FUNÇÃO DENSIDADE DE PROBABILIDADE APLICÁVEL À CIÊNCIA FLORESTAL

Eduardo Quadros da Silva¹
Sylvio Péllico Netto²
Sebastião do Amaral Machado³
Carlos Roberto Sanquetta⁴

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 tenha sido 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. Neste estudo procurou-se mostrar que o modelo é adaptável também a dados de diâmetro e situações onde há simetria. 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 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 testar a aplicabilidade em outras situações procurou-se ajustar o modelo a dados de diâmetros e, após a aplicação do teste de Kolmogorov-Smirnov, os resultados mostraram-se satisfatórios.

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

PROBABILITY DENSITY FUNCTION APPLICABLE TO FORESTRY

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 resulted distribution 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 an n -polynomial, preferably a 5th degree increasing curve, a positive decreasing curve tending to zero, and the integral converging to the infinite. 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

¹Prof. Adjunto da PUCPR; Doutorando em Engenharia Florestal pela UFPR. Email: eduardoq@bsi.com.br

²Prof. Titular da PUCPR; Prof. Sênior do Curso de Pós-Graduação em Engenharia Florestal da UFPR. Email: pellico@rla13.pucpr.br

³Prof. Sênior do Curso de Pós-Graduação em Eng. Florestal da UFPR. Pesquisador 1A do CNPq, email: sammac@floresta.ufpr.br

⁴Prof. Adjunto do Centro de Ciências Florestais e da Madeira da UFPR. Email: sanquetta@floresta.ufpr.br

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 mean height 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 other existing distributions.

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

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), a maioria das decisões humanas é tomada a partir da análise ou estudo de uma parte, apenas, do todo que envolve o problema. Pode-se diagnosticar uma doença a partir de uma pequena análise de tecido, pode-se analisar algumas árvores e verificar que muitas estão sendo atacadas pela vespa da madeira. 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.

Uma teoria que tem sido utilizada de forma significativa na ciência é a das probabilidades. Segundo Soong (1986), as idéias probabilísticas estão presentes em ampla diversidade de investigações científicas que envolvem aleatoriedade. 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.

O planejamento de um inventário florestal exige uma análise de diversos fatores que, muitas vezes, precisam ser estimados. Nem sempre é possível fazer uma completa enumeração ou censo devido a fatores de ordem econômica e ambiental.

Dentro do contexto de aleatoriedade, são formuladas distribuições probabilísticas que permitem prever

fenômenos a partir de modelos matemáticos, permitindo o estudo do comportamento de variáveis do meio físico ou biológico, denominadas funções densidade de probabilidade

Em Guidorizzi (2001) tem-se a informação que uma função definida para todo x real e integrável em todo intervalo $[a, b]$, com a e b reais e $a < b$ é uma função densidade de probabilidade se as seguintes condições forem satisfeitas: (i) $f(x) \geq 0$ para todo x e (ii)

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

Se X é uma variável aleatória contínua com função densidade de probabilidade f então define-se a esperança matemática de x por

$$E(X) = \int_{-\infty}^{\infty} xf(x)dx \quad (1)$$

e sua variância por

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

Neste trabalho elaborou-se uma função densidade de probabilidade constituída basicamente por três funções. A partir dos dados foi gerado um polinômio "mais próximo" dos valores das frequências.

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 onde 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 é preciso obter um valor aproximado da função em algum ponto fora do intervalo de tabelamento, ou seja, quando se quer extrapolar. Também não se deve utilizar interpolação quando 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.

Quando não é aconselhável utilizar a interpolação surge a necessidade de ajustar uma função que seja uma "aproximação razoável" para os dados e que permita fazer extrapolações.

Segundo 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, buscamos 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.

Segundo Piskounov (1993) a minimização da distância é feita com base nos conceitos estudados no Cálculo Diferencial e Integral. Em Cálculo estuda-se que, nos pontos críticos, (máximos ou mínimos) a derivada de primeira ordem deve ser nula. A partir dessas idéias são deduzidas as equações normais que fazem parte do método citado.

