UNIVERSIDADE FEDERAL DO CEARÁ
CENTRO DE TECNOLOGIA
DEPARTAMENTO DE ENGENHARIA DE TELEINFORMÁTICA
PROGRAMA DE PÓS-GRADUAÇÃO EM ENGENHARIA DE TELEINFORMTICA

LUIZA HELENA FÉLIX DE ANDRADE

AVANÇOS EM VARIEDADES ESTATÍSTICAS GENERALIZADAS: UMA EXTENSÃO PARA ARCOS EXPONENCIAL E MISTURA.

FORTALEZA

2018

LUIZA HELENA FÉLIX DE ANDRADE

AVANÇOS EM VARIEDADES ESTATÍSTICAS GENERALIZADAS: UMA EXTENSÃO PARA ARCOS EXPONENCIAL E MISTURA

Tese apresentada ao Programa de Pós-graduação em Engenharia de Teleinformática do Departamento de Engenharia de Teleinformática da Universidade Federal do Ceará, como parte dos requisitos necessários para a obtenção do título de Doutor em Engenharia de Teleinformática. Área de concentração: Sinais e Sistemas.
Orientador: Prof. Dr. Charles Casimiro Cavalcante
Coorientador: Prof. Dr. Rui Facundo Vigelis.

FORTALEZA

2018

Dados Internacionais de Catalogação na Publicação
Universidade Federal do Ceará
Biblioteca Universitária
Gerada automaticamente pelo módulo Catalog, mediante os dados fornecidos pelo(a) autor(a)

A568a Andrade, Luiza Helena Felix de.
Avanços em Variedades estatísticas generalizadas: uma extensão para arcos exponecial e mistura / Luiza Helena Felix de Andrade. – 2018.
82 f. : il. color.

Tese (doutorado) – Universidade Federal do Ceará, Centro de Tecnologia, Programa de Pós-Graduação em Engenharia de Teleinformática, Fortaleza, 2018.
Orientação: Prof. Dr. Charles Casimiro Cvalacante.
Coorientação: Prof. Dr. Rui Facundo Vigelis.

1. Variedades estatísticas generalizadas. 2. Famílias de exponenciais. 3. Arco exponencial. 4. Arco mistura. I. Título.

CDD 621.38

LUIZA HELENA FÉLIX DE ANDRADE

AVANÇOS EM VARIEDADES ESTATÍSTICAS GENERALIZADAS: UMA EXTENSÃO PARA ARCOS EXPONENCIAL E MISTURA

Tese apresentada ao Programa de Pós-graduação em Engenharia de Teleinformática do Departamento de Engenharia de Teleinformática da Universidade Federal do Ceará, como parte dos requisitos necessários para a obtenção do título de Doutor em Engenharia de Teleinformática. Área de concentração: Sinais e Sistemas.

Aprovoda em: 23/ 03/ 2018.

BANCA EXAMINADORA

Prof. Dr. Charles Casimiro Cavalcante (Orientador)
Universidade Federal do Ceará (UFC)

Prof. Dr. Rui Facundo Vigelis (Co-Orientador)
Universidade Federal do Ceará (UFC)

Profª. Dra. Sueli Irene Rodrigues Costa
Universidade Estadual de Campinas (UNICAMP)

Prof. Dr. João Eloir Strapasson
Universidade Estadual de Campinas (UNICAMP)

Prof. Jorge Herbert Soares de Lira
Universidade Federal do Ceará (UFC)

Prof. Dr. João César Moura Mota
Universidade Federal do Ceará (UFC)

Dedico este trabalho aos meus pais Maria de Lourdes e Osmar.

AGRADECIMENTOS

Agradeço primeiramente a Deus e à minha família por todo o apoio e dedicação para a realização desse grande sonho. Agradeço à minha mãe Maria de Lourdes, por sempre fazer até o impossível para que os estudos estivessem em primeiro lugar em nossas vidas, mesmo com todas as dificuldades financeiras. Ao meu pai por ser esse pai amigo e companheiro em todas as horas. A minha irmã Márcia Vanessa (Guell) por ser minha maior incentivadora e "terapeuta"em todas as etapas desse processo. Ao meu sobrinho Caio por ter feito com que eu quisesse fazer sempre o melhor pra ser um bom exemplo pra ele. Ao meu irmão Júnior e meu sobrinho Levi por todas as orações dedicadas a mim, na torcida que tudo desse certo. À minha nova família Fernandes Araújo pela torcida e compreensão, em especial à minha sogrinha D. Margarida por todas as orações e torcida. Um agradecimento especial ao meu amor, marido e companheiro, meu príncipe Silvio Roberto, por todo apoio, compreensão e dedicação, sem ele essa caminhada teria sido impossível. A todos os professores que me ajudaram em toda a minha formação acadêmica. Ao Professor João Montenegro de Miranda, por ter sido meu grande incentivador na graduação. Ao Professor Jorge Herbert Soares de Lira pela orientação no mestrado e por acreditar no meu potencial e me apresentar ao Professor Charles Casimiro Cavalcante, a quem tenho extrema gratidão pela excelente orientação ao longo desses 5 anos de doutorado e admiração pelo brilhante profissional que é e que me inspira a ser melhor a cada dia. Agradeço imensamente ao Professor Rui Facundo Vigelis, pela orientação e paciência comigo. Aos meus colegas de caminhada David de Souza Carneiro e Leidmar Vieira pelas longas discussões e pelos dias vividos no laboratório. Aos meus amigos que não fizeram doutorado comigo, mas foram grandes incentivadores em todo o processo, Fabricio de Figueredo, Natália Queiroz, Andréa Moura, Valdenize Lopes e às professoras gatas, em especial à Jusciane Costa por todo o apoio e amizade. Aos professores participantes da banca examinadora, Professora Sueli Irene Rodrigues Costa, Professor João Eloir Strapasson, Professor João César Moura Mota e ao Professor Jorge Herbert Soares de Lira pelo tempo, pelas valiosas colaborações e sugestões.

À CAPES, pelo apoio financeiro com a manutenção da bolsa de auxílio e à UFERSA pela liberação das minhas atividades docentes para que eu pudesse me dedicar ao doutorado.

"E lembre-se, com grandes poderes vêm grandes responsabilidades (Tio Ben em Homem-Aranha)."

RESUMO

A coleção de todas as densidades de probabilidade estritamente positivas, que são equivalentes a uma medida μ, $\mathcal{P}_\mu$, foi dotada com uma estrutura de C^∞-variedade de Banach. Essa estrutura é baseada em φ-famílias de distribuições de probabilidade. Conectar por arcos duas distribuições de probabilidade, ou seja, encontrar uma curva em que os extremos sejam duas distribuições e que esta curva esteja totalmente contida em $\mathcal{P}_\mu$, era uma questão em aberto para variedades estatísticas generalizadas. Nesta Tese, arcos em variedades estatísticas generalizadas são investigados. Os arcos exponencial e mistura, já bem conhecidos em Geometria da Informação, podem ser vistos como um caso especial desses arcos. Nós garantimos que o arco mistura generalizado está bem definido. Encontramos condições necessárias e suficientes para quaisquer duas distribuições de probabilidade serem conectadas por um arco exponencial generalizado, um φ-arco. Provamos ainda que, a partir de uma exponencial deformada e de duas distribuições de probabilidade fixadas, uma generalização da divergência de Rényi existe, sobre algumas condições essa generalização da divergência de Rényi é relacionada à φ-divergência, que pode ser vista como uma generalização da divergência de Kullback-Leibler. A função de normalização, em uma φ-família, é o análogo da função geradora de cumulantes. Foi ainda estudado o comportamento da função de normalização próximo ao bordo do domínio da φ-família, o que é um resultado necessário ao desenvolvimento dos arcos mistura generalizados, uma vez que esses arcos são dados a partir dos funcionais que pertencem ao subdiferencial da função de normalização. Foram encontradas condições para que os arcos generalizados possam ser tomadas sem que, necessariamente, as distribuições de probabilidade conectadas sejam os pontos extremos desses arcos, ou seja, os arcos são abertos. Um outro resultado importante desse trabalho foi provar que o conjunto de todas as distribuições conectadas, por um φ-arco aberto, a uma distribuição de probabilidade fixada, é a própria φ-família de distribuições de probabilidade. Garante-se ainda que conectar duas distribuições de probabilidade por arcos abertos é uma relação de equivalência para ambos os arcos generalizados.

Palavras-chave: Variedades estatísticas generalizadas. Famílias de exponenciais. Arco exponencial. Arco mistura.

ABSTRACT

The collection of all strictly positive probability densities, which are equivalent to a measure μ, $\mathcal{P}_\mu$, was endowed with a C^∞-Banach manifold structure. This structure is based on φ-families of probability distributions. To connect two probability distributions, that is, to find a curve where the ends are two distributions and that this curve is totally contained in $\mathcal{P}_\mu$, was an open question for generalized statistical manifold. In this thesis, arcs in the generalized statistical manifold are investigated. The exponential arcs and misture, already well-known in Information Geometry, can be seen as a special case of these arcs. We guarantee that the generalized misture arc is well defined. We found necessary and sufficient conditions for any two probability distributions to be connected by a generalized exponential arc, a φ-arco. We also prove that, from a deformed exponential and two fixed probability distributions, a generalization of the Rényi divergence exists, on some conditions this generalization of the Rényi divergence is related to φ-divergence, which can be seen as a generalization of the Kullback-Leibler divergence. The normalizing function, in a φ-family, is the analog of the cumulating generating function. We also studied the behavior of the normalization function near the boundary of the φ-family domain, which is a necessary result to the development of the generalized misture arcs, since these arcs are given from the functional ones that belong to the subdifferential of the normalizing function. Conditions were found so that generalized arcs can be taken without necessarily connecting the probability distributions to the extreme points of these arcs, that is, the arcs are opened. Another important result of this work has been to prove that the set of all the distributions connected by an open φ-arc, to a fixed probability distribution, is the φ-family of probability distributions. It is further ensured that connecting two probability distributions by open arcs is an equivalence relation for both generalized arcs.
Keywords: Generalized statistical manifolds. Exponential families. Exponential arcs. Mixture arcs.

LISTA DE FIGURAS

LISTA DE SÍMBOLOS

C^∞	Infinitamente diferenciável com derivada contínua
L^Φ	Espaço de Musielak-Orlicz relacionado a função Φ
$\tilde{L}^\Phi$	Classe de Musielak-Orlicz relacionada a função Φ
$(L^\Phi)^*$	Dual topológico de L^Φ
Φ^*	Função complementar da função de Musielak-Orlicz Φ
L^{Φ^*}	Espaço de Musielak-Orlicz da função Φ^*
$(L^\Phi)_s^\sim$	Coleção de todos funcionais puramente singular
$\Phi'_+(t,\cdot)$	derivada à direita de $\Phi(t,\cdot)$
$\mathcal{P}_\mu$	Conjunto de todas as densidades de probabilidade estritamente positivas
∂f	Subgradiente da função f
$D(\partial f)$	Domínio do subgradiente de f
$\mathcal{E}_p$	Família de exponenciais
$\mathcal{F}_c^\varphi$	φ-família de probabilidades

LISTA DE ABREVIATURAS E SIGLAS

q.t.p.	Quase todo ponto
s.c.i.	Semicontínua inferiormente

SUMÁRIO

1 INTRODUÇÃO

1.1 Estado da arte e motivação

Em modelos geométricos para estatística, Geometria da Informação (AMARI, 2001; CALIN and UDRIŞTE, 2014; AMARI, 2016) é o ramo da teoria de probabilidade dedicado a investigar funções de densidade de probabilidade equipadas com uma estrutura de geometria diferencial. Uma estrutura de geometria diferencial para as famílias de distribuições multiparamétricas foi fornecida em (AMARI, 1982). Em meados dos anos 80, outros tópicos relacionados a essas famílias, tais como fibrado e dualidade de conexões de modelos estatísticos, foram investigados em (AMARI, 1985) e (AMARI and NAGAOKA, 2000), respectivamente. No caso paramétrico, família exponencial, mistura e α-conexões, assim como sua estrutura dual, estão entre os mais importantes objetos geométricos (AMARI and NAGAOKA, 2000). A estrutura dual das α-conexões é o ponto chave que distingue uma variedade estatística contra variedades diferenciais arbitrárias.

A função de divergência é um tópico essencial em Geometria da Informação, para ambos os casos, paramétrico e não paramétrico, uma vez que uma métrica e conexões duais podem ser induzidas a partir de uma divergência (AMARI and CICHOCKI, 2010; AMARI, 2009; ZHANG, 2004; NIELSEN and NOCK, 2017). Dentre as divergências mais conhecidas e mais bem sucedidas na medida da dissimilaridade entre densidades de probabilidade está a divergência de Rényi (RÉNYI, 1961; ERVEN and HARREMOES, 2014), que é uma divergência dependente de um parâmetro α e no caso limite quando $\alpha \to 1$, temos a divergência de Kullback-Leibler (KULLBACK and LEIBLER, 1951) . Encontrar uma estrutura geométrica para famílias de distribuições multiparamétricas, com uma descrição mais geral, é um dos tópicos de interesse em geometria da informação (AMARI, OHARA, and MATSUZOE, 2012; HARSHA and MOOSATH, 2015; MATSUZOE, 2014; MATSUZOE and WADA, 2015).

Modelos estatísticos não paramétricos (GINÉ and NICKL, 2015) já se mostram importantes em pesquisas voltadas para psicologia, medição e percepção (TOWENSED, SOLOMON, and SMITH, 2001) e aplicações em finanças (TRIVELLATO, 2013). No caso paramétrico, a variedade de funções de densidade de probabilidade obtém uma topologia Euclidiana a partir do espaço dos parâmetros naturais. Para o caso não paramétrico, o maior desafio é definir uma topologia conveniente e uma noção de convergência. Pistone e Sempi (PISTONE and SEMPI, 1995) foram os primeiros a formular uma rigorosa extensão de dimensão infinita para as famílias de exponenciais. Eles dotaram o conjunto $\mathcal{P}_\mu$ de todas as densidades de probabilidade, com uma estrutura de variedade de Banach exponencial, usando espaços de Orlicz associados a uma função de Young (KRASNOSEL'SKI and RUTICKI, 1961). Em um trabalho posterior (PISTONE and ROGANTIN, 1999), mais propriedades das variedades estatísticas foram estudadas, especificamente com respeito a condição de ortogonalidade.

Assim como no caso paramétrico, em modelos não paramétricos as conexões exponencial e mistura estão entre os objetos mais importantes. Para encontrar estas conexões é necessário garantir a existência de arcos abertos. Usando a noção de convergência exponencial, os autores em (GIBILISCO and PISTONE, 1998) investigaram essas conexões. Neste trabalho os autores construíram as conexões exponencial e mistura no sentido em que a relação entre elas é a mesma que no caso paramétrico. Uma abordagem diferente foi usada em (GRASSELLI, 2010), na que o arco mistura foi adicionalmente estudado. Além disso, (GRASSELLI, 2010) provou que duas densidades de probabilidade na mesma vizinhança são conectadas por arcos mistura abertos se e somente se a diferença entre suas varáveis aleatórias é limitada.

A variedade estatística exponencial foi estudada posteriormente em (CENA and PISTONE, 2007), com outro sistema de cartas, o modelo estatístico $\mathcal{E}(p)$, chamado de modelo exponencial maximal. Os autores de (CENA and PISTONE, 2007) provaram que este modelo é o conjunto de todas as densidades positivas conectadas a uma dada densidade positiva p por um arco exponencial e vice e versa. Este modelo exponencial $\mathcal{E}(p)$ com os arcos exponencial e mistura abertos foram estudados mais recentemente em (SANTACROCE, SIRI, and TRIVELLATO, 2016, 2017), em que propriedades de dualidade de modelos estatísticos foram provadas. Aplicações de geometria da informação não paramétrica em Física estatística usando conexões por arcos foi estudada em (PISTONE, 2013a).

A generalização da variedade estatística exponencial tem sido um tópico ativo de pesquisa nos últimos anos (ZHANG and HÄSTÖ, 2006). Em (PISTONE, 2009) foi usada a κ-exponencial de Kaniadakis (KANIADAKIS, 2001) na construção de uma variedade estatística. Em (VIGELIS and CAVALCANTE, 2013b) foi proposta uma φ-família de distribuições de probabilidade $\mathcal{F}_c^\varphi$, que generaliza a família de exponencial $\mathcal{E}(p)$. Esta generalização é baseada na substituição da função exponencial pela exponencial deformada $\varphi(\cdot)$ que satisfaz algumas propriedades e fornece ao conjunto $\mathcal{P}_\mu$ uma estrutura de variedade de Banach, assim chamando variedade estatística generalizada. Em (PISTONE, 2013b) uma revisão de Geometria da Informação não paramétrica com questões específicas do conjunto infinito dimensional é fornecida. Neste trabalho, a deformada exponencial foi estudada com uma exponencial deformada definida em (NAUDTS, 2011) e um espaço modelo foi construído de acordo com o proposto em (VIGELIS and CAVALCANTE, 2013b).

Em (VIGELIS and CAVALCANTE, 2013b) o conjunto $\mathcal{P}_\mu$ foi dotado com uma estrutura de C^∞-variedade de Banach. Foi mostrado também que a função normalizadora ψ é convexa e Gâteaux-diferenciável e uma generalização da divergência de

Kullback-Leibler foi encontrada sendo denotada como φ-divergência. Em um outro trabalho (VIGELIS and CAVALCANTE, 2013a) os autores estudaram a condição Δ_2 e suas consequências nas φ-famílias. O comportamento da função normalizadora ψ próximo ao bordo do seu domínio foi também estudado. Em (SOUZA, VIGELIS, and CAVALCANTE, 2016) uma generalização da divergência de Rényi foi encontrada e uma geometria induzida por essa divergência foi fornecida.

Nosso objetivo nesse trabalho é encontrar arcos que conectem duas densidades de probabilidade na variedade estatística generalizada tanto arcos do tipo exponencial quanto arcos do tipo mistura e provar propriedades que ocorrem com essas conexões. Conseguimos construir e provar que estão bem definidas essas classes de arcos. Achamos condições necessárias e suficientes para que a generalização da divergência de Rényi obtida com a conexão por arcos, esteja bem definida. Estudamos o comportamento de ψ próximo ao bordo do seu domínio, o que é um resultado relevante para a construção dos arcos mistura generalizados.

1.2 Publicações decorrentes dessa tese

- Em (VIGELIS, ANDRADE, and CAVALCANTE, 2017) encontramos uma condição necessária e suficiente para a qual duas densidades de probabilidade na variedade estatística generalizada podem ser conectadas por um φ-arco.
- Em (ANDRADE *et al.*, 2017), considerando uma classe diferente de exponencial deformada, ou seja, funções $\varphi(\cdot)$ que não satisfazem todas as condições da definição, concluímos que a função normalizadora ψ próximo ao bordo do seu domínio se comporta de forma diferente.
- Em (ANDRADE *et al.*, 2018), definimos uma classe de arcos do tipo mistura na variedade estatística generalizada, além disso, garantimos que para exponenciais deformadas estritamente convexas, essa classe de arcos está bem definida.

Temos em processo de submissão os seguintes artigos:

- "Conditions for the Existence of a Generalization of Rényi Divergence" que encontra uma condição sobre a exponencial deformada para a qual podemos generalizar a divergência de Rényi.
- "A Family of Statistical Divergences Based on Quasiarithmetic Means", em que encontramos uma generalização para a entropia e divergência de Tsallis e provamos a desigualdade de Pinsker.

1.3 Organização da tese

A organização da tese é dada como segue. Na Seção 3 temos uma revisão necessária sobre os espaços de Musielak-Orlicz e a construção das φ-famílias de distribuições de probabilidade. Na Seção 4, encontramos condições necessárias e suficientes para que duas densidades de probabilidade sejam conectadas por um φ-arco, que pode ser visto como um arco exponencial generalizado. Ainda nessa seção encontramos condições para as quais a divergência de Rényi generalizada está bem definida. Encontramos também os φ-arcos abertos e provamos que a componente conectada a uma distribuição é igual a φ-família dessa distribuição. Na Seção 5 estudamos o comportamento da função normalizadora ψ próximo ao bordo do seu domínio, quando supomos que uma das condições na definição de exponencial deformada não ocorre. Na Seção 6 temos a construção de outra classe de arcos que pode ser vistos como arcos que generalizam os arcos mistura. Para a construção desses arcos utilizamos um ferramental de análise convexa, contido na Seção 6.1. Por fim, na Seção 7 temos nossas conclusões e perspectivas futuras.

2 VARIEDADES ESTATÍSTICAS EXPONENCIAIS

Nesta seção, vamos lembrar a construção das variedades estatísticas exponenciais (PISTONE and SEMPI, 1995; GIBILISCO and PISTONE, 1998; PISTONE and ROGANTIN, 1999; CENA and PISTONE, 2007; SANTACROCE, SIRI, and TRIVELLATO, 2016) e alguns resultados conhecidos em espaços de Orlicz (KRASNOSELl'SKI and RUTICKI, 1961), que são os espaços onde essas variedade estatísticas são modeladas.

2.1 Funções de Young e espaços de Orlicz

Seja (T, Σ, μ) o espaço de medida σ-finita e não atômica, ou seja, T um conjunto de pontos e Σ é uma σ-álgebra de seus subconjuntos em que a σ-aditiva função $\mu : \Sigma \to \mathbb{R}_+$ é dada (ISNARD, 2007). Para entendermos o significado de espaço de medida não atômica, é necessário saber a definição de átomo e de conjunto difuso.

Definição 2.1. Um conjunto $A \in \Sigma$ é uma átomo para a medida μou um μ-átomo se $\mu(A) > 0$ e para cada $B \subset A$, $B \in \Sigma$, ou $\mu(B) = 0$ ou $\mu(A - B) = 0$.

Ou seja, um átomo é um conjunto mensurável com medida positiva e que não tem nenhum subconjunto que tenha medida positiva. Um conjunto $D \in \Sigma$ é um conjunto difuso para μ se este conjunto não contém nenhum μ-átomo. Isto é, para $0 \leq \lambda \leq \mu(D)$, podemos encontrar um conjunto $D_1 \subset D$, $D \in \Sigma$ tal que $\mu(D_1) = \lambda$. Este conjunto D é chamado de não atômico. Uma medida que não possui átomos é chamada não atômica, caso contrário é chamada de medida atômica.

Seja $\mathcal{P}_\mu$ a família de todas as medidas de probabilidade em T que são equivalentes a uma medida μ, denotado por:

$$\mathcal{P}_\mu = \left\{ p \in L^0 : \; p > 0, \; \int_T p d\mu = 1 \right\},$$

em que L^0 é o conjunto das funções mensuráveis definidas em T . Em (PISTONE and SEMPI, 1995; GIBILISCO and PISTONE, 1998; PISTONE and ROGANTIN, 1999) o conjunto $\mathcal{P}_\mu$ foi dotado de uma estrutura de variedade de Banach, modelado em espaços de Orlicz, que são espaços relacionados a funções de Young. Vamos lembrar a definição e algumas propriedades de função de Young (RAO and REN, 1991). Além disso, vamos ver alguns exemplos importantes para a construção das famílias de exponenciais.

Definição 2.2. Uma função de Young é uma função convexa $\Phi : \mathbb{R} \to [0, \infty]$ tal que
(i) $\Phi(0) = 0$,
(ii) $\Phi(-x) = \Phi(x)$,
(iii) $\lim_{x\to\infty} \Phi(x) = \infty$.
Cada função Φ pode ser associada a uma outra função convexa $\Phi^* : \mathbb{R} \to [0, \infty]$ que tem propriedades similares e é definida por

$$\Phi^*(y) = \sup \{x|y| - \Phi(x); \; x \geq 0\}, \quad y \in \mathbb{R},$$

e é chamada de função complementar de Φ. Segue da definição que $\Phi^*(0) = 0$, $\Phi^*(-x) = \Phi^*(x)$ e $\lim_{x\to\infty} \Phi^*(x) = \infty$. O par (Φ, Φ^*) satisfaz a desigualdade de Fenchel-Young (RAO and REN, 1991)

$$xy \leq \Phi(x) + \Phi^*(y), \quad x, y \in \mathbb{R}.$$

Exemplo 2.3. A função $\Phi(x) = \frac{|x|^a}{a}$, em que $a > 1$, claramente é uma função de Young. Então, para $b > 1$ tal que $\frac{1}{b} + \frac{1}{a} = 1$ a função $\Phi^*(y) = \frac{|y|^b}{b}$ é a complementar de Φ.

Um conceito importante para a teoria de espaços de Orlicz é o conceito de funções de Young equivalentes.

Definição 2.4. Dizemos que duas funções de Young Φ_1 a Φ_2 são ditas equivalentes se existe $x_0 > 0$, e duas constantes positivas $c_1 < c_2$ tal que , $\forall x \geq x_0$,

$$\Phi_1(c_1 x) \leq \Phi_2(x) \leq \Phi(c_2 x).$$

Proposição 2.5 ((RAO and REN, 1991, Poposição 2))**.** *Se os pares de funções de Young (Φ_1, Φ_1^*) e (Φ_2, Φ_2^*) são de tal forma que Φ_i^* é a função complementar a Φ_i, $i = 1, 2$ e $\Phi_1(x) \leq \Phi_2(x)$, para $0 \leq x_0 \leq x$. Então $\Phi_2^*(y) \leq \Phi_1^*(y)$ para todo y, $0 \leq y_0 \leq y$, em que $y_0 = \Phi_2'(x_0)$, a derivada de Φ_2 em $x = x_0$.*

Como consequência da Proposição 2.5 se duas funções de Young são equivalentes, suas funções complementares também o são. Vimos isso no exemplo a seguir.

Exemplo 2.6. A função $\Phi_1(x) = \cosh(x)$-1 é uma função de Young, pois $\Phi(0) = 0$,

$\Phi(-x) = \Phi(x)$ e $\lim_{x\to\infty}\Phi(x) = \infty$. A função $\Phi_2(x) = e^{|x|} - |x| - 1$, é também uma função de Young e temos que Φ_1 e Φ_2 são equivalentes. A função complementar de Φ_1 é a função de Young $\Phi_1^*(y) = \int_0^y \sinh^{-1}(t)dt$ que por sua vez é equivalente a função de Young $\Phi_2^*(y) = (1+|y|)\log(1+|y|) - |y|$, que é a complementar da função Φ_2.

Agora, lembraremos um pouco sobre espaços de Orlicz (RAO and REN, 1991, Capítulo 3). Sejam Φ uma função de Young. O conjunto

$$L^{\Phi}(\mu) = \left\{ f \in L^0;\ \int_T \Phi(\alpha f) d\mu < \infty,\ \text{para algum } \alpha > 0 \right\} \tag{2.1}$$

é o espaço de Orlicz relativo a função de Young Φ. Espaços de Orlicz podem ser vistos como generalizações dos espaços de Lebesgue, uma vez que para as funções de Young do tipo $\Phi(x) = \frac{|x|^a}{a}$, em que $a > 1$ os espaços de Orlicz L^{Φ} são os espaços de Lebesgue L^a (BOTELHO, PELLEGRINO, and TEIXEIRA, 2015; BREZIS, 2011). O espaço de Orlicz (2.1) é um espaço vetorial (RAO and REN, 1991, Proposição 6) e é um espaço de Banach com a norma de Luxemburgo (RAO and REN, 1991, Teorema 3)

$$N_{\Phi}(f) = \inf\left\{ k > 0 :\ \int_T \Phi\left(\frac{f}{k}\right) d\mu \leq 1 \right\}$$

e com a norma de Orlicz

$$\|f\|_{\Phi} = \sup\left\{ \int_T |fg| d\mu;\ \int_T \Phi^*\left(|g|\right) d\mu \leq 1 \right\},$$

em que Φ^* é a função complementar de Φ. Além disso, as normas são equivalentes (RAO and REN, 1991, Proposição 4).

Na próxima seção, vamos lembrar a construção das famílias de exponenciais, como foi feito em (PISTONE and SEMPI, 1995) e estudado posteriormente em (GIBILISCO and PISTONE, 1998; PISTONE and ROGANTIN, 1999; CENA and PISTONE, 2007; PISTONE, 2009; SANTACROCE, SIRI, and TRIVELLATO, 2016).

2.2 Famílias de exponenciais

Em (PISTONE and SEMPI, 1995) o espaço de Orlicz relativo a função de Orlicz $\Phi(u) = \cosh x - 1$ foi utilizado para a construção da estrutura diferenciável para a variedade de medidas de probabilidades equivalentes $\mathcal{P}_{\mu}$. O espaço $L^{\Phi}(p)$ é equivalente ao conjunto de todas as funções $u \in L^0$cuja função geradora de momento $\hat{u}_p(\lambda) = E_p\left[e^{\lambda u}\right]$é finita em uma vizinhança de 0, em que $E_p[x] = \int_T xpd\mu$ é a esperança de x. Para cada função $u \in L^0$ definimos o funcional gerador de momento

$$M_p(u) = E_p\left[e^u\right],$$

e o funcional gerador de cumulante

$$K_p(u) = \log M_p(u). \tag{2.2}$$

O funcional gerador de cumulante (2.2) é convexo, semicontínuo inferiormente, infinitamente Gâteaux-diferenciável no interior do seu domínio próprio (PISTONE and SEMPI, 1995, Proposição 2.5). Denotemos por $\mathcal{K}_p$ o interior do conjunto de todas as funções $u \in L^\Phi(p)$ cujo funcional gerador de momento $M_p(u)$ é finito. Equivalentemente, a função $u \in L^\Phi(p)$ pertence a $\mathcal{K}_p$ se e somente se $M_p(\lambda u)$ é finito para todo λ em alguma vizinhança de $[0, 1]$. O subespaço fechado de p-centrada variáveis aleatórias

$$B_p = \left\{ u \in L^\Phi(p) : E_p[u] = 0 \right\},$$

é tomado sendo o espaço de Banach coordenado. O atlas da variedade é dado pela aplicação (parametrização exponencial) $e_p : \mathcal{B}_p \to \mathcal{E}_p$, que aplica o conjunto $\mathcal{B}_p = B_p \cap \mathcal{K}_p$ à família exponencial $\mathcal{E}_p = e_p(\mathcal{B}_p) \subset \mathcal{P}_\mu$, de acordo com

$$e_p(u) = e^{u - K_p(u)} p \in \mathcal{P}_\mu, \text{ para todo } u \in \mathcal{B}_p.$$

A aplicação e_p é uma bijeção de $\mathcal{B}_p$ para sua imagem $\mathcal{E}_p = e_p(\mathcal{B}_p)$, cuja aplicação inversa $e_p^{-1} : \mathcal{E}_p \to \mathcal{B}_p$ pode ser expressada como

$$e_p^{-1}(q) = \log\left(\frac{p}{q}\right) - E_p\left[\log\left(\frac{q}{p}\right)\right], \text{ para } q \in \mathcal{E}_p,$$

e a aplicação de transição $e_q^{-1} \circ e_p : \mathcal{B}_p \to \mathcal{B}_q$, pode ser escrita como

$$e_q^{-1} \circ e_p(u) = u + \log\left(\frac{p}{q}\right) - E_p\left[u + \log\left(\frac{q}{p}\right)\right], \text{ para todo } u \in \mathcal{B}_p.$$

Desta forma, $\mathcal{P}_\mu$ é dotado por uma estrutura de C^∞-variedade de Banach, e as famílias exponenciais $\mathcal{E}_p = e_p(\mathcal{B}_p)$ são os abertos cujo a união destes forma $\mathcal{P}_\mu$. Encontrar formas mais gerais para as famílias de exponenciais tem sido objeto de pesquisa. Em (PISTONE, 2009), um modelo geométrico foi encontrado utilizando a κ-exponencial. O mesmo ocorre em LOIZA and QUICENO (2013), mas desta vez a função utilizada foi a q-exponencial.

