



**UNIVERSIDADE FEDERAL DO CEARÁ  
CENTRO DE CIÊNCIA E TECNOLOGIA  
DEPARTAMENTO DE ENGENHARIA DE TELEINFORMÁTICA  
PROGRAMA DE PÓS-GRADUAÇÃO EM ENGENHARIA DE TELEINFORMÁTICA**

**FRANCISCA LEIDMAR JOSUÉ VIEIRA**

**EXPONENCIAIS DEFORMADAS EM VARIEDADES ESTATÍSTICAS:  
CARACTERIZAÇÃO E ANÁLISE DE FAMÍLIAS DE ENTROPIAS**

**FORTALEZA**

**2020**

FRANCISCA LEIDMAR JOSUÉ VIEIRA

EXPONENCIAIS DEFORMADAS EM VARIEDADES ESTATÍSTICAS:  
CARACTERIZAÇÃO E ANÁLISE DE FAMÍLIAS DE ENTROPIAS

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 Vigellis

FORTALEZA

2020

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)

---

- V715e Vieira, Francisca Leidmar Josué.  
EXPONENCIAIS DEFORMADAS EM VARIEDADES ESTATÍSTICAS: CARACTERIZAÇÃO E ANÁLISE DE FAMÍLIAS DE ENTROPIAS / Francisca Leidmar Josué Vieira. – 2020.  
87 f. : il.
- Tese (doutorado) – Universidade Federal do Ceará, Centro de Tecnologia, Programa de Pós-Graduação em Engenharia Elétrica, Fortaleza, 2020.  
Orientação: Prof. Dr. Charles Casemiro Cavalcante .  
Coorientação: Prof. Dr. Rui Facundo Vigelis .
1. Exponenciais deformadas. 2. Entropia. 3. Divergência. 4. Função Normalizadora. I. Título.  
CDD 621.3
-

FRANCISCA LEIDMAR JOSUÉ VIEIRA

EXPONENCIAIS DEFORMADAS EM VARIEDADES ESTATÍSTICAS:  
CARACTERIZAÇÃO E AVALIAÇÃO DE FAMÍLIAS DE ENTROPIAS

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.

Aprovada em: 23 / 01 / 2020.

BANCA EXAMINADORA

---

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

---

Prof(a). Dr(a). Fátima Nelsizeuma Sombra de Medeiros  
Universidade Federal do Ceará (UFC)

---

Prof. Dr. Juvêncio Santos Nobre  
Universidade Federal do Ceará (UFC)

---

Prof. Dr. Renato J. Cintra  
Universidade Federal de Pernambuco (UFPE)

---

Prof(a). Dr(a). Sueli Irene Rodrigues Costa  
Universidade de Campinas (Unicamp)

Ao meu filho, esposo, pais e irmãos.

## AGRADECIMENTOS

Primeiramente agradeço a Deus por todo amor e por permanecer sempre comigo.

Agradeço ao meu filho, Miguel, por todo amor e alegria que trouxe à minha vida.

Ao meu esposo, Tiago, por todo amor, carinho e apoio ao longo de todo esse tempo.

Aos meus pais, Antonio e Moça, por todo amor, dedicação, carinho, incentivo e por serem exemplos de determinação.

Aos meus irmãos, Francisco, Aurimar, Aparecida e Damiana e aos meus sobrinhos Lívia e Edmundo por todo amor, carinho, amizade e pelos momentos de descontração. Em especial, a Damiana por toda confiança e apoio.

Ao meu cunhado, Francisco, pelas conversas e brincadeiras.

Ao professor Charles Casimiro Cavalcante pela paciência e acessibilidade sempre presentes na sua orientação.

Ao professor Rui Facundo Vigelis pela co-orientação.

Aos professores, Dra. Fátima Nelsizeuma Sombra de Medeiros, Dr. Juvêncio Santos Nobre, Dr. Renato J. Cintra e Dra. Sueli Irene Rodrigues Costa por terem aceitado o convite em participar da banca examinadora.

Aos colegas de Pós-graduação pelas discussões matemáticas: Amilcar, André, Damiana, David, Fagner, Joseph(Ilhano), Luiza, Yosbel.

Agradecimentos também ao Renato, secretário da Pós-graduação, por toda competência e agilidade.

“A vida é uma viagem a três estações: ação, experiência e recordação. ”

(Julio Camargo)

## RESUMO

Nesta tese, são estudadas variedades estatísticas generalizadas no contexto de geometria da informação. O conjunto  $P_\mu$  das funções de densidade de probabilidade estritamente positivas pode ser munido com uma estrutura de variedade de Banach de classe  $C^\infty$ . Uma das contribuições deste trabalho está na generalização dessa construção que usa funções exponenciais deformadas injetivas e suas consequências. Mais precisamente, a parametrização construída nessa tese foi feita usando uma função exponencial deformada não injetiva a qual assume zero até certo ponto e dali em diante é estritamente crescente. Essa construção permite parametrizar  $P_\mu$  usando uma maior diversidade de funções (exponenciais deformadas que tem como casos particulares, por exemplo, a exponencial clássica e  $\kappa$ -exponencial). A partir dessas exponenciais deformadas não injetivas é possível definir uma divergência na vizinhança parametrizada e, conseqüentemente, a  $\varphi$ -divergência para  $P_\mu$ . Deriva dessa divergência a possibilidade de fazer geometria usando métrica e conexões. Além disso, estas estão relacionadas com  $q$ -divergência, métrica e conexões obtidas por meio da função  $q$ -exponencial já conhecidas na literatura. Provamos também que a  $\kappa$ -exponencial e a  $q$ -exponencial podem ser usadas na generalização da divergência de Rényi. Outra contribuição deste trabalho consiste em introduzir novas  $\varphi$ -entropias que recaem em entropias já conhecidas. Além disso, resultados análogos foram provados para as entropias relativas. Para as funções exponenciais deformadas injetivas, na construção da parametrização aparece uma função chamada de normalizadora. Ainda nesta tese, elucidamos o estudo do comportamento da função normalizadora próximo ao bordo do seu domínio no caso em que os pontos analisados estão na classe Musielak-Orlicz. Esse caso era o que faltava para completar todos os possíveis casos que compreendem a classe de funções estudadas. Diante desses fatos, a principal contribuição dessa tese consiste no estudo local das variedades estatísticas, usando exponenciais deformadas não injetivas, para obter novas  $\varphi$ -entropias com propriedades análogas às entropias já conhecidas e investigadas em geometria da informação. Tais resultados permitem adquirir um leque maior de ferramentas apropriadas no estudo problemas que envolvem processamento de sinais.

**Palavras-chave:** geometria da informação; exponencial deformada; função normalizadora; entropia; variedade estatística.



## ABSTRACT

In this thesis we study generalized statistical manifolds in the context of information geometry. The set of the strictly positive probability density functions,  $P_\mu$ , can be provided with a  $C^\infty$ -Banach manifold. One contribution is in the generalization of this construct that uses injective deformed exponential functions. More precisely, the parametrization constructed in this thesis was made using a non-injective deformed exponential function which assumes zero until a certain point and from then on is strictly increasing. This construction allowed us to parameterize  $P_\mu$  using a greater diversity of functions. From these non-injecting deformed exponentials it is possible to define a divergence in the parameterized neighborhood and consequently the  $\varphi$ -divergence for  $P_\mu$ . This divergence brings the possibility of a geometrical investigation using metric and connections. Moreover, they are related to  $q$ -divergence, metric, and connections obtained through the  $q$ -exponential function already known in the literature. We also prove that the  $\kappa$ -exponential and  $q$ -exponential function can be used to generalize the Rényi divergence. Another contribution of this work is the introduction of new  $\varphi$ -entropies that reduce, as particular cases, to already known entropies. Similar results were proved for the relative entropy. Going back to the study of the injected deformed exponential functions in the construction of the parametrization it appears a normalization function. Still in this thesis, we elucidate the study of the behavior of the normalizing function near the boundary of its domain in the case where the analyzed points are in the Musielak-Orlicz class. This case was missing to complete all possible cases regarding this class of deformed functions. Given these facts, the main contribution of this thesis is the local study of statistical manifolds, using non-injected deformed exponentials, to obtain new  $\varphi$ -entropies with properties similar to the entropy already known and investigated in information geometry. These results allow us to acquire a wider range of appropriate tools for studying problems such as signal processing, optimization, numerical analysis, among others.

**Keywords:** information geometry; deformed exponential; normalizing function; entropy; statistical manifold.

**LISTA DE FIGURAS**

Figura 1 – Variedade de Banach . . . . .	17
Figura 2 – Função de Young . . . . .	18
Figura 3 – Função 2-exponencial . . . . .	20
Figura 4 – Função $\kappa$ -exponencial . . . . .	24
Figura 5 – Imagem em nível de cinza e seu histograma (ASSIRATI, 2014) . . . . .	58

## SUMÁRIO

<b>1</b>	<b>INTRODUÇÃO</b>	<b>11</b>
<b>1.1</b>	<b>Motivação</b>	<b>11</b>
<b>1.2</b>	<b>Objetivos e contribuições</b>	<b>13</b>
<b>1.3</b>	<b>Organização do documento</b>	<b>14</b>
<b>2</b>	<b>VARIEDADES ESTATÍSTICAS: FUNDAMENTOS</b>	<b>16</b>
<b>2.1</b>	<b>Variedades estatísticas exponencial e <math>q</math>-exponencial</b>	<b>17</b>
<b>2.2</b>	<b>Espaços de Musielak-Orlicz</b>	<b>21</b>
<b>2.3</b>	<b>Construção das <math>\varphi</math>-famílias de distribuições de probabilidade</b>	<b>24</b>
<b>2.4</b>	<b>Construção de uma variedade exponencial deformada</b>	<b>27</b>
<b>2.4.1</b>	<b>Exponencial deformada e uma parametrização para <math>\mathcal{P}_\mu</math></b>	<b>27</b>
<b>2.4.2</b>	<b>Fibrado tangente</b>	<b>43</b>
<b>3</b>	<b>TEORIA DA INFORMAÇÃO EM VARIEDADES ESTATÍSTICAS</b>	<b>46</b>
<b>3.1</b>	<b>A <math>\varphi</math>-divergência e <math>q</math>-divergência</b>	<b>46</b>
<b>3.2</b>	<b>Generalização da divergência de Rényi e <math>\exp_\kappa</math></b>	<b>50</b>
<b>3.3</b>	<b>Entropia Generalizada</b>	<b>53</b>
<b>3.3.1</b>	<b><math>\varphi</math>-entropia generalizada</b>	<b>53</b>
<b>3.3.2</b>	<b><math>\varphi</math>-entropia relativa generalizada</b>	<b>58</b>
<b>4</b>	<b>EXPONENCIAL DEFORMADA E O COMPORTAMENTO DA FUNÇÃO NORMALIZADORA</b>	<b>64</b>
<b>4.1</b>	<b>A <math>\Delta_2</math>-Condição</b>	<b>64</b>
<b>4.2</b>	<b>A condição (8) sobre a exponencial deformada e suas consequências</b>	<b>73</b>
<b>4.3</b>	<b>Caso puramente atômico</b>	<b>76</b>
<b>5</b>	<b>CONCLUSÃO E PERSPECTIVAS FUTURAS</b>	<b>80</b>
	<b>REFERÊNCIAS</b>	<b>81</b>

# 1 INTRODUÇÃO

## 1.1 Motivação

A Teoria da Informação estuda a quantificação da informação. Essa teoria teve como um dos precursores Claude Shannon (SHANNON, 1948). O conceito de informação engloba diferentes significados e por isso é tão difícil definir de forma concisa. Shannon conseguiu formalizar esse e outros conceitos com aplicações em teoria da informação. Suas obras tem grande impacto, por exemplo, em campos da física teórica, biologia, economia e sensoriamento remoto (NASCIMENTO and PRUDENTE, 2016; BENEDITO-SILVA and PIQUEIRA, 1998; DIONÍSIO, 2006; SILVA and MARQUES, 2014; NASCIMENTO, CINTRA, and FRERY, 2009). Com a necessidade de estudar modelos estatísticos usando geometria diferencial surgiu a Geometria da Informação. O matemático C. R. Rao, em (RAO, 1945; BURBEA and RAO, 1982; RAO, 1992), foi um dos primeiros a publicar resultados envolvendo geometria diferencial e famílias de distribuições de probabilidade. Mais precisamente, usando a função exponencial, Rao equipou essa família de distribuições com uma métrica Riemanniana usando a matriz de informação de Fisher (FISHER, 1922).

Os modelos estatísticos podem ser divididos em dois tipos: paramétricos e não paramétricos (ZHANG, 2004; GINÉ and NICKL, 2015). Para o caso paramétrico, o conjunto das funções de distribuições de probabilidade possui uma topologia euclidiana mais tratável, obtida a partir de parâmetros naturais finitos. Além disso, nesse caso foi induzida uma estrutura geométrica usando funções exponenciais deformadas (NAUDTS, 2004). No caso não paramétrico, Pistone e Sempi (PISTONE and SEMPI, 1995) foram um dos primeiros a elaborar uma rigorosa construção de uma topologia usando os espaços de Orlicz (GIBILISCO and PISTONE, 1998; PISTONE and ROGANTIN, 1999) e a função exponencial para munir a coleção das funções de distribuições de probabilidade estritamente positivas com uma estrutura suave. Esse espaço munido dessa nova estrutura é denominado de variedade estatística (LAURITZEN, 1984; GINÉ and NICKL, 2015).

O conceito de variedade surgiu da necessidade de unificar o estudo das superfícies diferenciáveis com as propriedades do cálculo diferencial de várias variáveis (LEE, 2001; SPIVAK, 1970). As variedades vem se tornando cada vez mais estudados em diversos campos, como estatística, genética, robótica, econometria, computação gráfica, imagem biomédica, física teórica, entre outras (BREGMAN, 1967a; LESSA, 2010; CASTRO, 2002; SAIMINEN, 2010; ULLAH, 1996; HUANG, 1997; CARRILLO, 2015). As variedades estudadas em geometria da informação são modeladas em espaços de dimensão infinita. Assim, sua topologia sofre mudanças radicais. Por exemplo, resultados envolvendo compacidade podem não acontecer (ALAOGLU, 1940). Para entender melhor certos fenômenos, faz-se uso rotineiramente de

conceitos da análise funcional (BREZIS, 2011).

Um tópico essencial em geometria da informação é o conceito de divergência. Dentre as mais conhecidas está a de Kullback-Leibler (KULLBACK and LEIBLER, 1951) e a de Bregman (BREGMAN, 1967b). Essas divergências caracterizam a entropia (“teoria da informação”) em sistemas de informação e ganho de informação ao comparar com modelos estatísticos de inferência. Além disso, foi usada para induzir a métrica de Fisher e as conexões exponenciais e mistura (AMARI, 1982; AMARI and NAGAOKA, 2000; COSTA, SANTOS, and STRAPASSON, 2015). Por esse motivo, essas divergências equipam as variedades com uma estrutura geométrica. Outra divergência muito usada é a de Rényi (VAN ERVEN and HAREMOS, 2014; DE SOUZA, VIGELIS, and CAVALCANTE, 2016). Ela depende de um índice  $\alpha \in (0, 1)$ , chamado de índice entrópico, que permite realizar uma ligação direta dessa divergência com a divergência de Shannon quando  $\alpha$  tende a 1.

Em (VIGELIS and CAVALCANTE, 2013a), os autores apresentaram uma estrutura suave para o espaço das funções densidade de probabilidade usando uma função exponencial deformada injetiva  $\varphi$ . A principal diferença nessa construção é o surgimento de novos abertos em  $\mathcal{P}_\mu$ , modelados por conjuntos contidos em espaços Musielak–Orlicz (MUSIELAK, 1983), os quais chamaremos de  $\varphi$ -famílias. Nessa construção, funções não injetivas, como por exemplo a  $q$ -exponencial, não servem para munir  $\mathcal{P}_\mu$  com uma estrutura de variedade estatística.

Para o caso em que a exponencial deformada é a função  $q$ -exponencial, Loiza e Quiceno (LOAIZA and QUICENO, 2013b) construíram um atlas para o conjunto das funções de densidade de probabilidade usando o espaço das funções essencialmente limitadas para obter cartas. A característica principal dessa família é que quando  $q$  tende a 1, os modelos exponenciais não paramétricos são obtidos, recuperando assim, a variedade construída por Pistone e Sempi (PISTONE and SEMPI, 1995). Ainda neste trabalho, foi construído o espaço tangente e, conseqüentemente, o fibrado tangente. Logo depois, em (LOAIZA and QUICENO, 2013a) foi definido métrica, conexão, torção e curvatura. Além disso, foi verificado que a variedade é flat (DO CARMO, 1988).

A entropia é um conceito usado em termodinâmica que surgiu da necessidade de medir o grau de organização de um meio. Diante de necessidade de outras áreas esse conceito pôde ser ampliado. Em teoria da informação foi Shannon que introduziu o conceito de entropia para quantificação da informação (SHANNON, 1948). A entropia de Shannon teve notável sucesso para uma variedade de sistemas, especialmente aqueles nos quais dominam interações de curto alcance espacial ou temporal. Entretanto, surgiram limitações para usar essa entropia para outros tipos de sistemas, por exemplo turbulências e interações de longo alcance, como as gravitacionais, e a partir de necessidades desse tipo, surgiram outras entropias. Por exemplo, as entropias de Tsallis (TSALLIS, 1988) (que teve destaque em sistemas com interações de longo alcance) e Renyi (RÉNYI, 1961) (é utilizada no campo de informação quântica, como

medida de entrelaçamento) têm como caso particular a entropia de Shannon e também são usadas como ferramentas para o reconhecimento de padrões de imagens. Em (ASSIRATI, 2014) é realizada uma análise comparativa entre as técnicas de entropias em imagens. Vários campos da ciência produzem dados na forma de imagens. Por exemplo, em imagens histológicas o reconhecimento de padrões são usados na identificação de células cancerígenas (LESSA, 2010), imagens tomográficas no diagnóstico de tumores ou até imagens de satélite no controle de atividades agrícolas, desmatamento, derramamento de petróleo no oceano e aplicações em automação industrial envolvendo o uso de sensores visuais em robôs (SIMON, 2019; CASTRO, 2002; LESSA, 2010). Imagens com pouco detalhes possuem baixa entropia enquanto imagens ricas em detalhes possuem altos valores de entropia. Dessa forma, tomar conhecimento de onde a função entropia atinge o máximo determinam detalhes importantes.

## 1.2 Objetivos e contribuições

O objetivo central dessa tese consiste em estudar localmente o conjunto de todas as medidas de probabilidade em  $T$  as quais são equivalentes com relação a uma medida  $\mu$ , usando algumas exponenciais deformadas não injetivas. Além disso, vamos obter propriedades geométricas e novas entropias.

Essas exponenciais deformadas assumem o seguinte comportamento: são zero até certo valor  $e$ , dali em diante, têm o comportamento semelhante ao da função exponencial clássica, a qual é estritamente crescente. Casos particulares e bastante conhecidos são: a exponencial  $q$ -deformada (LOAIZA and QUICENO, 2013a) e a  $\kappa$ -exponencial (KANIADAKIS, 2001, 2002). Usando uma função exponencial deformada, vamos dotar  $\mathcal{P}_\mu$  com uma estrutura de variedade de Banach de classe  $C^\infty$  modelada numa quantidade maior de abertos em  $\mathcal{P}_\mu$ . Com a possibilidade de considerar uma classe maior de exponenciais deformadas, conseguimos várias alterações na construção dos abertos da parametrização e conseqüentemente na  $\varphi$ -família. O domínio apresentado em (VIGELIS and CAVALCANTE, 2013a) é refinado para um conjunto ainda menor de tal forma que ainda continuamos cobrindo a variedade estatística  $\mathcal{P}_\mu$ .

Como os espaços Musielak–Orlicz possuem dimensão infinita, o plano tangente de  $\mathcal{P}_\mu$  é um espaço vetorial de dimensão infinita, o que impossibilita o uso de coordenadas cartesianas para fazer geometria no plano tangente. Por esse motivo, vamos adotar uma interpretação, como em (LOAIZA and QUICENO, 2013b), que considera o plano tangente uma classe de equivalência onde seus elementos estão relacionados via derivada das transições de coordenadas. Esse é primeiro passo para introduzir métricas e conexões.

Uma vez parametrizada  $\mathcal{P}_\mu$ , observamos que a expressão local depende de uma função convexa chamada de normalizadora. A divergência de Bregman associada a essa norma-

lizadora dá origem a uma nova divergência chamada de  $\varphi$ -divergência. Em particular, quando  $\varphi$  é a  $q$ -exponencial, denotada por  $\exp_q$ , temos que a  $\exp_q$ -divergência está relacionada com a  $q$ -divergência definida em (LOAIZA and QUICENO, 2013a). Consequentemente a métrica e conexões obtida por meio da  $\exp_q$ -divergência se relacionam com a métrica e conexões obtidas da  $q$ -divergência.

Uma outra contribuição desta tese consiste em definir novas entropias relativas e expressar relações com entropias já conhecidas na literatura. Além disso, definimos entropias relativas nas  $\varphi$ -famílias. Apresentamos também uma entropia denominada  $\varphi$ -entropia e vimos que a  $\exp$ -entropia é a entropia de Shannon e a  $\exp_q$ -exponencial é a entropia de Tsallis. Outra propriedade é que um múltiplo da  $\exp$ -entropia é usada em (CHEN, WANG, and KROVETZ, 2005) e é uma medida mais abrangente que a pureza em um cluster, pois considera a distribuição em classes semânticas.

A estrutura diferenciável de  $\mathcal{P}_\mu$  permite transferir para as  $\varphi$ -famílias da variedade propriedades locais dos espaços Musielak–Orlicz. Uma propriedade que podemos destacar é a  $\Delta_2$ -condição. Essa condição é uma expressão que resgata um sentido geométrico para o fenômeno de funções Musielak–Orlicz não satisfazerem essa condição  $\Delta_2$ : o bordo do domínio da parametrização ser não-vazio. Isso permite estudar o comportamento da função normalizadora próximo ao bordo do domínio da parametrização (VIGELIS and CAVALCANTE, 2013b). Como parte da contribuição deste tese, mostramos que dada uma função na classe Musielak–Orlicz, a normalizadora converge próximo ao bordo do seu domínio. Além disso, no caso puramente atômico (em que consideremos uma medida de contagem sobre o conjunto  $\mathbb{N}$ ) encontramos uma consequência do fato da função Musielak–Orlicz não satisfazer a  $\delta_2$ -condição (o equivalente da condição  $\Delta_2$  do caso puramente atômico).

### 1.3 Organização do documento

Para uma melhor exposição do estudo e visando uma melhor compreensão dos conceitos e técnicas utilizadas, esta tese contém 5 capítulos.

- Capítulo 2 : Dividimos este capítulo em duas partes. Na tentativa de deixar a tese auto contida, essa primeira parte (que vai da Seção 2.1 até a Seção 2.3) é reservada a fazer uma breve introdução dos conceitos necessários para a compreensão dos resultados posteriores. Vamos apresentar, não nessa ordem, espaços e funções Musielak–Orlicz, definição de variedades estatísticas e construções (já feitas e que são casos particulares do que trataremos aqui) usando exponenciais deformadas injetivas. Na segunda parte, vamos expor alguns dos nossos resultados: construção de uma estrutura suave para  $\mathcal{P}_\mu$ , usando exponencias deformadas não injetivas, espaço tangente e fibrado tangente.

- Capítulo 3 : Nas Seções 3.1 e 3.2, definimos a  $\varphi$ -divergência, provamos que quando  $\varphi$  for a  $\exp_q$ , a  $\exp_q$ -divergência está relacionada com a  $q$ -divergência e que a  $\kappa$ -exponencial pode ser usada na generalização da divergência de Rényi. De modo análogo, as respectivas métricas e conexões estão relacionadas. Na Seção 3.3, definimos  $\varphi$ -entropia e vimos que quando  $\varphi$  é a exponencial e a  $q$ -exponencial, a  $\exp$ -entropia e a  $\exp_q$ -entropia coincidem com as entropias de Shannon e de Tsallis, respectivamente. Além disso, definimos uma  $\varphi$ -entropia de Rényi e provamos que está relacionada a  $\varphi$ -entropia por meio de um limite. Analogamente, estabelecemos as  $\varphi$ -entropia relativa e  $\varphi$ -entropia relativa de Rényi e provamos que estão relacionadas entre si por um limite e quando  $\varphi$  for a  $\exp$  e a  $\exp_q$ , segue que coincidem com as entropias relativas de Kullback-Leibler e de Tsallis, respectivamente.
- Capítulo 4 : Analisamos o comportamento da função normalizadora próximo ao bordo do seu domínio em pontos que estão na classe Musielak-Orlicz. Consideraremos também o caso puramente atômico, isto é,  $\mu$  é uma medida de contagem, e provaremos adaptações de resultados obtidos para o caso não-atômico.
- Capítulo 5 : Além de tecer algumas considerações finais, fizemos uma síntese das principais contribuições deste trabalho e propostas para pesquisas futuras.



