A distribuição t-assimétrica univariada: propriedades e inferência

Luciana Graziela de Godoi

Dissertação apresentada
ao
Instituto de Matemática e Estatística
da
Universidade de São Paulo
para
obtenção do título
de
Mestre em Ciências

Área de Concentração: **Estatística**

Orientadora: **Profa. Dra. Márcia D'Elia Branco**

Durante a elaboração deste trabalho a autora recebeu auxílio financeiro da CAPES e do CNPq.

- São Paulo, agosto de 2007 -

A distribuição t-assimétrica univariada: propriedades e inferência

Este exemplar corresponde à redação
final da dissertação devidamente corrigida
e defendida por Luciana Graziela de Godoi
e aprovada pela Comissão Julgadora.

São Paulo, agosto de 2007.

Banca Examinadora:

- Profa. Dra. Márcia D'Elia Branco (Orientadora) - IME-USP.
- Profa. Dra. Silvia Lopes de Paula Ferrari - IME-USP.
- Profa. Dra. Rosângela Helena Loschi - UFMG.

*A*os meus queridos pais,
Tânia e Benedito,
pelo incondicional
apoio e carinho.

ABSTRACT

In this work, we present several different results about the univariate skew-t distribution. This distribution includes as particular cases the t-Student, skew-normal and normal distributions. Initially, it is considered different ways to characterize this distribution and some properties. Inferential issues are argued from Bayesian and Classical approaches. Under the first perspective, we consider the methods of moments and maximum likelihood, and also we present graphical model diagnostics. Through a simulation study we evaluate the performance of these two types of estimation and we conclude that, in general, samples with great skewness present maximum likelihood estimates closer the true value of the parameter to those which were obtained by the method of the moments. However, we observe a great bias in both the estimates. The classical estimators had also presented theoretical problems. Under Bayesian point of view, we present possible prior distributions for the parameters of the skew-t distribution. In order to simplify the computational implementation of the Bayesian method, we present several hierarchical forms to represent the skew-t model. The Bayesian methodology revealed more efficient to estimate the shape and degrees of freedom parameters. Finally, the inferences proposals previously are applied in two data sets: notes of the pupils of the MAE-0116 course and one simulated sample.

Sumário

Lista de Tabelas

Lista de Figuras

Lista de Símbolos

$N(\mu, \sigma^2)$	Distribuição normal com parâmetro de posição μ e escala σ^2
t_ν	Distribuição t padrão com ν graus de liberdade
$t_\nu(\mu, \sigma^2)$	Distribuição t-Student com parâmetro de posição μ, escala σ^2 e graus de liberdade ν
$t_G(\mu, \sigma^2, \eta, \nu)$	Distribuição t-generalizada com parâmetro de posição μ e escala σ^2
$SN(\lambda)$	Distribuição normal-assimétrica padrão com parâmetro de assimetria λ
$SN(\mu, \sigma^2, \lambda)$	Distribuição normal-assimétrica com parâmetro de posição μ, escala σ^2 e assimetria λ
$ST(\nu, \lambda)$	Distribuição t-assimétrica padrão com parâmetro de assimetria λ e graus de liberdade ν
$ST(\mu, \sigma^2, \nu, \lambda)$	Distribuição t-assimétrica com parâmetro de posição μ, escala σ^2, graus de liberdade ν e assimetria λ
$C(\mu, \sigma^2)$	Distribuição Cauchy com parâmetro de posição μ e escala σ^2
$SC(\lambda)$	Distribuição Cauchy-assimétrica com parâmetro de assimetria λ
$GI(a, b)$	Distribuição gama invertida com parâmetros a e b

$Gama(a, b)$	Distribuição gama com parâmetro de escala a e forma b
χ^2_k	Distribuição qui-quadrado com k graus de liberdade
$F\text{-}Snedocor(a, b)$	Distribuição F-Snedocor com parâmetros a e b
$HN(0, 1)$	Distribuição normal-positiva (*half-normal*)
$HT(\nu)$	Distribuição t-positiva com parâmetro de curtose ν (*half-t*)
$Exp(\lambda)\ \mathrm{I}(a, \infty)$	Distribuição exponencial truncada à esquerda de a, com parâmetro λ
$t_2(\underset{\sim}{\mu}, \Sigma, \nu)$	Distribuição t-Student bivariada com vetor de posição $\underset{\sim}{\mu}_{(2 \times 1)}$, matriz de dispersão $\Sigma_{(2 \times 2)}$ e graus de liberdade ν
$t_{G_2}(\underset{\sim}{\mu}, \Sigma, \eta, \nu)$	Distribuição t-generalizada bivariada com vetor de posição $\underset{\sim}{\mu}_{(2 \times 1)}$ e matriz de dispersão $\Sigma_{(2 \times 2)}$
$\Gamma(\cdot)$	Função gama
$\psi(\cdot)$	Função digama
F_Z	Função de distribuição acumulada da variável aleatória Z
$X_n \xrightarrow{P} \mathrm{X}$	X_n converge em probabilidade para X
$X_n \xrightarrow{D} \mathrm{X}$	X_n converge em distribuição para X
$\mathrm{X} \perp \mathrm{W}$	X e W são variáveis aleatórias independentes
$\mathrm{X} \stackrel{d}{=} \mathrm{W}$	X e W têm a mesma distribuição
$\sim$	Distribuí-se como

Capítulo 1

Introdução

Em grande parte a análise estatística desenvolvida para o estudo de variáveis aleatórias contínuas está baseada no modelo normal. No entanto, supor que um conjunto de dados pode ser modelado através de uma distribuição normal requer alguns cuidados, tais como, assegurar a suposição de simetria e um determinado valor de curtose. Caso alguma destas condições sejam violadas, a razoabilidade das inferências obtidas ao se supor normalidade podem ficar comprometidas.

Dessa maneira, outras propostas de modelagem têm sido discutidas a fim de minimizar problemas como os descritos anteriormente. Como alternativa ao modelo normal, o modelo t-Student tem sido sugerido por diversos autores, entre eles, Arellano-Valle (1994). O uso deste modelo nos permite reduzir a influência de valores extremos nas inferências, sendo considerado, portanto, um modelo mais robusto que o normal (Lange, Little & Taylor, 1989).

Contudo, tanto a distribuição normal quanto a distribuição t-Student pressupõem a simetria dos dados. Em muitas situações, fazer tal suposição não é adequado. Embora tenhamos distribuições assimétricas bem conhecidas, como por exemplo, a distribuição gama, que é não invariante por transformação linear e cujo espaço amostral está definido apenas para valores positivos, seria interessante obter distribuições assimétricas menos restritivas e que tivessem, como caso particular, distribuições simétricas conhecidas. Nesse sentido, Azzalini (1985) propôs a distribuição normal-assimétrica como uma generalização do modelo normal. A partir de então, muitos trabalhos vêm sendo desenvolvidos nesta área e

novas distribuições assimétricas e propriedades têm sido relatadas na literatura (Branco & Arellano-Valle, 2004).

Um possível processo para assimetrizar uma distribuição simétrica, com função densidade de probabilidade f, consiste em considerar $f_Z(z) = 2\ f(z)\ Q(W(z))$, onde Q é uma função não-negativa tal que $Q(-z) = 1 - Q(z)$, para todo z e W é uma função ímpar, para todo z, então Z é uma variável aleatória distribuída assimetricamente.

Distribuições assimetrizadas segundo a proposta anterior tendem a preservar algumas propriedades da distribuição simétrica f. Dessa forma, conseguimos obter uma classe de famílias de distribuições paramétricas mais flexíveis, capazes de modelar de maneira mais adequada dados oriundos de fenômenos reais.

Para obtermos a função densidade de uma distribuição normal-assimétrica, por exemplo, basta considerar $W(z) = \lambda z$, $f = \phi$ e $Q = \Phi$, onde $\phi(.)$ e $\Phi(.)$ são as funções de densidade de probabilidade (f.d.p.) e de distribuição acumulada (f.d.a.) de uma distribuição normal padrão, respectivamente. Assim, a função densidade de probabilidade de uma distribuição normal assimetrizada é dada por $f_Z(z) = 2\ \phi(z)\ \Phi(\lambda z)$.

Observe que, quando consideramos $\lambda = 0$, obtemos a distribuição normal como caso particular. Maiores detalhes sobre a distribuição normal-assimétrica univariada podem ser vistos em Azzalini (1985) e Rodríguez (2005).

Nos casos em que a amostra tem um comportamento assimétrico e, mais do que isso, há a presença de valores extremos, a utilização da distribuição normal-assimétrica para a modelagem do conjunto de dados pode ser ineficaz. Em situações deste tipo, propõe-se como alternativa mais robusta ao modelo normal-assimétrico a versão assimétrica da distribuição t-Student.

A distribuição t-assimétrica é uma importante distribuição assimétrica que inclui não só a distribuição t-Student como caso particular, mas também as distribuições normal e normal-assimétrica em casos limites.

Os primeiros trabalhos a apresentar a distribuição t-assimétrica foram Branco & Dey (2001) e Azzalini & Capitanio (2003). Nesses trabalhos a abordagem

é multivariada. No primeiro, é proposta uma classe geral de distribuições assimétricas que generaliza os resultados apresentados e discutidos para a distribuição normal-assimétrica multivariada em Azzalini & Dalla Valle (1996) e Azzalini & Capitanio (1999). Propriedades e aspectos inferenciais da distribuição t-assimétrica multivariada foram discutidos em Azzalini & Capitanio (2002).

Do ponto de vista univariado, propostas alternativas à distribuição t-assimétrica foram feitas por Fernandes & Steel (1998), Jones (2001) e Jones & Faddy (2003), e não coincidem com o modelo discutido nesta dissertação. Como principal característica, a distribuição descrita por Fernandes & Steel (1998) preserva a moda e a de Jones (2001) desenvolvida em Jones & Faddy (2003) é baseada em uma transformação da distribuição beta. Uma forma multivariada da distribuição t-assimétrica foi proposta por Jones (2002), mas os aspectos inferenciais associados não foram discutidos.

1.1 Objetivos e apresentação dos capítulos

Um dos objetivos deste trabalho é mostrar que a distribuição t-assimétrica univariada pode ser caracterizada de diversas maneiras. Outro objetivo é discutir aspectos inferenciais desta distribuição, sob as abordagens clássica e bayesiana. Muitos dos resultados aqui apresentados foram obtidos em contextos mais amplos. No contexto multivariado, diferentes classes de distribuições elípticas assimétricas têm sido propostas na literatura (ver Branco & Arellano-Valle, 2004). A distribuição t-assimétrica univariada aqui considerada é um caso especial da distribuição elíptica assimétrica proposta por Branco & Dey (2001).

No segundo capítulo, apresentamos a distribuição t-assimétrica univariada padrão e geral, possíveis caracterizações e propriedades. Obtemos, também, a função geradora de momentos e, por conseguinte, momentos de ordem par e ímpar, coeficientes de assimetria e curtose.

No Capítulo 3, através da metodologia clássica, apresentamos uma discussão sobre a estimação dos parâmetros da distribuição t-assimétrica. Obtemos

os estimadores de momentos, as equações de verossimilhança e fazemos uma breve análise de modelos de diagnósticos. Por fim, analisamos os resultados de um estudo de simulação que nos permite comparar a eficiência dos estimadores obtidos via abordagem clássica.

No quarto capítulo, discutimos a metodologia bayesiana que, como veremos, nos permitirá controlar alguns problemas de estimação resultantes da utilização do método dos momentos (MM) e máxima verossimilhança (MV). Propomos e comparamos diferentes modelos bayesianos hierárquicos baseados nas caracterizações e propriedades apresentadas no capítulo 2.

No quinto capítulo, apresentamos uma aplicação baseada em um conjunto de dados relativos às notas da primeira prova dos alunos de Ciências Biológicas e Relações Públicas que cursaram a disciplina Noções de Estatística (MAE0116) no 1° semestre de 2006 no Instituto de Matemática e Estatística da Universidade de São Paulo. Também, é apresentada a análise de um conjunto de dados simulados através das abordagens bayesiana e clássica.

Finalmente, no sexto capítulo são apresentadas nossas conclusões e perspectivas futuras de trabalho.

São também apresentados cinco apêndices, a saber: A) Provas das propriedades; B) Obtenção dos estimadores de momentos; C) Obtenção das equações de verossimilhança; D) Tabelas e E) Figuras.

Capítulo 2

A Distribuição t-Student Assimetrizada

Neste capítulo apresentamos a distribuição t-Student assimetrizada univariada, algumas de suas propriedades e caracterizações. Analisaremos o caso em que a distribuição depende de apenas dois parâmetros, associados a assimetria e a curtose, a qual denominaremos distribuição t-assimétrica padrão. Uma generalização desta distribuição é obtida quando consideramos também os parâmetros de posição e escala. Por fim, apresentaremos a função geradora de momentos para o caso padrão e geral.

Para demonstrar mais facilmente algumas das caracterizações da distribuição t-assimétrica padrão, precisaremos recorrer a algumas propriedades da distribuição t-generalizada. Sendo assim, apresentamos a seguir a distribuição t-generalizada e um resumo das suas principais propriedades. Maiores detalhes sobre esta distribuição podem ser encontrados em Arellano-Valle (1994).

2.1 A distribuição t-generalizada univariada

Uma variável aleatória contínua X é denominada t-generalizada univariada, com parâmetro de posição ($\mu \in \mathbb{R}$), escala ($\sigma^2 \in \mathbb{R}_+^*$) e $\nu, \eta \in \mathbb{R}_+^*$, se a sua função densidade de probabilidade é dada por:

$$f_X(x) = \frac{\Gamma\left(\frac{\nu+1}{2}\right)}{\Gamma\left(\frac{\nu}{2}\right)} \frac{\eta^{\nu/2}}{\sqrt{\pi\sigma^2}} \left[\eta + \left(\frac{x-\mu}{\sigma}\right)^2\right]^{-\frac{\nu+1}{2}} \qquad (-\infty < x < \infty).$$

Notação: $X \sim t_G(\mu,\, \sigma^2,\, \eta,\, \nu)$.

Para obtermos uma variável aleatória $X \sim t_G(\mu, \sigma^2, \eta, \nu)$, basta considerarmos $X \stackrel{d}{=} \mu + \sigma\sqrt{V}\, Z$, onde $V \sim GI\left(\frac{\eta}{2}, \frac{\nu}{2}\right)$ e $Z \sim N(0,\, 1)$ são independentes.

A média desta distribuição é dada por μ, para $\nu > 1$. Sua variância por $\frac{\eta}{\nu-2}\,\sigma^2$, para $\nu > 2$. O coeficiente de curtose depende apenas do parâmetro ν, sendo igual a $3 + \frac{6}{\nu-4}$, para $\nu > 4$. Observe que o coeficiente de curtose da distribuição t-generalizada é maior do que o coeficiente de curtose da distribuição normal.

Quando $\eta = \nu$ temos a distribuição t-Student posição e escala, ou seja, $t_G(\mu,\, \sigma^2,\, \nu,\, \nu) \equiv t_\nu(\mu,\, \sigma^2)$. Quando $\mu = 0$ e $\sigma^2 = 1$ obtemos a distribuição t padrão.

Uma importante propriedade relacionada a esta distribuição, e que será utilizada para demonstrar as Proposições 2.3 e 2.4, diz respeito ao condicionamento de um vetor aleatório t-generalizado bivariado.

$$\text{Considere X} = \begin{pmatrix} X_1 \\ X_2 \end{pmatrix} \sim t_{G_2}\left[\begin{pmatrix} \mu_1 \\ \mu_2 \end{pmatrix}, \begin{pmatrix} \sigma_{11} & \sigma_{12} \\ \sigma_{21} & \sigma_{22} \end{pmatrix}, \eta, \nu\right], \text{ então}$$

- $X_1 \sim t_G(\mu_1, \sigma_{11}, \eta, \nu)$ e $X_2 \sim t_G(\mu_2, \sigma_{22}, \eta, \nu)$,
- $X_1 \mid X_2 = x_2 \sim t_G(\mu_1(x_2), \sigma_{11.2}, \eta_{q(x_2)}, \nu_1)$, com

$$\begin{cases} \mu_1(x_2) = \mu_1 + \sigma_{12}\,\sigma_{22}{}^{-1}(x_2 - \mu_2), \\ \sigma_{11.2} = \sigma_{11} - \sigma_{12}\,\sigma_{22}{}^{-1}\,\sigma_{21}, \\ q(x_2) = (x_2 - \mu_2)^2\,\sigma_{22}{}^{-1}, \\ \eta_{q(x_2)} = \eta + q(x_2), \\ \nu_1 = \nu + 1. \end{cases} \tag{2.1}$$

2.2 Distribuição t-assimétrica Padrão

2.2.1 Definição e caracterizações

Definição 2.1 *Uma variável aleatória contínua Z é denominada t-assimétrica padrão, com parâmetros de assimetria ($\lambda \in \mathbb{R}$) e curtose ($\nu \in \mathbb{R}_+^*$), se a sua função densidade de probabilidade é dada por:*

$$f_Z(z) = 2t_\nu(z)T_{\nu+1}\left(\lambda z\sqrt{\frac{1+\nu}{\nu+z^2}}\right) \qquad (-\infty < z < \infty), \tag{2.2}$$

onde t_ν denota a função densidade de probabilidade de uma distribuição t-Student padrão com ν graus de liberdade e $T_{\nu+1}$ a função de distribuição acumulada de uma distribuição t-Student padrão com ν+1 graus de liberdade.

Notação: $Z \sim ST(\nu, \lambda)$

O parâmetro λ regula a forma da distribuição t-assimétrica. Valores negativos de λ indicam assimetria negativa e valores positivos de λ indicam assimetria positiva. Quando λ= 0, a densidade definida anteriormente será simétrica e irá coincidir com a densidade da distribuição t padrão. Na Figura 2.1 apresentamos alguns gráficos que ilustram o comportamento da densidade descrita em (2.2) para λ igual a 0, 2, 10 e 50, fixado o valor de ν, e para valores de ν iguais a 2, 5 e 15, fixado o valor de λ.

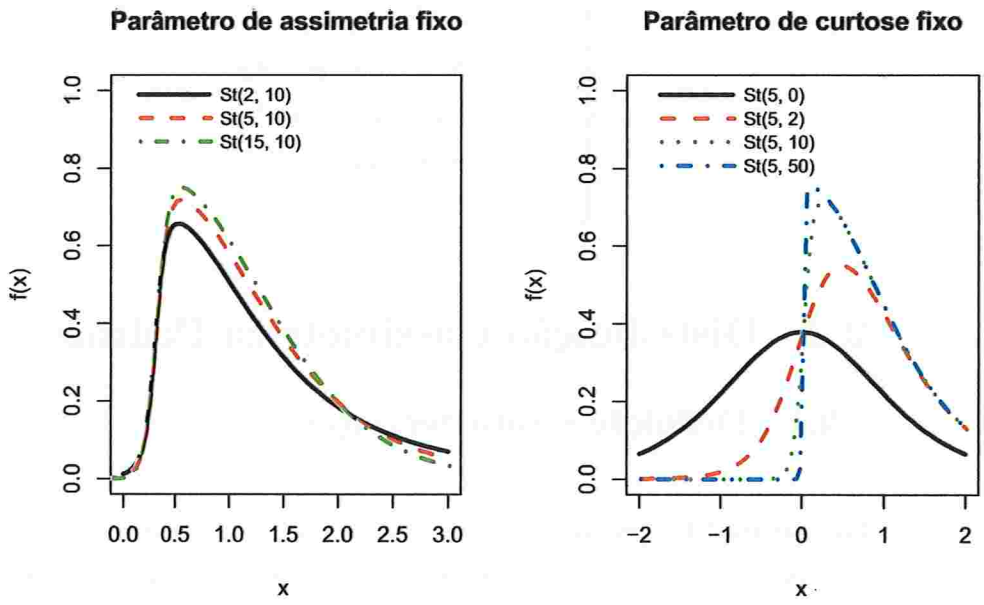


Figura 2.1: Funções de densidade para a distribuição t-assimétrica padrão, com diferentes valores para ν e λ

Apresentamos a seguir quatro maneiras diferentes de obter a distribuição t-assimétrica. A caracterização apresentada na Proposição 2.1 foi motivada pela maneira mais usual de obtenção da distribuição t padrão em inferência estatística, lembrando que a distribuição Gama$\left(\frac{\nu}{2}, \frac{1}{2}\right)$ é também conhecida como uma distribuição qui-quadrado com ν graus de liberdade.

Proposição 2.1 *Sejam X e W variáveis aleatórias independentes e admita que $X \sim SN(\lambda)$ e $W \sim Gama\left(\frac{\nu}{2}, \frac{1}{2}\right)$, então*

$$Z \stackrel{d}{=} \frac{X}{\sqrt{\frac{W}{\nu}}} \sim ST(\nu, \lambda). \tag{2.3}$$

Prova:

Seja $U = W$. Usando o método do Jacobiano para a transformação de variáveis, obtemos função de densidade conjunta de Z e U:

$$f_{Z,U}(z,u) = f_{X,W}\left(z\sqrt{\frac{u}{\nu}}, u\right)\left|\sqrt{\frac{u}{\nu}}\right|, \qquad \text{com } z \in \mathbb{R} \text{ e } u > 0.$$

Como X e W são variáveis aleatórias independentes, temos que

$$\begin{aligned} f_{Z,U}(z,u) &= f_X\left(z\sqrt{\frac{u}{\nu}}\right) f_W(u)\sqrt{\frac{u}{\nu}} \\ &= 2\phi\left(z\sqrt{\frac{u}{\nu}}\right)\Phi\left(\lambda z\sqrt{\frac{u}{\nu}}\right) f_W(u)\sqrt{\frac{u}{\nu}}. \end{aligned}$$

Integrando em u, resulta

$$\begin{aligned} f_Z(z) &= \int_0^\infty 2\,\phi\left(z\sqrt{\frac{u}{\nu}}\right)\Phi\left(\lambda z\sqrt{\frac{u}{\nu}}\right) f_W(u)\sqrt{\frac{u}{\nu}}\,du \\ &= \int_0^\infty 2\,\frac{1}{\sqrt{2\pi}}\,e^{-\frac{z^2 u}{2\nu}}\int_{-\infty}^{\lambda z\sqrt{\frac{u}{\nu}}}\frac{e^{-k^2/2}}{\sqrt{2\pi}}\,dk\,\frac{\left(\frac{1}{2}\right)^{\nu/2}}{\Gamma(\nu/2)}u^{\nu/2-1}e^{-u/2}\left(\frac{u}{\nu}\right)^{1/2}du. \end{aligned}$$

Sendo assim,

$$\begin{aligned} f_Z(z) = 2\,\frac{\Gamma\left(\frac{\nu+1}{2}\right)}{\Gamma\left(\frac{\nu}{2}\right)}\,\frac{1}{\sqrt{\nu\pi}}\left(1+\frac{z^2}{\nu}\right)^{-\frac{\nu+1}{2}}\int_0^\infty\int_{-\infty}^{\lambda z\sqrt{\frac{u}{\nu}}}\frac{1}{\Gamma\left(\frac{\nu+1}{2}\right)}\left(1+\frac{z^2}{\nu}\right)^{\frac{\nu+1}{2}} \times \\ \times\, e^{-\frac{1}{2}\left(\frac{z^2 u}{\nu}+k^2+u\right)}\,\frac{u^{\frac{\nu-1}{2}}}{\sqrt{2\pi}}\,2^{-\frac{\nu+1}{2}}dk\,du. \end{aligned}$$

Considere agora $x = k\sqrt{\left(\frac{\nu}{u}\right)\frac{1+\nu}{\nu+z^2}}$, então

$$f_Z(z) = 2\,t_\nu(z)\int_0^{\infty}\int_{-\infty}^{\lambda z\sqrt{\frac{1+\nu}{\nu+z^2}}} \frac{1}{\Gamma\left(\frac{\nu+1}{2}\right)}\left(1+\frac{z^2}{\nu}\right)^{\frac{\nu+1}{2}} e^{-u\left[\frac{1}{2}\left(1+\frac{z^2}{\nu}\right)\right]} e^{-\frac{1}{2}\left[\frac{u(\nu+z^2)}{\nu(1+\nu)}x^2\right]} \times$$

$$\times \frac{1}{\sqrt{2\pi}\,2^{\frac{\nu+1}{2}}}\,u^{\frac{\nu-1}{2}}\sqrt{\frac{u}{\nu}}\sqrt{\frac{\nu+z^2}{1+\nu}}\,dx\,du.$$

$$= 2\,t_\nu(z)\int_{-\infty}^{\lambda z\sqrt{\frac{1+\nu}{\nu+z^2}}} \frac{\Gamma\left(\frac{\nu+2}{2}\right)}{\Gamma\left(\frac{\nu+1}{2}\right)}\frac{1}{\sqrt{\pi}}\frac{(1+\nu)^{\frac{\nu+1}{2}}}{(1+\nu+x^2)^{\frac{\nu+2}{2}}}\int_0^{\infty}\frac{1}{\Gamma\left(\frac{\nu+2}{2}\right)}u^{\left(\frac{\nu}{2}+1\right)-1}\times$$

$$\times\; e^{-u\left[\frac{1}{2}\left(1+\frac{z^2}{\nu}+\frac{1}{\nu}\frac{(\nu+z^2)}{1+\nu}x^2\right)\right]}\left[\frac{1}{2}\left(1+\frac{z^2}{\nu}\right)\right]^{\frac{\nu}{2}+1}\frac{(1+\nu)^{-\left(\frac{\nu}{2}+1\right)}}{(1+\nu+x^2)^{-\left(\frac{\nu}{2}+1\right)}}\,du\,dx.$$

Portanto,

$$f_Z(z) = 2\,t_\nu(z)\int_{-\infty}^{\lambda z\sqrt{\frac{1+\nu}{\nu+z^2}}} \frac{\Gamma\left(\frac{\nu+2}{2}\right)}{\Gamma\left(\frac{\nu+1}{2}\right)}\frac{1}{\sqrt{\pi}}\frac{(1+\nu)^{\frac{\nu+1}{2}}}{(1+\nu+x^2)^{\frac{\nu+2}{2}}}\int_0^{\infty}\frac{1}{\Gamma\left(\frac{\nu+2}{2}\right)}u^{\left(\frac{\nu}{2}+1\right)-1}\times$$

$$\times\; e^{-u\left[\frac{1}{2}\left(1+\frac{z^2}{\nu}+\frac{1}{\nu}\frac{(\nu+z^2)}{1+\nu}x^2\right)\right]}\left[\frac{1}{2}\left(1+\frac{z^2}{\nu}\right)\left(1+\frac{x^2}{1+\nu}\right)\right]^{\frac{\nu}{2}+1}\,du\,dx.$$

Observe que podemos reescrever o termo $\left(1+\frac{z^2}{\nu}\right)\left(1+\frac{x^2}{1+\nu}\right)$ como $1+\frac{z^2}{\nu}+$ $+\frac{1}{\nu}\frac{(\nu+z^2)}{1+\nu}x^2$. Assim,

$$f_Z(z) = 2\,t_\nu(z)\int_{-\infty}^{\lambda z\sqrt{\frac{1+\nu}{\nu+z^2}}} \frac{\Gamma\left(\frac{\nu+2}{2}\right)}{\Gamma\left(\frac{\nu+1}{2}\right)}\frac{1}{\sqrt{\pi}}\frac{(1+\nu)^{\frac{\nu+1}{2}}}{(1+\nu+x^2)^{\frac{\nu+2}{2}}}\int_0^{\infty}\frac{1}{\Gamma\left(\frac{\nu+2}{2}\right)}u^{\left(\frac{\nu}{2}+1\right)-1}\times$$

$$\times\; e^{-u\left[\frac{1}{2}\left(1+\frac{z^2}{\nu}+\frac{1}{\nu}\frac{(\nu+z^2)}{1+\nu}x^2\right)\right]}\left[\frac{1}{2}\left(1+\frac{z^2}{\nu}+\frac{1}{\nu}\frac{(\nu+z^2)}{1+\nu}x^2\right)\right]^{\frac{\nu}{2}+1}\,du\,dx.$$

Note que, o integrando $f(u)$ é a densidade de uma distribuição gama com parâmetros $\frac{\nu}{2}+1$ e $\frac{1}{2}\left(1+\frac{z^2}{\nu}+\frac{1}{\nu}\frac{(\nu+z^2)}{1+\nu}\,x^2\right)$. Portanto, a integral em u é igual a um. Logo,

$$
\begin{aligned}
f_Z(z) &= 2\, t_\nu(z) \int_{-\infty}^{\lambda z\sqrt{\frac{1+\nu}{\nu+z^2}}} \frac{\Gamma\left(\frac{\nu+2}{2}\right)}{\Gamma\left(\frac{\nu+1}{2}\right)} \frac{1}{\sqrt{\pi}} \frac{(1+\nu)^{\frac{\nu+1}{2}}}{(1+\nu+x^2)^{\frac{\nu+2}{2}}}\, dx \\
&= 2\, t_\nu(z)\, T_{\nu+1}\left(\lambda z\sqrt{\frac{1+\nu}{\nu+z^2}}\right) \qquad (-\infty < z < \infty).
\end{aligned}
$$

♦

Uma segunda forma de obter a distribuição t-assimétrica consiste em representá-la como uma mistura no parâmetro de escala da distribuição normal-assimétrica. Esta representação é fundamental para a obtenção da função geradora de momentos da distribuição (2.2).

Proposição 2.2 *Se* $Z|\, W \sim SN\left(0, \frac{1}{W}, \lambda\right)$ *e* $W \sim Gama\left(\frac{\nu}{2}, \frac{\nu}{2}\right)$, *então* $Z \sim ST(\nu, \lambda)$.

