

**Luiz Medeiros de Araujo Lima Filho**

**Modelos simétricos transformados não-lineares  
com diferentes distribuições dos erros:  
aplicações em ciências florestais**

Recife

2009

**Luiz Medeiros de Araujo Lima Filho**

**Modelos simétricos transformados não-lineares  
com diferentes distribuições dos erros:  
aplicações em ciências florestais**

Dissertação submetida ao Programa de Pós-graduação em Biometria e Estatística Aplicada da Universidade Federal Rural de Pernambuco para a obtenção do título de Mestre em Biometria e Estatística Aplicada

Orientador:

José Antônio Aleixo da Silva

Co-orientadores: Gauss Moutinho Cordeiro

Rinaldo Luiz Caraciolo Ferreira

Universidade Federal Rural de Pernambuco  
Departamento de Estatística e Informática  
Mestrado em Biometria e Estatística Aplicada

Recife

2009

## FICHA CATALOGRÁFICA

L732m Lima Filho, Luiz Medeiros de Araujo  
Modelos simétricos transformados não - lineares com  
diferentes distribuições dos erros : aplicações em ciências  
florestais / Luiz Medeiros de Araújo Lima Filho. - 2009.  
90 f. : il.

Orientador : José Antônio Aleixo da Silva  
Dissertação (Mestrado em Biometria e Estatística Apli -  
cada) - Universidade Federal Rural de Pernambuco. Depar -  
tamento de Estatística e Informática.  
Inclui anexo, apêndice e bibliografia.

CDD 519.536

1. Modelos simétricos tranformados
  2. Modelos de crescimento
  3. Modelos não - lineares
  4. Chapman - Richaros
- I. Silva, José Antônio Aleixo da
  - II. Título

**UNIVERSIDADE FEDERAL RURAL DE PERNAMBUCO**  
**PRÓ-REITORIA DE PESQUISA E PÓS-GRADUAÇÃO**  
**PROGRAMA DE PÓS-GRADUAÇÃO EM BIOMETRIA E ESTATÍSTICA APLICADA**

Dissertação sob o título *Modelos simétricos transformados não-lineares com diferentes distribuições dos erros: aplicações em ciências florestais*, apresentada por Luiz Medeiros de Araujo Lima Filho e aprovada em 13 de fevereiro de 2009, em Recife, Estado de Pernambuco, pela banca examinadora constituída por:

Orientador:

---

Prof. Dr. José Antônio Aleixo da Silva  
Departamento de Ciências Florestais  
Universidade Federal Rural de Pernambuco

Banca Examinadora:

---

Prof. Dra. Cláudia Regina Oliveira de Paiva  
Lima  
Departamento de Estatística / CCEN  
Universidade Federal de Pernambuco

---

Prof. Dr. Francisco José de Azevêdo  
Cysneiros  
Departamento de Estatística / CCEN  
Universidade Federal de Pernambuco

---

Prof. Dra. Tatijana Stosic  
Pós-Graduação em Biometria e Estatística  
Aplicada  
Universidade Federal Rural de Pernambuco

## **Agradecimentos**

A Deus por me iluminar e fortalecer em todos os momentos da minha vida.

Aos meus pais pela criação maravilhosa, por todo amor dispensado e por estarem sempre presentes na minha vida, me orientando e apoiando.

À minha avó Reynildes pelo amor incondicional e por estar ao meu lado em todos os momentos.

À minha esposa Tarciana, por todo seu amor, paciência, carinho e companheirismo.

À minha irmã Pollyanna por toda força e carinho que tanto me ajudou.

A Agamenon, Vilma, Dayseane e Keyla por me acolherem, pelo crédito e por tudo que fizeram por mim.

A Kamyll e Bruna por todo o carinho com que sempre me receberam.

Ao professor Aleixo pela sua dedicada orientação, competência e atenção dispensada no desenvolvimento desta dissertação.

Ao professor Gauss Cordeiro pela confiança em mim depositada, pelo estímulo, dedicação e orientação incondicional.

Aos professores do Programa de Mestrado em Biometria e Estatística Aplicada/UFRPE pelos conhecimentos transmitidos.

Aos meus colegas de turma pelas experiências trocadas e pelas lições apreendidas com cada um, em especial a Andrea e Edileide.

Aos meus colegas de trabalho que me apoiaram e me ajudaram, em especial a Francisco Agostinho por toda a confiança em mim depositada.

# Resumo

Historicamente, a madeira do *Eucalyptus* é usada para os mais variados fins, tais como; lenha, carvão vegetal, celulose, dormentes ferroviários, postes para eletrificação, casca para curtir couro, óleos essenciais, construção civil, etc. O Pólo Gesseiro do Araripe em Pernambuco é um grande consumidor de madeira para produção de gesso. Devido à grande necessidade de se buscar uma alternativa econômica e ambiental para a região é de interesse obter uma produção sustentável para o *Eucalyptus*, uma vez que esta é uma árvore de rápido crescimento e grande versatilidade. No planejamento do manejo florestal sustentado uma variável é de extrema importância: o crescimento. Sua modelagem é fundamental na prognose da produtividade, qualidade do local e dinâmica de populações. Geralmente, as curvas de crescimento são estudadas por meio de modelos não-lineares desenvolvidos empiricamente para relacionar, por exemplo, altura e idade. Um modelo não-linear bastante utilizado na prática para modelar curvas de crescimento é o modelo de Chapman-Richards. Em estudos deste tipo, em geral, assume-se que os erros seguem distribuição normal. Contudo, a modelagem sob a suposição de erros com distribuição normal é bastante sensível a valores atípicos que por ventura possam ocorrer, podendo distorcer as estimativas dos parâmetros. Para corrigir esse problema Cordeiro et al. (2009) desenvolveram uma nova classe de modelos simétricos transformados considerando para os erros distribuições contínuas simétricas com caudas mais pesadas do que a distribuição normal e permitindo uma possível estrutura não-linear para a média. Dessa forma, com a expectativa de obter melhores estimativas de crescimento de *Eucalyptus*, aplicaram-se ao modelo de Chapman-Richards as seguintes distribuições dos erros: normal, t de Student, Cauchy, exponencial potência, logística I e logística II que apresentou a distribuição t de Student com 2 graus de liberdade com melhores estimativas de crescimento em altura e circunferência de *Eucalyptus* no Pólo Gesseiro de Pernambuco.

# Abstract

Historically, the wood of the eucalyptus is used for the most varied applications, such as; firewood, charcoal, cellulose, railway sleepers, posts for electrification, bark to tan leather, essential oils, civil construction, etc. The Gypsum Pole of Araripe in Pernambuco is a great firewood consumer for the gypsum production. Due to great need to find economical and environmental alternatives for the area, the sustainable production of eucalyptus that is a fast growth tree with great versatility has an important role. In the planning of the sustainable forest management there is a variable of extreme importance: the growth. To model the growth is fundamental in the prognosis of the productivity, site quality and dynamics of populations. Usually, the growth curves are fitted through nonlinear models developed empirically to relate, for instance, height and age. The Chapman-Richards model is a nonlinear model frequently used to model forest growth. In studies of this type, in general, it is assumed that the errors follow approximately the normal distribution. However, to model the growth assuming that the errors have a normal distribution is quite sensitive to atypical values that can happen, and generate bad estimates of the parameters. To correct that problem a new class of transformed symmetrical models was developed considering for the errors symmetrical continuous distributions with heavier tails than the normal distribution and allowing a possible nonlinear structure for the mean. With the expectation of obtaining better estimates of eucalyptus growth, it was applied to the Chapman-Richards model the following distributions of the errors: normal, t of Student, Cauchy, exponential potency, logistics I and logistics II. The t distribution of Student with 2 degrees of freedom was the most efficient to estimate height and circumference growth of eucalyptus in the Gypsum Pole of Pernambuco.

# Lista de Figuras

1	Densidade da distribuição normal padrão . . . . .	p. 20
2	Densidades da distribuição t de Student (3) e distribuição normal padrão. .	p. 24
3	Densidades da distribuição t de Student generalizada ( $s = 2, r = 3$ ) e distribuição normal padrão. . . . .	p. 26
4	Densidades da distribuição exponencial potência ( $l = 0, 3$ ) e distribuição Normal padrão. . . . .	p. 29
5	Densidades da distribuição Cauchy padrão e distribuição normal padrão. .	p. 32
6	Densidades da distribuição logística I padrão e distribuição normal padrão.	p. 34
7	Densidades da distribuição logística II padrão e distribuição normal padrão.	p. 35
8	Densidades da distribuição Kotz ( $N = 3, r = 1$ ) e distribuição normal padrão. . . . .	p. 36
9	Densidades da distribuição exponencial dupla padrão e distribuição normal padrão. . . . .	p. 39
10	Gráfico de linha das repetições ao longo do tempo . . . . .	p. 59
11	Graus de Liberdade da $t_2$ transformada versus AIC . . . . .	p. 61
12	<i>Erros percentuais para o modelo normal</i> . . . . .	p. 65
13	<i>Erros percentuais para o modelo <math>t_2</math> transformado</i> . . . . .	p. 65
14	<i>Erros percentuais para o modelo normal</i> . . . . .	p. 66
15	<i>Erros percentuais para o modelo <math>t_2</math> transformado na escala original</i> . . . .	p. 66
16	Gráfico de linha das repetições ao longo do tempo . . . . .	p. 67
17	Graus de liberdade da $t_2$ transformada versus AIC . . . . .	p. 68
18	<i>Erros percentuais para o modelo normal</i> . . . . .	p. 71
19	<i>Erros percentuais para o modelo <math>t_2</math> transformado</i> . . . . .	p. 71

20	<i>Erros percentuais para o modelo normal . . . . .</i>	p.72
21	<i>Erros percentuais para o modelo <math>t_2</math> transformado . . . . .</i>	p.72

# Lista de Tabelas

1	Expressões para $D^{(\lambda)}$ , $a$ e $b$ . . . . .	p. 52
2	Idade dos clones de <i>Eucalyptus</i> com relação a altura. . . . .	p. 56
3	Idade dos clones de <i>Eucalyptus</i> com relação à CAP. . . . .	p. 57
4	Estatísticas descritivas das alturas dos <i>Eucalyptus</i> . . . . .	p. 60
5	Estimativas dos parâmetros dos modelos transformados e não transformados. . . . .	p. 62
6	Estatísticas para seleção dos modelos. . . . .	p. 63
7	Comparação dos erros para os modelos ajustados. . . . .	p. 64
8	Comparação dos erros para os modelos ajustados. . . . .	p. 66
9	Estatísticas descritivas das circunferências dos <i>Eucalyptus</i> . . . . .	p. 68
10	Estimativas dos parâmetros dos modelos transformados e não transformados. . . . .	p. 69
11	Estatísticas para seleção dos modelos. . . . .	p. 70
12	Comparação dos erros para os modelos ajustados. . . . .	p. 70
13	Comparação dos erros para os modelos ajustados. . . . .	p. 72

# Sumário

<b>1</b>	<b>Introdução</b>	p. 11
<b>2</b>	<b>Distribuições Simétricas</b>	p. 14
2.1	Definição . . . . .	p. 15
2.2	Distribuição Normal . . . . .	p. 16
2.3	Distribuição t de Student . . . . .	p. 20
2.3.1	t de Student . . . . .	p. 21
2.3.2	t de Student Generalizada . . . . .	p. 24
2.4	Distribuição Exponencial Potência . . . . .	p. 27
2.5	Distribuição Cauchy . . . . .	p. 29
2.6	Distribuição Logística . . . . .	p. 32
2.6.1	Logística I . . . . .	p. 33
2.6.2	Logística II . . . . .	p. 34
2.7	Distribuição de Kotz . . . . .	p. 35
2.8	Distribuição de Laplace ou Exponencial Dupla . . . . .	p. 36
<b>3</b>	<b>Modelos Não-Lineares</b>	p. 40
3.1	Definição . . . . .	p. 41
3.1.1	Estimação . . . . .	p. 42
3.1.2	Resíduos . . . . .	p. 43
3.2	Modelos de Crescimento . . . . .	p. 43
3.2.1	Modelo de Chapman-Richards . . . . .	p. 46

<b>4 Modelos Simétricos Transformados</b>	p. 48
4.1 Definição . . . . .	p. 48
4.2 Inferência . . . . .	p. 50
<b>5 Material e Métodos</b>	p. 55
<b>6 Resultados</b>	p. 59
6.1 Modelo de Chapman-Richards para a variável altura . . . . .	p. 59
6.2 Modelo de Chapman-Richards para a variável circunferência . . . . .	p. 67
<b>7 Conclusão</b>	p. 73
<b>Referências</b>	p. 74
<b>Apêndice A - Algoritmos e programas</b>	p. 81

# 1 Introdução

A disseminação de sementes de *Eucalyptus* no mundo começou no início do século XIX na América do Sul. No Brasil, as primeiras mudas chegaram em 1824. O plantio experimental foi iniciado pelo trabalho pioneiro da Companhia Paulista de Vias Férreas e Fluviais, na pessoa de Navarro de Andrade, no horto florestal, localizado em Jundiá (ANDRADE, 1911; HASSE, 2006). Historicamente, a madeira do *Eucalyptus* é utilizada para os mais variados fins, tais como; lenha, carvão vegetal, celulose, dormentes ferroviários, postes para eletrificação, casca para curtir couro, óleos essenciais e construção civil (ANDRADE, 1911; LIMA, 1993). Hoje, o Brasil se configura como um dos maiores produtores de *Eucalyptus* do mundo.

O Polo Gesseiro do Araripe, localizado na microregião de Araripina, semi-árido Pernambucano, é um grande consumidor de biomassa vegetal que é usada na calcinação da gipsita. Essa microregião abrange 10 municípios e é responsável por 95% do gesso industrializado no Brasil (ALBUQUERQUE, 2002).

O bioma caatinga, no qual está localizado o Pólo Gesseiro do Araripe-PE, vem sofrendo pressão visto que é explorado de forma desordenada. Esse fato se deve, principalmente, a crescente demanda por recursos naturais renováveis, aumentando gradativamente a sua degradação. Uma alternativa econômica e ambiental viável é a implementação e o manejo sustentado de povoamentos florestais nativos ou o reflorestamento com florestas de rápido crescimento, com destaque para o *Eucalyptus* por sua elevada taxa de crescimento, a facilidade de reprodução, a rusticidade e o altíssimo nível de melhoramento genético em produtividade e qualidade da madeira.

Em face desse fato, torna-se de interesse quantificar o crescimento e a produção de florestas, promovendo um planejamento criterioso da produção através da prescrição de regimes de manejos adequados visando à qualidade do produto final. Sendo assim, pode-se dizer que a predição do crescimento e da produção é parte fundamental do processo de planejamento dos povoamentos florestais. Segundo Scolforo (1994), o conhecimento do crescimento e da produção presente e futura de árvores em povoamentos florestais

é elemento fundamental no manejo florestal sustentável, sendo necessário possuir como uma das fontes de informações mais importante a existência de relações quantitativas e modelos matemáticos que sejam consistentes e, numericamente, compatíveis para a predição do desenvolvimento do povoamento em qualquer idade.

As técnicas para prognosticar a dinâmica de um povoamento florestal são conhecidas como modelos de crescimento e produção (AVERY e BURKHART, 1994). Tais modelos são sínteses de observações e/ou fenômenos biológicos aplicados nas condições sob as quais foram coletados os dados. Na Engenharia Florestal, os modelos são imprescindíveis na prognose dos futuros recursos florestais, na escolha de metodologias silviculturais e apoio à tomada de decisões no manejo e na política florestal (SPATHELF e NUTTO, 2000).

Geralmente, as curvas de crescimento são estudadas por meio de modelos não-lineares. Com o desenvolvimento computacional na década de 70 alguns modelos que exigiam a utilização de processos iterativos para a estimação dos parâmetros começaram a ser mais utilizados. O modelo normal não-linear, por exemplo, que assume uma estrutura não-linear para os parâmetros teve um grande avanço. Esses modelos não-lineares têm tido uma ampla variedade de aplicações, com a vantagem de obter parâmetros que são facilmente interpretáveis. Um modelo não-linear bastante utilizado para descrever tais fenômenos em ciências florestais é o modelo de Chapman-Richards.

Durante muitos anos, modelos supondo erros normais foram utilizados para descrever a maioria dos fenômenos aleatórios, dado que a suposição de normalidade sempre foi muito atrativa para os erros dos modelos de regressão com resposta contínua. Contudo, com o passar do tempo, verificou-se que as estimativas obtidas para os coeficientes dos modelos normais se mostravam sensíveis à presença de observações extremas. Alternativas à suposição de erros normais têm sido proposta na literatura. A fim de tentar reduzir a influência dessas observações extremas nas estimativas dos coeficientes, uma das alternativas é assumir para os erros distribuições com caudas mais pesadas do que a normal. Cordeiro e Andrade (2009) propõem uma nova classe de modelos simétricos transformados com distribuições simétricas para a variável resposta e uma possível função de ligação não-linear para a resposta média. Esta classe de modelos estende os modelos de Box e Cox (1964) considerando distribuições contínuas simétricas com caudas mais leves e mais pesadas do que a normal.

Considerando a grande utilidade no uso de modelos não-lineares na análise do crescimento e produção de florestas e, a real necessidade de considerar para os erros distribuições com caudas mais leves e mais pesadas do que a normal os modelos simétri-

cos tende a dar uma grande contribuição para trabalhos futuros no manejo sustentado de povoamentos florestais.

O objetivo deste trabalho consiste em estimar a altura e circunferência dos *Eucalyptus*, através de modelos simétricos transformados não lineares baseados em erros com distribuições mais robustas que a distribuição normal.

A presente dissertação de mestrado está dividida em 7 Capítulos. No Capítulo 2, descreve-se a classe de distribuições simétricas e alguns resultados teóricos de algumas distribuições simétricas com suporte na reta real. O Capítulo 3 apresenta os modelos não-lineares e a definição do modelo de crescimento de Chapman-Richards. O Capítulo 4 aborda os modelos simétricos transformados e apresenta inferências estatísticas em modelos não-lineares com erros distribuídos na família simétrica. No Capítulo 5 é apresentado a descrição do experimento. Os resultados das aplicações são encontrados no Capítulo 6 e a conclusão encontra-se no capítulo 7.

Para realização desse trabalho foram utilizados os softwares R, SAS e a linguagem de tipografia  $\text{\LaTeX}$ . O software R é interpretado como uma linguagem computacional designada para análise de dados estatísticos que se caracteriza pelo compromisso entre a flexibilidade oferecida pelas linguagens compiladas, tais como C e C++, e a conveniência de softwares estatísticos tradicionais. Inclui uma ampla variedade de métodos estatísticos tradicionais e modernos, com a vantagem de ser gratuito. O SAS é um sistema integrado de aplicações para análise de dados que consiste de: recuperação de dados, gerenciamento de arquivos, análise estatística, acesso a banco de dados e geração de gráficos. Para elaboração do texto foi usada a linguagem de tipografia  $\text{\LaTeX}$  amplamente utilizada em trabalhos científicos. Esta linguagem se destaca pela flexibilidade e qualidade de apresentação.

## 2 Distribuições Simétricas

Usualmente em inferência estatística, considera-se a hipótese de que os erros aleatórios apresentam distribuição normal. Contudo, em estudos de fenômenos naturais essa hipótese não é sempre adequada. Uma solução é considerar modelos distribucionais mais adequados do que o Gaussiano para este tipo de problema.

Nesta dissertação será abordada a família de distribuições simétricas com suporte na reta real. Esta família gera uma classe geral de distribuições com a mesma simetria que a distribuição normal padrão. Pertencem a esta classe as seguintes distribuições: t de Student, Cauchy, t de Student generalizada, Kotz, Kotz generalizada, Exponencial potência, Logística tipo I, Logística tipo II, entre outras. Para maiores detalhes sobre a família de distribuições simétricas em modelo de regressão linear ver Cysneiros et al. (2005).

Nos últimos anos, esta classe de distribuições tem recebido crescente atenção na literatura. Em Chmielewski (1981), Fang et al. (1990), Fang e Zhang (1990), Fang e Anderson (1990) e Gupta e Varga (1993) pode ser encontrada uma revisão das áreas em que são aplicadas as distribuições simétricas. Lange et al. (1989) propuseram o modelo baseado na suposição de erros t de Student. Little (1988) e Yamaguchi (1990) utilizaram o modelo baseado na suposição de erros com distribuição normal contaminada. Nesses modelos adicionam-se parâmetros os quais permitem ajustar a curtose da distribuição dos dados. Taylor (1992) propôs o ajuste de um modelo de regressão linear supondo erros com distribuição exponencial potência com um parâmetro extra de forma. Arellano-Valle (1994) apresentou vários resultados sobre as propriedades, inferência e aplicações a modelos de regressão em distribuições elípticas. Ferrari e Uribe-Opazo (2001) estendem esses resultados para modelos de regressão linear simétricos. Cordeiro et al. (2000) obtiveram a correção do viés do estimador de máxima verossimilhança na classe de modelos não-lineares simétricos. Cordeiro (2004) desenvolveu uma correção de Bartlett para os modelos de regressão não-lineares simétricos generalizando os resultados de Ferrari e Uribe-Opazo (2001). Galea et. al. (2005) apresentaram alguns resultados sobre modelagem, em particular sobre o desenvolvimento da análise inferencial e de diagnóstico na classe não-lineares

com erros simétricos independentes. Cordeiro e Andrade (2009) introduziram uma nova classe de modelos simétricos transformados para estender o modelo de Box e Cox (1964).

## 2.1 Definição

Defini-se a seguir a classe simétrica univariada e alguns resultados teóricos.

Diz-se que a variável aleatória  $Y$  tem distribuição simétrica, com suporte em  $\mathfrak{R}$ , com parâmetros de locação  $\mu \in \mathfrak{R}$  e de escala  $\phi > 0$ , se sua função de densidade de probabilidade é dada por

$$f(y; \mu, \phi) = \frac{1}{\sqrt{\phi}} g \left\{ \frac{(y - \mu)^2}{\phi} \right\}, \quad y \in \mathfrak{R}, \quad (2.1)$$

para alguma função  $g(\cdot)$  denominada função geradora de densidade, em que  $u = \frac{(y - \mu)^2}{\phi}$ , com  $g(u) > 0$ , para  $u > 0$  e  $\int_0^\infty u^{-1/2} g(u) du = 1$ . Essa condição é necessária para que  $f(y; \mu, \phi)$  seja uma função densidade de probabilidade. Assim, denota-se por  $Y \sim S(\mu, \phi)$  e denomina-se de variável aleatória simétrica.

