

Universidade de Brasília
Instituto de Ciências Exatas
Departamento de Matemática

**Convergência em Informação de Fisher
Relativa de Somas Parciais para Variáveis
Aleatórias Estáveis**

por

Marta Lizeth Calvache Hoyos

Orientador: Cátia Regina Gonçalves

Brasília

2016

Universidade de Brasília
Instituto de Ciências Exatas
Departamento de Matemática

Convergência em Informação de Fisher Relativa de Somas Parciais para Variáveis Aleatórias Estáveis

por

Marta Lizeth Calvache Hoyos *

*Dissertação apresentada ao Departamento de Matemática da Universidade
de Brasília, como parte dos requisitos para obtenção do grau de*

MESTRE EM MATEMÁTICA

Brasília, 1 de agosto de 2016.

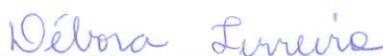
Comissão Examinadora:



Profa. Cátia Regina Gonçalves – MAT/UnB (Orientadora)



Profa. Daniele da Silva Baratela Martins Neto – MAT/UnB (Membro)



Profa. Débora Borges Ferreira - UFRN (Membro)

* O autor foi bolsista do CNPq durante a elaboração desta dissertação.

*Dedico este trabalho à minha família,
pelo amor que sempre me dedicaram
e pelo apoio incondicional.*

Agradecimentos

À Deus, por ser minha fortaleza e meu guia. Aos meus pais Fausto Calvache e Libia Hoyos, por dar-me seu amor, apoio incondicional e por compreender minha ausência. Aos meus irmãos, por acompanhar-me neste processo. À Wenison Silva, por sua compreensão e carinho.

À professora Cátia, pela sua dedicação, paciência e seus valorosos ensinamentos, que foram fundamentais para a realização deste trabalho.

Às professoras Daniele e Débora, por terem aceitado participarem da banca examinadora.

Aos meus amigos pelo apoio para continuar.

Ao CNPq, pelo apoio financeiro durante esse dois anos de mestrado.

Finalmente, a todas as pessoas que de uma ou outra forma estiveram ao meu lado ao longo desta etapa.

Resumo

Neste trabalho estudamos a convergência em Informação de Fisher Relativa, de uma sequência de somas estabilizadas de variáveis aleatórias i.i.d. para uma variável aleatória estável não-extremal, sob a hipótese de convergência em distribuição.

Palavras-chave: Informação de Fisher, Distribuição Estável, Teoremas de Limites e Convergência em Distribuição.

Abstract

In this work, we studied the convergence to a non-extremal stable variable in Relative Fisher Information for stabilized sums of i.i.d. random variables, under the hypothesis of convergence in distribution.

Key-words: Fisher Information, Stable Law, Limit Theorems and Distribution Convergence.

Sumário

Introdução	1
Lista de símbolos	5
1 Preliminares	6
1.1 Funções de Variação Lenta	7
1.2 Convoluções de Funções e Funções de Variação Limitada	8
1.3 Distribuições Estáveis e Teorema do Limite Central Generalizado	13
1.4 Distribuições Estáveis e Teoremas de Limites Uniformes	20
2 Informação de Fisher	25
2.1 Teoria da Informação e Informação de Fisher	25
2.2 Propriedades Gerais da Informação de Fisher	29
2.3 Informação de Fisher e Funções de Variação Limitada	36
2.4 Informação de Fisher e Convolução de Densidades	39
2.5 Convergência na Informação de Fisher	45
3 Convergência em Informação de Fisher Relativa e v.a.'s Estáveis	47
3.1 Resultados Auxiliares	49
3.2 Teoremas de Limites	66

Introdução

O estudo de uma grande variedade de modelos teóricos e aplicados, encontrados na literatura, basei-se no comportamento assintótico de seqüências de somas estabilizadas do tipo

$$Z_n = \frac{X_1 + X_2 + \cdots + X_n}{b_n} - a_n, \quad n \geq 1, \quad (1)$$

onde X_n , $n \geq 1$, são variáveis aleatórias (v.a.'s) independentes e identicamente distribuídas (i.i.d.), com função de distribuição F e $\{a_n\}_{n \geq 1}$ e $\{b_n\}_{n \geq 1}$ são seqüências de constantes reais, com $b_n > 0$.

Um dos principais teoremas da Teoria Clássica da Probabilidade, o Teorema do Limite Central (Básico) estabelece que: sob a condição $\mathbb{E}X_1^2 < +\infty$, com $VarX_1 > 0$, e considerando $a_n = \frac{\mathbb{E}X_1}{\sqrt{nVarX_1}}$ e $b_n = \sqrt{nVarX_1}$ temos a convergência fraca de Z_n à distribuição Normal padrão, ou seja,

$$Z_n \xrightarrow{\mathfrak{D}} Z, \quad (2)$$

com $Z \stackrel{\mathfrak{D}}{=} \mathcal{N}(0, 1)$, que é uma distribuição α -estável com $\alpha = 2$. Aqui, $\xrightarrow{\mathfrak{D}}$ indica convergência em distribuição e $\stackrel{\mathfrak{D}}{=}$ indica igualdade em distribuição.

Sem a hipótese da finitude do segundo momento é conhecido que as possíveis distribuições limites de Z são as distribuições α -estáveis, com $0 < \alpha < 2$, entre as quais incluem-se as distribuições de Cauchy ($\alpha = 1$) e a distribuição de Levy ($\alpha = 1/2$). Neste caso, diz-se que a função de distribuição F pertence ao domínio de atração da distribuição α -estável Z quando $Z_n \xrightarrow{\mathfrak{D}} Z$. Assim, as distribuições α -estáveis caracterizam-se como limites em distribuição de somas estabilizadas de v.a.'s i.i.d. e são utilizadas na modelagem de fenômenos aplicados nas mais diversas áreas, tais como economia, finanças, atuária, entre outras.

Generalizações do Teorema do Limite Central Básico (no caso i.i.d.) para distribuições limites α -estáveis têm sido amplamente investigadas na literatura (vide, por exemplo, [14] e [27]). Uma condição necessária e suficiente para que $Z_n \xrightarrow{\mathfrak{D}} Z$, onde Z é α -estável com

$0 < \alpha < 2$, é que a função de distribuição F possua cauda à direita e à esquerda de variação regular, com o mesmo índice caudal α , ou seja,

$$F(-x) \sim [c_0 + o(1)] x^{-\alpha} L(x) \quad \text{e} \quad 1 - F(x) \sim [c_1 + o(1)] x^{-\alpha} L(x), \quad \text{quando } x \longrightarrow +\infty, \quad (3)$$

onde $c_0 > 0$ e $c_1 > 0$ são constantes apropriadas e $L(x)$ é uma função lentamente variante no infinito, no sentido de Karamata. Neste caso, $b_n = n^{1/\alpha} h(n)$, onde h é uma função lentamente variante dada por $h(x) = (1 + o(1))L(x)$, quando $x \longrightarrow \infty$. Assim, as distribuições α -estáveis com $0 < \alpha < 2$ são distribuições de cauda pesada. O caso $\alpha = 2$ refere-se à distribuição Normal, que tem cauda leve.

Devido às inúmeras aplicações de teoremas de limites desta natureza, uma questão de interesse na literatura é se esta convergência fraca em (2) implica em convergência em um sentido mais forte para a mesma v.a. Z .

Neste trabalho, estamos interessados na convergência relativa à “distância” de Informação de Fisher ou também chamada Informação de Fisher Relativa, que é usada como “medida” de discrepância entre duas distribuições.

O conceito de Informação de Fisher foi introduzido por Fisher em 1925 no contexto de estimação de parâmetros e está fortemente relacionado com o conceito de entropia relativa (ou divergência de Kullback-Leibler). Neste contexto, e fazendo uso dos princípios da Teoria da Informação, podemos dizer que a Informação de Fisher é uma medida de discrepância local, que nos fornece a quantidade de “informação” fornecida pelo parâmetro de interesse de uma dada distribuição.

Num contexto geral, para X e Y v.a.’s absolutamente contínuas com densidades diferenciáveis p e q , respectivamente, a Informação de Fisher Relativa é definida por

$$I(X||Y) = I(p||q) = \int_{\mathbb{R}} \left(\frac{p'(x)}{p(x)} - \frac{q'(x)}{q(x)} \right)^2 p(x) dx, \quad (4)$$

ou equivalentemente,

$$I(X||Y) = \mathbb{E} (\rho_X(X) - \rho_X(Y))^2,$$

onde $\rho_X(x) = \frac{p'(x)}{p(x)}$ (para x tal que $p(x) \neq 0$) é chamada função escore.

oSob as hipóteses do Teorema do Limite Central, no caso i.i.d. com $\mathbb{E}X_1^2 < +\infty$, onde temos $Z_n \xrightarrow{\mathcal{D}} Z$, com $Z \stackrel{\mathcal{D}}{=} \mathcal{N}(0, 1)$, Barron e Johnson [3] mostram que $I(Z_n||Z) \longrightarrow 0$ se, e

somente se, $I(Z_{n_0}) < +\infty$ para algum $n_0 \geq 1$. Aqui, $I(Y)$ indica a Informação de Fisher da v.a. Y definida por

$$I(Y) = \int_{-\infty}^{+\infty} \left[\frac{p'(x)}{p(x)} \right]^2 p(x) dx, \quad (5)$$

onde p é a densidade de Y , com derivada p' .

Comparando com outros tipos de convergência, a convergência à uma distribuição normal na distância de Informação de Fisher é uma propriedade mais forte do que em variação total e do que em entropia relativa. Isto pode ser observado pelas conhecidas desigualdades

$$\frac{\sigma^2}{2} I(Z_n \parallel Z) \geq D(Z_n \parallel Z) \geq \frac{1}{2} \|p_n - \psi\|_{TV}^2$$

válidas se $EX_1 = EZ$ e $VarX_1 = VarZ = \sigma^2$, onde p_n indica a densidade de Z_n , ψ é a densidade Normal de Z , $D(Z_n \parallel Z)$ denota a entropia relativa dada por

$$D(Z_n \parallel Z) = \int_{\mathbb{R}} \ln \left[\frac{p_n(x)}{\psi(x)} \right] p_n(x) dx$$

e $\|p_n - \psi\|_{TV}$ é a distância induzida pela norma da variação total entre Z_n e Z . Em [14] podemos encontrar relações da Informação de Fisher com outros modos de convergência em situações mais gerais.

Nosso interesse neste trabalho é a extensão do resultado de Barron e Johnson para o caso em que Z tem distribuição α -estável não-extremal, ou seja, quando $\alpha = 2$ ou quando $0 < \alpha < 2$ e $-1 < \beta < 1$, sendo β o parâmetro de simetria. Esta extensão foi obtida em 2014 por Bobkov Chistyakov e Gotze [6] e é referência principal deste trabalho.

Assim, no Capítulo 1 apresentamos os conceitos e resultados preliminares sobre funções de variação lenta; convolução de funções, funções de variação limitada, distribuições estáveis e teoremas de limites relacionados.

O Capítulo 2 é dedicado ao estudo do conceito e das propriedades da Informação de Fisher. Iniciamos com uma motivação do conceito inserido no contexto da Teoria de Informação e uma descrição de propriedades gerais da Informação de Fisher. Em seguida, apresentamos propriedades específicas da Informação de Fisher relativas às funções de variação limitada e à convolução de densidades.

Finalmente, no Capítulo 3 estudamos o teorema principal de Bobkov Chistyakov e Gotze [6] o qual estabelece que: se $Z_n \xrightarrow{D} Z$, onde Z é α -estável não-extremal e Z_n , $n \geq 1$, é a sequência

de somas estabilizadas em (1), com X_1, X_2, \dots i.i.d. com densidade absolutamente contínua, então $I(Z_n||Z) \rightarrow 0$ se, e somente se, $I(Z_n) < +\infty$ para algum n . Uma breve justificativa para a exclusão do caso extremal $|\beta| = 1$ é apresentada no final do capítulo.

Lista de Símbolos

v.a.	Variável Aleatória.
f.d.	Função de Distribuição.
i.i.d.	Independentes e Identicamente Distribuídas.
$\xrightarrow{\mathcal{D}}$	Convergência em Distribuição.
$\stackrel{\mathcal{D}}{=}$	Identidade em Distribuição.
$f * g$	Convolução de f e g .
$f(x) \sim g(x)$ quando $x \rightarrow a$	$\lim_{x \rightarrow a} \frac{f(x)}{g(x)} = 1.$
$f(x) = o(g(x))$ quando $x \rightarrow a$	$\lim_{x \rightarrow a} \frac{f(x)}{g(x)} = 0.$
$f(x) = O(g(x))$ quando $x \rightarrow a$	$\limsup_{x \rightarrow a} \left \frac{f(x)}{g(x)} \right < \infty.$
\mathbb{R}	Conjunto dos números reais.
$\mathbb{E}X$	Esperança da v.a. X .
$VarX$	Variância da v.a. X .
$\int_A f(x)dx$	Integral de Lebesgue sobre $A \subset \mathbb{R}$.
1_A	Função Indicadora do conjunto A .
$[x]$	Parte Inteira de x .
$q.c.$	Quase Certamente.

Capítulo 1

Preliminares

O objetivo deste capítulo é apresentar os resultados que dão suporte à teoria desenvolvida nos segundo e terceiro capítulos.

Na Seção 1.1 apresentamos a definição e alguns teoremas importantes das funções de variação lenta, entre eles o Teorema de Representação e o Teorema de Convergência Uniforme. O estudo das funções de variação lenta é importante para nosso trabalho, pois com esta teoria podemos caracterizar os domínios de atração das distribuições estáveis.

Na Seção 1.2 apresentamos algumas propriedades das convoluções de funções definidas em \mathbb{R} e uma fórmula análoga ao binômio de Newton, chamada decomposição binomial para convoluções de funções, e obtemos a forma da função característica para a convolução de duas funções de densidade. Ao final da seção provamos uma desigualdade relacionada à variação total da n -ésima convolução de densidades absolutamente contínuas.

A Seção 1.3 é dedicada ao estudo das distribuições estáveis, conceito chave neste trabalho. Apresentamos algumas propriedades importantes dessas distribuições, a representação da função característica de uma lei estável, a definição de leis estáveis não-extremais e finalizamos com uma generalização do Teorema do Limite Central.

Na última seção apresentamos resultados sobre limites uniformes de densidades e das derivadas de densidades de sequências de somas estabilizadas convergindo em distribuição para uma lei estável.

As referências básicas deste capítulo são [13], [19], [22] e [27].

1.1 Funções de Variação Lenta

Usaremos a notação $h(x) \sim g(x)$ quando $x \rightarrow a$ para indicar $\lim_{x \rightarrow a} \frac{h(x)}{g(x)} = 1$.

Definição 1.1.1. *Seja $h : [0, +\infty) \rightarrow \mathbb{R}^+$. Dizemos que h é de variação lenta, se para todo $t \geq 0$*

$$\lim_{x \rightarrow \infty} \frac{h(tx)}{h(x)} = 1. \quad (1.1)$$

Definição 1.1.2. *Uma função $g : [0, +\infty) \rightarrow \mathbb{R}^+$ é dita de variação regular de índice α , se para todo $t > 0$*

$$\lim_{x \rightarrow \infty} \frac{g(tx)}{g(x)} = t^\alpha. \quad (1.2)$$

Claramente (1.2) é equivalente à escrever $g(x) = x^\alpha h(x)$, onde $h(x)$ é uma função de variação lenta.

Agora vamos apresentar dois teoremas básicos da teoria das funções de variação lenta, dos quais seguem a maioria das propriedades destas funções. A prova dos teoremas pode ser encontrada em [22].

Teorema 1.1.3 (Teorema da Convergência Uniforme). *Se h é uma função de variação lenta, então para qualquer intervalo fechado $[a, b]$, $0 < a < b < \infty$,*

$$\lim_{x \rightarrow \infty} \frac{h(tx)}{h(x)} = 1 \text{ uniformemente para todo } t \in [a, b].$$

Teorema 1.1.4 (Teorema de Representação de Karamata). *Uma função h é de variação lenta se, e somente se, pode ser representada na forma*

$$h(x) = c(x) \exp \left\{ \int_{x_0}^x \frac{w(y)}{y} dy \right\}, \quad (1.3)$$

para algum $x_0 > 0$, onde $c(x)$ é mensurável em $[x_0, \infty)$,

$$\lim_{x \rightarrow \infty} c(x) = c \neq 0,$$

e

$$\lim_{x \rightarrow \infty} w(x) = 0.$$

Corolário 1.1.5. *Se h é de variação lenta, então para todo $\alpha > 0$ temos*

$$\lim_{x \rightarrow \infty} x^\alpha h(x) = \infty$$

e

$$\lim_{x \rightarrow \infty} x^{-\alpha} h(x) = 0.$$

Teorema 1.1.6. *Seja h uma função de variação lenta em $[x_0, \infty)$, para algum $x_0 > 0$. Se f é uma função real tal que para $0 < \alpha < \infty$ e algum $\eta \geq 0$, a integral*

$$\int_{\alpha}^{\infty} x^\eta f(x) dx,$$

existe então a integral

$$\int_{\alpha}^{\infty} f(x) h(tx) dx$$

existe e

$$\int_{\alpha}^{\infty} f(x) h(tx) dx \sim h(t) \int_{\alpha}^{\infty} f(x) dx, \text{ quando } t \rightarrow \infty.$$

Ao longo deste trabalho, além das funções de variação lenta vamos utilizar seqüências de variação lenta, o qual definimos a seguir.

Definição 1.1.7. *Seja $\{h(n)\}_{n \in \mathbb{N}}$ uma seqüência de números reais positivos, dizemos que $h(n)$ é de variação lenta se para qualquer inteiro positivo k ,*

$$\lim_{n \rightarrow \infty} \frac{h(kn)}{h(n)} = 1.$$

1.2 Convoluções de Funções e Funções de Variação Limitada

Definição 1.2.1. *Sejam f e g funções reais dadas. A convolução de f e g é representada por $f * g$ e é definida pela integral*

$$(f * g)(x) = \int_{-\infty}^{\infty} f(y)g(x - y) dy \text{ para } x \in \mathbb{R}.$$

As seguintes propriedades das convoluções seguem diretamente da definição.

Propriedades

- (a) $f * g = g * f$ (Comutatividade)
- (b) $f * (g * h) = (f * g) * h$ (Associatividade)
- (c) $f * (g + h) = (f * g) + (f * h)$ (Distributividade)
- (d) $c(f * g) = (cf * g) = f * (cg)$ para $c \in \mathbb{R}$.
- (e) $\frac{d}{dx}(f * g) = \left(\frac{d}{dx}f\right) * g = f * \left(\frac{d}{dx}g\right)$ (Regra da Diferenciação)

Observação 1.2.2. Podemos ver que se X e Y são v.a.'s independentes com funções de distribuição F_1 e F_2 respectivamente, então $Z = X + Y$ tem função de distribuição $F_3 = F_1 * F_2$, isto é,

$$F_3(x) = \int_{-\infty}^{\infty} F_1(x - y)dF_2(y).$$

Além disso, se f_1, f_2 e f_3 são as densidades de X, Y e Z respectivamente, então

$$f_3(x) = f_1 * f_2(x) = \int_{-\infty}^{\infty} f_1(x - y)f_2(y)dy.$$

Proposição 1.2.3. (Decomposição Binomial para Convoluções) Sejam f e g funções reais. Então

$$(f + g)^{n*} = \sum_{k=0}^n \binom{n}{k} f^{k*} * g^{(n-k)*}, \quad (1.4)$$

onde h^{n*} indica a convolução de h n vezes, isto é, $h^{1*} = h$, $h^{2*} = h * h$ e para $n > 2$, $h^{n*} = h^{(n-1)*} * h$.

Demonstração. Claramente (1.4) é válida para $n = 1$. Das propriedades das convoluções, para $n = 2$,

$$(f + g)^{2*} = (f + g) * (f + g) = f^{2*} + 2(f * g) + g^{2*}.$$

Agora, suponhamos que (1.4) é válida para $n - 1$ e mostremos que também vale para n . De fato, da hipótese de indução e das propriedades de associatividade e distributividade segue

$$\begin{aligned} (f + g)^{n*} &= (f + g)^{(n-1)*} * (f + g) \\ &= \left[\sum_{k=0}^{n-1} \binom{n-1}{k} f^{k*} * g^{((n-1)-k)*} \right] * (f + g) \\ &= \sum_{k=0}^n \binom{n}{k} f^{k*} * g^{(n-k)*}. \end{aligned}$$

□

Desta proposição podemos obter uma caracterização da função característica associada à convolução de duas densidades de probabilidade, que apresentamos a seguir.

Corolário 1.2.4. *Se p e q são densidades de probabilidade com funções características associadas f e g respectivamente, então $f^k g^{n-k}$, $k = 1, \dots, n$, é a função característica associada a $p^{k*} * q^{(n-k)*}$.*

Demonstração. Denotemos por f^* a função característica de $p^{k*} * q^{(n-k)*}$. Usando o Teorema de Fubini obtemos,

$$\begin{aligned}
 f^*(t) &= \int_{-\infty}^{\infty} e^{itx} p^{k*} * q^{(n-k)*}(x) dx \\
 &= \int_{-\infty}^{\infty} e^{itx} \left[\int_{-\infty}^{\infty} p^{k*}(y) q^{(n-k)*}(x-y) dy \right] dx \\
 &= \int_{-\infty}^{\infty} p^{k*}(y) \left[\int_{-\infty}^{\infty} e^{itx} q^{(n-k)*}(x-y) dx \right] dy, \text{ para } z = x - y \\
 &= \left[\int_{-\infty}^{\infty} e^{ity} p^{k*}(y) dy \right] \left[\int_{-\infty}^{\infty} e^{itz} q^{(n-k)*}(z) dz \right] \\
 &= f^k(t) g^{n-k}(t).
 \end{aligned}$$

□

Uma propriedade importante para o desenvolvimento dos capítulos seguintes diz respeito à variação total da n -ésima convolução de densidades absolutamente contínuas, que será demonstrada na Proposição 1.2.8. Para isso, apresentamos inicialmente conceitos e propriedades de funções de variação limitada e variação total.

Definição 1.2.5. *Seja $f : I \rightarrow \mathbb{R}$, $I = [a, b]$. A variação de f sobre I , denotada por $\|f\|_{TV}$, é dada pela quantidade*

$$\|f\|_{TV} = \sup \sum_{k=1}^n |f(x_k) - f(x_{k-1})|, \tag{1.5}$$

onde o supremo é tomado sobre toda partição $P = \{a = x_0 < x_1 < \dots < x_n = b\}$ do intervalo I . Se $\|f\|_{TV} < \infty$ dizemos que f é de variação limitada e definimos o número $\|f\|_{TV}$ como a variação total de f em I .

