2} dx = 1$$

Existia uma função cujo gráfico era semelhante ao dos pontos do Jequitibá-Rosa e, o que era mais importante, havia possibilidade de fazer a integral convergir para 1. Poderiam também ser feitas generalizações sobre o modelo. A função tomaria a forma genérica

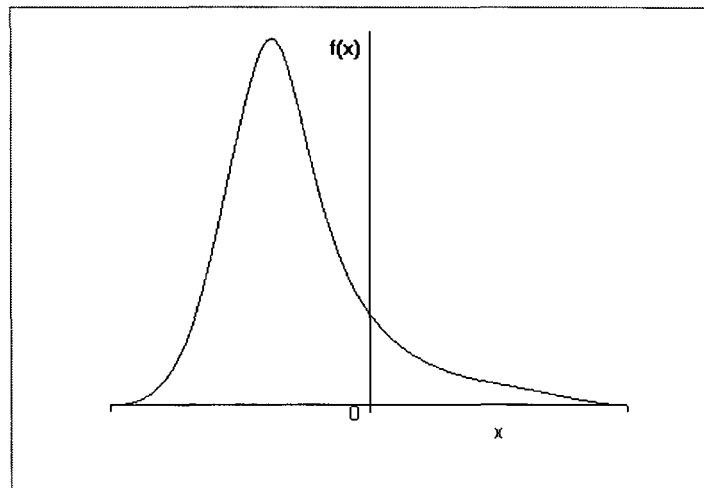
$$f(x) = \frac{a}{bx^2 + cx + d} \quad (29)$$

A partir da função elaborada com parâmetros foram feitos dezenas de gráficos para verificar qual a influência que os valores a, b, c e d exerciam sobre a sua forma.

A figura 12 mostra um gráfico, utilizando o comando “plot” do maple, no qual um dos parâmetros foi alterado.

> plot(1/(x^2+6*x+12), x=-infinity..infinity, color=black)

FIGURA 12 - GRÁFICO DE $f(x) = 1 / (x^2 + 6x + 12)$

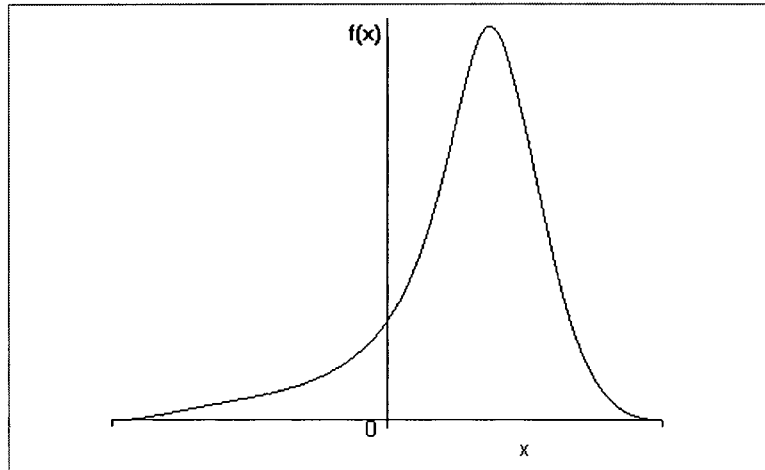


O fato de substituir o valor de c por 6 provocou alterações na assimetria.

A figura 13 mostra a forma da função quando se substitui o parâmetro c por -6 .

```
> plot(1/(x^2-6*x+12), x=-infinity..infinity, color=black)
```

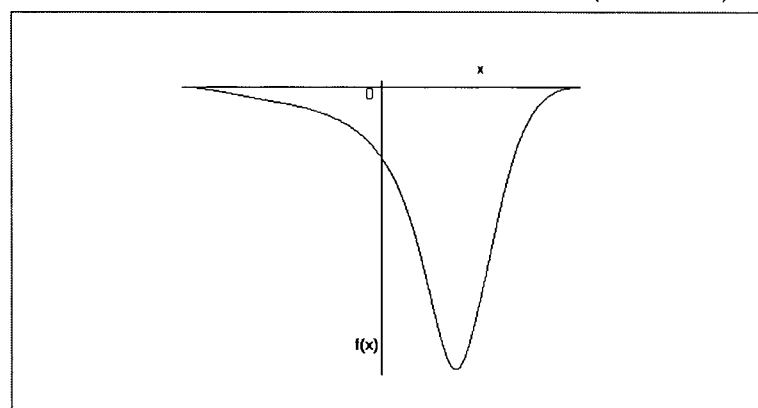
FIGURA 13 - GRÁFICO DE $f(x) = 1 / (x^2 - 6x + 12)$



A função que se começou a estudar teve de ser abandonada porque, após algumas observações, verificou-se que os parâmetros estariam sujeitos a muitas restrições. Este fato poderia comprometer a função, uma vez que ela seria proposta para atender a vários conjuntos de dados de diversas grandezas. Observou-se que a forma ficava determinada pela combinação de parâmetros. A figura 14 mostra um caso em que o parâmetro a (numerador) é negativo.

```
> plot(-1/(x^2-6*x+12), x=-infinity..infinity, color=black)
```

FIGURA 14 - GRÁFICO DA FUNÇÃO $f(x) = -1 / (x^2 - 2x + 2)$



Como uma função densidade de probabilidade deve ter seus valores sempre positivos, esta fórmula tornou-se muito restrita. Além disso, observou-se que o gráfico estava sujeito ao resultado do discriminante $\Delta = b^2 - 4ac$ no denominador. Se o discriminante fosse positivo a forma obtida não seria de nenhum interesse para a pesquisa.

Outras dezenas de funções foram submetidas a teste, mas muitos fatores deveriam ser considerados. Era necessário que o gráfico fosse assimétrico e que os valores de $f(x)$ sempre fossem positivos. Além disso, as integrais impróprias deveriam ser convergentes para atender às condições de uma função densidade de probabilidade definidas em 2.2.2 e para o cálculo da média e da variância como pode-se verificar nas fórmulas 3 e 4, respectivamente.

Os fatores citados eliminaram uma série de alternativas, pois o objetivo era atender a todos os requisitos de uma FDP e ainda dar uma contribuição significativa para o trabalho florestal, produzindo algo prático de ser trabalhado com a tecnologia de que se dispõe hoje.

Funções com gráficos assimétricos foram objeto de estudo por mais de dois anos. Muitas eram as formas, mas as análises de viabilidade conduziam à eliminação da função. Algumas chegaram a ser estudadas durante meses pelo fato de atenderem aos requisitos iniciais acima citados. Entretanto, foram descartadas porque não se encontrou uma forma prática para estabelecer seus parâmetros.

Na tentativa de estimar parâmetros procuraram-se métodos elementares da matemática, como sistema de equações. Também foram testados os estimadores de Zanakis, de Máxima Verossimilhança, o Método dos Momentos, o Método dos Percentis, ajuste por regressão e outros procedimentos.

Depois de diversas tentativas pôde-se constatar que para elaborar uma função densidade de probabilidade muitos fatores deveriam ser considerados para atender a todos os requisitos. Além disso, era necessário formular uma expressão que pudesse ser utilizada com qualquer conjunto de dados, principalmente da área florestal.

3.4 ELABORAÇÃO DA DISTRIBUIÇÃO POLINOMIAL

O modelo da distribuição polinomial 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.

Este modelo baseia-se na idéia de que a expressão matemática que define a função densidade de probabilidade pode ser elaborada a partir dos dados que se quer analisar.

3.4.1 Caracterização e Processo de Elaboração

Pode-se definir genericamente a Função Polinomial por:

$$f(x) = \frac{1}{k} \begin{cases} c_1 x^d & \text{se } 0 < x < \ell_1 \\ a_1 x^n + a_2 x^{n-1} + a_3 x^{n-2} + \dots + a_m, & \text{se } \ell_1 \leq x \leq \ell_2 \\ \frac{c_2}{x^h} & \text{se } x > \ell_2 \\ 0 & \text{e.o.c} \end{cases} \quad (30)$$

onde n, d e h são inteiros positivos;

$a_1, a_2, a_3, \dots, a_m$ são números reais;

c_1 e c_2 são números reais;

k é o valor da integral $\int_0^{\infty} [c_1 x^d + (a_1 x^n + a_2 x^{n-1} + a_3 x^{n-2} + \dots + a_m) + \frac{c_2}{x^h}] dx$;

x , variável;

ℓ_1 : limite superior da classe onde será ajustada a função $c_1 x^d$;

ℓ_2 : limite superior da última classe onde o polinômio produz um bom ajuste;

Conforme o exposto, a função polinomial é definida por três sentenças ou fórmulas, as quais devem atender aos requisitos de uma FDP, ou seja, deve ser contínua, com valores funcionais não negativos e convergentes em $(0, +\infty)$. Para elaborar a função polinomial alguns passos são importantes:

- **Primeira etapa:**

Inicialmente deve-se fazer um ajuste do polinômio que mais se aproxima dos dados. Este ajuste pode ser feito com auxílio do computador utilizando o *software* MATLAB. O polinômio é identificado como $g_2(x)$.

- **Segunda etapa:**

Deve ser feito um gráfico do polinômio produzido e desconsiderar as classes em que o polinômio assume valores negativos ou contraria a tendência dos dados observados.

- **Terceira etapa:**

Elaborar funções para as classes em que o polinômio não se ajusta aos dados. Normalmente essas classes são a(s) primeira(s) e a(s) última(s).

Para a primeira classe a função a ser ajustada é da forma $g_1(x) = c_1 x^d$. Inicialmente deve-se calcular a ordenada do valor ℓ_1 . Em seguida, descobrir qual deve ser o valor de c_1 para que a ordenada de ℓ_1 calculada pela função $g_1(x) = c_1 x^d$ tenha o mesmo valor que a ordenada de ℓ_1 calculada pelo polinômio. Este cuidado é indispensável para que a função seja contínua em ℓ_1 .

Para as últimas classes, é necessária uma função que, além de atender a tendência dos dados deve ser convergente no infinito; isto pode ser obtido

com uma função do tipo $g_3(x) = \frac{c_2}{x^h}$ em que c_2 é um número real. Como

na função anterior, deve-se obter o valor da ordenada de ℓ_2 no polinômio

e em seguida calcular c_2 de forma que a função $g_3(x) = \frac{c_2}{x^h}$ tenha o

mesmo valor da ordenada de ℓ_2 calculada pelo polinômio.

- **Quarta etapa:** Deve-se formar a função $g(x)$ com as três funções mencionadas $g_1(x) = c_1 x^d$, $g_2(x)$ (polinômio ajustado) e $g_3(x) = \frac{c_2}{x^h}$ e calcular $\int_0^{\infty} g(x) dx$ obtendo o valor de k , ou seja: a integral imprópria da função de 0 até infinito deve convergir para um número que será representado por k .
- **Quinta etapa:** Multiplicando a função $g(x)$ por $\frac{1}{k}$ obtém-se uma função $f(x)$ em que $\int_0^{\infty} f(x) dx = 1$. Portanto tem-se uma função contínua, com valores positivos e com integral convergente para 1, atendendo às exigências de uma função densidade de probabilidade.

Para exemplificar o método descrito, foi feita uma aplicação dos dados do Jequitibá-Rosa, seguindo-se as etapas descritas com o objetivo de reforçar a seqüência.

3.4.2 Aplicação para o Caso do Jequitibá-Rosa

Em estudo prévio demonstrado anteriormente verificou-se que os dados de altura total do Jequitibá-Rosa, tabela 3, têm assimetria positiva e que nas últimas classes seus valores vão diminuindo.

Para que o leitor tenha uma idéia mais clara das etapas já descritas, o exemplo seguirá rigorosamente os passos citados.

É importante salientar que nos trechos de programas apresentados Maple ou Matlab a notação para números segue o modelo do software. O ponto no software representa vírgula no sistema de numeração utilizado.

- **Primeira Etapa:**

Os dados analisados estão contidos na tabela 3.

Partindo dos pontos médios de classe e das freqüências absolutas, será feito o ajuste utilizando-se uma rotina do software matlab.

```

>> x=[2.5 7.5 12.5 17.5 22.5 27.5 32.5 37.5 42.5 47.5 52.5];
>> y=[2 89 141 85 69 55 27 12 6 4 3];
>> n=4;
>> p=polyfit(x, y, n)

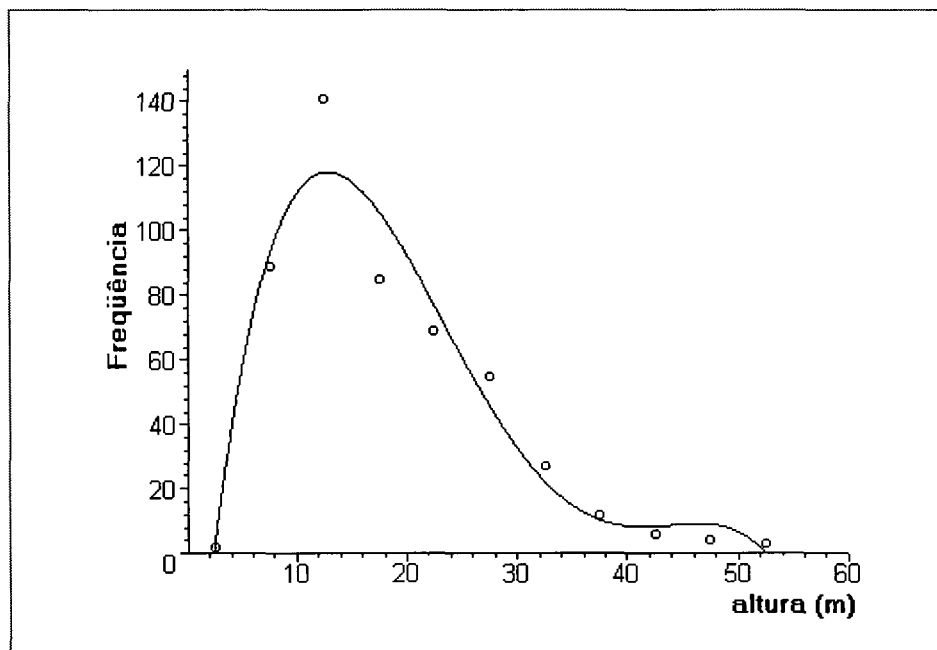
p = -0.00036969696970 0.04964413364413 -2.25747668997667 36.60301087801059
-74.59311261654885

```

Nas duas primeiras linhas entrou-se com os dados. Depois escolheu-se o grau do polinômio, no caso, o quarto, cujos coeficientes surgiram com o comando polyfit. O matlab oferece opções para fornecer os coeficientes com maior ou menor número de casas decimais.

A figura 15 mostra o gráfico do polinômio gerado pelo maple juntamente com os pontos referentes aos dados do Jequitibá-Rosa. Deve-se observar que, com exceção do ponto mais alto, todos os outros estão próximos do modelo.

