



Universidade de Brasília
Instituto de Ciências Exatas
Departamento de Matemática

**CONVERGÊNCIA EM DISTÂNCIA MALLOWS
PONDERADA COM APLICAÇÕES EM SOMAS
PARCIAIS E PROCESSOS EMPÍRICOS**

por

Wembesom Mendes Soares

Brasília, Julho de 2015.

Para minha Mônica.

A beleza não está nem na luz da manhã nem na sombra da noite, está no crepúsculo, nesse meio tom, nessa incerteza.

Lygia Fagundes Telles

Agradecimentos

A história de toda conquista é, antes de mais nada, um testemunho de esperança. As páginas de um trabalho escrito nunca contarão isso com a devida clareza, porque escrevo na tese os frutos da pesquisa científica, e mais nada. Nas próximas linhas, espero dar os devidos créditos para quem fez alguma diferença nessa trajetória, para quem traduziu em gestos a esperança.

Minha família respeitou e acolheu minha escolha pelo doutorado, sem que fosse preciso o completo entendimento do significado dessa decisão. Esse voto cego de confiança só poderia ser dado por pessoas que acompanharam e causaram minha trajetória até a vida adulta, até a saída da casa materna, e até aqui sempre se fazem presentes e disponíveis. Agradeço à minha mãe pelas doses cavалares de otimismo, força e determinação, que me ensinaram a ser firme e disciplinado; também agradeço a ela por ter me ensinado a ser independente com tamanha eficiência que, anos mais tarde, minha orientadora de mestrado citaria isso como uma característica positiva para um provável curso de doutorado.

Falando de minha orientadora de mestrado, agradeço a professora Cátia por ter me aceitado no mundo da pós-graduação sem muito conhecimento de causa e por conseguir, a duras penas, extrair uma dissertação interessante de um aluno que tinha mais disposição do que capacidade técnica. Graças a ela, aprendi o valor que exemplos simples e motivações claras possuem para exposições intelectuais de qualquer nível e para qualquer público: elas dão relevância ao que estamos fazendo e criam sementes incrivelmente férteis no coração de quem nos ouve ou nos lê. Espero que ela continue me chamando atenção todas as vezes que eu esquecer lições básicas como essa, seja como docente, seja como pesquisador, assim como ela fez no retorno sobre a leitura dessa tese.

Pela leitura cuidadosa da tese, quero agradecer à banca examinadora de meu trabalho. Dispor de tempo e paciência para entender e contribuir a um texto, em menos de um mês, é uma atitude de grande generosidade. Para além disso, agradeço, em particular, o reiterado interesse e apoio da professora Silvia, que me trouxe conforto e ânimo mesmo antes da conclusão desse trabalho, e que não deixou uma só página do trabalho sem a marca de sua revisão rigorosa e benevolente.

A propósito de Rigor e Benevolência, essas são palavras que, com o tempo, aprendi como sinônimos da professora Chang, orientadora deste doutorado. A precisão vigilante mesmo nos conceitos básicos, a simplicidade da exposição e a originalidade das ideias são características tão inerentes a ela quanto o processo de respiração. Não existe termo melhor que Benevolência para qualificar o ato de aceitar orientar a pesquisa de um aluno sem a devida intimidade com o mundo da Probabilidade e com um ritmo de aprendizado destacadamente lento. Minha orientadora foi paciente e incansável na difícil tarefa de me conduzir nesses anos, desde o aprendizado trivial dos assuntos preliminares até a tomada da pesquisa como algo que eu chamasse de meu, para criticar e defender, como sempre devemos fazer com aquilo que assumimos. A única parte triste de encerrar essa fase da vida é não ter mais a exclusividade de ser um aluno de doutorado dessa grande mulher. Agradeço imensamente por esse privilégio.

É justo e de bom tom que fique aqui registrada a minha gratidão aos meus amigos e amigas que me suportaram durante esses anos. Imagino que não tenha sido a melhor tarefa do mundo aguentar as lamúrias todas em vários momentos de fraqueza e pessimismo. Não é uma decisão muito prudente começar a citar nomes, pois alguém sempre acaba ficando de fora, mas seria um pouco injusto não tentar. Para mencionar algumas pessoas do mundo da matemática, um grande obrigado para Kellém, Mariana, Simone, Luciene, Tarcísio, Thaynara e Edimilson, que se envolveram em diferentes formas, níveis e momentos, mas sempre me passando a segurança que, por vezes, eu não tinha. Para mencionar algumas pessoas de fora da matemática que, ao contrário dos outros, não tinham nem o consolo de ouvir lamentações de sua própria área, deixo um grande obrigado para Beth, Laryssa, Suiane, Myller e Érica. A este último casal, Myller e Érica, agradeço pelos longos colóquios de motivação e pelo esforço logístico feito em prol de minha versão final da tese, agradeço até pelo seu pequeno Miguel, que assim como a também pequena Isabella, me permitiu esquecer um pouco minhas recorrentes preocupações durante essa pesquisa.

Como eu estava prevendo, alguns nomes acabaram ficando de fora. Não falei, por exemplo, dos professores Carlos Alberto, Liliane e Daniele, que fizeram de mim um matemático melhor. Nem mencionei outros funcionários do departamento de Matemática, onde sempre fui bem tratado. Também não falei nada de meus sogros, que trouxeram ao mundo um fator decisivo: Minha mulher, Mônica.

Todas as contribuições citadas anteriormente foram condições necessárias à realização desse sonho. Mas a condição necessária e suficiente foi mesmo a Mônica, a mulher que me entende melhor que ninguém, porque compartilha do meu mundo, dos meus anseios, e da minha vida. As madrugadas incontáveis usadas pela simples companhia, as inúmeras concessões pessoais e profissionais, os sacrifícios acadêmicos, a dedicação incondicional, o cansaço transformado em esperança e o amor incansavelmente traduzido em gestos não são termos que descrevem a mim, são termos que descrevem a mulher responsável por minha permanência nesse doutorado em cada momento de alívio, desespero, animação ou hesitação. Essa dura fase, não a primeira nem a última, mostrou o quanto somos fortes juntos e o quanto sua existência me faz bem. Assim, agradeço à minha amada Mônica por todo o apoio e por ser não a minha metade, mas meus três quartos. De certa forma, é o nome dela que deveria assinar essa tese...

Brasília, julho de 2015.

Resumo

Nesta tese, usamos a conexão entre a distância Mallows, $d_\alpha(F, G)$, e a convergência em distribuição para estender resultados assintóticos envolvendo somas parciais, mesmo quando $d_\alpha(F, G) = \infty$. Para isso, produzimos vários resultados matemáticos envolvendo as distribuições ponderadas e a distância Mallows ponderada, $d_{\alpha,w}(F, G)$. Provamos também a convergência em distância Mallows do processo empírico geral, um tipo particular de somas parciais baseadas numa distribuição empírica F_n , para uma variável aleatória Gaussiana e, por meio do processo quantil empírico, atestamos o mesmo modo de convergência da estatística $nd_{2,w}^2(F_n, F)$ para um funcional ponderado de Pontes Brownianas.

Palavras-chave: distribuições estáveis, distância Mallows ponderada, distribuições ponderadas, processos empíricos.

Abstract

The connection between Mallows distance, $d_\alpha(F, G)$, and the convergence in distribution has been successfully used to establish asymptotic results for partial sums. We further explore this connection by making use of the weighted Mallows distance, $d_{\alpha, w}(F, G)$, to extend the results for the extreme cases when $d_\alpha(F, G) = \infty$. Our results are applied to heavy-tailed partial sums and to partial sums that arise in the context of empirical processes and their related functionals.

Keywords: stable laws, weighted Mallows distance, weighted distributions, empirical processes, quantile processes.

Sumário

Introdução	i
1 Preliminares	1
1.1 Introdução	1
1.2 Distribuições Estáveis e Variação Regular	3
1.3 Processos Empíricos	7
1.4 Distância Mallows	11
2 Distância Mallows Ponderada	14
2.1 Introdução	14
2.2 Convergência de somas parciais na distância Mallows	16
2.3 Distribuições ponderadas e distância Mallows ponderada	18
2.4 Convergências entre sequências	23
3 Somas parciais e ponderação	28
3.1 Introdução	28
3.2 Convergências de somas parciais sob ponderação	29
3.3 Exemplos ilustrativos	35
4 Convergência em distância Mallows de processos empíricos	42
4.1 Introdução	42
4.2 Processo empírico geral	46
4.3 Processo quantil - Convergência de ordem 1	51
4.4 Processo quantil - Convergência de ordens superiores	54
5 Distância Mallows ponderada estabilizada	59
5.1 Introdução	59
5.2 Discussão de hipóteses	60
5.3 Convergência em distância Mallows de $nd_{2,w}^2(F_n, F)$	64
5.4 Generalização do caso de variação regular	67
Considerações Finais	71
Referências Bibliográficas	73

Introdução

Nesta tese, propomos o uso de ponderações para estabelecer convergências, na distância Mallows, de somas parciais produzidas por amostras de uma distribuição F para uma variável aleatória com distribuição de cauda pesada, α -estável G , $\alpha \geq 1$. Extensões para o caso de sequências que formam uma martingale também serão abordadas. Nossos resultados proveem análogos ponderados para conhecidas convergências na distância Mallows e para conhecidas equivalências entre tal modo de convergência e a convergência em distribuição. A vantagem deles reside, como veremos em detalhes, no fato de que o uso apropriado de ponderações pode contornar limitações inerentes a essas conhecidas conclusões. Além disso, dedicamos atenção particular sobre a convergência em distância Mallows do processo empírico $\beta_n(t)$, um tipo especial de somas parciais advindas da distribuição empírica. A partir do processo quantil (processo inverso de $\beta_n(t)$), estudamos o comportamento assintótico estabilizado da distância Mallows ponderada entre uma distribuição empírica F_n e a correspondente função de distribuição F , isto é, analisamos o comportamento assintótico da estatística $nd_{2,w}^2(F_n, F)$.

Para cumprir nosso objetivo, introduzimos as noções de ponderação w , de distribuição ponderada F_w e a distância Mallows ponderada $d_{\alpha,w}(F, G)$. A distância Mallows de ordem $\alpha > 0$ entre duas funções de distribuição F e G é dada pela expressão

$$d_{\alpha}(F, G) = \left(\inf_{(X,Y)} E|X - Y|^{\alpha} \right)^{\frac{1}{\alpha}}. \quad (1)$$

Nesse caso, o ínfimo é tomado sobre todos os vetores aleatórios (X, Y) com distribuições marginais F e G , ou seja, X tem distribuição F e Y tem distribuição G . Essa expressão é uma métrica no espaço das distribuições com α -ésimo momento finito, isto é,

$$\mathcal{L}^{\alpha} = \left\{ F; \int_{-\infty}^{\infty} |x|^{\alpha} dF(x) < \infty \right\}.$$

Em termos de variáveis aleatórias, é comum encontrarmos na literatura a expressão $d_{\alpha}(X, Y)$, que apenas denota $d_{\alpha}(F_X, F_Y)$. Uma representação útil de (1) em \mathcal{L}^{α} , conforme Dall'Aglio (1956), é dada por

$$d_{\alpha}(F, G) = \left(\int_0^1 |F^{-}(t) - G^{-}(t)|^{\alpha} dt \right)^{\frac{1}{\alpha}} \quad \alpha \geq 1, \quad (2)$$

em que F^- e G^- são, respectivamente, as inversas generalizadas de F e G . Essa representação foi reobtida por Bickel e Freedman (1981) e também por Dorea e Ferreira (2012), onde as restrições de $F \in \mathcal{L}^\alpha$ e $G \in \mathcal{L}^\alpha$ foram removidas.

O interesse pela métrica (1) não é recente. Uma versão primitiva dela, proposta por Gini, em 1914, provém de sua relação com o clássico problema de transporte de terra, proposto por Gaspar Monge, em 1781, no qual a preocupação reside no custo relativo ao transporte de certo volume de terra de um local a outro. Segundo Levina e Bickel (2001), a métrica (1) é a medida que descreve apropriadamente tal problema. Graças à formulação teórica, feita em Kantorovich (1940), a respeito do problema geral de transporte de massa (uma generalização do problema clássico de Monge), tanto a métrica quanto o problema tornaram-se acessíveis a muitas áreas da matemática. Uma recente referência histórica detalhada sobre o assunto é a obra de Rachev e Rüschendorf (1998).

Mallows (1972) utilizou, num contexto estatístico, as propriedades métricas de (1) para provar um Teorema do Limite Central e reobter a representação (2). Com base nesse trabalho, Bickel e Freedman (1981) investigaram as propriedades topológicas de (1) e mostraram aplicações estatísticas para a métrica no método *bootstrap*, dentre outros. No apêndice deste último trabalho, foi demonstrada uma equivalência entre a convergência na métrica (1) e a convergência de momentos que, a partir de então, viria a ser muito utilizada por todos os trabalhos interessados em convergências correlatas ao tema, situação que não é diferente conosco, motivo pelo qual enfatizaremos tal equivalência na Proposição 1.2.

Aplicações estatísticas mais recentes da distância Mallows podem ser encontradas nos trabalhos de Zhou et al. (2005), Kosmelj e Billard (2012) e também em Irpino e Verde (2014, 2014 e 2015), só para citar alguns exemplos.

Contribuições recentes sobre a convergência dessa métrica para somas parciais foram obtidas por Johnson e Samworth (2005), Barbosa e Dorea (2009) e também por Dorea e Ferreira (2012).

Em Barbosa e Dorea (2009), foi provada, sob certas condições, a convergência do tipo Limite Central dada por

$$d_\alpha \left(\frac{X_1 + \dots + X_n - c_n}{n^{1/\alpha}}, Y \right) \xrightarrow{n} 0, \quad \alpha \in [1, 2), \quad (3)$$

na qual Y é uma variável α -estável com distribuição G , e $\{X_n\}_{n \geq 1}$ é uma sequência de variáveis aleatórias independentes e identicamente distribuídas com distribuição F . Para $\alpha > 1$ é possível considerar $c_n = 0$. A principal imposição à validade de (3) é a condição

$$d_\alpha(F, G) < \infty. \quad (4)$$

Já em Dorea e Ferreira (2012), foi demonstrado que

$$d_\alpha \left(\frac{X_1 + \dots + X_n}{n^{1/\alpha}}, Y \right) \xrightarrow{n} 0 \quad \Leftrightarrow \quad \frac{X_1 + \dots + X_n}{n^{1/\alpha}} \xrightarrow{D} Y. \quad (5)$$

Novamente, a principal imposição à equivalência (5) é a condição (4).

Dado que G é α -estável, a finitude da distância Mallows disposta em (4) é uma restrição importante aos resultados descritos em (3) e (5). De fato, pode-se provar que essa condição falha mesmo em casos mais simples, como quando F tem segundo momento finito ou quando F está no domínio de atração de G . Ilustraremos essa afirmação em nosso Exemplo 3.2 e no Exemplo 2.1. Para estabelecer resultados similares a (3) e (5) mesmo quando

$$d_\alpha(F, G) = \infty, \tag{6}$$

escolhemos trabalhar com ponderações w , que nada mais são que densidades no intervalo $[0, 1]$, para contornarmos (6) através da obtenção de

$$d_{\alpha,w}(F, G) < \infty. \tag{7}$$

O uso de ponderações em associação com a distância Mallows apareceu indiretamente no contexto dos chamados testes de bondade do ajuste (aqui chamados apenas de testes de ajuste), que visam avaliar a proximidade da distribuição F relativa a uma amostra aleatória X_1, X_2, \dots com uma dada classe de distribuições \mathcal{F} .

Originalmente, os testes de ajuste mais simples consistiam da verificação, para dada distribuição F_0 , da hipótese $F = F_0$. As estatísticas de Kolmogorov (1933) e Smirnov (1939, 1941), dadas por

$$D_n = \sqrt{n} \sup_{-\infty < x < \infty} |F_n(x) - F(x)|,$$

$$D_n^+ = \sqrt{n} \sup_{-\infty < x < \infty} (F_n(x) - F(x)),$$

estão entre as pioneiras nesse campo de estudo. Neste caso, F_n é a distribuição empírica associada à amostra aleatória X_1, X_2, \dots, X_n , isto é,

$$F_n(x, \omega) = \frac{\sum_{i=1}^n \mathbb{I}_{(X_i(\omega) \leq x)}}{n}.$$

Foi também Smirnov (1936,1937) que propôs uma estatística mais interessante, com o uso de uma ponderação, descrita por

$$W_n^2(w) = n \int_{-\infty}^{\infty} w(F_0(x))(F_n(x) - F_0(x))dF_0(x).$$

Todas as estatísticas que têm sido obtidas, desde então, pela variação de w são chamadas genericamente na literatura de estatísticas do tipo Crámer-von Mises, graças às contribuições relevantes desses autores para o tema. O emprego de uma ponderação permite, por exemplo, ponderar os desvios de acordo com algum critério conveniente (como seu valor esperado). Uma descrição histórica da evolução dessas estatísticas durante todo o século XX pode ser encontrada em del Barrio et al. (2000). Uma nova abordagem para testes de ajuste envolvendo famílias de escala e locação, geradas por F_0 , foi proposta em del Barrio et al. (1999). Tal abordagem era baseada na distância Mallows entre a distribuição empírica F_n e a distribuição F_0 . Em termos precisos, a estatística utilizada seria

$$\sqrt{nd_2}(F_n, F_0). \tag{8}$$

Novamente, a utilização de ponderações w permitiu a melhor adequação de testes envolvendo a distância Mallows para classes mais abrangentes de distribuições F_0 . A inclusão de uma ponderação w em (8) foi proposta por de Wet (2002), fazendo com que a nova estatística recorrente no campo de testes de ajuste passasse a ser

$$nd_{2,w}^2(F_n, F_0). \quad (9)$$

Podemos encontrar, em trabalhos da década passada, algumas convergências em distribuição que descrevem o comportamento assintótico de (9). É o caso, por exemplo, de Csörgo e Szabó (2003) e del Barrio et al. (2005).

A recorrência da convergência da estatística (9) em testes de ajuste constitui uma motivação importante para explorarmos seu comportamento assintótico, com ou sem o fator de ajuste, em um modo de convergência mais forte que a modalidade em distribuição e para outras ordens distintas da ordem 2, o que pode permitir a confecção de testes de ajuste mais eficientes e mais gerais.

Em Dorea e Lopes (2014), foram propostos testes de similaridade envolvendo estatísticas baseadas em (9) sem restrições aos momentos das distribuições envolvidas. Isso foi feito por meio de distribuições ponderadas, que foram relacionadas à distância Mallows ponderada pela igualdade

$$d_{2,w}(F, G) = d_2(F_w, G_w). \quad (10)$$

Em nossa tese, estendemos a igualdade (10), para $\alpha \geq 1$, a fim de prover uma forma eficiente de obter a condição (7) e produzir versões ponderadas abrangentes de (3) e (5). Outras ferramentas são desenvolvidas, mas essa extensão é fundamental.

A propósito da distribuição empírica, o fato de F_n ser nada mais que uma soma parcial de indicadores instigou naturalmente nosso interesse em $\beta_n(t)$, o processo empírico definido por

$$\{\beta_n(t)\}_{n \geq 1} = \{\sqrt{n}(F_n(t) - F(t))\}_{n \geq 1}, \quad t \in \mathbb{R}.$$

Por isso, dedicamos parte dessa tese ao estudo da convergência em distância Mallows de $\beta_n(t)$ e de seu inverso, o processo quantil empírico $q_n(t)$, descrito por

$$\{q_n(t)\}_{n \geq 1} = \{\sqrt{n}(F_n^-(t) - F^-(t))\}_{n \geq 1}, \quad t \in (0, 1).$$

Vale destacar que nosso Lema 4.1, que está focado em $q_n(t)$, é a ferramenta essencial na obtenção da convergência em distância Mallows de (9) em nosso Teorema 5.2, tendo em vista a relação imediata entre esses entes que decorre da representação (2), a saber,

$$\sqrt{n}d_{2,w}(F_n, F) = \left(\int_0^1 q_n^2(t)w(t)dt \right)^{\frac{1}{2}}.$$

Certamente, nossos resultados pretendem ser os mais gerais possíveis. Contudo, nos deteremos preferencialmente em garantir nossas conclusões pelo menos para as distribuições α -estáveis, pois são estas as possíveis distribuições-limite não degeneradas para uma soma parcial estabilizada. O interesse em tal classe de distribuições se

justifica também porque, para além do caso Gaussiano, esse conjunto de distribuições (para $0 < \alpha < 2$) tem um maior apelo intuitivo para aplicações, já que elas não possuem a restritiva característica do segundo momento finito e conservam a invariância por escala e locação, condição que as define. Qualquer contribuição dessa tese que valha para estáveis não Gaussianas, por causa da variação regular de suas caudas, será naturalmente estendida para a classe de todas as distribuições com cauda de variação regular. Ressalte-se que a classe das distribuições com cauda de variação regular contém propriamente a classe das distribuições estáveis não Gaussianas, como mostram os exemplos da distribuição de Pareto e da distribuição maximal de Fréchet.

Uma vez que fizemos a descrição de nossos objetivos, nossas motivações e algumas estratégias, passamos agora a descrever a organização de nosso trabalho. Maiores detalhes sobre os capítulos serão fornecidos na seção de introdução em cada um deles.

No Capítulo 1, apresentamos o suporte teórico desta tese, com definições e resultados que darão apoio aos nossos raciocínios no decorrer da exposição. Tratamos das distribuições α -estáveis, com particular interesse nas propriedades ligadas à variação regular de suas caudas (no caso não Gaussiano). Abordamos, também, o processo empírico geral $\beta_n(t)$, o processo quantil empírico $q_n(t)$ e, finalmente, a distância Mallows. A convergência na distância Mallows será o conceito presente em praticamente todas as contribuições desta tese.

No Capítulo 2, iniciamos a apresentação de nossas contribuições. Discutimos os aspectos relevantes, para este trabalho, do teorema que estabelece a convergência descrita em (3). Definimos os conceitos dos entes matemáticos que usaremos, a saber: distribuições ponderadas e distância Mallows ponderada. Estabelecemos e demonstramos vários resultados fundamentais à elaboração das contribuições mais relevantes da tese. As primeiras consequências de tais resultados, nossas contribuições que versam sobre convergências em distância Mallows envolvendo sequências gerais, estão dispostas na última seção. O Lema 2.6 encerra o capítulo, apresentando nosso primeiro resultado de convergência na distância Mallows entre somas parciais, para o caso particular de martingales.

No Capítulo 3, apresentamos, usando os resultados do Capítulo 2, nossas principais contribuições. No Teorema 3.2, estabelecemos uma convergência em distância Mallows ponderada que vale mesmo quando não ocorre (4), condição restritiva à convergência (3). No Teorema 3.4, dispomos a versão da equivalência (5) que não depende da condição (4). O Teorema 3.3 provê uma equivalência da convergência em distância Mallows ponderada com o Teorema do Limite Central que, para além de sua relevância intrínseca, implica na convergência em distância Mallows do processo empírico $\beta_n(t)$, primeiro resultado do Capítulo 4.

No Capítulo 4, voltamos nossa atenção para a convergência em distância Mallows do processo empírico $\beta_n(t)$ e do processo quantil $q_n(t)$. Esse capítulo dialoga com os capítulos anteriores, porque $\beta_n(t)$ é uma soma parcial de muito interesse, como bem o mostra o contexto dos testes de ajuste. Ele também dialoga com o Capítulo 5, porque resultados desenvolvidos para $q_n(t)$ são úteis para a análise de $nd_{2,w}^2(F_n, F)$.

Finalmente, no Capítulo 5, tratamos do comportamento assintótico de $nd_{2,w}^2(F_n, F)$, estatística para a qual já foi provada a convergência em distribuição para um funcional de pontes Brownianas. Discutimos aspectos técnicos relevantes dessa convergência antes de estendê-la para a distância Mallows de ordem 1, no Teorema 5.2. Ainda nesse capítulo, provamos que as distribuições estáveis e, adicionalmente, as distribuições com cauda de variação regular, podem ser incluídas na conclusão do Teorema 5.2.

Capítulo 1

Preliminares

1.1 Introdução

Neste capítulo, estabeleceremos os conceitos e resultados matemáticos que servirão como suporte teórico a esta tese. As referências bibliográficas relativas a essas informações serão dadas no início de cada seção, a menos que sejam de resultados que somente constam de artigos ou que constituam referências isoladas.

Na Seção 1.2, apresentaremos as distribuições estáveis, com ênfase nas propriedades de seu domínio de atração e na variação regular de suas caudas. Convergências envolvendo elementos do domínio de atração das estáveis serão obtidas em nosso trabalho, quase sempre, por meio das propriedades advindas da variação regular.

Na Seção 1.3, abordaremos o processo empírico geral e o processo quantil. Os processos empíricos gerais surgem como um tipo particular de somas parciais, como enfatizaremos no Capítulo 4, por isso, são um objeto natural de interesse neste trabalho. Além disso, certos funcionais do processo quantil podem ser vistos, como ressaltaremos no Capítulo 5, como uma distância Mallows ponderada, de maneira que a abordagem deste processo também é pertinente ao nosso estudo.

Na Seção 1.4, trataremos da métrica cuja convergência associada é o principal pilar dessa tese. A convergência em distância Mallows e de sua versão ponderada serão objetivos sempre presentes em nossos resultados e compõem nosso principal foco.

A seguir, estão dispostas algumas notações que serão recorrentes no texto, por ordem de aparição. Destacamos essas abreviações aqui com o intuito de evitar pequenas confusões ou esforços desnecessários à leitura.

- X, Y, Z : Símbolos comumente empregados para representar variáveis aleatórias.
 F, G, H : Símbolos comumente empregados para representar funções de distribuição.
 $D(G)$: Domínio de Atração da função de distribuição G .
 \xrightarrow{D} : Convergência em distribuição.
 $X \stackrel{d}{=} Y$: As variáveis aleatórias X e Y possuem a mesma distribuição.
 Φ : Função de distribuição de uma variável aleatória Gaussiana padrão.
 $\varphi_X(t)$: Função característica da variável aleatória X .
 $N(\mu, \sigma)$: Variável aleatória Gaussiana com média μ e variância σ .
 $X \sim F$: A variável aleatória X possui distribuição F .
 $U \in RV_\alpha$: A função real não negativa U possui variação regular no infinito com índice α .
 \bar{F} : Cauda à direita da distribuição F .
 $F(-x)$: Cauda à esquerda da distribuição F .
 $f(x) \approx g(x)$: as funções reais f e g são tais que $\lim_{x \rightarrow \infty} \frac{f(x)}{g(x)} = 1$.
 $a \wedge b$: $\min\{a, b\}$.
 $a \vee b$: $\max\{a, b\}$.
 \mathbb{I}_A ou $\mathbb{I}(A)$: Função indicadora do conjunto A .
 $X_{k:n}$: k -ésima estatística de ordem da amostra X_1, X_2, \dots, X_n .
 $\xrightarrow{q.c.}$: Convergência quase certa.
Bernoulli(p) : Variável aleatória Bernoulli com parâmetro p .
 $\beta_n(x)$: Processo empírico geral.
 \tilde{W} : Movimento Browniano padrão.
 $B(t)$: Ponte Browniana padrão.
 $q_n(t)$: Processo quantil empírico.
 $\xrightarrow{d_\alpha}$: Convergência na α -ésima distância Mallows.
 \mathcal{L}^α : Espaço das funções de distribuição com α -ésimo momento finito.
 \mathcal{L}_w^α : Espaço das funções de distribuição com α -ésimo momento ponderado finito.
 F_w : Função de distribuição F ponderada por w .
 $a_n \approx b_n$: As sequências a_n e b_n são tais que $\lim_{n \rightarrow \infty} \frac{a_n}{b_n} = 1$.
 M_Z : Função geradora de momentos da variável aleatória Z .
 $M_Z^{(k)}$: k -ésima derivada da função geradora de momentos da variável aleatória Z .
 $X_n = O_P(1)$: $P \left\{ \limsup_{n \rightarrow \infty} |X_n| < \infty \right\} = 1$.

Quando não houver perigo de confusão, especialmente a partir do Capítulo 2, a notação $F \stackrel{d}{=} X$ poderá ser empregada para dizer que a variável aleatória X possui distribuição F .

1.2 Distribuições Estáveis e Variação Regular

Nesta seção, apresentaremos os conceitos de domínio de atração, distribuição estável e variação regular, além de estabelecer entre eles as ligações que são pertinentes a esta tese.

As definições e resultados desta seção podem ser encontrados no capítulo XVII de Feller (1971), no primeiro capítulo de Resnick (1987) e no capítulo 9 de Breiman (1992).

Definição 1.1. *Sejam F e G funções de distribuição. Diremos que F pertence ao Domínio de Atração de G , e escreveremos $F \in D(G)$, se existirem seqüências de constantes positivas $\{a_n\}_{n \geq 1}$ e de constantes reais $\{b_n\}_{n \geq 1}$ tais que*

$$\frac{X_1 + \dots + X_n}{a_n} - b_n \xrightarrow{D} Y, \quad (1.1)$$

em que $\{X_n\}_{n \geq 1}$ é uma seqüência de variáveis aleatórias independentes e identicamente distribuídas com função de distribuição F , e Y é uma variável aleatória com distribuição G .