Na próxima seção, veremos como essas famílias exponenciais foram generalizadas a partir da substituição da função exponencial, por uma função exponencial deformada.

3 ESPAÇOS DE MUSIELAK-ORLICZ E φ-FAMÍLIAS DE DISTRIBUIÇÕES DE PROBABILIDADE

Em (VIGELIS, 2011) φ-famílias de distribuições de probabilidades foram construídas, usando subconjuntos dos espaços de Musielak-Orlicz como conjuntos de coordenadas. Nesta seção vamos relembrar algumas propriedades dos espaços de Musielak-Orlicz e a construção das φ-famílias.

3.1 Espaços de Musielak-Orlicz

Seja (T, Σ, μ) um espaço de medida. Seja $X \subseteq T$, dizemos que uma determinada propriedade vale μ-q.t.p. em X, isto é em quase toda parte em X em relação a medida μ, se existe um conjunto $A \subseteq X$, em que $\mu(A) = 0$, tal que a propriedade em questão vale para todo $x \in X - A$. Entendemos assim, que o conjunto onde a propriedade não vale tem medida μ nula. Podemos dizer também que a propriedade vale para quase todo ponto de X. Quando $X = T$, então dizemos que a propriedade vale μ-q.t.p.. Uma função $\Phi : T \times [0, \infty]$ é dita ser uma função de *Musielak-Orlicz* se as seguintes condições forem satisfeitas (MUSIELAK, 1983, Definição 7.1):
(i) $\Phi(t, \cdot)$ é convexa e semicontínua inferiormente para μ-q.t.p. $t \in T$;
(ii) $\Phi(t, 0) = \lim_{u \downarrow 0} \Phi(t, u) = 0$ e $\Phi(t, \infty) = \lim_{u \uparrow \infty} \Phi(t, u) = \infty$ para μ-q.t.p. $t \in T$, e;
(iii) $\Phi(\cdot, u)$ é mensurável para todo $u \geq 0$.
Uma função de Musielak-Orlicz Φ é dita ser uma *função de Orlicz* se $\Phi(t, u) = \Phi(u)$ é independente da variável t . Pela continuidade de $\Phi(t, \cdot)$, a função $\Phi(t, \cdot)$ não é igual a 0 ou ∞ no intervalo $(0, \infty)$. Equivalentemente, denotando

$$a_\Phi(t) = \sup\{u \geq 0 : \Phi(t, u) = 0\}, \text{ e} \tag{3.1}$$

$$b_\Phi(t) = \sup\{u \geq 0 : \Phi(t, u) < \infty\}, \tag{3.2}$$

então $a_\Phi(t) < \infty$ e $b_\Phi(t) > 0$. Pela convexidade de $\Phi(t, \cdot)$, a função $\Phi(t, \cdot)$ é contínua para $u \in (0, b_\Phi(t))$.

A *função complementar* $\Phi^* : T \times [0, \infty] \to [0, \infty]$ à função de Musielak-Orlicz Φ é definida como

$$\Phi^*(t, v) = \sup_{u \geq 0}(uv - \Phi(t, u)), \quad \text{para todo } v \geq 0, \tag{3.3}$$

i. e., $\Phi^*(t, \cdot)$ é a Fenchel-conjugada de $\Phi(t, \cdot)$. A função complementar Φ^* é também uma função de Musielak-Orlicz, ou seja satisfaz as condições (i)-(iii) na definição de função de Musielak-Orlicz (VIGELIS, 2011). A função complementar de Φ^* é a função de Musielak-

Orlicz Φ. De fato, segue de (ROCKAFELLAR, 1970, Teorema 12.2) que uma função própria é igual a sua biconjugada (Fenchel-conjugada da Fenchel-conjugada) se e somente se a função é convexa e semicontínua inferiormente, o que segue da definição de função de Musielak-Orlicz. Em outras palavras:

$$\Phi(t,u) = \sup_{v\geq 0}(uv - \Phi^*(t,v)), \quad \text{para todo } u \geq 0.$$

Seja L^0 o espaço de todas as funções reais mensuráveis em T. Dada uma função de Musielak-Orlicz Φ, denotamos o funcional $I_\Phi(u) = \int_T \Phi(t,|u(t)|)d\mu$, para qualquer $u \in L^0$. O *espaço de Musielak-Orlicz*, a *classe de Musielak-Orlicz* e o *espaço Morse-Transue*, são definidos, respectivamente por :

$$L^\Phi = \{u \in L^0 : \ I_\Phi(\lambda u) < \infty \text{ para algum } \lambda > 0\},$$

$$\tilde{L}^\Phi = \{u \in L^0 : \ I_\Phi(u) < \infty\}$$

e

$$E^\Phi = \{u \in L^0 : I_\Phi(\lambda u) < \infty \text{ para todo } \lambda > 0\}.$$

Claramente, $E^\Phi \subset \tilde{L}^\Phi \subset L^\Phi$. O espaço de Musielak-Orlicz é um espaço de Banach quando é equipado com a *norma de Luxemburgo* dada por (MUSIELAK, 1983, Teorema 7.7)

$$\|u\|_\Phi = \inf \left\{\lambda > 0 : \ I_\Phi\left(\frac{u}{\lambda}\right) \leq 1\right\},$$

ou com a *norma de Orlicz*, representada por

$$\|u\|_{\Phi,0} = \sup\left\{\left|\int_T uv d\mu\right| \ : \ v \in \tilde{L}^{\Phi^*} \text{ e } I_{\Phi^*}(v) \leq 1\right\}.$$

Estas normas são equivalentes e as desigualdades $\|u\|_\Phi \leq \|u\|_{\Phi,0} \leq 2\|u\|_\Phi$ acontecem para todo $u \in L^\Phi$(MUSIELAK, 1983).

Definição 3.1. Sejam Φ e Ψ funções de Musielak-Orlicz. Escrevemos $\Psi \preceq \Phi$ ou $\Phi \succeq \Psi$ se existem constantes α, $\lambda > 0$, e uma função não-negativa $f \in \tilde{L}^\Phi$ para as quais a desigualdade

$$\alpha\Psi(t,u) \leq \Phi(t,\lambda u), \quad \text{para todo } u > f(t), \tag{3.4}$$

é satisfeita. Usaremos a notação $\Phi \sim \Psi$ para denotar que $\Psi \preceq \Phi$ e $\Psi \succeq \Phi$.

Queremos provar que $\Psi \preceq \Phi$ se e somente se $\Phi^* \preceq \Psi^*$, em que Φ^* e Ψ^* são as funções complementares de Φ e Ψ, respectivamente.

Proposição 3.2 ((VIGELIS, 2011, Proposição 2.3))**.** *Sejam Φ e Ψ funções de Musielak-Orlicz. Supondo que, para constantes α, $\lambda > 0$, existe uma função integrável $h : T \to$*

$[0, \infty)$ *tal que*

$$\alpha\Psi(t,u) \le \Phi(t,\lambda u) + h(t), \quad \textit{para todo } u{\ge}0. \tag{3.5}$$

Então, para constantes $\alpha^{'} \in (0,\alpha)$ *e* $\lambda^{'} = \lambda$, *ou* $\alpha^{'} = \alpha$ *e* $\lambda^{'} > \lambda$, *uma função não negativa* $f \in \tilde{L}^{\Psi}$ *pode ser encontrada tal que*

$$\alpha^{'}\Psi(t,u) \le \Phi(t,\lambda^{'}u), \quad \textit{para todo } u > f(t). \tag{3.6}$$

Demonstração. Seja $\Psi^{-1}(t,\cdot)$ denote a inversa contínua à esquerda de $\Psi(t,\cdot)$, a qual pode ser expressada por

$$\Psi^{-1}(t,v) = \begin{cases} a_{\Psi}(t), & \text{para } v = 0 \\ \Psi^{-1}(t,v), & \text{para } 0 < v \le \Psi(t, b_{\Psi}(t)), \\ b_{\Psi}(t), & \text{para } \Psi(t,b_{\Psi}(t)) < v. \end{cases}$$

A função $\Psi^{-1}(t,\cdot)$ satisfaz a inequação $\Psi(t,\Psi^{-1}(t,v)) \le v$ e $\Psi(t,u) \ge v$, para todo $v \ge 0$, e qualquer $u > \Psi^{-1}(t,v)$. Fixado qualquer $\alpha' \in (0,\alpha)$, definimos $f(t) = \Psi^{-1}(t, h(t)/(\alpha - \alpha'))$. Claramente, $f \in \tilde{L}^{\Psi}$. Tendo em vista (3.5), temos $\alpha'\Psi(t,u) = \Phi(t,\lambda u) = \infty$ para qualquer $u > b_{\Psi}(t)$. Desde que $h(t) \le (\alpha - \alpha')\Psi(t,u) < \infty$, para qualquer $u \in (f(t), b_{\Psi}(t))$, podemos escrever

$$\alpha'\Psi(t,u) \le \Phi(t,\lambda u) + h(t) - (\alpha - \alpha')\Psi(t,u) \le \Phi(t,\lambda u). \tag{3.7}$$

Pela continuidade à esquerda de $\Psi(t,\cdot)$ e $\Phi(t,\cdot)$, a inequação $\alpha'\Psi(t,u) \le \Phi(t,\lambda' u)$ é satisfeita para $u \in (f(t), b_{\Psi}(t)]$. Consequentemente (3.6) segue para qualquer $\alpha' \in (0,\alpha)$ e $\lambda' = \lambda$.

Agora seja $\alpha' = \alpha$ e $\lambda' > \lambda$. Pelo argumento usado anteriormente, podemos encontrar uma função não negativa $f \in \tilde{L}^{\Psi}$ tal que

$$\frac{\lambda}{\lambda'}\alpha\Psi(t,u) \le \Phi(t,\lambda u), \quad \text{para todo } u > f(t).$$

Então, para qualquer $u > f(t)$, podemos escrever

$$\alpha\Psi(t,u) = \frac{\lambda'}{\lambda}\left(\frac{\lambda}{\lambda'}\alpha\right)\Psi(t,u) \le \frac{\lambda'}{\lambda}\Phi(t,\lambda u) \le \Phi(t,\lambda' u),$$

o que termina a prova. □

Lema 3.3 ((VIGELIS, 2011, Lema 2.5))**.** *Sejam* Φ^* *e* Ψ^* *as funções complementares das funções de Musielak-Orlicz* Φ *e* Ψ, *respectivamente. Supondo que, para constantes* α,

$\lambda > 0$, *existe uma função não negativa* $f \in \tilde{L}^{\Psi}$ *tal que*

$$\alpha\Psi(t,u) \leq \Phi(t,\lambda u), \quad \textit{para todo } u > f(t). \tag{3.8}$$

Então, para constantes $\alpha^{'} = \frac{1}{\alpha}$ *e* $\lambda^{'} > \frac{\lambda}{\alpha}$, *ou* $\alpha^{'} \in (0, \frac{1}{\alpha})$ *e* $\lambda^{'} = \frac{\lambda}{\alpha}$, *uma função não negativa* $g \in \tilde{L}^{\Phi^*}$ *pode ser encontrada tal que*

$$\alpha^{'}\Phi^*(t,v) \leq \Psi^*(t,\lambda^{'}v), \quad \textit{para todo } v > g(t). \tag{3.9}$$

Demonstração. Definindo a função $h(t) = \Psi(t, f(t))$, podemos escrever

$$\alpha\Psi(t,u) \leq \Phi(t,\lambda u) + \alpha h(t), \quad \text{para todo } u \geq 0.$$

Calculando a Fenchel-conjugada das funções na inequação acima, obtemos

$$\frac{1}{\alpha}\Phi^*(t,v) \leq \Psi^*\left(t, \frac{\lambda}{\alpha}v\right) + h(t), \quad \text{para todo } v \geq 0.$$

A partir da Proposição 3.2, inferimos que (3.9) é satisfeita. □

Provamos assim, com o Lema 3.3 o seguinte corolário.

Corolário 3.4. *Sejam* Ψ *e* Φ *funções de Musielak-Orlicz e* Ψ^* *e* Φ^* *suas complementares. Então*$\Psi \preceq \Phi$ *se e somente se* $\Phi^* \preceq \Psi^*$.

Agora, vamos falar sobre duas condições importantes na teoria dos espaços de Musielak-Orlicz, a Δ_2-condição e a ∇_2-condição.

Definição 3.5. Uma função de Musielak-Orlicz é dita satisfazer a Δ_2-condição, ou pertencer a Δ_2-classe ($\Phi \in \Delta_2$) se podemos encontrar uma função não-negativa $f \in \tilde{L}^{\Phi}$ e uma constante $\alpha > 0$ tal que

$$\alpha\Phi(t,u) \leq \Phi\Big(t, \frac{1}{2}u\Big), \qquad \text{para todo } u \geq f(t), \quad \text{e } \mu\text{-q.t.p. } t \in T. \tag{3.10}$$

Quando a função de Musielak-Orlicz Φ satisfaz a Δ_2-condição, então $I_{\Phi}(u) < \infty$ para cada $u \in L^{\Phi}$ (MUSIELAK, 1983). Neste caso E^{Φ}, $\tilde{L}^{\Phi}$ e L^{Φ} são iguais como conjuntos. Além disso, se a função de Musielak-Orlicz não satisfaz a Δ_2-condição, então o conjunto E^{Φ} é um subconjunto próprio de L^{Φ}.

Proposição 3.6. *Toda função de Musielak-Orlicz* Φ *que satisfaz a* Δ_2*-condição é de valor finito.*

Demonstração. Supondo que a função Φ não é valor finito, ou seja, que $b_{\Phi}(t) < \infty$ obtemos $\infty = \alpha\Phi(t,u) > \Phi(t, \frac{1}{2}u)$ para todo $b_{\Phi}(t) < u < 2b_{\Phi}(t)$, o que implica que Φ não satisfaz a Δ_2-condição. □

Definição 3.7. Uma função de Musielak-Orlicz é dita satisfazer a ∇_2-condição, ou pertencer a ∇_2-classe ($\Phi \in \nabla_2$) se podemos encontrar uma constante $\gamma > 1$ e uma função não-negativa $f \in \tilde{L}^\Phi$ tal que

$$2\gamma\Phi(t,u) \leq \Phi\Big(t, \gamma u\Big), \qquad \text{para todo } u > f(t). \tag{3.11}$$

Proposição 3.8. *Seja Φ uma função de Musielak-Orlicz que satisfaz a ∇_2-condição, então $\lim_{u\to\infty} \frac{\Phi(t,u)}{u} = \infty$.*

Demonstração. Reescrevendo (3.11) como

$$\frac{\Phi(t,u)}{u} \leq \frac{1}{2}\frac{\Phi\,(t,\gamma u)}{\gamma u}, \qquad \text{para todo } u > f(t),$$

concluímos que

$$\lim_{u\to\infty} \frac{\Phi(t,u)}{u} \leq \frac{1}{2} \lim_{u\to\infty} \frac{\Phi(t,u)}{u}.$$

Consequentemente $\lim_{u\to\infty} \frac{\Phi(t,u)}{u} = \infty$. Além disso temos,

$$\lim_{u\to\infty} \frac{\Phi(t,u)}{u} = \lim_{u\to\infty} \Phi'_-(t,u) = \lim_{u\to\infty} \Phi'_+(t,u) = \infty.$$

□

Podemos dizer que a função de Musielak-Orlicz Φ que satisfaz a Δ_2-condição, não aumenta mais rapidamente que funções exponenciais. Como podemos ver na Figura 3.1, as funções $\Phi_1(u) = (1+u)\ln(1+u) - u$ e $\Phi_2(u) = \frac{u^2}{2}$, satisfazem a Δ_2-condição e não aumentam mais rapidamente que a função $\Phi_3(u) = \exp(u) - 1$, que por sua vez não satisfaz a Δ_2-condição.

Agora vamos entender o comportamento de uma função de Musielak-Orlicz Φ que satisfaz a ∇_2-condição. Para entendermos, precisamos de um resultado que envolve a função de Musielak-Orlicz Φ, sua complementar Φ^* e as duas condições. Para provar esse resultado faremos uso do seguinte lema.

Lema 3.9 ((VIGELIS, 2011, Lema 2.9))**.** *A Δ_2-condição é equivalente à afirmação que, para cada $\lambda \in (0,1)$, existe uma constante $\alpha_\lambda \in (0,1)$, e uma função não negativa $f_\lambda \in \tilde{L}^\Phi$ tal que*

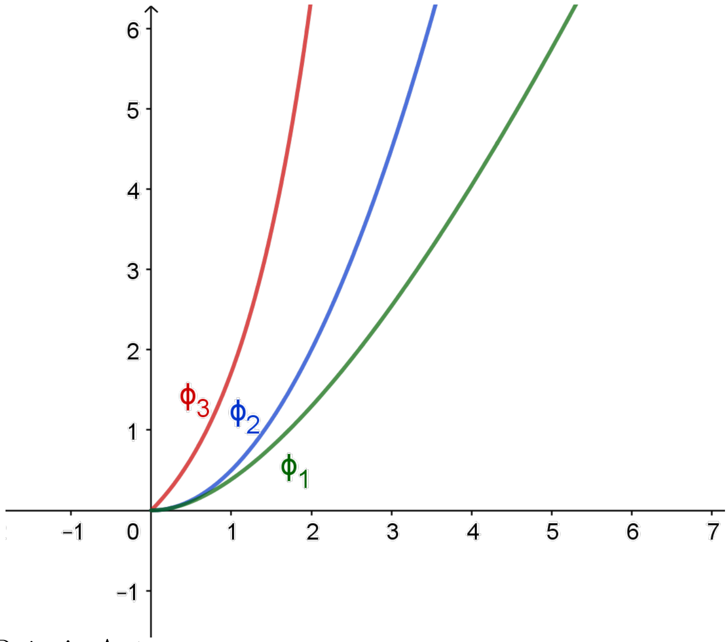
$$\alpha_\lambda\Phi(t,u) \leq \Phi(t,\lambda u), \quad \textit{para todo } u > f_\lambda(t). \tag{3.12}$$

A ∇_2-condição é equivalente à afirmação que, para qualquer $\lambda \in (0,1)$, existe uma constante $\gamma_\lambda > 1$, e uma função não negativa $f_\lambda \in \tilde{L}^\Phi$ tal que

$$\gamma_\lambda\Phi(t,u) \leq \Phi(t,\lambda_{\gamma_\lambda} u), \quad \textit{para todo } u > f_\lambda(t). \tag{3.13}$$

Demonstração. Suponha que (3.10) ocorre. Se o número natural $n \geq 1$ é tal que $2^{-n} \leq \lambda$, então $\alpha^n\Phi(t,u) \leq \Phi(t, 2^{-n}u) \leq \Phi(t,\lambda u)$, para todo $u > 2^{n-1}f(t)$. Por outro lado, se Φ

Figura 1 – Comportamento das funções em relação a Δ_2-condição.



Fonte: Próprio Autor.

satisfaz (3.12) e o número natural $n \geq 1$ é escolhido tal que $\lambda^n \leq \frac{1}{2}$, então $\alpha_\lambda^n \Phi(t,u) \leq \Phi(t, \lambda^n u) \leq \Phi(t, \frac{1}{2}u)$, para todo $u > \lambda^{-n+1} f_\lambda(t)$.

Assumindo que (3.11) é satisfeita. Seja $n \geq 1$ um número natural tal que $2^{-n} \leq \lambda$. Então $\gamma^n \Phi(t,u) \leq \Phi(t, 2^{-n}\gamma^n u) \leq \Phi(t, \lambda_{\gamma^n} u)$, para todo $u > f(t)$. Reciprocamente, se (3.13) ocorre e o número natural $n \geq 1$ é escolhido tal que $\lambda^n \leq \frac{1}{2}$, então $\gamma_\lambda^n \Phi(t,u) \leq \Phi(t, \lambda^n \gamma_\lambda^n u) \leq \Phi(t, \frac{1}{2}\gamma_\lambda^n u)$ para todo $u > f(t)$. □

Chegamos ao resultado que queremos.

Teorema 3.10. *Uma função de Musielak-Orlicz Φ satisfaz a Δ_2-condição se, e somente se, sua função complementar Φ^* satisfaz a ∇_2-condição.*

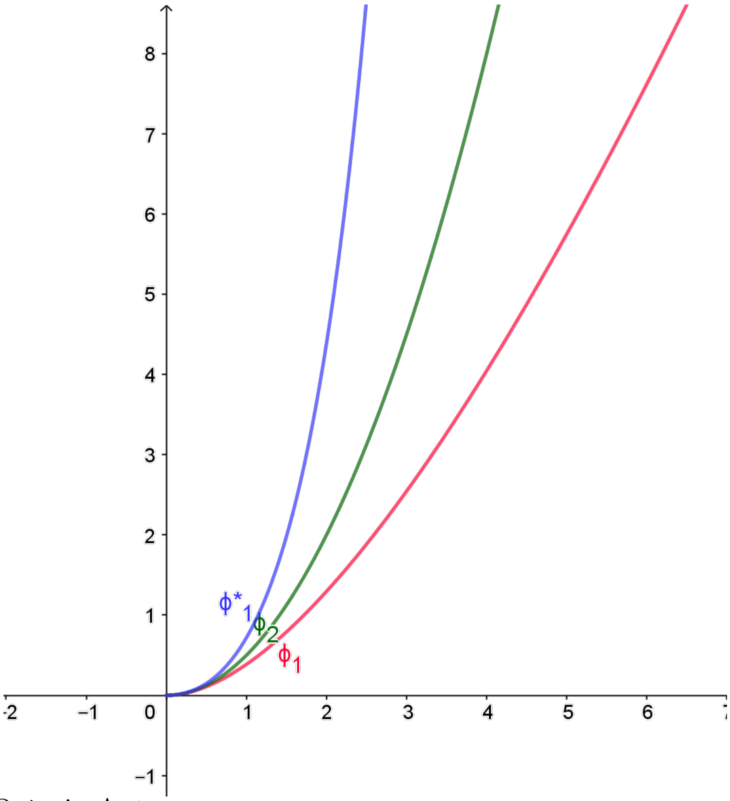
A demonstração do teorema segue dos Lemas 3.3 e 3.9.

Assim uma função de Musielak-Orlicz Φ^* satisfaz a ∇_2-condição se sua função complementar Φ satisfaz a Δ_2-condição. Temos portanto, que as funções$\Phi_1^*(u) = \exp(u) - u - 1$ e $\Phi_2^*(u) = \frac{u^2}{2}$, satisfazem a ∇_2-condição, uma vez que suas complementares $\Phi_1(u) = (1+u)\ln(1+u) - u$ e $\Phi_2(u) = \frac{u^2}{2}$, respectivamente, satisfazem a Δ_2-condição. Podemos observar na Figura 3.2 o comportamento das funções e temos que a função $\Phi_1(u) = (1+u)\ln(1+u) - u$, não satisfaz a ∇_2-condição.

Vemos assim que existem funções como, por exemplo, $\Phi_2(u) = \frac{u^2}{2}$ que satisfaz ambas as condições e existem funções como $\Phi_1^*(u) = \exp(x) - x - 1$, que satisfaz a ∇_2-condição e não satisfaz a Δ_2-condição. Funções como $\Phi_1^*(u) = \exp(x) - x - 1$, serão importantes para o nosso estudo.

O espaço dual de L^Φ, é denotado por $\left(L^\Phi\right)^*$ e é representado da seguinte forma (MUSIELAK, 1983; HUDZIK and ZBASZYNIAK, 1997; VIGELIS and CAVALCANTE,

Figura 2 – Comportamento das funções em relação a ∇_2-condição.



Fonte: Próprio Autor.

2014):

$$\left(L^{\Phi}\right)^{*} = L^{\Phi^{*}} \oplus \left(L^{\Phi}\right)^{\sim}_{s},$$

em que L^{Φ^*} é o conjunto dos funcionais contínuos em ordem e $\left(L^{\Phi}\right)^{\sim}_{s}$ é o conjunto formado pelas componentes singulares. Se a função de Musielak-Orlicz $\Phi \in \Delta_2$ então todo funcional em $\left(L^{\Phi}\right)^{*}$ é contínuo em ordem e é representado por

$$f_{v^*}(u) := \int_T uv^* d\mu, \quad \text{for all } u \in L^{\Phi}, \tag{3.14}$$

caso contrário, se $\Phi \notin \Delta_2$ podem existir funcionais f em $\left(L^{\Phi}\right)^{*}$ que podem ser unicamente expressados como

$$f = f_c + f_s, \tag{3.15}$$

no qual f_c é a componente contínua em ordem e f_sé a componente singular.

Outro resultado importante para este trabalho é o resultado sobre mergulhos entre espaços de Musielak-Orlicz e entre classes de Musielak-Orlicz. Se faz importante uma lema antes da proposição que trata dos mergulhos.

Lema 3.11 ((MUSIELAK, 1983, Lema 8.3))**.** *Considere uma medida não atômica e σ-finita μ. Se $\{u_n\}$ é uma sequência de funções mensuráveis, não negativas e de valor finito,* *e $\{\alpha_n\}$ é uma sequência de números reais positivos tais que*

$$\int_T u_n d\mu \geq 2^n \alpha_n, \qquad \textit{para todo } n \geq 1,$$

então podemos encontrar, uma sequência crescente de números naturais $\{n_i\}$ e uma sequência $\{A_i\}$ de conjuntos mensuráveis disjuntos dois a dois, tais que

$$\int_{A_i} u_{n_i} d\mu = \alpha_{n_i}, \qquad \textit{para todo } i \geq 1.$$

Proposição 3.12 ((MUSIELAK, 1983, Teorema 8.4))**.** *Seja Φ e Ψ duas funções de Musielak-Orlicz. Então $\tilde{L}^\Phi \subseteq \tilde{L}^\Psi$se, e somente se, existe uma constante $\alpha > 0$ e uma função não negativa $f \in \tilde{L}^\Psi$ tal que*

$$\alpha\Psi(t,u) \leq \Phi(t,u), \quad \textit{para todo } u > f(t).$$

Além disso, $L^\Phi \subset L^\Psi$ se, e somente se, existem constantes $\alpha, \lambda > 0$ e uma função não negativa $f \in \tilde{L}^\Psi$ tal que

$$\alpha\Psi(t,u) \leq \Phi(t,\lambda u), \quad \textit{para todo } u > f(t).$$

Na próxima seção, lembraremos como foram construídas as φ-famílias de distribuições de probabilidade, que são generalizações das famílias de exponenciais, substituindo a função exponencial por uma exponencial deformada.

3.2 φ-famílias de distribuições de probabilidades

As φ-famílias de distribuições de probabilidade foram propostas em (VIGELIS and CAVALCANTE, 2013b). Enquanto as famílias exponenciais são baseadas na função exponencial, as φ-famílias são baseadas em funções exponenciais deformadas.

Definição 3.13. Uma *exponencial deformada* é uma função $\varphi : \mathbb{R} \to (0,\infty)$ que satisfaz as seguintes propriedades:c

(a1) $\varphi(\cdot)$ é convexa,

(a2) $\lim_{u\to-\infty} \varphi(u) = 0$ e $\lim_{u\to\infty} \varphi(u) = \infty$,

(a3) Existe uma função mensurável $u_0 : T \to (0,\infty)$ tal que

$$\int_T \varphi(c + \lambda u_0) d\mu < \infty, \quad \text{para todo } \lambda > 0,$$

para uma função mensurável $c : T \to \mathbb{R}$ tal que $\int_T \varphi(c) d\mu = 1$.

Em (SOUZA, VIGELIS, and CAVALCANTE, 2016, Lema 1), foi mostrado que a restrição $\int_T \varphi(c) d\mu = 1$ pode ser trocada por $\int_T \varphi(c) d\mu < \infty$. Assim, a condição (a3) pode ser reescrita como:

(a3') Existe uma função mensurável $u_0 : T \to (0, \infty)$ tal que

$$\int_T \varphi(c + \lambda u_0) d\mu < \infty, \quad \text{para todo } \lambda > 0,$$

para uma função mensurável $c : T \to \mathbb{R}$ tal que $\int_T \varphi(c) d\mu < \infty$.
Assim, (a3) e (a3') são equivalentes.

Existem muitos exemplos de funções que satisfazem (a1) - (a3), ou seja, de funções exponenciais deformadas. Um exemplo relevante é a função exponencial $\varphi(x) = \exp(x)$ que satisfaz (a1) - (a3) para $u_0 = \mathbf{1}_T$. Outro exemplo é a κ-exponencial de *Kaniadakis* (KANIADAKIS, 2001).

Exemplo 3.14. A κ-exponencial de Kaniadakis $\exp_k : \mathbb{R} \to (0, \infty)$ para $\kappa \in [-1, 1]$ é definida como

$$\exp_\kappa(u) = \begin{cases} \left(\kappa u + \sqrt{1 + \kappa^2 u^2}\right)^{\frac{1}{\kappa}}, & \text{if } \kappa \neq 0, \\ \exp(u) & \text{if } \kappa = 0. \end{cases}$$

A inversa de $\exp_\kappa$é o κ-logaritmo de Kaniadakis, dada por

$$\ln_\kappa(u) = \begin{cases} \frac{u^\kappa - u^{-\kappa}}{2\kappa}, & \text{if } \kappa \neq 0, \\ \ln(u) & \text{if } \kappa = 0. \end{cases}$$

Pode ser facilmente notado que a κ-exponencial satisfaz (a1) - (a3) (SOUZA, VIGELIS, and CAVALCANTE, 2016; VIGELIS and CAVALCANTE, 2013b).