Se f não é de variação limitada, escrevemos, $\|f\|_{TV} = \infty$.

Teorema 1.2.6. *Suponhamos que f e f' são funções contínuas no intervalo $I = [a, b]$. Então f é de variação limitada em I e*

$$\|f\|_{TV} = \int_a^b |f'(t)| dt.$$

Demonstração. Ver [18], pág.313. □

De maneira análoga podemos definir variação total de uma função $f : \mathbb{R} \rightarrow \mathbb{R}$, isto é,

Definição 1.2.7. *Uma função $f : \mathbb{R} \rightarrow \mathbb{R}$ é de variação limitada se*

$$\|f\|_{TV} = \sup \sum_{k=-\infty}^{\infty} |f(x_k) - f(x_{k-1})| < \infty, \quad (1.6)$$

onde o supremo é tomado sobre todas as seqüências $\dots, x_{k-1}, x_k, x_{k+1}, \dots$ tais que $x_k \rightarrow \pm\infty$ quando $k \rightarrow \pm\infty$. Como acima, a quantidade $\|f\|_{TV}$ é chamada variação total de f em \mathbb{R} .

Agora, vamos demonstrar uma desigualdade que é outra consequência das propriedades das convoluções e do Teorema de Fubini.

Proposição 1.2.8. *Se p é uma densidade absolutamente contínua com derivada p' contínua, então*

$$\|p^{n*}\|_{TV} \leq \|p\|_{TV}, \quad \text{para } n = 1, 2, 3, \dots \quad (1.7)$$

Demonstração. Para $n = 2$, das regras de diferenciação de convoluções e do Teorema de Fubini segue

$$\begin{aligned}
\|p^{2*}\|_{TV} &= \int_{-\infty}^{\infty} |(p^{2*})'(x)| dx \\
&\leq \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} |p'(y)| |p(x-y)| dx dy \\
&= \left[\int_{-\infty}^{\infty} |p'(y)| dy \right] \left[\int_{-\infty}^{\infty} |p(z)| dz \right] \\
&\leq \|p\|_{TV}.
\end{aligned}$$

Fazendo indução sobre n , suponhamos que (1.7) é válida para $n-1$, então repetindo os mesmos argumentos anteriores segue

$$\begin{aligned}
\|p^{n*}\|_{TV} &= \int_{-\infty}^{\infty} |(p^{n*})'(x)| dx \\
&\leq \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} |(p^{(n-1)*})'(y)| |p(x-y)| dx dy \\
&= \left[\int_{-\infty}^{\infty} |(p^{(n-1)*})'(y)| dy \right] \left[\int_{-\infty}^{\infty} |p(z)| dz \right] \\
&\leq \|p^{(n-1)*}\|_{TV} \\
&\leq \|p\|_{TV}.
\end{aligned}$$

Logo, (1.7) é válida para todo $n \in \mathbb{N}$. □

1.3 Distribuições Estáveis e Teorema do Limite Central Generalizado

Nesta seção apresentamos uma síntese dos conceitos e principais propriedades de distribuições estáveis. Para maiores detalhes vide [13], [19] e [27].

Definição 1.3.1. Dizemos que uma v.a X segue uma lei estável, se para cada inteiro $n \geq 2$ e X_1, X_2, \dots, X_n v.a's independentes com a mesma distribuição de X , existem constantes $c_n > 0$ e d_n tais que

$$Z_n = X_1 + X_2 + \dots + X_n \stackrel{\mathcal{D}}{=} c_n X + d_n \quad (1.8)$$

onde $\stackrel{\mathcal{D}}{=}$ significa igualdade em distribuição, ou equivalentemente, em termos da função característica f de X ,

$$[f(t)]^n = e^{id_n t} f(c_n t). \quad (1.9)$$

Se X é estável, é possível mostrar que existem constantes $0 < \alpha \leq 2$, $\sigma > 0$, $|\beta| \leq 1$ e $\mu \in \mathbb{R}$ tais que

$$c_n = n^{1/\alpha} \quad e \quad d_n = \begin{cases} \mu(n - n^{1/\alpha}), & \text{se } \alpha \neq 1 \\ 2\sigma\beta \frac{n}{\pi} \log n, & \text{se } \alpha = 1. \end{cases}$$

No caso $\alpha > 1$, é possível tomar $\mu = \mathbb{E}X$.

Os seguintes exemplos mostram quatro casos de distribuições estáveis onde as suas densidades ou funções de probabilidade podem-se escrever por meio de funções elementares.

Exemplo 1. *Caso Degenerado*, isto é, $P(X = \mu) = 1$. Sabemos que a função característica de X é $f(t) = e^{it\mu}$. Logo, $[f(t)]^n = e^{i(nt)\mu} = f(nt)$. Assim, X é estável com $c_n = n$ e $d_n = 0$.

Exemplo 2. *Distribuição Normal*. Seja $X \stackrel{\mathcal{D}}{=} \mathcal{N}(\mu, \sigma^2)$, então a função característica de X é dada por $f(t) = \exp\left\{it\mu - \frac{t^2\sigma^2}{2}\right\}$. Logo, $[f(t)]^n = \exp\{i(\mu n - \mu\sqrt{n})t\} f(\sqrt{n}t)$. Portanto, X é estável com $c_n = n^{1/2}$ e $d_n = \mu(n - n^{1/2})$.

Exemplo 3. *Distribuição Cauchy*. $X \stackrel{\mathcal{D}}{=} \text{Cauchy}(a, b)$ ou seja, X com densidade

$$p(x) = \frac{1}{\pi} \left(\frac{a}{a^2 + (x - b)^2} \right), \quad x \in \mathbb{R}. \quad e \quad \text{função característica } f(t) = \exp\{ita - b|t|\}.$$

Logo, $[f(t)]^n = f(nt)$. Assim, X é estável com $c_n = n$ e $d_n = 0$.

Exemplo 4. *Distribuição Lévy. Seja $X \stackrel{\mathcal{D}}{=} \text{Lévy}(\gamma, \delta)$ então X tem densidade*

$$p(x) = \sqrt{\frac{\gamma}{2\pi}} \frac{1}{(x - \delta)^{3/2}} \exp\left\{-\frac{\gamma}{2(x - \delta)}\right\}, \quad x > \delta,$$

e função característica $f(t) = \exp\{it\gamma - \sqrt{-2it\delta}\}$. Logo, $[f(t)]^n = \exp it(\gamma n - \gamma n^2)f(n^2t)$ e portanto X é estável com $c_n = n^2$ e $d_n = \gamma(n - n^2)$.

Vamos nos reportar à v.a.'s estáveis apenas para o caso não-degenerado.

É conhecido que, com exceção das leis estáveis Normal, Cauchy e Levy, não temos expressões em termos de funções elementares para as densidades das distribuições estáveis (não degeneradas), o que temos é uma representação para suas correspondentes funções características, como mostra o teorema a seguir.

Teorema 1.3.2. *Uma v.a. X não-degenerada é estável, com parâmetros $\alpha, \sigma, \beta, \mu$, se, e somente se, sua função característica f tem a forma*

$$f(t) = \exp\{i\mu t - \sigma^\alpha |t|^\alpha [1 + i\beta \text{sign}(t)w(t, \alpha)]\}, \quad (1.10)$$

onde $0 < \alpha \leq 2$, $-1 \leq \beta \leq 1$, $\mu \in \mathbb{R}$, $\sigma > 0$ e

$$w(t, \alpha) = \begin{cases} \tan \frac{\pi\alpha}{2}, & \text{se } \alpha \neq 1 \\ \frac{2}{\pi} \log |t|, & \text{se } \alpha = 1. \end{cases}$$

Observação 1.3.3.

(a) *Por (1.10) temos que os parâmetros α , σ , β e μ caracterizam completamente a distribuição de X . Assim, costuma-se representar uma v.a. estável X por*

$$X \stackrel{\mathcal{D}}{=} S_\alpha(\sigma, \beta, \mu)$$

e denominar X como v.a. α -estável.

(b) *O parâmetro α é chamado índice de estabilidade e caracteriza a taxa de decrescimento da cauda da lei estável X .*

O número β é chamado parâmetro de simetria. Quando $\beta = 0$, as distribuições estáveis são simétricas, como é o caso de uma v.a com distribuição $\mathcal{N}(\mu, \sigma^2)$ ou com distribuição Cauchy (γ, δ) .

O número σ é o parâmetro escala e μ o de locação. Quando $\sigma = 1$ e $\mu = 0$ dizemos que X é α -estável padrão.

(c) De (1.10) segue que $|f(t)| = e^{-c|t|^\alpha}$, com $c = \sigma^\alpha$, e que f é integrável. Logo, pela Fórmula de Inversão, temos que a densidade ψ de X é dada por

$$\psi(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-itx} f(t) dt.$$

Podemos mostrar que as densidades α -estáveis são contínuas e infinitamente diferenciáveis.

(Vide por exemplo [19]). Além disso, segue que

$$|\psi^{(k)}(x)| \leq \int_{-\infty}^{\infty} |t|^k |f(t)| dt = \frac{\Gamma((k+1)/\alpha)}{\pi\alpha} c^{-(k+1)\alpha}, \quad k = 1, 2, 3, \dots$$

ou seja, tanto a densidade ψ como suas derivadas estão limitadas uniformemente em x .

(d) As densidades α -estáveis são unimodais. (Vide por exemplo, Teorema 2.5.3 em [13] ou [27]).

(e) Se $0 < \alpha \leq 1$, então f não é diferenciável em $t = 0$ e $\mathbb{E}(|X|) = \infty$. No caso $1 < \alpha < 2$, f é diferenciável em $t = 0$ apenas uma vez e assim $\mathbb{E}(|X|) < \infty$, mas $\mathbb{E}(X^2) = \infty$. Somente para $\alpha = 2$, ou seja, quando X segue uma distribuição normal, temos esperança e variância finitas.

(f) Pode-se mostrar que $\mathbb{E}X^\delta < \infty$ para todo $0 < \delta < \alpha$ (vide por exemplo [10], pág 215.)

No caso α -estável, com $0 < \alpha < 2$, embora não existam formas explícitas para densidades estáveis gerais, são conhecidas várias propriedades de suas densidades. Dentre elas destacamos, na proposição a seguir, a aproximação assintótica quando $x \rightarrow +\infty$ e quando $x \rightarrow -\infty$ por leis de potência com caudas pesadas, ou seja, regularmente variantes.

Proposição 1.3.4. *Seja $X \stackrel{\mathcal{D}}{=} S_\alpha(\sigma, \beta, \mu)$, $0 < \alpha < 2$ com densidade ψ . Então*

(a) *para $-1 < \beta \leq 1$, temos*

$$P(X > x) \sim C_0 x^{-\alpha} \quad \text{quando } x \rightarrow +\infty$$

e

$$\psi(x) \sim C_1 x^{-(\alpha+1)} \quad \text{quando } x \rightarrow +\infty$$

onde $C_0 = \sigma^\alpha c_\alpha (1 + \beta)$, $C_1 = \alpha C_0$ e $c_\alpha = \sin\left(\frac{\pi\alpha}{2}\right) \frac{\Gamma(\alpha)}{\pi}$.

(b) para $-1 \leq \beta < 1$, temos

$$P(X < -x) \sim D_0 x^{-\alpha} \text{ quando } x \rightarrow -\infty$$

e

$$\psi(x) \sim D_1 |x|^{-(\alpha+1)} \text{ quando } x \rightarrow -\infty$$

onde $D_0 = \sigma^\alpha c_\alpha (1 - \beta)$, e $D_1 = \alpha D_0$.

Observe que para $0 < \alpha < 2$ e $-1 < \beta < 1$ ambas as caudas da densidade, quando $x \rightarrow +\infty$ e $x \rightarrow -\infty$, podem ser aproximadas por leis de potência, ou seja,

$$\psi(x) \sim C_1 x^{-(\alpha+1)} \text{ quando } x \rightarrow +\infty \text{ e } \psi(x) \sim D_1 |x|^{-(\alpha+1)} \text{ quando } x \rightarrow -\infty. \quad (1.11)$$

Estas densidades, juntamente com a densidade normal, são chamadas de não-extremais como definimos a seguir.

Definição 1.3.5. Uma v.a. X α -estável é chamada não-extremal se X segue uma distribuição normal, ou se $0 < \alpha < 2$ e $-1 < \beta < 1$ em (1.10).

Observação 1.3.6. Se X é α -estável não-extremal, com $0 < \alpha < 2$, como toda densidade estável é unimodal, segue de (1.11) que a densidade ψ de X é positiva em toda a reta real (isto é, $\text{supp}(\psi) = \mathbb{R}$). Além disso, é possível mostrar que

$$\frac{|\psi'(x)|}{\psi(x)} \sim \frac{1}{|x|} \text{ quando } |x| \rightarrow +\infty \quad (1.12)$$

e que existe uma constante $C > 0$ tal que

$$\frac{|\psi'(x)|}{\psi(x)} \leq \frac{C}{1 + |x|} \text{ para } x \in \mathbb{R}. \quad (1.13)$$

Uma outra forma de caracterizar distribuições estáveis é como a distribuição limite de somas estabilizadas.

Definição 1.3.7. Uma sequência de v.a.'s dada por:

$$Z_n = \frac{S_n}{b_n} - a_n, \quad n \geq 1,$$

onde $a_n \in \mathbb{R}$, $b_n > 0$ e $S_n = X_1 + \dots + X_n$, com X_1, X_2, \dots v.a.'s i.i.d., é chamada uma sequência de somas estabilizadas.

Teorema 1.3.8. *Um v.a. Z é estável se, e somente se, existe uma sequência de somas estabilizadas que converge em distribuição para Z . Neste caso, se Z é não-degenerada, a sequência $\{b_n\}$ de números reais é dada por $b_n = n^{1/\alpha}h(n)$, $0 < \alpha \leq 2$, onde h é uma função de variação lenta no sentido de Karamata.*

Uma questão natural é: se X_1, X_2, \dots são v.a.'s i.i.d. com função de distribuição F , sob quais condições existem sequências $\{a_n\}$ e $\{b_n\}$, com $b_n > 0$, para as quais a sequência de somas estabilizadas $\{Z_n\}$ converge em distribuição para uma distribuição não-degenerada e qual seria esta distribuição limite?

Em outras palavras, deseja-se obter condições para as quais uma função de distribuição F esteja no domínio de atração de uma função de distribuição não-degenerada, conforme definimos a seguir.

Definição 1.3.9. *Seja F a função de distribuição de v.a.'s i.i.d. X_1, X_2, X_3, \dots . Dizemos que F pertence ao **domínio de atração** de uma função de distribuição não-degenerada G , se existem sequências de constantes $\{a_n\}$ e $\{b_n\}$, com $b_n > 0$ tais que a sequência de somas estabilizadas $\{Z_n\}$,*

$$Z_n = \frac{X_1 + X_2 + \dots + X_n}{b_n} - a_n, \quad n \geq 1,$$

converge fracamente para G . Ou seja,

$$Z_n \xrightarrow{\mathfrak{D}} Z, \tag{1.14}$$

onde Z é uma v.a. com f.d G .

Aqui $\xrightarrow{\mathfrak{D}}$ indica a convergência em distribuição, isto é,

$$Z_n \xrightarrow{\mathfrak{D}} Z \text{ se, e somente se, } P(Z_n \leq x) \xrightarrow[n \rightarrow \infty]{} G(x) = P(Z \leq x), \quad \forall x \in C(G),$$

onde $C(G)$ denota o conjunto de pontos de continuidade de G .

Note que, (1.14) é equivalente a

$$\lim_{n \rightarrow \infty} F^{n*}(b_n x + a_n b_n) = G(x), \quad \forall x \in C(G),$$

onde $F^{n*} = F * \dots * F$, a n -ésima convolução de F , é a função de distribuição de S_n .

Se $\mathbb{E}X_1^2 < \infty$, a resposta à pergunta inicial é dada por um dos resultados mais importantes da Teoria da Probabilidade, o Teorema do Limite Central, onde se mostra que para

$a_n = \frac{\mathbb{E}S_n}{\sqrt{\text{Var}S_n}}$ e $b_n = \sqrt{\text{Var}S_n}$ o limite segue uma distribuição Normal, ou seja, se F é uma função de distribuição com segundo momento finito, então F está no domínio de atração de uma distribuição Normal.

Generalizações do Teorema do Limite Central são apresentadas a seguir.

Primeiramente, apresentamos condições necessárias e suficientes para que uma função de distribuição F pertença ao domínio de atração de uma distribuição α -estável.

Teorema 1.3.10. *Uma função de distribuição F pertence ao domínio de atração de uma distribuição α -estável se, e somente se,*

$$F(x) \sim \frac{c_0 + o(1)}{(-x)^\alpha} L(-x) \text{ quando } x \rightarrow -\infty \quad (1.15)$$

e

$$1 - F(x) \sim \frac{c_1 + o(1)}{x^\alpha} L(x) \text{ quando } x \rightarrow \infty, \quad (1.16)$$

onde L é uma função lentamente variante no sentido de Karamata, para certas constantes $c_0 > 0$ e $c_1 > 0$.

Teorema 1.3.11 (Teorema do Limite Central Generalizado). *Sejam X_1, X_2, \dots v.a.'s i.i.d. com função de distribuição F satisfazendo*

$$\begin{cases} 1 - F(x) \sim cx^{-\gamma}, & \text{quando } x \rightarrow \infty, \\ F(x) \sim d|x|^{-\gamma}, & \text{quando } x \rightarrow -\infty, \end{cases} \quad (1.17)$$

onde $\gamma > 0$. Então F pertence ao domínio de atração de uma distribuição α -estável padrão, com parâmetros

$$\alpha = \begin{cases} \gamma & \text{se } \gamma \leq 2 \\ 2 & \text{se } \gamma > 2, \end{cases} \quad e \quad \beta = \frac{c - d}{c + d}.$$

Neste caso, as sequências normalizantes $\{a_n\}$ e $\{b_n\}$ são escolhidas de acordo com a tabela a seguir:

γ	α	a_n	b_n
$0 < \gamma < 1$	γ	0	$[\pi(c+d)]^{1/\alpha} [2\Gamma(\alpha) \sin(\alpha\pi/2)]^{-1/\alpha} n^{1/\alpha}$
$\gamma = 1$	γ	$\beta(c+d)n \ln n$	$(\pi/2)(c+d)n$
$1 < \gamma < 2$	γ	$n\mathbb{E}(X)$	$[\pi(c+d)]^{1/\alpha} [2\Gamma(\alpha) \sin(\alpha\pi/2)]^{-1/\alpha} n^{1/\alpha}$
$\gamma = 2$	2	$n\mathbb{E}(X)$	$(c+d)^{1/2} [n \ln n]^{1/2}$
$\gamma > 2$	2	$n\mathbb{E}(X)$	$[(1/2)\text{Var}X]^{1/2} n^{1/2}$

Tabela 1.1: Forma dos coeficientes a_n e b_n .

Finalizamos esta seção com um teorema que garante que: para que uma função de distribuição F esteja no domínio de atração de uma distribuição α -estável é necessário e suficiente que a respectiva função característica f_1 associada a F seja regularmente variante perto da origem.

Teorema 1.3.12. *Seja $Z = S_\alpha(\sigma, \beta, \mu)$ uma v.a α -estável com função característica dada por (1.10). Uma condição necessária e suficiente para que uma função de distribuição F esteja no domínio de atração de Z , ou seja, que $Z_n \xrightarrow{\mathcal{D}} Z$, onde $\{Z_n\}$ é uma sequência de somas estabilizadas associadas a F , é que sua respectiva função característica $f_1(t)$ satisfaça para t na vizinhança da origem*

$$f_1(t) = \exp \left\{ i\mu t - c|t|^\alpha h(|t|^{-1}) [1 + i\beta \text{sign}(t)w(t, \alpha)] \right\} \quad (1.18)$$

onde $c = \sigma^\alpha > 0$, $\alpha, \beta, \mu, \sigma, w(t, \alpha)$ como no Teorema 1.3.2 e $h(x)$ é uma função lentamente variante quando $x \rightarrow \infty$ tal que

$$\lim_{n \rightarrow \infty} \frac{nh(b_n)}{b_n^\alpha} = 1. \quad (1.19)$$

A demonstração do teorema anterior é dada em detalhes em [13] (Teorema 2.6.5.). Note que de (1.18) segue que numa vizinhança da origem

$$|f_1(t)| = \exp \left\{ -c|t|^\alpha h\left(\frac{1}{|t|}\right) \right\}.$$

No caso especial, quando $0 < \alpha < 2$, da prova do teorema pode-se concluir que a função $L(x)$, aparecendo nas relações assintóticas (1.15) e (1.16) do Teorema 1.3.10, e a função $h(x)$ estão relacionadas por

$$h(x) = (1 + o(1))L(x), \quad \text{quando } x \rightarrow \infty,$$

onde $o(1) \rightarrow 0$ quando $x \rightarrow \infty$. Assim de (1.15), (1.16) e (1.19) segue que

$$1 - F(b_n) + F(-b_n) \sim \frac{c}{n} \text{ quando } n \rightarrow \infty.$$

Logo, temos o seguinte resultado para o caso $0 < \alpha < 2$.

Corolário 1.3.13. *Se $Z_n \xrightarrow{\mathcal{D}} Z$, onde $Z = S_\alpha(\sigma, \beta, \mu)$, com $0 < \alpha < 2$, então*

$$|f_1(t)| = \exp \left\{ -c|t|^\alpha h(|t|^{-1}) \right\}, \quad (1.20)$$

onde $c > 0$ e $h(x)$ é uma função lentamente variante quando $x \rightarrow +\infty$ satisfazendo (1.19).

Mais ainda,

$$P(|X_1| > b_n) = 1 - F(b_n) + F(-b_n) \sim \frac{c}{n}. \quad (1.21)$$

Note que: se $\mathbb{E}(X_1^2) < \infty$, podemos tomar $h(x) = 1$ e, neste caso, $b_n \sim \sqrt{n}$, mas

$$P\{|X_1| > b_n\} = o\left(\frac{1}{n}\right),$$

Isto mostra que (1.21) não é verdade para $\alpha = 2$.

1.4 Distribuições Estáveis e Teoremas de Limites Uniformes

Nesta seção apresentamos teoremas de limites uniformes para a sequência de densidades e das derivadas das densidades de somas estabilizadas convergindo em distribuição para uma lei estável.

