Universidade de Brasília
Instituto de Ciências Exatas
Departamento de Estatística

Distribuições Inversíveis Simétricas e Assimétricas

Cláudia Raquel da Rocha Eirado

Dissertação de Mestrado

Orientador: Prof. Dr. Pushpa Narayan Rathie

Julho de 2011

Cláudia Raquel da Rocha Eirado

Distribuições Inversíveis Simétricas e Assimétricas

Dissertação apresentada ao Departamento de Estatística do Instituto de Ciências Exatas da Universidade de Brasília como requisito parcial à obtenção do título de Mestre em Estatística.

Universidade de Brasília

Brasília, Julho de 2011

Ao meu amado esposo Leonardo,

ao meu filho Mateus e

aos meus pais.

Agradecimentos

- Agradeço ao Soberano autor da vida, meu Deus, por ser meu refúgio e fortaleza.
- Ao meu esposo e filho, pela compreensão e incentivo em todos os momentos.
- Aos meus pais, por me orientarem ao longo da vida e me ensinarem a vencer desafios.
- E, finalmente, ao meu orientador Pushpa Narayan Rathie e Departamento de Estatística.

Sumário

Sumário **3**

Lista de Figuras **5**

Lista de Tabelas **6**

Resumo **7**

Abstract **8**

1 Introdução **9**

1.1 Motivação 9

2 Funções Matemáticas Especiais **10**

2.1 Introdução 10

2.2 A Função H(z) 10

2.2.1 Conceitos e fundamentos - [SPRINGER,1979] (18) 10

2.3 A Função Generalizada de Gauss ${}_pF_q$ 12

2.3.1 Conceitos e fundamentos - [SLATER,1966] (17) 12

2.3.2 A convergência da Função Generalizada de Gauss ${}_pF_q$ 13

2.3.3 A Função Hipergeométrica de Gauss ${}_2F_1$ 13

2.3.4 Convergência da Função de Gauss ${}_2F_1$ 13

2.3.5 Relações Funcionais de ${}_2F_1[a;b;c;z]$ 15

2.4 A Função G de Meijer 15

2.5 A Função Bivariada $H[x,y]$ 17

2.5.1 Conceitos e fundamentos - [MATHAI et al.,2010] (9) 17

2.6 Caso Especial da Função Bivariada $H[x,y]$ 19

2.6.1 Conceitos e fundamentos - [HAI & YAKUBOVICH, 1992] (8) 19

3 Uma Nova Distribuição 20

3.1 Introdução 20

3.2 Distribuição Acumulada Assimétrica 20

3.2.1 Propriedades da Distribuição Acumulada 21

3.2.2 Gráficos para a Distribuição Acumulada 22

3.2.3 Função Densidade de Probabilidade 25

3.2.4 Gráficos para a Função Densidade de Probabilidade 25

3.2.5 Distribuições de Mínimo e Máximo 30

3.2.6 Momentos para X 32

3.2.7 Função Característica para X 34

3.2.8 Estimadores para os parâmetros B, C, ρ de $f(x)$ 40

4 Versão Simétrica para a Distribuição 42

4.1 Introdução 42

4.2 Distribuição Acumulada Simétrica 42

4.2.1 Propriedades da Distribuição Acumulada Simétrica 42

4.2.2 Gráficos para a Distribuição Acumulada Simétrica 43

4.2.3 Função Densidade de Probabilidade Simétrica 47

4.2.4 Gráficos para a Função Densidade de Probabilidade Simétrica 47

4.2.5 Distribuições de Mínimo e Máximo 51

4.2.6 Momentos para X 52

4.2.7 Função Característica para X 54

4.2.8 Estimadores para os parâmetros B, C, ρ de $f(x)$ 55

5 Distribuições Assimétricas 57

5.1 Introdução 57

5.2 Distribuição Assimétrica $\mathrm{h}_1(x)$ 57

5.2.1 Construção de $h_1(x)$ 57

5.2.2 Gráfico para a Densidade Assimétrica $Sas(x)$ 58

5.3 Distribuição Assimétrica $h_2(x)$ 59

5.3.1 Construção de $h_2(x)$ 59

5.3.2 Gráfico para a Densidade Assimétrica $Sbs(x)$ 59

6 Aplicações **61**

6.1 Introdução 61

6.2 Precipitação 61

6.2.1 Análise Descritiva 61

6.2.2 Estimação dos Parâmetros 63

6.2.3 Ajuste da Curva 64

6.3 Retornos Diários de Taxa Câmbio 66

6.3.1 Análise Descritiva 66

6.3.2 Estimação dos Parâmetros 69

6.3.3 Ajuste da Curva 70

6.4 Consumo de Energia - kg de petróleo per capita 73

6.4.1 Análise Descritiva 73

6.4.2 Estimação dos Parâmetros 74

6.4.3 Ajuste da Curva 75

7 Conclusões **77**

Referências Bibliográficas **79**

Lista de Figuras

3.1 Formas para a Distribuição Acumulada Assimétrica $F(x)$ 22
3.2 Formas para a Distribuição Acumulada $F(x)$ - variação de ρ, com $B > 0$ 23
3.3 Formas para a Distribuição Acumulada $F(x)$ - variação de ρ, com $B < 0$ 23
3.4 Formas para a Distribuição Acumulada $F(x)$ - variação de B 24
3.5 Formas para a Distribuição Acumulada $F(x)$ - variação de C, com $B > 0$ 24
3.6 Formas para a Distribuição Acumulada $F(x)$ - variação de C, com $B < 0$ 24
3.7 Formas para a Densidade de Probabilidade Assimétrica $f(x)$ 25
3.8 Formas para a densidade Gama Generalizada 26
3.9 Formas para a Densidade de Probabilidade - variação de ρ, com $B > 0$ 27
3.10 Formas para a Densidade de Probabilidade - variação de ρ, com $B < 0$ 27
3.11 Formas para a Densidade de Probabilidade $f(x)$ - variação de B . . . 28
3.12 Formas para a Densidade de Probabilidade $f(x)$ - variação de C, com $B > 0$ 28
3.13 Formas para a Densidade de Probabilidade $f(x)$ - variação de C, com $B < 0$ 29
3.14 Comportamento da Densidade de Probabilidade $f(x)$ quando $x \to \infty$ 30

4.1 Formas para a Distribuição Acumulada Simétrica $G(x)$ 44
4.2 Formas para a Distribuição Simétrica Acumulada $G(x)$ - variação de ρ, com $B > 0$ 44
4.3 Formas para a Distribuição Simétrica Acumulada $G(x)$ - variação de ρ, com $B < 0$ 45
4.4 Formas para a Distribuição Simétrica Acumulada $G(x)$ - variação de B 45
4.5 Formas para a Distribuição Simétrica Acumulada $G(x)$ - variação de C, com $B > 0$ 46

4.6 Formas para a Distribuição Simétrica Acumulada $G(x)$ - variação de C, com $B < 0$ 46
4.7 Formas para a Densidade Simétrica $g(x)$ 48
4.8 Formas para a Densidade $g(x)$ - variação de ρ, com $B > 0$ 48
4.9 Formas para a Densidade $g(x)$ - variação de ρ, com $B < 0$ 49
4.10 Formas para a Densidade $g(x)$ - variação de B 49
4.11 Formas para a Densidade $g(x)$ - variação de C, com $B > 0$ 50
4.12 Formas para a Densidade $g(x)$ - variação de C, com $B < 0$ 50

5.1 Formas para a Densidade Assimétrica $Sas(x)$ 58
5.2 Formas para a Densidade Assimétrica $Sbs(x)$ 60

6.1 Boxplot para os Dados de Precipitação 62
6.2 Distribuição Empírica e Densidade Estimada por Kernel para os Dados de Precipitação 63
6.3 Análise de Bondade do Ajuste para os Dados de Precipitação 66
6.4 Boxplot para Retornos da Taxa de Câmbio - Real - Brasil 67
6.5 Boxplot para Retornos da Taxa de Câmbio - Dólar Canadense - Canadá 68
6.6 Distribuição Empírica e Densidade Estimada para Retornos da Taxa de Câmbio - Real - Brasil 68
6.7 Distribuição Empírica e Densidade Estimada para Retornos da Taxa de Câmbio - Dólar Canadense - Canadá 69
6.8 Análise de Bondade de Ajuste para Retornos Diários da Taxa de Câmbio - Real - Brasil 72
6.9 Análise de Bondade de Ajuste para Retornos Diários da Taxa de Câmbio - Dólar Canadense - Canadá 72
6.10 Boxplot para os Dados de Log-Consumo de Energia 74
6.11 Distribuição Empírica e Densidade Estimada por Kernel para os Dados de Log-Consumo de Energia 74
6.12 Análise de Bondade do Ajuste para os Dados de Log-Consumo de Energia 76

Lista de Tabelas

6.1 Estatísticas Descritivas para os Dados de Precipitação 62
6.2 E.M.V. e I.C. 95% Bootstrap para os Parâmetros B, C, ρ 64
6.3 Medidas de Acurácia de Ajuste para as Distribuição dos Dados de Precipitação 65
6.4 Estatísticas Descritivas para os Dados de Retornos Diários das Taxas de Câmbio do Brasil e Canadá 67
6.5 E.M.V. e I.C. Wald 95% para os Parâmetros B, C 70
6.6 Medidas de Acurácia de Ajuste para as Distribuições dos Retornos das Taxas Cambiais de Brasil e Canadá 71
6.7 Estatísticas Descritivas para os Dados de Log-Consumo de Energia . 73
6.8 E.M.V. para os Parâmetros $B, C, \gamma, \mu, \sigma$ 75
6.9 Medidas de Acurácia de Ajuste para as Distribuição dos Dados de Log-Consumo de Energia 75

Resumo

O estudo de novas distribuições se faz necessário à medida em que nos deparamos com dados que muitas vezes possuem peculiaridades que não se encaixam em distribuições já conhecidas.

O objetivo deste trabalho é apresentar novas distribuições assimétricas e simétricas, suas propriedades e aplicações com dados reais.

Palavras Chave: *Função H, função Hipergeométrica, funções de distribuição e densidades simétricas e assimétricas, estimadores de máxima verossimilhança.*

Abstract

The study of new distributions is necessary when we are faced with data that often do not fit known distributions.

The aim of this work is to present new skew symmetric and asymmetric distributions, their properties and applications with real data.

key words: *H-function, Hypergeometric function, skew and symmetric distributions and densities functions, maximum likelihood estimation.*

Capítulo 1

Introdução

1.1 Motivação

O trabalho trata de novas distribuições simétricas e assimétricas e suas propriedades. Por serem distribuições que podem assumir formas com caudas pesadas, podem ser utilizada para séries de dados financeiros e fenômenos naturais como índices pluviométricos e temperaturas em determinadas localidades ao longo do tempo. Será apresentada versão da distribuição apenas com valores positivos e outra versão para toda a reta real. Além de uma distribuição fortemente assimétrica gerada a partir da distribuição simétrica.

O objetivo é determinar propriedades da distribuição, encontrar a densidade, momentos, função característica e determinar estimadores para os parâmetros a fim de fazer inferência estatística a partir destas distribuições e modelar com dados reais.

O Capítulo 2 apresenta as funções matemáticas especias da família Hipergeométrica que serão utilizadas para cálculo de momentos e funções características.

O Capítulo 3 trata da distribuição assimétrica e suas propriedades. O Capítulo 4 trata da distribuição simétrica e suas propriedades.

O Capítulo 5 apresenta versão fortemente assimétrica gerada a partir da versão simétrica por meio de resultados encontrados em [Azzalini, 1985] (1) e [Fernandez et al., 1998](7).

O Capítulo 6 trata de aplicações com dados reais onde serão ajustadas as distribuições com estimação de máxima verossimilhança dos parâmetros.

Capítulo 2

Funções Matemáticas Especiais

2.1 Introdução

Serão apresentadas a função H(z) e a função Hipergeométrica Generalizada $({}_pF_q)$, pois alguns resultados para momentos e funções características das distribuições que serão estudadas adiante utilizam estes tipos de funções. A vantagem de utilizar estas funções é que muitos cálculos podem ser realizados por meio de implementações presentes em *softwares* matemáticos como o *Maple* e o *Mathematica*.

2.2 A Função H(z)

2.2.1 Conceitos e fundamentos - [SPRINGER,1979] (18)

Definição 2.2.1. *A Função $H(z)$ é definida por um tipo de integral de Mellin-Barnes*

$$H(z) = \mathbb{H}_{p,q}^{m,n}\left[z \left| \begin{array}{ccc} (a_1,\alpha_1), & \ldots, & (a_p,\alpha_p) \\ (b_1,\beta_1), & \ldots, & (b_q,\beta_q) \end{array} \right. \right] =$$

$$\frac{1}{2\pi i}\int_C \frac{\prod_{j=1}^{m}\Gamma(b_j-\beta_j s)\prod_{j=1}^{n}\Gamma(1-a_j+\alpha_j s)}{\prod_{j=m+1}^{q}\Gamma(1-b_j+\beta_j s)\prod_{j=n+1}^{p}\Gamma(a_j-\alpha_j s)} z^s ds, \tag{2.1}$$

onde

$0 \leq m \leq q$,

$0 \leq n \leq p$,

$m \neq 0$ *ou* $n \neq 0$,

$\alpha_j > 0, j = 1, 2, \ldots, p$,

$\beta_j > 0, j = 1, 2, \ldots, q$,

a_j *,com* $j = 1, 2, \ldots, p$, *e* b_j *, com* $j = 1, 2, \ldots, q$, *são números complexos tais que nenhum dos pólos de* $\Gamma(b_j - \beta_j s), j = 1, \ldots, m$, *coincidam com algum dos pólos de* $\Gamma(1 - a_j + \alpha_j s), j = 1, \ldots, n$. *Também, C é um contorno no plano complexo percorrendo* $\omega - i\infty \;\; a \;\; \omega + i\infty$, $i = \sqrt{-1}$, *para algum* ω *real tal que*

$$s = \frac{b_j + k}{\beta_j},$$

para $j = 1, \ldots, m$ *e* $k = 0, 1, \ldots$ *para o lado direito do contorno C e os pontos*

$$s = \frac{a_j - 1 - k}{\alpha_j},$$

para $j = 1, \ldots, n$ *e* $k = 0, 1, \ldots$ *para o lado esquerdo do contorno C com indentações, se for necessário para tal separação.*

O contorno C possui as seguintes definições:

i) Para $C = C_{-\infty}$ *que engloba estritamente os pólos de* $\Gamma(b_j - \beta_j s), j = 1, \ldots, m$, *a integral converge* $\forall z$ *se* $\mu > 0$ *e* $z \neq 0$ *ou* $\mu = 0$ *e* $0 < z < \eta$. *A integral também converge se* $\mu = 0, \; |z| = \eta$ *e* $Re(\delta) < -1$, *onde*

$$\eta = \left\{ \prod_{j=1}^{p} (\alpha_j)^{-\alpha_j} \right\} \left\{ \prod_{j=1}^{q} (\beta_j)^{\beta_j} \right\},$$

$$\mu = \sum_{j=1}^{q} \beta_j - \sum_{j=1}^{p} \alpha_j, \; e$$

$$\delta = \sum_{j=1}^{q} b_j - \sum_{j=1}^{p} a_j + \frac{p - q}{2}$$

ii) Para $C = C_{\infty}$ *que engloba estritamente os pólos de* $\Gamma(1 - a_j - \alpha_j s), j = 1, \ldots, n$, *a integral converge* $\forall z$ *se* $\mu < 0$ *e* $z \neq 0$ *ou* $\mu = 0$ *e* $|z| > \eta$.

iii) Para $C = C_{i\omega\infty}$ é um contorno iniciando no ponto $\omega - i\infty$ passando por $\omega + i\infty$, onde $\omega \in \mathbb{R} = (-\infty, +\infty)$ tal que todos os pólos de $\Gamma(b_j - \beta_j s), j = 1, \ldots, m$ são separados dos pólos de $\Gamma(1 - a_j - \alpha_j s), j = 1, \ldots, n$, a integral converge se $\upsilon < 0, |arg(z)| < \frac{1}{2}\pi\upsilon, \upsilon \neq 0$. A integral também converge se $\upsilon = 0, \omega + Re(\delta) < -1, arg(z) = 0$, e $z \neq 0$, onde

$$\upsilon = \sum_{j=1}^{n} \alpha_j - \sum_{j=n+1}^{p} \alpha_j + \sum_{j=1}^{m} \beta_j - \sum_{j=m+1}^{q} \beta_j$$

.

2.3 A Função Generalizada de Gauss ${}_pF_q$

2.3.1 Conceitos e fundamentos - [SLATER,1966] (17)

Definição 2.3.1. *A Função Generalizada de Gauss, denotada por ${}_pF_q$, é uma série de potências da forma*

$$1 + \frac{a_1 a_2 \ldots a_p}{b_1 b_2 \ldots b_q}\frac{z}{1!} + \frac{a_1(a_1+1)a_2(a_2+1)\ldots a_p(a_p+1)}{b_1(b_1+1)b_2(b_2+1)\ldots b_q(b_q+1)}\frac{z^2}{2!} + \ldots$$

$$\equiv \sum_{n=0}^{\infty} \frac{(a_1)_n (a_2)_n \ldots (a_p)_n}{(b_1)_n (b_2)_n \ldots (b_q)_n}\frac{z^n}{n!}, \tag{2.2}$$

onde $a_1, a_2, \ldots, a_q$ são parâmetros do numerador e $b_1, b_2, \ldots, b_q$ são parâmetros do denominador que não podem ser inteiros negativos, pois a série não é definida para estes valores, z é a variável e a e z podem assumir valores reais ou complexos.

Em particular, a função generalizada de Gauss é um caso especial da função H:

$${}_p\mathbb{F}_q\left[\begin{array}{c|c} a_1, \ \ldots, \ a_p & \\ b_1, \ \ldots, \ b_q & x \end{array}\right] =$$

$$\mathbb{H}^{1,p}_{p,q+1}\left[-x \left| \begin{array}{cccc} (1-a_1,1), & \ldots, & (1-a_p,1), & \\ (0,1), & (1-b_1,1), & \ldots, & (1-b_q,1) \end{array} \right.\right] \tag{2.3}$$

2.3.2 A convergência da Função Generalizada de Gauss ${}_pF_q$

A série ${}_pF_q$ converge para todos os valores de z real ou complexo quando $p \leq q$, se $u_n z^n$ é o n-ésimo termo de série

$$\left|\frac{u_{n+1}}{u_n}\right| \leq \frac{|a_1+n||a_2+n|\ldots|a_p+n||z|}{|b_1+n||b_2+n|\ldots|b_q+n|(1+n)}$$

$$\leq \frac{|z|n^{p-q-1}(1+|a_1|/n)(1+|a_2|/n)\ldots(1+|a_p|/n)}{(1+1/n)(1+|b_1|/n)(1+|b_2|/n)\ldots(1+|b_q|/n}, \tag{2.4}$$

e (2.4) tende a zero, quando $n \to \infty$, para $p \leq q$. Também, se $p = q+1$, a série converge desde que $|z| < 1$. Também converge quando $z = 1$, se $Re(\sum_{j=1}^{q} b_j - \sum_{j=1}^{p} a_j) > 0$, e $z = -1$, se $Re(\sum_{j=1}^{q} b_j - \sum_{j=1}^{p} a_j) > -1$.Se $p > q+1$, a série nunca converge, exceto para $z = 0$, e a função é definida apenas quando a série termina, isto é, quando um ou mais dos a parâmetros é zero ou um inteiro negativo.

2.3.3 A Função Hipergeométrica de Gauss ${}_2F_1$

Um caso em especial da função hipergeométrica, será do interesse de nossos estudos, ${}_2F_1$, também conhecida como série de Gauss.

Definição 2.3.2. *A série*

$$1 + \frac{ab}{c}\frac{z}{1!} + \frac{a(a+1)b(b+1)}{c(c+1)}\frac{z^2}{2!} + \frac{a(a+1)(a+2)b(b+1)(b+2)}{c(c+1)(c+2)}\frac{z^3}{3!} + \ldots \tag{2.5}$$

é simbolizada por ${}_2F_1[a;b;c;z]$, onde a, b e c são parâmetros e z a variável. Se a ou b for um inteiro negativo, a série tem um número finito de termos, e torna-se, de fato, um polinômio.

2.3.4 Convergência da Função de Gauss ${}_2F_1$

A série de Gauss, ${}_2F_1$, pode ser denotada por

$${}_2F_1[a;b;c;z] = \sum_{n=1}^{\infty} \frac{(a)_n(b)_n}{(c)_n n!}, \ \textit{onde} \tag{2.6}$$

$$(a)_n = a(a+1)(a+2)(a+3)\ldots(a+n-1). \tag{2.7}$$

Por exemplo, $(2)_5 = 2.3.4.5.6 = 720$. Em particular, $(a)_0 \equiv 1$ e $(1)_n = n!$. Logo,

$$(a)_n = \frac{\Gamma(a+n)}{\Gamma(a)}\ e \tag{2.8}$$

$$\lim_{n\to\infty}(a)_n = \frac{1}{\Gamma(a)}. \tag{2.9}$$

Se a é um inteiro negativo $-m$, então

$$(a)_n = (-m)_n,\ if\ m \geq n$$

e

$$(a)_n = 0,\ if\ m < n$$

.

Por exemplo, $(-2)_2 = (-2).(-1) = 2$, entretanto, $(-2)_5 = 0$.

Então em ${}_2F_1[a;b;c;z]$, se a ou b for um inteiro negativo, a série tem um número finito de termos, e torna-se, de fato, um polinômio, mas se c é um inteiro negativo, a função não é definida, uma vez que apenas um número finito de termos da série torna-se infinito. Temos que

$$\frac{d}{dz}({}_2F_1[a;b;c;z]) = \frac{ab}{c}\ {}_2F_1[a+1;b+1;c+1;z]. \tag{2.10}$$

2.3.5 Relações Funcionais de ${}_2F_1[a;b;c;z]$

Segundo [Temme, 1996] (19), a função hipergeométrica satisfaz às seguintes relações:

$$ {}_2F_1[a;b;c;z] = {}_2F_1[b;a;c;z] \tag{2.11}$$

$$ {}_2F_1[a;b;c;z] = (1-z)^{-a}\ {}_2F_1\left[a;c-b;c;\frac{z}{z-1}\right] \tag{2.12}$$

$$ {}_2F_1[a;b;c;z] = (1-z)^{-b}\ {}_2F_1\left[c-a;b;c;\frac{z}{z-1}\right] \tag{2.13}$$

$$ {}_2F_1[a;b;c;z] = (1-z)^{c-a-b}\ {}_2F_1\left[c-a;c-b;c;z\right] \tag{2.14}$$

Seja a forma abaixo a Integral de Contorno de ${}_2F_1[a;b;c;z]$. Temos que

$$ {}_2F_1[a;b;c;z] = \frac{\Gamma(c)}{\Gamma(a)\Gamma(b)}\frac{1}{2\pi i}\int_{\gamma-i\infty}^{\gamma+i\infty}\frac{\Gamma(-s)\Gamma(a+s)\Gamma(b+s)}{\Gamma(c+s)}(-z)^s ds \tag{2.15}$$

ou

$$ {}_2F_1[a;b;c;z] = \frac{\Gamma(c)}{\Gamma(a)\Gamma(b)}\frac{1}{2\pi i}\int_{-\gamma-i\infty}^{-\gamma+i\infty}\frac{\Gamma(s)\Gamma(a-s)\Gamma(b-s)}{\Gamma(c-s)}(-z)^{-s} ds \tag{2.16}$$

2.4 A Função G de Meijer

Definição 2.4.1. *A Função G de Meijer é definida pela integral*

$$ G(z) = \mathbb{G}_{p,q}^{m,n}\left[z\left|\begin{array}{ccc} a_1, & \ldots, & a_p \\ b_1, & \ldots, & b_q \end{array}\right.\right] =$$

$$\frac{1}{2\pi i}\int_C \frac{\prod_{j=1}^{m}\Gamma(b_j-s)\prod_{j=1}^{n}\Gamma(1-a_j+s)}{\prod_{j=m+1}^{q}\Gamma(1-b_j+s)\prod_{j=n+1}^{p}\Gamma(a_j-s)}z^s ds, \tag{2.17}$$

onde

$0 \leq m \leq q$,

$0 \leq n \leq p$,

$m \neq 0$ *ou* $n \neq 0$,

a_j *,com* $j = 1, 2, \ldots, p$, *e* b_j *, com* $j = 1, 2, \ldots, q$, *são números complexos tais que nenhum dos pólos de* $\Gamma(b_j - s), j = 1, \ldots, m$, *coincidam com algum dos pólos de* $\Gamma(1-a_j+s), j = 1, \ldots, n$. *Também, C é um contorno no plano complexo percorrendo* $\omega - i\infty$ *a* $\omega + i\infty$ *pra algum* ω *real tal que*

$$s = b_j + k,$$

para $j = 1, \ldots, m$ *e* $k = 0, 1, \ldots$ *para o lado direito do contorno C e os pontos*

$$s = a_j - 1 - k,$$

para $j = 1, \ldots, n$ *e* $k = 0, 1, \ldots$ *para o lado esquerdo do contorno C com indentações, se for necessário para tal separação. A integral (2.17) é chamada Mellin-Barnes e pode ser vista como uma Transformação Inversa de Mellin.*

A função G de Meijer é computável e está implementada em *softwares* matemáticos como o *Maple* e o *Mathematica*.