Prova:

A função densidade conjunta de Z e W, onde $w > 0$, é dada por:

$$
\begin{aligned}
f_{Z,W}(z,w) &= f_{Z|W}\,(z|w) f_W(w) \\
&= 2\,\phi(z\sqrt{w})\,\Phi(\lambda z\sqrt{w})\,\frac{\left(\frac{\nu}{2}\right)^{\nu/2}}{\Gamma\left(\frac{\nu}{2}\right)}\, e^{-\frac{\nu}{2}w}\, w^{\frac{\nu}{2}-\frac{1}{2}}.
\end{aligned}
$$

Então,

$$
\begin{aligned}
f_Z(z) &= \int_0^\infty 2\,\phi(z\sqrt{w})\,\Phi(\lambda z\sqrt{w})\,\frac{\left(\frac{\nu}{2}\right)^{\nu/2}}{\Gamma\left(\frac{\nu}{2}\right)}\, e^{-\frac{\nu}{2}w}\, w^{\frac{\nu}{2}-\frac{1}{2}}\, dw \\
&= 2\,\frac{\left(\frac{\nu}{2}\right)^{\nu/2}}{\Gamma\left(\frac{\nu}{2}\right)} \int_0^\infty \frac{1}{\sqrt{2\pi}}\, e^{-\frac{wz^2}{2}}\,\Phi(\lambda z\sqrt{w})\, e^{-\frac{\nu}{2}w}\, w^{\frac{\nu}{2}-\frac{1}{2}}\, dw \\
&= 2\,\frac{\Gamma\left(\frac{\nu+1}{2}\right)}{\Gamma\left(\frac{\nu}{2}\right)} \frac{\nu^{\nu/2}}{\sqrt{\pi}} \int_0^\infty \frac{1}{2^{\frac{\nu+1}{2}}} \frac{1}{\Gamma\left(\frac{\nu+1}{2}\right)}\, e^{-\frac{w}{2}(z^2+\nu)}\Phi(\lambda z\sqrt{w})\, w^{\frac{\nu}{2}-\frac{1}{2}}\, dw.
\end{aligned}
$$

Portanto,

$$f_{Z,W}(z,w) = 2\,\frac{\Gamma\left(\frac{\nu+1}{2}\right)}{\Gamma\left(\frac{\nu}{2}\right)}\frac{\nu^{\frac{\nu+1}{2}}}{\sqrt{\pi\nu}(\nu+z^2)^{\frac{\nu+1}{2}}}\int_0^{\infty}\left(\frac{\nu+z^2}{2}\right)^{\frac{\nu+1}{2}}\frac{1}{\Gamma\left(\frac{\nu+1}{2}\right)}\,e^{-\frac{w}{2}(z^2+\nu)}w^{\frac{\nu}{2}-\frac{1}{2}}\times$$

$$\times\int_{-\infty}^{\lambda z\sqrt{w}}\frac{1}{\sqrt{2\pi}}\,e^{-k^2/2}\,dk\,dw$$

$$= 2\,t_\nu(z)\int_0^{\infty}\left(\frac{\nu+z^2}{2}\right)^{\frac{\nu+1}{2}}\frac{1}{\Gamma\left(\frac{\nu+1}{2}\right)}\,e^{-\frac{w}{2}(z^2+\nu)}\,w^{\frac{\nu}{2}-\frac{1}{2}}\times$$

$$\times\int_{-\infty}^{\lambda z\sqrt{w}}\frac{1}{\sqrt{2\pi}}\,e^{-k^2/2}dk\,dw.$$

Utilizando a transformação $t=\frac{k}{\sqrt{w}}$ para resolver a integral em dk, temos

$$f_Z(z) = 2t_\nu(z)\int_0^{\infty}\left(\frac{\nu+z^2}{2}\right)^{\frac{\nu+1}{2}}\frac{1}{\Gamma\left(\frac{\nu+1}{2}\right)}\,w^{\nu/2}\int_{-\infty}^{\lambda z}\frac{1}{\sqrt{2\pi}}\,e^{-w\left(\frac{z^2+\nu+t^2}{2}\right)}dt\,dw.$$

Logo,

$$f_Z(z) = 2t_\nu(z)\int_{-\infty}^{\lambda z}\left(\frac{\nu+z^2}{2}\right)^{\frac{\nu+1}{2}}\frac{1}{\Gamma\left(\frac{\nu+1}{2}\right)}\frac{1}{\sqrt{2\pi}}\Gamma\left(\frac{\nu}{2}+1\right)\left(\frac{2}{z^2+\nu+t^2}\right)^{\frac{\nu}{2}+1}\times$$

$$\times\int_0^{\infty}\underbrace{\frac{1}{\Gamma\left(\frac{\nu}{2}+1\right)}\left(\frac{z^2+\nu+t^2}{2}\right)^{\frac{\nu}{2}+1}e^{-w\left(\frac{z^2+\nu+t^2}{2}\right)}\,w^{\nu/2}}_{Gama\left(\frac{\nu}{2}+1,\,\frac{z^2+\nu+t^2}{2}\right)}\,dw\,dt$$

$$= 2\,t_\nu(z)\int_{-\infty}^{\lambda z}\frac{\Gamma\left(\frac{\nu+2}{2}\right)}{\Gamma\left(\frac{\nu+1}{2}\right)}\frac{1}{\sqrt{\pi}}\frac{1}{(z^2+\nu)^{1/2}}\left(1+\frac{t^2}{z^2+\nu}\right)^{-\left(\frac{\nu}{2}+1\right)}dt.$$

Fazendo uma nova mudança de variável, $a = t\sqrt{\frac{1+\nu}{\nu+z^2}}$, temos que

$$
\begin{aligned}
f_Z(z) &= 2\,t_\nu(z)\int_{-\infty}^{\lambda z\sqrt{\frac{1+\nu}{\nu+z^2}}} \frac{\Gamma\left(\frac{\nu+2}{2}\right)}{\Gamma\left(\frac{\nu+1}{2}\right)}\frac{1}{\sqrt{\pi(\nu+1)}}\left(\frac{\nu+1+a^2}{\nu+1}\right)^{-\left(\frac{\nu}{2}+1\right)} da \\
&= 2\,t_\nu(z)\int_{-\infty}^{\lambda z\sqrt{\frac{1+\nu}{\nu+z^2}}} \frac{\Gamma\left(\frac{\nu+2}{2}\right)}{\Gamma\left(\frac{\nu+1}{2}\right)}\frac{1}{\sqrt{\pi(\nu+1)}}\left(1+\frac{a^2}{\nu+1}\right)^{-\left(\frac{\nu}{2}+1\right)} da \\
&= 2\,t_\nu(z)\,T_{\nu+1}\left(\lambda z\sqrt{\frac{1+\nu}{\nu+z^2}}\right) \qquad (-\infty < z < \infty).
\end{aligned}
$$

♦

A seguir, obtemos a caracterização da distribuição t-assimétrica padrão através do condicionamento de um vetor aleatório t-Student bivariado. Essa construção é denominada modelo de truncamento oculto (*hidden truncation model*) e pode ser entendida da seguinte maneira: suponha que um determinado fenômeno possa ser modelado através de uma distribuição t-Student bivariada mas que, no entanto, uma das variáveis aleatórias está restrita a alguma condição como, por exemplo, $X_1 > 0$. Se a quantidade de interesse for a variável X_2, observada a condição imposta por X_1, estamos diante de um processo de truncamento em X_1 que induz a assimetria em X_2.

Proposição 2.3 *Se* $X = \begin{pmatrix} X_1 \\ X_2 \end{pmatrix} \sim t_2\left[\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 & \delta \\ \delta & 1 \end{pmatrix}, \nu\right]$, *onde* $\delta = \frac{\lambda}{\sqrt{1+\lambda^2}}$, *então*

$$Z \stackrel{d}{=} X_2 \mid X_1 > 0 \sim ST(\nu, \lambda). \tag{2.4}$$

Prova:

Na Seção 2.1, vimos que toda distribuição t-Student pode ser entendida como uma distribuição t-generalizada. Dessa forma, temos que:

$$X = \begin{pmatrix} X_1 \\ X_2 \end{pmatrix} \sim t_{G_2}\left[\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 & \delta \\ \delta & 1 \end{pmatrix}, \nu, \nu\right].$$

Além disso, por propriedade, temos que

- $X_1 \sim t_G(0,\, 1,\, \nu,\, \nu) \equiv t_\nu(0,\, 1)$,
- $X_2 \sim t_G(0,\, 1, \nu,\, \nu) \equiv t_\nu(0,\, 1)$,
- $X_1 \mid X_2 = z \sim t_G(\delta z, 1-\delta^2, \nu + z^2, \nu + 1)$.

Portanto, $\mathrm{P}(X_1 > 0) = \frac{1}{2}$ e

$$P(X_1 > 0 \mid X_2 = z) = \int_0^\infty \frac{\Gamma\left(\frac{\nu+2}{2}\right)}{\Gamma\left(\frac{\nu+1}{2}\right)} \frac{(\nu+z^2)^{\frac{\nu+1}{2}}}{\sqrt{\pi(1-\delta^2)}} \left[\nu + z^2 + \left(\frac{x-\delta z}{\sqrt{1-\delta^2}}\right)^2\right]^{-\frac{\nu+2}{2}} dx.$$

Então,

$$\begin{aligned} f_Z(z) &= \frac{P(X_1 > 0 \mid X_2 = z) f_{X_2}(z)}{P(X_1 > 0)} \\ &= 2\,\frac{\Gamma\left(\frac{\nu+1}{2}\right)}{\Gamma\left(\frac{\nu}{2}\right)} \frac{1}{\sqrt{\pi\nu}} \left(\frac{\nu+z^2}{\nu}\right)^{-\frac{\nu+1}{2}} \int_0^\infty \frac{\Gamma\left(\frac{\nu+2}{2}\right)}{\Gamma\left(\frac{\nu+1}{2}\right)} \times \\ &\qquad \times \frac{(\nu+z^2)^{\frac{\nu+1}{2}}}{\sqrt{\pi(1-\delta^2)}} \left[\nu + z^2 + \left(\frac{x-\delta z}{\sqrt{1-\delta^2}}\right)^2\right]^{-\frac{\nu+2}{2}} dx \\ &= 2\, t_\nu(z) \int_0^\infty \frac{\Gamma\left(\frac{\nu+2}{2}\right)}{\Gamma\left(\frac{\nu+1}{2}\right)} \frac{(\nu+z^2)^{\frac{\nu+1}{2}}}{\sqrt{\pi(1-\delta^2)}} \left[\nu + z^2 + \left(\frac{x-\delta z}{\sqrt{1-\delta^2}}\right)^2\right]^{-\frac{\nu+2}{2}} dx. \end{aligned}$$

Fazendo a mudança de variável $t = \frac{x-\delta z}{\sqrt{1-\delta^2}}$, obtemos

$$f_Z(z) = 2\, t_\nu(z) \int_{-\frac{\delta z}{\sqrt{1-\delta^2}}}^{\infty} \frac{\Gamma\left(\frac{\nu+2}{2}\right)}{\Gamma\left(\frac{\nu+1}{2}\right)} \frac{(\nu+z^2)^{\frac{\nu+1}{2}}}{\sqrt{\pi}} \left[\nu + z^2 + t^2\right]^{-\frac{\nu+2}{2}} dt$$

$$= 2\, t_\nu(z) \int_{-\frac{\delta z}{\sqrt{1-\delta^2}}}^{\infty} \frac{\Gamma\left(\frac{\nu+2}{2}\right)}{\Gamma\left(\frac{\nu+1}{2}\right)} \frac{1}{\sqrt{\pi(\nu+z^2)}} \left[1 + \left(\frac{t}{\sqrt{\nu+z^2}}\right)^2\right]^{-\frac{\nu+2}{2}} dt.$$

Pela simetria da distribuição t padrão, notamos que

$$\int_{-\frac{\delta z}{\sqrt{1-\delta^2}}}^{\infty} f(t)\, dt = \int_{-\infty}^{\frac{\delta z}{\sqrt{1-\delta^2}}} f(t)\, dt.$$

Dessa forma,

$$f_Z(z) = 2\, t_\nu(z) \int_{-\infty}^{\frac{\delta z}{\sqrt{1-\delta^2}}} \frac{\Gamma\left(\frac{\nu+2}{2}\right)}{\Gamma\left(\frac{\nu+1}{2}\right)} \frac{1}{\sqrt{\pi(\nu+z^2)}} \left[1 + \left(\frac{t}{\sqrt{\nu+z^2}}\right)^2\right]^{-\frac{\nu+2}{2}} dt.$$

Ao fazer a mudança na variável $u = \frac{t}{\sqrt{\nu+z^2}}\sqrt{\nu+1}$, temos

$$f_Z(z) = 2\, t_\nu(z) \int_{-\infty}^{\delta z\sqrt{\frac{\nu+1}{(\nu+z^2)(1-\delta^2)}}} \frac{\Gamma\left(\frac{\nu+2}{2}\right)}{\Gamma\left(\frac{\nu+1}{2}\right)} \frac{1}{\sqrt{\pi(\nu+1)}} \left[1 + \left(\frac{u}{\sqrt{\nu+1}}\right)^2\right]^{-\frac{\nu+2}{2}} du$$

$$= 2\, t_\nu(z)\, T_{\nu+1}\left(\delta\, z\sqrt{\frac{\nu+1}{(\nu+z^2)(1-\delta^2)}}\right).$$

Considerando $\lambda = \frac{\delta}{\sqrt{1-\delta^2}}$, temos que

$$\delta\, z\sqrt{\frac{\nu+1}{(\nu+z^2)(1-\delta^2)}} = \frac{\delta}{\sqrt{1-\delta^2}}\, z\sqrt{\frac{\nu+1}{\nu+z^2}} = \lambda\, z\sqrt{\frac{\nu+1}{\nu+z^2}}.$$

Então,

$$f_Z(z) = 2\, t_\nu(z) T_{\nu+1}\left(\lambda z \sqrt{\frac{1+\nu}{\nu+z^2}}\right) \qquad (-\infty < z < \infty).$$

♦

Uma outra forma de construir a distribuição t-assimétrica consiste em obter Z como uma combinação linear de variáveis aleatórias dependentes. Tal construção se mostra útil na simulação da variável Z e ficou conhecida na literatura como representação estocástica.

Proposição 2.4 *Se* $X = \begin{pmatrix} X_1 \\ X_2 \end{pmatrix} \sim t_2\left[\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \nu\right]$, *onde* $\delta = \frac{\lambda}{\sqrt{1+\lambda^2}}$, *então*

$$Z \stackrel{d}{=} \delta\,|X_1| + \sqrt{1-\delta^2} X_2 \sim ST(\nu, \lambda). \tag{2.5}$$

Prova:

A distribuição t-Student é um caso particular da distribuição t-generalizada. Sendo assim, faremos uso das propriedades apresentadas na demonstração da Proposição 2.3, com $\delta = 0$, para obtermos as distribuições condicionais e marginais do vetor aleatório X.

Agora, considere $Z = \delta\,|X_1| + \sqrt{1-\delta^2}\, X_2$ e $T = X_1$. Vamos obter a função de densidade conjunta de T e Z através do Jacobiano da transformação, então

$$\begin{aligned} f_Z(z) &= \int_{-\infty}^{0} f_{T,Z}(t,z)\, dt + \int_{0}^{\infty} f_{T,Z}(t,z)\, dt \\ &= \int_{-\infty}^{0} f_{X_1,X_2}\left(t, \frac{z+\delta t}{\sqrt{1-\delta^2}}\right) \frac{1}{\sqrt{1-\delta^2}}\, dt + \\ &\qquad + \int_{0}^{\infty} f_{X_1,X_2}\left(t, \frac{z-\delta t}{\sqrt{1-\delta^2}}\right) \frac{1}{\sqrt{1-\delta^2}}\, dt. \end{aligned}$$

Resolvendo separadamente cada uma das integrais, temos:

1º Caso) Considere A $=\int_0^\infty f_{X_1,X_2}\left(t, \frac{z-\delta t}{\sqrt{1-\delta^2}}\right) \frac{1}{\sqrt{1-\delta^2}}\, dt.$

Reescrevendo a expressão acima, segue

$$A = \int_0^\infty f_{X_1|X_2}\left(t \,\middle|\, \frac{z-\delta t}{\sqrt{1-\delta^2}}\right) f_{X_2}\left(\frac{z-\delta t}{\sqrt{1-\delta^2}}\right) \frac{1}{\sqrt{1-\delta^2}}\, dt.$$

A fim de facilitar a notação, considere $x = \frac{z-\delta t}{\sqrt{1-\delta^2}}$. Por propriedade, conhecemos a distribuição de $X_1|X_2$, então:

$$A = \int_0^\infty \frac{\Gamma\left(\frac{\nu+2}{2}\right)}{\Gamma\left(\frac{\nu+1}{2}\right)} \frac{(\nu+x^2)^{\frac{\nu+1}{2}}}{\sqrt{\pi}} (\nu+x^2+t^2)^{-\frac{\nu+2}{2}} \frac{\Gamma\left(\frac{\nu+1}{2}\right)}{\Gamma\left(\frac{\nu}{2}\right)} \frac{1}{\sqrt{\nu\pi}} \left[\frac{\nu+x^2}{\nu}\right]^{-\frac{\nu+1}{2}} \times$$
$$\times \frac{1}{\sqrt{1-\delta^2}}\, dt.$$

Logo,

$$A = \frac{\Gamma\left(\frac{\nu+1}{2}\right)}{\Gamma\left(\frac{\nu}{2}\right)} \frac{1}{\sqrt{\pi}} \frac{\nu^{\nu/2}}{(\nu+z^2)^{\frac{\nu+1}{2}}} \int_0^\infty \frac{\Gamma\left(\frac{\nu+2}{2}\right)}{\Gamma\left(\frac{\nu+1}{2}\right)} \frac{1}{\sqrt{\pi}} (\nu+x^2+t^2)^{-\frac{\nu+2}{2}} (\nu+z^2)^{\frac{\nu+1}{2}} \times$$
$$\times \frac{1}{\sqrt{1-\delta^2}}\, dt.$$

Substituindo x por $\frac{z-\delta t}{\sqrt{1-\delta^2}}$, temos que:

$$A = t_\nu(z) \int_0^\infty \frac{\Gamma\left(\frac{\nu+2}{2}\right)}{\Gamma\left(\frac{\nu+1}{2}\right)} \frac{1}{\sqrt{\pi}} \left[\nu + \left(\frac{z-\delta t}{\sqrt{1-\delta^2}}\right)^2 + t^2\right]^{-\frac{\nu+2}{2}} (\nu+z^2)^{\frac{\nu+1}{2}} \frac{1}{\sqrt{1-\delta^2}}\, dt.$$

Observe que a parcela $\nu + \left(\frac{z-\delta t}{\sqrt{1-\delta^2}}\right)^2 + t^2$ é equivalente a $\nu + \frac{(t-\delta z)^2}{1-\delta^2} + z^2$. Portanto,

$$A = t_\nu(z) \int_0^\infty \frac{\Gamma\left(\frac{\nu+2}{2}\right)}{\Gamma\left(\frac{\nu+1}{2}\right)} \frac{1}{\sqrt{\pi}} \left[\nu + \frac{(t-\delta z)^2}{1-\delta^2} + z^2\right]^{-\frac{\nu+2}{2}} (\nu + z^2)^{\frac{\nu+1}{2}} \frac{1}{\sqrt{1-\delta^2}}\, dt.$$

Fazendo a mudança de variável $k = \frac{t-\delta z}{\sqrt{1-\delta^2}}$, obtemos

$$\begin{aligned} A &= t_\nu(z) \int_{-\frac{\delta z}{\sqrt{1-\delta^2}}}^\infty \frac{\Gamma\left(\frac{\nu+2}{2}\right)}{\Gamma\left(\frac{\nu+1}{2}\right)} \frac{1}{\sqrt{\pi}} (\nu + z^2 + k^2)^{-\frac{\nu+2}{2}} (\nu + z^2)^{\frac{\nu+1}{2}}\, dk \\ &= t_\nu(z) \int_{-\frac{\delta z}{\sqrt{1-\delta^2}}}^\infty \frac{\Gamma\left(\frac{\nu+2}{2}\right)}{\Gamma\left(\frac{\nu+1}{2}\right)} \frac{1}{\sqrt{\pi}} \frac{1}{\sqrt{\nu + z^2}} \left[1 + \frac{k^2}{\nu + z^2}\right]^{-\frac{\nu+2}{2}} dk. \end{aligned}$$

Mudando a variável u por $\frac{k}{\sqrt{\nu+z^2}}\sqrt{\nu+1}$ e substituindo $\frac{\delta}{\sqrt{1-\delta^2}}$ por λ, temos:

$$\begin{aligned} A &= t_\nu(z) \int_{-\lambda z\sqrt{\frac{\nu+1}{\nu+z^2}}}^\infty \frac{\Gamma\left(\frac{\nu+2}{2}\right)}{\Gamma\left(\frac{\nu+1}{2}\right)} \frac{1}{\sqrt{\pi}} \frac{(\nu+1)^{\frac{\nu+1}{2}}}{(\nu+1+u^2)^{\frac{\nu+2}{2}}}\, du \\ &= t_\nu(z) \int_{-\infty}^{\lambda z\sqrt{\frac{\nu+1}{\nu+z^2}}} \frac{\Gamma\left(\frac{\nu+2}{2}\right)}{\Gamma\left(\frac{\nu+1}{2}\right)} \frac{1}{\sqrt{\pi}} \frac{(\nu+1)^{\frac{\nu+1}{2}}}{(\nu+1+u^2)^{\frac{\nu+2}{2}}}\, du \\ &= t_\nu(z)\, T_{\nu+1}\left(\lambda\, z\sqrt{\frac{\nu+1}{\nu+z^2}}\right). \end{aligned}$$

2° Caso) Considere B $=\int_{-\infty}^{0} f_{X_1,X_2}\left(t, \frac{z+\delta t}{\sqrt{1-\delta^2}}\right) \frac{1}{\sqrt{1-\delta^2}}\, dt.$

A resolução desta integral é similar àquela apresentada no 1° *Caso* e seu resultado final é dado por:

$$B = \int_{-\infty}^{0} f_{X_1,X_2}\left(t, \frac{z+\delta t}{\sqrt{1-\delta^2}}\right) \frac{1}{\sqrt{1-\delta^2}}\, dt$$

$$= t_\nu(z)\, T_{\nu+1}\left(\lambda\, z\sqrt{\frac{\nu+1}{\nu+z^2}}\right).$$

Dessa forma,

$$f_Z(z) = \int_{0}^{\infty} f_{X_1,X_2}\left(t, \frac{z-\delta t}{\sqrt{1-\delta^2}}\right) \frac{1}{\sqrt{1-\delta^2}}\, dt + \\ + \int_{-\infty}^{0} f_{X_1,X_2}\left(t, \frac{z+\delta t}{\sqrt{1-\delta^2}}\right) \frac{1}{\sqrt{1-\delta^2}}\, dt$$

$$= t_\nu(z)\, T_{\nu+1}\left(\lambda\, z\sqrt{\frac{\nu+1}{\nu+z^2}}\right) + t_\nu(z)\, T_{\nu+1}\left(\lambda\, z\sqrt{\frac{\nu+1}{\nu+z^2}}\right).$$

Então, conclui-se que:

$$f_Z(z) = 2\, t_\nu(z)\, T_{\nu+1}\left(\lambda\, z\sqrt{\frac{\nu+1}{\nu+z^2}}\right) \qquad (-\infty < z < \infty).$$

♦

A Tabela 2.1 nos apresenta um resumo das diferentes caracterizações para a distribuição t-assimétrica descritas anteriormente.

Tabela 2.1: Resumo das Caracterizações

Proposição	Construção de $Z \sim ST(\nu, \lambda)$
I	Se $X \sim SN(\lambda)$, $W \sim \text{Gama}\left(\frac{\nu}{2}, \frac{1}{2}\right)$ e $X \perp W$, então $Z \stackrel{d}{=} \frac{X}{\sqrt{\frac{W}{\nu}}}$
II	$Z\|W \sim SN\left(0, \frac{1}{W}, \lambda\right)$ e $W \sim \text{Gama}\left(\frac{\nu}{2}, \frac{\nu}{2}\right)$
III	Se $X = \begin{pmatrix} X_1 \\ X_2 \end{pmatrix} \sim t_2\left[\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 & \delta \\ \delta & 1 \end{pmatrix}, \nu\right]$, então $Z \stackrel{d}{=} X_2 \mid X_1 > 0$, para $\delta = \frac{\lambda}{\sqrt{1+\lambda^2}}$
IV	Se $X = \begin{pmatrix} X_1 \\ X_2 \end{pmatrix} \sim t_2\left[\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \nu\right]$, então $Z \stackrel{d}{=} \delta\,\|X_1\| + \sqrt{1-\delta^2}\,X_2$, para $\delta = \frac{\lambda}{\sqrt{1+\lambda^2}}$

2.2.2 As Funções de Distribuição Acumulada e Geradora de Momentos

Nesta seção, apresentamos a função de distribuição acumulada e obtemos a função geradora de momentos da distribuição t-assimétrica padrão. A partir da função geradora de momentos conhecemos medidas importantes da distribuição t-assimétrica, tais como, média, variância e os coeficientes de assimetria e curtose.

Para obtermos a função de distribuição acumulada de uma variável aleatória com distribuição t-assimétrica, devemos considerar, a fim de minimizar cálculos, uma expressão equivalente a primeira caracterização (ver Proposição 2.1), cujo o enunciado pode ser descrito da seguinte maneira: se $X \sim SN(\lambda)$, $V \sim \text{Gama}(\frac{\nu}{2}, \frac{\nu}{2})$ e $X \perp V$ então $Z = V^{-1/2}X$ tem distribuição t-assimétrica padrão. Assim,

$$\begin{aligned} F_Z(z) &= P(Z \leq z) = P(V^{-1/2}X \leq z) \\ &= \int_0^\infty P(X \leq z\sqrt{v} \mid V = v)\, f_V(v)\, dv = \int_0^\infty P(X \leq z\sqrt{v})\, f_V(v)\, dv \\ &= \int_0^\infty F_X(z\sqrt{v})\, f_V(v)\, dv = E_V(F_X\,(z\sqrt{v})). \end{aligned}$$

Rodríguez (2005) apresenta uma expressão que nos permite calcular a função de distribuição de uma variável aleatória $X \sim SN(\lambda)$, a partir da função de distribuição acumulada de uma distribuição normal bivariada, da seguinte forma:

$$F_X(x) = 2\,\Phi_2(x, 0 \mid \Sigma), \text{ com } \Sigma = \begin{bmatrix} 1 & -\delta \\ -\delta & 1 \end{bmatrix}, \; \delta = \frac{\lambda}{\sqrt{1+\lambda^2}},$$

sendo $\Phi_2(. \mid \Sigma)$ a função de distribuição acumulada de uma distribuição normal bivariada com média zero e matriz de variância Σ.

Antes de apresentarmos a função geradora de momentos da distribuição t-assimétrica padrão, apresentaremos o Lema 2.1 (Rodríguez, 2005), que será útil na demonstração da Proposição 2.5.

Lema 2.1 *Se* $V \sim N_k(\eta, \Sigma)$, *então* $E[\Phi_m(a+AV|\gamma,\Gamma)] = \Phi_m(a|\gamma - A\eta, \Gamma + A\Sigma A^T)$.

Proposição 2.5 *A função geradora de momentos da distribuição t-assimétrica padrão é dada por*

$$M_Z(t) = \int_0^\infty h_W(w) M_X\left(\frac{t}{\sqrt{w}}\right) dw, \tag{2.6}$$

onde $h_W(w)$ *corresponde a função densidade de probabilidade de uma variável aleatória* W *com distribuição* $Gama(\frac{\nu}{2}, \frac{\nu}{2})$ *e* $M_X(t)$ *é a função geradora de momentos da distribuição normal-assimétrica padrão, dada por:*

$$M_X(t) = 2\, e^{\frac{t^2}{2}}\, \Phi(\delta t), \qquad com\ \delta = \frac{\lambda}{\sqrt{1+\lambda^2}}. \tag{2.7}$$

Prova:

Através da Proposição 2.2, cujo o resultado nos diz que Z pode ser representado como uma mistura da distribuição normal-assimétrica no parâmetro de escala, temos que:

$$\begin{aligned}
M_Z(t) &= E[e^{tz}] = \int_{-\infty}^{\infty} e^{tz} \int_0^\infty 2\, \phi(z\sqrt{w})\, \Phi(\lambda z\sqrt{w})\, \frac{\left(\frac{\nu}{2}\right)^{\nu/2}}{\Gamma\left(\frac{\nu}{2}\right)}\, e^{-\frac{\nu}{2}w}\, w^{\frac{\nu-1}{2}}\, dw\, dz \\
&= \int_{-\infty}^{\infty} e^{tz} \int_0^\infty 2\frac{1}{\sqrt{2\pi}} e^{-\frac{z^2 w}{2}} \Phi(\lambda z\sqrt{w})\, \frac{\left(\frac{\nu}{2}\right)^{\nu/2}}{\Gamma\left(\frac{\nu}{2}\right)}\, e^{-\frac{\nu}{2}w}\, w^{\frac{\nu-1}{2}}\, dw\, dz \\
&= \int_0^\infty \frac{\left(\frac{\nu}{2}\right)^{\nu/2}}{\Gamma\left(\frac{\nu}{2}\right)} \frac{e^{-\frac{\nu}{2}w}}{\sqrt{2\pi}}\, w^{\frac{\nu-1}{2}}\, 2 \int_{-\infty}^{\infty} e^{-\frac{1}{2}\left(z\sqrt{w} - \frac{t}{\sqrt{w}}\right)^2} e^{\frac{t^2}{2w}} \Phi(\lambda z\sqrt{w})\, dz\, dw \\
&= \int_0^\infty \frac{\left(\frac{\nu}{2}\right)^{\nu/2}}{\Gamma\left(\frac{\nu}{2}\right)} \frac{e^{-\frac{\nu}{2}w}}{\sqrt{2\pi}}\, w^{\frac{\nu-1}{2}}\, 2\, e^{\frac{t^2}{2w}} \int_{-\infty}^{\infty} e^{-\frac{1}{2}\left(z\sqrt{w} - \frac{t}{\sqrt{w}}\right)^2} \Phi(\lambda z\sqrt{w})\, dz\, dw.
\end{aligned}$$

Fazendo a mudança de variável $x = z\sqrt{w} - \frac{t}{\sqrt{w}}$, então

$$M_Z(t) = \int_0^\infty \frac{\left(\frac{\nu}{2}\right)^{\nu/2}}{\Gamma\left(\frac{\nu}{2}\right)} e^{-\frac{\nu}{2}w} w^{\frac{\nu-1}{2}} 2\, e^{\frac{t^2}{2w}} \int_{-\infty}^\infty \frac{e^{-\frac{1}{2}x^2}}{\sqrt{2\pi}} \Phi\left(\lambda x + \frac{\lambda t}{\sqrt{w}}\right) \frac{1}{\sqrt{w}}\, dx\, dw$$

$$= \int_0^\infty \frac{\left(\frac{\nu}{2}\right)^{\nu/2}}{\Gamma\left(\frac{\nu}{2}\right)} e^{-\frac{\nu}{2}w} w^{\frac{\nu}{2}-1} 2\, e^{\frac{t^2}{2w}} \underbrace{\int_{-\infty}^\infty \phi(x)\Phi\left(\lambda x + \frac{\lambda t}{\sqrt{w}}\right) dx}_{E\left[\Phi\left(\lambda x + \frac{\lambda t}{\sqrt{w}}\right)\right]}\, dw.$$

Utilizando o Lema 2.1, para $k = m = 1$, $\eta = 0$, $\Sigma = 1$, $a = \frac{\lambda t}{\sqrt{w}}$, $A = \lambda$, $\gamma = 0$ e $\Gamma = 1$, obtemos que

$$\begin{aligned} M_Z(t) &= \int_0^\infty \frac{\left(\frac{\nu}{2}\right)^{\nu/2}}{\Gamma\left(\frac{\nu}{2}\right)} e^{-\frac{\nu}{2}w} w^{\frac{\nu}{2}-1} 2\, e^{\frac{t^2}{2w}} \Phi\left(\frac{\lambda}{\sqrt{1+\lambda^2}} \frac{t}{\sqrt{w}}\right) dw \\ &= \int_0^\infty \frac{\left(\frac{\nu}{2}\right)^{\nu/2}}{\Gamma\left(\frac{\nu}{2}\right)} e^{-\frac{\nu}{2}w} w^{\frac{\nu}{2}-1} 2\, e^{\frac{t^2}{2w}} \Phi\left(\delta \frac{t}{\sqrt{w}}\right) dw \\ &= \int_0^\infty h_W(w) M_X\left(\frac{t}{\sqrt{w}}\right) dw, \end{aligned}$$

onde $h_W(.)$ é a função densidade de probabilidade de uma variável aleatória W com distribuição Gama$\left(\frac{\nu}{2}, \frac{\nu}{2}\right)$ e M_X é a função geradora de momentos de uma variável aleatória $X \sim SN(\lambda)$.