Algumas propriedades da distribuição normal podem ser estendidas para a classe de distribuições simétricas.

i) Se  $Y \sim S(\mu, \phi)$ , então a função característica de  $Y$  é dada por  $\psi(iZ) = e^{it\mu} \varphi(t^2\phi)$ ,  $t \in \mathfrak{R}$  para alguma função  $\varphi$ , com  $\varphi(u) \in \mathfrak{R}$  para  $u > 0$ . Quando existem,  $E(Y_i) = \mu_i$  e  $Var(Y_i) = \xi\phi$ , em que  $\xi > 0$  é uma constante dada por  $\xi = -2\varphi'(0)$ , com  $\varphi'(0) = \frac{\partial \varphi(u)}{\partial u} \Big|_{u=0}$  e que não depende dos parâmetros  $\mu$  e  $\phi$  (FANG et al., 1990). Se  $u^{-\frac{1}{2}(k+1)}g(u)$  for integrável então o  $k$ -ésimo momento de  $Y$  existe.

ii) Se  $Y \sim S(\mu, \phi)$ , então  $a + bY \sim S(a + b\mu, b^2\phi)$ , em que  $a, b \in \mathfrak{R}$  com  $b \neq 0$ , ou seja, a distribuição de qualquer transformação linear de uma variável aleatória com distribuição simétrica é também simétrica.

iii) Berkane e Bentler (1986) considerando uma distribuição simétrica padrão  $Z \sim S(0, 1)$  e a existência de seus momentos mostraram que a função característica de  $Z$  pode ser expandida como

$$\psi_Z(t) = \sum_{k=0}^{\infty} i^k \mu'_k \frac{t^k}{k!}, \quad (2.2)$$

em que  $\mu'_k = E(Z^k) = i^k \psi_z^{(k)}(0)$ , com  $\psi_z^{(k)}(0)$  denotando a  $k$ -ésima derivada de  $\psi_z^{(k)}(t)$  avaliada em  $t = 0$ . Então,

$$\mu'_k = \begin{cases} 0 & k \text{ ímpar} \\ \frac{(2m)!}{2^m m!} (\mu'_2)^m \{k(m) + 1\} & k = 2m, m = 1, 2, \dots \end{cases}$$

sendo

$$k(m) = \frac{\varphi^m(0)}{\{\varphi^{(1)}(0)\}^m} - 1, \quad (2.3)$$

em que  $\varphi^{(r)}(0)$  é a  $r$ -ésima derivada da função  $\varphi$ , avaliada em zero. Os coeficientes  $k(m)$ ,  $m = 1, 2, \dots$ , são conhecidos como parâmetros de momentos e generalizam o coeficiente de curtose  $\gamma_2 = 3\{k(2) + 1\}$  de uma distribuição  $S(\mu, \phi)$  (MUIRHEAD, 1982).

Cambanis et al. (1981) observaram que a família de distribuições simétricas coincide com a classe de distribuições elípticas univariadas. Na última década surgiram contribuições importantes a partir dos trabalhos de Kelker (1970) para as distribuições elípticas univariadas e multivariadas. Pode-se citar alguns trabalhos que discutem alguns aspectos dessas distribuições, tais como Berkane e Bentler (1986), Muirhead (1980 e 1982), Rao (1990), Cambanis et al. (1981), Anderson e Fang (1987) e Cysneiros et. al. (2005).

A seguir, são apresentadas algumas distribuições simétricas com suporte na reta real para  $Y \sim S(\mu, \phi)$  em que  $u = (y - \mu)^2 / \phi$ .

## 2.2 Distribuição Normal

A distribuição normal, conhecida também como distribuição de Gauss ou Gaussiana, é a distribuição mais utilizada, tanto por constituir uma aproximação razoável para um grande número de distribuições, bem como devido a todo desenvolvimento teórico e aplicado estabelecido no decorrer dos anos.

Os primeiros trabalhos consideraram a distribuição somente como uma aproximação conveniente para distribuição binomial. O reconhecimento da sua importância teórica foi propagado por Laplace e por Gauss no início do século XIX. A primeira derivação da distribuição normal, como uma aproximação da distribuição binomial, parece ter sido publicada por De Moivre em um folheto, escrito em latim, com data de 12 de novembro de 1733. Em 1738 De Moivre publicou uma tradução desse folheto com alguns acrésci-

mos. Em 1774 Laplace obteve a distribuição normal como aproximação da distribuição hipergeométrica e quatro anos depois defendeu a tabulação da probabilidade. Em 1809 e 1816, o trabalho de Gauss estabeleceu técnicas baseadas na distribuição normal, as quais tornaram-se métodos padrões adotados durante o século XIX.

A distribuição normal tornou-se amplamente aceita como a base da maioria dos trabalhos estatísticos. Quando não se conhece a distribuição dos dados, na prática, a distribuição normal pode ser frequentemente aplicada aos dados observados, com menor risco de se cometer erros graves na análise. Alguns resultados devidos a Muirhead (1982), Devlin et al. (1976) caracterizaram a distribuição normal, chamada de normal composta, dentro da classe de distribuições simétricas.

A variável aleatória  $Y$ , denotada por  $Y \sim N(\mu, \sigma^2)$ , é normalmente distribuída se sua função de densidade é definida por:

$$f(y; \mu, \sigma^2) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{y-\mu}{\sigma}\right)^2}, \quad (2.4)$$

com  $\mu \in \Re$  e  $\sigma > 0$ .

Sendo  $Y \sim N(\mu, \sigma^2)$  e  $Z = \frac{Y-\mu}{\sigma}$ ,  $Z$  é conhecida como distribuição normal padrão ou normal reduzida  $N(0, 1)$ . A função de densidade da normal padrão é dada por:

$$f(y; \mu, \sigma^2) = \frac{1}{\sqrt{2\pi}} e^{-\frac{y^2}{2}}. \quad (2.5)$$

Para o cálculo das probabilidades é necessário trabalhar com a distribuição acumulada de  $Z$ , usualmente representada por  $\Phi(z)$ , sendo

$$\Phi(z) = P[Z \leq z] = (\sqrt{2\pi})^{-1} \int_{-\infty}^z e^{-y^2/2} dy. \quad (2.6)$$

A distribuição normal possui algumas propriedades.

Se  $X$  e  $Y$  seguem uma distribuição normal, tem-se:

- i) Se  $Y \sim N(\mu, \sigma^2)$ , então,  $a + bY \sim N(a + b\mu, b^2\sigma^2)$ , em que  $a, b \in \Re$  com  $b \neq 0$ .
- ii) Se  $X$  e  $Y$  têm distribuições normais, então a soma  $U = X + Y$ , bem como a diferença  $V = X - Y$  também apresentam distribuições normais.
- iii) Se  $X$  e  $Y$  são independentes, então  $U$  e  $V$ , dadas acima, também serão independentes.

Seguem, a seguir, algumas relações da distribuição normal com outras distribuições.

i)  $R \sim \text{Rayleigh}(\sigma^2)$  é a distribuição de Rayleigh se  $R = \sqrt{X^2 + Y^2}$  em que  $X \sim N(0, \sigma^2)$  e  $Y \sim N(0, \sigma^2)$  são duas distribuições normais independentes.

ii)  $Y \sim \chi_\nu^2$  é a distribuição qui-quadrado com  $\nu$  graus de liberdade se  $Y = \sum_{k=1}^{\nu} X_k^2$  em que  $X_k \sim N(0, 1)$ , para  $k = 1, \dots, \nu$ , são distribuições normais padrões independentes.

iii)  $Z \sim \text{Cauchy}(0, 1)$  se  $Z = X/Y$  para  $X \sim N(0, 1)$  e  $Y \sim N(0, 1)$  distribuições normais padrões independentes.

iv) Se  $Z = e^X$  e  $X \sim N(\mu, \sigma^2)$ , então  $Z$  é a distribuição lognormal, isto é,  $Z \sim LN(\mu, \sigma^2)$

Na prática, a lognormal de dois parâmetros é uma importante distribuição deduzida da normal. Uma propriedade procedente da distribuição normal é que se os  $y_i$ 's são i.i.d. com  $y_i \sim LN(\mu, \sigma^2)$ , então a média geométrica  $\bar{z} = (\pi z_i)^{1/n}$  segue também uma lognormal, isto é,  $LN(\mu, \sigma^2/n)$ .

Os estimadores de máxima verossimilhança dos parâmetros da distribuição normal são dados por

$$\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i \quad (2.7)$$

e

$$s^2 = \frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})^2, \quad (2.8)$$

em que  $\bar{x}$  é o estimador da verdadeira média  $\mu$  e  $s^2$  é o estimador não viesado da verdadeira variância  $\sigma^2$ . Essas estimativas são conjuntamente suficientes para  $\mu$  e  $\sigma^2$ . Para a maioria dos estudos  $\bar{x}$  é o melhor estimador para  $\mu$ , para  $\sigma$  conhecido ou não.

Uma propriedade da distribuição normal é a independência das estatísticas amostrais  $\bar{x}$  e  $s^2$ . Para uma população de  $n$  eventos de uma distribuição normal,  $\bar{x}$  tem distribuição  $N(\mu, \sigma^2/n)$  e  $(n-1)s^2/\sigma^2$  tem distribuição aproximada qui-quadrada com  $n-1$  graus de liberdade. Usando a relação

$$\sum_{i=1}^n \left( \frac{x_i - \mu}{\sigma} \right)^2 = \frac{(n-1)s^2}{\sigma^2} + \left( \frac{\bar{x} - \mu}{\sigma/\sqrt{n}} \right)^2, \quad (2.9)$$

e calculando a função característica conjunta para a variável  $(n-1)s^2/\sigma^2$  e  $(\sqrt{n}(\bar{x}-\mu)/\sigma)^2$  é possível verificar a independência entre essas variáveis e portanto a independência entre  $\bar{x}$  e  $s^2$ .

Diz-se que  $Y \sim S(\mu, \phi)$  tem distribuição normal se sua função geradora de densidade  $g(\cdot)$  é da forma

$$g(u) = \frac{1}{\sqrt{2\phi}} \exp(-u/2), \quad u > 0, \quad (2.10)$$

então,  $Y$  tem distribuição normal denotada por  $Y \sim N(\mu, \phi)$ , e sua função característica é dada por

$$\psi_Y(t) = e^{it\mu} \exp(-t^2\phi/2), \quad t \in \mathfrak{R}. \quad (2.11)$$

Se  $Y \sim N(\mu, \phi)$ , os momentos centrais de ordem  $r$  são

$$\mu_r = E\{(Y - \mu)^r\} = \begin{cases} 0, & r \text{ ímpar} \\ \sigma^r r! / \{2^{r/2} (r/2)!\}, & r \text{ par} \end{cases}$$

e, portanto, o coeficiente de curtose é  $\gamma_2 = 3$ .

Implementações computacionais do Método de Monte Carlo normalmente precisam simular variáveis aleatórias normais. Uma forma rápida e prática de gerar variáveis normais é a partir da distribuição uniforme utilizando a transformação de Box-Muller. Sejam  $u_1$  e  $u_2$  valores independentes gerados pela distribuição uniforme entre 0 e 1. Então:

$$y_1 = \sqrt{-2 \ln(u_1)} \cdot \cos(2\pi u_2) \quad (2.12)$$

e

$$y_2 = \sqrt{-2 \ln(u_1)} \cdot \sin(2\pi u_2), \quad (2.13)$$

são normais padronizadas independentes.

Como ilustração, a Figura 1 apresenta o gráfico da função de densidade da distribuição normal padrão.

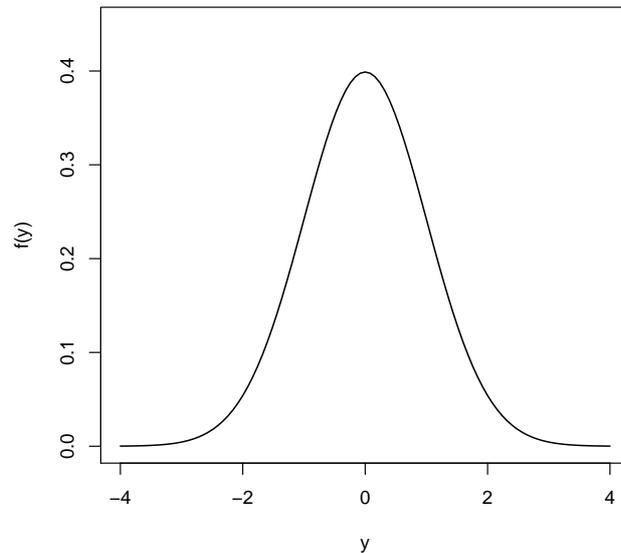


Figura 1: Densidade da distribuição normal padrão

## 2.3 Distribuição t de Student

A distribuição t de Student surgiu dos trabalhos de William Sealy Gosset que era um químico e matemático inglês. Quando se formou, em 1899, Gosset foi trabalhar para a destilaria Dublin de Arthur Guinness e Son, na qual iria aplicar os seus conhecimentos de estatística para a seleção das melhores espécies de cevada. Um outro funcionário da Guinness tinha publicado um trabalho que continha alguns segredos da Cervejeira. Para prevenir futuras revelações dos “segredos” da marca, a Guinness proibiu que os seus empregados pudessem publicar quaisquer trabalhos independentemente do conteúdo. Então, Gosset usou o pseudônimo “Student” para as suas publicações. O seu feito mais conhecido foi a distribuição t de Student, também conhecida como a distribuição t-Gosset.

A maior aplicação da distribuição t de Student é na construção de testes de hipóteses e intervalos de confiança relacionados a valores esperados da distribuição. A distribuição t de Student também é utilizada para modelar o comportamento de dados que provêm de uma distribuição com caudas mais pesadas que a distribuição normal, permitindo reduzir a influência de observações aberrantes. Essa distribuição permite ajustar a curtose dos dados através dos graus de liberdade “ $\nu$ ”. Para valores de  $n$  maiores do que 30 as diferenças entre as distribuições normal e t de Student são pequenas, ou seja com o aumento de  $n$  a distribuição t de Student se aproxima da distribuição normal.

Algumas referências relevantes são Praetz (1972), Praetz e Wilson (1978), Blattberg e Gonedes (1974), McLeay (1986) e Taylor e Kingsman (1979). Neste último artigo, uma  $t$  de Student com 3 parâmetros é ajustada para descrever mudanças em preços de commodity. Esta distribuição também pode ser aplicada como uma distribuição a priori para valores esperados de variáveis multinormais (ANGERS, 1992).

### 2.3.1 $t$ de Student

A distribuição  $t$  de Student com  $\nu$  graus de liberdade foi originada através da razão de duas variáveis aleatórias independentes. Suponha que  $Z$  tenha distribuição normal com média 0 e variância 1 e que  $V$  tenha distribuição qui-quadrado com  $\nu$  graus de liberdade. Suponha, também, que as variáveis  $Z$  e  $V$  sejam independentes. Assim,

$$t = \frac{Z}{\sqrt{V/\nu}}, \quad (2.14)$$

tem distribuição  $t$  de Student com  $\nu$  graus de liberdade.

A função densidade de probabilidade da  $t$  de Student é definida por:

$$f(y) = \frac{\Gamma\left(\frac{\nu+1}{2}\right)}{\sqrt{\nu\pi}\Gamma\left(\frac{\nu}{2}\right)} \left(1 + \frac{y^2}{\nu}\right)^{-\left(\frac{\nu+1}{2}\right)}, \quad (2.15)$$

em que  $\Gamma$  é a função gamma. Usando-se a função beta,  $B$ , a função densidade de probabilidade pode ser escrita como:

$$f(y) = \frac{1}{\sqrt{\nu}B\left(\frac{1}{2}, \frac{\nu}{2}\right)} \left(1 + \frac{y^2}{\nu}\right)^{-\left(\frac{\nu+1}{2}\right)}. \quad (2.16)$$

A primeira tabela publicada, apresentada por “student” (1908), apresentava os valores de  $P[z_\nu \leq z]$ , com  $z_\nu = t_\nu/\sqrt{\nu+1}$ . Após alguns anos, em 1925, o mesmo autor publicou a tabela para valores de  $P[t_\nu \leq t]$  com quatro casas decimais.

Após algumas álgebras, a função acumulada da distribuição  $t$  de Student é definida por

$$F(y) = \begin{cases} \frac{1}{2} - \frac{1}{2}I_{\frac{y^2}{\nu+y^2}}\left(\frac{1}{2}, \frac{\nu}{2}\right) & \text{para } -\infty < y < 0 \\ \frac{1}{2} + \frac{1}{2}I_{\frac{y^2}{\nu+y^2}}\left(\frac{1}{2}, \frac{\nu}{2}\right) & \text{para } 0 < y < \infty \end{cases}$$

Para determinar intervalos de confiança e testes de hipóteses usando a distribuição t de Student definimos a quantidade  $t_{\alpha,\nu}$  para

$$F(t_{\alpha,\nu}) = \int_{-\infty}^{t_{\alpha,\nu}} f(y, \nu) dt = 1 - \alpha, \quad (2.17)$$

isto é,  $\alpha$  é a probabilidade de uma variável com distribuição t de Student com  $\nu$  graus de liberdade exceder  $t_{\alpha,\nu}$ . Devido a simetria da distribuição t em torno do zero, temos que  $t_{\alpha,\nu} = -t_{1-\alpha,\nu}$ .

Para uma amostra normal o intervalo com  $(1 - \alpha)\%$  de confiança para  $\mu$  é dado por

$$\bar{x} - \frac{s}{\sqrt{n}} t_{\alpha/2,\nu-1} \leq \mu \leq \bar{x} + \frac{s}{\sqrt{n}} t_{\alpha/2,\nu-1}. \quad (2.18)$$

Nas situações em que  $\sigma^2$  é conhecido não será utilizado a distribuição t de Student. Nesse caso, a distribuição mais apropriada é a normal.

A estatística  $t$  pode ser utilizada para testar hipóteses sobre a média de uma população normal com variância desconhecida. Considere para uma amostra as seguintes hipóteses a serem testadas  $H_0 : \mu = \mu_0$  e  $H_1 : \mu \neq \mu_0$ . Assim, a estatística de teste é dada por  $t = \frac{\bar{x} - \mu_0}{s/\sqrt{\nu}}$ . Rejeita-se a hipótese nula se  $|t| > t_{\alpha/2,\nu-1}$ . No caso de duas amostras as hipóteses testadas são  $H_0 : \mu_x = \mu_y$  e  $H_1 : \mu_x \neq \mu_y$ . Para duas amostras, a estatística do teste é expressa como  $t = (\bar{x} - \bar{y})/s\sqrt{\frac{1}{m} + \frac{1}{\nu}}$  e a hipótese nula será rejeitada se  $t$  exceder  $t_{\alpha/2,\nu+m-2}$ .

A variável aleatória  $Y \sim S(\mu, \phi)$  tem distribuição t de Student se sua função geradora de densidade  $g(\cdot)$  é da forma

$$g(u) = \frac{\nu^{\nu/2}}{B(1/2, \nu/2)} (\nu + u)^{-\frac{\nu+1}{2}}, \quad \nu > 0, u > 0, \quad (2.19)$$

em que  $B(\cdot, \cdot)$  é a função Beta. Assim,  $Y$  é denotada por  $Y \sim t(\mu, \phi, \nu)$ . A função característica pode ser encontrada em Fang et al. (1990). Utilizando algumas propriedades tem-se que se  $Y = V^{1/2}Z$ , em que  $V \sim GI(\nu/2, \nu/2)$  é uma gama inversa,  $\nu > 0$  e  $Z \sim N(0, 1)$  sendo independentes, então  $Y \sim t(0, 1, \nu)$ .

Se  $Y \sim t(0, 1, \nu)$  tem-se que:

i) Se  $\nu > r$ , seus momentos de ordem  $r$  existem e são definidos por

$$E(Y^r) = \begin{cases} 0, & r \text{ ímpar} \\ \frac{\nu^{r/2} \Gamma(\frac{r+1}{2}) \Gamma(\frac{\nu-r}{2})}{\Gamma(\frac{1}{2}) \Gamma(\frac{\nu}{2})}, & r \text{ par} \end{cases}$$

em que  $\Gamma(\cdot)$  é a função gama. Assim, tem-se que,  $E(Y) = 0$  para  $\nu > 1$  e  $Var(Y) = \frac{\nu}{\nu-2}$  para  $\nu > 2$ . O momento de ordem  $r$  é infinito se  $r$  for par e  $r \geq \nu$ .

ii) O desvio médio é dado por

$$E(|Y|) = \frac{\nu^{1/2} \Gamma(\frac{\nu-1}{2})}{\Gamma(1/2) \Gamma(\nu/2)}. \quad (2.20)$$

iii) O coeficiente de curtose é  $\gamma_2 = 3 + \frac{6}{\nu-4}$ , para  $\nu > 4$ . Este coeficiente é maior que o coeficiente da distribuição normal.

iv)  $Y^2 \sim F(1, \nu)$ , sendo  $F$  a distribuição  $F$  de Snedecor com 1 e  $\nu$  graus de liberdade.

v) Se  $W = \left(\frac{\nu+1}{\nu+Y^2}\right)$ , então

$$E(Y^{2k} W^l) = \frac{(-\frac{\nu+1}{2})^l B[(2k+1)/2, \{\nu+2(l-k)\}/2]}{\nu^{l-k} B(1/2, \nu/2)}, \quad (2.21)$$

para  $l = 0, 1, 2$  e  $k = 1, 2, \dots$

vi) A função densidade de  $Y$  tem pontos de inflexão em  $\{-(\nu/(\nu+2))^{1/2}, (\nu/(\nu+2))^{1/2}\}$ .

vii) Segundo Manoukian (1985), a variável aleatória  $U = (1 + \nu/Y^2)^{-1}$  tem distribuição beta com parâmetros  $a = 1/2$  e  $b = \nu/2$ .

viii)  $Y/V = \nu \sim N(0, \nu)$ .

ix)  $V/Y = Y \sim GI((\nu+1)/2, (\nu+y^2)/2)$ .

De acordo com as propriedades citadas, pode-se verificar que a distribuição  $t(\mu, \phi, \nu)$  tende a uma distribuição normal  $N(\mu, \phi)$  quando  $\nu \rightarrow \infty$ . Quando  $\nu = 1$  a distribuição se reduz a distribuição de Cauchy com parâmetros  $\mu$  e  $\phi$ .

Para gerar números aleatórios da distribuição  $t$  de Student é necessário utilizar a relação:

$$t = \frac{z}{\sqrt{\chi_\nu^2/\nu}}, \quad (2.22)$$

em que  $z$  e  $\chi_\nu^2$  são números aleatórios da distribuição normal padrão e qui-quadrada com

$\nu$  graus de liberdade, respectivamente.

A Figura 2 a seguir apresenta os gráficos da função de densidade da distribuição t de Student e da função de densidade da distribuição normal padrão.

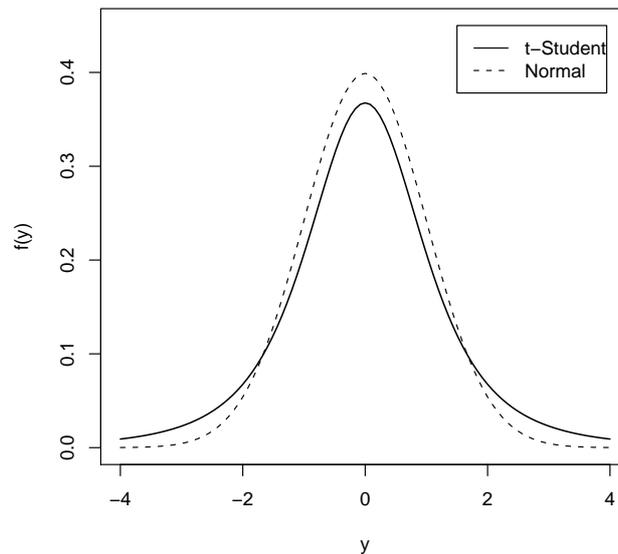


Figura 2: Densidades da distribuição t de Student (3) e distribuição normal padrão.