FIGURA 15 - POLINÔMIO AJUSTADO PARA OS DADOS DO JEQUITIBÁ-ROSA JUNTAMENTE COM VALORES OBSERVADOS



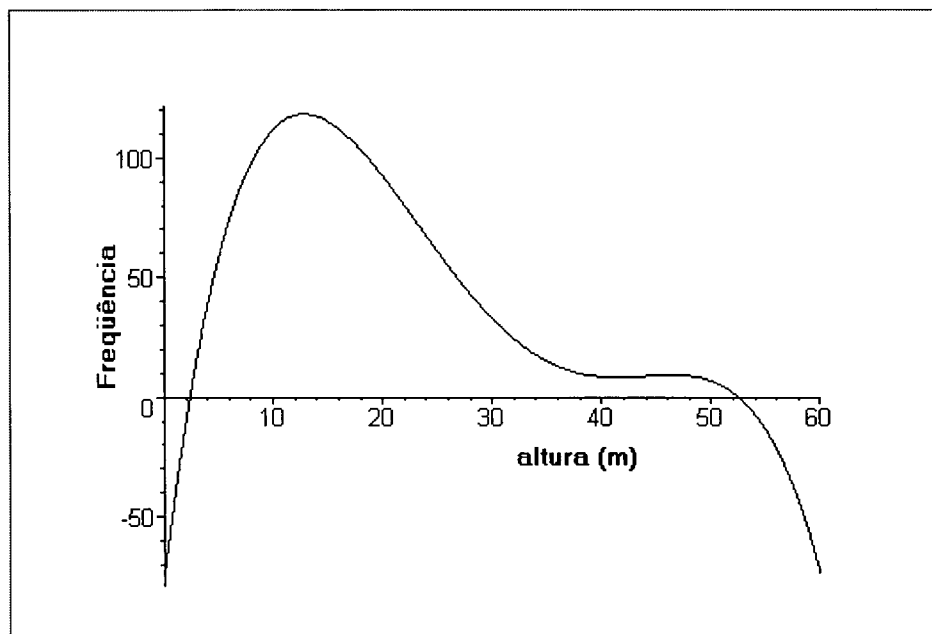
- **Segunda etapa:**

Nesta etapa os dados foram transferidos para o *software* Maple. Este *software* permite uma maior facilidade de interação com o usuário, principalmente quando for necessário unir as funções elaboradas nas etapas seguintes.

O comando “plot” no Maple gera o gráfico do polinômio no intervalo que for escolhido. Neste caso solicitou-se de $x=0$ até $x=60$ e o resultado está na figura 16.

```
>plot(-0.00036969696970*x^4+0.04964413364413*x^3-2.25747668997667*x^2+
36.60301087801059*x-74.59311261654885, x=0..60);
```

FIGURA 16 - GRÁFICO DO POLINÔMIO AJUSTADO MOSTRANDO AS RAÍZES E O INTERVALO ONDE É NEGATIVO



Portanto, tem-se o gráfico do polinômio $P(x) = -0.00036969696970 x^4 + 0.04964413364413 x^3 - 2.25747668997667 x^2 + 36.60301087801059 x - 74.59311261654885$.

Como pode-se perceber na figura 16, o polinômio tem duas raízes e partes onde as ordenadas são negativas.

Para descobrir as raízes do polinômio e ter certeza aos procedimentos seguintes, pode-se utilizar o comando solve (MAPLE), que fornece raízes reais e complexas do polinômio. Desta forma não se corre o risco de tomar iniciativas a partir de simples observação do gráfico:

```
> solve (-0.00036969696970*x^4+0.04964413364413*x^
3-2.25747668997667*x^2+ 36.60301087801059*x-74.59311261654885, x);
```

E se obtêm as raízes do polinômio

```
2,365315699; 39,70899831-6,929093004i; 39,70899831+6,929093004i ; 52,49999999
```

Verifica-se portanto que as raízes reais do polinômio (interceptos x) são $x_1 = 2,365315699$ e $x_2 = 52,49999999$. Os valores seguidos da letra I representam raízes imaginárias, que não podem ser representadas no gráfico, tendo em vista que os eixos são reais.

De posse dessas informações podem ser adotados certos procedimentos. O polinômio não serve para a primeira classe que tem limites 0 e 5, uma vez que antes da primeira raiz parte de seus valores são negativos, o que contraria a definição da função densidade de probabilidade. As duas últimas classes também não podem ser explicadas pelo polinômio, visto que este começa a ter uma característica crescente e os dados são sempre decrescentes.

Assim foram suprimidas a primeira classe, que compreende o intervalo (0,5] e as classes referentes aos intervalos (40,45] e (50,55]. Para essas classes foram elaboradas outras funções a partir do polinômio.

- **Terceira etapa:**

Inicialmente foi formulada uma função para a classe inicial (0,5]. Com base em estudos anteriores, optou-se por escolher uma função do tipo $g_1(x) = c_1 x^5$. Inicialmente calculou-se a ordenada do limite superior da classe ℓ_1 . No maple isso é possível utilizando o comando subs que retorna o valor da ordenada dada à abscissa.

```
> subs(x=5, -0.00036969696970*x^4+0.04964413364413*x^
x^2+ 36.60301087801059*x-74.59311261654885);
```

- Resultado

57,95948058

Como foi possível verificar para o valor $x=5$ o polinômio assume o valor $P(5)=57,95948058$. Portanto, deve-se calcular o valor de c_1 de forma que o ponto $(5; 57,95948058)$ pertença à curva da forma $g_1(x) = c_1x^5$, e resolver a equação $57,95948058 = c_1 5^5$. Ou $c_1 = \frac{57,95948058}{3125}$, que pode ser feito no maple com o comando "evalf".

```
> evalf(57.95948059/3125);
```

- Resultado

0,01854703379

A função para a primeira classe pode ser $g_1(x) = 0,01854703379x^5$. Para verificar se a função ficou razoável pode se verificar se o ponto $(2,5; 2)$ está próximo do modelo.

```
> subs(x=2.5, 0.01854703379*x^5);
```

- Resultado

1,811233769

Como 1,811233769 está próximo do valor observado, pode-se admitir $g_1(x) = 0,01854703379x^5$ como a função para a primeira classe.

Para elaborar a função g_3 inicialmente calcula-se a ordenada do limite ℓ_2 que no caso é 40, substituindo-a no polinômio ajustado.

```
> subs(x=40, -0.00036969696970*x^4+0.04964413364413*x^3-2.25747668997667*x^2+ 36.60301087801059*x-4.59311261654885);
```

- Resultado

8,36492938

O valor encontrado sinaliza que o ponto (40; 8, 36492938) deve pertencer à função $g_3(x) = \frac{c_2}{x^6}$. Então, $8, 36492938 = \frac{c_2}{40^6}$. Fazendo o cálculo para encontrar c_2 , obtém-se $c_2 = 3426275074$ e a função resulta em $g_3(x) = \frac{3426275074}{x^6}$.

> evalf(8.36492938*40^6);

▪ Resultado

0,342627507410¹¹

Para verificar se a função se aproxima dos dados observados, podem-se fazer alguns cálculos:

> subs(x=40.0, 34262750740/x^6);

▪ Resultado

8,364929380

> subs(x=42.5, 34262750740/x^6);

▪ Resultado

5,814182323

> subs(x=47.5, 34262750740/x^6);

▪ Resultado

2,983050249

> subs(x=52.5, 34262750740/x^6);

▪ Resultado

1,636313096

Para o valor observado 42,5 a frequência é 6, para o valor 47,5 tem-se 4 e para o valor 52,5 o valor é 3, o que significa que a função segue a tendência decrescente dos dados.

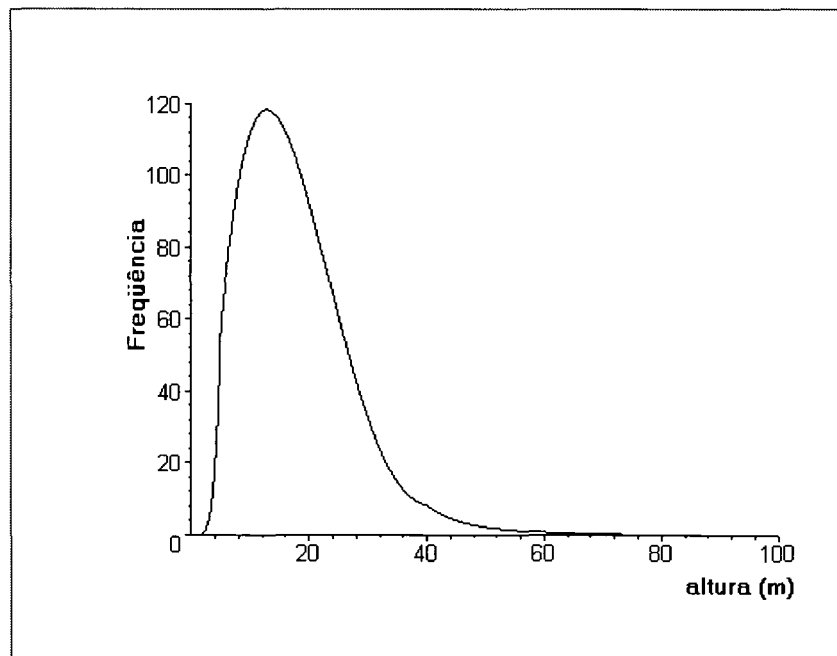
- **Quarta etapa:**

Nesta etapa será formada a função $g(x)$ a partir das três funções $g_1(x)$, $g_2(x)$ e $g_3(x)$ estimadas. Posteriormente será feito o cálculo da integral $\int_0^{\infty} g(x)dx$ e obtido o valor de k . Para unir as funções no maple será utilizado o comando `piecewise`, que permite escrever funções definidas por mais de uma sentença.

```
f:=x->piecewise(x>=0 and x<5, 0.01854703379*x^5, x<=40 and x>=5,
-0.00036969696970*x^4+0.04964413364413*x^3-2.25747668997667*x^2+
36.60301087801059*x-74.59311261654885, x>=40, 34262750740/x^6);
```

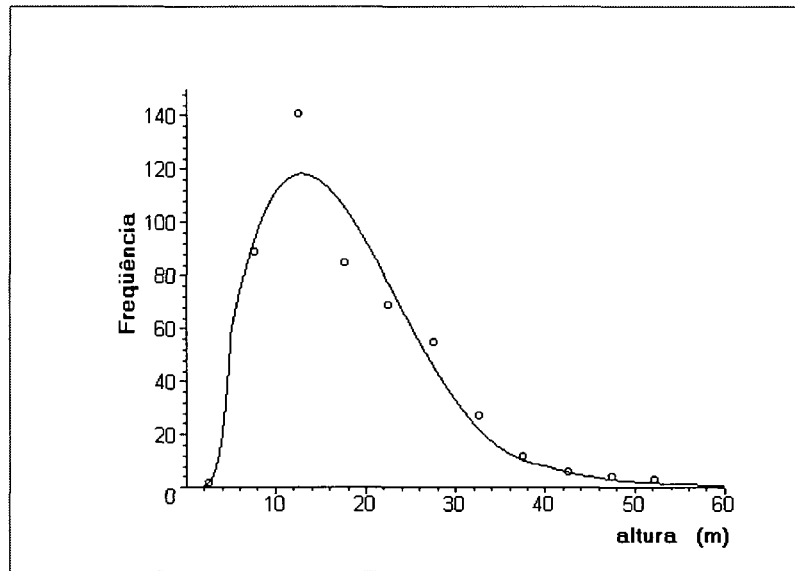
A figura 17 mostra o gráfico de $g(x)$.

FIGURA 17 - GRÁFICO DE $g(x)$, FUNÇÃO DEFINIDA POR TRÊS SENTENÇAS



Observe-se que os valores funcionais de $g(x)$ reproduzem a tendência dos dados observados. Na figura 18 pode-se observar o gráfico de $g(x)$ juntamente com os valores observados. Quase todos os pontos estão próximos da curva.

FIGURA 18 - GRÁFICO DE $g(x)$ COM OS VALORES OBSERVADOS DO JEQUITIBÁ-ROSA



Nesta etapa é conveniente verificar os números que o polinômio está fornecendo para os valores observados. Isto pode ser feito no maple pelo comando subs.

```
> subs(x=2.5, 0.01854703379*x^5);
```

Observado: 2

Calculado

1,811233769

```
>subs(x=7.5, 0-0.00036969696970*x^4+0.04964413364413*x^3-2.25747668997667*x^2+36.60301087801059*x-74.59311261654885);
```

Observado: 89

Calculado

92,72027917

```
>subs(x=12.5, 0-0.00036969696970*x^4+0.04964413364413*x^3-2.25747668997667*x^2+36.60301087801059*x-74.59311261654885);
```

Observado: 141

Calculado

118,1491842

> subs(x=17.5, 0-0.00036969696970*x^4+0.04964413364413*x^3-2.25747668997667*x^2+36.60301087801059*x-74.59311261654885);

Observado: 85

Calculado

105,9953380

> subs(x=22.5, 0-0.00036969696970*x^4+0.04964413364413*x^3-2.25747668997667*x^2+36.60301087801059*x-74.59311261654885);

Observado: 69

Calculado

76,8554781

> subs(x=27.5, 0-0.00036969696970*x^4+0.04964413364413*x^3-2.25747668997667*x^2+36.60301087801059*x-74.59311261654885);

Observado: 55

Calculado

45,7808856

> subs(x=32.5, 0-0.00036969696970*x^4+0.04964413364413*x^3-2.25747668997667*x^2+36.60301087801059*x-74.59311261654885);

Observado: 27

Calculado

22,277390

> subs(x=37.5, 0-0.00036969696970*x^4+0.04964413364413*x^3-2.25747668997667*x^2+36.60301087801059*x-74.59311261654885);

Observado: 12

Calculado

10,305362

> subs(x=42.5, 34262750740/x^6);

Observado: 6

Calculado

5,814182323

> subs(x=47.5, 34262750740/x^6);

Observado: 4

Calculado

2,98

> subs(x=52.5, 34262750740/x^6);

Observado: 3

Calculado

1,636313096

Na seqüência deve-se calcular o valor de $\int_0^{\infty} g(x)dx$ para obter o valor de k

como está previsto na quarta etapa. No maple foram calculadas cada integral separadamente e depois foi feita a soma.

Primeiro o cálculo da integral para a primeira classe [0;5]

> int(0.01854703379*x^5, x=0..5);

▪ Resultado

48,29956716

A integral do polinômio sera feita de 5 até 40.

> int(-0.00036969696970*x^4+0.04964413364413*x^3-2.25747668997667*x^2+36.60301087801059*x-74.59311261654885, x=5..40);

2341,996692

A última integral é imprópria, ou seja, seu limite superior é infinito. Por essa razão foi estudada uma função que fosse convergente no infinito.