Claramente, o Teorema do Limite Central implica que $F \in D(\Phi)$ sempre que F possuir segundo momento finito. Uma condição necessária e suficiente para que uma distribuição possua domínio de atração não vazio é que ela seja a distribuição de uma variável aleatória estável. Apresentaremos essa classe a seguir.

Definição 1.2. *Diremos que uma variável aleatória X é Estável se, para cada natural $n > 0$, existirem um número positivo c_n e um número real d_n tais que*

$$X_1 + \dots + X_n \stackrel{d}{=} c_n X + d_n,$$

em que X_1, \dots, X_n são cópias independentes de X . Se $d_n = 0$, então diremos que X é estritamente estável.

Nos dois exemplos a seguir, a coleção X_1, \dots, X_n denota um conjunto de cópias independentes de X .

Exemplo 1.1. *Seja X uma variável aleatória que se degenera em um ponto, isto é, existe $c \in \mathbb{R}$ tal que $P(X = c) = 1$. Neste caso, sua função característica é dada por*

$$\varphi_X(t) = e^{itc}.$$

Além disso, para $S_n = X_1 + \dots + X_n$, é verdade que

$$\varphi_{S_n}(t) = e^{itnc} \Rightarrow \varphi_{\frac{S_n}{n}}(t) = \varphi_{S_n}(t/n) = e^{itc}.$$

Portanto,

$$\frac{S_n}{n} \stackrel{d}{=} X \quad \text{ou} \quad S_n \stackrel{d}{=} nX,$$

e X é estritamente estável com $c_n = n$.

Exemplo 1.2. Para $X \stackrel{d}{=} N(0, 1)$, sabe-se que $\varphi_X(t) = e^{-t^2/2}$. Segue daí que

$$\varphi_{\frac{S_n}{\sqrt{n}}}(t) = \varphi_{S_n}(t/\sqrt{n}) = \exp \left\{ -\frac{(t/\sqrt{n})^2}{2} \right\}^n = e^{-t^2/2} = \varphi_X(t).$$

Portanto,

$$S_n \stackrel{d}{=} \sqrt{n}X,$$

e X é estritamente estável com $c_n = \sqrt{n}$.

Em geral, se X tem distribuição $N(\mu, \sigma^2)$, podemos ver que

$$S_n \stackrel{d}{=} \sqrt{n}X + (n\mu - \sqrt{n}\mu).$$

Observação 1.1. Se X é estável, é possível mostrar que existem $0 < \alpha \leq 2$, $\sigma > 0$, $|\beta| \leq 1$ e $\mu \in \mathbb{R}$, tais que

$$c_n = n^{\frac{1}{\alpha}},$$

e

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

No caso $\alpha > 1$, é possível tomar $\mu = E(X)$.

A constatação da Observação 1.1 motiva a representação de uma variável aleatória estável por

$$X = S_\alpha(\sigma, \beta, \mu),$$

e a denominação α -estável. O número α é chamado *estabilidade*, σ é chamado de *escala*, β é a *simetria* e μ a *locação*. Comumente, ao falarmos da distribuição de X , associamos as mesmas denominações, ou seja, a representação $F_X = S_\alpha(\sigma, \beta, \mu)$, ou simplesmente $F = S_\alpha(\sigma, \beta, \mu)$, também recorrente.

Exemplo 1.3.

1. $S_2(\sigma, 0, \mu)$ é a distribuição $N(\mu, 2\sigma^2)$.
2. $S_1(1, 0, 0)$ é a distribuição Cauchy padrão.
3. $S_{1/2}(\sigma, 0, 0)$ é a distribuição de Levy.

Observação 1.2. Somente as distribuições α -estáveis do Exemplo 1.3 possuem forma fechada conhecida. Em contrapartida, sabe-se que a função característica da variável aleatória $X = S_\alpha(\sigma, \beta, \mu)$ é dada por

$$\log \varphi_X(t) = \begin{cases} i\mu t - \sigma^\alpha |t|^\alpha \left[1 - i\beta \operatorname{sgn}(t) \tan \left(\frac{\pi\alpha}{2} \right) \right] & , \quad \alpha \neq 1, \\ i\mu t - \sigma |t| \left[1 + i\beta \frac{2}{\pi} \operatorname{sgn}(t) \log |t| \right] & , \quad \alpha = 1. \end{cases}$$

As únicas distribuições α -estáveis com segundo momento finito são a distribuição normal e a distribuição que se degenera em uma constante. Se $0 < \alpha < 2$, então o momento de ordem γ de uma variável α -estável, é finito quando $\gamma < \alpha$ e infinito quando $\gamma \geq \alpha$ (conforme Breiman, 1992, p. 213). Dessa forma, a única distribuição estável não degenerada com segundo momento finito é a Gaussiana. A distribuição de Levy e a distribuição de Cauchy não possuem esperança finita nem variância finita.

A seguir, introduzimos o conceito de variação regular, que será usado para a caracterização das caudas de distribuições estáveis não Gaussianas.

Definição 1.3. *Uma função mensurável $U : \mathbb{R}_+ \rightarrow \mathbb{R}_+$ é de variação regular no infinito com índice α , e escreve-se $U \in RV_\alpha$, se*

$$\lim_{t \rightarrow \infty} \frac{U(tx)}{U(t)} = x^\alpha, \quad x > 0.$$

Sempre que $\alpha = 0$, $U \in RV_0$ é dita lentamente variante, ou de variação lenta. Em termos práticos, dizer que $U \in RV_\alpha$ é afirmar que U comporta-se assintoticamente como x^α . De fato, toda função $U \in RV_\alpha$ pode ser representada na forma $U(x) = x^\alpha \ell(x)$, com ℓ sendo uma função de variação lenta, bastando que se tome $\ell(x) = U(x)/x^\alpha$. Define-se também a variação regular no zero, bastando que o limite acima (com $t \rightarrow 0$) funcione para $0 < x < 1$. Para a variação regular no 1, também basta que o limite funcione ($t \rightarrow 1$) em $0 < x < 1$, com a substituição de $U(tx)$ por $U(1 - (1 - t)x)$.

Exemplo 1.4. *Toda função para a qual $\lim_{t \rightarrow \infty} U(t)$ existe, e é finito, é de variação lenta. Assim, toda função de distribuição é de variação lenta.*

Exemplo 1.5. *Considere $\bar{F}(t) = 1 - F(t)$, em que $F = S_1(1, 0, 0)$ é a distribuição de Cauchy padrão. Neste caso,*

$$\bar{F}(t) = \frac{1}{2} - \frac{\arctan(t)}{\pi}.$$

Pela Regra de L'Hôpital,

$$\begin{aligned} \lim_{t \rightarrow \infty} \frac{\bar{F}(tx)}{\bar{F}(t)} &= \lim_{t \rightarrow \infty} \frac{\frac{1}{2} - \frac{\arctan(xt)}{\pi}}{\frac{1}{2} - \frac{\arctan(t)}{\pi}} = \lim_{t \rightarrow \infty} \frac{\frac{x}{\pi [1 + (xt)^2]}}{\frac{1}{\pi [1 + t^2]}} \\ &= \lim_{t \rightarrow \infty} \frac{x(1 + t^2)}{1 + (xt)^2} = \lim_{t \rightarrow \infty} \frac{x \left(\frac{1}{t^2} + 1 \right)}{\frac{1}{t^2} + x^2} = x^{-1}. \end{aligned}$$

Portanto, $\bar{F} \in RV_{-1}$.

Exemplo 1.6. *Seja $\bar{F}(t) = 1 - F(t)$, em que $F = S_{1/2}(1, 0, 0)$ é a distribuição de Levy, isto é,*

$$\bar{F}(t) = \frac{2}{\pi} \int_0^{1/\sqrt{2t}} e^{-s^2} dt.$$

O uso da regra de L'Hôpital e do Teorema Fundamental do Cálculo implicam em

$$\begin{aligned} \lim_{t \rightarrow \infty} \frac{\bar{F}(tx)}{\bar{F}(t)} &= \lim_{t \rightarrow \infty} \frac{\int_0^{1/\sqrt{2tx}} e^{-s^2} ds}{\int_0^{1/\sqrt{2t}} e^{-s^2} ds} \\ &= \lim_{t \rightarrow \infty} \frac{\frac{-1}{2\sqrt{2x}} \cdot t^{-3/2} \cdot e^{-1/(2tx)}}{\frac{-1}{2\sqrt{2}} \cdot t^{-3/2} \cdot e^{-1/(2t)}} \\ &= \lim_{t \rightarrow \infty} \frac{1}{\sqrt{x}} \cdot \frac{e^{-1/(2tx)}}{e^{-1/(2t)}} = x^{-1/2}. \end{aligned}$$

Portanto, $\bar{F} \in RV_{-\frac{1}{2}}$.

Para a derivada das integrais que são funções dos limites de integração, foi usado o seguinte corolário do Teorema Fundamental do Cálculo:

$$F(t) = \int_{g_1(t)}^{g_2(t)} f(s) ds \Rightarrow F'(t) = g_2'(t)f(g_2(t)) - g_1'(t)f(g_1(t)).$$

Exemplo 1.7. Se F é a distribuição de Fréchet, isto é, $F(x) = \exp\{-x^{-\alpha}\}$, $\alpha > 0$ e $x > 0$, então o uso da regra de L'Hôpital nos permite ver que $\bar{F} \in RV_{-\alpha}$, pois

$$\lim_{t \rightarrow \infty} \frac{\bar{F}(tx)}{\bar{F}(t)} = \lim_{t \rightarrow \infty} \frac{1 - F(tx)}{1 - F(t)} = \lim_{t \rightarrow \infty} \frac{xf'(tx)}{f'(t)} = x^{-\alpha} \lim_{t \rightarrow \infty} \frac{\exp\{-(tx)^{-\alpha}\}}{\exp\{-t^{-\alpha}\}} = x^{-\alpha}.$$

O Exemplo 1.5 e o Exemplo 1.6 sugerem que a cauda à direita das distribuições estáveis, $\bar{F} = 1 - F$, pode ser caracterizada pela variação regular, no sentido de que a variação regular pode ser inteiramente determinada pelo índice de estabilidade. De fato, o próximo resultado confirma essa impressão (também para a cauda à esquerda $F(-x)$, $x > 0$) e permite-nos estabelecer as distribuições estáveis não Gaussianas (e os elementos de seu domínio de atração) como um subconjunto próprio das distribuições com cauda de variação regular.

Teorema 1.1. Se F é uma distribuição que está no Domínio de Atração de uma distribuição α -estável ($0 < \alpha < 2$), então existem $C, \bar{C} > 0$ e $\ell \in RV_0$ tais que, para $x \rightarrow \infty$,

$$\bar{F}(x) = 1 - F(x) \approx \bar{C} \left(\frac{2 - \alpha}{\alpha} \right) x^{-\alpha} \ell(x)$$

e

$$F(-x) \approx C \left(\frac{2 - \alpha}{\alpha} \right) x^{-\alpha} \ell(x).$$

Ou, alternativamente, existem constantes positivas \bar{C}_α e C_α tais que

$$\bar{F}(x) = 1 - F(x) \approx \bar{C}_\alpha x^{-\alpha}$$

e

$$F(-x) \approx C_\alpha x^{-\alpha}.$$

Observação 1.3. *A rigor, a recíproca do Teorema 1.1 também é válida (conforme Breiman, 1992, Teorema 9.34).*

O próximo resultado associa a variação regular de uma função com sua derivada.

Teorema 1.2. *Se*

(i) $U(x) \approx cx^\alpha \ell(x)$ para algum $c \geq 0$, para $\alpha \in \mathbb{R}$ e $\ell \in RV_0$,

(ii)

$$U(x) = \int_x^\infty u(t)dt, \quad \text{ou} \quad U(x) = \int_0^x u(t)dt,$$

(iii) u é monótona em (z, ∞) , para algum $z > 0$,
então

$$u(x) \approx c\alpha x^{\alpha-1} \ell(x), \quad x \rightarrow \infty.$$

Por fim, listamos a seguir algumas propriedades operatórias envolvendo a variação regular.

Proposição 1.1. *Sejam $U_i \in RV_{\alpha_i}$, $i = 1, 2$.*

(1a) $(U_1 \cdot U_2) \in RV_{\alpha_1 + \alpha_2}$, $\left(\frac{U_1}{U_2}\right) \in RV_{\alpha_1 - \alpha_2}$ e $(U_1 + U_2) \in RV_{(\alpha_1 \vee \alpha_2)}$.

(1b) Se $\lim_{x \rightarrow \infty} U_2(x) = \infty$, então $(U_1 \circ U_2) \in RV_{\alpha_1 \cdot \alpha_2}$.

(1c) Se U_1 é não decrescente, $\lim_{x \rightarrow \infty} U_1(x) = \infty$, e $\alpha_1 \geq 0$, então $U_1^{-1} \in RV_{\frac{1}{\alpha_1}}$.

(1d) Se U_1 é não crescente e $\alpha_1 \leq 0$, então

$$U_1^*(t) = \inf \left\{ x; U_1(t) \leq \frac{1}{x} \right\} \in RV_{-\frac{1}{\alpha_1}}.$$

Em vista dessa proposição, é natural questionar porque essas relações não foram usadas nos exemplos anteriores, no cálculo da variação regular de \bar{F} . A relação dada por $\bar{F} = 1 + (-1)F$ mostra que não é possível dispor dessas propriedades para conclusões sobre a variação regular de \bar{F} , já que a função $g \equiv -1$ não é de \mathbb{R}_+ em \mathbb{R}_+ . Os cálculos dos exemplos, se aplicados às distribuições Normal e Gumbel, mostram que elas não possuem cauda de variação regular, isto é, suas caudas não “imitam” potências de x .

A partir do Capítulo 3, as distribuições estáveis serão associadas à convergência em distância Mallows e aos processos de somas parciais por meio de ponderações. Em particular, nos Capítulos 4 e 5 exploraremos processos gerados por amostras aleatórias de distribuições, com destacado interesse no caso α -estável.

1.3 Processos Empíricos

Nesta seção, abordaremos os conceitos e resultados que envolvem os processos empíricos. Cabe esclarecer que, aqui e no restante do texto, a expressão *amostra aleatória* de uma distribuição F denota uma sequência de variáveis aleatórias independentes e identicamente distribuídas com F . As definições e resultados dessa seção podem ser encontrados no livro de Csörgo e Horváth (1993) ou no livro de Shorack e Wellner (1986).

Definição 1.4. *Sejam X_1, \dots, X_n uma amostra aleatória da distribuição F . Para cada natural $n \geq 1$, define-se a n -ésima função de distribuição empírica da amostra X_1, X_2, \dots, X_n como*

$$F_n(x, \omega) = \frac{\sum_{i=1}^n \mathbb{I}_{(X_i(\omega) \leq x)}}{n}. \quad (1.2)$$

Para x fixado, a função acima é uma variável aleatória. O numerador representa o número de índices i 's que satisfazem $X_i(\omega) \leq x$. Por isso, $F_n(x)$ também pode ser vista como a proporção de valores da amostra que são menores ou iguais a x , o que justifica intuitivamente a seguinte definição alternativa da distribuição empírica

$$F_n(x) = \begin{cases} 0 & \text{se } x < X_{1:n}, \\ \frac{k}{n} & \text{se } X_{k:n} \leq x < X_{k:n+1}, \quad k = 1, \dots, n-1 \\ 1 & \text{se } x \geq X_{n:n}. \end{cases} \quad (1.3)$$

Nesse caso, $X_{k:n}$ é a k -ésima estatística de ordem da amostra, isto é,

$$\begin{aligned} X_{1:n} &= \min\{X_1, X_2, \dots, X_n\} \\ X_{2:n} &= \min\{\{X_1, X_2, \dots, X_n\} - \{X_{1:n}\}\} \\ X_{3:n} &= \min\{\{X_1, X_2, \dots, X_n\} - \{X_{1:n}, X_{2:n}\}\} \\ X_{4:n} &= \min\{\{X_1, X_2, \dots, X_n\} - \{X_{1:n}, X_{2:n}, X_{3:n}\}\} \\ &\vdots \\ X_{n:n} &= \max\{X_1, X_2, \dots, X_n\}. \end{aligned}$$

Como $\mathbb{I}_{(X_i(\omega) \leq x)} \stackrel{d}{=} \text{Bernoulli}(F(x))$, é imediato ver que, para cada $x_0 \in \mathbb{R}$,

$$E(F_n(x_0)) = F(x_0) \quad \text{e} \quad n \text{Var}(F_n(x_0)) = F(x_0)(1 - F(x_0)).$$

Consequentemente, a Lei Forte dos Grandes Números implica que

$$F_n(x_0) \xrightarrow{q.c.} F(x_0), \quad n \rightarrow \infty. \quad (1.4)$$

Em outras palavras, a sequência de variáveis aleatórias $F_n(x_0)$ é um bom estimador para o valor $F(x_0)$. Mas F_n , vista como função de distribuição, também é um bom estimador para a distribuição F , conforme atesta o Teorema de Glivenko-Cantelli, reproduzido a seguir.

Teorema 1.3. (Glivenko-1917, Cantelli-1933)

Se F_n é a distribuição empírica relativa a uma amostra aleatória X_1, \dots, X_n da distribuição F (não necessariamente conhecida), então

$$\sup_{x \in \mathbb{R}} |F_n(x) - F(x)| \xrightarrow{q.c.} 0, \quad n \rightarrow \infty. \quad (1.5)$$

Tendo em vista aspectos interessantes acerca de amostras, a exemplo de intervalos de confiança, só para citar uma recorrente investigação em Estatística, as relações (1.4) e (1.5) motivam naturalmente o estudo do seguinte processo estocástico:

Definição 1.5. *Define-se o n -ésimo Processo Empírico, ou simplesmente Processo Empírico, baseado na amostra aleatória X_1, \dots, X_n de F , como*

$$\{\beta_n(x), x \in \mathbb{R}\} = \{\sqrt{n}(F_n(x) - F(x)), x \in \mathbb{R}\}, \quad n \geq 1.$$

As considerações anteriores e a definição de processo empírico induzem ao questionamento em torno de possíveis relações entre $\beta_n(x)$ e a distribuição Gaussiana, via Teorema do Limite Central. De fato, o processo empírico está relacionado a uma variável aleatória Gaussiana particular, que faz parte da chamada Ponte Browniana.

Definição 1.6. *Um processo estocástico $\{B(t), 0 \leq t \leq 1\}$ é uma Ponte Browniana se*

$$B(t) = \tilde{W}(t) - t\tilde{W}(1), 0 \leq t \leq 1,$$

em que $\{\tilde{W}(t), 0 \leq t < \infty\}$ é um Movimento Browniano.

Trivialmente, $B(t)$ é um processo Gaussiano e $E(B(t)) = 0$. Além disso, para $s, t \in [0, 1)$,

$$\begin{aligned} \text{Cov}[B(s), B(t)] &= \text{Cov}[\tilde{W}(s) - s\tilde{W}(1), \tilde{W}(t) - t\tilde{W}(1)] \\ &= E\left\{\left[\tilde{W}(s) - s\tilde{W}(1)\right] \cdot \left[\tilde{W}(t) - t\tilde{W}(1)\right]\right\} \\ &\quad - E\left[\tilde{W}(s) - s\tilde{W}(1)\right] \cdot E\left[\tilde{W}(t) - t\tilde{W}(1)\right] \\ &= E\left[\tilde{W}(s)\tilde{W}(t) - t\tilde{W}(s)\tilde{W}(1) - s\tilde{W}(t)\tilde{W}(1) + st\tilde{W}(1)^2\right] \\ &= \min\{s, t\} - st = (s \wedge t) - st. \end{aligned}$$

Portanto, $\text{Var}(B(t)) = t(1 - t)$, o que permite a caracterização da Ponte Browniana como um processo Gaussiano com média zero e função covariância “ $(s \wedge t) - st$ ”.

Enunciamos a seguir, em forma de Lema, o Teorema do Limite Central para processos empíricos.

Lema 1.1. *Seja F uma função de distribuição. Se $x \in \mathbb{R}$ é tal que $0 < F(x) < 1$, então*

$$\beta_n(x) \xrightarrow{D} Z(x) \stackrel{d}{=} N(0, F(x)(1 - F(x))) \quad (1.6)$$

Essa convergência será peça fundamental em nosso estudo dos processos empíricos a ser feito no Capítulo 4.

A próxima definição inicia nossa abordagem do processo empírico quantil, um processo estocástico que pode ser visto como um inverso do processo empírico $\beta_n(x)$.

Definição 1.7. *Se X_1, \dots, X_n é uma amostra aleatória da distribuição F , então:*

1. *A inversa generalizada de F é a função dada por*

$$F^-(t) = \inf\{x; F(x) \geq t\}, \quad F^-(0) = F^-(0^+).$$

2. *A inversa generalizada empírica ou função quantil empírica é dada por*

$$F_n^-(t, \omega) = \inf\{x; F_n(x, \omega) \geq t\}$$

Naturalmente, por sua definição, a função quantil empírica pode ser vista como variável aleatória ou como função de distribuição. Para um fixado $t \in [0, 1)$, a definição acima permite a seguinte representação.

$$F_n^-(t) = \begin{cases} X_{1:n} & , \quad t = 0, \\ X_{k:n} & , \quad \frac{k-1}{n} < t < \frac{k}{n}. \end{cases} \quad (1.7)$$

Assim como ocorre com a função de distribuição empírica F_n , podemos associar à inversa empírica F_n^- um processo estocástico.

Definição 1.8. *Define-se o n -ésimo Processo Quantil Empírico, $n \geq 1$, ou simplesmente Processo Quantil, associado à amostra X_1, \dots, X_n de F , como*

$$\{q_n(t), t \in (0, 1)\} = \{\sqrt{n}(F_n^-(t) - F^-(t)), t \in (0, 1)\}, \quad n \geq 1.$$

Ao passo que F_n funciona como bom estimador para a distribuição F , a distribuição quantil empírica não funciona como bom estimador para a inversa generalizada F^- , exceto nos pontos de continuidade desta, conforme atesta o resultado a seguir.

Lema 1.2. (Conforme Csörgo e Horváth, 1993, p. 366)

Seja F uma função de distribuição.

1. *A menos que F tenha suporte limitado, é verdade que*

$$P \left\{ \lim_{n \rightarrow \infty} \sup_{0 < t < 1} |F^-(t) - F_n^-(t)| = \infty \right\} = 1.$$

2. *Se F^- é contínua em t_0 , então*

$$\lim_{n \rightarrow \infty} F_n^-(t_0) = F^-(t_0)$$

Este Lema já mostra que, na melhor das hipóteses, o processo quantil $q_n(t)$ só é bem comportado nos pontos de continuidade da inversa generalizada, o que já impõe certas limitações a qualquer analogia que possa ser feita com $\beta_n(t)$. O próximo lema confirma essa linha de pensamento, ao mostrar que o comportamento assintótico de $q_n(t)$ está bem definido, mas sob condições mais restritivas do que as que temos para o comportamento assintótico de $\beta_n(t)$.

Lema 1.3. (Conforme Shorack e Wellner (1986), p. 639).

Suponha que F é uma função de distribuição derivável em $F^-(t)$, $t \in (0, 1)$, que satisfaça a condição $F'(F^-(t)) = \frac{1}{f(F^-(t))} > 0$. Neste caso, é verdade que

$$q_n(t) \xrightarrow{D} \frac{B(t)}{f(F^-(t))} \stackrel{d}{=} N \left(0, \frac{t(1-t)}{f^2(F^-(t))} \right). \quad (1.8)$$

No Capítulo 4, para um fixado $t \in (0, 1)$, obteremos um modo de convergência mais forte tanto da sequência de variáveis aleatórias $\beta_n(t)$ para a ponte Browniana $B(t)$ como da sequência $q_n(t)$ para a variável aleatória limite de (1.8). Trata-se da convergência em distância Mallows, que está associada à métrica homônima em certo espaço de funções. Na próxima seção, abordaremos a distância Mallows e a convergência associada a ela.

1.4 Distância Mallows

Definição 1.9. Para $\alpha > 0$, define-se a *Distância Mallows de ordem α* , ou α -ésima *Distância Mallows*, ou simplesmente, *Distância Mallows entre as funções de distribuição F e G* , como

$$d_\alpha(F, G) = \left(\inf_{(X, Y)} E|X - Y|^\alpha \right)^{\frac{1}{\alpha}}. \quad (1.9)$$

Neste caso, o ínfimo é tomado sobre todos os vetores aleatórios (X, Y) com distribuições marginais F e G , ou seja, X tem distribuição F , e Y tem distribuição G .

Uma terminologia comumente empregada em estudos que envolvem este conceito é a *Distância Mallows de ordem α entre as variáveis aleatórias X e Y* , que nada mais é do que $d_\alpha(F_X, F_Y)$. Portanto, não faremos maiores diferenciações entre essas terminologias no desenvolvimento de nosso trabalho.

Segundo Bickel e Freedman (1981), a distância (1.9) é uma métrica no espaço das distribuições com α -ésimo momento finito, descrito por

$$\mathcal{L}^\alpha = \left\{ F \text{ é função de distribuição ; } \int_{-\infty}^{\infty} |x|^\alpha dF(x) < \infty \right\}. \quad (1.10)$$

Para $0 < \alpha < 1$, a distância d_α^α é uma métrica no mesmo espaço. Segue, portanto, que para $\alpha \geq 1$,

$$d_\alpha(F, G) \leq d_\alpha(F, H) + d_\alpha(H, G). \quad (1.11)$$

Observação 1.4. Apesar de ser possível que $d_\alpha(F, G) < \infty$ mesmo que as distribuições não estejam no espaço \mathcal{L}^α (basta considerar $F = G$), a imposição da finitude dos momentos para a adoção de d_α como métrica em \mathcal{L}^α não é, necessariamente, dispensável. No Capítulo 3, apresentaremos, no Exemplo 3.2, uma situação na qual a distância Mallows é infinita, apesar de uma das distribuições envolvidas possuir α -ésimo momento finito.

Nesta tese, as operações envolvendo a Distância Mallows serão feitas por meio de uma representação alternativa da expressão (1.9), cujo manuseio é mais conveniente aos nossos propósitos. Tal representação está disposta a seguir.

Teorema 1.4. (Dorea e Ferreira, (2012)).

Se $\alpha \geq 1$, então

$$d_\alpha^\alpha(F, G) = E\{|F^-(U) - G^-(U)|^\alpha\} = \int_0^1 |F^-(t) - G^-(t)|^\alpha dt \quad (1.12)$$

$$= E\{|X^* - Y^*|^\alpha\} = \int |x - y|^\alpha dH^*(x, y). \quad (1.13)$$

Neste caso, $(X^*, Y^*) \sim F \wedge G$, isto é $H^*(x, y) = P(X^* \leq x, Y^* \leq y) = F(x) \wedge G(y)$.

É possível mostrar que a representação da distância Mallows no Teorema 1.4 não pode ser assumida para $0 < \alpha < 1$, mesmo quando o α -ésimo momento das distribuições é finito, como mostra Exemplo 1 de Dorea e Ferreira (2012). Como essa representação

é a chave de alguns dos nossos resultados fundamentais, as contribuições de nosso trabalho sempre partirão do caso $\alpha \geq 1$.

A metricidade da distância Mallows motiva naturalmente alguma noção de convergência associada a ela. Na próxima definição, estabeleceremos o modo de convergência que permeará todo o texto deste trabalho.

Definição 1.10. *Sejam $\{F_n\}_{n \geq 1}$ e F funções de distribuição. Diremos que F_n converge para F em distância Mallows de ordem α , ou na α -ésima distância Mallows, se*

$$d_\alpha(F_n, F) \xrightarrow{n} 0.$$

A definição acima prescinde da finitude do α -ésimo momento, pois tal condição não é, a rigor, necessária. Para ver isso, basta considerar $F_n = F$ para $n \geq k \in \mathbb{N}$ com $F \notin \mathcal{L}^\alpha$.

A convergência em distância Mallows se relaciona com a convergência em distribuição por meio da equivalência disposta na próxima proposição. Esse resultado será útil em todos os capítulos deste trabalho, por permitir a obtenção da convergência em distância Mallows de forma indireta e simples. Um dos conceitos envolvidos no enunciado a seguir será explicado na sequência do texto.

Proposição 1.2. (Bickel e Freedman (1981)).

Se $F, \{F_n\}_{n \geq 1} \in \mathcal{L}^\alpha$, então a convergência $d_\alpha(F_n, F) \xrightarrow{n} 0$ é equivalente a qualquer das relações a seguir.

1. $F_n \xrightarrow{D} F$ e $\int_{-\infty}^{\infty} |x|^\alpha dF_n(x) \xrightarrow{n} \int_{-\infty}^{\infty} |x|^\alpha dF(x)$.

2. $F_n \xrightarrow{D} F$ e $\{F_n\}_{n \geq 1}$ é uniformemente α -integrável.