## 2 VARIEDADES ESTATÍSTICAS: FUNDAMENTOS

As variedades surgiram como generalizações de curvas e superfícies para dimensões maiores que três. Variedades euclidianas são conjuntos com uma estrutura suave de classe  $C^\infty$  localmente identificada com um espaço euclidiano de dimensão finita. Infelizmente nem todas as variedades podem ser modeladas localmente em espaços euclidianos. Por exemplo, os espaços de funções não são variedades euclidianas. Nesse caso, podemos modelar localmente em espaços de Banach (que são generalizações de espaços euclidianos). A desvantagem dessas variedades é a falta de intuição geométrica, porém ao longo do tempo, foi desenvolvido um vasto aparato técnico e conceitual que nos permitiram observar com mais detalhes esse ambiente invisível da geometria em dimensões infinitas.

Nessa Tese vamos considerar o espaço das funções de distribuições de probabilidade com uma estrutura de variedade de Banach de classe  $C^\infty$ . Essa variedade é chamada de variedade estatística. Para sua construção é necessário considerar certos conjuntos com propriedades específicas no intuito de definir parametrizações. O próximo passo é transferir essas propriedades para os conjuntos das funções densidade de probabilidade estritamente positivas. Por fim, obtemos uma estrutura de variedade e localmente conseguimos introduzir a noção de divergência, entropia, métrica, conexão entre outras. Pelo que foi mencionado, é natural se esperar que todas essas propriedades sejam derivadas da technicalidade presente na construção dessa variedade de Banach.

Para a construção é necessário usar exponenciais deformadas com algumas condições. Por exemplo em (PISTONE and SEMPI, 1995) foi considerada a exponencial clássica, em (LOAIZA and QUICENO, 2013a) foi considerado a  $q$ -exponencial e em (VIGELIS and CAVALCANTE, 2013a) foi considerada uma exponencial deformada mais geral que tem como caso particular a construção do Pistone. Vale ressaltar que na expressão da métrica e conexões aparecem essas exponenciais deformadas.

Para tornar este capítulo auto-contida dividiremos em duas partes: na primeira, apresentaremos o conceito de variedade estatística, construção, planos tangentes e exemplos. Na segunda parte, será apresentada parte das contribuições dessa tese.

Na próxima seção, vamos fazer uma breve revisão de resultados desenvolvidos em (PISTONE and SEMPI, 1995; GIBILISCO and PISTONE, 1998; PISTONE and ROGANTIN, 1999) e (LOAIZA and QUICENO, 2013a). De forma resumida, vamos munir o conjunto das funções de distribuições de probabilidade  $\mathcal{P}_\mu$  com uma estrutura de variedade de Banach de classe  $C^\infty$ . Ainda nesse capítulo, faremos também, de forma curta, uma construção mais geral feita por (VIGELIS and CAVALCANTE, 2013a). Além disso, revisaremos alguns conceitos que serão utilizados ao longo da tese como, por exemplo, espaços Musielak–Orlicz e funções de Musielak–Orlicz. Em seguida, na última seção, vamos apresentar uma construção que se

baseia numa nova classe de funções exponenciais deformadas.

## 2.1 Variedades estatísticas exponencial e $q$ -exponencial

Uma variedade de Banach  $M$  é um espaço topológico que possui a propriedade de que dado um ponto  $p$ , existe uma vizinhança desse ponto homeomorfa a um conjunto aberto de um espaço de Banach. Abaixo definiremos formalmente as condições para um conjunto ser um espaço de Banach.

**Definição 1** *Sejam  $M$  um espaço topológico,  $X$  um espaço de Banach e aplicações  $x_\alpha : U_\alpha \subset M \rightarrow x_\alpha(U_\alpha) \subset X$ . Um atlas de classe  $C^k$ ,  $k > 0$ , para  $M$  modelado em  $X$  é uma família de pares  $\{(U_\alpha, x_\alpha)\}_\alpha$  satisfazendo*

- $\bigcup_\alpha x_\alpha(U_\alpha) = M$ ;
- dado qualquer par de índices  $\alpha, \beta$  tais que  $x_\alpha(U_\alpha) \cap x_\beta(U_\beta) = W \neq \emptyset$ , os conjuntos  $x_\alpha^{-1}(W)$  e  $x_\beta^{-1}(W)$  são abertos em  $U_\alpha$  e  $U_\beta$ , respectivamente;
- a aplicação de transição  $x_\beta^{-1} \circ x_\alpha : x_\alpha^{-1}(W) \rightarrow x_\beta^{-1}(W)$  é um  $C^k$ -isomorfismo.

Cada aplicação  $x_\alpha$  é chamada uma carta para uma vizinhança de  $M$ .

Um espaço topológico  $M$  com um atlas de classe  $C^k$  modelado sobre um espaço de Banach  $X$  é chamado de variedade de Banach de classe  $C^k$  modelada sobre  $X$ .

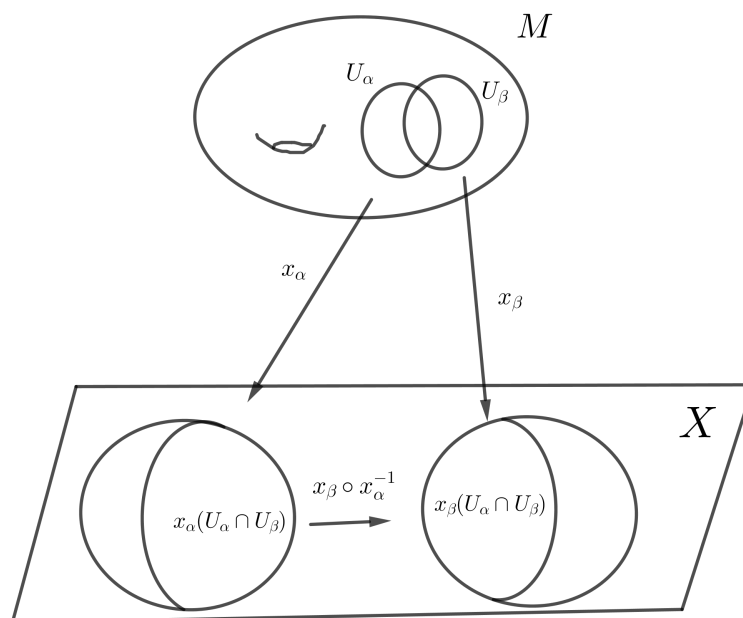


Figura 1: Variedade de Banach

O par  $\{(U_\alpha, x_\alpha)\}_\alpha$  com  $p \in x_\alpha(U_\alpha)$  é chamado de parametrização de  $M$  em  $p$  e  $x_\alpha(U_\alpha)$  é uma vizinhança de coordenada em  $p$ . Uma coleção de sistemas de coordenadas que satisfaz as propriedades acima é chamado de atlas.

O conjunto  $M$  pode ser induzido com uma topologia que torne  $x_\alpha(U_\alpha)$  aberto e as aplicações  $x_\alpha$  isomorfismos. Duas coleções  $\{(U_\alpha, x_\alpha)\}$  e  $\{(U_\beta, x_\beta)\}$  são compatíveis se  $\{(U_\alpha, x_\alpha)\} \cup \{(U_\beta, x_\beta)\}$  também satisfazem as propriedades  $a)$ ,  $b)$  e  $c)$ . Essa operação é uma relação de equivalência chamada de  $C^k$ -compatibilidade e uma classe de equivalência é dita ser uma estrutura  $C^k$ -diferenciável.

Seja  $L^0$  o espaço linear de todas as funções mensuráveis a valores-reais sobre  $T \subset \mathbb{R}$ , com igualdade  $\mu$ -q.t.p.. A família de todas as medidas de probabilidade sobre  $T$  que são equivalentes com relação a medida  $\mu$  é denotado por

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

**Exemplo 1** *Sejam  $\Sigma$  um subconjunto do subconjunto real  $T$  e  $(T, \Sigma, \mu)$  um espaço de medida não-atômico  $\sigma$ -finito. Em (GIBILISCO and PISTONE, 1998) a variedade  $\mathcal{P}_\mu = \mathcal{P}(T, \Sigma, \mu)$  a família de todas as medida de probabilidade em  $T$  que são equivalentes para a medida  $\mu$  foi modelada pelo espaço de Orlicz  $L^{\Phi_1}(p) = L^{\Phi_1}(T, \Sigma, p \cdot \mu)$  em que  $p \in \mathcal{P}_\mu$  e*

$$\begin{aligned} \Phi_1 : [0, \infty) &\longrightarrow [0, \infty) \\ u &\longmapsto e^u - 1 \end{aligned}$$

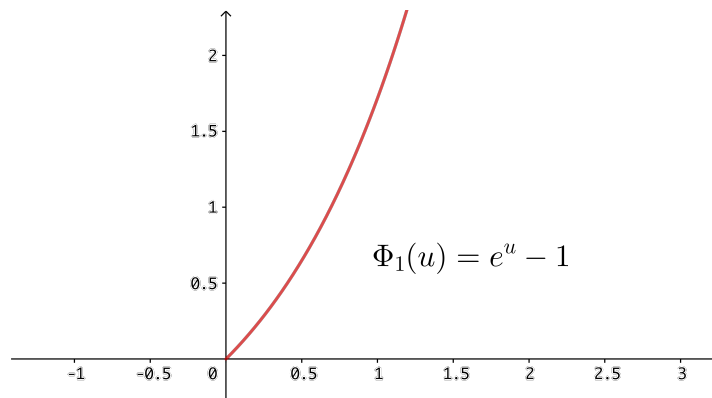


Figura 2: Função de Young

*Faremos uma breve apresentação sobre a expressão das suas cartas e a mudança de coordenadas.*

*Para cada função  $u \in L^0$ , definimos os funcionais*

$$\hat{u}_p(\lambda) = \int_T e^{\lambda u} d\mu, \quad K_p(u) = \log \hat{u}_p(1),$$

e o conjunto  $\mathcal{K}_p$  como sendo o interior de  $\{u \in L^0; \widehat{u}_p(1) < \infty\}$ .

A parametrização desse espaço é a aplicação  $e_p : \mathcal{B}_p \rightarrow \mathcal{E}_p$  dada por

$$e_p(u) = e^{u - K_p(u)} p, \quad (1)$$

em que  $\mathcal{B}_p = B_p \cap \mathcal{K}_p$ , com

$$B_p = \left\{ u \in L^{\Phi_1}(p); \int_T u p d\mu = 0 \right\}, \quad (2)$$

e  $\mathcal{E}_p = e_p(\mathcal{B}_p) \subseteq \mathcal{P}_\mu$ . A inversa  $e_p^{-1} : \mathcal{E}_p \rightarrow \mathcal{B}_p$  é dada por

$$e_p^{-1}(q) = \log\left(\frac{q}{p}\right) - \int_T \left[ \log\left(\frac{q}{p}\right) \right] d\mu, \quad (3)$$

para  $q \in \mathcal{E}_p$ . Por fim, a transição  $e_q^{-1} \circ e_p : \mathcal{B}_p \rightarrow \mathcal{B}_q$ , é dada por

$$e_q^{-1} \circ e_p(u) = u + \log\left(\frac{p}{q}\right) - \int_T \left[ u + \log\left(\frac{p}{q}\right) \right] q d\mu, \text{ para todo } u \in \mathcal{B}_p. \quad (4)$$

Observe que  $e_q^{-1} \circ e_p$  é de classe  $C^\infty$  e que  $\bigcup_{p \in \mathcal{P}_\mu} e_p(\mathcal{B}_p) = \mathcal{P}_\mu$ . Portanto,  $\mathcal{P}_\mu$  é uma variedade Banach de classe  $C^\infty$ .

**Exemplo 2** (LOAIZA and QUICENO, 2013b) Outra maneira de dotar  $\mathcal{P}_\mu$  com uma estrutura de Banach de classe  $C^\infty$ , é usando a função  $q$ -exponencial a qual é dada por

$$e_q^x = (1 + (1 - q)x)^{1/(1-q)}, \text{ em que } \frac{-1}{1-q} \leq x. \quad (5)$$

Nesse caso,  $\mathcal{P}_\mu$  é modelada pelo espaço  $L^\infty(p\mu)$ , em que  $p \in \mathcal{P}_\mu$ .

Da mesma forma do exemplo anterior, exibiremos as cartas e a transição entre elas.

Denotemos  $\mathcal{V}_p$  como sendo os elementos  $u \in L^\infty(p \cdot \mu)$  tais que  $\|u\|_\infty < 1$  e

$$\int_T u p d\mu = 0.$$

O modelo  $q$ -exponencial é a aplicação

$$e_{q,p} : \mathcal{V}_p \rightarrow \mathcal{P}_\mu,$$

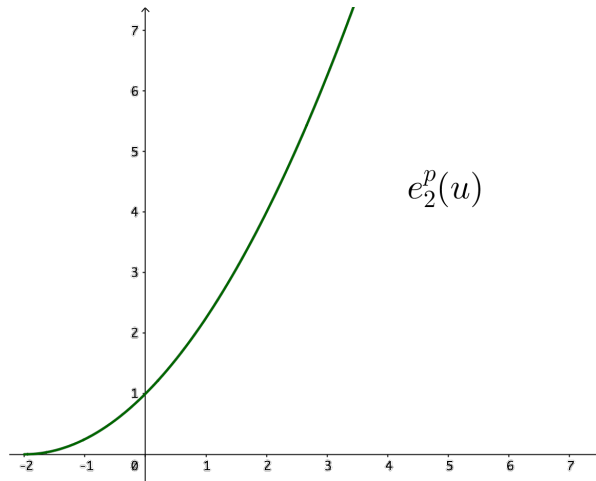


Figura 3: Função 2-exponencial

definida por

$$e_{q,p}(u) = e_q^{(u \ominus_q K_p^q(u))} p \quad (6)$$

em que

$$u \ominus_q K_p^q(u) = \frac{u - K_p^q(u)}{1 + (1 - q)K_p^q(u)}$$

e  $K_p^q$  é o funcional definido por

$$K_p^q(u) = \ln_q \left[ \int_T e_q^{(u)} d\mu \right]$$

para todo elemento unitário  $u \in L^\infty(p \cdot \mu)$  tal que

$$\int_T e_q^{(u)} d\mu < \infty.$$

A aplicação de transição  $e_{q,p_2}^{-1} \circ e_{q,p_1} : e_{q,p_1}^{-1}(\mathcal{U}_{p_1} \cap \mathcal{U}_{p_2}) \rightarrow e_{q,p_2}^{-1}(\mathcal{U}_{p_1} \cap \mathcal{U}_{p_2})$ , em que  $\mathcal{U}_p$  é a imagem de  $e_{q,p}$ , é expressa como

$$e_{q,p_2}^{-1}(e_{q,p_1}(u)) = \frac{u + [1 + (1 - q)u] \ln_q \left( \frac{p_1}{p_2} \right) - \int_T u + [1 + (1 - q)u] \ln_q \left( \frac{p_1}{p_2} \right) p_2 d\mu}{1 + (1 - q) \int_T u + [1 + (1 - q)u] \ln_q \left( \frac{p_1}{p_2} \right) p_2 d\mu} \quad (7)$$

em que  $p_1, p_2 \in \mathcal{P}_\mu$  com  $\mathcal{U}_{p_1} \cap \mathcal{U}_{p_2} \neq \emptyset$  e  $u \in e_{q,p_1}^{-1}(\mathcal{U}_{p_1} \cap \mathcal{U}_{p_2})$ .

Seja  $M$  uma variedade Banach de classe  $C^k$  e  $p \in M$ . Considere o trio  $(U, \varphi, v)$  em que  $(U, \varphi)$  é uma carta em  $p$  e  $v$  é um elemento de  $\varphi(U)$ . Dizemos que duas tripla  $(U, \varphi, v)$  e  $(V, \psi, w)$  são equivalentes (e representamos por  $(U, \varphi, v) \simeq (V, \psi, w)$ ) se a derivada da aplicação

$d(\psi \circ \varphi^{-1})_{\varphi_p}(v) = w$ . Esse terno é chamado de vetor tangente de  $M$  em  $p$  e é denotado por  $T_pM$ .

**Observação 1** Poderíamos expressar o plano tangente em função da parametrização, isto é, se a carta  $(U, \varphi)$  pode ser substituída pela parametrização  $(\varphi(U), \varphi^{-1})$ . A mudança que em essência aconteceria aparece na composição  $\psi^{-1} \circ \varphi$ . O restante permanece da mesma forma.

**Observação 2** Existem outras maneiras (clássicas) de definir o plano tangente. Por exemplo, podemos considerá-lo como o espaço dos vetores velocidade de uma curva em uma vizinhança de  $p$  ou como germe de funções. Temos um isomorfismo entre essas maneiras e considerar o plano tangente uma classe de equivalência. Mais informações veja (LEE, 2001) ou (LANG, 2002).

O fibrado tangente, definido por  $TM$ , é a união disjunta de todos os planos tangentes da variedade  $M$ , isto é,

$$TM = \bigsqcup_{p \in M} T_pM.$$

O fibrado tangente é uma variedade Banach de classe  $C^k$  e suas cartas são definidas a partir das cartas de  $M$ .

**Exemplo 3** Em (LOAIZA and QUICENO, 2013b), foi construído o plano tangente e, conseqüentemente, seu fibrado a partir de uma identificação natural com modelos  $q$ -exponenciais paramétricos de dimensão 1.

## 2.2 Espaços de Musielak-Orlicz

Os espaços de Musielak–Orlicz são generalizações dos espaços de Orlicz. Nesta tese os espaços Musielak–Orlicz desempenham um papel relevante na construção das variedades estatísticas. Esses espaços são normados (com a norma de Luxemburgo ou a de Orlicz), convexo e sua dimensão é infinita. Essas propriedades são ideias para servir de modelo local para o domínio das cartas da variedade estatística.

A aplicação  $\Phi : T \times [0, \infty) \rightarrow [0, \infty]$  é uma função Musielak–Orlicz para  $\mu$ -q.t.p. (quase sempre)  $t \in T$  se as seguintes condições ocorrem:

- (i)  $\Phi(t, \cdot)$  é convexa e semicontínua inferiormente para  $\mu$ -q.t.p.  $t \in T$ ;
- (ii)  $\Phi(t, 0) = \lim_{u \downarrow 0} \Phi(t, u) = 0$  e  $\lim_{u \uparrow \infty} \Phi(t, u) = \infty$  para  $\mu$ -q.t.p.  $t \in T$ ;
- (iii)  $\Phi(\cdot, u)$  é mensurável para cada  $u$  não negativo.

Uma consequência da definição é que segue dos itens (i) e (ii) que  $\Phi(t, \cdot)$  não é igual a 0 ou  $\infty$  no intervalo  $(0, \infty)$ .

Os conjuntos

$$L^\Phi = \left\{ u \in L^0; \int_T \Phi(t, |\lambda u(t)|) d\mu < \infty \text{ para algum } \lambda > 0 \right\},$$

$$\tilde{L}^\Phi = \left\{ u \in L^0; \int_T \Phi(t, |u(t)|) d\mu < \infty \right\}$$

e

$$E^\Phi = \left\{ u \in L^0; \int_T \Phi(t, |\lambda u(t)|) d\mu < \infty \text{ para todo } \lambda > 0 \right\}.$$

são, respectivamente, chamados de espaço Musielak–Orlicz, classe Musielak–Orlicz e espaço Morse–Transue associados a uma função Musielak–Orlicz  $\Phi$  (MUSIELAK, 1983; CHEN, 1996).

O espaço  $L^\Phi$  é Banach com a norma Luxemburgo

$$\|u\|_\Phi = \inf \left\{ \lambda > 0; \int_T \Phi \left( t, \left| \frac{u(t)}{\lambda} \right| \right) d\mu \leq 1 \right\},$$

ou com a norma Orlicz

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

em que  $\Phi^*(t, v) = \sup_{u \geq 0} (uv - \Phi(t, u))$  é o conjugado Fenchel de  $\Phi(t, \cdot)$ . Essas normas acima são equivalentes, pois

$$\|u\|_\Phi \leq \|u\|_{\Phi,0} \leq 2\|u\|_\Phi$$

ocorre para todo  $u \in L^\Phi$ . Para mais detalhes, ver (MUSIELAK, 1983).

Para construir a variedade exponencial deformada precisamos da relação estabelecida no próximo Lema, envolvendo convergência em norma e convergência em medida. Para tal, precisamos do seguinte:

Seja  $\{u_n\}_{n \in \mathbb{N}}$  uma sequência de funções mensuráveis. Dizemos que

a)  $u_n$  converge em norma (ou fortemente) para alguma função mensurável  $u$  se

$$\lim_{n \rightarrow \infty} \|u_n - u\| = 0;$$

b)  $u_n$  converge em medida para alguma função mensurável  $u$  se

$$\mu(\{t \in E; |u_n(t) - u(t)| > \varepsilon\}) \rightarrow 0, \text{ quando } n \rightarrow \infty,$$

é satisfeita para cada  $\varepsilon > 0$ , e para todo conjunto mensurável  $E$  com medida finita  $\mu(E) < \infty$ .

Os próximos dois resultados mostram como convergência em norma em  $L^\Phi$  e convergência em medida estão relacionadas.

**Lema 2.1** *Em  $L^\Phi$ , convergência em norma é mais forte que convergência em medida, ou seja, se  $u_n$  e  $u$  são sequências de funções mensuráveis, então  $u_n$  convergir em norma para  $u$  implica em  $u_n$  convergir em medida para  $u$ .*

**Demonstração:** Seja  $\{u_n\}$  uma sequência em  $L^\Phi$  convergindo em norma para  $u \in L^\Phi$ . Considere um conjunto mensurável  $A$  com medida finita. Dado  $\varepsilon > 0$ , denote para cada  $n \in \mathbb{N}$

$$A_n = \{t \in A; |u_n(t) - u(t)| > \varepsilon\}.$$

Desde que  $\lim_{n \rightarrow \infty} \Phi(t, u) = \infty$ , podemos encontrar  $\lambda > 0$  para o qual o conjunto

$$B_\lambda = \{t \in A; \Phi(t, \lambda \varepsilon) \geq 1\}$$

satisfaz  $\mu(A \setminus B_\lambda) \leq \varepsilon/2$ . A convergência  $\|u_n - u\|_\Phi \rightarrow 0$  garante que

$$I_\Phi(\lambda(u_n - u)) \rightarrow 0.$$

Então, existe  $n_0 \geq 1$  tal que para todo  $n \geq n_0$ ,

$$\begin{aligned} \frac{\varepsilon}{2} &\geq I_\Phi(\lambda(u_n - u)\chi_A) \geq I_\Phi(\lambda \varepsilon \chi_{A_n}) \\ &\geq I_\Phi(\lambda \varepsilon \chi_{A_n \cap B_\lambda}) \geq \mu(A_n \cap B_\lambda). \end{aligned}$$

Para qualquer  $n \geq n_0$ , segue que

$$\begin{aligned} \mu(A_n) &\leq \mu(A_n \cap B_\lambda) + \mu(A \setminus B_\lambda) \\ &\leq \frac{\varepsilon}{2} + \frac{\varepsilon}{2} \\ &= \varepsilon. \end{aligned}$$

Então,  $\{u_n\}$  converge em medida para  $u$ .  $\square$



## 2.3 Construção das $\varphi$ -famílias de distribuições de probabilidade

De agora em diante, consideraremos  $(T, \Sigma, \mu)$  um espaço de medida  $\sigma$ -finito não-atômico.

Uma exponencial deformada é uma aplicação mensurável

$$\varphi : T \times \mathbb{R} \rightarrow [0, \infty)$$

que satisfaz as seguintes condições:

1.  $\varphi$  é convexa;
2.  $\lim_{u \rightarrow -\infty} \varphi(t, u) = 0$ ;
3.  $\lim_{u \rightarrow \infty} \varphi(t, u) = \infty$ .

**Observação 3** Denote por  $\varphi$  o operador atuação sobre o conjunto das funções valores-reais  $u : T \rightarrow \mathbb{R}$  dados por  $\varphi(\mathbf{u})(\mathbf{t}) := \varphi(\mathbf{t}, \mathbf{u}(\mathbf{t}))$ .

**Exemplo 4** É claro que a exponencial  $\exp(u)$  é uma exponencial deformada. Outro exemplo clássico é a Kaniadakis  $\kappa$ -exponencial  $\exp_\kappa : \mathbb{R} \rightarrow (0, \infty)$  para  $\kappa \in [-1, 1]$ . Sua expressão é dada por

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

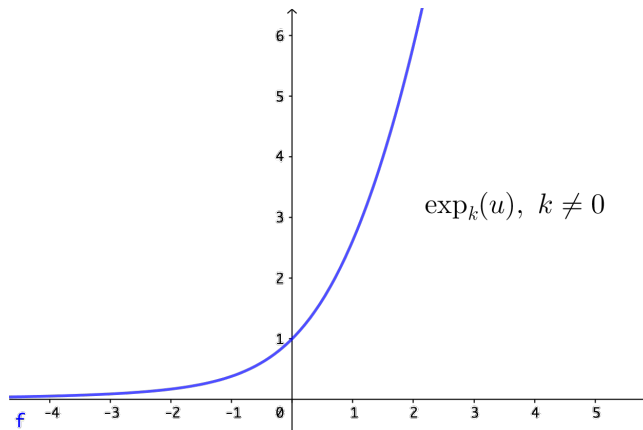


Figura 4: Função  $\kappa$ -exponencial

Em (VIGELIS and CAVALCANTE, 2013a), foi generalizada a construção da estrutura que torna  $\mathcal{P}_\mu$  uma variedade Banach de classe  $C^k$  discutida no Exemplo 1. Para isso, foi utilizada a função exponencial deformada com uma hipótese adicional:

**Condição 2.1**  $\varphi$  é injetiva e podemos encontrar uma função mensurável, positiva  $u_0 : T \rightarrow (0, \infty)$  tal que

$$\int_T \varphi(c + \lambda u_0) d\mu < \infty, \quad \forall \lambda > 0, \quad (8)$$

para cada função mensurável  $c : T \rightarrow \mathbb{R}$  tal que  $\varphi(c) \in \mathcal{P}_\mu$ .

