Universidade de Brasília
Instituto de Ciências Exatas
Departamento de Matemática

Testes de Similaridade na Distância de Mallows-Wasserstein Ponderada para Distribuições de Cauda Pesada

por

Luciene Pinheiro Lopes

Brasília, 2012.

Dedico essa tese a pessoa mais importante da minha vida, minha mamãe, Maria das Graças Lopes.

"Tenho medo de não responder, de fingir que eu não escutei, tenho medo de ouvir um chamado, um pedido ou um apelo e simplesmente virar para o outro lado e fingir que eu não sei".

Padre Zezinho.

Agradecimentos

Agradeço a ti Jesus por ter me dado forças para suportar todas as provações que surgiram nesse período de Doutorado que não foram poucas, por todos os livramentos e graças concedidas a fim de que eu pudesse finalizar essa tese. Entendo a data em que a minha defesa foi realizada como obra tua para me mostrar como a vida dá voltas e como a minha fé é pequena. A ti, toda honra, glória, louvor, domínio e gratidão.

A minha mãe, Maria das Graças, por me permitir conhecer o amor incondicional, por perdoar as coisas imperdoáveis, pelo desapego, pela compreensão e profunda paciência.

Ao meu amado irmão Gledson (meu melhor amigo), ao meu pai e os meus tios Divino, Maurício e Neusa, agradeço pela simplicidade, generosidade e cooperação nas dificuldades, agradecimentos que estendo aos meus primos Alex, Edvaldo e Vilma.

Em especial, dedico essa tese aos que estiveram comigo no período em que estive de licença da universidade. Obrigada por não terem desistido de mim no momento em que mais precisei de companhia, obrigada por terem me permitido, por meio da generosidade desse gesto raro que tiveram, finalizar esse projeto de Doutorado.

A minha orientadora, Chang Chung Yu Dorea, uma pessoa que além do profundo conhecimento matemático e generosidade em compartilhá-lo, consegue simultaneamente ser um ser humano extraordinário, que me fez amadurecer muito nesses anos de convivência, que me compreendeu quando isso não era algo trivial e me apoiou nos momentos em que seria completamente aceitável se não o fizesse, que sempre me estimulou a crescer, uma companhia muito agradável, das mais construtivas. Por tudo isso, tenho por ela muito mais do que uma profunda admiração e gratidão pelo trabalho acadêmico que fez comigo, mas um afeto forte, sincero e permanente.

A professora Cátia por ter despertado o meu interesse pela Probabilidade ao assistir suas excelentes aulas, por tudo que me ensinou quando me orientou no Mestrado, me dando uma base sólida para o Doutorado e sendo uma forte contribuidora para o sucesso do mesmo. Sou grata a ela por ter se mantido presente na minha vida acadêmica após o término do Mestrado e por ter me ajudado e tranquilizado sempre que recorri a ela, sempre tão divertida, linda, animada e querida por todos.

A professora Sílvia pelo exemplo de doçura e competência. Por todas as palavras de incentivo, pelos conselhos e o estímulo a continuidade da pesquisa. Uma pessoa pelo qual adquiri muito carinho.

Aos professores Ana Maria Gandulfo, Celius Magalhães e José Alfredo(in memoriam) pelo excelente trabalho que tanto me estimulou.

Aos colegas e amigos que fiz na universidade, pela troca de idéias, pelas discussões, pelos conselhos, sugestões, motivações e materiais que tanto contribuiram para o meu crescimento profissional e pessoal: Adriana, Alexandre, Anyelle, Bel, Débora, Eunice, Evander, Karise, Kéllem, Leandro, Magno, Manoel(in memoriam), Manuela, Mari, Martha, Miguel, Simone, Vivian, Walter e Wembesom.

Agradeço a Simone, ao Wembesom e a Mari por terem surgido na minha vida em um período de renovação, por terem sido tão presentes no desenvolvimento dessa tese e pelo suporte constante e parceria nos projetos de vida. E ao Walter que assim como a Simone, se disponibilizaram a me ouvir falar dos meus resultados e acreditaram neles, muitas vezes mais do que eu mesma.

Ao grupo da Probabilidade da UnB pelo espírito de equipe e gentileza.

As minhas amigas Dayane e Mônica pelo afeto a mim dirigido e pelo trabalho que tiveram revisando essa tese.

A minha amiga Eunice do qual só tenho boas recordações, obrigada pela torcida, pelas orações e palavras de otimismo ao longo desse Doutorado. Foram muitas risadas, passeios e lágrimas.

Ao meu amigo de infância Kleiton e a Dayane por fazerem tanta questão da minha amizade após tantos e tantos anos.

Aos membros da banca pelo posicionamento construtivo, por todo o incentivo dado para que eu continue a pesquisa e sugestões para tal, pelas correções, pela gentileza com a qual me trataram e a dedicação que tiveram ao meu trabalho durante todo processo de defesa.

Ao CNPq pelo suporte financeiro.

Resumo

Neste trabalho propomos testes não-paramétricos para classes de distribuições de cauda pesada, que incluem as α-estáveis e as extremais de Fréchet. As estatísticas apresentadas, funcionais do processo quantil empírico, permitem testar a pertinência da distribuição F à família de escala-locação gerada por uma distribuição de cauda pesada G, $F \in \mathcal{G}_G$, bem como a ε-similaridade, $d(F, \mathcal{G}_G) \leq \varepsilon$, em que d é uma métrica apropriada. Mediante o uso da distância Mallows-Wasserstein ponderada, determinamos, sob a hipótese nula, as distribuições assintóticas e mostramos que essas distribuições são os correspondentes funcionais das pontes Brownianas. Os resultados constituem uma extensão de similares, baseados na distância Wasserstein, aplicáveis a distribuições com segundo momento finito.