Em particular, a função G de Meijer é um caso especial da função H:

$$\mathbb{G}_{p,q}^{m,n}\left[\left.\begin{array}{cccc} a_1, & \ldots, & a_p \\ b_1, & \ldots, & b_q \end{array}\right| z\right] =$$

$$\mathbb{H}_{p,q}^{m,n}\left[z\left|\begin{array}{cccc} (a_1,1), & \ldots, & (a_p,1), \\ (b_1,1), & \ldots, & (b_q,1) \end{array}\right.\right] \tag{2.18}$$

Se os parâmetros $\alpha_j, j = 1, \ldots, p$ e $\beta_j, j = 1, \ldots, q$ da função Hipergeométrica Generalizada são racionais, a função H pode ser escrita como função G.

2.5 A Função Bivariada $H[x, y]$

2.5.1 Conceitos e fundamentos - [MATHAI et al.,2010] (9)

Definição 2.5.1. *A Função Hipergeométrica Generalizada Bivariada $H[x, y]$ é definida pela integral complexa*

$$H\left[\begin{array}{c} x \\ y \end{array}\right] = \mathbb{H}_{p_1,q_1:p_2,q_2;p_3,q_3}^{0,n_1:m_2,n_2;m_3,n_3}\left[\begin{array}{c|c} x & (a_i,\alpha_i,A_i)_{1,p_1} : (c_i,\gamma_i)_{1,p_2};(e_i,E_i)_{1,p_3} \\ y & (b_j,\beta_j,B_j)_{1,q_1} : (d_j,\delta_j)_{1,q_2};(f_j,F_j)_{1,q_3} \end{array}\right] = \tag{2.19}$$

$$-\frac{1}{4\pi^2}\int_{C_1}\int_{C_2}\phi(s,t)\phi_1(s)\phi_2(t)x^s y^t ds dt, \tag{2.20}$$

onde x e y são variáveis não-nulas. A notação $(a_i, \alpha_i, A_i)_{1,p_1}$ é a abreviação da sequência de parâmetros $(a_1, \alpha_1, A_1), \ldots, (a_{p_1}, \alpha_{p_1}, A_{p_1})$. De forma similar, deve ser feito com as outras sequências de parâmetros.Onde

$$\phi(s,t) = \frac{\prod_{i=1}^{n_1}\Gamma(1-a_i+\alpha_i s+A_i t)}{\prod_{i=n_1+1}^{p_1}\Gamma(a_i-\alpha_i s-A_i t)\prod_{j=1}^{q_1}\Gamma(1-b_j+\beta_j s+B_j t)}, \tag{2.21}$$

$$\phi_1(s) = \frac{\prod_{j=1}^{m_2}\Gamma(d_j-\delta_j s)\prod_{i=1}^{n_2}\Gamma(1-c_i+\gamma_i s)}{\prod_{j=m_2+1}^{q_2}\Gamma(1-d_j+\delta_j s)\prod_{i=n_2+1}^{p_2}\Gamma(c_i-\gamma_i s)}, \tag{2.22}$$

$$\phi_2(t) = \frac{\prod_{j=1}^{m_3}\Gamma(f_j-F_j t)\prod_{i=1}^{n_3}\Gamma(1-e_i+E_i t)}{\prod_{j=m_3+1}^{q_3}\Gamma(1-f_j+F_j t)\prod_{i=n_3+1}^{p_3}\Gamma(e_i-E_i t)}. \tag{2.23}$$

Assume-se que $a_i, b_j, c_i, d_j, e_i, f_j$ são números complexos e seus respectivos coeficientes $\alpha_i, A_i, \beta_j, B_j, \gamma_i, \delta_j, Ei, F_j$ são números reais e positivos, tais que os contornos C_1, C_2 existem se

$$\rho_1 = \sum_{i=1}^{p_1} \alpha_i + \sum_{i=1}^{p_2} \gamma_i - \sum_{j=1}^{q_1} \beta_j - \sum_{j=1}^{q_2} \delta_j \leq 0,$$

$$\rho_2 = \sum_{i=1}^{p_1} A_i + \sum_{i=1}^{p_2} E_i - \sum_{j=1}^{q_1} B_j - \sum_{j=1}^{q_2} F_j \leq 0,$$

$$\Omega_1 = -\sum_{i=n_1+1}^{p_1} \alpha_i - \sum_{j=1}^{q_1} \beta_j + \sum_{j=1}^{m_2} \delta_j - \sum_{j=m_2+1}^{p_2} \delta_j + \sum_{i=1}^{n_2} \gamma_i - \sum_{i=n_2+1}^{p_2} \gamma_i > 0,$$

$$\Omega_2 = -\sum_{i=n_1+1}^{p_1} A_i - \sum_{j=1}^{q_1} B_j + \sum_{j=1}^{m_3} F_j - \sum_{j=m_3+1}^{p_3} F_j + \sum_{i=1}^{n_3} E_i - \sum_{i=n_3+1}^{p_3} E_i > 0.$$

Sob as condições supracitadas, a integral dupla de Mellin-Barnes (2.20) converge e define uma função analítica de duas variáveis x e y dentro do setor dado por:

$$|arg(x)| < \frac{1}{2}\pi\Omega_1$$

$$|arg(y)| < \frac{1}{2}\pi\Omega_2$$

Mais informações podem ser encontradas em [Mathai et al.,2010] (9).

2.6 Caso Especial da Função Bivariada $H[x, y]$

2.6.1 Conceitos e fundamentos - [HAI & YAKUBOVICH, 1992] (8)

Definição 2.6.1. *Um caso especial da Função Hipergeométrica Generalizada Bivariada $H[x, y]$ é definida pela integral complexa a seguir*

$$H\left[\begin{array}{c} x \\ y \end{array}\right] = \mathbb{H}^{0,n_1:m_2,n_2;m_3,n_3}_{p_1,q_1:p_2,q_2;p_3,q_3}\left[\begin{array}{c|c} x & (\alpha^{(1)}_{p_1}, a^{(1)}_{p_1}) : (\alpha^{(2)}_{p_2}, a^{(2)}_{p_2}); (\alpha^{(3)}_{p_3}, a^{(3)}_{p_3}) \\ y & (\beta^{(1)}_{q_1}, b^{(1)}_{q_1}) : (\beta^{(2)}_{q_2}, b^{(2)}_{q_2}); (\beta^{(3)}_{q_3}, b^{(3)}_{q_3}) \end{array}\right] = \tag{2.24}$$

$$-\frac{1}{4\pi^2}\int_{C_1}\int_{C_2}\phi_1(s+t)\phi_2(s)\phi_3(t)x^{-s}y^{-t}dsdt, \tag{2.25}$$

onde x e y são variáveis não-nulas.Onde

$$\phi_1(s+t) = \frac{\prod_{j=1}^{n_1}\Gamma(1-\alpha^{(1)}_j - a^{(1)}_j(s+t))}{\prod_{j=n_1+1}^{p_1}\Gamma(\alpha^{(1)}_j + a^{(1)}_j(s+t))\prod_{j=1}^{q_1}\Gamma(1-\beta^{(1)}_j - b^{(1)}_j(s+t))}, \tag{2.26}$$

$$\phi_2(s) = \frac{\prod_{j=1}^{m_2}\Gamma(\beta^{(2)}_j - b^{(2)}_j s)\prod_{j=1}^{n_2}\Gamma(1-\alpha^{(2)}_j - a^{(2)}_j s)}{\prod_{j=m_2+1}^{p_2}\Gamma(\alpha^{(2)}_j + a^{(2)}_j s)\prod_{j=n_2+1}^{q_2}\Gamma(1-\beta^{(2)}_j - b^{(2)}_j s)}, \tag{2.27}$$

$$\phi_3(t) = \frac{\prod_{j=1}^{m_3}\Gamma(\beta^{(3)}_j - b^{(3)}_j t)\prod_{j=1}^{n_3}\Gamma(1-\alpha^{(3)}_j - a^{(3)}_j t)}{\prod_{j=m_3+1}^{p_3}\Gamma(\alpha^{(3)}_j + a^{(3)}_j t)\prod_{j=n_3+1}^{q_3}\Gamma(1-\beta^{(3)}_j - b^{(3)}_j t)}, \tag{2.28}$$

Assume-se que a_i, b_j são números reais e positivos. Mais informações sobre convergência podem ser encontradas em [HAI & YAKUBOVICH, 1992] (8).

Capítulo 3

Uma Nova Distribuição

3.1 Introdução

A distribuição que será apresentada a seguir tem como característica assimetria à direita acentuada. Como possui diversos parâmetros, a distribuição pode assumir várias formas, o que facilita o uso em diversos dados onde se queira modelar v.a's positivas.

3.2 Distribuição Acumulada Assimétrica

Definição 3.2.1. *Seja a forma*

$$F(x) = \frac{a^{\rho}x^{2\rho}}{(ax^2 + 2bx + c)^{\rho}} \tag{3.1}$$

a função de distribuição acumulada de uma v.a. X onde $x > 0$, $a > 0$, $c > 0$ e $b^2 < ac$, e $\rho > 0$.

A distribuição acumulada acima possui quatro parâmetros e pode ser reescrita como uma função de três parâmetros se o fator a^{ρ} no denominador (3.1) for colocado em evidência, então assumiremos que a função de distribuição acumulada assimétrica de interesse de nossos estudos é da forma abaixo.

Definição 3.2.2. *Seja a forma*

$$F(x) = \frac{x^{2\rho}}{(x^2 + 2Bx + C)^{\rho}} \tag{3.2}$$

a função de distribuição acumulada assimétrica de uma v.a. X onde $B = \frac{b}{a}$ *e* $C = \frac{c}{a}$, $x > 0$, $C > 0$, $B^2 < C$, *e* $\rho > 0$.

Uma vez que $F(x)$ existe, então encontraremos $f(x) = \frac{dF(x)}{dx}$ a seguir:

$$f(x) = \frac{d}{dx}(\frac{x^{2\rho}}{(x^2 + 2Bx + C)^{\rho}})$$

$$= \frac{(2\rho x^{2\rho-1})(x^2 + 2Bx + C)^{\rho} - ((x^{2\rho}\rho)(x^2 + 2Bx + C)^{\rho-1}(2x + 2B))}{(x^2 + 2Bx + C)^{2\rho}}$$

$$= \frac{2\rho x^{2\rho+1} + 4B^{2\rho} + 2C\rho x^{2\rho-1} - 2\rho x^{2\rho+1} - 2B\rho x^{2\rho}}{(x^2 + 2Bx + c)^{\rho+1}}$$

$$= \frac{2B\rho x^{2\rho} + 2C\rho x^{2\rho-1}}{(x^2 + 2Bx + C)^{\rho+1}}$$

$$= \frac{2\rho x^{2\rho-1}(Bx + C)}{(x^2 + 2Bx + C)^{\rho+1}}$$

3.2.1 Propriedades da Distribuição Acumulada

Vamos provar que $F(x)$ é função de distribuição acumulada:

i) $\lim_{x\to 0} F(x) = 0$

Demonstração:

$\lim_{x\to 0} F(x) = \lim_{x\to 0} \frac{x^{2\rho}}{(x^2+2Bx+C)^{\rho}} = 0$

ii) $\lim_{x\to \infty} F(x) = 1$

Demonstração:

$\lim_{x\to+\infty} F(x) = \lim_{x\to+\infty} \frac{x^{2\rho}}{(x^2+2Bx+C)^\rho} = \lim_{x\to+\infty}(\frac{1}{1+\frac{2B}{x}+\frac{C}{x^2}})^\rho = 1$

$iii) F(x) \leq F(y), \forall x < y; x, y \in \mathbb{R}$

Demonstração: Como $f(x) > 0, \forall x \in (0, \infty)$, então $F(x)$ é função não-decrescente, $\forall x \in (0, \infty)$.

$iv)\ 0 \leq F_X(x) \leq 1, \forall x \in (0, \infty)$;

Demonstração: De fato, pelas propriedades $i) - iii)$ conclui-se que $iv)$ também vale.

Para os valores fornecidos de $F(x)$, x pode ser obtido pela seguinte expressão:

$$x = \frac{BF^{1/\rho} + [B^2F^{2/\rho} + CF^{1/\rho}(1 - F^{1/\rho})]^{1/2}}{(1 - F^{1/\rho})}$$

Isto significa que F é inversível e os quantis podem ser obtidos facilmente.

3.2.2 Gráficos para a Distribuição Acumulada

Como provamos que $F(x)$ é distribuição, então apresentaremos algumas formas da distribuição, para diversos valores dos parâmetros B, C, ρ.

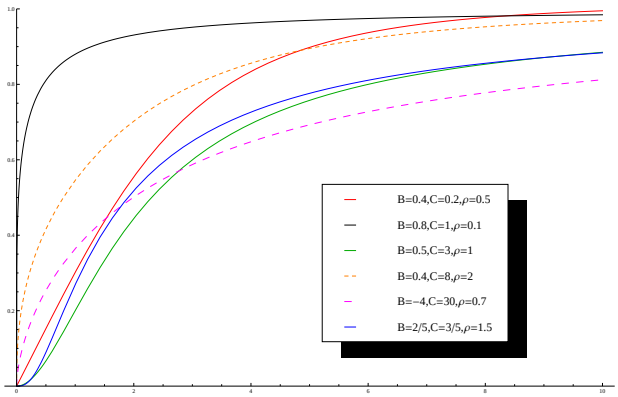

Figura 3.1: Formas para a Distribuição Acumulada Assimétrica $F(x)$

Para verificar o efeito de cada um dos parâmetros na curva, produzimos gráficos onde se fixa dois parâmetros enquanto um terceiro varia.

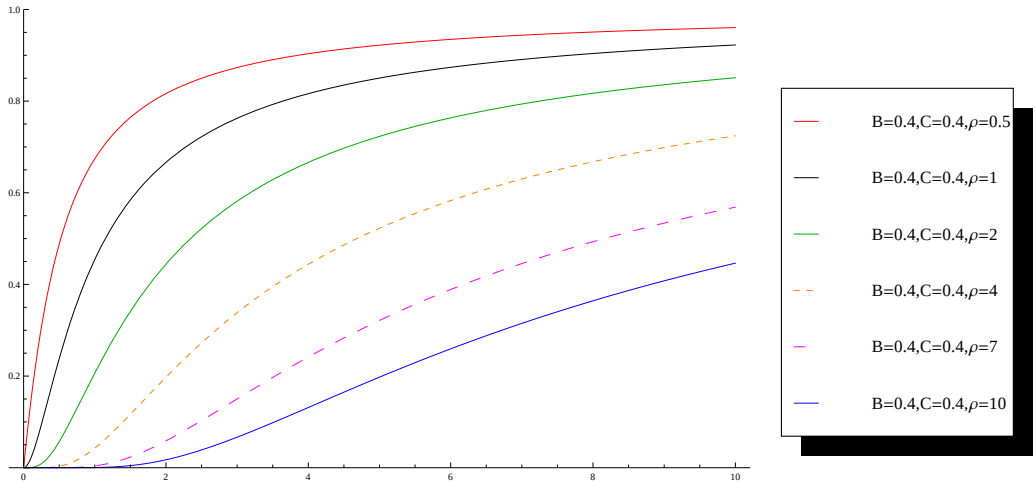

Figura 3.2: Formas para a Distribuição Acumulada $F(x)$ - variação de ρ, com $B > 0$

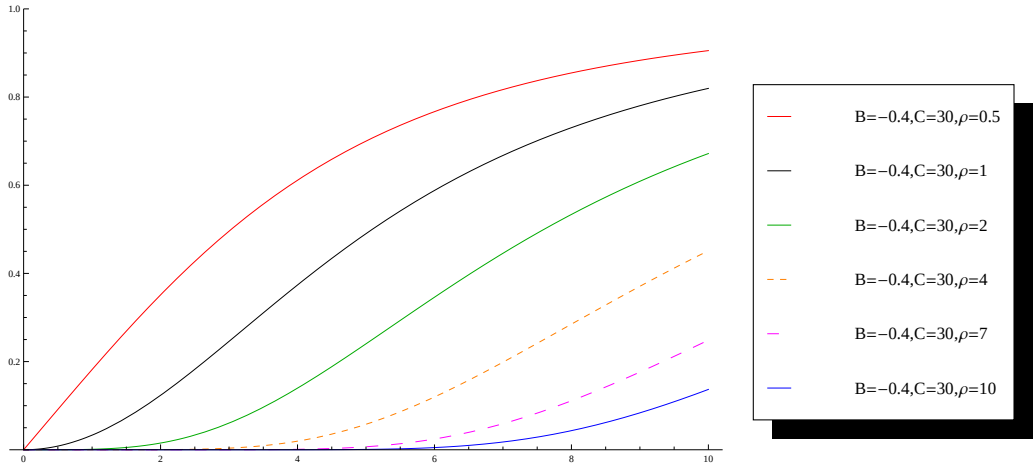

Figura 3.3: Formas para a Distribuição Acumulada $F(x)$ - variação de ρ, com $B < 0$

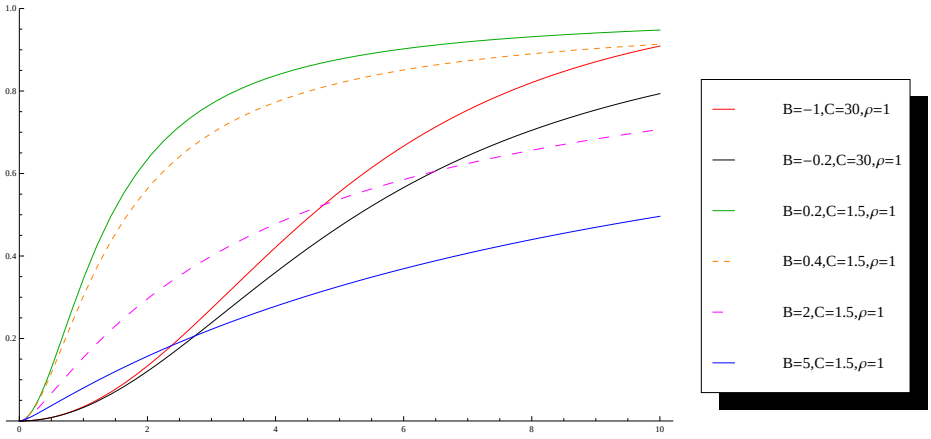

Figura 3.4: Formas para a Distribuição Acumulada $F(x)$ - variação de B

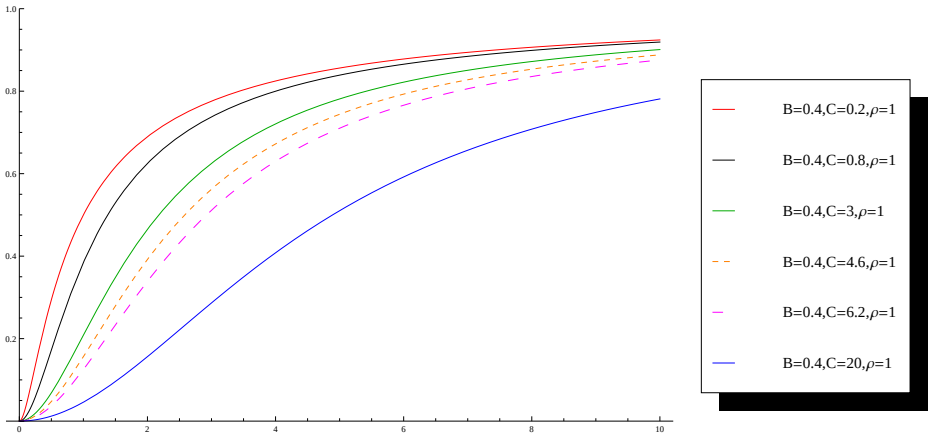

Figura 3.5: Formas para a Distribuição Acumulada $F(x)$ - variação de C, com $B > 0$

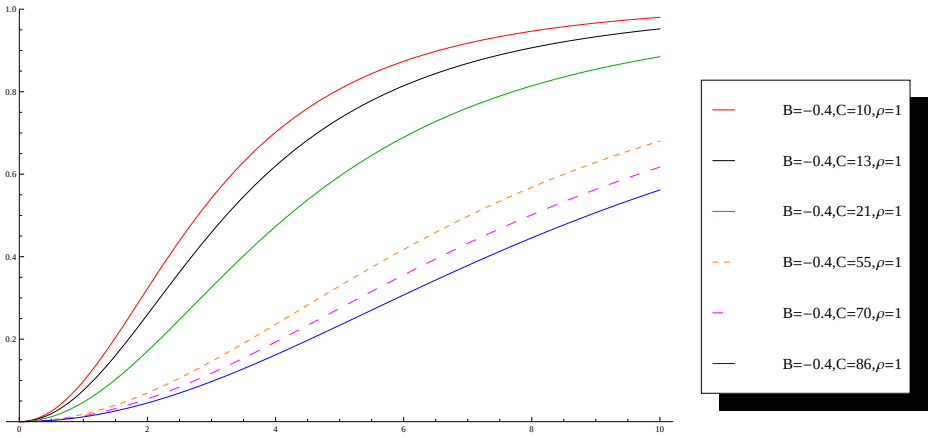

Figura 3.6: Formas para a Distribuição Acumulada $F(x)$ - variação de C, com $B < 0$

3.2.3 Função Densidade de Probabilidade

Definição 3.2.3. *Seja a forma*

$$f(x) = \frac{2\rho x^{2\rho-1}(Bx + C)}{(x^2 + 2Bx + C)^{\rho+1}} \tag{3.3}$$

a função densidade de probabilidade de uma v.a. X onde $x > 0$, $C > 0$, $B^2 < C$ e $\rho > 0$.