Esboçaremos o roteiro detalhado dessa construção. Os resultados que exibiremos nesta seção serão de certa forma adaptados para o caso em que  $\varphi$  não é injetiva. Isso por si só justifica a importância desta seção.

Seja  $c : T \rightarrow \mathbb{R}$  uma função mensurável tal que  $\varphi(c) = p \in \mathcal{P}_\mu$ . Assim, para a função Musielak–Orlicz

$$\Phi_c(t, u) = \varphi(t, c(t) + u) - \varphi(t, c(t)), \quad (9)$$

escrevamos  $L_c^\varphi$ ,  $\tilde{L}_c^\varphi$  e  $E_c^\varphi$ , no lugar de  $L^{\Phi_c}$ ,  $\tilde{L}^{\Phi_c}$  e  $E^{\Phi_c}$ , respectivamente. Dessa maneira, o espaço de Musielak–Orlicz pode ser escrito

$$L_c^\varphi = \left\{ u \in L^0; \int_T \varphi(t, c(t) + \lambda u(t)) d\mu < \infty \forall \lambda \in (-\varepsilon, \varepsilon), \exists \varepsilon > 0 \right\}.$$

O conjunto aberto

$$\mathcal{K}_c^\varphi = \left\{ u \in L_c^\varphi; \int_T \varphi(t, c(t) + \lambda u(t)) d\mu < \infty \forall \lambda \in (-\varepsilon, 1 + \varepsilon), \exists \varepsilon > 0 \right\}$$

é usado para definir a função normalizadora. Claramente, se  $u \in \mathcal{K}_c^\varphi$ , a função  $\varphi(c + u)$  não está necessariamente em  $\mathcal{P}_\mu$ . A função normalizadora é uma aplicação  $\psi : \mathcal{K}_c^\varphi \rightarrow \mathbb{R}$  que normaliza  $\varphi(c + u)$ , isto é

$$\varphi(c + u - \psi(u)u_0) \in \mathcal{P}_\mu$$

para qualquer  $u \in \mathcal{K}_c^\varphi$ .

**Observação 4** 1. Se  $u_0$  satisfaz a condição adicional 2.1 e  $\varphi(c + u - \alpha u_0)$  é integrável para algum  $\alpha \in \mathbb{R}$ , então  $\varphi(c + \lambda u)$  é  $\mu$ -integrável para qualquer  $\lambda \in (0, 1)$ . Desse modo, o domínio aberto, maximal de  $\psi$  está contido em  $\mathcal{K}_c^\varphi$ .

2. A aplicação  $\psi$  está bem definida, ou seja, dado  $u \in \mathcal{K}_c^\varphi$ , existe um único  $\psi(u) \in \mathbb{R}$  tal que  $\varphi(c + u - \psi(u)u_0) \in \mathcal{P}_\mu$ ;

Definindo o conjunto fechado

$$B_c^\varphi = \left\{ u \in L_c^\varphi; \int_T u \varphi'_+(c) d\mu = 0 \right\}, \quad (10)$$

associamos a parametrização

$$\varphi_c : \mathcal{B}_c^\varphi \rightarrow \mathcal{F}_c^\varphi, \quad u \mapsto \varphi(c + u - \psi(u)u_0)$$

em que  $\mathcal{B}_c^\varphi := \mathcal{B}_c^\varphi \cap \mathcal{K}_c^\varphi$  e  $\mathcal{F}_c^\varphi = \varphi_c(\mathcal{B}_c^\varphi)$ .

Note que

$$\bigcup_{\varphi(c) \in \mathcal{P}_\mu} \mathcal{F}_c^\varphi = \mathcal{P}_\mu$$

e para quaisquer  $c_1, c_2 : T \rightarrow \mathbb{R}$  tais que  $\varphi(c_1), \varphi(c_2) \in \mathcal{P}_\mu$ , os conjuntos

$$\varphi_{c_1}^{-1}(\mathcal{F}_{c_1}^\varphi \cap \mathcal{F}_{c_2}^\varphi) \quad \varphi_{c_2}^{-1}(\mathcal{F}_{c_1}^\varphi \cap \mathcal{F}_{c_2}^\varphi)$$

são abertos.

Além disso, a aplicação de transição  $\varphi_{c_2}^{-1} \circ \varphi_{c_1} : \varphi_{c_1}^{-1}(\mathcal{F}_{c_1}^\varphi \cap \mathcal{F}_{c_2}^\varphi) \rightarrow \varphi_{c_2}^{-1}(\mathcal{F}_{c_1}^\varphi \cap \mathcal{F}_{c_2}^\varphi)$  pode ser expressa como

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

Essa expressão é de classe  $C^\infty$ , pois  $w, c_1 - c_2 \in L_{c_2}^\varphi$ , e os espaços  $L_{c_1}^\varphi$  e  $L_{c_2}^\varphi$ , possuem normas equivalentes.

Por fim, note que a coleção  $\{(\mathcal{B}_c^\varphi, \varphi_c)\}_{\varphi(c) \in \mathcal{P}_\mu}$  satisfaz as propriedades da definição 1 o que torna  $\mathcal{P}_\mu$  com uma estrutura de variedade de Banach de classe  $C^\infty$ .

## 2.4 Construção de uma variedade exponencial deformada

Na construção feita em (VIGELIS and CAVALCANTE, 2013a), foram atribuídas duas condições essenciais para obter as  $\varphi$ -famílias: a injetividade da exponencial deformada e a condição (8). Para exponenciais deformadas não injetivas, a função normalizadora não fica bem definida. Com uma pequena modificação na definição dessa função é possível refinar os abertos dessa nova variedade.

### 2.4.1 Exponencial deformada e uma parametrização para $\mathcal{P}_\mu$

**Condição 2.2** *A aplicação  $\varphi : T \times \mathbb{R} \rightarrow [0, \infty)$  continua sendo uma exponencial deformada. Em (VIGELIS and CAVALCANTE, 2013a) foi adicionada sobre  $\varphi$  a hipótese da injetividade. Ela será removida e acrescentada uma outra:*

$$\varphi(t, x) = 0, \forall x < a_\varphi := \inf \{x \in \mathbb{R}; \varphi(x) > 0\}. \quad (11)$$

É claro que  $\varphi$  não é injetiva para valores  $x < a_\varphi$  e injetiva para o restante.

**Exemplo 5** *Um exemplo de uma função exponencial deformada que satisfaz a condição (2.2) é a função  $q$ -exponencial  $\exp_q : T \times \mathbb{R} \rightarrow [0, \infty)$  definida por*

$$\begin{aligned} \exp_q(u) &= [1 + (1 - q)u]_+^{\frac{1}{1-q}} \\ &= \max \left\{ 0, (1 + (1 - q)u)^{\frac{1}{1-q}} \right\} \end{aligned}$$

em que  $q : T \rightarrow (0, 1)$  é uma função mensurável. Não é difícil ver que a função  $q$ -exponencial satisfaz a condição (2.2) com  $a_\varphi = \frac{-1}{1-q}$ .

**Observação 5** *Observe que a aplicação trabalhada em (VIGELIS and CAVALCANTE, 2013a) satisfaz a hipótese (11). De fato, a injetividade implica que  $a_\varphi = -\infty$ .*

Diferentemente do que foi feito em (VIGELIS and CAVALCANTE, 2013a), precisamos fazer uma restrição no domínio. Desde que  $\varphi$  satisfaz (11), podemos ter  $\varphi(c + \lambda u) = 0$ , e dessa forma não temos  $\varphi(c + \lambda u) \in \mathcal{P}_\mu$ . Para solucionar esse problema precisaremos definir um outro conjunto para intersectar com o espaço Musielak-Orlicz  $L_c^\varphi$ .

**Lema 1** *O conjunto*

$$\begin{aligned} N_c^\varphi &= \{u \in L^0; \exists \lambda > 0, \text{ em que } |\lambda u| \leq c - a_\varphi\} \\ &= \{u \in L^0; \exists \varepsilon > 0, \text{ em que } \lambda u \leq c - a_\varphi, \text{ para cada } \lambda \in (-\varepsilon, \varepsilon)\}. \end{aligned}$$

é um espaço de Banach com a norma

$$\|u\|_{N_c^\varphi} = \inf \left\{ \lambda > 0; \left| \frac{1}{\lambda} u \right| \leq c - a_\varphi \right\}.$$

**Demonstração:** Primeiramente mostraremos que  $N_c^\varphi$  é um espaço vetorial.

- i) Claramente  $0 \in N_c^\varphi$ , pois  $c - a_\varphi > 0$ .
- ii) Sejam  $u, v \in N_c^\varphi$  e  $\alpha \in \mathbb{R}$ . Por definição existem  $\lambda_1, \lambda_2 > 0$ , tais que

$$|u| \leq \frac{1}{\lambda_1}(c - a_\varphi) \quad |v| \leq \frac{1}{\lambda_2}(c - a_\varphi).$$

Tomando  $\lambda = \min \left\{ \frac{\lambda_1}{2}, \frac{\lambda_2}{2} \right\}$ , temos que

$$\begin{aligned} |u + v| &\leq |u| + |v| \\ &\leq \frac{1}{\lambda_1}(c - a_\varphi) + \frac{1}{\lambda_2}(c - a_\varphi) \\ &\leq \frac{1}{2\lambda}(c - a_\varphi) + \frac{1}{2\lambda}(c - a_\varphi) \\ &= \frac{1}{\lambda}(c - a_\varphi). \end{aligned}$$

Assim,

$$u + v \in N_c^\varphi$$

- iii) Dado  $\alpha \in \mathbb{R}$ , obtemos

$$\begin{aligned} |\alpha u| &\leq |\alpha| |u| \\ &\leq \frac{|\alpha|}{\lambda_1}(c - a_\varphi), \end{aligned}$$

o que implica em

$$\alpha u \in N_c^\varphi$$

Dessa forma,  $N_c^\varphi$  é um subespaço.

Para o que resta, é suficiente mostrar que o espaço  $N_c^\varphi$  é fechado com a norma  $\|\cdot\|_{N_c^\varphi}$ .

Com efeito, seja  $(u_n) \subset N_c^\varphi$  uma sequência que converge para  $u \in L^0$ , isto é,  $\|u_n - u\|_{N_c^\varphi} \rightarrow 0$ . Por definição,

$$\forall \varepsilon > 0, \exists n_0 \in \mathbb{N}; n \geq n_0 \implies \|u_n - u\|_{N_c^\varphi} < \varepsilon.$$

Como  $\{u_n\} \subset N_c^\varphi$ , para cada  $n \in \mathbb{N}$  existe  $\lambda_n > 0$  tal que  $|\lambda_n u_n| \leq c - a_\varphi$ . Além disso  $\|u_{n_0} - u\|_{N_c^\varphi} < \varepsilon$ . Desse modo

$$\frac{1}{\varepsilon} |u_{n_0} - u| \leq c - a_\varphi$$

e portanto

$$|u_{n_0} - u| \leq \varepsilon(c - a_\varphi).$$

Tomando  $0 < \lambda_0 < \frac{\lambda_{n_0}}{2}$  temos

$$|2\lambda_0 u_{n_0}| \leq |\lambda_{n_0} u_{n_0}| < (c - a_\varphi)$$

$$\implies |\lambda_0 u_{n_0}| \leq \frac{c - a_\varphi}{2}.$$

Tome agora  $\varepsilon < \frac{1}{2\lambda_0}$ . Assim,

$$\begin{aligned} \lambda_0 |u| &\leq \lambda_0 |u_{n_0} - u| + \lambda_0 |u_{n_0}| \\ &\leq \lambda_0 \varepsilon (c - a_\varphi) + \frac{1}{2} (c - a_\varphi) \\ &\leq \frac{1}{2} (c - a_\varphi) + \frac{1}{2} (c - a_\varphi) \\ &= (c - a_\varphi) \end{aligned}$$

Isso prova que  $u \in N_c^\varphi$ .  $\square$

O próximo passo será refinar o conjunto  $N_c^\varphi$  com o espaço Musielak-Orlicz  $L_c^\varphi$ , de modo a obter um outro espaço de Banach de onde retiraremos o domínio da parametrização.

**Lema 2** O espaço  $\mathcal{N}_c^\varphi = N_c^\varphi \cap L_c^\varphi$  é de Banach com a norma

$$\|u\| = \|u\|_{L_c^\varphi} + \|u\|_{N_c^\varphi}. \quad (12)$$

**Demonstração:** Seja  $\{u_n\} \subset \mathcal{N}_c^\varphi$  uma sequência que converge para  $u$  na norma descrita na

equação (12). Dessa maneira, dado  $\varepsilon > 0$  existe  $n_0 \in \mathbb{N}$  tal que se  $n > n_0$  temos

$$\|u_n - u\|_{L_c^\varphi} < \varepsilon \quad \text{e} \quad \|u_n - u\|_{N_c^\varphi} < \varepsilon.$$

Como os espaços  $L_c^\varphi$  e  $N_c^\varphi$  são de Banach, segue que  $u \in L_c^\varphi \cap N_c^\varphi = \mathcal{K}_c^\varphi$   $\square$

**Lema 3** *O conjunto*

$$\mathcal{K}_c^\varphi = \left\{ u \in L_c^\varphi; \exists \varepsilon > 0, \text{ onde } \int_T \varphi(c + \lambda u) d\mu < \infty, \text{ para cada } \lambda \in (-\varepsilon, 1 + \varepsilon) \right\} \quad (13)$$

é convexo e aberto em  $L_c^\varphi$ .

**Demonstração:**

1)  $\mathcal{K}_c^\varphi$  é convexo. De fato, sejam  $u, v \in \mathcal{K}_c^\varphi$  e  $\alpha \in (0, 1)$ . Então, existem  $\varepsilon_1, \varepsilon_2 > 0$ , tais que

$$\int_T \varphi(c + \lambda u) d\mu < \infty, \text{ para cada } \lambda \in (-\varepsilon_1, 1 + \varepsilon_1)$$

e

$$\int_T \varphi(c + \lambda v) d\mu < \infty, \text{ para cada } \lambda \in (-\varepsilon_2, 1 + \varepsilon_2)$$

Tomando  $\varepsilon = \min\{\varepsilon_1, \varepsilon_2\}$  e considerando a convexidade de  $\varphi$ , temos

$$\int_T \varphi(c + \lambda[(1 - \alpha)u + \alpha v]) d\mu \leq (1 - \alpha) \int_T \varphi(c + \lambda u) d\mu + \alpha \int_T \varphi(c + \lambda v) d\mu < \infty,$$

para todo  $\lambda \in (-\varepsilon, 1 + \varepsilon)$ .

2) O conjunto  $\mathcal{K}_c^\varphi$  é aberto em  $L_c^\varphi$ . Com efeito, seja  $u \in \mathcal{K}_c^\varphi$ . Então, existe  $\varepsilon \in (0, 1)$ , tal que

$$\int_T \varphi(c + \alpha u) d\mu < \infty$$

para cada  $\alpha \in [-\varepsilon, 1 + \varepsilon]$ . Considerando  $\delta = \left[ \frac{2}{\varepsilon}(1 + \varepsilon) \left( 1 + \frac{\varepsilon}{2} \right) \right]^{-1}$  para qualquer  $v \in B_\delta(0) = \{w \in L_c^\varphi; I_{\Phi_c}(\frac{v}{\delta}) \leq 1\}$  e consequentemente

$$\int_T \varphi\left(c + \frac{1}{\delta}|v|\right) d\mu \leq 2.$$

Dado  $\alpha \in \left(0, 1 + \frac{\varepsilon}{2}\right)$  denotemos  $\lambda = \frac{\alpha}{1 + \varepsilon}$ . A desigualdade

$$\frac{\alpha}{1 - \lambda} = \frac{\alpha}{1 - \frac{\alpha}{1 + \varepsilon}} \leq \frac{1 + \frac{\varepsilon}{2}}{1 - \frac{1 + \frac{\varepsilon}{2}}{1 + \varepsilon}} = \frac{2}{\varepsilon}(1 + \varepsilon) \left( 1 + \frac{\varepsilon}{2} \right) = \frac{1}{\delta},$$

implica

$$\begin{aligned}\varphi(c + \alpha(u + v)) &\leq \varphi\left(\lambda\varphi\left(c + \frac{\alpha}{\lambda}u\right) + (1 - \lambda)\varphi\left(c + \frac{\alpha}{1 - \lambda}v\right)\right) \\ &\leq \lambda\varphi\left(c + \frac{\alpha}{\lambda}u\right) + (1 - \lambda)\varphi\left(c + \frac{\alpha}{1 - \lambda}v\right) \\ &\leq \lambda\varphi(c + (1 + \varepsilon)u) + (1 - \lambda)\varphi\left(c + \frac{1}{\delta}|v|\right).\end{aligned}$$

Para  $\alpha \in \left(-\frac{\varepsilon}{2}, 0\right)$ , podemos escrever

$$\begin{aligned}\varphi(c + \alpha(u + v)) &\leq \frac{1}{2}\varphi(c + 2\alpha u) + \frac{1}{2}\varphi(c + 2\alpha v) \\ &\leq \frac{1}{2}\varphi(c + 2\alpha u) + \frac{1}{2}\varphi(c + |v|).\end{aligned}$$

Então,

$$\int_T \varphi(c + \alpha(u + v)) d\mu < \infty,$$

para qualquer  $\alpha \in \left(-\frac{\varepsilon}{2}, 1 + \frac{\varepsilon}{2}\right)$ . Assim,  $u + v \in \mathcal{K}_c^\varphi$ . E desse modo,  $B_\delta(u)$  está contido em  $\mathcal{K}_c^\varphi$  e, portanto, o conjunto  $\mathcal{K}_c^\varphi$  é aberto  $\square$

Note que dado  $u \in \mathcal{K}_c^\varphi$ , temos que

$$\int_T \varphi(c + u) d\mu < \infty.$$

Note que se  $c + u \leq a_\varphi$  não temos  $\varphi(c + u) \in \mathcal{P}_\mu$ . Para resolver esse problema, precisamos definir um conjunto para intersectar com  $\mathcal{K}_c^\varphi$ .

$$\mathcal{M}_c^\varphi = \{u \in N_c^\varphi; \lambda u < c - a_\varphi \text{ para } \lambda \in (-\varepsilon, 1 + \varepsilon)\}.$$

**Proposição 1** *O conjunto  $\mathcal{M}_c^\varphi$  é aberto.*

**Demonstração:** Usaremos a mesma ideia do Lema 3.

Seja  $u \in \mathcal{M}_c^\varphi$ . Então, existe  $\varepsilon \in (0, 1)$ , tal que

$$\alpha u < c - a_\varphi$$



para todo  $\alpha \in [-\varepsilon, 1 + \varepsilon]$ . Em particular, temos

$$(1 + \varepsilon)u \leq c - a_\varphi.$$

Considere  $\delta = \left[ \frac{2}{\varepsilon}(1 + \varepsilon) \left(1 + \frac{\varepsilon}{2}\right) \right]^{-1}$ . Assim, dado

$$v \in B_\delta(0) = \{w \in N_c^\varphi; \|w\|_{N_c^\varphi} < \delta\},$$

obtemos

$$\frac{1}{\delta}|v| < c - a_\varphi.$$

Tomando  $\lambda = \frac{\alpha}{1 + \varepsilon}$  onde  $\alpha \in (0, 1 + \frac{\varepsilon}{2})$ , temos que

$$\frac{\alpha}{1 - \lambda} \leq \frac{1}{\delta}$$

e desse modo

$$\begin{aligned} \alpha(u + v) &= \lambda \frac{\alpha}{\lambda} u + (1 - \lambda) \frac{\alpha}{(1 - \lambda)} v \\ &\leq \lambda(1 + \varepsilon)u + (1 - \lambda) \frac{1}{\delta} v \\ &\leq \lambda(c - a_\varphi) + (1 - \lambda)(c - a_\varphi) \\ &\leq c - a_\varphi. \end{aligned}$$

No caso em que  $\alpha \in (-\frac{\varepsilon}{2}, 0)$ , obtemos

$$\begin{aligned} \alpha(u + v) &= \frac{1}{2}(2\alpha u) + \frac{1}{2}(2\alpha v) \\ &\leq \frac{1}{2}2\alpha u + \frac{1}{2\delta}|v| \\ &\leq \frac{1}{2}(c - a_\varphi) + \frac{1}{2}(c - a_\varphi) = c - a_\varphi \end{aligned}$$

e, dessa forma,

$$\alpha(u + v) < c - a_\varphi$$

para cada  $\alpha \in (-\varepsilon, 1 + \varepsilon)$ .

Portanto,  $B_\delta(u) \subset \mathcal{M}_c^\varphi$  e consequentemente  $\mathcal{M}_c^\varphi$  é aberto.  $\square$

Para o que se segue, precisamos fixar uma função mensurável  $c : T \rightarrow \mathbb{R}$  tal que  $\varphi(c) \in \mathcal{P}_\mu$ . Dado  $u \in \mathcal{E}_c^\varphi := \mathcal{H}_c^\varphi \cap \mathcal{M}_c^\varphi$ , é claro que nem sempre temos que  $\varphi(c+u) \in \mathcal{P}_\mu$ . Vamos ver logo abaixo que a função  $q$ -exponencial definida como no Exemplo 5 é tal que existe  $u_0 \in E_c^\varphi \cap N_c^\varphi$ .

**Exemplo 6** Dado  $\alpha \geq 1$ , consideremos dois casos:

Se  $u \leq 0$  temos

$$\begin{aligned} \exp_q(\alpha u) &\leq \exp_q(u) \\ &\leq \alpha^{\frac{1}{1-q}} \exp_q(u). \end{aligned}$$

Caso  $u > 0$ , obtemos

$$\begin{aligned} \exp_q(\alpha u) &= (1 + (1-q)\alpha u)^{\frac{1}{1-q}} \\ &= (\alpha\alpha^{-1} + (1-q)\alpha u)^{\frac{1}{1-q}} \\ &= \alpha^{\frac{1}{1-q}} (\alpha^{-1} + (1-q)u)^{\frac{1}{1-q}} \\ &\leq \alpha^{\frac{1}{1-q}} (1 + (1-q)u)^{\frac{1}{1-q}} \\ &= \alpha^{\frac{1}{1-q}} \exp_q(u). \end{aligned}$$

Pela convexidade de  $\exp_q(t, \cdot)$ , obtemos para qualquer  $\lambda \in (0, 1)$

$$\begin{aligned} \exp_q(c+u) &\leq \lambda \exp_q(\lambda^{-1}c) + (1-\lambda) \exp_q((1-\lambda)^{-1}u) \\ &\leq \lambda^{1-1/(1-q)} \exp_q(c) + (1-\lambda)^{1-1/(1-q)} \exp_q(u). \end{aligned}$$

Assim, qualquer função positiva  $u_0 : T \rightarrow (0, \infty)$  com

$$\int_T \exp_q(u_0) d\mu < \infty$$

satisfaz

$$\int_T \exp_q(c + \lambda u_0) d\mu < \infty$$

para todo  $\lambda > 0$ . Note que a função  $u_0$  é tal que  $c + \lambda u_0 > c > a_\varphi$  para todo  $\lambda > 0$ . Portanto,

$$u_0 \in E_c^\varphi \cap N_c^\varphi. \quad \square$$

**Exemplo 7** No exemplo acima encontramos uma função exponencial deformada satisfazendo a condição (11), tal que existe  $u_0 \in E_c^\Phi \cap N_c^\Phi$ . Agora ilustraremos a situação oposta. Considere a função

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

em que a medida  $\mu$  é  $\sigma$ -finita e não-atômica. Note que  $\varphi$  é convexa,  $\lim_{u \rightarrow \infty} \varphi(u) = \infty$  e  $\varphi(x) = 0$ , para todo  $x < a_\varphi$ . Encontraremos uma função mensurável  $c : T \rightarrow \mathbb{R}$  com

$$\int_T \varphi(c) d\mu < \infty,$$

mas

$$\int_T \varphi(c + u_0) d\mu = \infty.$$

Com efeito, para cada  $m \geq 1$ , consideremos

$$v_m(t) := \left( m \frac{\log(2)}{u_0} - \frac{u_0}{2} - 1 \right) \chi_{E_m}(t),$$

onde

$$E_m = \left\{ t \in T; m \frac{\log(2)}{u_0} - \frac{u_0}{2} - 1 > 0 \right\}$$

e

$$\chi_{E_m}(t) := \begin{cases} 1, & t \in E_m \\ 0, & t \notin E_m \end{cases}.$$

Desde que  $v_m \uparrow \infty$ , podemos encontrar uma subsequência  $\{v_{m_n}\}$  tal que

$$\int_{E_{m_n}} e^{(v_{m_n} + u_0 + 1)^2/2} d\mu \geq 2^n.$$

De acordo com (MUSIELAK, 1983), existe uma subsequência  $w_k = v_{m_{n_k}}$  e conjuntos mensuráveis dois a dois disjuntos  $A_k \subseteq E_{m_{n_k}}$  para os quais

$$\int_{A_k} e^{(w_k + u_0 + 1)^2/2} d\mu = 1.$$

Vamos definir

$$c = \bar{c} \mathbf{1}_{T \setminus A} + \sum_{k=1}^{\infty} w_k \mathbf{1}_{A_k},$$

em que  $A = \bigcup_{k=1}^{\infty} A_k$  e  $\bar{c}$  é qualquer função mensurável tal que  $\varphi(\bar{c}(t)) > 0$  para  $t \in T \setminus A$  e

$$\int_{T \setminus A} \varphi(\bar{c}) d\mu < \infty.$$

Para  $t \in A_k$

$$e^{(w_k(t)+u_0+1)^2/2} = 2^{m_{n_k}} e^{(w_k(t)+1)^2/2},$$

e temos que

$$\int_{A_k} e^{(w_k(t)+1)^2/2} d\mu = \frac{1}{2^{m_{n_k}}},$$

para todo  $m \geq 1$ . Assim, podemos escrever

$$\begin{aligned} \int_T \varphi(c) d\mu &= \int_{T \setminus A} \varphi(\bar{c}) d\mu + \sum_{k=1}^{\infty} \int_{A_k} e^{(w_k(t)+1)^2/2} d\mu \\ &= \int_{T \setminus A} \varphi(\bar{c}) d\mu + \sum_{k=1}^{\infty} \frac{1}{2^{m_{n_k}}} \\ &< \infty. \end{aligned}$$

Por outro lado, também temos

$$\begin{aligned} \int_T \varphi(c + u_0) d\mu &= \int_{T \setminus A} \varphi(\bar{c}) d\mu + \sum_{k=1}^{\infty} \int_{A_k} e^{(u_0+w_k(t)+1)^2/2} d\mu \\ &= \int_{T \setminus A} \varphi(\bar{c}) d\mu + \sum_{k=1}^{\infty} 1 \\ &= \infty, \end{aligned}$$

o qual mostra que a condição  $u_0 \in E_c^\varphi$  não é satisfeita.