Palavras-chave: testes de similaridade; leis estáveis; cauda pesada; distância Mallows-Wasserstein; processo quantil empírico; pontes Brownianas.

Abstract

In this work we derive the asymptotic null distribution of weighted quantile correlation tests statistics for the Fréchet family and the stable laws. This extends previous results for light-tailed distributions and is achieved by considering a special class of weight functions along with the use of weighted Mallows-Wasserstein distance. The test statistics for the location-scale family $\mathcal{G}_G$, generated by a heavy-tailed distribution G, when analyzed in the context of similarity of the distributions, shows that the distance between trimmed distributions according to the weight function is efficient in measuring the dissimilarity between heavy-tailed distributions.

Keywords: Similarity test; Stable laws; Heavy tail; Mallows-Wasserstein distance; Empirical quantile process; Brownian bridge.

Sumário

Introdução

Nesta tese estudamos a problemática de formular testes não paramétricos para distribuições de cauda pesada, tais como: as α-estáveis, a família das distribuições extremais de Fréchet e aquelas com caudas tipo Pareto. Mais especificamente, para a família de escala-locação $\mathcal{G}_G$, gerada por uma distribuição de cauda pesada G, propomos testes de ajuste de bondade para $F \in \mathcal{G}_G$, bem como testes de ε-similaridade para avaliar a dissimilaridade entre as distribuições F e $\mathcal{G}$ através da $d_2(F, \mathcal{G}_G) \leq \varepsilon$, em que d_2 é a distância-2 de Mallows-Wasserstein.

O interesse primordial deste trabalho são as distribuições estáveis $S_\alpha(\sigma, \beta, \mu)$, com $0 < \alpha \leq 2$, dada a sua importância central para análise estatística, já que representam os únicos limites assintóticos possíveis para somas parciais normalizadas de variáveis aleatórias independentes. Isso se deve a sua propriedade de infinita divisibilidade, isto é, a soma de cópias independentes de variáveis aleatórias estáveis ajustadas tem novamente a mesma distribuição estável. Trataremos do caso $\alpha < 2$, pois quando $\alpha = 2$, temos o caso da distribuição Gaussiana, já amplamente estudado. Como a nossa abordagem se baseia, essencialmente, no comportamento regularmente variante das caudas, os resultados com pequenas modificações também contemplam a família das extremais de Fréchet e outras distribuições com cauda tipo Pareto.

As estatísticas precursoras neste tipo de estudo são as renomadas estatísticas de Kolmorogov

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

e as de Kolmogorov-Smirnov

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

onde F_n é a distribuição empírica de uma amostra de F. Para ambos os casos, a distribuição limite pode ser expressa como funcionais da ponte Browniana $\{B(t) : 0 \leq t \leq 1\}$,

$$\lim_n P(D_n \leq x) = P\left(\sup_{0\leq t\leq 1} |B(t)| \leq x\right)$$

e

$$\lim_n P(D_n^+ > x) = P\left(\sup_{0\leq t\leq 1} B(t) > x\right).$$

Uma análise mais refinada, envolvendo o processo quantil empírico ajustado da distribuição F com densidade f,

$$\rho_{n,F}(t) := \sqrt{n}(F_n^{-1}(t) - F^{-1}(t))f(F^{-1}(t)),$$

permite provar que, sob condições de regularidade, temos uma aproximação precisa de $\rho_{n,F}(.)$ por uma sequência de pontes Brownianas (vide [5]):

$$n^{\frac{1}{2}-\nu} \sup_{\frac{1}{n}\leq t\leq 1-\frac{1}{n}} \frac{|\rho_{n,F}(t) - B_n(t)|}{(t(1-t))^\nu} = \begin{cases} O(\log n) & \text{se } \nu = 0 \\ O(1) & \text{se } 0 < \nu \leq \frac{1}{2}. \end{cases}$$

Essa aproximação bem como a identidade da distância Wasserstein

$$\sqrt{n}d_2(F_n, F) = \left(\int_0^1 \frac{\rho_{n,F}^2(t)}{f^2(F^{-1}(t))}dt\right)^{\frac{1}{2}},$$

válida para distribuições $F \in \mathcal{L}_2 = \left\{F : \int x^2 dF(x) < \infty\right\}$, serviram como base para os trabalhos [11] e [12], em que foram formulados testes de ajuste de bondade para avaliar a proximidade de uma dada distribuição F, relativa à família gerada pela distribuição normal padrão Φ,

$$\mathcal{G}_\Phi = \left\{G : \; G(x) = \Phi\left(\frac{x-\mu}{\sigma}\right), \; \mu \in R, \; \sigma > 0\right\},$$

e estendida, sob condições de regularidade, para distribuições com segundo momento finito. Para tal, a estatística central utilizada foi

$$R_n = 1 - \frac{\left(\int_0^1 F_n^{-1}(t)\Phi^{-1}(t)dt\right)^2}{S_n^2},$$

em que S_n^2 é a variância amostral. A estatística R_n mede o desvio da normalidade e pode ser interpretada como uma medida de correlação. Em [11], nota-se que existe uma sequência de constantes $\{a_n\}$, tal que, assintoticamente, $nR_n - a_n$ comporta-se como

$$\int_0^1 \frac{B^2(t) - EB^2(t)}{\phi^2(\Phi^{-1}(t))} dt - \left(\int_0^1 \frac{B(t)dt}{\phi(\Phi^{-1}(t))} \right)^2 - \left(\int_0^1 \frac{B(t)\Phi^{-1}(t)dt}{\phi(\Phi^{-1}(t))} \right)^2,$$

onde ϕ é a densidade da distribuição normal padrão.