Para isso, seja Z uma v.a. α -estável com densidade $\psi(x)$ e considere a sequência de somas estabilizadas

$$Z_n = \frac{X_1 + \cdots + X_n}{b_n} - a_n,$$

onde X_1, X_2, \dots são v.a.'s i.i.d. e $\{a_n\}$ e $\{b_n\}$, com $b_n > 0$, são sequências de constantes apropriadamente escolhidas tais que $Z_n \xrightarrow{\mathcal{D}} Z$.

Denotemos para cada $n \geq 1$ a densidade e a função característica de Z_n por $p_n(x)$ e $f_n(t)$, respectivamente. Em particular, $p_1(x)$ e $f_1(t)$ são, respectivamente, a densidade e a função característica de $X_1 = Z_1$.

O primeiro resultado apresenta condições suficientes para o limite uniforme de p_n à ψ .

Proposição 1.4.1. *Suponhamos que $Z_n \xrightarrow{\mathcal{D}} Z$, onde Z é uma v.a. α -estável. Se para algum $\nu > 0$*

$$\int_{-\infty}^{+\infty} |f_1(t)|^\nu dt < \infty, \quad (1.22)$$

então para todo n grande o suficiente, p_n é limitada e contínua. Além disso,

$$\lim_{n \rightarrow +\infty} \sup_x |p_n(x) - \psi(x)| = 0. \quad (1.23)$$

Demonstração. Ver [13], pág.126. □

Notemos que da condição (1.22), podemos obter que para todo $n \geq \nu$, a função característica f_n é integrável. Logo, as densidades para $n \geq \nu$ são dadas pela fórmula de inversão

$$p_n(x) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} e^{-itx} f_n(t) dt.$$

Utilizando as mesmas ideias da demonstração da Proposição 1.4.1, Bobkov, Chistyakov e Gotze obtiverem em [7] um resultado semelhante para o limite uniforme das sequências de derivadas $p'_n(x)$ que apresentamos a seguir e que será útil para o desenvolvimento do Capítulo 3.

Proposição 1.4.2. *Suponhamos que $Z_n \xrightarrow{\mathcal{D}} Z$, onde Z é uma v.a. α -estável. Se para algum $\nu > 0$*

$$\int_{-\infty}^{+\infty} |f_1(t)|^\nu |t| dt < \infty, \quad (1.24)$$

então para todo n grande o suficiente, p_n é continuamente diferenciável e sua derivada p'_n é limitada. Além disso,

$$\lim_{n \rightarrow +\infty} \sup_x |p'_n(x) - \psi'(x)| = 0. \quad (1.25)$$

Demonstração. Do comentário após a proposição anterior, temos que para todo $n \geq \nu$

$$p_n(x) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} e^{-itx} f_n(t) dt, \quad (1.26)$$

e da condição (1.24), obtemos que para todo $n \geq \nu$, a função $tf_n(t)$ é integrável. Assim, podemos diferenciar (1.26) e obter uma representação similar para p'_n , isto é,

$$p'_n(x) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} (-it) e^{-itx} f_n(t) dt. \quad (1.27)$$

Consequentemente, das propriedades da Transformada de Fourier segue que, para n suficientemente grande, p_n é continuamente diferenciável e sua derivada p'_n é limitada.

Agora, de (1.27) podemos escrever

$$p'_n(x) - \psi'(x) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} (-it)e^{-itx} (f_n(t) - f(t)) dt.$$

Nosso objetivo é limitar esta integral com expressões que tendam para zero. Para isso, vamos dividi-la em 3 partes: L_1 , L_2 , L_3 , correspondendo às integrais sobre as regiões $|t| \leq T_n$, $T_n < |t| < T'_n$ e $|t| \geq T'_n$, respectivamente, onde $(T_n)_{n \geq 1}$ e $(T'_n)_{n \geq 1}$ são sequências de números reais positivos que convergem para infinito suficientemente lento.

Primeiramente, consideremos o caso quando $|t| \leq T_n$. Pela hipótese da convergência fraca, temos que $f_n(t) \rightarrow f(t)$ uniformemente sobre todo intervalo limitado. Assim,

$$\delta_n = \sup_{|t| \leq T_n} |f_n(t) - f(t)| \rightarrow 0 \text{ quando } n \rightarrow +\infty.$$

Logo,

$$|L_1| = \left| \int_{|t| \leq T_n} (-it)e^{-itx} (f_n(t) - f(t)) dt \right| \leq \int_{|t| \leq T_n} |t| |f_n(t) - f(t)| dt \leq \delta_n T_n^2,$$

e temos que $L_1 \rightarrow 0$ quando $n \rightarrow +\infty$, como queríamos.

Consideremos o caso $T_n < |t| < T'_n$. Do Corolário 1.3.13 segue que dado $0 < \delta < \alpha$, existem números positivos $c(\delta)$ e ϵ , que não dependem de n , tal que para $|t| \leq \epsilon b_n$, a função característica f_n admite a cota

$$|f_n(t)| \leq e^{-c(\delta)|t|^\delta}. \tag{1.28}$$

De forma análoga, usando a representação (1.10) obtemos a mesma cota para f . Então, tomando $T'_n = \epsilon b_n$, $n \in \mathbb{N}$, temos

$$\begin{aligned}
|L_2| &= \left| \int_{T_n \leq |t| \leq T'_n} (-it)e^{-itx} (f_n(t) - f(t)) dt \right| \\
&\leq \int_{T_n \leq |t| \leq T'_n} |t| (|f_n(t)| + |f(t)|) dt \\
&\leq 2 \int_{T_n \leq |t| \leq T'_n} |t| e^{-c(\delta)|t|^\delta} dt \\
&\leq 2 \int_{|t| \geq T_n} |t| e^{-c(\delta)|t|^\delta} dt.
\end{aligned}$$

Logo, $L_2 \rightarrow 0$ quando $n \rightarrow +\infty$ pois $\lim_{n \rightarrow \infty} T_n = \infty$.

Finalmente, seja $c = \sup_{|t| \geq \epsilon} |f_1(t)|$. De (1.24) temos que $\lim_{t \rightarrow \infty} |f_1(t)|^\nu |t| = 0$, e assim, $\lim_{t \rightarrow \infty} f_1(t) = 0$ e $c < 1$. Consequentemente, lembrando que $T'_n = \epsilon b_n$, para todo $n \geq \nu$, segue

$$\begin{aligned}
\int_{|t| \geq T'_n} |t| |f_n(t)| dt &= \int_{|t| \geq T'_n} |t| \left| e^{-ita_n} \left(f_1 \left(\frac{t}{b_n} \right) \right)^n \right| dt \\
&= b_n^2 \int_{|t| \geq \epsilon} |t| |f_1(t)|^n dt \\
&= b_n^2 \int_{|t| \geq \epsilon} |t| |f_1(t)|^{n-\nu} |f_1(t)|^\nu dt \\
&\leq b_n^2 \int_{|t| \geq \epsilon} |t| c^{n-\nu} |f_1(t)|^\nu dt \\
&= \frac{T_n'^2}{\epsilon^2} c^{n-\nu} \int_{|t| \geq \epsilon} |t| |f_1(t)|^\nu dt.
\end{aligned}$$

Notemos que a última integral tende a zero quando $n \rightarrow \infty$. Por outro lado,

$$\begin{aligned} |L_3| &= \left| \int_{|t| \geq T'_n} (-it)e^{-itx} (f_n(t) - f(t)) dt \right| \\ &\leq \int_{|t| \geq T'_n} |t| |f_n(t) - f(t)| dt \\ &\leq \int_{|t| \geq T'_n} |t| |f_n| dt + \int_{|t| \geq T'_n} |t| |f(t)| dt \end{aligned}$$

e dado que estas duas últimas integrais tendem para zero, obtemos, $L_3 \rightarrow 0$ quando $n \rightarrow \infty$.

Isso completa a prova do resultado. □

Capítulo 2

Informação de Fisher

Neste capítulo apresentamos os conceitos e propriedades relacionadas à Informação de Fisher que serão fundamentais para o desenvolvimento do próximo capítulo.

Iniciamos, na Seção 2.1, com um breve histórico e uma breve discussão sobre os princípios básicos da Teoria da Informação e que servem de motivação para a definição da Informação de Fisher de uma variável aleatória absolutamente contínua a ser utilizada no restante do trabalho.

Na Seção 2.2, introduzimos os conceitos de Informação de Fisher e de Informação de Fisher Relativa (ou distância de Informação de Fisher) e suas propriedades básicas gerais.

Relações entre a Informação de Fisher e a norma da variação são apresentadas na Seção 2.3.

Na Seção 2.4 apresentamos os resultados de Bobkov, Chistyakov e Gotze [6] relativos à Informação de Fisher de densidades que podem ser representadas como convolução de duas densidades com Informação de Fisher finita e que serão úteis para a obtenção dos resultados auxiliares utilizados na demonstração do teorema principal do Capítulo 3.

Finalizamos o capítulo, apresentando na Seção 2.5 uma breve observação sobre as possíveis relações entre a convergência na Informação de Fisher e outros tipos de convergência conhecidos.

2.1 Teoria da Informação e Informação de Fisher

A teoria da informação é uma ramo da Matemática, em particular da Teoria da Probabilidade, que estuda a quantificação da informação. Esta teoria tem muitas aplicações em diversas áreas, como por exemplo, nos sistemas de comunicação, transmissão de dados, criptografia, codificação, teoria do ruído, correção de erros, compressão de dados, entre outras.

A origem da Teoria da Informação é atribuída ao matemático e engenheiro elétrico Claude Shannon no artigo “A mathematical theory of communication” de 1948 ([23]). Um dos conceitos principais da Teoria da Informação é o conceito de Entropia como sendo uma medida de incerteza do valor obtido por uma variável aleatória, ou seja, uma medida da quantidade de “informação” que ganhamos com o conhecimento do valor da variável aleatória. Na verdade, o conceito de entropia surgiu primeiramente no contexto de termodinâmica, em Mecânica Estatística, no século 19, como uma medida de “desordem”.

No contexto da Teoria da Informação, se A é um evento que ocorre com probabilidade $P(A)$, define-se a “informação” $I(A)$, que ganhamos com o conhecimento da ocorrência de A , como sendo

$$I(A) = -\log_2 P(A). \quad (2.1)$$

Na verdade, poderíamos escolher qualquer função crescente de $P(A)$ na definição de informação. Uma justificativa da escolha de \log_2 , usada desde 1928, pode ser encontrada na Seção 1.1.5 de [14]. A escolha do log na base 2 expressa a informação em “bits”, outras bases poderiam ser usadas como 10, e , entre outras, e a troca da base pode ser feita por $I_b(A) = (\log_b a)I_a(A)$. Aqui vamos assumir log na base e , indicado por $\ln x$.

Intuitivamente, (2.1) expressa a ideia de que ganhamos mais informação com o conhecimento da ocorrência de um evento raro, do que de um evento muito provável. Neste sentido, a Entropia de Shanon de uma v.a. X é a quantidade esperada de informação ganha sobre o conhecimento do valor de X . Se X é uma v.a. discreta com função de probabilidade $p(x) = P(X = x)$, então a entropia de X é definida por

$$H(X) = H(p) = -\mathbb{E}(\ln[p(X)]) = -\sum_x \ln[p(x)]p(x).$$

Se X é uma v.a. contínua com densidade $p(x)$, a entropia de X é dada por

$$H(X) = H(p) = -\mathbb{E}(\ln[p(X)]) = -\int_{\mathbb{R}} \ln[p(x)]p(x) dx. \quad (2.2)$$

Para maiores detalhes sobre as propriedades da entropia (de Shanon) vide [14]. Vamos nos concentrar neste trabalho ao caso contínuo.

O uso da Teoria da Informação em Estatística foi introduzido por Kullback e Leibler em 1951 ([16]) e mais tarde desenvolvido por Kullback em 1967 ([15]). Em Estatística a entropia é interpretada como uma medida de incerteza ou de “risco”. Neste sentido, o risco ou divergência

de **Kullback-Leibler** da densidade q relativo à densidade p , ou também chamada **entropia relativa** ou **divergência da Informação**, é definida por

$$D(p \parallel q) = \int_{\mathbb{R}} \ln \left(\frac{p(x)}{q(x)} \right) p(x) dx. \quad (2.3)$$

Em outras palavras, se X tem densidade $p(x)$ então

$$D(p \parallel q) = -\mathbb{E} \left[\ln \left(\frac{q(X)}{p(X)} \right) \right]. \quad (2.4)$$

Neste sentido, se uma observação vem de p , $D(p \parallel q)$ mede o risco de usar q no lugar de p . Lembrando que a entropia $H(p)$ mede o risco associado à p , então

$$D(p \parallel q) + H(p) = -\mathbb{E} [\ln q(X)] = \int_{-\infty}^{\infty} p(x) \ln \left(\frac{1}{q(x)} \right) dx$$

é o risco associado à q , também chamada de **entropia cruzada** (“cross entropy”).

Muitas vezes $D(p \parallel q)$ é chamada de distância. No entanto, formalmente ela não define uma métrica sobre as densidades, pois não é simétrica, nem satisfaz a desigualdade triangular. Propriedades da entropia relativa podem ser encontradas em [14].

Num outro caminho, no contexto da estimação de parâmetros, Ronald Fisher já tinha introduzido no artigo “Theory of Statistical Estimation” publicado em 1925 ([12]), o conceito, denominado posteriormente, Informação de Fisher que está fortemente relacionado com o conceito de entropia relativa. A Informação de Fisher tem inúmeras aplicações em outras diferentes áreas, especialmente na Física, como mostra o livro de Frieden, em [11].

Mais especificamente, seja uma distribuição com densidade $p_{\theta}(x) = p(x; \theta)$, com parâmetro desconhecido θ pertencente a um espaço paramétrico Θ , que aqui vamos assumir unidimensional, isto é, $\Theta \subset \mathbb{R}$. O problema usual em estimação estatística é estimar o valor de θ a partir de uma amostra de distribuição.

Assim, seja (X_1, \dots, X_n) uma amostra aleatória da densidade populacional p_{θ} , ou seja, X_1, \dots, X_n são v.a.’s i.i.d. com densidade $p_{\theta}(x)$. Um estimador $\hat{\theta}$ para θ é uma função da amostra (X_1, \dots, X_n) , ou seja, uma v.a. $\hat{\theta} = \hat{\theta}_n = \hat{\theta}(X_1, \dots, X_n)$.

Do ponto de vista estatístico, um bom estimador deve ser não-viesado (ou não-viciado), ou seja, $\mathbb{E}(\hat{\theta}) = \theta$ e que seja consistente, ou seja, $\hat{\theta} = \hat{\theta}_n \xrightarrow[n \rightarrow \infty]{} \theta$ em algum sentido probabilístico. Uma medida usual do erro desta aproximação é o erro médio quadrático $\mathbb{E}(\hat{\theta} - \theta)^2$, que para estimadores não-viesados é a sua variância, isto é, $Var(\hat{\theta}) = \mathbb{E}(\hat{\theta} - \theta)^2$.

Neste sentido, a Informação de Fisher está relacionada com um erro médio quadrático, através da cota de Cramér-Rao que veremos a seguir, e de certa forma fornece informação quanto à eficiência do estimador.

Neste contexto de estimação de parâmetros define-se a Informação de Fisher como

$$I(\theta) = I(p_\theta) = \mathbb{E} \left[\frac{\partial \ln p_\theta(X)}{\partial \theta} \right]^2 = \int_{\mathbb{R}} \frac{1}{p_\theta(x)} \left[\frac{\partial p_\theta(x)}{\partial \theta} \right]^2 dx. \quad (2.5)$$

A variável aleatória

$$V = \frac{\partial \ln p_\theta(X)}{\partial \theta}$$

é chamada função escore (“score”) e tem média $\mathbb{E}V = 0$. Assim, a Informação de Fisher é a variância da função escore.

Baseados nos resultados de Rao, 1962 [21], podemos obter uma justificativa intuitiva para o nome Informação de Fisher.

Para obter informação sobre como pequenas influências no parâmetro θ afeta a verossimilhança $p_\theta(x) = p(x; \theta)$, consideremos $\theta' = \theta + \Delta\theta$ e analisemos a discrepância entre $p(x; \theta)$ e $p(x; \theta')$. Como medida desta discrepância, ou distância, podemos considerar a entropia relativa, ou divergência de Kullback-Leibler entre $p_\theta(x) = p(x; \theta)$ e $p_{\theta'}(x) = p(x; \theta')$ condicionada ao valor de θ :

$$d(\theta : \theta') = D(p_\theta \parallel p_{\theta'}) = \mathbb{E}_\theta (\ln p_\theta(X) - \ln p_{\theta'}(X)) = \mathbb{E}_\theta \left(\ln \left[\frac{p_\theta(X)}{p_{\theta'}(X)} \right] \right),$$

onde E_θ indica a esperança condicionada ao valor dado θ .

Supondo que $p'_{\theta'}(x)$ é duas vezes diferenciável com respeito a θ , usando a expansão de Taylor até segunda ordem, é possível mostrar que para $\Delta\theta$ pequeno

$$d(\theta : \theta') \simeq \frac{1}{2} I(\theta) (\Delta\theta)^2.$$

Ou seja, $I(\theta)$ é uma medida da discrepância entre as duas distribuições. Pode-se verificar que o argumento anterior não depende da medida $d(\theta : \theta')$ de discrepância escolhida. (Vide por exemplo, [21]).

Assim, podemos dizer que a entropia relativa é uma medida de discrepância global, enquanto a Informação de Fisher é uma medida de discrepância local, que nos fornece a quantidade de “informação ” fornecida pelo parâmetro θ .

Voltando ao contexto de estimação do parâmetro θ , para um estimador $\hat{\theta}$ não-viesado temos a seguinte desigualdade, conhecida como desigualdade de Cramér-Rao,

$$\text{Var}(\hat{\theta}) = \mathbb{E}(\hat{\theta} - \theta)^2 \geq \frac{1}{I(\theta)}. \quad (2.6)$$

No caso da igualdade em (2.6), dizemos que $\hat{\theta}$ é um estimador eficiente. Assim, $I(\theta)$ é uma medida de eficiência do estimador $\hat{\theta}$.

No caso particular em que θ é um parâmetro de localização e temos que $p_\theta(x) = h(x - \theta)$, então

$$\frac{\partial h(x - \theta)}{\partial \theta} = -\frac{\partial h(x - \theta)}{\partial x}$$

e assim, neste caso, podemos escrever a Informação de Fisher da densidade p , dada em (2.5), como sendo

$$I(p) = \int_{-\infty}^{\infty} \frac{1}{p_\theta(x)} \left[\frac{dp_\theta}{dx}(x) \right]^2 dx. \quad (2.7)$$

Esta representação, neste caso especial, motiva a definição de Informação de Fisher para uma v.a. contínua arbitraria X com densidade $p(x)$, que será adotada e estudada nas próximas seções.

2.2 Propriedades Gerais da Informação de Fisher

Motivados pelas observações no final da seção anterior, podemos definir a Informação de Fisher de uma densidade p derivável.

Definição 2.2.1. *Se uma variável aleatória X tem densidade p absolutamente contínua com derivada p' , definimos a Informação de Fisher de X como sendo*

$$I(X) = I(p) = \int_{\{x: p(x) > 0\}} \frac{[p'(x)]^2}{p(x)} dx. \quad (2.8)$$

Quando $\tilde{p}(x) = p(x)$ q.c.x, escrevemos $I(\tilde{p}) = I(p)$. Em qualquer outro caso, $I(X) = \infty$.

A função $\rho_X(x) = \frac{d}{dx} \log(p(x)) = \frac{p'(x)}{p(x)}$ é chamada função escore. Assim,

$$I(X) = \mathbb{E}\rho_X^2(X) = \int_{\{x: p(x) > 0\}} \left[\frac{p'(x)}{p(x)} \right]^2 p(x) dx.$$

Exemplo 5. Se $X \stackrel{\mathcal{D}}{=} N(\mu, \sigma^2)$, então a função score $\rho_X(x) = \frac{-(x - \mu)}{\sigma^2}$ e $I(X) = \frac{1}{\sigma^2}$. Como

$p(x) = p(0) \exp \left\{ \int_0^x \rho_X(u) du \right\}$, então a função score é linear se, e somente se, $X \stackrel{\mathcal{D}}{=} N(\mu, \sigma^2)$.

Esta caracterização para as distribuições normais é muito usada na prova de diferentes resultados sobre a Informação de Fisher.

Observação 2.2.2.

(a) Da definição anterior, podemos definir I como um funcional sobre o espaço de todas as densidades de probabilidade. Se $I(X)$ é finita e p é a densidade de X , sempre assumimos que p é absolutamente contínua. Além disso, a derivada $p'(x)$ da densidade $p(x)$ existe e é finita a menos de um conjunto de medida de Lebesgue nula.

(b) Se $I(X) < \infty$ e p é diferenciável em x_0 , então $p'(x_0) = 0$ sempre que $p(x_0) = 0$, ver [5], Proposição 2.1. Como consequência, o domínio de integração em (2.8), pode ser estendido sobre a reta toda, ou seja,

$$I(X) = I(p) = \int_{-\infty}^{+\infty} \frac{p'(x)^2}{p(x)} dx.$$

(c) A Desigualdade de Cramér-Rao mostra que o funcional Informação de Fisher é minimizado pela densidade Normal. Em outras palavras, dada uma v.a. X com média μ e variância σ^2 , a informação de Fisher $I(X)$ é limitada inferiormente pelo inverso da variância, ou seja,

$$I(X) \geq \frac{1}{\sigma^2}.$$

A igualdade é atingida se, e somente se, $X \sim N(\mu, \sigma^2)$.

(d) Para $a, b \in \mathbb{R}$, $b \neq 0$ temos que: $I(a + bX) = \frac{1}{b^2} I(X)$. De fato: se $p(x)$ é a densidade de X , então a densidade de $a + bX$ é dada por $\frac{1}{b} p\left(\frac{1}{b}(y - a)\right)$ e daí por (2.8),

$$I(a + bX) = \frac{1}{b^3} \int_{-\infty}^{+\infty} p' \left(\frac{1}{b}(y - a) \right)^2 \left[p \left(\frac{1}{b}(y - a) \right) \right]^{-1} dy = \frac{1}{b^2} \int_{-\infty}^{+\infty} \frac{p'(x)^2}{p(x)} dx = \frac{1}{b^2} I(X),$$

onde na penúltima igualdade tomamos, $x = \frac{1}{b}(y - a)$.