♦

A seguir apresentamos os momentos de ordem par e ímpar da distribuição t-assimétrica padrão obtidos a partir da Proposição 2.5.

$$E(Z^{2k}) = 2^{-k} \frac{(2k)!}{k!} \left(\frac{\nu}{2}\right)^k \frac{\Gamma\left(\frac{\nu-2k}{2}\right)}{\Gamma\left(\frac{\nu}{2}\right)}, \qquad \text{para } \nu > 2k, \tag{2.8}$$

$$E(Z^{2k+1}) = \sqrt{\frac{2}{\pi}}\lambda(1+\lambda^2)^{-\left(k+\frac{1}{2}\right)}\, 2^{-k}\,(2k+1)!\sum_{j=0}^{k}\frac{j!\,(2\lambda)^{2j}}{(2j+1)!(k-j)!}\left(\frac{\nu}{2}\right)^{\frac{2k+1}{2}} \times$$

$$\times\frac{\Gamma\left(\frac{\nu-2k-1}{2}\right)}{\Gamma\left(\frac{\nu}{2}\right)}, \qquad \text{para } \nu > 2k+1. \tag{2.9}$$

Conseqüentemente, têm-se que a esperança e variância são dadas por:

$$E(Z) = \frac{\lambda}{\sqrt{1+\lambda^2}}\left(\frac{\nu}{\pi}\right)^{1/2}\frac{\Gamma\left(\frac{\nu-1}{2}\right)}{\Gamma\left(\frac{\nu}{2}\right)}, \qquad \text{para } \nu > 1,$$

$$Var(Z) = \frac{\nu}{\nu-2} - \beta^2, \qquad \text{para } \nu > 2 \text{ e } \beta = E(Z).$$

Seja X uma variável aleatória qualquer, com valor esperado μ e desvio-padrão σ. Supondo-se a existência do terceiro e do quarto momentos de X, os coeficientes de assimetria e curtose são definidos, respectivamente, por:

$$\gamma_1 = \frac{E[(X-\mu)^3]}{\sigma^3} \quad \text{e} \quad \gamma_2 = \frac{E[(X-\mu)^4]}{\sigma^4}.$$

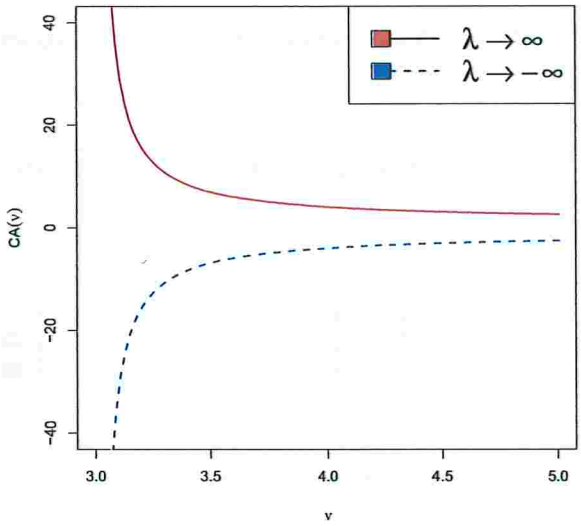
Dessa forma, considerando $\beta = E(Z)$ e utilizando-se dos resultados em (2.8) e (2.9) obtém-se o coeficiente de assimetria para a distribuição t-assimétrica padrão, apresentado a seguir:

$$\gamma_1 = \beta\left[(3-\delta^2)\frac{\nu}{\nu-3} - 3\frac{\nu}{\nu-2} + 2\beta^2\right]\left[\frac{\nu}{\nu-2} - \beta^2\right]^{-\frac{3}{2}}, \quad \text{para } \nu > 3.$$

Tal medida nos indica o quanto a distribuição t-assimétrica se afasta da condição de simetria. Observe que, para $|\lambda|$ infinito, γ_1 é uma função decrescente em ν.

A Figura 2.2 nos mostra que a medida que os graus de liberdade variam, o coeficiente de assimetria fica restrito a diferentes intervalos entre as curvas. Quando ν tende ao infinito, temos que $\gamma_1 \in [-0.99527, 0.99527]$. Além disso,

Figura 2.2: Coeficiente de assimetria



vemos que para valores pequenos de graus de liberdade, o coeficiente de assimetria varia num intervalo mais amplo. Observe que o menor intervalo para o coeficiente de assimetria corresponde justamente ao caso em os dados podem ser modelados através de uma distribuição normal-assimétrica ($\nu \to \infty$) e, portanto, a distribuição t-assimétrica tem uma maior flexibilidade em modelar amostras com grandes assimetrias do que a distribuição normal-assimétrica.

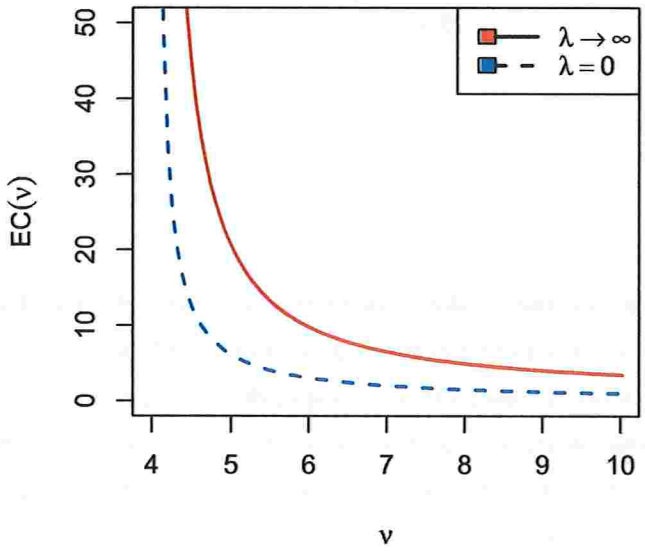
O coeficiente de curtose da distribuição t-assimétrica padrão, para $\beta = E(Z)$ e $\nu > 4$, é expresso por:

$$\gamma_2 = \left[\frac{3\nu^2}{(\nu-2)(\nu-4)} - \frac{4\beta^2\nu(3-\delta^2)}{\nu-3} + \frac{6\beta^2\nu}{\nu-2} - 3\beta^4\right]\left[\frac{\nu}{\nu-2} - \beta^2\right]^{-2}.$$

A curtose é uma medida que reflete o grau de achatamento de uma distribuição. Distribuições que possuem valor de curtose iguais a 3 são denominadas mesocúrticas. Aquelas distribuições que possuem $\gamma_2 < 3$ são conhecidas por distribuições de caudas leves (platicúrtica). As que possuem $\gamma_2 > 3$ são denominadas distribuições de caudas pesadas (leptocúrtica).

Uma distribuição normal padrão possui coeficiente de curtose dado por 3, indicando que o formato de sua curva é mesocúrtica. A seguir, apresentamos o gráfico do coeficiente de excesso de curtose para a distribuição t-assimétrica em relação a distribuição normal padrão, nos casos extremos em que $\lambda \rightarrow \infty$ e $\lambda = 0$.

Figura 2.3: Excesso de curtose em função de ν

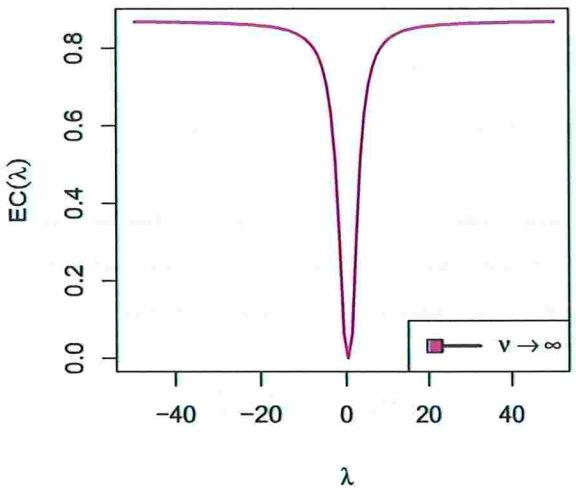


O estudo destas duas curvas é relevante pois, considerando ν fixo ($\nu > 4$), γ_2 será função par em λ e, sendo assim, os casos extremos do excesso de curtose relacionados a ν ficam representados por $\lambda \rightarrow \infty$ e $\lambda = 0$. No caso de $\lambda = 0$, o excesso de curtose da distribuição t-assimétrica corresponde ao excesso de curtose da distribuição t padrão que é $\frac{6}{\nu-4}$. Dessa forma, para qualquer $\nu > 4$ e λ, a distribuição t-assimétrica possui caudas pesadas.

Para cada número de graus de liberdade fixado, obtemos um intervalo para o excesso de curtose limitado pelas curvas apresentadas na Figura 2.3.

A Figura 2.4 apresenta a curva do excesso de curtose em função de λ, quando consideramos $\nu \rightarrow \infty$.

Figura 2.4: Excesso de curtose em função de λ



Ao analisar o gráfico de excesso de curtose para $\nu \to \infty$, vemos que coeficiente assume valores limitados a [0, 0.8691], coincidindo, assim, com o intervalo de excesso de curtose esperado para a distribuição normal-assimétrica. Note que, quanto maior a assimetria, maior será a curtose da distribuição t-assimétrica.

2.3 Distribuição t-assimétrica Posição e Escala

Com a introdução dos parâmetros de posição ($\mu \in \mathbb{R}$) e escala ($\sigma > 0$), conseguimos obter uma generalização da distribuição t-assimétrica padrão. Assim, uma variável aleatória Y será denominada t-assimétrica de posição e escala, ou simplesmente t-assimétrica, se $Y = \mu + \sigma Z$, onde $Z \sim ST(\nu, \lambda)$. A notação utilizada será $Y \sim ST(\mu, \sigma^2, \nu, \lambda)$.

Como conseqüência, a função densidade de Y será dada por:

$$f_Y(y) = \frac{2}{\sigma}\, t_\nu\left(\frac{y-\mu}{\sigma}\right) T_{\nu+1}\left[\lambda\left(\frac{y-\mu}{\sigma}\right)\sqrt{\frac{\nu+1}{\nu+\sigma^{-2}(y-\mu)^2}}\right].$$

Adotaremos a seguinte nomenclatura para os parâmetros:

- μ: parâmetro de posição,
- σ^2: parâmetro de escala,
- ν: graus de liberdade,
- λ: parâmetro de forma ou assimetria.

Assim como no caso padrão, a função geradora de momentos da distribuição t-assimétrica (Y) pode ser escrita como uma mistura da função geradora de momentos da distribuição normal-assimétrica. Então,

$$M_Y(t) = e^{t\mu} \int_0^{\infty} h(w) M_X\left(\frac{t\sigma}{\sqrt{w}}\right) dw, \tag{2.10}$$

onde $h(w)$ corresponde à função densidade de probabilidade de uma distribuição Gama$\left(\frac{\nu}{2}, \frac{\nu}{2}\right)$ e $M_X(t)$ corresponde à função geradora de momentos da distribuição normal-assimétrica padrão.

Para obtermos a expressão (2.10) basta considerar:

$$\begin{aligned} M_Y(t) &= E[e^{ty}] = \int_{-\infty}^{\infty} e^{ty} f_Y(y) dy \\ &= \int_{-\infty}^{\infty} e^{ty} \frac{2}{\sigma} t_\nu\left(\frac{y-\mu}{\sigma}\right) T_{\nu+1}\left[\lambda\left(\frac{y-\mu}{\sigma}\right)\sqrt{\frac{\nu+1}{\nu+\sigma^{-2}(y-\mu)^2}}\right] dy. \end{aligned}$$

Fazendo $z = \frac{y-\mu}{\sigma}$, temos que

$$\begin{aligned} M_Y(t) &= \int_{-\infty}^{\infty} 2\, e^{t(\mu+z\sigma)}\, t_v(z)\, T_{\nu+1}\left(\lambda z\sqrt{\frac{\nu+1}{\nu+z^2}}\right) dz \\ &= e^{t\mu} \int_{-\infty}^{\infty} 2\, e^{tz\sigma}\, t_v(z)\, T_{\nu+1}\left[\lambda z\sqrt{\frac{\nu+1}{\nu+z^2}}\right] dz \\ &= e^{t\mu}\, M_Z(t\sigma), \quad \text{onde } Z \sim ST(\nu, \lambda). \end{aligned}$$

Portanto,

$$M_Y(t) = e^{t\mu} \int_0^{\infty} h(w)\, M_X\left(\frac{t\sigma}{\sqrt{w}}\right) dw.$$

A partir de (2.10), obtemos a média e a variância de uma variável aleatória $Y \sim St(\mu,\, \sigma^2,\, \nu,\, \lambda)$.

$$E(Y) = \mu \,+\, \sigma \sqrt{\frac{\nu}{\pi}}\, \delta\, \frac{\Gamma\left(\frac{\nu-1}{2}\right)}{\Gamma\left(\frac{\nu}{2}\right)}, \quad \text{onde } \delta = \frac{\lambda}{\sqrt{1+\lambda^2}}\,,\ \nu > 1$$

e

$$Var(Y) = \sigma^2\, \nu \left[\frac{1}{\nu-2} - \frac{\delta^2}{\pi}\, \frac{\Gamma^2\left(\frac{\nu-1}{2}\right)}{\Gamma^2\left(\frac{\nu}{2}\right)}\right], \quad \text{onde } \delta = \frac{\lambda}{\sqrt{1+\lambda^2}} \text{ e } \nu > 2.$$

2.3.1 Propriedades

A seguir apresentamos algumas propriedades da distribuição t-assimétrica, que serão demonstradas no apêndice A. Dessa forma, considere $Z \sim ST(\nu,\, \lambda)$ e $Y \sim ST(\mu,\, \sigma^2,\, \nu,\, \lambda)$:

(P1) $|Z| \sim HT(\nu)$.

A notação $HT(\nu)$ indica que uma variável aleatória T tem distribuição t-positiva ou *half-t* com ν graus de liberdade. Sua função de densidade é dada por:

$$\begin{aligned} f_T(t) &= 2\, \frac{\Gamma\left(\frac{\nu+1}{2}\right)}{\Gamma\left(\frac{\nu}{2}\right)}\, \frac{1}{\sqrt{\nu\pi}} \left[1 + \frac{t^2}{\nu}\right]^{-\frac{\nu+1}{2}} \qquad (t \geq 0), \qquad (2.11) \\ &= 2\, f_U(t), \quad \text{onde } U \sim t_\nu \text{ e } t \geq 0. \end{aligned}$$

(P2) Se $\nu = 1$ então $Z \sim SC\,(\lambda)$.

A notação $SC\,(\lambda)$ é usada para identificar a distribuição Cauchy-assimétrica com parâmetro de assimetria λ. Sua função de densidade é dada por:

$$f_{SC}(x) = \frac{1}{\pi(1+x^2)}\left[1 + \frac{\lambda x}{\sqrt{1+(1+\lambda^2)x^2}}\right] \qquad (-\infty < x < \infty). \tag{2.12}$$

Observe que a distribuição de Cauchy padrão é um caso particular do modelo Cauchy-assimétrico, para $\lambda = 0$.

(P3) Quando $\lambda \longrightarrow \infty$, Z converge a uma $HT(\nu)$.

(P4) Quando $\nu \longrightarrow \infty$, Z converge a uma $SN(\lambda)$.

(P5) $-Z \sim ST(\nu, -\lambda)$.

(P6) $Z^2 \sim$ *F-Snedocor*$(1, \nu)$.

(P7) $F_Z(z; \nu, -\lambda) = 1 - F_Z(-z; \nu, \lambda)$.

(P8) Se $Y \sim St(\mu, \sigma^2, \nu, \lambda)$ e $Y_1 = a + bY$ então
$Y_1 \sim ST(a + b\mu;\ b^2\sigma^2,\ sinal(b)\lambda,\ \nu)$.

É importante notar que a propriedade (P8) nos permite dizer que a distribuição t-assimétrica é fechada para transformações lineares, ou seja, qualquer combinação linear de uma variável aleatória com distribuição t-assimétrica será também uma distribuição t-assimétrica.

Capítulo 3

Inferência Clássica

Nesta seção vamos considerar alguns métodos clássicos que possibilitam a obtenção de estimadores para os parâmetros da distribuição t-assimétrica.

De maneira geral, os procedimentos clássicos fundamentam-se na idéia de amostragem repetida. Segundo este princípio, "... *os métodos estatísticos devem apreciar-se através do respectivo comportamento num número indefinido de repetições - hipotéticas - efectuadas nas mesmas condições. Uma das faces do princípio reside precisamente na interpretação frequencista de probabilidade, isto é, na utilização de frequências como medidas de incerteza; a outra face reside na avaliação dos procedimentos estatísticos em termos de frequência com que produzem respostas correctas ou bons resultados*" (Paulino, Turkman & Murteira, 2003).

O primeiro método considerado é o método dos momentos em que os estimadores são obtidos igualando-se os momentos amostrais aos seus respectivos momentos populacionais. Na seção subseqüente, apresentamos o método de máxima verossimilhança em que os estimadores são obtidos a partir da maximização da função de verossimilhança. Em ambos os métodos, serão discutidas a funcionalidade e a existência de tais estimadores.

3.1 Método dos Momentos

Inicialmente, analisaremos o caso padrão. Seja $\mathbf{Z} = (Z_1, Z_2, ..., Z_n)$ uma amostra aleatória simples de tamanho n da variável aleatória Z com distribuição $ST(\nu, \lambda)$. O estimador de momentos de λ e ν é a solução da seguintes equações:

$$\frac{\lambda}{\sqrt{1+\lambda^2}} \sqrt{\frac{\nu}{\pi}} \frac{\Gamma\left(\frac{\nu-1}{2}\right)}{\Gamma\left(\frac{\nu}{2}\right)} = \overline{Z} \quad e \quad \frac{\nu}{\nu-2} = m_2, \tag{3.1}$$

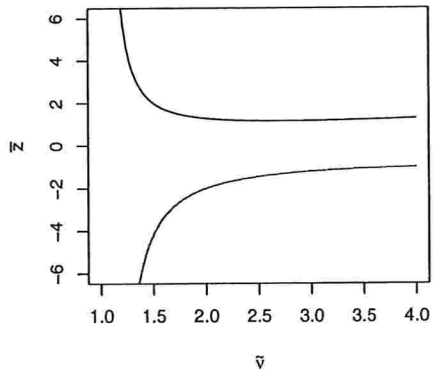
sendo $\overline{Z} = \frac{\sum_{i=1}^{n} Z_i}{n}$ e $m_2 = \frac{\sum_{i=1}^{n} Z_i^2}{n}$. Logo, temos que os estimadores para os parâmetros de assimetria e curtose obtidos via método dos momentos são dados, respectivamente, por:

$$\widetilde{\lambda} = \frac{\overline{Z}\,\Gamma\left(\frac{\widetilde{\nu}}{2}\right)}{\sqrt{\frac{\widetilde{\nu}}{\pi}\,\Gamma^2\left(\frac{\widetilde{\nu}-1}{2}\right) - \overline{Z}^2\,\Gamma^2\left(\frac{\widetilde{\nu}}{2}\right)}} \quad \text{e} \quad \widetilde{\nu} = \frac{2m_2}{m_2 - 1}.$$

Note que a existência do estimador $\widetilde{\lambda}$ está condicionada ao radicando ser maior do que zero, assim como o estimador de $\widetilde{\nu}$ só existe quando o denominador da fração é positivo e $m_2 > 0$. Dessa forma, estas condições de existência se traduzem em

$$\overline{Z} \in \left]-\sqrt{\frac{\widetilde{\nu}}{\pi}} \frac{\Gamma\left(\frac{\widetilde{\nu}-1}{2}\right)}{\Gamma\left(\frac{\widetilde{\nu}}{2}\right)};\ \sqrt{\frac{\widetilde{\nu}}{\pi}} \frac{\Gamma\left(\frac{\widetilde{\nu}-1}{2}\right)}{\Gamma\left(\frac{\widetilde{\nu}}{2}\right)}\right[\quad \text{e} \quad m_2 > 1.$$

Figura 3.1: Domínio da média amostral



A Figura 3.1 apresenta as curvas do limite superior e inferior do intervalo no qual a média amostral deverá estar contida para garantir a existência do estimador de momentos dos parâmetros da distribuição t-assimétrica padrão. Note que, quando ν tende ao infinito, Z tem distribuição $SN(\lambda)$ e, neste caso, $\overline{Z} \in \left[-\sqrt{\frac{2}{\pi}}, \sqrt{\frac{2}{\pi}}\right]$.

Vejamos, agora, o caso em que $\mathbf{Y} = (Y_1, Y_2, ..., Y_n)$ é uma amostra aleatória simples de tamanho n da variável aleatória Y com distribuição $ST(\mu, \sigma^2, \nu, \lambda)$. Os estimadores de momentos dos parâmetros ν, σ e $\delta = \frac{\lambda}{\sqrt{1+\lambda^2}}$ da distribuição t-assimétrica são apresentados abaixo:

$$\widetilde{\nu} = \frac{4\, a(\widetilde{\mu}) - 6\, b(\widetilde{\mu})^2}{a(\widetilde{\mu}) - 3\, b(\widetilde{\mu})^2},$$

$$\widetilde{\sigma}^2 = \frac{a(\widetilde{\mu})\, b(\widetilde{\mu})}{2\, a(\widetilde{\mu}) - 3\, b(\widetilde{\mu})^2},$$

$$\widetilde{\delta}^2 = \left[\frac{\Gamma\left(\frac{\widetilde{\nu}}{2}\right)}{\Gamma\left(\frac{\widetilde{\nu}-1}{2}\right)}\right]^2 \frac{\pi\, (m_1 - \widetilde{\mu})^2\, (a(\widetilde{\mu}) - 3\, b(\widetilde{\mu})^2)}{2\, a(\widetilde{\mu})\, b(\widetilde{\mu})},$$

sendo $m_j = \sum_{i=1}^{n} \frac{Y_i{}^j}{n}$, para $j = 1, ..., 4$ e

$$a(\widetilde{\mu}) = m_4 - 4\widetilde{\mu}^3 m_1 + 6\widetilde{\mu}^2 m_2 - 4\widetilde{\mu} m_3 + \widetilde{\mu}^4 \qquad \text{e} \qquad b(\widetilde{\mu}) = -2\widetilde{\mu} m_1 + m_2 + \widetilde{\mu}^2.$$

Observe que os estimadores de ν, δ^2, σ^2 dependem do estimador de momentos de μ. O estimador para o parâmetro de posição é obtido a partir da seguinte equação:

$$\widetilde{\sigma}^2\, (3 - \widetilde{\delta}^2)\, \frac{\widetilde{\nu}}{\widetilde{\nu} - 3}\, (m_1 - \widetilde{\mu}) = m_3 + 3\widetilde{\mu}^2 m_1 - 3\widetilde{\mu} m_2 - \widetilde{\mu}^3.$$

A solução desta equação pode ser obtida através de métodos numéricos ou aproximações. No apêndice B, encontram-se as demonstrações para a obtenção dos estimadores de momentos para os parâmetros da distribuição t-assimétrica padrão e geral.

3.2 Método de Máxima Verossimilhança

Inicialmente, vamos discutir o caso em que $\mathbf{Z} = (Z_1, Z_2, ..., Z_n)$ é uma amostra aleatória simples de tamanho n da variável aleatória Z com distribuição $ST(\nu, \lambda)$. Neste caso, a função de verossimilhança é dada por:

$$L(\nu, \lambda; \mathbf{z}) = \prod_{i=1}^{n} 2\, t_\nu(z_i) T_{\nu+1}\left(\lambda z_i \sqrt{\frac{1+\nu}{\nu + z_i^2}}\right).$$

Se considerarmos que o número de graus de liberdade é conhecido ($\nu > 0$), a função de verossimilhança dependerá apenas do parâmetro λ, logo:

$$L(\lambda; \nu, \mathbf{z}) \propto \prod_{i=1}^{n} T_{\nu+1}\left(\lambda z_i \sqrt{\frac{1+\nu}{\nu + z_i^2}}\right).$$

Considerando esta situação, se $\forall i\ z_i > 0$, então $L(\lambda; \mathbf{z}, \nu)$ é uma função monótona crescente em λ e, como conseqüência, o estimador que maximiza a função de verossimilhança será infinito. Utilizando o mesmo raciocínio, verificamos que se $\forall i\ z_i < 0$, então o e.m.v. será menos infinito.

Na Proposição 3.1, vamos determinar a probabilidade de se conseguir amostras cujo os e.m.v sejam infinitos. Para isso, precisamos enunciar o seguinte lema (Liseo & Loperfido, 2006):

Lema 3.1 *Seja* $\mathbf{X} = (X_1, X_2, \ldots, X_n)$ *uma amostra aleatória simples de tamanho* n *de* $X \sim SN(\lambda)$. *A probabilidade de se obter uma amostra com e.m.v infinito é dada por:*

$$P(X < 0)^n + P(X > 0)^n = \left(\frac{1}{2} - \frac{1}{\pi}\, arctan\, \lambda\right)^n + \left(\frac{1}{2} + \frac{1}{\pi}\, arctan\, \lambda\right)^n.$$

Proposição 3.1 *Considere* $\mathbf{Z} = (Z_1, Z_2, \ldots, Z_n)$ *uma amostra aleatória simples de tamanho* n *de* $Z \sim ST(\nu, \lambda)$, *com* $\nu > 0$, *conhecido. O e.m.v. do parâmetro de assimetria será infinito com probabilidade igual ao caso em que consideramos* $\mathbf{Z}$ *uma amostra aleatória de tamanho* n *de* $Z \sim SN(\lambda)$.

Prova:

Como $Z_1, Z_2, \ldots, Z_n$ é uma seqüência de variáveis aleatórias independentes e identicamente distribuídas de $Z \sim ST(\nu, \lambda)$, temos que

$$P(Z_1 > 0, Z_2 > 0, ..., Z_n > 0) = P(Z_1 > 0)P(Z_2 > 0)...P(Z_n > 0) = [P(Z_1 > 0)]^n.$$

Sabemos que Z_1 pode ser obtida através da caracterização 2, que expressa que se $Z_1|W \sim SN\left(0, \frac{1}{W}, \lambda\right)$ e $W \sim \text{Gama}\left(\frac{\nu}{2}, \frac{\nu}{2}\right)$, então $Z_1 \sim ST(\nu, \lambda)$. Dessa forma,

$$\begin{aligned}
P(Z_1 > 0) &= \int_0^\infty f_{Z_1}(z)\, dz = \int_0^\infty \int_0^\infty f_{Z_1,W}(z,w)\, dw\, dz \\
&= \int_0^\infty \int_0^\infty f_{Z_1|W}(z|w) f_W(w)\, dw\, dz \\
&= \int_0^\infty f_W(w) \int_0^\infty f_{Z_1|W}(z|w)\, dz\, dw \\
&= \int_0^\infty f_W(w)\, P(Z_1 > 0|W = w)\, dw \\
&= \int_0^\infty f_W(w)\, P\left(\frac{Z_1}{\sqrt{\frac{1}{w}}} > \frac{0}{\sqrt{\frac{1}{w}}}|W = w\right) dw \\
&= \int_0^\infty f_W(w)\, P\left(Y > 0|W = w\right)\, dw, \qquad \text{onde } Y|W \sim SN(\lambda).
\end{aligned}$$

Portanto,

$$
\begin{aligned}
P(Z_1 > 0) &= \int_0^{\infty} f_W(w) \left(\frac{1}{2} + \frac{1}{\pi}\, arctan\, \lambda\right) dw \\
&= \left(\frac{1}{2} + \frac{1}{\pi}\, arctan\, \lambda\right) \int_0^{\infty} \underbrace{f_W(w)}_{Gama\left(\frac{\nu}{2},\frac{\nu}{2}\right)} dw \\
&= \frac{1}{2} + \frac{1}{\pi}\, arctan\, \lambda.
\end{aligned}
$$

De maneira análoga, conseguimos mostrar que $\mathrm{P}(Z_1 < 0) = \frac{1}{2} - \frac{1}{\pi}\, arctan\, \lambda$. Portanto,

$$
P(Z < 0)^n + P(Z > 0)^n = \left(\frac{1}{2} - \frac{1}{\pi}\, arctan\, \lambda\right)^n + \left(\frac{1}{2} + \frac{1}{\pi}\, arctan\, \lambda\right)^n. \quad (3.2)
$$

♦

Na Tabela 3.1 apresentamos algumas destas probabilidades calculadas para valores específicos de λ e diferentes tamanhos de amostra.

Tabela 3.1: Probabilidade do e.m.v ser infinito

	$\lambda = 0$	$\lambda = 1$	$\lambda = 2$	$\lambda = 3$	$\lambda = 5$	$\lambda = 8$	$\lambda = 10$	$\lambda = 20$
$n = 5$	0.063	0.238	0.450	0.582	0.722	0.817	0.851	0.923
$n = 10$	0.002	0.056	0.202	0.339	0.522	0.667	0.724	0.851
$n = 20$	0.000	0.003	0.041	0.115	0.273	0.445	0.524	0.726
$n = 50$	0.000	0.000	0.003	0.004	0.038	0.132	0.199	0.449
$n = 80$	0.000	0.000	0.000	0.000	0.006	0.040	0.076	0.277
$n = 100$	0.000	0.000	0.000	0.000	0.001	0.017	0.039	0.201

Sartori (2006) apresenta o resultado (3.2) e cita que quando consideramos ambos os parâmetros da distribuição t-assimétrica padrão desconhecidos, a probabilidade de se obter o e.m.v do parâmetro de assimetria infinito é menor que

o mencionado na Proposição 3.1. Tal situação se deve a estimação do parâmetro de graus de liberdade.

Além disso, a estimativa para o parâmetro de graus de liberdade também pode ser infinita (ver Fonseca, 2004). Um valor de $\hat{\nu}$ infinito indica que a amostra foi gerada a partir do modelo normal-assimétrico. Da mesma forma que um valor $\hat{\lambda}$ infinito, indica que a amostra é proveniente de um modelo t-positivo. No entanto, existe uma probabilidade não desprezível de que distribuições com parâmetro de assimetria finito gerem amostras com estimativas de máxima verossimilhança infinitas, como pode ser observado na Tabela 3.1.

Nos restringiremos agora aos casos em que as amostras são mais regulares, apresentando os e.m.v. finitos, tanto para o parâmetro de assimetria, quanto para os graus de liberdade.