### 2.3.2 t de Student Generalizada

McDonald e Newey (1988) introduziram a distribuição t de Student generalizada com função densidade dada por

$$f(y; \sigma, p, q) = \frac{p}{2q^{1/p} B(p^{-1}, q) (1 + |y|^p/q)^{(q+p^{-1})}}, \quad -\infty < y < \infty, \quad p, q > 0, \quad (2.23)$$

em que  $\sigma$  é o parâmetro de escala, enquanto que  $p$  e  $q$  são os parâmetros de forma da densidade. Quanto maior os valores de  $p$  e  $q$  mais suaves serão as caudas da distribuição. Uma caso especial dessa distribuição é quando  $p = 2$  e  $q = 2\nu$ . Nesse caso a distribuição se reduz a distribuição t de Student. A densidade generalizada definida acima também inclui, como um caso especial, a densidade exponencial potência,

$$f(y) = \frac{pe^{-|y|^p}}{2\Gamma(p^{-1})}, \quad -\infty < y < \infty, \quad p > 0, \quad (2.24)$$

quando  $q \rightarrow \infty$ .

McDonald (1984) mostrou que a  $t$  de Student generalizada, de fato, é uma mistura da distribuição gamma generalizada e da distribuição exponencial potência de Box e Tiao. McDonald (1991) apresentou valores do coeficiente de curtose, para vários valores de  $p$  e  $q$ .

McDonald e Newey (1988) usaram a  $t$  de Student generalizada para desenvolver estimação parcialmente adaptada de modelos de regressão. Butler et al.(1990) discutiram a estimação robusta de modelos de regressão usando a  $t$  de Student generalizada. Similarmente, estimação parcialmente adaptada de modelos ARMA de séries temporais através da distribuição  $t$  de Student generalizada foi desenvolvida por McDonald (1989).

Se uma variável aleatória  $Y \sim S(\mu, \phi)$  e a função geradora de densidade  $g(\cdot)$  tem a forma

$$g(u) = \frac{s^{r/2}}{B(1/2, r/2)} (s + u)^{-\frac{r+1}{2}}, \quad s, r, u > 0, \quad (2.25)$$

então,  $Y$  possui distribuição  $t$  de Student generalizada com parâmetros  $(\mu, \phi, s, r)$ , ou seja,  $Y \sim tG(\mu, \phi, s, r)$  (DICKY, 1967). Quando  $s = r = \nu$  a distribuição coincide com a  $t \sim (\mu, \phi, \nu)$  e quando  $s = r = 1$  temos a distribuição Cauchy.

Suponha que  $Y/V = \nu \sim N(\mu, \nu\phi)$ , em que  $V$  é uma gama inversa  $V \sim GI(r/2, s/2)$ , independentes com  $s, r > 0$ , podendo não ser inteiro. Dessa forma, tem-se as seguintes propriedades:

i)  $Y \sim tG(\mu, \phi, s, r)$

ii)  $E(Y) = \mu$  para  $r > 1$ ,  $Var(Y) = \{s/(r - 2)\}\phi$  para  $r > 2$  e o coeficiente de curtose  $\gamma_2 = 3 + 6/(r - 4)$  para  $r > 4$ . Vale salientar que o coeficiente de curtose não depende do parâmetro  $s$ . O coeficiente de curtose da  $t$  de Student generalizada é maior do que o coeficiente de curtose da distribuição normal.

iii)  $V/Y \sim GI((r + 1)/2, (s + z^2)/2)$  em que  $z^2 = (y - \mu)^2/\phi$ .

iv)  $U^2 = rZ^2/s \sim F(1, r)$  em que  $Z \sim N(0, 1)$ .

v) Se  $W = (r + 1)/(s + Z^2)$  então,

$$E(Z^{2k}W^l) = \frac{\left(-\frac{r+1}{2}\right)^l B[(2k + 1)/2, (r + 2(l - k))/2]}{s^{(l-k)} B(1/2, r/2)}, \quad (2.26)$$

para  $l = 0, 1, 2$  e  $k = 1, 2, \dots$

vi) Os parâmetros  $s$  e  $r$ , apresentados a seguir, têm uma relação com o parâmetro de curtose e o segundo momento central (JOHNSON e KOTZ, 1970).

$$r = \frac{2(2\gamma_2 - 3)}{\gamma_2 - 3} \quad (2.27)$$

e

$$s = \frac{2\mu_2\gamma_2}{\gamma_2 - 3}. \quad (2.28)$$

vii) O  $l$ -ésimo momento existe se e somente se  $r > l$ .

viii) Para a variável aleatória  $Y = V^{-1/2}Z$ , com  $Z$  e  $V$  variáveis aleatórias independentes, em que  $Z \sim N(0, 1)$  e  $V \sim GI(r/2, s/2)$  então  $Y \sim tG(0, 1, s, r)$ .

Na Figura 3, apresentada a seguir, é possível observar os gráficos das funções de densidades das distribuições t de Student generalizada ( $s = 2, r = 3$ ) e da distribuição normal padrão.

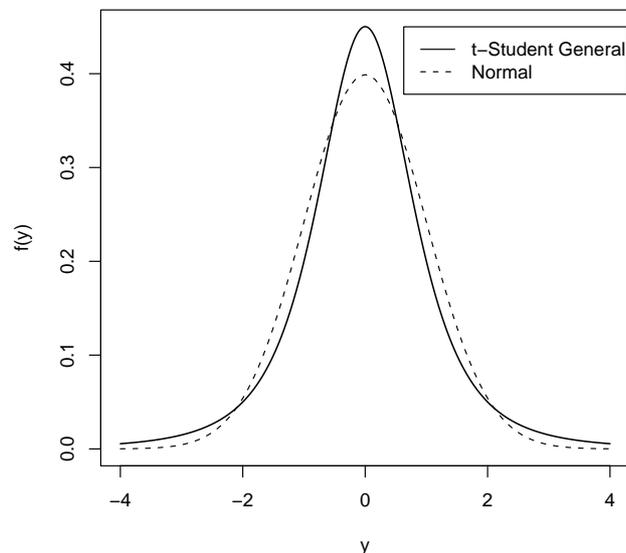


Figura 3: Densidades da distribuição t de Student generalizada ( $s = 2, r = 3$ ) e distribuição normal padrão.

## 2.4 Distribuição Exponencial Potência

A distribuição exponencial potência pode ser considerada como uma distribuição geral para erros aleatórios. A primeira formulação dessa distribuição pode ser atribuída para Subbotin (1923). Subbotin obteve a função densidade de probabilidade dada por

$$f(\varepsilon) = \frac{mh}{2\Gamma(1/m)} e^{-h^m |\varepsilon|^m}, \quad (2.29)$$

em que  $-\infty < \varepsilon < \infty$ ,  $h > 0$  e  $m \geq 1$ .

Seguindo o procedimento introduzido por Pearson (1895), Lunetta (1963) derivou uma diferente parametrização da distribuição exponencial potência resolvendo a seguinte equação diferencial:

$$\frac{\partial \log f}{\partial y} = p \left( \frac{\partial f - \log a}{y - c} \right), \quad (2.30)$$

onde  $a$  e  $c$  são constantes.

Assim, a função de densidade da distribuição exponencial potência é dada por

$$f(y) = \frac{1}{2\sigma p^{1/p} \Gamma(1 + 1/p)} \exp\left(-\frac{|y - \mu|^p}{p\sigma^p}\right), \quad (2.31)$$

em que  $-\infty < y < \infty$ ,  $-\infty < \mu < \infty$ ,  $\sigma > 0$  e  $p > 0$ .

É importante notar que, embora com uma diferente parametrização, a distribuição exponencial potência pode ser tratada também por métodos Bayesianos quando existe o problema para especificar uma distribuição a priori adequada (BOX e TIAO 1992, CHOY e SMITH 1997 e ACHCAR e PEREIRA 1999). Uma distribuição exponencial potência bivariada foi introduzida por De Simone (1968) e Taguchi (1978), ao passo que uma formulação multivariada dessa distribuição pode ser encontrada em Fang et al. (1990) e Krzanowski e Marriott (1994).

Na estimação dos parâmetros da distribuição exponencial potência, assumindo que o parâmetro de forma  $p$  é conhecido, os parâmetros de locação e escala devem ser facilmente obtidos pelo uso do método de estimação de máxima verossimilhança. Entretanto, a estimação do parâmetro de forma é, em geral, um problema. Vários procedimentos interessantes têm sido propostos na literatura, sendo um deles baseado no método de máxima verossimilhança. Em geral, a derivação de estimadores de máxima verossimilhança não

traz, formalmente, muitos problemas e os estimadores têm propriedades adequadas, pelo menos assintoticamente.

Uma variável aleatória  $Y \sim S(\mu, \phi)$  tem distribuição exponencial potência  $Y \sim EP(\mu, \phi, l)$  (BOX e TIAO, 1973) se a sua função geradora de densidade  $g(\cdot)$  é da forma

$$g(u) = C(l) \exp \left\{ -\frac{1}{2} u^{1/(1+l)} \right\}, \quad -1 < l \leq 1, u > 0, \quad (2.32)$$

em que  $C(l)^{-1} = \Gamma \left( 1 + \frac{1+l}{2} \right) 2^{1+(1+l)/2}$ .

Tem-se ainda que

$$E(Y) = \mu \quad (2.33)$$

e

$$Var(Y) = 2^{(1+l)} \left[ \frac{\Gamma \left\{ \frac{3(1+l)}{2} \right\}}{\Gamma \left( \frac{1+l}{2} \right)} \right] \phi. \quad (2.34)$$

O  $l$  pode ser analisado como uma medida de curtose, ou mesmo, como uma medida de não normalidade, pois quando  $l = 0$  obtém-se a distribuição normal. Quando  $l = 1$  tem-se a distribuição exponencial dupla. Se  $l$  tende a  $-1$ , a distribuição tende a uma uniforme  $(\mu - \sqrt{3\phi}, \mu + \sqrt{3\phi})$ . Quando  $l > 0$ , tem-se que  $\gamma_2 > 3$ , ou seja, a distribuição é leptocúrtica. Quando  $l < 0$ , tem-se que  $\gamma_2 < 3$ , ou seja, a distribuição é platicúrtica.

Se  $Y = (2W)^{1/r}V$  em que  $V \sim U(-1, 1)$ ,  $W \sim G(1 + 1/r, 1)$  supostas independentes e  $r = 2/(1 + l)$ , então  $Y \sim EP(0, 1, l)$  (DEVROYE, 1986). Essa relação pode ser utilizada para gerar amostras de uma distribuição  $EP(0, 1, l)$ .

A Figura 4 a seguir apresenta o gráfico da função de densidade da distribuição exponencial potência ( $l = 0, 3$ ) comparado com a função de densidade da distribuição normal padrão.

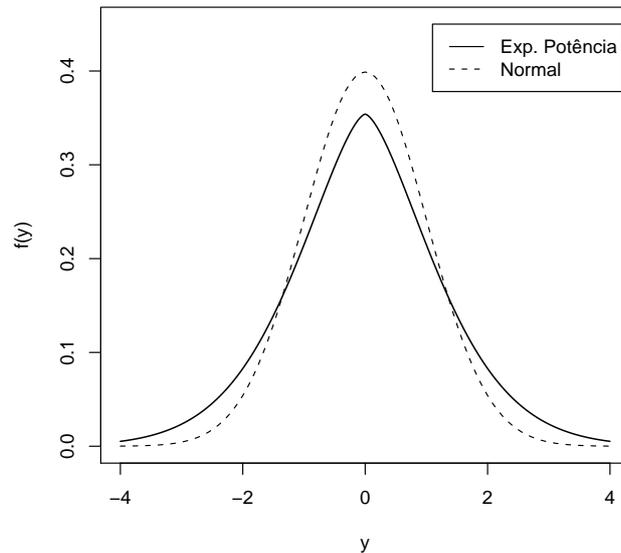


Figura 4: Densidades da distribuição exponencial potência ( $l = 0, 3$ ) e distribuição Normal padrão.

## 2.5 Distribuição Cauchy

A distribuição de Cauchy tem o nome do famoso matemático francês Augustin Louis Cauchy, um dos matemáticos mais produtivos que já existiu, professor na Escola Politécnica de Paris em 1816.

A função densidade de probabilidade é definida por

$$f(y) = \frac{1}{\pi(1+y^2)}, \quad -\infty < y < \infty. \quad (2.35)$$

Esta função densidade de probabilidade é peculiar pois os seus momentos e cumulantes não existem. Dessa forma, outras medidas bastante úteis nessa distribuição são a moda e a mediana.

A distribuição de Cauchy é frequentemente comparada com a distribuição normal  $N(\mu, \sigma^2)$  e com a exponencial dupla  $ED(\mu, \sigma^2)$ . A distribuição de Cauchy possui caudas mais pesadas que a exponencial dupla e essa, por sua vez, tem caudas mais pesadas que a distribuição normal.

A distribuição de Cauchy é relacionada com a distribuição t de Student com  $\nu$  graus de

liberdade

$$f(y, \nu) = \frac{\Gamma(\frac{\nu+1}{2})}{\sqrt{\nu\pi}\Gamma(\frac{\nu}{2})} \left(1 + \frac{y^2}{\nu}\right)^{-(\nu+1)/2} = \frac{(1 + \frac{y^2}{\nu})^{-(\nu+1)/2}}{\sqrt{\nu}B(\frac{1}{2}, \frac{\nu}{2})}. \quad (2.36)$$

Sendo  $\Gamma(x)$  a função gamma de Euler e  $B$  é a função beta definida em termos da função gamma como  $B(p, q) = \frac{\Gamma(p)\Gamma(q)}{\Gamma(p+q)}$ . A distribuição de Cauchy surge como um caso especial quando  $\nu = 1$ .

A equação de máxima verossimilhança é representada por um polinômio de grau  $n - 1$  em  $\theta$ , em que  $\theta$  são parâmetros, cujas soluções, em geral, incluem vários máximos e mínimos da log-verossimilhança. O processo iterativo para calcular  $\hat{\theta}$  é

$$\theta^{(m+1)} = \theta^m + \frac{4}{n} \sum_{i=1}^n \frac{y_i - \theta^{(m)}}{1 + (y_i - \theta^{(m)})^2}. \quad (2.37)$$

Para  $n$  pequeno esta equação pode não convergir. Porém, para  $n \geq 9$ ,  $\theta^{(m+1)}$  é quase sempre o ponto de máximo mais próximo da mediana amostral. Com o objetivo de obter a convergência, o processo iterativo deve ser iniciado em  $\theta^{(1)}$  igual a mediana amostral. Dessa forma, pode-se demonstrar que a estimativa obtida em uma iteração tem as mesmas propriedades assintóticas de  $\hat{\theta}$ .

Para trabalhar com as caudas pesadas da distribuição de Cauchy, as vezes, é necessário truncar a distribuição. Isto também pode resolver o problema da média indefinida e dos momentos divergentes. Para uma truncagem simétrica, ou seja,  $-X \leq x \leq X$ , a função densidade de probabilidade renormalizada é obtida por

$$f(x) = \frac{1}{2 \arctan X} \frac{1}{(1 + x^2)}. \quad (2.38)$$

A esperança e a variância são respectivamente  $E(x) = 0$  e  $V(x) = \frac{X}{\arctan X} - 1$ . O terceiro momento central é  $\mu_3 = 0$  e o quarto momento é  $\mu_4 = \frac{X}{\arctan X} \left(\frac{X^2}{3} - 1\right) + 1$ .

Em muitos casos é possível obter uma aproximação da distribuição da soma e da média de variáveis aleatórias através do teorema central do limite. Contudo, devido a variância divergente da distribuição de Cauchy o teorema não é válido para esta situação. Uma forma de obter essas distribuições é utilizar a função característica, uma vez que a função característica da soma de variáveis aleatórias independentes é igual ao produto das funções características individuais. Assim, se  $Y_1, Y_2, \dots, Y_n$  são variáveis aleatórias independentes e cada  $Y_j$  possui distribuição de Cauchy, então  $S = \sum_{j=1}^n Y_j$  tem distribuição

de Cauchy com parâmetro de locação  $\mu = \sum_{j=1}^n \mu_j$  e escala  $\sigma^2 = \sum_{j=1}^n \sigma_j^2$ . Um resultado interessante é que para  $a_j \neq 0$ ,  $\sum_{j=1}^n a_j Y_j$  independentes, tem-se uma distribuição de Cauchy com parâmetro de locação  $\sum_{j=1}^n a_j \mu_j$  e escala  $\sum_{j=1}^n |a_j| \sigma_j^2$ .

Um outro resultado importante é que através da função característica é possível verificar que a média de variáveis aleatórias independentes com distribuição de Cauchy é também distribuída de acordo com a distribuição de Cauchy, ou seja, se os  $Y_j$  são independentes e identicamente distribuídos,  $\bar{Y} = n^{-1} \sum_{j=1}^n Y_j \sim C(\mu, \sigma^2)$ . Dessa forma, a média amostral não é um estimador consistente para a mediana da população, pois tendo a mesma distribuição de Cauchy possui variância divergente. Portanto, a mediana amostral para uma amostra de  $n$  observações independentes é um estimador consistente para a verdadeira mediana.

Quando se tem  $C(0, 1)$  a distribuição Cauchy passa a ser chamada de Cauchy padrão ou  $t$  de Student central com um grau de liberdade.

A variável aleatória  $Y \sim S(\mu, \phi)$  tem distribuição de Cauchy se sua função geradora de densidade  $g(\cdot)$  é da forma

$$g(u) = \frac{1}{\pi(1+u)}, \quad u > 0. \quad (2.39)$$

Essa distribuição, denotada por  $Y \sim C(\mu, \phi)$ , é também conhecida como distribuição de Pearson Tipo VII. A sua função característica é dada por

$$\psi_y(t) = \exp \left\{ it\mu - |t| \sqrt{\phi} \right\}, \quad t \in \mathfrak{R}. \quad (2.40)$$

A distribuição de Cauchy é simétrica em torno da média e os pontos de inflexão da função densidade são  $(\mu \pm \phi\sqrt{3})$ . Os valores da função de distribuição acumulada nos pontos de inflexão são 0,273 e 0,723 que podem ser comparados com os valores da distribuição normal (0,159 e 0,841).

Uma alternativa para geração de números aleatórios seguindo uma distribuição de Cauchy é utilizar

$$y = \mu + \phi \frac{z_1}{z_2}, \quad (2.41)$$

em que  $z_1$  e  $z_2$  são valores de duas variáveis aleatória independente com distribuição normal padrão.

A Figura 5 apresenta uma comparação das densidades da distribuição Cauchy padrão e da distribuição normal padrão.

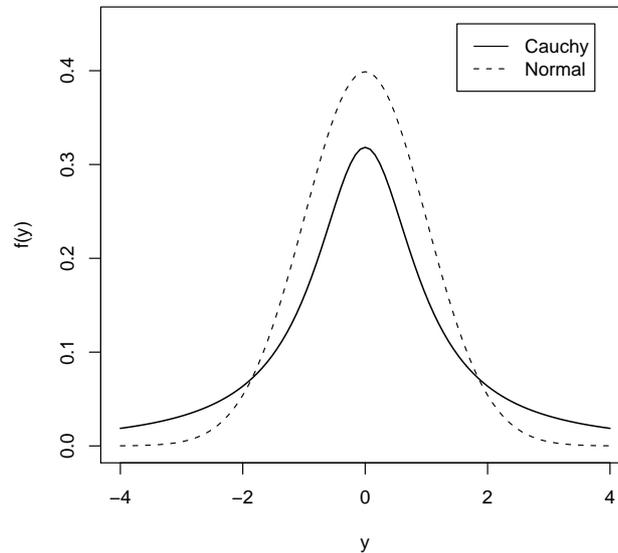


Figura 5: Densidades da distribuição Cauchy padrão e distribuição normal padrão.

## 2.6 Distribuição Logística

Uma primeira referência para o uso da função logística como uma curva de crescimento foi apresentada por Verhulst (1838,1845). O uso da curva para demografia econômica tem sido bastante popular desde o início do século XIX. Pearl e Reed (1920,1924), Pearl et al. (1940), e Schultz (1930) aplicaram o modelo logístico como modelo de crescimento em população humana, bem como em alguns organismo biológicos. Schultz (1930) usou a função logística para dados de produção agrícola. Outras aplicações bastante interessantes da função logística são nas áreas de análise de sobrevivência (PLACKEET, 1959) e em modelagem de distribuição de renda (FISK, 1961). Para mais detalhes e referências relevantes da distribuição logística ver Balakrishnan (1992).

Segundo Gumbel (1944), a distribuição logística surgiu de uma maneira puramente estatística como distribuição limite ( $n \rightarrow \infty$ ) de medidas padronizadas (média de maiores e menores valores) de amostras aleatórias de tamanho  $n$ . Gumbel e Keeney (1950) mostraram que a distribuição logística é obtida como a distribuição limite de um múltiplo apropriado do “quociente extremo” dado por *Maximo/Minimo*. Talacko (1956) mostrou que a logística é a distribuição limite ( $r \rightarrow \infty$ ) da variável padronizada correspondente

a  $\sum_{j=1}^r j^{-1} X_j$ , em que  $X_j$ 's são variáveis aleatórias independentes cada uma tendo distribuição valor extremo tipo I.

O fato adicional de que a distribuição logística tem forma similar a distribuição normal torna ela proveitosa em ocasiões apropriadas. Substituindo a normal pela logística obtém-se uma simplificação das análises sem trazer grandes discrepâncias na teoria. Contudo, tais substituições devem ser feitas com cuidado e entendimento da similaridade entre as duas distribuições.

### 2.6.1 Logística I

Uma variável aleatória  $Y \sim S(\mu, \phi)$  tem distribuição logística I se sua a função geradora de densidade  $g(\cdot)$  tem a forma

$$g(u) = c \frac{e^{-u}}{(1 + e^{-u})^2}, \quad u > 0, \quad (2.42)$$

em que  $c$  é a constante normalizadora obtida em  $\int_0^\infty u^{-1/2} g(u) du = 1$ , assim  $c = 1,484300-029$ . Esta distribuição é denotada por  $Y \sim LI(\mu, \phi)$ , com  $E(Y) = \mu$ ,  $Var(Y) \approx 0,79569\phi$  e  $\gamma_2 \approx 2,385165$ . Note que o coeficiente de curtose da distribuição normal é maior que o coeficiente de curtose da distribuição logística I.

Na Figura 6, apresentada a seguir, é possível observar o gráfico da função de densidade da distribuição logística I padrão comparado com a função de densidade da distribuição normal padrão.

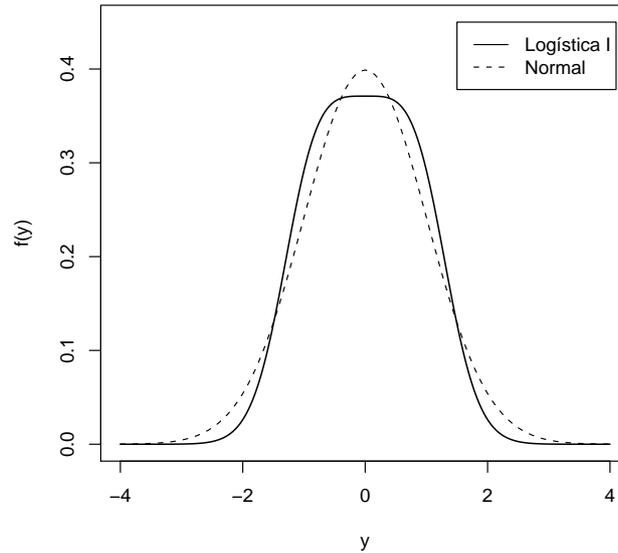


Figura 6: Densidades da distribuição logística I padrão e distribuição normal padrão.