Uma razão para a escolha da definição da exponencial deformada satisfazendo (a1)-(a3) é que existem também funções que satisfazem (a1) e (a2), mas não satisfazem (a3'). Um exemplo foi dado em (SOUZA, VIGELIS, and CAVALCANTE, 2016, Exemplo 2):

$$\varphi(u) = \begin{cases} e^{(u+1)^2/2}, & u \geq 0, \\ e^{(u+1/2)}, & u \leq 0. \end{cases} \tag{3.16}$$

Claramente $\lim_{u \to -\infty} \varphi(u) = 0$ e $\lim_{u \to \infty} \varphi(u) = \infty$. Foi mostrado em (SOUZA, VIGELIS, and CAVALCANTE, 2016) que existem funções $c : T \to \mathbb{R}$ e $u_0 : T \to (0, \infty)$ tal que $\int_T \varphi(c) d\mu < \infty$, mas $\int_T \varphi(c + u_0) d\mu = \infty$. A função (3.16) é uma exponencial deformada como discutida em (NAUDTS, 2011).

O fato de a função φ satisfazer a condição (a3) e (a3') é o ponto chave para a nossa pesquisa. Dizer que uma função φ não satisfaz a condição (a3) ou (a3'), é dizer que para uma função $c : T \to \mathbb{R}$ existe uma função $u_0 : T \to (0, \infty)$ tal que

$$\int_T \varphi(c + \lambda u_0) d\mu < \infty, \quad \text{para todo } \lambda > 0,$$

para uma outra função $\tilde{c} : T \to (0, \infty)$ para a mesma função u_0, podemos ter

$$\int_T \varphi(\tilde{c} + \lambda u_0) d\mu = \infty, \quad \text{para algum } \lambda > 0.$$

Definindo a função de Musielak-Orlicz a partir da exponencial deformada φ

$$\Phi_c(t, u) = \varphi(t, c(t) + u) - \varphi(t, c(t)), \tag{3.17}$$

para uma função mensurável $c : T \to \mathbb{R}$ tal que $\varphi(c(t))$ é μ-integrável, foi definida em (VIGELIS and CAVALCANTE, 2013b). Assim, os conjuntos L^{Φ_c}, $\tilde{L}^{\Phi_c}$ e E^{Φ_c} são denotados, respectivamente, por L_c^φ, $\tilde{L}_c^\varphi$ e E_c^φ, quando a função Φ_c é dada por (3.17). Desde que $\varphi(t, c(t))$ é μ-integrável, o espaço de Musielak-Orlicz L_c^φ corresponde ao conjunto de todas as funções $u \in L^0$ para as quais $\varphi(t, c(t) + \lambda u(t))$ é μ-integrável para todo λ contido em alguma vizinhança de 0.

Seja $\mathcal{K}_c^\varphi$ o conjunto de todas as funções tal que $\varphi(t, c(t) + \lambda u(t))$ é μ-integrável para cada λ em uma vizinhança de $[0, 1]$. In (VIGELIS and CAVALCANTE, 2013b, Lema 2) foi provado que $\mathcal{K}_c^\varphi$ é um subconjunto aberto de L_c^φ. Para $u \in \mathcal{K}_c^\varphi$ a função $\varphi(c + u)$ não necessariamente está em $\mathcal{P}_\mu$. A *função normalizadora* $\psi : \mathcal{K}_c^\varphi \to \mathbb{R}$ é introduzida com a finalidade de fazer a densidade

$$\varphi(c + u - \psi(u)u_0), \tag{3.18}$$

estar contida em $\mathcal{P}_\mu$, para qualquer $u \in \mathcal{K}_c^\varphi$. Essa função ψ, pode ser vista como uma generalização do funcional gerador de cumulantes(2.2). Para uma função $u \in \mathcal{K}_c^\varphi$, existe um único $\psi(u) \in \mathbb{R}$ para o qual $\varphi(c + u - \psi(u)u_0) \in \mathcal{P}_\mu$ (VIGELIS and CAVALCANTE, 2013b, Proposição 3).

A função $\psi : \mathcal{K}_c^\varphi \to \mathbb{R}$ pode assumir valores positivos e negativos. Então, seja o subespaço fechado

$$B_c^\varphi = \left\{ u \in L_c^\varphi : \int_T u \varphi'_+(c) d\mu = 0 \right\}, \tag{3.19}$$

em que $\varphi'_+(t, \cdot)$ é a derivada à direita de $\varphi(t, \cdot)$, e seja o conjunto $\mathcal{B}_c^\varphi = \mathcal{K}_c^\varphi \cap B_c^\varphi$. Pela convexidade de φ, temos que $\psi(u) \geq 0$, para $u \in \mathcal{B}_c^\varphi$. Assim temos a parametrização $\varphi_c : \mathcal{B}_c^\varphi \to \mathcal{F}_c^\varphi$, em que $\mathcal{F}_c^\varphi = \varphi_c(\mathcal{B}_c^\varphi) \subseteq \mathcal{P}_\mu$ de acordo com (3.18). Claramente, temos $\mathcal{P}_\mu = \bigcup \{\mathcal{F}_c^\varphi : \varphi(c) \in \mathcal{P}_\mu\}$ e tomando o domínio da parametrização no conjunto fechado B_c^φ, temos que a aplicação φ_c é uma bijeção e consequentemente uma densidade $q = \varphi(c + u - \psi(u)u_0)$ é representada de maneira única.

Sejam $\varphi_{c_1} : \mathcal{B}_{c_1}^\varphi \to \mathcal{F}_{c_1}^\varphi$ e $\varphi_{c_1} : \mathcal{B}_{c_1}^\varphi \to \mathcal{F}_{c_1}^\varphi$ parametrizações. A aplicação de transição

$$\varphi_{c_2}^{-1} \circ \varphi_{c_1} : \varphi_{c_1}^{-1} \left(\mathcal{F}_{c_1}^\varphi \cap \mathcal{F}_{c_2}^\varphi \right) \to \varphi_{c_2}^{-1} \left(\mathcal{F}_{c_1}^\varphi \cap \mathcal{F}_{c_2}^\varphi \right)$$

expressada como

$$\varphi_{c_2}^{-1} \circ \varphi_{c_1}(w) = c_1 - c_2 + w - \frac{\int_T (c_1 - c_2 + w)\varphi'_+(c_2)d\mu}{\int_T u_0 \varphi'_+(c_2)d\mu} u_0, \tag{3.20}$$

é de classe C^∞, se mostrarmos que w e $c_1 - c_2$ pertencem a $L^\varphi_{c_2}$ e que os espaços $L^\varphi_{c_1}$ e $L^\varphi_{c_2}$ tem normas equivalentes. A seguir vamos garantir que $c_1 - c_2$ pertence a $L^\varphi_{c_2}$ e que os espaços $L^\varphi_{c_1}$ e $L^\varphi_{c_2}$ são iguais como conjuntos.

Proposição 3.15 ((VIGELIS and CAVALCANTE, 2013b, Proposição 4))**.** *Assumindo que as funções mensuráveis $c_1, c_2 : T \to \mathbb{R}$ satisfazem $\int_T \varphi(t, c_1(t))d\mu < \infty$ e $\int_T \varphi(t, c_2(t))d\mu < \infty$. Então $L^\varphi_{c_1} \subseteq L^\varphi_{c_2}$ se e somente se $c_1 - c_2 \in L^\varphi_{c_2}$.*

Demonstração. Supondo que $c_1 - c_2$ não pertence a $L^\varphi_{c_2}$. Seja $A = \{t \in T : c_1(t) < c_2(t)\}$. Para $\lambda \in [0, 1]$, temos

$$\begin{aligned}\int_T \varphi(c_2 + \lambda(c_1 - c_2))d\mu &= \int_{T\setminus A} \varphi\left(c_2 + \lambda(c_1 - c_2)\right) d\mu + \int_A \varphi\left(c_2 + \lambda(c_1 - c_2)\right) d\mu \\ &\leq \int_{T\setminus A} \varphi\left(c_2 + (c_1 - c_2)\right) d\mu + \int_A \varphi(c_1)d\mu \\ &\leq \int_T \varphi(c_1)d\mu + \int_T \varphi(c_2)d\mu < \infty.\end{aligned}$$

Desde que $c_1 - c_2 \notin L^\varphi_{c_2}$, para qualquer $\lambda > 0$, ocorre $\int_T \varphi(c_2 - \lambda(c_1 - c_2))d\mu = \infty$. A partir de

$$\begin{aligned}\int_T \varphi\left(c_2 - \lambda(c_1 - c_2)\right) d\mu &= \int_{T\setminus A} \varphi\left(c_2 - \lambda(c_1 - c_2)\right) d\mu + \int_A \varphi(c_2 - \lambda(c_1 - c_2))d\mu \\ &\leq \int_A \varphi\left(c_2 + \lambda(c_1 - c_2)\right) d\mu,\end{aligned}$$

vemos que $(c_2 - c_1)\mathbf{1}_A$ não pertence a $L^\varphi_{c_2}$. Claramente, $(c_2 - c_1)\mathbf{1}_A \in L^\varphi_{c_1}$. Consequentemente, $L^\varphi_{c_1}$ não está contido em $L^\varphi_{c_2}$.

Reciprocamente, assuma que $c_1 - c_2 \in L^\varphi_{c_2}$. Seja w uma função qualquer em $L^\varphi_{c_1}$. Podemos encontrar $\epsilon > 0$ tal que $\int_T \varphi(c_1 + \lambda w)d\mu < \infty$, para cada $\lambda \in (-\epsilon, \epsilon)$. Considere a função convexa

$$g(\alpha, \lambda) = \int_T \varphi(c + \alpha(c_1 - c_2) + \lambda w)d\mu.$$

Esta função é finita para $\lambda = 0$ e α no intervalo $(-\eta, 1]$, para algum $\eta > 0$. Além disso, $g(1, \lambda)$ é finito para cada $\lambda \in (-\epsilon, \epsilon)$. Pela convexidade de g, vemos que g é finito no invólucro convexo do conjunto $1 \times (-\epsilon, \epsilon) \cup (-\eta, 1] \times 0$. Podemos encontrar que $g(0, \lambda)$ é finito para cada λ em alguma vizinhança de 0. Consequentemente, $w \in L^\varphi_{c_2}$. Desde que $w \in L^\varphi_{c_2}$ é arbitrário, segue a inclusão $L^\varphi_{c_1} \subseteq L^\varphi_{c_2}$. $\square$

Lema 3.16. *Se uma função u pertence a $\mathcal{K}_c^\varphi$ e denotamos $\tilde{c} = c + u - \psi(u)u_0$, então os espaços são iguais como conjuntos.*

Demonstração. A inclusão $L_c^\varphi \subseteq L_{\tilde{c}}^\varphi$ segue da Proposição 3.15. Desde que $u \in \mathcal{K}_c^\varphi$, temos

$$\int_T \varphi(\tilde{c} + \lambda u)d\mu \leq \int_T \varphi(c + (1+\lambda)u)d\mu < \infty,$$

para cada λ na vizinhança de 0. Assim, $c - \tilde{c} = -u + \psi(u)u_0$ pertence a $L_{\tilde{c}}^\varphi$. Da Proposição 3.15, obtemos $L_{\tilde{c}}^\varphi \subseteq L_c^\varphi$. $\square$

Temos que $L_{c_1}^\varphi$ e $L_{c_2}^\varphi$ são iguais como conjuntos e por (MUSIELAK, 1983, Teorema 8.5) suas normas são equivalentes. Como em (3.20) a função $w \in L_{c_2}^\varphi$ e pelo Lema 3.16, temos que $c_1 - c_2$ está em $L_{c_2}^\varphi$. Consequentemente a aplicação de transição é de classe C^∞. Outro resultado importante na construção das φ-famílias é provar que as φ-famílias $\mathcal{F}_c^\varphi$ são maximais no sentido que se duas φ-famílias $\mathcal{F}_{c_1}^\varphi$ e $\mathcal{F}_{c_2}^\varphi$ tem interseção diferente do vazio, então elas coincidem.

Lema 3.17. *Para uma função $u \in \mathcal{B}^\varphi$ denote $\tilde{c} = c + u - \psi(u)u_0$. Então $\mathcal{F}_c^\varphi = \mathcal{F}_{\tilde{c}}^\varphi$.*

Demonstração. Seja v uma função em $\mathcal{B}^\varphi$. Então existe $\epsilon > 0$ tal que , para cada $\lambda \in (-\epsilon, 1+\epsilon)$, $\int_T \varphi(c + \lambda v + (1-\lambda)u)d\mu < \infty$. Consequentemente, $\varphi(\tilde{c} + \lambda(v-u))$ é μ-integrável para todo $\lambda \in (-\epsilon, 1+\epsilon)$. Assim, a diferença $v - u$ está em $\mathcal{K}_{\tilde{c}}^\varphi$ e

$$w = v - u - \frac{\int_T (v-u)\varphi'_+(\tilde{c})d\mu}{\int_T u_0 \varphi'_+(\tilde{c})d\mu} u_0$$

pertence a $\mathcal{B}_{\tilde{c}}^\varphi$. Seja $\tilde{\psi} : \mathcal{B}_{\tilde{c}}^\varphi \to [0, \infty)$ a função de normalização associada a $\tilde{c}$. Então a densidade de probabilidade $\varphi(\tilde{c} + w - \tilde{\psi}(u)u_0)$ está em $\mathcal{F}_{\tilde{c}}^\varphi$. Esta densidade de probabilidade pode ser expressa como $\varphi(c + v - ku_0)$ para uma constante k. Pelo fato de existir um único $\psi(u) \in \mathbb{R}$ tal que a densidade de probabilidade $\varphi(c + v - \psi(v)u_0)$ está em $\mathcal{F}_c^\varphi$. Portanto $\mathcal{F}_c^\varphi \subseteq \mathcal{F}_{\tilde{c}}^\varphi$. Usando argumentos análogos podemos obter $\mathcal{F}_{\tilde{c}}^\varphi \subseteq \mathcal{F}_c^\varphi$. $\square$

Provando assim que $\mathcal{P}_\mu$ é equipado com uma estrutura C^∞-diferenciável.

A função normalizadora $\psi : \mathcal{K}_c^\varphi \to \mathbb{R}$ é convexa (VIGELIS and CAVALCANTE, 2013b). Assumindo que φ é continuamente diferenciável a função normalizadora é Gâteaux-diferenciável e a expressão para a Gâteaux derivada é (VIGELIS and CAVALCANTE, 2013b)

$$\partial\psi(u)v = \frac{\int_T v\varphi'(c + u - \psi(u)u_0)d\mu}{\int_T u_0\varphi'(c + u - \psi(u)u_0)d\mu}, \tag{3.21}$$

com $u \in \mathcal{K}_c^\varphi$ e $v \in L_c^\varphi$.

Utilizando a função normalizadora ψ uma divergência, baseada na divergência de Bregman associada a função normalizadora, foi encontrada. Divergência também conhecida com φ-divergência e dada pela expressão

$$D(p\|q) = \frac{\displaystyle\int_T \frac{\varphi^{-1}(p) - \varphi^{-1}(q)}{(\varphi^{-1})'(p)} d\mu}{\displaystyle\int_T u_0 \varphi'(c) d\mu}.$$

Esta divergência tem como um caso especial a divergência de Kullback-Leibler, quando $\varphi(u) = \exp(u)$ e $u_0 = 1$.

Uma propriedade importante aos estudos dos espaços de Musielak-Orlicz é a condição Δ_2. Vamos lembrar um pouco e ver algumas propriedades provadas em (VIGELIS and CAVALCANTE, 2013a) sobre essa condição e as φ-famílias de distribuição de probabilidades.

Seja a função de Musielak-Orlicz dada por:

$$\Phi_c(t, u) = \varphi(t, c(t) + u) - \varphi(t, c(t)), \tag{3.22}$$

em que a função exponencial deformada $\varphi(\cdot)$ satisfaz as condições (a1)-(a3') da Definição (3.18). Lembramos que, a função Φ_c satisfaz a condição Δ_2 ou $\Phi_c \in \Delta_2$ se uma constante $K > 0$ e uma função não negativa $f \in \tilde{L}_c^\varphi$ pode ser encontrada tal que

$$\alpha \Phi_c(t, 2u) \leq \Phi_c\Big(t, u\Big), \qquad \text{para todo } u \geq f(t), \quad \text{e } \mu\text{-q.t.p. } t \in T. \tag{3.23}$$

Sabemos também que, se $\Phi_c \in \Delta_2$, então $\int_T \varphi(c + u) d\mu < \infty$ para todo $u \in L_c^\varphi$, neste caso, $E_c^\varphi = L_c^\varphi = \tilde{L}_c^\varphi$, ou seja são iguais como conjuntos. Por outro lado se Φ_c não satisfaz a condição então E_c^φ é um subespaço próprio de L_c^φ, assim como $\tilde{L}_c^\varphi$ (MUSIELAK, 1983, Observação 7.3). Alguns resultados sobre a condição Δ_2 e a φ-família, foram provados em (VIGELIS and CAVALCANTE, 2013a).

Proposição 3.18 ((VIGELIS and CAVALCANTE, 2013a, Proposição 2))**.** *Dada qualquer exponencial deformada φ, podemos encontrar uma função mensurável $c : T \to \mathbb{R}$ em que $\int_T \varphi(c) d\mu = 1$ tal que a função de Musielak-Orlicz $\Phi_c(t, u(t)) = \varphi(t, c(t) + u) - \varphi(t, c(t))$ não satisfaz a condição Δ_2.*

Demonstração. Sejam A e B dois conjuntos mensuráveis disjuntos satisfazendo $0 < \mu(A) < \infty$ e $0 < \mu(B) < \infty$. Fixada qualquer função mensurável $\tilde{c}$ tal que $\int_T \varphi(\tilde{c}) d\mu = 1$, tomemos qualquer função não integrável f suportada em A tal que $\varphi(\tilde{c}) \mathbf{1}_A \leq f \mathbf{1}_A < \infty$. Seja $u : T \to [0, \infty)$ seja uma função mensurável suportada em A tal que $\varphi(\tilde{c} + u) \mathbf{1}_A = f \mathbf{1}_A$. Se $\beta > 0$ é tal que

$$\int_T \varphi(\tilde{c} - u) \mathbf{1}_A d\mu + \beta \mu(B) + \int_T \varphi(\tilde{c}) \mathbf{1}_{T \setminus (A \cup B)} d\mu = 1,$$

então definimos

$$c = (\tilde{c} - u)\mathbf{1}_A + \bar{c}\mathbf{1}_B + \tilde{c}\mathbf{1}_{T\setminus(A\cup B)},$$

em que $\bar{c} : T \to \mathbb{R}$ é uma função mensurável suportada em B tal que $\varphi(t, \bar{c}(t)) = \beta$, para μ-q.t.p. $t \in B$. Pelo fato de que a função u é suportada em A, podemos escrever

$$\int_T \varphi(c+u)d\mu = \int_T \varphi(\tilde{c})\mathbf{1}_A d\mu + \int_T \varphi(\bar{c})\mathbf{1}_B d\mu + \int_T \varphi(\tilde{c})\mathbf{1}_{T\setminus(A\cup B)} d\mu < \infty.$$

Por outro lado, desde que f é não integrável, temos

$$\int_T \varphi(c+2u)d\mu > \int_T \varphi(\tilde{c}+u)\mathbf{1}_A d\mu = \int_T f\, d\mu = \infty.$$

Portanto, a função de Musielak-Orlicz Φ_c não satisfaz a condição Δ_2. □

Proposição 3.19 ((VIGELIS and CAVALCANTE, 2013a, Proposiçao 3))**.** *Seja $b : T \to \mathbb{R}$ uma função mensurável tal que $\int_T \varphi(b)d\mu = 1$. Então $L_b^\varphi \subseteq L_c^\varphi$ para cada função $c : T \to \mathbb{R}$ tal que $\int_T \varphi(c)d\mu = 1$ se, e somente se, a função de Musielak-Orlicz $\Phi_c(t,u) = \varphi(t, b(t)+u) - \varphi(t, b(t))$ satisfaz a Δ_2-condição.*

O resultado principal em relação a condição Δ_2 e as φ-famílias de distribuições de probabilidade é dado pela proposição abaixo.

Proposição 3.20 ((VIGELIS and CAVALCANTE, 2013a, Proposição 4))**.** *Sejam $b, c : T \to \mathbb{R}$ funções mensuráveis tais que $\int_T \varphi(b)d\mu = 1$ e $\int_T \varphi(c)d\mu = 1$. Se as funções de Musielak-Orlicz $\Phi_b(t,u) = \varphi(t, b(t)+u) - \varphi(t, b(t))$ e $\Phi_c(t,u) = \varphi(t, c(t)+u) - \varphi(t, c(t))$ satisfazem a Δ_2-condição, então L_b^φ e L_c^φ são iguais como conjuntos. Além disso, $\mathcal{F}_b^\varphi = \mathcal{F}_c^\varphi$.*

Demonstração. A conclusão que L_b^φ e L_c^φ são iguais como conjuntos segue da Proposição 3.19. Pela Proposição 3.15, está claro que $(c-b) \in \mathcal{K}_b^\varphi$. Seja $\alpha \geq 0$ é tal que $u = (c-b) + \alpha u_0$ pertence a $\mathcal{B}_b^\varphi$. Se ψ_1 é a função normalizadora associada com $\mathcal{F}_b^\varphi$, então $\psi_1(u) = \alpha$ e $\varphi_b(u) = \varphi(b + u - \psi_1(u)u_0) = \varphi(c)$. Assim, as φ-famílias $\mathcal{F}_b^\varphi$ e $\mathcal{F}_c^\varphi$ tem interseção não vazia e portanto $\mathcal{F}_b^\varphi = \mathcal{F}_c^\varphi$. □

O resultado da Proposição 3.20 garante então que, para duas funções de Musielak-Orlicz tomadas a partir de uma exponencial deformada, que satisfazem a Δ_2-condição, então não só os espaços de Musielak-Orlicz obtidos a partir dessas funções são iguais como conjuntos, como também as φ-famílias de distribuição de probabilidade são iguais. Ainda iremos investigar sobre como a Δ_2-condição influencia o bordo do seu domínio.

Nesse texto, queremos encontrar condições sobre a função φ para a qual podemos conectar densidades de probabilidades por arcos, assim como foi feito em (CENA and PISTONE, 2007). Em outras palavras, queremos garantir a existência de arcos que generalizam os arcos exponenciais e arcos que generalizam os arcos mistura para que possamos conectar densidades de probabilidade na variedade estatística generalizada.

Na próxima seção, estudaremos as condições sobre φ, para as quais, garantirmos a existência de arcos na variedade estatística generalizada.

4 ARCOS EM VARIEDADES ESTATÍSTICAS GENERALIZADAS

Em (PISTONE and ROGANTIN, 1999; CENA and PISTONE, 2007; SANTACROCE, SIRI, and TRIVELLATO, 2016) foi mostrado que duas densidades de probabilidade são conectadas por um arco exponencial aberto (ou caminho exponencial aberto) se e somente se elas pertencem a mesma família exponencial. Arcos exponenciais são curvas auto-paralelas com respeito à conexão exponencial. Nesta seção, achamos condições necessárias e suficientes para que duas distribuições de probabilidade sejam conectadas por um arco (VIGELIS, ANDRADE, and CAVALCANTE, 2017). Nessa mesma linha, encontramos condições necessárias e suficientes para que possamos definir a generalização da divergência de Rényi definida em (SOUZA, VIGELIS, and CAVALCANTE, 2016).

4.1 Arcos exponencial e mistura em famílias exponenciais

Vamos lembrar as definições de arco mistura e arco exponencial em variedades estatística exponenciais e alguns resultados. Sejam $\Phi = \cosh(x) - 1$ uma função de Young e $L^{\Phi}(p)$ um espaço de Orlicz associado a essa função (KRASNOSELl'SKI and RUTICKI, 1961).

$$L^{\Phi}(p) = \left\{ u \in L^0 : \exists\, \alpha > 0 \text{ tal que } \int_T \Phi(\alpha u) p d\mu < +\infty \right\}.$$

Nesta seção, usaremos o símbolo $\propto$ como símbolo de proporcionalidade, ou seja, $a \propto b$, significa que a e b são proporcionais. Em (CENA and PISTONE, 2007), foram definidas densidades conectadas por arcos abertos mistura e exponencial abertos e a partir dessa conexão chegou se novamente no atlas que fornece uma estrutura de espaço de Banach para o conjunto das densidades equivalentes a uma medida μ, $\mathcal{P}_\mu$, dado pelo modelo maximal

$$\mathcal{E}(p) = \left\{ e^{u - k_p(u)} p | u \in \mathcal{K}_p \right\},$$

em que $\mathcal{K}_p$ é o interior do domínio do funcional gerador de cumulante $K_p(\cdot)$.

Definição 4.1. Duas densidades p e $q \in \mathcal{P}_\mu$ são conectados por um arco mistura aberto se existe um intervalo aberto I, tal que$[0,1] \subset I$ e $p(\alpha) = (1-\alpha)p + \alpha q$ pertence a $\mathcal{P}_\mu$, para cada $\alpha \in I$.

Definição 4.2. Duas densidades p e $q \in \mathcal{P}_\mu$ são conectadas por um arco exponencial aberto se existe um intervalo aberto I, tal que $[0,1] \subset I$ e $p(\alpha) \propto p^{1-\alpha} q^{\alpha}$ pertence a $\mathcal{P}_\mu$, para cada $\alpha \in I$.

Uma definição equivalente de conexão exponencial por arcos é.

Proposição 4.3 ((SANTACROCE, SIRI, and TRIVELLATO, 2016, Proposição 3.3)). *p, $q \in \mathcal{P}_\mu$ são conectadas por um arco exponencial aberto se, e somente se, existe um intervalo aberto $I \supset [0,1]$ e uma variável aleatória $u \in L^\Phi(p)$, tal que $p(\alpha) \propto e^{\alpha u}p$ pertence a $\mathcal{P}_\mu$, para cada $\alpha \in I$ e $p(0) = p$ e $p(1) = q$.*

Demonstração. Vamos assumir que $p, q \in \mathcal{P}_\mu$ são conectados por um arco exponencial aberto, isto é, $\int_T p^{(1-\alpha)} q^\alpha d\mu < +\infty$, para qualquer $\alpha \in I$. Desde que

$$\int_T p^{(1-\alpha)} q^\alpha d\mu = \mathbb{E}_p \left(\left(\frac{q}{p} \right)^\alpha \right) = \mathbb{E}_p(\mathrm{e}^{\alpha u}) \quad \text{com } u = \log \frac{q}{p},$$

então $u \in L^\Phi(p)$. Além disso, $p(\alpha) \propto \mathrm{e}^{\alpha u}$ pertence a $\mathcal{P}_\mu$, para cada $\alpha \in I$ e $p(0) = p$, $p(1) = q$. A recíproca segue imediatamente, observando que $q = p(1) = \mathrm{e}^u p$, isto é, $u = \log \frac{q}{p}$. $\square$

Estas conexões por arcos mistura abertos e arcos exponencial abertos são relações de equivalência (CENA and PISTONE, 2007). Em (CENA and PISTONE, 2007) alguns resultados a respeito dessa conexão por arcos exponencial foram provados. Entre esse resultados estão que, duas distribuições de probabilidade p e q são conectadas por uma arco exponencial aberto se, e somente se, $q \in \mathcal{E}(p)$. Além disso, como consequência dessa conexão temos que $\mathcal{E}(p) = \mathcal{E}(q)$ e $L^\Phi(p) = L^\Phi(q)$. Ou seja, a componente conectada a uma densidade p é igual ao modelo exponencial máximo $\mathcal{E}(p)$.

Dado $p \in \mathcal{P}_\mu$, denota-se por $\mathcal{M}(p)$ o conjunto de todas as densidades $q \in \mathcal{P}_\mu$ que são conectadas a p por um arco mistura aberto (SANTACROCE, SIRI, and TRIVELLATO, 2016).

Teorema 4.4 ((SANTACROCE, SIRI, and TRIVELLATO, 2016, Teorema 4.11)). *Sejam $p, q \in \mathcal{P}_\mu$. As seguintes afirmações são equivalentes.*

i) $q \in \mathcal{M}(p)$;

ii) $\mathcal{M}(p) = \mathcal{M}(q)$;

iii) $\frac{q}{p}, \frac{p}{q} \in L^\infty$.

Este teorema nos diz que duas densidades $p, q \in \mathcal{P}_\mu$ são conectadas por um arco mistura aberto se e somente se as razões $\frac{q}{p}$ e $\frac{p}{q}$ forem limitadas por constantes positivas. Outro resultado importante provado em (SANTACROCE, SIRI, and TRIVELLATO, 2016, Proposiçao 4.12) é que $\mathcal{M}(p) \subset \mathcal{E}(p)$, ou seja, duas densidades são conectadas por um arco mistura aberto, então são conectadas por uma arco exponencial aberto.

No mesmo sentido da ideia seguida em (CENA and PISTONE, 2007) e (SANTACROCE, SIRI, and TRIVELLATO, 2016) vamos tentar conectar densidades na variedade estatística generalizada, por arcos, tanto do tipo exponencial, como do tipo mistura.

Na próxima seção, baseado na estrutura que as φ-famílias de distribuições de probabilidade fornecem à $\mathcal{P}_\mu$, vamos investigar se é possível conectar por φ-arcos, densidades p e q em $\mathcal{P}_\mu$. Neste caso, os arcos exponencial são um caso especial desses φ-arcos.