Considere agora $\mathbf{Y} = (Y_1, Y_2, \ldots, Y_n)$ uma amostra aleatória simples de tamanho n de $Y \sim ST(\mu,\ \sigma^2,\ \lambda,\ \nu)$. Neste caso, a função de verossimilhança é dada por:

$$L(\mu, \sigma, \lambda, \nu; \mathbf{y}) = \prod_{i=1}^{n} \frac{2}{\sigma}\, t_v\left(\frac{y_i - \mu}{\sigma}\right) T_{v+1}\left[\lambda\left(\frac{y_i - \mu}{\sigma}\right)\sqrt{\frac{\nu + 1}{\nu + \sigma^{-2}(y_i - \mu)^2}}\right].$$

Aplicando o logaritmo natural em ambos os lados da equação acima, temos:

$$\begin{aligned} l(\mu, \sigma, \lambda, \nu; \mathbf{y}) &= \log L(\mu, \sigma, \lambda, \nu; \mathbf{y}) \\ &= n \log 2 - n \log \sigma + \sum_{i=1}^{n} \log\ t_v\left(\frac{y_i - \mu}{\sigma}\right) + \\ &\qquad \sum_{i=1}^{n} \log\ T_{v+1}\left[\lambda\left(\frac{y_i - \mu}{\sigma}\right)\sqrt{\frac{\nu + 1}{\nu + \sigma^{-2}(y_i - \mu)^2}}\right]. \end{aligned}$$

Ao derivar $l(\mu, \sigma, \lambda, \nu; \mathbf{y})$ em relação a cada parâmetro, obtemos os seguintes resultados:

$$\frac{\partial \ell}{\partial \mu} = \sum_{i=1}^{n} \left\{ \frac{1}{\sigma\left(\nu + a_i^2\right)} \left[\left(\nu + 1\right) a_i - \frac{t_{\nu+1}\left(\lambda a_i k_i\right)}{T_{\nu+1}\left(\lambda a_i k_i\right)} \lambda k_i \nu \right] \right\},$$

$$\frac{\partial \ell}{\partial \sigma} = -\frac{n}{\sigma} + \sum_{i=1}^{n} \left\{ \frac{a_i k_i}{\sigma} \left[\frac{a_i k_i}{\sigma} - \lambda \frac{\nu}{\nu + a_i^2} \left(\frac{t_{\nu+1}(\lambda a_i k_i)}{T_{\nu+1}(\lambda a_i k_i)} \right) \right] \right\},$$

$$\frac{\partial \ell}{\partial \lambda} = \sum_{i=1}^{n} a_i k_i \frac{t_{\nu+1}(\lambda a_i k_i)}{T_{\nu+1}(\lambda a_i k_i)},$$

$$\frac{\partial \ell}{\partial \nu} = \sum_{i=1}^{n} \left\{ \frac{1}{2} \left[\psi\left(\frac{\nu+1}{2}\right) - \psi\left(\frac{\nu}{2}\right) + 1 - \log\left(1 + \frac{a_i^2}{\nu}\right) - k_i^2 \right] + \right.$$
$$\left. + \frac{\partial}{\partial \nu} \log T_{\nu+1}\left(\lambda a_i k_i\right) \right\}.$$

Observe que $a_i = \frac{y_i - \mu}{\sigma}$, $k_i = \sqrt{\frac{\nu+1}{\nu + a_i^2}}$ e ψ corresponde a função digama, que se caracteriza por:

$$\psi(z) = \frac{\partial}{\partial z} \log \Gamma(z) = \frac{\Gamma'(z)}{\Gamma(z)}.$$

Assim como ocorre com a distribuição t-assimétrica padrão, os estimadores para o caso mais geral também devem ser obtidos através de métodos numéricos. Observe que os resultados apresentados acima coincidem com os obtidos por Azzalini & Capitanio (2002) no contexto multivariado para $d = 1$.

O programa R, disponível em *http://www.r-project.org*, possui a função st.mle no pacote SN, que utiliza, entre outros métodos, o quasi-Newton para a obtenção das estimativas de máxima verossimilhança. No entanto, alguns problemas com o uso deste método dizem respeito a convergência. Se os valores iniciais estiverem muito distantes do ponto de máximo, as estimativas podem não ser adequadas.

3.3 Procedimentos Gráficos

Em situações práticas, procuramos modelar um conjunto de dados através de uma distribuição de probabilidade que esteja de acordo com as características desta amostra. A partir desse modelo probabilístico, são discutidos os aspectos inferenciais (clássicos ou bayesianos) utilizados para a estimação dos parâmetros desta distribuição.

Uma vez obtidas estimativas para os parâmetros desta distribuição, nos cabe verificar a adequabilidade e a acurácia destes valores a distribuição proposta para modelar a amostra. Modelos de diagnósticos são fundamentais em análise estatística pois nos permitem verificar se é razoável considerar que um determinado conjunto de dados pode ser modelado através de uma específica distribuição de probabilidades.

Nesse sentido, procedimentos gráficos constituem-se em poderosos instrumentos para identificar se há alguma observação amostral discrepante ou má modelagem dos dados. Vários modelos gráficos são propostos na literatura com a finalidade de comparar distribuições empíricas e teóricas. Aqui discutiremos a construção dos gráficos de quantis, de probabilidades acumuladas, Healy-quantis e Healy-probabilidades acumuladas, bem como os prós e contras em se adotar cada um destes métodos.

Para isso, considere x_1, x_2, ...,x_n uma amostra aleatória e suponha que queremos verificar se estes dados seguem ou não uma distribuição teórica Y, com função de distribuição acumulada F_Y.

3.3.1 Gráficos de quantis

O gráfico de quantis é uma ferramenta visual muito conhecida por verificar se uma amostra está sendo adequadamente modelada ou não por uma particular distribuição conhecida. O uso mais comum desta metodologia consiste em

examinar se uma amostra provém de uma distribuição normal.

No entanto, a aplicabilidade deste método não se restringe apenas a distribuição normal. Os gráficos de quantis se mostram muito úteis para comparar distribuições empíricas e teóricas. Vejamos como construí-lo:

1. Ordenam-se os dados amostrais, obtendo $x_{(1)}$, $x_{(2)}$, ..., $x_{(n)}$.
2. Obtemos q_i, onde $\mathrm{P}(Y \leq q_i) = \frac{i-0.5}{n}$, para $i = 1, ..., n$.
3. O gráfico é obtido ao se desenhar, para $i = 1, ..., n$, os pares cartesianos formados por $(q_i, x_{(i)})$.

A inspeção deste tipo de gráfico pode nos fornecer um indicativo de valores discrepantes ou de diferentes padrões de assimetria ou achatamento da distribuição. É de se esperar que se a distribuição teórica e a empírica forem iguais, então os quantis de ambas as funções também deverão ser iguais e, portanto, o gráfico de quantis coincidirá com a reta bissetriz do primeiro quadrante. Toda a análise derivada do gráfico de quantis estará baseada nesta premissa.

Dessa forma, se o gráfico de quantis apresentar seus pares cartesianos praticamente sobre a reta bissetriz do primeiro quadrante, podemos dizer que as distribuições teórica e empírica coincidem.

Caso os pontos desenhados nos dê uma visão geral de uma reta não coincidente com a bissetriz do primeiro quadrante, podemos dizer que as distribuições teórica e a empírica estão relacionadas através de uma transformação linear. Para determinarmos os parâmetros desta transformação, basta considerarmos o coeficiente angular e linear de uma reta que melhor representa o gráfico de quantis.

Os gráficos de quantis descrevem bem regiões com baixas densidades, conseguindo ilustrar o comportamento nas caudas da distribuição. Entretanto, torna-se complicado comparar diferentes gráficos de quantis para diferentes distribuições teóricas, pois o conjunto de quantis q_i para cada distribuição pode estar em escalas diferentes.

3.3.2 Gráficos de probabilidades acumuladas

Um procedimento gráfico alternativo ao de quantis consiste em substituir os quantis observados e esperados pelas respectivas probabilidades acumuladas empíricas e teóricas.

1. Ordenam-se os dados amostrais, obtendo $x_{(1)}$, $x_{(2)}$, ..., $x_{(n)}$.
2. Calculamos as probabilidades acumuladas para a amostra ordenada, dadas por $F(x_{(1)})$, $F(x_{(2)})$, ..., $F(x_{(n)})$.
3. Obtemos $p_i = \frac{i-0.5}{n}$, para $i = 1, ..., n$, onde p_i corresponde a probabilidade acumulada empírica.
4. O gráfico é obtido ao se desenhar, para $i = 1, ..., n$, os pares cartesianos formados por $(F(x_{(i)}), p_i)$.

Os objetivos de se utilizar um gráfico de probabilidades acumuladas são bastante similares aos do gráfico de quantis, no entanto, a análise visual nos traz algumas informações diferentes daquelas apresentadas pelo gráfico de quantis.

Diferentemente dos gráficos de quantis, cuja escala gráfica depende da unidade do conjunto de dados analisado, os gráficos de probabilidades acumuladas apresentam sua escala limitada ao plano cartesiano de [0,1]x[0,1]. Por isso, torna-se mais fácil comparar uma amostra específica a diferentes distribuição teóricas. Se as distribuições teórica e empírica foram iguais, o gráfico de probabilidades acumuladas deverá nos mostrar uma linha reta na diagonal do quadrado [0,1]x[0,1].

No entanto, os gráficos de probabilidades acumuladas não são capazes de detectar transformações lineares que possam relacionar uma distribuição teórica a uma empírica. Além disso, a análise das caudas de uma distribuição através do gráfico de probabilidades não é tão sensível quanto o gráfico de quantis.

Através deste tipo de gráfico podemos determinar diferentes padrões de assimetria e/ou achatamento de uma distribuição. Distribuições assimétricas à

esquerda (direita) apresentam pontos extremos no lado superior (inferior) da diagonal e os pontos intermediários no lado inferior (superior).

Padrões específicos num gráfico de probabilidades acumuladas podem determinar se uma distribuição empírica tem, por exemplo, caudas mais pesadas do que a distribuição de referência. Distribuições com caudas leves (caudas pesadas) tendem a apresentar menores valores acumulados no lado superior (inferior) da diagonal, cruzando-a no centro e concentrando no lado inferior (superior) da diagonal os maiores valores. Casos acentuados dos padrões encontrados numa distribuição com caudas leves podem representar distribuições bimodais.

Na literatura, outras expressões para p_i, além daquela mencionada neste trabalho, são utilizadas para elaborar gráficos de quantis ou de probabilidades acumuladas. Alguns exemplos são $\frac{i}{n+1}$ e $\frac{i-0.375}{n+0.25}$. No entanto, para n suficientemente grande, estas expressões praticamente não se diferenciam.

Considere, nos dois métodos seguintes, x_1, x_2, ...,x_n uma amostra aleatória e suponha que estamos interessados em saber se é adequado modelar esta amostra através da distribuição $ST(\mu, \sigma^2, \nu, \lambda)$. Como notação, adotaremos $\bar{\mu}$, $\bar{\sigma}$, $\bar{\nu}$ e $\bar{\lambda}$ para indicar as estimativas dos parâmetros da distribuição t-assimétrica obtidas através do método de máxima verossimilhança ou momentos.

3.3.3 Gráficos Healy-quantis

Nas seções anteriores apresentamos dois métodos capazes de avaliar se uma amostra está sendo modelada adequadamente por uma distribuição univariada. Gráficos do tipo Healy também têm esta função, com a vantagem de poderem ser utilizados também em análises multivariadas.

Healy (1968) descreveu esta metodologia para o caso normal multivariado. Entretanto, esta proposta pode ser utilizada na análise de outras distribuições, por exemplo, as distribuições normal-assimétrica e t-assimétrica são utilizadas em Azzalini & Capitanio (2003).

A construção dos gráficos do tipo Healy dependem necessariamente da propriedade (P6), que nos diz que a forma quadrática de uma variável aleatória com distribuição t-assimétrica é F-Snedocor. Outra peculiaridade deste gráfico diz respeito ao uso da distância de Malahanobis que corresponde a distância quadrática de cada observação em relação a sua média amostral.

Vejamos como funciona este processo.

1. Calculamos as distâncias de Malahanobis, dadas por $D_i = \left(\frac{x_i - \bar{\mu}}{\bar{\sigma}}\right)^2$.
2. Ordenamos D_1, ..., D_n de forma crescente, obtendo $D_{(1)} \leq D_{(2)} \leq ... \leq D_{(n)}$.
3. Obtemos q_i, onde $\mathrm{P}(Y \leq q_i) = \frac{i}{n+1}$, para $i = 1, ..., n$ e $Y \sim$ F-Snedocor$(1, \bar{\nu})$.
4. O gráfico é obtido ao se desenhar, para $i = 1, ..., n$, os pares cartesianos formados por $(q_i, D_{(i)})$.

Observe que q_i corresponde ao quantil de ordem i e, portanto, vamos nos referir a este gráfico como Healy-quantis.

Da mesma forma que no gráfico de quantis, comparamos a distribuição teórica com a distribuição empírica, fornecida pelos dados ordenados, analisando se os n pares cartesianos estão em torno da reta bissetriz do primeiro quadrante. Outra semelhança diz respeito a escala gráfica, que em ambas metodologias dependem da unidade de medida do conjunto de dados analisado.

No entanto, os gráficos Healy-quantis, assim como os gráficos de probabilidades acumuladas, não são capazes de detectar transformações lineares que possam relacionar uma distribuição teórica a uma empírica.

3.3.4 Gráficos Healy-probabilidades acumuladas

Uma variante do método apresentado anteriormente consiste em utilizar as probabilidades acumuladas em substituição aos quantis. Denominaremos tal procedimento de Healy-probabilidades acumuladas.

A análise deste tipo de gráfico é semelhante àquelas feitas na análise de um gráfico de probabilidades acumuladas, lembrando apenas que agora não é possível detectar transformações lineares que possam relacionar uma distribuição teórica a uma empírica a partir de um gráfico Healy-probabilidades acumuladas. Dessa forma, nos limitaremos apenas em descrever o método.

1. Calculamos as distâncias de Malahanobis, dadas por $D_i = \left(\frac{x_i - \bar{\mu}}{\bar{\sigma}}\right)^2$.
2. Ordenamos D_1, ..., D_n de forma crescente, obtendo $D_{(1)} \leq D_{(2)} \leq ... \leq D_{(n)}$.
3. Calculamos as probabilidades acumuladas da amostra ordenada, através de $F_Y(D_{(i)})$, para i = 1, ..., n e $Y \sim$ F-Snedocor(1, $\bar{\nu}$).
4. Obtemos $p_i = \frac{i}{n+1}$, para $i = 1, ..., n$, onde p_i corresponde a probabilidade acumulada empírica.
5. O gráfico é obtido ao se desenhar, para $i = 1, ..., n$, os pares cartesianos formados por $(F_Y(D_{(i)}), p_i)$.

3.4 Estudos de Simulação

Inicialmente, vamos apresentar e comparar o método dos momentos (M.M.) com o método de máxima verossimilhança (M.V.) para a obtenção das estimativas dos parâmetros da distribuição t-assimétrica padrão. Posteriormente, iremos avaliar o desempenho do método de máxima verossimilhança no caso em que os quatro parâmetros da distribuição t-assimétrica são considerados desconhecidos.

Os resultados destes estudos de simulação encontram-se organizados nas tabelas do Apêndice D.

Consideramos $\mathbf{X} = (X_1, X_2, ..., X_n)$ uma amostra aleatória de tamanho n da distribuição t-assimétrica padrão. Para amostras com o n fixado em 20, 50, 80, 100 e 200, λ em 0, 1, 5, 20 e ν em 3, 8, 12, 100, obtivemos 500 replicações. Descartamos as amostras em que todas as observações fossem exclusivamente positivas ou negativas, obtendo apenas estimadores de máxima verossimilhança finitos.

3.4.1 Comparando M.M. com M.V.

Nesta seção, vamos apresentar um estudo de simulação que procura identificar em que situações o método de máxima verossimilhança se mostra mais eficaz que o método dos momentos na obtenção das estimativas dos parâmetros da distribuição t-assimétrica padrão.

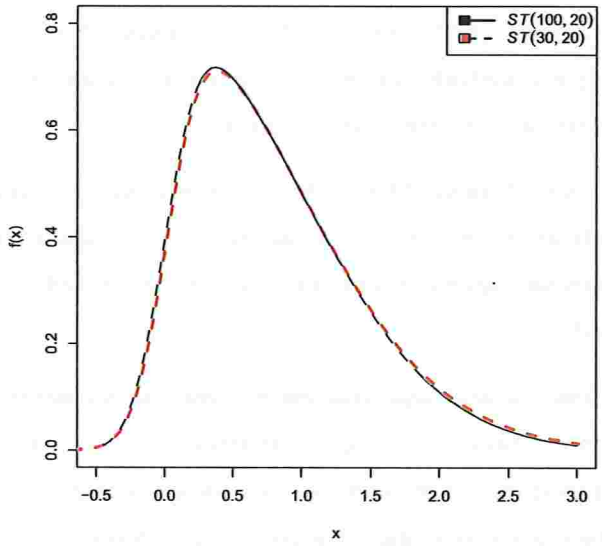
Nas Tabelas D.1 e D.2 apresentamos a mediana e o desvio padrão das estimativas do parâmetro de assimetria (λ) para ambos os métodos. Resultados similares são apresentados nas Tabelas D.3 e D.4, salientando que o parâmetro analisado é ν.

Para a obtenção das estimativas de momentos utilizamos a teoria apresentada na Seção 3.1. As estimativas de máxima verossimilhança foram obtidas através da maximização da função de log-verossimilhança da distribuição t-assimétrica apresentada na página 37, quando $\mu = 0$ e $\sigma = 1$.

As Tabelas D.1 e D.2 nos indicam que quando λ é inferior a 1 ambos os métodos apresentaram boas estimativas do parâmetro de assimetria, independente do número de graus de liberdade e do tamanho da amostra considerados. Vemos que, as estimativas de momentos para λ são péssimas quando a amostra possui grande assimetria. Em geral, observamos que as estimativas de máxima verossimilhança estão mais próximas do verdadeiro valor do parâmetro do que as estimativas de momentos, para quaisquer casos analisados.

Ao analisarmos as Tabelas D.3 e D.4, vemos que, para $\lambda < 5$ e $\nu < 8$, ambos os métodos apresentaram estimativas razoáveis para o número de graus de liberdade, independente do tamanho da amostra. Tal fato também se verifica quando o tamanho da amostra é superior a 100 e $\nu < 100$. Ambos os métodos não se comportam bem quando o verdadeiro valor de ν se enquadra no caso limite ($\nu = 100$), ainda que o tamanho da amostra seja grande. Vale ressaltar que, quando $n = 200$ e $\nu = 100$, as estimativas de máxima verossimilhança para ν variam numa faixa de 48 a 69, longe do verdadeiro valor do parâmetro. No entanto, valores de ν superiores a 30 geram funções densidade de probabilidade bastante similares, como podemos verificar na Figura 3.2.

Figura 3.2: Funções de Densidade



Em geral, as estimativas de máxima verossimilhança para ν têm um melhor desempenho do que as estimativas obtidas via método dos momentos. Além disso, por conta da condição de existência do estimador de momentos (ver Seção 3.1), vemos que o método de momentos tem um uso mais restrito do que o método de máxima verossimilhança, aplicando-se a poucas amostras.

3.4.2 Método de M.V - distribuição t-assimétrica posição e escala

As estimativas de máxima verossimilhança para os parâmetros da distribuição t-assimétrica posição e escala foram obtidas através da função *st.mle* existente no pacote SN do programa R. Lembramos que na Seção 3.4.1 não utilizamos a função *st.mle*, pois esta rotina não permite que se considere os parâmetros de posição e escala fixos e, por isso, apresenta como resultado as estimativas para os quatro parâmetros da distribuição t-assimétrica.

Analisemos, inicialmente, as Tabelas D.5 e D.6 que apresentam as médias e as medianas das estimativas de máxima verossimilhança para o parâmetro de assimetria e, em parênteses, os desvios padrões associados.

Nos casos em que $\lambda < 5$, tanto a média quando a mediana das estimativas de máxima verossimilhança para λ apresentaram resultados próximos ao verdadeiro valor do parâmetro, independente do número de graus de liberdade e do tamanho da amostra considerado. Tanto a média quanto a mediana das estimativas de λ apresentaram estimativas ruins, nos casos em que a amostra possui grande assimetria e tamanho pequeno. Em geral, a mediana das estimativas de λ tiveram um desempenho melhor do que as obtidas considerando a média.

Nas Tabelas D.7 e D.8 apresentamos a média, a mediana e, em parênteses, o desvio padrão das estimativas de máxima verossimilhança para ν. Nota-se que, tanto a média quanto a mediana das estimativas de ν super estimam o verdadeiro valor do parâmetro e possuem alta variabilidade. Em geral, vimos que a média das estimativas de ν estão muito distantes do verdadeiro valor do parâmetro, para quaisquer n e λ considerados. Para $\nu \leq 3$ e $n > 20$, a mediana das estimativas de ν apresentou resultados próximos do verdadeiro valor do parâmetro. Amostras pequenas e/ou com grandes valores para o número de graus de liberdade apresentaram péssimas estimativas para ν, tanto usando a média quanto a mediana.

Amostras com tamanho superior a 200 e $\nu < 100$ apresentaram estimativas razoáveis para o número de graus de liberdade quando consideramos a mediana das estimativas de ν, porém com variabilidade alta. Para $\nu = 100$, a mediana das estimativas para o número de graus de liberdade parece distante do verdadeiro valor do parâmetro. Porém, como observado no estudo de simulação anterior, valores de ν superiores a 30 apresentam funções densidade de probabilidade similares, enquadrando-se no caso limite da distribuição t-assimétrica.

Capítulo 4

Inferência Bayesiana

Inferência estatística é o processo pelo qual procuramos inferir propriedades de uma população, baseados em resultados obtidos de uma amostra dessa população. Tal raciocínio é intrinsecamente indutivo, pois procura fazer generalizações a partir de casos particulares.

Fazer inferência do ponto de vista clássico corresponde a estimar parâmetros de interesse, baseando-se exclusivamente na informação recolhida da amostra e no modelo probabilístico proposto.

Do ponto de vista bayesiano, considerar a informação oferecida por uma amostra para fazer inferência sobre os parâmetros de uma determinada população é uma condição necessária, entretanto, deve-se levar em conta também a natureza do parâmetro e possíveis informações *a priori* sobre ele. A questão que se coloca é como se deve incorporar esta informação. Aceitando-se a interpretação subjetivista de probabilidade e reconhecendo o parâmetro θ como desconhecido, a incerteza sobre esta quantidade pode ser quantificada em termos de probabilidade.

Nesse sentido, *"... os bayesianos defendem que a informação inicial ou a priori - anterior ou externa em relação à experiência mas demasiado importante para ser ignorada ou tratada ad hoc - pode traduzir-se formalmente por uma distribuição de probabilidade, geralmente subjectiva, para θ, seja $h(\theta)$, designada distribuição a priori..."* (Paulino, Turkman & Murteira, 2003).

Uma das principais ferramentas da teoria bayesiana é o Teorema de Bayes, que relaciona a informação *a priori* a respeito do parâmetro (expressa através de uma função de probabilidade) com aquela fornecida pela amostra (expressa através da função de verossimilhança). Assim, a distribuição *a posteriori* para θ é dada por:

$$h(\theta|y) = \frac{L(\theta;\ \mathrm{y})\, h(\theta)}{\int_\Theta L(\theta;\ \mathrm{y})\, h(\theta)\, d\theta}, \qquad \theta \in \Theta, \tag{4.1}$$

onde $h(\theta)$ é a distribuição *a priori* para θ, $L(\theta;\ \mathrm{y})$ corresponde a função de verossimilhança de θ, relacionada a uma amostra aleatória $\mathbf{Y} = (Y_1,\ Y_2,\ ...,\ Y_n)$ de tamanho n de $Y|\theta$, e $h(\theta|\mathrm{y})$ representa a distribuição *a posteriori*.

4.1 Escolha da Distribuição *a Priori*

Uma etapa fundamental nos procedimentos bayesianos diz respeito a escolha da distribuição que representará a informação *a priori* para os parâmetros. Diante de uma informação *a priori* mais ou menos substancial do fenômeno estudado, a abordagem bayesiana permite traduzir este conhecimento e incorporá-la a uma distribuição de probabilidade. Dessa forma, cada particular problema estabelece uma distribuição *a priori* específica. Distribuições *a priori* que possuem tais características são conhecidas como distribuições *a priori* subjetivas.

Quando não existe uma informação *a priori* palpável pode-se recorrer ao uso de distribuições *a priori* não informativas. Uma primeira técnica utilizada para gerar distribuições não informativas estabelece que *na ausência de razão suficiente para privilegiar umas possibilidades em detrimento de outras, decorrente da escassez informativa a priori, deve-se adoptar a equiprobabilidade* (Paulino, Turkman & Murteira, 2003). No entanto, tal método muitas vezes produz distribuições impróprias, ou seja, distribuições que não satisfazem ao axioma de probabilidade total e que podem resultar em distribuições *a posteriori* com o mesmo problema. Por exemplo, no caso do espaço paramétrico não ser limitado, a hipótese de equiprobabilidade induz a uma função constante cuja integral é infinita (distribuição *a priori* imprópria).

Vamos discutir, inicialmente, a determinação de distribuições *a priori* para os parâmetros da distribuição t-assimétrica padrão.

Com base no princípio de equiprobabilidade e considerando a parametrização $\delta = \frac{\lambda}{\sqrt{1+\lambda^2}}$, a qual é limitada em [-1,1], podemos assumir que a escolha mais intuitiva para a determinação da distribuição *a priori* para δ seria considerar $\delta \sim U(-1, 1)$, onde U refere-se a distribuição uniforme contínua. Esta distribuição *a priori* para δ induz no espaço paramétrico de λ a seguinte distribuição *a priori* $\lambda \sim t\left(0, \frac{1}{2}, 2\right)$.

Outra construção consiste em utilizar a distribuição *a priori* de Jeffreys, que é invariante por transformações injetivas. A distribuição *a priori* de Jeffreys para o modelo t-assimétrico não é simples de ser obtida. Portanto, usaremos a mesma distribuição *a priori* de Jeffreys adotada para λ quando a distribuição considerada é a normal-assimétrica, que é dada por:

$$\begin{aligned} f^J(\lambda) &\propto \sqrt{I(\lambda)} \\ &\propto \sqrt{\int_{-\infty}^{\infty} 2x^2\phi(x)\frac{\phi^2(\lambda x)}{\Phi(\lambda x)}dx}. \end{aligned} \tag{4.2}$$

Liseo e Loperfido (2006) mostraram que esta distribuição *a priori* é própria. No entanto, sua expressão não é fechada, sendo obtida apenas numericamente.

Rodríguez (2005) propõe utilizar a distribuição $t\left(0, \frac{\pi^2}{4}, \frac{1}{2}\right)$ como aproximação para a distribuição *a priori* (4.2). Em seu trabalho, fica demonstrado que a cauda desta aproximação tem a mesma ordem da distribuição *a priori* (4.2), que é $O(\lambda^{-3/2})$.

A Figura 4.1 apresenta as curvas representativas das distribuições *a priori* considerando as metodologias de Jeffreys e de equiprobabilidade.

Ao comparar as curvas das especificações *a priori* para λ na Figura 4.1, notamos que a distribuição t induzida pela distribuição uniforme é mais informativa do que a distribuição *a priori* da aproximada de Jeffreys.

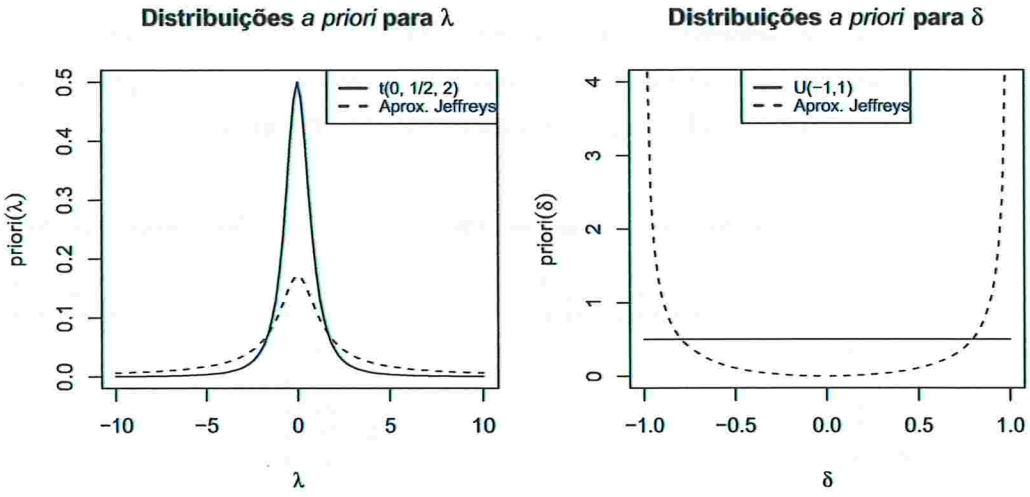


Figura 4.1: Especificações *a priori* para λ e δ

É interessante notar que as distribuições *a priori* anteriores são próprias, ou seja, são genuínas funções densidade de probabilidade. Este fato é importante pois garante a existência de distribuições *a posteriori* próprias.

Uma possível distribuição *a priori* própria para o número de graus de liberdade baseada parcialmente no princípio de equiprobabilidade consiste em adotar $\nu \sim U(1, 30)$. Observe que com esta distribuição *a priori*, estamos considerando que a nossa análise contemplará do modelo Cauchy-assimétrico ($\nu = 1$) ao modelo normal-assimétrico ($\nu = 30$), dado que, quando $\nu > 30$, temos que a função densidade de probabilidade da distribuição t-assimétrica pouco se altera, podendo ser aproximada pela função densidade de uma distribuição normal-assimétrica. Entretanto, tal distribuição *a priori* atribui probabilidade zero para os casos em que $0 < \nu < 1$ ou $\nu > 30$, o que pode ser restritivo.

Uma proposta alternativa de distribuição *a priori* para os graus de liberdade considera $\nu \sim Exp(0.1)\mathrm{I}(1,)$, ou seja, ν tem distribuição exponencial truncada com parâmetro 0.1. Utilizar esta especificação subjetiva indica que acreditamos ser mais provável adotar modelos com caudas mais grossas, uma vez que para a distribuição exponencial, quanto menor o grau de liberdade, maior é a chance de sua ocorrência.

Fonseca (2004) obteve a distribuição *a priori* não informativa de Jeffreys para ν, quando o modelo considerado é t-Student. Baseados nesta proposta, analisamos a forma desta densidade e observamos que uma possível aproximação para ela é obtida quando consideramos $\nu \sim Exp(0.5)\mathrm{I}(1,)$.