> int(34262750740.0/x^6, x=40..infinity);

66,91943504

De posse dos valores das três parciais pode-se obter o valor da integral que representa a área entre o eixo x e a curva de zero até infinito.

> evalf(48.29956716+2341.996692+66.91943504);

2457,215694

evalf((1/2457.215694)*(48.29956716+2341.996692+66.91943504));

1,000000000

Observa-se que $k = 2457,215694$ e que o inverso deste número faz com que o resultado da integral seja igual a 1.

- **Quinta etapa:**

A função densidade de probabilidade $f(x)$ foi formada multiplicando-se

$g(x)$ pelo resultado da integral $\int_0^{\infty} g(x)dx$, que é o valor de k .

Então pode-se escrever a função densidade de probabilidade para o Jequitibá-Rosa da seguinte maneira:

$$f(x) = \frac{1}{2457,215694} \begin{cases} 0,01854703379x^5 & \text{se } 0 \leq x < 5 \\ -0,00036969696970x^4 + 0,04964413364413x^3 - 2,25747668997667x^2 + \\ 36,60301087801059x - 74,59311261654885 & \text{se } 5 \leq x \leq 40 \\ \frac{34262750740}{x^6} & \text{se } x > 40 \\ 0 & \text{e.o.c.} \end{cases}$$

4 RESULTADOS

Nesta seção do trabalho foram feitos vários testes do modelo Polinomial e comparações com outras distribuições. Posteriormente foi feita uma análise dos resultados obtidos.

4.1 CÁLCULOS COM A FDP ELABORADA PARA O JEQUITIBÁ-ROSA

A expressão gerada na seção anterior foi submetida neste momento a testes e comparações. A FDP para o Jequitibá-Rosa resultou em:

$$f(x) = \frac{1}{2457,215694} \begin{cases} 0,01854703379x^5 & \text{se } 0 \leq x < 5 \\ -0,00036969696970x^4 + 0,04964413364413x^3 - 2,25747668997667x^2 + 36,60301087801059x - 74,59311261654885 & \text{se } 5 \leq x \leq 40 \\ \frac{34262750740}{x^6} & \text{se } x > 40 \\ 0 & \text{e.o.c.} \end{cases}$$

Podem ser feitos vários cálculos com a distribuição encontrada. A frequência esperada pode ser obtida da seguinte maneira: calcula-se a integral em cada intervalo e multiplica-se o resultado por 493, que é o valor observado.

$$> 493*(1/2457.215694)*\text{int}(0.01854703379*x^5, x=0..5);$$

9,690515434

$$> 493*(1/2457.215694)*\text{int}(-0.00036969696970*x^4+0.04964413364413*x^3-$$

$$2.25747668997667*x^2+36.60301087801059*x-74.59311261654885, x=5..40);$$

469,8831982

$$> 493*(1/2457.215694)*\text{int}(34262750740.0/x^6, x=40..infinity);$$

13,42628633

Somando-se os três valores pode-se encontrar:

```
> evalf(9.690515434+469.8831982+13.42628633);
```

```
492,9999999
```

Pode-se verificar que o valor esperado ficou muito próximo do observado.

4.1.2 Cálculo da Média pela Distribuição Polinomial

A media da distribuição deve ser calculada por $E(X) = \int_0^{\infty} xf(x)dx$.

Aplicando este conceito na função obtida para o Jequitibá-Rosa, obtiveram-se os seguintes resultados.

```
> int((1/2457.215694)*0.01854703379*x^6, x=0..5);
```

```
0,08424093395
```

```
> (1/2457.215694)*int(-0.00036969696970*x^5+0.04964413364413*x^4-
2.25747668997667*x^3+36.60301087801059*x^2-74.59311261654885*x, x=5..40);
```

```
16,46850697
```

```
> (1/2457.215694)*int(34262750740.0/x^5, x=40..infinity);
```

```
1,361692325
```

```
> evalf(.8424093395e-1+16.46850697+1.361692325);
```

```
17,91444022
```

Portanto a média calculada pela função densidade desenvolvida é 17,91444022 metros, ficando este valor muito próximo da média calculada pela fórmula convencional, que é 18 metros.

4.1.3 Cálculo da Variância pela Distribuição Polinomial

Para calcular a variância pela função densidade de probabilidade deve-se utilizar a expressão $Var(x) = \int_{-\infty}^{\infty} x^2 f(x)dx - [E(x)]^2$.

Inicialmente foi feito o cálculo da integral:

```
> int((1/2457.215694)*0.01854703379*x^7, x=0..5);
```

```
0,3685540860
```

```
> (1/2457.215694)*int(-0.00036969696970*x^6+0.04964413364413*x^5-  
2.25747668997667*x^4+36.60301087801059*x^3-74.59311261654885*x^2, x=5..40);
```

```
340,8462676
```

```
> (1/2457.215694)*int(34262750740.0/x^4, x=40..infinity);
```

```
72,62359070
```

Que somando resultou em:

```
> evalf(.3685540860+340.8462676+72.62359070);
```

```
413,8384124
```

Posteriormente foi feito o cálculo da variância,

```
> evalf(413.8384124-(17.91444022)^2);
```

```
92,9112440
```

Portanto a variância calculada pela FDP resultou em 92,9112440 m², valor próximo da variância calculada a partir dos dados observados 88,20996 m², o que mostra indícios de que o modelo é aderente aos dados.

4.2 MODELO ELABORADO PARA O JEQUITIBÁ-BRANCO

Para o Jequitibá-Branco (*Cariniana estrellensis*) o processo de elaboração do modelo polinomial seguiu as mesmas etapas.

No matlab o ajuste foi feito com um polinômio de 4º grau.

```
x = [2.5 7.5 12.5 17.5 22.5 27.5 32.5 37.5];
```

```
y = [1 22 63 39 20 14 6 2];
```

```
n = 4;
```

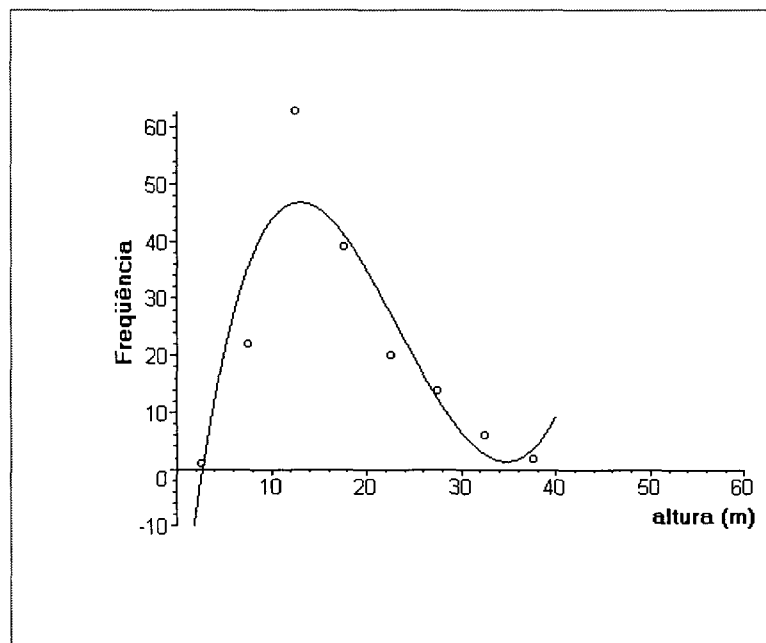
```
p = polyfit(x, y, n)
```

O comando polyfit forneceu os coeficientes do polinômio.

$p = -0.0000651515151515 \ 0.0150505050505051 \ -0.83941287878789 \ 14.75658369408377 \ -34.32092972132051$

Observa-se que os dados do Jequitibá-Branco também estão próximos do polinômio ajustado, com exceção do ponto mais alto. Este fato pode ser observado na figura 19.

FIGURA 19 - GRÁFICO DO POLINÔMIO GERADO PARA O JEQUITIBÁ-BRANCO JUNTAMENTE COM OS VALORES OBSERVADOS

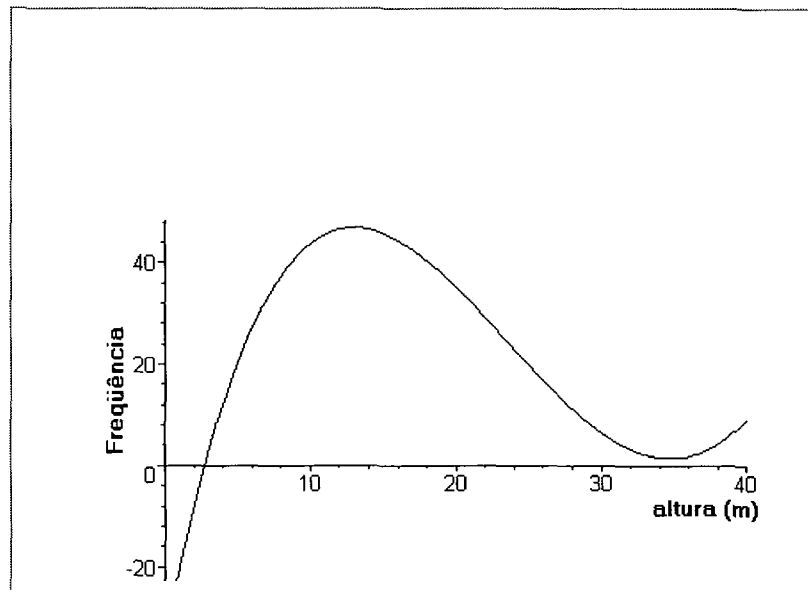


Depois o polinômio foi transferido para o maple para análise do gráfico e elaboração da FDP correspondente.

```
> plot(-0.0000651515151515*x^4+0.0150505050505051*x^3-0.83941287878789*x^2+
14.75658369408377*x-34.32092972132051, x=0..40);
```

A figura 20 mostra o gráfico do polinômio no software maple.

FIGURA 20 - GRÁFICO DO POLINÔMIO GERADO PARA O JEQUITIBÁ-BRANCO COM OS INTERVALOS ONDE DEVERÃO SER INSERIDAS NOVAS FUNÇÕES



Dos dados entendeu-se que deviam ser formuladas novas funções para a primeira e para as últimas classes.

As raízes do polinômio indicaram onde deve haver a intervenção.

```
> solve(-0.00006515151515*x^4+0.01505050505051*x^3-0.839412878789*x^2+
14.75658369408377*x-34.32092972132051, x);
```

- Raízes

```
2.728944492, 34.80249037 - 2.312887065 I, 34.80249037 + 2.312887065 I, 158.6738267
```

```
> subs(x=5, -0.00006515151515*x^4+0.01505050505051*x^3-0.839412878789*x^2+
14.75658369408377*x-34.32092972132051);
```

```
20,31726019
```

```
> subs(x=30, -0.00006515151515*x^4+0.01505050505051*x^3-0.839412878789*x^2+
14.75658369408377*x-34.32092972132051);
```

```
6,49589918
```

Ajuste da Função para a Primeira Classe

```
> evalf(20.31726019/(5^5));
```

```
0,006501523261
```

```
> subs(x=5, 0.0065015232612*x^5);
```

```
20,31726019
```

```
> subs(x=2.5, 0.006501523261*x^5);
```

```
0,6349143810
```

Ajuste da Função para as Últimas Classes

```
>subs(x=30, -0.00006515151515*x^4+0.01505050505051*x^3-0.83941287878789*x^2+
14.75658369408377*x-34.32092972132051);
```

```
> evalf(6.49589918*(30^6));
```

```
0,47355105021010
```

```
> subs(x=30, 4735510502/(x^6));
```

```
2367755251
364500000
```

```
> 6.495899180;
```

```
> subs(x=32.5, 4735510502/x^6);
```

```
4,018524664
```

```
> subs(x=37.5, 4735510502/(x^6));
```

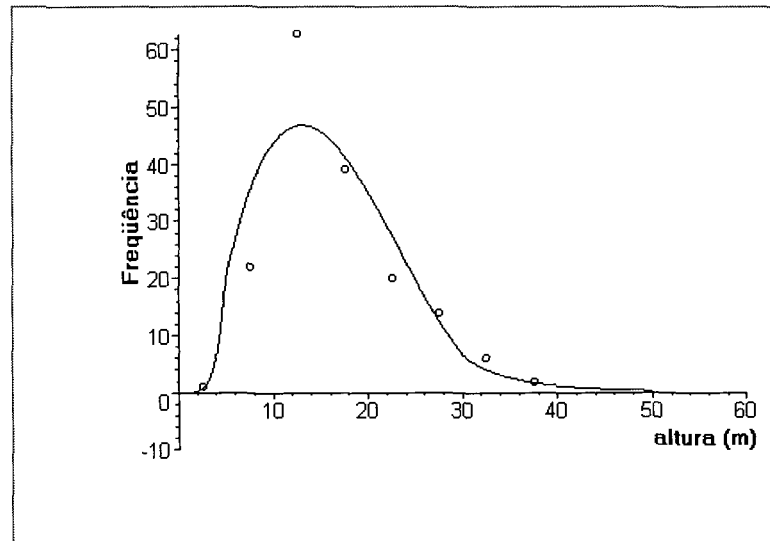
```
1,702860994
```

Preparação da Função que será Transformada em FDP

```
> f:=x->piecewise(x>=0 and x<5, 0.0065015232612*x^5, x<=30 and x>=5, -
0.0000651515151515*x^4+0.01505050505051*x^3-0.83941287878789*x^2+
14.75658369408377*x-34.32092972132051, x>=30, 4735510502/x^6);
f:= x -> piecewise( 0 ≤ x and x < 5, .0065015232612 x5, x ≤ 30 and 5 ≤ x,
-.00006515151515 x4 + .01505050505051 x3 - .83941287878789 x2
+ 14.75658369408377 x - 34.32092972132051, 30 ≤ x, 4735510502  $\frac{1}{x^6}$  )
plot(f, 0..100, discontin=true);
```

A figura 21 mostra o gráfico de $g(x)$ junto com os pontos.