A distância de Wasserstein (1970) é, de fato, um caso particular, $r = 1$, da distância de Mallows (1972) dada por:

$$d_r(F, G) = \inf_{(X,Y)} \{E(|X - Y|^r)\}^{1/r} \ , \quad r > 0,$$

onde o ínfimo é tomado sobre todos os vetores aleatórios (X, Y), com distribuições marginais F e G respectivamente. Denominaremos esta distância por Mallows-Wasserstein (M-W). A nossa tentativa de utilizar a distância d_α para desenvolver estatísticas análogas a R_n para distribuições α-estáveis apresentou dificuldades que não pudemos contornar pois temos resultados equivalentes aos do Teorema do Limite Central para distribuições estáveis que precisariam ser estendidos para incluir resultados quanto à convergência ao movimento de Lévy, e que seriam semelhantes ao Teorema de Donsker para o movimento Browniano. Além disso, seriam necessários resultados assintóticos semelhantes aos das pontes Brownianas (possivelmente, pontes de Levy).

Por outro lado, em [17] foi introduzida a noção de distância M-W ponderada e mostrada a aplicabilidade de testes de correlação ponderados para a família Weibull de escala. Também em [6], [8] e, mais recentemente em 2009, constatamos em ([9]) que estes testes ponderados podem ser utilizados para outras famílias de distribuições, incluindo a família Gamma, Lognormal e Gumbel de escala. Os métodos propostos são muito restritivos e seu uso não se estende a importantes classes de distribuição de cauda pesada, tal como as estáveis e as demais distribuições no domínio de atração para as extremais de Fréchet.

Diferentemente desses trabalhos, porém inspiradas neles, optaremos por trabalhar com distribuições ponderadas para contornar as dificuldades encontradas. O nosso Lema 2.2 exibe a equivalência destes procedimentos e permitirá a extensão destas estatísticas para distribuições estáveis e extremais de Fréchet. Exploraremos na Seção 2.3 as propriedades das distribuições ponderadas e determinaremos o comportamento caudal (variação regular) das funções de distribuições ponderadas, associadas a distribuições α-estáveis e extremais. Nosso Lema 2.1 ilustra o efeito da função peso sobre o comportamento caudal da distribuição por meio da estatística de teste

$$R_n(w) = 1 - \frac{\left(\int_0^1 F_n^{-1}(t)G^{-1}(t)w(t)dt\right)^2}{S^2_{F_n}(w)}$$

e do uso da distância M-W ponderada

$$d^2_{2,w}(F,G) = \int_0^1 (F^{-1}(t) - G^{-1}(t))^2 w(t)dt.$$

Apresentaremos na Condição 2.2 uma classe geral de funções peso w. Para essa classe, o Lema 2.2 mostrará a conexão entre distribuições ponderadas e a distância de M-W ponderada. Na Seção 2.5, propomos testes de ajuste de bondade para as distribuições α-estáveis e extremais de Fréchet. Os principais resultados são os Teoremas 2.3 e 2.4, que estabelecem resultados assintóticos e constituem uma extensão de resultados anteriores.

No Capítulo 3, apresentaremos a estatística para testar a similaridade entre duas distribuições F e G, que possuem caudas pesadas. Uma métrica intuitiva para avaliar a similaridade é a da variação total

$$d_{TV}(P_F, P_G) = \sup_B |P_F(B) - P_G(B)|,$$

onde P_F e P_G são, respectivamente, as medidas de probabilidade associadas às distribuições F e G. Na busca de similaridade, podemos utilizar procedimentos e testes que possam avaliar, para dado $0 \leq \varepsilon < 1$, a condição

$$d_{TV}(P_F, P_G) \leq \varepsilon \qquad (\varepsilon - \text{similaridade}).$$

A nossa abordagem tem como base os trabalhos [12], [14], [15] e [19]. Assim, motivadas por aplicações em que não são necessárias que as distribuições avaliadas coincidam, foram introduzidos vários conceitos de similaridade e ajustes parciais de medidas de probabilidade para possibilitar testar hipóteses do tipo:

$$H_0 : \tau_\gamma(F,G) > \delta > 0 \quad \text{vs} \quad H_1 : \tau_\gamma(F,G) \leq \delta,$$

onde $0 \leq \gamma < 1$ e, para $h_0(.)$ convenientemente escolhido

$$\tau_\gamma(F,G) = \left(\int_0^1 (F^{-1}(t) - G^{-1}(t))^2 h_0(t)dt\right)^{\frac{1}{2}}.$$

Em [14], sob condições que incluem o quarto momento finito, foi obtida a distribuição assintótica de $\sqrt{n}(\tau_\gamma(F_n, G) - \tau_\gamma(F, G))$.

A hipótese do quarto momento finito inviabiliza o uso de $\tau_\gamma(F_n, G)$ para distribuições α-estáveis, pois se $X \stackrel{d}{=} S_\alpha(\sigma, \beta, \mu)$ temos $E|X|^{\alpha'} = \infty$ para todo $\alpha' > \alpha$. Na Seção 3.2, apresentaremos vários conceitos e resultados relacionados à similaridade de distribuições e que motivaram a escolha da estatística que vai desempenhar o papel de $\tau_\gamma(F_n, G)$ para distribuições com caudas regularmente variantes. Apresentaremos na Condição 3.2 uma classe adequada de funções pesos, que garantirão a obtenção, no Teorema 3.2, da distribuição assintótica da estatística $\sqrt{n}(d_{2,w}(F_n, G) - d_{2,w}(F, G))$.