Em decorrência dos dois últimos exemplos, considere uma exponencial deformada satisfazendo (11) e para cada  $c$ , tal que  $\varphi(c) \in \mathcal{P}_\mu$ , seja  $u_0 \in E_c^\varphi \cap N_c^\varphi$  uma função mensurável e positiva. Em outras palavras, além da condição (11) precisamos adicionar as hipóteses:

”Fixada uma função mensurável  $c : T \rightarrow \mathbb{R}$ , existe uma função mensurável  $u_0 : T \rightarrow \mathbb{R}$ , tal que

$$\int_T \varphi(c + \lambda u_0) d\mu < \infty, \quad (14)$$

para todo  $\lambda > 0$  e

$$c - \varepsilon u_0 > a_\varphi, \quad (15)$$

para algum  $\varepsilon > 0$ .”

Dessa forma, podemos encontrar uma vizinhança da origem em  $\mathcal{E}_c^\varphi$ , onde existirá um único  $\psi(u)$ , tal que  $\varphi(c + u - \psi(u)u_0) \in \mathcal{P}_\mu$ . Na próxima proposição, daremos detalhes sobre essa construção.

**Proposição 2** *Podemos encontrar  $\delta > 0$  tal que, para cada  $u \in \mathcal{N}_c^\varphi$  com  $\|u\|_{\mathcal{N}_c^\varphi} < \delta$ , existe um único  $\psi(u) \in \mathbb{R}$  para o qual*

$$\varphi(c + u - \psi(u)u_0) \in \mathcal{P}_\mu.$$

**Demonstração:** Assuma que  $u \neq 0$ . Seja  $\varepsilon > 0$  tal que  $2\varepsilon u_0 < c - a_\varphi$ . Observe que

$$0 < \varepsilon \leq \frac{1}{2} \frac{1}{\|u_0\|_{\mathcal{N}_c^\varphi}}.$$

Seja  $\eta \in (0, 1)$  tal que

$$\int_T \varphi(c - 2\varepsilon u_0) d\mu = 1 - \eta. \quad (16)$$

Tome  $\delta \in (0, \frac{1}{2}\eta)$  e fixe qualquer  $u \in \mathcal{N}_c^\varphi$  com  $\|u\|_{\mathcal{N}_c^\varphi} < \delta$ . Desde que

$$\|u\|_{\mathcal{N}_c^\varphi} \leq \|u\|_{\mathcal{N}_c^\varphi} < \delta < \frac{1}{2},$$

temos

$$|2u| < c - a_\varphi.$$

Assim,

$$\begin{aligned} |u - \varepsilon u_0| &\leq \frac{1}{2}|2u| + \frac{1}{2}|2\varepsilon u_0| \\ &< \frac{1}{2}(c - a_\varphi) + \frac{1}{2}(c - a_\varphi) \\ &= c - a_\varphi. \end{aligned}$$

E, conseqüentemente,  $c + u - \varepsilon u_0 > a_\varphi$ .

Observando que  $\|u\|_{\Phi_c} \leq \|u\|_{\mathcal{N}_c^\varphi} < \delta < \frac{1}{2}\eta$ , obtemos

$$\begin{aligned} \int_T \Phi_c(2u) d\mu &\leq 2\delta \int_T \Phi_c\left(\frac{u}{\delta}\right) d\mu \\ &\leq 2\delta \\ &< \eta \end{aligned}$$

e desse modo

$$\begin{aligned} \int_T \varphi(c+2|u|) d\mu &= \int_T \varphi(c) d\mu + \int_T \Phi_c(2u) d\mu \\ &< 1 + \eta \end{aligned} \tag{17}$$

Das equações (16) e (17), segue que

$$\begin{aligned} \int_T \varphi(c+u-\varepsilon u_0) d\mu &\leq \int_T \varphi(c+|u|-\varepsilon u_0) d\mu \\ &= \int_T \varphi\left(\frac{1}{2}(c+2|u|) + \frac{1}{2}(c-2\varepsilon u_0)\right) d\mu \\ &\leq \frac{1}{2} \int_T \varphi(c+2|u|) d\mu + \frac{1}{2} \int_T \varphi(c-2\varepsilon u_0) d\mu \\ &< \frac{1}{2}(1+\eta) + \frac{1}{2}(1-\eta) = 1 \end{aligned}$$

e para qualquer  $\lambda > 0$  temos

$$\begin{aligned} \int_T \varphi(c+u+\lambda u_0) d\mu &\leq \int_T \varphi(c+|u|+\lambda u_0) d\mu \\ &= \int_T \varphi\left(\frac{1}{2}(c+2|u|) + \frac{1}{2}(c+2\lambda u_0)\right) d\mu \\ &\leq \frac{1}{2} \int_T \varphi(c+2|u|) d\mu + \frac{1}{2} \int_T \varphi(c+2\lambda u_0) d\mu \\ &< \infty. \end{aligned}$$

Pela continuidade de  $\varphi$  e pelo Teorema da Convergência Dominada, podemos con-

cluír que se  $u \in \mathcal{N}_c^\varphi$  com  $\|u\|_{\mathcal{N}_c^\varphi} < \delta$ , então existe um único número real  $\psi(u) < \varepsilon$  satisfazendo

$$\varphi(c + u - \psi(u)u_0) \in \mathcal{P}_\mu \quad \square$$

**Observação 6** Para  $\delta < \frac{1}{2}$ , denotemos por  $\mathcal{G}_c^\varphi$  a bola descrita na proposição 2. Note que dado  $u \in \mathcal{G}_c^\varphi$ , existe  $\varepsilon > 0$ , tal que

$$a_\varphi < c + \lambda u \quad e \quad \int_T \varphi(c + \lambda u) d\mu < \infty,$$

para todo  $\lambda \in (-\varepsilon, 1 + \varepsilon)$ . Em particular,  $\mathcal{G}_c^\varphi \subset \mathcal{E}_c^\varphi := \mathcal{K}_c^\varphi \cap \mathcal{M}_c^\varphi$ .

Considere agora

$$\begin{aligned} \widetilde{B}_c^\varphi &= B_c^\varphi \cap \mathcal{N}_c^\varphi \\ &= \left\{ u \in \mathcal{N}_c^\varphi; \int_T u \varphi'_+(c) d\mu = 0 \right\} \end{aligned}$$

e faça a interseção com a bola  $\mathcal{G}_c^\varphi$  :

$$\mathcal{A}_c^\varphi := \widetilde{B}_c^\varphi \cap \mathcal{G}_c^\varphi.$$

**Proposição 3** A restrição da função normalizadora  $\psi$  ao conjunto  $\mathcal{A}_c^\varphi$  é não negativa.

**Demonstração:** De fato, uma vez que  $\varphi(t, \cdot)$  é convexa temos

$$u \varphi'_+(t, c(t)) \leq \varphi(t, c(t) + u) - \varphi(t, c(t)), \quad \text{para todo } u \in \mathbb{R}.$$

Assim, para qualquer  $u \in \mathcal{A}_c^\varphi$  obtemos

$$\begin{aligned} 1 &= \int_T u \varphi'_+(c) d\mu + \int_T \varphi(c) d\mu \\ &\leq \int_T \varphi(c + u) d\mu \\ &< \infty. \end{aligned}$$

Portanto, para qualquer  $u \in \mathcal{A}_c^\varphi$  existe um único número real positivo  $\psi(u)$  tal que

$$\varphi_c(u) = \varphi(c + u - \psi(u)u_0) \in \mathcal{P}_\mu \quad \square$$

Outro fato imediato sobre  $\mathcal{P}_\mu$  é que a construção feita para a obtenção do conjunto

$\mathcal{A}_c^\varphi$  garante que

$$\mathcal{P}_\mu = \bigcup_{\varphi(c) \in \mathcal{P}_\mu} \varphi_c(\mathcal{A}_c^\varphi).$$

Os resultados seguintes servirão para definir a  $\varphi$ -família e garantir que quando duas  $\varphi$ -famílias se intersectam, a função  $u_0$  é a mesma. Além disso, teremos que a aplicação de transição é de classe  $C^\infty$ .

**Lema 4** Fixado  $u \in \mathcal{A}_c^\varphi$ , denote  $\tilde{c} = c + u - \psi(u)u_0$ . Então  $u_0 \in E_{\tilde{c}}^\varphi \cap N_{\tilde{c}}^\varphi$ .

**Demonstração:** É claro que  $u_0$  está em  $E_{\tilde{c}}^\varphi$ . Isso decorre de (18).

Falta mostrar que  $u_0 \in N_{\tilde{c}}^\varphi$ . Se  $\varepsilon > 0$  é o mesmo encontrado na prova da Proposição 2, então

$$\tilde{c} - [\varepsilon - \psi(u)]u_0 = c + u - \varepsilon u_0 > a_\varphi \quad \square$$

Na próxima proposição, encontramos uma condição de inclusão para os espaços  $N_c^\varphi$  e posteriormente vamos relacionar com classes de equivalência.

**Proposição 4** Assuma que as funções mensuráveis  $\tilde{c}, c : T \rightarrow \mathbb{R}$  são tais que  $\tilde{c} > a_\varphi$  e  $c > a_\varphi$ . Então  $N_{\tilde{c}}^\varphi \subseteq N_c^\varphi$  se, e somente se,  $\tilde{c} - c \in N_c^\varphi$ .

**Demonstração:** Suponha que  $N_{\tilde{c}}^\varphi \subseteq N_c^\varphi$  e defina o conjunto

$$A = \{t \in T : \tilde{c}(t) < c(t)\}.$$

Para  $\tilde{c} > a_\varphi$ , as desigualdades

$$(c - \tilde{c})\chi_A < (c - a_\varphi)\chi_A \leq c - a_\varphi$$

implicam que

$$(c - \tilde{c})\chi_A \in N_c^\varphi.$$

Pela desigualdade  $c > a_\varphi$ ,

$$(\tilde{c} - c)\chi_{T \setminus A} < (\tilde{c} - a_\varphi)\chi_{T \setminus A} \leq \tilde{c} - a_\varphi,$$

e portanto

$$(\tilde{c} - c)\chi_{T \setminus A} \in N_{\tilde{c}}^\varphi \subseteq N_c^\varphi.$$

Escrevendo

$$|\tilde{c} - c| = (c - \tilde{c})\chi_A + (\tilde{c} - c)\chi_{T \setminus A},$$

concluimos que  $\tilde{c} - c \in N_c^\varphi$ .



Reciprocamente, suponha que  $\tilde{c} - c \in N_c^\varphi$ . Para  $u \in N_{\tilde{c}}^\varphi$  fixado, seja  $\lambda > 0$  tal que

$$|\lambda(\tilde{c} - c)| \leq c - a_\varphi \quad \text{e} \quad |\lambda u| \leq \tilde{c} - a_\varphi.$$

As desigualdades

$$|\lambda u \chi_A| \leq (\tilde{c} - a_\varphi) \chi_A < c - a_\varphi$$

fornecem  $u \chi_A \in N_c^\varphi$ . Pela escolha de  $\lambda$ , tem-se

$$\begin{aligned} \left| \frac{\lambda^2}{1+\lambda} u \chi_{T \setminus A} \right| &\leq \frac{\lambda}{1+\lambda} (\tilde{c} - a_\varphi) \chi_{T \setminus A} \\ &= \frac{1}{1+\lambda} [\lambda(\tilde{c} - c) \chi_{T \setminus A} + \lambda(c - a_\varphi) \chi_{T \setminus A}] \\ &\leq \frac{1}{1+\lambda} [(c - a_\varphi) + \lambda(c - a_\varphi)] \\ &= c - a_\varphi. \end{aligned}$$

Assim,  $u \chi_{T \setminus A} \in N_c^\varphi$  e conseqüentemente  $u = u \chi_A + u \chi_{T \setminus A} \in N_c^\varphi$ .  $\square$

**Proposição 5** *Assuma que as funções mensuráveis  $\tilde{c}, c : T \rightarrow \mathbb{R}$  são tais que  $\tilde{c}, c > a_\varphi$ ,*

$$\int_T \varphi(\tilde{c}) d\mu < \infty \quad \text{e} \quad \int_T \varphi(c) d\mu < \infty.$$

*Se  $\tilde{c} - c \in \mathcal{N}_c^\varphi$  então  $\mathcal{N}_{\tilde{c}}^\varphi \subseteq \mathcal{N}_c^\varphi$ .*

**Demonstração:** Pela Proposição 4, a inclusão  $N_{\tilde{c}}^\varphi \subseteq N_c^\varphi$  é válida. Assim, resta mostrar que se  $u \in \mathcal{N}_{\tilde{c}}^\varphi$  tem-se  $u \in L_c^\varphi$ .

Como  $\tilde{c} - c \in \mathcal{N}_c^\varphi$ , podemos encontrar  $\eta > 0$  tal que

$$c + \alpha(\tilde{c} - c) > a_\varphi \quad \text{e} \quad \int_T \varphi(c + \alpha(\tilde{c} - c)) d\mu < \infty, \quad (18)$$

para todo  $\alpha \in (-2\eta, 1]$ .

Em adição, como  $u \in \mathcal{N}_{\tilde{c}}^\varphi \subseteq N_{\tilde{c}}^\varphi$ , para algum  $\varepsilon > 0$ , a função  $u$  satisfaz:

$$\begin{cases} c + \lambda u > a_\varphi \\ \tilde{c} + \lambda u > a_\varphi \end{cases} \quad \text{e} \quad \int_T \varphi(\tilde{c} + \lambda u) d\mu < \infty, \quad (19)$$

para todo  $\lambda \in (-2\varepsilon, 2\varepsilon)$ .

Pela combinação convexa envolvendo  $c + \lambda u > a_\varphi$  e  $\tilde{c} + \lambda u > a_\varphi$ , temos

$$c + \alpha(\tilde{c} - c) + \lambda u > a_\varphi, \quad (20)$$

para  $\alpha \in [0, 1]$  e  $\lambda \in (-2\varepsilon, 2\varepsilon)$ . Combinando  $c + \alpha(\tilde{c} - c) > a_\varphi$  e  $c + \lambda u > a_\varphi$ , podemos concluir que a desigualdade (20) é também satisfeita para  $\alpha \in (-\eta, \frac{1}{2}]$  e  $\lambda \in (-\varepsilon, \varepsilon)$ . Então (20) ocorre para  $\alpha \in (-\eta, 1]$  e  $\lambda \in (-\varepsilon, \varepsilon)$ . Vamos definir a função convexa

$$g(\alpha, \lambda) = \int_T \varphi(c + \alpha(\tilde{c} - c) + \lambda u) d\mu.$$

Expressões (18) e (19) implicam em  $g(\alpha, 0) < \infty$  para  $\alpha \in (-\eta, 1]$ , e  $g(1, \lambda) < \infty$  para  $\lambda \in (-\varepsilon, \varepsilon)$ . Como  $g$  é convexa, segue que  $g$  é finita na envoltória convexa da união de  $(-\eta, 1] \times 0$  e  $1 \times (-\varepsilon, \varepsilon)$ . Em particular,

$$g(0, \lambda) = \int_T \varphi(\tilde{c} + \lambda u) d\mu < \infty$$

para cada  $\lambda$  em alguma vizinhança de 0 e portanto  $u \in L_c^\varphi$ .  $\square$

O próximo Lema estabelece uma condição de igualdade para os espaços  $\mathcal{N}_c^\varphi$ .

**Lema 5** *Seja  $\tilde{c} = c + u - \psi(u)u_0$ , onde  $u \in \mathcal{A}_c^\varphi$ . Então  $\mathcal{N}_c^\varphi = \mathcal{N}_{\tilde{c}}^\varphi$ .*

**Demonstração:** Claramente  $\tilde{c} - c \in \mathcal{N}_c^\varphi$ . Isto significa que  $\mathcal{N}_{\tilde{c}}^\varphi \subseteq \mathcal{N}_c^\varphi$ . Resta mostrar a inclusão oposta.

Seja  $\varepsilon > 0$  tal que

$$c - a_\varphi > -\lambda u \quad \text{e} \quad \int_T \varphi(c + \lambda u) d\mu < \infty. \quad (21)$$

para  $\lambda \in (-\varepsilon, 1 + \varepsilon)$ . Como resultado temos

$$\lambda u - u < c - a_\varphi, \quad \text{para } \lambda \in (-\varepsilon, \varepsilon),$$

e

$$|\lambda u| < c + u - a_\varphi, \quad \text{para } \lambda \in (0, \varepsilon).$$

Pelo Lema 4, podemos encontrar  $\alpha > 0$  tal que

$$\alpha \psi(u)u_0 < \tilde{c} - a_\varphi.$$

Para qualquer  $\lambda \in (0, \varepsilon)$ , temos

$$\begin{aligned}
\left| \lambda \frac{\alpha}{1+\alpha} u \right| &< \frac{\alpha}{1+\alpha} (c+u-a_\varphi) \\
&= \frac{\alpha}{1+\alpha} [(c+u-\psi(u)u_0) - a_\varphi + \psi(u)u_0] \\
&= \frac{1}{1+\alpha} [\alpha(\tilde{c}-a_\varphi) + \alpha\psi(u)u_0] \\
&< \frac{1}{1+\alpha} [\alpha(\tilde{c}-a_\varphi) + (\tilde{c}-a_\varphi)] \\
&= \tilde{c}-a_\varphi.
\end{aligned} \tag{22}$$

Assim,  $u \in N_{\tilde{c}}^\varphi$ .

Denote  $\varepsilon' = \frac{\alpha}{1+\alpha}\varepsilon < \varepsilon$ . Por (22), temos  $\tilde{c} + \lambda'u > a_\varphi$  para todo  $\lambda' \in (-\varepsilon', \varepsilon')$ . A equação (21) implica

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

para todo  $\lambda' \in (-\varepsilon', \varepsilon')$  e conseqüentemente  $u \in L_{\tilde{c}}^\varphi$ . Portanto,

$$c - \tilde{c} = -u + \psi(u)u_0 \in \mathcal{N}_{\tilde{c}}^\varphi,$$

e a inclusão  $\mathcal{N}_c^\varphi \subseteq \mathcal{N}_{\tilde{c}}^\varphi$  é estabelecida.  $\square$

No conjunto

$$\mathcal{C} = \{c \in L^0; \varphi(c) \in \mathcal{P}_\mu\},$$

considere a relação de equivalência:

$$c \sim \tilde{c} \Leftrightarrow \mathcal{N}_c^\varphi = \mathcal{N}_{\tilde{c}}^\varphi.$$

Para cada classe de equivalência  $[c]$  fixamos alguma função positiva  $u_{0,[c]} \in \mathcal{N}_c^\varphi$  tal que

$$\int_T \varphi(\tilde{c} + \lambda u_{0,[c]}) d\mu < \infty \tag{23}$$

para todo  $\lambda > 0$  e  $\tilde{c} \in [c]$ .

Dado  $c \in \mathcal{C}$ , o conjunto

$$\mathcal{F}_c^\varphi = \{\varphi(c+u-\psi(u)u_{0,[c]}) : u \in \mathcal{A}_c^\varphi\}$$

é chamado de  $\varphi$ -família.

É claro que a união de todas as  $\varphi$ -famílias cobre o espaço  $\mathcal{P}_\mu$ . Além disso, dadas as funções mensuráveis  $c_1, c_2 : T \rightarrow \mathbb{R}$ , tais que  $\varphi(c_1), \varphi(c_2) \in \mathcal{P}_\mu$  e  $\mathcal{F}_{c_1}^\varphi \cap \mathcal{F}_{c_2}^\varphi \neq \emptyset$ , temos pelo Lema 5 que  $\mathcal{N}_{c_1}^\varphi = \mathcal{N}_{c_2}^\varphi$  e portanto  $u_0 = u_{0,[c_1]} = u_{0,[c_2]}$ .

E mais, sejam  $u \in \mathcal{A}_{c_1}^\varphi$  e  $v \in \mathcal{A}_{c_2}^\varphi$  tais que

$$c_1 + u - \psi_1(u)u_0 = c_2 + v - \psi_2(v)u_0.$$

A aplicação de transição é  $C^\infty$  e é dada pela expressão

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

## 2.4.2 Fibrado tangente

Na seção anterior, a expressão da aplicação de transição  $\varphi_{c_2}^{-1} \circ \varphi_{c_1}$  foi importante para garantir que  $\mathcal{P}_\mu$  pode ser equipado com uma estrutura  $C^\infty$ -Banach. Agora, usaremos a aplicação de transição para encontrar o espaço tangente de  $\mathcal{P}_\mu$  em um ponto  $p = \varphi(c)$  e o fibrado tangente (VIEIRA *et al.*, 2019b).

Dado  $p \in \mathcal{P}_\mu$ , consideremos a tripla  $(\mathcal{F}_c^\varphi; \varphi_c^{-1}; v)$ , onde  $\mathcal{F}_c^\varphi$  é a  $\varphi$ -família,  $\varphi_c$  é a parametrização e  $v$  é um vetor em  $\varphi_c^{-1}(\mathcal{F}_c^\varphi)$  o qual está contido no espaço vetorial  $\mathcal{N}_c^\varphi$ .

Nesse caso,

$$(\mathcal{F}_c^\varphi; \varphi_c^{-1}; v) \sim (\mathcal{F}_{\tilde{c}}^\varphi; \varphi_{\tilde{c}}^{-1}; w) \Leftrightarrow (\varphi_{\tilde{c}}^{-1} \circ \varphi_c)'(\varphi_c^{-1}(p))(v) = w.$$

O espaço tangente  $T_p(\mathcal{P}_\mu)$  é formado pelas classes  $[\mathcal{F}_c^\varphi; \varphi_c^{-1}; v]$ .