O uso da Proposição 1.2 para a obtenção da convergência em distância Mallows requer ferramentas que nos permitam provar a convergência dos momentos, já que isso não decorre naturalmente da convergência em distribuição. De fato, o Lema de Fatou e o Teorema do Mapeamento (conforme Billingsley, 1995, p. 338) mostram que a convergência em distribuição de F_n para F garante apenas que

$$\int_{-\infty}^{\infty} |x|^\alpha dF(x) \leq \liminf_n \int_{-\infty}^{\infty} |x|^\alpha dF_n(x), \quad \alpha \geq 1.$$

Uma situação na qual a convergência em distribuição pode levar à convergência dos momentos é aquela na qual temos a integrabilidade uniforme da sequência envolvida. Definimos esse conceito a seguir.

Definição 1.11. *Diremos que a sequência de variáveis aleatórias $\{X_n\}_{n \geq 1}$ é Uniformemente Integrável se*

$$\limsup_{b \rightarrow \infty} \frac{1}{n} \int_{(|X_n| \geq b)} |X_n| dP = 0. \quad (1.14)$$

Na expressão (1.14), o símbolo dP indica que a integração se dá com relação à medida de probabilidade do espaço de probabilidade (Ω, \mathcal{F}, P) onde estão definidas

todas as variáveis aleatórias da sequência $\{X_n\}_{n \geq 1}$. Uma condição suficiente para que tenhamos a integrabilidade uniforme é a limitação

$$\sup_n E|X_n|^{1+\varepsilon} < \infty, \quad \text{para algum } \varepsilon > 0. \quad (1.15)$$

De maneira geral, pode-se definir a α -integrabilidade uniforme da sequência de variáveis aleatórias $\{X_n\}_{n \geq 1}$ como a integrabilidade uniforme de $\{(X_n)^\alpha\}_{n \geq 1}$. Naturalmente podemos falar sem maiores explicações sobre integrabilidade uniforme de sequências compostas por funções de distribuições $\{F_n\}_{n \geq 1}$ pois, nesse caso, a Definição 1.11 refere-se a qualquer sequência de variáveis aleatórias $\{X_n\}_{n \geq 1}$ tal que $X_n \sim F_n$. Formalmente, a α -integrabilidade uniforme de uma sequência de distribuições $\{F_n\}_{n \geq 1}$ pode ser suficientemente representada pela expressão

$$\sup_n \int_{-\infty}^{\infty} |x|^\alpha dF_n(x) < \infty,$$

como usaremos por vezes nesta tese.

No desenvolvimento de nosso trabalho, usaremos, na maior parte do tempo, o teorema a seguir como principal meio de prova da convergência de momentos.

Teorema 1.5. (Conforme Billingsley, 1995, seção 25)

Se $X_n \xrightarrow{D} X$ e $\{X_n\}_{n \geq 1}$ é uniformemente integrável, então X é integrável e

$$E(X_n) \xrightarrow{n} E(X).$$

Por outro lado, se $E|X_n| \xrightarrow{n} E|X|$ e $X_n \xrightarrow{D} X$, então $\{X_n\}_{n \geq 1}$ é uniformemente integrável.

O Teorema 1.5 encerra o conjunto de resultados e conceitos básicos que precisamos para apresentar objetivamente as contribuições desta tese. No próximo capítulo, apresentaremos a distância Mallows ponderada e os conceitos associados a ela que nos interessam. Provaremos resultados matemáticos que, para além de seu valor teórico, serão úteis na confecção das principais contribuições desta tese, no Capítulo 3.

Capítulo 2

Distância Mallows Ponderada

2.1 Introdução

O Teorema do Limite Central clássico, em sua versão mais simples, assegura que uma sequência de variáveis aleatórias independentes e identicamente distribuídas (i.i.d) $\{X_n\}_{n \geq 1}$, tal que $E(X_1) = \mu < \infty$ e $0 < Var(X_1) = \sigma^2 < \infty$, possui, para suas somas parciais

$$S_n = X_1 + X_2 + \dots + X_n,$$

o comportamento assintótico descrito por

$$\frac{S_n - \mu}{\sigma\sqrt{n}} \xrightarrow{D} N(0, 1). \quad (2.1)$$

Intuitivamente, esse resultado assegura um comportamento palpável para qualquer situação que possa ser vista como a resultante de muitos fatores individualmente insignificantes e independentes. Mesmo em casos nos quais dispensamos a independência ou a identicabilidade das distribuições, certas imposições asseguram o mesmo comportamento assintótico de (2.1). A noção intuitiva que acabamos de citar é a ilustração mais simples do alcance das somas parciais em qualquer campo de estudo, dado que incontáveis situações podem ser descritas nos termos dessa interpretação. Em um caso mais geral do Teorema do Limite Central, quando mantemos a independência, mas não a identicabilidade, por exemplo, a convergência (2.1) se traduz em

$$\frac{S_n - E(S_n)}{s_n} \xrightarrow{D} N(0, 1), \quad (2.2)$$

com $(s_n)^2 = Var(X_1) + \dots + Var(X_n)$. Neste caso, mantemos a intuição inicial de (2.1), pois (2.2) requer a imposição da chamada Condição de Lindenberg (trivialmente satisfeita no caso i.i.d), pela qual asseguramos que as parcelas $\frac{X_k - \mu_k}{s_n}$ sejam uniformemente pequenas para grandes valores de n . A Condição de Lindeberg é assim descrita:

$$\forall \varepsilon > 0, \lim_{n \rightarrow \infty} \frac{1}{s_n^2} \sum_{k=1}^n \int_{|x - \mu_k| > \varepsilon s_n} (x - \mu_k)^2 dF_k(x) = 0. \quad (2.3)$$

No limite acima, $X_k \sim F_k$. Ao dispensarmos a limitação imposta por meio da condição $0 < Var(X_1) = \sigma^2 < \infty$, obtemos a classe G_α das distribuições estáveis não

Gaussianas no lugar do limite assintótico em (2.2). De fato, as distribuições estáveis são os únicos limites em distribuição possíveis para as somas parciais, conforme o Teorema 1.1 e a Observação 1.3.

Trabalhos recentes têm tido sucesso na obtenção de uma convergência para estáveis análoga ao Teorema do Limite Central quando o modo de convergência é a distância Mallows. Nesta esteira, podemos citar Johnson e Samworth (2005), Barbosa e Dorea (2009) e Dorea e Ferreira (2012). A obtenção de convergências do tipo

$$d_\alpha \left(\frac{S_n}{n^{1/\alpha}}, G_\alpha \right) \xrightarrow{n} 0, \quad (2.4)$$

não prescindem de alguma hipótese análoga à condição (2.3) em sua estruturação. Na versão i.i.d de (2.4), com $X_i \sim F$, a condição tipo-Lindeberg se descreve pela imposição

$$d_\alpha(F, G_\alpha) < \infty, \quad (2.5)$$

apesar de a finitude do segundo momento associado a F não ser necessário nessas formulações. Contudo, a finitude exigida em (2.5) pode não acontecer mesmo em casos mais simples, o que impede convergências do tipo (2.4).

Neste Capítulo e no Capítulo 3, voltamos nossa atenção à produção de convergências de somas parciais envolvendo a distância Mallows nos casos em que (2.5) falha. Para tanto, lançaremos mão, neste Capítulo, da distância Mallows ponderada e das distribuições ponderadas, para gerar resultados que serão fundamentais para a elaboração de teoremas de convergência em distância Mallows de somas parciais no Capítulo 3.

Na Seção 2.2, formalizaremos e discutiremos o teorema de Barbosa e Dorea (2009) que motivou nossa abordagem à distância Mallows ponderada. Apresentaremos um exemplo no qual, apesar da distribuição F pertencer ao domínio de atração de G , a distância Mallows entre F e G é infinita, a despeito de intuitivamente esperarmos que esse fato não ocorra.

Na Seção 2.3, definiremos as distribuições ponderadas F_w e os momentos ponderados. Mostraremos, em nossos primeiros resultados, as relações entre distribuições ponderadas, momentos ponderados e convergência em distribuição. Por fim, definiremos a distância Mallows ponderada e concluiremos a seção com o nosso Lema 2.3, importante ferramenta que relaciona $d_{\alpha,w}(F, G)$ e $d_\alpha(F_w, G_w)$. Esse Lema será indispensável no Capítulo 3, onde buscaremos resultados de convergência em distância Mallows que contornem a restrição $d_\alpha(F, G) = \infty$.

Na Seção 2.4, usamos os lemas da Seção 2.3 para produzirmos resultados de convergência em distância Mallows que envolvem duas sequências de distribuições. De certa forma, os teoremas dessa seção reproduzem para a convergência em distância Mallows uma relação clássica entre sequências que é válida para as convergências em probabilidade e em distribuição, conforme descrevemos a seguir, para duas sequências quaisquer de variáveis aleatórias $\{X_n\}_{n \geq 1}$ e $\{Y_n\}_{n \geq 1}$.

$$|X_n - Y_n| \xrightarrow{P} 0, Y_n \xrightarrow{D} Y \Rightarrow X_n \xrightarrow{D} Y. \quad (2.6)$$

Esta relação é identificada, por vezes, pela denominação Teorema de Slutsky, assim como inúmeras relações que envolvem operações entre sequências de variáveis aleatórias fracamente convergentes.

2.2 Convergência de somas parciais na distância Mallows

Nesta seção, apresentaremos um resultado de Barbosa e Dorea (2009) que prova uma convergência análoga ao Limite Central para a distância Mallows. Discutiremos alguns de seus aspectos e destacaremos a limitação de seu caso i.i.d, que serviu de motivação inicial às nossas abordagens neste Capítulo e no próximo.

O Teorema a seguir estabelece condições suficientes para que se aproxime, em distância Mallows, uma variável aleatória estritamente α -estável não Gaussiana por uma soma de variáveis aleatórias (não necessariamente identicamente distribuídas).

Teorema 2.1. (Barbosa e Dorea (2009))

Fixe $\alpha \in (0, 2)$. Considere $\{(X_n, Y_n)\}_{n \geq 1}$ uma sequência de vetores aleatórios independentes tal que Y_1, Y_2, \dots sejam cópias independentes de uma variável aleatória estritamente α -estável Y , satisfazendo, para cada $b > 0$,

$$\lim_{n \rightarrow \infty} \frac{1}{n} \sum_{i=1}^n \mathbb{E} \{ |X_i - Y_i|^\alpha \mathbb{I}(|X_i - Y_i| > bn^{(2-\alpha)/2\alpha}) \} = 0. \quad (2.7)$$

Neste caso, existe uma sequência de constantes $\{c_n\}_{n \geq 1}$ tal que

$$d_\alpha(F_n, G_\alpha) \xrightarrow{n} 0, \quad (2.8)$$

com $G_\alpha \stackrel{d}{=} Y$ e $F_n \stackrel{d}{=} \frac{X_1 + \dots + X_n - c_n}{n^{1/\alpha}}$. Além disso, para $1 < \alpha < 2$, podemos tomar

$$c_n = \sum_{j=1}^n EX_j - EY.$$

Observação 2.1. Há dois aspectos relevantes de interesse no Teorema 2.1.

1. A condição (2.7) equivale à clássica Condição de Lindeberg do Teorema do Limite Central. De fato, as identificações $X_i \leftrightarrow \frac{X_i}{s_n/n^{1/2}}$, $Y_i \leftrightarrow \frac{\mu_i}{s_n/n^{1/2}}$ e $b \leftrightarrow \varepsilon n^{1/2}$, mostram que

$$\begin{aligned} \frac{1}{s_n^2} \sum_{i=1}^n \int_{\{|x-\mu_i| \geq \varepsilon s_n\}} (x - \mu_i)^2 dF_{X_i}(x) &= \frac{1}{s_n^2} \sum_{i=1}^n E \{ |X_i - \mu_i|^2 \mathbb{I}(|X_i - \mu_i| \geq \varepsilon s_n) \} \\ &= \frac{1}{n} \sum_{i=1}^n E \left\{ \left| \frac{X_i - \mu_i}{\frac{s_n}{n^{1/2}}} \right|^2 \mathbb{I} \left(\frac{|X_i - \mu_i|}{s_n/n^{1/2}} \geq \varepsilon s_n \right) \right\} \\ &= \frac{1}{n} \sum_{i=1}^n E \{ |X_i|^\alpha \mathbb{I}(|X_i| \geq \varepsilon n^{1/\alpha}) \}. \end{aligned}$$

A última igualdade apenas considerou o caso particular $\mu_i = 0$ e $\sigma_i = 1$.

2. O Teorema 2.1 dispensa a condição de finitude do segundo momento exigida no Teorema do Limite Central, ao mesmo tempo que provê um modo de convergência

mais forte que a convergência em distribuição. Quando X_1, X_2, \dots são identicamente distribuídas com distribuição F , é possível mostrar que a condição (2.7) torna-se

$$d_\alpha(F, G_\alpha) < \infty. \quad (2.9)$$

Esse ponto de vista também surge diretamente como consequência da representação do Teorema 1.4, como é possível verificar no Lema 2 de Dorea e Ferreira (2012).

A versão identicamente distribuída do Teorema 2.1 pode ser assim descrita.

Teorema 2.2. (Versão i.i.d do Teorema 2.1)

Fixe $\alpha \in (0, 2)$. Considere $\{X_n\}_{n \geq 1}$ uma sequência de variáveis aleatórias independentes tal que $X_n \sim F$ para cada $n \geq 1$. Se existe uma variável aleatória estritamente α -estável $Y \sim G_\alpha$ satisfazendo

$$d_\alpha(F, G_\alpha) < \infty, \quad (2.10)$$

então existe uma sequência de constantes $\{c_n\}_{n \geq 1}$ tal que

$$d_\alpha(F_n, G_\alpha) \xrightarrow{n} 0, \quad (2.11)$$

com $F_n \stackrel{d}{=} \frac{X_1 + \dots + X_n - c_n}{n^{1/\alpha}}$. Para $1 < \alpha < 2$, podemos tomar $c_n = EX_1 - EY$.

Sob as mesmas notações do Teorema 2.2, com $1 \leq \alpha < 2$, pode-se ainda estabelecer uma equivalência entre a convergência na distância Mallows e a convergência em distribuição, conforme descrevemos a seguir.

Teorema 2.3. (Conforme Dorea e Ferreira (2012))

Seja $1 \leq \alpha < 2$. A condição (2.10) garante a equivalência dada por

$$d_\alpha(F_n, G_\alpha) \xrightarrow{n} 0 \quad \Leftrightarrow \quad F_n \xrightarrow{D} G_\alpha. \quad (2.12)$$

A condição (2.10) pode falhar mesmo em casos que, intuitivamente, esperaríamos que ela acontecesse naturalmente. Essa afirmação está ilustrada no próximo exemplo.

Exemplo 2.1. Considere a distribuição estritamente estável $G = S_1(1, 0, 0)$ e a distribuição F descrita por

$$F(-x) = 1 - F(x) = \begin{cases} \frac{1}{2\pi} \cdot \left(1 + \frac{1}{1 + \log 2}\right) & , \quad 0 \leq x < 2, \\ \frac{1}{\pi x} \cdot \left(1 + \frac{1}{1 + \log x}\right) & , \quad x \geq 2. \end{cases}$$

Usando o Teorema 1.1 e a Observação 1.3, podemos ver que $F \in D(G)$. Isso poderia conduzir-nos à expectativa de que a distância Mallows entre as distribuições seria finita, pois o conceito de domínio de atração estabelece alguma aproximação assintótica entre somas parciais provenientes de F e a distribuição G .

Contudo, pode-se mostrar que $d_1(F, G) = \infty$ (veja detalhes no Exemplo 2 de Dorea e Ferreira (2012)), de forma que não é possível, através do Teorema 2.2, aproximar em distância Mallows uma variável aleatória $Y \sim G$ por somas estabilizadas de amostras aleatórias da variável $X \sim F$. Não é possível, também, estabelecer a equivalência do Teorema 2.3

A fim de contornar os casos em que não ocorre a condição (2.10), nós proporemos, no Teorema 3.2, um resultado no qual a nova condição de convergência é a finitude da distância Mallows ponderada entre F e G . De fato, como será ilustrado em um momento posterior desta tese, o uso de uma ponderação na distância Mallows pode mudar a condição $d_\alpha(F, G) = \infty$ e permitir convergências (com os devidos ajustes) do tipo (2.11). Essa generalização da hipótese poderá implicar no estabelecimento de uma convergência de somas parciais com parcelas ponderadas para uma distribuição estável não Gaussiana ponderada. As próximas duas seções desenvolverão resultados e conceitos que, no terceiro capítulo, darão suporte à elaboração de conclusões sobre convergência em distância Mallows (ponderada ou não) para somas parciais.

2.3 Distribuições ponderadas e distância Mallows ponderada

Nesta seção, apresentaremos os conceitos de distribuição ponderada e distância Mallows ponderada, bem como proporemos as primeiras contribuições desta tese, que visam possibilitar a elaboração de resultados que contemplem a convergência em distância Mallows (ponderada ou não) de somas parciais.

O uso de distribuições ponderadas surge quando uma amostra aleatória de uma população de interesse não pode ser obtida ou não é desejada (seleção de modelos), só para citar uma recorrência. Alguns trabalhos recentes em matemática aplicada e matemática pura trataram dessa ferramenta, tais como Ma et al. (2005), Kim (2008) e Li et al. (2012). Como uma ponderação na distribuição pode representar uma ponderação em sua variável aleatória, esse procedimento também pode ser estabelecido, num contexto de somas parciais, para induzir médias ponderadas ou para fazer um *tapering* nos dados.

Definição 2.1.

1. Diremos que $w : [0, 1] \rightarrow [0, \infty)$ é uma função peso, ou ponderação, se ela for uma densidade em $[0, 1]$. Em outras palavras, se $w(\cdot) \geq 0$ e se

$$\int_0^1 w(u) du = 1. \quad (2.13)$$

2. Dada uma função de distribuição F , denotaremos sua distribuição ponderada por

$$F_w(x) = \int_{-\infty}^x w(F(y)) dF(y). \quad (2.14)$$

3. O espaço das distribuições com α -ésimo momento ponderado finito será descrito por

$$\mathcal{L}_w^\alpha = \left\{ F \text{ é função de distribuição ; } \int_{-\infty}^{\infty} |x|^\alpha w(F(x)) dF(x) < \infty \right\}.$$

Ressalte-se, para efeito de comparação, que o espaço definido no último item da Definição 2.3 é o equivalente ponderado do espaço

$$\mathcal{L}^\alpha = \left\{ F \text{ é função de distribuição ; } \int_{-\infty}^{\infty} |x|^\alpha dF(x) < \infty \right\}.$$

Nas definições acima, observe que a função peso pode gerar distribuições ponderadas independentemente da finitude dos momentos na distribuição original. Esse fato torna essa definição particularmente atraente, pois é recorrente, nos textos que tratam de distribuições ponderadas, a exigência de que algum momento envolvendo a ponderação seja finito, uma clara limitação. A definição de distribuição ponderada (2.14) foi proposta por Dorea e Lopes (2014).

O próximo exemplo ilustra um efeito, desejável aos nossos propósitos, do uso de ponderações nos moldes acima sobre uma distribuição com segundo momento infinito.

Exemplo 2.2. *Considere a função de distribuição F do Exemplo 2.1, isto é,*

$$F(-x) = 1 - F(x) = \begin{cases} \frac{1}{2\pi} \cdot \left(1 + \frac{1}{1 + \log 2} \right) & , \quad 0 \leq x < 2, \\ \frac{1}{\pi x} \cdot \left(1 + \frac{1}{1 + \log x} \right) & , \quad x \geq 2. \end{cases}$$

Para $X \sim F$, temos que

$$\sum_{n=1}^{\infty} P(|X| > n^{1/2}) = \sum_{n=1}^{\infty} [1 - P(|X| \leq n^{1/2})] = \sum_{n=1}^{\infty} [1 - F(n^{1/2}) + F(-n^{1/2})] = \infty,$$

o que equivale a dizer que o segundo momento de F é infinito, $F \notin \mathcal{L}^2$ (conforme Rohatgi e Saleh, 1976, p. 76). Agora, considere a função

$$w(t) = \begin{cases} 32t^3 & , \quad 0 \leq t < \frac{1}{2}, \\ 32(1-t)^3 & , \quad \frac{1}{2} \leq t \leq 1. \end{cases}$$

A função w é não negativa e satisfaz a igualdade (2.13), razão pela qual ela é uma função peso, ou ponderação. A partir de (2.14), podemos ver que a distribuição ponderada F_w é dada por

$$F_w(x) = \begin{cases} 8[F(x)]^4 & , \quad x \leq -2 \\ 8A^4 & , \quad -2 < x \leq 0 \\ 1 - 8A^4 & , \quad 0 < x < 2 \\ 1 - 8(1 - F(x))^4 & , \quad x \geq 2. \end{cases}$$

em que $A = \frac{1}{2\pi} \cdot \left(1 + \frac{1}{1 + \log 2} \right)$. Se $X_w \sim F_w$, então

$$\sum_{n=1}^{\infty} P(|X_w| > n^{1/2}) = \sum_{n=1}^{\infty} [1 - F_w(n^{1/2}) + F_w(-n^{1/2})] < \infty,$$

e o segundo momento ponderado é finito. Em outras palavras, $F \in \mathcal{L}_w^2$ ou $F_w \in \mathcal{L}^2$.

O efeito, descrito no Exemplo 2.2, da ponderação sobre a distribuição será muito útil na obtenção de alguma versão da condição (2.10) quando ela não ocorrer naturalmente. Veja, por exemplo, que se a ponderação que acabamos de usar garantir $G_w \in \mathcal{L}^2$, para a distribuição $G = S_1(1, 0, 0)$ do Exemplo 2.1, então, a desigualdade de Minkowski garante que

$$d_2(F_w, G_w) < \infty,$$

apesar de termos

$$d_2(F, G) = \infty.$$

Os resultados desta seção pretendem usar fenômenos como esse em proveito de nosso objetivo neste capítulo.

Antes de apresentarmos o primeiro lema desta tese, observemos que, quando F é a distribuição uniforme, a ponderação w define uma função de distribuição dada por

$$W(x) = \int_0^x w(u)du, \quad 0 \leq x \leq 1.$$

Lema 2.1. *Considere $\alpha > 0$. Nesse caso*

$$F_w \in \mathcal{L}^\alpha \Leftrightarrow F \in \mathcal{L}_w^\alpha. \quad (2.15)$$

Além disso,

$$F_w(x) = W(F(x)) \quad e \quad F_w^-(u) = F^-(W^-(u)) \quad (2.16)$$

Demonstração:

Primeiramente, provaremos a igualdade (2.16).

Claramente, a substituição $y = F^-(u)$ implica em

$$F_w(x) = \int_{-\infty}^x w(F(y))dF(y) = \int_0^{F(x)} w(u)du = W(F(x)).$$

Seja $0 < u < 1$. Como W é contínua pela direita, é verdade que $W(W^-(u)) \geq u$. Por outro lado, se x é tal que $F(x) \geq W^-(u)$, então

$$W(F(x)) \geq W(W^-(u)) \geq u.$$

Por isso,

$$\{x; F(x) \geq W^-(u)\} \subset \{x; W(F(x)) \geq u\}$$

ou

$$\{x; F(x) \geq W^-(u)\} \subset \{x; F_w(x) \geq u\}. \quad (2.17)$$

Se x é tal que $F_w(x) = W(F(x)) \geq u$, é imediato que $F(x) \geq W^-(u)$. Portanto,

$$\{x; F_w(x) \geq u\} = \{x; W(F(x)) \geq u\} \subset \{x; F(x) \geq W^-(u)\}. \quad (2.18)$$

Das inclusões (2.17) e (2.18) segue a igualdade

$$F_w^-(u) = F^-(W^-(u)).$$

Suponha agora que $F_w \in \mathcal{L}^\alpha$. Por (2.16), temos

$$\int |x|^\alpha dF_w(x) = \int |x|^\alpha dW(F(x)) = \int |x|^\alpha w(F(x)) dF(x),$$

donde segue que $F \in L_w^\alpha$. A mesma sequência de igualdades garante a recíproca e, conseqüentemente, a equivalência disposta em (2.15). □

Nosso próximo lema prova que a convergência fraca é invariante por ponderações. Isso nos permitirá usar ponderações para garantir convergências de uma forma indireta.

Observação 2.2. *Deve-se frisar aqui que todos os resultados nos quais a positividade da ponderação for imposta, podem prescindir dessa hipótese num âmbito mais geral. Na verdade, o que interessa em tais resultados é o crescimento estrito da função que possui a ponderação w como densidade, de forma que a positividade estrita de w é suficiente, mas não necessária.*

Lema 2.2. (Invariância da Convergência Fraca por Ponderações). *Sejam $\{F_n\}_{n \geq 1}$ e G funções de distribuição. Se $w(\cdot) > 0$, então*

$$F_n \xrightarrow{D} G \iff F_{n,w} \xrightarrow{D} G_w. \quad (2.19)$$

Demonstração:

W é estritamente crescente, pois $w(\cdot) > 0$. Primeiro mostraremos que são iguais os conjuntos dos pontos de continuidade $C(G)$ de G e de sua versão ponderada G_w , isto é,

$$C(G) = C(G_w). \quad (2.20)$$

Suponha que exista $x \in C(G)$ tal que $x \notin C(G_w)$. Neste caso, há pelo menos uma sequência $x_n \rightarrow x$ tal que

$$G_w(x_n) = W(G(x_n)) \not\rightarrow W(G(x)) = G_w(x). \quad (2.21)$$

Como $x \in C(G)$ é imperativo que $G(x_n) \rightarrow G(x)$, e a continuidade de W implica que $G_w(x_n) = W(G(x_n)) \rightarrow W(G(x)) = G_w(x)$, o que contradiz (2.21). Logo, vale que $C(G) \subset C(G_w)$.

Similarmente, suponha que exista $x \in C(G_w)$ tal que $x \notin C(G)$. Neste caso, podemos encontrar uma sequência $\bar{x}_n \rightarrow x$ com $G(\bar{x}_n) \not\rightarrow G(x)$. Como W é estritamente crescente, vale que

$$G_w(\bar{x}_n) = W(G_w(\bar{x}_n)) \not\rightarrow W(G(x)) = G_w(x),$$

uma contradição ao fato de que $x \in C(G_w)$. Logo $C(G_w) \subset C(G)$, e concluímos a igualdade (2.20).

Para provar a necessidade em (2.19), assumamos que $F_n \xrightarrow{D} G$ e tome $x \in C(G_w)$. Como $x \in C(G)$, é verdade que $F_n(x) \rightarrow G(x)$. Pela continuidade de W , chegamos a

$$F_{n,w}(x) = W(F_n(x)) \rightarrow W(G(x)) = G_w(x).$$

Para a suficiência em (2.19), assumamos que $F_{n,w} \xrightarrow{D} G_w$ e tome $x \in C(G)$. Suponha que temos $F_n(x) \rightarrow G(x)$. Nesse caso, para alguma subsequência vale que $F_k(x) \xrightarrow{k} Z$, digamos $Z < G(x)$. Como W é contínua e estritamente crescente, temos que

$$F_{k,w}(x) = W(F_k(x)) \xrightarrow{k} W(z) < W(G(x)) = G_w(x),$$

uma contradição ao fato de que $x \in C(G_w)$. Logo, $F_n(x) \rightarrow G(x)$ e temos a equivalência desejada. □

Uma implicação imediata do Lema 2.1 é a relação entre a distância Mallows e sua versão ponderada, cuja definição destacaremos a seguir.

A versão ponderada da distância Mallows tem sido utilizada com sucesso na proposição de testes de similaridade envolvendo essa métrica, nos quais alguns requisitos técnicos requerem a adoção de uma ponderação para serem válidos. Na última década, alguns trabalhos se destacaram nessa área de atuação, tais como de Wet (2002), Csörgo e Szabó (2003) e del Barrio et al. (2005). Um trabalho mais recente que usa a distância Mallows ponderada fora do contexto estatístico que pode ser citado aqui é o artigo de Lisini et al. (2012).

Definição 2.2. *A α -ésima distância Mallows Ponderada entre duas funções de distribuição F e G é definida como*

$$d_{\alpha,w}(F, G) = \left(\int_0^1 |F^-(t) - G^-(t)|^\alpha w(t) dt \right)^{\frac{1}{\alpha}}. \quad (2.22)$$

Anteriormente, citamos alguns trabalhos recentes que fizeram uso de distância Mallows ponderada ou distribuições ponderadas em seu desenvolvimento. Contudo, uma abordagem unificada de ambos os entes matemáticos para a resolução de algum problema só apareceu em Dorea e Lopes (2014). Nesse trabalho, as distribuições ponderadas e a distância Mallows ponderada foram usadas em conjunto para a proposição de testes de similaridade que incluíssem as distribuições estáveis. Tal abordagem inspirou nossa metodologia de obtenção de uma condição análoga a

$$d_\alpha(F, G_\alpha) < \infty, \quad (2.23)$$

por meio da distância Mallows ponderada, como delinearemos a seguir.

Nosso próximo resultado estende o Lema 2.2 de Dorea e Lopes (2014), onde o caso $\alpha = 2$ foi tratado.