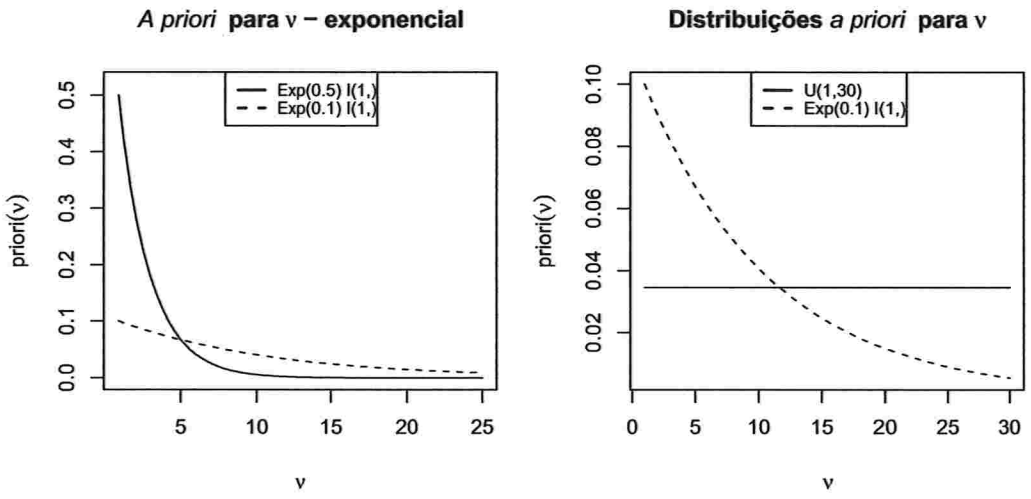


Figura 4.2: Especificações *a priori* para ν

Na Figura 4.2, podemos comparar as curvas das densidades *a priori* escolhidas para ν. Ao comparar as especificações baseadas na distribuição exponencial truncada, vemos que a distribuição *a priori* aproximada de Jeffreys atribui maior probabilidade de obtermos estimativas de graus de liberdade pequenos do que a distribuição *a priori* exponencial truncada de parâmetro 0.1. Análise similar a anterior pode ser descrita, quando comparamos as distribuições *a priori* exponencial truncada com parâmetro 0.1 e a distribuição uniforme.

No caso da distribuição t-assimétrica geral, precisamos especificar as distribuições *a priori* para os parâmetros de posição e escala.

Recorrendo novamente a Proposição 2.2, vemos que os parâmetros de posição e escala estão relacionados a distribuição normal-assimétrica de três parâmetros. Neste caso, Liseo & Loperfido (2006) utilizando o método de Berger &

Bernardo (1992) sugerem a seguinte distribuição *a priori* para os parâmetros da distribuição normal-assimétrica:

$$f(\mu, \sigma, \lambda) \propto \frac{1}{\sigma}\, g(\lambda)$$

onde $g(\lambda)$ pode ser substituída pela distribuição *a priori* de Jeffreys. Vale lembrar que a distribuição *a priori* de Jeffreys para os parâmetros de posição e de escala é dada por $\frac{1}{\sigma}$.

Utilizando os critérios discutidos anteriormente, podemos considerar a seguinte especificação para os parâmetros de uma distribuição t-assimétrica:

$$\begin{aligned} p(\mu, \sigma, \nu, \lambda) &= p(\mu, \sigma)\, p(\nu)\, p(\lambda) \\ &\propto \frac{1}{\sigma}\frac{1}{29}\left(1 + \frac{\lambda^2}{k{\sigma_t}^2}\right)^{-\frac{k+1}{2}} \\ &\propto \frac{1}{\sigma}\left(1 + \frac{\lambda^2}{k{\sigma_t}^2}\right)^{-\frac{k+1}{2}} . \end{aligned} \tag{4.3}$$

Quando consideramos $k = \frac{1}{2}$ e ${\sigma_t}^2 = \frac{\pi^2}{4}$ obtemos a aproximação da distribuição *a priori* de Jeffreys para λ e quando $k = 2$ e ${\sigma_t}^2 = \frac{1}{2}$ obtemos a distribuição *a priori* para λ induzida pela uniforme.

4.2 Modelos Hierárquicos

Nesta seção apresentamos diferentes maneiras de representar o modelo t-assimétrico de forma hierárquica, a fim de que possamos facilitar a implementação computacional das técnicas bayesianas. Inicialmente, vamos discutir o caso em que $\mathbf{Z} = (Z_1, Z_2, ..., Z_n)$ é uma amostra aleatória simples de tamanho n da variável aleatória Z com distribuição $ST(\nu, \lambda)$.

Podemos obter o modelo hierárquico através da Proposição 2.4 que apresenta a distribuição t-assimétrica padrão como uma combinação linear, explicitada aqui da seguinte forma:

$$Z_i = \frac{\lambda}{\sqrt{1+\lambda^2}}\, U_i + \frac{1}{\sqrt{1+\lambda^2}}\, X_i, \tag{4.4}$$

para $U_i \sim HT(\nu)$ e $X_i \sim t_v$ variáveis aleatórias dependentes. Então, o condicionamento em U_i resulta em

$$\begin{aligned} Z_i|U_i &\sim t\left(\frac{\lambda}{\sqrt{1+\lambda^2}}\, U_i,\ \frac{1}{1+\lambda^2},\ \nu\right), \\ U_i &\sim HT(\nu). \end{aligned} \tag{4.5}$$

Uma proposta mais simples que a anterior utiliza a estrutura descrita em (4.4) considerando a parametrização δ. Dessa maneira,

$$Z_i = \delta\, U_i + \sqrt{1-\delta^2}\, X_i, \tag{4.6}$$

onde $U_i \sim HT(\nu)$ e $X_i \sim t_v$ são variáveis aleatórias dependentes. Então,

$$\begin{aligned} Z_i|U_i &\sim t\left(\delta\, U_i, 1-\delta^2, \nu\right), \\ U_i &\sim HT(\nu). \end{aligned} \tag{4.7}$$

Apresentamos agora uma forma hierárquica alternativa construída a partir da caracterização apresentada na Proposição 2.2. Dessa forma,

$$\begin{aligned} Z_i|U_i &\sim SN\left(0,\ \frac{1}{U_i},\ \lambda\right), \\ U_i &\sim Gama\left(\frac{\nu}{2}, \frac{\nu}{2}\right). \end{aligned} \tag{4.8}$$

Observe que $Z_i|U_i = Y_i$ tem distribuição normal-assimétrica e, portanto, podemos representá-la segundo a seguinte combinação linear:

$$Y_i = \sqrt{\frac{1}{U_i}}\left(\frac{\lambda}{\sqrt{1+\lambda^2}}\, V_i + \frac{1}{\sqrt{1+\lambda^2}}\, T_i\right), \tag{4.9}$$

onde $V_i \sim HN(0,\ 1)$ e $T_i \sim N(0,1)$ são variáveis aleatórias independentes. Dessa forma, condicionando $Z_i|U_i = Y_i$ em V_i, obtemos o seguinte modelo hierárquico alternativo a (4.5):

$$\begin{aligned} Z_i|U_i, V_i &\sim N\left(\sqrt{\frac{1}{U_i}}\,\frac{\lambda}{\sqrt{1+\lambda^2}}\,V_i,\ \frac{1}{U_i}\,\frac{1}{1+\lambda^2}\right), \\ V_i &\sim HN(0,1), \\ U_i &\sim Gama\left(\frac{\nu}{2}, \frac{\nu}{2}\right). \end{aligned} \tag{4.10}$$

Observe que se utilizarmos a parametrização $\delta = \frac{\lambda}{\sqrt{1+\lambda^2}}$ teremos então $Z_i|U_i, V_i \sim N\left(\sqrt{\frac{1}{U_i}}\,\delta\, V_i,\ \frac{1}{U_i}\,(1-\delta^2)\right)$ em (4.10).

Utilizando as distribuições *a priori* discutidas na seção anterior e escrevendo a distribuição t-Student em forma hierárquica, temos:

$$\begin{aligned} \lambda|w &\sim N\left(0, \frac{{\sigma_t}^2}{w}\right), \\ w &\sim Gama\left(\frac{k}{2}, \frac{k}{2}\right), \\ \nu &\sim U(1,30) \text{ ou } \nu \sim Exp(0.1)\ I(1,) \text{ ou } \nu \sim Exp(0.5)\ I(1,). \end{aligned} \tag{4.11}$$

Lembramos que quando consideramos $k = \frac{1}{2}$ e ${\sigma_t}^2 = \frac{\pi^2}{4}$ obtemos a aproximação da distribuição *a priori* de Jeffreys e para $k = 2$ e ${\sigma_t}^2 = \frac{1}{2}$ obtemos a distribuição *a priori* induzida pela distribuição uniforme.

Todas as representações hierárquicas e respectivas especificações *a priori* são facilmente implementadas no *software* 'WinBUGS'. Este *software* utiliza o Algoritmo de Gibbs, e variações deste, para obter amostras da distribuição *a posteriori* e a partir delas obter medidas de interesse das distribuições *a posteriori* marginais.

Vejamos agora o caso em que $\mathbf{Y} = (Y_1, Y_2, ..., Y_n)$ é uma amostra aleatória simples da variável aleatória Y com distribuição $ST(\mu, \sigma^2, \nu, \lambda)$.

Os modelos hierárquicos propostos em (4.5), (4.7) e (4.10) para a distribuição t-assimétrica padrão podem ser facilmente generalizados para a distribuição t-assimétrica posição e escala se considerarmos $Z_i = \frac{Y_i - \mu}{\sigma}$. Além disso, podemos utilizar as distribuições *a priori* para λ e ν descritas anteriormente.

Independente do modelo hierárquico escolhido e das especificações *a priori* adotadas para λ e ν, escolhemos $\mu \sim N\left(0, \frac{10^5}{6}\right)$ e $\sigma^2 \sim Gama\left(\frac{1}{100}, \frac{1}{100}\right)$ como possíveis especificações *a priori* para os parâmetros de posição e de escala, respectivamente. Observe que a variabilidade destas duas distribuições é muito grande, implicando em distribuições *a priori* próximas das não informativas, porém próprias (distribuições *a priori* vagas). Tal consideração se faz relevante pois distribuições *a priori* impróprias não podem ser consideradas no 'WinBUGS'.

Capítulo 5

Aplicação

5.1 Notas de MAE-0116

Neste seção vamos considerar as notas da primeira prova de 87 alunos alunos de Ciências Biológicas (CB) e Relações Públicas (RP) que cursaram no período da manhã a disciplina Noções de Estatística (MAE-0116), 1° semestre de 2006. Procuraremos identificar a distribuição que melhor modela estes dados e, para isso, faremos uso da teoria apresentada nos Capítulos 3 e 4.

Na Figura 5.1, apresentamos os *box-plots* das notas de cada turma - RP e CB - separadamente e verificamos que o comportamento de ambas as amostras são similares, o que nos permite analisar estas duas amostras conjuntamente. Na Tabela 5.1 apresentamos as medidas resumo das notas da primeira prova (P1).

Tabela 5.1: Medidas-resumo das Notas da Primeira Prova

	Min.	1° Quartil	Mediana	Média	3° Quartil	Máx.	var
RP	0.000	4.650	7.100	6.618	8.400	9.700	5.564
CB	0.000	6.700	7.800	7.063	8.500	10.000	5.282
Total	0.000	6.300	7.700	6.951	8.500	10.000	5.327

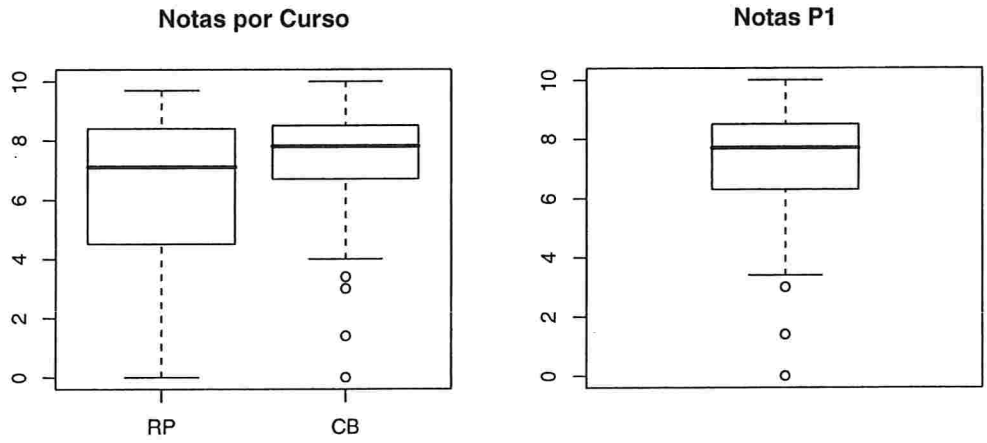


Figura 5.1: Boxplots das notas da primeira prova (P1) por curso e em conjunto

Na Tabela 5.2, apresentamos as estimativas de máxima verossimilhança dos parâmetros e os respectivos erros padrões quando consideramos as distribuições normal e t-Student. A Tabela 5.3 apresenta os resultados quando consideramos as distribuições normal e t assimetrizadas.

Tabela 5.2: Estimativas dos parâmetros das distribuições normal e t-Student

Método de Máxima Verossimilhança				
	Normal		t-Student	
	Estimativa	Erro Padrão	Estimativa	Erro Padrão
$\hat{\mu}$	6.951	0.246	7.647	0.211
$\hat{\sigma}$	2.295	0.174	1.299	0.233
$\hat{\nu}$	-	-	2.219	0.786

Tanto as estimativas dos parâmetros da distribuição normal quanto as da distribuição t-Student foram obtidas no programa R através da função *fitdistr* da biblioteca *MASS*. No caso das distribuições normal-assimétrica e t-assimétrica utilizamos a biblioteca *sn* e as funções *sn.mle* e *st.mle*, respectivamente.

Tabela 5.3: Estimativas dos parâmetros da normal-assimétrica e t-assimétrica

Método de Máxima Verossimilhança				
	Normal-assimétrica		t-assimétrica	
	Estimativa	Erro Padrão	Estimativa	Erro Padrão
$\hat{\mu}$	9.524	0.223	9.238	0.255
$\hat{\sigma}$	3.448	0.173	2.308	0.548
$\hat{\lambda}$	-7.197	0.049	-4.194	1.270
$\hat{\nu}$	-	-	3.653	2.084

Vale ressaltar que, a menos das estimativas dos parâmetros da distribuição normal, as estimativas dos parâmetros das outras distribuições foram obtidas através da maximização da função de verossimilhança, com o emprego de métodos numéricos.

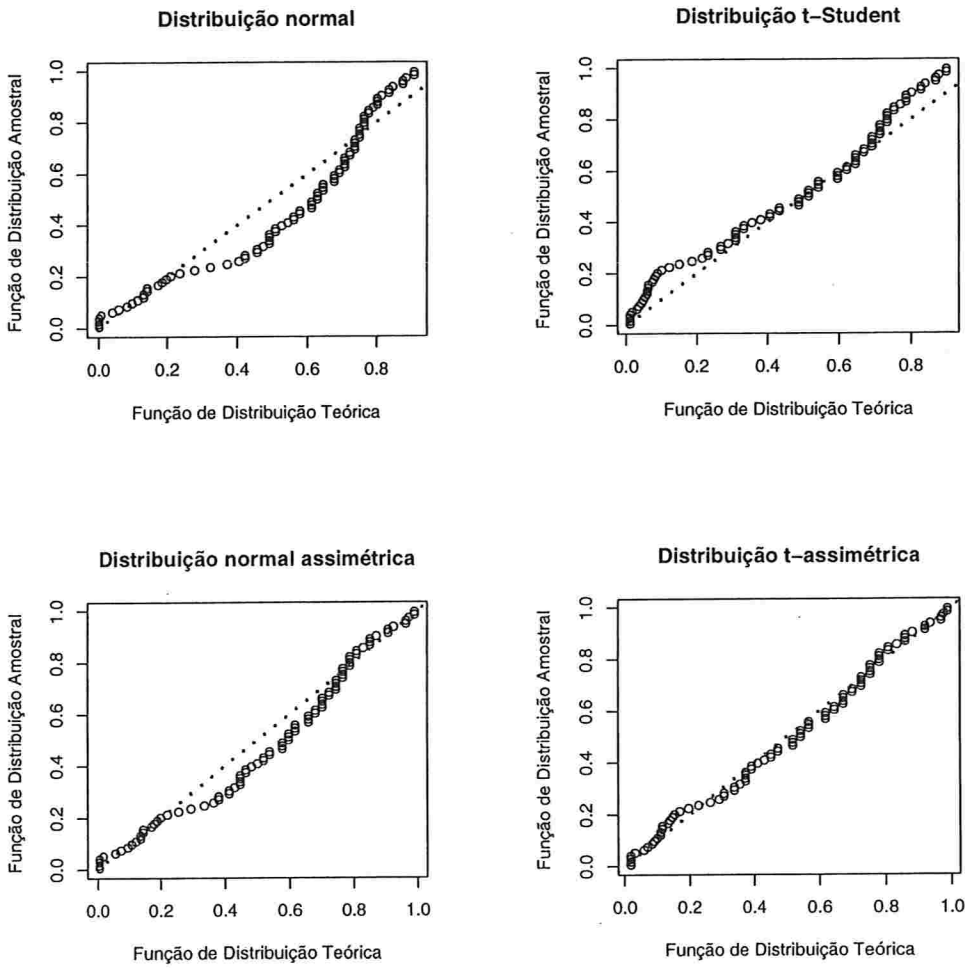
Comparando as Tabelas 5.2 e 5.3, vemos que os erros padrão das estimativas dos parâmetros da distribuição t-assimétrica tendem a ser superiores aos que foram apresentados por outras distribuições, principalmente no que diz respeito as estimativas dos parâmetros de assimetria e de graus de liberdade. Isto pode ser justificado pela complexidade do modelo t-assimétrico, quando comparado com as distribuições normal, t-Student e normal-assimétrica.

Tabela 5.4: Quartis e média das distribuições

Distribuição	1º Quartil	Mediana	Média	3º Quartil
Normal	5.403	6.951	6.951	8.499
t-Student	6.607	7.647	7.647	8.687
Normal-assimétrica	5.558	7.198	6.799	8.427
t-assimétrica	6.084	7.520	6.933	8.485

A partir dos resultados das Tabelas 5.2 e 5.3, obtivemos os quartis e a média para cada uma das distribuições consideradas, como pode ser visto na Tabela 5.4. Ao comparar a Tabela 5.4 com a Tabela 5.1, vemos que os quartis e a média amostral para o conjunto de notas estão muito próximos das medidas fornecidas pela distribuição t-assimétrica. Temos, portanto, um indício de que a

Figura 5.2: Gráficos de Probabilidades Acumuladas



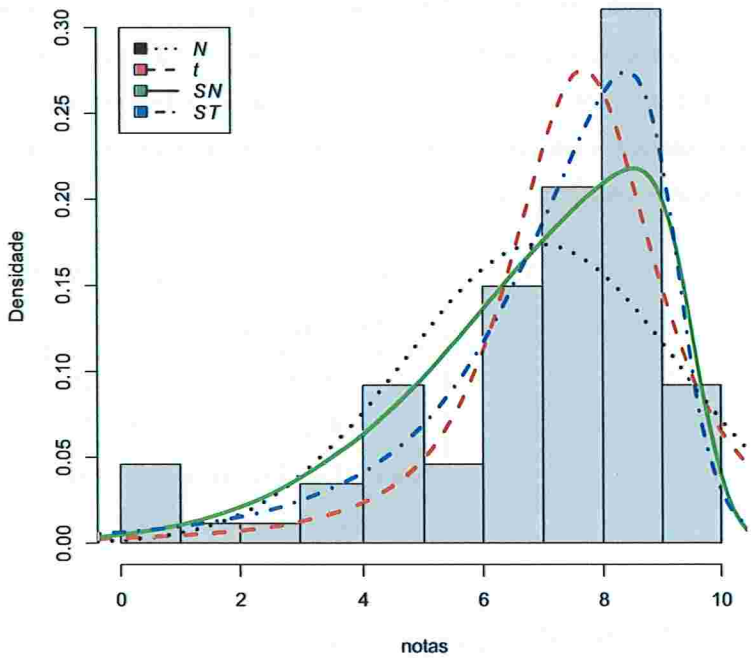
distribuição t-assimétrica pode ser a distribuição que melhor modela as notas da primeira prova. Vejamos se nossa intuição está correta, estudando os modelos de diagnósticos gráficos descritos na Seção 3.3.

Na Figura 5.2 são mostrados os gráficos de probabilidades acumuladas, considerando as diversas distribuições estudadas e as respectivas estimativas de seus parâmetros apresentadas nas Tabelas 5.2 e 5.3. No Apêndice E encontram-se os gráficos de quantis , Healy-quantis e Healy de probabilidades-acumuladas.

Comparando as distribuições através dos gráficos da Figura 5.2, vemos que a distribuição t-assimétrica é a distribuição que melhor ajusta os dados. Vale ressaltar que os outros métodos gráficos de diagnóstico também nos trazem essa informação, como pode ser observado nas figuras do Apêndice E. No entanto, o procedimento gráfico que melhor distingue as diferentes distribuições utilizadas para modelar o conjunto de notas é o de probabilidades acumuladas.

Tal afirmação pode ser confirmada quando analisamos simultaneamente os gráficos das Figuras 5.2 e 5.3 em que apresentamos o histograma das notas da primeira prova e as curvas de densidade associadas as estimativas obtidas anteriormente.

Figura 5.3: Histograma das Notas



No gráfico acima, vemos que a curva da distribuição t-assimétrica é a que melhor se adequa ao formato do histograma das notas.

Tabela 5.5: Probabilidade de se obter uma nota superior a 10

Distribuição (X)	P($X < 0$)	P($X > 10$)	Total
Normal	0.001	0.092	0.093
t-Student	0.011	0.099	0.110
Normal-assimétrica	0.006	0.009	0.015
t-assimétrica	0.019	0.011	0.030

Sabemos que as notas variam de 0 a 10. No entanto, todas as distribuições propostas para modelar os dados estão definidas em $\mathbb{R}$. Na Tabela 5.5, vamos obter a probabilidade de cada um dos modelos apresentar notas inferiores a 0 e superiores a 10. Notamos que as menores probabilidades ocorrem para as distribuições normal-assimétrica e t-assimétrica, indicando que o modelo a ser escolhido para modelar o conjunto de notas deve incorporar a assimetria dos dados.

Gupta & Chen (2001) propuseram o teste de Kolmogorov-Smirnov para verificar a aderência dos dados à distribuição normal-assimétrica. A seguir, apresentamos o resultado deste teste para as distribuições normal, t-Student, normal-assimétrica e t-assimétrica, baseados nas estimativas de máxima verossimilhança descritas nas Tabelas 5.2 e 5.3.

Tabela 5.6: Teste de Kolmogorov-Smirnov

Distribuição	P-valor
Normal	0.0136
t-Student	0.1717
Normal-assimétrica	0.1294
t-assimétrica	0.9234

Ao nível de significância de 5%, o teste de Kolmogorov-Smirnov aponta como aceitáveis a modelagem empregada através das distribuições t-Student, normal-assimétrica e t-assimétrica. No entanto, analisando exclusivamente os p-valores, vemos que o teste indica que a modelagem mais adequada é feita através da distribuição t-assimétrica.

As análises anteriores foram descritas apenas para o método de máxima verossimilhança pois, devido a condições de existência dos estimadores de momentos, não foi possível aplicar o método do momentos para a amostra em questão. Maiores detalhes sobre problemas de existência dos estimadores de momentos podem ser encontrados na Seção 3.1.

Do ponto de vista bayesiano, utilizamos o modelo hierárquico (4.5) para $Z_i = \frac{Y_i - \mu}{\sigma}$ e as distribuições *a priori* $\mu \sim N\left(0, \frac{10^5}{6}\right)$, $\sigma^2 \sim Gama\left(\frac{1}{100}, \frac{1}{100}\right)$ e $\lambda \sim t\left(0, \frac{1}{2}, 2\right)$. Para a distribuição *a priori* de ν utilizamos todas as especificações apresentadas no Capítulo 4.

As medidas *a posteriori* foram obtidas através do programa 'WinBUGS' que utiliza o algoritmo de Gibbs. Para a simulação geramos um *burn-in* de 5000 amostras e, devido a alta autocorrelação entre as cadeias, tomamos os resultados das iterações de 200 em 200 até obtermos uma amostra de Monte Carlo de tamanho 3000 aproximadamente independente. Resultados desta implementação são apresentados na Tabela 5.7.

Veja que para cada parâmetro de interesse o erro Monte Carlo é menor que 5% do seu respectivo desvio padrão amostral, indicando que a simulação após a convergência já atingiu um número suficiente de iterações necessário para obter amostras que podem ser usadas na inferência *a posteriori*.

Observe que, pelos resultados das Tabelas 5.3 e 5.5, as estimativas de máxima verossimilhança, a média e a mediana *a posteriori* para μ não diferem muito, independente da distribuição *a priori* para ν escolhida. No entanto, as estimativas dos outros parâmetros da distribuição t-assimétrica são sensíveis a escolha da especificação *a priori* ν, principalmente, a estimativa para o número de graus de liberdade.

Note que, quando utilizamos a distribuição aproximada de Jeffreys como distribuição *a priori* de ν, as estimativas pontuais do parâmetro de curtose se assemelham mais a estimativa de máxima verossimilhança de ν dada na Tabela 5.3.

Através da estatística bayesiana, conseguimos obter intervalos de credibilidade com probabilidade 0.95 para os parâmetros da distribuição t-assimétrica.

Tabela 5.7: Estimativas Bayesianas - Notas de MAE 0116

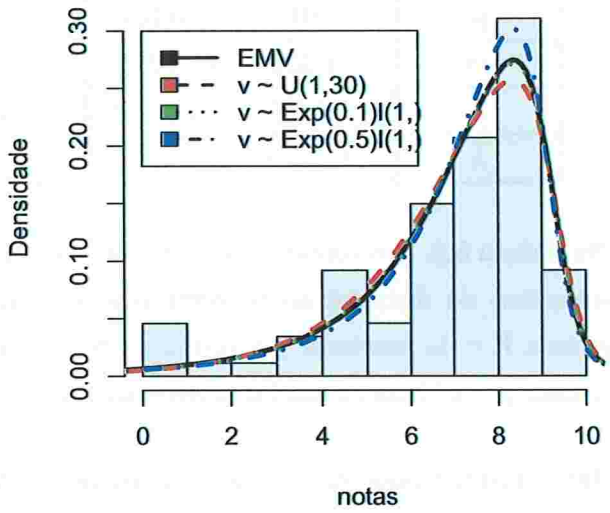
Priori ν	Parâ	Média	Desvio	2.5%	mediana	97.5%	Erro MCMC
$U(1, 30)$	μ	9.270	0.256	8.732	9.286	9.748	0.008
	σ	2.613	2.880	1.332	2.598	3.539	0.104
	ν	9.131	7.226	1.799	6.334	27.21	0.183
	λ	-4.783	1.947	-9.261	-4.543	-1.751	0.096
$Exp(0.1)$I(1,)	μ	9.224	0.266	8.682	9.236	9.708	0.009
	σ	2.431	2.660	1.249	2.377	3.398	0.096
	ν	6.241	5.751	1.690	4.383	22.830	0.135
	λ	-4.626	2.083	-9.769	-4.272	-1.623	0.094
$Exp(0.5)$I(1,)	μ	9.108	0.239	8.628	9.119	9.552	0.008
	σ	2.038	1.841	1.167	1.978	2.865	0.071
	ν	3.188	1.408	1.514	2.862	6.959	0.042
	λ	-3.836	1.530	-7.461	-3.579	-1.517	0.076

Observe que quando consideramos a distribuição *a priori* de ν uma distribuição exponencial truncada com parâmetro 0.1 ou uma distribuição uniforme, o intervalo de credibilidade para o parâmetro de graus de liberdade é pouco informativo (grande amplitude). O mesmo comportamento ocorre quando consideramos o intervalo de credibilidade para o parâmetro de assimetria. Vale ressaltar, ainda, que para quaisquer especificações *a priori* escolhidas para ν, os intervalos de credibilidade indicam que a estimativa de λ é negativa e, portanto, a distribuição modeladora dos dados é assimétrica negativa.

Na Figura 5.4, apresentamos o histograma das notas dos alunos, e as curvas de densidade da distribuição t-assimétrica quando consideramos as estimativas de máxima verossimilhança, a mediana *a posteriori* quando ν tem distribuição *a priori* $U(1, 30)$ e $Exp(0.1)$I(1,) e a média *a posteriori* para a distribuição aproximada de Jeffreys.

Observe que as curvas cujos parâmetros foram estimados através da mediana *a posteriori* se diferenciam principalmente no cume, mostrando a influência da estimativa do número de graus de liberdade. A curva associada a estimativa de máxima verossimilhança praticamente coincide com a curva em que a distribuição

Figura 5.4: Histograma das Notas e Curvas Associadas



a priori para ν é dada por $Exp(0.1)\mathrm{I}(1,)$. Note que a escolha da distribuição *a priori* de ν influi significativamente na modelagem das notas, sendo que a curva que melhor se adequa ao histograma, corresponde àquela gerada pelas estimativas da média *a posteriori*, quando ν tem uma distribuição *a priori* aproximada de Jeffreys.

5.2 Dados Simulados

Como discutido no Capítulo 3, é sabido que amostras que possuem todas as observações positivas ou todas negativas apresentam, com probabilidade positiva, o estimador para o parâmetro de assimetria infinito.

Neste capítulo, procuraremos explorar o aspecto inferencial clássico e bayesiano em situações como as descritas anteriormente. Para isso, simulamos uma amostra de tamanho 80 da distribuição $ST(0,1,\ 5,\ 20)$, onde todas as observações são positivas.

Tabela 5.8: Máxima Verossimilhança - Dados Simulados - $\nu = 5$ e $\lambda = 20$

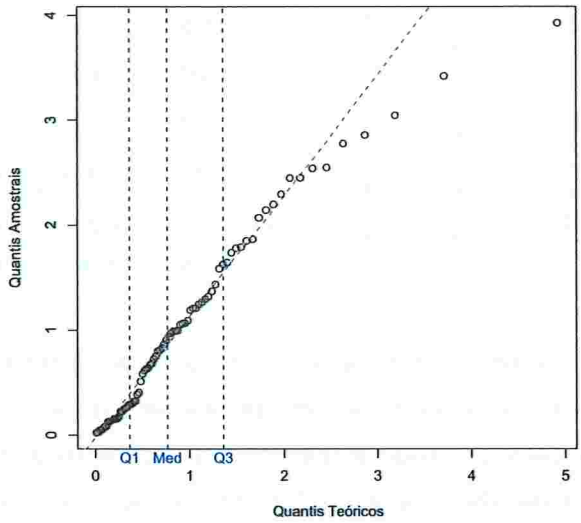
Parâmetro	MV - st.mle		MV - Padrão	
	Estimativas	Erro	Estimativas	Erro
$\hat{\mu}$	0.014	0.014	-	-
$\hat{\sigma^2}$	1.304	0.224	-	-
$\hat{\nu}$	20.207	122.776	4.028	1.491
$\hat{\lambda}$	122.776	131.572	77.365	81.948

Na Tabela 5.8, apresentamos as estimativas e os respectivos erros padrões dos parâmetros da distribuição t-assimétrica obtidas através da rotina *st.mle* do programa R e da maximização direta da função de verossimilhança quando consideramos μ e σ^2 fixos, o qual denotaremos *padrão.*

Inicialmente vamos analisar as estimativas obtidas através da rotina *st.mle* implementada no R. Observe que este processo nos fornece uma estimativa finita para o parâmetro de assimetria mas, na prática, vemos que a estimativa 122.78 é equivalente a considerar $\hat{\lambda}$ infinito. Quando consideramos os parâmetros μ e σ fixos e iguais a 0 e 1, respectivamente, vemos que a estimativa do número de graus de liberdade não está tão distante do verdadeiro valor do parâmetro, contrastando com a estimativa obtida pela rotina *st.mle* implementada no R. Como estamos interessados em estimar os parâmetros da distribuição t-assimétrica padrão, as análises subseqüentes do conjunto de dados simulados utilizarão apenas as estimativas de máxima verossimilhança obtidas via método *padrão.*

A Figura 5.5 apresenta o gráfico de quantis para a amostra simulada. Os símbolos Q_1, Med e Q_3 indicam, respectivamente, o primeiro quartil, a mediana (segundo quartil) e o terceiro quartil. Segundo o gráfico de quantis, vemos que parece razoável modelar a amostra simulada por uma distribuição t-assimétrica padrão, em que as estimativas dos parâmetros foram obtidas via maximização direta da função de verossimilhança.