3.2.4 Gráficos para a Função Densidade de Probabilidade

A seguir apresentaremos algumas formas da f.d.p., para diversos valores dos parâmetros B, C, ρ:

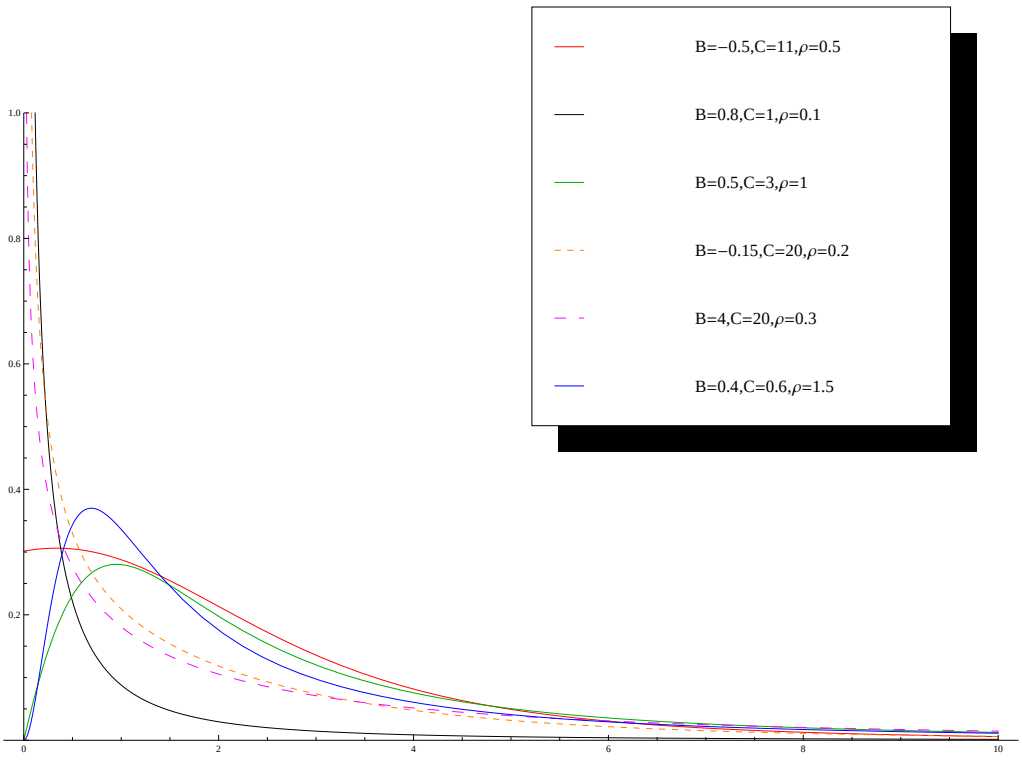

Figura 3.7: Formas para a Densidade de Probabilidade Assimétrica $f(x)$

Observando a figura (3.7), podemos ver certa semelhança com a distribuição Gama Generalizada. A função densidade de probabilidade da uma v.a. X com distribuição Gama Generalizada é da forma

$$h(x) = \frac{\alpha\beta^{\frac{\alpha}{\gamma}}}{\Gamma(\frac{\alpha}{\gamma})} x^{\alpha-1} e^{-\beta x^{\gamma}}, \tag{3.4}$$

onde $x > 0$, $\beta > 0$ é parâmetro de escala, $\alpha, \gamma > 0$ são parâmetros de forma. O gráfico da f.d.p com a v.a. $X \sim \Gamma(\alpha, \beta, \gamma)$pode ser observado abaixo:

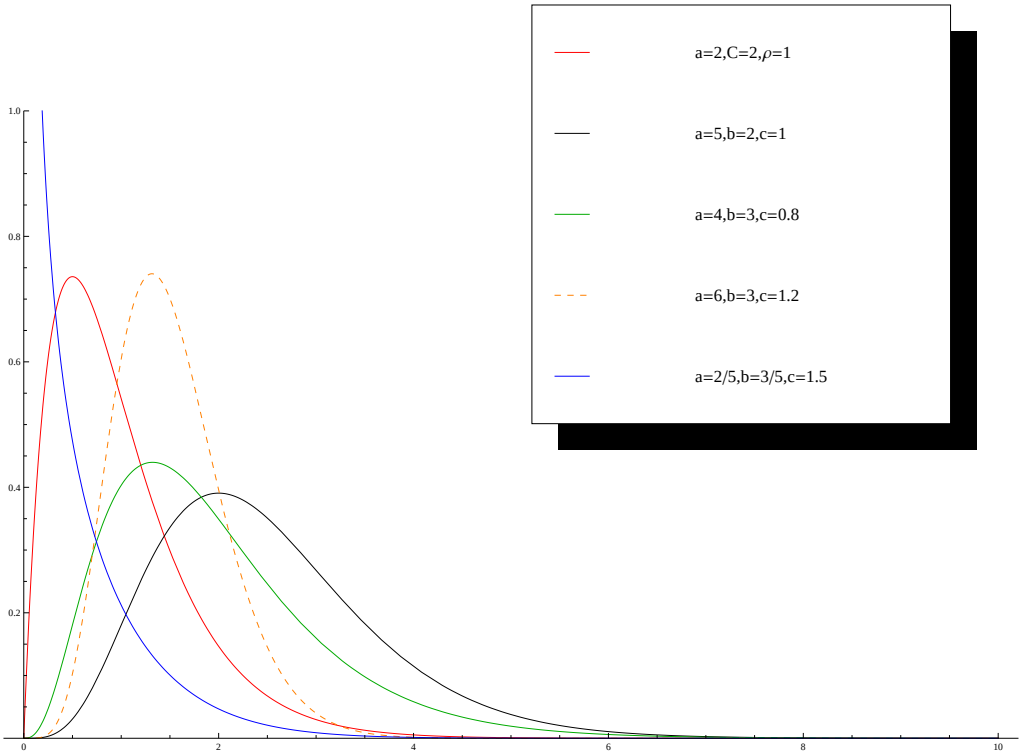

Figura 3.8: Formas para a densidade Gama Generalizada

Os gráficos das densidades presentes nas figuras (3.7) e (3.8) possuem diferenças nas caudas. No caso da nova densidade, ela pode assumir formas com caudas mais pesadas.

Para verificar o efeito de cada um dos parâmetros na curva, produzimos gráficos onde se fixa dois parâmetros enquanto o terceiro varia.

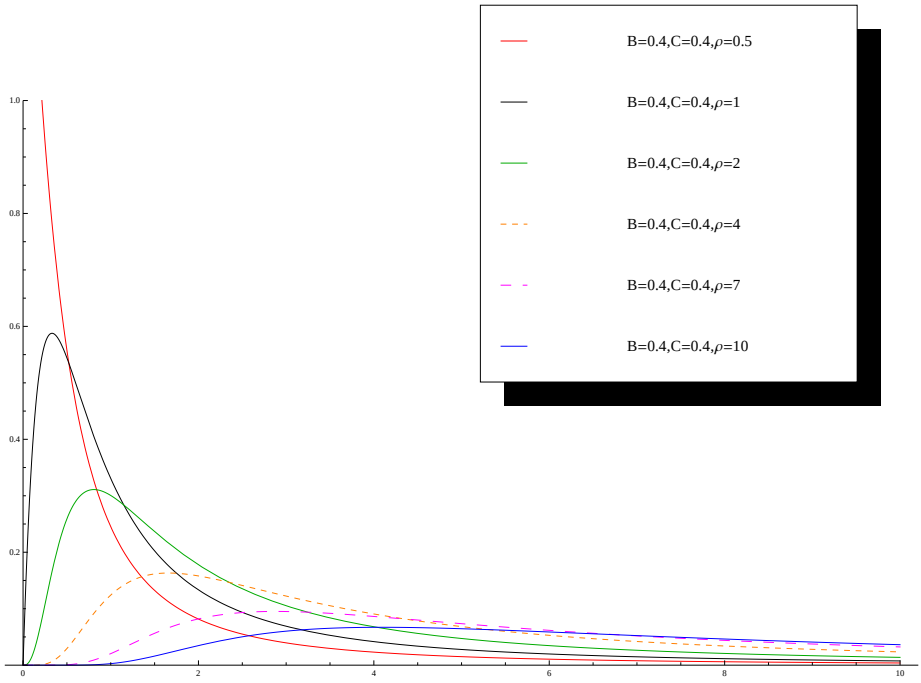

Figura 3.9: Formas para a Densidade de Probabilidade - variação de ρ, com $B > 0$

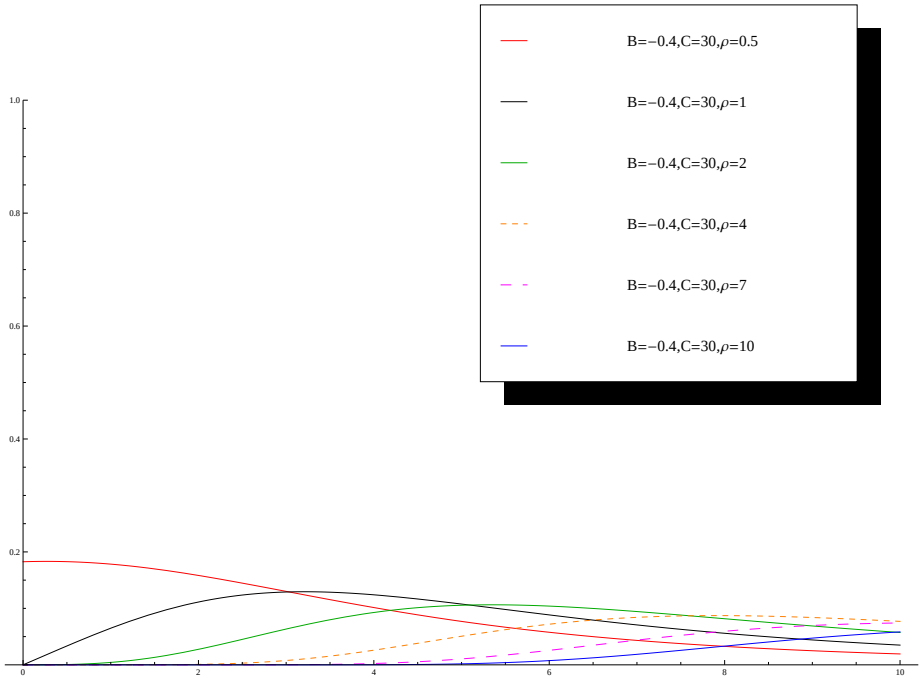

Figura 3.10: Formas para a Densidade de Probabilidade - variação de ρ, com $B < 0$

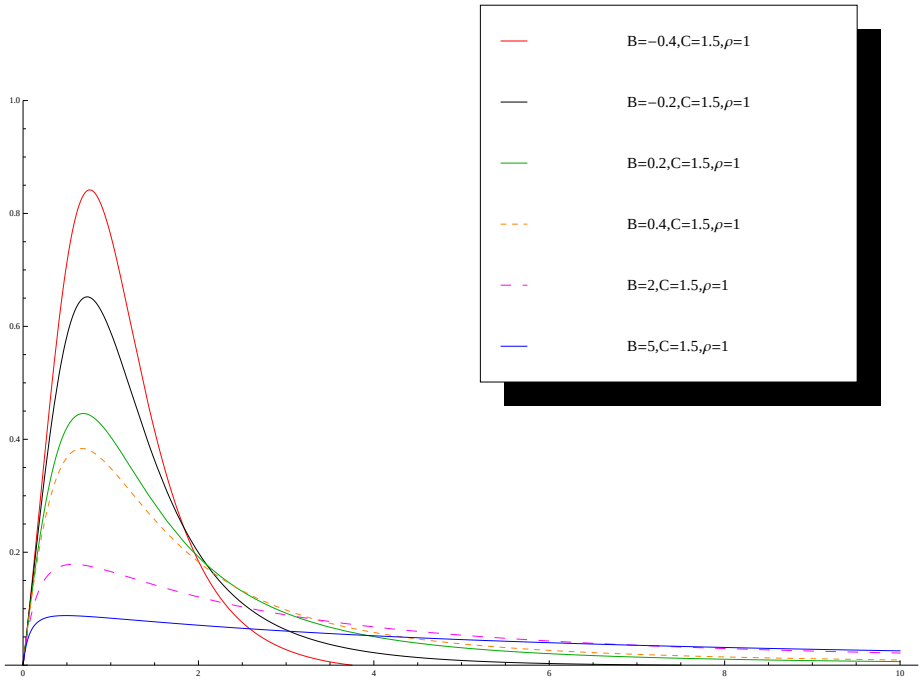

Figura 3.11: Formas para a Densidade de Probabilidade $f(x)$ - variação de B

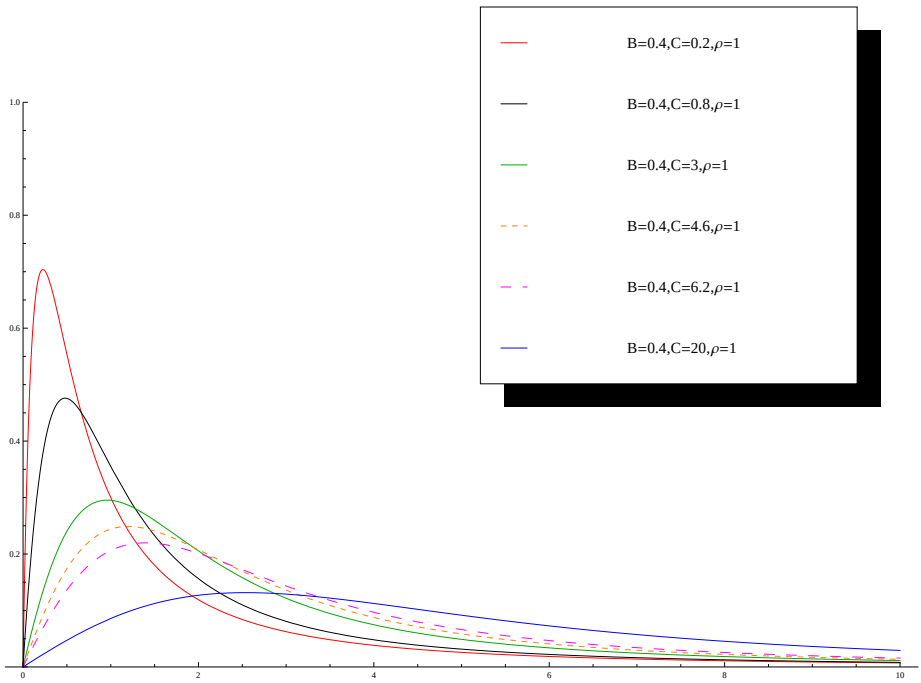

Figura 3.12: Formas para a Densidade de Probabilidade $f(x)$ - variação de C, com $B > 0$

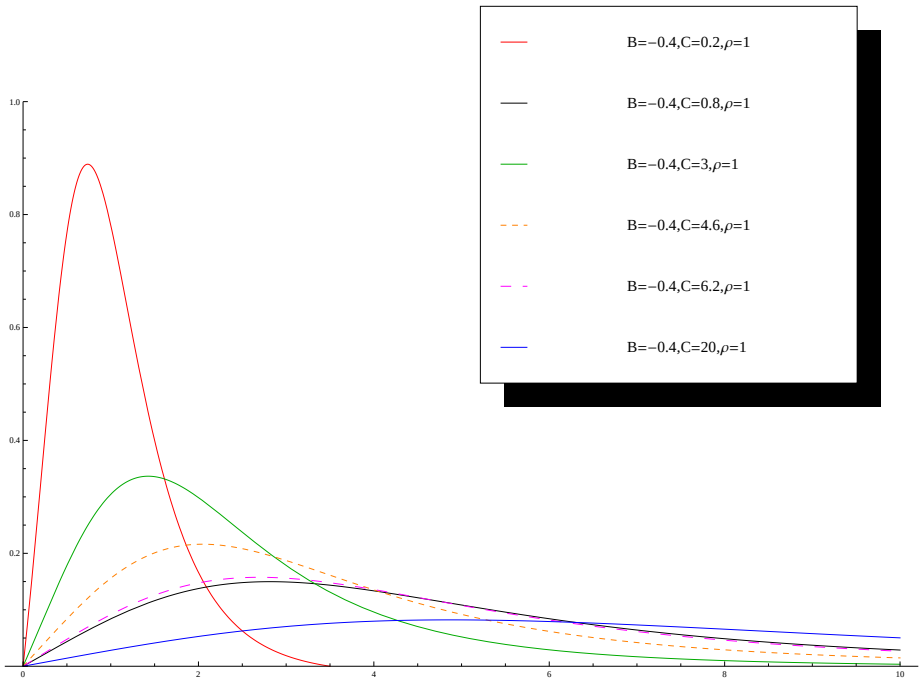

Figura 3.13: Formas para a Densidade de Probabilidade $f(x)$ - variação de C, com $B < 0$

Pela figura (3.9) observamos que o parâmetro ρ determina o comportamento da cauda da densidade. Quanto maior ρ, mais achatada é a curva e mais pesada é a cauda, para $B > 0$.

Pela figura (3.10) concluímos que o parâmetro ρ determina o comportamento da cauda da densidade. Mas com $B < 0$ associado a ρ, a curva fica mais achatada do que no caso em que $B > 0$.

Pela figura (3.11) verificamos que o parâmetro B determina o comportamento relacionado à assimetria. Quanto maior B, mais larga é a curva.

Pela figura (3.12) observamos que o parâmetro C determina o comportamento relacionado ao pico da função. Quanto maior C, menor o pico da curva, para $B > 0$.

Pela figura (3.13) concluímos que o parâmetro C determina o comportamento relacionado ao pico da função. Mas não existe a relação inversamente proporcional que encontramos quando $B > 0$.

Ao observarmos o comportamento da f.d.p. (3.3), quando $x \to \infty$, temos

$$\lim_{x\to\infty} f(x) \approx c\frac{x^{2\rho}}{x^{2\rho+2}} \approx \frac{c}{x^2},$$

onde c é uma constante, e possui gráfico com cauda assintótica:

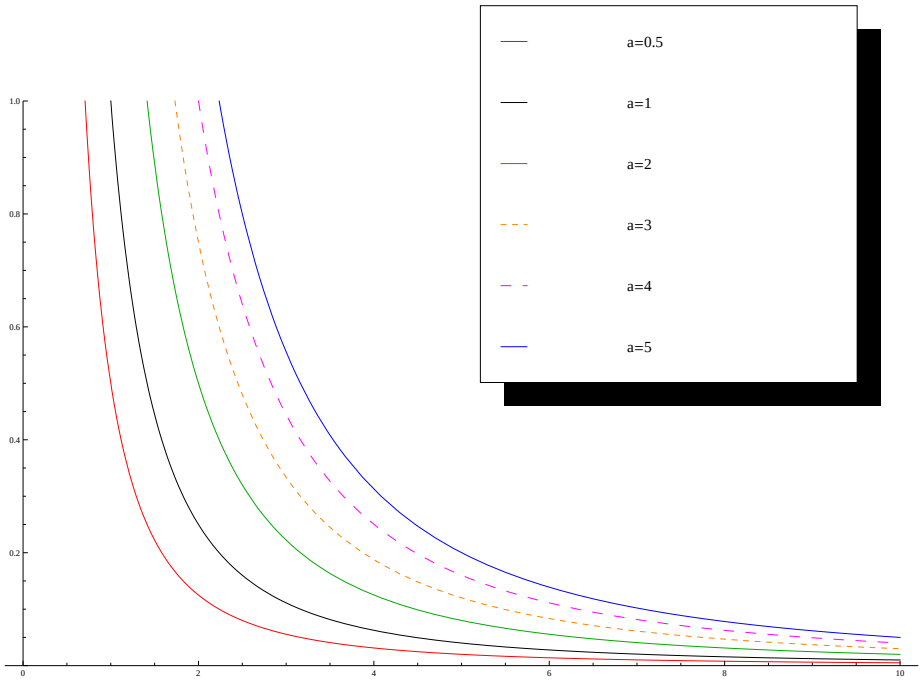

Figura 3.14: Comportamento da Densidade de Probabilidade $f(x)$ quando $x \to \infty$

3.2.5 Distribuições de Mínimo e Máximo

A distribuição e a densidade do Mínimo $X_{(1)} = min(X_1, \ldots, X_n)$ são dadas por:

$$P(X_{(1)} \leq z) = 1 - P(X_{(1)} > z) = 1 - P(min(X_1, \ldots, X_n) > z) = 1 - \prod_{i=1}^{n} [1 - F_{X_{(1)}}(z)]$$

$$F(z) = 1 - \left[1 - \frac{z^{2\rho}}{(z^2 + 2Bz + C)^{\rho}}\right]^n$$

$$f(z) = 2n\rho \frac{z^{2\rho-1}(Bz + C)}{(z^2 + 2Bz + C)^{\rho+1}} \left[1 - \frac{z^{2\rho}}{(z^2 + 2Bz + C)^{\rho}}\right]^{n-1}$$

A distribuição e a densidade do Máximo $X_{(n)} = max(X_1, \ldots, X_n)$ são dadas por:

$$P(X_{(n)} \leq w) = P(max(X_1, \ldots, X_n) \geq w) = \prod_{i=1}^{n} F_{X_{(1)}}(w)$$

$$F(w) = 1 - \left[\frac{w^{2\rho}}{(w^2 + 2Bw + C)^{\rho}}\right]^n$$

$$f(w) = 2n\rho \frac{w^{2\rho-1}(Bw + C)}{(w^2 + 2Bw + C)^{\rho+1}} \left[\frac{w^{2\rho}}{(w^2 + 2Bw + C)^{\rho}}\right]^{n-1}$$

3.2.6 Momentos para X

Calcularemos o k-ésimo momento μ'_k para a densidade definida em (3.3):

$$\mu'_k = EX^k = \int_0^\infty x^k dF(x) = \int_0^\infty x^k \frac{2\rho x^{2\rho-1}(Bx+C)}{(x^2+2Bx+C)^{\rho+1}} dx =$$

$$\int_0^\infty \frac{2\rho x^{k+2\rho-1}(Bx+C)}{(x^2+2Bx+C)^{\rho+1}} dx = \int_0^\infty \frac{2\rho B x^{k+2\rho}}{(x^2+2Bx+C)^{\rho+1}} dx + \int_0^\infty \frac{2\rho C x^{k+2\rho-1}}{(x^2+2Bx+C)^{\rho+1}} dx$$

Temos o seguinte resultado, encontrado em [PRUDNIKOV et al, 1986] (13):

$$\int_0^\infty \frac{x^{p-1}}{(ax^2+2bx+c)^q} dx = a^{-p/2} c^{p/2-q} B(p, 2q-p)\, {}_2F_1\left[\frac{p}{2}; q-\frac{p}{2}; q+\frac{1}{2}; 1-\frac{b^2}{ac}\right], \tag{3.5}$$

onde $B(a,b) = \frac{\Gamma(a)\Gamma(b)}{\Gamma(a+b)}$ para $a > 0, b^2 < ac, 0 < p < 2q$. Utilizando (2.13), obtivemos a seguinte expressão alternativa:

$$\int_0^\infty \frac{x^{p-1}}{(ax^2+2bx+c)^q} dx = a^{q-p} b^{p-2q} B(p, 2q-p)\, {}_2F_1\left[q-\frac{p}{2}+\frac{1}{2}; q-\frac{p}{2}; q+\frac{1}{2}; 1-\frac{ac}{b^2}\right]. \tag{3.6}$$

Usando o resultado (3.5) para resolver EX^k, temos:

$$EX^k = \underbrace{\int_0^\infty \frac{2\rho B x^{k+2\rho}}{(x^2+2Bx+C)^{\rho+1}} dx}_{I} + \underbrace{\int_0^\infty \frac{2\rho C x^{k+2\rho-1}}{(x^2+2Bx+C)^{\rho+1}} dx}_{II} =$$

i) Comparando (I) e a expressão (3.5), temos:

$$\begin{cases} p = k + 2\rho + 1 \\ q = \rho + 1 \end{cases}$$

Logo,

$$I = 2\rho B \left[C^{\frac{k-1}{2}} \frac{\Gamma(k+2\rho+1)\Gamma(1-k)}{\Gamma(2\rho+2)} \right] {}_2F_1 \left[\frac{k+2\rho+1}{2}; \frac{1-k}{2}; \rho + \frac{3}{2}; 1 - \frac{B^2}{C} \right]$$

ii) Comparando (II) e a expressão (3.5), temos:

$$\begin{cases} p = k + 2\rho \\ q = \rho + 1 \end{cases}$$

Logo,

$$II = 2\rho C \left[C^{\frac{k}{2}-1} \frac{\Gamma(k+2\rho)\Gamma(2-k)}{\Gamma(2\rho+2)} \right] {}_2F_1 \left[\rho + \frac{k}{2}; 1 - \frac{k}{2}; \rho + \frac{3}{2}; 1 - \frac{B^2}{C} \right]$$

Então,

$$\mu'_k = EX^k = (I) + (II)$$

$$= \Bigg\{ B \left[C^{\frac{k-1}{2}} \Gamma(k+2\rho+1)\Gamma(1-k) \right] {}_2F_1 \left[\frac{k+2\rho+1}{2}; \frac{1-k}{2}; \rho + \frac{3}{2}; 1 - \frac{B^2}{C} \right] +$$

$$\mathrm{C} \left[C^{\frac{k}{2}-1} \Gamma(k+2\rho)\Gamma(2-k) \right] {}_2F_1 \left[\rho + \frac{k}{2}; 1 - \frac{k}{2}; \rho + \frac{3}{2}; 1 - \frac{B^2}{C} \right] \Bigg\} \left[\frac{2\rho}{\Gamma(2\rho+2)} \right] \qquad (3.7)$$

onde $B^2 < C$, $0 < k < 1$. Observamos então que o k-ésimo momento de X só existe para $f(x)$, quando $k < 1$ (possui apenas momentos fracionários), isto é, a f.d.p. não possui Esperança, e consequentemente, não possui Variância, ou quaisquer momentos inteiros posteriores, utilizados frequentemente para estimação e intervalos de confiança. Como os momentos inteiros positivos não existem, então não será calculada a Função Geratriz de Momentos.