(e) O funcional I definido sobre o espaço de densidades

$$S = \{p : p \text{ é densidade de probabilidade sobre } \mathbb{R}\}$$

é convexo. Ou seja, para $p, q \in S$ e $\alpha \in [0, 1]$, temos $I(\alpha p + (1 - \alpha)q) \leq \alpha I(p) + (1 - \alpha)I(q)$. De fato, dado que a função $h(u, v) = \frac{u^2}{v}$ é convexa para $u, v \in \mathbb{R}$, $v > 0$, temos que se $(x, y), (u, v) \in \mathbb{R} \times (0, +\infty)$ e $\alpha \in [0, 1]$,

$$h(\alpha(u, v) + (1 - \alpha)(x, y)) \leq \alpha h(u, v) + (1 - \alpha)h(x, y).$$

Tomando $(u, v) = (p'(x), p(x))$ e $(x, y) = (q'(x), q(x))$, obtemos

$$\frac{(\alpha p'(x) + (1 - \alpha)q'(x))^2}{(\alpha p(x) + (1 - \alpha)q(x))^2} \leq \alpha \frac{p'(x)^2}{p(x)} + (1 - \alpha) \frac{q'(x)^2}{q(x)}.$$

O resultado segue usando a monotonicidade da integral.

Como consequência da convexidade obtemos a desigualdade de Jensen: para $p_1, \dots, p_n \in S$, $n \in \mathbb{N}$, temos:

$$I(\alpha_1 p_1 + \alpha_2 p_2 + \dots + \alpha_n p_n) \leq \sum_{k=1}^n \alpha_k I(p_k) \text{ onde } \alpha_k > 0, \sum_{k=1}^n \alpha_k = 1.$$

As desigualdades da próxima proposição nos permitirão obter uma cota superior para a Informação de Fisher da convolução de densidades, ou seja, de soma de v.a.'s independentes.

Proposição 2.2.3. *Sejam X e Y v.a.'s independentes. Então, para qualquer $\beta \in [0, 1]$:*

1. $I(X + Y) \leq \beta^2 I(X) + (1 - \beta)^2 I(Y)$,
2. $I(\sqrt{\beta}X + \sqrt{1 - \beta}Y) \leq \beta I(X) + (1 - \beta)I(Y)$.

Demonstração. Ver [14], pág. 26. □

Em particular, escolhendo adequadamente β na anterior proposição, obtemos o seguinte resultado:

Proposição 2.2.4 (Desigualdades de Stam). *Se X e Y são v.a.'s independentes, então*

$$I(X + Y) \leq I(Y)$$

e

$$\frac{1}{I(X+Y)} \geq \frac{1}{I(X)} + \frac{1}{I(Y)}. \quad (2.9)$$

Com igualdade se, e somente se, X e Y seguem distribuições normais. Generalizando, se X_1, X_2, \dots, X_k são v.a.'s independentes, então

$$\frac{1}{I(X_1 + \dots + X_k)} \geq \frac{1}{I(X_1)} + \dots + \frac{1}{I(X_k)}.$$

Demonstração. Basta considerar $\beta = 1$ para obter $I(X+Y) \leq I(Y)$ e $\beta = \frac{I(Y)}{I(X)+I(Y)}$ para obter (2.9). \square

Da proposição anterior segue que: se X e Y são v.a.'s independentes, então

$$\frac{1}{I(X+Y)} \geq \frac{1}{I(X)} + \frac{1}{I(Y)} \geq \frac{1}{I(X)}.$$

Portanto, $I(X+Y) \leq I(X)$. Similarmente, $I(X+Y) \leq I(Y)$. Logo,

$$I(X+Y) \leq \min \{I(X), I(Y)\}. \quad (2.10)$$

Assim, para a convolução de quaisquer duas densidades p e q

$$p * q(x) = \int_{-\infty}^{+\infty} p(x-y)q(y)dx,$$

temos que (2.10) pode ser escrita como

$$I(p * q) \leq \min \{I(p), I(q)\}. \quad (2.11)$$

Das desigualdades anteriores podemos ver que o funcional I decresce quando adicionamos uma parcela independente, por isso esta propriedade costuma ser chamada de monotonicidade da Informação de Fisher.

Para finalizar esta seção, definimos a distância de Informação de Fisher entre duas variáveis aleatórias absolutamente contínuas e apresentamos duas propriedades relacionadas à distribuições estáveis que serão utilizadas no próximo capítulo.

Definição 2.2.5. *Sejam X e Y v.a.'s com densidades p e q respectivamente. Definimos a distância de Informação de Fisher, ou Informação de Fisher relativa com sendo:*

$$I(X \parallel Y) = I(p \parallel q) = \int_{\mathbb{R}} \left(\frac{p'(x)}{p(x)} - \frac{q'(x)}{q(x)} \right)^2 p(x) dx. \quad (2.12)$$

Ou seja,

$$I(X \parallel Y) = \mathbb{E}(\rho_X(X) - \rho_X(Y))^2 = \int_{\mathbb{R}} \left(\frac{d}{dx} \log \left[\frac{p(x)}{q(x)} \right] \right)^2 p(x) dx, \text{ onde } \rho_X(x) \text{ é a função}$$

escore de X .

Notemos que, apesar de chamada de distância, na verdade $I(p \parallel q)$ não é uma métrica, pois não é simétrica e não satisfaz a desigualdade triangular. Porém, $I(p \parallel q) \geq 0$ e $I(p \parallel q) = 0$ se, e somente se, $p(x) = Cq(x)$ *q.c.x*, $C \in \mathbb{R}$.

Observação 2.2.6. *A Desigualdade de Cramér-Rao, motiva a definição de uma versão generalizada da Informação de Fisher: se X é uma v.a. com média μ , variância σ^2 e função escore ρ_X , definimos a informação de Fisher padrão ou normalizada de X por:*

$$I_p(X) = \sigma^2 I(X \parallel Z) = \sigma^2 \mathbb{E} \rho_X^2 - 1 = \sigma^2 \mathbb{E} \left(\rho_X + \frac{X - \mu}{\sigma^2} \right),$$

onde Z é uma v.a. com distribuição $N(\mu, \sigma^2)$.

As propriedades a seguir relacionam a Informação de Fisher com a distância da informação de Fisher, pois mostram, no caso que Z é estável não-extremal e não-normal, que $I(X \parallel Z)$ é finito se, e somente se, $I(X)$ é finito. No caso que Z é normal, $I(X \parallel Z) < \infty$ se, e somente se, $I(X) < \infty$ e $\mathbb{E}(X^2) < \infty$.

Proposição 2.2.7. *Se Z é uma v.a. com distribuição estável não extremal de índice α , $\alpha \in (0, 2)$, então para qualquer v.a. X ,*

$$I(X \parallel Z) \leq 2I(X) + c, \tag{2.13}$$

$$I(X) \leq 2I(X \parallel Z) + c, \tag{2.14}$$

onde a constante positiva c depende unicamente da v.a. Z . Além disso, temos $I(X \parallel Z) < +\infty$ se, e somente se, $I(X) < +\infty$.

Demonstração.

Se $I(X)$ ou $I(X \parallel Z) = +\infty$, obtemos trivialmente as desigualdades. Vamos assumir inicialmente $I(X) < +\infty$.

Pelas Observações 1.3.3(c) e 1.3.6 temos que a densidade ψ de uma distribuição estável não-extremal não-normal é positiva e suave, tal que para todo $k = 1, 2, \dots$,

$$|(\log \psi(x))^{(k)}| \sim \frac{(k-1)!}{|x|^k} \text{ quando } |x| \rightarrow \infty$$

e

$$\frac{|\psi'(x)|}{\psi(x)} \leq \frac{c}{1+|x|}, \text{ para } x \in \mathbb{R} \text{ e algum } c > 0.$$

Assim, como $(a - b)^2 \leq 2(a^2 + b^2)$ para $a, b \in \mathbb{R}$ e como $\frac{c}{1 + |x|} \leq c$, segue

$$\left(\frac{p'(x)}{p(x)} - \frac{\psi'(x)}{\psi(x)}\right)^2 \leq 2\left(\frac{p'(x)}{p(x)}\right)^2 + 2\left(\frac{\psi'(x)}{\psi(x)}\right)^2 \leq 2\left(\frac{p'(x)}{p(x)}\right)^2 + 2c^2.$$

Logo,

$$\int_{-\infty}^{+\infty} \left(\frac{p'(x)}{p(x)} - \frac{\psi'(x)}{\psi(x)}\right)^2 p(x) dx \leq \int_{-\infty}^{+\infty} 2\left(\frac{p'(x)}{p(x)}\right)^2 p(x) dx + \int_{-\infty}^{+\infty} 2c^2 p(x) dx.$$

Isto mostra (2.13) e segue que $I(X \parallel Z) < +\infty$.

Da mesma forma, assumindo $I(X \parallel Z) < +\infty$ podemos escrever

$$\left(\frac{p'(x)}{p(x)}\right)^2 \leq 2\left(\frac{p'(x)}{p(x)} - \frac{\psi'(x)}{\psi(x)}\right)^2 + 2\left(\frac{\psi'(x)}{\psi(x)}\right)^2 \leq \left(\frac{p'(x)}{p(x)} - \frac{\psi'(x)}{\psi(x)}\right)^2 + 2c^2,$$

e obter (2.14). De onde segue que $I(X) < +\infty$. □

Proposição 2.2.8. *Se Z tem uma distribuição normal, então $I(X \parallel Z) < \infty$ se, e somente se, $I(X) < \infty$ e $\mathbb{E}(X^2) < \infty$.*

Demonstração. Como $Z \stackrel{\mathcal{D}}{=} N(a, \sigma^2)$, então $\frac{\psi'(x)}{\psi(x)} = -\frac{x - a}{\sigma^2}$. Usando o mesmo argumento da proposição anterior podemos escrever

$$\left(\frac{p'(x)}{p(x)} - \frac{\psi'(x)}{\psi(x)}\right)^2 \leq 2\left(\frac{p'(x)}{p(x)}\right)^2 + 2\left(\frac{\psi'(x)}{\psi(x)}\right)^2 = 2\left(\frac{p'(x)}{p(x)}\right)^2 + \frac{2}{\sigma^4}(x - a)^2 \quad (2.15)$$

e

$$\left(\frac{p'(x)}{p(x)}\right)^2 - \frac{2}{\sigma^4}(x - a)^2 = \left(\frac{p'(x)}{p(x)}\right)^2 - 2\left(\frac{\psi'(x)}{\psi(x)}\right)^2 \leq 2\left(\frac{p'(x)}{p(x)} - \frac{\psi'(x)}{\psi(x)}\right)^2. \quad (2.16)$$

Assim, assumindo $I(X) < +\infty$ e $\mathbb{E}X^2 < +\infty$ podemos obter de (2.15) que

$$I(X \parallel Z) \leq 2I(X) + \frac{2}{\sigma^4}\mathbb{E}(X^2) - \frac{4a}{\sigma^4}\mathbb{E}(X) + \frac{2}{\sigma^4}a^2 < +\infty.$$

Por outro lado, assumindo $I(X \parallel Z) < \infty$ podemos obter de (2.16)

$$I(X) - \frac{2}{\sigma^2}\mathbb{E}(X - a)^2 \leq 2I(X \parallel Z) < +\infty.$$

Segue que $I(X) < +\infty$ e $\mathbb{E}(X - a)^2 < +\infty$. Logo $\mathbb{E}X^2 < +\infty$. □

2.3 Informação de Fisher e Funções de Variação Limitada

Nesta seção apresentamos desigualdades envolvendo a variação total e a Informação de Fisher. Em particular, obtemos cotas superiores para uma densidade p e sua respectiva função característica a partir da Informação de Fisher. No caso especial de convolução de três densidades de variação limitada podemos obter uma cota superior para a variação total e para a Informação de Fisher a partir da função característica.

Como vimos na Seção 1.2, denotamos a variação total da densidade p por $\|p\|_{TV}$.

Nesta primeira proposição provamos que a variação total de uma função de densidade p com Informação de Fisher finita é limitada superiormente por $\sqrt{I(p)}$. Uma consequência deste resultado é que a convergência na Informação de Fisher é mais forte do que na norma de variação total.

Proposição 2.3.1. *Se uma v.a. X tem densidade p absolutamente contínua com derivada p' , então*

$$\|p\|_{TV} = \int_{-\infty}^{+\infty} |p'(x)| dx \leq \sqrt{I(X)}. \quad (2.17)$$

Demonstração. Se $I(X) = \infty$, segue imediatamente (2.17).

Vamos supor que $I(X) < +\infty$. Segue da Definição 2.2.1 que as funções $\frac{p'(x)}{p(x)}$ e $\sqrt{p(x)}$ são quadrado integráveis. Assim, usando a desigualdade de Cauhy-Schwartz segue que

$$\|p\|_{TV} = \int_{-\infty}^{+\infty} |p'(x)| dx = \int_{-\infty}^{+\infty} |p'(x)| \frac{\sqrt{p(x)}}{\sqrt{p(x)}} dx \leq \left(\int_{-\infty}^{+\infty} \frac{p'(x)^2}{p(x)} dx \right)^{\frac{1}{2}} \left(\int_{-\infty}^{+\infty} p(x) dx \right)^{\frac{1}{2}} = \sqrt{I(X)}.$$

□

Observação 2.3.2. *Seja X uma v.a. com Informação de Fisher finita.*

(a) *A densidade p de X é uma função de variação limitada. Então os limites*

$$p(-\infty) = \lim_{x \rightarrow -\infty} p(x), \quad p(+\infty) = \lim_{x \rightarrow +\infty} p(x),$$

existem e são finitos. Como toda densidade é não negativa e integrável em \mathbb{R} segue que estes limites devem ser ambos zero.

(b) Seja f a função característica da v.a. X , ou seja, para $t \in \mathbb{R}$,

$$f(t) = \mathbb{E}(e^{itX}) = \int_{-\infty}^{+\infty} e^{itx} p(x) dx.$$

Integrando por partes e usando o fato que $\lim_{x \rightarrow +\infty} p(x) = 0$ e $\lim_{x \rightarrow -\infty} p(x) = 0$ obtemos para $t \neq 0$ que,

$$f(t) = - \int_{-\infty}^{+\infty} \frac{e^{itx}}{it} p'(x) dx.$$

Logo,

$$|f(t)| = \left| - \int_{-\infty}^{+\infty} \frac{e^{itx}}{it} p'(x) dx \right| \leq \int_{-\infty}^{+\infty} \left| \frac{e^{itx}}{it} p'(x) \right| dx = \frac{1}{|t|} \int_{-\infty}^{+\infty} |p'(x)| dx.$$

Então, para $t \neq 0$, segue de (2.17) que

$$|f(t)| \leq \frac{1}{|t|} \|p\|_{TV}. \quad (2.18)$$

(c) Tendo em vista que p é absolutamente contínua, obtemos

$$|p(x)| = \left| \int_{-\infty}^x p'(y) dy \right| \leq \int_{-\infty}^x |p'(y)| dy \leq \int_{-\infty}^{+\infty} |p'(y)| dy = \|p\|_{TV}.$$

Em outras palavras, p é limitada uniformemente por $\|p\|_{TV}$.

Das observações anteriores e usando (2.17) temos que se uma v.a. X tem Informação de Fisher finita, sua densidade p e sua função característica $f(t)$ possuem cotas superiores análogas às obtidas pela variação total, usando a Informação de Fisher, ou seja,

$$\sup_x p(x) \leq \sqrt{I(X)} \quad \text{e} \quad |f(t)| \leq \frac{\sqrt{I(X)}}{|t|} \quad \text{onde } t \neq 0. \quad (2.19)$$

A proposição a seguir apresenta uma cota superior da Informação de Fisher de uma soma de três v.a.'s i.i.d. a partir da norma de variação total. Em outras palavras, se p é a convolução de três densidades com variação limitada obtemos uma espécie de recíproca da desigualdade (2.17).

Proposição 2.3.3. *Sejam X_1, X_2 e X_3 v.a.'s independentes com densidades p_1, p_2 e p_3 respectivamente, de variação limitada. Então a v.a. $Y = X_1 + X_2 + X_3$ tem informação de Fisher finita e, além disso,*

$$I(Y) = I(p_1 * p_2 * p_3) \leq \frac{1}{2} [\|p_1\|_{TV} \|p_2\|_{TV} + \|p_1\|_{TV} \|p_3\|_{TV} + \|p_2\|_{TV} \|p_3\|_{TV}]. \quad (2.20)$$

Demonstração. Ver [24]. □

Observação 2.3.4. *A Proposição 2.3.3 não é válida para a soma de duas v.a.'s i.i.d. com densidades de variação limitada. pois se $Y = X_1 + X_2$, onde X_1, X_2 são v.a.'s independentes com densidades p_1, p_2 de variação limitada, pode acontecer que a Informação de Fisher de Y seja infinita. Por exemplo, consideremos X_1, X_2 v.a.'s independentes que seguem uma distribuição uniforme no intervalo $(-\frac{1}{2}, \frac{1}{2})$ e $Y = X_1 + X_2$. Claramente X_1, X_2 são de variação limitada e a densidade de Y é dada por*

$$g(x) = \begin{cases} 1+x, & \text{se } -1 \leq x < 0 \\ 1-x, & \text{se } 0 \leq x \leq 1 \\ 0, & \text{caso contrario.} \end{cases}$$

Mas,

$$I(Y) = \int_{-\infty}^{+\infty} \frac{g'(x)^2}{g(x)} dx = \int_{-1}^0 \frac{1}{1+x} dx + \int_0^1 \frac{1}{1-x} dx = +\infty.$$

Para finalizar esta seção apresentamos duas proposições que fornecem cotas superiores para a variação total de uma densidade p de uma v.a. X em termos de sua função característica. No caso que a v.a. X satisfaça as condições da Proposição 2.3.3 segue que podemos obter cotas para a Informação de Fisher de X a partir de sua função característica.

Vale a pena notar que as cotas apresentadas a seguir dependem especialmente das propriedades de integrabilidade de f e suas derivadas, as quais também podem depender da condição de finitude de alguns momentos de X .

Proposição 2.3.5. *Se a v.a. X tem segundo momento finito e*

$$\int_{-\infty}^{+\infty} |t| (|f(t)| + |f'(t)| + |f''(t)|) dt < \infty,$$

então X tem densidade p continuamente diferenciável com variação total finita tal que

$$\|p\|_{TV} \leq \frac{1}{2} \int_{-\infty}^{+\infty} (|tf''(t)| + 2|f'(t)| + |tf(t)|) dt.$$

Demonstração. Ver [3], Proposição 5.1, pág.44. □

Proposição 2.3.6. *Suponhamos que a função característica $f(t)$ de uma v.a. X tem derivada $f'(t)$ contínua para todo $t > 0$ e que*

$$\int_{-\infty}^{+\infty} t^2 (|f(t)|^2 + |f'(t)|^2) dt < \infty.$$

Então X tem densidade p absolutamente contínua com variação total finita tal que

$$\|p\|_{TV} \leq \left(\int_{-\infty}^{+\infty} |tf(t)|^2 dt \int_{-\infty}^{+\infty} |(tf(t))'|^2 dt \right)^{\frac{1}{4}}.$$

Demonstração. Ver [3], Proposição 5.2, pág.25. □

2.4 Informação de Fisher e Convolução de Densidades

Vimos na Observação 2.3.4 que a Proposição 2.3.3 não pode ser aplicada para a convolução de duas densidades de variação limitada. Nesta seção apresentamos os resultados obtidos por Bobkov, Chistyakov e Gotze [6], que obtiveram cotas superiores, similares às obtidas na Proposição 2.3.3, para a Informação de Fisher de densidades que podem ser representadas como convolução de duas densidades com Informação de Fisher finitas.

Para isto, vamos definir algumas notações a serem usadas.

Definição 2.4.1. *Dado um número real $I > 0$ denotamos por $\mathfrak{B}_2(I)$ a coleção de todas as funções sobre a reta real as quais podem ser representadas como convolução de duas densidades com Informação de Fisher no máximo I . Ou seja,*

$$\mathfrak{B}_2(I) = \{p : p = p_1 * p_2, \text{ com } p_1 \text{ e } p_2 \text{ densidades sobre } \mathbb{R}, \text{ tais que } I(p_i) \leq I, i = 1, 2\}.$$

Além disso, a coleção de todas as funções sobre a reta real representáveis como convolução de duas densidades de probabilidade com informação de Fisher finita será denotada por

$$\mathfrak{B}_2 = \cup_{I>0} \mathfrak{B}_2(I).$$

Observação 2.4.2. *Similarmente, dado um número real $I > 0$ e um inteiro $k \geq 1$ podemos definir $\mathfrak{B}_k(I)$ como a coleção de todas as funções sobre a reta real as quais podem ser representadas como convoluções de k densidades de probabilidade com Informação de Fisher no máximo I . Respectivamente,*

$$\mathfrak{B}_k = \cup_{I>0} \mathfrak{B}_k(I)$$

denota a coleção de todas as funções sobre a reta real representáveis como convolução de k densidades de probabilidade com informação de Fisher finita e \mathfrak{B}_1 será a coleção de todas as densidades com informação de Fisher finita.

Note que pela Desigualdade (2.9) da Proposição 2.2.4, segue que

$$p = p_1 * \cdots * p_k \in \mathfrak{B}_1 \left(\frac{I}{k} \right) \text{ sempre que } p_i \in \mathfrak{B}_1(I), \text{ para } i = 1, 2, \dots, k.$$

Em particular,

$$I(p) \leq \frac{I}{2} \text{ para qualquer } p \in \mathfrak{B}_2(I).$$

As densidades da coleção \mathfrak{B}_k satisfazem certas propriedades interessantes, como por exemplo, toda função $p \in \mathfrak{B}_k$ é $k-1$ vezes diferenciável e sua $(k-1)$ -ésima derivada é absolutamente contínua. Nós vamos ilustrar estas propriedades no caso $k=2$, pois é o caso de nosso interesse. Assim, se $X = X_1 + X_2$, com X_1, X_2 v.a.'s independentes com densidades p_1 e p_2 respectivamente, e $I(p_i) < +\infty$, $i = 1, 2$, então $p \in \mathfrak{B}_2$ e

$$p(x) = \int_{-\infty}^{+\infty} p_1(x-y)p_2(y)dy = \int_{-\infty}^{+\infty} p_1(y)p_2(x-y)dy. \quad (2.21)$$

Das propriedades das convoluções podemos derivar dentro do sinal da integral e obter a derivada da função p . Ou seja,

$$p'(x) = \int_{-\infty}^{+\infty} p'_1(x-y)p_2(y)dy = \int_{-\infty}^{+\infty} p'_1(y)p_2(x-y)dy. \quad (2.22)$$

Analogamente,

$$p''(x) = \int_{-\infty}^{+\infty} p'_1(x-y)p'_2(y)dy = \int_{-\infty}^{+\infty} p'_1(y)p'_2(x-y)dy \quad (2.23)$$

está bem definida para todo $x \in \mathbb{R}$ e podemos escrevê-la como convolução das funções p'_1 e p'_2 , que são integráveis, conforme a Proposição 2.3.1.