Lema 2.3. *Seja $\alpha \geq 1$. Assuma que $w(\cdot) > 0$. Nesse caso*

$$d_\alpha(F_w, G_w) = d_{\alpha,w}(F, G). \quad (2.24)$$

Demonstração:

Como $w(\cdot) > 0$, temos que $W^-(W(v)) = v$, $0 < v < 1$. A condição $\alpha \geq 1$ permite-nos usar a representação (1.12) do Teorema 1.4 e a relação (2.16) do Lema 2.1 para

escrever:

$$\begin{aligned}
 d_\alpha^\alpha(F_w, G_w) &= \int_0^1 |F_w^-(x) - G_w^-(x)|^\alpha dx \\
 &= \int_0^1 |F^-(W^-(u)) - G^-(W^-(u))|^\alpha du \\
 &= \int_0^1 |F^-(W^-(W(v))) - G^-(W^-(W(v)))|^\alpha dW(v) \\
 &= \int_0^1 |F^-(v) - G^-(v)|^\alpha w(v) dv \\
 &= d_{\alpha,w}^\alpha(F, G).
 \end{aligned}$$

Naturalmente, isso implica dizer que

$$d_\alpha(F_w, G_w) = d_{\alpha,w}(F, G).$$

□

Uma consequência imediata do Lema 2.3 é a possibilidade de definirmos, no caso $\alpha \geq 1$, a α -ésima distância Mallows ponderada como um ínfimo, nos moldes da Definição 1.9.

Definição 2.3. Para alguma ponderação w e para $\alpha \geq 1$, define-se a α -ésima distância Mallows entre as funções de distribuição F e G , como

$$d_{\alpha,w}(F, G) = \left(\inf_{(X_w, Y_w)} E|X_w - Y_w|^\alpha \right)^{\frac{1}{\alpha}}. \quad (2.25)$$

Nesse caso, o ínfimo é tomado sobre todos os vetores aleatórios (X_w, Y_w) com distribuições marginais F_w e G_w , ou seja, X_w tem distribuição F_w , e Y_w tem distribuição G_w .

Através do Lema 2.3, fica claro o papel fundamental das distribuições ponderadas na obtenção de uma condição análoga a (2.23) quando a distância Mallows entre as funções originais é infinita. Pelo Exemplo 2.2, se duas distribuições com segundo momento infinito podem ter segundo momento ponderado finito e, se a distância Mallows (de ordem menor ou igual a 2) entre as distribuições originais for infinita, a ponderação garantirá a finitude da distância Mallows entre as versões ponderadas (via desigualdade de Minkowski). Isso implicará na finitude da distância Mallows ponderada entre as distribuições originais, conforme o Lema 2.3.

Na próxima Seção, usaremos os lemas demonstrados até aqui para gerarmos resultados de convergência em distância Mallows entre sequências de funções de distribuição.

2.4 Convergências entre sequências

Nesta seção, serão apresentados alguns resultados que, de certa forma, reproduzem para a distância Mallows uma conhecida relação costumeiramente associada ao Teorema de Slutski, a saber,

$$|X_n - Y_n| \xrightarrow{P} 0, Y_n \xrightarrow{D} Y \Rightarrow X_n \xrightarrow{D} Y. \quad (2.26)$$

Lema 2.4. *Suponha $\alpha \geq 1$ e $w(\cdot) > 0$ uma ponderação limitada. Se $\{F_n\}_{n \geq 1}$ e $\{G_n\}_{n \geq 1}$ são sequências de funções de distribuição, então*

$$d_\alpha(F_n, G_n) \xrightarrow{n} 0 \Rightarrow d_{\bar{\alpha}, w}(F_n, G_n) \xrightarrow{n} 0, \quad 1 < \bar{\alpha} < \alpha. \quad (2.27)$$

Além disso, se $G_n \xrightarrow{D} G$ e

$$\sup_n \int_{-\infty}^{\infty} |x|^\alpha dG_n(x) < \infty, \quad (2.28)$$

então

$$d_\alpha(F_n, G_n) \xrightarrow{n} 0 \Rightarrow d_\alpha(F_n, G) \xrightarrow{n} 0. \quad (2.29)$$

Demonstração:

Suponha $\alpha > 1$. Se $\bar{\alpha} \in (1, \alpha)$, então para $p > 1$ e $q > 1$ tais que $\bar{\alpha}p = \alpha$ e $\frac{1}{p} + \frac{1}{q} = 1$, a Desigualdade de Hölder assegura que

$$\begin{aligned} d_{\bar{\alpha}, w}^\alpha(F_n, G_n) &= \int_0^1 |F_n^-(u) - G_n^-(u)|^{\bar{\alpha}} w(u) du \\ &\leq \left[\int_0^1 |F_n^-(u) - G_n^-(u)|^{\bar{\alpha}p} du \right]^{\frac{1}{p}} \cdot \left[\int_0^1 w^q(u) du \right]^{\frac{1}{q}} \end{aligned}$$

Como $\bar{\alpha}p = \alpha \geq 1$, a representação (1.12) do Teorema 1.4 e a limitação de $w(\cdot)$ atestam que, para alguma constante $C > 0$,

$$d_{\bar{\alpha}, w}^\alpha(F_n, G_n) \leq [d_\alpha^\alpha(F_n, G_n)]^{\frac{1}{p}} \cdot C \xrightarrow{n} 0,$$

e, conseqüentemente, resta provada a relação (2.27), pois

$$d_{\bar{\alpha}, w}(F_n, G_n) \xrightarrow{n} 0.$$

Considere $\alpha \geq 1$. Veja agora que a integrabilidade uniforme que decorre de (2.28) e a convergência $G_n \xrightarrow{D} G$ atestam, conforme a Proposição 1.2, que

$$d_\alpha(G_n, G) \xrightarrow{n} 0.$$

Pela relação métrica (1.11), temos que

$$d_\alpha(F_n, G) \leq d_\alpha(F_n, G_n) + d_\alpha(G_n, G),$$

o que completa a prova de (2.29), já que o primeiro termo vai a zero por hipótese. □

A hipótese de finitude do α -ésimo momento, disposta em (2.28), pode ser restritiva. Para $0 < \alpha < 2$, qualquer distribuição α -estável é tal que $G_\alpha \notin \mathcal{L}^\alpha$, por exemplo. Além disso, não podemos, a partir do Lema 2.4, obter a convergência em distribuição

$$F_n \xrightarrow{D} G, \quad (2.30)$$

pois precisaríamos garantir a finitude do α -ésimo momento de G , conforme a Proposição 1.2.

Neste caso, o uso de funções peso, em conjunto com a invariância da convergência fraca por ponderações do Lema 2.2 e com o Lema 2.3, provê uma abordagem alternativa ao Lema 2.4, na qual não só dispensamos a condição (2.28), como ainda obtemos a convergência em distribuição (2.30) sem a exigência $G \in \mathcal{L}^\alpha$.

Lema 2.5. *Considere $\alpha \geq 1$ e assuma que $w(\cdot)$ é uma ponderação limitada. Sejam $\{G_n\}_{n \geq 1}$ e G funções de distribuição tais que $G_n \xrightarrow{D} G$ e*

$$\sup_n \int |x|^\alpha dG_{n,w}(x) < \infty. \quad (2.31)$$

Neste caso, para qualquer sequência de distribuições $\{F_n\}_{n \geq 1}$, vale que

$$d_{\alpha,w}(F_n, G_n) \xrightarrow{n} 0 \Rightarrow F_n \xrightarrow{D} G. \quad (2.32)$$

Demonstração:

Pelo Lema 2.3, temos que

$$d_\alpha(F_{n,w}, G_{n,w}) = d_{\alpha,w}(F_n, G_n) \xrightarrow{n} 0. \quad (2.33)$$

Como $G_n \xrightarrow{D} G$, o Lema 2.2 implica em

$$G_{n,w} \xrightarrow{D} G_w. \quad (2.34)$$

A condição (2.31) atesta a α -integrabilidade uniforme de $\{G_{n,w}\}_{n \geq 1}$ e, conseqüentemente, a integrabilidade de G_w . Por isso, a convergência (2.33) e a convergência (2.34) são suficientes para usarmos o Lema 2.4 na conclusão de que $d_\alpha(F_{n,w}, G_w) \xrightarrow{n} 0$. Essa última convergência implica na α -integrabilidade uniforme de $\{F_{n,w}\}_{n \geq 1}$, o que nos leva a

$$F_{n,w} \xrightarrow{D} G_w.$$

Pelo Lema 2.2, temos $F_n \xrightarrow{D} G$, como queríamos.

□

Nossa meta, após a demonstração do Lema 2.5, é estabelecer convergências em distância Mallows de somas parciais a partir dos resultados produzidos até aqui.

Antes de encerrarmos nossa seção, abordaremos um caso peculiar no qual demonstraremos a convergência para zero de uma distância Mallows entre sequências de somas parciais. Esse caso mostrará uma situação na qual ocorre a convergência similar àquela que assumimos como hipótese nos dois lemas anteriores. Trata-se da situação em que a sequência das somas parciais é uma Martingale.

Definição 2.4. *Seja $\{\xi_n\}_{n \geq 1}$ uma sequência de variáveis aleatórias definidas num mesmo espaço de probabilidade (Ω, \mathcal{F}, P) e $\mathbb{F}_n \subset \mathcal{F}$ uma filtragem. Diremos que a sequência $\{\xi_n, \mathbb{F}_n\}_{n \geq 1}$ é uma Martingale se valem as duas propriedades a seguir.*

1. $E \{\xi_n\} < \infty, \forall n,$
2. $E \{\xi_n \mid \mathbb{F}_M\} = \xi_M, \text{ para cada } M \leq n.$

Agora, definiremos quando uma sequência de somas parciais de variáveis aleatórias é uma Martingale e, em seguida, citaremos o teorema que será usado na demonstração de nosso resultado.

Definição 2.5. *Seja $\{\xi_n\}_{n \geq 1}$ uma sequência de variáveis aleatórias e $\sigma(\xi_1, \dots, \xi_n)$ a sigma-álgebra gerada pela sequência, isto é, uma filtragem padrão. Diremos que a sequência $\left\{ \sum_{j=1}^n \xi_j, \sigma(\xi_1, \dots, \xi_n) \right\}_{n \geq 1}$ é uma Martingale se*

1. $E \left\{ \sum_{j=1}^n \xi_j \right\} < \infty, \forall n,$
2. $E \left\{ \sum_{j=1}^n \xi_j \mid \sigma(\xi_1, \dots, \xi_M) \right\} = \sum_{j=1}^M \xi_j, \text{ para cada } M \leq n.$

Teorema 2.4. (Conforme Hall e Heyde, 1980, Teorema 2.22)
 Se $1 \leq \alpha < 2$ e $\sup_{n \geq 1} \{E(|\xi_n|^\alpha)\} < \infty$, então

$$\frac{E \left\{ \left| \sum_{j=1}^n \xi_j \right|^\alpha \right\}}{n} \xrightarrow{n \rightarrow \infty} 0.$$

Nosso próximo lema encerra esta Seção e, conseqüentemente, este Capítulo.

Lema 2.6. *Seja $\{(X_n, Y_n)\}_{n \geq 1}$ uma sequência de vetores aleatórios com média zero. Assuma que*

$$\left\{ \sum_{j=1}^n (X_j - Y_j), \sigma((X_1, Y_1), \dots, (X_n, Y_n)) \right\}_{n \geq 1}$$

é uma martingale, que $F \stackrel{d}{=} X_n$ e $G \stackrel{d}{=} Y_n$. Se para algum $1 \leq \alpha < 2$, temos que

$$d_\alpha(F, G) < \infty, \tag{2.35}$$

então, para $F_n \stackrel{d}{=} X_1 + \dots + X_n$ e $G_n \stackrel{d}{=} Y_1 + \dots + Y_n$,

$$\frac{d_\alpha(F_n, G_n)}{n^{1/\alpha}} \xrightarrow{n \rightarrow \infty} 0. \tag{2.36}$$

Demonstração:

Como $\alpha \geq 1$, podemos usar a representação do Teorema 1.4. Tome $X_n^* \stackrel{d}{=} X_n, Y_n^* \stackrel{d}{=} Y_n$ e $(X_n^*, Y_n^*) \sim F \wedge G$. Temos, então, que

$$d_\alpha^\alpha(F_n, G_n) \leq E \left\{ \left| \sum_{j=1}^n (X_j^* - Y_j^*) \right|^\alpha \right\}.$$

A sequência $\{X_n^*, Y_n^*\}_{n \geq 1}$ pode ser escolhida de forma que se mantenha a estrutura de uma martingale. Por hipótese, temos que

$$d_\alpha^\alpha(F, G) = E(|X_1^* - Y_1^*|^\alpha) < \infty.$$

Por isso, é verdade que

$$\sup_n E(|X_n^* - Y_n^*|^\alpha) < \infty,$$

o que nos permite a aplicação do Teorema 2.4 para obtermos

$$\frac{d_\alpha^\alpha(F_n, G_n)}{n^{1/\alpha}} = d_\alpha^\alpha\left(\frac{X_1 + \dots + X_n}{n^{1/\alpha}}, \frac{Y_1 + \dots + Y_n}{n^{1/\alpha}}\right) \quad (2.37)$$

$$(2.38)$$

$$\leq \frac{E\left\{\left|\sum_{j=1}^n (X_j^* - Y_j^*)\right|^\alpha\right\}}{n} \xrightarrow{n \rightarrow \infty} 0. \quad (2.39)$$

□

O interesse no último Lema reside no fato de ele produzir uma ordem de convergência na distância Mallows entre somas parciais cujas diferenças formam uma Martingale. A partir disso, poderemos provar, no próximo capítulo, uma convergência em distribuição de somas parciais (no caso Martingale) para uma variável aleatória α -estável ($G_\alpha \notin \mathcal{L}^\alpha$), diretamente da condição (2.35).

Capítulo 3

Somas parciais e ponderação

3.1 Introdução

Neste capítulo, usaremos os resultados do Capítulo 2 para produzir algumas convergências em distância Mallows envolvendo somas parciais. O uso de ponderações estará presente em todos os teoremas, mesmo no primeiro deles, que é uma aplicação imediata do Lema 2.6, mas necessita do Lema 2.5 para sua demonstração.

Conforme será visto na Observação 3.1, nosso principal resultado neste capítulo, o Teorema 3.2, encerra a alternativa que encontramos para o Teorema 2.2, no caso em que

$$d_\alpha(F, G_\alpha) = \infty. \quad (3.1)$$

O Lema 2.3 é o ponto chave para a confecção desse resultado, pois ele contorna a situação na qual não podemos ter

$$d_\alpha(F, G) < \infty,$$

mas podemos encontrar uma ponderação tal que $F_w, G_w \in \mathcal{L}^\alpha$, pois assim

$$d_{\alpha,w}(F, G) = d_\alpha(F_w, G_w) < \infty. \quad (3.2)$$

Na Seção 3.2, apresentaremos, como já dito, o teorema que trata da convergência em distribuição de somas parciais para uma variável aleatória α -estável no caso Martingale (Teorema 3.1), além do Teorema 3.2, que apresenta alternativas à convergência de somas parciais na distância Mallows através da condição (3.2). Em seguida, tratamos de duas equivalências entre convergência em distância Mallows ponderada e convergência em distribuição ao estilo do Teorema do Limite Central, envolvendo somas parciais e distribuições α -estáveis tais que $1 \leq \alpha \leq 2$ (Teoremas 3.3 e 3.4). O corolário imediato do Teorema 3.3 será usado no resultado de abertura do Capítulo 4, em que trataremos de um tipo especial de soma parcial.

Na Seção 3.3, ilustraremos em detalhes, para G α -estável, algumas situações em que a condição (3.1) pode ser contornada por uma ponderação conveniente, de forma a obtermos a situação (3.2) e, conseqüentemente, podermos aplicar o Teorema 3.2. Faremos aí algumas generalizações de ponderações apropriadas para as classes de distribuições mais recorrentes em aplicações. Quase todas as ponderações sugeridas nessa seção não serão diferentes daquelas dispostas em Dorea e Lopes (2014). Ainda na Seção

3.3, exibiremos um caso no qual a ponderação genérica de Dorea e Lopes (2014) não pode funcionar para obtermos (3.2). Estabeleceremos, nesse caso, uma ponderação que implique diretamente na condição

$$d_{\alpha,w}(F, G) < \infty,$$

mesmo sem a finitude dos momentos ponderados. Isso mostrará que o Lema 2.3 provê uma abordagem suficiente, mas não necessária, para contornarmos (3.1).

3.2 Convergências de somas parciais sob ponderação

O teorema a seguir é o primeiro de alguns resultados análogos ao Teorema do Limite Central que envolvem a distância Mallows ou sua versão ponderada, e que são gerados a partir dos lemas do Capítulo 2. No enunciado adiante, a expressão *cauda pesada* denota uma distribuição com média zero e variância infinita.

Teorema 3.1. *Suponha que $\{X_n\}_{n \geq 1}$ seja uma sequência de variáveis aleatórias independentes e identicamente distribuídas com uma distribuição de cauda pesada F tal que*

$$\left\{ \sum_{j=1}^n X_j, \sigma(X_1, \dots, X_n) \right\}_{n \geq 1}$$

seja uma martingale. Assuma que, para algum $1 < \alpha < 2$, exista uma distribuição estritamente α -estável $G_\alpha = S_\alpha(1, 0, 0)$ tal que

$$d_\alpha(F, G_\alpha) < \infty. \tag{3.3}$$

Neste caso,

$$\frac{X_1 + \dots + X_n}{n^{1/\alpha}} \xrightarrow{D} G_\alpha.$$

Demonstração:

Tome $X^* \sim F$ e $(X^*, Y) \sim F \wedge G_\alpha$. Pela representação do Teorema 1.4, temos que

$$d_\alpha^\alpha(F, G_\alpha) = E|X^* - Y|^\alpha < \infty. \tag{3.4}$$

Agora, tome Y_1^*, Y_2^*, \dots cópias independentes de Y , X_1^*, X_2^*, \dots tais que $X_n^* \stackrel{d}{=} X_n$ e $(X_n^*, Y_n^*) \sim F \wedge G_\alpha$. Claramente $\{X_n^*\}_{n \geq 1}$ pode ser tomada de forma que

$$\left\{ \sum_{j=1}^n X_j^*, \sigma(X_1^*, X_2^*, \dots, X_n^*) \right\}_{n \geq 1}$$

seja uma Martingale. Para $\mathcal{F}_n = \sigma((X_1^*, Y_1^*), \dots, (X_n^*, Y_n^*))$, vale que

$$E \left\{ \sum_{j=1}^n (X_j^* - Y_j^*) | \mathcal{F}_{n-1} \right\} = \sum_{j=1}^{n-1} (X_j^* - Y_j^*) + E\{X_n^* | \mathcal{F}_{n-1}\} - E\{Y_n^* | \mathcal{F}_{n-1}\}.$$

Na igualdade acima, $E\{Y_n^*|\mathcal{F}_{n-1}\} = EY_n^* = 0$, pois Y_n^* é independente de \mathcal{F}_{n-1} . Da mesma forma, X_n^* é independente de Y_1^*, \dots, Y_{n-1}^* , o que implica em

$$E\{X_n^*|\mathcal{F}_{n-1}\} = E\{X_n^*|\sigma(X_1^*, X_2^*, \dots, X_n^*)\} = 0.$$

Logo

$$E\left\{\sum_{j=1}^n (X_j^* - Y_j^*)|\mathcal{F}_{n-1}\right\} = \sum_{j=1}^{n-1} (X_j^* - Y_j^*). \quad (3.5)$$

As relações (3.4) e (3.5) mostram que a sequência

$$\left\{\sum_{j=1}^n (X_j^* - Y_j^*)|\mathcal{F}_{n-1}\right\}_{n \geq 1}$$

é uma Martingale que satisfaz as hipóteses do Lema 2.6. Por isso, para

$$F_n \stackrel{d}{=} \frac{X_1 + \dots + X_n}{n^{1/\alpha}} \quad \text{e} \quad G_n \stackrel{d}{=} \frac{Y_1^* + \dots + Y_n^*}{n^{1/\alpha}},$$

é verdade que

$$d_\alpha^\alpha(F_n, G_n) = \frac{d_\alpha^\alpha(X_1 + \dots + X_n, Y_1^* + \dots + Y_n^*)}{n^{1/\alpha}} \xrightarrow{n} 0. \quad (3.6)$$

Como G_α é uma distribuição estritamente α -estável, segue da Definição 1.9 e da Observação 1.1 que $G_n = G_\alpha$ para cada $n \geq 1$. Uma consequência imediata disso é que

$$G_n \xrightarrow{D} G_\alpha. \quad (3.7)$$

Podemos escolher, como veremos na Observação 3.2, uma ponderação limitada $w(\cdot)$ que assegure a α -integrabilidade uniforme de $(G_\alpha)_w$ e, conseqüentemente, a α -integrabilidade de $\{G_{n,w}\}_{n \geq 1}$. Como tal ponderação é limitada, a convergência (3.6) implica em

$$d_{\alpha,w}(F_n, G_n) \xrightarrow{n} 0. \quad (3.8)$$

A α -integrabilidade uniforme de $\{G_{n,w}\}_n$, a convergência (3.7), e a convergência (3.8) são o que precisamos para usarmos o Lema 2.5 e concluirmos que

$$F_n \xrightarrow{D} G_\alpha.$$

□

Vimos na Seção 2.2, Teorema 2.2, que a condição de finitude da α -ésima distância Mallows entre F e G é uma importante limitação à convergência disposta naquele resultado. Nosso Lema 2.3 mostrou que, para uma ponderação positiva w , é verdade que

$$d_{\alpha,w}(F, G) = d_\alpha(F_w, G_w).$$

Isso contorna a situação na qual não podemos ter $d_\alpha(F, G) < \infty$ para algum $1 \leq \alpha < 2$, mas é possível introduzir uma função peso que assegure

$$d_{\alpha,w}(F, G) = d_\alpha(F_w, G_w) < \infty.$$

Seja $(X_1, Y_1), (X_2, Y_2), \dots$ uma seqüência de vetores aleatórios independentes e identicamente distribuídos tais que $F \stackrel{d}{=} X_n$ e $G \stackrel{d}{=} Y_n, n \geq 1$. Para uma dada ponderação w , considere $(X_{1,w}, Y_{1,w}), (X_{2,w}, Y_{2,w}), \dots$ tais que $F_w \stackrel{d}{=} X_{n,w}$ e $G_w \stackrel{d}{=} Y_{n,w}, n \geq 1$, e denote

$$F_{n,w}^* \stackrel{d}{=} \frac{X_{1,w} + \dots + X_{n,w}}{n^{1/\alpha}} \quad \text{e} \quad G_{n,w}^* \stackrel{d}{=} \frac{Y_{1,w} + \dots + Y_{n,w}}{n^{1/\alpha}}.$$

Teorema 3.2. *Seja $1 \leq \alpha < 2$ e assuma $w(\cdot) > 0$ limitada. Suponha que $E(X_{n,w}) = E(Y_{n,w}) = 0$ para $n \geq 1$, e que*

$$d_{\alpha,w}(F, G) < \infty. \tag{3.9}$$

Neste caso,

$$d_\alpha(F_{n,w}^*, G_{n,w}^*) \xrightarrow{n} 0. \tag{3.10}$$

Além disso, se existe uma função de distribuição G_w tal que $G_{n,w}^* \xrightarrow{D} G_w$ e

$$\sup_n \int |x|^{\bar{\alpha}} dG_{n,w}^*(x) < \infty,$$

para algum $1 < \bar{\alpha} \leq \alpha$, então

$$F_{n,w}^* \xrightarrow{D} G_w. \tag{3.11}$$

Demonstração:

Pelo Lema 2.3,

$$d_\alpha(F_w, G_w) = d_{\alpha,w}(F, G) < \infty. \tag{3.12}$$

Se tomarmos $\tilde{X}_n \stackrel{d}{=} X_{n,w}, \tilde{Y}_n \stackrel{d}{=} Y_{n,w}$ e $(\tilde{X}_n, \tilde{Y}_n) \sim (F_w \wedge G_w)$, a relação (3.12) implica que

$$\sup_n E \left\{ |\tilde{X}_n - \tilde{Y}_n|^\alpha \right\} < \infty, \tag{3.13}$$

pois $E|\tilde{X}_1 - \tilde{Y}_1|^\alpha = d_\alpha^\alpha(F_w, G_w) < \infty$. Usando os mesmos passos da relação (2.37), concluímos que

$$d_\alpha^\alpha(F_{n,w}^*, G_{n,w}^*) \leq \frac{E \left\{ \left| \sum_{j=1}^n \tilde{X}_j - \tilde{Y}_j \right|^\alpha \right\}}{n}.$$

A partir de (3.13), podemos usar o Teorema 2.4 para obter a convergência em (3.10), isto é,

$$d_\alpha(F_{n,w}^*, G_{n,w}^*) \xrightarrow{n} 0.$$

Suponha agora que $G_{n,w}^* \xrightarrow{D} G_w$ e $\sup_n \int |x|^{\bar{\alpha}} dG_{n,w}^*(x) < \infty$ para algum $1 < \bar{\alpha} \leq \alpha$. Nesse caso, o Lema 2.5 garante que

$$F_{n,w}^* \xrightarrow{D} G_w.$$

□

Observação 3.1. *Uma aplicação particularmente interessante do Teorema 3.2 é dada quando nós tomamos G como α -estável com $1 \leq \alpha < 2$. Nesse caso, sob a identicabilidade da distribuição, nós obtemos uma extensão do Teorema 2.2, em que a ponderação foi tomada como $w(u) = 1$, $0 \leq u \leq 1$.*

Como mostra nosso resultado, nós podemos obter (3.10), mas a convergência em (3.11) só pode ocorrer se G_w for uma distribuição α -estável, pois nenhum outro tipo de distribuição pode ser o limite assintótico de somas parciais estabilizadas.

Se $W(x) = \int_0^x w(u)du$, então (3.11) pode ser reescrita como

$$\frac{\sum_{j=1}^n F_w^-(F(X_j))}{n^{1/\alpha}} \xrightarrow{D} G_w^-(G(Y)),$$

onde $G \stackrel{d}{=} Y$. Note que para $G_w \stackrel{d}{=} Y_w$ temos

$$\begin{aligned} P(G_w^-(G(Y)) \leq y) &= P(G(Y) \leq G_w(y)) \\ &= P(Y \leq G^-(G_w(y))) \\ &= G(G^-(G_w(y))) \\ &= G_w(y). \end{aligned}$$

Portanto $G_w \stackrel{d}{=} G_w^-(G(Y))$.

Os Lemas 2.3 a 2.5 permitem-nos a obtenção de equivalências que reproduzem o Teorema do Limite Central Clássico para distribuições α -estáveis. A seguir, tratamos do caso $\alpha = 2$.

Sejam Y, Y_1, Y_2, \dots variáveis aleatórias independentes e identicamente distribuídas com distribuição comum $G = S_2 \left(\frac{\sigma}{\sqrt{2}}, 0, 0 \right) \stackrel{d}{=} N(0, \sigma^2)$. Sejam X, X_1, X_2, \dots variáveis aleatórias independentes e identicamente distribuídas com distribuição comum F . Assuma que $E(X) = 0$ e $E(X^2) = E(Y^2) = \sigma^2$, $0 < \sigma^2 < \infty$, e denote

$$F_n \stackrel{d}{=} \frac{X_1 + \dots + X_n}{\sqrt{n}}, \quad G_n \stackrel{d}{=} \frac{Y_1 + \dots + Y_n}{\sqrt{n}}.$$

Como $G_n = G$, temos que $G_n \xrightarrow{D} G$. A integrabilidade uniforme também é satisfeita, pois

$$E \left\{ \left(\frac{Y_1 + \dots + Y_n}{\sqrt{n}} \right)^2 \right\} = Var \left(\frac{Y_1 + \dots + Y_n}{\sqrt{n}} \right) = \sigma^2 < \infty.$$

Do Lema 2.4, se $d_2(F_n, G_n) \xrightarrow{n} 0$, temos

$$\frac{Y_1 + \dots + Y_n}{\sqrt{n}} \xrightarrow{D} N(0, \sigma^2). \tag{3.14}$$

Vale a pena notar que se (3.14) ocorre, é imediato que temos $d_2(F_n, G_n) \xrightarrow{n} 0$. De fato, se $d_2(F_n, G) \xrightarrow{n} 0$ e $d_2(G_n, G) \xrightarrow{n} 0$, então

$$d_2(F_n, G_n) \leq d_2(F_n, G) + d_2(G, G_n).$$

O teorema a seguir sintetiza em si o produto dessa reflexão e gera, a partir de nossos lemas, a equivalência, no caso Gaussiano, entre a convergência em distância Mallows ponderada e a convergência em distribuição.

Teorema 3.3. *Seja $\{X_n\}_{n \geq 1}$ uma seqüência de variáveis aleatórias independentes e identicamente distribuídas com distribuição F , e $w(\cdot) > 0$ uma ponderação limitada. Assuma que $E(X_1) = 0$ e $Var(X_1) = \sigma^2$, com $0 < \sigma^2 < \infty$. Para $\Phi \stackrel{d}{=} N(0, 1)$ e $F_n \stackrel{d}{=} \frac{X_1 + \dots + X_n}{\sigma\sqrt{n}}$, é verdade que*

$$d_{2,w}(F_n, \Phi) \xrightarrow{n} 0 \Leftrightarrow F_n \xrightarrow{D} \Phi.$$

Demonstração:

Suponha que $d_{2,w}(F_n, \Phi) \xrightarrow{n} 0$. A limitação de $w(\cdot)$ implica que $\Phi \in \mathcal{L}_w^2$ e $\{F_n\}_n \subset \mathcal{L}_w^2$. Do Lema 2.3, temos que $d_2(F_{n,w}, \Phi_w) \xrightarrow{n} 0$. Por isso, segue da Proposição 1.2 que $F_{n,w} \xrightarrow{D} \Phi_w$. Pelo Lema 2.2, concluímos que $F_n \xrightarrow{D} \Phi$.