Capítulo 1

Preliminares

1.1 Introdução

Os testes de ajuste de bondade e similaridade para distribuições α-estáveis e extremais de Fréchet, a serem propostos nesta tese, têm estatísticas baseadas em processos empíricos e fazem uso da distância de Mallows-Wasserstein (M-W), que constitui uma métrica no espaço das funções de distribuição.

Na Seção 1.2, apresentaremos as distribuições assintóticas de funcionais apropriados dos quantis de processos empíricos. Veremos que estas distribuições são funcionais das pontes Brownianas. Fato este é preponderante e fundamental para motivar a escolha das nossas estatísticas. Como referência básica sugerimos [41].

Na Seção 1.3, citaremos algumas propriedades da distância de M-W e o teorema de representação, que permite o cálculo dessa distância para distribuições que possuem média finita (ver, por exemplo, [3] ou [20]). Além disso, veremos que a distância de M-W, entre a função de distribuição empírica e a sua respectiva função de distribuição, pode ser expressa como um funcional do processo quantil empírico ([5]).

Na Seção 1.4, introduziremos as funções de variação regular, que são fundamentais para caracterizar as caudas das distribuições estáveis e extremais de Fréchet, e o Teorema de Karamata, que estabelece propriedades das integrais de função de variação regular. Apresentaremos também a relação assintótica envolvendo a cauda de distribuições regularmente variantes e a sua densidade correspondente. Como referências, citamos os textos [4] e [39].

A Seção 1.5 é dedicada à apresentação dos principais conceitos, propriedades e resultados clássicos das distribuições estáveis, tais como: a estabilidade ou invariância sob a adição, a natureza assintótica Pareto das caudas e o fato dessas distribuições serem os únicos limites assintóticos possíveis para soma de variáveis aleatórias. As principais referências bibliográficas utilizadas foram [21] e [40].

Na Seção 1.6, destacaremos alguns resultados referentes ao domínio de atração das distribuições extremais, em especial, da distribuição extremal Fréchet, que será explorada em maior detalhe no Capítulo 2 (Seção 2.5). Veremos que existe uma estreita relação entre os domínios de atração de uma distribuição estável e de uma distribuição extremal Fréchet. Como referências, podemos citar [28], [35] e [39].

1.2 Processo Quantil Empírico

Sejam $X_1, ..., X_n$ variáveis aleatórias (v.a.'s) independentes e identicamente distribuídas (i.i.d) com função de distribuição (f.d) base F e considere a sua inversa generalizada (ou função quantil) F^{-1}, bem como a sua correspondente função de distribuição empírica F_n, dadas respectivamente por

$$F^{-1}(t) = \inf\{x : F(x) \geq t\}, \quad 0 < t < 1$$

e

$$F_n(x) = \frac{\sum_{i=1}^{n} I_{\{X_i \leq x\}}}{n},$$

onde I_A denota a função indicadora de A. Naturalmente, se $X_{(1,n)} \leq X_{(2,n)} \leq ... \leq X_{(n,n)}$ são as estatísticas de ordem da amostra $\{X_1, ..., X_n\}$, podemos reescrever F_n como

$$F_n(x) = \begin{cases} 0 & \text{se } X_{(1,n)} > x \\ \dfrac{k}{n} & \text{se } X_{(k,n)} \leq x < X_{(k+1,n)},\ k = 1, 2, ..., n-1 \\ 1 & \text{se } X_{(n,n)} \leq x. \end{cases}$$

E, similarmente, temos também a função quantil empírica (ou amostral)

$$F_n^{-1}(t) = X_{(k,n)} \text{ se } \frac{k-1}{n} < t \leq \frac{k}{n}, \quad k = 1, 2, ..., n.$$

Com base nesses elementos, temos as renomadas estatísticas de Kolmogorov e de Kolmogorov-Smirnov

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

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

e

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

A grande vantagem das estatísticas D_n, D_n^+ e D_n^- reside no fato de serem livres da distribuição base F. Pois, por exemplo, se F for uma distribuição contínua e $U_{(1,n)} \leq U_{(2,n)} \leq ... \leq U_{(n,n)}$ forem as estatísticas de ordem de uma amostra da distribuição uniforme em $[0,1]$ temos

$$\begin{aligned} D_n^+ &= \max_{0\leq i\leq n} \sup_{X_{(i,n)}\leq x<X_{(i+1,n)}} \left[\frac{i}{n} - F(x)\right] \\ &= \max_{0\leq i\leq n} \left[\frac{i}{n} - \inf_{X_{(i,n)}\leq x<X_{(i+1,n)}} F(x)\right] \\ &= \max_{0\leq i\leq n} \left[\frac{i}{n} - F(X_{(i,n)})\right] \\ &= \max_{0\leq i\leq n} \left[\frac{i}{n} - U_{(i,n)}\right]. \end{aligned}$$

A última igualdade segue do fato de $F(X_i)$ ter distribuição uniforme para todo i. Note que, se U_n representa a distribuição empírica da distribuição uniforme U, temos pela Lei dos Grandes Numeros e pelo Teorema do Limite Central que

$$\sqrt{n}(U_n(t) - t) \xrightarrow{d} N(0, t(1-t)),$$

onde $\xrightarrow{d}$ representa convergência em distribuição e $\sqrt{n}(U_n(t)-t)$ é o processo empírico associado à distribuição U (processo empírico uniforme). Por outro lado, a ponte Browniana $B(t)$ tem distribuição $N(0,t(1-t))$.