Assim, usando o Teorema de Fubini e o fato que p_2 é absolutamente contínua, obtemos que $p'(x)$ também é absolutamente contínua, pois

$$p'(x) = \int_{-\infty}^x p''(z)dz.$$

As fórmulas anteriores podem ser usadas para derivar várias relações elementares para a classe \mathfrak{B}_2 , que apresentamos a seguir.

Proposição 2.4.3. Dada uma densidade $p \in \mathfrak{B}_2(I)$, para todo $x \in \mathbb{R}$, temos

$$|p'(x)| \leq I^{\frac{3}{4}} \sqrt{p(x)} \leq I. \quad (2.24)$$

Além disso, p' tem variação total finita tal que

$$\|p'\|_{TV} = \int_{-\infty}^{+\infty} |p''(x)| dx \leq I. \quad (2.25)$$

Demonstração. Sejam p_1, p_2 densidades de probabilidade tais que $p = p_1 * p_2$, $I(p_1) \leq I$ e $I(p_2) \leq I$.

Primeiramente, provemos (2.25) e assim obtemos que p' tem variação total finita. De fato, usando (2.23) e o Teorema de Fubini, temos

$$\|p'\|_{TV} = \int_{-\infty}^{+\infty} |p''(x)| dx \leq \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} |p'_1(y)| |p'_2(x-y)| dy dx = \int_{-\infty}^{+\infty} |p'_1(y)| dy \int_{-\infty}^{+\infty} |p'_2(z)| dz.$$

Mas, por (2.17),

$$\int_{-\infty}^{+\infty} |p'_1(y)| dy \int_{-\infty}^{+\infty} |p'_2(z)| dz \leq \sqrt{I(p_1)} \sqrt{I(p_2)} \leq I,$$

e (2.25) segue.

Para mostrar (2.24), sejam $u_1(x) = \frac{p'_1(x)}{\sqrt{p_1(x)}} 1_{\{x:p_1(x)>0\}}$ e $u_2(x) = \frac{p'_2(x)}{\sqrt{p_2(x)}} 1_{\{x:p_2(x)>0\}}$. Notemos que u_1, u_2 são funções quadrado integráveis, pois $I(p_i) \leq I$, $I = 1, 2$.

Por outro lado, de (2.22), temos

$$p'(x) = \int_{-\infty}^{+\infty} p'_1(x-y) p_2(y) dy = \int_{-\infty}^{+\infty} u_1(x-y) \sqrt{p_1(x-y)} p_2(y) dy.$$

Aplicando a Desigualdade de Cauchy-Schwartz e tomando $z = x - y$, obtemos

$$\begin{aligned} [p'(x)]^2 &= \left(\int_{-\infty}^{+\infty} u_1(x-y) \sqrt{p_1(x-y)} p_2(y) dy \right)^2 \\ &\leq \int_{-\infty}^{+\infty} u_1(z)^2 dz \int_{-\infty}^{+\infty} p_1(x-y) [p_2(y)]^2 dy \\ &= I(p_1) \int_{-\infty}^{+\infty} p_1(x-y) [p_2(y)]^2 dy. \end{aligned}$$

Agora, por (2.19), temos que $\sup_x p_i(x) \leq \sqrt{I(p_i)}$, $i = 1, 2$. Então,

$$\begin{aligned} [p'(x)]^2 &\leq I(p_1)\sqrt{I(p_2)} \int_{-\infty}^{+\infty} p_1(x-y)p_2(y)dy \\ &= I(p_1)\sqrt{I(p_2)}p(x). \end{aligned}$$

Mas, $p_1, p_2 \in \mathfrak{B}_2$, por hipótese. Então $I(p_i) \leq I$, $i = 1, 2$ e daí segue

$$[p'(x)]^2 \leq I^{3/2}p(x).$$

Finalmente, como $\sup_x p(x) \leq \sqrt{I(p)}$ e por (2.11), $I(p) = I(p_1 * p_2) \leq \min\{I(p_1), I(p_2)\} \leq I$, segue que

$$|p'(x)| \leq I^{3/4}\sqrt{p(x)} \leq I^{3/4}[I(p)]^{1/4} \leq I,$$

e (2.24) está provado. □

Agora, usando argumentos análogos à prova da proposição anterior, podemos obter uma cota similar para p'' .

Proposição 2.4.4. *Seja p qualquer função em \mathfrak{B}_2 . Se para algum x , $p(x) = 0$ então $p''(x) = 0$.*

Além disso,

$$\int_{\{x:p(x)>0\}} \frac{[p''(x)]^2}{p(x)} dx \leq I^2. \quad (2.26)$$

Demonstração. Por (2.23), temos que para $x \in \mathbb{R}$ fixo, podemos escrever

$$p''(x) = \int_{-\infty}^{+\infty} p'_1(y)p'_2(x-y)1_A(y)dy, \quad (2.27)$$

onde $A = \{y \in \mathbb{R} : p'_1(y)p'_2(x-y) > 0\}$.

Por outro lado, se $p(x) = 0$, ou seja, $\int_{-\infty}^{+\infty} p_1(y)p_2(x-y)dy = 0$ e então $p_1(y)p_2(x-y) = 0$ q.c. y .

Assim $1_A(y) = 0$ q.c. y . Logo, a integral em (2.27) é zero e portanto $p''(x) = 0$. Assim, por facilidade, podemos considerar a integral sobre \mathbb{R} .

Para provar (2.26) consideremos as funções u_1 e u_2 definidas na demonstração da proposição anterior. Agora, de (2.23) temos,

$$p''(x) = \int_{-\infty}^{+\infty} p'_1(x-y)p'_2(y)dy = \int_{-\infty}^{+\infty} u_1(x-y)u_2(y)\sqrt{p_1(x-y)p_2(y)}dy.$$

Pela Desigualdade de Cauchy-Schwartz segue

$$\begin{aligned}
[p''(x)]^2 &= \left(\int_{-\infty}^{+\infty} u_1(x-y)u_2(y)\sqrt{p_1(x-y)p_2(y)}dy \right)^2 \\
&\leq \int_{-\infty}^{+\infty} [u_1(x-y)]^2[u_2(y)]^2dy \int_{-\infty}^{+\infty} p_1(x-y)p_2(y)dy \\
&= [u(x)]^2p(x),
\end{aligned} \tag{2.28}$$

onde,

$$[u(x)]^2 = \int_{-\infty}^{+\infty} [u_1(x-y)]^2[u_2(y)]^2dy.$$

Agora, usando o Teorema de Fubini,

$$\int_{-\infty}^{+\infty} [u(x)]^2dx = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} [u_1(x-y)]^2[u_2(y)]^2dydx = \int_{-\infty}^{+\infty} [u_1(x-y)]^2dx \int_{-\infty}^{+\infty} [u_2(y)]^2dy = I(p_1)I(p_2) \leq I^2.$$

Deste e de (2.28) segue que

$$\int_{-\infty}^{\infty} \frac{[p''(x)]^2}{p(x)}dx \leq \int_{-\infty}^{\infty} [u(x)]^2dx \leq I^2.$$

□

O resultado final que obteremos nesta seção será uma estimativa “tipo-cauda” para a Informação de Fisher, que seguirá como corolário da seguinte proposição:

Proposição 2.4.5. *Seja X uma v.a. tal que sua densidade $p \in \mathfrak{B}_2$. Se*

$$\int_{-\infty}^{+\infty} |p''(x) \log p(x)|dx < +\infty,$$

então

$$I(X) = - \int_{-\infty}^{+\infty} p''(x) \log p(x)dx. \tag{2.29}$$

Demonstração. Ver [7], Proposição 8.1, pág.33.

□

Corolário 2.4.6. Se $p \in \mathfrak{B}_2(I)$, então para qualquer $T \in \mathbb{R}$,

$$\int_T^\infty \frac{[p'(x)]^2}{p(x)} dx \leq I^{\frac{3}{4}} \sqrt{p(T)} |\log p(T)| + I \left(\int_T^\infty p(x) [\log p(x)]^2 dx \right)^{\frac{1}{2}}. \quad (2.30)$$

Demonstração. Consideremos o aberto $G = \{x > T : p(x) > 0\}$. Então podemos escrever

$$G = \bigcup_{n \in \mathbb{N}} (a_n, b_n), \text{ onde } (a_n, b_n) \cap (a_m, b_m) = \emptyset \text{ se } n \neq m \text{ e } T \leq a_n < b_n \leq \infty.$$

Note que, se $a_n > T$, temos $p(a_n) = 0$. Assim por (2.24) e pela continuidade de p , segue

$$p'(x) \log p(x) \rightarrow 0 \text{ quando } x \downarrow a_n.$$

Similarmente, se $b_n < \infty$, $p(b_n) = 0$ e, além disso, $p(\infty) = \lim_{x \uparrow +\infty} p(x) = 0$.

Agora, sejam $n \in \mathbb{N}$ e T_1, T_2 números reais tais que $a_1 < T_1 < T_2 < b_n$. Usando integração por partes, a qual pode ser usada pois pela Proposição 2.4.3 obtemos que p' é uma função com variação total finita, e podemos obter

$$\int_{T_1}^{T_2} \frac{[p'(x)]^2}{p(x)} dx = \int_{T_1}^{T_2} p'(x) d(\log p(x)) = p'(T_2) \log p(T_2) - p'(T_1) \log p(T_1) - \int_{T_1}^{T_2} p''(x) \log p(x) dx.$$

Fazendo $T_1 \rightarrow a_n$ e $T_2 \rightarrow b_n$, no caso $a_n > T$ temos,

$$\int_{a_n}^{b_n} \frac{[p'(x)]^2}{p(x)} dx = - \int_{a_n}^{b_n} p''(x) \log p(x) dx,$$

e no caso $a_n = T$,

$$\int_{a_n}^{b_n} \frac{[p'(x)]^2}{p(x)} dx = -p(T) \log p(T) - \int_{a_n}^{b_n} p''(x) \log p(x) dx.$$

Note que, caso exista $n \in \mathbb{N}$ tal que $a_n = T$, este n é único. Logo,

$$\begin{aligned} \int_G \frac{[p'(x)]^2}{p(x)} dx &= \sum_{n=1}^{+\infty} \int_{a_n}^{b_n} \frac{[p'(x)]^2}{p(x)} dx \leq |p(T) \log p(T)| + \sum_{n=1}^{+\infty} \int_{a_n}^{b_n} |p''(x) \log p(x)| dx \\ &= |p(T) \log p(T)| + \int_G |p''(x) \log p(x)| dx. \end{aligned} \quad (2.31)$$

Só resta limitar adequadamente esta última expressão. Para isto de (2.24) temos

$$|p(T) \log p(T)| \leq I^{\frac{3}{4}} \sqrt{p(T)} |\log p(T)|. \quad (2.32)$$

Por outro lado, usando (2.26) e com a Desigualdade de Cauchy-Schwartz obtemos

$$\begin{aligned} \left(\int_G |p''(x) \log p(x)| dx \right)^2 &= \int_G \frac{|p''(x)|}{\sqrt{p(x)}} \sqrt{p(x)} |\log p(x)| dx \\ &\leq \left(\int_G \frac{[p''(x)]^2}{p(x)} dx \right) \left(\int_G p(x) [\log p(x)]^2 dx \right). \end{aligned}$$

Agora, usando (2.26) obtemos

$$\left(\int_G |p''(x) \log p(x)| dx \right)^2 \leq I^2 \int_G p(x) [\log p(x)]^2 dx. \quad (2.33)$$

Portanto, de (2.31), (2.32) e (2.33) obtemos

$$\int_T^\infty \frac{[p'(x)]^2}{p(x)} dx = \int_G \frac{[p'(x)]^2}{p(x)} dx \leq I^{\frac{3}{4}} \sqrt{p(T)} |\log p(T)| + I \left(\int_T^\infty p(x) [\log p(x)]^2 dx \right)^{\frac{1}{2}}.$$

e (2.30) está provado. □

2.5 Convergência na Informação de Fisher

O objetivo principal deste trabalho é demonstrar o Teorema 3.2.1 de Barron e Johnson [6], a ser apresentado no Capítulo 3, que estabelece condições necessárias e suficientes para que uma sequência de somas estabilizadas que converge em distribuição para uma v.a. estável Z convirja também na distância de Informação de Fisher (Definição 2.2.5) para a mesma v.a. Z .

As relações (2.17) e (2.19), apresentadas na Seção 2.3, a saber

$$\|p\|_{TV} \leq \sqrt{I(p)} \quad \text{e} \quad \sup_x p(x) \leq \sqrt{I(p)},$$

já nos dão indicação da razão do interesse em teoremas de limites na distância de Informação de Fisher. Estas relações e as que apresentaremos na proposição a seguir, mostram que convergência na Informação de Fisher é um resultado mais forte e implica outras formas conhecidas de convergência. Maiores detalhes sobre estas e outras relações da Informação de Fisher e outras formas de convergência podem ser encontradas em [14] (Seção 5.4 e Apêndice E).

Proposição 2.5.1. *Se h é uma densidade log-côncava e simétrica em torno de zero, então, existe uma constante $K = K(h)$ tal que para qualquer densidade p temos*

$$(a) \int_{-\infty}^{\infty} |p(x) - h(x)| dx \leq 2K \sqrt{I(p \parallel h)}.$$

$$(b) \sup_x |p(x) - h(x)| \leq (1 + Kh(0)) \sqrt{I(p \parallel h)}.$$

Capítulo 3

Convergência em Informação de Fisher Relativa e v.a.'s Estáveis

Sejam X_1, X_2, \dots variáveis aleatórias independentes e identicamente distribuídas com função de distribuição comum F , e suponhamos que existam sequências de constantes reais $\{a_n\}_{n \geq 1}$ e $\{b_n\}_{n \geq 1}$, com $b_n > 0$, para as quais a sequência de somas estabilizadas

$$Z_n = \frac{X_1 + X_2 + \dots + X_n}{b_n} - a_n, \quad n \geq 1 \quad (3.1)$$

converge em distribuição para uma v.a. Z com distribuição estável não-degenerada.

Conforme já foi lembrado na Seção 1.4, no Capítulo 1, o Teorema do Limite Central garante que sob a condição $\mathbb{E}X_1^2 < +\infty$, com $VarX_1 > 0$, temos para $a_n = \frac{\mathbb{E}(X_1)}{\sqrt{nVarX_1}}$ e $b_n = \sqrt{nVarX_1}$ que $Z_n \xrightarrow{\mathcal{D}} Z$, com Z tendo distribuição Normal-padrão, que é uma distribuição α -estável com $\alpha = 2$.

Também apresentamos na mesma seção um Teorema do Limite Central Generalizado (Teorema 1.3.11), que apresenta condições suficientes sobre F para que existam constantes $\{a_n\}$ e $\{b_n\}$, apropriadamente escolhidas, para as quais $Z_n \xrightarrow{\mathcal{D}} Z$, onde Z é uma v.a. α -estável, com $0 < \alpha \leq 2$.

Devido às inúmeras aplicações de teoremas de limites desta natureza, uma questão de grande interesse na literatura é saber se esta convergência fraca implica em convergência em um sentido mais forte para a mesma v.a. Z .

Neste trabalho, estamos interessados em estabelecer condições para que a convergência fraca produza a convergência na distância de Informação de Fisher, a qual já observamos no

Capítulo 2 é mais forte que outros tipos de convergência conhecidos como por exemplo, na norma da variação total (Seção 2.5).

Assim, vamos assumir que a f.d. F , das v.a.'s i.i.d. X_1, X_2, \dots , possui densidade p , que também é absolutamente contínua com derivada p' . Vamos indicar por p_n a densidade de Z_n e por ψ a densidade de Z .

Neste sentido, no caso clássico do Teorema do Limite Central, onde $Z_n \xrightarrow{\mathfrak{D}} Z$ com Z tendo distribuição $\mathcal{N}(0, 1)$, Barron e Jonhson [3] provaram que esta convergência fraca implica na convergência na Informação de Fisher Relativa

$$I(Z_n \parallel Z) = I(p_n \parallel \psi) \xrightarrow[n \rightarrow \infty]{} 0$$

se, e somente se, $I(Z_{n_0})$ é finito para algum $n_0 \geq 1$.

Recentemente, Bobkov, Chistyakov e Gotze [7] estabeleceram, sob a hipótese de finitude de $\mathbb{E}X_1^2$, condições equivalentes à $I(Z_{n_0}) < +\infty$ para algum $n_0 \geq 1$, além de estudarem a validade do Teorema do Limite Central clássico por meio das propriedades da Informação de Fisher sobre convolução de densidades. Esse trabalho serviu de base para a obtenção de um resultado semelhante no caso não-normal, ou seja, quando $Z_n \xrightarrow{\mathfrak{D}} Z$ e Z tem distribuição α -estável, com $0 < \alpha < 2$ e $-1 < \beta < 1$, publicado recentemente pelos mesmos autores em [6], e que será apresentado neste capítulo .

Cabe ressaltar, que no livro [14], Oliver Johnson já apresentou a conjectura de um resultado, em certo sentido similar ao obtido por Barron e Johnson [3], (referente à convergência na Informação de Fisher relativa no caso normal), também seria possível para o caso α -estável, com $0 < \alpha < 2$.

Assim, para a apresentação em detalhes do trabalho de Bobkov, Chistyakov e Gotze [6], apresentamos na Seção 3.1 alguns lemas auxiliares, que serão necessários para a demonstração do Teorema principal (Teorema 3.2.2) o qual será apresentado e demonstrado na Seção 3.2 finalizando o capítulo.

3.1 Resultados Auxiliares

Considere X_1, X_2, \dots v.a.'s i.i.d. tendo função de distribuição comum $F(x)$ com densidade $p(x)$ e suponha que existem sequências de números reais $\{a_n\}_{n \geq 1}$ e $\{b_n\}_{n \geq 1}$, com $b_n > 0$, tais que

$$Z_n \xrightarrow{\mathcal{D}} Z,$$

onde $\{Z_n\}_{n \geq 1}$ é a sequência de somas estabilizadas dada por (3.1) e Z uma v.a. α -estável com $0 < \alpha \leq 2$.

Por todo este capítulo vamos utilizar a seguinte notação:

$p_n(x)$: densidade de Z_n .

$\psi(x)$: densidade de Z .

$f_n(t) = \mathbb{E}e^{itZ_n}$: função característica de Z_n .

$g(t) = \mathbb{E}e^{itZ}$: função característica de Z .

$f(t) = \mathbb{E}e^{itX_1}$: função característica de X_1 .

Note que, como X_1, X_2, \dots são v.a.'s i.i.d. com função característica comum $f(t)$, então

$$f_n(t) = e^{ita_n} [f(t/b_n)]^n. \quad (3.2)$$

Como Z é α -estável, pelo Teorema 1.3.8 temos que necessariamente

$$b_n = n^{1/\alpha} h(n), \quad (3.3)$$

onde h é uma função de variação lenta no sentido de Karamata e pelo Teorema 1.3.2 a função característica de Z é da forma

$$g(t) = \exp \{ i\mu t - \sigma^\alpha |t|^\alpha [1 + i\beta \operatorname{sign}(t)w(t, \alpha)] \},$$

onde $0 < \alpha \leq 2$, $-1 \leq \beta \leq 1$, $\mu \in \mathbb{R}$, $\sigma > 0$ e

$$w(t, \alpha) = \begin{cases} \tan\left(\frac{\pi\alpha}{2}\right), & \text{se } \alpha \neq 1 \\ \frac{2}{\pi} \log |t|, & \text{se } \alpha = 1. \end{cases}$$

Assim, denotamos $Z \stackrel{\mathcal{D}}{=} S_\alpha(\sigma, \beta, \mu)$.

Se a densidade p_n é absolutamente contínua com derivada $p'_n(x)$, a Informação de Fisher de Z_n é dada por

$$I(Z_n) = I(p_n) = \int_{-\infty}^{\infty} \frac{[p'_n(x)]^2}{p_n(x)} dx, \quad (3.4)$$

e em qualquer outra situação onde a integral não esteja bem definida, definimos $I(Z_n) = +\infty$. Notemos que, pela Observação 2.2.2 (d), temos

$$I(Z_n) = I(Z_n + a_n).$$

Assim, por facilidade, vamos assumir que $a_n = 0$, pois, caso contrário, basta considerar

$$Z'_n = Z_n + a_n = \frac{X_1 + \cdots + X_n}{b_n}.$$

No caso $\alpha = 2$, ou seja, Z tem distribuição Normal e $\mathbb{E}X_1^2 = \sigma^2 < +\infty$, escolhemos $b_n = \sqrt{n}\sigma$. Neste caso, se $I(Z_1) < +\infty$, então segue da Proposição 2.2.4 (Desigualdade de Stam) e da Observação 2.2.2 (d) que para todo $n > 1$,

$$I(Z_n) = b_n^2 I(X_1 + \cdots + X_n) \leq \frac{b_n^2 I(X_1)}{n} = \sigma^2 I(X_1) = \sigma^2 I(Z_1) < \infty.$$

Ou seja, $\sup_{n>1} I(Z_n) < +\infty$.

Nosso objetivo nesta seção é provar que esta propriedade se estende para Z α -estável não extremal. Ou seja, vamos provar:

Lema 3.1.1. *Suponhamos que $Z_n \xrightarrow{\mathcal{D}} Z$, onde Z segue uma lei estável não extremal. Se $I(Z_{n_0}) < \infty$ para algum n_0 , então $\sup_{n>n_0} I(Z_n) < \infty$.*

Lembrando a Definição 1.3.5: dizemos que Z é α -estável não-extremal se Z tem distribuição Normal ($\alpha = 2$) ou se $0 < \alpha < 2$ e $-1 < \beta < 1$.

Para provar o Lema 3.1.1, necessitamos de vários resultados auxiliares. Por facilidade de argumentação vamos assumir que $n_0 = 1$. Ou seja, $I(X_1) = I(p) < +\infty$.