Assuma que $F_n \xrightarrow{D} \Phi$. Pelo Lema 2.2, $F_{n,w} \xrightarrow{D} \Phi_w$. Novamente a limitação de $w(\cdot)$ garante que $\Phi \in \mathcal{L}_w^2$ e $\{F_n\}_n \subset \mathcal{L}_w^2$. Para concluirmos que $d_{2,w}(F_n, \Phi) \xrightarrow{n} 0$, precisamos antes mostrar, por meio da Proposição 1.2, que $d_2(F_{n,w}, \Phi_w) \xrightarrow{n} 0$, a fim de usarmos o Lema 2.3 em nosso favor. Isso exige que verifiquemos a integrabilidade uniforme de $\{F_{n,w}\}_n$, o que é imediato pela limitação de w , já que

$$\sup_n E \left\{ \left(\frac{X_1 + \dots + X_n}{\sigma\sqrt{n}} \right)^2 \right\} = \sup_n Var \left(\frac{X_1 + \dots + X_n}{\sigma\sqrt{n}} \right) = \sup_n \left(\frac{n\sigma^2}{\sigma^2 n} \right) = 1.$$

□

A equivalência do Teorema 3.3 permite, por exemplo, a obtenção da convergência em distribuição de uma soma parcial via distância Mallows ponderada. De fato, há uma infinidade de ponderações que podem ser usadas na distância Mallows ponderada da Definição 2.2, o que sugere mais formas de se trabalhar o lado esquerdo da equivalência do Teorema supracitado. Reciprocamente, situações nas quais já temos convergência em distribuição podem implicar num modo de convergência mais robusto, que é o caso da distância Mallows, conforme teremos oportunidade de ilustrar neste trabalho. Com efeito, a produção do corolário a seguir se justifica por nosso objetivo em usar o Teorema 3.3 para o primeiro resultado do próximo capítulo, no qual trataremos dos processos empíricos, um tipo particular de soma parcial.

Corolário 3.1. *Seja $\{X_n\}_{n \geq 1}$ uma seqüência de variáveis aleatórias independentes e identicamente distribuídas com distribuição F . Se $E(X_1) = 0$ e $Var(X_1) = \sigma^2$ com $0 < \sigma < \infty$, então as seguintes convergências são equivalentes:*

1. $\frac{X_1 + X_2 + \dots + X_n}{\sqrt{n}} \xrightarrow{D} N(0, \sigma).$

$$2. d_2 \left(\frac{X_1 + X_2 \dots + X_n}{\sqrt{n}}, N(0, \sigma) \right) \xrightarrow{n} 0.$$

Demonstração:

Basta usar $w \equiv 1$ no Teorema 3.3 e fazer exatamente os mesmos passos de sua demonstração, substituindo $N(0, 1)$ por $N(0, \sigma)$.

□

No próximo teorema, estabeleceremos a mesma convergência do Teorema 3.3 no caso α -estável com $1 \leq \alpha < 2$. Dispensaremos as imposições acerca da finitude do segundo momento. Ao invés disso, contaremos com uma ponderação apropriada para garantir α -ésimo momento ponderado finito. Essa rotina garante uma alternativa, que ao mesmo tempo é uma extensão, ao Teorema 2.3 quando falha a condição $d_\alpha(F, G_\alpha) = \infty$.

Teorema 3.4. *Considere $\{X_n\}_{n \geq 1}$ uma sequência de variáveis aleatórias independentes e identicamente distribuídas com distribuição F e assuma $E(X_1) = 0$. Sejam $1 \leq \alpha < 2$, $G = S_\alpha(1, 0, 0)$, e $w(\cdot)$ uma ponderação limitada tal que $F \in \mathcal{L}_w^\alpha$ e $G \in \mathcal{L}_w^\alpha$. Neste caso,*

$$d_{\alpha,w}(F_n, G) \xrightarrow{n} 0 \Leftrightarrow F_n \stackrel{d}{=} \frac{X_1 + \dots + X_n}{n^{1/\alpha}} \xrightarrow{D} G.$$

Demonstração:

Para provar a necessidade, considere uma sequência de variáveis aleatórias independentes $\{Y_n\}_{n \geq 1}$, distribuídas com G e $G_n \stackrel{d}{=} \frac{Y_1 + \dots + Y_n}{n^{1/\alpha}}$. Como $G_n \stackrel{d}{=} G$ pela definição de distribuição α -estável, temos que $G_n \xrightarrow{D} G$. Além disso, por hipótese, vale que

$$\sup_n \int |x|^\alpha dG_{n,w}(x) < \infty,$$

e que $d_{\alpha,w}(F_n, G_n) = d_{\alpha,w}(F_n, G) \xrightarrow{n} 0$. Por fim, o Lema 2.5 atesta que $F_n \xrightarrow{D} G$.

Para a suficiência, veja que o Lema 2.2 implica em $F_{n,w} \xrightarrow{D} G_w$. Além disso, temos que

$$\sup_n \int |x|^\alpha dF_{n,w}(x) < \infty,$$

o que é suficiente para que $d_\alpha(F_{n,w}, G_w) \xrightarrow{n} 0$, pela Proposição 1.2. Usando o Lema 2.3, concluímos que

$$d_{\alpha,w}(F_n, G) \xrightarrow{n} 0.$$

□

No Capítulo 2 e nas primeiras seções do Capítulo 3, temos nos debruçado, basicamente, sobre a produção de resultados de convergência nos quais a infinitude de momentos ou da distância Mallows sejam contornados pelo uso de alguma ponderação conveniente. Em termos simples, nos concentramos em situações nas quais

$$d_\alpha(F, G) = \infty, \tag{3.15}$$

mas, para alguma ponderação,

$$d_{\alpha,w}(F, G) < \infty. \quad (3.16)$$

Graças ao nosso Lema 2.3, é possível obter (3.16) por meio das condições

$$F \in \mathcal{L}_w^\alpha, \quad G \in \mathcal{L}_w^\alpha. \quad (3.17)$$

Uma particular atenção tem sido dada ao caso $G = S_\alpha(1, 0, 0)$, por ser essa uma classe de distribuições cuja relevância prática e teórica sempre atrai grande interesse. Na próxima seção, exibiremos em detalhes alguns exemplos em que (3.15) pode ser contornada por (3.16) para $G = S_\alpha(1, 0, 0)$, além de propor algumas generalizações acerca de certas classes de F . No que tange aos exemplos explícitos, escolheremos $G = S_1(1, 0, 0)$ por causa da limitação que temos a respeito da forma fechada das distribuições α -estáveis, conforme vimos na Observação 1.2. Quanto às generalizações que proporemos, sua validade será estabelecida para $G = S_\alpha(1, 0, 0)$ com $1 \leq \alpha < 2$.

3.3 Exemplos ilustrativos

Nesta seção, exibiremos, para $G = S_\alpha(1, 0, 0)$, alguns exemplos detalhados nos quais temos

$$d_\alpha(F, G) = \infty, \quad (3.18)$$

mas, para alguma ponderação w , temos

$$d_{\alpha,w}(F, G) < \infty. \quad (3.19)$$

Adicionalmente, proporemos a generalização adequada para o uso da ponderação w para algumas classes de distribuições F . Conforme dito na Observação 2.2, o que interessa na ponderação w é que ela garanta que sua primitiva W seja estritamente crescente. Parte de nossos resultados exigiu $w(\cdot) > 0$ porque essa condição é suficiente para que tenhamos W estritamente crescente, apesar de outras formatações da ponderação (sendo nula num conjunto de medida nula, por exemplo) garantirem o mesmo efeito.

Exemplo 3.1. *Vimos, no Exemplo 2.1, que $d_1(F, G) = \infty$, para $G = S_1(1, 0, 0)$ e F descrita por*

$$F(-x) = 1 - F(x) = \begin{cases} \frac{1}{2\pi} \cdot \left(1 + \frac{1}{1 + \log 2}\right) & , \quad 0 \leq x < 2, \\ \frac{1}{\pi x} \cdot \left(1 + \frac{1}{1 + \log x}\right) & , \quad x \geq 2. \end{cases}$$

Agora, considere a ponderação

$$w(t) = \begin{cases} 32t^3 & , \quad 0 \leq t < \frac{1}{2}, \\ 32(1-t)^3 & , \quad \frac{1}{2} \leq t \leq 1. \end{cases}$$

De acordo com o Exemplo 2.2, é verdade que $F_w \in \mathcal{L}^2$. A distribuição ponderada G_w é dada por

$$G_w(x) = \begin{cases} 8[G(x)]^4 & , \quad x < 0, \\ 1 - 8[1 - G(x)]^4 & , \quad x \geq 0. \end{cases}$$

Essa distribuição possui segundo momento finito, pois, para $Y_w \sim G_w$, temos que

$$\sum_{n=1}^{\infty} P(|Y_w| > n^{1/2}) = \sum_{n=1}^{\infty} 8 \left\{ [1 - G(\sqrt{n})]^4 + [G(-\sqrt{n})]^4 \right\} < \infty.$$

A finitude da expressão acima segue do teste da integral para séries e do comportamento das caudas de G , dado pelo Teorema 1.1, pois

$$8 \lim_{t \rightarrow \infty} \int_1^t [\bar{G}(x)]^4 dx + \lim_{t \rightarrow \infty} \int_1^t [G(-x)]^4 dx \approx \lim_{t \rightarrow \infty} \int_1^t \bar{C}_1 x^{-4} dx + \lim_{t \rightarrow \infty} \int_1^t C_1 x^{-4} dx.$$

O Lema 2.3, a representação da distância Mallows dada pelo Teorema 1.4, a desigualdade de Jensen e a desigualdade de Minkowski asseguram que

$$d_{1,w}^2(F, G) = d_1^2(F_w, G_w) = \{E|X^* - Y^*|\}^2 \leq E|X^* - Y^*|^2 = d_2^2(F_w, G_w) < \infty. \quad (3.20)$$

Na expressão acima, $(X^*, Y^*) \sim F_w \wedge G_w$.

Observação 3.2. De forma geral, é sempre possível ponderar uma distribuição α -estável ($0 < \alpha < 2$) de forma que se tenha como resultado uma distribuição ponderada com segundo momento finito. De fato, conforme Dorea e Lopes (2014), se $G = S_\alpha(\sigma, 0, \mu)$, então $G_w \in \mathcal{L}^2$ para

$$w(t) = \begin{cases} kt^{-\beta} & , \quad 0 < t < t_*, \\ k(1-t)^{-\beta} & , \quad t_* \leq t < 1. \end{cases}$$

em que $\beta < -\frac{2}{\alpha}$ e $k = \frac{1-\beta}{t_*^{1-\beta} + (1-t_*)^{1-\beta}}$.

A eficiência da ponderação acima na obtenção de momentos ponderados está ligada à variação regular da cauda de distribuições estáveis não Gaussianas (Teorema 1.1). Isso é perceptível pela expressão do segundo momento ponderado finito, dada por

$$\mu_w^2(G) = \int_{-\infty}^{\infty} x^2 w(G(x)) g(x) dx.$$

Como já se conhece a variação regular de $G'(x) = g(x)$ (pelo Teorema 1.2) e de x^α , basta escolher w de forma que $x^2 w(G(x)) g(x) \in RV_{-(1+\varepsilon)}$ para $\varepsilon > 0$. Isso é possível através das operações envolvendo a variação regular, descritas na Proposição 1.1. Pela Observação 3.2, percebemos que o sucesso da ponderação w do Exemplo 3.1, na obtenção de $F_w \in \mathcal{L}^2$, é uma consequência de $\bar{F}(x) \in RV_{-1}$ e $F(-x) \in RV_{-1}$, fatos verificáveis pela Definição 1.3. Por isso, se F é qualquer distribuição com caudas de variação regular com índice $-\gamma$, $0 < \gamma < 2$, e $G = S_\alpha(1, 0, 0)$, então podemos sempre garantir $d_{2,w}(F, G) = d_2(F_w, G_w) < \infty$, usando $\beta < -\frac{2}{(\alpha \wedge \gamma)}$ na ponderação da

Observação 3.2. Além das distribuições estáveis não Gaussianas, fazem parte do conjunto das distribuições F com cauda de variação regular exemplos como a distribuição maximal de Frechét de parâmetro $0 < \gamma < 2$, a distribuição de Pareto e a distribuição loggama.

A finitude da distância Mallows de ordem $\alpha > 0$ não exige, a rigor, que as distribuições envolvidas tenham α -ésimo momento finito (basta tomar as distribuições iguais). Em contrapartida, é possível que a mesma distância seja infinita ainda que uma das distribuições tenha α -ésimo momento finito. Ilustraremos esse fato no próximo Exemplo.

Exemplo 3.2. Considere a distribuição $G = S_1(1, 0, 0)$ e a distribuição F dada por

$$F(x) = \begin{cases} 0 & , \quad x < 1, \\ 1 - \frac{1}{x^3} & , \quad x \geq 1. \end{cases}$$

Podemos verificar aqui que $d_1(F, G) = \infty$. De fato, se for assumida a representação do Teorema 1.4 e a identificação

$$H(t) = |F^-(t) - G^-(t)| = \left| \left(\frac{1}{1-t} \right)^3 - \tan \left(t\pi - \frac{\pi}{2} \right) \right|, \quad 0 < t < 1,$$

é verdade que,

$$\begin{aligned} d_1(F, G) &= \int_0^1 H(t) dt \\ &= \int_0^a H(t) dt + \int_a^b H(t) dt + \int_b^1 H(t) dt \\ &> \int_b^1 \frac{1}{(1-t^5)} dt = \infty. \end{aligned}$$

Na segunda igualdade, em que $0 < a < b < 1$, temos que $H(t) < t^{-2}$ se $t \in (0, a)$, e $H(t) > (1-t^5)^{-1}$ se $t > b$, de forma que as duas primeiras integrais são positivas e finitas. Vale notar que o segundo momento de F é finito, de forma que a ponderação do Exemplo 3.1, por ser limitada, garante também que $F_w \in \mathcal{L}^2$, pois $t \in (0, 1)$ implica em $0 < t^{-\beta} \vee (1-t)^{-\beta} < 1$, já que $-\beta > 1$. Consequentemente:

$$\begin{aligned} \mu_F^2(w) &= \int_{-\infty}^{\infty} x^2 w(F(x)) dF(x) \\ &= \int_0^1 [F^-(t)]^2 w(t) dt \\ &= \int_0^{\frac{1}{2}} kt^{-\beta} [F^-(t)]^2 dt + \int_{\frac{1}{2}}^1 k(1-t)^{-\beta} [F^-(t)]^2 dt \\ &< k \int_0^{\frac{1}{2}} [F^-(t)]^2 dt + k \int_{\frac{1}{2}}^1 [F^-(t)]^2 dt = k \int_0^1 [F^-(t)]^2 dt < \infty \end{aligned}$$

Dessa forma, os mesmos passos de (3.20) garantem que

$$d_{1,w}(F, G) < \infty.$$

Pelo Exemplo 3.2, podemos garantir a condição $d_{\alpha,w}(F, G) < \infty$ sempre que F possuir segundo momento finito com a mesma ponderação da Observação 3.2, pois tal escolha não interfere na finitude dos momentos de F . As distribuições maximais de Weibull e Gumbel, a distribuição Burr XII, lognormal e Gaussiana são casos relevantes que atendem a essa observação.

As generalizações desta seção cobrem a maior parte das distribuições F mais recorrentes na estatística e na teoria de probabilidade. Dificilmente uma aplicação envolvendo distribuições de probabilidade toma elementos fora do conjunto das distribuições subexponenciais.

As distribuições subexponenciais dividem-se entre o Domínio de Atração Maximal da distribuição de Fréchet e o Domínio de Atração Maximal da distribuição Gumbel (conforme Embrechts et al., 1997, p. 451). Na primeira situação, F terá cauda de Variação Regular (Embrechts et al., 1997, Teorema 3.3.7), caso tratado após a Observação 3.2. Na segunda situação, apesar de não termos encontrado uma generalização interessante para a ponderação w , vimos que alguns casos, como as distribuições Burr XII e lognormal, foram alcançados pelo Exemplo 3.2.

A seguir, apresentamos um exemplo no qual obtemos $d_{\alpha,w}(F, G) < \infty$ através de uma ponderação que não garante, simultaneamente, $G \in \mathcal{L}_w^\alpha$ e $F \in \mathcal{L}_w^\alpha$. Por isso, a convergência (3.10) no Teorema 3.2 é mais forte do que nossos outros resultados, uma vez que não requer a finitude de momentos ponderados. A demonstração de (3.10), de fato, usa o Lema 2.3 (que não precisa de momentos ponderados finitos) e o Teorema 2.4, que é atendido pela hipótese

$$d_\alpha(F_w, G_w) < \infty.$$

Exemplo 3.3. Uma distribuição de Variação Lenta.

Sejam $G = S_1(1, 0, 0)$ e F a distribuição cuja densidade é

$$f(x) = \begin{cases} \frac{1}{2|x|(\log|x|)^2} & , \quad |x| > e, \\ 0 & , \quad cc. \end{cases}$$

Através do já citado critério da página 76 de Rohatgi e Saleh (1976), podemos ver que distribuição F não possui momentos de ordem alguma. De fato, para $X \sim F$, $\alpha > 0$ e $x > e$, temos

$$x^\alpha P(|X| > x) = x^\alpha [1 - F(x) + F(-x)] = \frac{x^\alpha}{2 \log(x)} \rightarrow \infty, \quad x \rightarrow \infty.$$

Com relação à distância Mallows de ordem 1, usando as expressões das inversas gene-

realizadas de F e G , vemos que:

$$\begin{aligned}
 d_1(F, G) &= \int_0^1 |F^-(t) - G^-(t)| dt \\
 &= \int_0^{1/2} \left| -e^{\frac{1}{2t}} - \tan[\pi(t - 1/2)] \right| dt + \int_{1/2}^1 \left| e^{\frac{1}{2(1-t)}} - \tan[\pi(t - 1/2)] \right| dt \\
 &= \int_0^{1/2} \{e^{1/2t} + \tan[\pi(t - 1/2)]\} dt + \int_{1/2}^1 \{e^{1/2(1-t)} - \tan[\pi(t - 1/2)]\} dt \\
 &= \lim_{\varepsilon \rightarrow 0} \left(\int_{\varepsilon}^{1/2} \{e^{1/2t} + \tan[\pi(t - 1/2)]\} dt + \int_{1/2}^{1-\varepsilon} \{e^{1/2(1-t)} - \tan[\pi(t - 1/2)]\} dt \right) \\
 &= \lim_{\varepsilon \rightarrow 0} \left(\int_{\varepsilon}^{1/2} \tan[\pi(t - 1/2)] dt + \int_{1/2}^{1-\varepsilon} \tan[\pi(1/2 - t)] dt \right) \\
 &\quad + \lim_{\varepsilon \rightarrow 0} \left(\int_{\varepsilon}^{1/2} e^{1/2t} dt + \int_{1/2}^{1-\varepsilon} e^{1/2(1-t)} dt \right).
 \end{aligned}$$

A terceira igualdade se justifica porque o primeiro integrando é não positivo em $(0, 1/2]$ e o segundo integrando é estritamente positivo em $(1/2, 1)$. Como integrais de funções contínuas são funções contínuas de seus limites de integração, vale a quarta igualdade. Agora, a substituição $u = 1 - t$ na segunda e na quarta integral nos levam a

$$\begin{aligned}
 d_1(F, G) &= \lim_{\varepsilon \rightarrow 0} 2 \int_{\varepsilon}^{1/2} \tan[\pi(u - 1/2)] du + \lim_{\varepsilon \rightarrow 0} 2 \int_{\varepsilon}^{1/2} e^{1/2u} du \\
 &= \lim_{\varepsilon \rightarrow 0} 2 \int_{\varepsilon}^{1/2} \{e^{1/2u} + \tan[\pi(u - 1/2)]\} du \\
 &> \lim_{\varepsilon \rightarrow 0} 2 \int_{\varepsilon}^{1/2} \frac{1}{u} du = \infty.
 \end{aligned}$$

A última desigualdade decorre das informações que podemos apurar, usando técnicas de otimização, sobre o comportamento no intervalo $(0, 1)$ da função

$$h(x) = e^{1/2x} + \tan\left[\pi\left(x - \frac{1}{2}\right)\right] - \frac{1}{x}.$$

Tentaremos agora uma ponderação que implique na finitude da expressão

$$d_{1,w}(F, G) = \lim_{\varepsilon \rightarrow 0} 2 \int_{\varepsilon}^{1/2} \{e^{1/2t} + \tan[\pi(t - 1/2)]\} w(t) dt.$$

Um cálculo direto mostra que a cauda de F possui variação lenta, o que nos afasta de uma tentativa nos moldes da Observação 3.2, já que o índice de variação regular aqui é nulo. Considere a ponderação

$$w(t) = \begin{cases} k \arctan(t) & , \quad 0 \leq t \leq c, \\ \frac{\tilde{k}}{e^{1/2(1-t)}} & , \quad c < t \leq 1. \end{cases}$$

Na função acima, k e \tilde{k} são constantes convenientemente escolhidas ($\tilde{k} > k$), e $c \in (\frac{2}{5}, \frac{1}{2})$ é a única raiz de $h(t) = \tilde{k}e^{-1/2(1-t)} - k \arctan(t)$ no intervalo $(0, 1)$. A relação entre \tilde{k} e k é natural, pois o decréscimo brusco da parte exponencial de $w(t)$ impõe essa necessidade. Uma escolha possível para tais constantes pode ser

$$(k, \tilde{k}) = \left(\frac{1}{2 \int_0^c \arctan(t) dt}, \frac{1}{2 \int_c^1 e^{-1/2(1-t)} dt} \right).$$

Passemos agora ao cálculo da distância Mallows entre F e G , ponderada por $w(t)$.

$$\begin{aligned} d_{1,w}(F, G) &= \lim_{\varepsilon \rightarrow 0} 2k \underbrace{\int_{\varepsilon}^{1/2} \arctan(t) \cdot \tan[\pi(u - 1/2)] dt}_{(I)} \\ &\quad + 2\tilde{k} \underbrace{\int_c^{1/2} \{e^{-1/2(1-t)} \cdot \tan[\pi(t - 1/2)]\} dt}_{(II)} \\ &\quad + \lim_{\varepsilon \rightarrow 0} 2\tilde{k} \int_{1/2}^{1-\varepsilon} e^{1/2(1-t)} \cdot e^{-1/2(1-t)} dt. \end{aligned}$$

Através de métodos numéricos e de otimização, prova-se que $\frac{2k(c-\varepsilon)}{5} \leq (I) < 0$ ou, no caso limite, a primeira parcela da expressão é negativa, mas limitada inferiormente por $-\frac{k}{5}$. Da mesma forma, a segunda parcela é negativa, mas limitada inferiormente por $-\frac{3k}{200}$. Assim, as duas primeiras parcelas resultam em um valor negativo, mas limitado inferiormente por $-\frac{k}{2}$. Como a terceira integral é igual a $\frac{\tilde{k}}{2}$, a relação entre as constantes \tilde{k} e k garante que $0 < d_{1,w}(F, G) < \infty$.

Diferentemente do que fizemos antes, não usamos no Exemplo 3.3 a sequência de resultados que garantiu a relação (3.20). Tal procedimento é impossível agora, pois a ponderação usada não garante a finitude do segundo momento ponderado de F . De fato,

$$\begin{aligned} \mu_F^2(w) &= k \int_0^c \arctan(t) \cdot e^{1/t} dt + \tilde{k} \int_c^{1/2} e^{1/t} \cdot e^{-1/(1-t)} dt + \tilde{k} \int_{1/2}^1 e^{1/2(1-t)} \cdot e^{-1/2(1-t)} dt \\ &> k \int_0^c \arctan(t) \cdot e^{1/t} dt \\ &> k \int_0^d \arctan(t) \cdot e^{1/t} dt > k \int_0^d \frac{1}{t} dt = \infty. \end{aligned}$$

A penúltima desigualdade se apoia em métodos computacionais e de otimização, que atestam a existência de $d \in (0, c)$ tal que $\arctan(d) \cdot e^{1/d} = \frac{1}{d}$. A última desigualdade é válida, porque $\arctan(t) \cdot e^{1/t} > \frac{1}{t}$ se $t \in (0, d)$. Dessa forma, a obtenção de uma

condição alternativa ao fato $d_\alpha(F, G) = \infty$ na construção de teoremas de convergência para somas parciais tem no Lema 2.3 uma ferramenta suficiente, mas não necessária.

No Exemplo 3.3, podemos conseguir a finitude dos segundos momentos ponderados de F e G (por consequência, restam finitos os momentos ponderados de primeira ordem) com a seguinte escolha:

$$w(t) = \begin{cases} \frac{\tilde{C}}{e^{1/t}} & , \quad 0 \leq t \leq \frac{1}{2}, \\ \frac{\tilde{C}}{e^{1/(1-t)}} & , \quad \frac{1}{2} < t \leq 1, \end{cases}$$

em que \tilde{C} é convenientemente escolhida para garantir que w seja uma densidade em $[0, 1]$. A rigor, a constante \tilde{C} pode ser encontrada a partir da expressão

$$1 = \int_0^1 w(t) dt = \tilde{C} \int_0^{1/2} e^{-1/t} dt + \tilde{C} \int_{1/2}^1 e^{-1/(1-t)} dt.$$

O uso da substituição $u = \frac{1}{t}$ na primeira integral e da substituição $u = \frac{1}{1-t}$ na segunda integral implicam que

$$\tilde{C} = \frac{1}{2 \int_2^\infty \left(\frac{e^{-u}}{u^2} \right) du}.$$

Capítulo 4

Convergência em distância Mallows de processos empíricos

4.1 Introdução

Neste capítulo, trataremos da convergência em distância Mallows dos processos empíricos. Como vimos no Capítulo 1, o processo empírico $\beta_n(t)$, que é baseado na distribuição empírica F_n , nada mais é que uma soma parcial, na qual as parcelas dependem de funções indicadoras. Por isso, face ao desenvolvimento que dedicamos à convergência na distância Mallows até o Capítulo 3, o interesse nesse processo surge naturalmente. Estabeleceremos as condições sob as quais é possível melhorar o modo de convergência do Lema 1.1, dado por

$$\beta_n(t) \xrightarrow{D} Z(t) \stackrel{d}{=} N(0, F(t)(1 - F(t))). \quad (4.1)$$

O processo de renovação é o inverso do processo de somas parciais e, de maneira similar, a função quantil empírica F_n^- é o inverso da distribuição empírica F_n . Por isso, somas parciais e processos de renovação podem ser utilizados na obtenção de representações eficientes para F_n e F_n^- , conforme podemos ver em detalhes no Capítulo 3 de Csörgö e Horváth (1993). Além disso, muitos resultados assintóticos estabelecem relações entre os processos de renovação e somas parciais por meio de distribuições Gaussianas. Por isso, se estabelecemos qualquer avanço em relação à convergência (4.1), é imediato o questionamento em torno de uma melhora similar em torno de seu processo inverso, $q_n(t)$, que é baseado na função quantil empírica F_n^- . Assim, o mesmo expediente a respeito de $\beta_n(t)$ será adotado com relação ao modo de convergência do Lema 1.3, dado por

$$q_n(t) \xrightarrow{D} \frac{B(t)}{f(F^-(t))}. \quad (4.2)$$

Nosso interesse na convergência em distância Mallows de $q_n(t)$ não se origina apenas em sua relação com $\beta_n(t)$ (o que seria suficiente), mas também no papel que $q_n(t)$ fará em nossa retomada da convergência em distância Mallows ponderada, no Capítulo 5. De fato, naquela parte de nosso trabalho, versaremos sobre o comportamento assintótico da estatística

$$nd_{2,w}^2(F_n, F), \quad (4.3)$$

gerando certa unificação dos temas abordados em nosso trabalho, especialmente porque a estatística (4.3) relaciona-se com o processo quantil por

$$\sqrt{n}d_{2,w}(F_n, F) = \left(\int_0^1 [q_n(t)]^2 w(t) dt \right)^{\frac{1}{2}},$$

conforme justificaremos na introdução de nosso último capítulo.

O estabelecimento da convergência na distância Mallows em (4.1) será mais simples e direto. Em verdade, para a convergência em ordem 2, não exigiremos nenhuma hipótese adicional àquelas necessárias ao Lema 1.1, além de dispor apenas de um de nossos resultados do Capítulo 3. Para o caso da convergência na distância Mallows em (4.2), precisaremos ser um pouco mais exigentes, além de não podermos contar diretamente com os resultados do Capítulo 3.