Pelo fato de se dispor de tecnologia que pudesse otimizar o trabalho utilizou-se para ajustar o polinômio pelo método dos quadrados mínimos o software Matlab, mais especificamente o comando polyfit. Segundo Matsumoto (2001) dados uma série de valores x , y e o grau do polinômio n , o comando polyfit retorna coeficientes do

polinômio de grau n que melhor aproxima a curva dada pela série em x , y (método dos mínimos quadrados)". Também se pode encontrar exemplos de ajuste de polinômios pela função polyfit em (Penny E Lindfield, 1999). Naquela obra o assunto é tratado no capítulo referente à regressão polinomial.

Outro software utilizado foi o Maple pelo fato de se ter uma maior facilidade em plotar gráficos e uma facilidade em realizar cálculos. A obra de Taneja (1997) especifica os comandos e respectivas funções no Maple.

Ao ajustar o polinômio percebeu-se que havia um problema para ser resolvido quando houvessem raízes reais. Tais raízes indicam que valores funcionais anteriores ou posteriores podem ser negativos o que não satisfaz uma das condições de uma função densidade de probabilidade. Por este motivo, para as classes onde o polinômio resultou em valores negativos foram elaboradas outras funções. Desta forma o modelo elaborado ficou constituído basicamente por três funções.

O objetivo geral do trabalho foi contribuir com a Ciência Florestal utilizando as estruturas matemáticas para a formulação de uma nova função densidade de probabilidade.

Para se atingir o objetivo geral as seguintes metas foram estabelecidas: encontrar um polinômio de grau n que pudesse ser transformado a fim de atender às condições de uma função probabilística, propor fórmulas para o cálculo da média e da variância da distribuição e desenvolver uma função que fosse formulada a partir da característica dos dados, atendendo aos requisitos de uma função densidade de probabilidade e princípios da Matemática.

MATERIAIS E MÉTODOS

Foram realizados testes com várias funções, utilizando dados diversos de altura de árvores, DAP, circunferência e outras medidas.

Pode-se definir genericamente a função polinomial por:

$$f(x) = \frac{1}{k} \begin{cases} c_1 x^d & \text{se } 0 < x < \ell_1 \\ a_1 x^n + a_2 x^{n-1} + a_3 x^{n-2} + \dots + a_m, & \text{se } \ell_1 \leq x \leq \ell_2 \\ \frac{c_2}{x^h} & \text{se } x > \ell_2 \\ 0 & \text{e.o.c} \end{cases} \quad (3)$$

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

Conforme o exposto, a função polinomial é definida por três sentenças ou fórmulas. Estas devem atender aos requisitos de uma fdp ou seja, formarem um todo contínuo, com valores funcionais não negativos e convergente no infinito. 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 onde o polinômio assume valores negativos ou contraria a tendência do modelo;

Terceira etapa:

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

Para a primeira classe a função a ser ajustada é da forma $g_1(x) = c_1 x^5$. Inicialmente deve se calcular a ordenada do limite ℓ_1 . Em seguida deve se descobrir

qual deve ser o valor de c_1 para que a ordenada de ℓ_1 calculada pela função $g_1(x) = c_1 x^5$ tenha o mesmo valor que a ordenada de ℓ_1 calculada pelo polinômio. Este cuidado deve ser tomado para que a função seja contínua em ℓ_1 .