O vetor  $v \in \varphi_c^{-1}(\mathcal{F}_c^\varphi)$  é o vetor velocidade de uma curva no domínio da parametrização. De fato, considere  $(\mathcal{F}_{c_1}^\varphi, \varphi_{c_1}^{-1})$  e  $(\mathcal{F}_{c_2}^\varphi, \varphi_{c_2}^{-1})$  sejam as cartas sobre  $p \in \mathcal{P}_\mu$  e

$$g : I \subset T \rightarrow \mathcal{P}_\mu$$

uma curva tal que  $g(t_0) = p$ , para algum  $t_0 \in T$ . Tomando

$$g(t) = \varphi_{c_1}(u_1) = \varphi(c_1 + u_1 - \psi(u_1)u_0),$$

temos que  $u_1(t) = \varphi_{c_1}^{-1}(g(t))$ . Além disso,  $g(t) = \varphi_{c_1}(u_1)$  e  $u_2(t) = \varphi_{c_2}^{-1}(g(t))$ . Usando variáveis aleatórias temos que

$$u_2(t_0) = \varphi_{c_2}^{-1}(g(t_0)) = (\varphi_{c_2}^{-1} \circ \varphi_{c_1})(u_1(t_0)).$$

Portanto, pela regra da cadeia podemos escrever

$$u'_2(t_0) = (\varphi_{c_2}^{-1} \circ \varphi_{c_1})'(u_1(t_0))u'_1(t_0) = (\varphi_{c_2}^{-1} \circ \varphi_{c_1})'(\varphi_{c_1}^{-1}(p))u'_1(t_0).$$

**Proposição 6** *A representação local do fibrado tangente*

$$\tau(\mathcal{P}_\mu) = \bigsqcup_{p \in \mathcal{P}_\mu} T_p(\mathcal{P}_\mu).$$

é da forma

$$(u_1, v_1) \mapsto \left( \varphi_{c_2}^{-1} \circ \varphi_{c_1}(u_1), v_1 - \frac{\int_T v_1(\varphi)'_+(c_2)d\mu}{\int_T (\varphi)'_+(c_2)d\mu} \right) \in \mathcal{E}_{c_2}^\varphi \times \mathcal{N}_{c_2}^\varphi. \quad (24)$$

**Demonstração:** Dado  $w \in \varphi_{c_1}^{-1}(\mathcal{F}_{c_1}^\varphi \cap \mathcal{F}_{c_2}^\varphi)$ , temos que a derivada da aplicação  $\varphi_{c_2}^{-1} \circ \varphi_{c_1}$  avaliada em  $w$  na direção de  $v \in \mathcal{N}_c^\varphi$  é da forma

$$(\varphi_{c_2}^{-1} \circ \varphi_{c_1})'(w)v = v - \frac{\int_T v(\varphi)'_+(c_2)d\mu}{\int_T (\varphi)'_+(c_2)d\mu}. \quad (25)$$

De fato, pela convexidade de  $\varphi$ , temos que

$$\int_T (c_1 - c_2 + w)(\varphi)'_+(c_2)d\mu \leq \int_T [\varphi(c_1 + w) + \varphi(c_2)]d\mu.$$

Como  $w \in \varphi_{c_1}^{-1}(\mathcal{F}_{c_1}^\varphi \cap \mathcal{F}_{c_2}^\varphi) \subset \mathcal{E}_{c_1}^\varphi$ , temos que  $\varphi(c_1 + w)$  is  $\mu$ -integrável, and consequentemente,  $\int_T (c_1 - c_2 + w)(\varphi)'_+(c_2)d\mu$  is  $\mu$ -integrável. Então, do Teorema da Convergência Dominada segue que (25) ocorre.

O fibrado tangente é então denotado por

$$\tau(\mathcal{P}_\mu) = \{(\varphi_c(u), v); \varphi_c(u) \in \mathcal{F}_c^\varphi \subset \mathcal{P}_\mu \text{ e } v \text{ é um vetor tangente em } \varphi_c(u)\}.$$

Suas cartas são representadas como

$$(v, u) \in \tau(\mathcal{F}_c^\varphi) \mapsto \left( \varphi_{c_2}^{-1}(v), v - \frac{\int_T v(\varphi)'_+(c_2)d\mu}{\int_T (\varphi)'_+(c_2)d\mu} \right),$$

a qual foi definida na coleção de subconjuntos abertos  $\mathcal{F}_{c_1}^\varphi \times \mathcal{E}_{c_1}^\varphi$  de  $\mathcal{P}_\mu \times \mathcal{N}_{c_1}^\varphi$ . Então, como a equação (25) ocorre, a aplicação de transição é dada para  $(u_1, v_1) \in \mathcal{E}_{c_1}^\varphi \times \mathcal{N}_{c_1}^\varphi$  por

$$(u_1, v_1) \mapsto \left( \varphi_{c_2}^{-1} \circ \varphi_{c_1}(u_1), v_1 - \frac{\int_T v_1(\varphi)'_+(c_2) d\mu}{\int_T (\varphi)'_+(c_2) d\mu} \right) \in \mathcal{E}_{c_2}^\varphi \times \mathcal{N}_{c_2}^\varphi. \quad \square \quad (26)$$

### 3 TEORIA DA INFORMAÇÃO EM VARIEDADES ESTATÍSTICAS

A divergência de Bregman foi definida e utilizada pelo matemático Bregman em (BREGMAN, 1967b) para solucionar problemas de otimização convexa (BREGMAN, 1967a). Dentre estes, destacamos a generalização do método de ponto proximal (KIWIEL, 1997) e distribuições de famílias exponenciais (BANERJEE *et al.*, 2005). Essa divergência tem muita relevância também em outras áreas como, por exemplo, análise numérica, computação neural, estimação e inferência (CAYTON, 2008; BANERJEE *et al.*, 2005; LIU, 2011). Em teoria da informação, geralmente é usada como medida de similaridade no contexto de recuperação de imagens (XU *et al.*, 2012). Na construção dessa nova parametrização local de  $\mathcal{P}_\mu$ , surge uma função convexa que será chamada de função normalizadora. A divergência de Bregman associada a normalizadora dá origem a uma nova divergência.

Seguiremos esse roteiro no desenvolvimento deste capítulo. A primeira seção vamos definir a  $\varphi$ -divergência para o caso em que  $\varphi$  é a exponencial deformada e uma outra divergência usando a  $q$ -exponencial (VIEIRA *et al.*, 2019b). Na segunda parte, provamos que as funções  $q$ -exponencial e  $\kappa$ -exponencial podem ser usadas na generalização da divergência de Rényi (VAN ERVEN and HAREMOS, 2014; DE SOUZA, VIGELIS, and CAVALCANTE, 2016; VIEIRA *et al.*, 2019b). E na terceira parte, definiremos generalizações das entropias de Shannon, Tsallis e Rényi e mostramos como elas podem estar relacionadas.

#### 3.1 A $\varphi$ -divergência e $q$ -divergência

Seja  $X$  um espaço de Banach e  $f : S \rightarrow \mathbb{R}$  uma função convexa no subespaço convexo  $S$  contido em  $X$ . A divergência de Bregman é aplicação  $B_f : S \times S \rightarrow [0, \infty)$  dada pela expressão

$$B_f(y, x) = f(x) - f(y) - \partial_+ f(x)(y - x)$$

em que

$$\partial_+ f(x)h = \lim_{t \downarrow 0} \frac{f(x + th) - f(x)}{t}.$$

A divergência de Bregman  $B_\psi : \mathcal{A}_c^\varphi \times \mathcal{A}_c^\varphi \rightarrow [0, \infty)$  associada a função normalizadora  $\psi : \mathcal{A}_c^\varphi \rightarrow [0, \infty)$  é dada por (BREGMAN, 1967b; ZHANG, 2004; KORBEL, HANEL, and THURNER, 2019)

$$B_\psi(v, u) = \psi(v) - \psi(u) - \partial_+ \psi(u)(v - u). \quad (27)$$

A divergência  $\mathcal{D}_\psi : \mathcal{A}_c^\varphi \times \mathcal{A}_c^\varphi \rightarrow [0, \infty)$  relacionada a  $\varphi$ -família  $\mathcal{F}_c^\varphi$  é definida como

$$\mathcal{D}_\psi(u, v) = B_\psi(v, u).$$

Dados  $u, v \in \mathcal{A}_c^\psi$ , temos que

$$\varphi(c + u - \psi(u)u_0), \varphi(c + v - \psi(v)u_0) \in \mathcal{P}_\mu.$$

Portanto,  $c + u - \psi(u)u_0, c + v - \psi(v)u_0 > a_\varphi$  e supondo  $\varphi$  continuamente diferenciável, segue que a divergência  $\mathcal{D}_\psi$  não depende da parametrização de  $\mathcal{F}_c^\psi$  (VIGELIS and CAVALCANTE, 2013a). Isto permite definir a  $\varphi$ -divergência entre as densidades de probabilidade  $p = \varphi_c(u)$  e  $z = \varphi_c(v)$ , para  $u, v \in \mathcal{A}_c^\psi$  como

$$\mathcal{D}(p \parallel z) = \mathcal{D}_\psi(u, v) = \frac{\int_T \frac{\varphi^{-1}(p) - \varphi^{-1}(z)}{(\varphi^{-1})'(p)} d\mu}{\int_T \frac{1}{(\varphi^{-1})'(p)} d\mu}. \quad (28)$$

Note que a divergência está bem definida dentro da mesma  $\varphi$ -família. A condição  $\mathcal{D}(p \parallel z) = \infty$  se  $p$  e  $z$  não estão na mesma  $\varphi$ -família estende a divergência para  $\mathcal{P}_\mu$ . Denotaremos essas divergências por  $\mathcal{D}_\varphi$  e denominamos de  $\varphi$ -divergência (VIGELIS and CAVALCANTE, 2013a).

Dados  $u, v \in \mathcal{A}_c^\psi$ , por definição,  $p = \varphi(c + u - \psi(u)u_0) > 0$  e  $z = \varphi(c + v - \psi(v)u_0) > 0$ . Além disso, como  $\varphi(t, \cdot)$  é estritamente convexa para valores maiores que  $a_\varphi$ , segue que  $\mathcal{D}_\varphi$  é sempre não-negativa e  $\mathcal{D}_\varphi(p \parallel z)$  é igual a zero se e somente se  $p = z$ . No seguinte exemplo, encontramos a  $\varphi$ -divergência para o caso no qual a função exponencial deformada  $\varphi$  é a função  $q$ -exponencial (VIEIRA *et al.*, 2019b).

**Exemplo 8** A inversa da função  $q$ -exponencial  $\varphi(t, u) = \exp_q(t, u) = \exp_{q(t)}(u)$  é a função  $q$ -logaritmo

$$\ln_q(t, u) = \ln_{q(t)}(u) = \ln_q(u) = \frac{u^{1-q} - 1}{1 - q}.$$

Dado  $p, z \in \mathcal{P}_\mu$ , a  $q$ -divergência (divergência de Tsallis) de  $p$  com respeito a  $z$  é dada por

$$\mathcal{D}(p \parallel z) = \frac{\int_T \frac{\ln_q(p) - \ln_q(z)}{\ln'_q(p)} d\mu}{\int_T \frac{u_0}{\ln'_q(p)} d\mu},$$

em que  $\ln_q(p)$  denota  $\ln_{q(t)}(p(t))$ . Como a derivada da função  $q$ -logaritmo é  $\ln'_q(u) = \frac{1}{u^q}$ , temos



que

$$\begin{aligned} \int_T \frac{\ln_q(p) - \ln_q(z)}{\ln'_q(p)} d\mu &= \int_T \frac{p^{1-q} - 1}{1-q} - \frac{z^{1-q} - 1}{1-q} \frac{1}{p^q} d\mu \\ &= \int_T \frac{p^q(p^{1-q} - z^{1-q})}{1-q} d\mu \end{aligned}$$

e

$$\begin{aligned} \int_T \frac{u_0}{\ln'_q(p)} d\mu &= \int_T \frac{u_0}{\frac{1}{p^q}} d\mu \\ &= \int_T u_0 p^q d\mu. \end{aligned}$$

Portanto,

$$\mathcal{D}(p \parallel z) = \frac{\int_T \frac{p^q(p^{1-q} - z^{1-q})}{1-q} d\mu}{\int_T u_0 p^q d\mu}. \quad (29)$$

Definamos agora a métrica

$$g : \Sigma(\mathcal{P}_\mu) \times \Sigma(\mathcal{P}_\mu) \rightarrow F(\mathcal{P}_\mu) \quad g(u, v) = \frac{q \int_T \frac{uv}{z} d\mu}{\int_T u_0 z^q d\mu}, \quad (30)$$

onde  $\Sigma(\mathcal{P}_\mu)$  é o conjunto de campos de vetores  $u : \mathcal{A}_c^\mathcal{P} \rightarrow T_p(\mathcal{A}_c^\mathcal{P})$  e  $F(\mathcal{P}_\mu)$  o conjunto de funções  $C^\infty$  dadas por  $f : \mathcal{A}_c^\mathcal{P} \rightarrow \mathbb{R}$ . Essa aplicação está bem definida, pois  $\left(\frac{\partial}{\partial u}\right)_p \mathcal{D}(p \parallel z)|_{p=z} = 0$  e  $\left(\frac{\partial}{\partial v}\right)_p \mathcal{D}(p \parallel z)|_{p=z} = 0$ .

A partir da métrica, podemos definir também a conexão (derivada covariante)

$$\nabla_w u = \frac{\partial}{\partial w} u - \frac{(1-q)}{z} u w + \frac{u}{A^2} B - w \frac{C}{A^2}, \quad (31)$$

onde  $A = \int_T u_0 z^q d\mu$ ,  $B = \left(\frac{\partial}{\partial w}\right)_p A|_{p=z}$  e  $C = \left(\frac{\partial}{\partial u}\right)_p A|_{p=z}$ . A notação  $\left(\frac{\partial}{\partial w}\right)_p A|_{p=z}$  significa a derivada de  $A$  na direção de  $w$  no ponto  $z$  quando  $p = z$ .

Em (LOAIZA and QUICENO, 2013a), foi definida uma outra divergência chamada de  $q$ -divergência. Em seguida, foi definida a métrica e um par de conexões obtidas a partir dessa divergência. No próximo exemplo vamos apresentá-las e em seguida iremos relacioná-las com as equações (29), (30) e (31).

**Exemplo 9** Seja  $f$  uma função, definida para todo  $t \neq 0$  e  $0 < q < 1$  por

$$f(t) = -t \ln_q \left( \frac{1}{t} \right)$$

e para  $p, z \in \mathcal{P}_\mu$ , a  $q$ -divergência de  $z$  com respeito a  $p$  é dada por

$$I^q(p||z) := \int_T z f \left( \frac{p}{z} \right) d\mu.$$

Através de algumas manipulações, podemos escrever

$$I^q(p||z) = \frac{1}{1-q} \left[ 1 - \int_T p^q z^{1-q} d\mu \right], \quad (32)$$

o qual é o funcional divergência de Tsallis.

O funcional  $q$ -divergência  $I^q(p||z)$  induz uma métrica riemanniana  $g$  e um par de conexões, dadas por

$$g(u, v) = -(d_u)_p (d_v)_z I^q(p||z)|_{p=z} \quad (33)$$

e

$$\langle \nabla_w^{(q)} u, v \rangle = -(d_w)_p (d_u)_p (d_v)_z I^q(p||z)|_{p=z}, \quad (34)$$

em que  $v \in T_z(\mathcal{P}_\mu)$  e  $u \in T_p(\mathcal{P}_\mu)$ .

Sejam  $p, z \in \mathcal{P}_\mu$  e  $u, v$  campos de vetores, o tensor métrica é definido como

$$g : \Sigma(\mathcal{P}_\mu) \times \Sigma(\mathcal{P}_\mu) \rightarrow F(\mathcal{P}_\mu) \quad g(u, v) = q \int_T \frac{uv}{z} d\mu. \quad (35)$$

Dessa forma, a família de derivadas covariantes (conexões) é dado por

$$\nabla_w^{(q)} u : \Sigma(\mathcal{P}_\mu) \times \Sigma(\mathcal{P}_\mu) \rightarrow \Sigma(\mathcal{P}_\mu) \quad \nabla_w^{(q)} u = d_w u - \left( \frac{1-q}{z} \right) uw. \quad (36)$$

A divergência  $\mathcal{D}(p || z)$ , dada pela equação (29) pode ser relacionada com a  $q$ -

divergência pela igualdade

$$\begin{aligned}
I^{(q)}(p \parallel z) &= \int_T z f\left(\frac{p}{z}\right) d\mu = \int_T z \left(\frac{p}{z} \ln_q\left(\frac{z}{p}\right)\right) d\mu \\
&= - \int_T p \left(\frac{\ln_q(z) - \ln_q(p)}{1 + (1-q)\ln_q(p)}\right) d\mu \\
&= - \int_T p \left(\frac{(z^{1-q} - p^{1-q}) / (1-q)}{1 + (1-q)\frac{p^{1-q}-1}{(1-q)}}\right) d\mu \\
&= \int_T p^q \left(\frac{(p^{1-q} - z^{1-q})}{(1-q)}\right) d\mu.
\end{aligned}$$

Portanto,

$$\mathcal{D}(p \parallel z) = \frac{I^{(q)}(p \parallel z)}{\int_T u_0 p^q d\mu}.$$

Observe que, considerando

$$\int_T u_0 p^q d\mu = 1,$$

teremos que a  $\varphi$ -divergência, métrica e conexão definida pelas equações (29), (30), (31) coincidem com a divergência, métrica e conexão definida em (32), (35), (36), respectivamente.

### 3.2 Generalização da divergência de Rényi e $\exp_\kappa$

A divergência de Rényi é uma medida de informação a qual depende de um parâmetro  $\alpha$  (denominado ordem de divergência) e tem várias aplicações, veja por exemplo (CSISZÁR, 1995; HARREMOËS, 2006). No artigo (DE SOUZA, VIGELIS, and CAVALCANTE, 2016) foi desenvolvida uma generalização da Divergência de Rényi a partir de uma exponencial deformada. Essa generalização está relacionada com a  $\varphi$ -divergência e a divergência de Kullback-Leibler. Nesta seção, mostramos que as funções  $q$ -exponencial e  $\kappa$ -exponencial podem ser usadas na generalização da divergência de Rényi.

**Definição 2** *Seja  $\varphi : T \times \mathbb{R} \rightarrow [0, \infty)$  uma função exponencial deformada. Dizemos que  $p$  e  $z$  em  $\mathcal{P}_\mu$  são  $\varphi$ -conectadas por um arco aberto, se existem um intervalo aberto  $I \supset [0, 1]$ , uma*

constante  $\kappa(\alpha) := \kappa(\alpha; p, z)$  e uma função mensurável  $u_0 : T \rightarrow (0, \infty)$  tal que

$$\int_T \varphi((1 - \alpha)\varphi^{-1}(p) + \alpha\varphi^{-1}(z) - \kappa(\alpha)u_0) d\mu = 1, \quad (37)$$

para cada  $\alpha \in I$ .

Através da função normalizadora, podemos definir uma generalização da divergência de Rényi de ordem  $\alpha \in (0, 1)$  como

$$\mathcal{D}_{R,\varphi}^{(\alpha)}(p \parallel z) = \frac{\kappa(\alpha)}{\alpha(1 - \alpha)}, \quad (38)$$

onde  $\kappa(\alpha)$  satisfaz (37).

Essa generalização no caso  $\alpha \in \{0, 1\}$  é definida como um limite

$$\mathcal{D}_{R,\varphi}^{(0)}(p \parallel z) = \lim_{\alpha \rightarrow 0} \mathcal{D}_{R,\varphi}^{(\alpha)}(p \parallel z) \quad (39)$$

e

$$\mathcal{D}_{R,\varphi}^{(1)}(p \parallel z) = \lim_{\alpha \rightarrow 1} \mathcal{D}_{R,\varphi}^{(\alpha)}(p \parallel z). \quad (40)$$

Os limites em (39) e (40), sob algumas condições, tem valores finitos e convergem para a  $\varphi$ -divergência.

$$\mathcal{D}_{R,\varphi}^{(0)}(z \parallel p) = \mathcal{D}_{R,\varphi}^{(1)}(p \parallel z) = \mathcal{D}_\varphi(p \parallel z) < \infty.$$

Na próxima proposição veremos uma condição necessária e suficiente para conectar duas densidades de probabilidade em  $\mathcal{P}_\mu$ .

**Proposição 7 ((VIGELIS, de ANDRADE, and CAVALCANTE, 2017))** *Seja  $\mu$  uma medida não-atômica. Sejam  $\varphi : \mathbb{R} \rightarrow [0, \infty)$  uma função exponencial deformada positiva e  $u_0 : T \rightarrow (0, \infty)$  uma função mensurável. Considere  $\alpha \in (0, 1)$ . Para cada par de distribuições de probabilidade  $p$  e  $z$  em  $\mathcal{P}_\mu$ , existe uma constante  $\kappa(\alpha) := \kappa(\alpha; p, z)$  satisfazendo (37) se, e somente se,*

$$\int_T \varphi(c + \lambda u_0) d\mu < \infty, \text{ para todo } \lambda > 0, \quad (41)$$

para cada função mensurável  $c : T \rightarrow \mathbb{R}$  satisfazendo  $\int_T \varphi(c) d\mu = 1$ .

No Exemplo 6 provamos que a função  $q$ -exponencial satisfaz a condição (41). Desse modo, pela Proposição 7 e equação (38), concluímos que esta função pode ser usada na generalização da divergência de Rényi. Analogamente, a função dada no Exemplo 7 não pode ser usada na generalização da divergência de Rényi.

Na próxima proposição, apresentamos um critério equivalente para uma função exponencial deformada  $\varphi$  satisfazer a condição (41).

**Proposição 8 ((VIGELIS, de ANDRADE, and CAVALCANTE, 2017))** *Seja  $\varphi : \mathbb{R} \rightarrow [0, \infty)$  uma função exponencial deformada. Então, (41) é satisfeita se, e somente se,*

$$\limsup_{u \rightarrow \infty} \frac{\varphi(u)}{\varphi(u - \lambda_0)} < \infty, \text{ para algum } \lambda_0 > 0.$$

No próximo exemplo, mostraremos uma classe de funções exponenciais deformadas que podem ser usadas na generalização da divergência de Rényi (VIEIRA *et al.*, 2019b).

**Exemplo 10** *Mostraremos que a função  $\kappa$ -exponencial  $\exp_\kappa : \mathbb{R} \rightarrow (0, \infty)$  para  $\kappa \in [-1, 1]$  definida como (KANIADAKIS, 2001; PISTONE, 2009)*

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

*satisfaz a condição (41). De fato, sua inversa denominada  $\kappa$ -logaritmo  $\log_\kappa : (0, \infty) \rightarrow \mathbb{R}$  é dada por*

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

*Verificaremos que existem  $\alpha \in (0, 1)$  e  $\lambda > 0$  para o qual*

$$\lambda \leq \log_\kappa(v) - \log_\kappa(\kappa v), \text{ para todo } v > 0. \quad (42)$$

*Após algumas manipulações é possível mostrar que a derivada de  $\log_\kappa(v) - \log_\kappa(\alpha v)$  é negativa para  $0 < v \leq v_0$  e positiva para  $v \geq v_0$ , em que*

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

*Consequentemente, a diferença  $\log_\kappa(v) - \log_\kappa(\alpha v)$  atinge um mínimo em  $v_0$ . Dado  $\alpha \in (0, 1)$ , a desigualdade (42) é satisfeita para algum  $\lambda > 0$ . Inserindo  $v = \exp_\kappa(u)$  sobre (42), podemos escrever*

$$\alpha \exp_\kappa(u) \leq \exp_\kappa(u - \lambda), \text{ para todo } u \in \mathbb{R}. \quad (43)$$

*Se  $n \in \mathbb{N}$  é tal que  $n\lambda \geq 1$ , então uma repetida aplicação de (43) nos fornece*

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

Logo,

$$\limsup_{u \rightarrow \infty} \frac{\varphi(u)}{\varphi(u - \lambda_0)} = \limsup_{u \rightarrow \infty} \frac{\exp_{\kappa}(u)}{\exp_{\kappa}(u - 1)} \leq \limsup_{u \rightarrow \infty} \frac{1}{\alpha^n} < \infty.$$

E, portanto, pela Proposição 8  $\kappa$ -exponencial  $\exp_{\kappa}(\cdot)$  satisfaz a condição (41).

Como consequência do exemplo anterior, Proposição 7 e equação (38), temos que  $\exp_{\kappa}(u)$  pode ser usada na generalização da divergência de Rényi.

### 3.3 Entropia Generalizada

A entropia de uma variável aleatória é definida como o valor esperado da incerteza de seus resultados. Mais precisamente, em teoria da informação a entropia mede a quantidade esperada de informações transmitidas, identificando o resultado de um estudo aleatório. Três entropias são bastante conhecidas e utilizadas nessa área. São elas: a entropia de Shannon, Tsallis e Rényi. Vamos generalizar essas entropias usando uma função exponencial deformada.

#### 3.3.1 $\varphi$ -entropia generalizada

Seja  $\Delta_n$  o conjunto de todas funções de distribuições de probabilidade em  $I_n = \{x \in \mathbb{N}; 1 \leq x \leq n\} = [1, n] \cap \mathbb{N}$ , isto é,

$$\Delta_n = \left\{ p = (p_1, \dots, p_n) \mid \sum_{i=1}^n p_i = 1, p_i \geq 0, \forall i = 1, \dots, n \right\}.$$

A entropia de Shannon (SHANNON, 1948), denotada por  $H^S(p)$ , em que  $p$  é uma distribuição de probabilidade, é definida como sendo

$$H^S(p) = \sum_{i=1}^n p_i \ln \frac{1}{p_i}.$$

Em (TSALLIS, 1988), foi generalizada essa entropia usando a função  $q$ -logaritmo. Dado  $q \in \mathbb{R}$ , a função  $q$ -logaritmo é a inversa da  $q$ -exponencial de um número real  $x > 0$

$$\ln_q x = \begin{cases} \ln x, & \text{se } q = 1, \\ \frac{x^{1-q} - 1}{1 - q}, & \text{caso contrário.} \end{cases}$$

Note que por continuidade temos  $\lim_{q \rightarrow 1} \ln_q x = \ln x$ . Por esse motivo, a entropia de Tsallis definida como sendo

$$H_q^T(p) = \sum_{i=1}^n p_i^q \ln_q p_i$$

generaliza a entropia de Shannon (TSALLIS, 1994). Por fim, Rényi em (RÉNYI, 1961), generalizou a entropia de Shannon:

$$H_{\alpha}^R(p) = \frac{1}{1-\alpha} \ln \sum_{i=1}^n p_i^{\alpha},$$

em que  $\alpha \in (0, \infty) \setminus \{1\}$ .