Definição 1.1. *Um processo Gaussiano $\{B(t) : \ 0 \leq t \leq 1\}$ é chamado de uma ponte Browniana se a covariância é dada por cov$\{B(t), B(s)\} = \min\{s,t\} - st$. Ou, equivalentemente, $\{B(t) \overset{d}{=} W(t) - tW(1) : \ 0 \leq t \leq 1\}$, onde $\{W(t) : \ t \geq 0\}$ é o movimento Browniano padrão.*

As considerações acima nos levam a fazer uso do processo quantil empírico da distribuição F, convenientemente ajustado, e "justificam" a aproximação dos processos empíricos por pontes Brownianas.

Definição 1.2. *Dada uma função de distribuição F com função de densidade f, o processo quantil empírico ajustado é definido por*

$$\rho_{n,F}(t) = \sqrt{n}(F_n^{-1}(t) - F^{-1}(t))f(F^{-1}(t)), \ \ 0 \leq t \leq 1. \tag{1.1}$$

Enunciaremos a seguir os resultados de [5], que garantem a aproximação e fornecem taxas de convergência de processos quantis empíricos por pontes Brownianas.

Condição 1.1. *Seja F uma f.d. com a_F e b_F definidos por*

$$a_F = \sup\{x : F(x) = 0\} \quad e \quad b_F = \inf\{x : F(x) = 1\}. \tag{1.2}$$

Assuma que:

(a) F é duas vezes diferenciável em (a_F, b_F).

(b) $F^{'}(x) = f(x) > 0$, $\forall\ x \in (a_F, b_F)$, onde $' := \dfrac{d}{dx}$.

(c) Para algum $\gamma > 0$ temos

$$\sup_{0\leq t\leq 1} \frac{t(1-t)|f^{'}(F^{-1}(t))|}{f^2(F^{-1}(t))} \leq \gamma < \infty. \tag{1.3}$$

Teorema 1.1. *(Csorgo (1983) ([5]) Assuma que F satisfaça Condição 1.1 e seja $\rho_{n,F}(t)$ definido por (1.1). Então existe uma sequência de pontes Brownianas $\{B_n(t) : 0 \leq t \leq 1\}$, tal que*

$$n^{\frac{1}{2}-\nu} \sup_{\frac{1}{n}\leq t\leq 1-\frac{1}{n}} \frac{|\rho_{n,F}(t) - B_n(t)|}{(t(1-t))^{\nu}} = \begin{cases} O(\log n) & se\ \nu = 0 \\ O(1) & se\ 0 < \nu \leq \frac{1}{2}. \end{cases}$$

Dadas funções g e h, dizemos que $g = O(h)$ se $\limsup_{x\to\infty} \dfrac{|g(x)|}{h(x)} < \infty$.

Corolário 1.1. *Suponha que as hipóteses do Teorema 1.1 estejam satisfeitas, então existe uma sequência de pontes Brownianas $\{B_n(t) : 0 \leq t \leq 1\}$, tal que se F é uma f.d com suporte limitado então*

$$\sup_{0\leq t\leq 1} |\rho_{n,F}(t) - B_n(t)| \xrightarrow{p} 0,$$

caso contrário,

$$\sup_{\frac{1}{n+1}\leq t\leq \frac{n}{n+1}} |\rho_{n,F}(t) - B_n(t)| \xrightarrow{p} 0,$$

onde $\xrightarrow{p}$ simboliza a convergência em probabilidade.

Teorema 1.2. *(Csorgo (1983) ([5]) Suponha que as hipóteses do Teorema 1.1 estão satisfeitas e que* $\{B(t) : 0 \leq t \leq 1\}$ *é uma ponte Browniana. Então*

$$\int_0^1 \frac{B(t)dt}{f^2(F^{-1}(t))} < \infty \ \ se \ \ \int_0^1 \frac{t(1-t)dt}{f^2(F^{-1}(t))} < \infty$$

e

$$\int_0^1 \frac{B(t)dt}{f^2(F^{-1}(t))} = \infty \ \ se \ \ \int_0^1 \frac{t(1-t)dt}{f^2(F^{-1}(t))} = \infty.$$

1.3 Distância de Mallows-Wasserstein

A distância de Mallows (1972) entre duas funções de distribuições F e G generaliza a considerada por Wasserstein em 1970 (caso: $r = 1$). Assim, na literatura, o nome distância de Wasserstein também tem sido usado ao invés de Mallows. Neste texto usaremos a denominação distância de M-W (Mallows-Wasserstein).

Definição 1.3. *Para* $r > 0$*, a distância-r de M-W entre as distribuições* F *e* G *é definida por*

$$d_r(F, G) = \inf_{(X,Y)} \{E(|X-Y|^r)\}^{1/r}, \ \ X \stackrel{d}{=} F, \ Y \stackrel{d}{=} G, \tag{1.4}$$

onde $\stackrel{d}{=}$ *simboliza igualdade em distribuição, isto é, o ínfimo é tomado sobre todos os vetores aleatórios, tais que as distribuições marginais são* F *e* G*.*

Esta distância constitui uma métrica para o espaço das distribuições de probabilidade, pois valem as seguintes relações métricas

$$d_r(F, G) \leq d_r(F, F') + d_r(F', G), \ \ r \geq 1$$

e

$$d_r^r(F, G) \leq d_r^r(F, F') + d_r^r(F', G), \ \ 0 < r < 1.$$

O teorema a seguir mostra que o cálculo desta distância é simples quando $r \geq 1$.