## 2.6.2 Logística II

Uma variável aleatória  $Y \sim S(\mu, \phi)$  tem distribuição logística II se sua função geradora de densidade  $g(\cdot)$  é da forma

$$g(u) = \frac{e^{-u^{1/2}}}{(1 + e^{-u^{1/2}})^2}, \quad u > 0. \quad (2.43)$$

Dessa forma, a variável aleatória é denotada por  $Y \sim LII(\mu, \phi)$ .

A sua função característica é dada por

$$\psi_y(t) = \frac{2(e^{it\mu}\pi\phi^{1/2}t)}{(e^{\pi\phi^{1/2}t} - e^{-\pi\phi^{1/2}t})}, \quad t \in \mathfrak{R}. \quad (2.44)$$

Tem-se que  $E(Y) = \mu$ ,  $Var(Y) = \pi^2\phi/3$  e  $\gamma_2 = 4, 2$ , que é maior que as curtoses das distribuições normal e logística I.

Para geração de números aleatórios com distribuição  $LII(\mu, \phi)$ , uma relação bastante útil, definida em Hastings e Peacock (1975), é dada por

$$y = u + \sqrt{\phi} \log\{u/(1-u)\}, \quad (2.45)$$

em que  $u$  são números aleatórios com distribuição uniforme entre 0 e 1.

A Figura 7, apresenta o gráfico da função de densidade da distribuição logística II padrão comparado com a função de densidade da distribuição normal padrão.

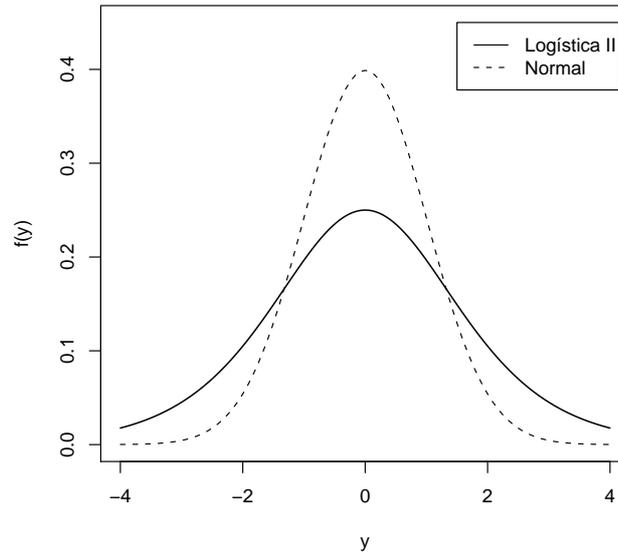


Figura 7: Densidades da distribuição logística II padrão e distribuição normal padrão.

## 2.7 Distribuição de Kotz

A distribuição de Kotz, na classe de distribuições elípticas univariadas, pode ser interpretada como sendo gerada por uma distribuição de Weibull ou uma distribuição valor extremo tipo III. Essa distribuição tem caudas mais pesadas do que a distribuição normal e um fato interessante é que os estimadores de máxima verossimilhança são as medianas espaciais generalizadas.

A distribuição de Kotz tem função de densidade de probabilidade dada por

$$f(y) = \frac{sr^{(2q-1)/2s}}{\Gamma((2q-1)/2s)} y^{2(q-1)} e^{-ry^{2s}}, \quad r, s > 0 \quad e \quad q > 1/2. \quad (2.46)$$

Uma variável aleatória  $Y \sim S(\mu, \phi)$  tem distribuição de Kotz se sua função geradora de densidade  $g(\cdot)$  é dada por

$$g(u) = \frac{r^{(2N-1)/2}}{\Gamma\left(\frac{2N-1}{2}\right)} u^{N-1} e^{-ru}, \quad r > 0, N \geq 1, u > 0, \quad (2.47)$$

e assim denotada por  $Y \sim K(\mu, \phi, N, r)$ . Tem-se que  $E(Y) = \mu$ ,  $Var(Y) = \{(2N - 1)/2r\}\phi$  e  $\gamma_2 = (2N + 1)/(2N - 1)$  e os momentos de ordem central  $2m$  são dados por

$$\mu_{2m} = \frac{\Gamma\{(2N + 2m - 1)/2\}}{r^m \Gamma\{(2N - 1)/2\}} \phi^m, m > 0. \quad (2.48)$$

Se  $N = 1$  a distribuição de Kotz se reduz a  $N(\mu, \phi/(2r))$ . Se  $N > 1$  a distribuição é bimodal com modas  $Y = \mu \pm \sqrt{(N - 1)/(r\phi)}$ . Além disso, tem-se que, se  $Z^2 = (y - \mu)^2/\phi$  então  $Z^2 \sim G(\{2N - 1\}/2, r)$ . Em um caso particular, quando  $N = 1$  e  $r = 1/2$ , tem-se que  $Z^2 \sim \chi_1^2$ .

Na Figura 8 apresentada a seguir, é possível observar os gráficos das densidades da distribuição de Kotz ( $N = 3, r = 1$ ) e da distribuição normal padrão.

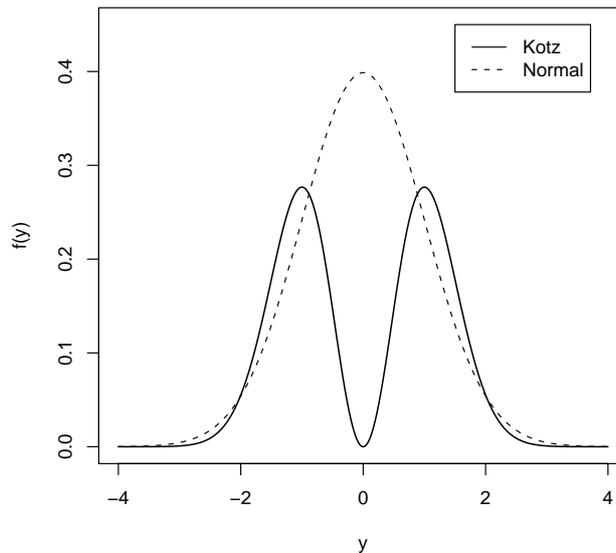


Figura 8: Densidades da distribuição Kotz ( $N = 3, r = 1$ ) e distribuição normal padrão.

## 2.8 Distribuição de Laplace ou Exponencial Dupla

A distribuição de Laplace foi desenvolvida pelo astrônomo Francês Laplace (1774), como a distribuição para qual a função de verossimilhança é maximizada pelo conjunto de parâmetros de locação iguais as medianas dos valores observados de um número de variáveis aleatórias independentes e identicamente distribuídas. Este resultado, conhecido como 1º lei de Laplace, apareceu no artigo fundamental de Laplace sobre distribuições simétricas para descrever erros de medidas. Dessa forma, uma característica importante

dessa distribuição é que o melhor estimador da média populacional é a mediana amostral e não a média amostral.

Embora conduzisse bastante pesquisa sobre física, outro tema principal era a teoria das probabilidades. Laplace desenvolveu um sistema matemático de raciocínio indutivo baseado em probabilidades, que hoje coincide com as idéias Bayesianas. Em 1812, Laplace publicou sua teoria analítica das probabilidades. O método de estimar a proporção do número de casos favoráveis comparada ao número total de casos possíveis já havia sido indicado por Laplace em um artigo escrito em 1779. Ele consiste em tratar os valores sucessivos de qualquer função como coeficientes na expansão de outra função com referência a uma variável diferente. Assim, Laplace mostra como, por meio de interpolação, esses coeficientes podem ser determinados a partir da função geradora.

Esta distribuição é conhecida sob diferentes nomes, sendo “Exponencial Dupla” um dos mais comuns, uma vez que pode ser considerada como a distribuição da diferença de duas variáveis aleatórias i.i.d. com distribuição exponencial. Outras denominações da distribuição são: exponencial cauda dupla (GREENWOOD et al., 1962), exponencial bilateral (FELLER, 1966) e primeira lei de erro Poisson (WEIDA, 1935). Um breve resumo de vários desenvolvimentos significantes dessa distribuição foi dado por Ord em 1983.

Andrews et al. (1972) e Hoaglin et al. (1985) utilizaram a distribuição de Laplace como alternativa para a distribuição normal em estudos de modelos robustos. Em 1976, Manly apresentou alguns exemplos de funções baseadas na distribuição de Laplace e Easterling em 1978, considerou um modelo com resposta exponencial com erro de medida exponencial dupla. Hsu (1979) sugeriu a utilização da distribuição de Laplace para estudos de navegação e Okubo e Narita (1980) usaram a distribuição de Laplace para distribuição de ventos extremos no Japão.

A função densidade de Laplace é dada por

$$f(y) = \frac{1}{2\sigma^2} e^{-|y-\mu|/\sigma^2}, \quad -\infty < y <, \quad \sigma^2 > 0. \quad (2.49)$$

A função acumulada da distribuição de Laplace é dada por

$$F(y) = \begin{cases} \frac{1}{2} e^{-\lambda(\mu-y)} & \text{se } y \leq \mu, \\ 1 - \frac{1}{2} e^{-\lambda(y-\mu)} & \text{se } y > \mu, \end{cases}$$

Um resultado interessante entre a distribuição normal e a de Laplace foi estabelecido por Nyquist et al. (1954). Eles mostraram que se  $X_1, X_2, X_3$  e  $X_4$  são variáveis aleatórias

independentes com distribuição normal, então

$$|\mathbf{D}| = \begin{vmatrix} X_1 & X_2 \\ X_3 & X_4 \end{vmatrix} = X_1X_4 - X_2X_3,$$

tem distribuição de Laplace com  $\mu = 0$  e  $\sigma^2 = 2$ . Uma demonstração heurística de que  $X_1X_4 + X_2X_3$  também tem distribuição de Laplace foi apresentada por Mantel e Pasternack (1966).

Algumas relações importantes encontradas na literatura são:

Se  $Y \sim Laplace(0, b)$  então  $|Y| \sim Exponencial(b^{-1})$ .

Se  $Y \sim Exponencial(\lambda)$  e  $X \sim Bernoulli(0,5)$ , sendo  $X$  e  $Y$  independentes, então  $Y(2X - 1) \sim ED(0, \lambda^{-1})$ .

Se  $Y_1 \sim Exponencial(\lambda_1)$  e  $Y_2 \sim Exponencial(\lambda_2)$ , ambas independentes, então  $(\lambda_1Y_1 - \lambda_2Y_2) \sim ED(0, 1)$ .

Uma variável aleatória  $Y \sim S(\mu, \phi)$  tem distribuição de Laplace ou exponencial dupla se sua função geradora de densidade  $g(\cdot)$  tem a forma

$$g(u) = \frac{1}{2}e^{-\sqrt{u}}, \quad u > 0, \quad (2.50)$$

em que a variável aleatória  $Y$  é denotada por  $Y \sim ED(\mu, \phi)$ .

A função característica dessa distribuição é dada por

$$\psi_y(t) = \frac{e^{it\mu}}{1 + t^2\phi}, \quad t \in \Re. \quad (2.51)$$

Se  $Z \sim ED(0, 1)$  os momentos  $\mu'_r$  são dados por

$$\mu'_r = E(z^r) = \begin{cases} 0 & r \text{ ímpar} \\ r! & r \text{ par} \end{cases}$$

Assim,  $E(Y) = \mu$ ,  $Var(Y) = 2\phi$  e  $\gamma_2 = 6$ . Os quartis da distribuição são  $(\mu \pm 0,534\sqrt{\phi})$ .

Para geração de números aleatórios de uma distribuição de Laplace é necessário resolver a equação  $F(y) = \epsilon$ , em que  $\epsilon$  é um número aleatório uniforme entre 0 e 1 e  $y$  é expresso como

$$y = \mu + \ln(2\epsilon)/\phi, \quad \text{se } \epsilon \leq \frac{1}{2}, \quad (2.52)$$

e

$$y = \mu - \ln(2 - 2\epsilon)/\phi, \quad \text{se } \epsilon > \frac{1}{2}. \quad (2.53)$$

Na Figura 9 é possível observar os gráficos das densidades da distribuição exponencial dupla padrão e da distribuição normal padrão.

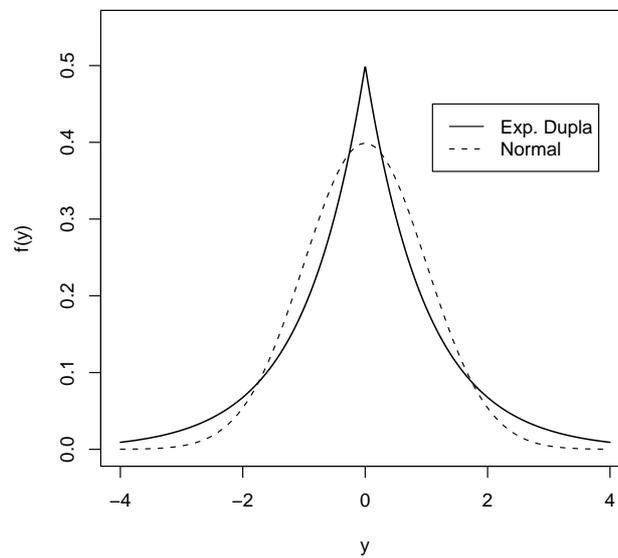


Figura 9: Densidades da distribuição exponencial dupla padrão e distribuição normal padrão.

## 3 Modelos Não-Lineares

Os modelos lineares são bastante utilizados para um grande número de aplicações em regressão, com o objetivo de relacionar uma variável resposta a uma ou mais variáveis explicativas. Porém, em algumas situações um modelo não-linear pode ser mais apropriado. A principal característica dos modelos não-lineares é que eles são deduzidos a partir de suposições teóricas e os parâmetros resultantes são interpretáveis. Assim, aproximá-los por modelos lineares, mesmo que sejam alcançados ajustes satisfatórios, prejudicaria bastante a obtenção de estimativas mais realistas dos parâmetros de interesse.

Para aplicar os modelos não-lineares é necessário supor que os erros tenham variância constante, propriedade conhecida como homocedasticidade. Os modelos não-lineares são aplicados em diversas áreas, tais como Ecologia, Biologia, Ciência Florestal, entre outras. Uma aplicação frequente dos modelos não-lineares pode ser para explicar a altura de árvores pela sua idade ou para explicar o comprimento de um peixe pela sua idade. Fenômenos produzindo curvas sigmoidais na forma de S, frequentemente encontrados na Agricultura, em Biologia, Ecologia, Engenharia e Economia, são explicados através dos modelos não-lineares.

As principais técnicas desenvolvidas para os modelos não-lineares, até o início dos anos 70, eram restritas à suposição de normalidade para a variável resposta. Embora as técnicas de diagnóstico do modelo normal não-linear sejam simples extensões das técnicas do modelo linear, as interpretações não são diretamente aplicadas, particularmente em virtude dos resíduos ordinários não terem mais distribuição aproximadamente normal quando a variância é desconhecida. Técnicas específicas de diagnóstico para os modelos normais não-lineares foram desenvolvidas por Cook e Tsai (1985). Similarmente, as propriedades das somas de quadrados contidas nas tabelas clássicas de análise de variância não são estendidas diretamente para o caso não-linear.

### 3.1 Definição

O modelo não-linear é definido por

$$y_i = h(x_i, \beta) + \varepsilon_i = \eta_i(\beta) + \varepsilon_i, \quad (3.1)$$

em que  $h(x_i, \beta)$  é uma função não-linear diferenciável, os erros aleatórios,  $\varepsilon_i$ , são não-correlacionados com média zero e variância constante,  $\beta$  é um vetor  $p \times 1$  de parâmetros desconhecidos e  $x_i = (x_1, \dots, x_q)$  representa os valores de  $q$  variáveis explicativas.

Uma diferença importante do modelo não-linear é que o número de parâmetros regressores não são necessariamente diretamente relacionados com o número de variáveis explicativas no modelo. No modelo linear, se existem  $p-1$  variáveis explicativas então existem  $p$  parâmetros regressores no modelo, o que pode não acontecer no modelo não-linear. Por exemplo, considerando o modelo de regressão exponencial dado por

$$Y_i = \beta_0 + \beta_1 \exp(\beta_2 x_i) + \varepsilon_i, \quad (3.2)$$

é possível verificar que existe um única variável regressora e três coeficientes de regressão.

Em alguns casos, é possível linearizar a variável resposta através de uma transformação. Uma função não-linear que pode ser linearizada por uma transformação é denominada de função intrinsecamente linear. Por exemplo, a função exponencial

$$f(x, \gamma) = \gamma_0 \exp(\gamma_1 x), \quad (3.3)$$

é uma função intrinsecamente linear devido ao fato de ser linearizada através de uma transformação logaritma

$$\log f(x, \gamma) = \log(\gamma_0) + \gamma_1 x. \quad (3.4)$$

Esta função transformada pode ser representada no modelo linear na forma

$$g(x, \gamma) = \beta_0 + \beta_1 x, \quad (3.5)$$

em que  $g(x, \gamma) = \log f(x, \gamma)$ ,  $\beta_0 = \log \gamma_0$  e  $\beta_1 = \gamma_1$ .

O fato da função não-linear ser intrinsecamente linear não implica necessariamente que a regressão linear é apropriada. A razão é que a transformação para linealizar a resposta pode afetar o termo de erro no modelo.

Na literatura, Draper e Smith (1981) e Montgomery e Peck (1982) apresentam de forma introdutória o modelo de regressão não-linear. Por outro lado, duas referências que abrangem de forma completa o tema são os livros de Bates e Watts (1988) e Seber e Wild (2003).

### 3.1.1 Estimação

A estimação dos parâmetros de um modelo não-linear é usualmente obtida pelo método de mínimos quadrados ou pelo método de máxima verossimilhança, assim como no modelo linear. A função de mínimos quadrados do modelo não-linear será dada por

$$S(\beta) = \sum_{i=1}^n \{y_i - \eta_i(\beta)\}^2. \quad (3.6)$$

Diferenciando a equação anterior em relação a cada elemento  $\beta_j = (j = 1, \dots, p)$  do vetor  $\beta$  e igualando a zero, obtém-se um conjunto de  $p$  equações normais expressas por

$$\sum_{i=1}^n [y_i - \eta_i(\beta)] \left[ \frac{\partial \eta_i(\beta)}{\partial \beta_j} \right]_{\beta=\hat{\beta}} = 0. \quad (3.7)$$

Entretanto, no modelo não-linear  $\frac{\partial \eta_i(\beta)}{\partial \beta_j}$  será função dos parâmetros desconhecidos e assume também uma forma não-linear. Assim, a solução das equações normais pode ser extremamente difícil de ser obtida, fazendo-se necessária a utilização de métodos iterativos para obtenção das estimativas do vetor de parâmetros  $\beta$ .

Sejam  $y_1, \dots, y_n$  variáveis aleatórias independentes com estrutura não-linear. O algoritmo de Newton-Raphson para obtenção da estimativa de mínimos quadrados, que coincide com a estimativa de máxima verossimilhança, é apresentado a seguir. Essa estimativa é obtida minimizando a função quadrática dada em 3.6.

Expandindo  $\eta(\beta)$  por série de Taylor em torno de um valor  $\beta^{(0)}$  até a segunda ordem, chega-se ao seguinte processo iterativo:

$$\beta^{(m+1)} = \beta^{(m)} + \{\tilde{X}^{(m)T} \tilde{X}^{(m)}\}^{-1} \tilde{X}^{(m)T} \{y - \eta(\beta^{(m)})\}, \quad (3.8)$$

$m = 0, 1, \dots$ , em que  $\tilde{X}$  é a matriz Jacobiana da transformação de  $\eta(\beta^{(m)})$  em  $\beta$ . A convergência em geral depende de valores iniciais para os parâmetros do vetor  $\beta$ . Isso pode evitar que problemas relacionados com a estrutura paramétrica do modelo, tais como não linearidade acentuada e/ou mal condicionamento da matriz  $\tilde{X}$ , prejudiquem a convergência do processo iterativo. Existem algumas técnicas para se obter valores iniciais para os parâmetros de  $\beta$ . Bates e Watts (1988) apresentaram e discutiram de forma mais detalhada técnicas para se obter boas escolhas para tais valores iniciais.

### 3.1.2 Resíduos

Uma etapa importante na análise de regressão é a verificação de possíveis afastamentos das suposições feitas para o modelo, bem como a existência de observações extremas com alguma interferência desproporcional nos resultados do ajuste. Tal etapa, conhecida como análise de diagnóstico, iniciou-se com a análise de resíduos para detectar a presença de pontos extremos e avaliar a adequação da distribuição proposta para a variável resposta. Uma referência importante é o artigo de Cox e Snell (1968) em que é apresentada uma forma bastante geral de definir resíduos, usada até os dias atuais.

No modelo de regressão linear, como os erros apresentam distribuição normal e são não correlacionados, testes estatísticos e intervalos de confiança baseados nas distribuições *t* de Student e *F* podem ser usados para realizar inferência sobre as estimativas dos parâmetros que apresentam ainda propriedades interessantes como: não-viés, mínima variância e distribuição Normal.

O resíduos ordinários não-lineares são definidos por

$$r_i = y_i - \eta_i(\hat{\beta}), \quad i = 1, \dots, n \quad (3.9)$$

Em geral, mesmo assintoticamente, os resíduos ordinários são viesados e possuem distribuição não normal dificultando a verificação dos modelos pelos métodos tradicionais.

## 3.2 Modelos de Crescimento

A análise de dados de crescimento é importante em muitas áreas de estudo. Biólogos estão interessados na descrição do crescimento e em tentar entender seu mecanismo subjacente. Químicos estão interessados na formulação do produto de uma reação química

sobre o tempo. Médicos estão interessados no crescimento infantil, bem como no crescimento de tumores e o efeito de tratamentos sobre tais crescimento. Na área social o interesse pode estar no crescimento de populações ou crescimento de demanda de energia.

Estudos de crescimento e produção tratam do desenvolvimento de mecanismos de predição das características quantitativas de um povoamento florestal em condições específicas. Essa predição é importante para a avaliação econômica das alternativas de manejo.

A análise do crescimento de seres vivos, utilizando modelos matemáticos e estatísticos, não é algo novo. Esse tipo de análise está cada vez mais sofisticada no setor florestal e, segundo Clutter et al. (1983) e Davis e Johnson (1987), uma das principais tarefas atribuídas ao profissional da área é entender como ocorre o crescimento das árvores de um povoamento, qual tratamento silvicultural adotar e prever qual será a época de corte e a produtividade do plantio.

O desenvolvimento de novos modelos de crescimento tem tido avanços consideráveis nas ciências florestais, como ocorreu com a generalização de Chapman-Richards para o modelo de Bertalanffy (1957) e trabalhos de Prodan (1968) dando um tratamento compreensivo de várias funções de crescimento e leis de crescimento estudadas com referência ao crescimento florestal. Esses modelos são formulações não-lineares que, através de curvas de crescimento procuram explicar as relações de crescimento inerentes a várias partes do organismo vegetal, até a sua maturidade, em situações ambientais distintas. Na literatura, são propostos vários modelos não-lineares para descrever curvas de crescimento tais como: Logístico, Gompertz, Bertalanffy, Monomolecular, Weibull, Chapman-Richards, Mitschirlich, Schumacher, entre outros. Os modelos de crescimento têm uma hipótese associada a causa ou função do fenômeno descrito pela variável resposta e têm um significado a partir de parâmetros florestais.