Para as últimas classes é necessário uma função que, além de atender a tendência dos dados seja convergente no infinito, isto pode ser obtido com uma função do tipo $g_3(x) = \frac{c_2}{x^h}$ onde c_2 é um número real. Como na função anterior deve-se calcular o valor da ordenada de ℓ_2 no polinômio. Em seguida deve-se 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 $g_1(x) = c_1 x^5$, $g_2(x)$ e

$g_3(x) = \frac{c_2}{x^h}$ e calcular $\int_0^{\infty} g(x) dx$ para a obtenção do

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-se a função $g(x)$ por $\frac{1}{k}$ obtém-se uma função $f(x)$ onde $\int_{-\infty}^{\infty} f(x)dx = 1$. Portanto tem-se uma função contínua, com valores positivos e com integral convergente para um, atendendo as exigências de uma função densidade de probabilidade.

RESULTADOS

Para verificar se o modelo elaborado tinha consistência foram feitos testes com dados de diversas características. Os dados de DAP com casca, apresentados na tabela 1 foram extraídos de uma parcela de 600 m² medida sobre uma floresta de Pinus com 12 anos de idade. Foram estabelecidas classes com amplitude de 2cm.

Tabela 1: Dados de diâmetros de pinus - análise exploratória

Table 1: DBH pine data – prospective analysis

CLASSES (cm)	x_i	f_i	$f_i \cdot x_i$	x_i^2	$x_i^2 f_i$
7,3 I---9,3	8,3	3	24,9	68,89	206,67
9,3 I---11,3	10,3	2	20,6	106,09	212,18
11,3I---13,3	12,3	2	24,6	151,29	302,58
13,3I---15,3	14,3	4	57,2	204,49	817,96
15,3I---17,3	16,3	8	130,4	265,69	2125,52
17,3I---19,3	18,3	9	164,7	334,89	3014,01
19,3I---21,3	20,3	9	182,7	412,09	3708,81
21,3I---23,3	22,3	12	267,6	497,29	5967,48
23,3I---25,3	24,3	8	194,4	590,49	4723,92
25,3I---27,3	26,3	7	184,1	691,69	4841,83
27,3I---29,3	28,3	5	141,5	800,89	4004,45
29,3I---31,3	30,3	2	60,6	918,09	1836,18
31,3I---33,3	32,3	1	32,3	1043,29	1043,29
		72	1485,6		32804,88
	μ	20,6333			
	σ^2	30,3099			
	σ	5,50544			

A elaboração do modelo Polinomial para os dados da tabela 1 seguiu as etapas:

Primeira Etapa:

Os dados que foram analisados estão contidos na tabela 1.

Na tabela 1, partindo dos pontos médios de classe e das frequências absolutas foi feito o ajuste utilizando-se o comando polyfit do Matlab.

$x=[8.3 \ 10.3 \ 12.3 \ 14.3 \ 16.3 \ 18.3 \ 20.3 \ 22.3 \ 24.3 \ 26.3 \ 28.3 \ 30.3 \ 32.3];$

$y=[3 \ 2 \ 2 \ 4 \ 8 \ 9 \ 9 \ 12 \ 8 \ 7 \ 5 \ 2 \ 1];$

$n=4;$

$p=\text{polyfit}(x,y,n)$

$p = 0.00061167386535$

$0.05243598313451 \ 1.53618831842178$

-17.70209157203100

71.30115396731084

Os números que o comando *polyfit* fornece são os coeficientes do polinômio de 4º grau que melhor se aproxima dos dados. A figura 1 mostra o gráfico, elaborado no *software Maple*, do polinômio $p(x) = 0,00061167386535 x^4 - 0,05243598313451 x^3 + 1,53618831842178 x^2 -$

$17,70209157203100x + 71,30115396731084$.

Figura 1: Gráfico do polinômio ajustado juntamente com pontos observados

Figure 1: Adjusted polynomial line on the observed data

Segunda etapa:

Na figura 1 observou-se que o polinômio não era adequado para as classes [7,3; 9,3) e [9,3; 11,3) pois a área gerada de zero até o valor 11,3 iria distorcer os resultados uma vez que estes iniciam em 7,3.

Na última classe [31,3; 33,3), foi constatado que o polinômio tinha uma tendência crescente, o que não poderia acontecer pois a integral deve convergir para um número.