Teorema 1.3. *(Teorema da Representação da Distância de M-W) Seja* $r \geq 1$ *e assuma que* $\int |x|^r dF(x) < \infty$ *e* $\int |y|^r dG(y) < \infty$*. Então*

$$d_r^r(F, G) = E\{|F^{-1}(U) - G^{-1}(U)|^r\},$$

onde U *é a distribuição uniforme no intervalo* $(0, 1)$*.*

Estaremos particularmente interessados na métrica em

$$\mathcal{L}_2 = \left\{ \mathrm{F} : \int x^2 dF(x) < \infty \right\} \tag{1.5}$$

dada por

$$d_2^2(F, G) = \int_0^1 (F^{-1}(t) - G^{-1}(t))^2 dt = E|F^{-1}(U) - G^{-1}(U))|^2. \tag{1.6}$$

Note que $d_2(F_n, F)$, distância entre a função de distribuição empírica e a função de distribuição base F, pode ser expressa por

$$\sqrt{n} d_2(F_n, F) = \left(n \int_0^1 (F_n^{-1}(t) - F^{-1}(t))^2 dt \right)^{\frac{1}{2}}$$

e reescrita em termos de processo quantil empírico ajustado

$$\rho_{n,F}(t) := \sqrt{n}(F_n^{-1}(t) - F^{-1}(t)) f(F^{-1}(t))$$

como

$$\sqrt{n} d_2(F_n, F) = \left(\int_0^1 \frac{\rho_{n,F}^2(t)}{f^2(F^{-1}(t))} dt \right)^{\frac{1}{2}}. \tag{1.7}$$

As estatísticas que analisaremos nos próximos capítulos têm $\sqrt{n} d_2(F_n, F)$ como sua principal componente e o Corolário 1.1 nos permite inferir que a distribuição assintótica de $\sqrt{n} d_2(F_n, F)$ seja dada em termos de pontes Brownianas, mais precisamente, de acordo com o Teorema de Donsker, esperamos obter como distribuição assintótica

$$\left(\int_0^1 \frac{B^2(t)}{f^2(F^{-1}(t))} dt \right)^{\frac{1}{2}}.$$

De acordo com o Teorema de Donsker, a distribuição de um funcional contínuo de um processo empírico uniforme converge para o correspondente funcional de uma ponte Browniana, para mais detalhes vide [41].

Para os nossos testes de ajuste de bondade faremos uso de famílias de escala-locação, geradas pela distribuição base G_0

$$\mathcal{G}_{G_0} = \left\{ G \,:\, G(x) = G_0\left(\frac{x - \mu}{\sigma} \right), \mu \in \mathbb{R},\ \sigma > 0 \right\}.$$

E podemos definir a distância-2 de M-W por:

Definição 1.4. *A distância-2 de M-W entre uma função de distribuição F e uma família de escala-locação de distribuições gerada por G_0 é dada por:*

$$d_2(F, \mathcal{G}_{G_0}) = \inf\{d_2(F, G) \,:\, G \in \mathcal{G}_{G_0}\}. \tag{1.8}$$

1.4 Variação Regular

A teoria de funções de variação regular é a principal ferramenta analítica para o estudo de distribuições estáveis, pois essas distribuições têm caudas cujo comportamento pode ser descrito com o uso dessas funções. Além disso, o Teorema de Karamata e similares têm papel preponderante para provar os nossos resultados relacionados aos testes de ajuste de bondade para dados empíricos, com respeito a uma família de escala-locação de distribuições geradas por uma v.a. α-estável e por extremais de Fréchet.

Definição 1.5. *Dizemos que uma função $\ell : \mathbb{R}^+ \longmapsto \mathbb{R}^+$ varia lentamente no infinito (lentamente variante ou de variação lenta) se*

$$\lim_{t\to\infty} \frac{\ell(tx)}{\ell(t)} = 1, \ x > 0.$$

Por exemplo, toda função de distribuição varia lentamente no infinito. Pela próxima definição, vemos que as funções de variação lenta são funções de variação regular com expoente $\rho = 0$.

Definição 1.6. *Dizemos que uma função $U : \mathbb{R}^+ \longmapsto \mathbb{R}^+$ varia regularmente no infinito (regularmente variante ou ρ-variante) com expoente $\rho \in \mathbb{R}$ constante se:*

$$\lim_{t\to\infty} \frac{U(tx)}{U(t)} = x^\rho, \ x > 0.$$

Neste caso, escrevemos $U \in RV_\rho$.

Similarmente, podemos definir variação regular no zero substituindo $t \to \infty$ por $t \to 0$, pois uma função U é de variação regular no zero se a função $U\left(\frac{1}{t}\right)$ for de variação regular no infinito ($t \to 0$). Para qualquer outro ponto $a > 0$, podemos definir a variação regular de U em a quando $U\left(\frac{1}{a-t}\right)$ for de variação regular no infinito ($t \to a$).

Nas proposições a seguir serão apresentadas propriedades operacionais das funções regularmente variantes e também relações entre a monotonicidade e o expoente de variação.