A propósito da convergência do processo quantil, é relevante citar aqui uma ferramenta estatística que pareceria ser o meio natural de investigação assintótica em torno de $q_n(t)$ e seus momentos, mas não foi usada em nossa pesquisa por limitações técnicas que nos pareceram inconvenientes. Trata-se do *Método Delta*, cujo formato citamos na próxima proposição, em uma de suas mais conhecidas versões.

Proposição 4.1. (Conforme Rohatgi e Saleh, 1976, seção 7.5)

Suponha que g seja uma função real diferenciável, e que $\{Y_n\}_{n \geq 1}$ seja uma sequência de variáveis aleatórias para qual existem um número real μ e uma sequência de números reais positivos σ_n tais que

$$\frac{Y_n - \mu}{\sigma_n} \xrightarrow{D} N(0, 1).$$

Se $\sigma_n \rightarrow 0$ e g é diferenciável em μ , com $g'(\mu) \neq 0$, então

$$\frac{g(Y_n) - g(\mu)}{g'(\mu)\sigma_n} \xrightarrow{D} N(0, 1). \quad (4.4)$$

Além disso,

$$E [g(Y_n) - g(\mu)]^2 \approx [g'(\mu)]^2 E[Y_n - \mu]^2. \quad (4.5)$$

O método delta foi originalmente proposto por Dorfman (1938) sob o pretexto de se conseguir uma variância aproximada para funções não lineares de variáveis aleatórias. Desde então, tem aparecido na literatura estatística também como ferramenta para aproximar os vários momentos de uma variável aleatória, a exemplo de Oehlert (1992); como um meio de correção do viés para a esperança de uma função de variável aleatória, como em Hoef (2012); e como um teorema que estabelece o limite em distribuição para uma função de um estimador assintoticamente normal, conforme seção 5.5 d livro de Casella e Berger (2002).

Suponha que uma função de distribuição F possua derivada positiva. Neste caso, para cada $t_0 \in (0, 1)$, existe um único $x \in \mathbb{R}$ tal que $F(x) = t_0$. O Lema 1.1 garante, então, que

$$\frac{F_n(x) - F(x)}{\sqrt{\frac{F(x) \cdot [1 - F(x)]}{n}}} \xrightarrow{D} N(0, 1),$$

ou

$$\frac{F_n(F^-(t_0)) - t_0}{\sqrt{\frac{t_0(1-t_0)}{n}}} \xrightarrow{D} N(0, 1). \quad (4.6)$$

As identificações $Y_n = F_n(F^-(t_0))$, $\mu = t_0$ e $\sigma_n = \sqrt{\frac{t_0(1-t_0)}{n}}$ implicam, pela Proposição 4.1, em

$$\frac{g(F_n(F^-(t_0))) - g(t_0)}{g'(t_0)\sigma_n} \xrightarrow{D} N(0, 1), \quad (4.7)$$

e

$$E [g(F_n(F^-(t_0))) - g(t_0)]^2 \approx [g'(t_0)]^2 E[F_n(F^-(t_0)) - t_0]^2. \quad (4.8)$$

A escolha dada por

$$\begin{aligned} g: (0, 1) &\rightarrow \mathbb{R} \\ z &\mapsto F^-(z), \end{aligned}$$

poderia diretamente produzir a convergência (4.2) a partir de (4.7) e a respectiva convergência dos momentos a partir de (4.8), bastando, para isso, que seja possível a troca assintótica

$$F^-(F_n(\cdot)) = F_n^-(F(\cdot)). \quad (4.9)$$

A despeito do apelo intuitivo da igualdade (4.9), já que ambas as sequências convergem quase certamente para $F^-[F(x)]$, não nos parece razoável usá-la arbitrariamente, pois tal procedimento teria como base apenas a coincidente convergência quase certa das sequências, o que não é o mesmo que dizer que elas se equivalem. De fato, podemos ver com detalhes, em Doss e Gill (1992), que equivalências assintóticas nesse sentido possuem restrições não elementares. Em todo caso, a convergência (4.2) pode ser estabelecida por outras vias, como o mostra a referência do Lema 1.3. Por isso, a discussão sobre (4.9) pôde ser evitada numa possível reobtenção da convergência em distribuição de $q_n(t)$.

No caso de (4.8), que nos daria a convergência dos momentos em (4.2), entendemos que o problema já se encontra no estabelecimento de (4.5). Vejamos a intuição que conduz à conclusão (4.5) e suas versões não sequenciais: a partir da hipótese da Proposição 4.1, temos que $Y_n - \mu \xrightarrow{q.c.} 0$. O desenvolvimento de Taylor para $g(Y_n)$ em torno de μ é dado por

$$g(Y_n) - g(\mu) = g'(\mu)(Y_n - \mu) + o(|Y_n - \mu|^2).$$

Isso implica, pela elevação ao quadrado, em

$$[g(Y_n) - g(\mu)]^2 = [g'(\mu)(Y_n - \mu)]^2 + 2g'(\mu)(Y_n - \mu)o(|Y_n - \mu|^2) + [o(|Y_n - \mu|^2)]^2.$$

$$\frac{[g(Y_n) - g(\mu)]^2}{[g'(\mu)(Y_n - \mu)]^2} = 1 + \frac{2}{g'(\mu)}(Y_n - \mu)\frac{o(|Y_n - \mu|^2)}{|Y_n - \mu|^2} + \frac{1}{[g'(\mu)]^2}(Y_n - \mu)^2\frac{o(|Y_n - \mu|^4)}{|Y_n - \mu|^4}.$$

$$\frac{[g(Y_n) - g(\mu)]^2}{[g'(\mu)(Y_n - \mu)]^2} = 1 + \frac{2(Y_n - \mu)}{g'(\mu)}o(1) + \frac{(Y_n - \mu)^2}{[g'(\mu)]^2}o(1).$$

Portanto,

$$\lim_{n \rightarrow \infty} \frac{[g(Y_n) - g(\mu)]^2}{[g'(\mu)(Y_n - \mu)]^2} = 1$$

ou

$$[g(Y_n) - g(\mu)]^2 \approx [g'(\mu)]^2(Y_n - \mu)^2.$$

Tomando a esperança dos dois lados, teríamos

$$E[g(Y_n) - g(\mu)]^2 \approx [g'(\mu)]^2 E[Y_n - \mu]^2. \quad (4.10)$$

Imediatamente antes de (4.10), o que se obtém é uma aproximação no sentido da convergência quase certa. Contudo, o que o último passo faz, matematicamente, é concluir a convergência dos momentos a partir disso, o que é equivocado de maneira geral, como mostra o exemplo a seguir.

Exemplo 4.1. Considere $\Omega = \{1, 2, 3, \dots\}$, $\mathcal{F} = \mathcal{P}(\Omega)$, e $P(\{n\}) = \frac{1}{n(n+1)}$. Defina

$$X_n(\omega) = \begin{cases} 2^n & , \omega = n \\ 0 & , \omega \neq n \end{cases}$$

Como $X_n(\omega) = 0$ para n suficientemente grande, segue que $X_n \xrightarrow{n} 0$ q.c.

Mas $E(|X_n - 0|^p) = \frac{2^{np}}{n(n+1)} \xrightarrow{n} \infty$, e não vale a convergência dos momentos.

Em geral, artigos e livros que apresentam ou usam a aproximação (4.5), ou sua versão não sequencial, evitam estabelecer precisamente as hipóteses matemáticas que asseguram tais aproximações dos momentos. Em geral, afirmam que a aproximação torna-se melhor à medida que a função g for mais suave, ou apenas afirmam vagamente que estão sendo assumidas “condições convenientes”, sem especificá-las (veja Casella e Berger, 2002, p. 240; Rohatgi e Saleh, 1976, p. 337; Hoef, 2012). Nesses casos, simulações é que detêm o poder de decisão sobre a eficácia do método.

As melhores exposições encontradas no sentido de estabelecer hipóteses claras sobre g em técnicas de aproximação dos momentos via desenvolvimento de Taylor são aquelas feitas em Oehlert (1992) ou no livro de Crámer (1946). Contudo, essas obras, como os próprios autores esclarecem, não abordam o comportamento assintótico dos momentos em casos como a Proposição 4.1, mas apenas a computação dos momentos de uma função de variável aleatória.

Ainda que impuséssemos alguma hipótese razoável para a ocorrência de (4.10), como, por exemplo, a integrabilidade uniforme das sequências, ainda dependeríamos da troca (4.9) para chegarmos à convergência de momentos em (4.2). Portanto, não usaremos o método delta em nossa investigação sobre a convergência em distância Mallows do processo quantil.

Na Seção 4.2, usaremos o Corolário 3.1, que estabelece a convergência de segunda ordem d_2 em caso Gaussiano (Teorema do Limite Central), para concluir a convergência em distância Mallows de ordem 2 no lugar da convergência em distribuição de (4.1). Em seguida, abordaremos o mesmo modo de convergência em ordens superiores.

Na Seção 4.3, estabeleceremos a convergência em distância Mallows de ordem 1 no lugar da convergência em distribuição de (4.2) de uma forma mais simples e direta. Contudo, tal metodologia gerará um resultado menos robusto do que seu equivalente da Seção 4.2, por exigir a finitude do primeiro momento de F .

Por fim, na Seção 4.4, trataremos da convergência em distância Mallows de $q_n(t)$ para ordens superiores e sem a restrição sobre os momentos de F . Um dos resultados obtidos nessa seção será de muita valia no próximo capítulo, onde retomaremos a convergência em distância Mallows ponderada em associação com os processos empíricos.

4.2 Processo empírico geral

Nesta seção, provaremos a convergência em distância Mallows de ordem 2 do processo empírico $\beta_n(t)$ para a variável aleatória $Z(t) \stackrel{d}{=} N(0, F(t)(1 - F(t)))$, usando o Corolário 3.1. Em seguida, mostraremos que essa convergência ocorre para qualquer ordem $1 \leq r \leq 5$. Finalmente, estabeleceremos uma hipótese suficiente para que a convergência em distância Mallows aconteça em qualquer ordem maior que 5.

No teorema a seguir, usamos o Corolário 3.1 para melhorarmos a convergência do Lema 1.1 para convergência em distância Mallows sem nenhuma hipótese adicional.

Teorema 4.1. *Seja F uma função de distribuição e $Z(t) \stackrel{d}{=} N(0, F(t)(1 - F(t)))$. Se $t \in \mathbb{R}$ é tal que $0 < F(t) < 1$, então*

$$d_2(\beta_n(t), Z(t)) \xrightarrow{n} 0. \quad (4.11)$$

Demonstração:

Considere $\{Y_n\}_{n \geq 1}$ uma amostra aleatória de F . Para um fixado $t \in \mathbb{R}$, sabemos que

$$\begin{aligned} \beta_n(t) &= \sqrt{n} \left(\frac{\mathbb{I}(Y_1 \leq t) + \mathbb{I}(Y_2 \leq t) \dots + \mathbb{I}(Y_n \leq t)}{n} - F(t) \right) \\ &= \sqrt{n} \left(\frac{\mathbb{I}(Y_1 \leq t) - F(t) + \mathbb{I}(Y_2 \leq t) - F(t) + \dots + \mathbb{I}(Y_n \leq t) - F(t)}{n} \right) \\ &= \frac{X_1 + X_2 \dots + X_n}{\sqrt{n}}. \end{aligned}$$

Na última igualdade, foi usada a identificação $X_i = \mathbb{I}(Y_i \leq t) - F(t)$, para $i = 1, 2, \dots, n$. A sequência $\{X_n\}_{n \geq 1}$ é constituída, naturalmente, de variáveis aleatórias independentes e identicamente distribuídas. Além disso, temos que

$$E(X_1) = E(\mathbb{I}(Y_1 \leq t) - F(t)) = 0,$$

$$\sigma_t^2 = Var(X_1) = Var(\mathbb{I}(Y_1 \leq t) - F(t)) = Var(\mathbb{I}(Y_1 \leq t)) = F(t)[1 - F(t)].$$

O Teorema do Limite Central assegura, nesse caso, que

$$\frac{X_1 + X_2 \dots + X_n}{\sqrt{n}} \xrightarrow{D} N(0, \sigma_t). \quad (4.12)$$

Portanto, temos satisfeitas as hipóteses do Corolário 3.1, razão pela qual podemos concluir que

$$d_2(\beta_n(t), Z(t)) \xrightarrow{n} 0.$$

□

Sob as mesmas hipóteses do Teorema 4.1, podemos provar a convergência descrita em (4.11) para distância Mallows de ordens superiores. Contudo, tal prova não utiliza o Corolário 3.1 (nem outros resultados do Capítulo 3), já que ele limita-se à ordem 2. Ainda assim, nossa abordagem mimetiza as estratégias que permearam as demonstrações dos capítulos anteriores, que usaram fortemente a integrabilidade uniforme e a Proposição 1.2. Estabeleceremos a convergência em distância Mallows de $\beta_n(t)$ em ordens superiores a partir de uma condição suficiente para certa integrabilidade uniforme. Antes disso, exibiremos um teorema que garante a convergência (4.11) para ordem 5 de forma direta.

Teorema 4.2. *Seja F uma função de distribuição. Se $t \in \mathbb{R}$ é tal que $0 < F(t) < 1$, então*

$$d_5(\beta_n(t), Z(t)) \xrightarrow{n} 0. \quad (4.13)$$

Demonstração:

A condição $0 < F(t) < 1$ assegura, pelo Lema 1.1, que

$$\beta_n(t) \xrightarrow{D} Z(t),$$

com $Z(t) \stackrel{d}{=} N(0, F(t)(1 - F(t)))$. Esse fato, aliado à Proposição 1.2, indica que (4.13) será verdadeira se pudermos provar a convergência

$$E[\beta_n(t)]^5 \xrightarrow{n} E[Z(t)]^5. \quad (4.14)$$

Vamos provar (4.14) através da função geradora de momentos. Se considerarmos $c = F(t)$, então a função geradora de momentos de $Z(t)$ será dada por

$$M_z(x) = \exp\left(\frac{c(1-c)x^2}{2}\right), \quad x \in \mathbb{R}.$$

Essa função é infinitamente derivável e bem definida em qualquer vizinhança da origem, de maneira que os primeiros cinco momentos de $Z(t)$ podem ser computados como

$$E(Z(t))^k = M_z^{(k)}(0) = \begin{cases} 0 & , \quad k = 1, 3, 5 \\ (k-1)[c(1-c)]^{k/2} & , \quad k = 2, 4. \end{cases}$$

Para um número real t fixo, o processo empírico $\beta_n(t)$ é uma soma de variáveis aleatórias independentes e identicamente distribuídas, isto é,

$$\beta_n(t) = \sqrt{n} \left\{ \frac{T_1}{n} + \dots + \frac{T_n}{n} - c \right\}, \quad T_i \stackrel{d}{=} \text{Bernoulli}(c).$$

Por isso, a função geradora de momentos de $\beta_n(t)$ é descrita por

$$M_{\beta_n}(x) = P(x) \cdot Q(x),$$

em que

$$P(x) = \exp(-\sqrt{nc}x)$$

e

$$Q(x) = (1 - c + c \exp(x/\sqrt{n}))^n.$$

Aqui, omitimos t por simplicidade de notação. Algumas derivações mostram que as derivadas de Q são dadas por

$$P^{(k)}(0) = (-1)^k n^{k/2} c^k, \quad k = 0, 1, 2, \dots$$

Sucessivas derivações pela Regra da Cadeia e pela Regra do produto mostram que

$$Q^{(1)}(0) = c\sqrt{n},$$

$$Q^{(2)}(0) = c + (n-1)c^2,$$

$$Q^{(3)}(0) = \frac{c}{\sqrt{n}} + c^2 n^{3/2} \left[\frac{3}{n} - \frac{3}{n^2} + c - \frac{3c}{n} + \frac{2c}{n^2} \right],$$

$$Q^{(4)}(0) = \frac{c}{n} + 7c^2 \left[1 - \frac{1}{n} \right] + 6c^3 \left[n - 3 + \frac{2}{n} \right] + c^4 (n-2)(n-3) \left[1 - \frac{1}{n} \right],$$

$$\begin{aligned} Q^{(5)}(0) &= \frac{c}{n^{3/2}} + 15c^2 \left(\frac{1}{\sqrt{n}} - \frac{1}{n^{3/2}} \right) + 25c^3 \left(\sqrt{n} - \frac{3}{\sqrt{n}} + \frac{2}{n^{3/2}} \right) \\ &\quad + 10c^4 \left(n^{3/2} - 6\sqrt{n} + \frac{11}{\sqrt{n}} - \frac{6}{n^{3/2}} \right) \\ &\quad + c^5 \left(n^{5/2} - 10n^{3/2} + 35\sqrt{n} - \frac{50}{\sqrt{n}} + \frac{24}{n^{3/2}} \right). \end{aligned}$$

A regra da derivação para um produto de funções implica que

$$\frac{d^k}{dx^k} M_{\beta_n}(0) = \sum_{i=0}^k \binom{k}{k-i} P^{(i)}(0) \cdot Q^{(k-i)}(0),$$

de maneira que

$$\begin{aligned} M_{\beta_n}^{(5)}(0) &= P(0)Q^{(5)}(0) + 5P^{(1)}(0)Q^{(4)}(0) + 10P^{(2)}(0)Q^{(3)}(0) \\ &\quad + 10P^{(3)}(0)Q^{(2)}(0) + 5P^{(4)}(0)Q^{(1)}(0) + P^{(5)}(0)Q(0). \end{aligned}$$

A substituição dos valores acima, a computação dos produtos e a realização de todas as simplificações possíveis conduzem à seguinte expressão.

$$\begin{aligned} M_{\beta_n}^{(5)}(0) &= \frac{c}{n^{3/2}} + 15c^2 \left(\frac{1}{\sqrt{n}} - \frac{1}{n^{3/2}} \right) + 25c^3 \left(\frac{2}{n^{3/2}} - \frac{3}{\sqrt{n}} \right) \\ &\quad - \frac{5c^3}{n} \left(\frac{1}{c} - 7 + 12c - 6c^2 \right) + 10c^4 \left(\frac{11}{\sqrt{n}} - \frac{6}{n^{3/2}} \right) + 2c^5 \left(\frac{12}{n^{3/2}} - \frac{25}{\sqrt{n}} \right). \end{aligned}$$

Assim, vale a convergência (4.14), pois

$$E[\beta_n(t)]^5 = M_{\beta_n}^{(5)}(0) \xrightarrow{n} 0 = M_z^{(5)}(0) = E(Z(t))^5.$$

Agora, a aplicação da Proposição 1.2 é suficiente para a conclusão do Teorema. \square

Um corolário imediato do Teorema 4.2 é a convergência de $\beta_n(t)$, na distância Mallows de ordem $r \in [1, 5]$, para uma distribuição normal, como enunciaremos a seguir.

Corolário 4.1. *Sob as mesmas condições do Teorema 4.2, é verdade que, para cada $1 \leq r \leq 5$,*

$$d_r(\beta_n(t), Z(t)) \xrightarrow{n} 0. \quad (4.15)$$

Demonstração:

A desigualdade de Minkowski e a representação (1.12) são ferramentas suficientes para assegurarmos que

$$0 \leq d_r(\beta_n(t), Z(t)) \leq d_5(\beta_n(t), Z(t)) \xrightarrow{n} 0,$$

o que atesta nosso corolário. \square

Os aspectos computacionais indigestos das derivadas da função geradora de momentos do processo $\beta_n(t)$ impedem, por ora, o estabelecimento do Corolário 4.1 para $r > 5$ apenas sob a hipótese $0 < F(t) < 1$. Ainda que essa ideia seja intuitivamente válida, não podemos garantir a convergência de momentos em (4.14) para $r > 5$ apenas com base na existência dos momentos de $\beta_n(t)$, conforme ilustra o exemplo a seguir.

Exemplo 4.2. *Seja $X \stackrel{d}{=} N(0, 1)$ e $\{Y_n\}_{n \geq 1}$ uma sequência de variáveis aleatórias independentes de X tais que*

$$P(Y_n = x) = \begin{cases} \frac{1}{n} & , \quad x = \sqrt{n}, \\ 0 & , \quad x \neq \sqrt{n}. \end{cases}$$

Se considerarmos $X_n = X + Y_n$, então $X_n \xrightarrow{D} X$. Note que, assim como seu limite em distribuição, X_n possui todos os momentos. Pela independência, podemos concluir que $E(X_n^2) = 2$ para cada $n \geq 1$. Contudo, $E(X^2) = 1$, o que mostra claramente que não temos a convergência do segundo momento.

O Exemplo 4.2 também é uma evidência contra uma suposta estratégia de se conseguir a convergência (4.14) via processo de indução para os números inteiros positivos, pois, apesar de $E(X_n) \xrightarrow{n} E(X)$, não vale a convergência do próximo momento inteiro.

Em nosso próximo resultado, estabeleceremos uma condição suficiente para que a convergência (4.15) ocorra para qualquer ordem $r > 5$. Tal condição permitirá, assim

como fizemos nos dois capítulos anteriores, o uso da integrabilidade uniforme como ferramenta de obtenção da convergência em distância Mallows.

A seguir, a exemplo do que fizemos na demonstração do Teorema 4.2 para t fixo, representaremos a j -ésima derivada da função geradora de momentos de $\beta_n(t)$ por

$$M_{\beta_n(t)}^{(j)}(x),$$

em que $x \in \mathbb{R}$, e $j \in \mathbb{N}$. Para $r \in \mathbb{R}$, o símbolo $[r]$ denotará o menor inteiro maior ou igual a r .

Teorema 4.3. *Suponha que $r, t \in \mathbb{R}$ são tais que $r > 5$ e $0 < F(t) < 1$. Se vale a condição*

$$\sup_n M_{\beta_n(t)}^{([r]+1)}(0) < \infty, \tag{4.16}$$

então

$$d_r(\beta_n(t), Z(t)) \xrightarrow{n} 0.$$

Demonstração:

Por simplicidade, omitiremos a letra t durante a demonstração. Nesse caso, devemos provar que $d_r(\beta_n, Z) \xrightarrow{n} 0$. Suponha que $r \in \mathbb{N}$. Sob as hipóteses do Teorema, já sabemos que $\beta_n \xrightarrow{D} Z$. Isso implica, pelo Teorema do Mapeamento, que

$$\beta_n^r \xrightarrow{D} Z^r. \tag{4.17}$$

Além disso, a condição (4.16) implica na integrabilidade uniforme de β_n^r , já que

$$\int_{\{|\beta_n|^r \geq \alpha\}} |\beta_n|^r dP \leq \frac{1}{\alpha} E \{|\beta_n|^{r+1}\} = \frac{1}{\alpha} M_{\beta_n}^{(r+1)}(0),$$

e, por isso,

$$\lim_{\alpha \rightarrow \infty} \sup_n \int_{\{|\beta_n|^r \geq \alpha\}} |\beta_n|^r dP = 0.$$

Segundo o Teorema 1.5, a integrabilidade uniforme de β_n^r e a convergência (4.17) bastam para que valha a convergência

$$E(\beta_n^r) \xrightarrow{n} E(Z^r). \tag{4.18}$$

Pela Proposição 1.2, as convergências (4.17) e (4.18) garantem nosso Teorema.

Para o caso em que $r \notin \mathbb{N}$, a condição (4.16) e os mesmos passos de antes, asseguram que

$$d_{[r]}(\beta_n(t), Z(t)) \xrightarrow{n} 0.$$

Pela Desigualdade de Minkowski e pela representação (1.12), teremos diretamente que

$$d_r(\beta_n(t), Z(t)) \xrightarrow{n} 0.$$

□

Observação 4.1. *Para garantirmos a integrabilidade uniforme na demonstração do Teorema 4.3, é suficiente que a condição (4.16) seja substituída por*

$$\sup_n E \{ |\beta_n(t)|^{r+\varepsilon} \} < \infty$$

para algum $\varepsilon > 0$. Optamos pela condição descrita no Teorema com um expoente inteiro, porque a análise da condição via geradora de momentos da variável $\beta_n(t)$ parece mais factível, de maneira geral.

Na próxima seção, voltaremos nossa atenção para a convergência em distância Mallows do processo quantil. Veremos que a obtenção de uma convergência mais forte que a do Lema 1.3, por vias diretas, requer uma hipótese mais exigente e, adicionalmente, gera um resultado um pouco restritivo.

4.3 Processo quantil - Convergência de ordem 1

Nesta seção, mostraremos que o processo quantil converge, em algum ponto de $(0, 1)$, na distância Mallows de ordem 1 para uma variável aleatória Gaussiana. Vimos, no primeiro Capítulo, que o processo quantil é definido com base na inversa generalizada da distribuição F , a saber,

$$\{q_n(t), t \in (0, 1)\} = \{\sqrt{n} (F_n^-(t) - F^-(t)), t \in (0, 1)\}, \quad n \geq 1.$$

Além disso, se for válida a diferenciabilidade da distribuição F com densidade positiva, o Lema 1.3 garante a convergência em distribuição de $q_n(t)$ para uma variável Gaussiana, isto é,

$$q_n(t) \xrightarrow{D} \frac{B(t)}{f(F^-(t))}, \quad (4.19)$$

em que $B(t) \stackrel{d}{=} N(0, t(1-t))$. A fim de provarmos por métodos diretos que a convergência acima vale também em distância Mallows, passamos pela necessidade de trabalhar com uma expressão mais simples para $F_n^-(t)$, do que aquela apresentada na Definição 1.7. Usaremos aqui a representação dada na igualdade (1.7), ou seja,

$$F_n^-(t) = \begin{cases} X_{1:n} & , \quad t = 0, \\ X_{k:n} & , \quad \frac{k-1}{n} < t < \frac{k}{n}. \end{cases}$$

Na expressão acima, $X_{k:n}$ é a k -ésima estatística de ordem de uma amostra de F com n elementos. Através da distribuição associada a k -ésima estatística de ordem (veja Galambos, 1987, Seção 2.8, ou Rohatgi e Saleh, 1976, Seção 4.7), podemos computar o primeiro momento de $X_{k:n}$ como

$$E(X_{k:n}) = \int_{-\infty}^{\infty} x \frac{n!}{(k-1)!(n-k)!} [F(x)]^{k-1} [1-F(x)]^{n-k} f(x) dx, \quad 1 \leq k \leq n. \quad (4.20)$$

Essa representação se vale do fato de que as estatísticas de ordem se referem a uma amostra de F , isto é, X_1, X_2, \dots, X_n são independentes e identicamente distribuídas

com F . A igualdade acima pressupõe a finitude do primeiro momento associado à distribuição F . De maneira geral, se $X \sim F$, o primeiro momento de $g(X_{k:n})$ pode ser computado no formato acima, bastando que g seja mensurável e que $g(X)$ seja finito. Sob as hipóteses do Lema 1.3 e a finitude do primeiro momento associado à distribuição F , mostraremos agora que a convergência em (4.19) pode ser melhorada para a convergência em distância Mallows de ordem 1, pelo menos em algum ponto do intervalo $(0, 1)$. No enunciado a seguir, considere a identificação

$$B_t = \frac{B(t)}{f(F^-(t))} \stackrel{d}{=} N\left(0, \frac{t(1-t)}{f^2(F^-(t))}\right). \quad (4.21)$$

Teorema 4.4. *Se F é uma função de distribuição com primeiro momento finito e densidade positiva, então existe $y \in (0, 1)$ tal que*

$$d_1(q_n(y), B_y) \xrightarrow{n} 0. \quad (4.22)$$

Demonstração:

Como a distribuição F possui densidade positiva em todo seu domínio e a função $g(x) = x - F^-(t)$ é mensurável para cada $0 < t < 1$, podemos usar a igualdade (4.20) para escrever

$$E(X_{k:n} - F^-(t)) = \int_0^1 [F^-(z) - F^-(t)] \frac{n!}{(k-1)!(n-k)!} z^{k-1} [1-z]^{n-k} dz,$$

o que nos leva a

$$E(q_n(t)) = \int_0^1 [F^-(z) - F^-(t)] \frac{n! \sqrt{n}}{(k-1)!(n-k)!} z^{k-1} [1-z]^{n-k} dz.$$

Para estudar o comportamento da expressão acima (quando $k < n$), vamos verificar a convergência da série de potências dada por

$$\sum_{n>k}^{\infty} \frac{n! \sqrt{n}}{(n-k)!} (1-z)^{n-k}. \quad (4.23)$$

Usando o Teste da Razão para séries de potências, vemos que

$$\frac{|a_{n+1}|}{|a_n|} = \frac{\left| \frac{(n+1)! \sqrt{n+1}}{(n-k+1)!} (1-z)^{n-k+1} \right|}{\left| \frac{n! \sqrt{n}}{(n-k)!} (1-z)^{n-k} \right|} = \frac{1}{\left(\frac{n-k}{n+1} + \frac{1}{n+1} \right)} \cdot |1-z| \sqrt{1 + \frac{1}{n}}.$$

Como $|1-z| < 1$, segue que

$$\lim_{n \rightarrow \infty} \frac{|a_{n+1}|}{|a_n|} = 0 < 1,$$

e a série em (4.23) é convergente. Por isso, seu termo geral tende a zero quando n cresce, o que significa dizer que

$$E(q_n(t)) = \int_0^1 M_n [F^-(z) - F^-(t)] dz,$$

em que $M_n \xrightarrow{n} 0$, já que $0 < z^{k-1} \leq 1$. De fato, a convergência de M_n não depende de z , pois as hipóteses e a mudança de variável que fizemos impõem a localização de z no intervalo aberto $(0, 1)$.