Terceira etapa:

Foi elaborada uma função $g_1(x)$ para as duas primeiras classes de forma que a ordenada do valor 11,3 obtida no polinômio coincidissem com a ordenada na função elaborada. Para agilizar o processo utilizou-se comandos do Maple.

```
> subs (x=11.3,0.00061167386535*x^4-
0.05243598313451*x^3+
1.53618831842178*
17.70209157203100*x+
71.301153967310);
```

1.73686377

```
> evalf (1.73686377/ (11.3)^5);
```

0,942700068810⁻⁵

```
> subs (x=11.3,0.9427000688e-5*x^5);
```

1,736863770

Como foi verificado na última linha do maple que a função reproduzia a ordenada do polinômio, determinou-se

$$g_1(x) = 0,9427000688 \cdot 10^{-5} x^5.$$

Para a última classe foi elaborada uma função utilizando a igualdade $g_2(x) = \frac{c_2}{x^h}$.

```
> subs (x=31.3,0.00061167386535*x^4-
0.05243598313451*
x^3+1.53618831842178*x^2-
17.70209157203100*x+ 71.301153967310);
```

1,38144687

```
> evalf (1.38144687* (31.3)^20);
```

0,112517723710³¹

```
> subs (x=31.3,0.1125177237e31/x^20);
```

1,381446870

Como a ordenada da função para o valor 31,3 é o mesmo do polinômio pôde-se concluir que

$$g_3(x) = \frac{0,1125177237 \cdot 10^{31}}{x^{20}}.$$

Quarta etapa:

Nesta etapa foi formada a função $g(x)$ com a composição das três parciais $g_1(x)$, $g_2(x)$ e $g_3(x)$. Posteriormente foi

feito o cálculo da integral $\int_0^{\infty} g(x)dx$ e obtido

Figura 2: Gráfico da função $g(x)$

Figure 2: Graphic of the function $g(x)$

Para o cálculo do valor de k utilizou-se o comando `int` do Maple que fornece a integral em um dado intervalo.

> int (0.9427000688e-5*x^5,x=0..11.3);

3,271093434

> int (0.00061167386535*x^4-0.05243598313451*x^3+1.53618831842178*x^2-17.70209157203100*x+71.301153967310,x=11.3..31.3);

132,5274348

> int (0.1125177237e31/x^20,x=31.3..infinity);

o valor de k . Para unir as funções no maple foi 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);

Com o comando `plot` elaborou se o gráfico de $g(x)$ como mostra a figura 2.

2,275751948

> evalf (3.271093434+132.5274348+2.275751948);

138,0742801

Portanto, o valor que deve multiplicar a função para que sua integral no campo dos reais seja um é o inverso de k ou seja, $1/138,0742801$

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 .

Finalmente pôde-se escrever a função densidade de probabilidade para os dados do

$$f(x) = \frac{1}{138,0742801} \begin{cases} 0,9427000688 \cdot 10^{-5} x^5 & \text{se } 0 \leq x < 11,3 \\ 0,00061167386535x^4 - 0,05243598313451x^3 + 1,53618831842178x^2 - \\ 17,70209157203100x + 71,301153967310 & \text{se } 11,3 \leq x \leq 31,3 \\ \frac{1125177237 \cdot 10^{21}}{x^{20}} & \text{se } x > 31,3 \\ 0 & \text{e.o.c.} \end{cases}$$

Média Obtida pelo Modelo Polinomial ajustado

Para calcular esta medida utilizou-se a fórmula (1), ou seja, o cálculo foi feito por meio da integral.