Vamos definir e discutir uma generalização dessas entropias acima. Essas generalizações dependem de uma função exponencial deformada  $\varphi : \mathbb{R} \rightarrow [0, \infty)$  satisfazendo as condições dadas pelas equações (11), (14), (15).

Lembrando que  $a_{\varphi} := \inf_{x \in \mathbb{R}} \{\varphi(x) > 0\}$ , a restrição

$$\varphi : \{x \in \mathbb{R} \mid x \geq a_{\varphi}\} \rightarrow [0, \infty)$$

é injetiva e, portanto, a inversa  $\varphi^{-1}$  está bem definida. Para  $p_i > 0$ , o quociente

$$\frac{\varphi^{-1}(p_i)}{(\varphi^{-1})'(p_i)}$$

é não positivo.

O interior de  $\Delta_n$  é o conjunto

$$\text{int } \Delta_n = \left\{ (p_1, \dots, p_n) \mid \sum_{i=1}^n p_i = 1, p_i > 0, \forall i = 1, \dots, n \right\}.$$

Caso  $p = (p_1, \dots, p_n) \notin \text{int } \Delta_n$ , então existe pelo menos um  $i \in I_n$  tal que  $p_i = 0$ . Para esta situação, adote a convenção

$$\frac{\varphi^{-1}(p_i)}{(\varphi^{-1})'(p_i)} = 0.$$

Agora que foi definido o quociente  $\frac{\varphi^{-1}(p_i)}{(\varphi^{-1})'(p_i)}$  para  $p_i \geq 0$ , a  $\varphi$ -entropia generalizada na densidade de probabilidade  $p = (p_1, \dots, p_n) \in \Delta_n$  é definida como

$$H_{\varphi}(p) = - \sum_{i=1}^n \frac{\varphi^{-1}(p_i)}{(\varphi^{-1})'(p_i)}. \quad (44)$$

**Exemplo 11** Note que se  $\varphi = \exp$  na equação (44), temos a entropia de Shannon

$$\begin{aligned} H_{\exp}(p_1, p_2, \dots, p_n) &= - \sum_{i=1}^n \frac{\ln(p_i)}{\ln'(p_i)} \\ &= - \sum_{i=1}^n \frac{\ln p_i}{\frac{1}{p_i}} \\ &= - \sum_{i=1}^n p_i \ln p_i \\ &= H^S(p_1, \dots, p_n). \end{aligned}$$

**Exemplo 12** Fazendo agora  $\varphi = \exp_q$  na equação (44), temos a entropia de Tsallis. De fato, como a derivada da  $q$ -logaritmo é  $\ln'_q x = x^{-q}$ , temos que

$$\frac{\ln_q x}{\ln'_q x} = \frac{x^q - x}{q - 1}.$$

Como  $\sum_{i=1}^n p_i = 1$ , a entropia de Tsallis pode ser escrita

$$\begin{aligned} H_{\exp_q}(p) &= - \sum_{i=1}^n \frac{\ln_q p_i}{\ln'_q p_i} \\ &= H_q^T(p). \end{aligned}$$

Vamos definir agora uma generalização da entropia de Rényi. Para isto, consideramos uma classe de funções exponenciais deformadas que satisfazem a igualdade  $\varphi(0) = 1$ . A  $\varphi$ -entropia de Rényi é definida como

$$H_{\alpha, \varphi}^R(p) = \frac{\varphi^{-1} \left( \sum_{i=1}^n \varphi(\alpha \varphi^{-1}(p_i)) \right)}{1 - \alpha}. \quad (45)$$

**Exemplo 13** Tomando  $\varphi = \exp$  na  $\varphi$ -entropia de Rényi, definida acima, obtemos a entropia de Rényi

$$\begin{aligned} H_{\alpha, \exp}^R(p) &= \frac{\ln \left( \sum_{i=1}^n \exp(\alpha \ln(p_i)) \right)}{1 - \alpha} \\ &= H_{\alpha}^R(p). \end{aligned}$$



**Proposição 9** A  $\varphi$ -entropia de Rényi e a  $\varphi$ -entropia generalizada estão relacionadas pelas igualdade

$$\lim_{\alpha \rightarrow 1} H_{\alpha, \varphi}^R(p) = (\varphi^{-1})'(1) \cdot H_{\varphi}(p). \quad (46)$$

**Demonstração:** Definindo as aplicações

$$f(\alpha) = \varphi^{-1} \left( \sum_{i=1}^n \varphi(\alpha \varphi^{-1}(p_i)) \right) \quad \text{e} \quad g(\alpha) = 1 - \alpha$$

temos que

$$\lim_{\alpha \rightarrow 1} f(\alpha) = \varphi^{-1} \left( \sum_{i=1}^n \varphi \left( \lim_{\alpha \rightarrow 1} \alpha \varphi^{-1}(p_i) \right) \right) = \varphi^{-1}(1) = 0$$

e

$$\lim_{\alpha \rightarrow 1} g(\alpha) = 0.$$

Dessa maneira, podemos usar a regra de L'Hôpital

$$\begin{aligned} \lim_{\alpha \rightarrow 1} H_{\alpha, \varphi}^R(p) &= \lim_{\alpha \rightarrow 1} \frac{\varphi^{-1} \left( \sum_{i=1}^n \varphi(\alpha \varphi^{-1}(p_i)) \right)}{1 - \alpha} \\ &= \lim_{\alpha \rightarrow 1} \frac{f(\alpha)}{g(\alpha)} = \lim_{\alpha \rightarrow 1} \frac{f'(\alpha)}{g'(\alpha)}. \end{aligned}$$

Portanto,

$$\begin{aligned} \lim_{\alpha \rightarrow 1} H_{\alpha, \varphi}^R(p) &= \lim_{\alpha \rightarrow 1} \frac{\varphi^{-1} \left( \sum_{i=1}^n \varphi(\alpha \varphi^{-1}(p_i)) \right)}{1 - \alpha} \\ &= \lim_{\alpha \rightarrow 1} \frac{(\varphi^{-1})' \left( \sum_{i=1}^n \varphi(\alpha \varphi^{-1}(p_i)) \right) \sum_{i=1}^n \varphi'(\alpha \varphi^{-1}(p_i)) \varphi^{-1}(p_i)}{-1} \\ &= -(\varphi^{-1})' \left( \sum_{i=1}^n p_i \right) \sum_{i=1}^n \varphi'(\varphi^{-1}(p_i)) \varphi^{-1}(p_i) \\ &= -(\varphi^{-1})'(1) \sum_{i=1}^n \frac{\varphi^{-1}(p_i)}{(\varphi^{-1})'(p_i)}. \end{aligned}$$

**Exemplo 14** Para  $\varphi = \exp$ , a equação (46) possui como limite

$$\lim_{\alpha \rightarrow 1} H_{\alpha}^R(p) = H^S(p).$$

**Proposição 10** Se a fração

$$\frac{\varphi^{-1}(x)}{(\varphi^{-1})'(x)}$$

for côncava para todo  $x \in (0, \infty)$ , então a  $\varphi$ -entropia generalizada  $H_{\varphi}$  é côncava e atinge seu máximo global na distribuição uniforme

$$p = \left( \frac{1}{n}, \dots, \frac{1}{n} \right) \in \text{int } \Delta_n.$$

**Demonstração:** Como a inversa e a soma finita de funções côncavas ainda é uma função côncava, temos que  $\sum_{i=1}^n \frac{\varphi^{-1}(x)}{(\varphi^{-1})'(x)}$  é côncava. Isso prova que  $H_{\varphi}$  também é côncava.

Para o que falta, sabemos da análise convexa que os pontos críticos de uma função côncava são máximos globais. Portanto, a demonstração da proposição acaba se provarmos:

**Afirmção 1** Seja  $p = \left( \frac{1}{n}, \dots, \frac{1}{n} \right) \in \text{int } \Delta_n$ . O funcional  $dH_{\varphi,p}$  é identicamente nulo, isto é, o ponto  $p$  é crítico de  $H_{\varphi}$ .

De fato, primeiramente note que o plano tangente de  $\Delta_n$  pode ser interpretado como o conjunto dos vetores velocidades  $v = (v_1, \dots, v_n) = \lambda'(0)$  do caminho suave

$$\lambda : (-\delta, \delta) \rightarrow \Delta_n, \quad t \mapsto \left( \frac{1}{n} + tv_1, \frac{1}{n} + tv_2, \dots, \frac{1}{n} + tv_n \right).$$

E mais, observe que  $v_1 + v_2 + \dots + v_n = 0$ , pois

$$\sum_{i=1}^n \left( \frac{1}{n} + tv_i \right) = 1 \implies \sum_{i=1}^n v_i = 0.$$

Por fim,

$$\begin{aligned}
 (H_\varphi \circ \lambda)'(0) &= \left. \frac{d}{dt} \right|_{t=0} H_\varphi(p + tv) \\
 &= \sum_{i=1}^n \left. \frac{d}{dx} \right|_{x=1/n} \left( -\frac{\varphi^{-1}(x)}{(\varphi^{-1})'(x)} \right) v_i \\
 &= \left. \frac{d}{dx} \right|_{x=1/n} \left( -\frac{\varphi^{-1}(x)}{(\varphi^{-1})'(x)} \right) \left( \sum_{i=1}^n v_i \right) \\
 &= 0
 \end{aligned}$$

■

**Exemplo 15 ((ASSIRATI, 2014))** *Vamos discutir como aplicar a  $\varphi$ -entropia em imagens. Mais precisamente, vamos considerar a caso particular  $\varphi = \exp$  para analisar imagens e reconhecimentos de padrões através de um histograma de probabilidades de uma imagem.*

*Consideremos uma imagem em níveis de cinza de dimensões  $L_x \times L_y$  onde  $L_x$  e  $L_y$  são valores inteiros pertencentes ao  $[0, 255]$ . O histogramas  $\tilde{p}(x)$  de níveis de cinza de uma imagem são obtidos por meio de contagem do número de pixels com uma determinada intensidade  $\tilde{p}_{ij}$ .*

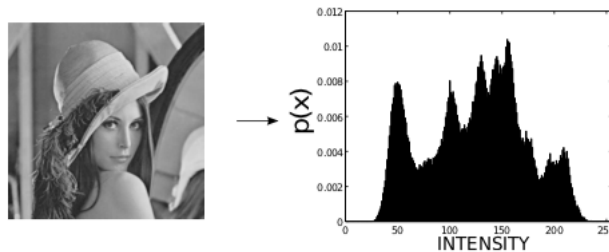


Figura 5: Imagem em nível de cinza e seu histograma (ASSIRATI, 2014)

*Imagens com poucos detalhes produzem histogramas vazios e conseqüentemente geram uma baixa entropia. Já imagens com muitos detalhes produzem histogramas melhores preenchidos e altos valores de entropia.*

### 3.3.2 $\varphi$ -entropia relativa generalizada

Temos como exemplos bem conhecidos de entropias relativas: a entropia relativa de Shannon ou Divergência de Kullback-Leibler, entropia relativa de Tsallis e entropia de Rényi. Nesta seção, vamos usar exponenciais deformadas para generalizar essas entropias.

Vamos considerar a exponencial deformada  $\varphi$  com as mesmas condições que na

seção anterior. A única mudança será na restrição:  $\varphi : \{x \in \mathbb{R} \mid x > a_\varphi\} \rightarrow (0, \infty)$ . Sejam  $p = (p_1, \dots, p_n)$ ,  $q = (q_1, \dots, q_n)$  duas distribuições de probabilidades no interior de  $\Delta_n$  e  $\alpha \in (0, 1)$ . Vamos definir três divergências e relacioná-las entre si.

a)  $\varphi$ -divergência ou  $\varphi$ -entropia relativa

$$D_\varphi(p : q) = \sum_{i=1}^n \frac{\varphi^{-1}(p_i) - \varphi^{-1}(q_i)}{(\varphi^{-1})'(p_i)}.$$

b)  $\varphi$ -divergência de Tsallis

$$D_{\alpha, \varphi}^T(p : q) = \frac{\sum_{i=1}^n \varphi(\varphi^{-1}(q_i) + \alpha(\varphi^{-1}(p_i) - \varphi^{-1}(q_i))) - 1}{\alpha - 1};$$

c)  $\varphi$ -divergência de Rényi (adicionando a condição  $\varphi(0) = 1$ )

$$D_{\alpha, \varphi}^R(p : q) = \frac{\varphi^{-1} \left( \sum_{i=1}^n \varphi(\varphi^{-1}(q_i) + \alpha(\varphi^{-1}(p_i) - \varphi^{-1}(q_i))) \right)}{\alpha - 1}.$$

**Exemplo 16** Fazendo  $\varphi = \exp$  temos que a  $\varphi$ -divergência,  $\varphi$ -divergência de Tsallis e a  $\varphi$ -divergência de Rényi são, respectivamente, a divergência de Kullback-Leibler, de Tsallis e de Rényi. Em outras palavras,

$$D_{\exp}(p : q) = \sum_{i=1}^n \frac{\ln(p_i) - \ln(q_i)}{\ln'(p_i)} = \sum_{i=1}^n p_i \ln \left( \frac{p_i}{q_i} \right) = D^{K,L}(p : q); \quad (47)$$

$$D_{\alpha, \exp}^T(p : q) = \frac{\sum_{i=1}^n p_i^\alpha q_i^{1-\alpha} - 1}{\alpha - 1} = D_\alpha^T(p : q); \quad (48)$$

$$D_{\alpha, \exp}^R(p : q) = \frac{\ln \sum_{i=1}^n p_i^\alpha q_i^{1-\alpha}}{\alpha - 1} = D_\alpha^R(p : q). \quad (49)$$

Podemos provar que

$$\lim_{\alpha \uparrow 1} D_\alpha^R(p : q) = (\varphi^{-1})'(1) D_\varphi(p : q).$$

De fato, aplicando a regra de L'Hôpital ao limite

$$\lim_{\alpha \rightarrow 1} D_{\alpha, \varphi}^R(p : q) = \lim_{\alpha \rightarrow 1} \frac{\varphi^{-1} \left( \sum_{i=1}^n \varphi(\varphi^{-1}(q_i) + \alpha(\varphi^{-1}(p_i) - \varphi^{-1}(q_i))) \right)}{\alpha - 1}$$

temos

$$\begin{aligned} \lim_{\alpha \uparrow 1} D_{\alpha, \varphi}^R(p : q) &= (\varphi^{-1})' \left( \sum_{i=1}^n p_i \right) \sum_{i=1}^n \varphi'(\varphi^{-1}(p_i)) (\varphi^{-1}(p_i) - \varphi^{-1}(q_i)) \\ &= (\varphi^{-1})'(1) \sum_{i=1}^n \frac{\varphi^{-1}(p_i) - \varphi^{-1}(q_i)}{(\varphi^{-1})'(p_i)}. \end{aligned}$$

De modo análogo, podemos provar que

$$\lim_{\alpha \uparrow 1} D_{\alpha}^T(p : q) = D_{\varphi}(p : q).$$

Na observação seguinte provaremos que a  $\varphi$ -entropia relativa generalizada é uma entropia relativa. As demonstrações para a  $\varphi$ -divergência de Tsallis e a  $\varphi$ -divergência de Rényi seguem de forma semelhante.

**Observação 7** *Temos que*

1.  $D^{\varphi}(p||q) > 0$ .

*A concavidade da inversa  $\varphi^{-1}$  fornece a desigualdade*

$$\frac{\varphi^{-1}(p_i) - \varphi^{-1}(q_i)}{(\varphi^{-1})'(p_i)} \geq p_i - q_i \quad \forall i \in [1, n] \cap \mathbb{N}.$$

*Desse modo,*

$$D^{\varphi}(p||q) = \sum_{i=1}^n \frac{\varphi^{-1}(p_i) - \varphi^{-1}(q_i)}{(\varphi^{-1})'(p_i)} \geq \sum_{i=1}^n (p_i - q_i) = \sum_{i=1}^n p_i - \sum_{i=1}^n q_i = 0.$$

2.  $D^{\varphi}(p||q) = 0 \Leftrightarrow p = q$ .

$\Leftrightarrow$  *É claro que se  $p = q$  então  $D^{\varphi}(p||q) = 0$ .*

$\Rightarrow$ ) Suponha agora que  $D^\varphi(p||q) = 0$ . Portanto,

$$0 = \sum_{i=1}^n \frac{\varphi^{-1}(p_i) - \varphi^{-1}(q_i)}{(\varphi^{-1})'(p_i)} = \sum_{i=1}^n (p_i - q_i)$$

$$\implies (p_i - q_i)(\varphi^{-1})'(p_i) = \varphi^{-1}(p_i) - \varphi^{-1}(q_i),$$

para todo  $i \in [1, n] \cap \mathbb{N}$ . Essa igualdade só acontece quando  $p_i = q_i$  para todo  $i \in [1, n] \cap \mathbb{N}$ .

Para cada distribuição de probabilidade  $p = (p_1, \dots, p_n) \in \text{int } \Delta_n$  vamos considerar a  $\varphi$ -família centrada em  $p$ , isto é, para cada  $p = (p_1, \dots, p_n)$ , existe  $c = (c_1, \dots, c_n)$  tal que  $p_i = \varphi(c_i)$  para cada  $i \in [1, n] \cap \mathbb{N}$ .

Suponha que existe uma sequência  $u_{0,c} = (u_{0,c_1}, \dots, u_{0,c_n})$  pertencente ao conjunto

$$n_c^\varphi = \{(u_1, \dots, u_n); |\lambda u_i| \leq c_i - a_\varphi \text{ para algum } \lambda > 0\}$$

satisfazendo a igualdade

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

Observe que para cada  $c = (c_1, \dots, c_n)$  expressado acima, teremos um  $u_{0,c}$  correspondente.

Defina a relação de equivalência

$$\tilde{c} \sim c \iff \eta_c^\varphi = \eta_{\tilde{c}}^\varphi,$$

onde  $\eta_c^\varphi = n_c^\varphi \cap \ell_c^\varphi$  e  $\ell_c^\varphi = \{(u_1, \dots, u_n); \sum_{i=1}^n \varphi(c_i + \lambda u_i) < \infty \text{ para algum } \lambda > 0\}$ . Além disso, denote a  $\varphi$ -família centrada em  $p$ , por  $\mathcal{F}_c^\varphi = \{\varphi_c(u) = \varphi(c + u - \lambda u_{0,[c]})\}$ .

A aplicação  $\varphi_c$ , parametriza o interior de  $\Delta_n$ . Mais especificamente, o domínio está contido no subespaço

$$B_c^\varphi = \left\{ (u_1, \dots, u_n) \in \eta_c^\varphi; \sum_{i=1}^n u_i \varphi'(c_i) = 0 \right\}.$$

A imagem desta aplicação é da forma

$$q = \varphi_c(u_1, \dots, u_n) = (\varphi(c_1 + u_1 - \psi_c(u)u_{0,c_1}), \dots, \varphi(c_n + u_n - \psi_c(u)u_{0,c_n})). \quad (50)$$

A aplicação  $\psi$  é chamada de função normalizadora e é definida de tal forma que ocorre

$$\sum_{i=1}^n \varphi(c_i + u_i - \psi_c(u)u_{0,c_i}) = 1.$$

Seu domínio é, obviamente, o mesmo domínio da parametrização  $\varphi_c$ .

Note que  $\psi(u) \geq 0$ , pois o domínio da parametrização está contido no conjunto  $B_c^\varphi$ . De fato, pela convexidade da exponencial deformada temos que

$$q_i \varphi'(p_i) \leq \varphi(p_i + q_i) - \varphi(p_i).$$

Assim,

$$1 = \sum_{i=1}^n u_i \varphi'(c_i) + \sum_{i=1}^n \varphi(c_i) \leq \sum_{i=1}^n \varphi(c_i + u_i).$$

Para mostrar que  $\psi$  está bem definida, basta utilizar o Teorema do valor intermediário para a aplicação

$$g(\lambda) = \sum_{i=1}^n \varphi(c_i + u_i - \lambda u_{0,c_i}).$$

Usando a função normalizadora, podemos definir a divergência dada pela expressão

$$D(p||q) = \frac{\sum_{i=1}^n \frac{\varphi^{-1}(p_i) - \varphi^{-1}(q_i)}{(\varphi^{-1})'(p_i)}}{\sum_{i=1}^n \frac{u_{0,c_i}}{(\varphi^{-1})'(p_i)}},$$

onde  $p$  e  $q$  são duas densidades de probabilidade na mesma  $\varphi$ -família. E adicionando a condição

$$\sum_{i=1}^n u_{0,c_i} \varphi'(c_i) = 1$$

temos que  $D^\varphi(p||q) = D(p||q)$ .

Outro fato é que podemos provar que a  $\varphi$ -divergência  $D_\varphi(\cdot||\cdot)$  está intimamente ligada a função normalizadora. Mais precisamente,

$$\psi_c(u) = D_\varphi(p||q) = \sum_{i=1}^n \frac{\varphi^{-1}(p_i) - \varphi^{-1}(q_i)}{(\varphi^{-1})'(p_i)}$$

onde  $p_i = \varphi(c_i)$  e  $q_i = \varphi(c_i + u_i - \psi_c(u)u_{0,c_i})$ . Veremos mais detalhes na seguinte proposição.

**Proposição 11** *Se a função normalizadora  $\psi_c(u)$  é não-negativa, então  $\psi_c(u)$  é igual a entropia relativa generalizada.*

**Demonstração:** Temos que

$$\varphi^{-1}(q_i) = c_i + u_i - \psi_c(u)u_{0,c_i}$$

ou seja,

$$\psi_c(u)u_{0,c_i} = \varphi^{-1}(p_i) - \varphi^{-1}(q_i) + u_i$$

Multiplicando por  $\varphi'(c_i)$  e aplicando o somatório no numerador e denominador obtemos

$$\psi_c(u) = \frac{\sum_{i=1}^n [\varphi^{-1}(p_i) - \varphi^{-1}(q_i) + u_i] \varphi'(c_i)}{\sum_{i=1}^n u_{0,c_i} \varphi'(c_i)}.$$

Usando  $\sum_{i=1}^n u_{0,c_i} \varphi'(c_i) = 1$ , podemos escrever

$$\begin{aligned} \psi_c(u) &= \sum_{i=1}^n [\varphi^{-1}(p_i) - \varphi^{-1}(q_i)] \varphi'(c_i) \\ &= \sum_{i=1}^n [\varphi^{-1}(p_i) - \varphi^{-1}(q_i)] \cdot \frac{1}{(\varphi^{-1})'(p_i)} \\ &= D^\varphi(p||q). \quad \blacksquare \end{aligned}$$



## 4 EXPONENCIAL DEFORMADA E O COMPORTAMENTO DA FUNÇÃO NORMALIZADORA

Nesse capítulo, estudaremos o comportamento da função normalizadora próximo ao bordo do domínio da parametrização construída por Vigelis e Cavalcante (VIGELIS and CAVALCANTE, 2013a). No artigo (VIGELIS and CAVALCANTE, 2013b), foi estudado o comportamento da função normalizadora no caso em que a função exponencial deformada satisfaz a condição (8). Posteriormente, em (DE ANDRANDE *et al.*, 2017), foi analisada a situação em que a condição (8) não ocorre para pontos não pertencentes a classe Musielak-Orlicz. Na primeira seção desse capítulo, faremos uma breve revisão de resultados envolvendo a  $\Delta_2$ -condição e revisaremos esses casos já provados anteriormente. Na segunda seção veremos o caso em que a condição (8) não ocorre para pontos pertencentes a classe Musielak-Orlicz. Na terceira seção, provaremos resultados envolvendo a condição (62) e a  $\delta_2$ -condição no caso puramente atômico (VIEIRA *et al.*, 2019a).

### 4.1 A $\Delta_2$ -Condição

Uma função Musielak-Orlicz  $\Phi$  é dita satisfazer a  $\Delta_2$ -condição, ou pertencer a  $\Delta_2$ -classe (denotado por  $\Delta_2$ ), se uma constante  $K > 0$  e uma função não-negativa  $f \in \tilde{L}^\Phi$  podem ser encontradas tais que

$$\Phi(t, 2u) \leq K\Phi(t, u), \text{ para todo } u \geq f(t) \text{ e } \mu\text{-q.t.p. } t \in T. \quad (51)$$

Pode ser verificado que a  $\Delta_2$ -condição dada pela equação (51) é equivalente a existência de uma constante  $\alpha > 0$ , e uma função não-negativa  $f \in \tilde{L}^\Phi$  tal que

$$\alpha\Phi(t, u) \leq \Phi\left(t, \frac{1}{2}u\right), \text{ para todo } u > f(t). \quad (52)$$

É fácil ver que, se uma função Musielak-Orlicz  $\Phi$  satisfaz a  $\Delta_2$ -condição, então

$$I_\Phi(u) < \infty$$

para cada  $u \in L^\Phi$ . De fato, caso contrário teríamos

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

e conseqüentemente,

$$\infty = \alpha \Phi(t, u) > \Phi\left(t, \frac{1}{2}u\right),$$

para todo

$$b_{\Phi}(t) < u < 2b_{\Phi}(t),$$

o qual implica que  $\Phi$  não pode satisfazer a  $\Delta_2$ -condição.