FIGURA 21 - GRÁFICO DA FUNÇÃO $g(x)$ DEFINIDA PARA O JEQUITIBÁ-BRANCO



Cálculo da Integral

```
> int(0.0065015232612*x^5, x=0..5);
```

16,93105016

```
> int(-0.0000651515151515*x^4+0.01505050505051*x^3-0.83941287878789*x^2+
14.75658369408377*x-34.32092972132051, x = 5..30);
```

807,0217380

```
> int(0.4735510502/x^6, x=30..infinity);
```

```
0,389753950810-8
```

```
> evalf((1/823.9527882)*(16.93105016+807.0217380+.3897539508e-8));
```

```
1,000000000
```

4.2.1 Cálculo da Média para o Jequitibá-Branco

```
> (1/823.9527882)*int(0.0065015232612*x^6, x=0..5);
```

```
0,08806529281
```

```
> (1/823.9527882)*int(-0.00006515151515*x^5+0.01505050505051*x^4-
```

```
0.83941287878789*x^3+14.75658369408377*x^2-34.32092972132051*x, x=5..30);
```

```
15,23066355
```

```
> (1/823.9527882)*int(0.4735510502/x^5, x=30..infinity);
```

```
0,177386051310-9
```

```
> evalf(.8806529281e-1+15.23066355+.1773860513e-9);
```

```
15,31872884
```

Portanto a média calculada pela FDP é 15,3187284 m, valor próximo da média calculada com os valores observados, que é 16,42215569 m.

4.2.2 Cálculo da Variância para o Jequitibá-Branco

```
> (1/823.9527882)*int(0.0065015232612*x^7, x=0..5);
```

```
0,3852856560
```

```
> (1/823.9527882)*int(-0.00006515151515*x^6+0.01505050505051*x^5-
```

```
0.83941287878789*x^4+14.75658369408377*x^3-34.32092972132051*x^2, x=5..30);
```

```
272,4284051
```

```
> (1/823.9527882)*int(0.4735510502/x^4, x=30..infinity);
```

```
0,709544205310-8
```

```
> evalf(.3852856560+272.4284051+.7095442053e-8);
```

```
272,8136908
```

```
> eval(272.8136908-(15.31872884)^2);
```

```
38,1502375
```

A variância pela fdp é $38,1502375 m^2$, e pelos dados observados é de $47,32522906 m^2$.

4.3 TESTES COMPARATIVOS DO JEQUITIBÁ-BRANCO E DO JEQUITIBÁ-ROSA

Neste item serão feitas comparações entre a distribuição polinomial e as distribuições Gama, Beta e Weibull dois parâmetros. Além disso, para cada função será aplicado um teste de aderência. A obra de COSTA NETO (1999) traz informações sobre testes de aderência que podem ser utilizados.

Um dos testes de aderência é o do qui-quadrado.

$$\chi_v^2 = \sum_{i=1}^k \frac{(f_{oi} - f_{ei})^2}{f_e} \quad (31)$$

onde:

χ_v^2 = é a estatística de teste, com v graus de liberdade;

f_{oi} = é a freqüência observada de uma determinada classe ou valor da variável;

f_{ei} = é a freqüência esperada, segundo o modelo testado, dessa classe ou valor da variável;

n = número de elementos da amostra;

k = número de classes ou valores considerados.

Pearson mostrou que se o modelo testado for verdadeiro e se as frequências esperadas forem maiores ou iguais a cinco, a expressão anterior terá aproximadamente distribuição qui-quadrado com $\nu = k - m - 1$ graus de liberdade, sendo k o número de classes ou parcelas somadas e m o número de parâmetros estimados a partir da amostra. A subtração de 1 ao valor de k deve-se à existência da restrição $\sum_{i=1}^k f_{oi} = n$ entre frequências observadas.

Caso existam classes que não satisfaçam à condição $f_{ei} \geq 5$, estas deverão ser agrupadas às classes adjacentes, para em seguida se proceder ao cálculo. A condição $f_{ei} \geq 5$ é, em geral, conservadora, podendo-se, em muitos casos, realizar o teste com boa precisão, mesmo com alguma f_{ei} da ordem de 1, 5. A condição anterior é obrigatória apenas quando $k = 2$.

Em muitas das situações estudadas o número de classes foi inferior a cinco unidades motivo que conduziu a utilizar outra estatística para testar a aderência. O teste de Kolmogorov-Smirnov onde:

$F_o(X)$; freqüência observada acumulada.

$F_e(X)$; freqüência esperada acumulada obtida a partir do modelo.

ABS; valor absoluto da diferença entre $F_o(X)$ e $F_e(X)$.

Kolmogorov e Smirnov desenvolveram um método, em geral mais poderoso que o do qui-quadrado, no qual a variável de teste é dada pelo quociente entre a maior diferença absoluta e o número de observações n , ou seja:

$$d_{\text{calc}} = (\max |F_o(X) - F_e(X)|) / n \quad (32)$$

Se d_{calc} for menor do que o valor crítico tabelado em função do nível de significância α e do tamanho da amostra n , aceita-se o ajuste. Para $n > 50$, os valores críticos, para $\alpha = 5\%$ e 1% são dados por, respectivamente:

$$\frac{1,36}{\sqrt{n}} \text{ e } \frac{1,63}{\sqrt{n}}. \quad (33)$$

A seguir será feito o ajuste da distribuição Polinomial e de outras distribuições já conhecidas objetivando uma comparação entre os diferentes modelos. Também será apresentado o teste de Kolmogorov-Smirnov para testar a

aderência dos dados a cada modelo estudado. O teste do qui-quadrado também poderia ser apresentado, mas para efeito de comparação com a distribuição polinomial seria difícil estabelecer o número de graus de liberdade atribuído em cada caso.

4.3.1 Teste dos Dados do Jequitibá-Rosa com a Distribuição Polinomial

A distribuição polinomial, como já foi definida pela fórmula (30), é dada por

$$f(x) = \frac{1}{k} \begin{cases} c_1 x^d & \text{se } x < \ell_1 \\ a_1 x^n + a_2 x^{n-1} + a_3 x^{n-2} + \dots + a_m, & \text{se } \ell_1 \leq x \leq \ell_2 \\ \frac{c_2}{x^h} & \text{se } x > \ell_2 \\ 0 & \text{e.o.c.} \end{cases}$$

onde n, d e h são inteiros positivos

$a_1, a_2, a_3, \dots, a_m$ são números reais.

c_1 e c_2 são números reais

k valor da integral $\int_0^{\infty} [c_1 x^d + (a_1 x^n + a_2 x^{n-1} + a_3 x^{n-2} + \dots + a_m) + \frac{c_2}{x^h}] dx$

x , variável

ℓ_1 : limite superior da classe onde será ajustada a função $c_1 x^d$

ℓ_2 : limite superior da última classe onde o polinômio produz um bom ajuste

A função que foi desenvolvida e será analisada na tabela 4 é

$$f(x) = \frac{1}{2457,215694} \begin{cases} 0,01854703379x^5 & \text{se } 0 \leq x < 5 \\ -0,000369696970x^4 + 0,04964413364413x^3 - 2,25747668997667x^2 + \\ 36,60301087801059x - 74,59311261654885 & \text{se } 5 \leq x \leq 40 \\ \frac{34262750740}{x^6} & \text{se } x > 40 \\ 0 & \text{e.o.c.} \end{cases}$$

TABELA 4 - AJUSTE DOS DADOS DO JEQUITIBÁ-ROSA À DISTRIBUIÇÃO POLINOMIAL

x_i	f_o	$f(x)$	f_e	FO	FE	ABS
2,5	2	0,0010	2,4650	2	2,465	0,465
7,5	89	0,0377	93,0140	91	95,4789999	4,479
12,5	141	0,0481	118,5234	232	214,00244	17,998
17,5	85	0,0431	106,3311	317	320,333501	3,334
22,5	69	0,0313	77,0988	386	397,432345	11,432
27,5	55	0,0186	45,9258	441	443,358103	2,358
32,5	27	0,0091	22,3477	468	465,705843	2,294
37,5	12	0,0042	10,3377	480	476,043555	3,956
42,5	6	0,0020	4,9300	486	480,973555	5,026
47,5	4	0,0010	2,4650	490	483,438555	6,561
52,0	3	0,0010	2,4650	493	485,903555	7,096
	493				dcalc	0,036
					dtab	0,061

Como $d_{calc} < d_{tab}$ conclui-se pelo teste de Kolmogorov-Smirnov que os dados são aderentes ao modelo Polinomial. Nesta distribuição não existem parâmetros a serem estimados.

Nas tabelas que serão utilizadas para os testes de aderência $f(x)$ significa o valor da função no ponto, f_e a frequência esperada, FO a frequência observada acumulada, FE a frequência esperada acumulada e ABS o valor absoluto da diferença entre FO e FE .

4.3.2 Teste dos Dados do Jequitibá-Rosa com a Distribuição Gama

A distribuição gama apresenta a seguinte fórmula já mencionada, para o ajuste. Esta fórmula foi utilizada para fazer o teste na tabela 5.

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

TABELA 5 - AJUSTE DOS DADOS DO JEQUITIBÁ-ROSA AO MODELO GAMA

x_i	f_o	$f(x)$	f_e	FO	FE	ABS
2,5	2	0,0043	10,4957	2	10,36	8,3588
7,5	89	0,0336	82,8133	91	93,17	2,1721
12,5	141	0,0497	122,4814	232	215,65	16,3466
17,5	85	0,0447	110,2446	317	325,90	8,8980
22,5	69	0,0316	77,8575	386	403,76	17,7556
27,5	55	0,0193	47,5987	441	451,35	10,3543
32,5	27	0,0107	26,4426	468	477,80	9,7969
37,5	12	0,0056	13,7209	480	491,52	11,5178
42,5	6	0,0027	6,7651	486	498,28	12,2828
47,5	4	0,0013	3,2057	490	501,49	11,4885
52,5	3	0,0006	1,4716	493	502,96	9,9601
	493		503,0970		dcalc	0,0332
BETA	4,70024				dtab	0,0613
ALFA	3,84836					

Os parâmetros α e β da distribuição são obtidos a partir das expressões da média e da variância que são dadas pelas fórmulas (15) e (16).

4.3.3 Teste dos Dados do Jequitibá-Rosa com a Distribuição Beta

Para a distribuição beta, será utilizada a fórmula simplificada

$$f(x) = \frac{\Gamma(\alpha + \beta)}{\Gamma(\alpha)\Gamma(\beta)} x^{\alpha-1} (1-x)^{\beta-1} \quad \text{com } 0 < x < 1. \quad \text{Para que se possa trabalhar com essa}$$

fórmula, os valores serão corrigidos pela expressão que faz com que os valores sejam colocados entre zero e 1. A tabela 6 mostra o ajuste do modelo Beta.

$$x_{\text{corrigido}} = \frac{x_i - x_{\text{mínimo}}}{x_{\text{máximo}} - x_{\text{mínimo}}} \quad (34)$$

TABELA 6 - AJUSTE DOS DADOS DO JEQUITIBÁ-ROSA AO MODELO BETA

x_i	$x_i\text{-corr.}$	f_o	$f(x)$	f_e	FO	FE	ABS
2,5	0,05	2	0,63	28,37	2	28,37	26,3735
7,5	0,14	89	1,75	78,47	91	106,84	15,8435
12,5	0,23	141	2,23	99,77	232	206,61	25,3857
17,5	0,32	85	2,16	96,96	317	303,57	13,4294
22,5	0,41	69	1,77	79,38	386	382,96	3,0448
27,5	0,50	55	1,25	55,96	441	438,92	2,0849
32,5	0,59	27	0,75	33,51	468	472,42	4,4243
37,5	0,68	12	0,36	16,23	480	488,65	8,6514
42,5	0,77	6	0,13	5,66	486	494,31	8,3065
47,5	0,86	4	0,02	1,03	490	495,34	5,3405
52,5	0,95	3	0,00	0,02	493	495,36	2,3629
	-	493	-	495,36	-	dcalc	0,0535
						dtab	0,0613
E(X) =	0,328877		$\alpha =$	2,2538462			
V(X) =	0,0281055		$\beta =$	4,5993121			

Os parâmetros α e β foram calculados a partir de operações matemáticas elementares nas fórmulas da média e da variância já definidas:

$$E(x) = \frac{\alpha}{\alpha + \beta} \text{ e } \sigma^2 = \frac{\alpha \cdot \beta}{(\alpha + \beta)^2(\alpha + \beta + 1)}$$

Neste caso é importante informar que as fórmulas para os parâmetros são

$$\alpha = \frac{\mu^2 - \mu^3 - \sigma^2 \mu}{\sigma^2} \quad (35)$$

$$\text{e } \beta = \frac{\alpha - \alpha \cdot \mu}{\mu} \quad (36)$$

4.3.4 Teste dos Dados do Jequitibá-Rosa com a Distribuição Weibull Dois Parâmetros

Na distribuição Weibull a fórmula já foi definida como:

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

O ajuste do modelo está na tabela 7.

TABELA 7 - AJUSTE DOS DADOS DO JEQUITIBÁ ROSA AO MODELO WEIBULL DOIS PARÂMETROS

x_i	f_o	$f(x)$	f_e	FO	FE	ABS
2,5	2	0,002693	6,64	2	6,64	4,6374
7,5	89	0,017905	44,14	91	50,77	40,2263
12,5	141	0,037701	92,93	232	143,71	88,2937
17,5	85	0,049489	121,99	317	265,70	51,3031
22,5	69	0,045167	111,34	386	377,03	8,9668
27,5	55	0,029110	71,76	441	448,79	7,7891
32,5	27	0,013077	32,23	468	481,02	13,0237
37,5	12	0,003996	9,85	480	490,87	10,8749
42,5	6	0,000807	1,99	486	492,86	6,8640
47,5	4	0,000104	0,26	490	493,12	3,1211
52,5	3	0,000008	0,02	493	493,14	0,1417
-	493	-	493,14		dcalc	0,1791
					dtab	0,0613
c =	2,7688	E(X) =	19,55		V(Y) =	0,1208254
b =	21,965	E(Y) =	0,89		V(X) =	58,291918

A média da distribuição é $E(x) = a + b \cdot \Gamma\left(1 + \frac{1}{c}\right)$ e a sua variância é dada por $\sigma_x^2 = b^2 \cdot \sigma_y^2$, onde $y = \frac{x-a}{b}$.

Segundo PÉLLICO NETTO (1993) os parâmetros da distribuição Weibull podem ser obtidos pelas fórmulas 37 e 38, considerando $a = 0$.

$$c = \frac{\left[\sum_{i=1}^k (f \cdot \ln(x_i - a)) \right]}{\sum_{i=1}^k f} \quad (37)$$

e

$$b = \left[\frac{\sum_{i=1}^k [f \cdot (x_i - a)^c]}{\sum_{i=1}^k f} \right]^{\frac{1}{c}} \quad (38)$$

onde f representa a frequência dos dados observados e a será considerado como zero.