3.2.7 Função Característica para X

Calcularemos o k-ésimo momento $\varphi_X(t)$ para a densidade definida em (3.3):

$$\varphi_X(t) = Ee^{itX} = \int_0^\infty e^{itx} dF(x) = \int_0^\infty e^{itx} \frac{2\rho x^{2\rho-1}(Bx+C)}{(x^2+2Bx+C)^{\rho+1}} dx =$$

$$\int_0^\infty e^{itx} \frac{2\rho B x^{2\rho}}{(x^2+2Bx+C)^{\rho+1}} dx + \int_0^\infty e^{itx} \frac{2\rho C x^{2\rho-1}}{(x^2+2Bx+C)^{\rho+1}} dx =$$

$$\int_0^\infty cos(tx)+isen(tx)) \frac{2\rho B x^{2\rho}}{(x^2+2Bx+C)^{\rho+1}} dx + \int_0^\infty (cos(tx)+isen(tx)) \frac{2\rho C x^{2\rho-1}}{(x^2+2Bx+C)^{\rho+1}} dx =$$

$$\underbrace{\int_0^\infty cos(tx) \left[\frac{2\rho B x^{2\rho}}{(x^2+2Bx+C)^{\rho+1}} + \frac{2\rho C x^{2\rho-1}}{(x^2+2Bx+C)^{\rho+1}} \right] dx}_{I} +$$

$$\underbrace{i \int_0^\infty sen(tx) \left[\frac{2\rho B x^{2\rho}}{(x^2+2Bx+C)^{\rho+1}} + \frac{2\rho C x^{2\rho-1}}{(x^2+2Bx+C)^{\rho+1}} \right] dx}_{II}$$

Temos os seguintes resultados, disponíveis em *http://functions.wolfram.com* (20):

$$cos(tx) = \frac{\sqrt{\pi}}{2\pi i} \int_{\gamma-i\infty}^{\gamma+i\infty} \frac{\Gamma(-s)}{\Gamma(\frac{1}{2}+s)} \left[\frac{tx}{2} \right]^{2s} ds;\ 0 < \gamma < \frac{1}{2}, x > 0 \tag{3.8}$$

e

$$sen(tx) = \frac{\sqrt{\pi}}{2\pi i} \int_{\gamma-i\infty}^{\gamma+i\infty} \frac{\Gamma(-s)}{\Gamma(\frac{3}{2}+s)} \left[\frac{tx}{2} \right]^{2s-1} ds;\ 0 < \gamma < 1, x > 0 \tag{3.9}$$

Usando os resultados (3.8) e para resolver a expressão (I), temos:

$$I = \frac{\sqrt{\pi}}{2\pi i}\int_{\gamma-i\infty}^{\gamma+i\infty}\frac{\Gamma(-s)}{\Gamma(\frac{1}{2}+s)}\left[\frac{t^2}{4}\right]^s\int_0^\infty\frac{2\rho Bx^{2\rho+2s}}{(x^2+2Bx+C)^{\rho+1}}dxds+$$

$$\frac{\sqrt{\pi}}{2\pi i}\int_{\gamma-i\infty}^{\gamma+i\infty}\frac{\Gamma(-s)}{\Gamma(\frac{1}{2}+s)}\left[\frac{t^2}{4}\right]^s\int_0^\infty\frac{2\rho Cx^{2\rho+2s-1}}{(x^2+2Bx+C)^{\rho+1}}dxds$$

i) Comparando (I) e a expressão (3.5), temos:

$$I = \frac{2\rho B\sqrt{\pi}}{\Gamma(2\rho+2)\sqrt{C}}\frac{1}{2\pi i}\int_{\gamma-i\infty}^{\gamma+i\infty}\frac{\Gamma(-s)}{\Gamma(\frac{1}{2}+s)}\left[\frac{t^2C}{4}\right]^s *$$

$$\Gamma(2\rho+1+2s)\Gamma(1-2s)_2F_1\left[\rho+\frac{1}{2}+s;\frac{1}{2}-s;\rho+\frac{3}{2};1-\frac{B^2}{C}\right]ds+$$

$$\frac{2\rho\sqrt{\pi}}{\Gamma(2\rho+2)}\frac{1}{2\pi i}\int_{\gamma-i\infty}^{\gamma+i\infty}\frac{\Gamma(-s)}{\Gamma(\frac{1}{2}+s)}\left[\frac{t^2C}{4}\right]^s *$$

$$\Gamma(2\rho+2s)\Gamma(2-2s)_2F_1\left[\rho+s;1-s;\rho+\frac{3}{2};1-\frac{B^2}{C}\right]ds$$

ii) Usando as relações (2.13) e (2.11), temos que:

$$I = \frac{2\rho\sqrt{\pi}}{\Gamma(2\rho+2)}\frac{1}{2\pi i}\int_{\gamma-i\infty}^{\gamma+i\infty}\frac{\Gamma(-s)}{\Gamma(\frac{1}{2}+s)}\left[\frac{4}{t^2B^2}\right]^{-s} *$$

$$\Gamma(2\rho+1+2s)\Gamma(1-2s)_2F_1\left[\frac{1}{2}-s;1-s;\rho+\frac{3}{2};1-\frac{C}{B^2}\right]ds+$$

$$\frac{2\rho\sqrt{\pi}C}{\Gamma(2\rho+2)B^2}\frac{1}{2\pi i}\int_{\gamma-i\infty}^{\gamma+i\infty}\frac{\Gamma(-s)}{\Gamma(\frac{1}{2}+s)}\left[\frac{4}{t^2B^2}\right]^{-s}*$$

$$\Gamma(2\rho+2s)\Gamma(2-2s)_2F_1\left[1-s;\frac{3}{2}-s;\rho+\frac{3}{2};1-\frac{C}{B^2}\right]ds$$

iii)Usando a definição (2.16) da Integral de Contorno de ${}_2F_1[a;b;c:z]$ e a expressão (I), temos:

$$I=\frac{2\rho\sqrt{\pi}\Gamma(\rho+3/2)}{\Gamma(2\rho+2)}\frac{-1}{4\pi^2}\int_{\gamma-i\infty}^{\gamma+i\infty}\int_{\delta-i\infty}^{\delta+i\infty}\frac{\Gamma(-s)\Gamma(1-2s)\Gamma(2\rho+1+2s)}{\Gamma(\frac{1}{2}+s)\Gamma(1-s)\Gamma(\frac{1}{2}-s)}*$$

$$\frac{\Gamma(s_1)\Gamma(\frac{1}{2}-s-s_1)\Gamma(1-s-s_1)}{\Gamma(\rho+\frac{3}{2}-s_1)}\left[\frac{C}{B^2}-1\right]^{-s_1}\left[\frac{4}{t^2B^2}\right]^{-s}ds_1ds+$$

$$\frac{2\rho\sqrt{\pi}\Gamma(\rho+3/2)C}{\Gamma(2\rho+2)B^2}\frac{-1}{4\pi^2}\int_{\gamma-i\infty}^{\gamma+i\infty}\int_{\delta-i\infty}^{\delta+i\infty}\frac{\Gamma(-s)\Gamma(2-2s)\Gamma(2\rho+2s)}{\Gamma(\frac{1}{2}+s)\Gamma(\frac{3}{2}-s)\Gamma(1-s)}*$$

$$\frac{\Gamma(s_1)\Gamma(\frac{3}{2}-s-s_1)\Gamma(1-s-s_1)}{\Gamma(\rho+\frac{3}{2}-s_1)}\left[\frac{C}{B^2}-1\right]^{-s_1}\left[\frac{4}{t^2B^2}\right]^{s}ds_1ds$$

iv) Usando a definição (2.24) de $H\left[\begin{array}{c} x \\ y \end{array}\right]$, temos que:

$$I = c_1 \mathbb{H}^{0,2:1,2;1,0}_{0,0:1,3;0,2}\left[\begin{array}{c|c} \frac{4}{t^2B^2} & \left(0,1\right)\left(\frac{1}{2},1\right) : (1,1)\,(0,2)\left(\frac{1}{2},1\right);-- \\ \frac{C}{B^2}-1 & --:(0,1)\,(2\rho+1,2)\left(\frac{1}{2},1\right):(0,1)\left(-\frac{1}{2}-\rho,1\right) \end{array}\right]+$$

$$c_2 \mathbb{H}^{0,2:2,1;1,0}_{0,0:2,4;0,2}\left[\begin{array}{c|c} \frac{4}{t^2B^2} & \left(0,1\right)\left(-\frac{1}{2},1\right) : (1,1)\,(-1,2)\left(\frac{1}{2},1\right);-- \\ \frac{C}{B^2}-1 & --:(0,1)\,(2\rho,2)\left(-\frac{1}{2},1\right):(0,1)\left(-\frac{1}{2}-\rho,1\right) \end{array}\right]$$

em que $c_1 = \frac{2\rho\sqrt{\pi}\Gamma\left(\rho+\frac{3}{2}\right)}{\Gamma(2\rho+2)}$ e $c_2 = \frac{2\rho C\sqrt{\pi}\Gamma\left(\rho+\frac{3}{2}\right)}{B^2\Gamma(2\rho+2)}$.

Usando o resultado (3.9) para resolver a expressão (II), temos:

$$II = \frac{\sqrt{\pi}}{2\pi i}\int_{\gamma-i\infty}^{\gamma+i\infty}\frac{\Gamma(-s)}{\Gamma(\frac{3}{2}+s)}\left[\frac{t}{2}\right]^{2s-1}\int_0^\infty \frac{2\rho B x^{2\rho+2s-1}}{(x^2+2Bx+C)^{\rho+1}}dxds+$$

$$\frac{\sqrt{\pi}}{2\pi i}\int_{\gamma-i\infty}^{\gamma+i\infty}\frac{\Gamma(-s)}{\Gamma(\frac{3}{2}+s)}\left[\frac{t}{2}\right]^{2s-1}\int_0^\infty \frac{2\rho C x^{2\rho+2s-2}}{(x^2+2Bx+C)^{\rho+1}}dxds$$

i) Comparando (II) e a expressão (3.5), temos:

$$II = \frac{4\rho B\sqrt{\pi}i}{\Gamma(2\rho+2)Ct}\frac{1}{2\pi i}\int_{\gamma-i\infty}^{\gamma+i\infty}\frac{\Gamma(-s)}{\Gamma(\frac{3}{2}+s)}\left[\frac{t^2C}{4}\right]^s *$$

$$\Gamma(2\rho+2s)\Gamma(2-2s)_2F_1\left[\rho+s;1-s;\rho+\frac{3}{2};1-\frac{B^2}{C}\right]ds+$$

$$\frac{4\rho\sqrt{\pi}i}{\Gamma(2\rho+2)\sqrt{C}t}\frac{1}{2\pi i}\int_{\gamma-i\infty}^{\gamma+i\infty}\frac{\Gamma(-s)}{\Gamma(\frac{3}{2}+s)}\left[\frac{t^2C}{4}\right]^s *$$

$$\Gamma(2\rho-1+2s)\Gamma(3-2s)_2F_1\left[\rho-\frac{1}{2}+s;\frac{3}{2}-s;\rho+\frac{3}{2};1-\frac{B^2}{C}\right]ds$$

ii) Usando as relações (2.13) e (rel1), temos que:

$$II=\frac{4\rho\sqrt{\pi}i}{tB\Gamma(2\rho+2)}\frac{1}{2\pi i}\int_{\gamma-i\infty}^{\gamma+i\infty}\frac{\Gamma(-s)}{\Gamma(\frac{3}{2}+s)}\left[\frac{4}{t^2B^2}\right]^{-s}*$$

$$\Gamma(2\rho+2s)\Gamma(2-2s)_2F_1\left[1-s;\frac{3}{2}-s;\rho+\frac{3}{2};1-\frac{C}{B^2}\right]ds+$$

$$\frac{4\rho\sqrt{\pi}C}{\Gamma(2\rho+2)tB^3}\frac{1}{2\pi i}\int_{\gamma-i\infty}^{\gamma+i\infty}\frac{\Gamma(-s)}{\Gamma(\frac{3}{2}+s)}\left[\frac{4}{t^2B^2}\right]^{-s}*$$

$$\Gamma(2\rho+1+2s)\Gamma(3-2s)_2F_1\left[\frac{3}{2}-s;2-s;\rho+\frac{3}{2};1-\frac{C}{B^2}\right]ds$$

iii)Usando a definição da Integral de Contorno de ${}_2F_1[a;b;c:z]$ (2.16) e a expressão (II), temos:

$$II=\frac{4\rho\sqrt{\pi}\Gamma(\rho+3/2)i}{\Gamma(2\rho+2)tB}\frac{-1}{4\pi^2}\int_{\gamma-i\infty}^{\gamma+i\infty}\int_{\delta-i\infty}^{\delta+i\infty}\frac{\Gamma(-s)\Gamma(2-2s)\Gamma(2\rho+2s)}{\Gamma(\frac{3}{2}+s)\Gamma(\frac{3}{2}-s)\Gamma(1-s)}*$$

$$\frac{\Gamma(s_1)\Gamma(\frac{3}{2}-s-s_1)\Gamma(1-s-s_1)}{\Gamma(\rho+\frac{3}{2}-s_1)}\left[\frac{C}{B^2}-1\right]^{-s_1}\left[\frac{4}{t^2B^2}\right]^{-s}ds_1ds+$$

$$\frac{4\rho C\sqrt{\pi}\Gamma(\rho+3/2)i}{\Gamma(2\rho+2)tB^3}\frac{-1}{4\pi^2}\int_{\gamma-i\infty}^{\gamma+i\infty}\int_{\delta-i\infty}^{\delta+i\infty}\frac{\Gamma(-s)\Gamma(3-2s)\Gamma(2\rho+1+2s)}{\Gamma(\frac{3}{2}+s)\Gamma(2-s)\Gamma(\frac{3}{2}-s)}*$$

$$\frac{\Gamma(s_1)\Gamma(2-s-s_1)\Gamma(\frac{3}{2}-s-s_1)}{\Gamma(\rho+\frac{3}{2}-s_1)}\left[\frac{C}{B^2}-1\right]^{-s_1}\left[\frac{4}{t^2B^2}\right]^{-s}ds_1ds$$

iv) Usando a definição (2.24) de $H\left[\begin{array}{c} x \\ y \end{array}\right]$ (2.19), temos que:

$$II = c_3\mathbb{H}^{0,2:1,2;1,0}_{0,0:1,3;0,2}\left[\begin{array}{c|c} \frac{4}{t^2B^2} & \left(0,1\right)\left(-\frac{1}{2},1\right):(1,1)\left(-1,2\right)\left(\frac{3}{2},1\right);-- \\ \frac{C}{B^2}-1 & --:(0,1)\left(2\rho,2\right)\left(-\frac{1}{2},1\right):(0,1)\left(-\frac{1}{2}-\rho,1\right) \end{array}\right]+$$

$$c_4\mathbb{H}^{0,2:2,1;1,0}_{0,0:2,4;0,2}\left[\begin{array}{c|c} \frac{4}{t^2B^2} & \left(-1,1\right)\left(-\frac{1}{2},1\right):(1,1)\left(-2,2\right)\left(\frac{3}{2},1\right);-- \\ \frac{C}{B^2}-1 & --:(0,1)\left(2,2\right)\left(\frac{1}{2},1\right)\left(1-\rho,1\right);(0,1)\left(-\frac{1}{2}-\rho,1\right) \end{array}\right]$$

em que $c_3=\frac{4\rho\sqrt{\pi}\Gamma\left(\rho+\frac{3}{2}\right)i}{tB^3\Gamma(2\rho+2)}$ e $c_4=\frac{4\rho C\sqrt{\pi}\Gamma\left(\rho+\frac{3}{2}\right)i}{tB^3\Gamma(2\rho+2)}$.

Mas, $\varphi_X(t)=(I)+(II)$, então

$$\varphi_X(t) = c_1\mathbb{H}^{0,2:1,2;1,0}_{0,0:1,3;0,2}\left[\begin{array}{c|c} \frac{4}{t^2B^2} & \left(0,1\right)\left(\frac{1}{2},1\right):(1,1)\left(0,2\right)\left(\frac{1}{2},1\right);-- \\ \frac{C}{B^2}-1 & --:(0,1)\left(2\rho+1,2\right)\left(\frac{1}{2},1\right):(0,1)\left(-\frac{1}{2}-\rho,1\right) \end{array}\right]+$$

$$c_2\mathbb{H}^{0,2:2,1;1,0}_{0,0:2,4;0,2}\left[\begin{array}{c|c} \frac{4}{t^2B^2} & \left(0,1\right)\left(-\frac{1}{2},1\right):(1,1)\left(-1,2\right)\left(\frac{1}{2},1\right);-- \\ \frac{C}{B^2}-1 & --:(0,1)\left(2\rho,2\right)\left(-\frac{1}{2},1\right):(0,1)\left(-\frac{1}{2}-\rho,1\right) \end{array}\right]$$

$$c_3\mathbb{H}^{0,2:1,2;1,0}_{0,0:1,3;0,2}\left[\begin{array}{c|c} \frac{4}{t^2B^2} & \left(0,1\right)\left(-\frac{1}{2},1\right):(1,1)\left(-1,2\right)\left(\frac{3}{2},1\right);-- \\ \frac{C}{B^2}-1 & --:(0,1)\left(2\rho,2\right)\left(-\frac{1}{2},1\right):(0,1)\left(-\frac{1}{2}-\rho,1\right) \end{array}\right]+$$

$$c_4 \mathbb{H}_{0,0:2,4;0,2}^{0,2:2,1;1,0} \left[\begin{array}{c|c} \frac{4}{t^2 B^2} & (-1,1)\left(-\frac{1}{2},1\right) : (1,1)\,(-2,2)\left(\frac{3}{2},1\right);-- \\ \frac{C}{B^2}-1 & --:(0,1)\,(2,2)\left(\frac{1}{2},1\right)(1-\rho,1);(0,1)\left(-\frac{1}{2}-\rho,1\right) \end{array} \right]$$

,

onde $c_1 = \frac{2\rho\sqrt{\pi}\Gamma\left(\rho+\frac{3}{2}\right)}{\Gamma(2\rho+2)}$, $c_2 = \frac{2\rho\sqrt{\pi}\Gamma\left(\rho+\frac{3}{2}C\right)}{\Gamma(2\rho+2)B^2}$, $c_3 = \frac{4\rho\pi\Gamma\left(\rho+\frac{3}{2}\right)i}{tB\Gamma(2\rho+2)}$ e $c_4 = \frac{4\rho\sqrt{\pi}\Gamma\left(\rho+\frac{3}{2}\right)Ci}{tB^3\Gamma(2\rho+2)}$.

3.2.8 Estimadores para os parâmetros B, C, ρ de $f(x)$

O método utilizado é o de Máxima Verossimilhança. Os estimadores do Método dos Momentos não serão calculados, pois existem apenas os momentos fracionários entre 0 e 1, que podem não fornecer boas estimativas. Desta forma, calcularemos o E.M.V., onde $\underline{\theta} = (B, C, \rho), \underline{\theta} \in \Theta = C > 0, B^2 < C, \rho > 0$ como a seguir:

$$L(\underline{\theta}; x) = \prod_{i=1}^{n} f(x_i|\theta) = \prod_{i=1}^{n} \frac{2\rho x_i^{2\rho-1}(Bx_i + C)}{(x_i^2 + 2Bx_i + C)^{\rho+1}} =$$

$$L(\underline{\theta}; x) = (2\rho)^n \prod_{i=1}^{n} \frac{Bx_i^{2\rho} + Cx_i^{2\rho-1}}{(x_i^2 + 2Bx_i + C)^{\rho+1}} \tag{3.10}$$

e

$$l(\underline{\theta}; x) = logL(\underline{\theta}; x) = n\log 2 + n\log\rho + \sum_{i=1}^{n} \log\left(\frac{Bx_i^{2\rho} + Cx_i^{2\rho-1}}{(x_i^2 + 2Bx_i + C)^{\rho+1}}\right) =$$

$$n\log 2 + n\log\rho + \sum_{i=1}^{n} \log((Bx_i + C)x_i^{2\rho-1}) - \sum_{i=1}^{n} \log(x_i^2 + 2Bx_i + C)^{\rho+1} =$$

$$n\log 2 + n\log\rho + \sum_{i=1}^{n} \log(Bx_i + C) + (2\rho - 1)\sum_{i=1}^{n} \log x_i - (\rho + 1)\sum_{i=1}^{n} \log(x_i^2 + 2Bx_i + C) \tag{3.11}$$

Para estimar valores $\underline{\theta} = (B, C, \rho)$, resolveremos a $l'(\underline{\theta}; x) = \frac{\partial l(\underline{\theta}; x)}{\partial \underline{\theta}} = 0$:

I)

$$\frac{\partial l(\underline{\theta}; x)}{\partial\ \rho} = 0 \Rightarrow \frac{n}{\rho} + 2\sum_{i=1}^{n} \log x_i - \sum_{i=1}^{n} \log(x_i^2 + 2Bx_i + C) = 0$$

II)

$$\frac{\partial l(\underline{\theta}; x)}{\partial B} = 0 \Rightarrow \sum_{i=1}^{n} \frac{x_i}{Bx_i + C} - (\rho + 1)\sum_{i=1}^{n} \frac{2x_i}{(x_i^2 + 2Bx_i + C)} = 0$$

III)

$$\frac{\partial l(\underline{\theta}; x)}{\partial C} = 0 \Rightarrow \sum_{i=1}^{n} \frac{1}{Bx_i + C} - (\rho + 1)\sum_{i=1}^{n} \frac{1}{(x_i^2 + 2Bx_i + C)} = 0$$

Resolvendo as equações I)-III) acima, obtemos o E.M.V. para $\underline{\theta} = (B, C, \rho)$. Utilizaremos o Intervalo de Confiança descrito no Apêndice I.

Capítulo 4

Versão Simétrica para a Distribuição

4.1 Introdução

Apresentaremos a seguir distribuição simétrica correspondente àquela apresentada no Capítulo 3. Como possui diversos parâmetros, a distribuição pode assumir várias formas, o que facilita o uso em diversos dados onde se queira modelar v.a's em toda a reta.

4.2 Distribuição Acumulada Simétrica

Definição 4.2.1. *Seja a forma*

$$G(x) = \frac{1}{2} + \frac{1}{2}\frac{x|x|^{2\rho-1}}{(x^2 + 2B|x| + C)^{\rho}} \tag{4.1}$$

a função de distribuição simétrica correspondente a (6.2.3) de uma v.a. X onde $-\infty < x < \infty, C > 0, B^2 < C$ e $\rho > 0$.

4.2.1 Propriedades da Distribuição Acumulada Simétrica

Vamos provar que $G(x)$ é distribuição:

$i) \lim_{x\to-\infty} G(x) = 0$

Demonstração:

$\lim_{x\to-\infty} G(x) = \lim_{x\to-\infty} \frac{1}{2}+\frac{1}{2}\frac{x|x|^{2\rho-1}}{(x^2+2B|x|+C)^\rho} = \lim_{x\to-\infty} \frac{1}{2}+sinal[x]\frac{1}{2}(\frac{1}{1+\frac{2B}{|x|}+\frac{C}{|x|^2}})^\rho = \frac{1}{2}-\frac{1}{2}=0$,

onde $sinal[x] = \frac{x}{|x|} = \frac{|x|}{x}$.

$ii) \lim_{x\to\infty} G(x) = 1$

Demonstração:

$\lim_{x\to+\infty} G(x) = \lim_{x\to+\infty} \frac{1}{2}+\frac{1}{2}\frac{x|x|^{2\rho-1}}{(x^2+2B|x|+C)^\rho} = \lim_{x\to+\infty} \frac{1}{2}+sinal[x]\frac{1}{2}(\frac{1}{1+\frac{2B}{|x|}+\frac{C}{|x|^2}})^\rho = \frac{1}{2}+\frac{1}{2}=1$

$iii) G(x) \leq G(y), \forall x < y; x, y \in \mathbb{R}$

Demonstração: Como $G(x)$ é função não-decrescente, $\forall x \in \mathbb{R}$, logo atende à propriedade iv).

$iv) 0 \leq G_X(x) \leq 1, \forall x;$

Demonstração: De fato, pelas propriedades $i) - iii)$ conclui-se que $iv)$ também vale.

4.2.2 Gráficos para a Distribuição Acumulada Simétrica

Como provamos que $G(x)$ é distribuição, então apresentaremos algumas formas da distribuição, para diversos valores dos parâmetros B, C, ρ:

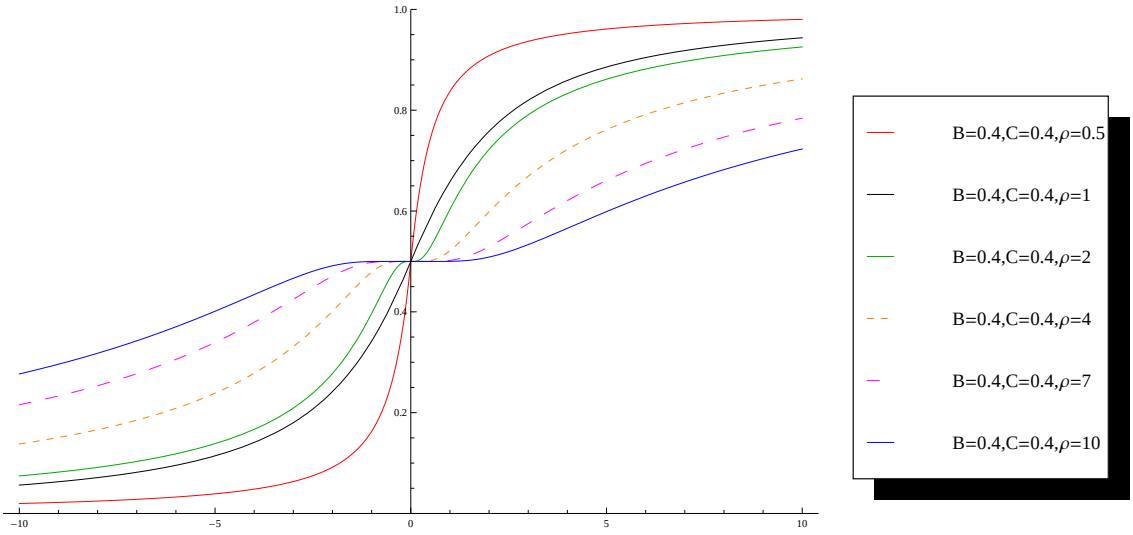

Figura 4.1: Formas para a Distribuição Acumulada Simétrica $G(x)$

Para verificar o efeito de cada um dos parâmetros na curva, produzimos gráficos onde se fixa dois parâmetros enquanto um terceiro varia.