Figura 5.5: Gráfico de Quantis - Dados Simulados



Vamos agora analisar a amostra simulada sob o ponto de vista bayesiano. Dessa forma, para obter as estimativas bayesia- nas, utilizamos o modelo hierárquico (4.5) e as distribuições *a priori* $\lambda \sim \mathrm{t}\left(0, \frac{1}{2}, 2\right)$ e $\nu \sim U(3,\ 30)$ ou $\nu \sim Exp(\theta)\mathrm{I}(1,)$, onde $\theta = 0.1,\ 0.5$.

As simulações foram implementadas no programa "Winbugs", com *burn-in* de 5000 amostras e, devido a alta correlação entre as cadeias, tomamos os resultados das iterações de 400 em 400 até obtermos uma amostra de Monte Carlo de tamanho 6000 aproximadamente independente.

Na Tabela 5.9, apresentamos as estatísticas obtidas através da análise bayesiana. Segundo esta tabela, a média e principalmente a mediana *a posteriori* para o número de graus de liberdade não diferem muito do verdadeiro valor do parâmetro, quando consideramos as distribuições *a priori* para $\nu \sim U(1,\ 30)$ e $\nu \sim Exp(0.1)\mathrm{I}(1,)$. Note que, no caso em que consideramos a aproximação da distribuição *a priori* de Jeffreys para ν, a média *a posteriori* gera uma estimativa mais razoável do parâmetro de curtose do que a mediana *a posteriori*.

Tabela 5.9: Estimativas bayesianas - Dados Simulados - $\nu = 5$, $\lambda = 20$

Priori ν	Parâ	Média	Desvio	2.5%	mediana	97.5%	Erro MCMC
$U(1, 30)$	λ	35.040	33.160	6.935	24.180	129.500	1.365
	ν	6.545	3.818	2.652	5.463	17.930	0.038
$Exp(0.1)$I(1,)	λ	36.23	35.45	6.925	24.68	143.8	1.685
	ν	5.475	2.745	2.457	4.837	12.28	0.035
$Exp(0.5)$I(1,)	λ	33.940	29.010	7.221	25.300	117.900	1.303
	ν	4.136	1.382	2.199	3.892	7.625	0.019

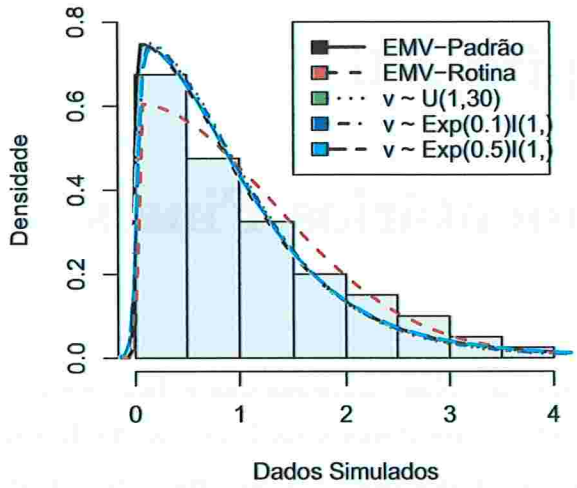
Quando consideramos o parâmetro de assimetria, vemos que tanto a média quanto a mediana *a posteriori* apresentam estimativas mais precisas do que aquelas obtidas via método de máxima verossimilhança, independente da especificação *a priori* escolhida para ν. No entanto, as estimativas bayesianas são bastante sensíveis a escolha da distribuição *a priori*.

Analisando os intervalos de credibilidade, vemos que a estimativa de λ é positiva, indicando que a modelagem dos dados simulados será descrita por uma distribuição assimétrica positiva. Veja que o intervalo de credibilidade para ν torna-se mais preciso a medida que utilizamos distribuições *a priori* para ν mais informativas. Vale ressaltar que o menor intervalo de credibilidade para ν e λ ocorre quando a especificação *a priori* para ν é dada pela aproximação da distribuição *a priori* de Jeffreys.

Lembramos que, não foi possível obter as estimativas de momentos para a amostra simulada, devido às condições de existência destes estimadores. Maiores detalhes na Seção 3.1.

Apresentamos na Figura 5.6 o histograma dos dados simulados e as curvas de densidade da distribuição t-assimétrica padrão quando consideramos as estimativas de máxima verossimilhança obtidas via rotina *st.mle* do programa R e maximização direta, as estimativas da mediana *a posteriori* quando ν tem distribuição à priori $U(1, 30)$ e $Exp(0.1)$I(1,) e a média *a posteriori* para a distribuição aproximada de Jeffreys.

Figura 5.6: Histograma dos Dados Simulados e Curvas Associadas



Na Figura 5.6, observe que as curvas geradas via metodologia bayesiana e maximização direta da função de log-verossimilhança praticamente coincidem, ainda que algumas das estimativas não pareçam tão próximas do verdadeiro valor do parâmetro.

Capítulo 6

Comentários Finais

Neste trabalho caracterizamos a distribuição t-assimétrica de diversas maneiras, obtendo resultados semelhantes às da distribuição t-Student. Apresentamos diversas propriedades, em especial, a função geradora de momentos que nos permitiu obter os momentos de ordem par e ímpar da distribuição t-assimétrica padrão. Muitos dos resultados apresentados nesta dissertação foram obtidos em contextos mais complexos, tais como a classe de distribuições elípticas assimétricas. Aqui procuramos estudar de maneira mais detalhada a distribuição t-assimétrica univariada, tanto do ponto de vista inferencial quanto de propriedades.

Em inferência clássica, obtivemos o estimador de momentos e as equações de verossimilhança para a distribuição t-assimétrica univariada. Através de um estudo de simulação comparamos as estimativas dos parâmetros da distribuição t-assimétrica padrão obtidas via método de máxima verossimilhança e método dos momentos. Um segundo estudo de simulação nos permitiu avaliar o desempenho das estimativas de máxima verossimilhança no caso em que os quatro parâmetros da distribuição t-assimétrica são considerados desconhecidos.

As estimativas de máxima verossimilhança para os parâmetros da distribuição t-assimétrica padrão foram obtidas através da maximização direta da função de verossimilhança e para a distribuição t-assimétrica posição e escala foram obtidas via função função *st.mle* do programa R. As estimativas de momentos foram obtidas através da metodologia apresentada na Seção 3.3.

Considerando a distribuição t-assimétrica padrão vimos que, tanto o método de máxima verossimilhança quanto o método de momentos apresentaram boas estimativas para λ quando a amostra possui pouca assimetria. Amostras com grandes assimetrias apresentaram péssimas estimativas de momento para λ. Ambos os métodos apresentaram estimativas razoáveis para ν, quando o valor verdadeiro dos parâmetros foram $\lambda < 5$ e $\nu \leq 8$. As estimativas de ν obtidas por ambos os métodos não se comportam bem quando o verdadeiro valor de ν se enquadra no caso limite ($\nu = 100$), ainda que o tamanho da amostra seja grande. Porém, para valores de ν superiores a 30, as funções densidade de probabilidade são bastante similares.

Para a distribuição t-assimétrica de posição e escala notamos que, quando a amostra possui pouca assimetria, tanto a média quando a mediana das estimativas de λ apresentaram bons resultados, independente do número de graus de liberdade e do tamanho da amostra considerado. De um modo geral, a mediana das estimativas de λ apresentou resultados mais próximos do verdadeiro valor do parâmetro do que a média das estimativas. Nota-se que, em geral, tanto a média quanto a mediana das estimativas de ν super estimam o verdadeiro valor do parâmetro, com alta variabilidade. Em geral, considerar a mediana das estimativas para ν é muito melhor do que considerar a sua média. Observamos que estimar ν em amostras pequenas com grande assimetria pode não ser adequado.

Do ponto de vista bayesiano, sugerimos algumas distribuições *a priori* para os parâmetros da distribuição t-assimétrica. Apresentamos diversas maneiras de representar a distribuição t-assimétrica de maneira hierárquica, como forma de facilitar a implementação computacional das técnicas bayesianas.

Através das aplicações do Capítulo 5, vimos que a distribuição t-assimétrica se mostra útil em modelar amostras com valores discrepantes e comportamento assimétrico, tendo, por isso, um desempenho superior às modelagens baseadas nas distribuições t-Student, normal e normal-assimétrica.

Nas aplicações, vimos que as estimativas bayesianas são bastante sensíveis a escolha das distribuições *a priori* para ν. No entanto, em situações onde as estimativas de máxima verossimilhança foram ruins, como no exemplo de dados

simulados, o desempenho das estimativas bayesianas se mostraram mais satisfatórios, para quaisquer distribuições *a priori* para ν escolhidas.

Através da metodologia bayesiana, conseguimos determinar intervalos de credibilidade para os parâmetros de interesse. Em especial, notamos que, independente da distribuição *a priori* escolhida para ν, os intervalos de credibilidade para λ conseguiram determinar com precisão se a distribuição utilizada na modelagem é assimétrica positiva ou negativa.

A seguir, apresentamos algumas possibilidades de futuras pesquisas relacionadas a distribuição t-assimétrica.

- Estender os resultados obtidos neste trabalho para famílias de distribuições mais gerais.
- Fazer um estudo de simulação para comparar estimativas bayesianas às de máxima verossimilhança.
- As estimativas bayesianas dos parâmetros da distribuição t-assimétrica mostraram-se muito dependentes da escolha da distribuição *a priori* para ν. Nesse sentido, caberia um estudo de simulação para diagnosticar quais distribuições *a priori* geram estimativas sensíveis a esta escolha.
- Obter, do ponto de vista bayesiano, distribuições *a priori* de referência para os parâmetros da distribuição t-assimétrica.
- Estudar o modelo de regressão linear simples com erros t-assimétricos.

Apêndice A

Provas das Propriedades

(P1) Se $Z \sim \mathrm{ST}(\nu, \lambda)$ então $|Z| \sim \mathrm{HT}(\nu)$.

Prova:

Considere $T = |Z|$, onde $t \geq 0$. Então,

$$F_T(t) = P(T \leq t) = P(|Z| \leq t) = P(-t \leq Z \leq t) = F_Z(t) - F_Z(-t).$$

Derivando com relação a t em ambos os lados da igualdade e sabendo que $Z \sim \mathrm{ST}(\nu, \lambda)$, temos

$$\begin{aligned} f_T(t) &= f_Z(t) + f_Z(-t) \\ &= 2\, t_\nu(t) T_{\nu+1}\left(\lambda t \sqrt{\frac{1+\nu}{\nu+t^2}}\right) + 2\, t_\nu(-t) T_{\nu+1}\left(\lambda(-t)\sqrt{\frac{1+\nu}{\nu+t^2}}\right). \\ &= 2\frac{\Gamma\left(\frac{\nu+1}{2}\right)}{\Gamma\left(\frac{\nu}{2}\right)} \frac{1}{\sqrt{\nu\pi}} \left(1+\frac{t^2}{\nu}\right)^{-\frac{\nu+1}{2}} T_{\nu+1}\left(\lambda t \sqrt{\frac{1+\nu}{\nu+t^2}}\right) + 2\frac{\Gamma\left(\frac{\nu+1}{2}\right)}{\Gamma\left(\frac{\nu}{2}\right)} \frac{1}{\sqrt{\nu\pi}} \times \\ &\qquad \times \left(1+\frac{t^2}{\nu}\right)^{-\frac{\nu+1}{2}} T_{\nu+1}\left(\lambda(-t)\sqrt{\frac{1+\nu}{\nu+t^2}}\right). \end{aligned}$$

Portanto,

$$
\begin{aligned}
f_T(t) &= 2\,\frac{\Gamma\left(\frac{\nu+1}{2}\right)}{\Gamma\left(\frac{\nu}{2}\right)}\,\frac{1}{\sqrt{\nu\pi}}\left(1+\frac{t^2}{\nu}\right)^{-\frac{\nu+1}{2}}\left[T_{\nu+1}\left(\lambda t\sqrt{\frac{1+\nu}{\nu+t^2}}\right)+1-\right.\\
&\qquad\qquad\left.-\,T_{\nu+1}\left(\lambda t\sqrt{\frac{1+\nu}{\nu+t^2}}\right)\right]\\
&= 2\,\frac{\Gamma\left(\frac{\nu+1}{2}\right)}{\Gamma\left(\frac{\nu}{2}\right)}\,\frac{1}{\sqrt{\nu\pi}}\left(1+\frac{t^2}{\nu}\right)^{-\frac{\nu+1}{2}}, \quad \text{para} \quad t>0.
\end{aligned}
$$

Concluí-se então que $T = |Z| \sim \mathrm{HT}(\nu)$.

♦

(P2) Se $Z \sim \mathrm{ST}(\nu,\,\lambda)$ e $\nu = 1$ então $Z \sim SC\,(\lambda)$.

Prova:

Substituindo ν por 1 em (2.2), obtemos:

$$
\begin{aligned}
f_Z(z) &= 2\,t_1(z)\,T_2\left(\lambda z\sqrt{\frac{2}{1+z^2}}\right)\\
&= 2\,\frac{\Gamma(1)}{\Gamma\left(\frac{1}{2}\right)}\frac{1}{\sqrt{\pi}}(1+z^2)^{-1}\int_{-\infty}^{\lambda z\sqrt{\frac{2}{1+z^2}}}\frac{\Gamma\left(\frac{3}{2}\right)}{\Gamma(1)}\frac{1}{\sqrt{2\pi}}\left(1+\frac{a^2}{2}\right)^{-3/2}da.
\end{aligned}
$$

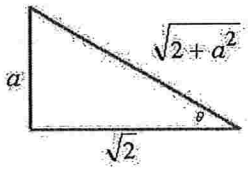
Alguns valores específicos da função gama são conhecidos, tais como $\Gamma(1) = 1$, $\Gamma\left(\frac{1}{2}\right) = \sqrt{\pi}$ e $\Gamma\left(\frac{3}{2}\right) = \frac{\sqrt{\pi}}{2}$. Dessa forma,

$$
\begin{aligned}
f_Z(z) &= \frac{1}{\pi(1+z^2)\sqrt{2}}\int_{-\infty}^{\lambda z\sqrt{\frac{2}{1+z^2}}}\left(1+\frac{a^2}{2}\right)^{-3/2}da\\
&= \frac{2^{3/2}}{\pi(1+z^2)\sqrt{2}}\int_{-\infty}^{\lambda z\sqrt{\frac{2}{1+z^2}}}\frac{1}{\sqrt{[2+a^2]^3}}\,da.
\end{aligned}
$$

Fazendo a mudança da variável a para $\sqrt{2}\ tg\ \theta$, temos:

- $2 + a^2 = 2 + (\sqrt{2}\ tg\ \theta)^2 = 2 + 2\ tg^2\ \theta = 2\ (1 + tg^2\ \theta) = 2\ sec^2\ \theta$,
- $da = \sqrt{2}\ sec^2\ \theta\ d\theta$,
- $tg\ \theta = \frac{a}{\sqrt{2}} \Rightarrow sen\ \theta = \frac{a}{\sqrt{2+a^2}}$, pois

Figura A.1: Relação trigonométrica



Sendo assim, resolvendo exclusivamente a integral, temos

$$\int \frac{1}{\sqrt{[2+a^2]^3}}\ da = \int \frac{\sqrt{2}\ sec^2\ \theta}{\sqrt{[2\ sec^2\ \theta]^3}}\ d\theta = \int \frac{1}{2\ sec\ \theta}\ d\theta = \int \frac{cos\theta}{2}\ d\theta = \frac{sen\theta}{2} + C.$$

Mas, $sen\ \theta = \frac{a}{\sqrt{2+a^2}}$, então

$$\begin{aligned} f_Z(z) &= \frac{2^{3/2}}{\pi(1+z^2)\sqrt{2}}\ \frac{1}{2}\left[\frac{a}{\sqrt{2+a^2}} + C'\right]\Bigg|_{-\infty}^{\lambda z\sqrt{\frac{2}{1+z^2}}} \\ &= \frac{1}{\pi(1+z^2)}\left[\frac{\lambda z\sqrt{\frac{2}{1+z^2}}}{\sqrt{2+\left(\lambda z\sqrt{\frac{2}{1+z^2}}\right)^2}} - \lim_{a\to-\infty}\frac{a}{\sqrt{2+a^2}}\right] \\ &= \frac{1}{\pi(1+z^2)}\left[\frac{\lambda z}{\sqrt{1+z^2+(\lambda z)^2}} - (-1)\right]. \end{aligned}$$

Dessa forma, temos que

$$f_Z(z) = \frac{1}{\pi\left(1+z^2\right)}\left[1+\frac{\lambda z}{\sqrt{1+\left(1+\lambda^2\right)z^2}}\right], \quad \text{para} \quad -\infty < z < \infty.$$

Portanto, $Z \sim SC\left(\lambda\right)$, quando $\nu = 1$.

♦

(P3) Quando $\lambda \to \infty$, $Z \sim \mathrm{ST}(\nu,\, \lambda)$ converge a uma $HT(\nu)$.

Prova:

Pela Proposição 2.1, podemos representar Z como $\frac{X}{\sqrt{\frac{W}{\nu}}}$, onde $X \sim SN(\lambda)$, $W \sim \text{Gama}\left(\frac{\nu}{2}, \frac{1}{2}\right)$ e X independente de W.

Sabe-se que, se $X \sim SN(\lambda)$ e $\lambda \to \infty$ então X converge a uma $HN(0,1)$ (ver Rodríguez, 2005).

Dessa forma, para $\lambda \to \infty$, temos que Z converge para Z^*:

$$Z^* = \frac{X^*}{\sqrt{\frac{W}{\nu}}}, \quad \text{onde} \quad \begin{cases} X^* \sim HN(0,1), \\ W \sim Gama\left(\frac{\nu}{2},\, \frac{1}{2}\right), \\ X^* \perp W. \end{cases}$$

Vamos obter a função de densidade conjunta de Z^* e $U = W$, através do Jacobiano da transformação. Dessa forma,

$$f_{Z^*,U}(z,u) = f_{X^*,W}\left(z\sqrt{\frac{u}{\nu}}, u\right)\left|\sqrt{\frac{u}{\nu}}\right|.$$

Pela independência das variáveis X^* e W, temos

$$f_{Z^*,U}(z,u) = f_{X^*}\left(z\sqrt{\frac{u}{\nu}}\right) f_W(u)\sqrt{\frac{u}{\nu}}.$$

A distribuição marginal Z é obtida a partir da distribuição conjunta de Z e W, da seguinte maneira:

$$
\begin{aligned}
f_{Z^*}(z) &= \int_0^\infty f_{Z^*,U}(z,u)du \\
&= \int_0^\infty 2\,\frac{1}{\sqrt{2\pi}}\,e^{-\frac{z^2 u}{2\nu}}\,\frac{\left(\frac{1}{2}\right)^{\nu/2}}{\Gamma(\nu/2)}u^{\nu/2-1}e^{-u/2}\left(\frac{u}{\nu}\right)^{1/2}du \\
&= 2\,\frac{\Gamma\left(\frac{\nu+1}{2}\right)}{\Gamma\left(\frac{\nu}{2}\right)}\,\frac{1}{\sqrt{\nu\pi}}\left(1+\frac{z^2}{\nu}\right)^{-\frac{\nu+1}{2}}\int_0^\infty \frac{\left[\frac{1}{2}\left(1+\frac{z^2}{\nu}\right)\right]^{\frac{\nu+1}{2}}}{\Gamma\left(\frac{\nu+1}{2}\right)}\times \\
&\qquad \underbrace{\times\; e^{-u\left[\frac{1}{2}\left(1+\frac{z^2}{\nu}\right)\right]}\,u^{\frac{\nu+1}{2}-1}}_{Gama\left(\frac{\nu+1}{2},\,\frac{1}{2}\left(1+\frac{z^2}{\nu}\right)\right)}\,du \\
&= 2\,t_\nu(z), \quad \text{para} \quad z>0.
\end{aligned}
$$

Portanto, $Z \sim ST(\nu,\,\lambda)$ converge a $HT(\nu)$, quando $\lambda \to \infty$.

♦

(P4) Quando $\nu \to \infty$, $Z \sim ST(\nu,\,\lambda)$ converge a uma $SN(\lambda)$.

Prova:

Considerando a caracterização apresentada na Proposição 2.1, temos que

$$
Z = \frac{X}{\sqrt{\frac{W}{\nu}}} \sim ST(\nu,\lambda)\,,\ \text{onde } X \sim SN(\lambda),\ W \sim \text{Gama}\left(\tfrac{\nu}{2},\tfrac{1}{2}\right) \text{ e } X \perp W.
$$

Vejamos o que ocorre com a variável aleatória $\frac{W}{\nu}$, para $W \sim \text{Gama}\left(\frac{\nu}{2},\frac{1}{2}\right)$, quando $\nu \to \infty$. Sabemos que $E(W) = \nu$ e $Var(W) = 2\,\nu$, então $\forall\ \epsilon > 0$, temos:

$$
P\left(\left|\frac{W}{\nu}-1\right| \geq \epsilon\right) = P\left(|W-\nu| \geq \epsilon\nu\right) \leq \frac{Var(W)}{\epsilon^2\nu^2} = \frac{2}{\epsilon^{\,2}\nu} \to 0, \text{ quando } \nu \to \infty.
$$

Ou seja, quando $\nu \to \infty$ temos que $\frac{W}{\nu} \xrightarrow{P} 1$. Dessa forma, como $X \sim SN(\lambda)$ e utilizando o teorema de Slutsky , temos

$$\frac{X}{\sqrt{\frac{W}{\nu}}} \xrightarrow{D} X.$$

Portanto, $Z \sim ST(\nu, \lambda)$ converge a uma $SN(\lambda)$, quando $\nu \to \infty$.

♦

(P5) Se $Z \sim ST(\nu, \lambda)$, então $-Z \sim ST(\nu, -\lambda)$.

Prova:

Considere $Z \sim ST(\nu, \lambda)$ e $T = -Z$. Então:

$$F_T(t) = P(T \leq t) = P(-Z \leq t) = P(Z \geq -t) = 1 - F_Z(-t).$$

Derivando com relação a t em ambos os lados e sabendo que $Z \sim ST(\nu, \lambda)$, temos

$$\begin{aligned}
f_T(t) &= -f_Z(-t)(-1) = f_Z(-t) \\
&= 2\, t_\nu(-t) T_{\nu+1}\left(\lambda(-t)\sqrt{\frac{1+\nu}{\nu+(-t)^2}}\right) \\
&= 2\,\frac{\Gamma\left(\frac{\nu+1}{2}\right)}{\Gamma\left(\frac{\nu}{2}\right)}\,\frac{1}{\sqrt{\nu\pi}}\left[1+\frac{(-t)^2}{\nu}\right]^{-\frac{\nu+1}{2}} T_{\nu+1}\left((-\lambda)\, t\sqrt{\frac{1+\nu}{\nu+t^2}}\right) \\
&= 2\,\frac{\Gamma\left(\frac{\nu+1}{2}\right)}{\Gamma\left(\frac{\nu}{2}\right)}\,\frac{1}{\sqrt{\nu\pi}}\left[1+\frac{t^2}{\nu}\right]^{-\frac{\nu+1}{2}} T_{\nu+1}\left((-\lambda)\, t\sqrt{\frac{1+\nu}{\nu+t^2}}\right) \\
&= 2\, t_\nu(t)\, T_{\nu+1}\left((-\lambda)\, t\sqrt{\frac{1+\nu}{\nu+t^2}}\right), \quad \text{para} \quad -\infty < t < \infty.
\end{aligned}$$

Portanto, $T = -Z \sim ST(\nu, -\lambda)$.

♦

(P6) Se $Z \sim ST(\nu, \lambda)$, então $Z^2 \sim$ *F-Snedocor*$(1, \nu)$.

Prova:

Considere $T = Z^2$, onde $0 < t < \infty$. Então,

$$F_T(t) = P(T \leq t) = P(Z^2 \leq t) = P(-\sqrt{t} \leq Z \leq \sqrt{t}) = F_Z(\sqrt{t}) - F_Z(-\sqrt{t}).$$

Derivando com relação a t, temos

$$\begin{aligned}
f_T(t) &= f_Z(\sqrt{t})\frac{1}{2\sqrt{t}} - f_Z(-\sqrt{t})\left(-\frac{1}{2\sqrt{t}}\right) \\
&= \frac{1}{2\sqrt{t}}\left(f_Z(\sqrt{t}) + f_Z(-\sqrt{t})\right) \\
&= \frac{1}{2\sqrt{t}}\left[2\, t_\nu(\sqrt{t})T_{\nu+1}\left(\lambda\sqrt{t}\sqrt{\frac{1+\nu}{\nu+t}}\right) + 2\, t_\nu(-\sqrt{t})T_{\nu+1}\left(-\lambda\sqrt{t}\sqrt{\frac{1+\nu}{\nu+t}}\right)\right] \\
&= \frac{1}{\sqrt{t}}\left\{\frac{\Gamma\left(\frac{\nu+1}{2}\right)}{\Gamma\left(\frac{\nu}{2}\right)}\frac{1}{\sqrt{\nu\pi}}\left[1+\frac{t}{\nu}\right]^{-\frac{\nu+1}{2}} T_{\nu+1}\left(\lambda\sqrt{\frac{t\,(1+\nu)}{\nu+t}}\right) + \frac{\Gamma\left(\frac{\nu+1}{2}\right)}{\Gamma\left(\frac{\nu}{2}\right)} \times \right. \\
&\qquad \left. \times \frac{1}{\sqrt{\nu\pi}}\left[1+\frac{t}{\nu}\right]^{-\frac{\nu+1}{2}} T_{\nu+1}\left(-\lambda\sqrt{\frac{t\,(1+\nu)}{\nu+t}}\right)\right\} \\
&= \frac{1}{\sqrt{t}}\left\{\frac{\Gamma\left(\frac{\nu+1}{2}\right)}{\Gamma\left(\frac{\nu}{2}\right)}\frac{1}{\sqrt{\nu\pi}}\left[1+\frac{t}{\nu}\right]^{-\frac{\nu+1}{2}}\left[T_{\nu+1}\left(\lambda\sqrt{\frac{t\,(1+\nu)}{\nu+t}}\right) + \right.\right. \\
&\qquad \left.\left. + T_{\nu+1}\left(-\lambda\sqrt{\frac{t\,(1+\nu)}{\nu+t}}\right)\right]\right\}. \qquad \text{(A.1)}
\end{aligned}$$

Como a função densidade de probabilidade de uma distribuição t padrão é simétrica ao redor de zero, temos que:

$$T_{\nu+1}\left(-\lambda\sqrt{\frac{t\,(1+\nu)}{\nu+t}}\right) = 1 - T_{\nu+1}\left(\lambda\sqrt{\frac{t\,(1+\nu)}{\nu+t}}\right)$$

$$T_{\nu+1}\left(-\lambda\sqrt{\frac{t\,(1+\nu)}{\nu+t}}\right) + T_{\nu+1}\left(\lambda\sqrt{\frac{t\,(1+\nu)}{\nu+t}}\right) = 1. \qquad \text{(A.2)}$$

Substituindo (A.2) em (A.1) e lembrando que $\sqrt{\pi} = \Gamma\left(\frac{1}{2}\right)$, obtemos:

$$\begin{aligned} f_T(t) &= \frac{1}{\sqrt{t}}\left\{\frac{\Gamma\left(\frac{\nu+1}{2}\right)}{\Gamma\left(\frac{\nu}{2}\right)}\frac{1}{\sqrt{\nu\pi}}\left[1+\frac{t}{\nu}\right]^{-\frac{\nu+1}{2}}\right\} \\ &= \frac{\Gamma\left(\frac{\nu+1}{2}\right)}{\Gamma\left(\frac{\nu}{2}\right)\Gamma\left(\frac{1}{2}\right)}\left(\frac{1}{\nu}\right)^{1/2} t^{\frac{1-2}{2}}\,\frac{1}{\left(1+\frac{t}{\nu}\right)^{\frac{\nu+1}{2}}}, \quad \text{para} \quad t>0. \end{aligned}$$

Esta última expressão é a função densidade de probabilidade de uma variável aleatória com distribuição $F_{1,\nu}$. Portanto, $Z^2 \sim F_{1,\nu}$, quando $Z \sim ST(\nu,\lambda)$.

♦

(P7) F_Y(y; ν, -λ) = 1 - F_Y(-y; ν, λ).

Prova:

Consideremos, inicialmente, o caso em que Z tem distribuição t-assimétrica padrão. Por definição,

$$F_Z(z;\nu,-\lambda) = \int_{-\infty}^{z} 2\,t_\nu(x)\,T_{\nu+1}\left(-\lambda x\sqrt{\frac{1+\nu}{\nu+x^2}}\right)dx.$$

Fazendo uma mudança de variável em $x=-u$, temos

$$\begin{aligned} F_Z(z;\nu,-\lambda) &= -\int_{\infty}^{-z} 2\,t_\nu(-u)\,T_{\nu+1}\left(\lambda u\sqrt{\frac{1+\nu}{\nu+u^2}}\right)du \\ &= \int_{-z}^{\infty} 2\,t_\nu(-u)\,T_{\nu+1}\left(\lambda u\sqrt{\frac{1+\nu}{\nu+u^2}}\right)du. \end{aligned}$$

Sendo assim,

$$\begin{aligned} F_Z(z;\nu,-\lambda) &= 1-\int_{-\infty}^{-z} 2\, t_\nu(u)\, T_{\nu+1}\left(\lambda u\sqrt{\frac{1+\nu}{\nu+u^2}}\right) du \\ &= 1-F_Z(-z;\ \nu,\ \lambda). \end{aligned}$$

Logo, F_Z(z; ν, -λ) = 1 - F_Z(-z; ν, λ).