Proposição 1.1. *Sejam* $U_1 : \mathbb{R}^+ \longmapsto \mathbb{R}^+$ ρ_1*-variante e* $U_2 : \mathbb{R}^+ \longmapsto \mathbb{R}^+$ ρ_2*-variante. Então temos:*

(i) $U_1.U_2$ *é* $(\rho_1 + \rho_2)$*-variante.*

(ii) $\dfrac{U_1}{U_2}$ *é* $(\rho_1 - \rho_2)$*-variante.*

(iii) Se $\lim\limits_{x\to\infty} U_2(x) = \infty$ *então* $U_1 o U_2$ *é* $\rho_1.\rho_2$*-variante.*

(iv) $U_1 + U_2$ *é* $\max\{\rho_1, \rho_2\}$*-variante.*

(v) $U_1 - U_2$ *é* $\min\{\rho_1, \rho_2\}$*-variante.*

Proposição 1.2. *Seja* $U : \mathbb{R}^+ \longmapsto \mathbb{R}^+ \rho$*-variante. Temos:*

(i) Se U *é crescente então* $\rho \geq 0$.

(ii) Se U *é decrescente então* $\rho \leq 0$.

(iii) Se U *é não decrescente e* $\rho \in [0, \infty)$ *então* U^* *é* $\dfrac{1}{\rho}$*-variante, onde*

$$U^*(x) := \inf\{y | U(y) \geq x\}. \tag{1.9}$$

(iv) Se U *é não crescente e* $\rho \in (-\infty, 0]$ *então* U^{**} *é* $-\dfrac{1}{\rho}$*-variante, onde*

$$U^{**}(x) := \inf\left\{ y \,\middle|\, U(y) \leq \frac{1}{x} \right\}. \tag{1.10}$$

Além disso,

(v) Se $U : \mathbb{R}^+ \longmapsto \mathbb{R}^+$ *é* ρ*-variante para algum* $\rho \neq 0$*, então,*

$$\lim_{x\to\infty} U(x) = \begin{cases} \infty & se \;\; \rho > 0 \\ 0 & se \;\; \rho < 0. \end{cases}$$

Os teoremas sobre integrais e derivadas de funções de variação regular são centrais para deduzir os nossos resultados. A seguinte notação será utilizada: para funções g e h

$$g \sim h \quad \Leftrightarrow \quad \lim_{x\to\infty} \frac{g(x)}{h(x)} = 1. \tag{1.11}$$

Teorema 1.4. *(Teorema de Karamata) Seja* $U : \mathbb{R}^+ \longmapsto \mathbb{R}^+$ *uma função* ρ*-variante com* $\rho < -1$*. Então*

$$\lim_{x\to\infty} \frac{xU(x)}{\displaystyle\int_x^\infty U(t)dt} = -\rho - 1$$

e

$$\int_x^\infty U(t)dt < \infty \quad e \quad \int_x^\infty U(t)dt \in RV_{\rho+1}. \tag{1.12}$$

De acordo com o próximo resultado, toda função regularmente variante admite uma representação em termos de uma função lentamente variante.

Proposição 1.3. *Seja $U : \mathbb{R}^+ \longmapsto \mathbb{R}^+$ ρ-variante no infinito então para alguma função lentamente variante ℓ temos que*

$$U(x) = x^\rho \ell(x), \;\; x > 0. \tag{1.13}$$

O resultado a seguir é crucial para a diferenciabilidade de funções regularmente variantes.

Teorema 1.5. *Seja $U(x) = \int_0^x u(y)dy$ $\left(ou \int_x^\infty u(y)dy\right)$ onde $u(.)$ é monótona em (z,∞) para algum $z > 0$. Se para algum $c \geq 0, \rho \in \mathbb{R}$ e $\ell \in RV_0$ temos*

$$U(x) \sim cx^\rho \ell(x) \quad quando \; x \to \infty,$$

então

$$u(x) \sim c\rho x^{\rho-1}\ell(x) \quad quando \; x \to \infty,$$

com $\sim$ no sentido de (1.11).

Como aplicação do Teorema de Karamata (Teorema 1.4), obtemos a relação assintótica entre caudas regularmente variantes de distribuições e suas respectivas densidades.

Proposição 1.4. *Seja U positiva e limitada com $U \in RV_\rho$ e seja γ uma constante tal que $\gamma + \rho < 0$, então*

$$\lim_{y\to\infty} \frac{\int_y^\infty t^\gamma dU(t)}{y^\gamma U(y)} = \frac{-\rho}{\gamma+\rho}.$$

1.5 Distribuições Estáveis

O uso de distribuições estáveis não Gaussianas na modelagem de dados financeiros foi, pioneiramente, apresentado por Mandelbrot em 1963. Constatou-se que distribuições α-estáveis com $0 < \alpha < 2$ se ajustam melhor a dados financeiros que apresentam alta variabilidade do que o modelo normal, comumente usado. Em [22], observamos que a mudança de preço dos ativos em qualquer intervalo de tempo pode ser considerada como a soma das mudanças de preço ocorridas de transação para transação durante aquele intervalo de tempo. Se as transações consideradas ocorrem uniformemente ao longo de qualquer intervalo de tempo e se as mudanças de preço entre uma transação e outra podem ser consideradas independentes e identicamente distribuídas, temos que a mudança de preço diária, semanal ou mensal será a mesma. Veremos que as distribuições estáveis têm esta característica da estabilidade, isto é, a soma de cópias independentes de variáveis aleatórias estáveis, feita alguma correção de escala e locação, é novamente a mesma distribuição estável. Em suma, as distribuições α-estáveis, $S_\alpha(\sigma, \beta, \mu)$ com $0 < \alpha \leq 2$, possuem a propriedade de divisibilidade infinita.