$$\begin{aligned} &> (1/138.0742801)*\text{int} \\ &(0.9427000688e-5*x^6, x=0..11.3); \\ &\boxed{0,2294625500} \\ &> (1/138.0742801)*\text{int} \\ &(0.00061167386535*x^5 - \\ &0.05243598313451*x^4 + \\ &1.53618831842178*x^3 - \\ &17.70209157203100*x^2 + \\ &71.301153967310*x, x=11.3..31.3); \\ &\boxed{20,42058086} \\ &> (1/138.0742801)*\text{int} \\ &(0.1125177237e31/x^19, x=31.3..infinity); \\ &\boxed{0,5445497576} \\ &> \text{evalf} \\ &(0.2294625500+20.42058086+0.544549757 \\ &6); \\ &\boxed{21,19459317} \end{aligned}$$

A média obtida é 21,19459317cm.

Variância Obtida pelo Modelo Polinomial ajustado

Para calcular a variância foi utilizada a fórmula (2).

$$\begin{aligned} &> (1/138.0742801)*\text{int} \\ &(0.9427000688e-5*x^7, x=0..11.3); \end{aligned}$$

DAP, mostrados na tabela 1.

$$\begin{aligned} &2.268810963 \\ &> (1/138.0742801)*\text{int} (0.0006116738653 \\ &5*x^6 - 0.05243598313451*x^5 + \\ &1.53618831842178*x^4 - \\ &17.70209157203100*x^3 + 1.301153967310* \\ &x^2, x=11.3..31.3); \\ &\boxed{453,9603611} \\ &> (1/138.0742801)*\text{int} \\ &(0.1125177237e31/x^18, x=31.3..infinity); \\ &\boxed{18,04701961} \\ &> \text{evalf} \\ &(2.268810963+453.9603611+18.04701961); \\ &\boxed{474,2761917} \\ &> \text{evalf} (474.2761917 - (21.19459317)^2); \\ &\boxed{25,0654121} \end{aligned}$$

A variância é 25,0654121 cm^2

Teste de Aderência do modelo Polinomial ajustado.

Segundo COSTA NETO (1999) Kolmogorov e Smirnov desenvolveram um método, em geral mais poderoso que o do Qui-Quadrado, para testar a aderência, em que a variável de teste é o quociente entre a maior diferença absoluta observada entre a frequência acumulada da amostra e a estimada pelo modelo pelo número de observações n , ou seja:

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

Se d_{cal} 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:

$$\frac{1,36}{\sqrt{n}} \text{ e } \frac{1,63}{\sqrt{n}}.$$

Tabela 2: Teste de kolmogorov-smirnov
Table 2: Kolmogorow-smirnov test

classes	x_i	f_i	$f(x)$	f_e	FO	FE	ABS
7,3I---9,3	8,3	3	0,0027	0,3873	3	0,3873	2,6127
9,3I---11,3	10,3	2	0,0079	1,1397	5	1,5270	3,4730
11,3I---13,3	12,3	2	0,0174	2,5023	7	4,0293	2,9707
13,3I---15,3	14,3	4	0,0329	4,7354	11	8,7647	2,2353
15,3I---17,3	16,3	8	0,0507	7,2990	19	16,0638	2,9362
17,3I---19,3	18,3	9	0,0656	9,4422	28	25,5060	2,4940
19,3I---21,3	20,3	9	0,0740	10,6590	37	36,1650	0,8350
21,3I---23,3	22,3	12	0,0742	10,6883	49	46,8533	2,1467
23,3I---25,3	24,3	8	0,0661	9,5140	57	56,3673	0,6327
25,3I---27,3	26,3	7	0,0511	7,3651	64	63,7324	0,2676
27,3I---29,3	28,3	5	0,0327	4,7152	69	68,4476	0,5524
29,3I---31,3	30,3	2	0,0159	2,2833	71	70,7309	0,2691
31,3I---33,3	32,3	1	0,0053	0,7681	72	71,4990	0,0179
		72		71,4990		dcalc	0,0482
						dtab	0,1603

Na tabela 2, x_i é o ponto médio da classe, f_i a frequência, $f(x)$ é o valor obtido na função para cada x_i , f_e é a frequência esperada, FO é a frequência observada acumulada, FE é a frequência esperada acumulada e ABS é o valor absoluto da diferença entre FO e FE.