4.3.5 Teste dos Dados do Jequitibá-Branco com a Distribuição Polinomial

O modelo polinomial que será utilizado para o teste com o Jequitibá-Branco é o mesmo já definido.

$$f(x) = \frac{1}{k} \begin{cases} c_1 x^d & \text{se } x < \ell_1 \\ a_1 x^n + a_2 x^{n-1} + a_3 x^{n-2} + \dots + a_m, & \text{se } \ell_1 \leq x \leq \ell_2 \\ \frac{c_2}{x^h} & \text{se } x > \ell_2 \\ 0 & \text{e.o.c.} \end{cases}$$

onde n, d e h são inteiros positivos

$a_1, a_2, a_3, \dots, a_m$ são números reais.

c_1 e c_2 são números reais

k valor da integral $\int_0^{\infty} [c_1 x^d + (a_1 x^n + a_2 x^{n-1} + a_3 x^{n-2} + \dots + a_m) + \frac{c_2}{x^h}] dx$

x , variável

ℓ_1 : limite superior da classe onde será ajustada a função $c_1 x^d$

ℓ_2 : limite superior da última classe onde o polinômio produz um bom ajuste

Com a tabela de valores observados do Jequitibá-Branco calculou-se o valor da constante k que tornou a integral igual a 1.

Os coeficientes do polinômio foram obtidos por ajuste pelo polyfit do matlab quadrados mínimos não lineares e as funções das classes dos extremos foram ajustadas a partir das ordenadas do polinômio no ponto limítrofe da classe considerada. A função a seguir já foi desenvolvida no capítulo em 4.2 e resultou em:

$$f(x) = \frac{1}{823,952} \begin{cases} 0,0065015232612 \cdot x^5 & \text{se } 0 \leq x < 5 \\ -0,000065151x^4 + 0,015050x^3 - 0,83941x^2 + 14,75658x - 34,32092 & \text{se } 5 < x \leq 40 \\ \frac{0,473551}{x^6} & \text{se } x \geq 40 \\ 0 & \text{e.o.c.} \end{cases}$$

Para fazer o teste de aderência elaborou-se a tabela 8.

TABELA 8 - AJUSTE DO JEQUITIBÁ-BRANCO À DISTRIBUIÇÃO POLINOMIAL

x_i	f_o	$f(x)$	f_e	FO	FE	ABS
2,5	1	0,00077	0,64	1	0,64	0,3570
7,5	22	0,04282	35,75	23	36,40	13,3958
12,5	63	0,05678	47,41	86	83,81	2,1940
17,5	39	0,05025	41,95	125	125,76	0,7610
22,5	20	0,03336	27,85	145	153,62	8,6160
27,5	14	0,01508	12,59	159	166,20	7,2041
32,5	6	0,01183	9,88	165	176,08	11,0843
37,5	2	0,00501	4,19	167	180,27	13,2710
	167		180,27		dcalc	0,0802
					dtab	0,1052

4.3.6 Teste dos Dados do Jequitibá-Branco com a Distribuição Gama

Para a distribuição gama procedeu-se da mesma forma que com o Jequitibá-Rosa, ou seja, os parâmetros foram calculados com base nas expressões da média e da variância, os quais foram depois substituídos na expressão geral dada por

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

Para fazer os testes de aderência elaborou-se a tabela 9.

TABELA 9 - AJUSTE DO JEQUITIBÁ-BRANCO PELA DISTRIBUIÇÃO GAMA

x_i	f_o	$f(x)$	f_e	F0	FE	ABS
2,5	1	0,0010	0,86	1	0,8631	0,14
7,5	22	0,0318	26,58	23	27,45	4,45
12,5	63	0,0619	51,69	86	79,14	6,86
17,5	39	0,0530	44,29	125	123,43	1,57
22,5	20	0,0305	25,43	145	148,86	3,86
27,5	14	0,0138	11,51	159	160,37	1,37
32,5	6	0,0053	4,45	165	164,82	0,18
37,5	2	0,0018	1,54	167	166,35	0,65
	167		166,35		dcalc	0,04
ALFA	5,70				dtab	0,11
BETA	2,88					

4.3.7 Teste dos Dados do Jequitibá-Branco com a Distribuição Beta

Na distribuição beta o procedimento foi o mesmo que com o Jequitibá-Rosa. É importante salientar que novamente foi feita a correção dos dados para que ficassem no intervalo de zero até 1. Caso contrário teria sido necessário utilizar a fórmula da distribuição beta generalizada, que é um pouco mais complexa. A tabela 10 contém os dados de frequência e os testes de aderência para esse modelo.

$$f(x) = \frac{\Gamma(\alpha + \beta)}{\Gamma(\alpha)\Gamma(\beta)} x^{\alpha-1} (1-x)^{\beta-1} \quad \text{com } 0 < x < 1.$$

TABELA 10 - AJUSTE DO JEQUITIBÁ-BRANCO PELA DISTRIBUIÇÃO BETA

x_i	f_o	$f(x)$	f_e	FO	FE	ABS
0,06	1	0,2299264	4,6077246	1	4,61	3,6077
0,19	22	1,3296093	26,645371	23	31,25	8,2531
0,31	63	2,048715	41,056249	86	72,31	13,6907
0,44	39	2,0614547	41,311553	125	113,62	11,3791
0,56	20	1,5125612	30,311726	145	143,93	1,0674
0,69	14	0,7334523	14,698384	159	158,63	0,3690
0,81	6	0,2066641	4,1415495	165	162,77	2,2274
0,94	2	0,0067214	0,1346967	167	162,91	4,0927
-	167	-	162,90725	-	dcalc	0,0820
					dtab	0,1052
E(X) =	0,4103593	$\alpha =$	2,9385395			
V(X) =	0,0296493	$\beta =$	4,2223549			

4.3.8 Teste dos Dados do Jequitibá-Branco com a Distribuição Weibull

A expressão da função Weibull é, como já foi citado,

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

Devido ao fato de os parâmetros serem calculados por fórmulas extensas, costuma-se elaborar planilhas que vão fornecendo valores parciais dos parâmetros e da distribuição. Como tais planilhas foram muito extensas, optou-se por apresentar somente o valor funcional e os valores de interesse para o teste de aderência, como se pode observar na tabela 11.

TABELA 11 - AJUSTE DO JEQUITIBÁ-BRANCO PELA DISTRIBUIÇÃO WEIBULL

X_i	f_o	$f(x)$	f_e	FO	FE	ABS
2,5	1	0,0045722	3,8178021	1	3,8178021	2,8178021
7,5	22	0,0276908	23,121798	23	26,939601	3,9396005
12,5	63	0,0517273	43,192281	86	70,131881	15,868119
17,5	39	0,056052	46,803385	125	116,93527	8,064733
22,5	20	0,0383829	32,049753	145	148,98502	3,98502
27,5	14	0,0165689	13,835062	159	162,82008	3,8200815
32,5	6	0,0043822	3,6591175	165	166,4792	1,4791991
37,5	2	0,0006836	0,570774	167	167,04997	0,049973
-	167	-	167,04997		dcalc	0,0950187
					dtab	0,10524
c =	2,7113529	E(X) =	16,686105		V(X) =	44,086419
b =	18,76084	E(Y) =	0,8894114		V(Y) =	0,1252565

4.3.9 Análise Exploratória dos Dados do Jequitibá-Rosa

A tabela 12 tem como objetivo mostrar algumas medidas da estatística descritiva, como média, desvio padrão, variância e outras.

TABELA 12 - JEQUITIBÁ-ROSA – ANÁLISE EXPLORATÓRIA

x_i	f_o	$f_o \cdot x_i$	$(x_i)^2$	$(x_i)^2 f_o$	$f_o(x-u)^3$
2,5	2	5	6,25	12,50	-7575,67
7,5	89	667,5	56,25	5006,25	-105648,01
12,5	141	1762,5	156,25	22031,25	-24606,18
17,5	85	1487,5	306,25	26031,25	-17,30
22,5	69	1552,5	506,25	34931,25	5924,95
27,5	55	1512,5	756,25	41593,75	45853,79
32,5	27	877,5	1056,25	28518,75	80819,25
37,5	12	450	1406,25	16875,00	87776,04
42,5	6	255	1806,25	10837,50	87286,79
47,5	4	190	2256,25	9025,00	101770,76
52,5	3	157,5	2756,25	8268,75	122248,04
	493	8917,5	11068,75	203131,25	393832,45
MÉDIA	18,0882		COEF.VARIAÇÃO	0,51	
VARIÂNCIA	85,0191		ASSIMETRIA	1,02	
DESVIO PADRÃO	9,22058		MODA	12,41	

4.3.10 Análise Exploratória dos Dados do Jequitibá-Branco

Como os dados do Jequitibá-Branco também serão objeto de discussão na seqüência, é importante fazer uma exploração das características dos dados para que o leitor possa fazer associações com as análises feitas e com o que foi desenvolvido anteriormente, principalmente no modelo polinomial. A tabela 13 mostra a análise exploratória dos dados do Jequitibá-Branco.

TABELA 13 - JEQUITIBÁ-BRANCO - ANÁLISE EXPLORATÓRIA

x_i	f_o	$F_o x_i$	x_i^2	$x_i^2 f_o$	$f_o(x-u)^3$
2,50	1,00	2,50	6,25	6,25	-2698,48
7,50	22,00	165,00	56,25	1237,50	-15625,46
12,50	63,00	787,50	156,25	9843,75	-3801,16
17,50	39,00	682,50	306,25	11943,75	48,83
22,50	20,00	450,00	506,25	10125,00	4490,33
27,50	14,00	385,00	756,25	10587,50	19032,39
32,50	6,00	195,00	1056,25	6337,50	24936,43
37,50	2,00	75,00	1406,25	2812,50	18728,73
	167,00	2742,50	4250,00	52893,75	45111,61
MÉDIA	16,42		CV	0,42	
VARIÂNCIA	47,33		M3	270,13	
DESVIO PADRÃO	6,88		ASSIMETRIA	0,83	

4.4 AJUSTE DO MODELO POLINOMIAL A DADOS DE CIRCUNFERÊNCIA A 60 cm DE ALTURA DE PLANTAS DE PALMITO (*Euterpe edulis*)

Com o objetivo de mostrar que o modelo polinomial não foi elaborado somente para ajuste a dados de altura, será apresentada sua aplicação a dados de medidas de circunferência de plantas de palmito, a 60 cm de altura.

O estudo começa com uma análise exploratória dos dados, para se ter uma idéia inicial sobre suas características.

4.4.1 Caracterização do Conjunto de Dados

Inicialmente elaborou-se a tabela 14 onde constam medidas empíricas. Tais medidas de estatística descritiva foram obtidas por meio de fórmulas estudadas em estatística básica.

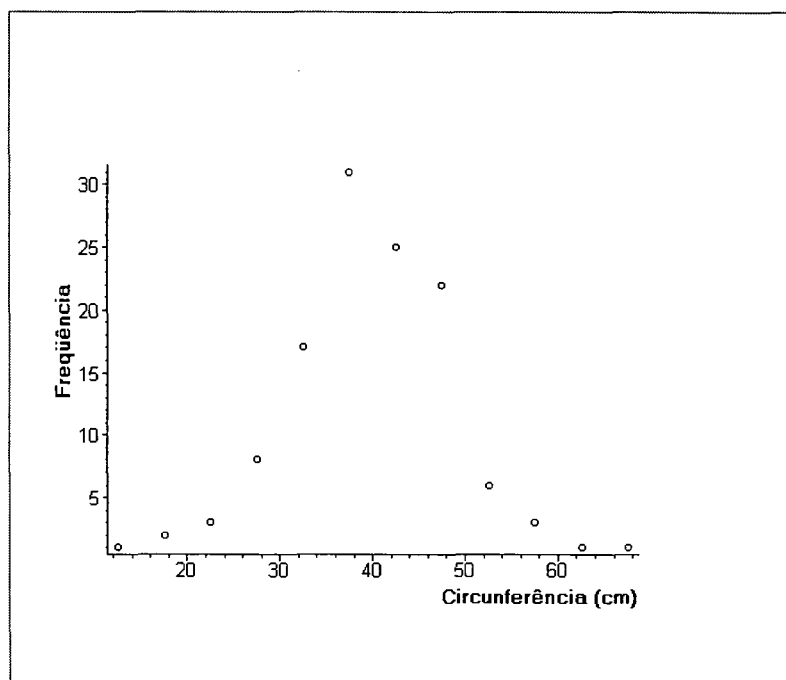
TABELA 14 - MEDIDAS EMPÍRICAS DOS DADOS DE CIRCUNFERÊNCIAS A 60 CM DE ALTURA DE PLANTAS DE PALMITO

x_i	f_o	$x_i \cdot f_o$	$X_i^2 \cdot f_o$	$f_o \cdot (x_i - u)^3$	$f_o \cdot (x_i - u)^4$
12,5	1	12,5	156,3	1953,13	24414,0625
17,5	2	35,0	612,5	10718,75	187578,125
22,5	3	67,5	1518,8	34171,88	768867,1875
27,5	8	220,0	6050,0	166375,00	4575312,5
32,5	17	552,5	17956,3	583578,13	18966289,06
37,5	31	1162,5	43593,8	1634765,63	61303710,94
42,5	25	1062,5	45156,3	1919140,63	81563476,56
47,5	22	1045,0	49637,5	2357781,25	111994609,4
52,5	6	315,0	16537,5	868218,75	45581484,38
57,5	3	172,5	9918,8	570328,13	32793867,19
62,5	1	62,5	3906,3	244140,63	15258789,06
67,5	1	67,5	4556,3	307546,88	20759414,06
	120	4842,5	204156,3	9006265,63	414537226,6
Média	39,79		CV=	0,27	
Desvio	10,90		M3	75052,21	
Variância	118,90		Assimetria	0,04	

Para que se tenha uma idéia da disposição dos dados, elaborou-se o gráfico no Maple com suas freqüências observadas como estão na tabela 14. A figura 22 ilustra o gráfico, obtido com o comando "pointplot".

```
>pointplot([12.5,1],[17.5,2],[22.5,3],[27.5,8],[32.5,17],[37.5,31],[42.5,25],[47.5,22],[52.5,6],[57.5,3],[62.5,1],[67.5,1]),symbol=circle,color=black);
```

FIGURA 22 - FREQUÊNCIA DOS DADOS DE CIRCUNFERÊNCIA À 60 cm DE ALTURA DE PLANTAS DE PALMITO



O gráfico de pontos e as medidas da tabela 14 foram uma referência para o ajuste do modelo polinomial e a comparação com os modelos já existentes na literatura.

4.4.2 Elaboração do Modelo Polinomial para os Dados de Circunferências a 60 cm de Altura de Plantas de Palmito

Inicialmente, fez-se um estudo no *software* Matlab para verificar qual é o polinômio que melhor se aproxima dos valores observados de frequência .

Colocados os dados no Matlab, foi dado o comando "polyfit", que fez o ajuste de um polinômio pelo método dos quadrados mínimos não lineares.