♦

(P8) Se $Y \sim$ St(μ, σ^2, ν, λ) e Y_1 = a + bY então

$$Y_1 \sim ST(a+b\mu,\ b^2\sigma^2,\ sinal(b)\lambda,\ \nu).$$

Prova:

Analisaremos a distribuição de Y_1 em duas situações. O primeiro caso considera $b > 0$. Então,

$$F_{Y_1}(r) = P(Y_1 \le r) = P(a+bY \le r) = P\left(Y \le \frac{r-a}{b}\right) = F_Y\left(\frac{r-a}{b}\right).$$

Derivando com relação a r em ambos os lados da igualdade, obtemos:

$$\begin{aligned} f_{Y_1}(r) &= f_Y\left(\frac{r-a}{b}\right)\frac{1}{b} \\ &= \frac{2}{\sigma b}\, t_\nu\left(\frac{r-(a+b\mu)}{b\sigma}\right) T_{\nu+1}\left[\lambda\left(\frac{r-(a+b\mu)}{b\sigma}\right)\times\right. \\ &\qquad \left.\times\sqrt{\frac{\nu+1}{\nu+(\sigma b)^{-2}(r-(a+b\mu))^2}}\right]. \end{aligned}$$

Portanto, $Y_1 \sim (a+b\mu,\ b^2\sigma^2,\ \lambda,\ \nu)$.

Para $b < 0$, temos

$$
\begin{aligned}
F_{Y_1}(r) &= P(Y_1 \leq r) = P(a + bY \leq r) = P(bY \leq r - a) = P(-bY \geq -(r-a)) \\
&= P\left(Y \geq \frac{-(r-a)}{-b}\right) = P\left(Y \geq \frac{r-a}{b}\right) = 1 - P\left(Y < \frac{r-a}{b}\right) \\
&= 1 - F_Y\left(\frac{r-a}{b}\right).
\end{aligned}
$$

Derivando com relação a r em ambos os lados da igualdade, obtemos:

$$
\begin{aligned}
f_{Y_1}(r) &= -f_Y\left(\frac{r-a}{b}\right)\frac{1}{b} \\
&= -\frac{2}{\sigma b}\, t_\nu\left(\frac{r-a-b\mu}{b\sigma}\right) \times \\
&\qquad \times T_{\nu+1}\left[\lambda\left(\frac{r-a-b\mu}{b\sigma}\right)\sqrt{\frac{\nu+1}{\nu+(\sigma b)^{-2}(r-a-b\mu)^2}}\right] \\
&= \frac{2}{\sigma(-b)}\, t_\nu\left(\frac{r-(a+b\mu)}{b\sigma}\right) \times \\
&\qquad \times T_{\nu+1}\left[-\lambda\left(\frac{r-(a+b\mu)}{-b\sigma}\right)\sqrt{\frac{\nu+1}{\nu+(\sigma b)^{-2}(r-(a+b\mu))^2}}\right].
\end{aligned}
$$

Portanto, para $b < 0$, temos que $Y_1 \sim (a + b\mu, b^2\sigma^2, -\lambda, \nu)$. Dessa forma, concluímos que $Y_1 \sim (a + b\mu, b^2\sigma^2, sinal(b)\lambda, \nu)$.

♦

Apêndice B

Obtenção dos Estimadores de Momentos

Proposição B.1 *Seja* $\mathbf{Z} = (Z_1, Z_2, \ldots, Z_n)$ *uma amostra aleatória simples de tamanho* n *da variável aleatória* Z *com distribuição* $ST(\nu, \lambda)$. *Então, para* $\overline{Z} = \frac{\sum_{i=1}^{n} Z_i}{n}$ *e* $m_2 = \frac{\sum_{i=1}^{n} {Z_i}^2}{n}$, *o estimador de momentos para os parâmetros da distribuição t-assimétrica padrão são:*

$$\widetilde{\lambda} = \frac{\overline{Z}\,\Gamma\left(\frac{\widetilde{\nu}}{2}\right)}{\sqrt{\frac{\widetilde{\nu}}{\pi}\,\Gamma^2\left(\frac{\widetilde{\nu}-1}{2}\right) - \overline{Z}^2\,\Gamma^2\left(\frac{\widetilde{\nu}}{2}\right)}} \quad e \quad \widetilde{\nu} = \frac{2m_2}{m_2 - 1}.$$

onde

$$\overline{Z} \in \left]-\sqrt{\frac{\widetilde{\nu}}{\pi}}\,\frac{\Gamma\left(\frac{\widetilde{\nu}-1}{2}\right)}{\Gamma\left(\frac{\widetilde{\nu}}{2}\right)};\ \sqrt{\frac{\widetilde{\nu}}{\pi}}\,\frac{\Gamma\left(\frac{\widetilde{\nu}-1}{2}\right)}{\Gamma\left(\frac{\widetilde{\nu}}{2}\right)}\right[\quad e \quad m_2 > 1.$$

Prova:

Igualando os momentos amostrais aos populacionais, temos que

$$E(Z) = \overline{Z} \quad \Leftrightarrow \quad \frac{\widetilde{\lambda}}{\sqrt{1+\widetilde{\lambda}^2}}\sqrt{\frac{\widetilde{\nu}}{\pi}}\,\frac{\Gamma\left(\frac{\widetilde{\nu}-1}{2}\right)}{\Gamma\left(\frac{\widetilde{\nu}}{2}\right)} = \overline{Z} \quad \Leftrightarrow$$

$$\frac{\widetilde{\lambda}}{\sqrt{1+\widetilde{\lambda}^2}} = \frac{\overline{Z}}{\sqrt{\frac{\widetilde{\nu}}{\pi}}}\,\frac{\Gamma\left(\frac{\widetilde{\nu}}{2}\right)}{\Gamma\left(\frac{\widetilde{\nu}-1}{2}\right)} \quad \Leftrightarrow \quad \widetilde{\lambda} = \frac{\overline{Z}\,\Gamma\left(\frac{\widetilde{\nu}}{2}\right)}{\sqrt{\frac{\widetilde{\nu}}{\pi}\,\Gamma^2\left(\frac{\widetilde{\nu}-1}{2}\right) - \overline{Z}^2\,\Gamma^2\left(\frac{\widetilde{\nu}}{2}\right)}}, \tag{B.1}$$

$$E(Z^2) = m_2 \Leftrightarrow \frac{\widetilde{\nu}}{\widetilde{\nu}-2} = m_2 \Leftrightarrow \widetilde{\nu} = \frac{2\,m_2}{m_2-1}. \tag{B.2}$$

Em (B.1) e (B.2) obtivemos os estimadores de momentos para os parâmetros da distribuição t-assimétrica padrão. Note que a existência dos estimadores está condicionada as seguintes situações:

I. $\frac{\widetilde{\nu}}{\pi}\,\Gamma^2\left(\frac{\widetilde{\nu}-1}{2}\right) - \overline{Z}^2\,\Gamma^2\left(\frac{\widetilde{\nu}}{2}\right) > 0 \Leftrightarrow \overline{Z} \in \left]-\sqrt{\frac{\widetilde{\nu}}{\pi}}\,\frac{\Gamma\left(\frac{\widetilde{\nu}-1}{2}\right)}{\Gamma\left(\frac{\widetilde{\nu}}{2}\right)};\ \sqrt{\frac{\widetilde{\nu}}{\pi}}\,\frac{\Gamma\left(\frac{\widetilde{\nu}-1}{2}\right)}{\Gamma\left(\frac{\widetilde{\nu}}{2}\right)}\right[,$

II. $\widetilde{\nu} > 0 \Leftrightarrow \frac{2\,m_2}{m_2-1} > 0$ e $m_2 > 0 \Leftrightarrow m_2 > 1$.

♦

Proposição B.2 *Seja* $\mathbf{Y} = (Y_1, Y_2, \ldots, Y_n)$ *uma amostra aleatória simples de tamanho* n *da variável aleatória* Y *com distribuição* $ST(\mu, \sigma^2, \nu, \lambda)$. *Considere* $\delta = \frac{\lambda}{\sqrt{1+\lambda^2}}$ *e os momentos amostrais dados por* $m_j = \sum_{i=1}^{n} \frac{Y_i^j}{n}$, *para* $j = 1, \ldots, 4$. *Então, o estimador de momentos para o parâmetro de posição é solução da seguinte equação:*

$$\widetilde{\sigma}^2\,(3-\widetilde{\delta}^2)\,\frac{\widetilde{\nu}}{\widetilde{\nu}-3}\,(m_1-\widetilde{\mu}) = m_3 + 3\widetilde{\mu}^2 m_1 - 3\widetilde{\mu} m_2 - \widetilde{\mu}^3.$$

e os estimadores de ν, δ^2, σ^2 *são dados em função do estimador de momentos de* μ, *como segue:*

$$\widetilde{\nu} = \frac{4\,a(\widetilde{\mu}) - 6\,b(\widetilde{\mu})^2}{a(\widetilde{\mu}) - 3\,b(\widetilde{\mu})^2},$$

$$\widetilde{\sigma}^2 = \frac{a(\widetilde{\mu})\,b(\widetilde{\mu})}{2\,a(\widetilde{\mu}) - 3\,b(\widetilde{\mu})^2},$$

$$\widetilde{\delta}^2 = \left[\frac{\Gamma\left(\frac{\widetilde{\nu}}{2}\right)}{\Gamma\left(\frac{\widetilde{\nu}-1}{2}\right)}\right]^2 \frac{\pi\,(m_1-\widetilde{\mu})^2\,(a(\widetilde{\mu}) - 3\,b(\widetilde{\mu})^2)}{2\,a(\widetilde{\mu})\,b(\widetilde{\mu})},$$

onde

$$a(\widetilde{\mu}) = m_4 - 4\widetilde{\mu}^3 m_1 + 6\widetilde{\mu}^2 m_2 - 4\widetilde{\mu} m_3 + \widetilde{\mu}^4 \quad e \quad b(\widetilde{\mu}) = -2\widetilde{\mu} m_1 + m_2 + \widetilde{\mu}^2.$$

Prova:

Igualando os momentos populacionais aos amostrais e observando que Y pode ser reescrito como $\mu + \sigma\ Z$, onde $Z \sim ST(\nu, \lambda)$, obtemos as seguintes igualdades:

I. $E(Y) = m_1 \Leftrightarrow \widetilde{\sigma}\ E(Z) = m_1 - \widetilde{\mu}$,

II. $E(Y^2) = m_2 \Leftrightarrow \widetilde{\sigma}^2\ E(Z^2) = \underbrace{-2\widetilde{\mu}m_1 + m_2 + \widetilde{\mu}^2}_{b(\widetilde{\mu})}$,

III. $E(Y^3) = m_3 \Leftrightarrow \widetilde{\sigma}^3\ E(Z^3) = m_3 - 3\widetilde{\mu}m_2 + 3\widetilde{\mu}^2 m_1 - \widetilde{\mu}^3$,

IV. $E(Y^4) = m_4 \Leftrightarrow \widetilde{\sigma}^4\ E(Z^4) = \underbrace{m_4 - 4\widetilde{\mu}m_3 + 6\widetilde{\mu}^2 m_2 - 4\widetilde{\mu}^3 m_1 + \widetilde{\mu}^4}_{a(\widetilde{\mu})}$.

Através da função geradora de momentos, obtemos em (2.8) e (2.9) os momentos de ordem par e ímpar de uma variável aleatória Z com distribuição t-assimétrica padrão. Sendo assim,

$$E(Z) = \frac{\lambda}{\sqrt{1+\lambda^2}} \left(\frac{\nu}{\pi}\right)^{1/2} \frac{\Gamma\left(\frac{\nu-1}{2}\right)}{\Gamma\left(\frac{\nu}{2}\right)}, \text{ para } \nu > 1 \text{ e } E(Z^2) = \frac{\nu}{\nu-2}, \text{ para } \nu > 2.$$

$$E(Z^3) = E(Z)(3-\delta^2)\frac{\nu}{\nu-3}, \text{ para } \nu > 3 \text{ e } E(Z^4) = \frac{3\ \nu^2}{(\nu-2)(\nu-4)}, \text{ para } \nu > 4.$$

Por (II), temos que

$$\widetilde{\sigma}^2 E(Z^2) = b(\widetilde{\mu}) \Leftrightarrow \widetilde{\sigma}^2 \frac{\widetilde{\nu}}{\widetilde{\nu}-2} = b(\widetilde{\mu}) \Leftrightarrow \widetilde{\sigma}^2 \widetilde{\nu} = b(\widetilde{\mu})\ (\widetilde{\nu}-2). \tag{B.3}$$

Substituindo o quarto momento da variável aleatória Z e o resultado (B.3) em (IV), obtemos

$$\widetilde{\nu} = \frac{4\ a(\widetilde{\mu}) - 6\ b(\widetilde{\mu})^2}{a(\widetilde{\mu}) - 3\ b(\widetilde{\mu})^2}. \tag{B.4}$$

Substituindo (B.4) em (B.3) obtemos,

$$\widetilde{\sigma}^2 = \frac{a(\widetilde{\mu})\ b(\widetilde{\mu})}{2\ a(\widetilde{\mu}) - 3\ b(\widetilde{\mu})^2}.$$

Em (I), substituímos o primeiro momento da variável Z e os estimadores $\widetilde{\nu}$ e $\widetilde{\sigma}^2$, encontrando, assim,

$$\widetilde{\delta}^2 = \left[\frac{\Gamma\left(\frac{\widetilde{\nu}}{2}\right)}{\Gamma\left(\frac{\widetilde{\nu}-1}{2}\right)}\right]^2 \frac{\pi\ (m_1 - \widetilde{\mu})^2\ (a(\widetilde{\mu}) - 3\ b(\widetilde{\mu})^2)}{2\ a(\widetilde{\mu})\ b(\widetilde{\mu})}.$$

Substituindo os estimadores encontrados, o terceiro momento da variável Z e observando que $E(Z) = \frac{m_1 - \widetilde{\mu}}{\widetilde{\sigma}}$ em (III) conseguimos visualizar a equação, cuja a solução nos permitirá obter o estimador de momentos para o parâmetro de posição $\widetilde{\mu}$. Esta equação é dada por:

$$\widetilde{\sigma}^2\ (3 - \widetilde{\delta}^2)\ \frac{\widetilde{\nu}}{\widetilde{\nu} - 3}\ (m_1 - \widetilde{\mu}) = m_3 + 3\widetilde{\mu}^2 m_1 - 3\widetilde{\mu} m_2 - \widetilde{\mu}^3.$$

♦

Apêndice C

Obtenção das Equações de Verossimilhança

Proposição C.1 *Considere* $\mathbf{Y} = (Y_1, Y_2, \ldots, Y_n)$ *uma amostra aleatória simples de tamanho* n *de* $Y \sim ST(\mu,\ \sigma^2,\ \lambda,\ \nu)$. *Então as equações de verossimilhança são dadas por:*

$$\frac{\partial \ell}{\partial \mu} = \sum_{i=1}^{n} \left\{ \frac{1}{\sigma\left(\nu + a_i^2\right)} \left[(\nu+1)\, a_i - \frac{t_{\nu+1}\left(\lambda a_i k_i\right)}{T_{\nu+1}\left(\lambda a_i k_i\right)} \lambda k_i \nu \right] \right\} = 0,$$

$$\frac{\partial \ell}{\partial \sigma} = -\frac{n}{\sigma} + \sum_{i=1}^{n} \left\{ \frac{a_i k_i}{\sigma} \left[\frac{a_i k_i}{\sigma} - \lambda\, \frac{\nu}{\nu + a_i^2} \left(\frac{t_{\nu+1}(\lambda a_i k_i)}{T_{\nu+1}(\lambda a_i k_i)} \right) \right] \right\} = 0,$$

$$\frac{\partial \ell}{\partial \lambda} = \sum_{i=1}^{n} a_i k_i\, \frac{t_{\nu+1}(\lambda a_i k_i)}{T_{\nu+1}(\lambda a_i k_i)} = 0,$$

$$\frac{\partial \ell}{\partial \nu} = \sum_{i=1}^{n} \left\{ \frac{1}{2} \left[\psi\left(\frac{\nu+1}{2}\right) - \psi\left(\frac{\nu}{2}\right) + 1 - \log\left(1 + \frac{a_i^2}{\nu}\right) - k_i^2 \right] + \right.$$
$$\left. + \frac{\partial}{\partial \nu} \log\, T_{\nu+1}\left(\lambda a_i k_i\right) \right\} = 0,$$

onde $a_i = \frac{y_i-\mu}{\sigma}$, $k_i = \sqrt{\frac{\nu+1}{\nu+a_i^2}}$ *e* ψ *corresponde a função digama, que se caracteriza por:* $\psi(z) = \frac{\partial}{\partial z}\, ln\, \Gamma(z) = \frac{\Gamma'(z)}{\Gamma(z)}$.

Prova:

A função de log-verossimilhança é dada por:

$$\begin{aligned} l(\mu,\sigma,\lambda,\nu;\mathbf{y}) &= \log L(\mu,\sigma,\lambda,\nu;\mathbf{y}) \\ &= n\log 2 - n\log\sigma + \sum_{i=1}^{n}\log\, t_v\left(\frac{y_i-\mu}{\sigma}\right) + \\ &\qquad \sum_{i=1}^{n}\log\, T_{v+1}\left[\lambda\left(\frac{y_i-\mu}{\sigma}\right)\sqrt{\frac{\nu+1}{\nu+\sigma^{-2}(y_i-\mu)^2}}\right]. \end{aligned}$$

Derivada em relação ao parâmetro de posição:

$$\frac{\partial\ell}{\partial\mu} = \sum_{i=1}^{n}\left\{\frac{\partial}{\partial\mu}\log\, t_\nu\left(\frac{y_i-\mu}{\sigma}\right) + \frac{\partial}{\partial\mu}\log\, T_{\nu+1}\left[\lambda\left(\frac{y_i-\mu}{\sigma}\right)\sqrt{\frac{\nu+1}{\nu+\sigma^{-2}(y_i-\mu)^2}}\right]\right\}.$$

Resolvendo separadamente a primeira parcela entre chaves, temos:

$$\begin{aligned} \frac{\partial}{\partial\mu}\log\, t_\nu\left(\frac{y_i-\mu}{\sigma}\right) &= \frac{\partial}{\partial\mu}\log\left\{\frac{\Gamma\left(\frac{\nu+1}{2}\right)}{\Gamma\left(\frac{\nu}{2}\right)}\,\frac{1}{\sqrt{\pi}}\,\frac{\nu^{\frac{\nu}{2}}}{\left[\nu+\sigma^{-2}(y_i-\mu)^2\right]^{\frac{\nu+1}{2}}}\right\} \\ &= \frac{\partial}{\partial\mu}\left\{-\frac{\nu+1}{2}\,\log\,\left[\nu+\sigma^{-2}(y_i-\mu)^2\right]\right\} \\ &= \frac{\nu+1}{\sigma}\,\left(\frac{y_i-\mu}{\sigma}\right)\frac{1}{\nu+\sigma^{-2}(y_i-\mu)^2}. \end{aligned} \tag{C.1}$$

Agora, resolvendo a segunda parcela entre chaves, temos:

$$\begin{aligned} \frac{\partial}{\partial\mu}\log T_{\nu+1}\left[\lambda\left(\frac{y_i-\mu}{\sigma}\right)\sqrt{\frac{\nu+1}{\nu+\sigma^{-2}(y_i-\mu)^2}}\right] &= \frac{t_{\nu+1}\left[\lambda\left(\frac{y_i-\mu}{\sigma}\right)\sqrt{\frac{\nu+1}{\nu+\sigma^{-2}(y_i-\mu)^2}}\right]}{T_{\nu+1}\left[\lambda\left(\frac{y_i-\mu}{\sigma}\right)\sqrt{\frac{\nu+1}{\nu+\sigma^{-2}(y_i-\mu)^2}}\right]} \times \\ &\times\left\{\frac{\lambda}{\sigma}\sqrt{\frac{\nu+1}{\nu+\sigma^{-2}(y_i-\mu)^2}}\,\left[-1+\left(\frac{y_i-\mu}{\sigma}\right)^2\frac{1}{\nu+\sigma^{-2}(y_i-\mu)^2}\right]\right\}. \end{aligned} \tag{C.2}$$

Considere $a_i = \frac{y_i-\mu}{\sigma}$ e $k_i = \sqrt{\frac{\nu+1}{\nu+\sigma^{-2}(y_i-\mu)^2}}$ em (C.1) e (C.2). Então,

$$\begin{aligned}\frac{\partial \ell}{\partial \mu} &= \sum_{i=1}^{n}\left\{\frac{\nu+1}{\sigma}\frac{a_i}{\nu+a_i^2}+\frac{t_{\nu+1}(\lambda a_i k_i)}{T_{\nu+1}(\lambda a_i k_i)}\left[\frac{\lambda}{\sigma}k_i\left(-1+\frac{a_i^2}{\nu+a_i^2}\right)\right]\right\}\\ &= \sum_{i=1}^{n}\left\{\frac{1}{\sigma(\nu+a_i^2)}\left[(\nu+1)a_i-\frac{t_{\nu+1}(\lambda a_i k_i)}{T_{\nu+1}(\lambda a_i k_i)}\lambda k_i \nu\right]\right\}.\end{aligned}$$

Derivada em relação ao parâmetro de escala:

$$\begin{aligned}\frac{\partial \ell}{\partial \sigma} &= -\frac{\partial}{\partial \sigma}n\log\sigma+\sum_{i=1}^{n}\left\{\frac{\partial}{\partial \sigma}\log t_\nu\left(\frac{y_i-\mu}{\sigma}\right)+\right.\\ &\qquad\left.+\frac{\partial}{\partial \sigma}\log T_{\nu+1}\left[\lambda\left(\frac{y_i-\mu}{\sigma}\right)\sqrt{\frac{\nu+1}{\nu+\sigma^{-2}(y_i-\mu)^2}}\right]\right\}.\end{aligned}$$

Resolvendo a primeira parcela, temos que $-\frac{\partial}{\partial\sigma}(n\log\sigma) = -\frac{n}{\sigma}$. Para a primeira parcela entre chaves, temos:

$$\begin{aligned}\frac{\partial}{\partial \sigma}\log t_\nu\left(\frac{y_i-\mu}{\sigma}\right) &= \frac{\partial}{\partial \sigma}\log\left\{\frac{\Gamma\left(\frac{\nu+1}{2}\right)}{\Gamma\left(\frac{\nu}{2}\right)}\frac{1}{\sqrt{\pi}}\frac{\nu^{\frac{\nu}{2}}}{\left[\nu+\sigma^{-2}(y_i-\mu)^2\right]^{\frac{\nu+1}{2}}}\right\}\\ &= \frac{\partial}{\partial \sigma}\left\{-\frac{\nu+1}{2}\log\left[\nu+\sigma^{-2}(y_i-\mu)^2\right]\right\}\\ &= \frac{\nu+1}{\sigma}\left(\frac{y_i-\mu}{\sigma}\right)^2\frac{1}{\nu+\sigma^{-2}(y_i-\mu)^2}.\end{aligned} \tag{C.3}$$

Derivando a segunda parcela entre chaves, temos:

$$\begin{aligned}\frac{\partial}{\partial \sigma}\log T_{\nu+1}\left[\lambda\left(\frac{y_i-\mu}{\sigma}\right)\sqrt{\frac{\nu+1}{\nu+\sigma^{-2}(y_i-\mu)^2}}\right] &= \frac{t_{\nu+1}\left[\lambda\left(\frac{y_i-\mu}{\sigma}\right)\sqrt{\frac{\nu+1}{\nu+\sigma^{-2}(y_i-\mu)^2}}\right]}{T_{\nu+1}\left[\lambda\left(\frac{y_i-\mu}{\sigma}\right)\sqrt{\frac{\nu+1}{\nu+\sigma^{-2}(y_i-\mu)^2}}\right]}\times\\ &\times\left\{\frac{\lambda}{\sigma}\left(\frac{y_i-\mu}{\sigma}\right)\sqrt{\frac{\nu+1}{\nu+\sigma^{-2}(y_i-\mu)^2}}\left[-1+\left(\frac{y_i-\mu}{\sigma}\right)^2\frac{1}{\nu+\sigma^{-2}(y_i-\mu)^2}\right]\right\}.\end{aligned} \tag{C.4}$$

Considere em (C.3) e (C.4) $a_i = \frac{y_i-\mu}{\sigma}$ e $k_i = \sqrt{\frac{\nu+1}{\nu+\sigma^{-2}(y_i-\mu)^2}}$. Logo,

$$\begin{aligned}\frac{\partial \ell}{\partial \sigma} &= -\frac{n}{\sigma} + \sum_{i=1}^{n}\left\{\frac{a_i^2k_i^2}{\sigma} + \frac{t_{\nu+1}(\lambda a_i k_i)}{T_{\nu+1}(\lambda a_i k_i)}\left[\frac{\lambda a_i k_i}{\sigma}\left(-1+\frac{a_i^2}{\nu+a_i^2}\right)\right]\right\} \\ &= -\frac{n}{\sigma} + \sum_{i=1}^{n}\left\{\frac{a_i k_i}{\sigma}\left[\frac{a_i k_i}{\sigma} - \lambda\frac{\nu}{\nu+a_i^2}\left(\frac{t_{\nu+1}(\lambda a_i k_i)}{T_{\nu+1}(\lambda a_i k_i)}\right)\right]\right\}.\end{aligned}$$

Derivada em relação ao parâmetro de assimetria:

$$\begin{aligned}\frac{\partial \ell}{\partial \lambda} &= \sum_{i=1}^{n}\frac{\partial}{\partial\lambda}\log T_{\nu+1}\left[\lambda\left(\frac{y_i-\mu}{\sigma}\right)\sqrt{\frac{\nu+1}{\nu+\sigma^{-2}(y_i-\mu)^2}}\right] \\ &= \sum_{i=1}^{n}\frac{t_{\nu+1}\left[\lambda\left(\frac{y_i-\mu}{\sigma}\right)\sqrt{\frac{\nu+1}{\nu+\sigma^{-2}(y_i-\mu)^2}}\right]}{T_{\nu+1}\left[\lambda\left(\frac{y_i-\mu}{\sigma}\right)\sqrt{\frac{\nu+1}{\nu+\sigma^{-2}(y_i-\mu)^2}}\right]}\left[\left(\frac{y_i-\mu}{\sigma}\right)\sqrt{\frac{\nu+1}{\nu+\sigma^{-2}(y_i-\mu)^2}}\right].\end{aligned}$$

Para $a_i = \frac{y_i-\mu}{\sigma}$ e $k_i = \sqrt{\frac{\nu+1}{\nu+\sigma^{-2}(y_i-\mu)^2}}$, temos

$$\frac{\partial \ell}{\partial \lambda} = \sum_{i=1}^{n} a_i k_i \frac{t_{\nu+1}(\lambda a_i k_i)}{T_{\nu+1}(\lambda a_i k_i)}.$$

Derivada em relação aos graus de liberdade:

$$\frac{\partial \ell}{\partial \nu} = \sum_{i=1}^{n}\left\{\frac{\partial}{\partial\nu}\log t_\nu\left(\frac{y_i-\mu}{\sigma}\right) + \frac{\partial}{\partial\nu}\log T_{\nu+1}\left[\lambda\left(\frac{y_i-\mu}{\sigma}\right)\sqrt{\frac{\nu+1}{\nu+\sigma^{-2}(y_i-\mu)^2}}\right]\right\}.$$

Inicialmente, vamos resolver a primeira parcela entre chaves, então

$$\begin{aligned}
\frac{\partial}{\partial\nu}\log\, t_\nu\left(\frac{y_i-\mu}{\sigma}\right) &= \frac{\partial}{\partial\nu}\log\left\{\frac{\Gamma\left(\frac{\nu+1}{2}\right)}{\Gamma\left(\frac{\nu}{2}\right)}\,\frac{1}{\sqrt{\pi}}\frac{\nu^{\frac{\nu}{2}}}{[\nu+\sigma^{-2}(y_i-\mu)^2]^{\frac{\nu+1}{2}}}\right\} \\
&= \frac{\partial}{\partial\nu}\left[\log\,\Gamma\left(\frac{\nu+1}{2}\right)-\log\,\Gamma\left(\frac{\nu}{2}\right)+\frac{\nu}{2}\,\log\nu+\right. \\
&\qquad\left.+\left(-\frac{\nu+1}{2}\right)\log\left(\nu+\sigma^{-2}(y_i-\mu)^2\right)\right].
\end{aligned}$$

Mas, observe que ψ corresponde a função digama, que se caracteriza por:

$$\psi(z)=\frac{\partial}{\partial z}\,\log\,\Gamma(z)=\frac{\Gamma'(z)}{\Gamma(z)}.$$

Então,

$$\begin{aligned}
\frac{\partial}{\partial\nu}\log\, t_\nu\left(\frac{y_i-\mu}{\sigma}\right) &= \frac{1}{2}\psi\left(\frac{\nu+1}{2}\right)-\frac{1}{2}\psi\left(\frac{\nu}{2}\right)+\frac{1}{2}\log\nu+\frac{1}{2}- \\
&\quad -\frac{1}{2}\log\left(\nu+\sigma^{-2}(y_i-\mu)^2\right)-\left(\frac{\nu+1}{2}\right)\frac{1}{\nu+\sigma^{-2}(y_i-\mu)^2} \\
&= \frac{1}{2}\left\{\psi\left(\frac{\nu+1}{2}\right)-\psi\left(\frac{\nu}{2}\right)+1-\log\left[1+\frac{1}{\nu}\left(\frac{y_i-\mu}{\sigma}\right)^2\right]-\right. \\
&\qquad\left.-\frac{\nu+1}{\nu+\sigma^{-2}(y_i-\mu)^2}\right\}. \qquad \text{(C.5)}
\end{aligned}$$

Não conseguimos desenvolver a segunda parcela entre chaves, então para $a_i=\frac{y_i-\mu}{\sigma}$ e $k_i=\sqrt{\frac{\nu+1}{\nu+\sigma^{-2}(y_i-\mu)^2}}$, temos

$$\begin{aligned}
\frac{\partial\ell}{\partial\nu} &= \sum_{i=1}^{n}\left\{\frac{1}{2}\left[\psi\left(\frac{\nu+1}{2}\right)-\psi\left(\frac{\nu}{2}\right)+1-\log\left(1+\frac{a_i^2}{\nu}\right)-k_i^2\right]+\right. \\
&\qquad\left.+\frac{\partial}{\partial\nu}\log\,T_{\nu+1}\left(\lambda a_i k_i\right)\right\}.
\end{aligned}$$