Bertalanffy (1934) definiu crescimento como um aumento mensurável de um sistema orgânico como função da assimilação de material proveniente de seu ambiente. Nas ciências biológicas, as funções de crescimento têm sido usadas por muitos anos, principalmente, para prover dados de crescimento de um organismo em um determinado tempo.

Schumacher (1939) desenvolveu um modelo para inicialmente estimar índice de sítio, mas que por sua amplitude de aplicação também é usado nas estimativas de crescimento e sobrevivência/mortalidade.

Em 1930, Mitschirlich utilizou um princípio usado na economia (a produção se apro-

xima de um limite e o incremento na produção tende a zero), traduzindo-o em termos de crescimento biológico como "Lei dos efeitos dos fatores de crescimento"(SILVA, 1986).

Essa lei diz que a taxa de crescimento  $\frac{\partial W}{\partial t}$  é proporcional a diferença entre o tamanho máximo (último nível de crescimento ou termo assintótico) e o próprio parâmetro de crescimento,

$$\frac{\partial W}{\partial t} = c(U - W), \quad (3.10)$$

resultando em

$$W = U[1 - \exp(-c't)]^n, \quad (3.11)$$

em que

W = parâmetro de crescimento

U = termo assintótico que significa o último nível de crescimento

$c' = c_1 = c_2 = \dots = c_n$

t = tempo

Bertalanffy (1934, 1957) derivou seu modelo de crescimento a partir de estudos das relações alométricas dos organismos. Estudos extensivos de muitos organismos aquáticos e terrestres levaram Bertalanffy a conclusão de que para uma classe principal de organismos a relação alométrica entre a área da superfície e volume total pode ser adequadamente expressa por:

$$S = \alpha V^{2/3}, \quad (3.12)$$

em que  $S$  é a área da superfície,  $V$  é o volume do organismo,  $\alpha$  é o parâmetro a ser calculado e  $2/3$  é a constante alométrica.

Após algumas álgebras, a equação original foi expressa por

$$\frac{\partial W}{\partial t} = \eta W^{2/3} - kW. \quad (3.13)$$

Richards (1959), estudando o crescimento de plantas e Chapman (1961) o crescimento de peixes consideraram que a constante alométrica de  $2/3$  do modelo de Berta-

lanffy era muito restrita. Nas ciências biológicas, Richards foi o primeiro a aplicar o modelo de crescimento desenvolvido por Bertalanffy para descrever o crescimento de animais (FRANCE e THORNLEY, 1984). A generalização de Chapman-Richards do modelo de crescimento de Bertalanffy é uma expressão matemática de uma hipótese concernente a causas essenciais do fenômeno de crescimento, de tal modo que os parâmetros no modelo têm pelo menos uma total interpretação biológica ou fisiológica (TURNBULL, 1963; PIENAAR e TURNBULL, 1973). Nesse trabalho utilizou-se o modelo de Chapman-Richards, definido a seguir.

### 3.2.1 Modelo de Chapman-Richards

Através da modificação sugerida por Chapman-Richards, obtém-se

$$\frac{\partial W}{\partial t} = \alpha W^m - \gamma W, \quad (3.14)$$

resultando em

$$W = U[1 - \beta \exp(-k(t - t_0))]^{1/(1-m)}. \quad (3.15)$$

Quando  $t_0 = 0$

$$W = U[1 - \beta \exp(-kt)]^{1/(1-m)}. \quad (3.16)$$

Este modelo foi nomeado como modelo de crescimento de Chapman-Richards por Turnbull (1963). Devido ao valor do parâmetro  $\beta$  ser muito próximo de 1 esse parâmetro é retirado do modelo.

Dependendo do valor de  $m$ , o modelo de Chapman-Richards representa várias leis de crescimento. O modelo de Bertalanffy ( $m=2/3$ ), a função de crescimento monomolecular ( $m=0$ ), a função de crescimento Gompertz ( $m=1$ ) e a função de crescimento logística ( $m=2$ ) são casos especiais da função de crescimento de Chapman-Richards (Machado, 1978).

Clutter e Jones (1980) desenvolveram um procedimento em que o valor assintótico  $U$ , que é constante no modelo, pode ser eliminado e substituído por um valor que corresponde a um determinado valor de  $W$  no tempo  $i$ . Para isto utilizou-se o método das diferenças de equações.

Considerando o modelo anterior, em um tempo inicial  $i$ , tem-se

$$W_i = U[1 - \exp(-kt_i)]^{1/(1-m)}. \quad (3.17)$$

Considerando agora um tempo final  $f$ , o modelo é

$$W_f = U[1 - \exp(-kt_f)]^{1/(1-m)}. \quad (3.18)$$

Como  $U$  é constante nos tempos  $i$  e  $f$ , pode-se conhecer seu valor, por exemplo, no tempo  $i$

$$U = W_i(1 - \exp(-kt_i))^{-1/(1-m)}. \quad (3.19)$$

Substituído a equação (3.26) na equação (3.25), tem-se

$$W_f = W_i \left( \frac{1 - \exp(-kt_f)}{1 - \exp(-kt_i)} \right)^{1/(1-m)}. \quad (3.20)$$

Para estimação da curva de crescimento, a partir desses resultados, pode-se utilizar a equação (3.27) com o valor assintótico eliminado do modelo para dar lugar a um valor real  $W_i$ .

Contudo, na prática, é possível estimar o crescimento  $W$  em função da idade  $t$  a partir da equação (3.24). Denotando por  $W(t)$  a altura da árvore na idade  $t$ , a curva de Chapman-Richards é dada por

$$W(t) = U(1 - \exp(-kt))^{1/(1-m)} + e, \quad (3.21)$$

em que  $e$  é um erro aleatório.

## 4 Modelos Simétricos Transformados

### 4.1 Definição

O uso de transformações em análise de regressão é muito comum e pode ser útil quando o modelo original não satisfaz as suposições usuais. A família de transformações proposta por Box e Cox (1964) é frequentemente usada para obter um modelo linear normal. Os modelos de Box e Cox adicionam ferramentas úteis para a estatística aplicada, pertinente aos aspectos da homogeneidade da variância, aditividade do modelo e normalidade. Essa transformação raramente satisfaz as suposições básicas de linearidade, normalidade e homocedasticidade simultaneamente. Dessa forma, a nova classe de modelos proposta por Cordeiro e Andrade (2009) trabalharam com uma família geral de transformações monótonas e combina a idéia em modelos de transformação de Box e Cox a variável resposta com a componente sistemática do modelo linear generalizado, embora assumam uma família geral de distribuições simétricas para a resposta. É importante citar que as demais seções desse capítulo são baseadas no artigo de Cordeiro e Andrade (2009).

Considere uma família de transformação paramétrica geral para a variável resposta  $Y$

$$Y^{(\lambda)} = \Lambda(Y, \lambda), \quad (4.1)$$

em que  $\lambda$  é um parâmetro escalar definindo uma transformação particular. Assume-se que para cada  $\lambda$ ,  $Y^{(\lambda)}$  é uma função monótona de  $Y$ . Geralmente, considera-se a transformação potência de Box e Cox (1964),  $Y^{(\lambda)} = (Y^\lambda - 1)/\lambda$  quando  $\lambda \neq 0$  ou  $Y^{(\lambda)} = \log(Y)$  quando  $\lambda = 0$ , e assume-se que existe um valor de  $\lambda$  para a variável resposta tal que  $Y^{(\lambda)}$  segue um modelo de regressão não-linear  $\mu = h(X, \beta)$  com erro normal e variância constante. É frequente assumir a transformação potência quando  $Y$  é positivo, por outro lado, se  $Y$  for negativo muitos valores de  $\lambda$  poderão ser claramente inadmissíveis.

Manly (1976) propôs a transformação exponencial para ser usada com  $Y$ 's negativos da forma:  $Y^{(\lambda)} = (e^{\lambda Y} - 1)/\lambda$  quando  $\lambda \neq 0$ , ou  $Y^{(\lambda)} = Y$  quando  $\lambda = 0$ . Essa transformação parece ser eficaz em transformar distribuições unimodais para uma distribuição simétrica quase normal. Transformações alternativas para a transformação potência são revistas por Sakia (1992).

Considere as variáveis aleatórias transformadas  $Y_1^\lambda, \dots, Y_n^\lambda$  com  $Y^\lambda$  definida em (4.1) independentes e cada  $Y_i^\lambda$  seguindo uma distribuição simétrica contínua com parâmetros de locação  $\mu_i \in \mathfrak{R}$  e parâmetro de dispersão  $\phi > 0$

$$f(y; \mu, \phi) = \frac{1}{\sqrt{\phi}} g\{\phi^{-1}(y - \mu)^2\}, \quad y \in \mathfrak{R}. \quad (4.2)$$

A função  $g(\cdot)$  (conhecida como densidade geradora) satisfaz  $\int_0^\infty U^{-1/2} g(y) dy = 1$ , em que  $U = \frac{(y-\mu)^2}{\phi}$ . A função densidade de probabilidade da variável aleatória padronizada  $Z^{(\lambda)} = (Y_i^\lambda - \mu_i)/\sqrt{\phi}$  é  $f(v, 0, 1) = g(v^2)$ ,  $v \in \mathfrak{R}$ , isto é,  $Z^\lambda \sim S(0, 1)$ . A família simétrica de densidades de locação-dispersão guarda a estrutura da distribuição normal mas elimina a forma específica da densidade normal para incluir densidades simétricas que tem caudas mais curtas ou mais longas do que as caudas da normal.

A função característica de  $Y_i^\lambda$  é  $\psi(t) = e^{it\mu_i} \varphi(t^2\phi)$ ,  $t \in \mathfrak{R}$ , para alguma função  $\varphi(x) \in \mathfrak{R}$  e  $x > 0$ . Mostrando que eles existem,  $E(Y_i^{(\lambda)}) = \mu_i$  e  $Var(Y_i^{(\lambda)}) = k_2\phi$ , em que  $k_2 = -2\varphi'(0)$  é uma constante e  $\varphi' = \varphi(x)/dx |_{x=0}$ .

Para introduzir uma estrutura regressora na classe de distribuições (4.2), toma-se a componente sistemática do modelo linear generalizado para o vetor da média  $\mu = E(Y^{(\lambda)})$  dado por

$$g(\mu) = \eta_i(\beta) = h(x_i, \beta), \quad (4.3)$$

em que  $g(\cdot)$  é conhecida e duas vezes diferenciável,  $\eta_i(\beta)$  é o preditor não-linear,  $X$  é uma matriz  $n \times p$  de posto completo e  $\beta = (\beta_1, \dots, \beta_p)^T$  é um conjunto de parâmetros não-lineares desconhecidos a serem estimados.

Os modelos simétricos transformados assumem que existe algum valor de  $\lambda$  em (4.1) tal que as variáveis aleatórias transformadas  $Y_1^{(\lambda)}, \dots, Y_n^{(\lambda)}$  podem ser tratada como distribuídas independentemente seguindo a componente aleatória (4.2) e a componente sis-

temática (4.3).

A forma generalizada (4.1) é usada para determinar a forma específica dentre uma particular classe de funções de transformações que é otimizada em referência ao critério de máxima verossimilhança. O valor ótimo de  $\lambda$  é obtido para produzir um modelo simétrico ajustado mais aproximado para os dados transformados.

## 4.2 Inferência

Sejam  $X$  a matriz do modelo, os dados originais  $y = (y_1, \dots, y_n)^T$  e assumindo que as observações transformadas  $y^\lambda = (y_1^{(\lambda)}, \dots, y_n^{(\lambda)})^T$ , para alguma transformação desconhecida do parâmetro  $\lambda$  em 4.1, satisfazem as suposições (4.2) e (4.3). O principal objetivo na análise de modelos transformados simétricos é fazer inferências no vetor de parâmetros  $\beta$  e nos escalares  $\phi$  e  $\lambda$ . Seja  $J(\lambda, y)$  o Jacobiano da transformação de  $y$  para  $y^{(\lambda)}$ . A log-verossimilhança para os parâmetros do modelo pode ser expressa como

$$l(\beta, \phi, \lambda) = -\frac{n}{2} \log \phi + \sum_{i=1}^n \log g \left\{ \phi^{-1}(y_i^{(\lambda)} - \mu_i)^2 \right\} + \sum_{i=1}^n \log \{J(\lambda, y_i)\}, \quad (4.4)$$

em que

$$J(\lambda, y_i) = \left| \frac{d\Lambda(y_i, \lambda)}{dy_i} \right|. \quad (4.5)$$

Os modelos simétricos potência com transformação de Box e Cox produzem  $J(\lambda, y_i) = |y_i|^{(\lambda-1)}$ . A log-verossimilhança apresentada pode ser maximizada incondicionalmente usando alguns software como o SAS, Matlab, R ou a linguagem de programação 0x.

Alternativamente, pode-se assumir  $\lambda$  fixo e assim obter as equações de log-verossimilhança condicionadas para estimar  $\beta$  e  $\phi$ . Seja  $\hat{\beta}^{(\lambda)}, \hat{\eta}^{(\lambda)} = h(X_i, \hat{\beta}^{(\lambda)}), \hat{\mu}^{(\lambda)} = g^{-1}(\hat{\eta}^{(\lambda)})$  e  $\hat{\phi}^{(\lambda)}$  as estimativas de máxima verossimilhança de  $\beta, \eta, \mu$  e  $\phi$ , respectivamente, para dado  $\lambda$ . As estimativas condicionais de  $\hat{\beta}^{(\lambda)}$  e  $\hat{\phi}^{(\lambda)}$  são obtidas facilmente por iteração reponderando o quadrado médio.

Seja a função  $t(x) = \log\{g(x)\}$  e sua primeira derivada  $t'(x)$ , a matriz diagonal  $H = \text{diag}\{\mu'_i\}$   $n \times n$  e  $D^{(\lambda)} = \text{diag}\{d_i^{(\lambda)}\}$ , em que  $\mu'_i = d\mu_i/\mu_{i(\beta; x_i)}$  é a primeira derivada da função de ligação inversa,  $d_i^{(\lambda)} = -2t'(u_i^{(\lambda)})$ ,  $u_i^{(\lambda)} = \phi^{-1}(y_i^{(\lambda)} - \mu_i)^2$  é a raiz da  $i$ -ésima observação transformada padronizada  $y_i^{(\lambda)}$  e  $Q(\mu^{(\lambda)}, y^\lambda) = (y^{(\lambda)} - \mu)^T D^{(\lambda)} (y^{(\lambda)} - \mu)$ . A

função escore para o vetor de parâmetro  $\beta$  e parâmetro de dispersão  $\phi$  para dado  $\lambda$  é

$$U_{\beta}^{(\lambda)} = \phi^{-1} X^T H D^{(\lambda)} (y^{(\lambda)} - \mu) \quad (4.6)$$

e

$$U_{\phi}^{(\lambda)} = (2\phi)^{-1} \{ \phi^{-1} Q(\mu^{(\lambda)}, y^{(\lambda)}) - n \}. \quad (4.7)$$

Condicionando em  $\lambda$ , pode-se ver facilmente que os parâmetros  $\beta$  e  $\phi$  são globalmente ortogonais e então as estimativas de máxima verossimilhança de  $\hat{\beta}^{(\lambda)}$  e  $\hat{\phi}^{(\lambda)}$  são assintoticamente independentes devido a sua normalidade assintótica e a estrutura bloco diagonal da matriz de informação conjunta, isto é,  $K = \text{diag}(K_{\beta}, k_{\phi})$ , em que  $K_{\beta} = 4a\phi^{-1} X^T H^2 X$  e  $k_{\phi} = n(4b - 1)/(4\phi^2)$  são, respectivamente, as matrizes de informação para  $\beta$  e  $\phi$ . As constantes  $a$  e  $b$ , independentes em  $\mu$  e  $\phi$ , podem ser obtidas da variável aleatória  $Z^{(\lambda)} \sim S(0, 1)$ , através de  $a = E\{t'(Z^{(\lambda)2})^2 Z^{(\lambda)2}\}$  e  $b = E\{t'(Z^{(\lambda)2})^2 Z^{(\lambda)4}\}$

Os estimadores de máxima verossimilhança de  $\beta$  e  $\phi$  são obtidos através da solução do sistema de equações dado por

$$U_{\beta}^{(\lambda)} = 0 \quad (4.8)$$

e

$$U_{\phi}^{(\lambda)} = 0. \quad (4.9)$$

As formulações acima são não-lineares e não podem ser resolvidas explicitamente. No entanto, podem ser resolvidas por meio de um método iterativo, como por exemplo o método escore de Fisher. Assim, o processo iterativo reduz para

$$\beta^{(\lambda)(m+1)} = (X^T H^{(\lambda)(m)2} X)^{-1} X^T H^{(\lambda)(m)2} \delta^{(\lambda)(m)} \quad (4.10)$$

e

$$\phi^{(\lambda)(m+1)} = \frac{1}{n} Q(\mu^{(\lambda)(m)}, y^{(\lambda)}), \quad (4.11)$$

em que

$$\delta^{(\lambda)(m)} = \eta^m + (4a)^{-1} H^{(\lambda)(m)(-1)} D^{(\lambda)(m)} (y^{(\lambda)} - \mu^{(\lambda)(m)}). \quad (4.12)$$

Aproximações iniciais  $\beta^{(\lambda)(1)}$  e  $\phi^{(\lambda)(1)}$  escolhidas, por exemplo, pelo ajuste do modelo Box-Cox, são usadas para avaliar  $H^{(\lambda)(1)}$ ,  $D^{(\lambda)(1)}$  e  $\delta^{(\lambda)(1)}$  e assim produzirem a próxima estimativa para  $\beta^{(\lambda)(2)}$ . Então, atualiza-se  $\mu^{(\lambda)(2)}$  e  $Q(\mu^{(\lambda)(2)}, y^{(\lambda)})$  para encontrar  $\phi^{(\lambda)(2)}$  e assim continuando as iterações até que as estimativas  $\beta^{(\lambda)}$  e  $\phi^{(\lambda)}$  sejam obtidas.

As expressões de  $D^{(\lambda)}$ ,  $a$  e  $b$  são facilmente derivadas para muitas distribuições simétricas e podem ser vistas na tabela 1.

Tabela 1: Expressões para  $D^{(\lambda)}$ ,  $a$  e  $b$ .

Distribuição	$D^{(\lambda)}$	$a$	$b$
Normal	$I$	$\frac{1}{4}$	$\frac{3}{4}$
$t$ -Student	$diag \left\{ \frac{\nu+1}{(\nu+u_i^{(\lambda)})} \right\}$	$\frac{(\nu+1)}{4(\nu+3)}$	$\frac{3(\nu+1)}{4(\nu+3)}$
$t$ -Student generalizada	$diag \left\{ \frac{(r+1)}{(s+u_i^{(\lambda)})} \right\}$	$\frac{r(r+1)}{4s(r+3)}$	$\frac{3(r+1)}{4(r+3)}$
Logística I	$2diag \left\{ \tanh \left( \frac{u_i^{(\lambda)}}{2} \right) \right\}$	0,369310044	1,003445984
Logística II	$diag \left\{ \frac{\exp(-\sqrt{u_i^{(\lambda)}}) - 1}{\sqrt{u_i^{(\lambda)}}(1 + \exp(-\sqrt{u_i^{(\lambda)}}))} \right\}$	$\frac{1}{12}$	0,60749
Exponencial potência	$diag \left\{ \frac{1}{(1+k)u_i^{k/(k+1)}(\lambda)} \right\}$	$\frac{\Gamma\{(3-k)/2\}}{4(2^{k-1})(1+k)^2\Gamma\{(k+1)/2\}}$	$\frac{(k+3)}{4(k+1)}$

Substituindo as estimativas de máxima verossimilhança de  $\hat{\beta}^{(\lambda)}$  e  $\hat{\phi}^{(\lambda)}$  em (4.4) obtém-se a log-verossimilhança perfilada para  $\lambda$

$$l_p(\lambda) = -\frac{n}{2} \log\{\hat{\phi}^{(\lambda)}\} + \sum_{i=1}^n \log[g\{\hat{\phi}^{(\lambda)(-1)}(y_i^{(\lambda)} - \hat{\mu}_i^{(\lambda)})^2\}] + \sum_{i=1}^n \log\{J(\lambda, y_i)\}. \quad (4.13)$$

Seja  $s_i^{(\lambda)2} = g\{\hat{\phi}^{(\lambda)(-1)}(y_i^{(\lambda)} - \hat{\mu}_i^{(\lambda)})^2\}$  e definindo as médias geométricas

$$\tilde{s}^{(\lambda)2} = \left( \prod_{i=1}^n s_i^{(\lambda)2} \right)^{1/n} \quad (4.14)$$

e

$$\tilde{J}^{(\lambda)} = \left( \prod_{i=1}^n J(\lambda, y_i) \right)^{1/n}. \quad (4.15)$$

Então, a log-verossimilhança perfilada para  $\lambda$  reduz para

$$l_p(\lambda) = n \log \left( \frac{\tilde{s}^{(\lambda)2} \tilde{J}^{(\lambda)}}{\sqrt{\hat{\phi}^{(\lambda)}}} \right). \quad (4.16)$$

A log-verossimilhança perfilada  $l_p(\lambda)$  plotada versus  $\lambda$  para um série de ensaios de valores determina numericamente o valor da estimativa de verossimilhança  $\hat{\lambda}$ . Para maximizar a log-verossimilhança perfilada só é necessário encontrar um valor de  $\lambda$  que minimize a seguinte taxa

$$\hat{\lambda} = \arg \min_{\lambda} \left( \frac{\sqrt{\hat{\phi}^{(\lambda)}}}{\tilde{s}^{(\lambda)2} \tilde{J}^{(\lambda)}} \right). \quad (4.17)$$

Para modelos simétricos potência,  $\tilde{J}^{(\lambda)} = \tilde{y}^{\lambda-1}$ , em que  $\tilde{y}$  é a média geométrica dos dados originais.

Uma vez que a estimativa de máxima verossimilhança  $\hat{\lambda}$  é obtida, ela pode ser substituída dentro das equações (4.10) e (4.11) para produzir as estimativas irrestritas  $\hat{\beta} = \hat{\beta}^{(\hat{\lambda})}$  e  $\hat{\phi} = \hat{\phi}^{(\hat{\lambda})}$  e então a log-verossimilhança maximizada  $\hat{l} = l_p(\hat{\lambda})$  segue da equação (4.16) para o modelo considerado.

Essencialmente faz-se inferências sobre  $\beta$  e  $\phi$  condicionado em  $\lambda = \hat{\lambda}$ . Então,  $\hat{\lambda}$  é considerada conhecida, e intervalos de confiança para os parâmetros  $\beta$ ,  $\eta$ ,  $\mu$  e  $\phi$ , testes de hipóteses, resíduos e medidas de diagnóstico podem ser obtidas em um contexto similar de modelos lineares generalizados para valores ajustados  $\hat{\beta}$ ,  $\hat{\eta}$ ,  $\hat{\mu}$  e  $\hat{\phi}$ . A covariância assintótica estimada de  $\hat{\beta}$  e  $\hat{\phi}$  dada por  $(4a)^{-1} \hat{\phi} (X^T \hat{H}^2 X)^{-1}$  e  $(4\hat{\phi}^2)/n(4b-1)$  pode ser usada para obter intervalos normais assintóticos para esses parâmetros.