A partir do conjunto de dados da tabela 14 os dados foram colocados da mesma maneira que nos outros exemplos e foram escolhidas as condições que produzissem o polinômio mais adequado para o estudo.

```
p = 0,00007193501563 -0,01151034070928 0,60492347471042 -11,55958835553016
71,35015569780330
```

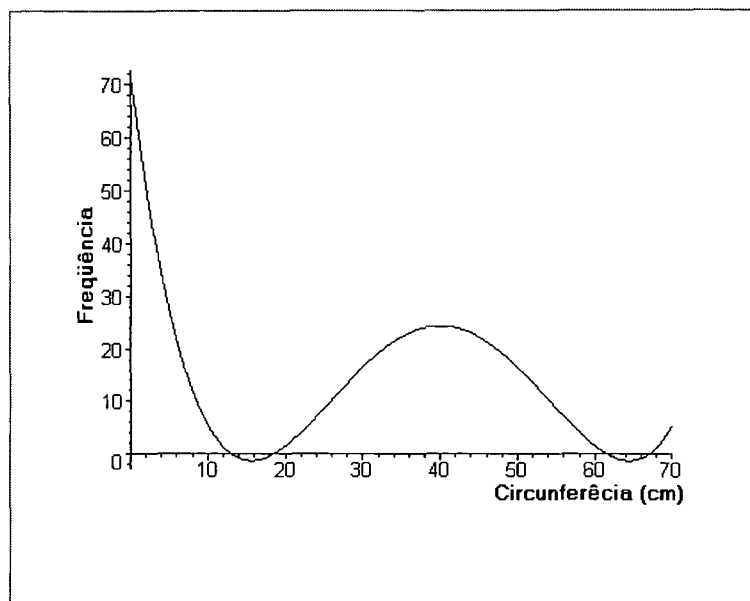
O polinômio obtido foi do quarto grau, uma vez que na terceira linha de comandos foi especificado n=4. É importante ressaltar novamente que se deve optar no Matlab pelo estilo "long", que fornece os coeficientes do polinômio com um número maior de casas decimais. Se os valores tiverem menos casas decimais, o gráfico terá outra forma, com esboço mais grosseiro dos dados.

Na seqüência os dados do polinômio de quarto grau foram transferidos para o *software* maple e ali foram plotados e tratados de acordo com o objetivo de construir o modelo polinomial.

No maple inicialmente fez-se um gráfico para verificar as classes nas quais o polinômio resultava negativo. A figura 23 mostra o resultado.

```
plot(0.00007193501563*x^4-0.01151034070928*x^3+0.60492347471042*x^2 -
11.55958835553016*x+71.35015569780330,x=0..70,color=black);
```

FIGURA 23 - GRÁFICO DO POLINÔMIO AJUSTADO PARA OS DADOS DE CIRCUNFERÊNCIAS DE PLANTAS DE PALMITO A 60cm DE ALTURA



A partir do gráfico da figura 23 e com os coeficientes obtidos no Matlab, pode-se começar a construção da função densidade de probabilidade Polinomial.

Em algumas classes o polinômio será substituído por outras funções .

Para se ter uma idéia melhor dos pontos de intersecção do polinômio com o eixo x, utilizou-se o comando "solve" do maple, que fornece as raízes.

```
>solve(0.00007193501563*x^4-0.01151034070928*x^3+0.60492347471042*x^2 -
11.55958835553016*x+71.35015569780330,x);
```

As raízes são as seguintes:

13,07254338; 18,40210796; 61,58045219; 66,955115863

A análise das raízes, juntamente com o gráfico da figura 23, determinou o procedimento que foi adotado.

Elaborou-se uma função $g_1(x)$ para as duas primeiras classes [10, 15) e [15, 20). Também se ajustou uma função $g_2(x)$ para as últimas classes [60, 65) e [65, 70).

Para elaborar a função $g_1(x)$ deve-se calcular a ordenada do ponto em que o polinômio foi seccionado, ou seja a ordenada de $x=20$. Isto foi feito utilizando-se o comando "subs" do maple, que fornece o valor numérico de uma expressão algébrica.

```
> subs(x=20,0.00007193501563*x^4-0.01151034070928*x^3+0.60492347471042*x^2 -
11.55958835553016*x+71.35015569780330);
```

O valor fornecido foi 1,5546552. A partir deste valor foi ajustada uma curva crescente do tipo $y = c_1 x^d$, conforme a expressão (48) que determina a estrutura básica do modelo polinomial.

Após algumas tentativas chegou-se à conclusão de que $d=2$ foi o melhor expoente para este caso. A partir daí foram feitos cálculos utilizando-se o comando "evalf" do maple para se obter a função.

```
> evalf(1.55465520/(20^2));
```

Assim, obteve-se $c_1=0,003886638000$, e para verificar se a função elaborada realmente produzia a ordenada do valor 20, fez-se novo cálculo com o comando "subs" do maple.

```
> subs(x=20,0.003886638000*x^2);
```

• Resultado

1,5546552

Portanto a função obtida foi

$$g_1(x) = 0,003886638x^2.$$

Para a função das últimas classes adotou-se o mesmo procedimento. É importante salientar que as funções destas classes foram escolhidas seguindo-se a tendência dos dados, ou seja, a tendência decrescente.

A função $g_3(x)$ foi elaborada de forma a permitir o cálculo das integrais

$$\int_{\ell_2}^{\infty} g_3(x) dx, \int_{\ell_2}^{\infty} x \cdot g_3(x) dx \text{ e } \int_{\ell_2}^{\infty} x^2 \cdot g_3(x) dx. \text{ A primeira foi utilizada para estabelecer a}$$

função densidade de probabilidade e as duas últimas para o cálculo da média e da variância, respectivamente.

Como o valor de ℓ_2 , que para esse caso era 60, foi procurada a ordenada desse valor no polinômio de forma a garantir a continuidade da função.

```
> subs(x=60,0.00007193501563*x^4-0.01151034070928*x^3+0.60492347471042*x^2 -
11,55958835553016*x+71.35015569780330);
```

1,54357310

Com o valor obtido ajustou-se uma função do tipo $y = \frac{c_2}{x^h}$, conforme a fórmula (30) do modelo Polinomial. A partir de testes sucessivos ficou determinado que o valor $h=6$ produziu um bom ajuste.

Calculou-se c_2 utilizando o commando "evalf" do Maple, que realizou as operações de aritmética.

```
> evalf(1.54357310*(60^6));
> 72016946550
```

Com o valor obtido pôde-se definir $g_3(x) = \frac{72016946550}{x^6}$, e para verificar se

realmente esta função fornecia a ordenada de $x=20$ igual à gerada pelo polinômio isto é, 1,54357310, fez-se novo cálculo com o comando "subs".

```
> subs(x=60,72016946550/x^6);
> 1.543573100;
```

Depois de vários cálculos julgou-se oportuno fazer um resumo da função obtida:

$$g(x) = \begin{cases} 0,003886638x^2 & \text{se } 0 \leq x < 20 \\ 0,00007193501563x^4 - 0,01151034070928x^3 + 0,60492347471042x^2 - \\ 11,55958835553016x + 71,35015569780330 & \text{se } 20 \leq x \leq 60 \\ \frac{60,72016946550}{x^6} & \text{se } x > 60 \\ 0 & \text{e.o.c.} \end{cases}$$

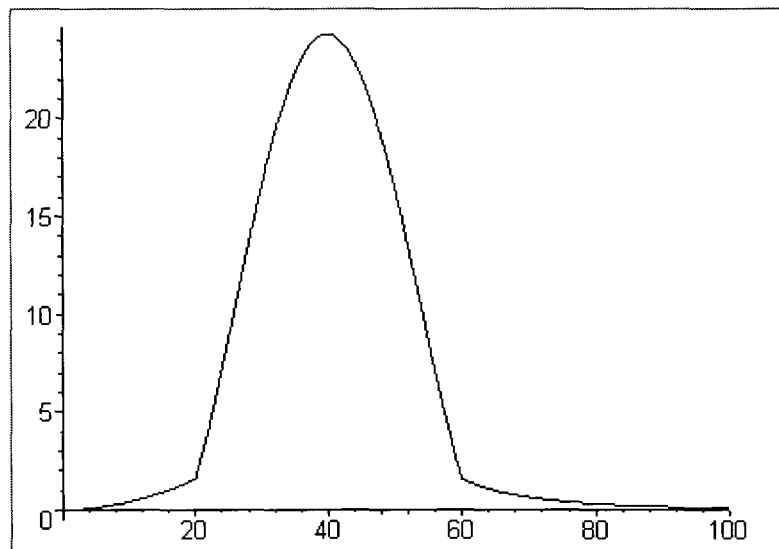
A figura 24 mostra a função $g(x)$ definida no intervalo $0 \leq x < 20$ pela função $g_1(x)$, no intervalo $20 \leq x \leq 60$ pelo polinômio ajustado e para $x > 60$ pela função $g_3(x)$.

O gráfico de $g(x)$ pôde ser obtido pelo comando "piecewise" do Maple, que elabora curvas para funções definidas por mais de uma sentença.

```
> f:=x->piecewise(x>=0 and x<20,0.003886638000*x^2,x<=60 and
x>=20,0.00007193501563*x^4-0.01151034070928*x^3+0.60492347471042*x^2 -
11.55958835553016*x+71.35015569780330,x>=60,72016946550/x^6).
```

Na figura 24 pode-se verificar que houve uma tentativa de acompanhar os pontos apresentados na figura 22.

FIGURA 24 - GRÁFICO DA FUNÇÃO $g(x)$ ELABORADA PARA OS DADOS DE CIRCUNFERÊNCIA DE PLANTAS DE PALMITO A 60 cm DE ALTURA.



Na seqüência, calculou-se a integral da função $g(x)$ no intervalo de zero até infinito. Este procedimento foi feito com os comandos "int" e "evalf" do maple.

```
> int(0.003886638000*x^2,x=0..20);
```

```
10,36436800
```

```
> int(0.00007193501563*x^4-
0.01151034070928*x^3+0.60492347471042*x^2 -
11.55958835553016*x+71.35015569780330,x=20..60);
```

```
608,2307234
```

```
> int(72016946550/x^6,x=60..infinity);
```

```
160037659
8640000
```

```
> evalf(10.36436800+608.2307234+160037659/8640000);
```

```
637,1179686
```

Portanto, de acordo com a fórmula (30), $k = 637.1179686$.

Multiplicando a função $g(x)$ pelo inverso de k obteve-se a integral unitária, outra condição para que se tenha uma função densidade de probabilidade.

```
> (1/637.1179686)*int(0.003886638000*x^2,x=0..20);
```

```
0,01626758075
```

```
> (1/637.1179686)*int(0.00007193501563*x^4-
0.01151034070928*x^3+0.60492347471042*x^2 -
11.55958835553016*x+71.35015569780330,x=20..60);
```

```
0,9546595032
```

```
> (1/637.1179686)*int(72016946550/x^6,x=60..infinity);
```

```
0,02907291602
```

```
> evalf(0.1626758075e-1+0.9546595032+0.2907291602e-1);
```

```
1,000000000
```

A função densidade de probabilidade Polinomial elaborada para os dados de circunferência de plantas de palmito à 60 cm de altura ficou determinada da seguinte forma:

$$f(x) = \frac{1}{k} \cdot g(x)$$

$$f(x) = \frac{1}{637,1179686} \cdot \begin{cases} 0,003886638 x^2 & \text{se } 0 \leq x < 20 \\ 0,0000719350 1563 x^4 - 0,0115103407 0928 x^3 + 0,6049234747 1042 x^2 - \\ 11,5595883555 3016 x + 71,3501556978 0330 & \text{se } 20 \leq x \leq 60 \\ \frac{60,7201694655 0}{x^6} & \text{se } x > 60 \\ 0 & \text{e.o.c} \end{cases}$$

4.4.3 Cálculo da Média e da Variância pelo Modelo Polinomial

Para calcular a média foi utilizada a fórmula (3). Os cálculos foram feitos com o comando "int" do maple. Calculou-se a integral de cada uma das componentes separadamente e depois os resultados foram somados.

```
> (1/637.1179686)*int(0.003886638000*x^3,x=0..20);
```

```
0,2440137112
```

```
> (1/637.1179686)*int(0.00007193501563*x^5-
0.01151034070928*x^4+0.60492347471042*x^3 -
11.55958835553016*x^2+71.35015569780330*x,x=20..60);
```

```
38,18504953
```

```
> (1/637.1179686)*int(72016946550/x^5,x=60..infinity);
```

```
2,180468702
```

```
> evalf(0.2440137112+38.18504953+2.180468702);
```

```
40,60953194
```

O resultado para a média 40,60953194 m está próximo de 39,79 m obtido na tabela 14 de maneira empírica.

A variância pôde ser obtida pela formula (4) e também foi calculada utilizando-se comandos do Maple.

```
> (1/637.1179686)*int(0.003886638000*x^4,x=0..20);
```

```
3.904219379
```

```

>(1/637.1179686)*int(0.00007193501563*x^6-
0.01151034070928*x^5+0.60492347471042*x^4-
11.55958835553016*x^3+71.35015569780330*x^2,x=20..60);
1600.099701
> (1/637.1179686)*int(72016946550/x^4,x=60..infinity);
174,4374961
> evalf(3.904219379+1600.099701+174.4374961);
1778,441416
> evalf(1778.441416-(40.60953194)^2);
129,307332

```

O resultado obtido para a variância 129,307332 m² está de certa forma próximo do valor 118,9 m² obtido na tabela 14. Isto indica que o modelo desenvolvido reproduz razoavelmente as medidas de posição e dispersão calculadas de forma empírica.

4.4.4 Testes de Aderência do Modelo Polinomial aos Dados de Circunferência a 60cm de Altura de Plantas de Palmito

Os resultados obtidos no ajuste do modelo polinomial aos dados de circunferência à 60 cm de altura de dados de plantas de palmito estão na tabela 15.

A comparação entre os resultados está na seção 4.5.

TABELA 15 -TESTES DE ADERÊNCIA DOS DADOS DO PALMITO AO MODELO POLINOMIAL

x_i	f_o	$cx^2 e \frac{d}{6}$	b_4x^4	b_3x^3	b_2x^2	b_1x	b_0	$f(x)$	f_e	FO	FE	ABS
12,500	1	0,001						0,001	0,572	1	0,570	0,430
17,500	2	0,002						0,002	1,121	3	1,691	1,309
22,500	3		18,436	-131,110	306,243	-260,091	71,350	0,008	4,547	6	6,238	0,238
27,500	8		41,141	-239,379	457,473	-317,889	71,350	0,020	11,957	14	18,194	4,194
32,500	17		80,255	-395,128	638,950	-375,687	71,350	0,031	18,591	31	36,785	5,785
37,500	31		142,254	-606,991	850,674	-433,485	71,350	0,037	22,416	62	59,201	2,799
42,500	25		234,691	-883,598	1092,643	-491,283	71,350	0,037	22,416	87	81,618	5,382
47,500	22		366,197	-1233,585	1364,859	-549,080	71,350	0,031	18,590	109	100,208	8,792
52,500	6		546,484	-1665,582	1667,320	-606,878	71,350	0,020	11,954	115	112,162	2,838
57,500	3		786,342	-2188,224	2000,028	-664,676	71,350	0,008	4,540	118	116,702	1,298
62,500	1	0,002						0,002	1,138	119	117,840	1,160
67,500	1	0,001						0,001	0,717	120	118,557	1,443
	120								118,559		dcalc	0,073
											dtab	0,124

4.4.5 Ajuste de Outras Distribuições Contínuas aos Dados de Circunferência a 60 cm de Altura de Plantas de Palmito.

Os cálculos para o modelo gama estão na tabela 16.