No caso em que $k = n$, a igualdade (4.20) e o uso da função mensurável dada por $g(x) = x - F^-(t)$ garantem diretamente que

$$E(q_n(t)) = \int_0^1 n z^{n-1} \sqrt{n} [F^-(z) - F^-(t)] dz = \int_0^1 M_n [F^-(z) - F^-(t)] dz.$$

Esta relação vale para qualquer $t \in (0, 1)$. Naturalmente a sequência M_n não é a mesma nas duas esperanças acima, mas usamos a mesma notação por simplicidade, já que a única informação importante sobre M_n , nos dois casos, é sua convergência.

Cabe ressaltar agora que é sempre possível escolher $\varepsilon > 0$ tal que $\varepsilon < t < 1 - \varepsilon$, com $\varepsilon < \frac{1}{2}$, seja qual for o número $t \in (0, 1)$. Como F é contínua com densidade positiva, sua inversa também é contínua. Pelo Teorema do Valor Médio para integrais, aplicado a F^- no intervalo fechado $[\varepsilon, 1 - \varepsilon]$, podemos garantir que existe pelo menos um número $y \in (\varepsilon, 1 - \varepsilon)$ tal que

$$\int_0^1 F^-(z) dz = F^-(y)(1 - \varepsilon - \varepsilon) = (1 - 2\varepsilon) \cdot F^-(y).$$

Portanto,

$$E(q_n(y)) = \int_0^1 M_n [F^-(z) - F^-(y)] dz = -2\varepsilon \cdot F^-(y) \cdot M_n,$$

$$\lim_{n \rightarrow \infty} E(|q_n(y)|) = 2\varepsilon \cdot F^-(y) \lim_{n \rightarrow \infty} |M_n| = 0,$$

ou

$$\lim_{n \rightarrow \infty} E(|q_n(y)|) = E(B_y).$$

A relação acima e a convergência em distribuição de (4.19) asseguram, pela Proposição 1.2, a convergência em distância Mallows disposta em (4.22).

□

Observação 4.2. *Ao contrário do que aconteceu com os Teoremas 4.1 e 4.2 na obtenção de um modo de convergência mais forte em torno do processo empírico geral $\beta_n(t)$, o Teorema 4.4 não se traduz num aperfeiçoamento completo do Lema 1.3 para o processo quantil, no sentido de que não garante a convergência proposta para qualquer ponto do intervalo $(0, 1)$. Qualquer abordagem das possíveis convergências em distância Mallows de ordens superiores que fizer uso de representações como aquela que se vê em (4.20) não poderá prescindir da finitude dos momentos nem gerará uma convergência que valha para todo o intervalo $(0, 1)$, já que a conclusão estará restrita às garantias do Teorema do Valor Médio para integrais.*

Tal quadro constitui-se numa boa motivação para fazermos nossa abordagem por outra metodologia, a fim de que possamos garantir a convergência em distância Mallows para todo o intervalo $(0, 1)$.

Na próxima seção, exploraremos a relação entre o processo quantil $q_n(t)$ e seu caso particular F uniforme, a fim de poder transportar para o processo quantil geral vantagens distribucionais de seu caso uniforme. Essa abordagem tem sido a chave no desenvolvimento de teorias assintóticas em torno dos processos empíricos, conforme podemos observar claramente na obra de Csörgo e Horváth (1993), só para citar um canône desse campo de estudo.

4.4 Processo quantil - Convergência de ordens superiores

Nesta seção, provaremos a convergência em distância Mallows de ordem 2 do processo quantil $q_n(t)$ para a variável aleatória Gaussiana

$$B_t = \frac{B(t)}{f(F^-(t))} \stackrel{d}{=} N\left(0, \frac{t(1-t)}{f^2(F^-(t))}\right),$$

para cada $t \in (0, 1)$. Isso representará um avanço em relação ao Teorema 4.4, pois aquele resultado garante essa convergência apenas para algum ponto $y \in (0, 1)$, somente para a ordem 1 e sob a hipótese de finitude do primeiro momento, que dispensaremos aqui.

A seguir, exploraremos a relação entre o processo quantil e o caso uniforme do processo empírico, pois ela será fundamental na obtenção de nosso resultado.

Considere o caso uniforme do processo empírico geral, isto é,

$$\{\alpha_n(t), t \in [0, 1]\} = \{\sqrt{n}(U_n(t) - t), x \in \mathbb{R}\}, \quad n \geq 1. \quad (4.24)$$

Neste caso, U_n representa a distribuição empírica uniforme. O processo empírico $\alpha_n(t)$ possui uma teoria assintótica bem desenvolvida na literatura em geral (veja, por exemplo, Shorack e Wellner, 1986, ou Csörgo e Horváth, 1993). Na prática, sua simplicidade foi a chave para o desenvolvimento satisfatório do estudo voltado ao processo $\beta_n(t)$, quando a distribuição F é contínua. Isso foi possível, porque $F(X)$ possui distribuição uniforme sempre que X é uma variável aleatória com distribuição contínua F e, conseqüentemente, as estatísticas $X_{1:n}, \dots, X_{n:n}$ da amostra X_1, \dots, X_n induzem as estatísticas $U_{1:n} = F(X_{1:n}), \dots, U_{n:n} = F(X_{n:n})$ da amostra $U_1 = F(X_1), \dots, U_n = F(X_n)$.

As representações alternativas para a função de distribuição empírica e para a função quantil empírica que foram apresentadas em (1.3) e (1.7) geram as relações

$$\beta_n(x) = \alpha_n(F(x)),$$

$$\beta_n(F^-(t)) = \alpha_n(t).$$

A função quantil empírica uniforme pode ser representada, segundo a linha de raciocínio acima, por

$$U_n^-(t) = F(F_n^-(t)), \quad 0 \leq t \leq 1. \quad (4.25)$$

Contudo, a relação com o processo quantil geral $q_n(t)$ e o conseqüente aproveitamento de resultados teóricos não é tão imediata. De fato, se usarmos a representação (4.25) e o Teorema do Valor Médio, veremos que, para cada $t \in (0, 1)$,

$$\begin{aligned} q_n(t) &= \sqrt{n} (F_n^-(t) - F^-(t)) \\ q_n(t) &= \sqrt{n} (F^-(U_n^-(t)) - F^-(t)) \\ q_n(t) &= \sqrt{n} (U_n^-(t) - t) \cdot [F^-(\xi_n)]' \\ q_n(t) &= \sqrt{n} (U_{k:n} - t) \cdot \frac{1}{f(F^-(\xi_n))}, \end{aligned} \quad (4.26)$$

em que $t \wedge U_n^-(t) \leq \xi_n \leq t \vee U_n^-(t)$ e $\frac{(k-1)}{n} < t \leq \frac{k}{n}$, $1 \leq k \leq n$. A relação acima exige que a distribuição F seja absolutamente contínua com densidade estritamente positiva. Essa imposição já é necessária à convergência em distribuição de $q_n(t)$ para B_t e, naturalmente, para a convergência em distância Mallows de ordem 1 que provamos no Teorema 4.4.

Note que a positividade estrita da densidade ainda permite que (4.26) seja escrita da seguinte forma:

$$q_n(t) = \frac{1}{f(F^-(t))} \sqrt{n} (U_{k:n} - t) \cdot \frac{f(F^-(t))}{f(F^-(\xi_n))}. \quad (4.27)$$

Dentro da teoria dos processos empíricos, a limitação uniforme da taxa

$$\frac{f(F^-(t))}{f(F^-(\xi_n))} \quad (4.28)$$

é uma forma conveniente de se prover resultados para $q_n(t)$ que já valham para o caso F uniforme, como bem ilustra a exposição do capítulo 6 de Csörgo e Horváth (1993). No próximo capítulo, abordaremos as condições naturais nas quais tal limitação uniforme pode ocorrer. Para estabelecer uma convergência em distância Mallows de ordem 2 de $q_n(t)$ que dispense a finitude de momentos de F , acrescentaremos uma hipótese de limitação uniforme sobre o denominador de (4.26) ou sobre a taxa (4.28). No enunciado a seguir, a medida de probabilidade P é aquela do espaço de probabilidade onde estão definidas as variáveis aleatórias da amostra geradora dos processos quantis envolvidos.

Lema 4.1. *Seja F uma distribuição contínua com derivada positiva. Suponha que*

$$\inf_{0 < t < 1} f(F^-(t)) > 0, \quad (4.29)$$

ou

$$\sup_{\frac{1}{n} \leq t \leq 1 - \frac{1}{n}} \frac{f(F^-(t))}{f(F^-(\xi_n))} = O_p(1), \quad (4.30)$$

para $t \wedge U_n^-(t) \leq \xi_n \leq t \vee U_n^-(t)$.

Neste caso, para cada $t \in (0, 1)$, temos que

$$E[q_n(t)]^2 \xrightarrow{n} E[B_t]^2. \quad (4.31)$$

Demonstração:

Já sabemos do Lema 1.3 e do Teorema do Mapeamento que

$$[q_n(t)]^2 \xrightarrow{D} (B_t)^2. \quad (4.32)$$

Para $1 \leq k \leq n$, a distribuição associada à estatística de ordem $U_{k:n}$ possui terceiro momento finito, o que nos permite reescrever a igualdade (4.20) como

$$E[\sqrt{n}(U_{k:n} - t)]^3 = \int_0^1 (z - t)^3 \frac{n!n^{3/2}}{(k-1)!(n-k)!} z^{k-1}(1-z)^{n-k} dz,$$

pela mensurabilidade da função cúbica. Uma análise exatamente igual àquela feita na demonstração do Teorema 4.4 mostra que

$$E[\sqrt{n}(U_{k:n} - t)]^3 = \int_0^1 (z - t)^3 M_n dz, \quad (4.33)$$

com M_n convergindo a zero independentemente de z . A limitação óbvia da integral acima implica dizer que

$$E[\sqrt{n}(U_{k:n} - t)]^3 < \infty, \quad (4.34)$$

para cada $n \geq 1$. A hipótese descrita em (4.29) garante que existe $\delta > 0$ tal que $f(F^-(\xi_n)) > \frac{1}{\delta}$ para cada $\xi_n \in (t \wedge U_n^-(t), t \vee U_n^-(t))$, isto é

$$\frac{1}{f^3(F^-(\xi_n))} < \delta^3.$$

A desigualdade acima, combinada com (4.34) e com a relação (4.26), nos assegura que

$$\sup_n E\{|q_n(t)|^3\} < \infty, \quad (4.35)$$

de forma que

$$\lim_{a \rightarrow \infty} \sup_n \int_{\{|q_n(t)|^2 \geq a\}} |q_n(t)|^2 dP \leq \lim_{a \rightarrow \infty} \sup_n \frac{E\{|q_n(t)|^3\}}{a} = \lim_{a \rightarrow \infty} \frac{\sup_n E\{|q_n(t)|^3\}}{a} = 0. \quad (4.36)$$

Em outras palavras, temos assegurada a integrabilidade uniforme de $[q_n(t)]^2$.

Por outro lado, a limitação uniforme indicada em (4.30) e a expressão (4.27) também garantem a relação (4.35). Consequentemente, fica válida, também por essa via, a convergência (4.36).

A integrabilidade uniforme indicada em (4.36) e a convergência (4.32) garantem, segundo o Teorema 1.5, que

$$E[q_n(t)]^2 \xrightarrow{n} E[B_t]^2.$$

□

Teorema 4.5. *Seja F uma distribuição contínua com derivada positiva. Se vale uma das condições do Lema 4.1, então*

$$d_2(q_n(t), B_t) \xrightarrow{n} 0, \quad (4.37)$$

para cada $t \in (0, 1)$.

Demonstração:

Para cada $t \in (0, 1)$, vale que

$$[q_n(t)]^2 \xrightarrow{D} (B_t)^2, \quad (4.38)$$

pelo Lema 1.3 e pelo Teorema do Mapeamento. Pelo Lema 4.1, sabemos que

$$E[q_n(t)]^2 \xrightarrow{n} E[B_t]^2. \quad (4.39)$$

Portanto, a Proposição 1.2 garante o resultado. □

Observação 4.3. *A Desigualdade de Minkowski mostra que o Teorema 4.5 representa uma evolução ao Teorema 4.4, ao implicar também na convergência em distância Mallows de ordem 1. Mais que isso, o presente Teorema não só garante a mesma convergência na ordem 2 para qualquer ponto de $(0, 1)$, como também não impõe a finitude do segundo momento de F , já que a limitação nesse sentido se transfere para a distribuição uniforme, que a satisfaz sem nenhuma hipótese adicional sobre F .*

O Teste da Razão para séries de potências mostra que, a exemplo do que vimos na demonstração do Teorema 4.4, para cada $r \in \mathbb{R}$, vale a convergência de

$$\sum_{n>k}^{\infty} \frac{n!n^{r/2}}{(n-k)!} (1-z)^{n-k} \quad (4.40)$$

para qualquer $z \in (0, 1)$. Portanto, a sequência

$$R_n = \frac{n!n^{r/2}}{(n-k)!} (1-z)^{n-k}$$

converge para zero independentemente de $z \in (0, 1)$. Por isso, analogamente à igualdade (4.33), podemos escrever

$$E[\sqrt{n}(U_{k:n} - t)]^{r+1} = \int_0^1 (z-t)^{r+1} R_n dz. \quad (4.41)$$

Estas observações permitem-nos estabelecer uma consequência imediata ao Teorema 4.5.

Corolário 4.2. *Seja F uma distribuição contínua com derivada positiva. Se é válida uma das hipóteses do Lema 4.1, então*

$$d_r(q_n(t), B_t) \xrightarrow{n} 0, \quad (4.42)$$

para cada $t \in (0, 1)$ e para cada $r > 2$.

Demonstração:

Do Lema 1.3 e do Teorema do Mapeamento, temos que

$$q_n^r(t) \xrightarrow{D} B_t^r.$$

A integrabilidade uniforme da sequência $\{q_n^r(t)\}_n$ segue imediatamente da relação (4.41). Por isso, o Teorema 1.5 garante a convergência dos momentos e a convergência (4.42) fica estabelecida como consequência da Proposição 1.2.

□

A ausência de uma imposição sobre os momentos da distribuição F no Lema 4.1 permitirá, no próximo capítulo, o uso desse resultado na produção da convergência em distância Mallows de um funcional ponderado do processo quantil quando $q_n(t)$ é gerado a partir de classes de distribuições mais restritivas quanto aos momentos, como a classe das distribuições estáveis.

Capítulo 5

Distância Mallows ponderada estabilizada

5.1 Introdução

O processo quantil

$$q_n(t) = \sqrt{n} (F_n^-(t) - F^-(t)), \quad t \in (0, 1), \quad (5.1)$$

é o componente básico de muitas estatísticas interessantes, algumas das quais foram abordadas na introdução desta tese. Uma quantidade substancial de obras se dedica a essa abordagem. Referências fundamentais nesse sentido são os textos de Shorack e Wellner (1986) e de Csörgo e Horváth (1993). Uma estatística derivada de (5.1) é o segundo momento de $q_n(t)$ com respeito à medida de Lebesgue, e, em geral, com respeito a qualquer medida com densidade $w(t)$ em $(0, 1)$, dado por

$$\|q_n\|_{2,w} := \left(\int_0^1 [q_n(t)]^2 w(t) dt \right)^{1/2}. \quad (5.2)$$

O funcional (5.2) é o principal componente de estatísticas usadas, por exemplo, em testes de ajuste baseados no coeficiente de correlação, como em Lockhart e Stephens (1998), e testes de ajuste baseados na distância Mallows, como em Barrio et al. (1999). Tais classes de testes contêm importantes membros, como os testes de normalidade de Shapiro-Wilk.

A expressão da distância Mallows ponderada de ordem 2 entre a função de distribuição F e sua empírica associada F_n é dada, conforme a Definição 2.2, por

$$d_{2,w}(F_n, F) = \left(\int_0^1 |F_n^-(t) - F^-(t)|^2 w(t) dt \right)^{1/2}. \quad (5.3)$$

Claramente, (5.2) e (5.3) estão relacionadas por

$$\sqrt{n} d_{2,w}(F_n, F) = \|q_n\|_{2,w},$$

ou, alternativamente,

$$n d_{2,w}^2(F_n, F) = \|q_n\|_{2,w}^2. \quad (5.4)$$

A convergência em distribuição da estatística $nd_{2,w}^2(F_n, F)$ tem sido objeto de interesse em muitos trabalhos recentes, especialmente, porque sua convergência permite a proposição de testes de ajuste para famílias de escala e locação geradas por uma distribuição especificada. Uma descrição teórica completa da convergência em distribuição de $\|q_n\|_{2,w}^2$ e algumas de suas variações pode ser encontrada em del Barrio et al. (2005), sob a imposição do segundo momento finito de F . Por outro lado, podemos encontrar, em Dorea e Lopes (2014), a proposição de testes de ajuste baseados em (5.4) quando F é uma distribuição estável, sob a conveniente escolha de $w(t)$ que, neste caso, não depende intrinsecamente da distribuição desconhecida F .

Sob certas condições de regularidade, provou-se em del Barrio et al. (2005) que, se $B(t)$ é uma Ponte Browniana, então

$$\|q_n(t)\|_{2,w}^2 \xrightarrow{D} \|B_0\|_{2,w}^2, \quad (5.5)$$

com $B_0(t) = \frac{B(t)}{f(F^-(t))}$, ou seja,

$$nd_{2,w}^2(F_n, F) \xrightarrow{D} \int_0^1 \frac{B^2(t)}{f^2(F^-(t))} w(t) dt. \quad (5.6)$$

Neste capítulo, provaremos que a convergência em distribuição disposta em (5.6) pode ser melhorada para convergência em distância Mallows, isto é,

$$nd_{2,w}^2(F_n, F) \xrightarrow{d_1} \int_0^1 \frac{B^2(t)}{f^2(F^-(t))} w(t) dt. \quad (5.7)$$

Além disso, mostraremos que a escolha da ponderação $w(t)$, feita em Dorea e Lopes (2014), satisfaz tal convergência sempre que F possuir cauda de variação regular, independentemente dos momentos da distribuição, o que terminará por incluir, em (5.7), classes de distribuições mais restritivas quanto aos momentos, como as estáveis não Gaussianas.

A obtenção de (5.7) disporá da correspondente convergência em distribuição. Por isso, faremos, na próxima seção, uma breve discussão sobre dois aspectos técnicos importantes na produção de (5.6).

5.2 Discussão de hipóteses

A convergência em distribuição listada em (5.6) se baseia no processo quantil $q_n(t)$. Por isso, sua produção passa pelo Lema 1.3 e pela relação entre $q_n(t)$ e seu caso particular F uniforme. Nesta seção, discutiremos algumas ressalvas a esses dois aspectos, que permitirão uma melhor compreensão das hipóteses que serão impostas à obtenção da convergência (5.7).

Segundo o Teorema de Donsker, a distribuição de um funcional contínuo de um processo de somas parciais converge para o correspondente funcional de um movimento Browniano $B(t)$. Basta que o processo de somas parciais em questão seja obtido de variáveis aleatórias independentes e identicamente distribuídas com segundo momento

finito. Esse é o caso, por exemplo, do processo empírico uniforme $\alpha_n(t)$, uma soma de variáveis aleatórias indicadoras com segundo momento finito. Por isso, se considerarmos $\alpha_n(t)$ e $B(t)$ como elementos aleatórios tomando valores no espaço $D[0, 1]$ sob a topologia de Skorohod¹, é verdade que a convergência $\alpha_n(t) \xrightarrow{D} B(t)$ implica em

$$\|\alpha_n\|_\infty = \sup_{0 < t < 1} \alpha_n(t) \xrightarrow{D} \sup_{0 < t < 1} B(t) = \|B\|_\infty,$$

pois o funcional $x \mapsto \|x\|_\infty$ é contínuo na topologia de Skorohod a menos de um conjunto com medida de Wiener nula. Em outras palavras, o funcional é B -contínuo.

No caso do processo quantil $q_n(t)$, a convergência do Lema 1.3, dada por

$$q_n(t) \xrightarrow{D} \frac{B(t)}{f(F^-(t))}, \quad (5.8)$$

e a B -continuidade, na topologia de Skorohod, do funcional

$$x \mapsto \int_0^1 x(t)^2 dt,$$

poderiam facilmente nos induzir à conclusão de que

$$nd_2^2(F_n, F) = \int_0^1 [q_n(t)]^2 dt \xrightarrow{D} \int_0^1 \frac{B^2(t)}{f^2(F^-(t))} dt, \quad (5.9)$$

sem maiores exigências. Contudo, a convergência (5.9) não pode seguir automaticamente da convergência (5.8). A razão disso está na finitude do processo limite, condição que está intimamente ligada à finitude de sua função covariância, isto é,

$$\int_0^1 \frac{t(1-t)}{f^2(F^-(t))} dt < \infty. \quad (5.10)$$

O próximo Lema deixa clara essa relação.

Lema 5.1. (Csörgo e Horváth, 1993, Lema 5.2.3).

Suponha que $0 \leq p < \infty$, que h é positiva em $(0, 1)$, e que $B(t)$ é uma Ponte Browniana. As condições a seguir são equivalentes:

1. $\int_0^1 \frac{(t(1-t))^{p/2}}{h(t)} dt < \infty.$
2. $\int_0^1 \frac{|B(t)|^p}{h(t)} dt < \infty.$

A situação da convergência (5.9) é o caso $p = 2$ e $h(t) = f^2(F^-(t))$. Os exemplos a seguir ilustram as limitações que o Lema 5.1 impõe à convergência (5.9).

¹Um estudo apropriado da convergência em distribuição do processo empírico exige que o mesmo seja considerado no Espaço $D[0, 1]$, chamado Càdlàg (do francês “continue à droite avec limits à gauche”). A topologia de Skorohod torna esse espaço separável e completo, além de garantir a mensurabilidade do processo empírico.

Exemplo 5.1. *Seja F uma distribuição α -estável não Gaussiana. Pelo comportamento assintótico das caudas e da densidade de F , descrito pelos Teoremas 1.1 e 1.2, é verdade que*

$$\int_0^1 \frac{t(1-t)}{f^2(F^-(t))} dt = \int_{-\infty}^{\infty} \frac{F(x)\bar{F}(x)}{f^2(x)} dF(x) = \int_{-\infty}^{\infty} \frac{F(x)\bar{F}(x)}{f(x)} dx = \infty,$$

pois

$$\frac{\bar{F}(x)}{f(x)} \approx \frac{x^{-\alpha}}{x^{-(\alpha+1)}} = x.$$

Assim, mesmo que uma distribuição α -estável não Gaussiana satisfaça o Lema 1.3, a consequência natural do Lema 5.1 é que não vale a convergência (5.9).

Exemplo 5.2. *Seja F a distribuição maximal de Fréchet, com parâmetro $\alpha > 0$. Como essa distribuição possui cauda à direita de variação regular com índice $-\alpha$, os cálculos que demonstram a infinitude do limite em (5.9) são similares ao Exemplo 5.1, de forma que temos novamente a mesma conclusão. Vale observar aqui que, para $\alpha > 2$, F possui segundo momento finito, pois para $z \rightarrow \infty$,*

$$\int_z^{\infty} x^2 f(x) dx \approx \int_z^{\infty} \alpha x^{2-(\alpha+1)} dx.$$

Exemplo 5.3. *Suponha que F seja a distribuição normal padrão. Neste caso, conforme página 963 de Bickel e Zwet (1978), é verdade que*

$$\int_{1/n}^{1-1/n} \frac{t(1-t)}{f^2(F^-(t))} dt = \log(\log(n)) + \log(2) + \gamma + o(1),$$

em que γ é a constante de Euler, isto é,

$$\gamma = \lim_{k \rightarrow \infty} \left[\sum_{j=1}^k \frac{1}{j} - \log(k) \right].$$

Mesmo em casos supostamente mais simples, como é o caso das distribuições com segundo momento finito, o Lema 5.1 aponta a impossibilidade da convergência (5.9). De fato, não só o Exemplo 5.3 ilustra essa afirmação, mas o Exemplo 5.2 também confirma esse ponto de vista, pois para $\alpha > 2$ temos segundo momento finito.

Portanto, a adoção de uma ponderação conveniente para assegurar a condição

$$\int_0^1 \frac{t(1-t)}{f^2(F^-(t))} w(t) dt < \infty,$$

é indispensável na obtenção da convergência (5.7), pois nos valeremos da convergência em distribuição.

A igualdade

$$\sqrt{nd}_{2,w}(F_n, F) = \|q_n\|_{2,w}, \tag{5.11}$$

mostra que uma análise de convergência acerca de $nd_{2,w}^2(F_n, F)$ é, fundamentalmente, uma investigação sobre a convergência do processo quantil $q_n(t)$. A exemplo do que

buscamos fazer para a convergência em distância Mallows de $q_n(t)$, no capítulo anterior, exigiremos aqui alguma condição que nos permita aproveitar as boas propriedades do processo quantil no caso F uniforme, em virtude da relação (4.26), que pode ser escrita como

$$q_n(t) = \frac{1}{f(F^-(t))} \sqrt{n} (U_{k:n} - t) \cdot \frac{f(F^-(t))}{f(F^-(\xi_n))}. \quad (5.12)$$

Na relação acima, como já vimos, $U_{k:n} = F^-(U_n^-(t))$, $\frac{(k-1)}{n} < t \leq \frac{k}{n}$, $1 \leq k \leq n$. Além disso, $t \wedge U_n^-(t) < \xi_n < t \vee U_n^-(t)$, em que $U_n^-(t)$ é a inversa empírica uniforme proveniente da amostra $U_{k:n} = F(X_{k:n})$, na qual $X_{k:n}$ é a k -ésima estatística de ordem da amostra de F .

Contanto que possamos impor algum controle uniforme sobre a taxa

$$\frac{f(F^-(t))}{f(F^-(\xi_n))}, \quad (5.13)$$

é possível transferir a teoria que envolve o caso uniforme de $q_n(t)$ para o caso geral, a exemplo do que fizemos em nosso Lema 4.1. Essa transferência é interessante, pois o caso F uniforme não apresenta os problemas listados nos exemplos anteriores. Conforme o Capítulo 6 de Csörgo e Horváth (1993), a condição a seguir é uma imposição natural na obtenção do controle da taxa (5.13).

Definição 5.1. Condição Fundamental.

Considere

$$a_F = \sup\{x; F(x) = 0\}, \quad b_F = \inf\{x; F(x) = 1\}.$$

A função de distribuição F satisfaz a Condição Fundamental se:

1. F é duas vezes diferenciável em (a_F, b_F) .
2. $F'(x) = f(x) > 0$, $\forall x \in (a_F, b_F)$.
3. Para algum $\gamma > 0$, é verdade que

$$\sup_{a_F < x < b_F} F(x)[1 - F(x)] \frac{|f'(x)|}{f^2(x)} \leq \gamma,$$

ou equivalentemente,

$$\sup_{0 < t < 1} t[1 - t] \frac{|f'(F^-(t))|}{f^2(F^-(t))} \leq \gamma.$$

No contexto de proposição dos mais variados testes de ajuste que sejam baseados na distância Mallows ponderada e no processo quantil, a Condição Fundamental não pode ser dispensada. Por isso, ao projetarmos uma extensão da convergência em (5.6), é imperativa a adoção dessa hipótese. Apresentamos a seguir um lema que sintetiza o poder da Condição Fundamental.

Lema 5.2. *(Conforme Csörgo e Horváth, 1993, pp. 369 e 378)*

Se F satisfaz a Condição Fundamental, então

$$\frac{f(F^-(t_1))}{f(F^-(t_2))} \leq \left\{ \frac{t_1 \vee t_2}{t_1 \wedge t_2} \frac{1 - (t_1 \wedge t_2)}{1 - t_1 \vee t_2} \right\}^\gamma,$$

para cada par $t_1, t_2 \in (0, 1)$. Além disso, para ξ_n como em (5.12),

$$\sup_{\frac{1}{n} \leq t \leq 1 - \frac{1}{n}} \frac{f(F^-(t))}{f(F^-(\xi_n))} = O_p(1).$$

Na próxima seção, exibiremos o resultado de del Barrio et al. (2005) que garante, sob certas condições, a convergência em distribuição dada por

$$nd_{2,w}^2(F_n, F) \xrightarrow{D} \|B_0\|_{2,w}^2.$$

Em seguida, estenderemos tal convergência para a distância Mallows.

5.3 Convergência em distância Mallows de $nd_{2,w}^2(F_n, F)$

A seguir, transcrevemos o resultado de del Barrio et al. (2005), cujo modo de convergência estenderemos. Usaremos no enunciado a terminologia da igualdade (5.4).