Para realizar testes de hipóteses sobre o parâmetro de transformação da família definida

em (4.1) é possível obter de (4.16) e (4.17) através da estatística  $w = 2\{l_p(\hat{\lambda}) - l_p(\lambda^{(0)})\}$  para testar  $\lambda = \lambda^{(0)}$  a qual tem distribuição assintótica  $\chi_1^2$ . Baseado nesse teste, um intervalo de confiança para  $\lambda$  é facilmente obtido para grandes amostras. A precisão desta aproximação segue do fato de que  $Pr\{w \geq \chi_1^2(\alpha)\} = \alpha + O(n^{-1/2})$ .

Procedimentos de identificação e comparação de modelos têm sido propostos na literatura. Uma nova perspectiva do problema de identificação é obtida pela análise do método de máxima verossimilhança. O fato de que sob certas condições de regularidade o estimador de máxima verossimilhança é assintoticamente eficiente mostra que a função de verossimilhança tende a ser um critério mais sensível a pequenos desvios dos parâmetros do modelo de seus valores verdadeiros. Baseado nesta idéia, Akaike (1972) apresentou um método de identificação de modelos. Em 1974, Akaike descreveu como o problema de seleção de modelos pode sistematicamente ser manuseado pelo uso do critério de informação introduzido em 1972. Esse critério de informação de Akaike (AIC) é uma estatística bem conhecida e de fácil interpretação para seleção de modelos de regressão. Desta forma, para comparar todos os modelos não transformados e transformados ajustados aos dados, pode-se usar o critério de informação de Akaike definido por:

$$AIC = -2\hat{l} + 2r, \quad (4.18)$$

em que  $\hat{l}$  é a log-verossimilhança maximizada e  $r = p + 1$  ou  $r = p + 2$  para os modelos não transformados e transformados, respectivamente. O modelo com o menor valor do AIC, entre todos os modelos ajustados, pode ser considerado como o que melhor explica os dados.

Com o intuito de avaliar melhor os ajustes obtidos calculou-se o erro quadrático médio (EQM) e o erro percentual absoluto médio (MAPE) dados por

$$EQM = \frac{1}{n} \sum_{i=1}^n (y_i^{(\hat{\lambda})} - \hat{\mu}_i)^2 \quad (4.19)$$

e

$$MAPE = \frac{100\%}{n} \sum_{i=1}^n \left| \frac{y_i^{(\hat{\lambda})} - \hat{\mu}_i}{y_i^{(\hat{\lambda})}} \right|. \quad (4.20)$$

## 5 Material e Métodos

No Pólo Gesseiro do Araripe não existiam florestas plantadas de *Eucalyptus*. Em março de 2002 foi implantado na Estação Experimental do Instituto Agrônomo de Pernambuco na Chapada do Araripe - PE, o Módulo de Experimentação Florestal que utilizou 15 clones do gênero *Eucalyptus* em uma área de 2,352 ha, entre híbridos e espécies do gênero *Eucalyptus*. Visando atender a carência de informações, este experimento servirá de base para o aprofundamento na seleção de clones de *Eucalyptus* na região, através de estudos de modelagem de crescimento, mortalidade, sobrevivência, volumetria, calorimetria, entre outros. Nessa dissertação a análise é realizada para um dos clones referente ao tratamento Híbrido de *Eucalyptus urophylla* x *Eucalyptus tereticornis* x *Eucalyptus pellita* (polinização controlada).

A região na qual o experimento foi realizado possui as seguintes coordenadas geográficas de posição 07°29'00 S e 40°36'00 W e altitude de 816 metros. O clima regional predominante é do tipo BShw' de Koppen, quente e seco das baixas latitudes, com chuvas de verão. A presença da chapada confere uma diferenciação no clima regional, elevando a pluviometria e registrando cerca de 735 mm anuais. A temperatura média é 24,6° C, a máxima é de 33° C e a mínima de 15,9° C, com cerca de 2°C de amplitude térmica. O solo é do tipo latossolo vermelho-amarelo (ENCARNAÇÃO, 1980; ARAUJO, 2004).

Foram utilizadas para este trabalho 83 árvores, sobreviventes das 100 árvores plantadas no início do experimento, sendo 25 árvores em cada uma das quatro repetições. A variável altura(H) foi medida em todas as árvores ao longo do tempo, durante 6 anos e meio, sendo denominada como H1 a H14, indicando as 14 medidas obtidas conforme expresso na Tabela 2. O tempo inicial considerado foi de 2 meses que corresponde à idade em que as mudas foram plantadas no campo. As mudas dos clones de *Eucalyptus* foram provenientes da Comercial Agrícola Paineiras LTDA, localizada no município de Urbano Santos, no estado do Maranhão.

Tabela 2: Idade dos clones de *Eucalyptus* com relação a altura.

Medidas	$H_i$	t (idade em meses)
01	$H_1$	2
02	$H_2$	6
03	$H_3$	12
04	$H_4$	18
05	$H_5$	24
06	$H_6$	30
07	$H_7$	36
08	$H_8$	42
09	$H_9$	48
10	$H_{10}$	54
11	$H_{11}$	60
12	$H_{12}$	66
13	$H_{13}$	72
14	$H_{14}$	78

A variável circunferência foi medida à altura do peito (CAP), convencionado como sendo a 1,30 m. Essa variável foi medida em todas as árvores ao longo do tempo, durante 5 anos e meio, sendo denominada como C1 a C10, indicando as 10 medidas obtidas conforme expresso na Tabela 3. Para a variável circunferência o tempo inicial considerado foi de 12 meses, pois se considerou um limite mínimo de medição igual a 6,0 centímetros de CAP.

Tabela 3: Idade dos clones de *Eucalyptus* com relação à CAP.

Medidas	$C_i$	t (idade em meses)
01	$C_1$	12
02	$C_2$	18
03	$C_3$	24
04	$C_4$	30
05	$C_5$	36
06	$C_6$	42
07	$C_7$	48
08	$C_8$	54
09	$C_9$	60
10	$C_{10}$	66

Para ter êxito no reflorestamento deve-se seguir um planejamento florestal no qual são necessárias as informações sobre o crescimento e a produtividade dos povoamentos florestais (RIBEIRO et al.,2001). O ajuste de curvas de crescimento altura-idade ou circunferência-idade é uma ferramenta fundamental no planejamento do inventário florestal, pois permite classificar e prever a produção futura de sítios florestais.

Para estimar a altura e circunferência dos *Eucalyptus* foi utilizado o modelo de Chapman-Richards dado por:

$$W(t)^{(\lambda)} = U(1 - \exp(-kt))^{\theta} + e, \quad (5.1)$$

em que  $U$ ,  $k$  e  $\theta$  ( $\theta = \frac{1}{1-m}$ ) são parâmetros a serem estimados e  $t$  corresponde a idade da árvore em meses e  $e$  o erro associado. Os modelos foram ajustados supondo diferentes distribuições simétricas para os erros (normal, t de Student, Cauchy, exponencial potência, logística I e logística II). Segue abaixo a estrutura dos erros considerando a distribuição t de Student

$$g(u) = \frac{\nu^{\nu/2}}{B(1/2, \nu/2)} (\nu + u)^{-\frac{\nu+1}{2}}, \quad \nu > 0, u > 0, \quad (5.2)$$

em que  $B(\cdot, \cdot)$  é a função Beta. Para demais distribuições da estrutura dos erros ver Capítulo 2.

Inicialmente foram calculadas as estimativas dos parâmetros através dos modelos

simétricos transformados e não transformados supondo diferentes distribuições simétricas. O parâmetro  $\lambda$  da transformação de Box-Cox foi estimado usando a verossimilhança perfilada.

## 6 Resultados

### 6.1 Modelo de Chapman-Richards para a variável altura

Inicialmente foi realizada uma análise descritiva dos dados com o intuito de compreender melhor o comportamento da variável em estudo. A Figura 10 apresenta a curva de crescimento para cada uma das quatro repetições. Observa-se um comportamento semelhante para as 4 repetições.

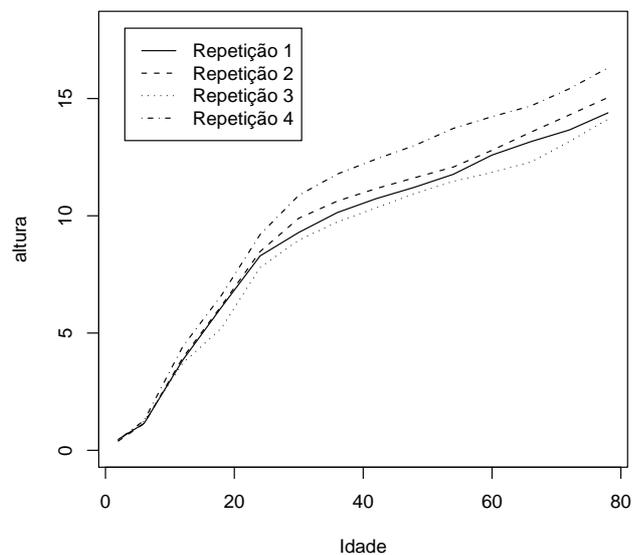


Figura 10: Gráfico de linha das repetições ao longo do tempo

Na Tabela 4 observa-se algumas estatísticas descritivas, entre elas a média, o desvio padrão e o coeficiente de variação.

Tabela 4: Estatísticas descritivas das alturas dos *Eucalyptus*.

Medidas	t (idade em meses)	Média	D.P.	C.V.
01	2	0,4308	0,0340	7,8954
02	6	1,2011	0,0672	5,5937
03	12	3,9855	0,3094	7,7634
04	18	6,0204	0,5868	9,7468
05	24	8,4405	0,5820	6,8959
06	30	9,7553	0,8393	8,6039
07	36	10,5718	0,8850	8,3715
08	42	11,1614	0,9027	8,0874
09	48	11,6985	0,9108	7,7856
10	54	12,2620	1,0080	8,2205
11	60	12,8648	0,9945	7,7302
12	66	13,4244	0,9922	7,3913
13	72	14,1362	0,9725	6,8793
14	78	14,9660	0,9757	6,5196

Para os dados da variável altura a transformação de Box-Cox utilizada foi  $Y^{(\lambda)} = (Y + c)^\lambda$ . Devido a natureza dos dados optou-se por adicionar uma constante ( $c = 2$ ) aos dados, uma vez que uma constante não interfere na dispersão da variável em estudo.

A Figura 11 apresenta os valores de AIC versus os graus de liberdade para o modelo transformado supondo erro com distribuição t de Student (equação 5.1), de modo que é possível visualizar para que grau de liberdade obtém-se o menor AIC. Para os dados de altura o modelo transformado supondo erros com distribuição t de Student com 2 graus de liberdade ( $t_2$ ) obteve o menor valor do critério AIC.

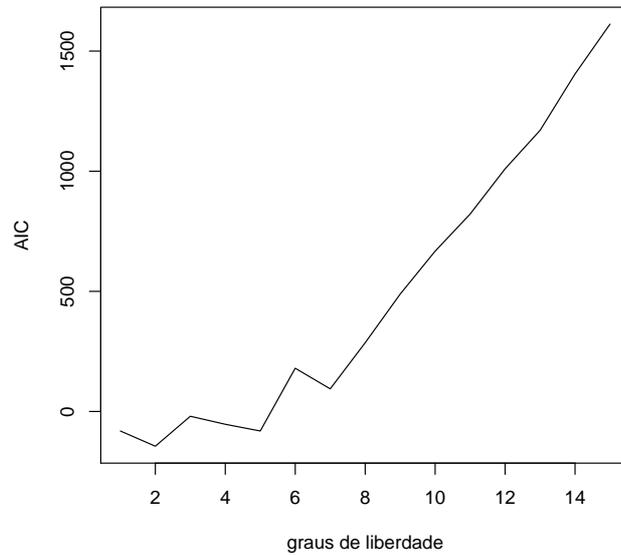


Figura 11: Graus de Liberdade da  $t_2$  transformada versus AIC

A Tabela 5 apresenta as estimativas dos parâmetros dos modelos transformados e não transformados ( $\lambda = 1$ ), bem como do parâmetro de dispersão  $\phi$  condicionado em  $\lambda$  fixo ou estimado. Observa-se que o único parâmetro não significativo foi o parâmetro  $k$  para o modelo exponencial potência transformado.

Tabela 5: Estimativas dos parâmetros dos modelos transformados e não transformados.

Distribuições	$\lambda$	U	k	$\theta$	$\phi$
Normal	1	15,009354	0,041073	1,396629	0,644726
	0,349525	20,863111	0,000000548	0,201941	0,006263
Student $t_2$	1	14,571636	0,044017	1,471376	0,335500
	0,292516	6,456838	0,000019222	0,158514	0,001648
Exp. Potência (0,5)	1	14,795644	0,042753	1,443735	0,264620
	0,556131	29,043895	0,000033368 <sup>(*)</sup>	0,29888	0,015517
Cauchy	1	14,424669	0,044146	1,466475	0,207607
	0,621003	18,629501	0,000392	0,33124	0,023132
Logística I	1	15,387640	0,037299	1,273507	0,832586
	0,491116	18,499327	0,000064238	0,283212	0,029651
Logística II	1	14,793052	0,042823	1,443189	0,213413
	0,503790	19,571692	0,000052663	0,278970	0,008376

(\*) parâmetro não significativo ao nível de 5%.

Para escolher o modelo que melhor se ajustou aos dados apresentamos na Tabela 6 a log-verossimilhança maximizada  $\hat{l}$  e o critério AIC. O modelo  $t_2$  transformado foi o que apresentou a maior log-verossimilhança (77, 2998) e o menor AIC (-144, 5996), sendo assim o modelo mais indicado para representar os dados de acordo com esses critérios.

Tabela 6: Estatísticas para seleção dos modelos.

Distribuições	$\lambda$	$\hat{l}$	AIC
Normal	1	-15,7099	39,4198
	0,349525	42,8705	-75,7410
Student $t_2$	1	6,2366	-4,4732
	0,292516	77,2998	-144,5996
Exp. Potência (l=0,5)	1	-67,9845	143,9690
	0,556131	-37,7118	85,4236
Cauchy	1	-12,5898	33,1796
	0,621003	9,9590	-9,9180
Logística I	1	-66,9973	141,9946
	0,491116	-28,8508	67,7016
Logística II	1	-67,9859	143,9718
	0,503790	-31,6358	73,2716

Na Tabela 7 apresentamos os valores das estatísticas EQM e MAPE. De acordo com os resultados obtidos o modelo  $t_2$  transformado continua sendo o mais adequado pois apresenta menores valores para os dois critérios apresentados. É possível observar que em comparação com os modelos não transformados, os modelos transformados apresentaram menores valores para esses critérios.

Tabela 7: Comparação dos erros para os modelos ajustados.

Distribuições	$\lambda$	EQM	MAPE (%)
Normal	1	0,6447	9,5224
	0,349525	0,0064	3,1550
Student $t_2$	1	0,6658	9,1782
	0,292516	0,0035	2,4372
Exp. Potência ( $l=0,5$ )	1	0,6487	9,1143
	0,556131	0,0387	4,6516
Cauchy	1	0,6966	9,1625
	0,621003	0,0664	5,1585
Logística I	1	0,6564	12,2361
	0,491116	0,0235	4,4055
Logística II	1	0,6428	9,1355
	0,503790	0,0256	4,3226

Na área de Ciências Florestais um dos mais importantes elementos para avaliação de um modelo é a visualização dos seus resíduos percentuais dados por

$$E\% = \left( \frac{y_i^{(\hat{\lambda})} - \hat{\mu}_i}{y_i^{(\hat{\lambda})}} \right). \quad (6.1)$$

De forma geral erros percentuais em torno de 20% são considerados aceitáveis. Considerando que, em geral, o modelo normal é o mais utilizado para dados de crescimento, nas Figuras 12 e 13 são apresentados os erros percentuais versus o índice das observações para os ajustes através das distribuições normal e  $t_2$  transformada. Verifica-se que para o  $t_2$  transformado os erros percentuais estão mais próximos de zero e o maior valor apresentado foi de aproximadamente  $-10\%$  ao passo que para o normal o maior valor encontrado foi de aproximadamente  $-60\%$ .

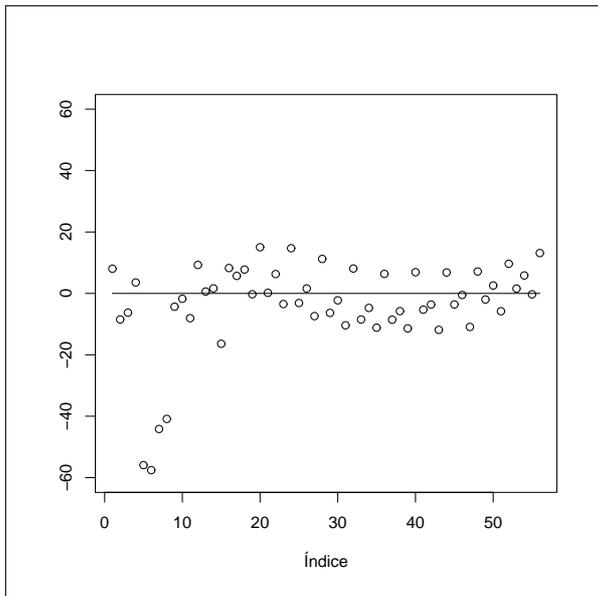


Figura 12: *Erros percentuais para o modelo normal*

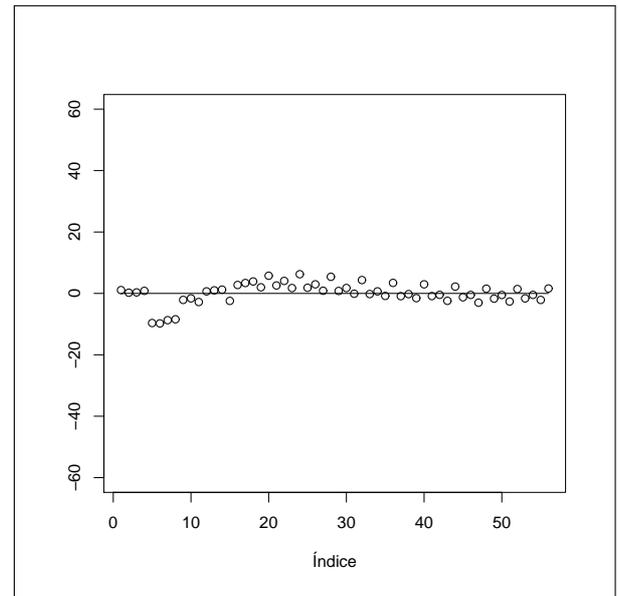


Figura 13: *Erros percentuais para o modelo  $t_2$  transformado*

Na prática, é de interesse que as estimativas estejam na escala original. Nas Figuras 14 e 15 são apresentados os erros percentuais versus o índice das observações para os ajustes através das distribuições normal e  $t_2$  transformada na escala original, ou seja, para a transformação utilizada  $Y^{(\lambda)} = (Y + c)^\lambda$  retorna-se para a escala original  $Y = (Y^{(\lambda)})^{1/\lambda} - c$ .

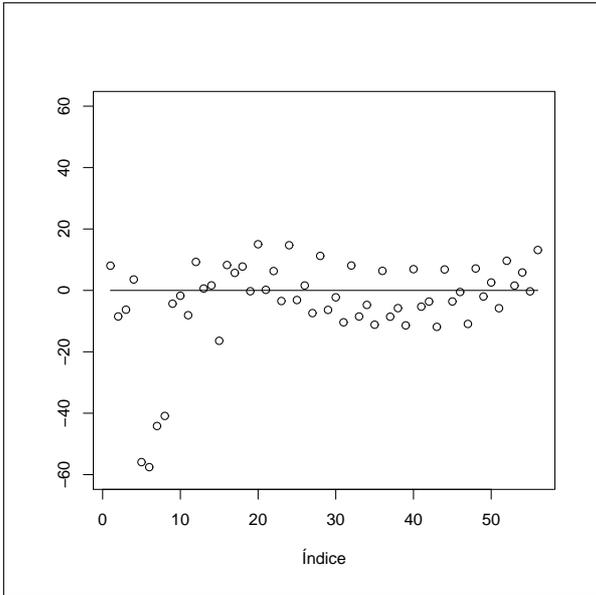


Figura 14: Erros percentuais para o modelo normal

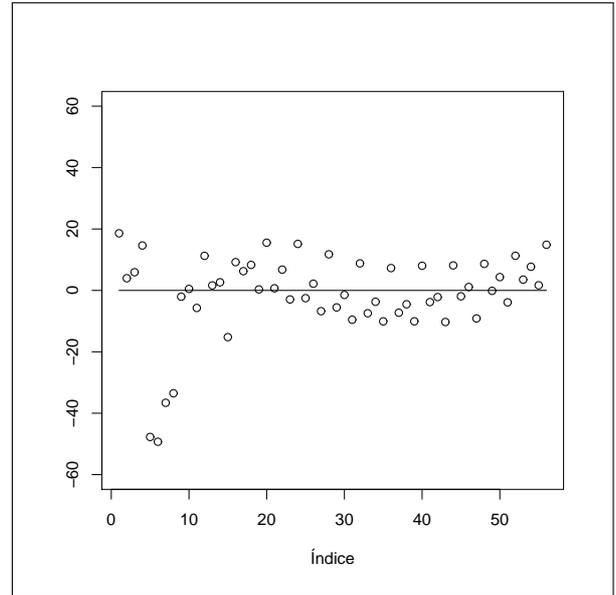


Figura 15: Erros percentuais para o modelo  $t_2$  transformado na escala original

Na tabela 8 são apresentados os valores das estatísticas EQM e MAPE para os ajustes através das distribuições normal e  $t_2$  transformada na escala original. Observa-se com os resultados obtidos que o modelo  $t_2$  transformado na escala original obteve menor valor para o MAPE, enquanto o modelo normal obteve menor valor para o EQM. O modelo  $t_2$  transformado na escala original apresentar maior EQM pode ser explicado pelo fato do erro quadrático médio levar em consideração o quadrado do erro. Isto é, como o modelo  $t_2$  é menos sensível a observações atípicas a tendência é que essas observações se distanciem da estimativa fazendo com que o erro quadrático médio penalize os modelos com caudas mais pesadas.

Tabela 8: Comparação dos erros para os modelos ajustados.

Distribuições	EQM	MAPE (%)
Normal	0,6447	9,5224
Student	0,6658	9,1782

## 6.2 Modelo de Chapman-Richards para a variável circunferência

A Figura 16 apresenta o crescimento da circunferência dos *Eucalyptus* para cada uma das quatro repetições. Observa-se um comportamento semelhante para as 4 repetições, no entanto, a repetição 4 apresenta um crescimento superior as demais repetições para todos os tempos estudados, fato que pode influenciar nas estimativas dos parâmetros quando utilizado o modelo de Chapman-Richards supondo para o erro distribuição normal.