Pelas observações anteriores, basta analisarmos o caso $0 < \alpha < 2$. Para isto, para cada $n \geq 1$, vamos decompor a densidade $p(x)$, de X_1 , da seguinte forma

$$p(x) = (1 - \delta_n)\hat{p}_n(x) + \delta_n\hat{q}_n(x),$$

onde $\delta_n = \int_{|x|>b_n} p(x)dx$ e $\hat{p}_n(x)$ e $\hat{q}_n(x)$ são as densidades truncadas

$$\hat{p}_n(x) = \frac{b_n}{1 - \delta_n} p(b_n x) 1_{\{|x| \leq 1\}} \quad \text{e} \quad \hat{q}_n(x) = \frac{b_n}{\delta_n} p(b_n x) 1_{\{|x| > 1\}}, \quad (3.5)$$

com respectivas funções características

$$\hat{f}_n(t) = \frac{1}{1 - \delta_n} \int_{-b_n}^{b_n} e^{itx/b_n} p(x) dx \quad \text{e} \quad \hat{g}_n(t) = \frac{1}{\delta_n} \int_{\{|x|>b_n\}} e^{itx/b_n} p(x) dx. \quad (3.6)$$

Note que, pelo Corolário 1.3.13, segue de (1.21) que

$$\delta_n \sim \frac{c}{n}, \quad \text{quando } n \rightarrow \infty, \quad (3.7)$$

para uma certa constante $c > 0$, que depende somente de p .

Agora, como $p_n(x) = p^{n*}(x)$, ou seja, a n -ésima convolução de p , podemos obter da Proposição 1.2.3 e das propriedades de convoluções de densidades a seguinte decomposição binomial

$$p_n = ((1 - \delta_n)\hat{p}_n + \delta_n\hat{q}_n)^{n*} = \sum_{k=0}^n \binom{n}{k} (1 - \delta_n)^k \delta_n^{n-k} \hat{p}_n^{k*} * \hat{q}_n^{(n-k)*}. \quad (3.8)$$

Assim, da propriedade de convexidade do funcional I , demonstrada na Observação 2.2.2 (e), segue que

$$I(Z_n) = I(p_n) \leq \sum_{k=0}^n \binom{n}{k} (1 - \delta_n)^k \delta_n^{n-k} I(\hat{p}_n^{k*} * \hat{q}_n^{(n-k)*}). \quad (3.9)$$

Observe que cada convolução $\hat{p}_n^{k*} * \hat{q}_n^{(n-k)*}$, que aparece na soma, está representando uma densidade de probabilidade com função característica $\hat{f}_n^k(t) \hat{g}_n^{(n-k)}(t)$.

Por (3.9), para provar o Lema 3.1.1, basta obtermos cotas superiores para $I(\hat{p}_n^{k*} * \hat{q}_n^{(n-k)*})$, para cada $0 \leq k \leq n$, em termos de $I(X_1) = I(p)$. Lembrando que, por facilidade de notação, estamos assumindo que $n_0 = 1$, ou seja, $I(X_1) = I(p) < +\infty$.

Faremos isto em vários passos, que serão apresentados por meio de lemas.

No primeiro lema, obtemos cotas superiores para $k \leq n - 3$, com $n \geq 3$.

Lema 3.1.2. *Se $k \leq n - 3$, então*

$$I(\hat{p}_n^{k*} * \hat{q}_n^{(n-k)*}) \leq C(nb_n)^2 I(p), \quad (3.10)$$

onde a constante C depende unicamente de p .

Demonstração. Por (2.11) segue que

$$I(\hat{p}_n^{k*} * \hat{q}_n^{(n-k)*}) \leq I(\hat{q}_n^{(n-k)*}). \quad (3.11)$$

Por outro lado, da Proposição 2.3.3 se $n - k \geq 3$, como as densidades \hat{q}_n são de variação limitada, segue que

$$I(\hat{q}_n^{(n-k)*}) \leq \frac{1}{2} \left(\left\| \hat{q}_n^{[(n-k)/3]*} \right\|_{TV}^2 + 2 \left\| \hat{q}_n^{[(n-k)/3]*} \right\|_{TV} \left\| \hat{q}_n^{n-k-2[(n-k)/3]*} \right\|_{TV} \right).$$

Mas, das propriedades de variação total temos que

$$\left\| \hat{q}_n^{s*} \right\|_{TV} \leq \left\| \hat{q}_n \right\|_{TV} \quad \text{para } s = 1, 2, 3, 4, \dots$$

e então

$$I(\hat{q}_n^{(n-k)*}) \leq \frac{3}{2} \left\| \hat{q}_n \right\|_{TV}^2. \quad (3.12)$$

Da definição de \hat{q}_n em (3.5) e como $\|p\|_{TV} \leq \sqrt{I(p)}$, (2.3.1) temos

$$\left\| \hat{q}_n \right\|_{TV} = \frac{b_n}{\delta_n} \left\| p \mathbf{1}_{\{|x| > b_n\}} \right\|_{TV} \leq \frac{b_n}{\delta_n} \|p\|_{TV} \leq \frac{b_n}{\delta_n} \sqrt{I(p)}. \quad (3.13)$$

Logo, como $\delta_n \sim \frac{c}{n}$, com c dependendo de p , segue de (3.11) a (3.13) que

$$I(\hat{p}_n^{k*} * \hat{q}_n^{(n-k)*}) \leq I(\hat{q}_n^{(n-k)*}) \leq \frac{3}{2} \left\| \hat{q}_n \right\|_{TV}^2 \leq \frac{3}{2} \left(\sqrt{I(p)} \frac{b_n}{\delta_n} \right)^2 \leq \frac{3}{2} \hat{C} (b_n n)^2 I(p),$$

onde \hat{C} depende de p . Assim fazendo $C = \frac{3}{2} \hat{C}$, obtemos o resultado. \square

Para as outras parcelas de (3.9), quando $n - k < 3$, vamos novamente utilizar (2.11) para obter

$$I(\hat{p}_n^{k*} * \hat{q}_n^{(n-k)*}) \leq I(\hat{p}_n^{k*}). \quad (3.14)$$

Assim, basta encontrarmos uma cota para $I(\hat{p}_n^{k*})$.

Por facilidade, consideremos a densidade centrada

$$r_n(x) = \hat{p}_n(x + d_n), \quad (3.15)$$

onde

$$d_n = \int_{-1}^1 x \hat{p}_n(x) dx = \frac{1}{b_n(1 - \delta_n)} \int_{-b_n}^{b_n} x p(x) dx. \quad (3.16)$$

Note que $r_n(x)$ é a densidade de uma v.a. $Y - d_n$, onde Y tem densidade $\hat{p}_n(x)$. Assim, será suficiente mostrar que

$$I(\hat{p}_n^{k*}) = I(r_n^{k*}) \leq C, \quad (3.17)$$

para alguma constante C dependendo somente da densidade p .

Provaremos (3.17) utilizando cotas superiores para a função característica associada a $r_n(x)$ e de suas derivadas.

Assim, denotemos $\varphi_n(t)$ a função característica associada à $r_n(x)$, ou seja, $\varphi_n(t) = \mathbb{E}e^{it(Y-d_n)}$, onde Y tem densidade $\hat{p}_n(x)$. Então,

$$\varphi_n(t) = e^{-itd_n} \hat{f}_n(t), \quad (3.18)$$

onde \hat{f}_n é a função característica associada a $\hat{p}_n(x)$.

No lema a seguir, obtemos uma cota para $\varphi'_n(t)$ dependendo somente de p .

Lema 3.1.3. *Para todo $t \in \mathbb{R}$,*

$$|\varphi'_n(t)| \leq \frac{C}{n}|t|, \quad (3.19)$$

para alguma constante $C > 0$ dependendo unicamente de p .

Demonstração.

Por (3.18) segue que $\varphi'_n(t) = e^{-itd_n}(\hat{f}'_n(t) - id_n\hat{f}_n(t))$ e como $\varphi'_n(0) = \mathbb{E}(Y - d_n) = 0$, podemos escrever

$$\begin{aligned} \varphi'_n(t) &= \varphi'_n(t) - \varphi'_n(0) = \int_{-\infty}^{\infty} i(x - d_n) (e^{it(x-d_n)} - 1) \hat{p}_n(x) dx \\ &= \frac{ib_n}{1 - \delta_n} \int_{-1}^1 (x - d_n) (e^{it(x-d_n)} - 1) p(b_n x) dx. \end{aligned}$$

onde a última igualdade segue da definição de \hat{p}_n em (3.5).

Agora, fazendo uma mudança de variáveis obtemos

$$\varphi'_n(t) = \frac{1}{1 - \delta_n} \int_{-b_n}^{b_n} \left(\frac{x}{b_n} - d_n \right) (e^{it(x/b_n - d_n)} - 1) dF(x).$$

onde F é a função de distribuição de X_1 , que tem densidade $p(x)$.

Usando a conhecida desigualdade $|e^{is} - 1| \leq |s|$, $\forall s \in \mathbb{R}$ e em seguida a desigualdade $(a - b)^2 \leq 2a^2 + 2b^2$, obtemos

$$\begin{aligned}
|\varphi'_n(t)| &\leq \frac{1}{1-\delta_n} \int_{-b_n}^{b_n} \left| \frac{x}{b_n} - d_n \right| \left| t \left(\frac{x}{b_n} - d_n \right) \right| dF(x) \\
&= \frac{|t|}{1-\delta_n} \int_{-b_n}^{b_n} \left(\frac{x}{b_n} - d_n \right)^2 dF(x) \\
&\leq \frac{|t|}{1-\delta_n} \int_{-b_n}^{b_n} 2 \left(\frac{x}{b_n} \right)^2 dF(x) + \frac{|t|}{1-\delta_n} \int_{-b_n}^{b_n} 2d_n^2 dF(x) \\
&\leq \frac{2|t|}{b_n^2(1-\delta_n)} \int_{-b_n}^{b_n} x^2 dF(x) + \frac{2|t|}{1-\delta_n} d_n^2.
\end{aligned}$$

Então, usando (3.16), podemos escrever

$$|\varphi'_n(t)| \leq \frac{2|t|}{b_n^2(1-\delta_n)} \left[\int_{-b_n}^{b_n} x^2 dF(x) + \frac{1}{(1-\delta_n)^2} \left(\int_{-b_n}^{b_n} |x| dF(x) \right)^2 \right]. \quad (3.20)$$

Integrando por partes com $u = x^2$ e $dv = dF(x)$, obtemos,

$$\begin{aligned}
\int_{-b_n}^{b_n} x^2 dF(x) &= -b_n^2 (1 - F(b_n) + F(-b_n)) + 2 \int_0^{b_n} x (1 - F(x) + F(-x)) dx \\
&\leq 2 \int_0^{b_n} x (1 - F(x) + F(-x)) dx.
\end{aligned} \quad (3.21)$$

Por outro lado, integrando também por partes, tomando $u = x$, $dv = dF(x)$,

$$\begin{aligned}
\int_{-b_n}^{b_n} |x| dF(x) &= -b_n (1 - F(b_n) + F(-b_n)) + \int_0^{b_n} (1 - F(x) + F(-x)) dx \\
&\leq \int_0^{b_n} (1 - F(x) + F(-x)) dx.
\end{aligned} \quad (3.22)$$

Agora, como $\delta_n \sim \frac{c}{n}$, com c dependendo somente de p , temos que $\frac{1}{1-\delta_n} \xrightarrow{n \rightarrow \infty} 1$ e assim é limitada por uma constante que depende somente de p . Então de (3.20), (3.21) e (3.22) segue

$$\begin{aligned}
|\varphi'_n(t)| &\leq \frac{C|t|}{b_n^2} \int_0^{b_n} x (1 - F(x) + F(-x)) dx \\
&\quad + \frac{C^3|t|}{b_n^2} \left(\int_0^{b_n} (1 - F(x) + F(-x)) dx \right)^2,
\end{aligned} \quad (3.23)$$

para alguma constante $C > 0$ dependendo unicamente de p .

Assim, para obter o resultado desejado devemos limitar as duas integrais em (3.23). Para isto, lembremos que pelo Corolário 1.3.13 temos $|f(t)| = \exp\{-c|t|^\alpha h(1/|t|)\}$ onde $c > 0$ e $h(x)$ é uma função de variação lenta quando $x \rightarrow \infty$ e pelo Teorema 1.3.10,

$$F(x) = \frac{(c_0 + o(1))}{(-x)^\alpha} h(-x) \quad x < 0 \quad \text{e} \quad F(x) = 1 - \frac{(c_1 + o(1))}{x^\alpha} h(x) \quad x > 0. \quad (3.24)$$

Por um lado, considerando a primeira integral em (3.23), vamos provar que para alguma constante $C_1 > 0$,

$$\frac{1}{b_n^2} \int_0^{b_n} x [1 - F(x) + F(-x)] dx \leq \frac{C_1}{n}. \quad (3.25)$$

De fato, podemos escrever para a função $h(x)$

$$\int_0^{b_n} \frac{h(x)}{x^{\alpha-1}} dx = b_n^{2-\alpha} h(b_n) \int_0^1 \frac{h(sb_n)}{h(b_n)} \frac{ds}{s^{\alpha-1}}.$$

Agora, como h é de variação lenta e $b_n = n^{1/\alpha} h(n)$, segue do Corolário 1.1.5 que $b_n \rightarrow \infty$ e daí, pelo Teorema 1.1.3, temos $\frac{h(b_n s)}{h(b_n)} \xrightarrow{n \rightarrow \infty} 1$ uniformemente para todo $s \in [0, 1]$. Logo

$$\int_0^1 \frac{h(sb_n)}{h(b_n)} \frac{ds}{s^{\alpha-1}} \xrightarrow{n \rightarrow \infty} \int_0^1 \frac{ds}{s^{\alpha-1}} = \frac{1}{2-\alpha}.$$

Assim, existe $M > 0$ tal que

$$\left| \int_0^1 \frac{h(sb_n)}{h(b_n)} \frac{ds}{s^{\alpha-1}} \right| \leq M. \quad (3.26)$$

Agora, usando (3.24) e fazendo a mudança de variáveis $x = sb_n$, obtemos

$$\begin{aligned} \frac{1}{b_n^2} \int_0^{b_n} x (1 - F(x) + F(-x)) dx &= \frac{1}{b_n^2} \int_0^{b_n} (c_1 + c_0 + o(1)) \frac{h(x)}{x^{\alpha-1}} dx \\ &= \frac{1}{b_n^2} (c_1 + c_0 + o(1)) b_n^{2-\alpha} h(b_n) \int_0^1 \frac{h(sb_n)}{h(b_n)} \frac{ds}{s^{\alpha-1}}. \end{aligned}$$

Logo, por (3.26) e como $h(b_n) \sim \frac{b_n^\alpha}{n}$ (pelo Teorema 1.3.12) e obtemos (3.25) como queríamos. Por outro lado, para obtermos uma cota superior para a segunda integral em (3.23), vamos considerar três casos separadamente.

Caso $0 < \alpha < 1$. Utilizando (3.24), temos

$$\frac{1}{b_n^2} \left(\int_0^{b_n} (1 - F(x) + F(-x)) dx \right)^2 \leq \frac{1}{b_n^2} (c_1 + c_0 + o(1))^2 \left[\int_0^{b_n} \frac{h(x)}{x^\alpha} dx \right]^2. \quad (3.27)$$

Agora, usando os mesmos argumentos utilizados anteriormente para a prova de (3.25), obtemos que

$$\int_0^{b_n} \frac{h(x)}{x^\alpha} dx = b_n^{1-\alpha} h(b_n) \int_0^1 \frac{h(sb_n)}{h(b_n) s^\alpha} ds.$$

Mas, quando $n \rightarrow \infty$,

$$\int_0^1 \frac{h(sb_n)}{h(b_n) s^\alpha} ds \rightarrow \int_0^1 \frac{ds}{s^\alpha} = \frac{1}{1-\alpha},$$

e como $\lim_{n \rightarrow \infty} \frac{nh(b_n)}{b_n^\alpha} = 1$, segue que

$$\int_0^{b_n} \frac{h(x)}{x^\alpha} dx = b_n^{1-\alpha} h(b_n) \int_0^1 \frac{h(sb_n)}{h(b_n) s^\alpha} ds \sim \frac{b_n}{(1-\alpha)n}, \quad \text{quando } n \rightarrow \infty.$$

Logo, existe $C_2 > 0$ dependendo de p e α , tal que

$$\left| \int_0^{b_n} \frac{h(x)}{x^\alpha} dx \right| \leq C_2 \frac{b_n}{(1-\alpha)n},$$

e de (3.27) segue que

$$\frac{1}{b_n^2} \left(\int_0^{b_n} (1 - F(x) + F(-x)) dx \right)^2 \leq \frac{C_2}{n^2},$$

Caso $\alpha = 1$. Com h é de variação lenta, pelo Corolário 1.1.5, segue que para todo $\epsilon > 0$ dado, $x^{-\epsilon}h(x) \rightarrow 0$ quando $x \rightarrow \infty$.

Então, dado $\epsilon = \frac{1}{4}$, existe uma constante $C_1 > 0$, dependendo somente de $\epsilon = \frac{1}{4}$, tal que $h(x) \leq C_1 x^{1/4}$, para todo $x \geq 1$. Logo,

$$\int_1^{b_n} \frac{h(x)}{x} dx \leq C_1 (b_n^{1/4} - 1) \leq C_1 b_n^{1/4}. \quad (3.28)$$

Agora como $1 - F(x) + F(-x) \leq 1$, usando (3.24) e (3.28), obtemos

$$\begin{aligned} \frac{1}{b_n^2} \left(\int_0^{b_n} [1 - F(x) + F(-x)] dx \right)^2 &\leq \frac{1}{b_n^2} \left(1 + \int_1^{b_n} [1 - F(x) + F(-x)] dx \right)^2 \\ &\leq \frac{1}{b_n^2} \left(1 + (c_0 + c_1 + o(1)) \int_1^{b_n} \frac{h(x)}{x} dx \right)^2 \\ &\leq \frac{1}{b_n^2} (1 + C_1 b_n^{1/4})^2. \end{aligned}$$

Agora, como $b_n \rightarrow +\infty$, segue que $(1 + C_1 b_n^{1/4}) \sim C_1 b_n^{1/4}$ quando $n \rightarrow \infty$, e daí segue

$$\frac{1}{b_n^2} \left(\int_0^{b_n} [1 - F(x) + F(-x)] dx \right)^2 \leq \frac{\tilde{C}_1}{b_n^{3/2}},$$

onde \tilde{C}_1 depende unicamente de p .

Por outro lado, como $b_n = nh(n)$ (pois $\alpha = 1$), temos que $b_n^{3/2} = n^{3/2} \tilde{h}(n)$, onde $\tilde{h}(n) = [h(n)]^{3/2}$ também é de variação lenta. Então $\frac{b_n^{3/2}}{n} = n^{1/2} \tilde{h}(n) \xrightarrow{n \rightarrow \infty} +\infty$ pelo Corolário 1.1.5.

Assim, segue que $\frac{\tilde{C}_1}{b_n^{3/2}} = O\left(\frac{1}{n}\right)$ e daí,

$$\frac{1}{b_n^2} \left(\int_0^{b_n} [1 - F(x) + F(-x)] dx \right)^2 \leq C_3 \frac{1}{n},$$

onde C_3 é uma constante que depende somente de p .

Caso $1 < \alpha < 2$. Neste situação, temos

$$\int_0^{+\infty} (1 - F(x) + F(-x)) dx = \int_0^{\infty} P(|X_1| > x) dx = \mathbb{E}|X_1| < +\infty.$$

Agora, $b_n^2 = n^{2/\alpha} \tilde{h}(n)$, com $\tilde{h}(n) = (h(n))^2$ de variação lenta e daí $\frac{b_n^2}{n} = n^{(2/\alpha)-1} \tilde{h}(n) \rightarrow 0$ pelo Corolário 1.1.5, pois $\frac{2}{\alpha} - 1 > 0$. Então segue

$$\frac{1}{b_n^2} \left(\int_0^{b_n} (1 - F(x) + F(-x)) dx \right)^2 \leq \frac{C}{n},$$

com C uma constante dependendo somente de p .

Portanto, em todos os três casos considerados, podemos concluir que

$$\frac{1}{b_n^2} \left(\int_0^{b_n} [1 - F(x) + F(-x)] dx \right)^2 \leq \frac{C}{n}, \quad (3.29)$$

para alguma constante C que depende somente de p .

Finalmente, usando (3.25) e (3.29) em (3.23) obtemos (3.19). \square

No lema a seguir vamos obter uma cota superior para a função $[\varphi_n(t)]^k$ análoga à cota (1.28) obtida para a derivada da função característica de Z_n .

Lema 3.1.4. *Sejam $\delta \in (0, \alpha)$ e $\eta \in (0, 1)$ fixos. Então, existem constantes positivas ϵ, c, C que dependem de p, δ, η com a seguinte propriedade: se $k > \eta n$, então*

$$|\varphi_n(t)|^k = |\hat{f}_n(t)|^k \leq C e^{-c|t|^\delta} \text{ sempre que } |t| \leq \epsilon b_n, \quad (3.30)$$

onde $\hat{f}_n(t)$ é a função característica associada à $\hat{p}_n(x)$, dada em (3.6).

Demonstração. Por conveniência, vamos assumir $|t| \geq 1$. Primeiramente notemos que para todo $t \in \mathbb{R}$,

$$\hat{f}_n(t) = \frac{1}{1 - \delta_n} (f_1(t/b_n) - \delta_n \hat{g}_n(t)), \quad (3.31)$$

onde $\hat{g}_n(t)$ é a função característica associada à $\hat{q}_n(t)$, dada em (3.6).

Por outro lado, do Corolário 1.3.13 temos

$$|f(t)| = \exp \{ -c|t|^\alpha h(1/|t|) \}, \quad (3.32)$$

com $c > 0$ e $h(x)$ é uma função de variação lenta quando $x \rightarrow \infty$. Sendo assim, pelo Teorema de Representação de Karamata (Teorema 1.1.4), podemos representar $h(x)$ como

$$h(x) = c(x) \exp \left\{ \int_{x_0}^x \frac{w(y)}{y} dy \right\}, \quad (3.33)$$

onde $x_0 > 0$, $c(x) \rightarrow c > 0$ e $w(x) \rightarrow 0$ quando $x \rightarrow \infty$.

Como $\frac{c(b_n/|t|)}{c(b_n)} \rightarrow 1$ quando $n \rightarrow \infty$, existe $c_0 > 0$ tal que $\frac{c(b_n/|t|)}{c(b_n)} \geq c_0$.