Além disso, uma função Musielak-Orlicz  $\Phi$  satisfaz a  $\Delta_2$ -condição ou pertence a  $\Delta_2$ -classe (denotado por  $\Delta_2$ ) se e somente se 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

$$\alpha_{\lambda} \Phi(t, u) \leq \Phi(t, \lambda u), \text{ para todo } u > f_{\lambda}(t). \quad (53)$$

Observe que se  $\Phi$  satisfaz a  $\Delta_2$ -condição, então  $L^{\Phi}$ ,  $\tilde{L}^{\Phi}$  e  $E^{\Phi}$  são iguais como conjuntos. Por outro lado, se a função Musielak-Orlicz  $\Phi$  não satisfaz a  $\Delta_2$ -condição, então  $E^{\Phi}$  é um subespaço próprio de  $L^{\Phi}$ . Como consequência temos o seguinte lema.

**Lema 4.1** ((VIGELIS and CAVALCANTE, 2014)) *Seja  $\Phi$  uma função Musielak-Orlicz não satisfazendo a  $\Delta_2$ -condição. Assuma que*

$$\Phi(t, b_{\Phi}(t)) = \infty$$

para  $\mu$ -q.t.p.  $t \in T$ . Então, podemos encontrar uma seqüência estritamente crescente  $\{\lambda_n\}$  em  $(0, 1)$ , tal que  $\lambda_n \uparrow 1$ , seqüências  $\{u_n\}$  e  $\{A_n\}$  de funções mensuráveis a valores-finitos e conjuntos mensuráveis dois a dois disjuntos, respectivamente, tais que

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

**Observação 8** *Seja  $\Phi$  uma função Musielak-Orlicz não satisfazendo a  $\Delta_2$ -condição e tal que  $\Phi(t, b_{\Phi}(t)) = \infty$  para  $\mu$ -q.t.p.  $t \in T$ . Então, podemos encontrar funções  $u_*$  e  $u^*$  em  $L^{\Phi}$  tais que*

$$\begin{cases} I_{\Phi}(\lambda u_*) < \infty, & \text{for } 0 \leq \lambda \leq 1, \\ I_{\Phi}(\lambda u_*) = \infty, & \text{for } 1 < \lambda, \end{cases} \quad (54)$$

e

$$\begin{cases} I_{\Phi}(\lambda u^*) < \infty, & \text{for } 0 \leq \lambda < 1, \\ I_{\Phi}(\lambda u^*) = \infty, & \text{for } 1 \leq \lambda. \end{cases} \quad (55)$$

Essas funções foram construídas usando as seqüências  $\{\lambda_n\}$ ,  $\{u_n\}$   $\{A_n\}$  do Lema 4.1. Definamos

$$u_* = \sum_{n=1}^{\infty} \lambda_n u_n \chi_{A_n} \quad e \quad u^* = \sum_{n=1}^{\infty} u_n \chi_{A_n}. \quad (56)$$

Para qualquer  $0 \leq \lambda \leq 1$ , temos

$$\begin{aligned}
 I_{\Phi}(\lambda u_*) &\leq I_{\Phi}(u_*) \\
 &= \sum_{i=1}^{\infty} I_{\Phi}(\lambda u_n \chi_{A_n}) \\
 &\leq \sum_{i=1}^{\infty} 2^{-n} \\
 &= 1.
 \end{aligned}$$

E para qualquer  $\lambda > 1$ , considere um número natural  $n_0 \geq 1$ , tal que  $\lambda \lambda_{n_0} \geq 1$ . Assim, podemos escrever

$$\begin{aligned}
 I_{\Phi}(\lambda u_*) &= \sum_{n=1}^{\infty} I_{\Phi}(\lambda \lambda_n u_n \chi_{A_n}) \\
 &= \sum_{n=1}^{n_0-1} I_{\Phi}(\lambda \lambda_n u_n \chi_{A_n}) + \sum_{n=n_0}^{\infty} I_{\Phi}(\lambda \lambda_n u_n \chi_{A_n}) \\
 &\geq \sum_{n=1}^{n_0-1} I_{\Phi}(\lambda \lambda_n u_n \chi_{A_n}) + \sum_{n=n_0}^{\infty} I_{\Phi}(u_n \chi_{A_n}) \\
 &= \infty.
 \end{aligned}$$

Com relação a  $u^*$ , é claro que  $I_{\Phi}(\lambda u^*) = \infty$  para qualquer  $\lambda \geq 1$ . De fato,

$$\begin{aligned}
 I_{\Phi}(\lambda u^*) &= \sum_{n=1}^{\infty} I_{\Phi}(\lambda u_n \chi_{A_n}) \\
 &\geq \sum_{n=1}^{\infty} I_{\Phi}(u_n \chi_{A_n}) \\
 &= \infty.
 \end{aligned}$$

Se  $\lambda < 1$  e o número natural  $n_0 \geq 1$  é tal que  $\lambda \leq \lambda_{n_0}$ , obtemos

$$\begin{aligned}
I_{\Phi}(\lambda u^*) &= \sum_{n=1}^{\infty} I_{\Phi}(\lambda u_n \chi_{A_n}) \\
&= \sum_{n=1}^{n_0-1} I_{\Phi}(\lambda u_n \chi_{A_n}) + \sum_{n=n_0}^{\infty} I_{\Phi}(\lambda u_n \chi_{A_n}) \\
&\leq \sum_{n=1}^{n_0-1} I_{\Phi}(\lambda u_n \chi_{A_n}) + \sum_{n=n_0}^{\infty} I_{\Phi}(\lambda_n u_n \chi_{A_n}) \\
&\leq \sum_{n=1}^{n_0-1} I_{\Phi}(\lambda u_n \chi_{A_n}) + \sum_{n=n_0}^{\infty} 2^{-n} \\
&< \infty.
\end{aligned}$$

Dessa forma, as funções  $u_*$  e  $u^*$  satisfazem (67).

Em (VIGELIS and CAVALCANTE, 2013b) foi provado que dada uma função exponencial deformada  $\varphi$ , podemos encontrar um  $c : T \rightarrow \mathbb{R}$ , com  $\varphi(c) \in \mathcal{P}_{\mu}$ , tal que a função Musielak-Orlicz, definida por

$$\Phi_c(t, u) = \varphi(t, c(t) + u) - \varphi(t, c(t)) \quad (57)$$

não satisfaz a  $\Delta_2$ -condição. Em outras palavras, sempre podemos encontrar uma  $\varphi$ -família modelada sobre um espaço Musielak-Orlicz gerado por uma função Musielak-Orlicz não satisfazendo a  $\Delta_2$ -condição.

Desde que uma função  $u \in L^0$  está em  $\mathcal{K}_c^{\varphi}$  se existe  $\varepsilon > 0$ , tal que

$$\int_T \varphi(c + \lambda u) d\mu < \infty$$

para todo  $\lambda \in (-\varepsilon, 1 + \varepsilon)$  e

$$\mathcal{B}_c^{\varphi} := \mathcal{K}_c^{\varphi} \cap B_c^{\varphi}$$

é aberto em  $B_c^{\varphi}$ , dizemos que uma função  $u \in B_c^{\varphi}$  pertence ao bordo de  $\mathcal{B}_c^{\varphi}$  se e somente se

$$\int_T \varphi(c + \lambda u) d\mu < \infty$$

para todo  $\lambda \in (0, 1)$  e

$$\int_T \varphi(c + \lambda u) d\mu = \infty$$

para cada  $\lambda > 1$ . A condição  $\Delta_2$  é de extrema importância para o estudo do comportamento da

função normalizadora próximo ao bordo do seu domínio. Mais precisamente, se uma função Musielak-Orlicz  $\Phi_c$  não satisfaz a  $\Delta_2$ -condição, então o bordo de  $\mathcal{B}_c^\Phi$  é não vazio.

(VIGELIS and CAVALCANTE, 2013b) consideraram que a condição (8) ocorre e supondo que a  $\Delta_2$ -condição não ocorre puderam exibir funções pertencentes ao bordo de  $\mathcal{B}_c^\Phi$ , sendo que uma delas está na classe Musielak-Orlicz e a outra está fora da classe Musielak-Orlicz. Para construir essas funções, foram usadas as funções dadas em (56). Em outras palavras, temos as seguintes funções no bordo de  $\mathcal{B}_c^\Phi$

$$w_* = u_* - \frac{u_* \phi'_+(c)}{u_0 \phi'_+(c)} u_0$$

e

$$w^* = u^* - \frac{u^* \phi'_+(c)}{u_0 \phi'_+(c)} u_0,$$

satisfazendo

$$\int_T \varphi(c + w_*) d\mu < \infty$$

e

$$\int_T \varphi(c + w^*) d\mu = \infty.$$

Ainda nesse artigo, foi provado que o comportamento da função normalizadora próximo ao bordo do seu domínio depende de quando a função  $\varphi(c + u)$  é  $\mu$ -integrável ou não. Mais especificamente:

**Proposição 4.1** *Seja  $u$  uma função no bordo de  $\mathcal{B}_c^\Phi$ . Para  $\lambda \in [0, 1)$ , denote  $\psi_u(\lambda) := \psi(\lambda u)$  cuja derivada à direita indicamos por  $(\psi_u)'_+(\lambda)$ . Se*

$$\int_T \varphi(c + u) d\mu < \infty,$$

*então  $\psi_u(\lambda) = \psi(\lambda u)$  converge para algum  $\alpha \in (0, \infty)$ , quando  $\lambda \uparrow 1$ . Por outro lado, se*

$$\int_T \varphi(c + u) d\mu = \infty,$$

*então  $(\psi_u)'_+(\lambda)$  tende ao infinito quando  $\lambda \uparrow 1$ .*

**Demonstração:** Observando que a função normalizadora  $\psi$  é convexa com  $\psi(0) = 0$ , concluímos que  $\psi_u(\lambda) = \psi(\lambda u)$  é não-decrescente e contínua em  $[0, 1)$ . Além disso,  $(\psi_u)'_+(\lambda)$  é não-decrescente em  $[0, 1)$ . Fixe qualquer função  $u$  no bordo de  $\mathcal{B}_c^\Phi$  tal que  $\int_T \varphi(c + u) d\mu < \infty$  e assumamos que  $\psi(\lambda u)$  tende ao  $\infty$  quando  $\lambda \uparrow 1$ . Assim, temos que

$$\varphi(c + \lambda u - \psi(\lambda u)u_0) \leq \varphi(c + u1_{u>0} - \psi(\lambda u)u_0) \rightarrow 0$$

quando  $\lambda \uparrow 1$ .

Desde que  $\varphi(c + \lambda u - \psi(\lambda u)u_0) \leq \varphi(c + u1_{u>0})$ , pelo Teorema da Convergência Dominada obtemos

$$\int_T \varphi(c + \lambda u - \psi(\lambda u)u_0) d\mu \rightarrow 0$$

quando  $\lambda \uparrow 1$ . Isso é uma contradição pois

$$\int_T \varphi(c + \lambda u - \psi(\lambda u)u_0) d\mu = 1.$$

Então,  $\psi(\lambda u)$  é limitado em  $[0, 1)$  e  $\psi(\lambda u)$  converge para algum  $\alpha \in (0, \infty)$  quando  $\lambda \uparrow 1$ .

Agora considere qualquer função  $u$  no bordo de  $\mathcal{B}_c^\varphi$  satisfazendo

$$\int_T \varphi(c + u) d\mu = \infty.$$

Suponha que  $(\psi_u)'_+(\lambda)$  converge para algum  $\alpha \in (0, \infty)$  quando  $\lambda \uparrow 1$ . Do Lema de Fatou obtém-se

$$\int_T \varphi(c + u - \alpha u_0) d\mu \leq \liminf_{\lambda \uparrow 1} \int_T \varphi(c + \lambda u - \psi(\lambda u)u_0) d\mu = 1.$$

Desde que  $\varphi(t, \cdot)$  é convexa, para qualquer  $\lambda \in (0, 1)$ , podemos escrever

$$\begin{aligned} \varphi(c + \lambda u - \psi(\lambda u)u_0) &= \varphi\left(\lambda(c + u - \alpha u_0) + (1 - \lambda)\left(c - \alpha u_0 + \frac{\alpha - \psi(\lambda u)}{1 - \lambda}u_0\right)\right) \\ &\leq \lambda \varphi(c + u - \alpha u_0) + (1 - \lambda) \varphi\left(c - \alpha u_0 + \frac{\alpha - \psi(\lambda u)}{1 - \lambda}u_0\right). \end{aligned}$$

Observando que

$$\begin{aligned} \beta &= \lim_{\lambda \uparrow 1} (\psi_u)'_+(\lambda) \\ &= \lim_{\lambda \uparrow 1} \frac{\alpha - \psi(\lambda u)}{1 - \lambda}, \end{aligned}$$

podemos concluir

$$\varphi(c + \lambda u - \psi(\lambda u)u_0) \leq \varphi(c + u - \alpha u_0) + \varphi(c - \alpha u_0 - \beta u_0)$$

e desse modo  $\varphi(c + \lambda u - \psi(\lambda u)u_0)$  é dominada por uma função integrável. Então, pelo Teo-

rema da Convergência Dominada, segue que

$$\begin{aligned} \int_T \varphi(c+u-\alpha u_0) d\mu &= \int_T \left[ \lim_{\lambda \uparrow 1} \varphi(c+\lambda u-\psi(\lambda u)u_0) \right] d\mu \\ &= \lim_{\lambda \uparrow 1} \int_T \varphi(c+\lambda u-\psi(\lambda u)u_0) d\mu \\ &= 1. \end{aligned}$$

A definição de  $u_0$  nos dá que  $\int_T \varphi(\tilde{c}+\lambda u_0) d\mu < \infty$  para todo  $\lambda \in \mathbb{R}$  e qualquer função mensurável  $\tilde{c}$  tal que  $\int_T \varphi(\tilde{c}) d\mu = 1$ . Em particular, considerando  $\tilde{c} = c+u-\alpha u_0$  e  $\lambda = \alpha$ , temos  $\int_T \varphi(c+u) d\mu < \infty$ . Isso contradiz a suposição que  $\int_T \varphi(c+u) d\mu = \infty$ . Portanto

$$\lim_{\lambda \uparrow 1} (\psi_u)'_+(\lambda) = \infty \quad \square$$

Posteriormente, em (DE ANDRANDE *et al.*, 2017), provamos que no caso em que

$$\int_T \varphi(c+u) d\mu = \infty,$$

temos também que  $\psi(\alpha u) \rightarrow \infty$  quando  $\alpha \uparrow 1$ .

No próximo Lema, temos uma condição de inclusão para as classes Musielak–Orlicz.

**Lema 4.2 ((MUSIELAK, 1983), Teorema 8.4)** *Sejam  $\Psi$  e  $\Phi$  funções Musielak–Orlicz valores-finitos. Então, a inclusão  $\tilde{L}^\Phi \subset \tilde{L}^\Psi$  é satisfeita se e somente se existem  $\alpha > 0$  e uma função não-negativa  $f \in \tilde{L}^\Phi$  tais que*

$$\alpha \Psi(t, u) \leq \Phi(t, u)$$

para todo  $u > f(t)$ .

Como consequência do lema anterior, temos a seguinte proposição que estabelece uma equivalência para a condição (8) ocorrer.

**Proposição 12** (DE ANDRANDE *et al.*, 2017, Proposition 2) *Dizemos que uma função exponencial deformada  $\varphi$  e uma função mensurável  $u_0 : T \rightarrow (0, \infty)$  satisfazem a condição (8) se, e somente se, para alguma função mensurável  $c : T \rightarrow \mathbb{R}$  tal que  $\varphi(c)$  é  $\mu$ -integrável, podemos*

encontrar constantes  $\bar{\lambda}, \alpha > 0$  e uma função não-negativa  $f \in \tilde{L}^{\Phi_c}$  tais que

$$\alpha \Phi_c(t, u) \leq \Phi_{c-\bar{\lambda}u_0}(t, u) \text{ para todo } u > f(t), \quad (58)$$

onde  $\Phi_c(t, u) = \varphi(t, c(t) + u(t)) - \varphi(c(t))$  é uma função Musielak-Orlicz.

Para o que segue definiremos a integral

$$I_{\Phi_c}(u(t)) = \int_T \Phi_c(t, |u(t)|) d\mu$$

para qualquer  $u \in L^0$ .

**Lema 4.3** (MUSIELAK, 1983) *Seja  $\mu$  uma medida  $\sigma$ -finita não-atômica. Se  $\{\alpha_n\}$  é uma sequência de números reais, positivos, e  $\{u_n\}$  é uma sequência de funções mensuráveis não-negativas tais que*

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

*Então existem uma sequência crescente  $\{n_i\}$  de números naturais e uma sequência  $\{A_i\}$  de conjuntos mensuráveis tais que*

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

**Lema 6** (DE ANDRANDE et al., 2017, Lemma 3) *Considere as funções mensuráveis  $c : T \rightarrow [0, \infty)$  e  $u_0 : T \rightarrow (0, \infty)$  tais que  $\int_T \varphi(c) d\mu < \infty$  e*

$$\int_T \varphi(c + \lambda u_0) d\mu < \infty,$$

*para todo  $\lambda > 0$ . Suponha que para cada  $\lambda > 0$  não podemos encontrar  $\alpha > 0$  e  $f \in \tilde{L}^{\Phi_c}$  tais que*

$$\alpha \Phi_c(t, u) \leq \Phi_{c-\lambda u_0}(t, u) \text{ para todo } u > f(t). \quad (59)$$

*Então, podemos encontrar uma sequência estritamente decrescente  $0 < \lambda_n \downarrow 0$ , uma sequência  $\{u_n\}$  de funções mensuráveis a valores-finitos e uma sequência  $\{A_n\}$  de conjuntos mensuráveis dois a dois disjuntos tais que*

$$I_{\Phi_c}(u_n \chi_{A_n}) = 1 \text{ and } I_{\Phi_{c-\lambda_n u_0}}(u_n \chi_{A_n}) \leq 2^{-n} \text{ para todo } n \geq 1. \quad (60)$$



**Demonstração:** Seja  $\{\lambda'_m\}$  uma sequência estritamente decrescente tal que  $0 < \lambda'_m$ . Defina 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  $\sup \emptyset = 0$ . Desde que (8) não é satisfeita temos  $I_{\Phi_c}(f_m) = \infty$  para cada  $m \geq 1$ . Para cada número racional  $r > 0$  defina os conjuntos mensuráveis

$$A_{m,r} = \{t \in T; 2^{-m}\Phi_c(t, u) > \Phi_{c-\lambda'_m u_0}(t, u)\}$$

e as funções simples  $u_{m,r} = r\chi_{A_{m,r}}$ . Para  $r = 0$  obtemos  $u_{m,r} = 0$ . Seja  $\{r_i\}$  uma enumeração de números racionais não-negativos com  $r_1 = 0$ . Defina as funções simples não-negativas  $v_{m,k} = \max_{1 \leq i \leq k} u_{m,r_i}$ , para cada  $m, k \geq 1$ . Pela continuidade de  $\Phi_c(t, u)$  e  $\Phi_{c-\lambda'_m u_0}(t, \cdot)$ , segue que  $v_{m,k} \uparrow f_m$ , quando  $k \rightarrow \infty$ . Em virtude do Teorema da Convergência Monótona, 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  $\Phi(t, v_m(t)) < \infty$  e  $2^{-m}\Phi_c(t, v_m(t)) \geq \Phi_{c-\lambda'_m u_0}(t, v_m(t))$ . Pelo Lema 4.3 existe uma sequência crescente  $\{m_n\}$  de índices e uma sequência  $\{A_n\}$  de conjuntos mensuráveis dois a dois disjuntos tais que  $I_{\Phi_c}(v_{m_n}\chi_{A_n}) = 1$ . Tomando  $\lambda_n = \lambda'_{m_n}$ ,  $u_n = v_{m_n}$  e  $A_n$ , obtemos (60).  $\square$

Além disso, considerando que a condição (8) não ocorre e  $u_0 \in E_c^\Phi$ , foi provado que existe  $u \in \partial \mathcal{B}_c^\Phi$  onde  $\psi$  é convergente em valores próximos a  $u$ . Mais precisamente

**Proposição 4.2** *Seja  $c : T \rightarrow \mathbb{R}$  uma função mensurável, tal que  $\varphi(c) \in \mathcal{P}_\mu$ . Assumindo que a condição (8) não ocorre e que  $u_0 : T \rightarrow \mathbb{R}$  é uma função mensurável, tal que*

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

então existe  $u \in \partial \mathcal{B}_c^\Phi$  tal que

$$\int_T \varphi(c + u) d\mu = \infty,$$

mas  $\psi(\alpha u) \rightarrow \beta$ , com  $\beta \in (0, \infty)$ , quando  $\alpha \uparrow 1$ .

Resta analisar o caso em que a condição (8) não ocorre e que a função  $u$  não está na classe Musielak-Orlicz. Esse caso será explicitado na próxima seção.

## 4.2 A condição (8) sobre a exponencial deformada e suas consequências

**Observação 9** Considere a função exponencial deformada  $\varphi$  e  $u$  tal que

$$\int_T \varphi(c+u)d\mu < \infty.$$

Não é difícil perceber que  $\psi(\alpha u) \rightarrow \beta$ , com  $\beta \in (0, \infty)$  quando  $\alpha \uparrow 1$ . De fato, temos

$$\int_T \varphi(c+u-\lambda u_0)d\mu < \infty$$

para todo  $\lambda > 0$ . Suponha que  $\psi(\alpha u) \uparrow \infty$ , quando  $\alpha \uparrow 1$ . Então, para todo  $A > 0$ , existe  $\delta > 0$ , tal que  $0 < 1 - \alpha < \delta \Rightarrow \psi(\alpha u) > A$ . Desde que

$$\int_T \varphi(c+u-\lambda u_0)d\mu < \infty$$

para todo  $\lambda > 0$  existe  $\gamma > \lambda$  tal que

$$\int_T \varphi(c+u-\gamma u_0)d\mu < 1.$$

Em particular considere  $A = \gamma$ . Dessa forma, pelo Teorema da Convergência Dominada obtemos

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

o que é um absurdo. Portanto, obtemos a desigualdade desejada.

A observação acima fornece o seguinte fato: Dada uma função exponencial deformada  $\varphi$ , independente da condição (8) ocorrer, dado qualquer  $u \in \partial \mathcal{B}_c^\varphi$ , tal que

$$\int_T \varphi(c+u)d\mu < \infty,$$

tem-se que  $\psi(\alpha u)$  converge quando  $\alpha \uparrow 1$ .

No próximo resultado provamos a existência de uma relação entre a condição (8) e

a  $\Delta_2$ -condição. Para isso usamos o fato de que se  $\Phi_c \in \Delta_2$  então  $L_c^\varphi = \tilde{L}_c^\varphi$ , e conseqüentemente  $I_{\Phi_c}(u) < \infty$  para todo  $u \in L_c^\varphi$ .

**Proposição 13** *Se a função exponencial deformada  $\varphi$  não satisfaz a condição (8), então  $\Phi_c \notin \Delta_2$ .*

**Demonstração:** Suponha que (8) não ocorre. Pela Proposição 12 e Lema 6, dado  $\lambda > 0$  existe  $n_0 \in \mathbb{N}$  tal que  $\lambda > \lambda_n$  para todo  $n \geq n_0$ . Considere  $u = \sum_{n=n_0}^{\infty} u_n \chi_{A_n}$ . Dessa forma  $u \in L_c^\varphi$ ,  $I_{\Phi_c}(u) = \infty$  e daí segue o resultado  $\square$

A recíproca da Proposição 13 não é válida, pois a função exponencial satisfaz a condição (8) e não satisfaz a  $\Delta_2$ -condição.

A proposição anterior garante que supondo que a condição (8) não ocorre, temos que o  $\partial \mathcal{B}_c^\varphi$  é não-vazio. Desse modo, como consequência do Lema 6, podemos provar que se a função mensurável  $u_0 : T \rightarrow (0, \infty)$  satisfaz

$$\int_T \varphi(c + \lambda u_0) d\mu < \infty,$$

para todo  $\lambda > 0$ , existe  $w \in \partial \mathcal{B}_c^\varphi$  tal que  $w \in \tilde{L}_c^\varphi$  e  $w + \lambda u_0 \notin \tilde{L}_c^\varphi$  para todo  $\lambda > 0$ . Além disso  $\psi$  será convergente em valores próximos a  $w$ .

**Proposição 14** *Suponha que a exponencial deformada  $\varphi$  não satisfaz (8) e  $u_0 \in E_c^\varphi$ . Dessa forma, existe  $w \in \partial \mathcal{B}_c^\varphi$  satisfazendo*

$$\int_T \varphi(c + w) d\mu < \infty \text{ e } \int_T \varphi(c + w + \bar{\lambda} u_0) d\mu = \infty$$

para todo  $\bar{\lambda} > 0$  e  $\psi(\alpha w) \rightarrow \beta$  com  $\beta \in (0, \infty)$  e  $\alpha \uparrow 1$ .