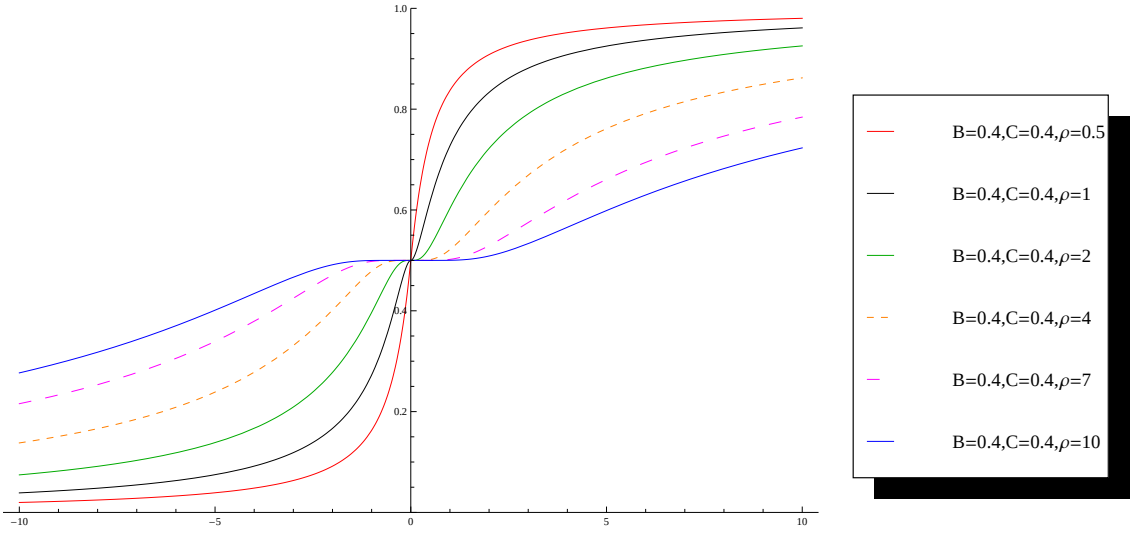

Figura 4.2: Formas para a Distribuição Simétrica Acumulada $G(x)$ - variação de ρ, com $B > 0$

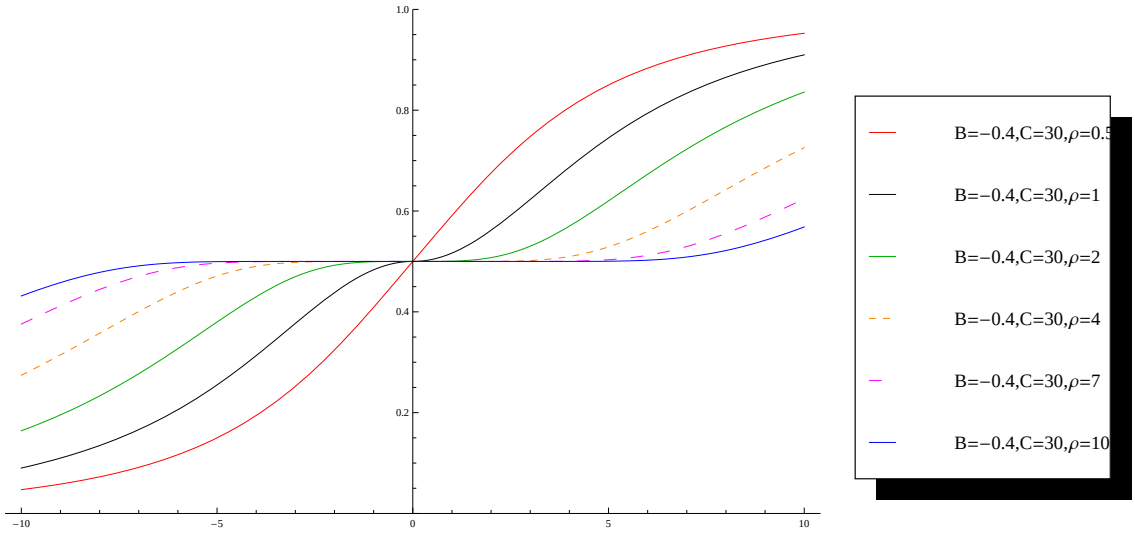

Figura 4.3: Formas para a Distribuição Simétrica Acumulada $G(x)$ - variação de ρ, com $B < 0$

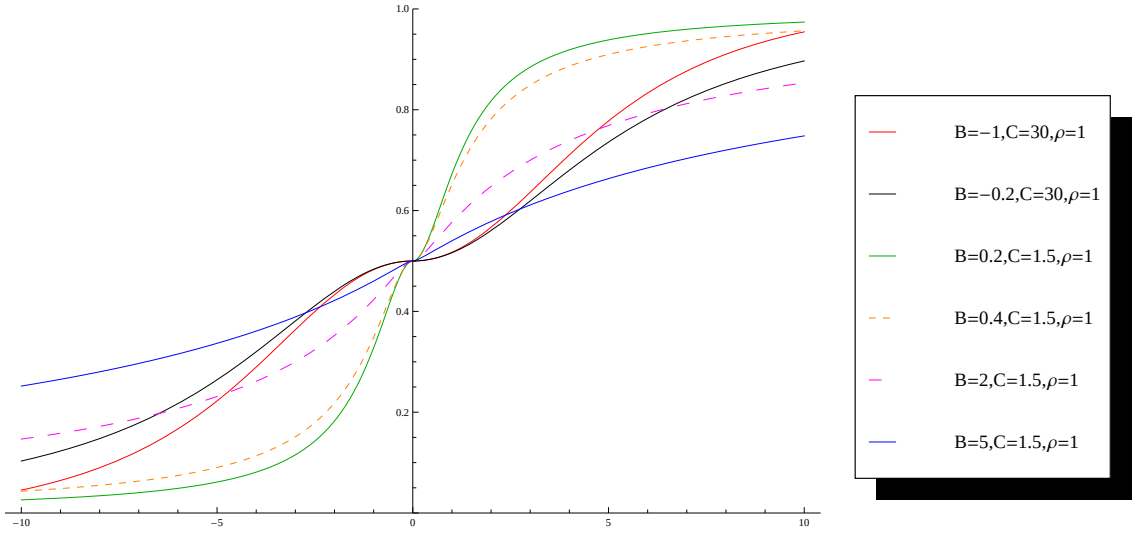

Figura 4.4: Formas para a Distribuição Simétrica Acumulada $G(x)$ - variação de B

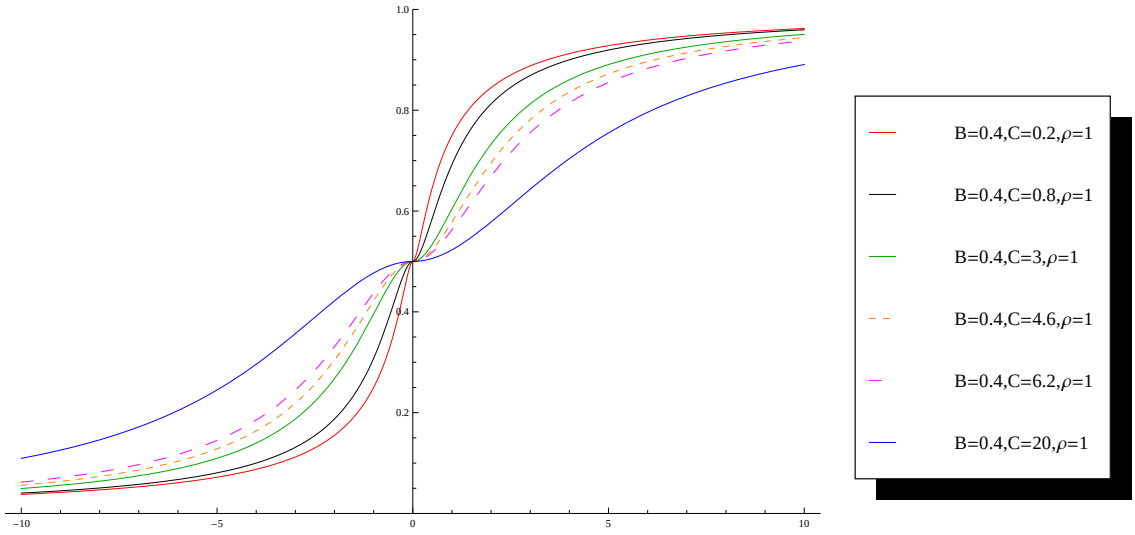

Figura 4.5: Formas para a Distribuição Simétrica Acumulada $G(x)$ - variação de C, com $B > 0$

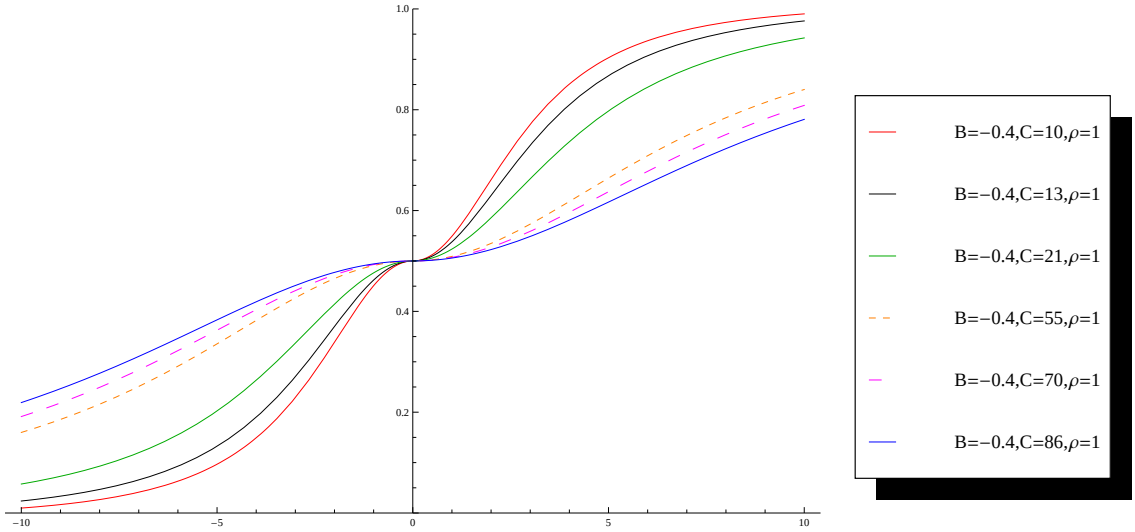

Figura 4.6: Formas para a Distribuição Simétrica Acumulada $G(x)$ - variação de C, com $B < 0$

4.2.3 Função Densidade de Probabilidade Simétrica

Uma vez que $G(x)$ existe, então encontraremos $g(x) = \frac{dG(x)}{dx}$ a seguir:

$$g(x) = \frac{d}{dx}(\frac{1}{2} + \frac{1}{2}\frac{x|x|^{2\rho-1}}{(x^2+2B|x|+C)^{\rho}})$$

$$= \frac{(2\rho|x|^{2\rho-1})(x^2+2B|x|+C)^{\rho} - ((x|x|^{2\rho-1}\rho)(x^2+2B|x|+C)^{\rho-1}(2x+\frac{|x|}{x}2B))}{2(x^2+2B|x|+C)^{2\rho}}$$

$$= \frac{(2\rho|x|^{2\rho-1})(x^2+2B|x|+C)^{\rho} - 2(x^2+2B|x|+C)^{\rho-1}[(x^2|x|^{2\rho-1}\rho) + (|x|^{2\rho}\rho)]}{(x^2+2B|x|+C)^{2\rho}}$$

$$= \frac{\rho|x|^{2\rho+1} + 2B\rho|x|^{2\rho} + C\rho|x|^{2\rho-1} - \rho|x|^{2\rho+1} - B\rho|x|^{2\rho}}{(x^2+2B|x|+C)^{\rho+1}}$$

$$= \frac{B\rho|x|^{2\rho} + C\rho|x|^{2\rho-1}}{(x^2+2B|x|+C)^{\rho+1}}$$

Definição 4.2.2. *Seja a forma*

$$g(x) = \frac{\rho|x|^{2\rho-1}(B|x|+C)}{(x^2+2B|x|+C)^{\rho+1}} \tag{4.2}$$

a função densidade de probabilidade de uma v.a. X onde $-\infty < x < \infty$, $C > 0$, $B^2 < C$ e $\rho > 0$.

4.2.4 Gráficos para a Função Densidade de Probabilidade Simétrica

A seguir apresentaremos algumas formas da f.d.p., para diversos valores dos parâmetros B, C, ρ:

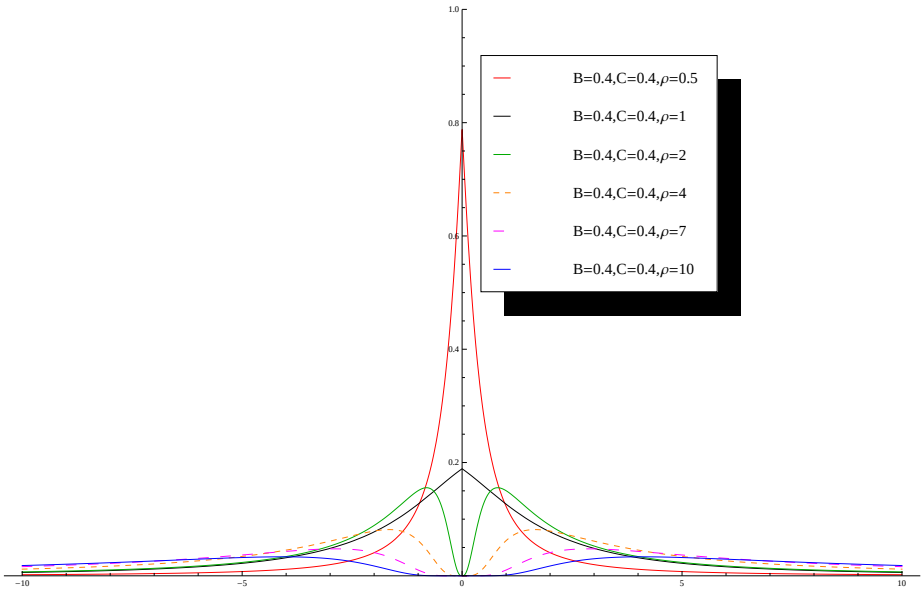

Figura 4.7: Formas para a Densidade Simétrica $g(x)$

Observamos que ao fixar $\rho = \frac{1}{2}$, obtemos densidade simétrica unimodal e para todos os outros valores de ρ, criamos densidades bimodais. Para verificar o efeito de cada um dos parâmetros na curva, produzimos gráficos onde se fixa dois parâmetros enquanto um terceiro varia.

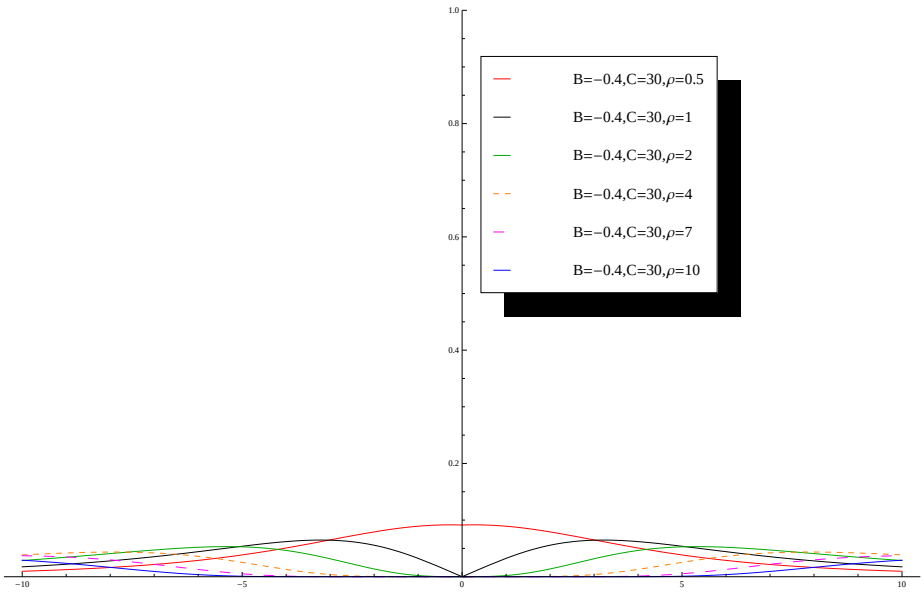

Figura 4.8: Formas para a Densidade $g(x)$ - variação de ρ, com $B > 0$

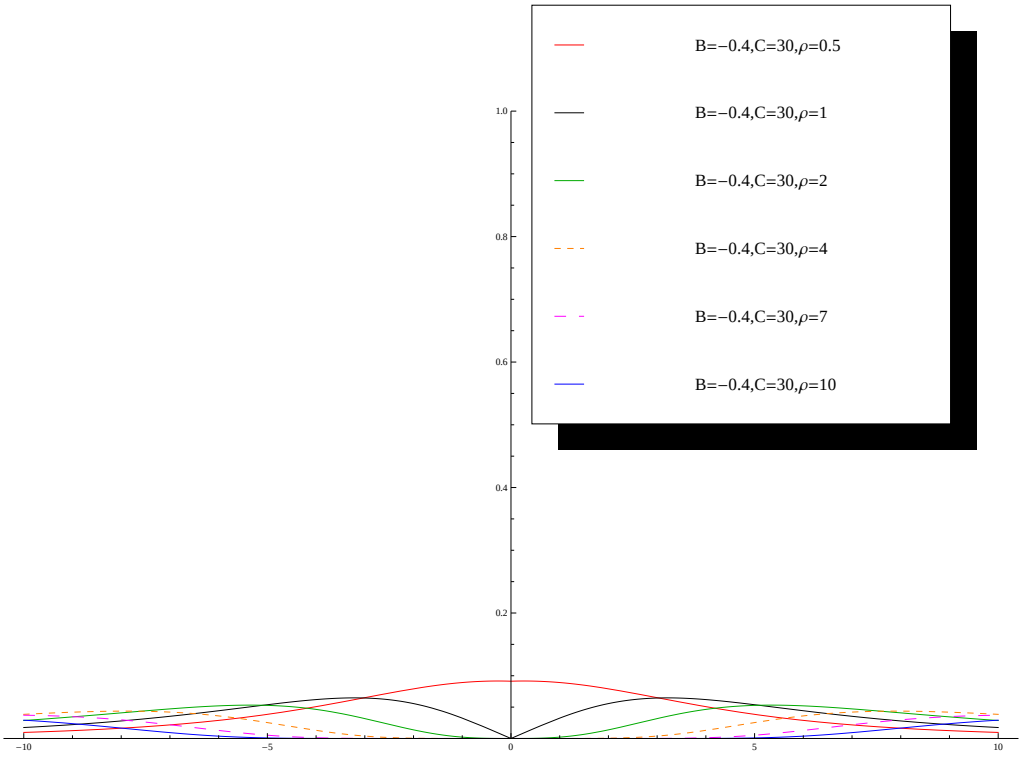

Figura 4.9: Formas para a Densidade $g(x)$ - variação de ρ, com $B < 0$

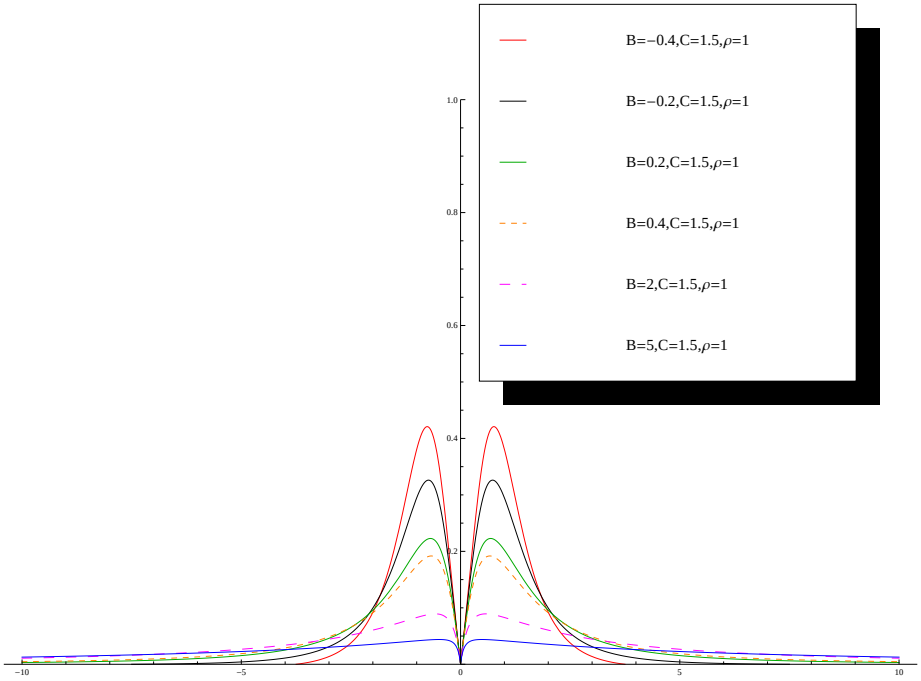

Figura 4.10: Formas para a Densidade $g(x)$ - variação de B

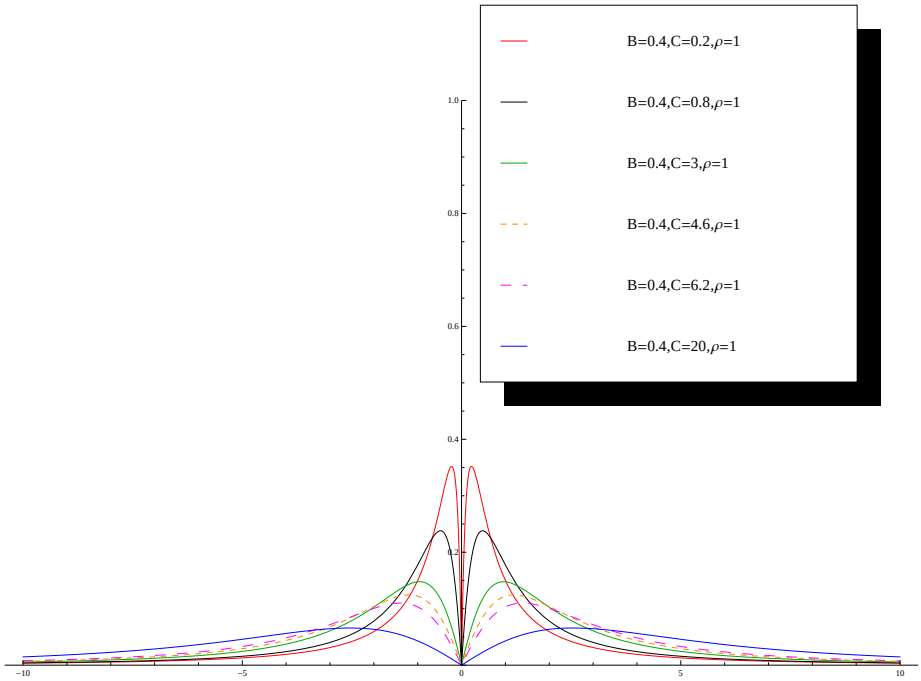

Figura 4.11: Formas para a Densidade $g(x)$ - variação de C, com $B > 0$

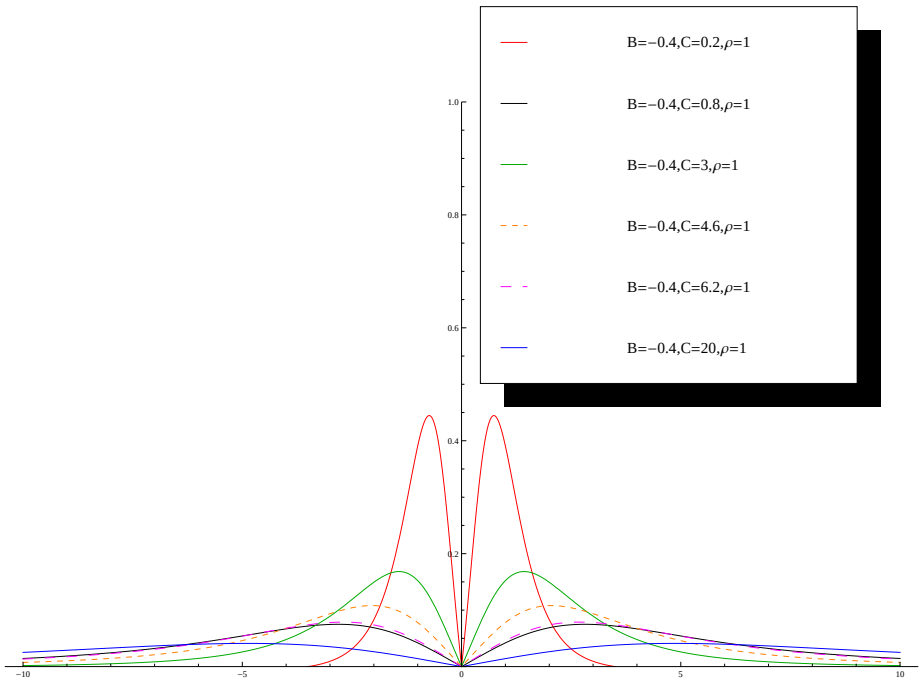

Figura 4.12: Formas para a Densidade $g(x)$ - variação de C, com $B < 0$

As figuras (4.8) e (4.9) apontam que o parâmetro ρ determina o comportamento da cauda da densidade.