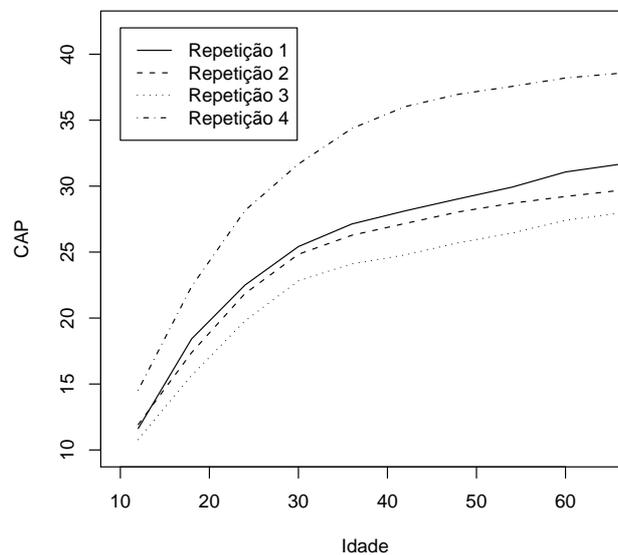


Figura 16: Gráfico de linha das repetições ao longo do tempo

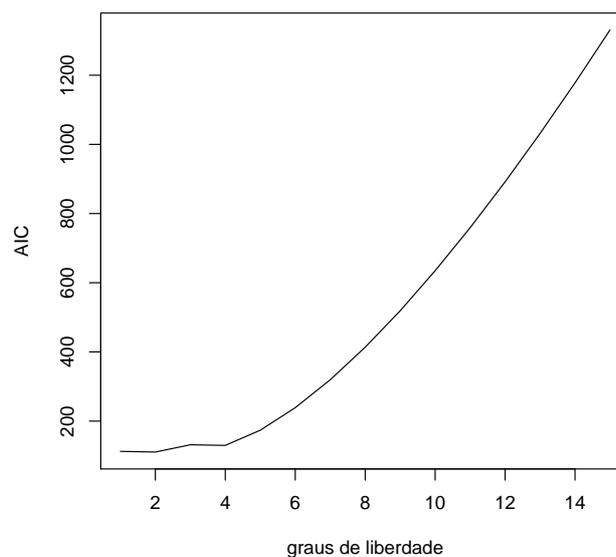
Na Tabela 9 observa-se algumas estatísticas descritiva, são elas a média, o desvio padrão e o coeficiente de variação.

Tabela 9: Estatísticas descritivas das circunferências dos *Eucalyptus*.

Medidas	t (idade em meses)	Média	D.P.	C.V.
01	12	12,2188	1,6183	13,2441
02	18	18,4588	2,8597	15,4924
03	24	23,0743	3,5796	15,5133
04	30	26,1893	3,8235	14,5996
05	36	27,9723	4,4496	15,9072
06	42	29,0300	4,8742	16,7901
07	48	29,9440	4,8869	16,3201
08	54	30,6583	4,8234	15,7328
09	60	31,4760	4,7254	15,0127
10	66	31,9565	4,6509	14,5538

Para os dados da variável circunferência a transformação de Box-Cox utilizada foi  $Y^\lambda = (Y^\lambda - 1)/\lambda$  quando  $\lambda \neq 0$  ou  $\log(Y)$  quando  $\lambda = 0$ .

A Figura 17 apresenta os valores de AIC versus os graus de liberdade para o modelo  $t$  de Student transformado. Para os dados de circunferência o modelo  $t_2$  transformado obteve o menor valor do critério AIC.

Figura 17: Graus de liberdade da  $t_2$  transformada versus AIC

A Tabela 10 apresenta as estimativas dos parâmetros dos modelos transformados e não transformados, bem como do parâmetro de dispersão  $\phi$  condicionado em  $\lambda$  fixo ou estimado.

Tabela 10: Estimativas dos parâmetros dos modelos transformados e não transformados.

Distribuições	$\lambda$	U	k	$\theta$	$\phi$
Normal	1	31,976838	0,073007	1,762786	13,035169
	0,683641	14,101273	0,075066	1,431948	1,519147
Student $t_2$	1	30,670557	0,073250	1,701564	7,398160
	0,716904	14,606708	0,076263	1,466910	0,611377
Exp. Potência (0,1)	1	31,721566	0,072329	1,740883	10,954037
	0,836213 <sup>(*)</sup>	19,771051	0,072441	1,562938	1,628747
Cauchy	1	29,767306	0,079154	1,898140	1,589942
	0,880181 <sup>(*)</sup>	21,341953	0,081329	1,871679	0,741396
Logística II	1	31,1358	0,073349	1,728695	4,051503
	0,562828	10,451776	0,076126	1,261740	0,209691

(\*) parâmetros não significativo ao nível de 5%.

Para a seleção dos modelos ajustados, apresentamos na Tabela 11 a log-verossimilhança maximizada  $\hat{l}$  para os valores de  $\lambda$  correspondente e o critério AIC. Assim como os dados de altura, para os dados de circunferência o modelo supondo erro com distribuição  $t_2$  transformado obteve maior valor da log-verossimilhança  $\hat{l}$  (-51,0629) e um menor valor do AIC (112,1258).

Tabela 11: Estatísticas para seleção dos modelos.

Distribuições	$\lambda$	$\hat{l}$	AIC
Normal	1	-71,3530	150,7060
	0,683641	-69,3773	148,7546
Student $t_2$	1	-85,3460	178,6920
	0,716904	-51,0629	112,1258
Exp. Potência (I=0,1)	1	-108,1370	224,2740
	0,836213	-106,4585	222,9170
Cauchy	1	-61,3277	130,6554
	0,880181	-61,1704	132,3408
Logística II	1	-108,3575	224,7150
	0,562828	-106,1884	222,3768

De acordo com a Tabela 12 o ajuste do modelo  $t_2$  transformado apresentou menor valor para o MAPE enquanto o modelo normal transformado apresentou um menor valor para o EQM.

Tabela 12: Comparação dos erros para os modelos ajustados.

Distribuições	$\lambda$	EQM	MAPE (%)
Normal	1	13,0352	10,6389
	0,683641	1,5464	10,0586
Student $t_2$	1	13,9751	9,3747
	0,716904	2,2085	8,3034
Exp. Potência (I=0,1)	1	13,0945	10,1965
	0,836213	8,0369	11,4087
Cauchy	1	15,3466	9,0134
	0,880181	10,1753	16,5417
Logística II	1	13,4132	9,7222
	0,562828	1,5773	13,1988

Nas Figuras 18 e 19 são apresentados os erros percentuais versus o índice das observações. É possível verificar que ambos os ajustes obtiveram erros percentuais aceitáveis (< 20%). Entretanto verifica-se que o ajuste para o modelo  $t_2$  transformado apresenta erros percentuais menores que o ajuste para o modelo normal.

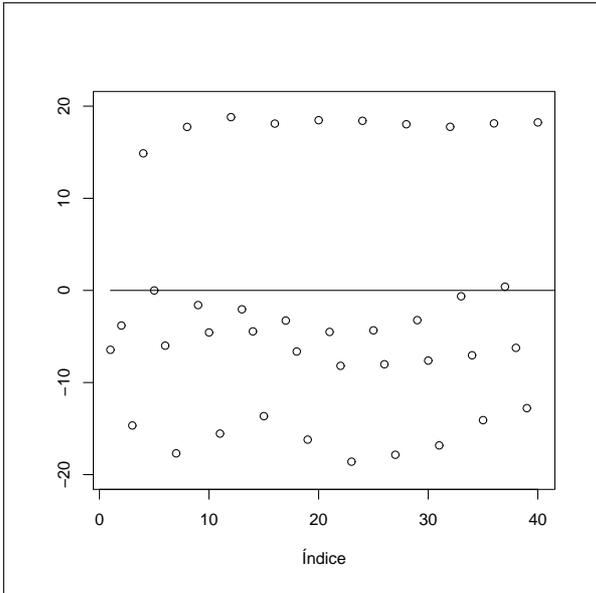


Figura 18: *Erros percentuais para o modelo normal*

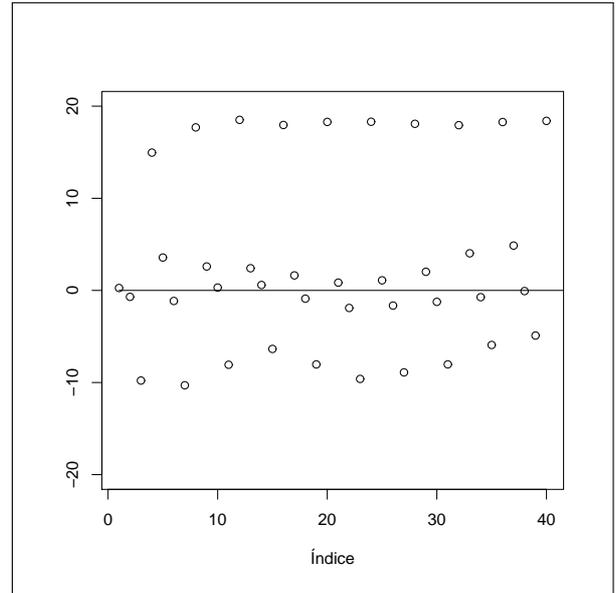


Figura 19: *Erros percentuais para o modelo  $t_2$  transformado*

Nas Figuras 20 e 21 são apresentados os erros percentuais versus o índice das observações para os ajustes através das distribuições normal e  $t_2$  transformada na escala original, ou seja, para a transformação utilizada  $Y^{(\lambda)} = (Y^\lambda + 1)/\lambda$  retorna-se para a escala original  $Y = (\lambda \cdot Y^{(\lambda)} - 1)^{1/\lambda}$ . Observa-se que o modelo  $t_2$  transformada na escala original obteve melhores estimativas quando comparado com o modelo normal.

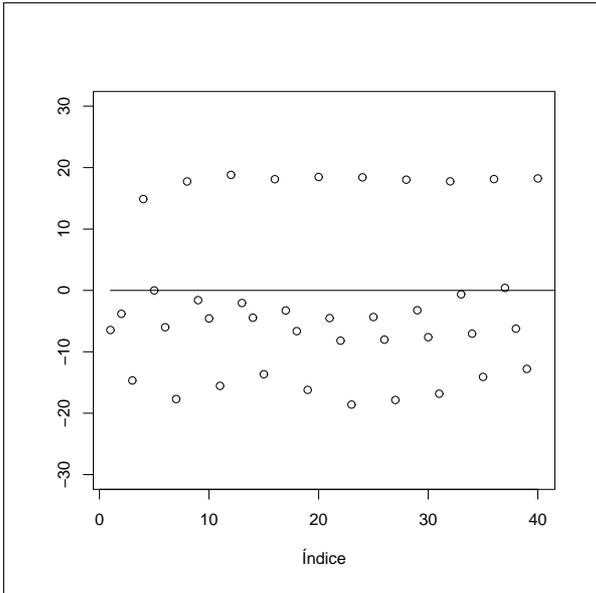


Figura 20: Erros percentuais para o modelo normal

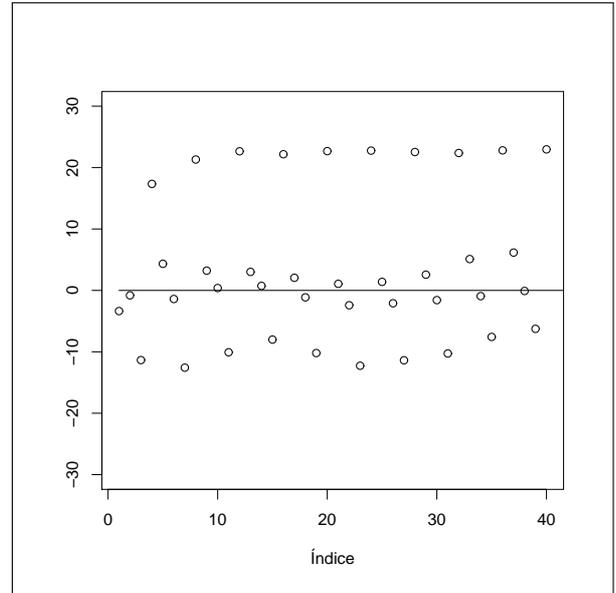


Figura 21: Erros percentuais para o modelo  $t_2$  transformado na escala original

Na tabela 13 são apresentados os valores das estatísticas EQM e MAPE para os ajustes através das distribuições normal e  $t_2$  transformada na escala original. Observa-se com os resultados obtidos que o modelo  $t_2$  transformado na escala original obteve menor valor para o MAPE, enquanto o modelo normal obteve menor valor para o EQM. Esse fato ocorre devido ao mesmo motivo exposto na análise dos dados de altura.

Tabela 13: Comparação dos erros para os modelos ajustados.

Distribuições	EQM	MAPE (%)
Normal	13,0352	10,6389
Student $t_2$	15,0853	9,0857

## 7 Conclusão

Sabendo da importância ambiental, econômica e social do *Eucalyptus*, estudos nessa área são fundamentais na prognose da produtividade, qualidade do local e dinâmica de populações. A modelagem da altura e circunferência do *Eucalyptus* são feitas através dos modelos de crescimento. Na prática, esses modelos não-lineares, desenvolvidos empiricamente para relacionar altura e idade ou circunferência e idade, têm-se mostrado adequados para descrever a curva de crescimento. O modelo de Chapman-Richards com distribuição normal para os erros é um modelo bastante utilizado para este fim. Contudo, alternativas à suposição de erros normais têm sido propostas na literatura ao longo dos anos com o intuito de tentar reduzir a influência de outliers nas estimativas dos coeficientes. Dessa forma, nessa tese foi apresentada a transformação de Box e Cox para o modelo de Chapman-Richards supondo diferentes distribuições simétricas para os erros.

Na primeira aplicação foi estudada a altura das árvores em função da sua idade. Conclui-se que para os dados de altura o modelo  $t_2$  transformado foi o modelo que melhor se ajustou aos dados. De acordo com os critérios utilizados o modelo  $t_2$  transformado obteve AIC e MAPE inferiores ao obtido pelo modelo normal.

Na segunda aplicação foi estudada a circunferência das árvores em função da sua idade. Pelas estatísticas de comparação dos erros o modelo normal obteve o menor valor para o EQM enquanto que o modelo  $t_2$  transformado obteve o menor valor para o MAPE. De acordo com o gráfico dos erros percentuais fica evidente que o modelo  $t_2$  transformado obteve melhores estimativas.

Após todas as análises dos modelos, observou-se que para as duas variáveis estudadas o modelo  $t_2$  transformado obteve melhores resultados do que o modelo normal. Dessa forma, é possível concluir que os modelos simétricos transformados são bastantes relevantes para os estudos de modelos de crescimento de forma prática e bastante úteis para análise de dados reais. Espera-se que essa dissertação possa contribuir de forma efetiva no sentido de ampliar as possibilidades de análise para os modelos de crescimento adotados em Ciências Florestais objetivando assim estimativas mais precisas.

## Referências

- ACHGAR, J. A.; PEREIRA, G. A. Use of exponential power distribution for mixture models in the presence of covariates. **Journal of Applied Statistics**, v. 26, n. 6, p. 669-679, 1999.
- AKAIKE, H. Use of an information theoretic quantity for statistical model identification. **Proceedings of 5th Hawaii International Conference on System Sciences**, p. 249-250, 1972.
- AKAIKE, H. A new look at the statistical model identification. **IEEE Trans. Automatic Control**, v. 19, p. 716-723, 1974.
- ALBUQUERQUE, J.L. **Diagnóstico ambiental e questões estratégicas: Uma análise considerando o Pólo Gesseiro do Sertão do Araripe - Estado de Pernambuco**. 2002. Tese (Doutorado em Ciências Florestais) - Universidade Federal do Paraná, Brazil.
- ANDERSON, T.W.; FANG, K.T. Cochran's theorem for elliptically contoured distributions. **Sankhya**, v. 49, p. 305-315, 1987.
- ANDRADE, E.N. **Manual do plantador de eucalyptos**. São Paulo: Typografia Brazil de Rothschild & Comp., p. 335, 1911.
- ANDREWS, D.F.; BICKEL, P.J.; HAMPEL, F.R.; HUBER, P.J.; ROGERS, W.H.; TUKEY, J.W. **Robust Estimates of Location**. NJ: Princeton University Press, 1972.
- ANGERS, J.F. Use of Student-t Prior for the Estimation of Normal Means : A Computational Approach. **Bayesian Statistics**, v. 4, p. 567-575, 1992.
- ARAUJO, S.M.S. **O Pólo Gesseiro do Araripe: unidades geo-ambientais e impactos da mineração**. 2004. Tese Doutorado em Ciências - Área de Administração e Política de Recursos Minerais - Universidade de Campinas, Brazil.
- ARALLANO-VALLE, R.B. **Distribuições elípticas: propriedades, inferência e aplicações a modelos de regressão**. 1994. Tese (Doutorado em Estatística) - Universidade de São Paulo, Brazil.
- AVERY, T.E.; BURKHART, H. **Forest management**. 4ª Edição, New York: McGraw-Hill, p. 432, 1994.
- BALAKRISHNAN, N. **Handbook of the Logistic Distribution**. New York: Dekker, 1992.
- BATES, D.M.; WATTS, D.G. **Nonlinear regression analysis and its applications**. New York: John Wiley, 1988.
- BEALE, E.M.L. Confidence regions in nonlinear estimation. **Journal of the Royal Statistical Society**. Series B, v. 22, p. 41-88, 1960.

BERKANE, M.; BENTLER, P.M. Moments of elliptical distributed random variates. **Statistics and Probability Letters**, v. 4, p. 333-335, 1986.

BERTALANFFY, L. von. A quantitative theory of organic growth. **Human Biology**, v. 10, p. 181-213, 1934.

BERTALANFFY, L. von. Quantitative laws for metabolism and growth. **Quarterly Review of Biology**, v. 32, p. 217-231, 1957.

BLATTBERG, R.C.; GONEDDES, N.J. A comparison of the stable and Student distributions as statistical models for stock prices. **Journal of Business**, v. 47, p. 244-280, 1974.

BOX, G.E.P.; COX, D.R. An analysis of transformation. **Journal of the Royal Statistical Society**. Series B, v. 26, n. 2, p. 211-252, 1964.

BOX, M.J.; TIAO, G.C. **Bayesian inference in statistical analysis**. London: Addison-Wesley, 1973.

BOX, M.J.; TIAO, G.C. **Bayesian Inference in Statistical Analysis**. New York: John Wiley, 1992.

BUTLER, R.J.; MCDONALD, J.B.; NELSON, R.D.; WHITE, S.B. Robust and partially adaptive estimation of regression models. **Review of Economics and Statistics**, v. 72, p. 321-326, 1990.

CAMBANIS, S.; HUANG, S.; SIMONS, G. On the theory of elliptically contoured distributions. **Journal of Multivariate Analysis**, v. 11, p. 368-385, 1981.

CHAPMAN, D.E. Statistical problems in dynamics of exploited fishing populations. **Berkeley Symposium on Mathematical Statistics and Probability**, p. 153-158, 1961.

CHMIELEWSKI, M.A. Elliptically symmetric distributions: a review and bibliography. **International Statistical Review**, v. 49, p. 67-74, 1981.

CHOY, S.T.B.; SMITH, A.F.M. On Robust Analysis of a Normal Location Parameter. **Journal of the Royal Statistical Society**. Series B, v. 59, p. 463-474, 1997.

CLUTTER, J.L. et al. **Timber management: a quantitative approach**. New York: John Wiley, 1983.

CLUTTER, J.L.; JONES, E.P. **Prediction of growth after thinning in old-field slash pine plantation**. Washington, D.C. USDA, 1980.

COOK, R.D.; TSAI, C.L. Residuals in nonlinear regression. **Biometrika**, v. 72, p. 23-29, 1985.

CORDEIRO, G.M. Corrected LR tests in symmetric nonlinear regression models. **Journal of Statistical Computation and Simulation**, v. 74, n. 5, p. 609-620, 2004.

CORDEIRO, G.M.; FERRARI, S.L.P.; OPAZO, M.U.; VASCONCELLOS, K.L.P. Corrected maximum likelihood estimation in a class of symmetric nonlinear regression models. **Statistics and Probability Letters**, v. 46, p. 317-328, 2000.

CORDEIRO, G.M.; ANDRADE, M.G. Transformed symmetric models. **Statistical Modelling** (artigo submetido), 2009.

COX, D.R.; SNELL, E.J. A general definition of residuals. **Journal of the Royal Statistical Society**. Series B, v. 30, p. 248-275, 1968.

CYSNEIROS, F.J.A.; PAULA, G.A. Restricted Methods in Symmetrical Linear Regression Models. **Computational Statistics and Data Analysis**. V. 49, n. 3, p. 689-708, 2005.

DAVIS, L.S.; JOHNSON, K.N. **Forest management**. 3<sup>a</sup> ed. New York:McGraw-Hill, 1987.

DEVLIN, S.J.; GNANADESIKAN, R.; KETTENRING, J.R. Some multivariate applications of elliptical distributions. **Essays in Probability and Statistics**, v. 24, p. 365-393, 1976. Tokyo: Shinko Tsusho Co.

DEVROYE, L. **Non-uniform random variable generator**. New York: Springer-Verlag, 1986.

DICKEY, J.M. Multivariate generalizations of the multivariate t distribution and the inverted multivariate t distribution. **Annals of Mathematical Statistics**, v. 38, p. 511-518, 1967.

DRAPER, N.R.; SMITH, H. **Applied regression analysis**. New York: John Wiley, 1981.

EASTERLING, R.G. Exponential responses with double exponential measurement error - A model for steam generator inspection. **Proceedings of the DOE Statistical Symposium**, U.S. Department of Energy, p. 90-110, 1978.

ENCARNAÇÃO, C.R.F. **Observações meteorológicas e tipos climáticos das unidades e campos experimentais da empresa IPA**. Recife: IPA, 1980.

FANG, K.T.; ANDERSON, T.W. **Statistical inference in elliptical contoured and related distributions**. New York: Allerton Press, 1990.

FANG, K.T.; KOTZ, S.; NG, K.W. **Symmetric multivariate and related distributions**. London: Chapman and Hall, 1990.

FANG, K.T.; ZHANG, Y.T. **Generalized multivariate analysis**. New York: Springer-Verlag, 1990.

FELLER, W. **An Introduction to Probability Theory and Its Applications**, v. 2, New York: Wiley, 1966.

FERRARI, S.L.P.; URIBE-OPAZO, M.A. Corrected likelihood ratio tests in a class of symmetric linear regression models. **Brazilian Journal of Probability and Statistics**, v. 15, p. 49-67, 2001.

FISK, P.R. The graduation of income distributions. **Econometrica**, v. 29, p. 171-185, 1961b.

FRANCE, J.; THORNLEY, J.H.M. **Mathematical models in agriculture**. London: Butterworths, 1984.

GALEA, M.; PAULA, G.A.; CYSNEIROS, F.J.A. On Diagnostic in Symmetrical Nonlinear Models. **Statistics and Probability Letters**. V. 73, n. 4, p. 459-467, 2005.