4.2 Distribuições de probabilidade conectadas

A partir de agora, vamos discutir a importância de uma exponencial deformada φ ser tomada como na Definição 3.13 para a conexão por arcos na variedade estatística generalizada. Sejam φ_1 e φ_2 satisfazendo as condições (a1) e (a2) na Definição 3.13. Fixemos uma função mensurável e positiva $u_0 : T \to (0, \infty)$. Dadas duas distribuições de probabilidade p e q pertencentes a $\mathcal{P}_\mu$, um φ_1/φ_2-*arco*(ou φ_1/φ_2-caminho) é uma curva em $\mathcal{P}_\mu$ definida por $\alpha \mapsto \varphi_1(\alpha\varphi_2^{-1}(p) + (1-\alpha)\varphi_2^{-1}(q) + \kappa(\alpha)u_0)$. A constante $k(\alpha) \colon = k(\alpha; p, q) \in \mathbb{R}$ é introduzida para que

$$\int_T \varphi_1(\alpha\varphi_2^{-1}(p) + (1-\alpha)\varphi_2^{-1}(q) + \kappa(\alpha)u_0)d\mu = 1. \tag{4.1}$$

Usamos φ-arco ao invés de φ/φ-arco, se $\varphi = \varphi_1 = \varphi_2$. O caso $\varphi = \varphi_1 = \varphi_2$ e $u_0 = 1$ foi analisado por (EGUCHI and KOMORI, 2015). O arco exponencial corresponde ao φ-arco com $\varphi_1(\cdot)$ e $\varphi_2(\cdot)$ iguais a $\exp(\cdot)$ e $u_0 = 1$. Um φ_1/φ_2-arco pode ser visto como φ_1-arco conectando $\varphi_1(\varphi_2^{-1}(p) + k(1)u_0)$ e $\varphi_1(\varphi_2^{-1}(q) + \kappa(0)u_0)$. A menos que $\varphi = \varphi_1 = \varphi_2$, um φ_1/φ_2-arco não conecta p e q. Podemos usar $\kappa(\alpha)$ para definir a divergência

$$\mathcal{D}^{(\alpha)}(p \parallel q) = -\frac{1}{\alpha}\kappa(0) - \frac{1}{1-\alpha}\kappa(1) + \frac{1}{\alpha(1-\alpha)}\kappa(\alpha). \tag{4.2}$$

Esta divergência para $\varphi = \varphi_1 = \varphi_2$ é relacionada à generalização da divergência de Rényi definida em (SOUZA, VIGELIS, and CAVALCANTE, 2016), que vamos estudar mais detalhadamente na próxima seção. Se $\varphi_1(\cdot)$ e $\varphi_2(\cdot)$ são iguais à $\exp(\cdot)$ e $u_0 = 1$, então a divergência $\mathcal{D}^{(\alpha)}(\cdot \parallel \cdot)$ se reduz à divergência de Rényi. A pergunta que queremos responder nessa seção é sobre que condições sobre as funções φ_1 e φ_2, existe $\kappa(\alpha)$ em (4.1) para cada $p, q \in \mathcal{P}_\mu$ e $\alpha \in [0, 1]$. A proposição a seguir nos ajuda a elucidar essa questão.

Proposição 4.5. *Assumindo que a medida μ é não atômica, sejam $\varphi_1, \varphi_2 : \mathbb{R} \to (0, \infty)$ duas funções positivas, satisfazendo (a1) e (a2) na Definição3.13 e seja $u_0 : T \to (0, \infty)$ uma função mensurável positiva. Fixando qualquer $\alpha \in (0, 1)$, para cada par de distribuições de probabilidades p e q em $\mathcal{P}_\mu$, existe uma constante $\kappa(\alpha) \colon = \kappa(\alpha; p, q)$ satisfazendo (4.1) se, e somente se,*

$$\int_T \varphi_1(c + \lambda u_0)d\mu < \infty, \qquad \textit{para todo } \lambda \geq 0, \tag{4.3}$$

para cada função mensurável $c : T \to \mathbb{R}$ satisfazendo $\int_T \varphi_2(c)d\mu < \infty$.

Quando $\varphi = \varphi_1 = \varphi_2$, a condição (4.3) é a mesma que a condição (a3') da Definição 3.13. A fim de provar a Proposição 4.5 precisamos de alguns resultados intermediários.

Lema 4.6. *Supondo que, para cada $\lambda > 0$, não podemos encontrar $\alpha \in (0,1)$ e uma função mensurável $c : T \to \mathbb{R} \cup \{-\infty\}$ tal que $\int_T \varphi_1(c)d\mu < \infty$ e*

$$\alpha\varphi_1(u) \leq \varphi_2(u - \lambda u_0(t)), \quad \textit{para todo } u \geq c(t). \tag{4.4}$$

Então existem sequências $\{\lambda_n\}$, $\{c_n\}$ e $\{A_n\}$ de números positivos, onde λ_n é uma sequência monótona decrescente convergindo para zero, $\lambda_n \downarrow 0$, funções mensuráveis, e conjuntos mensuráveis dois a dois disjuntos, respectivamente, tal que

$$\int_T \varphi_1(c_n)d\mu = 1 \quad e \quad \int_{A_n} \varphi_2(c_n - \lambda_n u_0)d\mu \leq 2^{-n}, \quad \textit{para todo } n \geq 1. \tag{4.5}$$

Demonstração. Seja $\{\lambda'_m\}$ uma sequência de números positivos $\lambda'_m \downarrow 0$. Para cada $m \geq 1$, definimos a função

$$f_m(t) = \sup\{u \in \mathbb{R} : 2^{-m}\varphi_1(u) > \varphi_2(u - \lambda'_m u_0(t))\},$$

na qual usamos a convenção $\sup \emptyset = -\infty$. Vamos verificar que f_m é mensurável. Para cada número racional r, definimos os conjuntos mensuráveis

$$E_{m,r} = \{t \in T : 2^{-m}\varphi_1(r) > \varphi_2(r - \lambda'_m u_0(t))\}$$

e a função simples $u_{m,r} = r\chi_{E_{m,r}}$. Seja $\{r_i\}$ uma enumeração de números racionais. Pra cada $m, k \geq 1$, considere a função simples não negativa $v_{m,k} = \max_{1 \leq i \leq k} u_{m,r_i}$. Além disso, denote $B_{m,k} = \bigcup_{i=1}^k E_{m,r_i}$. Pela continuidade de $\varphi_1(\cdot)$ e $\varphi_2(\cdot)$, segue que $v_{m,k}\chi_{B_{m,k}} \uparrow f_m$ quando $k \to \infty$, o que mostra que f_m é mensurável. Uma vez que (4.4) não é satisfeita, temos que $\int_T \varphi_1(f_m)d\mu = \infty$ para todo $m \geq 1$. Em virtude do Teorema da Convergência Monótona (ISNARD, 2007, Teorema 5.34), para cada $m \geq 1$, podemos encontrar algum $k_m \geq 1$ tal que a função $v_m = v_{m,k_m}$ e o conjunto $B_m = B_{m,k_m}$ satisfaçam $\int_{B_m} \varphi_1(v_m)d\mu \geq 2^m$. Claramente, temos que $\varphi_1(v_m)\chi_{B_m} < \infty$ e $2^{-m}\varphi_1(v_m)\chi_{B_m} \geq \varphi_2(v_m - \lambda'_m u_0)\chi_{B_m}$. Pelo Lema 8.3 em (MUSIELAK, 1983), existe uma sequência crescente $\{m_n\}$ de índices e uma sequência $\{A_n\}$ de conjuntos mensuráveis disjuntos tal que $\int_{A_n} \varphi_1(v_{m_n})d\mu = 1$. Claramente, $\int_{A_n} \varphi_2(v_{m_n} - \lambda'_{m_n} u_0)d\mu \leq 2^{-m_n}$. Denotando $\lambda_n = \lambda'_{m_n}$, $c_n = v_{m_n}$, obtemos (4.5). $\square$

A proposição seguinte nos fornece uma condição equivalente à condição (4.3).

Proposição 4.7. *Duas funções φ_1 e φ_2, satisfazendo (a1) e (a2) na Definição 3.13,φ_1/φ_2 : $\mathbb{R} \to [0,\infty)$ e uma função mensurável $u_0 : T \to (0,\infty)$ satisfazem a condição (4.3) se, e somente se, para cada $\lambda > 0$ podemos encontrar $\alpha \in (0,1)$ e uma função mensurável*

$c : T \to \mathbb{R} \cup \{-\infty\}$ tal que $\int_T \varphi_1(c)d\mu < \infty$ e

$$\alpha\varphi_1(u) \leq \varphi_2(u - \lambda u_0(t)), \quad \textit{para todo } u \geq c(t), \tag{4.6}$$

para μ-q.t.p. $t \in T$.

Demonstração. Assumindo que $\varphi_1(\cdot)$, $\varphi_2(\cdot)$ e u_0 satisfazem a condição (4.3) suponha que (4.4) não acontece. Sejam ainda $\{\lambda_n\}, \{c_n\}$ e $\{A_n\}$ como consideradas no Lema 4.6. Então definimos $c = c_0\chi_{T\setminus A} + \sum_{n=1}^{\infty}(c_n - \lambda_n u_0)\chi_{A_n}$, em que $A = \cup_{n=1}^{\infty} A_n$ e $c_0 : T \to \mathbb{R}$ é uma função mensurável qualquer tal que $\int_{T\setminus A} \varphi_2(c_0)d\mu < \infty$. Tendo em vista (4.5), temos que

$$\begin{aligned}
\int_T \varphi_2(c)d\mu &= \int_{T\setminus A} \varphi_2(c_0)d\mu + \sum_{n=1}^{\infty} \int_{A_n} \varphi_2(c_n - \lambda_n u_0)d\mu \\
&\leq \int_{T\setminus A} \varphi_2(c_0)d\mu + \sum_{n=1}^{\infty} 2^{-n} < \infty.
\end{aligned}$$

Dado qualquer $\lambda > 0$, tomemos $n_0 \geq 1$ tal que $\lambda \geq \lambda_n$ para todo $n \geq n_0$. Então podemos escrever,

$$\begin{aligned}
\int_T \varphi_1(c + \lambda u_0)d\mu &\geq \sum_{n=n_0}^{\infty} \int_{A_n} \varphi_1(c_n + (\lambda - \lambda_n)u_0)d\mu \\
&\geq \sum_{n=n_0}^{\infty} \int_{A_n} \varphi_1(c_n)d\mu = \sum_{n=1}^{\infty} 1 = \infty,
\end{aligned} \tag{4.7}$$

o que é uma contradição à condição (4.3).

Por outro lado, suponha que a expressão (4.6) ocorre para um dado $\lambda > 0$. Seja $\tilde{c} : T \to \mathbb{R}$ qualquer função mensurável satisfazendo $\int_T \varphi_2(\tilde{c})d\mu < \infty$. Denote $A = \{t : \tilde{c}(t) + \lambda u_0 \geq c(t)\}$. Usamos a inequação (4.6) para escrever

$$\begin{aligned}
\alpha \int_T \varphi_1(\tilde{c} + \lambda u_0)d\mu &\leq \alpha \int_A \varphi_1(\tilde{c} + \lambda u_0)d\mu + \alpha \int_{T\setminus A} \varphi_1(c)d\mu \\
&\leq \int_A \varphi_2(\tilde{c})d\mu + \int_{T\setminus A} \varphi_2(c - \lambda u_0)d\mu < \infty.
\end{aligned}$$

Assim, segue a condição (4.3).

Precisaremos de mais um resultado preliminar antes de demonstrar a Proposição (4.5). □

Lema 4.8. *Sejam $\varphi_1, \varphi_2 : \mathbb{R} \to (0, \infty)$ funções positivas, satisfazendo as condições (a1) e (a2) da Definição 3.13 e $\tilde{c} : T \to \mathbb{R}$ uma função mensurável tal que*

$$\int_A \varphi_i(\tilde{c})d\mu < 1, \quad \textit{para } i = 1, 2,$$

em que A e $B = T \setminus A$ são conjuntos mensuráveis tais que $\mu(A) > 0$ e $\mu(B) > 0$. Fixando qualquer $\alpha \in (0,1)$, então, podemos encontrar funções mensuráveis $b_1, b_2 : T \to \mathbb{R}$ para as quais $p = \varphi_2(c_1)$ e $q = \varphi_2(c_1)$ estão em $\mathcal{P}_\mu$, em que $c_1 = \widetilde{c}\chi_A + b_1\chi_B$ e $c_2 = \widetilde{c}\chi_A + b_2\chi_B$, e

$$\int_T \varphi_1(\alpha\varphi_2^{-1}(p) + (1-\alpha)\varphi_2^{-1}(q)) < 1. \tag{4.8}$$

Além do mais, assumimos que $b_1\chi_B \neq b_2\chi_B$.

Demonstração. Seja $\{B_n\}$ uma sequência de conjuntos mensuráveis tais que $B = \cup_{n=1}^\infty B_n$ e $0 < \mu(B_n) < \infty$. Para cada $n \geq 1$, selecionamos conjuntos mensuráveis C_n e D_n tal que $B_n = C_n \cup D_n$ e $\mu(C_n) = \mu(D_n) = \frac{1}{2}\mu(B_n)$. Sejam ainda $\{\gamma_n^{(1)}\}$ e $\{\gamma_n^{(2)}\}$ sequências de números positivos satisfazendo

$$\sum_{n=1}^\infty \gamma_n^{(1)} < 1 - \int_A \varphi_1(\widetilde{c})d\mu, \quad \text{e} \quad \sum_{n=1}^\infty \gamma_n^{(2)} = 1 - \int_A \varphi_2(\widetilde{c})d\mu.$$

Então, tomamos $\beta_n \in \mathbb{R}$ e $\theta_n > 0$ tais que

$$\varphi_2(\beta_n) + \varphi_2(-\theta_n) = 2\frac{\gamma_n^{(2)}}{\mu(B_n)} \tag{4.9}$$

e

$$\varphi_1(\alpha\beta_n - (1-\alpha)\theta_n) + \varphi_1(-\alpha\theta_n + (1-\alpha)\beta_n) \leq 2\frac{\gamma_n^{(1)}}{\mu(B_n)}. \tag{4.10}$$

Números β_n e θ_n, satisfazendo (4.9) e (4.10) existem pelo fato de $\varphi_1(\cdot)$ e $\varphi_2(\cdot)$ são positivas e $\beta_n < \varphi_2^{-1}(2\gamma_n^{(2)}/\mu(B_n))$. Definamos

$$b_1 = \sum_{n=1}^\infty \beta_n\chi_{C_n} - \theta_n\chi_{D_n}$$

e

$$b_2 = \sum_{n=1}^\infty -\theta_n\chi_{C_n} + \beta_n\chi_{D_n}.$$

A partir dessas escolhas segue que

$$\begin{aligned}
\int_B \varphi_2(b_1)d\mu &= \sum_{n=1}^\infty \varphi_2(\beta_n)\mu(C_n) + \varphi_2(-\theta_n)\mu(D_n) \\
&= \sum_{n=1}^\infty [\varphi_2(\beta_n) + \varphi_2(-\theta_n)]\frac{\mu(B_n)}{2} \\
&= \sum_{n=1}^\infty \gamma_n^{(2)} = 1 - \int_A \varphi_1(\widetilde{c})d\mu,
\end{aligned}$$

o que implica que $\int_T \varphi_2(c_1) = 1$, em que $c_1 = \tilde{c}\chi_A + b_1\chi_B$. Da mesma forma, temos que

$\int_T \varphi_2(c_2)d\mu = 1$, em que $c_2 = \widetilde{c}\chi_A + b_2\chi_B$. Por outro lado, podemos escrever

$$\begin{aligned}
&\int_B \varphi_1(\alpha b_1 + (1-\alpha)b_2) \\
&= \sum_{n=1}^{\infty} \varphi_1(\alpha\beta_n - (1-\alpha)\theta_n)\mu(C_n) + \varphi_1(-\alpha\theta_n + (1-\alpha)\beta_n)\mu(D_n) \\
&= \sum_{n=1}^{\infty} [\varphi_1(\alpha\beta_n - (1-\alpha)\theta_n) + \varphi_1(-\alpha\theta_n + (1-\alpha)\beta_n)]\frac{\mu(B_n)}{2} \\
&\leq \sum_{n=1}^{\infty} \gamma_n^{(1)} < 1 - \int_A \varphi_1(\widetilde{c})d\mu,
\end{aligned}$$

do qual segue a expressão (4.8). □

Finalmente, podemos demonstrar a Proposição 4.5.

Demonstração. [Proposição 4.5]. Pelo fato de $\varphi_2(\cdot)$ ser convexa, segue que $\int_T \varphi_2(c)d\mu < \infty$, em que $c = \alpha\varphi_2^{-1}(p) + (1-\alpha)\varphi_2^{-1}(q)$. A condição (4.3) juntamente com Teorema da convergência Monótona e a continuidade de $\varphi_1(\cdot)$ implicam a existência e unicidade de $k(\alpha)$.

Reciprocamente, assumindo a existência de $\kappa(\alpha)$ em (4.1) para todo $p, q \in \mathcal{P}_\mu$. Começamos mostrando que

$$\int_T \varphi_1(c - \lambda u_0)d\mu < \infty, \quad \text{para todo } \lambda \geq 0, \tag{4.11}$$

para toda função mensurável $c : T \to \mathbb{R}$ tal que $\int_T \varphi_2(c)d\mu < \infty$. Se a expressão (4.11) não ocorre, então para alguma função mensurável $c : T \to \mathbb{R}$ com $\int_T \varphi_2(c)d\mu < \infty$, e algum $\lambda_0 \geq 0$, temos

$$\begin{cases} \int_T \varphi_1(c - \lambda u_0)d\mu < \infty, & \text{for } \lambda_0 \leq \lambda, \\ \int_T \varphi_1(c - \lambda u_0)d\mu = \infty, & \text{for } 0 \leq \lambda < \lambda_0, \end{cases} \tag{4.12}$$

ou

$$\begin{cases} \int_T \varphi_1(c - \lambda u_0)d\mu < \infty, & \text{for } \lambda_0 < \lambda, \\ \int_T \varphi_1(c - \lambda u_0)d\mu = \infty, & \text{for } 0 \leq \lambda \leq \lambda_0. \end{cases} \tag{4.13}$$

Note que a expressão (4.12) com $\lambda_0 = 0$ corresponde à (4.11), então, em (4.12) assumimos que $\lambda_0 > 0$. Seja $\{T_n\}$ uma sequência de conjuntos mensuráveis não decrescentes com $0 < \mu(T_n) < \mu(T)$ e $\mu(T \setminus \bigcup_{n=1}^{\infty} T_n) = 0$. Defina $E_n = T_n \cap \{c - \lambda_0 u_0 \leq n\}$, para cada $n \geq 1$. Claramente, a sequência $\{E_n\}$ é não decrescente e satisfaz $\mu(E_n) < \infty$ e $\mu(T \setminus \cup_{n=1}^{\infty} E_n) = 0$.

Se a expressão (4.12) é satisfeita para $\lambda_0 > 0$, selecionamos um suficientemente grande $n_0 \geq 1$ tal que $\int_{T\setminus E_{n_0}} \varphi_i(c - \lambda_0 u_0)d\mu < 1$, para $i = 1, 2$. Denote $A\colon = T \setminus E_{n_0}$ e $B\colon = E_{n_0}$. De acordo com o Lema 4.8, podemos encontrar funções mensuráveis para quais $p = \varphi_2(c_1)$ e $q = \varphi_2(c_2)$ estão em $\mathcal{P}_\mu$, em que $c_1 = (c - \lambda_0 u_0)\chi_A + b_1\chi_B$ e $c_2 = (c - \lambda_0 u_0)\chi_A + b_2\chi_B$, e a inequação (4.8) é satisfeita. Para qualquer $\lambda > 0$, podemos escrever

$$\begin{aligned}\int_T \varphi_1(\alpha\varphi_2^{-1}(p) + (1-\alpha)\varphi_2^{-1}(q) + \lambda u_0) &\geq \int_B \varphi_1(c - (\lambda_0 - \lambda)u_0)d\mu \\ &= \int_T \varphi_1(c - (\lambda_0 - \lambda)u_0)d\mu - \int_{A_{n_0}} \varphi_1(c - (\lambda_0 - \lambda)u_0)d\mu = \infty.\end{aligned}$$

Pela expressão acima e a inequação (4.8), concluímos que a constante $\kappa(\alpha)$ como definida por (4.1) não pode ser encontrada.

Agora supondo que (4.13) é satisfeita. Seja $\{\lambda_n\}$ uma sequência em (λ_0, ∞) tal que $\lambda_n \downarrow \lambda_0$. Definimos indutivamente uma sequência crescente $\{k_n\} \subset \mathbb{N}$ como segue. Escolha $k_0 \geq 1$ tal que $\int_{T\setminus E_{k_0}} \varphi_1(c - \lambda_1 u_0)d\mu \leq 2^{-2}$. Dado k_{n-1} selecionamos algum $k_n > k_{n-1}$ tal que

$$\int_{E_{k_n}\setminus E_{k_{n-1}}} \varphi_1(c - \lambda_0 u_0)d\mu \geq 1$$

e

$$\int_{T\setminus E_{k_n}} \varphi_1(c - \lambda_{n+1}u_0)d\mu \leq 2^{-(n+2)}.$$

Vamos denotar $A_n = E_{k_n} \setminus E_{k_{n-1}}$ para $n \geq 1$. Note que os conjuntos A_n são disjuntos dois a dois. Tome $n_0 > 1$ tal que $\int_A \varphi(c)d\mu < 1$, em que $A = \cup_{n=n_o}^{\infty} A_n$. Agora definimos $\widetilde{c} = \sum_{n=n_0}^{\infty}(c - \lambda_n u_0)\chi_{A_n}$. Como um resultado dessas escolhas, segue que

$$\int_A \varphi_1(\widetilde{c})d\mu = \sum_{n=n_0}^{\infty} \int_{A_n} \varphi(c - \lambda_n u_0)d\mu \leq \sum_{n=n_0}^{\infty} 2^{-n_0} < 1$$

e

$$\int_A \varphi_2(\widetilde{c})d\mu < \int_A \varphi_2(c)d\mu < 1.$$

Denote $B = T\setminus A$. Tendo em vista o Lema 4.8, existem funções mensuráveis $b_1, b_2 : T \to \mathbb{R}$ tal que $p = \varphi_2(c_1)$ e $q = \varphi_2(c_2)$ estão em $\mathcal{P}_\mu$, em que $c_1 = \widetilde{c}\chi_A + b_1\chi_B$ e $c_2 = \widetilde{c}\chi_A + b_2\chi_B$, e a inequação (4.8) é satisfeita. Reciprocamente, se a constante $\kappa(\alpha)$ como dada em (4.1) existe, então $\kappa(\alpha) > 0$. Fixado um $\lambda > 0$ arbitrário, tomamos $n_1 \geq n_0$ tal que $\lambda_n - \lambda \leq \lambda_0$

para todo $n \geq n_1$. Observando que $\int_T \varphi_1(c - \lambda_0 u_0) d\mu \geq 1$, podemos escrever

$$\begin{aligned}\int_T \varphi_1(\alpha\varphi_2^{-1}(p) + (1-\alpha)\varphi_2^{-1}(q) + \lambda u_0) d\mu &\geq \int_A \varphi_1(\widetilde{c} + \lambda u_0) d\mu \\ &\geq \sum_{n=n_1}^{\infty} \int_{A_n} \varphi_1(c - (\lambda_n - \lambda) u_0) d\mu \geq \sum_{n=n_1}^{\infty} 1 = \infty,\end{aligned}$$

o que mostra que $\kappa(\alpha)$ não pode ser encontrado.

Suponha agora que a condição (4.3) não é satisfeita. Pela Proposição 4.7 e Lema 4.6, podemos encontrar sequências $\{\lambda_n\}$, $\{c_n\}$ e $\{A_n\}$ de números positivos $\lambda_n \downarrow 0$, funções mensuráveis e conjuntos mensuráveis disjuntos, respectivamente, tal que

$$\int_{A_n} \varphi_1(c_n) d\mu = 1 \quad \text{and} \quad \int_{A_n} \varphi_2(c_n - \lambda_n u_0) d\mu \leq 2^{-n}, \quad \text{for all } n \geq 1.$$

Pela expressão (4.11), podemos concluir que $\sum_{n=n_0}^{\infty} \int_{A_n} \varphi_1(c_n - \lambda_n u_0) d\mu < \infty$. Então, podemos pegar algum $n_0 > 1$ para o qual a função $\widetilde{c} = \sum_{n=n_0}^{\infty} (c_n - \lambda_n u_0) \chi_{A_n}$satisfaz $\int_A \varphi_i(\widetilde{c}) d\mu < 1$, para $i = 1, 2$, em que $A = \cup_{n=n_0}^{\infty} A_n$. Vamos denotar $B = T \setminus A$. De (4.3), existem funções mensuráveis $b_1, b_2 : T \to \mathbb{R}$ tal que $p = \varphi_2(c_1)$ e $q = \varphi_2(c_2)$ estão em $\mathcal{P}_\mu$, em que $c_1 = \widetilde{c}\chi_A + b_1\chi_B$ e $c_2 = \widetilde{c}\chi_A + b_2\chi_B$, e a inequação (4.8) é satisfeita. Dado qualquer $\lambda > 0$, tome $n_1 \geq n_0$ tal que $\lambda \geq \lambda_n$ para todo $n \geq n_1$. Então, podemos escrever

$$\begin{aligned}\int_T \varphi_1(\alpha\varphi_2^{-1}(p) + (1-\alpha)\varphi_2^{-1}(q) + \lambda u_0) d\mu &\geq \int_A \varphi_1(\widetilde{c} + \lambda u_0) d\mu \\ &\geq \sum_{n=n_1}^{\infty} \int_{A_n} \varphi_1(c_n + (\lambda - \lambda_n) u_0) d\mu \geq \sum_{n=n_1}^{\infty} \int_{A_n} \varphi_1(c_n) d\mu = \infty.\end{aligned}$$

Esta expressão e a inequação (4.8) implicam que a constante $\kappa(\alpha)$ como definida por (4.1) pode ser encontrada. Portanto, a condição (4.3) deve ser satisfeita. $\square$

Encontramos assim, condição necessária e suficiente para qual $\kappa(\alpha)$ existe para cada $p, q \in \mathcal{P}_\mu$ e $\alpha \in [0, 1]$. Como consequência da Proposição 4.7 temos o seguinte resultado para funções φ_1 e φ_2 com $\varphi = \varphi_1 = \varphi_2$.

Proposição 4.9. *Seja uma função $\varphi : \mathbb{R} \to [0, \infty)$. Então podemos encontrar uma função mensurável $u_0 : \mathbb{R} \to (0, \infty)$ para a qual a condição (4.3) acontece para $\varphi = \varphi_1 = \varphi_2$ se, e somente se,*

$$\limsup_{u \to \infty} \frac{\varphi(u)}{\varphi(u - \lambda_0)} < \infty, \qquad \textit{para algum } \lambda_0 > 0. \tag{4.14}$$

Demonstração. Pela Proposição 4.7 podemos concluir que a existência de u_0 implica (4.14). Reciprocamente, assumindo que a expressão (4.14) acontece para algum $\lambda_0 > 0$ existe $M \in (1, \infty)$ e $\bar{c} \in \mathbb{R}$ tal que $\frac{\varphi(u)}{\varphi(u-\lambda_0)} \leq M$ para todo $u \geq \bar{c}$. Seja $\{\lambda_n\}$ uma sequência qualquer em $(0, \lambda_0]$ tal que $\lambda_n \downarrow 0$. Para cada $n \geq 1$, defina

$$c_n = \sup\{u \in \mathbb{R} : \alpha\varphi(u) > \varphi(u - \lambda_n)\}, \tag{4.15}$$

na qual $\alpha = \frac{1}{M}$ e adotamos a convenção $\sup \emptyset = -\infty$. A partir da escolha de $\{\lambda_n\}$ e α, segue que $-\infty \leq c_n \leq \bar{c}$. Afirmamos que $\varphi(c_n) \downarrow 0$. Se a sequência converge para algum $c > -\infty$, a igualdade $\alpha\varphi(c_n) = \varphi(c_n - \lambda_n)$ implica que $\alpha\varphi(c) = \varphi(c)$ e então $\varphi(c) = 0$. No caso de $c_n \downarrow -\infty$, está claro que $\varphi(c_n) \downarrow 0$. Seja $\{T_n\}$ a sequência de conjuntos mensuráveis disjuntos com $\mu(T_k) < \infty$ e $\mu(T \setminus \bigcup_{k=1}^{\infty} T_k) = 0$. Assim, podemos selecionar uma subsequência $\{c_n\}$ tal que $\sum_{k=1}^{\infty} \varphi(c_{n_k})\mu(T_k) < \infty$. Vamos definir $c = \sum_{k=1}^{\infty} c_{n_k}\chi_{T_k}$ e $u_0 = \sum_{k=1}^{\infty} \lambda_{n_k}\chi_{T_k}$. De (4.15) segue que $\alpha\varphi(u) \leq \varphi(u - u_0(t))$, para todo $u \geq c(t)$. A Proposição 4.7 implica que $\varphi(\cdot)$ e u_0 satisfaz a condição (4.3). $\square$

Assim, concluímos que, só é possível conectar duas distribuições de probabilidades por um φ- arco, com $\varphi = \varphi_1 = \varphi_2$, se e somente se φ é uma exponencial deformada dada como na Definição 3.13. Daí a importância da definição de exponencial deformada para o desenvolvimento da nossa pesquisa. Na tentativa de conectar duas densidades em $\mathcal{P}_\mu$ por um φ- arco, vimos que essa conexão gera um valor $\kappa(\alpha) \geq 0$ que depende das densidades p e q e do valor de α. Esse valor pode ser relacionado com a divergência de Rényi no caso quando $\varphi(\cdot) = \exp(\cdot)$ e $u_0 = 1$. Sendo assim, utilizando esse valor $\kappa(\alpha)$ foi obtida uma generalização para a divergência de Rényi (SOUZA, VIGELIS, and CAVALCANTE, 2016).

Na próxima seção, vamos investigar essa divergência.

4.3 Generalização da divergência de Rényi

A divergência de Rényi (RÉNYI, 1961) é um das mais bem sucedidas medidas de dissimilaridade entre duas distribuições de probabilidade, tendo encontrado muitas aplicações (ERVEN and HARREMOES, 2014), aparece como ferramenta crucial em provas da convergência do comprimento mínimo da descrição e estimadores Bayesianos, em modelos paramétricos e não paramétricos (ZHANG, 2006), (HAUSSLER and OPPER, 1997), (ERVEN, 2010). Uma generalização da divergência de Rényi foi proposta em (SOUZA, VIGELIS, and CAVALCANTE, 2016). Como resultado deste trabalho, daremos nesta seção condições necessárias e suficientes para a existência dessa generalização, para ambos os casos, não atômico e o puramente atômico.