Pela figura (3.11) verificamos que o parâmetro B determina o comportamento relacionado à distância entre as curvas do lado negativo e positivo. Quanto maior B, mais próximas ficam as curvas.

Pelas figuras (4.11) e (4.12) observamos que o parâmetro C determina o comportamento relacionado ao pico da função. Quanto maior C, menor o pico da curva.

4.2.5 Distribuições de Mínimo e Máximo

A distribuição e a densidade do Mínimo $X_{(1)} = min(X_1, \ldots, X_n)$ são dadas por:

$$P(X_{(1)} \leq z) = 1 - P(X_{(1)} > z) = 1 - P(min(X_1, \ldots, X_n) > z) = 1 - \prod_{i=1}^{n} [1 - G_{X_{(1)}}(z)]$$

$$G(z) = 1 - \left[\frac{1}{2} - \frac{1}{2}\frac{z|z|^{2\rho-1}}{(z^2 + 2B|z| + C)^{\rho}}\right]^n$$

$$g(z) = n\rho\frac{|z|^{2\rho-1}(B|z| + C)}{(z^2 + 2B|z| + C)^{\rho} + 1}\left[\frac{1}{2} - \frac{1}{2}\frac{z|z|^{2\rho-1}}{(z^2 + 2B|z| + C)^{\rho}}\right]^{n-1}$$

A distribuição e a densidade do Máximo $X_{(n)} = max(X_1, \ldots, X_n)$ são dadas por:

$$P(X_{(n)} \leq w) = P(max(X_1, \ldots, X_n) \geq w) = \prod_{i=1}^{n} F_{X_{(1)}}(w)$$

$$G(w) = \left[\frac{1}{2} + \frac{1}{2}\frac{w|w|^{2\rho-1}}{(w^2 + 2B|w| + C)^{\rho}}\right]^n$$

$$g(w) = n\rho \frac{|w|^{2\rho-1}(B|w|+C)}{(z^2+2B|z|+C)^{\rho+1}} \left[\frac{1}{2} + \frac{1}{2}\frac{w|w|^{2\rho-1}}{(z^2+2B|z|+C)^{\rho}}\right]^{n-1}$$

4.2.6 Momentos para X

Calcularemos o k-ésimo momento μ'_k para a densidade definida em (4.2), utilizando (3.5):

$$\mu'_k = EX^k = \int_{-\infty}^{\infty} x^k dG(x) = \int_{-\infty}^{\infty} x^k \frac{\rho|x|^{2\rho-1}(B|x|+C)}{(x^2+2B|x|+C)^{\rho+1}}dx =$$

$$\int_{-\infty}^{0} \frac{\rho x^k|x|^{2\rho-1}(B|x|+C)}{(x^2+2B|x|+C)^{\rho+1}}dx + \int_{0}^{\infty} \frac{\rho x^k|x|^{2\rho-1}(B|x|+C)}{(x^2+2B|x|+C)^{\rho+1}}dx =$$
$$\int_{-\infty}^{0} \rho B \frac{x^k(-x)^{2\rho}}{(x^2+2B(-x)+C)^{\rho+1}}dx + \int_{-\infty}^{0} \rho C \frac{x^k(-x)^{2\rho-1}}{(x^2-2Bx+C)^{\rho+1}}dx +$$

$$\int_{0}^{\infty} \rho B \frac{(x)^{k+2\rho}}{(x^2+2Bx+C)^{\rho+1}}dx + \int_{0}^{\infty} \rho C \frac{(x)^{k+2\rho-1}}{(x^2+2Bx+C)^{\rho+1}}dx =$$

Fazendo a mudança de variável $x = -y$, temos:

$$-\int_{\infty}^{0} \rho B \frac{(-y)^k(y)^{2\rho}}{(x^2+2By+C)^{\rho+1}}dy - \int_{\infty}^{0} \rho C \frac{(-y)^k(y)^{2\rho-1}}{(y^2+2By+C)^{\rho+1}}dy +$$

$$\int_{0}^{\infty} \rho B \frac{(x)^{k+2\rho}}{(x^2+2Bx+C)^{\rho+1}}dx + \int_{0}^{\infty} \rho C \frac{(x)^{k+2\rho-1}}{(x^2+2Bx+C)^{\rho+1}}dx =$$

$$\int_0^\infty \rho B \frac{(-y)^k (y)^{2\rho}}{(x^2+2By+C)^{\rho+1}} dy + \int_0^\infty \rho C \frac{(-y)^k (y)^{2\rho-1}}{(y^2+2By+C)^{\rho+1}} dy +$$

$$\int_0^\infty \rho B \frac{(x)^{k+2\rho}}{(x^2+2Bx+C)^{\rho+1}} dx + \int_0^\infty \rho C \frac{(x)^{k+2\rho-1}}{(x^2+2Bx+C)^{\rho+1}} dx =$$

$$(-1)^k \int_0^\infty \rho B \frac{(y)^{k+2\rho}}{(y^2-2By+C)^{\rho+1}} dy + (-1)^k \int_0^\infty \rho C \frac{(y)^{k+2\rho-1}}{(x^2+2By+C)^{\rho+1}} dy +$$

$$\int_0^\infty \rho B \frac{(x)^{k+2\rho}}{(x^2+2Bx+C)^{\rho+1}} dx + \int_0^\infty \rho C \frac{(x)^{k+2\rho-1}}{(x^2+2Bx+C)^{\rho+1}} dx$$

Se k é ímpar, $\mu'_k = 0$. Se k é par, temos o resultado idêntico ao caso assimétrico do Capítulo 3, então:

$$\mu'_k = \Bigg\{ B\left[C^{\frac{k-1}{2}}\Gamma(k+2\rho+1)\Gamma(1-k)\right] {}_2F_1\left[\frac{k+2\rho+1}{2}; \frac{1-k}{2}; \rho+\frac{3}{2}; 1-\frac{B^2}{C}\right] +$$

$$\mathrm{C}\left[C^{\frac{k}{2}-1}\Gamma(k+2\rho)\Gamma(2-k)\right] {}_2F_1\left[\rho+\tfrac{k}{2}; 1-\tfrac{k}{2}; \rho+\tfrac{3}{2}; 1-\tfrac{B^2}{C}\right] \Bigg\} \left[\tfrac{2\rho}{\Gamma(2\rho+2)}\right] \qquad (4.3)$$

onde $B^2 < C$, $0 < k < 1$. Observamos então que o k-ésimo momento de X só existe para $f(x)$, quando $k < 1$ e é par(possui apenas momentos fracionários), isto é, a f.d.p. não possui Esperança, e consequentemente, não possui Variância, ou quaisquer momentos inteiros posteriores, utilizados frequentemente para estimação e intervalos de confiança. Como os momentos inteiros positivos não existem, então não será calculada a Função Geratriz de Momentos.

4.2.7 Função Característica para X

Calcularemos o $\varphi_X(t)$ para a densidade definida em (4.2):

$$\varphi_X(t) = Ee^{itX} = \int_{-\infty}^{\infty} e^{itx} dG(x) = \int_{-\infty}^{\infty} (cos(tx) + isen(tx)) dG(x) =$$

Usando a fórmula $\int_{-\infty}^{\infty} e^{itx} dG(x) = 2\int_{0}^{\infty} cos(tx) dG(x)$,temos:

$$\int_0^{\infty} 2cos(tx) dG(x) = \int_0^{\infty} 2cos(tx) \frac{\rho|x|^{2\rho-1}(B|x| + C)}{(x^2 + 2B|x| + C)^{\rho+1}} dx =$$

$$\int_0^{\infty} cos(tx) \frac{2\rho x^{2\rho-1}(Bx + C)}{(x^2 + 2Bx + C)^{\rho+1}} dx$$

Mas, o resultado acima é idêntico à expressão (I) da função característica da distribuição assimétrica apresentada no Capítulo 3, então

$$\varphi_X(t) = c_1 \mathbb{H}_{0,0:1,3;0,2}^{0,2:1,2;1,0} \left[\begin{array}{c|c} \frac{4}{t^2B^2} & \left(0,1\right)\left(\frac{1}{2},1\right) : (1,1)\,(0,2)\left(\frac{1}{2},1\right);-- \\ \frac{C}{B^2} - 1 & --:(0,1)\,(2\rho+1,2)\left(\frac{1}{2},1\right):(0,1)\left(-\frac{1}{2}-\rho,1\right) \end{array} \right] +$$

$$c_2 \mathbb{H}_{0,0:2,4;0,2}^{0,2:2,1;1,0} \left[\begin{array}{c|c} \frac{4}{t^2B^2} & \left(0,1\right)\left(-\frac{1}{2},1\right) : (1,1)\,(-1,2)\left(\frac{1}{2},1\right);-- \\ \frac{C}{B^2} - 1 & --:(0,1)\,(2\rho,2)\left(-\frac{1}{2},1\right):(0,1)\left(-\frac{1}{2}-\rho,1\right) \end{array} \right]$$

em que $c_1 = \frac{2\rho\sqrt{\pi}\Gamma\left(\rho+\frac{3}{2}\right)}{\Gamma(2\rho+2)}$ e $c_2 = \frac{2\rho C\sqrt{\pi}\Gamma\left(\rho+\frac{3}{2}\right)}{B^2\Gamma(2\rho+2)}$.

4.2.8 Estimadores para os parâmetros B, C, ρ de $f(x)$

O método utilizado é o Máxima Verossimilhança. Os estimadores do Método dos Momentos não serão calculados,pois existem apenas os momentos fracionários entre 0 e 1, que podem não fornecer boas estimativas. Desta forma, calcularemos o E.M.V., onde $\underline{\theta} = (B, C, \rho), \underline{\theta} \in \Theta = C > 0, B^2 < C, \rho > 0$ como a seguir:

$$L(\underline{\theta}; x) = \prod_{i=1}^{n} g(x_i|\theta) = \prod_{i=1}^{n} \frac{\rho|x_i|^{2\rho-1}(B|x_i| + C)}{(x_i^2 + 2B|x_i| + C)^{\rho+1}} =$$

$$L(\underline{\theta}; x) = (\rho)^n \prod_{i=1}^{n} \frac{B|x_i|^{2\rho} + C|x_i|^{2\rho-1}}{(x_i^2 + 2B|x_i| + C)^{\rho+1}} \tag{4.4}$$

e

$$l(\underline{\theta}; x) = logL(\underline{\theta}; x) = n \log \rho + \sum_{i=1}^{n} \log \left(\frac{B|x_i|^{2\rho} + C|x_i|^{2\rho-1}}{(x_i^2 + 2B|x_i| + C)^{\rho+1}} \right) =$$

$$n \log \rho + \sum_{i=1}^{n} \log((B|x_i| + C)|x_i|^{2\rho-1}) - \sum_{i=1}^{n} \log(x_i^2 + 2B|x_i| + C)^{\rho+1} =$$

$$n \log \rho + \sum_{i=1}^{n} \log(B|x_i| + C) + (2\rho - 1) \sum_{i=1}^{n} \log |x_i| - (\rho + 1) \sum_{i=1}^{n} \log(x_i^2 + 2B|x_i| + C) \tag{4.5}$$

Para estimar valores $\underline{\theta} = (B, C, \rho)$, resolveremos a $l'(\underline{\theta}; x) = \frac{\partial l(\underline{\theta}; x)}{\partial \underline{\theta}} = 0$:

I)

$$\frac{\partial l(\underline{\theta}; x)}{\partial\ \rho} = 0 \Rightarrow \frac{n}{\rho} + 2 \sum_{i=1}^{n} \log |x_i| - \sum_{i=1}^{n} \log(x_i^2 + 2B|x_i| + C) = 0$$

II)

$$\frac{\partial l(\underline{\theta};x)}{\partial B}=0\Rightarrow\sum_{i=1}^{n}\frac{|x_i|}{B|x_i|+C}-(\rho+1)\sum_{i=1}^{n}\frac{2|x_i|}{(x_i^2+2B|x_i|+C)}=0$$

III)

$$\frac{\partial l(\underline{\theta};x)}{\partial C}=0\Rightarrow\sum_{i=1}^{n}\frac{1}{B|x_i|+C}-(\rho+1)\sum_{i=1}^{n}\frac{1}{(x_i^2+2B|x_i|+C)}=0$$

Resolvendo as equações I)-III) acima, obtemos o E.M.V. para $\underline{\theta}=(B,C,\rho)$. Utilizaremos o Intervalo de Confiança descrito no Apêndice I.

Capítulo 5

Distribuições Assimétricas

5.1 Introdução

Serão geradas distribuições assimétricas (conhecidas como *skew*) a partir daquelas apresentadas no Capítulo 4. Os resultados que geram distribuições não simétricas a partir de distribuições simétricas podem ser encontrados em [Azzalini, 1985] (1) e [Fernandez et al, 1998](7).

5.2 Distribuição Assimétrica $\mathbf{h}_1(x)$

5.2.1 Construção de $h_1(x)$

Utilizaremos o seguinte resultado de [Azzalini, 1985] para gerar uma distribuição assimétrica a partir de uma distribuição simétrica:

$h_1(x) = 2g(x)G(w(x)), -\infty < x < \infty$, onde $g(x)$ é a função de densidade simétrica em torno da origem dada por (4.2), $G(x)$ é dada por (4.1), e $w(x)$ é função ímpar de x.

Fazendo a escolha da função ímpar $w(x) = \gamma x$, em que $-\infty < \gamma < \infty$, temos:

$$h_1(x) = 2\left[\frac{\rho|x|^{2\rho-1}(B|x|+C)}{(x^2+2B|x|+C)^{\rho+1}}\right]\left[\frac{1}{2}+\frac{1}{2}\frac{\gamma x|\gamma x|^{2\rho-1}}{((\gamma x)^2+2B|\gamma x|+C)^{\rho}}\right]$$

$$h_1(x) = \left[\frac{\rho|x|^{2\rho-1}(B|x|+C)}{(x^2+2B|x|+C)^{\rho+1}}\right] + \left[\frac{\rho\gamma|\gamma|^{2\rho-1}x|x|^{4\rho-2}(B|x|+C)}{[(x^2+2B|x|+C)^{\rho+1}][((\gamma x)^2+2B|\gamma x|+C)^{\rho}]}\right] \tag{5.1}$$

Utilizando um parâmetro de locação μ e parâmetro de escala $\sigma > 0$, temos:

$$Sas(x) = \left[\frac{\rho|\left(\frac{x-\mu}{\sigma}\right)|^{2\rho-1}(B|\left(\frac{x-\mu}{\sigma}\right)|+C)}{(\left(\frac{x-\mu}{\sigma}\right)^2+2B|\left(\frac{x-\mu}{\sigma}\right)|+C)^{\rho+1}}\right] +$$

$$\left[\frac{\rho\gamma|\gamma|^{2\rho-1}\left(\frac{x-\mu}{\sigma}\right)|\left(\frac{x-\mu}{\sigma}\right)|^{4\rho-2}(B|\left(\frac{x-\mu}{\sigma}\right)|+C)}{[(\left(\frac{x-\mu}{\sigma}\right)^2+2B|\left(\frac{x-\mu}{\sigma}\right)|+C)^{\rho+1}][((\gamma\left(\frac{x-\mu}{\sigma}\right))^2+2B|\gamma\left(\frac{x-\mu}{\sigma}\right)|+C)^{\rho}]}\right] \tag{5.2}$$

5.2.2 Gráfico para a Densidade Assimétrica $Sas(x)$

Notamos que a densidade pode assumir a forma bimodal, além de possuir curvas com caudas pesadas. Segue o gráfico para a densidade $h_1(x)$ para alguns valores de $\Theta = (B, C, \rho, \gamma, \mu, \sigma)$:

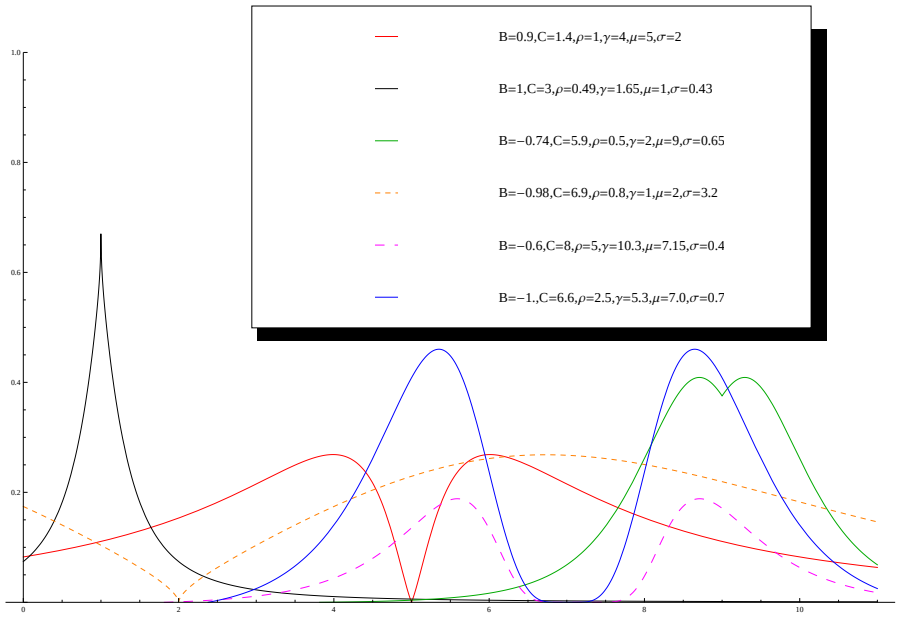

Figura 5.1: Formas para a Densidade Assimétrica $Sas(x)$

5.3 Distribuição Assimétrica $h_2(x)$

5.3.1 Construção de $h_2(x)$

Utilizaremos o seguinte resultado para gerar uma distribuição assimétrica a partir de uma distribuição simétrica:

$h_2(x) = \frac{2}{\alpha+\frac{1}{\alpha}} g(x\alpha^{sinal[x]})$, onde $sinal[x] = \frac{|x|}{x}$, $\alpha > 0, -\infty < x < \infty$, $g(x)$ é dada por (4.2). Logo,

$$h_2(x) = \left[\frac{2\alpha}{\alpha^2+1}\right]\left[\frac{\rho|x\alpha^{sinal[x]}|^{2\rho-1}(B|x\alpha^{sinal[x]}|+C)}{(x\alpha^{sinal[x]})^2+2B|x\alpha^{sinal[x]}|+C)^{\rho+1}}\right] \tag{5.3}$$

Utilizando um parâmetro de posição μ e parâmetro de escala $\sigma > 0$ em (5.3), temos:

$$Sbs(x) = \left[\frac{2}{\alpha+\frac{1}{\alpha}}\right]\left[\frac{\rho|\left(\frac{x-\mu}{\sigma}\right)\alpha^{sinal[\left(\frac{x-\mu}{\sigma}\right)]}|^{2\rho-1}(B|\left(\frac{x-\mu}{\sigma}\right)\alpha^{sinal[\left(\frac{x-\mu}{\sigma}\right)]}|+C)}{(\left(\frac{x-\mu}{\sigma}\right)\alpha^{sinal[\left(\frac{x-\mu}{\sigma}\right)]})^2+2B|x\alpha^{sinal[\left(\frac{x-\mu}{\sigma}\right)]}|+C)^{\rho+1}}\right] \tag{5.4}$$

5.3.2 Gráfico para a Densidade Assimétrica $Sbs(x)$

Notamos que a densidade pode assumir a forma bimodal, além de possuir curvas com caudas pesadas também. Segue o gráfico para a densidade $h_2(x)$ para alguns valores de $\Theta = (B, C, \rho, \alpha, \mu, \sigma)$:

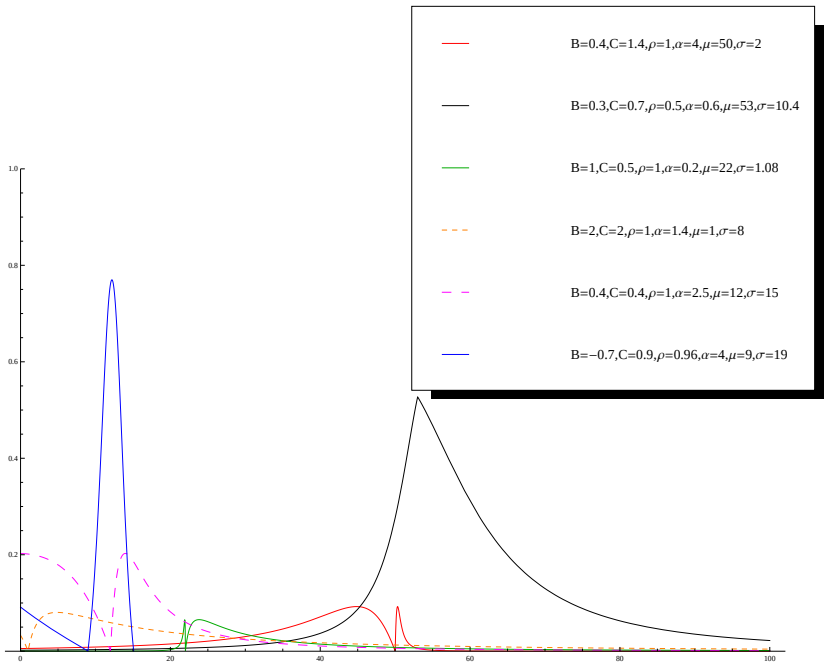

Figura 5.2: Formas para a Densidade Assimétrica $Sbs(x)$

Capítulo 6

Aplicações

6.1 Introdução

Serão apresentados conjuntos de dados onde foram aplicadas as distribuições. Serão realizados os ajustes das distribuições e densidades, logo serão calculados os valores dos parâmetros. A primeira aplicação trata de dados de precipitação, onde utilizaremos a forma assimétrica da densidade introduzida em (3.3). O segundo exemplo trará dados de retornos de taxas de câmbio do Brasil e Canadá, onde realizaremos o ajuste para a versão simétrica apresentada em (4.2). O terceiro exemplo trará dados de log-consumo de energia, onde aplicaremos a densidade assimétrica derivada a partir da forma de [Azzalini,1985] presente em (5.2). A análise estatística foi realizada no *software R*.

6.2 Precipitação

6.2.1 Análise Descritiva

Os dados foram obtidos no sítio *http://www.weather.gov/* (12). São dados anuais de precipitação(chuva) entre 1878 e 1998, para o centro da cidade de Los Angeles, com 121 observações. Segundo o sítio (12), o ano hidrológico termina em 30 de junho de cada ano. A precipitação observada é um subproduto do National Weather Service (NWS) dos Estados Unidos da América. Este conjunto de dados foi utilizado por [Rathie & Coutinho,2011] (15) e os resultados serão comparados.

Foram calculadas as estatísticas descritivas, bem como o *boxplot*. A distribuição empírica dos dados, assim como a densidade foram construídas com *software R* por meio dos comandos *ecdf* e *density*.