- GAUSS, C.F. **Theoria motus corporum coelestium**. Hamburg: Perthes and Besser, 1809.
- GAUSS, C.F. Bestimmung der Genauigkeit der Beobachtungen. **Zeitschrift Astronomi**, v. 1, p. 185-197, 1816.
- GREENWOOD, J.A.; OLKIN, I.; SAVAGE, I.R. **Index to Annals of Mathematical Statistics**. Volumes 1-31, p. 1930-1960. University of Minnesota, Minneapolis St. Paul: North Central Publishing.
- GUMBEL, E.J. Ranges and midranges. **Annals of Mathematical Statistics**, v. 15, p. 414-422, 1944.
- GUMBEL, E.J.; KEENEY, R.D. The extremal quotient. **Annals of Mathematical Statistics**, v. 21, p. 523-538, 1950.
- GUPTA, A.K.; VARGA, T. Elliptically contoured models in statistics. **Kluwer Academic Publishers**, 1993.
- HASSE, G. **Eucalypto: histórias de um imigrante vegetal**. Porto Alegre: Já Editores, p. 127, 2006.
- HASTINGS, N.A.J.; PEACOCK, J.B. **Statistical distributions**. New York: John Wiley, 1975.
- HOAGLIN, D.C.; MOSTELLER, F.; TUKEY, J.W. **Exploring Data Tables, Trends and Shapes, Chapter 9**. New York: Wiley, 1985.
- HSU, D.A. Long-tailed distributions for position errors in navigation. **Applied Statistics**, v. 28, p. 62-72, 1979.
- JENRICH, R.I. Asymptotic properties of nonlinear least squares estimation. **Annals of Mathematical Statistics**, v. 40, p. 633-643, 1969.
- JOHANSEN, S. Some topics in regression. **Scand J. Stat**, v. 10, p. 161-194, 1983.
- JOHNSON, R.; KOTZ, S. **Continuous univariate distributions**. V.2. Boston: Houghton Mifflin, 1970.
- KELKER, D. Distribution theory of spherical distributions and a location scale parameter generalization. **Sankhya**, v. 32, p. 419-430, 1970.
- KRZANOWSKI, W.J.; MARRIOTT, F.H.C. **Multivariate Analysis**. London: Edward Arnold, 1994.
- LANGE, K.L.; LITTLE, R.J.A.; TAYLOR, J.M.G. Robust statistical modeling using the t distribution. **Journal of the American Statistical Association**, v. 84, p. 881-896, 1989.
- LAPLACE, P.S. Déterminer le milieu que l'on doit prendre entre trois observations données d'un même phénomène. **Mémoires de Mathématique et Physique présentées à l'Académie Royale des Sciences par divers Savans**, v. 6, p. 621-625, 1774.
- LAPLACE, P.S. **Théorie analytique des probabilités**. Paris: Veuve Courcier, 1812.

- LIMA, W.P. **Impacto ambiental do eucalipto**. 2ª Edição, São Paulo: EDUSP, p. 301, 1993.
- LITTLE, R.J.A. Robust estimation of the mean and covariance matrix from data with missing values. **Applied Statistics**, v. 37, p. 23-39, 1988.
- LUNETTA, G. Di una Generalizzazione dello Schema della Curva Normale. **Annali della Facoltà di Economia e Commercio di Palermo**, v. 17, p. 237-244, 1963.
- MACHADO, S.A. **Studies in growth and yield estimation for *Pinus taeda* L. plantations in the State of Parana - Brazil**. 1978. Thesis (Doctor of Philosophy) - University of Washington, Washington.
- MANLY, B.F. Exponential data transformation. **The Statistician**, v. 25, p. 37-42, 1976.
- MANOUKIAN, E.B. **Modern concepts and theorems of mathematical statistics**. New York:Springer-Verlag, 1985.
- MANTEL, N.; PASTERNAK, B.S. Light bulb statistics. **Journal of the American Statistical Association**, v. 61, p. 633-639, 1966.
- MCDONALD, J.B. Some generalized functions for the size distribution of income. **Econometrica**, v. 52, p. 647-663, 1984.
- MCDONALD, J.B.; NEWAY, W.K. Partially adaptive estimation of regression models via the generalized  $t$  distribution. **Econometric Theory**, v. 4, p. 428-457, 1988.
- MCDONALD, J.B. Partially adaptive estimation of ARMA time series models. **International Journal of Forecasting**, v. 5, p. 217-230, 1989.
- MCDONALD, J.B. Parametric models for partially adaptive estimation with skewed and leptokurtic residuals. **Economics Letters**, v. 37, p. 237-278, 1991.
- MCLEAY, S. Student's  $t$  and the distribution of financial ratios. **Journal of Business Finance and Accounting**, v. 13, p. 209-222, 1986.
- MOIVRE, A. de. Approximatio ad summam ferminorum Binomii in seriem expansi. **Supplementum II to Miscellanae Analytica**, p. 1-7, 1733.
- MOIVRE, A. de. **The doctrine of chances**. London: Frank Cass and Co, 1738.
- MONTGOMERY, D.C.; PECK, E.A. **Introduction to linear regression analysis**. New York: John Wiley, 1982.
- MUIRHEAD, R. The effects of symmetric distributions on some standard procedures involving correlation coefficients. **Multivariate Statistical Analysis North-Holland**, p. 143-159, 1980.
- MUIRHEAD, R. **Aspects of multivariate statistical theory**. New York: John Wiley, 1982.
- NYQUIST, H.; RICE, S.O.; RIORDAN, J. The distribution of random determinants. **Quarterly of Applied Mathematics**, v. 42, p. 97-104, 1954.
- OKUBO, T.; NARITA, N. On the distribution of extreme winds expected in Japan. **National Bureau of Standards Special Publication 560-1**, 12 pp, 1980.

ORD, J.K. Laplace distribution, **In Encyclopedia of Statistical Sciences**, v. 4, New York: Wiley, 1983.

PEARL, R.; REED, L.J. On the rate of growth of the population of the United States since 1970 and its mathematical representation. **Proceedings of the National Academy of Sciences**, v. 6, p. 275-288, 1920.

PEARL, R.; REED, L.J. **Studies in Human Biology**. Baltimore: Williams and Wilkins, 1924.

PEARL, R.; REED, L.J.; KISH, J.F. The logistic curve and the census count of 1940. **Science**, v. 92, p. 486-488, 1940.

PEARSON, K. Contributions to the mathematical theory of evolution. II . skew variation in homogeneous material. **Philosophical Transactions of the Royal Society of London**, v. 186, p. 343-414, 1895.

PIENAAR, L.V.; TURNBULL, K.J. The Chapman-Richards generalization of von Bertalanffy's growth model for basal area growth and yield in even-aged stands. **Forest Science**, v. 19, p. 02-22, 1973.

PLACKETT, R.L. The analysis of life-test data. **Technometrics**, v. 1, p. 9-19, 1959.

PRAETZ, P.D. The distribution of share price changes. **Journal of Business**, v. 45, p. 49-55, 1972.

PRAETZ, P.D.; WILSON, E.J.G. The distribution of stock market returns: 1958-1973. **Australian Journal of Management**, v. 3, p. 79-90, 1978.

PRODAN, M. **Forest biometrics**. [S.l]: Pergamon Press, 1968.

RAO, B.L.S.P. Remarks on univariate symmetric distributions. **Statistics and Probability Letters**, v. 10, p. 307-315, 1990.

RIBEIRO, C.A.S. et al. Seleção de modelos volumétricos para leucena no agreste de Pernambuco. **Brasil Florestal**. Brasília, DF , ano 20, v.72, 2001.

RICHARDS, F.J. A flexible growth function of empirical use. **Journal of Experimental Botany**, v. 10, p. 290-300, 1959.

SAKIA, R.M. The Box-Cox transformation technique: a review. **The Statistician**, v. 76, v. p. 259-265, 1992.

SCHULTZ, H. The standard error of a forecast from a curve. **Journal of the American Statistical Association**, v. 25, p. 139-185, 1930.

SCOLFORO, J.R. **Modelos para expressar o crescimento e a produção florestal: Parte 1**. Lavras: ESAL / FAEPE, 1994.

SEBER, G.A.F.; WILD, C.J. **Nonlinear regression**. New York: John Wiley, 2003.

SILVA, J.A.A. **Dynamics of stand structure in fertilized slash pine plantations**. 1986. Thesis (Doctor of Philosophy) - University of Georgia, Athens.

SIMONE, S. de. Su una estensione dello schema delle curve normali di ordine  $r$  alle variabili doppie. **Statistica**, v. 37, p. 447-474, 1968.

SPATHELF, P; NUTTO, L. **Modelagem aplicada ao crescimento florestal**. Santa Maria:[s.n], 2000. p. 70.

"STUDENT". On the probable error of the mean. **Biometrika**, v. 6, p. 1-25, 1908.

"STUDENT". New tables for testing the significance of observations. **Metron**, v. 5, p. 105-108, p. 114-120, 1925.

SCHUMACHER, F.X.A. A new growth curve and its applications to timber-yield studies. **Journal of Forestry**, Bethesda, v. 37, p. 819-820, 1939.

SUBBOTIN, M.T. On the law of frequency of errors. **Matematicheskii Sbornik**, v. 31, p. 296-301, 1923.

TAGUCHI, T. On a generalization of gaussian distribution. **Annals of the Institute of Statistical Mathematics**, v. 30, p. 211-242, 1978.

TALACKO, J. Perks' distributions and their role in the theory of Wiener's stochastic variables. **Trabajos de Estadística**, v. 7, p. 159-174, 1956.

TAYLOR, S.J.; KINGSMAN, B.G. An analysis of the variance and distribution of commodity price-changes. **Australian Journal of Management**, v. 4, p. 135-149, 1979.

TAYLOR, J.M.G. Properties of modelling the error distribution with an extra shape parameter. **Computational statistics and data analysis**, v. 13, p. 33-46, 1992.

TURNBULL, K.J. **Population dynamics em mixed forest stands: A system of mathematical models of mixed stand growth and structure**. University of Washington, Washington, 1963.

VERHULST, P.J. Notice sur la loi que la population suit dans sons accroissement. **Corr. Math. et Physique**, v. 10, p. 113-121, 1838.

VERHULST, P.J. Recherches mathématiques sur la loi d'accroissement de la population. **Académie de Bruxelles**, v. 18, p. 1-38, 1845.

WEIDA, F.M. On certain distribution functions when the law of the universe is Poisson's first law of error. **Annals of Mathematical Statistics**, v. 6, p. 102-110, 1935.

YAMAGUCHI, K. Generalized EM algorithm for model with contaminated error term. **Proceedings of the seven Japan and Korea joint conference of statistical**, p. 107-114, 1990.

## APÊNDICE A - Algoritmos e programas

### PROGRAMA DO MODELO DE CHAPMAN-RICHARDS PARA OS DADOS DE ALTURA

```
data arvore;
input rep t h6;
datalines;
```

```
1 2 0.4700
```

```
.
```

```
.
```

```
.
```

```
.
```

```
;
```

```
run;
```

```
*/PROC NLP NORMAL NÃO TRANSFORMADO;
```

```
proc nlp data=arvore cov=2 pcov vardef=n covariance=h maxiter=10000;
max loglikO;
parms BETA2=20, BETA3=1, BETA4=1, phi=1;
media=beta2*(1-exp(-beta3*t))**beta4;
u=((h6-media)**2)/phi;
loglikO=(-0.5)*log(phi)-u/2;
run;
```

```
*/PROC NLP t STUDENT NÃO TRANSFORMADO;
```

```

proc nlp data=arvore cov=2 pcov vardef=n covariance=h;
max loglikO;
parms BETA2=20, BETA3=1, BETA4=1, phi=1;
media=beta2*(1-exp(-beta3*t))**beta4;
u=(h6-media)**2/phi;
m=2;
den=gamma((m+1)/2)/gamma(1/2)*gamma(m/2);
t1=m**(m/2);
t2= t1*(m+u)**(-(m+1)/2);
ger=t2/den;
loglikO=-0.5*log(phi)+log(ger);
run;

```

\*/PROC NLP EXPONENCIAL POTENCIA NÃO TRANSFORMADO;

```

proc nlp data=arvore cov=2 pcov vardef=n covariance=h;
max loglikO;
parms BETA2=20, BETA3=1, BETA4=1, phi=1;
media=beta2*(1-exp(-beta3*t))**beta4;
u=(h6-media)**2/phi;
m=0.5;
t=1+0.5*(1+m);
kon=gamma(t)*2**t;
rr=(1+m)**(-1);
ger= kon**(-1)*exp(-0.5*u**rr);
loglikO=-0.5*log(phi)+log(ger);
run;

```

\*/PROC NLP CAUCHY NÃO TRANSFORMADO;

```

proc nlp data=arvore cov=2 pcov vardef=n covariance=h;
max loglikO;
parms BETA2=20, BETA3=1, BETA4=1, phi=1;
media=beta2*(1-exp(-beta3*t))**beta4;

```

```

u=(h6-media)**2/phi;
ger=1/(1+u);
loglikO=-0.5*log(phi)+log(ger);
run;

```

```

*/PROC NLP LOGISTICA I NÃO TRANSFORMADO;

```

```

proc nlp data=arvore cov=2 pcov vardef=n covariance=h;
max loglikO;
parms BETA2=20, BETA3=1, BETA4=1, phi=1;
media=beta2*(1-exp(-beta3*t))**beta4;
u=(h6-media)**2/phi;
ger= 1.484300029*exp(-u)/(1+exp(-u))**2;
loglikO=-0.5*log(phi)+log(ger);
run;

```

```

*/PROC NLP LOGISTICA II NÃO TRANSFORMADO;

```

```

proc nlp data=arvore cov=2 pcov vardef=n covariance=h;
max loglikO;
parms BETA2=20, BETA3=1, BETA4=1, phi=1;
media=beta2*(1-exp(-beta3*t))**beta4;
u=(h6-media)**2/phi;
ger= exp(-u**0.5)/(1+exp(-u**0.5))**2;
loglikO=-0.5*log(phi)+log(ger);
run;

```

```

*/PROC NLP NORMAL TRANSFORMADO;

```

```

proc nlp data=arvore cov=2 pcov vardef=n covariance=h maxiter=10000;
max loglikO;
parms BETA2=20, BETA3=1, BETA4=1, phi=1, lambda=1;
media=beta2*(1-exp(-beta3*t))**beta4;
h6t =((h6+2)**lambda);

```

```

u=((h6t-media)**2)/phi;
loglikO=(-0.5)*log(phi)-u/2+(lambda-1)*log(h6);
run;

```

```

*/PROC NLP t STUDENT TRANSFORMADO;

```

```

proc nlp data=arvore cov=2 pcov vardef=n covariance=h;
max loglikO;
parms BETA2=20, BETA3=1, BETA4=1, phi=1, lambda=1;
media=beta2*(1-exp(-beta3*t))**beta4;
h6t =(h6+2)**lambda;
u=(h6t-media)**2/phi;
m=2;
den=gamma((m+1)/2)/gamma(1/2)*gamma(m/2);
t1=m**(m/2);
t2= t1*(m+u)**(-(m+1)/2);
ger=t2/den;
loglikO=-0.5*log(phi)+log(ger)+(lambda-1)*log(h6);
run;

```

```

*/PROC NLP EXPONENCIAL POTENCIA TRANSFORMADO;

```

```

proc nlp data=arvore cov=2 pcov vardef=n covariance=h;
max loglikO;
parms BETA2=20, BETA3=1, BETA4=1, phi=1, lambda=1;
media=beta2*(1-exp(-beta3*t))**beta4;
h6t =(h6+2)**lambda;
u=(h6t-media)**2/phi;
m=0.5;
t=1+0.5*(1+m);
kon=gamma(t)*2**t;
rr=(1+m)**(-1);
ger= kon**(-1)*exp(-0.5*u**rr);
loglikO=-0.5*log(phi)+log(ger)+(lambda-1)*log(h6);

```

```
run;
```

```
*/PROC NLP CAUCHY TRANSFORMADO;
```

```
proc nlp data=arvore cov=2 pcov vardef=n covariance=h;
max loglikO;
parms BETA2=20, BETA3=1, BETA4=1, phi=1, lambda=1;
media=beta2*(1-exp(-beta3*t))**beta4;
h6t =(h6+2)**lambda;
u=(h6t-media)**2/phi;
ger=1/(1+u);
loglikO=-0.5*log(phi)+log(ger)+(lambda-1)*log(h6);
run;
```

```
*/PROC NLP LOGISTICA I TRANSFORMADO;
```

```
proc nlp data=arvore cov=2 pcov vardef=n covariance=h;
max loglikO;
parms BETA2=20, BETA3=1, BETA4=1, phi=1, lambda=1;
media=beta2*(1-exp(-beta3*t))**beta4;
h6t =(h6+2)**lambda;
u=(h6t-media)**2/phi;
ger= 1.484300029*exp(-u)/(1+exp(-u))**2;
loglikO=-0.5*log(phi)+log(ger)+(lambda-1)*log(h6);
run;
```

```
*/PROC NLP LOGISTICA II TRANSFORMADO;
```

```
proc nlp data=arvore cov=2 pcov vardef=n covariance=h;
max loglikO;
parms BETA2=20, BETA3=1, BETA4=1, phi=1, lambda=1;
media=beta2*(1-exp(-beta3*t))**beta4;
h6t =(h6+2)**lambda;
u=(h6t-media)**2/phi;
```

```

ger= exp(-u**0.5)/(1+exp(-u**0.5))**2;
loglikO=-0.5*log(phi)+log(ger)+(lambda-1)*log(h6);
run;

```

## PROGRAMA DO MODELO DE CHAPMAN-RICHARDS PARA OS DADOS DE CIRCUNFERÊNCIA

```

data arvore;
input rep t c6;
datalines;

1 12 11.625 . . . .
;
run;

*/PROC NLP NORMAL NÃO TRANSFORMADO;

proc nlp data=arvore cov=2 pcov vardef=n covariance=h maxiter=10000;
max loglikO;
parms BETA2=50, BETA3=1, BETA4=1, phi=1;
media=beta2*(1-exp(-beta3*t))**beta4;
u=((c6-media)**2)/phi;
loglikO=(-0.5)*log(phi)-u/2;
run;

*/PROC NLP t STUDENT NÃO TRANSFORMADO;

proc nlp data=arvore cov=2 pcov vardef=n covariance=h;
max loglikO;
parms BETA2=50, BETA3=1, BETA4=1, phi=1;
media=beta2*(1-exp(-beta3*t))**beta4;
u=(c6-media)**2/phi;
m=4;
den=gamma((m+1)/2)/gamma(1/2)*gamma(m/2);
t1=m**(m/2);

```

```

t2= t1*(m+u)**(-(m+1)/2);
ger=t2/den;
loglikO=-0.5*log(phi)+log(ger);
run;

```

```

*/PROC NLP EXPONENCIAL POTENCIA NÃO TRANSFORMADO;

```

```

proc nlp data=arvore cov=2 pcov vardef=n covariance=h;
max loglikO;
parms BETA2=50, BETA3=1, BETA4=1, phi=1;
media=beta2*(1-exp(-beta3*t))**beta4;
u=(c6-media)**2/phi;
m=0.1;
t=1+0.5*(1+m);
kon=gamma(t)*2**t;
rr=(1+m)**(-1);
ger= kon**(-1)*exp(-0.5*u**rr);
loglikO=-0.5*log(phi)+log(ger);
run;

```

```

*/PROC NLP CAUCHY NÃO TRANSFORMADO;

```

```

proc nlp data=arvore cov=2 pcov vardef=n covariance=h;
max loglikO;
parms BETA2=50, BETA3=1, BETA4=1, phi=1;
media=beta2*(1-exp(-beta3*t))**beta4;
u=(c6-media)**2/phi;
ger=1/(1+u);
loglikO=-0.5*log(phi)+log(ger);
run;

```

```

*/PROC NLP LOGISTICA I NÃO TRANSFORMADO;

```

```

proc nlp data=arvore cov=2 pcov vardef=n covariance=h;

```

```

max loglikO;
parms BETA2=50, BETA3=1, BETA4=1, phi=1;
media=beta2*(1-exp(-beta3*t))**beta4;
u=(c6-media)**2/phi;
ger= 1.484300029*exp(-u)/(1+exp(-u))**2;
loglikO=-0.5*log(phi)+log(ger);
run;

```

```

*/PROC NLP LOGISTICA II NÃO TRANSFORMADO;

```

```

proc nlp data=arvore cov=2 pcov vardef=n covariance=h;
max loglikO;
parms BETA2=50, BETA3=1, BETA4=1, phi=1;
media=beta2*(1-exp(-beta3*t))**beta4;
u=(c6-media)**2/phi;
ger= exp(-u**0.5)/(1+exp(-u**0.5))**2;
loglikO=-0.5*log(phi)+log(ger);
run;

```

```

*/PROC NLP NORMAL - TRANSFORMADO;

```

```

proc nlp data=arvore cov=2 pcov vardef=n covariance=h maxiter=10000;
max loglikO;
parms BETA2=50, BETA3=1, BETA4=1, phi=1, lambda=1;
media=beta2*(1-exp(-beta3*t))**beta4;
c6t =(c6**lambda-1)/lambda;
u=((c6t-media)**2)/phi;
loglikO=(-0.5)*log(phi)-u/2+(lambda-1)*log(c6);
run;

```

```

*/PROC NLP t STUDENT - TRANSFORMADO;

```

```

proc nlp data=arvore cov=2 pcov vardef=n covariance=h;
max loglikO;

```

```

parms BETA2=50, BETA3=1, BETA4=1, phi=1, lambda=1;
media=beta2*(1-exp(-beta3*t))**beta4;
c6t =(c6**lambda-1)/lambda;
u=(c6t-media)**2/phi;
m=2;
den=gamma((m+1)/2)/gamma(1/2)*gamma(m/2);
t1=m**(m/2);
t2= t1*(m+u)**(-(m+1)/2);
ger=t2/den;
loglikO=-0.5*log(phi)+log(ger)+(lambda-1)*log(c6);
run;

```

\*/PROC NLP EXPONENCIAL POTENCIA - TRANSFORMADO;

```

proc nlp data=arvore cov=2 pcov vardef=n covariance=h;
max loglikO;
parms BETA2=50, BETA3=1, BETA4=1, phi=1, lambda=1;
media=beta2*(1-exp(-beta3*t))**beta4;
c6t =(c6**lambda-1)/lambda;
u=(c6t-media)**2/phi;
m=0.1;
t=1+0.5*(1+m);
kon=gamma(t)*2**t;
rr=(1+m)**(-1);
ger= kon**(-1)*exp(-0.5*u**rr);
loglikO=-0.5*log(phi)+log(ger)+(lambda-1)*log(c6);
run;

```

\*/PROC NLP CAUCHY - TRANSFORMADO;

```

proc nlp data=arvore cov=2 pcov vardef=n covariance=h;
max loglikO;
parms BETA2=50, BETA3=1, BETA4=1, phi=1, lambda=1;
media=beta2*(1-exp(-beta3*t))**beta4;

```

```

c6t =(c6**lambda-1)/lambda;
u=(c6t-media)**2/phi;
ger=1/(1+u);
loglikO=-0.5*log(phi)+log(ger)+(lambda-1)*log(c6);
run;

```

```

*/PROC NLP LOGISTICA I - TRANSFORMADO;

```

```

proc nlp data=arvore cov=2 pcov vardef=n covariance=h;
max loglikO;
parms BETA2=50, BETA3=1, BETA4=1, phi=1, lambda=1;
media=beta2*(1-exp(-beta3*t))**beta4;
c6t =(c6**lambda-1)/lambda;
u=(c6t-media)**2/phi;
ger= 1.484300029*exp(-u)/(1+exp(-u))**2;
loglikO=-0.5*log(phi)+log(ger)+(lambda-1)*log(c6);
run;

```

```

*/PROC NLP LOGISTICA II - TRANSFORMADO;

```

```

proc nlp data=arvore cov=2 pcov vardef=n covariance=h;
max loglikO;
parms BETA2=50, BETA3=1, BETA4=1, phi=1, lambda=1;
media=beta2*(1-exp(-beta3*t))**beta4;
c6t =(c6**lambda-1)/lambda;
u=(c6t-media)**2/phi;
ger= exp(-u**0.5)/(1+exp(-u**0.5))**2;
loglikO=-0.5*log(phi)+log(ger)+(lambda-1)*log(c6);
run;

```