TABELA 16 - AJUSTE DOS DADOS DO PALMITO PELA DISTRIBUIÇÃO GAMA

x_i	f_o	$f(x)$	f_e	FO	FE	ABS
12,5	1	0,0002	0,1336	1	0,13	0,8700
17,5	2	0,0026	1,5701	3	1,70	1,2999
22,5	3	0,0108	6,4699	6	8,17	2,1701
27,5	8	0,0238	14,2927	14	22,46	8,4627
32,5	17	0,0348	20,8740	31	43,34	12,3367
37,5	31	0,0378	22,7007	62	66,04	4,0374
42,5	25	0,0330	19,7959	87	85,83	1,1667
47,5	22	0,0242	14,5412	109	100,37	8,6255
52,5	6	0,0155	9,3123	115	109,69	5,3132
57,5	3	0,0089	5,3309	118	115,02	2,9823
62,5	1	0,0046	2,7795	119	117,80	1,2028
67,5	1	0,0022	1,3391	120	119,14	0,8637
	120				dcalc	0,1028
ALFA	13,315762				dtab	0,1242
BETA	2,98818799					

Na tabela 17 os resultados obtidos para o modelo beta.

TABELA 17 - AJUSTE DOS DADOS DO PALMITO PELA DISTRIBUIÇÃO BETA

x_i	x_i -corr.	f_o	$f(x)$	f_e	FO	FE	ABS
12,5	0,04	1	0,002	0,02	1	0,02	0,9842
17,5	0,13	2	0,092	0,92	3	0,94	2,0612
22,5	0,21	3	0,483	4,83	6	5,77	0,2329
27,5	0,29	8	1,192	11,92	14	17,68	3,6836
32,5	0,38	17	1,968	19,68	31	37,36	6,3605
37,5	0,46	31	2,460	24,59	62	61,95	0,0517
42,5	0,54	25	2,431	24,30	87	86,25	0,7520
47,5	0,63	22	1,899	18,98	109	105,23	3,7718
52,5	0,71	6	1,120	11,19	115	116,42	1,4214
57,5	0,79	3	0,440	4,39	118	120,82	2,8156
62,5	0,88	1	0,080	0,80	119	121,62	2,6202
67,5	0,96	1	0,001	0,01	120	121,63	1,6328
	-	120	-	121,63	-	dcalc	0,0530
						dtab	0,1242
a =	5,0455999		E(X) =	0,4965278			
b =	5,1161677		V(X) =	0,0223968			

Os valores dos modelos Weibull e Normal estão nas tabelas 18 e 19.

TABELA 18 - AJUSTE DOS DADOS DO PALMITO PELA DISTRIBUIÇÃO WEIBULL

x_i	f_o	$f(x)$	f_e	FO	FE	ABS
12,5	1	0,0033665	2,0198967	1	2,0198967	1,0198967
17,5	2	0,0079961	4,7976803	3	6,817577	3,817577
22,5	3	0,0146754	8,8052324	6	15,622809	9,6228094
27,5	8	0,0224409	13,464555	14	29,087364	15,087364
32,5	17	0,0293555	17,613306	31	46,700671	15,700671
37,5	31	0,032996	19,797628	62	66,498299	4,4982991
42,5	25	0,0316298	18,977879	87	85,476178	1,523822
47,5	22	0,0254727	15,28362	109	100,7598	8,2402018
52,5	6	0,0168887	10,133231	115	110,89303	4,1069704
57,5	3	0,0089974	5,398466	118	116,2915	1,7085044
62,5	1	0,003747	2,2481937	119	118,53969	0,4603107
67,5	1	0,0011834	0,7100157	120	119,2497	0,750295
-	120	-	119,2497		dcalc	0,1308389
					dtab	0,1241504
c =	3,6548417	E(X) =	38,117772	V(X) =	134,59329	
b =	42,26637	E(Y) =	0,9018464	V(Y) =	0,0753414	

TABELA 19 - AJUSTE DOS DADOS DO PALMITO PELA DISTRIBUIÇÃO NORMAL

x_i	$x_i(\text{transf.})$	f_o	$1/\text{raiz}(2\pi)*\sigma$	$(-1/2)((x-\mu)/\sigma)^2$	$e^{((-1/2)(x-\mu)/\sigma)^2}$	$f(x)$	f_e	F0	FE	ABS
12,5	0,042	1	2,665765356	-4,6122	0,0099	0,0265	0,2638	1	0,2638	0,7362
17,5	0,125	2	2,665765356	-3,0815	0,0459	0,1224	1,2188	3	1,4826	1,5174
22,5	0,208	3	2,665765356	-1,8585	0,1559	0,4157	4,1403	6	5,6229	0,3771
27,5	0,292	8	2,665765356	-0,9339	0,3931	1,0478	10,4363	14	16,0592	2,0592
32,5	0,375	17	2,665765356	-0,3297	0,7192	1,9171	19,0943	31	35,1535	4,1535
37,5	0,458	31	2,665765356	-0,0331	0,9674	2,5789	25,6857	62	60,8392	1,1608
42,5	0,542	25	2,665765356	-0,0462	0,9549	2,5455	25,3534	87	86,1926	0,8074
47,5	0,625	22	2,665765356	-0,3685	0,6918	1,8442	18,3685	109	104,5611	4,4389
52,5	0,708	6	2,665765356	-0,9984	0,3685	0,9824	9,7845	115	114,3456	0,6544
57,5	0,792	3	2,665765356	-1,9490	0,1424	0,3797	3,7820	118	118,1276	0,1276
62,5	0,875	1	2,665765356	-3,1978	0,0409	0,1089	1,0850	119	119,2126	0,2126
67,5	0,958	1	2,665765356	-4,7542	0,0086	0,0230	0,2289	120	119,4415	0,5585
		120					119,4414		dcalc	0,0370
Média=	0,49652778	Desvio=	0,14965563						dtab	0,1242

4.5 ANÁLISE DE RESULTADOS

4.5.1 Análise dos Resultados da Variável Altura Obtidos para o Jequitibá-Rosa

O Jequitibá-Rosa apresentou na Distribuição Polinomial um $d_{calc} = 0,036$ e $d_{tab} = 0,061$, portanto, $d_{calc} < d_{tab}$. Isto indica que os dados são aderentes ao modelo polinomial.

Para o modelo Gama observou-se na tabela 5 que $d_{calc} = 0,03$ e $d_{tab} = 0,06$.

Se $d_{calc} < d_{tab}$ então o modelo gama é aderente aos dados.

No modelo Beta obteve-se $d_{calc} = 0,0535$ e $d_{tab} = 0,0613$, portanto, $d_{calc} < d_{tab}$ o que indica que os dados são aderentes ao modelo.

No modelo Weibull testado (dois parâmetros) obteve-se $d_{calc} = 0,17$ e $d_{tab} = 0,061$. Como $d_{calc} > d_{tab}$, o modelo não é aderente aos dados.

A média obtida diretamente dos dados observados para o Jequitibá-Rosa é de aproximadamente 18,08 m e a média obtida por meio da fórmula (3) no modelo Polinomial, calculada em 4.1.2 resultou em aproximadamente 17,91 m, ou seja, extremamente próximos.

A tabela 12 mostra uma variância de aproximadamente 85,0191 m² para os dados do Jequitibá-Rosa. Pela distribuição Polinomial, calculada por meio da fórmula (4) em 4.1.3, a variância resultou em 92,9112 m² o que vem confirmar o bom ajuste indicado pela estatística de Kolmogorov-Smirnov.

Pelos cálculos efetuados pelas distribuições gama, beta e Weibull, usando-se respectivamente as fórmulas (15), (20) e (23), os seguintes resultados foram obtidos para a média: gama: 18,08 m, beta: 18,087 m (0,32 convertido) e Weibull: 19,55 m.

Para calcular a variância pelas distribuições citadas, utilizaram-se as fórmulas (16) para a Gama, (21) para a Beta e (24) para a Weibull. Os resultados obtidos para a variância foram: Gama: 83,42 m²; Beta: 84,7 m² (0,028 convertido) e Weibull: 58,29 m².

Portanto para o Jequitibá Rosa somente o modelo Weibull não foi aderente. Mesmo assim a média ainda resultou em valor próximo, apenas a variância (58,29 m²) ficou distante do valor obtido na análise exploratória.

4.5.2 Análise dos Resultados Obtidos para o Jequitibá-Branco

Para os dados do Jequitibá-Branco, pode-se observar nas tabelas 8, 9, 10 e 11 que, segundo o teste de Kolmogorov-Smirnov os modelos Polinomial, Gama, Beta e Weibull de dois parâmetros, mostraram que têm aderência para os dados. Isto pode ser confirmado, por exemplo, pela variância da Weibull (44,08 m²), tabela 11, valor próximo de 47,33 m² da análise exploratória.

A média dos dados observados segundo a tabela 13, foi de 16,42 m e o valor obtido por meio da integral $E(X) = \int_{-\infty}^{\infty} xf(x)dx$ resultou em 15,31 m.

A variância, segundo a citada tabela do Jequitibá-Branco foi de 47,32 m². Já nos cálculos com a fórmula $\sigma_x^2 = \int_{-\infty}^{\infty} x^2 f(x)dx - [E(x)]^2$ resultou em 38,15m².

4.5.3 Análise dos Resultados dos Testes de Aderência do Modelo Polinomial e de outros Modelos aos Dados de Circunferência à 60 cm de Altura de Plantas de Palmito

A média e a variância do modelo Polinomial, como já foi comentado, resultaram em estimadores muito bons se comparados com aqueles da tabela 14: a média na tabela foi 39,79 m, e de aproximadamente 40 m no modelo, a variância na tabela resultou em 118,9 m² e no modelo 129m².

Com os dados das plantas de palmito verificou-se que o modelo polinomial pode ser aplicado também a outras medidas e não somente a dados de altura.

Com relação aos dados da tabela 15, pode-se verificar que o dcalc. foi 0,073 e o dtab. teve resultado 0,124, o que significa aderência pelo teste de Kolmogorov-Smirnov.

Para os dados da tabela 16, da distribuição gama resultou dcalc=0,1028 e dtab=0,1242, o que significa que o modelo gama é aderente aos dados do palmito.

Para o modelo Beta a tabela 17 mostrou $d_{calc.}=0,054$ e $d_{tab}=0,124$, o que significa que os dados do palmito são aderentes ao modelo Beta.

Para o modelo Weibull com dois parâmetros, o teste de Kolmogorov-Smirnov mostrou que $d_{calc.}=0,1308$ e $d_{tab.}=0,1242$, o que significa que o modelo é razoavelmente aderente aos dados tomados para o palmito. Esta pequena diferença pode ser verificada na variância, tabela 18, $134,59m^2$ que ultrapassou o valor $118,90m^2$ da análise exploratória.

Para o modelo normal, tabela 19, foi feita uma correção na escala dos valores. Pela característica dos pontos da figura 22 pode-se inferir que há tendência para uma curva simétrica neste conjunto de dados. Esta tendência foi confirmada nos valores $d_{calc}=0,003$ e $d_{tab}=0,1242$, o que indica que o modelo normal é aderente aos dados.

É importante salientar que, no caso do palmito os dados da figura 22 mostram tendência para uma simetria, o que significa que a distribuição polinomial não é apenas para situações onde há assimetria. Pode-se dizer que para casos com simetria o modelo normal e outros modelos são muito mais indicados, motivo pelo qual constam de várias publicações. A idéia da distribuição polinomial é oferecer uma alternativa para os casos em que há assimetria e não há possibilidade de ajuste pelos modelos existentes.

CONCLUSÕES

O presente trabalho possibilitou um estudo das funções matemáticas e suas fórmulas para verificar a possibilidade de transformá-las em funções densidade de probabilidade. Na revisão de literatura pôde-se constatar o grande número de conceitos matemáticos e expressões envolvidos nas distribuições contínuas. As fórmulas das funções já existentes, seus gráficos e as expressões que calculam a média e a variância revelaram uma dimensão mais prática e foram gradativamente balizando a formação de um modelo que pudesse suprir condições peculiares de dados biológicos amostrados em florestas naturais.

O estudo parcial do modelo desenvolvido permite as seguintes conclusões.

- a) o ajuste de um polinômio utilizando-se aplicativo computacional atende ao propósito do projeto de pesquisa, que, desde o início, tinha como objetivo encontrar um polinômio de grau n , que pudesse ser transformado em função densidade de probabilidade;
- b) na comparação com outras funções já existentes, pode-se observar na discussão dos resultados, que o objetivo foi atingido, ou seja, um modelo polinomial foi construído e, devido à sua flexibilidade, demonstra ser aderente a dados que não se ajustam bem em outras distribuições.
- c) os testes realizados indicam que, mesmo com dados nos quais a frequência atinge abruptamente, valores muito altos ou muito baixos é possível traçar uma curva que mais se aproxime dos dados observados e, a partir daí, fazer os ajustes para transformar a equação da curva em função densidade de probabilidade;
- d) observando as distribuições já existentes pode-se perceber que a lógica de elaboração é semelhante, ou seja, um parâmetro é responsável pela forma e outros pela escala. Entretanto, a tentativa de

elaborar uma curva que vá ao encontro dos dados e se aproxime ao máximo deles demonstra ser mais eficiente que o método de determinar parâmetros que combinados produzam a melhor curva;

- e) os conceitos do Cálculo Diferencial e Integral, como funções, continuidade e integrais impróprias, contribuem para que se elabore uma função que, além de ser não negativa seja convergente no infinito. A convergência no infinito norteou a elaboração da função para as classes finais. Entretanto, vale salientar a importância do aplicativo computacional que permite verificar instantaneamente se uma função é convergente ou não;
- f) o modelo polinomial é, em sua concepção, formado por três funções. A maior parte dos dados é explicada pelo polinômio, e as funções das classes dos extremos têm o objetivo de eliminar os intervalos em que a expressão ajustada tenha valores negativos ou assuma comportamentos que contrariem a tendência dos dados. Isto não quer dizer que o número de funções ou o seu grau sejam fixos. Podem haver adaptações resultantes de estudos, e, desde que se respeitem os princípios matemáticos, novos modelos podem ser elaborados dentro dessa idéia;
- g) as comparações com as principais distribuições que constam da literatura nacional e internacional mostram que o modelo desenvolvido apresenta vantagens, principalmente se houver aplicativos como Maple ou Matlab para a elaboração de curvas. As vantagens mais expressivas são uma maior flexibilidade na curva e a formação de um modelo com funções simples de trabalhar.
- h) a comparação com dados do palmito mostra que a função densidade de probabilidade polinomial não é aderente apenas a dados de altura, pois pode ser utilizada também para outras variáveis; além disso pode-se constatar que se os dados tiverem tendência para simetria o modelo polinomial também pode ser ajustado.