Teorema 5.1. (del Barrio et al. (2005))

Sejam F uma distribuição contínua derivável e w uma ponderação que satisfaçam as seguintes hipóteses:

a) F satisfaz a Condição Fundamental;

b) $a_F > -\infty$ ou

$$\liminf_{x \rightarrow 0^+} \frac{|f'(F^-(x))|x}{f^2(F^-(x))} > 0;$$

c) $b_F < \infty$ ou

$$\liminf_{x \rightarrow 0^+} \frac{|f'(F^-(1-x))|x}{f^2(F^-(1-x))} > 0;$$

d)

$$\lim_{x \rightarrow 0^+} \frac{x \int_0^x w(t) dt}{f^2(F^-(x))} = 0, \quad \lim_{x \rightarrow 0^+} \frac{x \int_{1-x}^1 w(t) dt}{f^2(F^-(1-x))} = 0;$$

e)

$$\int_0^1 \frac{t(1-t)}{f^2(F^-(t))} w(t) dt < \infty.$$

Neste caso, temos que

$$q_n(t) \sqrt{w(t)} \xrightarrow{D} B_0(t) \sqrt{w(t)}. \tag{5.14}$$

Além disso,

$$nd_{2,w}^2(F_n, F) \xrightarrow{D} \|B_0\|_{2,w}^2. \tag{5.15}$$

Observe que, conforme discutido na seção anterior, o quinto item das hipóteses assegura a finitude do limite em (5.15). A seguir, estendemos a convergência do Teorema 5.1 apenas com a adição de uma hipótese do Lema 4.1.

Teorema 5.2. *Seja F uma distribuição contínua com densidade positiva. Se F satisfaz as hipóteses do Teorema 5.1 e do Lema 4.1, então*

$$nd_{2,w}^2(F_n, F) \xrightarrow{d_1} \|B_0\|_{2,w}^2. \quad (5.16)$$

Demonstração:

As hipóteses do Teorema 5.1 asseguram que

$$nd_{2,w}^2(F_n, F) \xrightarrow{D} \|B_0\|_{2,w}^2. \quad (5.17)$$

As hipóteses do Lema 4.1 implicam em

$$E[q_n(t)]^2 \xrightarrow{n} E[B_t]^2.$$

Disso, para cada $t \in (0, 1)$, vale que

$$E[(q_n(t))^2 w(t)] \xrightarrow{n} E[(B_t)^2 w(t)].$$

Pelo Teorema de Tonelli, podemos escrever

$$E\{\|q_n\|_{2,w}\} \xrightarrow{n} E\{\|B_0\|_{2,w}\},$$

ou

$$E\{nd_{2,w}^2(F_n, F)\} \xrightarrow{n} E\{\|B_0\|_{2,w}^2\}. \quad (5.18)$$

As relações (5.17) e (5.18) garantem nosso resultado, conforme a Proposição 1.2. □

Ilustraremos o Teorema 5.2 no próximo exemplo, através de uma distribuição cujas propriedades de variação regular motivarão um caso particular do resultado.

Exemplo 5.4. *Considere $F \sim S_1(1, 0, 0)$, a distribuição de Cauchy padrão. A densidade da distribuição é*

$$f(x) = \frac{1}{\pi(1+x^2)}, x \in \mathbb{R}.$$

Como $a_F = -\infty$ e $b_F = \infty$, é verdade que F é duas vezes diferenciável em (a_F, b_F) , com densidade estritamente positiva no mesmo intervalo. Além disso

$$\begin{aligned} \sup_{0 \leq t \leq 1} \frac{t(1-t)|f'(F^-(t))|}{f^2(F^-(t))} &= \sup_{x \in \mathbb{R}} \frac{F(x)(1-F(x))|f'(x)|}{f^2(x)} \\ &= \sup_{x \in \mathbb{R}} 2\pi|x|F(x)\bar{F}(x) \\ &= 2\pi \sup_{x \in \mathbb{R}} \frac{F(x)\bar{F}(x)}{\frac{1}{|x|}} \\ &= 2\pi. \end{aligned}$$

Na última igualdade foi usada a regra de L'Hôpital para o cálculo das indeterminações $\lim_{x \rightarrow \infty} |x| \bar{F}(x)$ e $\lim_{x \rightarrow -\infty} |x| F(x)$. Essas verificações atestam que F satisfaz a Condição Fundamental. Para verificar o item b) das hipóteses do Teorema 5.1, observe que a troca $x = F(t)$ implica em

$$\liminf_{x \rightarrow 0^+} \frac{|f'(F^{-1}(x))| x}{f^2(F^{-1}(x))} = \liminf_{t \rightarrow -\infty} \frac{|f'(t)| F(t)}{f^2(t)}.$$

Mas, usando novamente a regra de L'Hôpital,

$$\lim_{t \rightarrow -\infty} \frac{|f'(t)| F(t)}{f^2(t)} = \lim_{t \rightarrow \infty} \frac{2| -t| F(t)}{\pi} = \lim_{t \rightarrow \infty} \frac{t^2}{1 + t^2} = 1.$$

Como o limite acima é 1, vale o mesmo para o limite inferior. Quanto ao item c), veja que

$$\liminf_{x \rightarrow 0^+} \frac{|f'(F^{-1}(1-x))| x}{f^2(F^{-1}(1-x))} = \liminf_{t \rightarrow 1^-} \frac{|f'(F^{-1}(t))| (1-t)}{f^2(F^{-1}(t))} = \liminf_{y \rightarrow \infty} \frac{|f'(y)| [1 - F(y)]}{f^2(y)}.$$

Do mesmo modo que antes,

$$\lim_{y \rightarrow \infty} \frac{|f'(y)| [1 - F(y)]}{f^2(y)} = \lim_{y \rightarrow \infty} \frac{2y[1 - F(y)]}{\pi} = \lim_{y \rightarrow \infty} \frac{-y^2}{-(1 + y^2)} = 1.$$

Agora, considere a mesma ponderação do Exemplo 3.1, isto é,

$$w(t) = \begin{cases} 32t^3 & , \quad 0 \leq t < \frac{1}{2}, \\ 32(1-t)^3 & , \quad \frac{1}{2} \leq t \leq 1. \end{cases}$$

Trocas como as anteriores resultam na igualdade

$$\lim_{x \rightarrow 0^+} \frac{x \int_0^x w(t) dt}{f^2(F^{-1}(x))} = \lim_{x \rightarrow 0^+} \frac{8x^5}{f^2(F^{-1}(x))} = 8 \lim_{t \rightarrow \infty} \frac{\left[\frac{1}{2} + \frac{\arctan(-t)}{\pi} \right]^5}{\frac{1}{\pi^2(1+t^2)^2}} = 0,$$

e na igualdade

$$\lim_{x \rightarrow 0^+} \frac{x \int_{1-x}^1 w(t) dt}{f^2(F^{-1}(1-x))} = \lim_{x \rightarrow 0^+} \frac{8x^5}{f^2(F^{-1}(1-x))} = 8 \lim_{t \rightarrow \infty} \frac{\left[\frac{1}{2} - \frac{\arctan(-t)}{\pi} \right]^5}{\frac{1}{\pi^2(1+t^2)^2}} = 0,$$

o que significa que está satisfeito o item d) do Teorema 5.1. Podemos assegurar também o item e), pois

$$\int_0^1 \frac{t(1-t)}{f^2(F^{-1}(t))} w(t) dt = 32 \int_0^{\frac{1}{2}} \frac{t^4(1-t)}{f^2(F^{-1}(t))} w(t) dt + 32 \int_{\frac{1}{2}}^1 \frac{t(1-t)^4}{f^2(F^{-1}(t))} w(t) dt < \infty.$$

A finitude dessas duas integrais segue de fatos similares, de maneira que justificaremos apenas a finitude da segunda delas. Já provamos, na Seção 1.3, que $\bar{F} \in RV_{-1}$, de forma que $\bar{F}(x)^4 \approx cx^{-4}$ para alguma constante c . Logo, para z suficientemente grande

$$\int_z^\infty \frac{F(x)\bar{F}(x)^4}{f(x)} dx \leq \int_z^\infty \frac{\bar{F}(x)^4}{f(x)} dx \approx \int_z^\infty \frac{\pi(1-x^2)}{x^4} dx < \infty.$$

Como a distribuição satisfaz a Condição Fundamental, o Lema 5.2 implica na limitação (4.30), uma das hipóteses do Lema 4.1, como queríamos. Portanto, para a distribuição de Cauchy F e para a ponderação escolhida, resta garantido que

$$nd_{2,w}^2(F_n, F) \xrightarrow{d_1} \|B_0\|_{2,w}^2.$$

Exemplo 5.5. Considere $F \sim S_{1/2}(1, 0, 0)$, a distribuição de Levy. Neste caso,

$$F(x) = \frac{2}{\pi} \int_{1/\sqrt{2x}}^\infty e^{-t^2} dt.$$

Aqui, todos os itens do Teorema 5.2 são satisfeitos por causa da variação regular de \bar{F} , provada na Seção 1.3, e da variação regular de $F(-x)$, diretamente verificada por cálculos imediatos. A escolha conveniente da ponderação w pode ser

$$w(t) = \begin{cases} 192t^5 & , \quad 0 \leq t < \frac{1}{2}, \\ 192(1-t)^5 & , \quad \frac{1}{2} \leq t \leq 1. \end{cases}$$

O Exemplo 5.4 e o Exemplo 5.5 sugerem uma caracterização mais geral de classes de distribuições F e ponderações w para as quais se aplica o Teorema 5.2. De fato, as hipóteses do resultado funcionaram com base em características de F que são comuns a muitas outras distribuições e com ponderações que não dependem essencialmente de F .

Na próxima seção, provaremos que o Teorema 5.2 vale para todas as distribuições com caudas de variação regular, contanto que a ponderação apropriada seja utilizada.

5.4 Generalização do caso de variação regular

A fim de propor testes de similaridade que incluíssem as distribuições α -estáveis ($0 < \alpha < 2$), foi apresentada, em Dorea e Lopes (2014), uma classe de ponderações w construída com base no comportamento das caudas da função de distribuição. Tal classe de ponderações, quando aplicada de forma conveniente nas distribuições α -estáveis, transformava-as em distribuições com segundo momento finito, conforme ilustramos no Exemplo 3.1. Reescrevemos a seguir a estrutura dessa classe de ponderações.

Definição 5.2. (Conforme Dorea e Lopes (2014), p. 8)

Suponha que F é uma distribuição α -estável ($0 < \alpha < 2$) e $\beta < -\frac{2}{\alpha}$. Neste caso,

$$w(t) = \begin{cases} Ct^{-\beta} & , \quad 0 < t < \tilde{C}, \\ C(1-t)^{-\beta} & , \quad \tilde{C} \leq t < 1, \end{cases}$$

em que $0 < \tilde{C} < 1$ e $C = \frac{1-\beta}{\tilde{C}^{1-\beta} + (1-\tilde{C})^{1-\beta}}$.

Em nosso próximo resultado, provaremos que a classe de ponderações da Definição 5.2 é suficiente para aplicarmos o Teorema 5.2, quando F pertence ao domínio de atração de uma distribuição α -estável não Gaussiana.

Teorema 5.3. *Seja G uma distribuição α -estável ($0 < \alpha < 2$). Se $F \in D(G)$ e w é a ponderação da Definição 5.2, então*

$$nd_{2,w}^2(F_n, F) \xrightarrow{d_1} \|B_0\|_{2,w}^2.$$

Demonstração:

Provaremos que estão satisfeitas, neste caso, as hipóteses do Teorema 5.2. A validade da Condição Fundamental e da relação

$$\int_0^1 \frac{t(1-t)}{f^2(F^-(t))} w(t) dt < \infty,$$

estão detalhadas pelo Lema 3.1 de Dorea e Lopes (2014) quando F possui caudas de variação regular com índice $0 < \alpha < 2$, que é justamente o nosso caso, conforme o Teorema 1.1. Por isso, não repetiremos essas verificações aqui. Vamos verificar as outras hipóteses do Teorema 5.1.

Para provar o item *c*), observe que as trocas $x = F(t)$ e $t = -y$, combinadas com os Teoremas 1.1 e 1.2, implicam em

$$\begin{aligned} \liminf_{x \rightarrow 0^+} \frac{|f'(F^-(x))|x}{f^2(F^-(x))} &= \liminf_{t \rightarrow -\infty} \frac{|f'(t)|F(t)}{f^2(t)} \\ &= \liminf_{y \rightarrow \infty} \frac{|f'(-y)|F(-y)}{f^2(-y)} \\ &\approx \lim_{y \rightarrow \infty} C^* \frac{y^{-1-2} \cdot y^{-1}}{y^{-4}} \\ &= C^* > 0. \end{aligned}$$

Os mesmos teoremas e as trocas $t = 1 - x$ e $F(y) = t$ garantem o item *d*), pois

$$\begin{aligned} \liminf_{x \rightarrow 0^+} \frac{|f'(F^-(1-x))|x}{f^2(F^-(1-x))} &= \liminf_{t \rightarrow 1^-} \frac{|f'(F^-(t))|(1-t)}{f^2(F^-(t))} \\ &= \liminf_{y \rightarrow \infty} \frac{|f'(y)|[1-F(y)]}{f^2(y)} \\ &\approx C' \lim_{y \rightarrow \infty} \frac{y^{-1-2} \cdot y^{-1}}{y^{-4}} \\ &= C' > 0. \end{aligned}$$

Agora, considere a escolha da Definição 5.2. As trocas $x = F(-t)$ e $x = 1 - F(t)$ resultam, respectivamente, na igualdade

$$\lim_{x \rightarrow 0^+} \frac{x \int_0^x w(t) dt}{f^2(F^-(x))} = \lim_{x \rightarrow 0^+} \frac{8x^5}{f^2(F^-(x))} = 8 \lim_{t \rightarrow \infty} \frac{[F(-t)]^5}{f^2(-t)} \approx \lim_{t \rightarrow \infty} t^{-1} = 0,$$

e na igualdade

$$\lim_{x \rightarrow 0^+} \frac{x \int_{1-x}^1 w(t) dt}{f^2(F^-(1-x))} = \lim_{x \rightarrow 0^+} \frac{8x^5}{f^2(F^-(1-x))} = 8 \lim_{t \rightarrow \infty} \frac{[1 - F(t)]^5}{f^2(t)} \approx \lim_{t \rightarrow \infty} t^{-1} = 0,$$

o que significa que está satisfeito o item *d*) do Teorema 5.1. Essa verificação e o Teorema 5.2 encerram nossa demonstração, pois uma das hipóteses do Lema 4.1 já fica garantida pela Condição Fundamental, conforme o Lema 5.2.

□

A demonstração do Teorema 5.3 foi baseada em propriedades decorrentes da variação regular de F , por isso, um questionamento automático diante do resultado é sobre sua validade para qualquer caso em que F possua caudas de variação regular. Para os casos em que temos somente $\bar{F} \in RV_\alpha$ ou $F(-x) \in RV_{-\alpha}$ com $\alpha > 2$, temos a finitude do segundo momento (conforme ilustra o Exemplo 5.2), a validade da Condição Fundamental e a falha da condição (5.10). Para contornar esse problema, duas variações da Definição 5.2 são válidas, a saber:

Definição 5.3. *Considere, a seguir, C e \tilde{C} convenientemente escolhidos para que w seja uma densidade em $(0, 1)$.*

1. *Para distribuições com cauda de variação regular apenas à esquerda:*

$$w(t) = \begin{cases} Ct^2 & , \quad 0 \leq t < \tilde{C}, \\ 1 & , \quad \tilde{C} \leq t \leq 1, \end{cases}$$

2. *Para distribuições com cauda de variação regular apenas à direita:*

$$w(t) = \begin{cases} 1 & , \quad 0 \leq t < \tilde{C}, \\ C(1-t)^2 & , \quad \tilde{C} \leq t \leq 1, \end{cases}$$

Observe que, se F possui cauda de variação regular apenas de um lado com índice $-\alpha$ ($0 < \alpha < 2$), a verificação das hipóteses do Teorema 5.1 torna-se mais simples, pela troca $-\beta = 2$ na estrutura da ponderação. As escolhas de ponderações da Definição 5.3 são suficientes para que possamos incluir no Teorema 5.1 as distribuições com caudas de variação regular com qualquer índice não nulo $-\alpha$, pois as verificações dos itens *b*), *c*) e *d*) daquele Teorema seguem inteiramente válidas, nesse caso, bastando que se use, conforme necessário, a Definição 5.2 ou 5.3. Resumimos esses fatos no Corolário a seguir.

Corolário 5.1. (do Teorema 5.1)

Seja F uma distribuição contínua tal que $\bar{F} \in RV_{-\alpha}$ e/ou $F(-x) \in RV_{-\alpha}$, $\alpha > 0$. Neste caso,

$$nd_{2,w}^2(F_n, F) \xrightarrow{D} \|B_0\|_{2,w}^2,$$

em que w é uma ponderação da Definição 5.2 ou da Definição 5.3.

Quanto aos Teoremas 5.2 e 5.3, a inclusão dos casos acima depende do atendimento da distribuição F a uma das hipóteses do Lema 4.1.

O Corolário 5.1 encerra as contribuições de nosso trabalho. Na próxima seção, disporemos as considerações finais desta tese.

Considerações Finais

Cumprimos os objetivos aos quais nos propusemos desde o início deste texto.

Derivamos extensões para conhecidas convergências na distância Mallows envolvendo somas parciais, usando resultados de nossa autoria que se apoiaram em ponderações.

Expandimos nossa abordagem para os processos empíricos, um tipo particular de somas parciais, mimetizando as técnicas de demonstração usadas nos primeiros capítulos. Continuamos nosso estudo em direção aos processos quantis gerais, obtendo a convergência em distância Mallows através do processo quantil uniforme.

Chegamos, por fim, à unificação dos elementos norteadores da exposição por meio do Teorema 5.2, um resultado que estabeleceu o comportamento assintótico em distância Mallows da estatística $nd_{2,w}^2(F_n, F)$, uma medida de dissimilaridade entre a distribuição F e sua correspondente distribuição empírica. Indicamos, ainda, a aplicabilidade desse Teorema para a classe das distribuições com cauda de variação regular, conjunto que contém propriamente as distribuições estáveis não Gaussianas.

Os lemas que propomos na Seção 2.3 foram os ingredientes fundamentais na produção de nossos principais resultados. Os entes matemáticos ponderados que foram trabalhados naqueles lemas, permitiram-nos usar, com mais desenvoltura, as ferramentas que auxiliam a relação entre a convergência em distribuição e a convergência em distância Mallows, tais como a Proposição 1.2 e o Teorema 1.5. Apenas para ilustrar um exemplo, pudemos usar a invariância da convergência fraca por ponderações, disposta no Lema 2.2, como um conveniente atalho em demonstrações dos Capítulos 2 e 3, pois a relação

$$F_n \xrightarrow{D} G \Leftrightarrow F_{n,w} \xrightarrow{D} G_w,$$

garante a adoção e a dispensa de uma ponderação sem maiores exigências.

O sucesso dos lemas da Seção 2.3 na obtenção de nossos mais importantes resultados, instiga, naturalmente, o interesse em testar sua eficiência na abordagem de assuntos presentes nos artigos que citamos para ilustrar a relevância das distribuições ponderadas e da distância Mallows ponderada. Essa verificação, possivelmente longa, será objeto de trabalhos futuros.

Os processos empíricos, $\beta_n(t)$ e $q_n(t)$, surgiram em nosso trabalho por constituírem, ao mesmo tempo, uma ilustração de somas parciais e uma ligação clara com o Capítulo 5. Para t fixado, a convergência na r -ésima distância Mallows de $\beta_n(t)$ para uma ponte

Browniana precisou de uma imposição sobre a função geradora de momentos dada por

$$\sup_n M_{\beta_n(t)}^{([r]+1)}(0) < \infty. \quad (5.19)$$

Sob nosso ponto de vista, a imposição (5.19) pareceu verdadeira desde o início da pesquisa no tema. Contudo, não encontramos uma forma eficiente e direta de verificá-la, exceto para $r = 5$. Essa verificação é parte de novos projetos.

Em se tratando de processos empíricos, há um relevante aspecto do estudo que pode ser tratado em trabalhos vindouros: trata-se do caso em que $\beta_n(t)$ e $q_n(t)$ surgem de uma sequência de variáveis aleatórias não necessariamente independentes. A hipótese de independência pode ser razoavelmente admitida em aplicações práticas, mas prescindir dela torna qualquer modelagem mais realista. Nessa situação, todos os resultados clássicos em torno da distribuição empírica, e dos entes matemáticos derivados dela, precisam de alguma revisão a respeito de sua validade.

A convergência em distância Mallows da estatística $nd_{2,w}^2(F_n, F)$ reuniu os tópicos de interesse de nossa apresentação, mas um questionamento sobre sua validade para ordens superiores (tanto na estatística quanto no modo de convergência) é automático e justificável, visto nossa preocupação com esse aspecto no Capítulo 4. A discussão que fizemos na Seção 5.2 ilustra bem o quão delicada pode ser uma extensão nesse sentido. De fato, há limitações técnicas acerca do funcional integral envolvido na expressão de $d_{r,w}^r(F_n, F)$ que escolhemos não abordar aqui, especialmente porque muitos novos conceitos e suposições seriam necessários, o que justifica uma ramificação futura de nossa pesquisa, mas não a produção de mais um capítulo nesta tese.

Essas considerações encerram a exposição de nosso trabalho.

Referências Bibliográficas

- [1] BARBOSA, E. G. and DOREA, C. C. Y. *A note on Lindeberg condition for Convergence to stable laws in Mallows distance*. Bernoulli, Vol. **15** (3), pp. 922-924, 2009.
- [2] BICKEL, P.J. and FREEDMAN, D.A. *Some Asymptotic Theory for the Bootstrap*. The Annals of Statistics, Vol. **9**, pp. 1196-1217, 1981.
- [3] BICKEL, P.J. and van ZWET, W.R. *Asymptotic Expansions for the Power of Distribution free Tests in the Two-Sample Problem*. The Annals of Statistics, Vol. **6**, pp. 937-1004, 1978.
- [4] BILLINGSLEY, P. *Probability and Measure*. Third Edition. New York, John Wiley & Sons, 1995.
- [5] BREIMAN, L. *Probability*. Philadelphia, Classics in Applied Mathematics - 7, 1992.
- [6] CASELLA, J. and BERGER, R. *Statistical Inference*. Second Edition. USA, Duxbury Thomson Learning, 2002.
- [7] CRAMÉR, H. *Mathematical Methods of Statistics*. USA, Princeton University Press, 1946.
- [8] CSÖRGO, M. and HORVÁTH, L. *Weighted Approximations in Probability and Statistics*. New York, John Wiley & Sons, 1993.
- [9] CSÖRGO, S. and SZABÓ, T. *Weighted Correlation Tests for Gamma and Lognormal Families*. Tatra Mountains Mathematical Publications, Vol. **26**, pp. 337-356, 2003.
- [10] DALL'AGLIO, G. *Sugli Estremi dei Momenti delle Funzioni di Ripartizione Doppia*. Annali della Scuola Normale Superior de Pisa, CL. Sci. **3**(1), pp. 33-74, 1956.
- [11] del BARRIO, E., CUESTA-ALBERTOS, J., MATRÁN, C. and RODRÍGUEZ-RODRÍGUEZ, J. *Tests of Goodness of fit based on the L_2 -Wasserstein Distance*. Ann. Statist, **27**, pp. 1230-1239, 1999.

- [12] del BARRIO, E., CUESTA-ALBERTOS, J. and MATRÁN, C. *Contributions of empirical and quantile processes to the asymptotic theory of goodness-of-fit tests*. Sociedad de Estadística e Investigación Operativa, Vol. **9(1)**, pp. 1-96, 2000.
- [13] del BARRIO, E., GUINÉ, E. and UTZET, F. *Asymptotics for L_2 Functionals of the empirical quantile processes, with applications to tests of fit based on weighted Wasserstein distances*. Bernoulli, Vol. **11**, pp. 131-189, 2005.
- [14] de WET, T. *Goodness of Fit Tests for Location and Scale Families Based on a Weighted L_2 Wasserstein Distance Measure*. Sociedad de Estadística e Investigación Operativa, Vol. **11**, pp. 89-107, 2002.
- [15] DOREA, C. C. Y. and FERREIRA, D. B. *Conditions for Equivalence Between Mallows Distance and Convergence to Stable Laws*. Acta Mathematica Hungarica, Vol. **134(1-2)**, pp. 1-11, 2012.
- [16] DOREA, C. C. Y. and LOPES, L. P. *Weighted Similarity Tests for Location-Scale Families of Stable Distributions*. Communications in Statistics: Theory and Methods, a aparecer, 2014.
- [17] DORFMAN, R. *A Note on the δ -Method for Finding Variance Formulae*. The Biometric Bulletin, Vol. **1**, pp. 129-137, 1938.
- [18] DOSS, H. and GILL, R. *An Elementary Approach to Weak Convergence for Quantile Processes, With Applications to Censored Survival Data*. Journal of the American Statistical Association, Vol. **87** N_o 419, pp. 869-877, 1992.
- [19] EMBRECHTS, P., KLÜPPELBERG, C. and MIKOSCH, T. *Modeling Extremal Events for Insurance and Finance*. Springer-Verlang, Berlim, 1997.
- [20] FELLER, W. *An Introduction to Probability Theory and its Applications, Vol. II*. Second Edition. New York, John Wiley & Sons, 1971.
- [21] GALAMBOS, J., *The asymptotic theory of extreme order statistics*. New York, John Wiley, 1978.
- [22] HALL, P. and HEYDE, C. C., *Martingale Limit Theory and its Application*. Academic. New York, 1980.
- [23] HOEF, JAY M. Ver. *Who Invented the Delta Method?*. The American Statistician, Vol. **66:2**, pp. 124-127, 2012.
- [24] IRPINO, A. and VERDE, R. *Basic statistics for distributional symbolic variables: a new metric-based approach*. Advances in Data Analysis and Classification, a aparecer, 2014.
- [25] IRPINO, A. and VERDE, R. *Dynamic clustering of histogram data based on adaptive squared Wasserstein distances*. Expert Systems with Applications, Vol. **41(7)**, pp. 3351-3366, 2014.

- [26] IRPINO, A. and VERDE, R. *Linear Regression for numeric symbolic variables: a least squares approach based on Wasserstein Distance*. Advances in Data Analysis and Classification, Vol. **9(1)**, pp. 81-106, 2015.
- [27] JOHNSON, O. and SAMWORTH, R. *Central limit theorem and convergence to stable laws in Mallows distance*. Bernoulli, Vol. **11**, pp. 829-845, 2005.
- [28] KANTOROVICH, L.V. *On one Effective Method of Solving Certain Classes of Extremal Problems*. Dokl. Akad. Nauk., USSR, Vol. **28**, pp. 212-215, 1940.
- [29] KIM, Hea-Jung. *A class of flexible weighted distributions for modelling and analyzing non-normal data*. Journal of the Korean Statistical Society, Vol. **37**, pp. 107-118, 2008.
- [30] KOSMELJ, K. and BILLARD, L. *Mallows L^2 -Distance in Some Multivariate Methods and its Application to Histogram-Type Data*. Metodoloski Zvezki, Vol. **9(2)**, pp. 107-118, 2012.
- [31] LEVINA, E., BICKEL, P. *The Earth Mover's distance is the Mallows Distance: Some insights from statistics*. Proceedings of IEEE International Conference on Computer Vision, Vol. **2**, pp. 251-256. Vancouver, BC, Canada, 2001.
- [32] LI, Ying. LI, Yu. and HU, TAIZHONG. *Probability inequalities for weighted distributions*. Journal of Statistical Planning and Inference, **142**, pp. 1272-1278, 2012.
- [33] LISINI, S., MATTHES, D. SAVARÉ, G. *Cahn-Hilliard and thin film equations with nonlinear mobility as gradient flows in weighted-Wasserstein metrics*. Journal of Differential Equations, **253**, pp. 814-850, 2012.
- [34] LOCKHART, R.A. and STEPHENS, M.A. *The probability plot: test of fit based on the correlation coefficient*. In N. Balakrishnan and C.R Rao (eds), *Orders Statistic: Applications*, Handbook of Statistics, **17**, pp. 453-473, Amsterdam: North-Holland, 1998.
- [35] MA, Y., GENTON, M.G. and TSIATIS, A.A. *Locally efficient semiparametric estimators for generalized skew-elliptical distributions*. Journal of the American Statistical Association, Vol. **100**, pp. 980-989, 2005.
- [36] MALLOWS, C.L. *A Note on asymptotic joint normality*. Annals of Mathematical Statistics, Vol. **43(2)**, pp. 508-515, 1972.
- [37] OEHLERT, G. *A Note on the Delta Method*. The American Statistician, Vol. **46**, pp. 27-29, 1992.
- [38] RACHEV, S.T. and RÜSCHENDORF, L., *Mass Transportation Problems, Volume I: Theory*. Springer series in Statistics. New York, 1998.
- [39] RESNICK, S. I. *Extremes Values, Regular Variation, and Point Processes*. New York, Springer-Verlag, 1987.

-
- [40] ROHATGI, V.K. and SALEH, A.K., *An Introduction to Probability and Statistics*. Second Edition. New York, John Wiley & Sons, 1976.
- [41] SHORACK, Galen G. and WELLNER, Jon A. *Empirical Processes with Applications to Statistics*. New York, Wiley & Sons, 1986.
- [42] ZHOU, D., LI, J., ZHA, H. *A New Mallows Distance Based Metric for Comparing Clustering*. Proceedings of the 22nd International Conference on Machine Learning, Bonn, Germany, 2005.