4.3.1 *Caso não atômico*

Neste caso a medida μ é não atômica. A divergência de Rényi, de ordem $\alpha \in (0,1)$, entre distribuições de probabilidade p e q em $\mathcal{P}_\mu$ é definida como

$$\mathcal{D}^{(\alpha)}(p\|q) = \frac{1}{\alpha(\alpha-1)} \log\left(\int_T p^\alpha q^{1-\alpha} d\mu\right),$$

que pode ser reescrita como

$$\mathcal{D}^{(\alpha)}(p\|q) = \frac{\kappa(\alpha)}{\alpha(1-\alpha)}, \tag{4.16}$$

em que

$$\kappa(\alpha) = -\log\left(\int_T p^\alpha q^{1-\alpha} d\mu\right).$$

Para $\alpha \in \{0,1\}$, a divergência de Rényi é definida tomando um limite:

$$\mathcal{D}^{(0)}(p\|q) = \lim_{\alpha \downarrow 0} \mathcal{D}^{(\alpha)}(p\|q),$$
$$\mathcal{D}^{(1)}(p\|q) = \lim_{\alpha \uparrow 1} \mathcal{D}^{(\alpha)}(p\|q).$$

A expressão (4.16) pode ser usada para definir a divergência de Rényi para todo $\alpha \in \mathbb{R}$. Contudo, para $\alpha \notin (-1,1)$ esta expressão talvez não seja valor finito para todo p e q em $\mathcal{P}_\mu$. Para evitar algumas minúcias, assumimos que $\alpha \in [-1,1]$. A forma padrão da divergência de Rényi encontrada na literatura difere de (4.16) por um fator $\frac{1}{\alpha}$. Escolhemos definir $\mathcal{D}^{(\alpha)}(\cdot\|\cdot)$ como em (4.16) de modo que alguma simetria poderia ser preservada quando os limites $\alpha \downarrow 0$ e $\alpha \uparrow 1$ são tomados.

A generalização da divergência de Rényi é baseada na interpretação alternativa de $\kappa(\alpha)$. Fixado $\alpha \in (0,1)$, e dado qualquer p e q em $\mathcal{P}_\mu$ a função $\kappa(\alpha) \colon = \kappa(\alpha; p, q)$ é o único número não negativo tal que

$$\int_T \exp(\alpha \ln(p) + (1-\alpha)\ln(q) + \kappa(\alpha)) d\mu = 1.$$

Para generalizar a divergência de Rényi, consideramos uma exponencial deformada $\varphi(\cdot)$ no lugar da função exponencial. Dada qualquer p e q em $\mathcal{P}_\mu$, tomamos $\kappa(\alpha) = \kappa(\alpha; p, q) \geq 0$ de modo que

$$\int_T \varphi(\alpha\varphi^{-1}(p) + (1-\alpha)\varphi^{-1}(q) + \kappa(\alpha)u_0) d\mu = 1, \tag{4.17}$$

em que $u_0 : T \to \mathbb{R}$ é a função mensurável positiva. Vamos analisar sobre quais condições sobre a função u_0 essa generalização da divergência de Rényi está bem definida.

Definimos a generalização da divergência de Rényi de ordem $\alpha \in (0,1)$ por

$$\mathcal{D}_\varphi^{(\alpha)}(p\|q) = \frac{\kappa(\alpha)}{\alpha(1-\alpha)},$$

com $\kappa(\alpha)$ dado como em (4.17). Para $\alpha \in \{0,1\}$, a generalização é definida tomando o limite:

$$\mathcal{D}_{\varphi}^{(0)}(p\|q) = \lim_{\alpha\downarrow 0} \mathcal{D}_{\varphi}^{(\alpha)}(p\|q), \tag{4.18}$$

$$\mathcal{D}_{\varphi}^{(1)}(p\|q) = \lim_{\alpha\uparrow 1} \mathcal{D}_{\varphi}^{(\alpha)}(p\|q). \tag{4.19}$$

Estes limites são relacionados com uma generalização da divergência de Kullback-Leibler (KULLBACK and LEIBLER, 1951), a chamada φ-divergência, a qual foi introduzida em (VIGELIS and CAVALCANTE, 2013b). A φ-divergência é dada por

$$\mathcal{D}_{\varphi}(p\|q) = \frac{\displaystyle\int_T \varphi^{-1}(p) - \varphi^{-1}(q)}{\displaystyle\int_T \frac{u_0}{(\varphi^{-1})'(p)} d\mu} d\mu.$$

No caso em que $\varphi(\cdot)$ é a função exponencial e $u_0 = 1$, a φ-divergência se reduz a divergência de Kullback-Leibler. Sobre algumas condições, os limites (4.18) e (4.19) são de valor finito e convergem para a φ-divergência:

$$\mathcal{D}_{\varphi}^{(0)}(q\|p) = \mathcal{D}_{\varphi}^{(1)}(p\|q) = \mathcal{D}_{\varphi}(p\|q) < \infty.$$

Estas condições são determinadas na Proposição 4.10 para o caso envolvendo a divergência de Rényi generalizada.

Proposição 4.10 ((SOUZA, VIGELIS, and CAVALCANTE, 2016, Proposição 3)). *Assumindo que $\varphi(\cdot)$ é continuamente diferenciável, considere a condição*

$$\int_T \varphi(\alpha\varphi^{-1}(p) + (1-\alpha)\varphi^{-1}(q)) d\mu < \infty. \tag{4.20}$$

Se a expressão (4.20) é satisfeita para todo $\alpha \in [\alpha_0, 0)$ e algum $\alpha_0 < 0$, então

$$\mathcal{D}_{\varphi}^{(\alpha)}(p\|q) = \frac{\partial\kappa}{\partial\alpha}(0) = \mathcal{D}_{\varphi}(p\|q) < \infty.$$

Se a expressão (4.20) é satisfeita para todo $\alpha \in (1, \alpha_0]$ e algum $\alpha_0 > 1$, então

$$\mathcal{D}_{\varphi}^{(1)}(p\|q) = -\frac{\partial\kappa}{\partial\alpha}(1) = \mathcal{D}_{\varphi}(p\|q) < \infty.$$

Desde que $\varphi(\cdot)$ seja convexa, a expressão (4.20) sempre ocorre para $\alpha \in [0,1]$. Para uma demonstração da Proposição 4.10 nos referimos a (SOUZA, VIGELIS, and CAVALCANTE, 2016, Lema 2, Proposição 3).

Para generalizarmos a divergência de Rényi precisamos que $\kappa(\alpha)$ esteja bem definida. Para garantir a existência e unicidade de $\kappa(\alpha)$ como definido em (4.17) assumimos que existe $u_0 : T \to (0, \infty)$ tal que

$$\int_T \varphi(c + \lambda u_0) d\mu < \infty, \qquad \text{para todo } \lambda{>}0, \tag{4.21}$$

para cada função mensurável $c : T \to \mathbb{R}$ satisfazendo $\int_T \varphi(c) d\mu < \infty$. Em outras palavras, a existência e unicidade de $\kappa(\alpha)$ é equivalente à condição (a3') da definição de exponencial deformada.

O próximo resultado mostra que a condição (a3') e a existência de $\kappa(\alpha)$, são equivalentes.

Proposição 4.11. *Assuma que a medida μ é não atômica. Fixe qualquer $\alpha \in (0, 1)$. Uma função $\varphi : \mathbb{R} \to [0, \infty)$, satisfazendo (a1) e (a2) na Definição 3.13 e uma função mensurável $u_0 : T \to (0, \infty)$ satisfazem a condição (a3') se, e somente se, para cada distribuição de probabilidade p e q em $\mathcal{P}_\mu$, existe uma constante $\kappa(\alpha) \colon = \kappa(\alpha; p, q)$ tal que*

$$\int_T \varphi(\alpha \varphi^{-1}(p) + (1 - \alpha)\varphi^{-1}(q) + \kappa(\alpha) u_0) d\mu = 1.$$

Esta proposição é exatamente um caso particular da Proposição 4.5 para $\varphi = \varphi_1 = \varphi_2$.

Como vimos no Exemplo (3.16) nem toda função que satisfaz as condições (a1) e (a2) na Definição 3.13, aceita a existência de uma função u_0 satisfazendo a condição (a3'). Apresentaremos um critério equivalente a uma função φ e uma função u_0 que satisfazem a condição (a3').

Proposição 4.12. *Uma função $\varphi : \mathbb{R} \to [0, \infty)$, que satisfaz as condições (a1) e (a2) na Definição 3.13, satisfaz a condição (a3') se, e somente se, para alguma constante $\alpha \in (0, 1)$, podemos encontrar uma função mensurável $c : T \to \mathbb{R} \cup \{-\infty\}$ tal que $\int_T \varphi(c) d\mu < \infty$ e*

$$\alpha \varphi(u) \leq \varphi(u - u_0(t)), \qquad \textit{para todo } u \geq c(t), \tag{4.22}$$

para μ-q.t.p. $t \in T$.

Esta proposição por sua vez é um caso particular da Proposição 4.7 com $\varphi = \varphi_1 = \varphi_2$ e $\lambda = 1$. Assim, encontramos um critério que nos diz que, para que exponenciais deformadas $\varphi(\cdot)$, $k(\alpha)$ existe e é única. Agora, vamos mostrar que a função dada no Exemplo 3.14 é uma exponencial deformada, ou seja, satisfaz as condições (a1)-(a3'), utilizando a Proposição 4.12. Dada a função

$$\exp_\kappa(u) = \begin{cases} \left(\kappa u + \sqrt{1 + \kappa^2 u^2}\right)^{\frac{1}{\kappa}}, & \text{if } \kappa \neq 0, \\ \exp(u) & \text{if } \kappa = 0. \end{cases}$$

e sua inversa

$$\ln_\kappa(u) = \begin{cases} \frac{u^\kappa - u^{-\kappa}}{2\kappa}, & \text{if } \kappa \neq 0, \\ \ln(u) & \text{if } \kappa = 0. \end{cases}$$

Claramente a função satisfaz (a1) e (a2). Vamos verificar que existe $\alpha \in (0,1)$ e $\lambda > 0$ para os quais

$$\lambda \leq \log_\kappa(v) - \log_\kappa(\alpha v), \qquad \text{para todo } v > 0. \tag{4.23}$$

Seja

$$f(v) = \log_\kappa(v) - \log_\kappa(\alpha v),$$

a derivada de $f(\cdot)$ é dada por

$$f'(v) = \frac{v^{\kappa-1}[1-\alpha^\kappa] - v^{-\kappa-1}[\alpha^{-\kappa}-1]}{2}.$$

Encontramos um ponto crítico de $f(\cdot)$, o ponto

$$v_0 = \left(\frac{\alpha^{-\kappa}-1}{1-\alpha^\kappa}\right)^{\frac{1}{2\kappa}} > 0.$$

Pelo fato de que a derivada de $f(\cdot)$ é negativa para $0 < v \leq v_0$ e positivo para $v \geq v_0$, então a diferença $\log_\kappa(v) - \log_\kappa(\alpha v)$ atinge um mínimo em v_0; dado $\alpha \in (0,1)$, a desigualdade (4.23) é satisfeita para algum $\lambda > 0$. Inserindo $v = \exp_\kappa(u)$ em (4.23), podemos escrever

$$\alpha \exp_\kappa(u) \leq \exp_\kappa(u-\lambda), \qquad \text{para todo } u \in \mathbb{R}. \tag{4.24}$$

Se $n \in \mathbb{N}$ é tal que $n\lambda \geq 1$, então uma aplicação repetida de (4.24) produz

$$\alpha^n \exp_\kappa(u) \leq \exp_\kappa(u-n\lambda) \leq \exp_\kappa(u-1), \qquad \text{para todo } u \in \mathbb{R}.$$

A Proposição 4.12 implica que $u_0 = 1$ satisfaz a condição (a3').

Podemos concluir pela Proposição 4.9, que para toda função que satisfaz (4.14), então $k(\alpha)$ está bem definida.

Para a função $\varphi(\cdot)$ dada no Exemplo (3.16), segue que

$$\begin{aligned}\limsup_{u\to\infty} \frac{\varphi(u)}{\varphi(u-\lambda_0)} &= \limsup_{u\to\infty} \frac{e^{(u+1)^2/2}}{e^{(u-\lambda_0+1)^2/2}} \\ &= \limsup_{u\to\infty} e^{u\lambda_0 - (\lambda_0^2-\lambda_0)/2} = \infty,\end{aligned}$$

o que mostra que $\varphi(\cdot)$ não pode ser usado na generalização da divergência de Rényi.

Concluímos que, a Definição 3.13 de exponencial deformada é adequada não apenas para fornecermos uma estrutura de C^∞-variedade de Banach para $\mathcal{P}_\mu$, mas também a condição (a3') da definição é a condição necessária e suficiente para definirmos as geodésicas, ou seja, os arcos que conectam duas densidades de probabilidade. A partir dessa conexão, definimos uma generalização para a divergência de Rényi no caso não atômico. Vamos analisar agora o caso puramente atômico, ou seja, μ é uma medida contável no conjunto dos números naturais $T = \mathbb{N}$.

4.3.2 *Caso puramente atômico*

Neste caso, a notação muda um pouco. Teremos sequências e somatórios em vez de funções e integrais, respectivamente. Assim, a condição (a3') da definição de exponencial deformada é reescrita como.

(a3') Existe uma sequência $\{u_{0,i}\} \subset (0,\infty)$, tal que

$$\sum_{i=1}^{\infty} \varphi(c_i + \lambda u_{0,i}) < \infty, \quad \text{para todo } \lambda>0, \tag{4.25}$$

para cada sequência $\{c_i\} \subset \mathbb{R}$ tal que $\sum_{i=1}^{\infty} \varphi(c_i) < \infty$.

Podemos encontrar, assim como no caso não atômico, uma condição equivalente à condição (a3'). Para encontrar essa condição equivalente precisamos de um resultado antes.

Lema 4.13. *Supondo que não podemos encontrar $\alpha \in (0,1)$, $\epsilon > 0$ e uma sequência $\{c_i\} \subset \mathbb{R} \cup \{-\infty\}$ tal que $\sum_{i=1}^{\infty} \varphi(c_i) < \infty$ e*

$$\alpha\varphi(u) \leq \varphi(u - \lambda u_{0,i}), \quad \textit{para todo } u > c_i \textit{ com } \varphi(u - u_{0,i}) < \epsilon, \tag{4.26}$$

então existem sequências de números reais de valor finito $\{c_{n,i}\}$ e sequências de conjuntos disjuntos $\{A_n\}$ em $\mathbb{N}$, tal que

$$\frac{1}{2} \leq \sum_{i \in A_n} \varphi(c_{n,i}) \quad \textit{e} \quad \sum_{i \in A_n} \varphi(c_{n,i} - u_{0,i}) \leq 2^{-n}, \tag{4.27}$$

para cada $n \geq 1$.

Demonstração. Para cada $m \geq 1$, definimos as sequências $\{f_{m,i}\} \subset \mathbb{R} \cup \{-\infty\}$ por

$$f_{m,i} = \sup\{u \in \mathbb{R} : \ 2^{-m}\varphi(u) > \varphi(u - u_{0,i}) \ \text{ e } \varphi(u - u_{0,i}) \leq 2^{-m-1}\},$$

em que usamos a convenção $\sup \emptyset = -\infty$. Desde que (4.26) não é satisfeita, temos que $\sum_{i=1}^{\infty} \varphi(f_{m,i}) = \infty$ para cada $m \geq 1$. Vamos dividir em casos.

Caso 1. Existe uma sequência estritamente crescente $\{m_n\} \subset \mathbb{N}$ para qual o conjunto $B_n = \{i : \ \varphi(f_{m,i} - u_{0,i}) = 2^{-m_n-1}\}$ tem um número infinito de elementos. Então, podemos selecionar uma sequência estritamente crescente $\{i_n\} \subset \mathbb{N}$ tal que

$$2^{-m_n}\varphi(f_{m_n,i_n}) \geq \varphi(f_{m_n,i_n} - u_{0,i_n}) = 2^{-m_n-1},$$

o que implica $\varphi(f_{m_n,i_n}) \geq 1/2$. A expressão (4.27) segue com $c_{n,i} = f_{m_n,i}$ e $A_n = \{i_n\}$.

Caso 2. Existe uma sequência estritamente crescente $\{m_n\} \subset \mathbb{N}$ para qual o conjunto $B_n = \{i : \varphi(f_{m,i} - u_{0,i}) = 2^{-m_n-1}\}$ tem um número finito de elementos. Denote por $C_n = \mathbb{N} \setminus B_n = \{i : \varphi(f_{m_n,i} - u_{0,i}) < 2^{-m_n-1}\}$. Pela continuidade de $\varphi(\cdot)$, temos que $2^{-m_n}\varphi(f_{m_n,i_n}) = \varphi(f_{m_n,i_n} - u_{0,i_n})$ para todo $i \in C_n$. Pelo fato de $\varphi(f_{m_n,i_n}) \leq 1/2$ para cada $i \in C_n$, e $\sum_{i-1}^{\infty} \varphi(f_{m_n,i}) - \infty$ para todo $n \geq 1$, podemos encontrar uma sequência estritamente crescente $\{k_n\} \subset \mathbb{N}$ para qual o conjunto $A_n = C_n \cap \{k_{n-1}, \ldots, k_n - 1\}$ satisfaz

$$\frac{1}{2} \leq \sum_{i \in A_n} \varphi(f_{m_n,i}) \leq 1.$$

A segunda inequação acima juntamente com $2^{-m_n}\varphi(f_{m_n,i_n}) = \varphi(f_{m_n,i_n} - u_{0,i_n})$ implica que

$$\sum_{i \in A_n} \varphi(f_{m_n,i} - u_{0,i}) \leq 2^{-m_n}.$$

Assim, a expressão (4.27) segue com $c_{n,i} = f_{m_n,i}$. □

O Lema 4.13 acima é a contrapartida no caso puramente atômico do Lema 4.6.

Proposição 4.14. *Uma função $\varphi : \mathbb{R} \to [0, \infty)$, que satisfaz as condições (a1) e (a2) na Definição 3.13, e uma sequência $\{u_{0,i}\}$ satisfazem a condição (4.25) se, e somente se, para alguma constante $\alpha \in (0, 1)$ e $\epsilon > 0$, uma sequência $\{c_i\} \subset \mathbb{R}\{-\infty\}$ pode ser encontrada tal que $\sum_{i=1}^{\infty} \varphi(c_i) < \infty$ e*

$$\alpha\varphi(u) \leq \varphi(u - \lambda u_{0,i}), \quad \textit{para todo } u > c_i \textit{ com } \varphi(u - u_{0,i}) < \epsilon. \tag{4.28}$$

Demonstração. Para provar que a condição (4.25) implica a condição (4.28), basta utilizar a Proposição 4.7 com $\varphi_1 = \varphi_2$, usando o Lema 4.13 em vez de Lema 4.6.

Supondo que a inequação (4.28) é satisfeita. Seja $\{\tilde{c}_i\}$ uma sequência qualquer de números reais tal que $\sum_{i=1}^{\infty} \varphi(\tilde{c}_i) < \infty$. Denote $A = \{i : c_i + \tilde{u_{0,i}} \geq c_i\}$ e $B = \{i \in A : \varphi(\tilde{c}_i) \leq \epsilon\}$. A partir da inequação (4.22) podemos escrever

$$\begin{aligned} \alpha \sum_{i=1}^{\infty} \varphi(\tilde{c}_i + u_{0,i}) &\leq \alpha \sum_{i \in A \cap B} \varphi(\tilde{c}_i + u_{0,i}) + \alpha \sum_{i \in A \setminus B} \varphi(\tilde{c}_i + u_{0,i}) + \alpha \sum_{i \in T \setminus A} \varphi(c_i) \\ &\leq \sum_{i \in A} \varphi(\tilde{c}_i) + \alpha \sum_{i \in A \setminus B} \varphi(\tilde{c}_i + u_{0,i}) + \alpha \sum_{i \in T \setminus A} \varphi(c_i). \end{aligned} \tag{4.29}$$

Temos que a segunda soma em (4.29) é finita pois, o conjunto $T \setminus B$ é finito. Assim, temos que $\sum_{i=1}^{\infty} \varphi(\tilde{c}_i + nu_{0,i}) < \infty$ para todo $n \geq 1$ e portanto $\sum_{i=1}^{\infty} \varphi(\tilde{c}_i + \lambda u_{0,i}) < \infty$ para todo $\lambda > 0$. □

Com a proposição anterior, temos uma desigualdade que é equivalente a condição (a3') para medida μ puramente atômica, utilizando essa desigualdade vamos provar o próximo resultado. Vamos mostrar que, no caso de medida μ puramente atômica qualquer função que satisfaz as condições (a1) e (a2) na Definição 3.13, satisfaz a condição (4.25) e assim, pode ser usada para encontrar uma generalização da divergência de Rényi.

Proposição 4.15. *Seja $\varphi : \mathbb{R} \to [0, \infty)$ uma função que satisfaz as condições (a1) e (a2) na Definição 3.13. Podemos encontrar uma sequência $\{u_{0,i}\}$ para qual a condição (4.25) ocorre.*

Demonstração. Seja $\{\lambda_n\} \subset (0, \infty)$ uma sequência qualquer, decrescente convergindo para 0. Fixemos qualquer $\alpha \in (0, 1)$ e $\eta \in \mathbb{R}$ tal que $\alpha\varphi(\eta) < \varphi(\eta - \lambda_1)$. Denotando $\epsilon = \varphi(\eta - \lambda_1)$, definimos

$$\tilde{c}_n = \sup\{u \in \mathbb{R} : \ \alpha\varphi(u) > \varphi(u - \lambda_n) \ \text{ e } \varphi(u - \lambda_n) \leq \epsilon\}, \qquad \text{para cada } n \geq 1,$$

adotamos a convenção $\sup \emptyset = -\infty$. Claramente, a sequência $\{\tilde{c}_n\} \subset [-\infty, \eta)$ é decrescente. Queremos provar que $\varphi(\tilde{c}_n) \downarrow 0$. Se a sequência $\{\tilde{c}_n\}$ converge para algum $c > -\infty$, a inequação, $\alpha\varphi(\tilde{c}_n) \geq \varphi(\tilde{c}_n - \lambda_n)$ implica $\alpha\varphi(c) \geq \varphi(c)$ e então $\varphi(c) = 0$. No caso de $\tilde{c}_n \downarrow -\infty$, está claro que $\varphi(\tilde{c}_n) \downarrow 0$. Assim, podemos selecionar uma subsequência $c_i = \tilde{c}_{n_i}$ tal que $\sum_{i=1}^{\infty} \varphi(c_i) < \infty$ e

$$\alpha\varphi(u) \leq \varphi(u - u_{0,i}), \qquad \text{para todo } u > c_i \text{ com } \varphi(u) < \epsilon,$$

em que $u_{0,i} = \lambda_i$. Pela Proposição 4.14, segue que $\{u_{0,i}\}$ satisfaz a condição (4.25). □

Provamos que no caso de medida puramente atômica qualquer função que satisfaz as condições (a1) e (a2), satisfaz a condição (4.25) e pode ser usada para definir uma generalização da divergência de Rényi. No caso de medida não atômica é necessário e suficiente que a função $\varphi(\cdot)$ satisfaça as condições (a1)-(a3') da Definição 3.13.

Na próxima seção, vamos estudar os arcos abertos, ou seja, definidos em um intervalo I aberto de modo que $[0, 1] \subset I$.

4.4 Arcos exponenciais abertos generalizados

Em (CENA and PISTONE, 2007), o modelo exponencial maximal em $p \in \mathcal{P}_\mu$, definido por

$$\mathcal{E}(p) \colon = \{\mathrm{e}^{u - K_p(u)} p | u \in \mathcal{K}_p\},$$

em que $\mathcal{K}_p$ é o interior do domínio da função geradora de cumulante $K_p(\cdot)$, foi provado ser a componente de $\mathcal{P}_\mu$ conectada a p, ou seja, o conjunto de todas as distribuições q em $\mathcal{P}_\mu$ que são conectadas a p por um arco exponencial aberto. Nesse mesmo sentido, nesse texto, vamos definir arcos abertos em variedade estatística generalizada e relacionar com

φ-família de distribuições de probabilidade $\mathcal{F}_c^\varphi$ (ANDRADE *et al.*, 2017). Além disso, vamos provar algumas das propriedades envolvendo esses arcos abertos.

Definição 4.16. Para uma exponencial deformada $\varphi(\cdot)$ dizemos que p e q em $\mathcal{P}_\mu$ são φ-conectadas por um arco aberto se existe um intervalo aberto $I \supset [0,1]$ e uma constante $\kappa(\alpha)$ tal que

$$p(\alpha) = \varphi((1-\alpha)\varphi^{-1}(p) + \alpha\varphi^{-1}(q) + \kappa(\alpha)u_0)$$

pertence a $\mathcal{P}_\mu$ para cada $\alpha \in I$, em que $\kappa(\alpha)$ depende de α, p e q.

Para $\alpha \in (0,1)$, a constante $\kappa(\alpha)$ é a mesma encontrada em (4.17). Na proposição a seguir daremos uma definição equivalente a definição de distribuições φ-conectados por um arco aberto.

Proposição 4.17. *p, $q \in \mathcal{P}_\mu$ são φ-conectadas por um arco aberto se, e somente se, existem um intervalo aberto $I \supset [0,1]$ e uma variável aleatória $v \in L_c^\varphi$, tal que $p(\alpha) \propto \varphi(c+\alpha v)$ pertence a $\mathcal{P}_\mu$, para todo $\alpha \in I$ e $p(0) = p$ e $p(1) = q$.*

Demonstração. Vamos assumir que p, q são φ-conectadas por um arco aberto, i.e. , $\int_T \varphi((1-\alpha)\varphi^{-1}(p) + \alpha\varphi^{-1}(q))d\mu < \infty$, para todo $\alpha \in I$. Desde que

$$\begin{aligned}\int_T \varphi((1-\alpha)\varphi^{-1}(p) + \alpha\varphi^{-1}(q))d\mu &= \int_T \varphi(\alpha[\varphi^{-1}(q) - \varphi^{-1}(p)] + \varphi^{-1}(p))d\mu \\ &= \int_T \varphi(c+\alpha v)d\mu,\end{aligned}$$

em que $v = \varphi^{-1}(q) - \varphi^{-1}(p)$ e $\varphi(c) = p$, então $v \in L_c^\varphi$. Além disso, $p(\alpha) \propto \varphi(c+\alpha v)$ pertence a $\mathcal{P}_\mu$, para cada $\alpha \in I$ e $p(0) = \varphi(c) = p$ e $p(1) = q$. A recíproca segue imediatamente. Suponha que $q = p(1)$, temos $\varphi(c+v) = q$, então $v = \varphi^{-1}(q) - \varphi^{-1}(p)$, com $\varphi(c) = p = p(0)$. $\square$

A necessidade de definir o arco aberto, segue do fato de $v \in L_c^\varphi$. Como consequência da Proposição 4.17, temos que se $p, q \in \mathcal{P}_\mu$ são φ-conectadas por um arco aberto, então a variável aleatória v pertence a $\mathcal{K}_c^\varphi$, desde que $\int_T \varphi(c+\alpha v)d\mu < \infty$ para todo $\alpha \in (-\epsilon, 1+\epsilon)$. Com isso, podemos provar o seguinte resultado.

Corolário 4.18. *Seja $p, q \in \mathcal{P}_\mu$, em que $p = \varphi(c)$. Temos que $q \in \mathcal{F}_c^\varphi$ se, e somente se, p e q são φ-conectados por um arco aberto.*

Demonstração. Supondo que $q \in \mathcal{F}_c^\varphi$, então $q = \varphi(c + v - \psi(v)u_0)$ em que $v \in \mathcal{B}_c^\varphi$. Então, temos $\int_T \varphi(c+\alpha v)d\mu < \infty$ para todo $\alpha \in (-\epsilon, 1+\epsilon)$, deduzimos que $p(\alpha) \propto \varphi(c+\alpha v)$ é um arco aberto contendo p e q. Reciprocamente, supondo que p e q são φ-conectados por um arco aberto, pela Proposição 4.17 existe um intervalo aberto $I \supset [0,1]$ e $v \in \mathcal{K}_c^\varphi$ tal que $p(\alpha) \propto \varphi(c+\alpha v)$ pertence a $\mathcal{P}_\mu$ com $q = p(1)$. Se $v \in \mathcal{B}_c^\varphi$, então $q = \varphi(c+v) \in \mathcal{F}_c^\varphi$ e a prova acaba. Caso contrário, seja w tal que

$$w = v - \frac{\int_T v\varphi'(c)d\mu}{\int_T u_0\varphi'(c)d\mu}u_0,$$

assim, $\int_T w\varphi'(c)d\mu = 0$ e $w \in \mathcal{B}_c^\varphi$. Consequentemente, temos $q = \varphi(c+v) = \varphi(c+w)$ e $q \in \mathcal{F}_c^\varphi$. □

Com isso, provamos que para $\varphi(c) = p$, a φ-família de distribuições de probabilidade $\mathcal{F}_c^\varphi$ é o conjunto de todas as distribuições $q \in \mathcal{P}_\mu$ tal que q é φ-conectada à p por um arco aberto.

Corolário 4.19. *Seja $p = \varphi(c)$ e $q = \varphi(\tilde{c})$ tal que $p, q \in \mathcal{P}_\mu$ são φ-conectadas por um arco aberto. Então os espaços de Musielak-Orlicz L_c^φ e $L_{\tilde{c}}^\varphi$ são iguais como conjuntos.*

Demonstração. Segue a partir do Corolário 4.18 que p e q estão na mesma φ-família, então $\tilde{c} = c + u - \psi(u)u_0$ e por (VIGELIS and CAVALCANTE, 2013b, Lema 5) o resultado segue. □

Agora, vamos mostrar que conexão por arcos abertos generalizados é uma relação de equivalência.

Proposição 4.20. *A relação Definição 4.16 é uma relação de equivalência.*

Demonstração. As propriedades reflexividade e simetria seguem da definição e agora, vamos provar a transitividade. Considere $p, q, r \in \mathcal{P}_\mu$

$$p(t) \propto \varphi(c+tu), \qquad r(t) \propto \varphi(c+tv), \qquad t \in (-\varepsilon, 1+\varepsilon),$$

com $p(0) = \varphi(c) = p$, $p(1) = \varphi(c+u) = q$, $r(0) = \varphi(c) = p$, $r(1) = \varphi(c+v) = r$ com $u, v \in L_c^\varphi$. Temos que p é φ-conectada à q e r, respectivamente. Precisamos provar que q e r são φ-conectadas também. Considere

$$q(t) \propto \varphi(c + (1-t)u + tv) \propto \varphi(c + u + t(v-u))$$

definido com $c+u = \tilde{c}, p(t) \propto \varphi(\tilde{c}+t(v-u)), v-u \in L_c^\varphi$, tal que $q(0) = \varphi(\tilde{c}) = \varphi(c+u) = q$, $q(1) = \varphi(\tilde{c} + (v-u)) = \varphi(c+v) = r$. Portanto, q e r são φ-conectados. □

Por fim queremos provar que uma φ-família $\mathcal{F}_c^\varphi$ é convexa para alguma exponencial deformada φ.