Tabela 6.1: Estatísticas Descritivas para os Dados de Precipitação

Estatística	*Valor*
Mínimo	4,85
1o. Quartil	10,40
Mediana	13,07
Média	15,11
Variância	46,98
Desvio-Padrão	6,85
3o. Quartil	19,28
Máximo	38,18

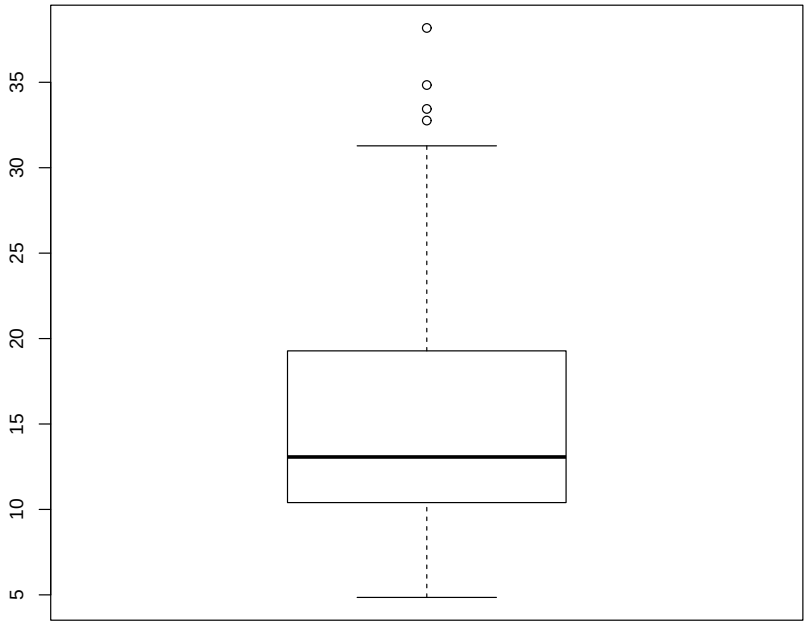

Figura 6.1: Boxplot para os Dados de Precipitação

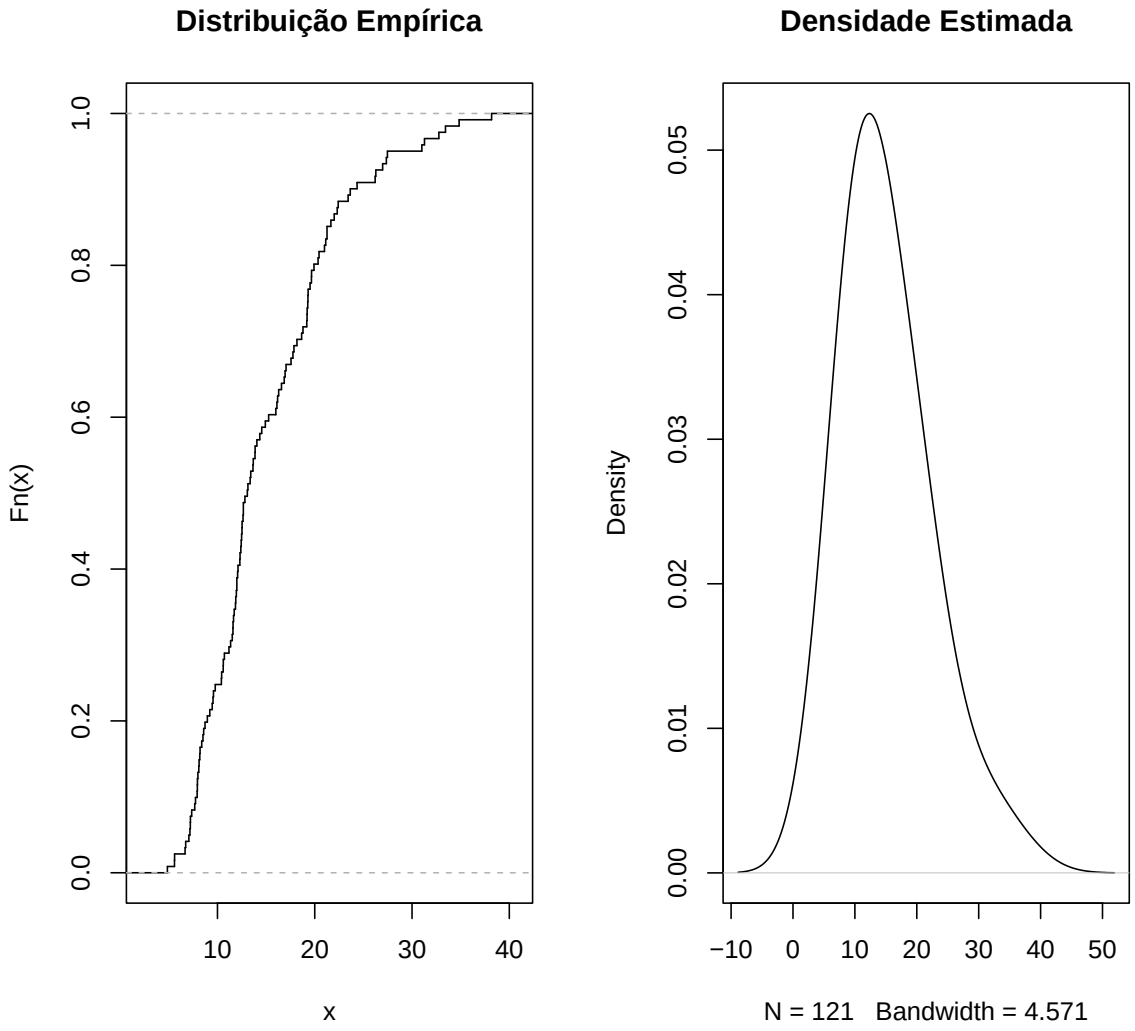

Figura 6.2: Distribuição Empírica e Densidade Estimada por Kernel para os Dados de Precipitação

Observamos pelo *boxplot* (6.1), que a distribuição é muito assimétrica à direita. Foi verificada a hipótese de normalidade dos dados por meio do teste Shapiro-Wilks, que apontou $W = 0,9211$ e $p - valor = 0,00000256$. Logo, existem fortes evidências de que os dados não possuem distribuição Normal.

6.2.2 Estimação dos Parâmetros

O Método da Máxima Verossimilhança foi empregado para estimação dos parâmetros B, C, ρ. A otimização da função de log-verossimilhança foi realizada com o auxílio do pacote *DEoptim* (11) no *software R*. Outros métodos como o *nlm*,*nlminb*, *optim* e o

pacote *rgenoud* foram utilizados, mas não encontraram soluções satisfatórias.

Para a estimação do intervalo de confiança a 95%, utilizamos o método *Bootstrap* não-paramétrico baseado em 1.000 amostras com reposição, por meio do pacote *boot*. A seguir apresentamos os resultados:

Tabela 6.2: E.M.V. e I.C. 95% Bootstrap para os Parâmetros B, C, ρ

Parâmetro	*Estimativa*	I.C.Bootstrap Normal 95%	I.C.Bootstrap Básico 95%
B	-0,35	[-0,3844; -0,3427]	[-0,3816; -0,3447]
C	28,00	[27,34; 29,41]	[27,44; 29,28]
ρ	7,27	[6,759; 7,798]	[6,816; 7,752]

Para garantir que $\hat{B}, \hat{C}, \hat{\rho}$ é E.M.V., calculamos a matriz Hessiana, denotada por $H(\underline{\theta})$ e verificamos no *software Mathematica* que é definida negativa:

$$\begin{bmatrix} -65,00 & -3,16 & -16,13 \\ -3,16 & -0,19 & -0,72 \\ -16,13 & -0,72 & -2,28 \end{bmatrix}$$

6.2.3 Ajuste da Curva

Calculamos o teste de bondade de ajuste Kolmogorov-Smirnov, onde a máxima distância ou máximo desvio encontrado foi de $D = 0,0496$ e $p - valor = 0,9984$, indicando que houve um bom ajuste, não se rejeita a hipótese de que os dados seguem a distribuição simétrica apresentada em . A distribuição e densidade são, respectivamente, da forma:

$$F(x) = \frac{x^{14,54}}{(x^2 - 0.7x + 28)^{7,27}} \tag{6.1}$$

$$f(x) = \frac{14,54x^{13,54}(-0,35x + 28)}{(x^2 - 0,7x + 28)^{8,27}} \tag{6.2}$$

Foram calculadas medidas de acurácia do ajuste: Erro Quadrático Médio (MSE), Desvio Médio Absoluto (MAD) e Desvio Máximo Absoluto (MD). A medida MSE (Mean Square Error) consiste na média do quadrado da diferença entre os valores da distribuição empírica e valores da distribuição ajustada.A medida MAD (Mean Absolute Deviation) consiste na média do módulo da diferença entre os valores da distribuição empírica e valores da distribuição ajustada. E, finalmente, a medida MD (Maximum Deviation) consiste no valor máximo entre as diferenças absolutas para a distribuição empírica e a distribuição ajustada. Apresentamos as medidas:

Tabela 6.3: Medidas de Acurácia de Ajuste para as Distribuição dos Dados de Precipitação

Medidas de Acurácia	Desvios
MSE	0,0006507454
MAD	0,02188007
MD	0,04724667

Observamos um bom ajuste dos dados à distribuição apresentada no Capítulo 3. Ao compararmos os resultados encontrados neste trabalho e aqueles apresentados em Rathie e Coutinho(2011) (15), vemos que a acurácia do ajustamento foi semelhante, já que os valores encontrados lá foram MSE=0.000589018, MAD=0.0192187 e MD=0.0547488.

A seguir apresentamos os gráficos com a distribuição e densidade ajustadas aos dados de precipitação do centro de Los Angeles:

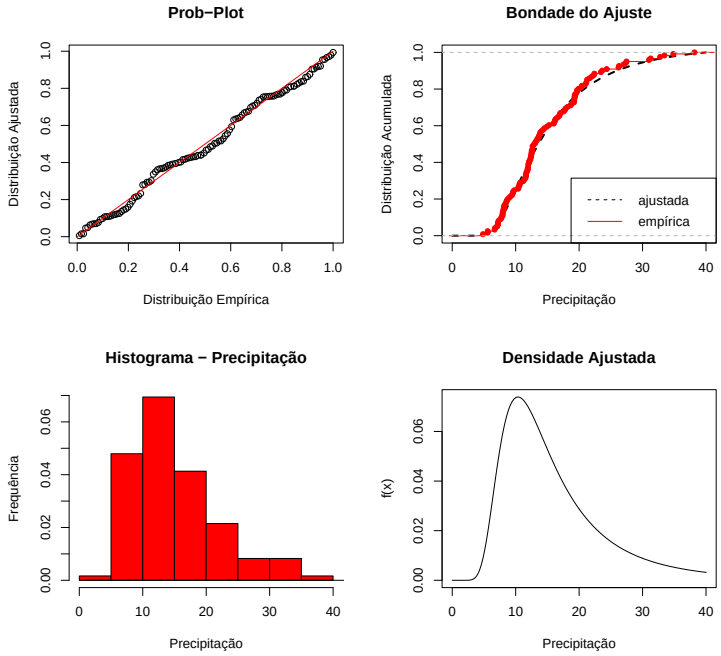

Figura 6.3: Análise de Bondade do Ajuste para os Dados de Precipitação

6.3 Retornos Diários de Taxa Câmbio

6.3.1 Análise Descritiva

Os dados de taxas de câmbio foram obtidos no sítio *http://www.federalreserve.gov/* (5). A informação é baseada em dados coletados e certificados pelo Federal Reserve Bank of New York a partir de uma amostra de participantes no mercado. O período das taxas de cambio diárias encontram-se entre 03 de janeiro de 2000 e 20 de maio de 2011. As moedas pesquisadas foram o Real e Dólar Canadense que pertencem aos países Brasil e Canadá, respectivamente. Segundo o FDR (5), os dados são mensurados ao meio-dia pelas taxas de compra em Nova York para transferências a pagar em moeda estrangeira. Os retornos diários r_i foram calculados pela relação:

$$r_i = \frac{P_i}{P_{i-1}} - 1 \tag{6.3}$$

onde

P_i é o preço ou taxa de câmbio do dia i

P_{i-1} é o preço ou taxa de câmbio do dia anterior i-1

Foram calculadas as estatísticas descritivas, bem como o *boxplot*. A distribuição empírica dos dados, assim como a densidade foram construídas com *software R* por meio dos comandos *ecdf* e *density*, para Brasil e Canadá:

Tabela 6.4: Estatísticas Descritivas para os Dados de Retornos Diários das Taxas de Câmbio do Brasil e Canadá

Estatística	País	
	Brasil	Canadá
Mínimo	-9,2156	-4,9451
1o. Quartil	-0,5143	-0,3363
Mediana	-0,0232	-0,0129
Média	0,0019	0,0119
Variância	1,1570	0,3765
Desvio-Padrão	1,0756	0,6136
3o. Quartil	0,4623	0,3006
Máximo	7,8440	3,8803

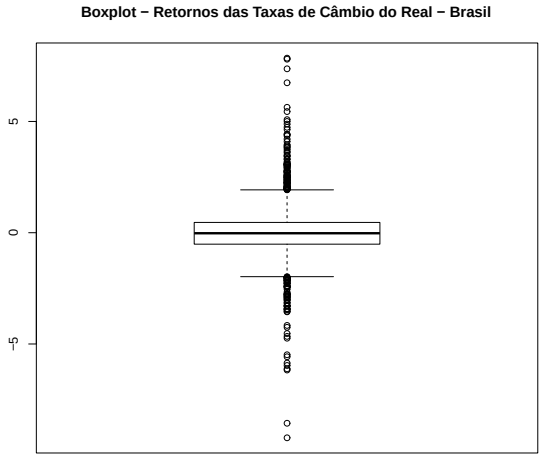

Figura 6.4: Boxplot para Retornos da Taxa de Câmbio - Real - Brasil

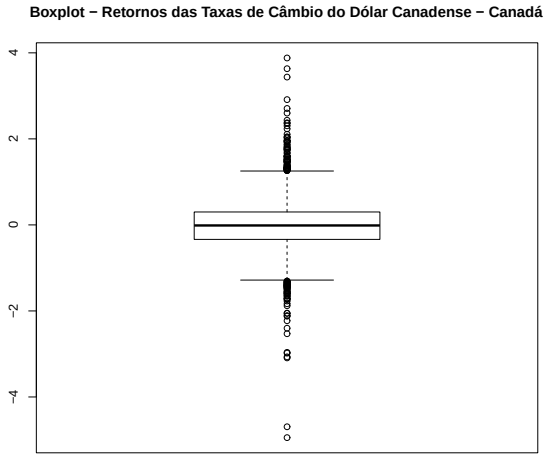

Figura 6.5: Boxplot para Retornos da Taxa de Câmbio - Dólar Canadense - Canadá

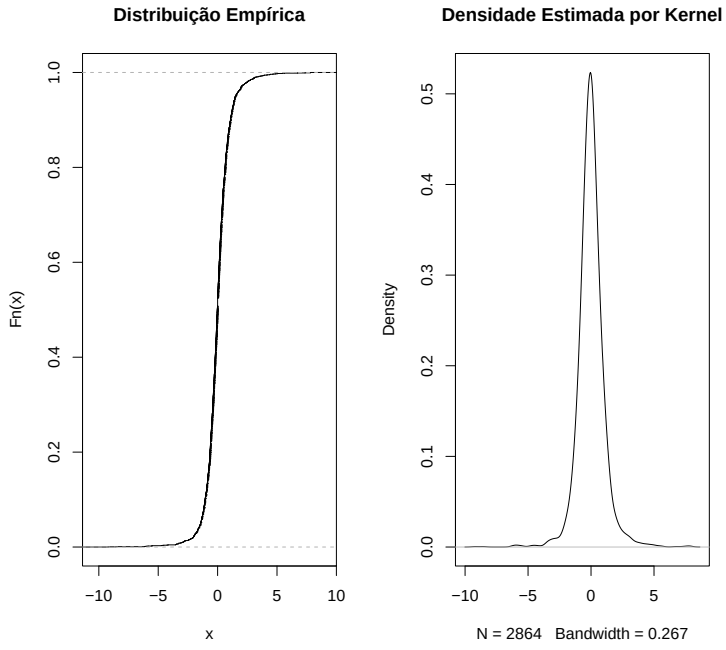

Figura 6.6: Distribuição Empírica e Densidade Estimada para Retornos da Taxa de Câmbio - Real - Brasil

Foi verificada a hipótese de normalidade dos retornos de taxas cambiais por meio do teste Shapiro-Wilks, que apontou $W = 0,8898$ e $p - valor < 2,2.10^{-16}$ para o Brasil (real) e $W = 0,9492$ e $p - valor < 2.2.10^{-16}$ para o Canadá (dólar). Logo, existem evidências de que os dados não possuem distribuição Normal para ambos os casos.

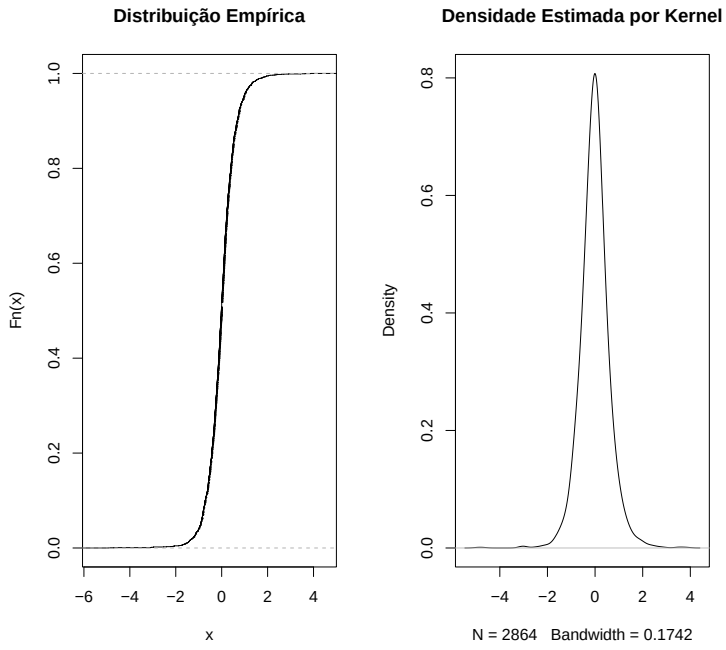

Figura 6.7: Distribuição Empírica e Densidade Estimada para Retornos da Taxa de Câmbio - Dólar Canadense - Canadá

6.3.2 Estimação dos Parâmetros

O Método da Máxima Verossimilhança foi empregado para estimação dos parâmetros B, C. Fixaremos o parâmetro $\widehat{\rho} = \frac{1}{2}$, pois pela análise gráfica da densidade estimada por kernel dos dados de retornos de taxas de câmbio, observamos que os dados possuem curva unimodal, tanto para o Brasil quanto para o Canadá. Para a análise gráfica da densidade simétrica apresentada em (4.2), a única hipótese em que a densidade apresenta forma unimodal é quando $\rho = \frac{1}{2}$.

A otimização da função de log-verossimilhança foi realizada com o auxílio do pacote *DEoptim* no *R*. Para garantir que $\hat{B}, \hat{C}$ é E.M.V., calculamos a matriz Hessiana e verificamos no *software Mathematica* que é definida negativa, para Brasil e Canadá, respectivamente:

$$\begin{bmatrix} 8478,36 & 6757,35 \\ 6757,35 & 3283,16 \end{bmatrix}$$

e

$$\begin{bmatrix} 50537,63 & 34867,27 \\ 34867,27 & 21929,15 \end{bmatrix}$$

Utilizamos intervalos de confiança de Wald, baseados na Matrizes Inversas de Informação de Fisher, apresentadas para o real e o dólar canadense, respectivamente:

$$\begin{bmatrix} 0.00018 & -0.0003 \\ -0.0003 & 0.0004 \end{bmatrix}$$

e

$$\begin{bmatrix} 0.0002 & -0.0003 \\ -0.0003 & 0.0004 \end{bmatrix}$$

A seguir apresentamos os resultados do E.M.V., com quatros casas de precisão:

Tabela 6.5: E.M.V. e I.C. Wald 95% para os Parâmetros B, C

Parâmetro	*Câmbio*	*Estimativa*	I.C.Wald 95%
B	Real	0,05	[0,0234;0,0765]
	Dólar Canadense	-0,01	[-0.0379;0.0179]
C	Real	0.6193	[0,5792;0,6647]
	Dólar Canadense	0.2810	[0.2385; 0,3235]

6.3.3 Ajuste da Curva

Calculamos o teste de bondade de ajuste Kolmogorov-Smirnov, onde a máxima distância encontrada entre a distribuição empírica e a distribuição ajustada foi de

$D = 0,0258$ com $p - valor = 0,2946$ para o Brasil e $D = 0,0269$ com $p - valor = 0.2518$ para o Canadá. Isto é, não existem evidências de que a distribuição estimada e a distribuição empírica sejam diferentes. Apresentamos adiante medidas de acurácia do ajuste, a saber: Erro Quadrático Médio (MSE), Desvio Médio Absoluto (MAD) e Desvio Máximo Absoluto (MD).

Tabela 6.6: Medidas de Acurácia de Ajuste para as Distribuições dos Retornos das Taxas Cambiais de Brasil e Canadá

Câmbio	*MedidasdeAcurácia*	*Desvios*
Real	MSE	0,0001986364
	MAD	0,01209738
	MD	0,02546275
DólarCanadense	MSE	0,0001775864
	MAD	0,01138909
	MD	0,02671749

A seguir apresentamos os gráficos com a distribuição e densidade ajustadas aos dados dos retornos das taxas cambiais de Brasil e Canadá, respectivamente:

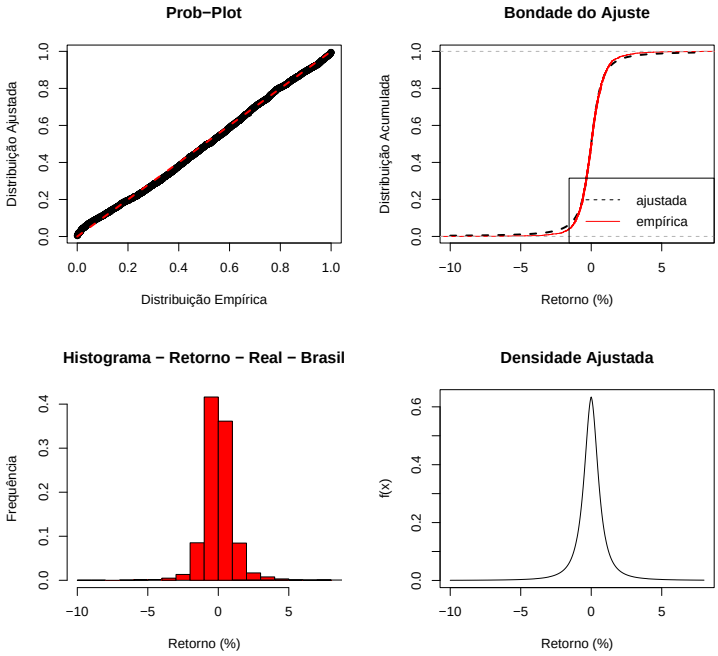

Figura 6.8: Análise de Bondade de Ajuste para Retornos Diários da Taxa de Câmbio - Real - Brasil

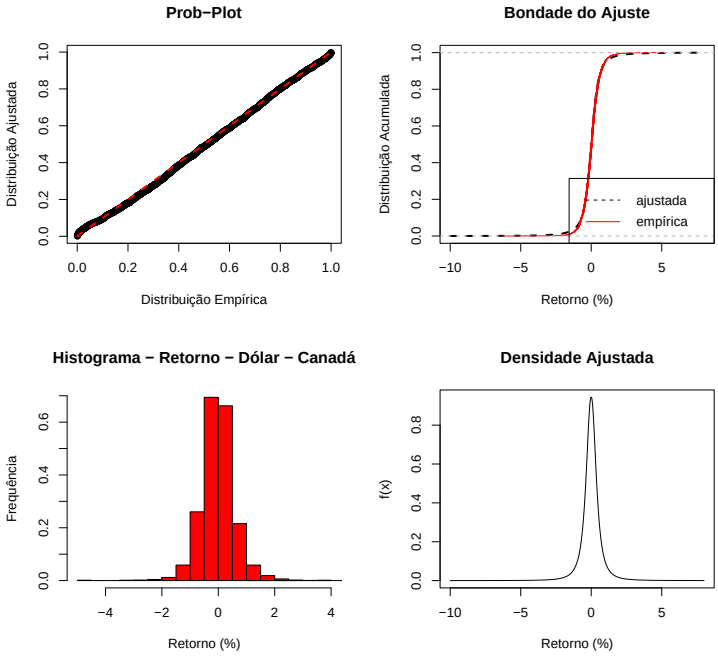

Figura 6.9: Análise de Bondade de Ajuste para Retornos Diários da Taxa de Câmbio - Dólar Canadense - Canadá

6.4 Consumo de Energia - kg de petróleo per capita

6.4.1 Análise Descritiva

O sítio *http://data.worldbank.org/indicator/EG.USE.PCAP.KG.OE* (21) contém os dados de consumo de energia. O consumo de energia refere-se a utilização da energia primária antes da transformação para outros combustíveis de uso final, que é igual à produção indígena, mais as importações e as mudanças de estoque, menos as exportações e combustíveis fornecidos a navios e aeronaves em tráfego internacional para 136 países no ano de 1997. Os dados estão na escala de quilos de petróleo *per capita*. Para nossa análise, transformamos a escala para log-consumo de energia.

Foram calculadas as estatísticas descritivas, bem como o *boxplot*. A distribuição empírica dos dados, assim como a densidade foram construídas com *software R* por meio dos comandos *ecdf* e *density*.