Agora, para $\epsilon \in (0, 1]$ fixo, para $x_0 = \min_{n \leq 1} \{b_n : 1 \leq |t| \leq \epsilon b_n\}$, de (3.33) segue

$$\begin{aligned} \frac{h(b_n/|t|)}{h(b_n)} &= \frac{c(b_n/|t|)}{c(b_n)} \exp \left\{ - \int_{b_n}^{b_n/|t|} \frac{w(y)}{y} dy \right\} \\ &\geq c_0 \exp \left\{ - \int_{1/\epsilon}^{b_n/|t|} \frac{w(y)}{y} dy \right\}. \end{aligned}$$

Como $w(x) \rightarrow 0$, quando $x \rightarrow \infty$, podemos tomar $\gamma = \gamma(\epsilon) = \sup_{y \geq 1/\epsilon} |w(y)|$ e obter

$$\frac{h(b_n/|t|)}{h(b_n)} \geq c_0 \exp \left\{ -\gamma \int_{1/\epsilon}^{b_n/|t|} \frac{1}{y} dy \right\} \geq c_0 |t|^{-\gamma}. \quad (3.34)$$

Daí, usando $\lim_{n \rightarrow \infty} \frac{nh(n)}{b_n^\alpha} = 1$ e (3.34) e (3.32) obtemos

$$|f(t)| \leq \exp \left\{ -c_1 |t|^{\alpha-\gamma}/n \right\},$$

para alguma constante $c_1 > 0$.

Agora, vamos escolher $\epsilon > 0$ de modo que $\gamma < \alpha - \delta$. Então usando a desigualdade anterior em (3.31), para $1 \leq |t| \leq \epsilon b_n$, temos

$$\begin{aligned} |\hat{f}_n(t)| &= \frac{1}{1 - \delta_n} |f_1(t/b_n) - \delta_n \hat{g}_n(t)| \\ &\leq \frac{1}{1 - \delta_n} (|f_1(t/b_n)| + \delta_n) \\ &\leq \frac{1}{1 - \delta_n} \left(\exp \left\{ \frac{-c_1 |t|^{\alpha-\gamma}}{n} \right\} + \delta_n \right) \\ &\leq \exp \left\{ \frac{-c_1 |t|^{\alpha-\gamma}}{n} \right\} + \delta_n. \end{aligned} \quad (3.35)$$

Notemos que para $1 \leq |t| \leq \epsilon b_n$, como $\alpha - \gamma > \delta > 0$, temos

$$\frac{c_1 |t|^{\alpha-\gamma}}{n} \leq \frac{c_1 (\epsilon b_n)^{\alpha-\gamma}}{n}.$$

Mas como $b_n = n^{1/\alpha}h(n)$, então $\frac{b_n^{\alpha-\gamma}}{n} = n^{-\alpha/\gamma}\tilde{h}(n) \xrightarrow[n \rightarrow \infty]{} 0$, pelo Corolário 1.1.5, pois $\tilde{h} = (h(n))^{\alpha-\gamma}$ é de variação lenta e $\frac{\gamma}{\alpha} > 0$.

Então, existe $K > 0$ constante tal que:

$$\frac{c_1|t|^{\alpha-\gamma}}{n} \leq \frac{c_1|b_n|^{\alpha-\gamma}}{n} < K.$$

Agora, usando

$$\log x \leq x - 1 \text{ para } x > 0, \quad e^{-x} \leq 1 - \frac{1}{K}(1 - e^{-K})x \text{ para } 0 \leq x \leq K,$$

e como $\frac{c_1|t|^{\alpha-\gamma}}{n} < K$, segue que

$$\begin{aligned} \log \left(\exp \left\{ \frac{-c_1|t|^{\alpha-\gamma}}{n} \right\} + \delta_n \right) &\leq \exp \left\{ \frac{-c_1|t|^{\alpha-\gamma}}{n} \right\} + \delta_n - 1 \\ &\leq -\frac{1}{K} (1 - e^{-K}) \frac{c_1|t|^{\alpha-\gamma}}{n} + \delta_n. \end{aligned}$$

Daí, como $\delta_n \sim \frac{c}{n}$, com $c > 0$ constante segue que

$$\log \left(\exp \left\{ \frac{-c_1|t|^{\alpha-\gamma}}{n} \right\} + \delta_n \right) \leq \frac{c_2}{n} - \frac{c_3|t|^{\alpha-\gamma}}{n},$$

com c_3 constante positiva independente de n . Sendo assim, de (3.35) obtemos

$$|\hat{f}_n(t)| \leq \exp \left\{ \frac{1}{n} (c_2 - c_3|t|^{\alpha-\gamma}) \right\}.$$

Como esta desigualdade é também válida para qualquer potência k , temos

$$\begin{aligned} |\hat{f}_n(t)|^k &\leq \exp \left\{ \frac{k}{n} (c_2 - c_3|t|^{\alpha-\gamma}) \right\} \\ &= C \exp \left\{ -\frac{k}{n} c_3|t|^{\alpha-\gamma} \right\}. \end{aligned}$$

Portanto, para $k > \eta n$ segue

$$|\hat{f}_n(t)|^k \leq C \exp \left\{ -\eta c_3|t|^\delta \right\}$$

e (3.30) está provada com $c = \eta c_3$. □

Notemos que nas demonstrações dos lemas anteriores não usamos o fato que $I(p)$ é finito. Sendo assim, os resultados são válidos para quaisquer distribuições que estão no domínio de atração dessas leis estáveis.

A seguir, vamos apresentar uma aplicação dos dois lemas anteriores, onde incluímos a hipótese $I(p) < +\infty$ que é de nosso interesse para provarmos (3.17) no próximo lema.

Primeiramente lembremos que se $I(p)$ é finito, a densidade p tem variação limitada e se anula no infinito. Assim, da Proposição 2.3.1 temos $\|p\|_{TV} \leq \sqrt{I(p)}$ e daí podemos obter

$$\|r_n\|_{TV} = \|\hat{p}_n\|_{TV} = \frac{b_n}{1 - \delta_n} \|p1_{\{|x| \leq b_n\}}\|_{TV} \leq \frac{b_n}{1 - \delta_n} \|p\|_{TV} \leq \frac{b_n}{1 - \delta_n} \sqrt{I(p)}. \quad (3.36)$$

Por outro lado, de (2.18) temos $|f(t)| \leq \frac{\|p\|_{TV}}{|t|}$, ($t \neq 0$) onde f é a função característica de p . Então, usando esta desigualdade para as funções características \hat{f}_n e φ_n , associadas às densidades de \hat{p}_n e r_n respectivamente, obtemos de (3.36)

$$|\varphi_n(t)| = |\hat{f}_n(t)| \leq \frac{\|\hat{p}_n\|_{TV}}{|t|} \leq \frac{1}{1 - \delta_n} \frac{\sqrt{I(p)}}{|t|}, \quad (t \neq 0),$$

e como $\frac{\delta_n}{n} \sim c$, com c dependendo somente de p , segue que

$$|\varphi_n(t)| \leq C \frac{b_n}{|t|}, \quad (t \neq 0) \quad (3.37)$$

para alguma constante C dependendo somente de p .

Corolário 3.1.5. *Suponhamos que $I(p) < \infty$. Sejam $\delta \in (0, 1)$ e $\eta \in (0, 1)$ fixos. Então para $k \geq 4$ existe uma constante C dependendo de p , δ e η tal que*

$$\int_{-\infty}^{\infty} t^2 |(\varphi_n^k)'(t)| dt \leq C. \quad (3.38)$$

$$\int_{-\infty}^{\infty} (1 + |t|) |\varphi_n(t)|^k dt \leq C. \quad (3.39)$$

Demonstração. Como $(\varphi_n^k)'(t) = k\varphi_n^{k-1}\varphi'(t)$, temos de (3.19), no Lema 3.1.3, que

$$\int_{-\infty}^{\infty} t^2 |(\varphi_n^k)'(t)|^2 dt = k^2 \int_{-\infty}^{\infty} t^2 |(\varphi_n)'(t)|^2 |\varphi_n(t)|^{2(k-1)} dt \leq \frac{C^2}{n^2} \int_{-\infty}^{\infty} t^4 |\varphi_n(t)|^{2(k-1)} dt. \quad (3.40)$$

Basta limitar a última integral por alguma constante \bar{C} . Para isso, se $|t| \leq \epsilon b_n$, de (3.30), no Lema 3.1.4, segue

$$\int_{|t| \leq \epsilon b_n} t^4 |\varphi_n(t)|^{2(k-1)} dt \leq \int_{|t| \leq \epsilon b_n} t^4 \hat{C} e^{-c|t|^\delta} dt \leq \int_{|t| \leq \epsilon b_n} \hat{C} (\epsilon b_n)^4 e^{-c|t|^\delta} dt \leq \hat{C} (\epsilon b_n)^4 \int_{|t| \leq \epsilon b_n} dt \leq \bar{C}. \quad (3.41)$$

No caso $|t| > \epsilon b_n$, observemos que

$$\hat{f}_n(b_n t) = \frac{1}{1 - \delta_n} \int_{-b_n}^{b_n} e^{itx} p(x) dx.$$

Agora, usando cotas de separação gerais para funções características discutidas em [4], podemos obter

$$\sup_{|t| \geq \epsilon} |\psi_n(b_n t)| = \sup_{|t| \geq \epsilon} |\hat{f}_n(b_n t)| \leq e^{-c}, \quad (3.42)$$

Assim, fazendo a mudança de variáveis $t = b_n y$ e usando (3.1) e (3.42), temos para $k \geq 4$

$$\begin{aligned} \int_{|t| \geq \epsilon b_n} t^4 |\varphi_n(t)|^{2(k-1)} dt &= b_n^5 \int_{|y| \geq \epsilon} y^4 |\varphi_n(b_n y)|^6 |\varphi_n(b_n y)|^{2(k-4)} dy \\ &\leq b_n^5 \int_{|y| \geq \epsilon} y^4 |\varphi_n(b_n y)|^6 e^{-2c(k-4)} dy \\ &\leq b_n^5 e^{-2c(k-4)} \int_{|y| \geq \epsilon} \frac{c}{y^2} dy \\ &\leq \hat{C} b_n^5 e^{-2ck}, \quad \text{com } \hat{C} \in \mathbb{R}. \end{aligned} \quad (3.43)$$

Assim, usando (3.41) e (3.43) em (3.40) obtemos

$$\int_{-\infty}^{\infty} t^2 |(\varphi_n^k)'(t)|^2 dt \leq \frac{C^2}{n^2} \left(\bar{C} + \hat{C} b_n^5 e^{-2ck} \right),$$

e como $\frac{1}{n^2} \rightarrow 0$ e $\frac{b_n^5}{n^2} = n^{(5/2-2)} \tilde{h}(n) \xrightarrow{n \rightarrow \infty} 0$, pois $\tilde{h}(n) = [h(n)]^5$ é de variação lenta e $\frac{5}{2} - 2 > 0$, segue (3.38).

Usando o mesmo raciocínio, podemos provar (3.39). □

Agora, usando as cotas obtidas no Corolário 3.1.5, vamos obter cotas para as parcelas $I\left(\hat{p}_n^{k*} * \hat{q}_n^{(n-k)*}\right)$ em (3.9), para $k > n - 3$, não incluídas no Lema 3.1.2.

Para isso, vamos fazer alguns cálculos preliminares que simplificarão a prova do lema a seguir. Recordemos que $r_n(x) = \hat{p}_n(x + d_n)$ é uma densidade com função característica φ_n e que $[\varphi_n]^k$ é função característica da densidade r_n^{k*} . Sendo assim, podemos usar as fórmulas de inversão para $k \geq 4$, pois as cotas obtidas no Corolário 3.1.5 garantem as condições de integrabilidade que precisamos.

Para $x \in \mathbb{R}$ podemos escrever,

$$r_n^{k*}(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-itx} [\varphi_n(t)]^k dt, \quad (3.44)$$

e derivando obtemos

$$(r_n^{k*})'(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-itx} (-it) [\varphi_n(t)]^k dt. \quad (3.45)$$

Note que pelo Corolário 3.1.5, para $k \geq 4$,

$$\int_{-\infty}^{\infty} |\varphi_n(t)|^k dt \leq \int_{-\infty}^{\infty} (1 + |t|) |\varphi_n(t)|^k dt \leq C$$

e

$$\int_{-\infty}^{\infty} |t| |\varphi_n(t)|^k dt \leq \int_{-\infty}^{\infty} (1 + |t|) |\varphi_n(t)|^k dt \leq C, \quad (3.46)$$

logo, as igualdades (3.44) e (3.45) estão bem definidas.

Por outro lado, podemos escrever

$$r_n^{k*}(x) + x (r_n^{k*})'(x) = -\frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-itx} t [(\varphi_n(t))^k]' dt, \quad (3.47)$$

pois, integrando por partes, fazendo $u = e^{itx} t$ e $dv = [(\varphi_n(t))^k]' dt$ e usando (3.44) e (3.45), temos

$$\begin{aligned} -\frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-itx} t [(\varphi_n(t))^k]' dt &= x \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-itx} (-it) [\varphi_n(t)]^k dt + \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-itx} [\varphi_n(t)]^k dt \\ &= r_n^{k*}(x) + x (r_n^{k*})'(x). \end{aligned}$$

Note que, pelo Corolário 3.1.5, (3.47) está bem definida.

Logo, segue de (3.47) e de (3.44) que

$$x (r_n^{k*})' (x) = -\frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-itx} (\varphi_n(t)^k + tk\varphi_n(t)^{k-1}\varphi_n'(t)) dt.$$

Isto mostra que a função $x (r_n^{k*})' (x) \in L_{\mathbb{R}}^2$, pois é a Transformada de Fourier da função $\varphi_n(t)^k + tk\varphi_n(t)^{k-1}\varphi_n'(t) \in L_{\mathbb{R}}^2 \cap L_{\mathbb{R}}^1$. Portanto, podemos escrever

$$| (r_n^{k*})' (x) | \leq \frac{u_{nk}(x)}{|x|}$$

onde $u_{nk}(x) \in L_{\mathbb{R}}^2$, com $\|u_{nk}\|_2^2 = \int_{-\infty}^{\infty} (u_{nk})^2 dx \leq C$.

Além disso, por (3.46) temos para uma constante $C > 0$, independente de k e n , que $\|(-it)(\psi_n(t)^k)\|_1 \leq C$ e de (3.44) temos

$$\sup_x |(r_n^{k*})(x)| \leq C \text{ para todo } n.$$

Logo, após algumas manipulações podemos concluir que

$$| (r_n^{k*})' (x) | \leq \frac{v_{nk}(x)}{1 + |x|}, \text{ para alguma função } v_{nk}(x) \in L^2. \quad (3.48)$$

Assim, estamos prontos para provar o seguinte resultado.

Lema 3.1.6. *Se $15 \leq \eta n \leq k \leq n$, então*

$$I(\hat{p}_n^{k*} * \hat{q}_n^{(n-k)*}) \leq C, \quad (3.49)$$

onde C depende unicamente de p e η .

Demonstração. Lembrando (3.14), temos

$$I(\hat{p}_n^{k*} * \hat{q}_n^{(n-k)*}) \leq I(\hat{p}_n^{k*}),$$

e pela definição de r_n em (3.15), é suficiente provar (3.17), ou seja,

$$I(\hat{p}_n^{k*}) = I(r_n^{k*}) \leq C.$$

Assumamos primeiramente que $\eta_0 n \leq k \leq n$, onde $\eta_0 \leq \eta$. Notemos que, usando (3.36) como $\delta_n \sim \frac{c}{n}$, então para alguma constante $C > 0$,

$$\|r_n\|_{TV} \leq C b_n \sqrt{I(p)} < +\infty,$$

ou seja, a variação total de r_n é finita. Logo, usando os mesmos argumentos para obter (3.12) na prova do Lema 3.1.2, podemos concluir para $k \geq 3$,

$$I(r_n^{k*}) \leq \frac{3}{2} \|r_n\|_{TV}^2 < +\infty.$$

Por outro lado, de (3.48), usando a Desigualdade de Cauchy-Schwartz, obtemos

$$\|r_n^{k*}\|_{TV} \leq \int_{-\infty}^{\infty} |(r_n^{k*})'(x)| dt \leq \left(\int_{-\infty}^{\infty} u_{nk}^2(x) dx \right)^{1/2} \left(\int_{-\infty}^{\infty} \frac{dx}{1+|x|} \right)^{1/2} = \bar{C} \|u_{nk}\|_2 = C,$$

onde a constante C pode depender de p e η_0 .

Agora, novamente aplicando a Proposição 2.3.3 para convoluções de quaisquer três densidades r_n^{k*} , isto é, se $\eta_0 n \leq k_j \leq n$ e $k_j \geq 5$, $j = 1, 2, 3$ obtemos que,

$$\begin{aligned} I(r_n^{(k_1+k_2+k_3)*}) &\leq \frac{1}{2} \left(\|r_n^{[(k_1+k_2+k_3)/3]*}\|_{TV}^2 + 2 \| |r_n^{[(k_1+k_2+k_3)/3]*} \|_{TV} \| |r_n^{(k_1+k_2+k_3)-2[(k_1+k_2+k_3)/3]*} \|_{TV} \right) \\ &\leq \frac{3}{2} \|r_n^{k*}\|_{TV} \\ &\leq \frac{3}{2} C^2. \end{aligned} \tag{3.50}$$

Assim, se $k \geq 15$, basta escolher $k_1 = k_2 = \lceil \frac{k}{3} \rceil$, $k_3 = k - (k_1 + k_2)$ e teríamos $k_j \geq 5$, $j = 1, 2, 3$, $k = k_1 + k_2 + k_3$ e o resultado segue de (3.50).

Da mesma forma, se $k \geq \eta n$, temos $k_j \geq \lceil \frac{\eta n}{3} \rceil \geq \frac{\eta n}{6}$ e basta escolher $\eta_0 = \frac{\eta}{6}$, e (3.49) segue também de (3.50). \square

Finalmente, podemos agora provar o Lema 3.1.1.

Prova do Lema 3.1.1. Seja $0 < \eta < 1$. No caso, $15 \leq \eta n \leq n - 3$, então podemos por (3.9) escrever

$$\begin{aligned} I(p_n) &\leq \sum_{k=0}^n \binom{n}{k} (1 - \delta_n)^k \delta_n^{n-k} I(\hat{p}_n^{k*} * \hat{q}_n^{(n-k)*}) \\ &= \sum_{0 \leq k < \eta n} \binom{n}{k} (1 - \delta_n)^k \delta_n^{n-k} I(\hat{p}_n^{k*} * \hat{q}_n^{(n-k)*}) + \sum_{\eta n \leq k \leq n} \binom{n}{k} (1 - \delta_n)^k \delta_n^{n-k} I(\hat{p}_n^{k*} * \hat{q}_n^{(n-k)*}). \end{aligned}$$

Agora, na primeira soma como $k < \eta n \leq n - 3$ podemos usar a cota obtida no Lema 3.1.2 e na segunda soma, como $15 \leq \eta n \leq k \leq n$, podemos usar a cota obtida no Lema 3.1.6. Assim

obtemos

$$\begin{aligned}
I(p_n) &\leq \sum_{0 \leq k < \eta n} \binom{n}{k} (1 - \delta_n)^k \delta_n^{n-k} \bar{C} I(p) (nb_n)^2 + \sum_{\eta n \leq k \leq n} \binom{n}{k} (1 - \delta_n)^k \delta_n^{n-k} C \\
&= \tilde{C} (nb_n)^2 \sum_{0 \leq k < \eta n} \binom{n}{k} (1 - \delta_n)^k \delta_n^{n-k} + C \sum_{\eta n \leq k \leq n} \binom{n}{k} (1 - \delta_n)^k \delta_n^{n-k} \\
&\leq \tilde{C} (nb_n)^2 2^n \delta_n^{(1-\eta)n} + C.
\end{aligned}$$

Mas, como $\delta_n \sim \frac{c}{n}$, podemos escolher $0 < \eta < 1$ (por exemplo $\eta = \frac{1}{2}$) tal que

$$\lim_{n \rightarrow \infty} (nb_n)^2 2^n \delta_n^{(1-\eta)n} = 0.$$

Logo, segue $I(p_n) \leq C$, para alguma constante $C > 0$ e o lema está provado. \square

3.2 Teoremas de Limites

Consideremos, como antes, X_1, X_2, \dots v.a.'s i.i.d. com função de densidade comum $p(x)$ e assumamos que p é absolutamente contínua com derivada $p'(x)$.

Suponha que existem sequências de constantes $\{a_n\}_{n \geq 1}$ e $\{b_n\}_{n \geq 1}$, $b_n > 0$, tais que $Z_n \xrightarrow{\mathfrak{D}} Z$, onde Z é uma v.a. α -estável, com $0 < \alpha \leq 2$ e $Z_n = \frac{S_n}{b_n} - a_n$ com $S_n = X_1 + \dots + X_n$.

Usando as mesmas notações da seção anterior, nosso objetivo é apresentar o Teorema Principal de [6] que estabelece condições para que tenhamos $I(Z_n \parallel Z) \xrightarrow[n \rightarrow \infty]{} 0$ quando Z é α -estável não-extremal.

Lembrando que $I(Z_n \parallel Z)$ denota a Informação de Fisher relativa ou a distância de Informação de Fisher que, de acordo com a Definição 2.2.5, é dada por

$$I(Z_n \parallel Z) = I(p_n \parallel \psi) = \int_{-\infty}^{+\infty} \left(\frac{p'_n(x)}{p_n(x)} - \frac{\psi'(x)}{\psi(x)} \right)^2 p_n(x) dx, \quad (3.51)$$

onde $p_n(x)$ denota a densidade de Z_n e ψ a densidade de Z .

Vimos na Proposição 2.2.7 que se Z é α -estável não-extremal e não-normal, ou seja, $0 < \alpha < 2$ e $-1 < \beta < 1$, então

$$I(Z_n \parallel Z) < +\infty \Leftrightarrow I(Z_n) < +\infty.$$

No caso de Z possuir distribuição Normal, na Proposição 2.2.8 temos que

$$I(Z_n \parallel Z) < +\infty \Leftrightarrow I(Z_n) < +\infty \text{ e } \mathbb{E}Z_n^2 < +\infty,$$

onde $\mathbb{E}Z_n^2 < +\infty$ é equivalente a $\mathbb{E}X_1^2 < +\infty$.

Além disso, pelo Lema 3.1.1, vimos que para Z não-extremal temos que : se $I(Z_{n_0}) < +\infty$ para algum $n_0 \geq 1$ então $\sup_{n \geq n_0} I(Z_n) < +\infty$.

Antes de enunciarmos o resultado principal deste trabalho, necessitamos demonstrar um lema auxiliar que é na verdade um corolário das Proposições 1.4.1 e 1.4.2.