- i) para a obtenção da distribuição polinomial é necessário usar o aplicativo computacional, porém em muitas outras distribuições existe a necessidade de recorrer a tabelas prontas para se calcular parâmetros, que fazem parte da composição da distribuição;
- j) um outro ponto verificado neste trabalho foi a dificuldade de, em alguns casos, estimar parâmetros. Como a função densidade Polinomial não tem parâmetros, não há esse problema. Há também possibilidade de se controlar melhor a função adaptando-a para situações especiais.
- k) os aplicativos computacionais permitem que haja uma maior flexibilidade tanto na visualização de gráficos, quanto na otimização de cálculos. Isso permite que no modelo polinomial seja feita de forma instantânea a escolha do grau do polinômio que melhor se ajusta aos dados. Este fato permitiu que fossem elaboradas dezenas de funções e que fosse analisada a sua viabilidade em pouco espaço de tempo.
- l) A maneira proposta para a obtenção da distribuição polinomial pode resultar em benefícios para a evolução da Ciência Florestal. Fica reafirmada a idéia de que as ciências quando trabalham juntas na resolução de problemas do mundo trazem benefícios para todos.

REFERÊNCIAS

- CORAIOLA, Márcio. **Caracterização estrutural de uma floresta estacional semidecidual localizada no Município de Cássia MG**. Curitiba, 1997. 186 f. Dissertação (Mestrado) Universidade Federal do Paraná.
- COSTA NETO, Pedro Luiz de Oliveira. **Estatística**. São Paulo: Edgard Blücher, 1999. p.130-135.
- CUNHA, Cristina. **Métodos numéricos**. 2.ed. Campinas, Editora da Unicamp, 1990. p.93-120.
- GIOTTO, Enio. **Análise da dinâmica de evolução da cobertura florestal e sua quantificação por métodos e processos de amostragem em múltiplas ocasiões** Curitiba, 1986. 318 f. Tese (Doutorado) Universidade Federal do Paraná.
- GMURMAN, V. E. **Teoria das probabilidades e estatística matemática**. Moscou: Mir, 1983. p.139-150.
- GUIDORIZZI, Hamilton Luiz. **Um curso de cálculo**. 5.ed. São Paulo: LTC, 1994. 566p.
- GUIDORIZZI, Hamilton Luiz. **Um curso de cálculo**. São Paulo: LTC, 1987. 566p.
- IEZZI, Gelson, MURAKAMI, Carlos. **Fundamentos da matemática elementar**. 5.ed. São Paulo: Atual, 1993. 269p. V.8
- LEITHOLD, Louis. **O cálculo com geometria analítica**. 3.ed. São Paulo: Harbra, 1994. p39. V1.
- MAGALHÃES, Marcos Nascimento; LIMA, Antonio Carlos Pedroso. **Noções de probabilidade e estatística**. São Paulo: Edusp, 2002. p.165-194.
- MATSUMOTO, Élia Yathie. **Matlab 6: fundamentos de programação**. São Paulo: Erica, 2001. p.232-239.
- MEYER, Paul L. **Probabilidade: aplicações à estatística**. 2.ed. Rio de Janeiro: LTC, 1984. p.214-274.
- PÉLLICO NETTO, Sylvio. **Biometria: teoria de probabilidades**. Curitiba: UFPR, 1993. 278p. (Apostila).
- PÉLLICO NETTO, Sylvio; BRENA, Doádi Antonio. **Inventário florestal**. Curitiba: Gráfica da UFPR, 1997.
- PENNY, John; LINDFIELD, George. **Numerical methods using matlab**. New Jersey: Prentice Hall, 2000. p.282-330.
- PISKOUNOV, N. **Cálculo diferencial e integral**. 16.ed. Porto, Portugal: Lopes da Silva, 1993. p.60-62. V.1.
- RADAM BRASIL. **Levantamento de recursos naturais**. [snt], 1978. 492p. V.32.

ROSS, Sheldon M. **A first course in probability**. 5.ed. New Jersey: Prentice Hall, 1998. p.192-226.

ROUSSAS, George G. **A course in mathematical statistics**. 2.ed. Londres: Adisson Wesley Publishing Company, 1997. 536p

RUGGIERO, Márcia A. Gomes; LOPES, Vera Lúcia da Rocha. **Cálculo numérico – aspectos teóricos e computacionais**. 2.ed. São Paulo: Makron Books, 1996. p.268-291.

SOONG, T. T. **Modelos probabilísticos em engenharia e ciências**. Rio de Janeiro: LTC, 1986. p.172-216.

SWOKOWSKI, Earl Willian. **Cálculo com geometria analítica**. 2.ed. São Paulo: Makron Books 1994. p.619-652.

TANEJA, Inder Jeet. **MAPLE V: uma abordagem computacional no ensino de cálculo**. Florianópolis: UFSC, 1997. 327p.

GLOSSÁRIO

dap	Diâmetro à altura do peito.
evalf	Comando do Maple que executa cálculos aritméticos.
f.d.p	Função densidade de probabilidade.
int	Comando do Maple que resolve a integral indefinida ou definida conforme a entrada de informações.
piecewise	Comando do Maple que permite fazer o gráfico de uma função definida por mais de uma sentença.
plot	Comando do Maple que faz o gráfico de uma função de uma ou duas variáveis livres.
pointplot	Comando do Maple que fornece o gráfico dos pontos.
polyfit	Comando do Matlab que retorna os coeficientes de um polinômio de grau n a partir de um conjunto de dados fornecidos.
solve	Comando do Maple que determina as raízes reais e complexas de um polinômio.
subs	Comando do Maple que permite obter o valor numérico de uma expressão algébrica.

**ANEXO 1 - AJUSTE DOS DADOS DE ALTURAS DE *Eucalyptus saligna* PELA
DISTRIBUIÇÃO WEIBULL COM DOIS PARÂMETROS**

CLASSES (m)	x_i	f_o	$x_i \cdot f_o$	$x_i^2 \cdot f_o$	$f_o \cdot \ln x_i$	$f_o \cdot (x_i)^c$	$f(x)$	f_e	FO	FE	ABS
2--4	3	1,00	3	9	1,10	19,40	0,01	9,11	1,00	9,11	8,11
4--6	5	5,00	25	125	8,05	385,01	0,02	20,95	6,00	30,06	24,06
6--8	7	7,00	49	343	13,62	1336,60	0,04	34,66	13,00	64,71	51,71
8--10	9	13,00	117	1053	28,56	4891,31	0,05	47,53	26,00	112,24	86,24
10--12	11	22,00	242	2662	52,75	14227,26	0,06	56,83	48,00	169,07	121,07
12--14	13	62,00	806	10478	159,03	62936,57	0,07	60,51	110,00	229,58	119,58
14--16	15	122,00	1830	27450	330,38	182224,69	0,07	57,91	232,00	287,49	55,49
16--18	17	168,00	2856	48552	475,98	351776,59	0,06	49,97	400,00	337,47	62,53
18--20	19	37,00	703	13357	108,94	104600,26	0,04	38,89	437,00	376,36	60,64
20--22	21	3,00	63	1323	9,13	11111,31	0,03	27,24	440,00	403,60	36,40
	-	440,00	6694	105352	1187,55	733509,01	-	403,60	-	dcalc	0,28
										dtab	0,06
	c =	2,70		E(X) =	13,89		V(Y)	0,13			
	b =	15,62		E(Y) =	0,89		V(X)	30,81			
MÉDIA E VARIÂNCIA CALCULADAS DIRETAMENTE A PARTIR DOS DADOS											
				média	15,21		var.	7,99			

**ANEXO 2 - AJUSTE DOS DADOS DE DAP COM CASCA DE *pinus* sp PELA
DISTRIBUIÇÃO BETA**

CLASSES	x_i	f_o	$f(x)$	f_e	F_o	F_e	ABS
7,3 ----- 9,3	8,3	3	0,0045	1	3	1	2,3450
9,3 ----- 11,3	10,3	2	0,0178	3	5	3	1,7854
11,3 ----- 13,3	12,3	2	0,0314	5	7	8	0,7398
13,3 ----- 15,3	14,3	4	0,0435	6	11	14	3,0094
15,3 ----- 17,3	16,3	8	0,0531	8	19	22	2,6573
17,3 ----- 19,3	18,3	9	0,0595	9	28	30	2,2277
19,3 ----- 21,3	20,3	9	0,0624	9	37	39	2,2079
21,3 ----- 23,3	22,3	12	0,0614	9	49	48	0,9472
23,3 ----- 25,3	24,3	8	0,0566	8	57	56	0,7940
25,3 ----- 27,3	26,3	7	0,0480	7	64	63	0,8761
27,3 ----- 29,3	28,3	5	0,0360	5	69	68	0,6938
29,3 ----- 31,3	30,3	2	0,0211	3	71	71	0,3497
31,3 ----- 33,3	32,3	1	0,0052	1	72	72	0,0944
soma	-	72		71,6	-	dcalc =	0,0418

$$d_{\text{tab}} = \frac{1,36}{\sqrt{72}} = 0,1603$$

$$d_{\text{calc}} < d_{\text{tab}}$$

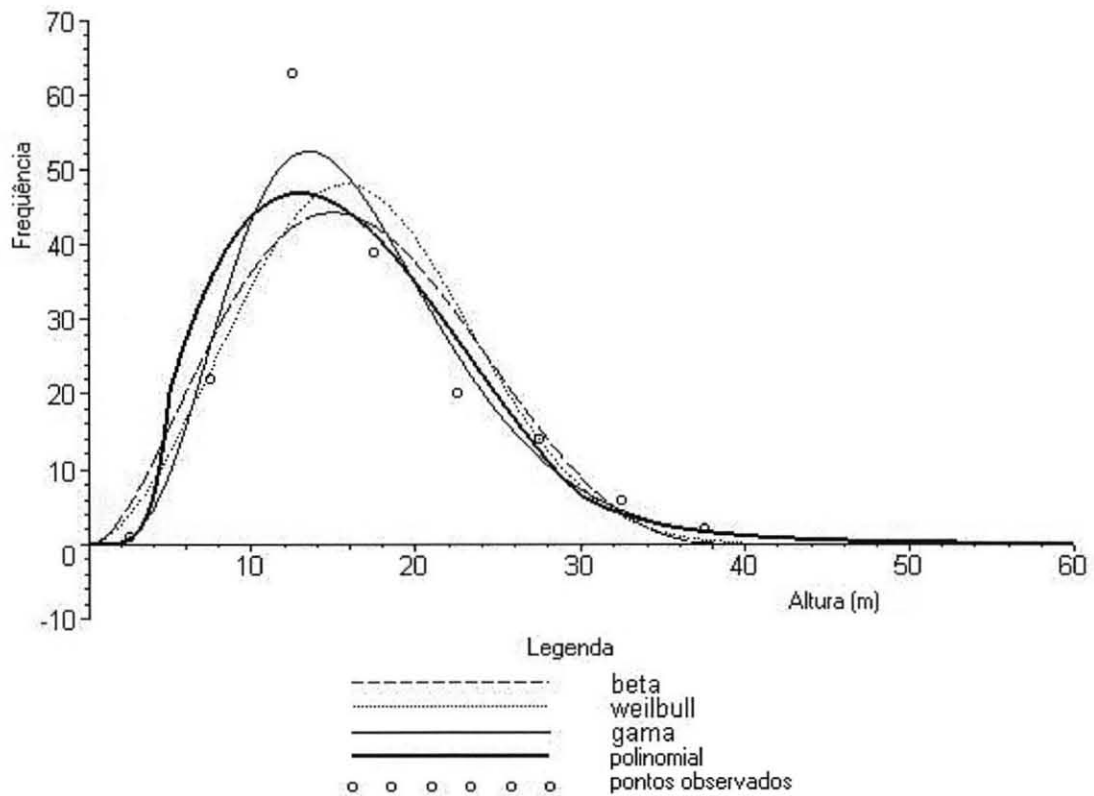
**ANEXO 3 - AJUSTE DOS DADOS DE DAP COM CASCA DE *pinus* sp PELA
DISTRIBUIÇÃO WEIBULL COM DOIS PARÂMETROS**

CLASSES	x_i	f_o	$f(x)$	f_e	F_o	F_e	ABS
7,3 ----- 9,3	8,3	3	0,0051	1	3	1	2,2621
9,3 ----- 11,3	10,3	2	0,0111	2	5	2	2,6626
11,3 ----- 13,3	12,3	2	0,0206	3	7	5	1,6990
13,3 ----- 15,3	14,3	4	0,0337	5	11	10	0,8496
15,3 ----- 17,3	16,3	8	0,0493	7	19	17	1,7494
17,3 ----- 19,3	18,3	9	0,0646	9	28	27	1,4422
19,3 ----- 21,3	20,3	9	0,0752	11	37	37	0,3914
21,3 ----- 23,3	22,3	12	0,0766	11	49	48	0,5820
23,3 ----- 25,3	24,3	8	0,0667	10	57	58	1,0205
25,3 ----- 27,3	26,3	7	0,0483	7	64	65	0,9801
27,3 ----- 29,3	28,3	5	0,0282	4	69	69	0,0403
29,3 ----- 31,3	30,3	2	0,0127	2	71	71	0,1264
31,3 ----- 33,3	32,3	1	0,0043	1	72	72	0,0000
Soma	-	72		71,6	-	dcalc =	0,0370

$$d_{\text{tab}} = \frac{1,36}{\sqrt{72}} = 0,1603 \text{ e } d_{\text{calc}} = 0,0370$$

$$d_{\text{calc}} < d_{\text{tab}}$$

**ANEXO 4 - GRÁFICO DE ALTURA TOTAL DO JEQUITIBÁ BRANCO
JUNTAMENTE COM PONTOS OBSERVADOS E AJUSTES FEITOS
PELAS DISTRIBUIÇÕES BETA, WEIBULL DOIS PARÂMETROS,
GAMA E POLINOMIAL**



ANEXO 5 - GRÁFICO DE ALTURA TOTAL DO JEQUITIBÁ ROSA JUNTAMENTE COM PONTOS OBSERVADOS E AJUSTES FEITOS PELAS DISTRIBUIÇÕES POLINOMIAL, GAMA, BETA E WEIBULL DE DOIS PARÂMETROS