Lema 4.21. *Seja φ uma exponencial deformada fixada. Assumindo que $(\varphi^{-1})''(x)$ é contínua a*

$$\frac{\alpha\varphi''(\alpha\varphi^{-1}(x)+k)}{\varphi''(\alpha\varphi^{-1}(x)+k)} \geq \frac{\varphi''(\varphi^{-1}(x))}{\varphi''(\varphi^{-1}(x))}, \tag{4.30}$$

então $F(x) = \varphi(\alpha\varphi^{-1}(x)+k)$, para algum $\alpha > 1$ fixado e $k \in \mathbb{R}$, é uma função convexa.

Demonstração. Sabemos que se $F''(x) \geq 0$, $\forall\alpha > 1$ e $\forall x$, então $F(x)$ é uma função convexa. Temos

$$F''(x) = \frac{\alpha^2\varphi''(\alpha\varphi^{-1}(x)+k)\varphi'(\varphi^{-1}(x)) - \alpha\varphi'(\alpha\varphi^{-1}(x)+k)\varphi''(\varphi^{-1}(x))}{[\varphi'(\alpha\varphi^{-1}(x))]^3},$$

pelo fato de que φ é uma função crescente $[\varphi'(\alpha\varphi^{-1}(x))]^3 > 0$. Consequentemente, temos $F''(x) \geq 0$ se, e somente se,

$$\alpha^2\varphi''(\alpha\varphi^{-1}(x)+k)\varphi'(\varphi^{-1}(x)) - \alpha\varphi'(\alpha\varphi^{-1}(x)+k)\varphi''(\varphi^{-1}(x)) \geq 0,$$

o que segue de (4.30). □

Proposição 4.22. *Seja $p \in \mathcal{P}_\mu$ tal que $\varphi(c) = p$. Assumindo que*

$$\frac{\alpha\varphi''(\alpha\varphi^{-1}(x)+k)}{\varphi''(\alpha\varphi^{-1}(x)+k)} \geq \frac{\varphi''(\varphi^{-1}(x))}{\varphi''(\varphi^{-1}(x))}$$

para algum $\alpha > 1$ fixado e $k \in \mathbb{R}$. Então, a φ-família de probabilidades $\mathcal{F}_c^\varphi$ é convexa.

Demonstração. Notemos que para qualquer $\varphi(\tilde{c}) = r \in \mathcal{F}_c^\varphi$, $\mathcal{F}_c^\varphi = \mathcal{F}_{\tilde{c}}^\varphi$. Suponha $q \in \mathcal{F}_c^\varphi$, e considere $p(\lambda) = \lambda p + (1-\lambda)q$ para qualquer $\lambda \in [0,1]$. Mostraremos que $p(\lambda) \in \mathcal{F}_c^\varphi$, $\forall\lambda \in [0,1]$ provando que $\int_T \varphi((1-\alpha)\varphi^{-1}(p) + \alpha\varphi^{-1}(p(\lambda))d\mu < \infty$ para $\alpha \in (-\varepsilon, 1+\varepsilon)$. Em outras palavras, vamos mostrar que $p(\lambda)$ e p são φ-conectados por um arco aberto para todo $\lambda \in [0,1]$.

Para $\alpha \in (0,1)$, devido a convexidade de φtemos

$$\begin{aligned}
\int_T \varphi((1-\alpha)\varphi^{-1}(p) + \alpha\varphi^{-1}(p(\lambda)))d\mu &\leq \int_T (1-\alpha)\varphi(\varphi^{-1}(p)) + \alpha\varphi(\varphi^{-1}(p(\lambda)))d\mu \\
&= \int_T (1-\alpha)p + \alpha p(\lambda)d\mu \\
&= (1-\alpha)\int_T p d\mu + \alpha\int_T p(\lambda)d\mu \\
&= 1.
\end{aligned}$$

Para $\alpha \in (-\epsilon, 0)$, de acordo com a convexidade de $\alpha\varphi^{-1}(x)$ a $\varphi(x)$, temos

$$\begin{aligned}
&\int_T \varphi((1-\alpha)\varphi^{-1}(p) + \alpha\varphi^{-1}(p(\lambda)))d\mu \\
&\qquad \leq \int_T \varphi(\lambda\alpha\varphi^{-1}(p) + (1-\lambda)\alpha\varphi^{-1}(q) + (1-\alpha)\varphi^{-1}(p))d\mu \\
&\qquad = \int_T \varphi(\lambda[\alpha\varphi^{-1}(p) + (1-\alpha)\varphi^{-1}(p)] + (1-\lambda)[\alpha\varphi^{-1}(q) + (1-\alpha)\varphi^{-1}(p)])d\mu \\
&\qquad \leq \int_T \lambda\varphi(\alpha\varphi^{-1}(p) + (1-\alpha)\varphi^{-1}(p)) + (1-\lambda)\varphi(\alpha\varphi^{-1}(q) + (1-\alpha)\varphi^{-1}(p))d\mu \\
&\qquad = \lambda \int_T \varphi(\varphi^{-1}(p))d\mu + (1-\lambda)\int_T \varphi(\alpha\varphi^{-1}(q) + (1-\alpha)\varphi^{-1}(p))d\mu \\
&\qquad = \lambda + (1-\lambda)\int_T \varphi(\alpha\varphi^{-1}(q) + (1-\alpha)\varphi^{-1}(p))d\mu,
\end{aligned}$$

desde que $q \in \mathcal{F}_c^{\varphi}$, temos pelo Corolário 4.18 que q e p são φ-conectadas. Consequentemente,

$$\int_T \varphi((1-\alpha)\varphi^{-1}(p) + \alpha\varphi^{-1}(p(\lambda)))d\mu < \infty$$

então $p(\lambda)$ e p são φ-conectadas por um arco aberto, para todo $\alpha \in (-\epsilon, 0)$.

Agora, se $\alpha \in (1, 1+\epsilon)$, pelo Lema 4.21, $F(x) = \varphi(\alpha\varphi^{-1}(x) + k)$ é uma função convexa, então

$$\varphi(\alpha\varphi^{-1}(\lambda x + (1-\lambda)y) + k) \leq \lambda\varphi(\alpha\varphi^{-1}(x) + k) + (1-\lambda)\varphi(\alpha\varphi^{-1}(y) + k), \tag{4.31}$$

em que $\lambda \in [0, 1]$ e k uma constante. Tomando $k = (1-\alpha)\varphi^{-1}(p)$, temos

$$\begin{aligned}
\int_T \varphi(\alpha\varphi^{-1}(p(\lambda)) + (1-\alpha)\varphi^{-1}(p))d\mu &\leq \lambda \int_T \varphi(\alpha\varphi^{-1}(p) + (1-\alpha)\varphi^{-1}(p)d\mu \\
&\quad + (1-\lambda)\int_T \varphi(\alpha\varphi^{-1}(q) + (1-\alpha)\varphi^{-1}(p)d\mu \\
&= \lambda + (1-\lambda)\int_T \varphi(\alpha\varphi^{-1}(q) + (1-\alpha)\varphi^{-1}(p))d\mu \\
&< \infty,
\end{aligned}$$

desde que $q \in \mathcal{F}_c^{\varphi}$ e, portanto, p e q são φ-conectadas por um arco aberto. $\square$

Nessa seção, generalizamos os arcos abertos exponencias definidos em (CENA and PISTONE, 2007), assim como demos uma generalização da divergência de Rényi para exponencias deformadas $\varphi(\cdot)$ como na Definição 3.13. Provamos que a φ-família $\mathcal{F}_c^{\varphi}$ coincide com a componente conectada a p, com $p = \varphi(c)$ e que sobre certas condições em $\varphi(\cdot)$ a φ-família $\mathcal{F}_c^{\varphi}$ é convexa, assim como foi feito para famílias de exponenciais em (CENA and PISTONE, 2007) e estudado mais recentemente em (SANTACROCE, SIRI, and TRIVELLATO, 2016). Lembrando que, no caso não atômico, esses resultados só foram possíveis para funções que satisfazem a condição (a3') na definição de exponencial deformada. No caso puramente atômico, qualquer função que satisfaça as condições (a1) e (a2), é uma exponencial deformada e pode ser usada para conectar duas densidades por um arco e como consequência, para generalizar a divergência de Rényi.

Queremos encontrar uma forma de conectar duas distribuições na variedade estatística generalizada, no mesmo sentido que foi feito em (CENA and PISTONE, 2007, Definição 13), por arcos mistura abertos. Em (CENA and PISTONE, 2007) uma variedade mistura foi obtida utilizando o Gâteaux-gradiente do funcional gerador de cumulantes (2.2), sendo suportada pelas funções com integral igual a 1, não necessariamente, densidades positivas. Dessa forma, o caminho que encontramos para essa conexão por arcos, nos quais os arcos mistura são um caso especial, depende do Gâteaux-gradiente da função normalizadora ψ que aparece na equação (3.18). A partir de agora, vamos concentrar nossos esforços em estudar propriedades dessa função de normalização ψ. Começaremos na próxima seção, vamos estudar o comportamento da função normalizadora ψ, considerando se a função φ satisfaz a condição (a3') da Definição (3.13) ou não.

5 O COMPORTAMENTO DA FUNÇÃO NOMALIZADORA ψ

Sabemos que existem funções que satisfazem as condições (a1) e (a2) da Definição 3.13, mas não satisfazem a condição (a3'), por exemplo, a função dada em (3.16). Em (VIGELIS and CAVALCANTE, 2013a), o comportamento da função normalizadora ψ, que aparece em (3.18), foi estudado em relação à condição Δ_2, considerando exponenciais deformadas. Nessa seção, vamos analisar como a função normalizadora ψ se comporta, considerando que a função φ satisfaz (a1) e (a2), mas não satisfaz (a3'). A nossa conjectura é que a função normalizadora se comporta de forma diferente próximo ao bordo do seu domínio com relação a condição (a3').

5.1 A condição Δ_2 e o comportamento da função normalizadora ψ próximo ao bordo do seu domínio

Para estudarmos o comportamento da função normalizadora ψ próximo ao bordo do seu domínio, precisamos relembrar alguns resultados obtidos em (VIGELIS and CAVALCANTE, 2013a) sobre o comportamento de ψ em relação a condição Δ_2.

Seja a função de Musielak-Orlicz dada por:

$$\Phi_c(t,u) = \varphi(t, c(t) + u) - \varphi(t, c(t)), \tag{5.1}$$

em que a função $\varphi(\cdot)$ é uma exponencial deformada, ou seja, satisfaz as condições (a1)-(a3') da Definição (3.18). Lembramos que, a função Φ_c satisfaz a condição Δ_2 ou $\Phi_c \in \Delta_2$ se uma constante $K > 0$ e uma função não negativa $f \in \tilde{L}_c^\varphi$ pode ser encontrada tal que

$$\alpha\Phi_c(t, 2u) \leq \Phi_c\Big(t, u\Big), \qquad \text{para todo } u \geq f(t), \quad \text{e } \mu\text{-q.t.p. } t \in T. \tag{5.2}$$

Sabemos também que, se $\Phi_c \in \Delta_2$, então $\int_T \varphi(c+u)d\mu < \infty$ para todo $u \in L_c^\varphi$, neste caso, $E_c^\varphi = L_c^\varphi = \tilde{L}_c^\varphi$ e então o conjunto $\mathcal{B}_c^\varphi = \mathcal{K}_c^\varphi \cap B_c^\varphi$ tem bordo $\partial\mathcal{B}_c^\varphi$ vazio (VIGELIS and CAVALCANTE, 2013a). Uma função $u \in \mathcal{B}_c^\varphi$ pertence ao bordo de $\mathcal{B}_c^\varphi$ se, e somente se, $\int_T \varphi(c+\alpha u)d\mu < \infty$, para todo $\alpha \in (0,1)$ e $\int_T \varphi(c+\alpha u)d\mu = \infty$, para cada $\alpha > 1$. Se uma função de Musielak-Orlicz não satisfaz a condição Δ_2 ou $\Phi_c \notin \Delta_2$, temos que $E_c^\varphi \subsetneq L_c^\varphi \subsetneq \tilde{L}_c^\varphi$ e portanto o bordo $\partial\mathcal{B}_c^\varphi$ de $\mathcal{B}_c^\varphi$ é diferente do vazio (VIGELIS and CAVALCANTE, 2013a, Proposição 5). Além disso, podemos encontrar funções $u_* \in \partial\mathcal{B}_c^\varphi$ tais que $\int_T \varphi(c+u_*)d\mu < \infty$ e funções $u^* \in \partial\mathcal{B}_c^\varphi$ tais que $\int_T \varphi(c+u^*)d\mu = \infty$ (VIGELIS and CAVALCANTE, 2013a, Proposição 5).

A função normalizadora ψ aparece em (3.18), pelo fato de que para $u \in \mathcal{K}_c^\varphi$ a função $\varphi(c+u)$ não necessariamente pertence a $\mathcal{P}_\mu$. Assim, se faz necessário uma função $\psi : \mathcal{K}_c^\varphi \to \mathbb{R}$, tal que

$$\varphi(c + u - \psi(u)u_0) \tag{5.3}$$

pertença a $\mathcal{P}_\mu$. Quando restringirmos ψ ao conjunto $\mathcal{B}_c^\varphi$ temos que $\psi(u) \geq 0$ (VIGELIS and CAVALCANTE, 2013b). Para uma função $u \in \partial\mathcal{B}_c^\varphi$, o comportamento da função normalizadora $\psi(\alpha u)$ quando $\alpha \uparrow 1$ depende de se $\int_T \varphi(c+u)d\mu < \infty$ ou se $\int_T \varphi(c+u)d\mu = \infty$. Em (VIGELIS and CAVALCANTE, 2013a, Proposição 6) esse comportamento foi parcialmente elucidado.

Proposição 5.1 ((VIGELIS and CAVALCANTE, 2013a, Proposição 6))**.** *Seja u uma função no bordo de $\mathcal{B}_c^\varphi$. Para $\alpha \in [0,1)$, denotamos $\psi_u(\alpha) := \psi(\alpha u)$, cuja derivada à direita é indicada por $(\psi_u)'_+(\alpha)$. Se $\int_T \varphi(c+u)d\mu < \infty$ então $\psi_u(\alpha) = \psi(\alpha u)$ converge para algum $\beta \in (0,\infty)$ quando $\alpha \uparrow 1$. Por outro lado, se $\int_T \varphi(c+u)d\mu = \infty$ então $(\psi_u)'_+(\alpha)$ tende a ∞ quando $\alpha \uparrow 1$.*

Como nosso primeiro resultado nessa seção, elucidamos totalmente o comportamento da ψ próximo ao bordo do seu domínio, para a parametrização (5.3) a partir de uma exponencial deformada φ.

Proposição 5.2. *Seja uma função $u \in \partial\mathcal{B}_c^\varphi$, tal que $\int_T \varphi(c+u)d\mu = \infty$, onde φ é uma exponencial deformada. Então $\psi(\alpha u) \to \infty$ quando $\alpha \uparrow 1$.*

Demonstração. Suponhamos que, para algum $\bar{\lambda} > 0$ a função $\psi(\alpha u) \leq \bar{\lambda}$ para todo $\alpha \in [0,1)$. Denotemos $A = \{u \geq 0\}$. Observemos que,

$$\int_A \varphi(c+\alpha u - \bar{\lambda}u_0)d\mu \leq \int_T \varphi(c+\alpha u - \bar{\lambda}u_0)d\mu \leq \int_T \varphi(c+\alpha u - \psi(\alpha u)u_0)d\mu = 1,$$

obtemos que $\int_A \varphi(c+u-\overline{\lambda}u_0)d\mu < \infty$. Adicionalmente está claro que

$$\int_{T\setminus A} \varphi(c+u-\overline{\lambda}u_0)d\mu \leq \int_{T\setminus A} \varphi(c)d\mu \leq 1.$$

Como um resultado temos que $\int_T \varphi(c+u-\overline{\lambda}u_0)d\mu < \infty$. Pela condição (a3') segue que $\int_T \varphi(c+u)d\mu < \infty$, o que é uma contradição. □

Portanto, para uma função de Musielak-Orlicz Φ_c, que não satisfaz a condição Δ_2, dada como em (5.1), para uma exponencial deformada $\varphi(\cdot)$. A função normalizadora se comporta da seguinte forma:

(i) para toda função $u \in \partial\mathcal{B}_c^{\varphi}$ tal que $\int_T \varphi(c+u)d\mu < \infty$, então $\psi(\alpha u) \to \beta$, com $\beta \in (0,\infty)$ quando $\alpha \uparrow 1$.

(ii) para toda função $u \in \partial\mathcal{B}_c^{\varphi}$ tal que $\int_T \varphi(c+u)d\mu = \infty$, então $\psi(\alpha u) \to \infty$ quando $\alpha \uparrow 1$.

A seguir, faremos o estudo do comportamento da função normalizadora ψ próximo ao bordo de $\mathcal{B}_c^{\varphi}$considerando uma função que satisfaz as condições (a1) e (a2) da Definição 3.13, mas não satisfaz a condição (a3') (ANDRADE *et al.*, 2017).

5.2 A definição da exponencial deformada φ e suas consequências

Para iniciar esse estudo, vamos provar que a condição (a3'), da definição de exponencial deformada, é equivalente à existência de constantes $\overline{\lambda}, \alpha > 0$ e uma função não negativa $f \in \tilde{L}_c^{\varphi}$ tal que

$$\alpha\Phi_c(t,u) \leq \Phi_{c-\overline{\lambda}u_0}(t,u), \qquad \text{for all } u > f(t). \tag{5.4}$$

Proposição 5.3. *Uma função mensurável u_0 satisfaz a condição (a3') da definição de exponencial deformada se, e somente se, para alguma função mensurável $c\colon T \to \mathbb{R}$ tal que $\int_T \varphi(c)d\mu < \infty$, podemos encontrar constantes $\overline{\lambda}, \alpha > 0$ e uma função não negativa $f \in \tilde{L}_c^{\varphi}$ tal que*

$$\alpha\Phi_c(t,u) \leq \Phi_{c-\overline{\lambda}u_0}(t,u), \qquad \textit{for all } u > f(t). \tag{5.5}$$

Demonstração. Suponhamos que u_0 satisfaz a condição (a3'). Seja $c : T \to \mathbb{R}$ uma função mensurável qualquer tal que $\int_T \varphi(c)d\mu < \infty$. Como u é uma função mensurável com $\int_T \varphi(c-\overline{\lambda}u_0+u)d\mu < \infty$ então

$$\int_T \varphi(c+u)d\mu = \int_T \varphi(c-\overline{\lambda}u_0+u+\overline{\lambda}u_0)d\mu < \infty.$$

Este resultado implica que $\tilde{L}^{\Phi_{c-\overline{\lambda}u_0}} \subset \tilde{L}^{\Phi_c}$. A inequação (5.5) segue a partir da Proposição 3.12.

Agora, supondo que a inequação (5.5) é satisfeita. Pela Proposição 3.12 temos que $\tilde{L}^{\Phi_{c-\lambda u_0}} \subset \tilde{L}^{\Phi_c}$. Portanto, $u \in \tilde{L}^{\Phi_c}$ implica que $u + \overline{\lambda} u_0 \in \tilde{L}^{\Phi_{c-\overline{\lambda} u_0}} \subset \tilde{L}^{\Phi_c}$. Ou, equivalentemente, se u é uma função mensurável tal que $\int_T \varphi(c+u)d\mu < \infty$, então $\int_T \varphi(c+u+\overline{\lambda}u_0)d\mu < \infty$. Como um resultado, concluímos que $\int_T \varphi(c+u+\lambda u_0)d\mu < \infty$ para todo $\lambda > 0$. Seja $\widetilde{c}\colon T \to \mathbb{R}$ uma função mensurável qualquer satisfazendo $\int_T \varphi(\widetilde{c})d\mu < \infty$ e denotamos $A = \{\widetilde{c} > c\}$. Assim, para cada $\lambda > 0$, segue que

$$\int_T \varphi(\widetilde{c} + \lambda u_0)d\mu = \int_T \varphi(c + (\widetilde{c} - c) + \lambda u_0)d\mu \leq \int_T \varphi(c + (\widetilde{c} - c)\chi_A + \lambda u_0)d\mu < \infty,$$

o que mostra que u_0 é dado como na Definição 3.13. □

Desta proposição, concluímos que, a condição (a3') não é satisfeita se, e somente se, existe uma função mensurável $u : T \to \mathbb{R}$ tal que $\int_T \varphi(c+u)d\mu = \infty$, mas $\int_T \varphi(c+u-\overline{\lambda}u_0)d\mu < \infty$ para algum $\bar{\lambda} > 0$. Queremos mostrar que, supondo uma função φ que não satisfaça (a3'), então podemos encontrar uma função $u \in \partial\mathcal{B}_c^{\varphi}$ com $\int_T \varphi(c+u)d\mu = \infty$, mas $\psi(\alpha u) \to \beta$, com $\beta \in (0,\infty)$, com $\alpha \uparrow 1$, ou seja, a função normalizadora se comporta de uma forma se φ satisfaz (a3'), como vimos na seção anterior, e de outra forma se a função φ não satisfaz (a3'). Para provar isso precisamos de uma resultado anterior.

Lema 5.4. *Considere $c : T \to [0,\infty)$ uma função mensurável tal que $\int_T \varphi(c)d\mu < \infty$. Suponha que, para cada $\lambda > 0$, não podemos encontrar $\alpha > 0$ e $f \in \tilde{L}^{\Phi_c}$ tal que*

$$\alpha\Phi_c(t,u) \leq \Phi_{c-\lambda u_0}(t,u), \quad \textit{para todo } u > f(t). \tag{5.6}$$

Então uma sequência estritamente decrescente $0 < \lambda_n \downarrow 0$, e sequências $\{u_n\}$ e $\{A_n\}$ de funções mensuráveis de valor finito e conjuntos mensuráveis disjuntos dois a dois, respectivamente, podem ser encontrados tais que

$$I_{\Phi_c}(u_n\chi_{A_n}) = 1, \quad \textit{e} \quad I_{\Phi_{c-\lambda_n u_0}}(u_n\chi_{A_n}) \leq 2^{-n}, \qquad \textit{para todo } n \geq 1. \tag{5.7}$$

Demonstração. Sejam $\{\lambda_m\}$ uma sequência estritamente decrescente tal que $0 < \lambda_n \downarrow 0$. Definimos as funções não negativas

$$f_m(t) = \sup\{u > 0 : 2^{-m}\Phi_c(t,u) > \Phi_{c-\lambda_m u_0}(t,u)\}, \text{ para todo } m \geq 1,$$

em que adotamos a convenção que $\sup\emptyset = 0$. Desde que (5.6) não é satisfeita, temos que $I_{\Phi_c}(f_m) = \infty$ para cada $m \geq 1$. Para cada número racional $r > 0$, defina os conjuntos mensuráveis

$$A_{m,r} = \{t \in T : 2^{-m}\Phi_c(t,r) > \Phi_{c-\lambda_m u_0}(t,r)\},$$

e a função simples $u_{m,r} = r\chi_{A_{m,r}}$. Para $r = 0$, estabeleça $u_{m,r} = 0$. Seja $\{r_i\}$ uma enumeração de números racionais não negativos com $r_1 = 0$. Defina a função simples não negativa $v_{m,k} = \max_{1\leq i\leq k} u_{m,r_i}$, para cada $m, k \geq 1$. Pela continuidade de $\Phi_c(t,\cdot)$ e $\Phi_{c-\lambda_m u_0}(t,\cdot)$, segue que $v_{m,k} \uparrow f_m$ quando $k \to \infty$. Em virtude do Teorema da Convergência Monótona (ISNARD, 2007, Teorema 5.34), para cada $m \geq 1$, podemos encontrar algum $k_m \geq 1$ tal que a função $v_m = v_{m,k_m}$ satisfaz $I_{\Phi_c}(v_m) \geq 2^m$. Claramente, temos que $\Phi_c(t, v_m(t)) < \infty$ e $2^{-m}\Phi_c(t, v_m(t)) \geq \Phi_{c-\lambda_m u_0}(t, v_m(t))$. Pelo Lema 3.11 existe uma sequência crescente $\{m_n\}$ de índices e uma sequência de conjuntos mensuráveis, disjuntos dois a dois $\{A_n\}$ tais que $I_{\Phi_c}(v_{m_n}\chi_{A_n}) = 1$. Tomando $\lambda_n = \lambda_m$, $u_n = v_{m_n}$ e A_n, obtemos (5.7). □

Agora podemos provar o principal resultado dessa seção.

Proposição 5.5. *Assuma que φ é uma função que satisfaz as condições (a1) e (a2), mas não satisfaz a condição (a3') na Definição 3.13, então existe $u \in \partial\mathcal{B}_c^\varphi$ tal que $\int_T \varphi(c+u)d\mu = \infty$, mas $\psi(\alpha u) \to \beta$, com $\beta \in (0,\infty)$, com $\alpha \uparrow 1$.*

Demonstração. Sejam $\{\lambda_n\}$, $\{u_n\}$ e $\{A_n\}$ como no Lema 5.4. Dado qualquer $\lambda > 0$, tome $n_0 \geq 1$ tal que $\lambda \geq \lambda_n$ para todo $n \geq n_0$. Denote o conjunto $B = T \setminus \bigcup_{n=n_0}^\infty A_n$, então definimos $u = \sum_{n=n_0}^\infty u_n\chi_{A_n}$. De (5.7), segue que

$$\begin{aligned}
\int_T \varphi(c+u-\lambda u_0)d\mu &= \int_B \varphi(c-\lambda u_0)d\mu + \sum_{n=n_0}^\infty \int_{A_n} \varphi((c-\lambda u_0)+u_n)d\mu \\
&= \int_B \varphi(c-\lambda u_0)d\mu \\
&\quad + \sum_{n=n_0}^\infty \left\{ \int_{A_n} \varphi(c-\lambda u_0)d\mu + I_{\Phi_{c-\lambda u_0}}(u_n\chi_{A_n}) \right\} \\
&\leq \int_T \varphi(c-\lambda u_0)d\mu + \sum_{n=n_0}^\infty 2^{-n} < \infty.
\end{aligned}$$

Consequentemente, para $\alpha \in (0,1)$, podemos escrever

$$\begin{aligned}
\int_T \varphi(c+\alpha u)d\mu &= \int_T \varphi\Big(c+\alpha(u-\lambda u_0) + (1-\alpha)\frac{\alpha\lambda}{1-\alpha}u_0\Big)d\mu \\
&\leq \alpha \int_T \varphi(c+u-\lambda u_0)d\mu + (1-\alpha)\int_T \varphi\Big(c+\frac{\alpha\lambda}{1-\alpha}u_0\Big)d\mu \\
&< \infty.
\end{aligned}$$

Por outro lado, para $\alpha \geq 1$, segue que

$$\begin{aligned}\int_T \varphi(c+\alpha u)d\mu &\geq \int_B \varphi(c)d\mu + \sum_{n=n_0}^{\infty}\int_{A_n}\varphi(c+u_n)d\mu \\ &\geq \int_B \varphi(c)d\mu + \sum_{n=n_0}^{\infty}\left\{\int_{A_n}\varphi(c)d\mu + I_{\Phi_c}(u_n\chi_{A_n})\right\} \\ &= \int_T \varphi(c)d\mu + \sum_{n=n_0}^{\infty} 1 = \infty.\end{aligned}$$

Podemos encontrar $\lambda' < 0$ tal que

$$w = \lambda' u_0 \chi_B + \sum_{n=n_0}^{\infty} u_n \chi_{A_n}$$

satisfaz $\int_T w\varphi'_+(c)d\mu = 0$. Claramente, $\int_T \varphi(c+w)d\mu = \infty$, $\int_T \varphi(c+\alpha w)d\mu < \infty$ para $\alpha \in (0,1)$ e $\int_T \varphi(c+\alpha w)d\mu = \infty$ para $\alpha > 1$, ou seja, $w \in \partial\mathcal{B}_c^{\varphi}$ e temos também que

$$\int_T \varphi(c+w-\lambda u_0)d\mu < \infty \quad \text{para algum } \lambda > 0 \text{ fixado.}$$

Suponha que $\psi(\alpha w) \uparrow \infty$, então para todo $K > 0$ existe $\delta > 0$ tal que $0 < |\alpha - 1| < \delta$ implica que $\psi(\alpha w) > K$. Seja $\lambda'' > \lambda$ tal que $\int_T \varphi(c+w-\lambda'' u_0)d\mu < 1$, tomando $K = \lambda''$ temos

$$\varphi(c+\alpha w - \psi(w)u_0) < \varphi(c+\alpha w_{\{w>0\}} - \lambda'' u_0) < \varphi(c+w_{\{w>0\}} - \lambda'' u_0),$$

que é uma função μ-integrável. Portanto, pelo teorema da convergência dominada (ISNARD, 2007, Teorema 7.22) temos

$$\lim_{\alpha\uparrow 1}\int_T \varphi(c+\alpha w - \lambda'' u_0)d\mu = \int_T \varphi(c+w-\lambda'' u_0)d\mu,$$

então

$$\begin{aligned}1 = \lim_{\alpha\uparrow 1}&\int_T \varphi(c+\alpha w - \psi(\alpha w)u_0)d\mu \\ &\leq \lim_{\alpha\uparrow 1}\int_T \varphi(c+\alpha w - \lambda'' u_0)d\mu = \int_T \varphi(c+w-\lambda'' u_0)d\mu < 1,\end{aligned}$$

o que é uma contradição. □

Concluímos nessa seção, que se a função φ não satisfaz a condição (a3') da Definição 3.13, então a função normalizadora ψ próximo ao bordo de $\mathcal{B}_c^{\varphi}$ se comporta de forma diferente. Enquanto que para um exponencial deformada $\varphi(\cdot)$, temos que

(i) $\forall\, u \in \partial\mathcal{B}_c^\varphi$ tal que $\int_T \varphi(c+u)d\mu = \infty$, então $\psi(\alpha u) \to \infty$, quando $\alpha \to 1$.

Para uma função $\varphi(\cdot)$ que satisfaz as condições (a1) e (a2), mas não satisfaz a condição (a3') da Definição 3.13, temos que existe $w \in \partial\mathcal{B}_c^\varphi$ tal que $\int_T \varphi(c+w)d\mu = \infty$ e $\psi(\alpha w)$ converge para um valor finito, quando $\alpha \to 1$.

Saber como a função normalizadora ψ se comporta próximo ao bordo de $\mathcal{B}_c^\varphi$ é crucial para garantirmos a existência de arcos mistura na variedade estatística generalizada.

Na próxima seção, chegamos a um dos resultados principais desse trabalho. Iremos conectar densidades de uma mesma φ-família por arcos que são, na verdade, um caso mais geral dos arcos mistura e iremos garantir que estes arcos estão bem definidos em variedades estatística generalizadas.