Tabela 6.7: Estatísticas Descritivas para os Dados de Log-Consumo de Energia

Estatística	*Valor*
Mínimo	4,832
1o. Quartil	6,347
Mediana	7,185
Média	7.225
Variância	1,153
Desvio-Padrão	1,073
3o. Quartil	8,091
Máximo	9,847

Observamos pela figura 6.11 pela densidade que a distribuição é bimodal.

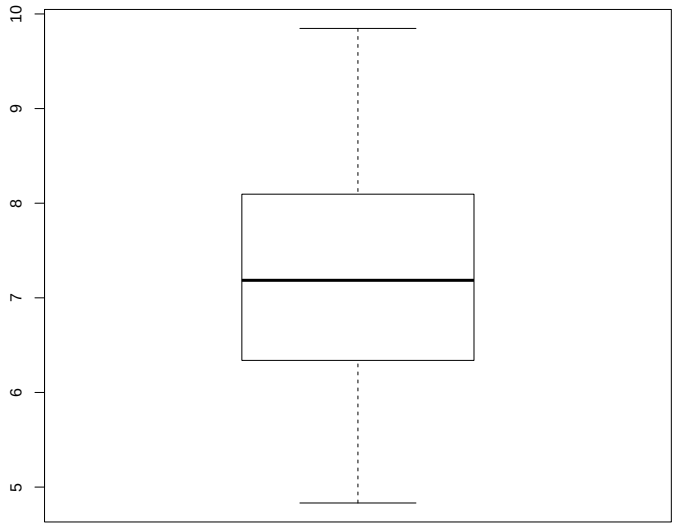

Figura 6.10: Boxplot para os Dados de Log-Consumo de Energia

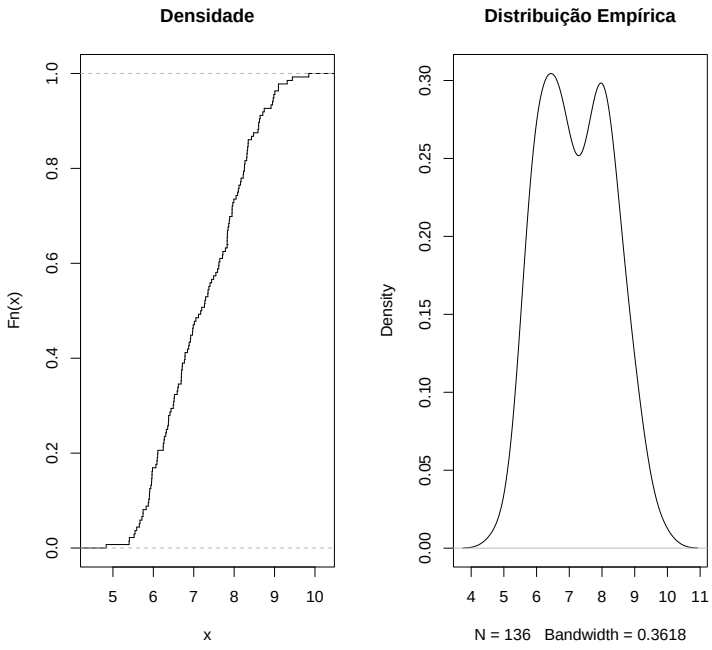

Figura 6.11: Distribuição Empírica e Densidade Estimada por Kernel para os Dados de Log-Consumo de Energia

6.4.2 Estimação dos Parâmetros

O Método da Máxima Verossimilhança foi empregado para estimação dos parâmetros $B, C, \gamma, \mu, \sigma$. Fixamos o parâmetro $\rho = \frac{1}{2}$. A otimização da função de log-verossimilhança

foi realizada com o auxílio do pacote *DEoptim* (11) no *software R*. Outros métodos como o *nlm*,*nlminb*, *optim* e o pacote *rgenoud* foram utilizados, mas não encontraram soluções satisfatórias. A estimação realizada para os parâmetros foi pontual. A seguir apresentamos os resultados:

Tabela 6.8: E.M.V. para os Parâmetros $B, C, \gamma, \mu, \sigma$

Parâmetro	*Estimativa*
B	-1,00
C	5.90
γ	0,81
μ	7,00
σ	0,64

6.4.3 Ajuste da Curva

Calculamos o teste de bondade de ajuste Kolmogorov-Smirnov, onde a máxima distância encontrada entre a distribuição empírica e a distribuição ajustada foi de $D = 0,0809$ com $p-valor = 0,7653$, o que indica que a hipótese de semelhança entre as distribuições empírica e ajustada não foi rejeitada. Foram calculadas medidas de acurácia do ajuste: Erro Quadrático Médio (MSE), Desvio Médio Absoluto (MAD) e Desvio Máximo Absoluto (MD). Apresentamos as medidas:

Tabela 6.9: Medidas de Acurácia de Ajuste para as Distribuição dos Dados de Log-Consumo de Energia

Medidas de Acurácia	Desvios
MSE	0,001026046
MAD	0,0251874
MD	0,07521916

A seguir apresentamos os gráficos com a distribuição e densidade ajustadas aos dados de consumo de energia:

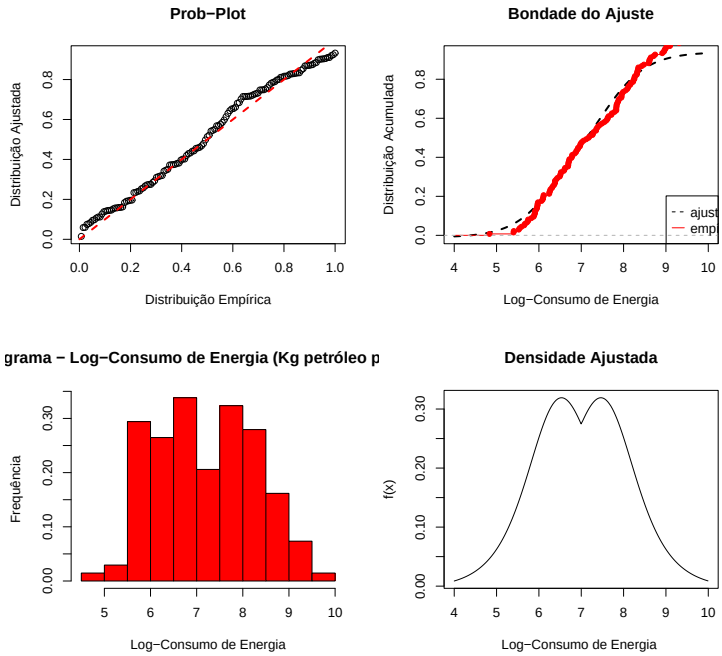

Figura 6.12: Análise de Bondade do Ajuste para os Dados de Log-Consumo de Energia

Capítulo 7

Conclusões

Os capítulos anteriores caracterizam as distribuições assimétricas e simétricas. Pela quantidade de parâmetros, pode-se observar pelos gráficos que as distribuições são maleáveis e podem assumir diversos formatos, inclusive podem assumir formas com caudas grossas.

Um fato interessante é que ambas não possuem momentos inteiros positivos, geralmente utilizados na Inferência Estatística Clássica. Os momentos e funções características possuem resultados em funções especias que podem ser computados em *softwares* matemáticos como *Maple* e *Mathematica*.

Observamos que tanto para o caso simétrico, quanto para o caso assimétrico, foram encontrados dados reais em que as distribuições tiveram aplicações. A distribuição ajustada obteve resultado excelente para os dados de precipitação, fortemente assimétricos, apesar das poucas observações. Para os dados de retornos percentuais de taxa de câmbio do real e dólar canadense em relação ao dólar americano, o resultado do ajuste foi ótimo.

A distribuição assimétrica do Capítulo 5 obtida a partir da distribuição simétrica do Capítulo 4 também gerou bons resultados e tem a vantagem de modelar dados bimodais e fortemente assimétricos, com caudas pesadas. A diferença entre a distribuição empírica e distribuição ajustada poderia ser menor se houvessem mais observações.

As distribuições apresentadas mostraram grande aplicabilidade aos dados reais, principalmente em áreas com um comportamento típico de caudas pesadas

como dados financeiros e dados assimétricos e bimodais como dados ecológicos e ambientais.

Referências Bibliográficas

[1] Azzalini, A., 1985. "A class of distrubutions which includes the normal ones." *Scandinavian J. Statistics,* ***12***, 171-178.

[2] Bolfarine, H. e Sandoval, M. C., 2002. "Introdução à Inferência Estatística." *Sociedade Brasileira de Matemática.*

[3] Canty, A. and Ripley,B. D., 2011. "boot: Bootstrap R (S-Plus) Functions" R package version 1.3-1.

[4] Davison, A. C. and Hinkley, D. V., 1997 "Bootstrap Methods and Their Applications" *Cambridge University Press.*

[5] Federal Reserve web site. Disponível em *http://www.federalreserve.gov/.*

[6] Feller, W., 1976. "An Introduction to Probability Theory and Its Applications, vol I, II." *John Wiley & Sons, New York.*

[7] Fernandez, C. and Steel, M. F. J., 1998. "On Bayesian modeling of fat tails and skewness." *JASA 93*, 359-371.

[8] Hai, N. T. and Yakubovich, S. B., 1992. "The Double Mellin-Barnes Type Integrals and their Applications to Convolution Theory." *World Scientific Publishing Co. Pte. Ltd, Singapore.*

[9] Mathai, A. M., Saxena, R. K. and Haubold, H. J., 2010. "H-Function - Theory and Applications". *Springer Science and Business, New York.*

[10] Mood, A. M., Graybill, F. A. and Boes, D. C., 1974. "Introduction to the Theory of Statistics". *McGraw-Hill Inc., New York.*

[11] Mullen, K., Ardia, D., Gil, D.,Windover, D., and Cline, J., 2011. "DEoptim: An R Package for Global Optimization by Differential Evolution", *Journal of Statistical Software,* v. 40, p. 1-26. Disponível em *http://www.jstatsoft.org/v40/i06/.*

[12] National Oceanic and Atmospheric Administration's Weather Service web site *http://www.weather.gov/.*

[13] Prudnikov, A.P., Brychkov, Y.A. and Marichev, O.I., 1986. "Integrals and Series (3 vol.)". *Gorgon & Breach Science Publishers, Amsterdam.*

[14] Rathie, P. N., Swamee, P. K., Matos, G.G. , Coutinho, M. and Carrijo, T.B., 2008. "H-Functions and Statistical Distributions". *Ganita*, v. 59, p. 23-37.

[15] Rathie, P. N., and Coutinho, M., 2011. "A new Skew Generalized Logistic distribution and Approximations to Skew Normal Distribution. *The Aligarh Journal of Statistics*. No Prelo.

[16] Rathie, P. N., Dorea, C. C. Y., Otiniano, C. E. G. and Matsushita, R . L, 2007. "Lévy flight approximations for scaled transformations of random walks". *Computational Statistics & Data Analysis*, v. 51, p. 6343-6354.

[17] Slater, L. J., 1966. "Generalized Hypergeometric Function". *Cambridge University Press, London.*

[18] Springer, M. D., 1979. "The Algebra of Random Variables". *John Wiley & Sons, New York.*

[19] Temme, N. M., 1996. "Special Functions: An introduction to the classical functions of mathematical physics". *John Wiley & Sons, New York.*

[20] The Wolfram Functions web site *http://functions.wolfram.com.*

[21] World Bank Group web site. Disponível em http://data.worldbank.org/indicator/EG.USE.PCAP.KG.OE.

Apêndice I

Conceitos Básicos em Probabilidade e Inferência Estatística

Para analisar o problema proposto serão necessários alguns conceitos de Probabilidade e Inferência Estatística. No estudo de Probabilidade serão introduzidas definições de Distribuição de Probabilidade, Momentos, Função Geratriz de Momentos e Função Característica. Estes conceitos serão necessários para montar várias propriedades das distribuições que serão apresentadas.

Por outro lado, após apresentação das propriedades das novas distribuições é indispensável a estimação dos parâmetros destas, onde serão necessários conceitos a Máxima Verossimilhança, Método do Momentos e Intervalos de Confiança. Breves definições e conceitos serão apresentados a seguir.

Probabilidade

A palavra probabilidade deriva do Latim *probare* (provar ou testar). O estudo da probabilidade está interessado em observar fenômenos cotidianos aleatórios. Note que ao realizar experimentos aleatórios, não se sabe qual será o resultado exato antes de realizá-lo, mas sim quais são os resultados prováveis.

Distribuição de Probabilidade Acumulada

Definição 7.0.1. *Seja X uma variável aleatória (v.a.) em um espaço de probabilidade. A função de distribuição de probabilidade acumulada de X, denotada por F_X ou simplesmente F, é definida por*

$$F_X(x) = P(X \leq x),\ x \in \mathbb{R} \tag{7.1}$$

Proposição 7.1. ***Propriedades da Função de Distribuição Acumulada***

i) $0 \leq F_X(x) \leq 1, \forall x$*;*

ii) $\lim_{x \to -\infty} F_X(x) = 0$ *e* $\lim_{x \to \infty} F_X(x) = 1$*;*

iii) F_X *é contínua à direita;*

iv) F_X *não é decrescente, isto é,* $F_X(x) \leq F_X(y), \forall x < y; x, y \in \mathbb{R}$*.*

Função Densidade de Probabilidade

Definição 7.0.2. *Seja X uma variável aleatória em um espaço de probabilidade. X é denominada* ***discreta*** *se assume um número finito ou enumerável de valores. Neste caso, a função densidade de probabilidade de uma v.a. discreta, denotada por* $p(x_i)$*, é definida por*

$$p(x_i) = P(X = x_i)\ e$$

$$F_X(x) = \sum_{i:x_i \leq x} P(X = x_i) = \sum_{i:x_i \leq x} p(x_i) \tag{7.2}$$

Proposição 7.2. ***Propriedades da Função Densidade de Probabilidade de uma V.A. Discreta***

i) $p(x_i) \geq 0, x \in \mathbb{R}$*;*

ii) $\{x_i : p(x_i) \neq 0\}$ *é subconjunto finito ou infinito enumerável de R;*

iii) $\sum_i p(x_i) = 1$*.*

Definição 7.0.3. *Seja X uma variável aleatória em um espaço de probabilidade. X é denominada* ***contínua*** *se sua função densidade de probabilidade (****f.d.p.****), denotada por* $f(x)$*, é tal que*

$$F_X(x) = \int_{-\infty}^{x} f(t)dt; \forall x \in \mathbb{R} \tag{7.3}$$

Proposição 7.3. ***Propriedades da Função Densidade de Probabilidade de uma V.A. Contínua***

i) $f(x) \geq 0, x \in \mathbb{R}$*;*
ii) $\int_{-\infty}^{\infty} f(x)dx = 1$*;*
iii) $f(x) = \frac{dF(x)}{dx}$.

Esperança Matemática

Definição 7.0.4. *Seja X uma variável aleatória qualquer e F sua distribuição em um espaço de probabilidade. A* ***esperança*** *de X, denotada por EX, é definida por*

$$EX = \int_{-\infty}^{\infty} xdF(x), \tag{7.4}$$

sempre que a integral de Riemann-Stieltjes é bem definida e $E|X|$ *existe. Como* $EX = \int_{-\infty}^{\infty} xdF(x) = \int_{-\infty}^{0} xdF(x) + \int_{0}^{\infty} xdF(x)$*, então a esperança estará bem definida se* $\int_{-\infty}^{0} xdF(x)$ *e* $\int_{0}^{\infty} xdF(x)$ *forem finitas.*

Momentos

Definição 7.0.5. *Seja X uma varável aleatória em um espaço de probabilidade. O* ***k-ésimo momento de X****, denotado por* μ_k'*, é definido por*

$$\mu_k' = EX^k, \tag{7.5}$$

desde que $E|X|^k$ *exista. O* ***k-ésimo momento central de X*** *(em torno da média), denotado por* μ_k*, é definido por*

$$\mu_k = E(X - EX)^k, \tag{7.6}$$

desde que $E|X - EX|^k$ *exista.*

Função Geratriz de Momentos

Definição 7.0.6. *Seja X uma varável aleatória em um espaço de probabilidade. A* ***função geratriz de momentos*** *de X, denotada por $M_X(t)$, é definida por*

$$M_X(t) = Ee^{tX}, \tag{7.7}$$

desde que a esperança exista.

$M_X(t)$ é chamada de função geratriz de momentos pois $EX^k = \frac{d^k}{dt^k}(M_X(t))|_{t=0}$.

Função Característica

Definição 7.0.7. *Seja X uma varável aleatória em um espaço de probabilidade. A* ***função característica****, denotada por $\varphi_X(t)$, é a função $\varphi : \mathbb{R} \to \mathbb{C}$ definida por*

$$\varphi_X(t) = Ee^{itX}; \tag{7.8}$$

onde i é o número imaginário $\sqrt{-1}$.

Proposição 7.4. ***Propriedades da Função Característica***
i) $\varphi_X(t)$ é limitada, isto é, $|\varphi_X(t)| \leq 1, \forall t \in \mathbb{R}$;
ii) $\varphi_X(0) = 1$;
iii) $\varphi_X(t)$ é uniformemente contínua na reta;
iv) A função característica de uma v.a. X determina a distribuição de X unicamente;
v) Se $E|X| < \infty$, então φ_X possui k derivadas contínuas e $\frac{d^k}{dt^k}(\varphi_X(t))|_{t=0} = i^k EX^k$.

Note que, pelas propriedades acima apresentadas, a *função característica* sempre existe.

Inferência Estatística

A inferência estatística tem por objetivo fazer afirmações sobre uma população ou universo a partir de uma amostra. Tal tipo de afirmação deve sempre vir acompanhada de uma medida de precisão. Os estimadores estão ligados à idéia de fazer

afirmações de sobre a população e intervalos de confiança estão associados à idéia de precisão. A seguir serão apresentados conceitos relacionados à estimação de parâmetros e intervalos de confiança. O leitor deverá estar familiarizado com as ferramentas da Inferência como espaço paramétrico, estimadores, estatísticas suficientes e completas. Tais fundamentos podem ser encontrados em Bolfarine[2002] etc.

Métodos de Estimação

Serão apresentados métodos que possibilitam a estimação dos parâmetros nas distribuições que serão estudadas: o Método dos Momentos e o Método da Máxima Verossimilhança.

Método dos Momentos

O **Método dos Momentos** consiste em igualar momentos amostrais aos momentos populacionais.

Definição 7.0.8. *Seja o* ***k-ésimo momento amostral****, denotado por* m_k*, definido por*

$$m_k = \frac{1}{n}\sum_{i=1}^{n} X_i^k, \tag{7.9}$$

onde n é o número de elementos de uma amostra $X_1, X_2, \ldots, X_n$*.*

Logo, o Método dos Momentos assume que

$$m_k = \mu_k$$

$$\frac{1}{n}\sum_{i=1}^{n} X_i^k = EX^k \tag{7.10}$$

Método da Máxima Verossimilhança

Definição 7.0.9. *Sejam $X = (X_1, X_2, \ldots, X_n)$ uma amostra aleatória da v.a. X com f.d.p. $f(x|\theta)$ em que $\theta \in \Theta$ é um vetor $(\theta_1, \theta_2, \ldots, \theta k)$, onde Θ é o espaço paramétrico e k é a quantidade de parâmetros que a f.d.p. possui. A* ***função de verossimilhança****, denotada por $L(\theta; X)$, é definida por*

$$L(\theta; X) = \prod_{i=1}^{n} f(x_i|\theta) \tag{7.11}$$

Definição 7.0.10. *O* ***estimador de máxima verossimilhança (E.M.V.)****, denotado pelo vetor $\hat{\theta}$, é alcançado por meio da maximização de $L(\theta; X)$.*

$$L(\hat{\theta}) = max_{\hat{\theta} \in \Theta} L(\theta; X) \tag{7.12}$$

A seguir é apresentado o procedimento para obtenção do E.M.V.

i) Obtém-se o logaritmo natural de $L(\theta; X)$

$$l(\theta; X) = logL(\theta; X) \tag{7.13}$$

ii) Deriva-se a expressão $l(\theta; X)$ em relação a cada θ_i e iguala-se a 0

$$U_i(\theta) = \frac{\partial l(\theta; X)}{\partial \theta_i} = 0, \tag{7.14}$$

onde $i = 1, \ldots, k$ e $U_i(\theta)$ é conhecida como função escore. Logo teremos

$$\frac{\partial l(\theta; X)}{\partial \theta_1} = 0$$

$$\frac{\partial l(\theta; X)}{\partial \theta_2} = 0$$

.

.

.

$$\frac{\partial l(\theta; X)}{\partial \theta_k} = 0$$

iii) Resolve-se as k equações (7.14), isolando cada $\hat{\theta}_i$.

iv) Verifica-se se a matriz quadrada KxK das derivadas segundas de $l(\theta)$ – $\frac{\partial^2 l(\theta;X)}{\partial\theta\partial\theta^T}$, *denominada matriz Hessiana, definida por*

$$H(l(\theta)) = \begin{bmatrix} \frac{\partial^2 l(\theta;X)}{\partial\theta_1^2} & \frac{\partial^2 l(\theta;X)}{\partial\theta_1\partial\theta_2} & \cdots & \frac{\partial^2 l(\theta;X)}{\partial\theta_1\partial\theta_k} \\ \frac{\partial^2 l(\theta;X)}{\partial\theta_2\partial\theta_1} & \frac{\partial^2 l(\theta;X)}{\partial\theta_2^2} & \cdots & \frac{\partial^2 l(\theta;X)}{\partial\theta_2\partial\theta_k} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial^2 l(\theta;X)}{\partial\theta_k\partial\theta_1} & \frac{\partial^2 l(\theta;X)}{\partial\theta_k\partial\theta_2} & \cdots & \frac{\partial^2 l(\theta;X)}{\partial\theta_k^2} \end{bmatrix} \tag{7.15}$$

é negativa definida.

Note que o E.M.V pode não existir ou pode não ser único. Também observe que em situações em que o espaço paramétrico Θ é discreto ou que o máximo de $l(\theta; X)$ ocorre na fronteira de Θ, o E.M.V deve ser obtido examinando a função de verossimilhança, ao invés de utilizar o procedimento acima descrito.

A propriedade da **Invariância** presente nos E.M.V. é muito vantajosa e será enunciada a seguir.

Teorema 7.0.1. ***(O princípio da Invariância)*** *Seja g(.) função real e inversível definida no espaço paramétrico* Θ. *Então, se* $\underline{\hat{\theta}}$ *é E.M.V. de* θ, *então* $g(\hat{\theta})$ *é o E.M.V. para* $g(\theta)$.

Proposição 7.5. ***Propriedades Assintóticas do E.M.V.***

$$\hat{\theta} \xrightarrow{d} N(\theta, \mathbb{I}^{-1}(\theta))$$

Em que $\mathbb{I}(\theta) = -E\left[\frac{\partial^2 l(\theta;X)}{\partial\theta\partial\theta^T}\right]$ *é denominada a Matriz de Informação de Fisher. Observe que* $\mathbb{I}(\theta)$ *é a matriz hessiana com sinais trocados.*

Intervalos de Confiança (I.C.)

O Intervalo de Confiança de Wald é definido por:

$$\hat{\theta} - z_{\alpha/2}\sqrt{diag(\mathbb{I}^{-1}(\hat{\theta}))}$$

$$\hat{\theta} + z_{1-\alpha/2}\sqrt{diag(\mathbb{I}^{-1}(\hat{\theta}))},$$

em que *diag* é a diagonal da Matriz Inversa de Informação de Fisher e $z_{\alpha/2}, z_{1-\alpha/2}$ são os quantis da distribuição $N(0,1)$ com nível de confiança α.

Outro tipo de I.C. está baseado no método não-paramétrico de *Bootstrap*, descrito a seguir: i)Estimar $\hat{\theta}$ utilizando os dados reais.
ii) Retirar B amostras *bootstrap* de tamanho m, com reposição dos dados reais $(x_1, x_2, \ldots, x_n)$.
iii) Estimar $\hat{\theta}_1, \ldots, \hat{\theta}_B$ das B amostras.
iv) Para estimar o I.C. *Bootstrap* Percentil, ordenar $\hat{\theta}_1, \ldots, \hat{\theta}_B$ e selecionar o percentis $(K_1(\frac{\alpha}{2}), K_2(1-\frac{\alpha}{2}))$, onde K_1 é o percentil de ordem $\frac{\alpha}{2}$ e K_2 é o percentil de ordem $(1-\frac{\alpha}{2})$.
v) Para estimar o I.C. *Bootstrap* Normal, calcular a média $\overline{\hat{\theta}}$ e desvio-padrão $dp(\hat{\theta})$de $\hat{\theta}_1, \ldots, \hat{\theta}_B$ e calcular

$$\overline{\hat{\theta}} - z_{\alpha/2}dp(\hat{\theta})$$

$$\overline{\hat{\theta}} + z_{1-\alpha/2}dp(\hat{\theta})$$

.

Para mais detalhes sobre o método *Bootstrap*, veja (4).