**Demonstração:** Sejam  $\{\lambda_n\}$ ,  $\{u_n\}$  e  $\{A_n\}$  bem definidas como no Lema 6. Dado  $\alpha > 1$ , existe  $n_1 \in \mathbb{N}$ , tal que  $\alpha(u_n - \lambda_n u_0) > u_n$  para todo  $n \geq n_1$ . Dado qualquer  $\bar{\lambda} > 0$ , existe  $n_2 \in \mathbb{N}$  tal que  $\bar{\lambda} > \lambda_n$  for all  $n \geq n_2$ . Tome  $n_0 = \max\{n_1, n_2\}$ . Considere

$$B = \bigcup_{n=n_0}^{\infty} A_n \quad \text{e} \quad u = \sum_{n=n_0}^{\infty} (u_n - \lambda_n u_0) \chi_{A_n}$$

. Deste modo,

$$\begin{aligned}
\int_T \varphi(c+u)d\mu &= \int_T \varphi\left(c + \sum_{n=n_0}^{\infty} (u_n - \lambda_n u_0) \chi_{A_n}\right) d\mu \\
&= \int_{T \setminus B} \varphi(c)d\mu + \sum_{n=n_0}^{\infty} \left\{ \int_{A_n} \varphi(c)d\mu + I_{\Phi_c}(u_n - \lambda_n u_0) \chi_{A_n} \right\} \\
&\leq \int_T \varphi(c)d\mu + \sum_{n=n_0}^{\infty} \left\{ I_{\Phi_{c-\lambda_n u_0}}(u_n \chi_{A_n}) \right\} \\
&< \infty.
\end{aligned} \tag{61}$$

Para  $\alpha \in (0, 1)$  temos

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

e para  $\alpha > 1$  temos

$$\begin{aligned}
\int_T \varphi(c + \alpha u) d\mu &= \int_{T \setminus B} \varphi(c) d\mu + \sum_{n=n_0}^{\infty} \int_{A_n} \varphi(c + \alpha(u_n - \lambda_n u_0)) d\mu \\
&\geq \int_T \varphi(c) d\mu + \sum_{n=n_0}^{\infty} I_{\Phi_c}(u_n \chi_{A_n}) d\mu \\
&= \infty.
\end{aligned}$$

Assim, por (61) e Observação 9 temos  $\psi(\alpha u) \rightarrow \beta$ , com  $\beta \in (0, \infty)$  e  $\alpha \uparrow 1$ . Considere  $\lambda > 0$ , tal que  $w = \sum_{n=n_0}^{\infty} \alpha(u_n - \lambda_n u_0) - \lambda u_0 \chi_{T \setminus B} \in B_c^\varphi$ . 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  $\alpha > 1$ . Portanto,  $w \in \partial B_c^\varphi$  e dado  $\bar{\lambda} > 0$  obtemos

$$\begin{aligned}
\int_T \varphi(c + w + \bar{\lambda} u_0) d\mu &> \int_{T \setminus B} \varphi(c - \lambda u_0) d\mu + \sum_{n=n_0}^{\infty} \int_{A_n} \varphi(c + u_n - \lambda_n u_0 + \bar{\lambda} u_0) d\mu \\
&> \int_{T \setminus B} \varphi(c - \lambda u_0) d\mu + \sum_{n=n_0}^{\infty} \left( I_{\Phi_c}(u_n) - \int_{A_n} \varphi(c) d\mu \right) \\
&= \infty. \quad \square
\end{aligned}$$

### 4.3 Caso puramente atômico

Nesta seção vamos considerar o caso puramente atômico (VIEIRA *et al.*, 2019a). A condição  $\Delta_2$  agora é denotada por  $\delta_2$ -condição. Vamos apresentar novos resultados envolvendo a condição (8) e a  $\delta_2$ -condição. Neste caso, as integrais presentes na definição da condição (8) são substituídas por somas e as funções são substituídas por sequências de funções. Assuma que existe uma sequência  $\{u_{0,i}\} \subset (0, \infty)$  tal que

$$\sum_{i=1}^{\infty} \varphi(c_i + \lambda u_{0,i}) < \infty \text{ para todo } \lambda > 0, \quad (62)$$

para cada sequência  $\{c_i\} \subset \mathbb{R}$  tal que

$$\sum_{i=1}^{\infty} \varphi(c_i) < \infty.$$

Denotaremos, nesse caso, o espaço Musielak-Orlicz e a classe Musielak-Orlicz, respectivamente, por  $\ell^\Phi$  e  $\tilde{\ell}^\Phi$ . No próximo lema encontramos uma condição necessária e suficiente para que possamos ter a inclusão entre duas classes Musielak-Orlicz.

**Lema 7** *Sejam  $\Psi$  e  $\Phi$  funções Musielak-Orlicz valores-finitos. Então, a inclusão  $\tilde{\ell}^\Phi \subseteq \tilde{\ell}^\Psi$  ocorre se, e somente se, existem  $\varepsilon$ ,  $\alpha > 0$  e uma sequência de números reais não-negativos  $f = \{f_i\} \in \tilde{\ell}^\Phi$  tais que*

$$\alpha \Psi_i(u) \leq \Phi_i(u), \text{ para todo } u > f_i \text{ com } \Phi_i(u) < \varepsilon. \quad (63)$$

A prova do Lema 7 é análoga a prova fornecida em (MUSIELAK, 1983, Theorem 8.4).

Na próxima proposição, encontramos uma condição equivalente a uma sequência de funções  $u_0 = \{u_{0,i}\}$  e uma função exponencial deformada  $\varphi$  satisfazerem a condição (62).

**Proposição 15** *Uma sequência  $u_0 = \{u_{0,i}\}$  e uma função exponencial deformada  $\varphi$  satisfazem a condição (62) se, e somente se, para alguma sequência  $c = \{c_i\}$  de funções mensuráveis tais que  $\sum_{i=1}^{\infty} \varphi(c_i) = 1$ , podemos encontrar constantes  $\varepsilon$ ,  $\lambda$ ,  $\alpha > 0$  e uma sequência  $f = \{f_i\} \in \tilde{\ell}^{\Phi_c}$  de funções reais não-negativas tais que*

$$\alpha \Phi_{c,i}(u) \leq \Phi_{c-\lambda u_{0,i}}(u), \text{ para todo } u > f_i \text{ com } \Phi_{c-\lambda u_{0,i}}(u) < \varepsilon. \quad (64)$$

**Demonstração:** Seja  $u_0 = \{u_{0,i}\}$  uma sequência de funções satisfazendo a condição (62). Pelo

Lema 7, é suficiente provarmos que  $\tilde{\ell}^{\Phi_{c-\bar{\lambda}u_0}} \subseteq \tilde{\ell}^{\Phi_c}$ . De fato, seja  $u = \{u_i\} \in \tilde{\ell}^{\Phi_{c-\bar{\lambda}u_0}}$ . Então,

$$\sum_{i=1}^{\infty} \varphi(c_i + u_i - \bar{\lambda}u_{0,i}) < \infty.$$

Como vale a condição (62), obtemos

$$\begin{aligned} \sum_{i=1}^{\infty} \varphi(c_i + u_i) &< \sum_{i=1}^{\infty} \varphi((c_i + u_i - \bar{\lambda}u_{0,i}) + \bar{\lambda}u_{0,i}) \\ &< \infty \end{aligned}$$

e como consequência temos  $u = \{u_i\} \in \tilde{\ell}^{\Phi_c}$ .

Reciprocamente, suponha que a desigualdade (64) ocorre. Pelo Lema 7, temos que  $\tilde{\ell}^{\Phi_{c-\bar{\lambda}u_0}} \subseteq \tilde{\ell}^{\Phi_c}$ . Dessa forma, dado  $u \in \tilde{\ell}^{\Phi_c}$ , tem-se que  $u + \bar{\lambda}u_0 \in \tilde{\ell}^{\Phi_{c-\bar{\lambda}u_0}} \subseteq \tilde{\ell}^{\Phi_c}$ . E desse modo  $\sum_{i=1}^{\infty} \varphi(c_i + u_i + \bar{\lambda}u_{0,i}) < \infty$  para cada  $\bar{\lambda} > 0$ .

Seja  $\tilde{c} = \{\tilde{c}_i\}$  uma sequência de funções mensuráveis satisfazendo  $\sum_{i=1}^{\infty} \varphi(\tilde{c}_i) < \infty$ . Considere o conjunto  $A_i = \{\tilde{c}_i > c_i; \text{ para cada } i \in \mathbb{N}\}$ . Como consequência para cada  $\lambda > 0$  temos

$$\begin{aligned} \sum_{i=1}^{\infty} \varphi(\tilde{c}_i + \lambda u_{0,i}) &= \sum_{i=1}^{\infty} \varphi(c_i + (\tilde{c}_i - c_i) + \lambda u_{0,i}) \\ &< \sum_{i=1}^{\infty} \varphi(c_i + (\tilde{c}_i - c_i)\chi_{A_i} + \lambda u_{0,i}) \\ &< \infty, \end{aligned}$$

e portanto  $u_0 = \{u_{0,i}\}$  satisfaz a condição (62).

Os autores em (VIGELIS and CAVALCANTE, 2014), demonstraram que se uma função Musielak-Orlicz  $\Phi$  não satisfaz a  $\Delta_2$ -condição, então podemos encontrar funções  $u^*$  e  $u_*$  admitindo as propriedades em (67). Através dessas funções, no artigo (VIGELIS and CAVALCANTE, 2013b), os autores encontraram elementos pertencentes ao  $\partial \mathcal{B}_c^\varphi$ .

A sequência  $\Phi = \{\Phi_i\}$  não satisfaz a  $\delta_2$ -condição se, e somente se, para cada  $\lambda \in (0, 1)$ , existem constantes  $\varepsilon > 0$ ,  $\alpha \in (0, 1)$ , e uma sequência não-negativa  $f = \{f_i\}$  com  $I_\Phi(f) < \infty$  tal que

$$\alpha \Phi_i(u) \leq \Phi_i(\lambda u), \text{ para todo } u > f_i \text{ com } \Phi_i(u) < \varepsilon. \quad (65)$$

**Lema 8** *Seja  $\Phi = \{\Phi_i\}$  uma sequência de funções Musielak-Orlicz a valores-finitos a qual não satisfaz a  $\delta_2$ -condição. Então podemos encontrar uma sequência estritamente crescente  $\{\lambda_n\}$  em  $(0, 1)$  convergindo para cima para 1, e sequências  $\{u_n\}$  e  $\{A_n\}$  de números reais valores-*

finitos, e conjuntos dois a dois disjuntos em  $\mathbb{N}$ , respectivamente, tais que

$$1 - 2^{-n} \leq I_{\Phi}(u_n \chi_{A_n}) \leq 1 \text{ e } I_{\Phi}(\lambda_n u_n \chi_{A_n}) \leq 2^{-n}, \quad (66)$$

para todo  $n \geq 1$ .

**Demonstração:** Suponha que a função Musielak-Orlicz  $\Phi$  não satisfaz a  $\delta_2$ -condição. Seja  $\{\lambda_n\}$  uma sequência estritamente crescente em  $(0, 1)$  tal que  $\lambda_n \uparrow 1$ . Para cada  $n \geq 1$ , definimos a sequência não-negativa  $u_n = \{u_{n,i}\}$  por

$$u_{n,i} = \sup\{u > 0; 2^{-n} \Phi_i(u) \leq \Phi_i(\lambda_n u) \text{ e } \Phi_i(u) < 2^{-n}\},$$

onde adotamos a convenção  $\sup \emptyset = 0$ . Se (65) não for satisfeita, então temos que  $I_{\Phi}(u_n) = \infty$  para cada  $n \geq 1$ . Como  $\Phi_i(u_{n,i}) \leq 2^{-n}$ , podemos encontrar uma sequência crescente  $\{k_n\} \subset \mathbb{N}$  tal que

$$1 - 2^{-n} \leq \sum_{i=k_{n-1}}^{k_n} \Phi_i(u_{n,i}) \leq 1.$$

A segunda desigualdade acima em conjunção com

$$2^{-n} \Phi_i(u_{n,i}) \geq \Phi_i(\lambda_n u_{n,i})$$

implica em

$$\sum_{i=k_{n-1}}^{k_n} \Phi_i(\lambda_n u_{n,i}) \leq 2^{-n}.$$

Assim, a expressão (66) segue com  $A_n = [k_{n-1}, k_n - 1] \cap \mathbb{N}$ .  $\square$

Semelhante ao que foi feito em (VIGELIS and CAVALCANTE, 2014, Remark 3.12), usando o Lema 8 seja  $\Phi = \{\Phi_i\}$  uma sequência de funções Musielak-Orlicz valores-finitos não satisfazendo a  $\delta_2$ -condição. Desse modo, podemos encontrar funções

$$u_* = \sum_{n=1}^{\infty} \lambda_n u_n \chi_{A_n}$$

e

$$u^* = \sum_{n=1}^{\infty} u_n \chi_{A_n}$$

em  $\ell^{\Phi}$  tais que

$$\begin{cases} I_{\Phi}(\lambda u_*) < \infty, & \text{para } 0 \leq \lambda \leq 1, \\ I_{\Phi}(\lambda u_*) = \infty, & \text{para } 1 < \lambda, \end{cases} \quad (67)$$

e

$$\begin{cases} I_{\Phi}(\lambda u^*) < \infty, & \text{para } 0 \leq \lambda < 1, \\ I_{\Phi}(\lambda u^*) = \infty, & \text{para } 1 \leq \lambda. \end{cases} \quad (68)$$



---

## 5 CONCLUSÃO E PERSPECTIVAS FUTURAS

Nesse trabalho, construímos uma parametrização da variedade estatística de Banach usando uma função exponencial deformada. Temos encontrado o espaço tangente de  $\mathcal{P}_\mu$  e também construímos seu fibrado tangente. Definimos uma  $\varphi$ -divergência, métrica, conexões e provamos que estão relacionadas com as definidas em LOAIZA and QUICENO (2013b) e LOAIZA and QUICENO (2013a). Outra importante contribuição é que as funções  $\kappa$ -exponencial e  $q$ -exponencial podem ser usadas para generalizar a divergência de Rényi.

Usando a exponencial deformada satisfazendo a condição (8), definimos uma  $\varphi$ -entropia e uma generalização da entropia de Rényi e mostramos que estão relacionadas. Analogamente, foi possível fazer o mesmo no caso das entropias relativas. Através desses resultados, vimos relações entre as já conhecidas entropias e divergências de Rényi, Shannon e Tsallis.

No caso não-atômico, independente da condição (8) ocorrer, vimos que a função normalizadora converge para um valor-finito próximo ao bordo do seu domínio no caso em que as funções  $u$  pertencem a classe Musielak–Orlicz. Provamos que se a condição (8) não ocorre, então a função Musielak–Orlicz não satisfaz a  $\Delta_2$ -condição. Além disso, no caso puramente atômico encontramos uma equivalência para a ocorrência da condição (62) e dada uma função Musielak–Orlicz não satisfazendo a  $\delta_2$ -condição, encontramos funções no espaço Musielak–Orlicz satisfazendo as equações em (67).

As entropias encontradas na tese podem ser usadas para fazer análise comparativa para o reconhecimento entre padrões de imagens. Mais precisamente, quais entropias extraem mais informações numa imagem.

## REFERÊNCIAS

- ALAOGLU, L. Weak topologies of normed linear spaces. **Annals of Mathematics**, p. 252–267, 1940.
- 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.; 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.
- ASSIRATI, L. **Entropia aplicada ao reconhecimento de padrões em imagens**. Universidade de São Paulo, 2014.
- BANERJEE, A.; MERUGU, S.; DHILLON, I. S.; GHOSH, J. Clustering with Bregman divergences. **Journal of machine learning research**, v. 6, n. Oct, p. 1705–1749, 2005.
- BENEDITO-SILVA, A. A. ; PIQUEIRA, J. R. C. **Entropia informacional e cronobiológica**. 1998.
- BREGMAN, L. M. The relaxation method of finding the common point of convex sets and its application to the solution of problems in convex programming. **USSR computational mathematics and mathematical physics**, v. 7, n. 3, p. 200–217, 1967a.
- BREGMAN, L. M. The relaxation method of finding the common point of convex sets and its application to the solution of problems in convex programming. **USSR computational mathematics and mathematical physics**, v. 7, n. 3, p. 200–217, 1967b.
- BREZIS, H. **Functional analysis, Sobolev spaces and partial differential equations**. Universitext. Springer, New York, 2011, xiv+599 p.
- BURBEA, J.; RAO, C. R. Entropy differential metric, distance and divergence measures in probability spaces: A unified approach. *Journal of Multivariate Analysis*, v. 12, n. 4, p. 575–596, 1982.
- CARRILLO, P.; KUMAR V.; CASTELLANOS J. A., H.; DAMES. Autonomous robotic exploration using occupancy grid maps and graph slam based on shannon and rényi entropy. *2015 IEEE international conference on robotics and automation (ICRA)*. IEEE, 2015, p. 487–494.

CASTRO, A. F. **Modelagem e desenvolvimento de um banco de dados geográficos: aplicação à elaboração de mapas de sensibilidade ambiental ao derramamento de óleo na área costeira entre Galinhos e São Bento do Norte-RN.** Universidade Federal do Rio Grande do Norte, 2002.

CAYTON, L. Fast nearest neighbor retrieval for bregman divergences. **Proceedings of the 25th international conference on Machine learning.** ACM, 2008, p. 112–119.

CHEN, S. *Geometry of Orlicz spaces.* 1996.

CHEN, Y.; WANG, J. Z.; KROVETZ, R. CLUE: cluster-based retrieval of images by unsupervised learning. **IEEE transactions on Image Processing**, v. 14, n. 8, p. 1187–1201, 2005.

COSTA, S. I.; SANTOS, S. A.; STRAPASSON, J. E. Fisher information distance: a geometrical reading. **Discrete Applied Mathematics**, v. 197, p. 59–69, 2015.

CSISZÁR, I. Generalized cutoff rates and Rényi's information measures. **IEEE Transactions on information theory**, v. 41, n. 1, p. 26–34, 1995.

DE ANDRANDE, L. H. F.; VIGELIS, R. F.; VIEIRA, F. L. J.; CAVALCANTE, C. C. *Normalization and  $\phi$ -function: Definition and Consequences*, Springer International Publishing, p. 231–238. 2017.

DE 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.

DIONÍSIO, R. MENDES D., A. MENEZES. *O princípio da entropia máxima.* 2006.

DO CARMO, M. P. **Geometria riemanniana.** Instituto de Matemática Pura e Aplicada, 1988.

FISHER, R. A. On the mathematical foundations of theoretical statistics. **Philosophical Transactions of the Royal Society of London. Series A, Containing Papers of a Mathematical or Physical Character**, v. 222, n. 594-604, p. 309–368, 1922.

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 University Press, 2015.

HARREMOËS, P. Interpretations of Rényi entropies and divergences. **Physica A: Statistical Mechanics and its Applications**, v. 365, n. 1, p. 57–62, 2006.

HUANG, R.; SHIFFMAN C. A., M.; AARON. Maximum entropy method for magnetoencephalography. **IEEE transactions on biomedical engineering**, v. 44, n. 1, p. 98–102, 1997.

KANIADAKIS, G. Non-linear kinetics underlying generalized statistics. **Physica A: Statistical mechanics and its applications**, v. 296, n. 3-4, p. 405–425, 2001.

KANIADAKIS, G. Statistical origin of quantum mechanics. **Physica A: Statistical Mechanics and its Applications**, v. 307, n. 1-2, p. 172–184, 2002.

KIWIEL, K. C. Proximal minimization methods with generalized Bregman functions. **SIAM journal on control and optimization**, v. 35, n. 4, p. 1142–1168, 1997.

KORBEL, J.; HANEL, R.; THURNER, S. Information geometric duality of  $\varphi$ -deformed exponential families. **Entropy**, v. 21, n. 2, p. 112, 2019.

KULLBACK, S.; LEIBLER, R.. A. On information and sufficiency. **The annals of mathematical statistics**, v. 22, n. 1, p. 79–86, 1951.

LANG, S. **Introduction to differentiable manifolds**. Springer Science & Business Media, 2002.

LAURITZEN, S. L. **Statistical Manifolds: Paper Presented at the NATO Advanced Workshop on "Differential Geometry in Statistical Inference"**. London, April 8-11, 1984. Institute of Electronic Systems, Aalborg University Center, 1984.

LEE, J. M. **Introduction to smooth manifolds**. Springer, 2001.

LESSA, V. S. Classificação de imagens de ultrassom de câncer de mama baseada em informações híbridas utilizando teoria da informação. 2010.

LIU, M. **Total Bregman divergence, a robust divergence measure, and its applications**. University of Florida, 2011.

LOAIZA, G; QUICENO, H. R. A Riemannian geometry in the q-Exponential Banach manifold induced by q-Divergences. **International Conference on Geometric Science of Information**. Springer, 2013a, p. 737–742.

LOAIZA, G.; QUICENO, H.R. A q-exponential statistical Banach manifold. **Journal of**

**Mathematical Analysis and Applications**, v. 398, n. 2, p. 466–476, 2013b.

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

NASCIMENTO, A. DC; CINTRA, R. J.; FRERY, A. C. Hypothesis testing in speckled data with stochastic distances. **IEEE Transactions on Geoscience and Remote Sensing**, v. 48, n. 1, p. 373–385, 2009.

NASCIMENTO, W. S.; PRUDENTE, F. V. Sobre um estudo da entropia de Shannon no contexto da mecânica quântica: uma aplicação ao oscilador harmônico livre e confinado. **Química Nova**, v. 39, n. 6, p. 757–764, 2016.

NAUDTS, J. Estimators, escort probabilities, and phi-exponential families in statistical physics. **J. Ineq. Pure Appl. Math**, v. 5, n. 4, p. 102, 2004.

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

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, C. R. Information and the accuracy attainable in the estimation of statistical parameters. *Reson. J. Sci. Educ*, v. 20, p. 78–90, 1945.

RAO, C. R. Information and the accuracy attainable in the estimation of statistical parameters. **Breakthroughs in statistics**, Springer, p. 235–247. 1992.

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.

SAIMINEN, K., A.; KAARNIRANTA. Genetics vs. entropy: longevity factors suppress the NF- $\kappa$ B-driven entropic aging process. **Ageing research reviews**, v. 9, n. 3, p. 298–314, 2010.

SHANNON, C. E. A mathematical theory of communication. **Bell system technical journal**, v. 27, n. 3, p. 379–423, 1948.

SILVA, H. T. A.; MARQUES, P. M. A. Validação clínica do uso da entropia de Tsallis no

co-registro de neuroimagens. 2014.

SIMON, F. L. Algoritmo de processamento de imagem para ordenha robótica. 2019.

SPIVAK, M. D. **A comprehensive introduction to differential geometry**. Publish or perish, 1970.

TSALLIS, C. Possible generalization of Boltzmann-Gibbs statistics. **Journal of statistical physics**, v. 52, n. 1-2, p. 479–487, 1988.

TSALLIS, C. What are the numbers that experiments provide. **Quimica Nova**, v. 17, n. 6, p. 468–471, 1994.

ULLAH, A. Entropy, divergence and distance measures with econometric applications. **Journal of Statistical Planning and Inference**, v. 49, n. 1, p. 137–162, 1996.

VAN ERVEN, T.; HAREMOS, P. Rényi divergence and Kullback-Leibler divergence. **IEEE Transactions on Information Theory**, v. 60, n. 7, p. 3797–3820, 2014.

VIEIRA, J. L. F.; de ANDRADE, F. H. L.; VIGELIS, R. F.; CAVALCANTE, C. C. Deformed Exponential and the Behavior of the Normalizing Function. **International Conference on Geometric Science of Information**. Springer, 2019a, p. 271–278.

VIEIRA, J. L. F.; de ANDRADE, F. H. L.; VIGELIS, R. F.; CAVALCANTE, C. C. A Deformed Exponential Statistical Manifold. **Entropy**, v. 21, n. 5, p. 496, 2019b.

VIGELIS, R. F.; CAVALCANTE, C. C. **Journal of Theoretical Probability**, v. 26, n. 3, p. 870–884, 2013a.

VIGELIS, R. F.; CAVALCANTE, C. C. The  $\Delta$  2-condition and  $\varphi$ -families of probability distributions. **International Conference on Geometric Science of Information**. Springer, 2013b, p. 729–736.

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

VIGELIS, R. F.; de ANDRADE, L. H.F.; CAVALCANTE, C. C. On the existence of paths connecting probability distributions. **International Conference on Geometric Science of Information**. Springer, 2017, p. 801–808.

XU, B.; BU, J.; CHEN, C.; CAI, D. A Bregman divergence optimization framework for ranking on data manifold and its new extensions. **Twenty-Sixth AAAI Conference on**

**Artificial Intelligence.** 2012.

ZHANG, J. Divergence function, duality, and convex analysis. **Neural Computation**, v. 16, n. 1, p. 159–195, 2004.