6 ARCOS MISTURA GENERALIZADOS

Com o intuito de encontrar condições para o estudo da dualidade representacional na variedade estatística generalizada, queremos agora definir outra classe de arcos que pode nos ajudar nesse caminho da dualidade no sentido da representação conjugada $u^* = \partial\psi(u) \leftrightarrow u = (\partial\psi^*)(u^*) = (\partial\psi)^{-1}(u)$, como refletida por $\psi \leftrightarrow \psi^*$ (ZHANG, 2004, 2013). Assim, vamos utilizar o Gâteaux-gradiente da função de normalização ψ para definir essa nova classe de arcos. A função normalizadora $\psi : \mathcal{K}_c^\varphi \to \mathbb{R}$ é convexa e Gâteaux-diferenciável (VIGELIS and CAVALCANTE, 2013b, Lema 10) e sua Gâteaux derivada é dada por (3.21).

6.1 O subdiferencial de uma função convexa

Nesta seção, vamos lembrar algumas propriedades de funções convexas de valor real estendida em espaços de Banach, isto é, funções com valores em $\mathbb{R} \cup \{\pm\infty\}$. Principalmente, discutiremos os subdiferenciais de funções convexas semicontínuas inferiormente e suas propriedades.

Seja E um espaço de Banach. Uma função f é uma função convexa em E, com o epigrafo (ASPLUND and ROCKAFELLAR, 1969)

$$\text{epi } f = \{(x,\alpha) : x \in E,\ \alpha \in \mathbb{R},\ \alpha \geq f(x)\}.$$

Se $f(x) > -\infty$ para todo x e $f(x) < +\infty$ para pelo menos um valor de x, chamamos f uma função própria. O conjunto

$$\text{dom } f = \{x \in E : f(x) < \infty\}$$

denota o *domínio efetivo* de f. Uma função $f : E \to (-\infty, \infty]$ é dita ser *semicontínua inferiormente* (s.c.i.) se para todo $\lambda \in \mathbb{R}$ o conjunto

$$[f \leq \lambda] = \{x \in E : f(x) \leq \lambda\}$$

é fechado.

Seja E^* o espaço dual de E. Um vetor $x^* \in E^*$ é dito ser um *subgradiente* de f em $x \in E$ se

$$(x^*, z) \leq f(x+z) - f(x) \quad \text{for all } z \in E.$$

Denotamos por $\partial f(x)$ o conjunto de subgradientes de f em x e o *subdiferencial* de f é a aplicação multivalor $x \mapsto \partial f(x)$ de E para E^*. Pela definição $\partial f(x)$ é sempre um subconjunto fechado e convexo de E^* para cada x. Supondo que f é uma função convexa finita em x. Tem se $x^* \in \partial f(x)$ se, e somente se,

$$(z, x^*) \leq f'(x; z), \quad \forall z \in E,$$

em que

$$f'(x; z) = \lim_{t \to 0^+} \frac{f(x+tz) - f(x)}{t}$$

é a derivada direcional de f em x na direção $z \in E$. O subdiferencial pode ser vazio nos pontos do dom f, então denotamos por

$$D(\partial f) = \{x \in E : \partial f(x) \neq \emptyset\},$$

o domínio de ∂f e temos que $D(\partial f) \subset \text{dom } f$. Dizemos que f é subdiferenciável em x para todo $x \in D(\partial f)$.

Uma função convexa $f : E \to (-\infty, \infty]$ é dita ser semicontínua inferiormente se para cada $\lambda \in \mathbb{R}$ o conjunto

$$[f \leq \lambda] = \{x \in E;\ f(x) \leq \lambda\}$$

é fechado (BREZIS, 2011). Seja f uma função própria, convexa e semicontínua inferiormente, então int dom $f \subset D(\partial f)$ (BARBU and PRECUPANU, 2012, Coralário 2.38). A conjugada de f é a função $f^* : E^* \to \overline{\mathbb{R}}$ definida por

$$f^*(x^*) = \sup\{(x, x^*) - f(x) : x \in E\}, \quad x^* \in E^*. \tag{6.1}$$

Observe que se f é própria, então "sup" em (6.1) pode ser restrito ao pontos $x \in \text{dom } f$. A conjugada f^* é uma função convexa e semicontínua inferiormente em E^* e juntamente com f satisfaz a bem conhecida desigualdade de Young

$$(x, x^*) \leq f(x) + f^*(x^*), \tag{6.2}$$

com a igualdade ocorrendo se, e somente se, $x^* \in \partial f(x)$. Se f é uma função semicontínua inferiormente, o subdiferencial ∂f^* da função conjugada f^* coincide com $(\partial f)^{-1}$(BARBU and PRECUPANU, 2012, Proposição 2.33).

É conhecido que se f é uma função convexa, própria e semicontínua inferiormente então

$$\text{int dom } f \subset D(\partial f) \subset \text{dom } f,$$

e foi mostrado em (BRØNDSTED and ROCKAFELLAR, 1965) que $D(\partial f)$ é, de fato, denso em dom f.

Fato 6.1 ((BAUSCHKE, BORWEIN, and COMBETTES, 2001, Corolário 2.19),(BORWEIN and VANDERWERFF, 2010, Corolário 7.2.3))**.** *Suponha $x \in D(\partial f)$. Então, $x \in$ int dom f se, e somente se, ∂f é localmente limitado em x.*

Fato 6.2 ((BAUSCHKE, BORWEIN, and COMBETTES, 2001, Lema 2.20),(BORWEIN and VANDERWERFF, 2010, Lema 7.2.4))**.** *Se int dom $f \neq \emptyset$ e $x \in D(\partial f) \setminus$ int dom f, então $\partial f(x)$ é ilimitado.*

Uma função $f : E \to E^*$ é dita ser Gâteaux-diferenciável em $x_0 \in E$ se existe uma aplicação linear $A : E \to E^*$ tal que

$$\lim_{t \to 0} \frac{1}{t} \| f(x_0 + tv) - f(x_0) - Av \| = 0,$$

para todo $v \in E$. A Gâteaux derivada ou Gâteaux-gradiente de f em x_0 é denotado por $A = \text{grad } f(x_0)$. O subdiferencial de uma função convexa está intimamente relacionado ao Gâteaux-gradiente. Podemos ver essa relação no seguinte fato.

Fato 6.3 ((BARBU and PRECUPANU, 2012, Proposição 2.40))**.** *Se a função convexa f é Gâteaux-diferenciável em $x_0 \in E$, então $\partial f(x_0)$ consiste de um único elemento $x^* =$ grad $f(x_0)$.*

Na próxima seção, nós investigamos o subdiferencial da função de normalização ψ, que aparece na equação (3.18). Este resultado será útil para garantir que o arco mistura generalizado está bem definido.

6.2 Subdiferencial da função de normalização ψ

Como já vimos anteriormente, a variedade mistura (CENA and PISTONE, 2007), (PISTONE, 2013b), é baseada no gradiente da função geradora de cumulantes. Queremos fornecer uma expressão para a qual podemos conectar duas densidades em uma mesma φ-família, nesse mesmo sentido, para φ uma exponencial deformada. Então, definimos os arcos pela expressão dada por (ANDRADE *et al.*, 2018):

$$p(t) = F^{-1}((1-t)F(p) + tF(q)), \tag{6.3}$$

em que

$$F(p) = \frac{\varphi'(\varphi^{-1}(p))}{\int_T u_0 \varphi'(\varphi^{-1}(p)) d\mu}, \tag{6.4}$$

$t \in [0,1]$ e para p, q pertencem a uma φ-família de distribuição de probabilidades. Podemos reescrever o funcional $F(p)$ como

$$\frac{\varphi'(c + u - \psi(u)u_0)}{\int_T u_0 \varphi'(c + u - \psi(u)u_0) d\mu}, \qquad u \in \mathcal{B}_c^{\varphi}, \tag{6.5}$$

com $p = \varphi_c(c + u - \psi(u)u_0)$ e a expressão (6.5) é o Gâteaux-gradiente de ψ em uma função $u \in \mathcal{B}_c^{\varphi}$. Lembrando que para a função exponencial $\varphi(u) = \exp(u)$, com $u_0 = 1$, temos que o funcional $F(p) = p$ e o arco (6.3) se torna o arco mistura $p(t) = (1 - t)p + tq$, já conhecido e estudado anteriormente nas famílias exponenciais (CENA and PISTONE, 2007),(SANTACROCE, SIRI, and TRIVELLATO, 2016),(SANTACROCE, SIRI, and TRIVELLATO, 2017) e (PISTONE, 2013b). Por esse fato, chamaremos esses arcos (6.3) de arcos mistura generalizados.

Para os arcos definidos pela expressão (6.3) estarem bem definidos é necessário que o conjunto dos funcionais (6.5) que pertencem ao subdiferencial da função de normalização ψ seja convexo. A partir de agora, vamos trabalhar para provar a convexidade desse conjunto de funcionais (6.5). Como pelo Fato 6.3, o Gâteaux-gradiente e o subdiferencial são relacionados, então como passo inicial vamos investigar o conjunto dos subgradientes da função de normalização ψ.

A partir da investigação feita na Seção 4.2, vamos considerar nessa busca pelos arcos mistura generalizados, funções que satisfazem as condições (a1)-(a3') da definição de exponencial deformada. Vamos considerar também, que as funções de Musielak-Orlicz dada da forma (3.17), não satisfazem a condição Δ_2. Consequentemente, temos que o bordo do domínio da φ-família $\partial\mathcal{B}_c^{\varphi}$ é não vazio (VIGELIS and CAVALCANTE, 2013a).

O domínio efetivo da função de normalização ψ é o conjunto

$$\operatorname{dom} \psi = \{u \in B_c^{\varphi} : \psi(u) < \infty\}.$$

Precisamos encontrar as funções que formam o domínio, então vamos provar algumas propriedades de ψ. É importante lembrar que vamos investigar a convexidade desse conjunto de funcionais (6.5) com a função ψ definida no conjunto

$$B_c^{\varphi} = \left\{ u \in L_c^{\varphi} : \int_T u \varphi'_+(c) d\mu = 0 \right\}, \tag{6.6}$$

pois neste conjunto garantimos a unicidade da representação de um elemento q pertencente a φ-família de distribuições de probabilidade $\mathcal{F}_c^\varphi$. Temos assim, o nosso primeiro resultado nessa seção.

Proposição 6.4. *A função de normalização $\psi : B_c^\varphi \to \mathbb{R} \cup \{\infty\}$ é semicontínua inferiormente.*

Demonstração. Dado $\alpha \in \mathbb{R}$, seja $C_\alpha = \{u \in B_c^\varphi : \psi(u) \leq \alpha\}$. Para provar a afirmação, é suficiente mostrar que C_α é fechado. Definimos um conjunto

$$B = \left\{ u \in B_c^\varphi : \int_T \varphi(c + u - \alpha u_0) d\mu \leq 1 \right\},$$

e vamos provar que B é um conjunto fechado e que $B = C_\alpha$. Seja $\{u_n\}$ uma sequência pertencente a B, tal que $\| u_n - u \|_{\Phi_c} \to 0$. Deste modo, $u_n \to u$, μ-q.t.p. Desde que φ é uma função contínua, temos que $\varphi(c + u_n - \alpha u_0) \to \varphi(c + u + \alpha u_0)$, μ-q.t.p. . Pelo lema de Fatou (ISNARD, 2007), segue que

$$\begin{aligned}
\int_T \varphi(c + u - \alpha u_0) d\mu &= \int_T \liminf_{n\to\infty} \varphi(c + u_n - \alpha u_0) d\mu \\
&\leq \liminf_{n\to\infty} \int_T \varphi(c + u_n - \alpha u_0) d\mu \\
&\leq 1
\end{aligned}$$

assim, $u \in B$ e B é um conjunto fechado. Agora, vamos provar que $B = C_\alpha$. Seja u uma função que pertence a C_α, então $\psi(u) < \alpha$. A função φ é uma função estritamente crescente, de modo que

$$\int_T \varphi(c + u - \alpha u_0) d\mu \leq \int_T \varphi(c + u - \psi(u) u_0) d\mu = 1,$$

assim, $u \in B$.

Supondo que existe $w \in B \setminus C_\alpha$, então $w \in B$, o que implica que $\int_T \varphi(c + w - \alpha u_0) d\mu \leq 1$ e $w \notin C_\alpha$, o que implica que $\psi(w) > \alpha$. Então

$$\int_T \varphi(c + w - \alpha u_0) d\mu > \int_T \varphi(c + w - \psi(w) u_0) d\mu = 1,$$

logo $\int_T \varphi(c + w - \alpha u_0) d\mu > 1$. Isto contradiz a suposição de que $w \in B$. Portanto $B = C_\alpha$ e C_α é fechado. $\square$

O resultado da proposição anterior é muito importante na nossa investigação, pois existem várias propriedades, extremamente úteis, de funções convexas, semicontínuas inferiormente. Pela investigação na Seção 5, sobre o comportamento de ψ no bordo, juntamente com o fato de a função $\varphi(\cdot)$ é uma exponencial deformada, ou seja, satisfaz as condições (a1)-(a3') e ainda ψ ser semicontínuo inferiormente, pela Proposição 6.4, então temos que existem funções u em $\partial\mathcal{B}_c^\varphi$, tais que $\psi(u) < \infty$. Consequentemente, o domínio efetivo da função de normalização é o conjunto

$$\text{dom } \psi = \mathcal{B}_c^\varphi \cup \{\partial\mathcal{B}_c^\varphi\}_{<\infty} \tag{6.7}$$

em que $\{\partial\mathcal{B}_c^\varphi\}_{<\infty}$ é o conjunto dos pontos no bordo de $\mathcal{B}_c^\varphi$ tal que $\psi(u) < \infty$.

O subdiferencial de ψ em uma função $u \in \text{dom } \psi$ é o conjunto

$$\partial\psi(u) = \left\{ u^* \in (L^{\Phi_c})^* : \int_T u^* v d\mu \leq \psi(u+v) - \psi(u), \text{ for all } v \in B_c^\varphi \right\}, \tag{6.8}$$

em que $(L^{\Phi_c})^*$ denota o espaço dual de L^Φ. Lembrando que o espaço dual é dado pela soma direta (3.15) da contínua em ordem $L^{\Phi_c^*}$ com a componente singular $\left(L^{\Phi_c}\right)_s^\sim$. Nós sabemos que , para todo $u \in \mathcal{B}_c^\varphi$ a função de normalização ψ é Gâteaux-diferenciável a o Gâteaux-gradiente é dado por (6.5) (VIGELIS and CAVALCANTE, 2013b, Lema 10). Consequentemente, pelo Fato 6.3 $\partial\psi(u)$ consiste de um único elemento, o Gâteaux-gradiente dado por

$$\partial\psi(u) = \frac{\varphi'(c + u - \psi(u)u_0)}{\int_T u_0\varphi'(c + u - \psi(u)u_0 d\mu}, \qquad u \in \mathcal{B}_c^\varphi$$

De fato, nós provamos a seguir que (6.5) pertence a $\partial\psi(u)$, para todo $u \in \mathcal{B}_c^\varphi$.

Proposição 6.5. *Seja u uma função pertencente ao dom ψ. Supondo que o funcional*

$$\frac{\varphi'(c + u - \psi(u)u_0)}{\int_T u_0\varphi'(c + u - \psi(u)u_0)d\mu} \tag{6.9}$$

pertence a $L^{\Phi_c^*}$, então (6.9) pertence a $\partial\psi(u)$.

Demonstração. Temos que o funcional (6.9) pertence a $L^{\Phi_c^*}$. Seja v uma função em B_c^φ tal que $\int_T \varphi(c + u + v)d\mu < \infty$. Em outras palavras, $u + v \in \text{dom } \psi$, então temos que $\int_T \varphi(c+u-\psi(u)u_0)d\mu = 1$ e $\int_T \varphi(c+u+v-\psi(u+v)u_0)d\mu = 1$. Assim, pela convexidade de φ temos

$$\int_T [v + (\psi(u+v) - \psi(u))u_0]\varphi'(c + u - \psi(u)u_0)d\mu \leq$$
$$\int_T \varphi(c + u - \psi(u)u_0)d\mu - \int_T \varphi(c + u + v - \psi(u+v)u_0)d\mu = 0.$$

Então

$$\int_T v\varphi'(c + u - \psi(u)u_0)d\mu \leq \int_T u_0\varphi'(c + u - \psi(u)u_0)d\mu(\psi(u+v) - \psi(u))d\mu$$

e portanto

$$\frac{\int_T v\varphi'(c+u-\psi(u)u_0)d\mu}{\int_T u_0\varphi'(c+u-\psi(u)u_0)d\mu} \leq \psi(u+v)-\psi(u). \tag{6.10}$$

Se $u+v \in B_c^{\varphi} \setminus \operatorname{dom} \psi$, então $\psi(u+v)=\infty$, então

$$\frac{\int_T v\varphi'(c+u-\psi(u)u_0)d\mu}{\int_T u_0\varphi'(c+u-\psi(u)u_0)d\mu} < \psi(u+v)-\psi(u).$$

Consequentemente, a inequação (6.10) ocorre para todo $v \in B_c^{\varphi}$ e o resultado segue. $\square$

Conhecemos assim, o subdiferencial de ψ das funções $u \in \mathcal{B}_c^{\varphi}$. Precisamos encontrar o subdiferencial de ψ for u no conjunto $\{\partial\mathcal{B}_c^{\varphi}\}_{<\infty}$. Sabemos que ψ é uma função convexa, própria e semicontínua inferiormente, então

$$\operatorname{int} \operatorname{dom} \psi \subset D(\partial\psi) \subset \operatorname{dom} \psi,$$

em que $\operatorname{int} \operatorname{dom} \psi = \mathcal{B}_c^{\varphi}$ e $D(\partial\psi) \setminus \operatorname{int} \operatorname{dom} \psi = \{\partial\mathcal{B}_c^{\varphi}\}_{<\infty}$. Desde que $\operatorname{int} \operatorname{dom} \psi \neq \emptyset$, então pelo Fato 6.2, para $u \in D(\partial\psi) \setminus \operatorname{int} \operatorname{dom} \psi$, $\partial\psi(u)$ é ilimitado.

Na demostração da Proposição 6.8, temos que todo funcional do tipo (6.9) que pertence ao espaço $L^{\Phi_c^*}$, pertence ao subdiferencial. Na verdade, podem existir funcionais do tipo (6.9) que não pertencem ao espaço dual $\left(L^{\Phi_c}\right)^*$ e consequentemente não pertencem ao subdiferencial. Desde que, estamos interessados em provar que o conjunto dos funcionais (6.5) é convexo e estes funcionais são contínuos em ordem. Precisamos analisar somente a parte contínua em ordem do subdiferencial, isto é, a parte do subdiferencial que pertence a $L^{\Phi_c^*}$.Precisamos investigar sobre que condições o funcional (6.9) pertence ou não a $L^{\Phi_c^*}$, para $u \in \{\partial\mathcal{B}_c^{\varphi}\}_{<\infty}$. Para isto, precisamos de alguns resultados preliminares.

Lema 6.6 ((VIGELIS and CAVALCANTE, 2014, Lema 3.11))**.** *Seja Φ_c uma função de Musielak-Orlicz que não satisfaz a condição Δ_2. Adicionalmente, assumindo que $\Phi_c(t, b_\Phi(t)) = \infty$ para μ-q.t.p. $t \in T$. Então existe uma sequência estritamente crescente $0 < \lambda_n \uparrow 1$, e sequências $\{u_n\}$ e $\{A_n\}$ de funções não negativas de valor finito, mensuráveis, e conjuntos mensuráveis disjuntos dois a dois, respectivamente, tais que*

$$I_{\Phi_c}(u_n\chi_{A_n}) = 1, \quad e \quad I_{\Phi_c}(\lambda_n u_n \chi_{A_n}) \leq 2^{-n}, \qquad \textit{para todo } n \geq 1.$$

Agora vamos provar que se a função de Musielak-Orlicz Φ_c não satisfaz a condição Δ_2, podemos encontrar uma função u que pertence a classe de Musielak-Orlicz $\tilde{L}^{\Phi_c}$, mas o funcional $\Phi'_{c_+}(t, u(t)$ não pertence ao espaço $L^{\Phi_c^*}$. Esse resultado é crucial para acharmos condições sobre as quais o funcional (6.9) pertence ou não ao espaço $L^{\Phi_c^*}$.

Proposição 6.7. *Seja Φ_c uma função de Musielak-Orlicz que não satisfaz a condição Δ_2 e que $\Phi_c(t, b_\Phi(t)) = \infty$ para μ-q.t.p. $t \in T$. Então podemos encontrar uma função não negativa $u \in \tilde{L}^{\Phi_c}$ tal que $I_{\Phi_c^*}(\Phi'_{c_+}(t, u(t))) = \infty$.*

Demonstração. Sejam $\{\lambda_n\}$, $\{u_n\}$ e $\{A_n\}$ dado como no Lema 6.6. Selecione uma subsequência $\{\lambda_{n_k}\} \subset \{\lambda_n\}$ para qual a série $\sum_{k=1}^{\infty}(1-\lambda_{n_k})$ converge, e $(1-\lambda_{n_k})+2^{-n_k}<1$ para todo $k \geq 1$. Pelo fato de que $\lambda \to I_{\Phi_c}(\lambda u_{n_k}\chi_{A_{n_k}})$ é contínua para $\lambda \in [0,1]$, podemos encontrar $\lambda'_k \in (\lambda_{n_k},1)$ tal que $I_{\Phi_c}(\lambda'_k u_{n_k}\chi_{A_{n_k}}) = (1-\lambda_{n_k})+2^{-n_k}$. Defina $u=\sum_{k=1}^{\infty}\lambda'_k u_{n_k}\chi_{A_{n_k}}$. Então podemos escrever

$$I_{\Phi_c}(u) = \sum_{k=1}^{\infty} I_{\Phi_c}(\lambda'_k u_{n_k}\chi_{A_{n_k}}) = \sum_{k=1}^{\infty}[(1-\lambda_{n_k})+2^{-n_k}] < \infty, \tag{6.11}$$

e

$$\begin{aligned}
\int_T u(t)\Phi'_{c_+}(t,u(t)))d\mu &= \sum_{k=1}^{\infty}\int_{A_{n_k}} \lambda'_k u_{n_k}(t)\Phi'_{c+}(t,\lambda'_k u_{n_k}(t))d\mu \\
&\geq \sum_{k=1}^{\infty}\lambda'_k \frac{1}{\lambda'_k-\lambda_{n_k}}[I_{\Phi_c}(\lambda'_k u_{n_k}\chi_{A_{n_k}}) - I_{\Phi_c}(\lambda_{n_k}u_{n_k}\chi_{A_{n_k}})] \\
&\geq \sum_{k=1}^{\infty}\frac{\lambda'_k}{1-\lambda_{n_k}}[(1-\lambda_{n_k})+2^{-n_k}-2^{-n_k}] \\
&= \sum_{k=1}^{\infty}\lambda'_k = \infty.
\end{aligned}$$

Consequentemente segue que

$$I_{\Phi_c^*}(\Phi'_{c_+}(t,u(t))) = \int_T u(t)\Phi'_{c_+}(t,u(t)))d\mu - I_{\Phi_c}(u) = \infty,$$

o que conclui a prova. $\square$

A proposição anterior deixa claro que podemos encontrar um $u \in \tilde{L}^{\Phi_c}$, em que $\Phi'_{c_+}(t,u(t)) \notin \tilde{L}^{\Phi_c^*}$. Seja u como na Proposição 6.7, vamos provar que para $\lambda \in [0,1]$, $I_{\Phi_c}(\lambda u) < \infty$ e para $\lambda > 1$, $I_{\Phi_c}(\lambda u) = \infty$. De fato, seja $u = \sum_{k=1}^{\infty}\lambda'_k u_{n_k}\chi_{A_{n_k}}$.

Para $\lambda \in (0,1)$, pela equação (6.11) temos

$$I_{\Phi_c}(\lambda u) \leq I_{\Phi_c}(u) = \sum_{k=1}^{\infty} I_{\Phi_c}(\lambda'_k u_{n_k}\chi_{A_{n_k}}) < \infty.$$

Para $\lambda > 1$, tomando um número natural $k_0 \geq 1$ tal que $\lambda\lambda'_{k_0} \geq 1$. Então podemos escrever

$$I_{\Phi_c}(\lambda u) = \sum_{k=1}^{\infty} I_{\Phi_c}(\lambda\lambda'_k u_{n_k}\chi_{A_{n_k}}) \geq \sum_{k=1}^{\infty} I_{\Phi_c}(u_{n_k}\chi_{A_{n_k}}) = \infty.$$

A partir do resultado da Proposição 6.7 podemos finalmente chegar a condição para a qual podemos encontrar uma função que pertence ao bordo de $\mathcal{B}_c^{\varphi}$, mas o funcional do tipo (6.9) dessa função não pertence ao espaço $L^{\Phi_c^*}$.

Proposição 6.8. *Seja Φ_c uma função de Musielak-Orlicz tal que , Φ_c satisfaz a condição ∇_2, não satisfaz a condição Δ_2 e $\Phi_c(t, b_\Phi(t)) = \infty$. Então podemos encontrar $w \in \partial\mathcal{B}_c^\varphi$ tal que*

$$\frac{\varphi'(c + w - \psi(w)u_0)}{\int_T u_0\varphi'(c + w - \psi(w)u_0)d\mu} \notin L^{\Phi_c^*} \tag{6.12}$$

Demonstração. Tomando $k_0 \geq 1$ e denotando o conjunto $B = T \setminus \bigcup_{k=k_0}^\infty A_{n_k}$, então definimos $\tilde{u} = \sum_{k=k_0}^\infty \lambda_k' u_{n_k}\chi_{A_{n_k}}$. Podemos escolher $\lambda' < 0$ tal que

$$w = \lambda' u_0 \chi_B + \tilde{u}$$

satisfaz $\int_T w\varphi'(c)d\mu = 0$. Em outras palavras, $w \in B_c^\varphi$. É fácil ver que $\int_T \varphi(c + \alpha w)d\mu < \infty$ para $\alpha \in (0, 1)$ e $\int_T \varphi(c + \alpha w)d\mu = \infty$ para $\alpha > 1$, então $w \in \partial\mathcal{B}_c^\varphi$. Ainda resta mostrar (6.12). A partir da Proposição 6.7 temos que

$$\int_T w\Phi_{c+}'(t, w(t))d\mu = \int_B \lambda' u_0(t)\Phi_{c+}'(t, \lambda' u_0)d\mu + \sum_{k=k_0}^\infty \int_{A_{n_k}} I_{\Phi_c}(\lambda_k' u_{n_k}\chi_{A_{n_k}})d\mu = \infty,$$

desde que

$$\begin{aligned}\int_B \lambda' u_0(t)\Phi_{c+}'(t, \lambda' u_0)d\mu &\leq \lambda' \int_B \Phi_c((\lambda' + 1)u_0) - \Phi_c(\lambda' u_0)d\mu \\ &\leq \lambda' \int_B \varphi(c + (\lambda' + 1)u_0) - \varphi(c + \lambda' u_0)d\mu \\ &< \infty.\end{aligned}$$

Assim,

$$I_{\Phi_c^*}\left(\Phi_{c+}'(t, w(t))\right) = \int_T w(t)\Phi_{c+}'(t, w(t))d\mu - I_{\Phi_c}(w) = \infty,$$

consequentemente $\varphi'(c + w) \notin \tilde{L}^{\Phi_c^*}$. Pelo Teorema 3.10 desde que $\Phi_c \in \nabla_2$, temos que $\Phi_c^* \in \Delta_2$ e portanto $\tilde{L}^{\Phi_c^*} = L^{\Phi_c^*}$. Concluímos que $\varphi'(c + w) \notin L^{\Phi_c^*}$. Como $L^{\Phi_c^*}$ é um conjunto linear , temos que (6.12) ocorre. $\square$

Como uma consequência da Proposição 6.8 temos que é possível encontrar $u \in \{\partial\mathcal{B}_c^\varphi\}_{<\infty}$ tal que

$$\frac{\varphi'(c + u - \psi(u)u_0)}{\int_T u_0\varphi'(c + u - \psi(u)u_0)d\mu} \notin L^{\Phi_c^*},$$

e portanto o funcional (6.9) não pertence ao subdiferencial da função normalizadora em u, ou seja,

$$\frac{\varphi'(c + u - \psi(u)u_0)}{\int_T u_0\varphi'(c + u - \psi(u)u_0)d\mu} \notin \partial\psi(u)$$

.

Nesta seção, concluímos que para $u \in$ dom ψ nem todo o funcional do tipo

$$\frac{\varphi'(c+u-\psi(u)u_0)}{\displaystyle\int_T u_0\varphi'(c+u-\psi(u)u_0)d\mu} \tag{6.13}$$

pertence a $L^{\Phi_c^*}$, mas se o funcional pertence ao espaço $L^{\Phi_c^*}$, então o funcional pertence ao subdiferencial $\partial\psi(u)$.

Na próxima seção, finalmente provaremos que o conjunto formado pelos funcionais pertencentes ao subdiferencial da função de normalização ψ é convexo. Garantimos assim, que o arco mistura generalizado como definido em (6.3) está bem definido.

6.3 Convexidade do conjunto de funcionais

Vamos mostrar que o conjunto dos funcionais

$$\left\{\frac{\varphi'(c+u-\psi(u)u_0)}{\int_T u_0\varphi'(c+u-\psi(u)u_0)d\mu}, \quad u \in \text{dom } \psi\right\} \bigcap L^{\Phi_c^*} \tag{6.14}$$

é convexo. Da Proposição 6.5, o conjunto (6.14) está contido na imagem de $\partial\psi(\cdot)$, o conjunto dado por

$$\text{Im } \partial\psi = \bigcup \{\partial\psi(u) : u \in \text{dom } \psi\}. \tag{6.15}$$

Sabemos que nem sempre o conjunto (6.15) é convexo.

Em (ROCKAFELLAR, 1970, Teorema 5) foi mostrado que sejam X um espaço de Banach reflexivo e uma função convexa, própria, semicontínua inferiormente, Fréchet diferenciável e que satisfaz a condição que o domínio do subdiferencial está contido no interior do do domínio da função f. Então a imagem da aplicação grad $f : D \to X^*$ é um conjunto convexo, em que D é o domínio do subgradiente. Mas no nosso caso, o espaço de Musielak-Orlicz L^{Φ_c} não é reflexivo para qualquer função Φ_c, na verdade o espaço L^{Φ_c} só é reflexivo se a função Φ_c satisfazer a condição Δ_2 e a condição ∇_2 ao mesmo tempo, o que não é caso em que estamos trabalhando. Além disso, não temos que a função de normalização ψ é Fréchet diferenciável. Tivemos que encontrar outro meio de chegar a essa convexidade do conjunto (6.14).