♦

Apêndice D

Tabelas

Tabela D.1: Mediana das E.M. e M.V. (Padrão) para λ

n	ν	Parâmetro de Assimetria							
		0		1		5		20	
		EMV-P	EM	EMV-P	EM	EMV-P	EM	EMV-P	EM
20	3	0.015	0.016	1.033	1.117	4.931	2.980	12.030	2.997
		(1.37)	(0.46)	(1.59)	(1.66)	(7.73)	(6.42)	(13.40)	(7.04)
	8	0.010	-0.035	1.040	1.094	5.068	3.163	13.240	3.507
		(0.83)	(0.38)	(2.08)	(1.44)	(6.17)	(5.83)	(24.36)	(5.74)
	12	-0.034	-0.015	1.024	1.180	4.765	4.032	13.790	3.396
		(1.65)	(0.36)	(2.09)	(1.82)	(6.32)	(3.49)	(18.42)	(7.28)
	100	-0.002	-0.026	1.105	1.329	5.209	3.291	14.050	3.596
		(1.60)	(26.17)	(1.488)	(1.58)	(6.60)	(5.02)	(17.31)	(7.71)
50	3	0.005	0.009	1.017	1.071	5.469	3.382	16.250	3.723
		(0.19)	(0.25)	(0.36)	(0.89)	(5.49)	(11.0)	(12.40)	(6.10)
	8	0.002	0.010	1.010	1.133	5.192	3.560	17.180	4.257
		(1.28)	(0.19)	(0.99)	(0.71)	(9.46)	(3.87)	(19.87)	(10.9)
	12	0.008	0.011	1.018	1.091	5.473	3.607	19.210	4.110
		(0.18)	(0.19)	(0.35)	(0.92)	(7.52)	(3.81)	(32.67)	(6.01)
	100	-0.007	-0.001	1.050	1.150	5.223	4.161	19.300	4.871
		(0.19)	(0.18)	(2.93)	(0.54)	(5.86)	(5.03)	(54.14)	(5.71)

Tabela D.2: Mediana das E.M. e M.V. (Padrão) para λ - Cont.

n	ν	Parâmetro de Assimetria							
		0		1		5		20	
		EMV-P	EM	EMV-P	EM	EMV-P	EM	EMV-P	EM
80	3	-0.005	0.007	1.018	1.078	5.14	3.534	19.020	4.230
		(0.15)	(0.18)	(0.29)	(0.68)	(3.05)	(5.28)	(13.25)	(19.22)
	8	-0.003	-0.002	1.014	1.071	5.274	3.525	20.000	4.239
		(1.01)	(0.14)	(1.44)	(0.33)	(4.13)	(6.76)	(22.19)	(5.34)
	12	0.005	0.006	1.020	1.049	5.260	3.820	21.790	4.684
		(0.15)	(0.14)	(0.25)	(0.33)	(3.12)	(6.76)	(18.09)	(5.33)
	100	-0.009	0.017	1.040	1.111	5.269	4.606	20.030	5.370
		(0.16)	(0.14)	(0.26)	(0.37)	(3.19)	(8.42)	(46.31)	(6.49)
100	3	-0.001	0.012	0.986	1.089	5.068	3.840	18.730	4.145
		(0.12)	(0.16)	(0.23)	(0.48)	(2.27)	(4.01)	(11.63)	(12.85)
	8	-0.001	0.001	0.998	1.043	5.183	3.660	20.260	4.457
		(0.13)	(0.13)	(0.21)	(0.31)	(2.83)	(5.43)	(18.74)	(5.32)
	12	-0.001	-0.001	0.999	1.054	5.078	4.052	20.110	5.041
		(0.13)	(0.13)	(0.22)	(0.29)	(2.99)	(8.51)	(21.79)	(8.78)
	100	-0.011	-0.006	1.024	1.095	5.147	4.732	20.990	5.689
		(0.13)	(0.13)	(0.21)	(0.28)	(3.23)	(11.70)	(23.08)	(6.19)
200	3	-0.005	0.000	0.998	1.050	5.032	4.089	21.220	4.805
		(0.09)	(0.11)	(0.16)	(0.25)	(1.51)	(8.48)	(13.58)	(6.10)
	8	-0.001	0.001	1.008	1.019	5.037	4.195	21.300	5.015
		(0.09)	(0.09)	(0.15)	(0.19)	(1.37)	(4.89)	(16.23)	(57.57)
	12	0.000	0.002	1.006	1.001	5.078	4.416	20.920	5.648
		(0.088)	(0.94)	(0.15)	(0.19)	(1.34)	(6.01)	(21.12)	(39.82)
	100	-0.005	0.001	1.027	1.059	5.290	5.023	20.740	6.456
		(0.09)	(0.09)	(0.139)	(0.19)	(1.17)	(11.7)	(15.12)	(18.3)

Tabela D.3: Mediana das E.M. e M.V. (Padrão) para ν

n	λ	Graus de Liberdade							
		3		8		12		100	
		EMV-P	EM	EMV-P	EM	EMV-P	EM	EMV-P	EM
20	0	3.172	3.703	7.182	6.513	11.990	7.918	17.910	10.820
		(23.9)	(9.33)	(41.17)	(17.65)	(42.86)	(19.09)	(51.54)	(26.17)
	1	3.368	3.983	8.447	7.083	8.770	8.014	19.060	11.321
		(16.48)	(10.09)	(27.29)	(14.61)	(33.70)	(25.88)	(37.70)	(23.93)
	5	3.512	4.850	12.030	8.669	30.220	10.830	40.370	14.120
		(17.82)	(15.99)	(34.11)	(25.05)	(32.44)	(26.48)	(39.68)	(30.93)
	20	3.824	5.022	15.110	8.393	47.560	10.300	64.690	15.050
		(29.69)	(15.99)	(46.25)	(25.05)	(48.92)	(26.48)	(51.03)	(30.93)
50	0	3.182	3.565	8.402	7.492	10.360	6.533	37.040	14.910
		(10.84)	(6.20)	(31.22)	(17.95)	(33.11)	(20.77)	(27.52)	(29.69)
	1	3.179	3.662	8.500	7.679	15.030	9.900	28.580	15.320
		(6.901)	(4.13)	(22.44)	(17.01)	(24.99)	(21.90)	(41.93)	(27.24)
	5	3.303	4.413	10.210	9.534	16.440	12.680	43.830	19.050
		(5.225)	(11.00)	(30.81)	(20.48)	(32.30)	(26.57)	(35.32)	(31.49)
	20	3.364	4.049	9.228	9.090	18.120	12.610	66.130	21.720
		(13.62)	(6.15)	(38.77)	(27.17)	(53.83)	(24.14)	(61.4)	(34.99)
80	0	3.059	3.409	7.965	7.727	15.590	10.170	54.400	18.670
		(5.64)	(1.83)	(31.7)	(21.00)	(34.12)	(26.32)	(40.65)	(31.26)
	1	3.131	3.347	7.998	7.747	13.050	11.110	39.750	20.140
		(3.10)	(1.55)	(25.08)	(16.30)	(25.2)	(26.32)	(31.33)	(32.48)
	5	3.124	3.718	9.540	9.437	14.860	12.710	43.860	22.570
		(3.5)	(4.04)	(23.41)	(24.85)	(29.52)	(26.46)	(47.79)	(32.61)
	20	3.189	3.740	8.862	9.913	15.950	13.360	78.110	25.790
		(2.31)	(4.31)	(34.83)	(16.05)	(42.96)	(27.82)	(51.15)	(36.24)
100	0	2.994	3.376	7.496	7.940	14.400	10.530	56.160	20.660
		(1.35)	(1.49)	(4.36)	(13.38)	(30.29)	(24.15)	(40.57)	(33.40)
	1	3.054	3.385	8.643	8.194	13.410	10.940	42.960	20.220
		(1.18)	(1.77)	(18.82)	(14.89)	(26.04)	(21.03)	(37.97)	(33.47)
	5	3.006	3.706	8.734	9.223	15.330	12.960	45.720	25.670
		(1.4)	(2.53)	(18.44)	(17.72)	(28.92)	(24.20)	(37.37)	(35.01)
	20	2.824	3.642	9.113	9.571	13.590	13.620	71.880	28.900
		(3.28)	(5.50)	(29.45)	(17.71)	(34.32)	(25.61)	(65.36)	(39.07)

Tabela D.4: Mediana das E.M. e M.V. (Padrão) para ν - Cont.

n	λ	Graus de Liberdade							
		3		8		12		100	
		EMV-P	EM	EMV-P	EM	EMV-P	EM	EMV-P	EM
200	0	3.044	3.287	8.236	8.370	12.170	11.250	58.330	25.460
		(0.67)	(0.64)	(12.43)	(8.91)	(22.36)	(17.97)	(32.09)	(38.75)
	1	3.014	3.276	8.236	8.092	12.550	11.440	54.640	23.980
		(0.66)	(0.63)	(8.57)	(6.39)	(19.88)	(20.22)	(38.53)	(32.08)
	5	3.077	3.386	8.174	9.075	13.350	13.050	48.560	33.340
		(1.41)	(0.88)	(10.19)	(14.26)	(22.25)	(27.86)	(37.19)	(39.05)
	20	3.065	3.326	8.630	9.284	13.450	13.840	68.050	36.200
		(0.95)	(0.82)	(13.3)	(13.04)	(30.26)	(24.69)	(54.03)	(41.61)

Tabela D.5: Estimativas de Máxima Verossimilhança para λ

n	ν	Parâmetro de Assimetria			
		0	1	5	20
		Mediana - Média (Desvio-padrão)	Mediana - Média (Desvio-padrão)	Mediana - Média (Desvio-padrão)	Mediana - Média (Desvio-padrão)
20	3	-0.092 -0.058 (2.52)	1.230 1.3670 (2.75)	3.684 4.398 (2.97)	5.352 6.207 (3.77)
	8	0.073 0.090 (2.79)	1.084 1.201 (3.35)	3.405 3.783 (3.35)	5.797 6.502 (4.07)
	12	-0.022 0.020 (3.42)	0.702 0.720 (3.03)	3.497 4.085 (3.46)	6.010 6.846 (4.08)
	100	-0.028 0.033 (3.22)	0.159 0.344 (3.079)	3.225 3.747 (3.52)	6.671 7.286 (4.25)
50	3	-0.086 -0.066 (1.51)	1.174 1.528 (1.77)	5.428 6.630 (4.45)	12.42 13.140 (6.41)
	8	-0.054 -0.014 (2.89)	1.166 1.383 (2.69)	4.96 6.590 (5.22)	12.66 14.260 (7.91)
	12	-0.004 -0.049 (1.98)	1.097 1.045 (1.92)	5.035 6.749 (5.50)	13.030 14.740 (8.09)
	100	-0.030 0.029 (2.22)	0.803 1.028 (2.77)	5.021 6.455 (5.30)	12.900 14.670 (8.26)

Tabela D.6: Estimativas de Máxima Verossimilhança para λ - Cont.

n	ν	Parâmetro de Assimetria			
		0	1	5	20
		Mediana - Média (Desvio-padrão)	Mediana - Média (Desvio-padrão)	Mediana - Média (Desvio-padrão)	Mediana - Média (Desvio-padrão)
80	3	0.038 0.016 (0.74)	1.045 1.186 (0.90)	5.540 6.816 (5.09)	16.020 18.250 (9.39)
	8	-0.05 0.037 (1.17)	1.077 1.0990 (1.24)	5.750 7.104 (5.03)	16.632 19.370 (10.49)
	12	0.037 0.063 (2.06)	1.122 1.018 (1.24)	5.231 7.003 (6.12)	16.480 25.800 (11.76)
	100	-0.180 -0.078 (2.35)	0.708 0.622 (1.73)	5.032 6.719 (6.44)	17.500 20.070 (11.8)
100	3	-0.020 -0.021 (0.76)	1.010 1.179 (0.81)	5.082 6.215 (4.10)	17.520 19.990 (10.51)
	8	-0.080 -0.054 (1.18)	1.100 1.095 (1.14)	5.233 6.374 (4.10)	18.750 22.180 (13.04)
	12	0.090 0.061 (1.06)	1.130 1.095 (1.09)	5.529 6.781 (5.14)	18.400 22.080 (13.11)
	100	-0.123 -0.075 (1.33)	0.830 0.719 (1.41)	4.983 6.055 (4.2)	21.340 26.640 (12.91)
200	3	0.039 0.031 (0.40)	1.024 1.082 (0.47)	5.411 5.989 (3.41)	20.210 25.120 (16.36)
	8	0.025 0.002 (0.72)	1.071 1.097 (0.64)	5.186 6.032 (5.42)	21.570 27.950 (20.27)
	12	0.003 0.023 (0.84)	0.997 1.001 (0.69)	5.297 6.007 (6.8)	21.270 27.640 (16.63)
	100	-0.035 0.005 (0.967)	0.871 0.729 (0.86)	4.994 5.373 (1.80)	21.39 27.460 (20.10)

Tabela D.7: Estimativas de Máxima Verossimilhança para ν

n	λ	Graus de Liberdade			
		3	8	12	100
		Mediana - Média (Desvio-padrão)	Mediana - Média (Desvio-padrão)	Mediana - Média (Desvio-padrão)	Mediana - Média (Desvio-padrão)
20	0	5.509 10^{15} (10^{17})	6302 10^{30} (10^{32})	10570 10^{11} (10^{12})	11020 10^{8} (10^{10})
	1	7.209 10^{9} (10^{11})	7077 10^{42} (10^{43})	10060 10^{11} (10^{13})	11260 10^{29} (10^{23})
	5	4.628 10^{4} (10^{5})	9200 10^{12} (10^{14})	10270 10^{10} (10^{12})	10810 10^{21} (10^{11})
	20	3.702 10^{28} (10^{29})	13970 10^{40} (10^{41})	54030 10^{59} (10^{60})	140600 10^{31} (10^{33})
50	0	3.493 10^{4} (10^{7})	23.730 10^{27} (10^{28})	2628 10^{6} (10^{8})	9998 10^{15} (10^{16})
	1	3.474 10^{4} (10^{6})	19.170 10^{18} (10^{19})	2134 10^{16} (10^{17})	11060 10^{11} (10^{12})
	5	3.382 10^{3} (10^{5})	11.950 10^{4} (10^{6})	170.300 10^{7} (10^{9})	10005 10^{43} (10^{44})
	20	3.016 10^{6} (10^{9})	10.020 10^{16} (10^{18})	3.1970 10^{17} (10^{19})	10630 10^{33} (10^{34})
80	0	3.248 10^{3} (10^{6})	10.470 10^{4} (10^{6})	77.030 10^{12} (10^{13})	9925 10^{7} (10^{9})
	1	3.375 10^{3} (10^{5})	12.550 10^{4} (10^{6})	31.250 10^{4} (10^{6})	9417 10^{19} (10^{21})
	5	3.240 312.600 (10^{4})	13.610 10^{46} (10^{48})	27.260 10^{20} (10^{21})	10190 10^{14} (10^{15})
	20	3.140 312.400 (10^{4})	9.405 10^{21} (10^{22})	24.790 10^{12} (10^{13})	10160 10^{30} (10^{31})
100	0	3.212 219.100 (10^{4})	10.480 10^{4} (10^{6})	27.520 10^{3} (10^{7})	7641 10^{6} (10^{8})
	1	3.174 10^{3} (10^{6})	10.910 10^{4} (10^{7})	25.960 10^{9} (10^{11})	7514 10^{6} (10^{8})
	5	3.121 10^{13} (10^{14})	9.969 10^{15} (10^{17})	21.020 10^{59} (10^{80})	10120 10^{10} (10^{11})
	20	3.156 10^{3} (10^{6})	9.857 10^{9} (10^{12})	15.790 10^{22} (10^{24})	10470 10^{17} (10^{18})

Tabela D.8: Estimativas de Máxima Verossimilhança para ν - Cont.

n	λ	Graus de Liberdade			
		3	8	12	100
		Mediana (Desvio-padrão)	Mediana (Desvio-padrão)	Mediana (Desvio-padrão)	Mediana (Desvio-padrão)
200	0	5.057 3.096 (10^5)	9.334 10^3 (10^5)	16.67 10^6 (10^9)	28.17 10^5 (10^7)
	1	3.11 4.190 (15.93)	9.797 10^3 (10^5)	14.33 10^4 (10^6)	6292 10^4 (10^6)
	5	3.113 3.398 (1.10)	8.638 10^3 (10^5)	13.58 10^{19} (10^{21})	7379 10^4 (10^6)
	20	3.019 3.267 (1.07)	8.9 10^8 (10^{11})	13.3 10^{16} (10^{17})	10140 10^{12} (10^{13})

Apêndice E

Figuras

Figura E.1: Gráficos de Quantis

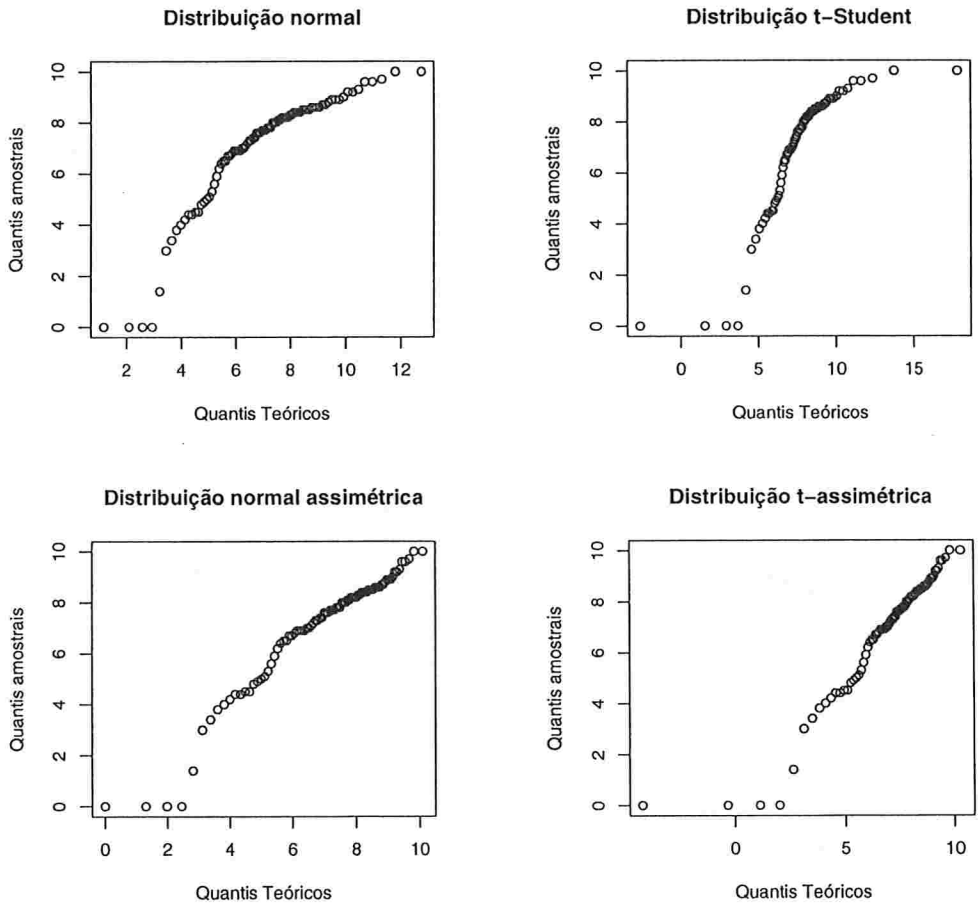


Figura E.2: Gráficos Healy-quantis

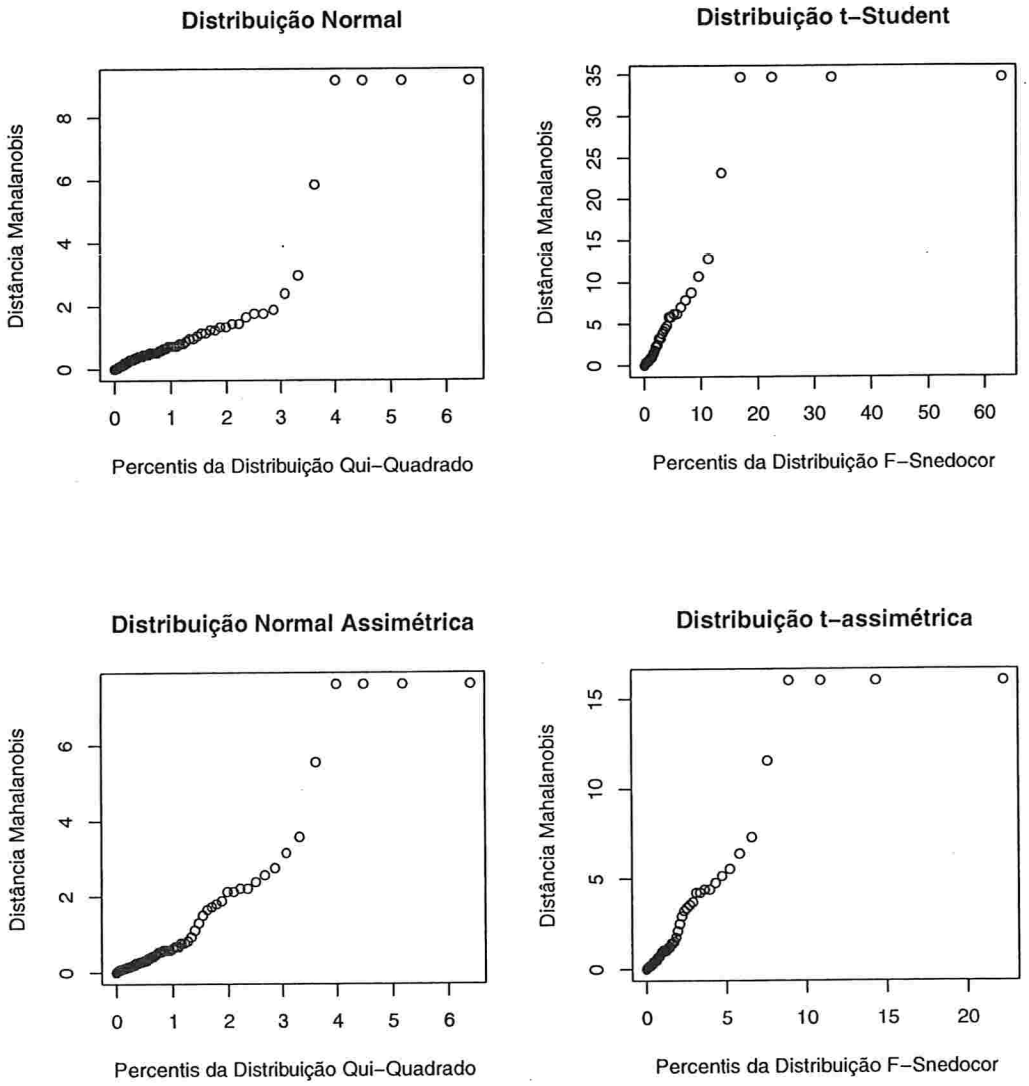
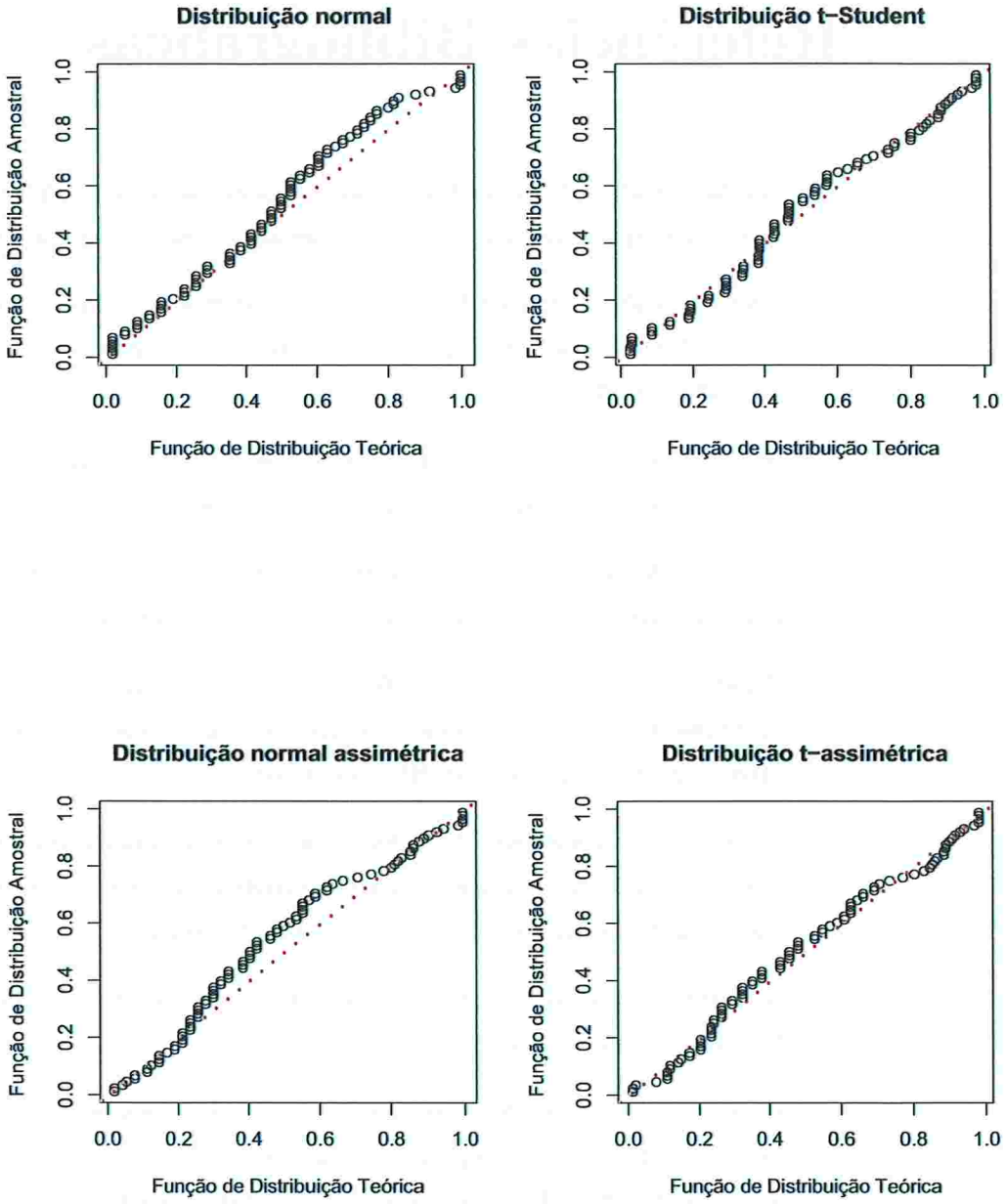


Figura E.3: Gráficos Healy-probabilidades Acumuladas



Referências Bibliográficas

[1] Arellano-Valle, R. B. (1994). *Distribuições Elípticas: Propriedades, Inferência e Aplicações a Modelos de Regressão*, Tese de Doutorado, IME-USP.

[2] Arellano-Valle, R.B., del Pino, G. & San Martín, E. (2002). Definition and probabilistic properties of skew-distributions, *Statistics and Probability Letters* **58**: 111-121.

[3] Azzalini, A.(1985). A class of distributions which includes the normal ones, *Scandinavian Journal of Statistics* **12**: 171-178.

[4] Azzalini, A.(2005). The skew-normal distribution and related multivariate families, *Scandinavian Journal of Statistics* **32**: 159-188.

[5] Azzalini, A. & Capitanio, A. (2002). Distributions generated by perturbation of symmetry with emphasis on a multivariate skew t distribution (full paper). http://azzalini.stat.unipd.it/SN/se-ext.pdf.

[6] Azzalini, A. & Capitanio, A. (2003). Distributions generated by perturbation of symmetry with emphasis on a multivariate skew t-distribution, *Journal of the Royal Statistical Society: Series B* **65(2)**: 367-389.

[7] Azzalini, A. & Dalla Valle, A. (1996). The multivariate skew-normal distribution, *Biometrika* **83**: 715-726.

[8] Berger, J.O. & Bernardo, J.M. (1992). On the development of reference priors (with Discussion), *In Bayesian Statistics 4 (ed. Bernardo, J.M., Berger, J.O., Dawid, A.P. & Smith, F.M.)*, pp. 35-60, Oxford University Press.

[9] Branco, M.D. & Arellano-Valle, R.B. (2004). *Distribuições Elípticas Assimétricas*, 16º SINAPE, ABE.

[10] Branco, M.D. & Dey, D.K. (2001). A general class of multivariate skew-elliptical distributions, *Journal of Multivariate Analysis*, **79**: 99-113.

[11] Branco, M.D. & Dey, D.K. (2002). Regression model under skew elliptical error distribution, *The Journal of Mathematical Sciences, Delhi, New Series*, **1**: 151-169.

[12] Cabras, S. & Castellanos, M.E. (2005). Testing the skew-normal model from a Bayesian perspective. Technical Report. http://bayes.escet.urjc.es/publicaciones/wp05-11.pdf.

[13] Cabras, S. & Castellanos, M.E. (2005). Comparing skewness in two populations using Bayesian inference on skew normal and skew t models. Technical Report. http://bayes.escet.urjc.es/publicaciones/wp05-12.pdf.

[14] Chang, F.C., Gupta, A.K. & Huang, W. J. (2005). Some skew-symmetric models. http://www.stat.nuk.edu.tw/huangwj/paper_new/2002-3.pdf.

[15] Dalla Valle, A. (2004). The skew normal distribution, *In Skew-elliptical distributions and their applications (ed. M.G. Genton)*, chapter 1, 3-24. London: Chapman and Hall/CRC.

[16] Fernández, C. & Steel, M.F.J. (1998). *On bayesian modeling of fat tails and skewness,* Journal of the American Statistical Association **93**: *359-371.*

[17] Ferreira, D.F. (2005). *Estatística Básica.* Lavras: Editora UFLA.

[18] Fonseca, T.C.O. (2004). *Análise Bayesiana de Referência para a Classe de Distribuições Hiperbólicas Generalizadas*, Dissertação de Mestrado, UFRJ.

[19] Gupta, A.K & Chen, T. (2001). Goodness-of-fit tests for the skew-normal distribution, *Communications in Statistics: Simulation and Computation* **30(4)**: 907-930.

[20] Healy, M.J.R. (1968). Multivariate normal plotting, *Applied Statistics* **17**: 157-161.

[21] Jones, M. C. (2001). A skew t distribution, *In Probability and Statistical Models with Applications (eds C. A. Charalambides, M. V. Koutras and N. Balakrishnan)*, 269-278. London: Chapman and Hall.

[22] Jones, M. C. (2002). Multivariate t and beta distributions associated with the multivariate F distributions, *Metrika* **54**: 215-231.

[23] Jones, M. C. & Faddy, M. J. (2003). A skew extension of the t-distribution, with applications, *Journal of the Royal Statistical Society: Series B* **65**: 159-174.

[24] Lange, K. L., Little, R. J. & Taylor, J.M.G. (1989). Robust Statistical Modelling Using the t-Distribution , *Journal of the American Statistical Association* **84**: 881-896.

[25] Liseo, B. & Loperfido, N. (2006). A note on reference priors for the scalar skew-normal distribution, *Journal of Statistical Planning and Inference* **136**: 373-389.

[26] Paulino, C., Turkman, A. & Murteira, B. (2003). *Estatística Bayesiana.* Lisboa: Fundação Calouste Gulbenkian.

[27] Rodríguez, C. L. B. (2005). *Inferência bayesiana no modelo normal assimétrico*, Dissertação de Mestrado, IME-USP.

[28] Sartori, N. (2006). Bias prevention of maximum likelihood estimates for scalar skew normal and skew t distributions, *Journal of Statistical Planning and Inference* **136**: 4259-4275.

[29] Severini, T. A. (2000). *Likelihood Methods in Statistcs.* New York: Oxford University Press.