Lema 3.2.1. *Suponhamos que $Z_n \xrightarrow{\mathfrak{D}} Z$, onde Z segue uma lei estável não extremal. Se $I(Z_{n_0}) < \infty$ para algum n_0 , então, para todo n grande o suficiente, as densidades p_n são continuamente diferenciáveis. Além disso, para todo $T > 0$ fixo,*

$$\int_{-T}^T \left(\frac{p'_n(x)}{p_n(x)} - \frac{\psi'(x)}{\psi(x)} \right)^2 p_n(x) dx = o(1) \text{ quando } n \rightarrow \infty. \quad (3.52)$$

Demonstração. Segue do Lema 3.1.1 que para todo $n \geq n_0$, $I(Z_n) < \infty$. Assim, para $n \geq n_0$, a expressão (3.52) fica bem definida. Por outro lado, tomando $C = \sqrt{I(Z_{n_0})}$ segue de (2.19) que

$$|f_{n_0}(t)| \leq \frac{C}{|t|}, \quad t \neq 0.$$

Por outro lado, como $f_{n_0}(t) = \mathbb{E}(e^{itZ_{n_0}}) = e^{ia_{n_0}} \left[f \left(\frac{t}{n_0} \right) \right]^{n_0}$, segue que

$$|f_{n_0}(t)| \leq \left(\frac{C}{b_{n_0}} \right)^{1/n_0} |t|^{-1/n_0}.$$

Assim, temos

$$\int_{-\infty}^{+\infty} |f_{n_0}(t)|^{2n_0} dt \leq \left(\frac{C}{b_{n_0}} \right)^2 \int_{-\infty}^{\infty} |t|^{-2} dt < +\infty,$$

e

$$\int_{-\infty}^{+\infty} |f_{n_0}(t)|^{3n_0} |t| dt \leq \left(\frac{C}{b_{n_0}} \right)^3 \int_{-\infty}^{\infty} |t|^{-2} dt < +\infty.$$

Ou seja, as condições (1.22) e (1.24) das Proposições 1.4.1 e 1.4.2 são satisfeitas e assim segue (1.23) e (1.25). Além disso, temos para n suficientemente grande que as densidades p_n são continuamente diferenciáveis. Fica assim mostrada a primeira parte da proposição.

Agora, já vimos que a densidade ψ de uma lei estável não extremal é positiva e contínua em todo seu domínio. Logo existem a e A em $[-T, T]$ tais que $0 < \psi(a) \leq \psi(x) \leq \psi(A)$ para todo $x \in [-T, T]$.

Assim, de (1.23), tomando $\epsilon = \frac{\psi(a)}{2} > 0$, para n suficientemente grande

$$0 < \frac{\psi(a)}{2} \leq -\frac{\psi(a)}{2} + \psi(x) < p_n(x), \quad \text{para todo } x \in [-T, T],$$

ou seja, fazendo $k = \frac{\psi(a)}{2} > 0$ temos para n suficientemente grande que $p_n(x) \geq k$, se $|x| \leq T$.

Por outro lado, pelas convergência uniformes em (1.23) e (1.25) existem seqüências de números reais não negativos $(A_n)_{n \geq 1}$ e $(B_n)_{n \geq 1}$ que convergem para zero, tais que para n grande o suficiente e para todo $x \in \mathbb{R}$,

$$|p_n(x) - \psi(x)| \leq A_n \quad \text{e} \quad |p'_n(x) - \psi'(x)| \leq B_n. \quad (3.53)$$

Além disso, por (1.12) na Observação 1.3.6, segue que

$$\frac{|\psi'(x)|}{\psi(x)} \leq C, \quad \text{para } x \in \mathbb{R}. \quad (3.54)$$

Finalmente, podemos escrever

$$\begin{aligned} \int_{-T}^T \left(\frac{p'_n(x)}{p_n(x)} - \frac{\psi'(x)}{\psi(x)} \right)^2 p_n(x) dx &= \int_{-T}^T \left(\frac{p'_n(x)\psi(x) - p_n(x)\psi'(x)}{p_n(x)\psi(x)} \right)^2 p_n(x) dx \\ &= \int_{-T}^T \frac{(p'_n(x) - \psi'(x))^2}{p_n(x)} + \frac{2\psi'(x)(p'_n(x) - \psi'(x))(p_n(x) - \psi(x))}{\psi(x)p_n(x)} dx \\ &\quad + \int_{-T}^T \left(\frac{\psi'(x)}{\psi(x)} \right)^2 \frac{(\psi(x) - p_n(x))^2}{p_n(x)} dx. \end{aligned}$$

Agora, usando (3.53), (3.54) e que $p_n(x) \geq k$ se $|x| \leq T$ para n suficiente grande, segue que

$$\begin{aligned} \int_{-T}^T \left(\frac{p'_n(x)}{p_n(x)} - \frac{\psi'(x)}{\psi(x)} \right)^2 p_n(x) dx &\leq \int_{-T}^T \frac{A_n^2}{k} + \frac{2}{k} A_n B_n C + \frac{1}{k} (C B_n)^2 dx \\ &= 2T \left(\frac{A_n^2}{k} + \frac{2}{k} A_n B_n C + \frac{1}{k} (C B_n)^2 \right). \end{aligned}$$

A última expressão tende para zero quando $n \rightarrow \infty$, pois $A_n \rightarrow 0$ e $B_n \rightarrow 0$, e o resultado segue. \square

Finalmente, estamos em condições de demonstrar o Teorema Principal deste trabalho.

Teorema 3.2.2. *Suponhamos que $Z_n \xrightarrow{\mathfrak{D}} Z$, onde Z segue uma lei α -estável não-extremal. Então $I(Z_n \| Z) \rightarrow 0$ quando $n \rightarrow \infty$ se, e somente se, $I(Z_n \| Z) < \infty$ para algum n .*

Demonstração. Suponha que $Z_n \xrightarrow{\mathfrak{D}} Z$, com Z α -estável não-extremal.

Se $I(Z_n \| Z) \xrightarrow{n \rightarrow \infty} 0$, é imediato que $I(Z_n \| Z) < +\infty$ para algum n .

Assim, só temos que mostrar a recíproca.

Suponhamos que $I(Z_{n_0} \| Z) < +\infty$ para algum $n_0 \geq 1$. Então, como observamos anteriormente, das Proposições 2.2.7 e 2.2.8 segue que $I(Z_{n_0}) < +\infty$. Vale notar que no caso Normal, também implica $\mathbb{E}X_1^2 < +\infty$. Daí, pelo Lema 3.1.1 segue que

$$I' = \sup_{n \geq n_0} I(Z_n) < +\infty.$$

Agora, como $Z_n = \frac{S_n}{b_n} - a_n$, da Observação 2.2.2 (e), segue

$$I(Z_n) = \frac{1}{b_n^2} I(S_n).$$

Logo, para todo $n \geq n_0$

$$I(p_n) = I(S_n) = b_n^2 I(Z_n) \leq b_n^2 I' < +\infty.$$

Mas, para $n \geq 2n_0$, temos que $p_n = p_{n_0} * p_{n-n_0}$ e daí, segue da Definição 2.4.1 que $p_n \in \mathfrak{B}_2$, onde \mathfrak{B}_2 indica o conjunto de todas as densidades que são representáveis pela convolução de duas densidades com Informação de Fisher finita. Assim, pela Proposição 2.4.3 segue que $p'(x)$ é contínua e de variação limitada.

Então para $n \geq 2n_0$, escrevemos $n = n_1 + n_2$ com $n_1 = \lfloor \frac{n}{2} \rfloor$ e $n_2 = n - n_1$. Notemos que $n_1, n_2 \geq n_0$ e então

$$I(S_{n_1}) \leq b_{n_1}^2 I' \quad \text{e} \quad I(S_n - S_{n_1}) \leq b_{n_2}^2 I'.$$

Agora, como por (3.3), $b_n = n^{1/\alpha} h(n)$, segue que existe uma constante $I > 0$ tal que $I(S_{n_1}) \leq b_n^2 I$ e $I(S_n - S_{n_1}) \leq b_n^2 I$. Daí para $n \geq 2n_0$, temos que

$$Z_n = \left(\frac{S_{n_1}}{b_n} - a_n \right) + \frac{S_n - S_{n_1}}{b_n}$$

representa a soma de duas v.a.'s independentes e ambas com Informação de Fisher no máximo I . Ou seja, $p_n \in \mathfrak{B}_2(I)$ para todo $n \geq 2n_0$ e pelo Lema 3.2.1 segue para $T > 0$

$$\int_{|x| \leq T} \left(\frac{p'_n(x)}{p_n(x)} - \frac{\psi'(x)}{\psi(x)} \right)^2 p_n(x) dx = o(1) \text{ quando } n \rightarrow \infty. \quad (3.55)$$

Assim, resta mostrar que para um $T > 0$ adequado

$$J = \int_{|x| \geq T} \left(\frac{p'_n(x)}{p_n(x)} - \frac{\psi'(x)}{\psi(x)} \right)^2 p_n(x) dx = o(1) \text{ quando } n \rightarrow \infty. \quad (3.56)$$

Notemos que $J \leq 2J_1 + 2J_2$, onde

$$J_1 = \int_{|x| \geq T} \frac{p'_n(x)^2}{p_n(x)} dx \quad \text{e} \quad J_2 = \int_{|x| \geq T} \left(\frac{\psi'(x)}{\psi(x)} \right)^2 p_n(x) dx. \quad (3.57)$$

Primeiramente, vamos mostrar que podemos escolher $T > 0$ tal que $J_2 < \epsilon$ para n suficientemente grande. Para isto, vamos considerar dois casos separadamente.

Caso 1. $0 < \alpha < 2$.

Dado que Z é estável não-extremal, por (1.12) sua densidade ψ satisfaz,

$$\frac{|\psi'(x)|}{\psi(x)} \leq \frac{c}{1 + |x|},$$

para alguma constante c dependendo unicamente de ψ . Logo,

$$J_2 \leq \int_{|x| \geq T} \left(\frac{c}{1 + |x|} \right)^2 p_n(x) dx \leq \int_{|x| \geq T} \left(\frac{c}{1 + T} \right)^2 p_n(x) dx \leq \left(\frac{c}{1 + T} \right)^2,$$

e assim podemos escolher $T > 0$ suficientemente grande tal que $J_2 < \epsilon, \forall n$.

Caso 2. $\alpha = 2$ e $\mathbb{E}X_1^2 < \infty$.

Podemos assumir, sem perda de generalidade, que $\mathbb{E}X_1 = 0, \mathbb{E}X_1^2 = 1$. Neste caso, ψ é a densidade de uma Normal padrão e podemos escrever

$$J_2 = \int_{|x| \geq T} x^2 p_n(x) dx.$$

Agora, como $\mathbb{E}Z_n^2 = 1, \forall n \geq 1$ temos que $q(x) = x^2 p_n(x)$ é uma densidade de probabilidade na reta, com função característica $\varphi(t) = \int_{-\infty}^{\infty} e^{itx} x^2 p_n(x) dx = -f_n''(t)$.

Agora, aplicando a conhecida Desigualdade de truncação (Vide [1], pág. 303) para a densidade q , temos

$$\int_{|x| \geq T} q(x) dx \leq T \int_0^{2/T} (1 - \operatorname{Re} \varphi(t)) dt.$$

Ou seja,

$$J_2 \leq T \int_0^{2/T} (1 + \operatorname{Re} f_n''(t)) dt.$$

Em nosso caso, a função característica de Z é $g(t) = e^{-t^2/2}$, então $1 + g''(t) \rightarrow 0$ quando $t \rightarrow 0$ e para qualquer constante $c > 0$, $\sup_{|t| \leq c} |f_n''(x) - g''(x)| \rightarrow 0$ quando $n \rightarrow \infty$.

Assim, dado $\epsilon > 0$, podemos escolher $T > 0$ tal que

$$\begin{aligned} J_2 &\leq T \int_0^{2/T} |1 + g''(t)| dt + T \int_0^{2/T} |f_n''(t) - g''(t)| dt \\ &\leq \frac{\epsilon}{2} + 2 \sup_{|t| \leq \frac{2}{T}} |f_n''(t) - g''(t)| dt, \end{aligned}$$

e daí, segue que $J_2 < \epsilon$ para n suficientemente grande e T apropriadamente escolhido.

Por outro lado, para estimar J_1 , como $p_n \in \mathfrak{B}_2$ aplicaremos o Corolário 2.4.6 para obter que para qualquer $T \in \mathbb{R}$,

$$\begin{aligned} J_1 &\leq I^{3/4} \left(\sqrt{p_n(T)} |\log p_n(T)| + \sqrt{p_n(-T)} |\log p_n(-T)| \right) \\ &\quad + 2I \left(\int_{|x| \geq T} p_n(x) [\log(p_n(x))]^2 dx \right)^2. \end{aligned} \quad (3.58)$$

Usando (1.22) e a relação assintótica para $\psi(x)$ quando $x \rightarrow \infty$ da Proposição 1.3.4 temos para todo n suficiente grande e para todo $T \geq T_0$,

$$\sqrt{p_n(\pm T)} |\log p_n(\pm T)| \leq \bar{c} \frac{\log T}{\sqrt{T}} + \bar{\epsilon}_n, \quad (3.59)$$

onde $\bar{\epsilon}_n \rightarrow 0$ quando $n \rightarrow \infty$, $\bar{c} > 0$ e T_0 suficientemente grande que dependem unicamente de ψ .

Agora, nosso objetivo é limitar a integral J_1 com expressões que possam ser tão pequenas como desejemos para n grande. Note que de (3.59) o problema fica resolvido para a primeira parcela de (3.58). Assim, só resta limitar a integral

$$\int_{|x| \geq T} p_n(x) \log^2(p_n(x)) dx.$$

Para isto, vamos escrever

$$\{x \in \mathbb{R} : |x| \geq T\} = A \cup B,$$

onde $A = \{x \in \mathbb{R} : |x| \geq T, p_n(x) \leq |x|^{-4}\}$ e $B = A^c$. Da definição do conjunto A segue

$$\int_A p_n(x) [\log(p_n(x))]^2 dx \leq \int_A \frac{1}{|x|^4} \left[\log \frac{1}{|x|^4} \right]^2 dx = 16 \int_A \frac{1}{|x|^4} [\log |x|]^2 dx.$$

De onde,

$$\int_A p_n(x) [\log(p_n(x))]^2 dx \leq \frac{32}{T}. \quad (3.60)$$

Por outro lado, já vimos que $p_n \in \mathfrak{B}_2(I)$ sempre que $n \geq 2n_0$. Daí, p_n está uniformemente limitada, pois, $\sup_x p_n(x) \leq \sqrt{I}$ para todo $n \geq 2n_0$. Logo, sobre o conjunto B , temos

$$\begin{aligned} |\log p_n(x)| &\leq |\log \sqrt{I}| \leq |\log \sqrt{I}| + \log \frac{\sqrt{I}}{p_n(x)} \leq |\log \sqrt{I}| + |x|^4 \log \sqrt{I} \\ &= \frac{1}{2} \log I + 4 \log |x| + \frac{1}{2} |\log I| \leq \frac{1}{2} |\log I| + 4 \log |x| + \frac{1}{2} |\log I| \\ &= 4 \log |x| + |\log I|. \end{aligned}$$

Agora, note que $(4 \log |x| + |\log I|)^2 = (\log |x|^4 + |\log I|)^2 = \begin{cases} 16 \log^2(I^{\frac{1}{4}}|x|) & \text{se } I \geq 1 \\ 16 \log^2(I^{-\frac{1}{4}}|x|) & \text{se } 0 < I < 1. \end{cases}$

Portanto,

$$\int_B p_n(x) [\log(p_n(x))]^2 dx \leq C \int_{|x| \geq T} p_n(x) [\log |x|]^2 dx, \quad (3.61)$$

onde a constante C depende unicamente de I .

Finalmente, pela Observação 1.3.3 (f), os momentos $\mathbb{E}|Z_n|^\delta$ estão uniformemente limitados em n , sempre que $0 < \delta < \alpha$. Fazendo $\delta = \frac{\alpha}{2}$, seja $K > 0$ tal que para todo n , $\mathbb{E}|Z_n|^{\alpha/2} \leq K$. Então usando a cota elementar $|x|^{\alpha/4} \geq c_\alpha [\log |x|]^2$, para $|x| \geq T_0$, onde c_α é uma constante que depende de α , podemos obter,

$$\begin{aligned} K &\geq \mathbb{E}|Z_n|^{\alpha/2} = \int_{\mathbb{R}} |x|^{\alpha/2} p_n(x) dx \geq T^{\alpha/4} \int_{|x| \geq T} |x|^{\alpha/4} p_n(x) dx \\ &\geq c_\alpha T^{\alpha/4} \int_{|x| \geq T} p_n(x) [\log |x|]^2 dx. \end{aligned}$$

Assim, de (3.61)

$$\int_B p_n(x) [\log p_n(x)]^2 dx \leq CT^{-\alpha/4}, \quad (3.62)$$

com alguma constante C independente de n .

Usando (3.60) e (3.62) em (3.58), obtemos

$$J_1 \leq 2I^{3/4} \left(\bar{c} \frac{\log T}{\sqrt{T}} + \bar{\epsilon}_n \right) + 2I \left(\frac{32}{T} + cT^{-\alpha/4} \right)^2 \leq CT^{-\alpha/8} + \epsilon_n, \quad (3.63)$$

onde a constante C é independente de n e $\epsilon_n \rightarrow 0$ quando $n \rightarrow \infty$. Assim, isto completa a prova de (3.56), que juntamente com (3.55) resulta que $I(Z_n \parallel Z) \rightarrow 0$ quando $n \rightarrow \infty$. \square

Observação 3.2.3.

a) Vale notar que no caso extremal $|\beta| = 1$, a densidade $\psi(x)$ tem um comportamento diferente do caso não-extremal dado em (1.11). Por exemplo, quando $0 < \alpha < 1$ e $\beta = 1$, a densidade é estritamente positiva somente num intervalo do tipo $(x_0, +\infty)$ ou $(-\infty, x_0)$ e quando $x \rightarrow x_0$, $\psi(x) \rightarrow 0$ extremadamente rápido. Assim, para garantir a finitude de $I(Z_n \parallel Z)$ em (3.51) seria necessário estabelecer condições adicionais sobre $p(x)$ e $p_n(x)$ na vizinhança de x_0 . Um comportamento similar é observado no caso $1 \leq \alpha < 2$ e $|\beta| = 1$. Para maiores detalhes sobre o comportamento das densidades α -estáveis não-extremais veja, por exemplo, [13] ou [27].

b) Em [7] é provado que as seguintes duas condições são equivalentes:

(i) $I(Z_n) < \infty$ para algum $n \geq 1$.

(ii) Para algum $n \geq 1$, Z_n tem densidade p_n continuamente diferenciável tal que

$$\int_{-\infty}^{\infty} |p'_n(x)| dx < \infty.$$

Ou seja, para algum n , a densidade p_n é de variação limitada.

Sendo assim, o Teorema 3.2.2 é válido quando (ii) é satisfeita.

Referências Bibliográficas

- [1] ASH, R.B. and DOLÉANS- DADE. C. *Probability and Measure Theory*. Academic Press, 2000.
- [2] BARRON, A.R. *Monotonic central limit theorem for densities*. Technical report 50. Dept. Statistics. Stanford University. 1984, 1-31.
- [3] BARRON, A. and JOHNSON, O. *Fisher information inequalities and the central limit theorem*. Probab. Theory Related Fields, 2004, 391-409.
- [4] BOBKOV, S.G., CHISTYAKOV, G.P. and GOTZE, F. *Bounds for characteristic functions in terms of quantiles and entropy*. Electron. Commun, Probab, 2012, 1-9.
- [5] BOBKOV, S.G., CHISTYAKOV, G.P. and GOTZE, F. *Convergence to stable laws in relative entropy*. J. Theoret. Probab, 2013, 803-818.
- [6] BOBKOV, S.G., CHISTYAKOV, G.P. and GOTZE, F. *Fisher information and convergence to stable laws*. BERNOULLI, 2014, 1620-1646.
- [7] BOBKOV, S.G., CHISTYAKOV, G.P. and GOTZE, F. *Fisher information and the central limit theorem*. Probab. Theory. Relat. Fields, 2013, 1-59.
- [8] CSISZÁR, I. *Information-type measures of difference of probability distributions and indirect observations*. Studia Sci. Math. Hungar, 1967, 299-308.
- [9] COVER, T.M. and THOMAS, J.A. *Elements of Information Theory*. Wiley-interscience. New Jersey, 2006.
- [10] FELLER, W. *An introduction to probability theory and its applications II*. Wiley. New York, 1971.

- [11] FRIEDEN, A. *Physics from Fisher information, A unification*. Cambridge University Press. Cambridge, 1998.
- [12] FISHER R.A. *Theory and Statistical Estimation*. Cambridge Philosophical Society, 1925, 700-725.
- [13] IBRAGIMOV, I.A. and LINNIK, J.V. *Independent and Stacionary Sequences of Random Variables*. Wolters-Noordhoff, Groningen, 1971.
- [14] JOHNSON, O. *Information Theory and The Central Limit Theorem* Imperial College Press. London, 2004.
- [15] KULLBACK, S. *A lower bound for discrimination in terms of variation*. IEEE Trans.Inform, 1967, 126-127.
- [16] KULLBACK, S. and LEIBLER, R. *On information and sufficiency*. Ann. Math-Statist.,1951, 79-86.
- [17] LINNIK, J.V. *An information-theoretic proof of the central limit theorem with Lindeberg conditions*. . Theory Probab. Appl, 1959, 288-299.
- [18] MORREY, C.B. and PROTTER, M.H *A First Course in Real Analysis*. Springer-Verlag. New york, 1991.
- [19] NOLAN, J. *Stable Distributions Models for Heavy Tailed Data*. Birkhauser, 2007.
- [20] PINSKER, M.S. *Information and information stability of random variables and processes* Translated and Edited by Amiel Feinstein San Francisco, CA: Holden-Day, 1964.
- [21] RAO, C.R. *Efficient estimates and optimum inference procedures in large samples*. Wiley, 1962, 46-72.
- [22] SENETA, E. *Regularly varying functions*. Springer, 1976.
- [23] SHANON, C.E. *A mathematical theory of communication*. The Bell System Technical Journal,1948, 379-423, 623-656.
- [24] STAM, A.J. *Some inequalities satisfied by the quantities of information of Fisher and Shannon*. Information and Control 2, 1959, 101-112.

- [25] VUCHAIKIN, V.V. and ZOLOTAREV, V.M. *Chance and stability: Stable Distributions and their Applications*. Modern Probability and Statistics, VSP. volume 3, 1999.
- [26] YAMAZATO, M. *Unimodality of infinitely divisible distribution functions of class L*. Ann. Probab. 6, 1978, 523-531.
- [27] ZOLOTAREV, V.M. *One-dimensional stable distributions*. Probability Theory and mathematical statistics. Moscow: Nauka, 1983.