Discussão dos resultados

A função densidade de probabilidade Polinomial apresentou uma frequência esperada de 71,499 ou seja, próxima da frequência observada 72, conforme tabela 2.

Na figura 1, pode-se constatar que o polinômio, responsável pelo ajuste na maioria das classes, está bem próximo dos pontos que representam os dados observados.

A média obtida no modelo foi 21,19459317cm, valor próximo a 20,6333 da tabela 1, obtido empiricamente a partir dos dados observados.

A variância no modelo resultou em 25,0654121 cm² valor relativamente próximo a 30,3098 cm² da tabela 1.

No teste de Kolmogorov-Smirnov os resultados obtidos foram $d_{72;0,05} = 0,1603$ pela fórmula 4 e

O teste do qui-quadrado não foi realizado porque com a distribuição polinomial ficaria difícil estabelecer o número de graus de liberdade que seria atribuído em cada caso. (4)

Na tabela 2 pode-se verificar os resultados obtidos no teste de aderência.

$d_{calc} = 0,0482$ pela fórmula 3. Portanto, como $d_{calc} < d_{tab}$ o modelo é aderente aos dados.

A figura 2 mostra que os dados apresentam uma certa simetria o que significa que talvez sejam aderentes ao modelo normal.

CONCLUSÕES

- Após a análise dos resultados pôde-se tirar as seguintes conclusões:
- O gráfico da figura 1 mostrou que o polinômio sempre fornece uma boa aproximação dos dados observados, desde que se escolha corretamente o seu grau. Esta escolha pode ser feita no Matlab alterando os valores de n.
- Como a distribuição foi ajustada a partir dos dados, não ficou sujeita a combinações de parâmetros como acontece em distribuições contínuas existentes.
- Os conceitos do Cálculo Diferencial e Integral como funções, continuidade e integrais impróprias foram importantes para que se elaborasse uma função que, além de ser não negativa devia ser convergente no infinito. A convergência no infinito norteou a elaboração da função para as classes finais.

- À primeira vista o ajuste da função polinomial parece ser mais complexo que o uso das funções densidade já existentes. Na realidade, para se obter resultados em tempo razoável é necessário utilizar aplicativos computacionais o que hoje, não é um problema.
- 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 onde 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.

BIBLIOGRAFIA CITADA

- COSTA NETO, P. L. O. **Estatística**. São Paulo: Edgard Blücher 1999. p.130-135.
- CUNHA, C. Métodos numéricos. 2.ed. Campinas, **Editora da Unicamp**, 1990. p.93-120.
- GUIDORIZZI, H. L. Um curso de cálculo. São Paulo: **LTC**, 2001. p.45-60.
- MATSUMOTO, E. Y. **Matlab 6 - Fundamentos de programação**. São Paulo: Erica, 2001. p.232-239.
- PÉLLICO NETTO, S.; BRENA, D. **Inventário florestal**. Curitiba: Editorado pelos autores, 1997. p13.
- PENNY, J.; LINDFIELD, G. 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 Editora**, 1993. p.60. 516p. V.1.
- RUGGIERO, M. A. G.; LOPES, V. L. R.. Cálculo numérico - aspectos teóricos e computacionais. 2.ed. São Paulo: **Makron Books**, 1994. p.268-291.
- SALAS, S. L.; HILLE, E. Calculus - one and several variables. 6th ed. New York: **John Wiley and Sons**, 1990. p.569-609.
- SOONG, T. T. **Modelos probabilísticos em engenharia e ciências**. Rio de Janeiro: LTC, 1986. p.172-216.
- TANEJA, I. J. MAPLE V: **Uma abordagem computacional no ensino de cálculo**. Florianópolis: UFSC, 1997. 327p.