Seja ψ^* a função conjugada de ψ. Pelo fato que ψ^* é uma função convexa, semicontínua inferiormente e própria, int dom ψ^* e dom ψ^* são conjuntos convexos e a imagem de $\partial\psi$ é o domínio efetivo de $\partial\psi^*$, desde que $(\partial\psi)^{-1} = \partial\psi^*$ (BARBU and PRECUPANU, 2012, Proposição 2.33). Assim,

$$\text{int dom } \psi^* \subset D(\partial\psi^*) \subset \text{dom } \psi^* \tag{6.16}$$

é o mesmo que

$$\text{int dom } \psi^* \subset \text{Im } \partial\psi \subset \text{dom } \psi^*. \tag{6.17}$$

A fim de provar que o conjunto (6.14) é convexo, analisamos o conjunto (6.15) em três casos. Sejam u^*, v^* elementos em (6.15) tais que

Caso 1. u^*,v^* $\in$ int dom ψ^*, então pela convexidade de int dom ψ^*, para $\lambda \in (0,1)$, temos $\lambda u^* + (1-\lambda)v^* \in$ int dom ψ^*.

Caso 2. Se $u^* \in$ int dom ψ^* e $v^* \in D(\partial\psi^*) \setminus$ int dom ψ^*, então $\lambda u^* + (1-\lambda)v^* \in$ int dom ψ^*, para $\lambda \in (0,1)$ (BAUSCHKE, BORWEIN, and COMBETTES, 2001, Fato 2.1).

Caso 3. Sejam u^*, v^*elementos em (6.15) pertencentes a $D(\partial\psi^*) \setminus$ int dom ψ^*.

Fica claro aqui, que nos Casos 1 e 2 a convexidade é satisfeita, pois a combinação convexa de elementos do conjunto 6.15 é ainda um elemento desse conjunto. Ficamos assim com o Caso 3 pra trabalhar essa convexidade. Queremos provar que , para $\lambda \in (0,1)$, $\lambda u^* + (1-\lambda)v^*$ pertence a (6.15). A fim de resolver este problema nós vamos provar que $D(\partial\psi^*) =$ int dom ψ^*. Assumindo que φ é uma função estritamente convexa, então ψ é uma função estritamente convexa. Na próxima proposição mostramos que $\partial\psi^*(u^*)$ é um conjunto unitário.

Proposição 6.9. *Seja ψ uma função estritamente convexa, então $\partial\psi^*(u^*)$ é um conjunto unitário, em que $u^* \in \partial\psi(u)$, com $u \in D(\partial\psi)$.*

Demonstração. Assumindo que ψ é uma função estritamente convexa temos que para $\lambda \in (0,1)$ e $\forall u_1 \neq u_2 \in$ dom ψ

$$\psi(\lambda u_1 + (1-\lambda)u_2) < \lambda\psi(u_1) + (1-\lambda)\psi(u_2). \tag{6.18}$$

Supondo que $\partial\psi^*(u^*)$ não é um conjunto unitário, isto é, $\partial\psi^*(u^*) = \{u_1, u_2, \ldots\}$, em que $u_i^* \in D(\partial\psi)$, $i = 1, 2, \ldots$. Tomando $u_1, u_2 \in \partial\psi^*(u^*)$. Pela desigualdade de Young (6.2)

$$(\lambda u_1 + (1-\lambda)u_2, u^*) \leq \psi(\lambda u_1 + (1-\lambda)u_2) + \psi^*(u^*), \tag{6.19}$$

em que $\lambda \in (0,1)$ e como consequência de $u_1, u_2 \in \partial\psi^*(u^*)$, temos

$$\psi(u_1) + \psi^*(u^*) = (u_1, u^*), \tag{6.20}$$

e

$$\psi(u_2) + \psi^*(u^*) = (u_2, u^*). \tag{6.21}$$

Tomando o produto de (6.20) por λ, o produto de (6.21) por $(1-\lambda)$ e adicionando as duas equações obtidas, temos

$$\lambda\psi(u_1) + (1-\lambda)\psi(u_2) + \psi^*(u^*) = (\lambda u_1 + (1-\lambda)u_2, u^*). \tag{6.22}$$

De (6.19) e (6.22) obtemos

$$\lambda\psi(u_1) + (1-\lambda)\psi(u_2) \leq \psi(\lambda u_1 + (1-\lambda)u_2),$$

o que é uma contradição por (6.18). O que implica que $\partial\psi^*(u^*)$ é um conjunto unitário e isto completa a prova. $\square$

Assim, o conjunto $\partial\psi^*(u^*)$ é unitário, então $\partial\psi^*$ é localmente limitado em $u^* \in D(\partial\psi^*)$ e portanto pelo Fato 6.1 concluímos que $u^* \in \text{int dom } \psi^*$ o que implica que $D(\partial\psi^*) \subset \text{int dom } \psi^*$, por (6.16) temos que

$$\text{Im } \partial\psi = D(\partial\psi^*) = \text{int dom } \psi^*. \tag{6.23}$$

Portanto, pelo Fato 6.2, não existe funcional u^* em (6.15) tal que $u^* \in D(\partial\psi^*)\backslash\text{int dom } \psi^*$. Assim, (6.15) é um conjunto convexo e como consequência, o arco mistura generalizado está bem definido, desde que o conjunto (6.14) é um conjunto convexo. De fato, sejam u, v funções no dom ψ tal que

$$u^* = \frac{\varphi'(c + u - \psi(u)u_0)}{\int_T u_0\varphi'(c + u - \psi(u)u_0)d\mu} \tag{6.24}$$

e

$$v^* = \frac{\varphi'(c + v - \psi(v)u_0)}{\int_T u_0\varphi'(c + v - \psi(v)u_0)d\mu} \tag{6.25}$$

pertencem a (6.14). Claramente,

$$\int_T u_0 u^* d\mu = \frac{\int_T u_0\varphi'(c + u - \psi(u)u_0)d\mu}{\int_T u_0\varphi'(c + u - \psi(u)u_0)d\mu} = 1$$

e

$$\int_T u_0 v^* d\mu = \frac{\int_T u_0\varphi'(c + v - \psi(v)u_0)d\mu}{\int_T u_0\varphi'(c + v - \psi(v)u_0)d\mu} = 1.$$

Notamos que, os funcionais (6.14) são os únicos elementos em (6.15) que satisfazem $\int_T u_0 u^* d\mu = 1$. Para $\lambda \in (0,1)$ temos

$$\int_T u_0((1-\lambda)u^* + \lambda v^*)d\mu = 1,$$

então existem funções $w_\lambda \in \text{dom } \psi$ tal que

$$\frac{\varphi'(c + w_\lambda - \psi(w_\lambda)u_0)d\mu}{\int_T u_0\varphi'(c + w_\lambda - \psi(w_\lambda)u_0)d\mu} = (1-\lambda)u^* + \lambda v^*, \text{ para cada } \lambda \in (0,1).$$

Assim, o conjunto (6.14) é um conjunto convexo.

Nesta seção, provamos que o arco mistura generalizado está bem definido para uma exponencial deformada φ estritamente convexa e para uma função de Musielak-Orlicz Φ_c que satisfaz a condição ∇_2 e não satisfaz a condição Δ_2. Na próxima seção, discutimos como podemos trabalhar com esse arco mistura generalizado aberto, ou seja, sem que necessariamente as densidades sejam os extremos do arco.

6.4 Arco aberto mistura generalizado

Na seção anterior, provamos que o arco mistura generalizado dado por

$$p(\alpha) = F^{-1}\left((1-\alpha)F(p) + \alpha F(q)\right), \tag{6.26}$$

está bem definido para $\alpha \in [0,1]$. Nesta seção nosso objetivo é duplo: primeiramente, garantir que o arco aberto também está bem definido; e em segundo lugar, fornecer algumas propriedades desses arcos. Para tais objetivos, usamos (6.23), que estabelece que $D(\partial\psi^*) = \operatorname{Im}\,\partial\psi$ é um conjunto aberto, então nós podemos estender a combinação convexa em (6.26) entre $F(p)$ e $F(q)$ além destes pontos extremos enquanto mantemos a positividade de $(1-\alpha)F(p) + \alpha F(q)$. De fato, pelo conjunto $D(\partial\psi^*) = \operatorname{Im}\,\partial\psi$ ser aberto, existe $\epsilon_1 > 0$ tal que $B(F(p), \epsilon_1)$ é a bola aberta de raio ϵ_1 centrada em $F(p)$ com $B(F(p), \epsilon_1) \subset D(\partial\psi^*)$. Similarmente, existe $\epsilon_2 > 0$ tal que $B(F(q), \varepsilon_2) \subset D(\partial\psi^*)$. Tomando $\epsilon = \min\{\epsilon_1, \epsilon_2\}$ garantimos que a combinação $(1-\alpha)F(p) + \alpha F(q)$ pode ser estendida a $\alpha \in I = (-\varepsilon, 1+\varepsilon) \supset [0,1]$.

Definição 6.10. Para um exponencial deformada φ, dizemos que p e q estão em $\mathcal{P}_\mu$ são φ-conectados por um arco mistura aberto se existe um intervalo $I \supset [0,1]$ tal que

$$p(\alpha) = F^{-1}\left((1-\alpha)F(p) + \alpha F(q)\right) \tag{6.27}$$

pertence a $\mathcal{P}_\mu$ para cada $\alpha \in I$, em que $F(p) = \frac{\varphi'(\varphi^{-1}(p))}{\int_T u_0 \varphi'(\varphi^{-1}(p))d\mu}$.

Em (CENA and PISTONE, 2007) foi mostrado que densidades conectadas por arcos mistura abertos tem razão limitada por constantes positivas. Em (SANTACROCE, SIRI, and TRIVELLATO, 2016) os autores mostraram a implicação recíproca, fornecendo uma caracterização de modelos mistura abertos. Aqui, pode se ver que o papel fundamental por ser conectado por arcos mistura abertos generalizados é dado pela razão $\frac{F(p)}{F(q)}$ a qual tem que ser limitada. O funcional $F(p)$ na definição de arco mistura aberto satisfaz $F(p) > 0$. Assim, a combinação $(1-\alpha)F(p) + \alpha F(q)$ em (6.27) tem que satisfazer a mesma propriedade, isto é, $(1-\alpha)F(p) + \alpha F(q) > 0$. Assuma que p e q são conectadas por um arco mistura aberto dado de acordo com (6.27) pertencente a $\mathcal{P}_\mu$ para todo $\alpha \in (-\varepsilon_1, 1+\varepsilon_2) \supset [-\varepsilon, 1+\varepsilon]$ com $\epsilon > 0$. Desde que $p(-\epsilon)$ e $p(1+\epsilon) \in \mathcal{P}_\mu$, então

$$F(p(-\varepsilon)) = (1+\varepsilon)F(p) + (-\varepsilon)F(q) > 0,$$

o que implica que

$$\frac{F(p)}{F(q)} > \frac{\varepsilon}{1+\varepsilon} \tag{6.28}$$

e

$$F(p(1+\varepsilon)) = (-\varepsilon)F(p) + (1+\varepsilon)F(q) > 0,$$

que nos dá

$$\frac{F(p)}{F(q)} < \frac{1+\varepsilon}{\varepsilon}. \tag{6.29}$$

Combinando ambas desigualdades (6.28) e (6.29), temos

$$\frac{\varepsilon}{1+\varepsilon} < \frac{F(p)}{F(q)} < \frac{1+\varepsilon}{\varepsilon}. \tag{6.30}$$

Reciprocamente, se temos (6.30), então $(1-\alpha)F(p) + \alpha F(q) > 0$ e (6.27) pertence a $\mathcal{P}_\mu$, Assim, temos que p e q são φ-conectadas por um arco mistura aberto se, e somente se, a razão $\frac{F(p)}{F(q)}$ é limitada. Pelo fato de que Im $\partial\psi$ é um conjunto aberto, então existe um intervalo $I \supset [0,1]$ tal que $(1-\alpha)F(p) + \alpha F(q)$ pertence a Im $\partial\psi$ e temos

$$\int_T u_0 \left[(1-\alpha)F(p) + \alpha F(q)\right] d\mu = 1, \text{ para todo } \alpha \in I \supset [0,1],$$

para todo $\alpha \in I$. Então, existem funções $w_\alpha \in$ dom ψ tal que

$$(1-\alpha)F(p) + \alpha F(q) = \frac{\varphi'(c + w_\alpha - \psi(w_\alpha)u_0)}{\int_T u_0\varphi'(c + w_\alpha - \psi(w_\alpha)u_0)d\mu},$$

com

$$p(\alpha) = F^{-1}\left(\frac{\varphi'(c + w_\alpha - \psi(w_\alpha)u_0)}{\int_T u_0\varphi'(c + w_\alpha - \psi(w_\alpha)u_0)d\mu}\right), \text{ for all } \alpha \in I \supset [0,1],$$

isto é, a combinação convexa (6.27) é também uma função do tipo (6.4) para todo $\alpha \in I$. Então, o arco mistura aberto está bem definido. Outra propriedade dessa conexão por arcos mistura abertos generalizados é que esta é uma relação de equivalência.

Proposição 6.11. *A relação na Definição 6.10 é uma relação de equivalência.*

Demonstração. As propriedades, reflexividade e simetria seguem da definição. Quanto a transitividade, considere p, q e $r \in \mathcal{P}_\mu$ tal que

$$p(\lambda) = F^{-1}\left((1-\lambda)F(p) + \lambda F(q)\right) \in \mathcal{P}_\mu \text{ e } q(\beta) = F^{-1}\left((1-\beta)F(q) + \beta F(r)\right) \in \mathcal{P}_\mu$$

com $\lambda, \beta \in [-\varepsilon, 1+\varepsilon]$ para algum $\epsilon > 0$. Podemos tomar

$$p(-\varepsilon) = F^{-1}\left((1+\varepsilon)F(p) + (-\varepsilon)F(q)\right) \text{ e } q(-\varepsilon) = F^{-1}\left((1+\varepsilon)F(q) + (-\varepsilon)F(r)\right),$$

e defina a distribuição de probabilidade

$$\begin{aligned} p_1 =& F^{-1}\Big(\Big(1-\frac{\varepsilon}{1+2\varepsilon}\Big)F(p(-\varepsilon))+\frac{\varepsilon}{1+2\varepsilon}F(q(-\varepsilon))\Big)\\ =& F^{-1}\Big(\frac{(1+\varepsilon)^2}{1+2\varepsilon}F(p)-\frac{\varepsilon^2}{1+2\varepsilon}F(r)\Big). \end{aligned}$$

Se temos

$$p(1+\varepsilon)=F^{-1}\left((-\varepsilon)F(p)+(1+\varepsilon)F(q)\right) \text{ e } q(1+\varepsilon)=F^{-1}\left((-\varepsilon)F(q)+1+\varepsilon)F(r)\right),$$

podemos definir uma distribuição de probabilidade como

$$\begin{aligned} p_2 =& F^{-1}\Big(\frac{-\varepsilon}{-1-2\varepsilon}F(p(1+\varepsilon))+\frac{-1-\varepsilon}{-1-2\varepsilon}F(q(1+\varepsilon))\Big)\\ =& F^{-1}\Big(\frac{-\varepsilon^2}{1+2\varepsilon}F(p)+\frac{(1+\varepsilon)^2}{1+2\varepsilon}F(r)\Big). \end{aligned}$$

O arco mistura aberto generalizado,

$$r(\alpha)=F^{-1}\left((1-\alpha)F(p_1)+\alpha F(p_2)\right),\ \alpha\in(0,1),$$

conecta as distribuições de probabilidade $r\big(\frac{(1+\varepsilon)^2}{2\varepsilon^2+2\varepsilon+1}\big)=p$ e $r\big(\frac{\varepsilon^2}{2\varepsilon^2+2\varepsilon+1}\big)=r$. O que garante a transitividade. $\square$

Nessa seção, propomos uma classe de arcos na variedade estatística generalizada, aos quais chamamos de arcos mistura generalizados pelo fato de os já conhecidos arcos mistura serem um caso especial desses. Para definir esses arcos, pensamos na dualidade no sentido da dualidade representacional. Assim, os arcos mistura generalizados são definidos a partir de funcionais no espaço de Musielak-Orlicz $L^{\Phi_c^*}$, que é o espaço dos funcionais contínuos em ordem definidos em L^{Φ}. Encontramos condições para obter esses arcos de forma que as densidades não sejam os pontos extremos desses arcos e garantimos que conectar duas densidades por um arco mistura generalizado é uma relação de equivalência.

7 CONCLUSÕES E PERSPECTIVAS FUTURAS

Variedade estatística generalizada é uma linha de pequisa em Geometria da Informação que vem se desenvolvendo nos últimos anos. Nesse trabalho, uma importante contribuição para o desenvolvimento dessas variedades foi dado. A construção de arcos no mesmo sentido dos arcos exponenciais e mistura. Com a construção desses arcos, um passo importante para o estudo da dualidade nas variedades estatísticas generalizadas foi dado.

Os arcos mistura foram construídos baseados no Gâteaux-gradiente da função normalizadora. Uma vez que a conexão por arcos é definida a partir de exponenciais deformadas, foi de extrema relevância, para a conexão por arcos, entender como a função de normalização ψ se comporta próximo ao bordo do seu domínio, considerando que a função de Musielak-Orlicz Φ_c é dada por uma função φ que satisfaz as condições (a1) e (a2), mas não satisfaz a condição (a3'), juntamente com o fato de ela não satisfazer a condição Δ_2.

Durante a investigação desses arcos generalizados, provamos que a generalização da divergência de Rényi proposta em (SOUZA, VIGELIS, and CAVALCANTE, 2016) está bem definida. Provamos que essa generalização está bem definida para arcos gerados a partir de exponenciais deformadas.

Investigações futuras irão iluminar como podemos definir transportes paralelos $\tau^{(1)}_{p_1,p_2}$ e $\tau^{(-1)}_{p_1,p_2}$ na variedade estatística generalizada, para estudarmos dualidade nessa variedade. Ajudarão também a entender como a divergência de Rényi generalizada pode ser relacionada com $\left(\tau^{(-1)}_{p_1,p_2}, \tau^{(1)}_{p_1,p_2}, \langle \cdot, \cdot \rangle\right)$. Podemos ainda pensar em como construir uma estrutura C^∞-diferenciável em $\mathcal{P}_\mu$ baseada em funções que não satisfazem a condição $(a3')$ ou ainda, baseada em funções que perdem a injetividade a partir de um certo valor, como acontece por exemplo em, q-exponencial de Tsallis. Uma ramo de investigação é utilizar a exponencial deformada para generalizar outras divergências, como a divergência gerada pelo q-logaritmo de Tsallis. Aplicações dessa variedade estatística generalizada são um dos principais objetivos de investigação.

REFERÊNCIAS

AMARI, S. Differential Geometry of Curved Exponential Families-Curvatures and Information Loss. **The Annals of Statistics**, v. 10, n. 2, p. 357–385, 1982.

AMARI, S. **Differential-geometrical methods in statistics**, v. 28 of *Lecture Notes in Statistics*. Springer-Verlag, New York, 1985, v+290 p.

AMARI, S. Information geometry on hierarchy of probability distributions. **IEEE Transactions on Information Theory**, v. 47, n. 5, p. 1701–1711, 2001.

AMARI, S. α -Divergence Is Unique, Belonging to Both f-Divergence and Bregman Divergence Classes. **IEEE Transactions on Information Theory**, v. 55, n. 11, p. 4925–4931, 2009.

AMARI, S. **Information geometry and its applications**, v. 194 of *Applied Mathematical Sciences*. Springer, [Tokyo], 2016, xiii+374 p.

AMARI, S.; CICHOCKI, A. Information geometry of divergence functions. **Bulletin of the Polish Academy of Sciences: Technical Sciences**, v. 58, n. 1, p. 183–195, 2010.

AMARI, S.; NAGAOKA, H. **Methods of information geometry**, v. 191 of *Translations of Mathematical Monographs*. American Mathematical Society, Providence, RI; Oxford University Press, Oxford, 2000, x+206 p. Translated from the 1993 Japanese original by Daishi Harada.

AMARI, S.; OHARA, A.; MATSUZOE, H. Geometry of deformed exponential families: Invariant, dually-flat and conformal geometries. **Physica A: Statistical Mechanics and its Applications**, v. 391, n. 18, p. 4308 – 4319, 2012.

ANDRADE, L. H. F.; VIEIRA, F. V. J.; VIGELIS, R. F.; CAVALCANTE, C. C. Mixture and Exponential Arcs on Generalized Statistical Manifold. **Entropy**, v. 20, n. 3, p. 147, 2018.

ANDRADE, L. H. F.; VIGELIS, R. F.; VIEIRA, F. V. J.; CAVALCANTE, C. C. *Normalization and ϕ -function: Definition and Consequences*, Cham: Springer International Publishing, p. 231–238. 2017.

ASPLUND, E.; ROCKAFELLAR, R. T. Gradients of convex functions. **Trans. Amer. Math. Soc.**, v. 139, p. 443–467, 1969.

BARBU, V.; PRECUPANU, T. **Convexity and optimization in Banach spaces**. Springer Monographs in Mathematics. Springer, Dordrecht, 4. ed., 2012, xii+368 p.

BAUSCHKE, H. H.; BORWEIN, J. M.; COMBETTES, P. L. Essential smoothness, essential strict convexity, and Legendre functions in Banach spaces. **Commun. Contemp. Math.**, v. 3, n. 4, p. 615–647, 2001.

BORWEIN, J. M.; VANDERWERFF, J. D. **Convex functions: constructions, characterizations and counterexamples**, v. 109 of *Encyclopedia of Mathematics and its Applications*. Cambridge University Press, Cambridge, 2010, x+521 p.

BOTELHO, G. M. A.; PELLEGRINO, D. M.; TEIXEIRA, E. V. **Introdução à Análise Funcional**. SBM, 2. ed., 2015.

BREZIS, H. **Functional analysis, Sobolev spaces and partial differential equations**. Universitext. Springer, New York, 2011, xiv+599 p.

BRØNDSTED, A.; ROCKAFELLAR, R. T. On the subdifferentiability of convex functions. **Proc. Amer. Math. Soc.**, v. 16, p. 605–611, 1965.

CALIN, O.; UDRIŞTE, C. **Geometric modeling in probability and statistics**. Springer, Cham, 2014, xxiv+375 p.

CENA, A.; PISTONE, G. Exponential statistical manifold. **Ann. Inst. Statist. Math.**, v. 59, n. 1, p. 27–56, 2007.

EGUCHI, S.; KOMORI, O. Path Connectedness on a Space of Probability Density Functions. Frank Nielsen; Frédéric Barbaresco (Ed.), **Geometric Science of Information**. Cham: Springer International Publishing, 2015, p. 615–624.

ERVEN, T. A. L. V. **When data compression and statistics disagree : two frequentist challenges for the minimum description length principle**. Mathematical Institute, Faculty of Science, Leiden University, 2010.

ERVEN, T. V.; HARREMOES, P. Renyi Divergence and Kullback-Leibler Divergence. **IEEE Transaction on Information Theory**, v. 60, n. 7, p. 3797–3820, 2014.

GIBILISCO, P.; PISTONE, G. Connections on Non-Parametric Statistical Manifolds by Orlicz Space Geometry. **Infinite Dimensional Analysis, Quantum Probability and Related Topics**, v. 01, n. 02, p. 325–347, 1998.

GINÉ, E.; NICKL, R. **Mathematical Foundations of Infinite-Dimensional Statistical Models**. Cambridge Series in Statistical and Probabilistic Mathematics. Cambridge University Press, 2015.

GRASSELLI, M. R. Dual connections in nonparametric classical information geometry. **Ann. Inst. Statist. Math.**, v. 62, n. 5, p. 873–896, 2010.

HARSHA, K. V.; MOOSATH, K. S. S. Dually flat geometries of the deformed exponential family. **Physica A:Statistical Mechanics and its applications**, v. 433, p. 136–147, 2015.

HAUSSLER, D.; OPPER, M. Mutual information, metric entropy and cumulative relative entropy risk. **Ann. Statist.**, v. 25, n. 6, p. 2451–2492, 1997.

HUDZIK, H.; ZBASZYNIAK, Z. Smoothness in Musielak-Orlicz spaces equipped with the Orlicz norm. **Collect. Math.**, v. 48, n. 4-6, p. 543–561, 1997. Fourth International Conference on Function Spaces (Zielona Góra, 1995).

ISNARD, C. **Introdução à Medida e Integração**. IMPA, 2007.

KANIADAKIS, G. Non-linear kinetics underlying generalized statistics. **Physica A: Statistical Mechanics and its Applications**, v. 296, n. 3, p. 405 – 425, 2001.

KRASNOSELĺSKI, M. A.; RUTICKI, J. B. **Convex functions and Orlicz spaces**. Translated from the first Russian edition by Leo F. Boron. P. Noordhoff Ltd., Groningen, 1961.

KULLBACK, S.; LEIBLER, R. A. On Information and Sufficiency. **Ann. Math. Statist.**, v. 22, n. 1, p. 79–86, 1951.

LOIZA, G.; QUICENO, H. R. A q-exponential statistical Banach manifold. **J. Math. Anal. Appl.**, v. 398, n. 2, p. 466–476, 2013.

MATSUZOE, H. Hessian structures on deformed exponential families and their conformal structures. **Differential Geometry and its applications**, v. 35, n. 1, p. 323–333, 2014.

MATSUZOE, H.; WADA, T. Deformed Algebras and Generalizations of Independence on Deformed Exponential Families. **Entropy**, v. 17, n. 8, p. 5729–5751, 2015.

MUSIELAK, J. **Orlicz spaces and modular spaces**, v. 1034 of *Lecture Notes in Mathematics*. Springer-Verlag, Berlin, 1983, iii+222 p.

NAUDTS, J. **Generalised thermostatistics**. Springer-Verlag London, Ltd., London, 2011, x+201 p.

NIELSEN, F.; NOCK, R. On w-mixtures: Finite convex combinations of prescribed component distributions., 2017. Disponível em: <https://arxiv.org/abs/1708.00568>. Acesso em: 09 jul. 2018.

PISTONE, G. kappa-exponential models from the geometrical viewpoint. **European Physical Journal B**, v. 70, n. 1, p. 29–37, 2009.

PISTONE, G. Examples of the application of nonparametric information geometry to statistical physics. **Entropy**, v. 15, n. 10, p. 4042–4065, 2013a.

PISTONE, G. Nonparametric information geometry. **Geometric science of information**, Springer, Heidelberg, v. 8085 of *Lecture Notes in Comput. Sci.*, p. 5–36. 2013b.

PISTONE, G.; ROGANTIN, M. P. The exponential statistical manifold: mean parameters, orthogonality and space transformations. **Bernoulli**, v. 5, n. 4, p. 721–760, 1999.

PISTONE, G.; SEMPI, C. An infinite-dimensional geometric structure on the space of all the probability measures equivalent to a given one. **Annals of Statistics**, v. 23, n. 5, p. 1543–1561, 1995.

RAO, M. M.; REN, Z. D. **Theory of Orlicz spaces**, v. 146 of *Monographs and Textbooks in Pure and Applied Mathematics.* Marcel Dekker, Inc., New York, 1991, xii+449 p.

RÉNYI, A. On measures of entropy and information. **Proc. 4th Berkeley Sympos. Math. Statist. and Prob., Vol. I**, Univ. California Press, Berkeley, Calif., p. 547–561. 1961.

ROCKAFELLAR, R. T. **Convex analysis**. Princeton Mathematical Series, No. 28. Princeton University Press, Princeton, N.J., 1970, xviii+451 p.

SANTACROCE, M.; SIRI, P.; TRIVELLATO, B. New results on mixture and exponential models by Orlicz spaces. **Bernoulli**, v. 22, n. 3, p. 1431–1447, 2016.

SANTACROCE, M.; SIRI, P.; TRIVELLATO, B. *On Mixture and Exponential Connection by Open Arcs*, Cham: Springer International Publishing, p. 577–584. 2017.

SOUZA, D. C.; VIGELIS, R. F.; CAVALCANTE, C. C. Geometry Induced by a Generalization of Rényi Divergence. **Entropy**, v. 18, n. 11, p. 407, 2016.

TOWENSED, J. T.; SOLOMON, B.; SMITH, J. S. The perfect gestalt: Infinite dimensional Riemannian face spaces and other aspects of face perception. Wenger, MJ and Townsend, JT (Ed.), **Computacional, Geometric and Process Perspectives on Facial Cognition: Contexs and Challenges**. Soc Math Psychol, 2001, p. 39–82.

TRIVELLATO, B. Deformed exponentials and applications to finance. **Entropy**, v. 15, n. 9, p. 3471–3489, 2013.

VIGELIS, R. F. **On Musielak-Orlicz Spaces and Applications to Information Geometry.** Department of Teleinformatics Engineering, Federal University of Ceará, Fortaleza-Brazil, 2011.

VIGELIS, R. F.; ANDRADE, L. H. F.; CAVALCANTE, C. C. *On the Existence of Paths Connecting Probability Distributions*, Cham: Springer International Publishing, p. 801–808. 2017.

VIGELIS, R. F.; CAVALCANTE, C. C. The Δ_2-condition and ϕ-families of probability distributions. **Geometric science of information**, Springer, Heidelberg, v. 8085 of *Lecture Notes in Comput. Sci.*, p. 729–736. 2013a.

VIGELIS, R. F.; CAVALCANTE, C. C. On ϕ-families of probability distributions. **J. Theoret. Probab.**, v. 26, n. 3, p. 870–884, 2013b.

VIGELIS, R. F.; CAVALCANTE, C. C. Smoothness of the Orlicz norm in Musielak-Orlicz function spaces. **Math. Nachr.**, v. 287, n. 8-9, p. 1025–1041, 2014.

ZHANG, J. Divergence Function, Duality, and Convex Analysis. **Neural Comput.**, v. 16, n. 1, p. 159–195, 2004.

ZHANG, J. From ϵ -entropy to KL-entropy: Analysis of minimum information complexity density estimation. **Ann. Statist.**, v. 34, n. 5, p. 2180–2210, 2006.

ZHANG, J. Nonparametric information geometry: from divergence function to referential-representational biduality on statistical manifolds. **Entropy**, v. 15, n. 12, p. 5384–5418, 2013.

ZHANG, J.; HÄSTÖ, P. Statistical manifold as an affine space: a functional equation approach. **J. Math. Psych.**, v. 50, n. 1, p. 60–65, 2006.