Veremos que as três características mais importantes das distribuições α-estáveis são a estabilidade ou invariância sob a adição, a natureza assintótica Pareto das caudas $(0 < \alpha < 2)$ e o fato dessas distribuições serem os únicos limites possíveis para soma de variáveis aleatórias.

Definição 1.7. *Para $0 < \alpha \leq 2$, dizemos que $X \sim S_\alpha(\sigma, \beta, \mu)$ é uma distribuição α-estável com índice de estabilidade α, parâmetro de escala $\sigma > 0$, parâmetro de assimetria $|\beta| \leq 1$ e parâmetro de locação $\mu \in \mathbb{R}$, se para qualquer $n \geq 2$ temos que*

$$X_1 + X_2 + ... + X_n \overset{d}{=} n^{1/\alpha} X + \mu(n - n^{1/\alpha}),$$

se $\alpha \neq 1$, e

$$X_1 + X_2 + ... + X_n \overset{d}{=} nX + \frac{2\sigma\beta n \ln n}{\pi},$$

se $\alpha = 1$, onde $X_1, ..., X_n$ são cópias independentes da v.a. X. Neste caso, dizemos que X é α-estável.

Note que, se $\beta = 0$ e $\alpha = 2$, temos o caso Gaussiano com média μ e variância σ^2. Outras distribuições familiares que são estáveis: $S_1(\sigma, 0, 0)$ que é a distribuição de Cauchy e $S_{1/2}(\sigma, 0, 0)$, que é a distribuição de Lévy. Outra definição de variável α-estável é dada pela sua função característica.

Definição 1.8. *A v.a. X é α-estável se a sua função característica $\Phi_X(t)$ é dada por:*

$$\Phi_X(t) = \exp\{it\mu - \sigma|t|^\alpha(1 - i\beta sinal(t) z(t, \alpha))\}, \tag{1.14}$$

onde $\alpha \in (0, 2]$, $\mu \in \mathbb{R}$, $\sigma > 0$, $|\beta| \leq 1$ e

$$Z(t, \alpha) = \begin{cases} \tan \dfrac{\pi\alpha}{2} & se \ \ \alpha \neq 1 \\ \dfrac{-2ln|t|}{\pi} & se \ \ \alpha = 1. \end{cases}$$

Proposição 1.5. *Seja X uma v.a. com distribuição $S_\alpha(\sigma, \beta, \mu)$.*

(i) Se $\beta = \mu = 0$. Então a v.a. X tem distribuição simétrica em torno do zero (α-estável simétrica). Nesse caso, escrevemos $S\alpha S$.

(ii) Se $0 < \alpha < 2$ então a distribuição tem cauda pesada, pois $E(|X|^{\alpha'}) < \infty$ para $0 < \alpha' < \alpha$ e $E(|X|^{\alpha'}) = \infty$ para $\alpha' \geq \alpha$.

(iii) A distribuição α-estável possui função densidade infinitamente diferenciável.

Conforme a proposição a seguir, as distribuições α-estáveis com $0 < \alpha < 2$ têm um comportamento assintótico caudal do tipo Pareto que dependem dos parâmetros α e β.

Proposição 1.6. *Seja $F_\alpha = S_\alpha(\sigma, \beta, \mu)$ e assuma que $0 < \alpha < 2$. Então existem uma função lentamente variante ℓ e $0 \leq p, q \leq 1$ com $p + q = 1$, tais que para $x \to \infty$*

$$\bar{F}_\alpha(x) = 1 - F_\alpha(x) \sim p\frac{2-\alpha}{\alpha}x^{-\alpha}\ell(x) \tag{1.15}$$

e

$$F_\alpha(-x) \sim q\frac{2-\alpha}{\alpha}x^{-\alpha}\ell(x). \tag{1.16}$$

Ou alternativamente, existem constantes C_α e C'_α, tais que para $x \to \infty$

$$\bar{F}_\alpha(x) = 1 - F_\alpha(x) \sim C_\alpha x^{-\alpha}$$

e

$$F_\alpha(-x) \sim C'_\alpha x^{-\alpha},$$

com $\sim$ no sentido de (1.11).

Teorema 1.6. *Uma variável aleatória X tem distribuição α-estável se e somente se X for o limite em distribuição de somas normalizadas de v.a.'s i.i.d, ou seja, existem v.a.'s $X_1, ..., X_n$ i.i.d e sequências de constantes $a_n > 0$ e $b_n \in \mathbb{R}$, tais que*

$$\frac{X_1 + ... + X_n - b_n}{a_n} \xrightarrow{d} X.$$

Definição 1.9. *Dizemos que uma função de distribuiçao F pertence ao domínio de atração de uma distribuição α-estável G_α se existirem sequências de constantes $a_n \in \mathbb{R}$ e $b_n > 0$, tais que*

$$\frac{(X_1 + ... + X_n) - b_n}{a_n} \xrightarrow{d} Y,$$

onde $Y \stackrel{d}{=} G_\alpha$ *e* $X_1, ..., X_n$ *são cópias independentes de* $X \stackrel{d}{=} F$*. Nesse caso, usamos a notação* $F \in \mathcal{D}(G_\alpha)$ *(ou* $X \in \mathcal{D}(G_\alpha)$*).*

A noção de domínio de atração nos permite formular um novo enunciado para o Teorema 1.6 e também explicitar o comportamento caudal das distribuições que pertencem ao domínio de atração das estáveis.

Teorema 1.7. *Uma função de distribuição possui domínio de atração não vazio se e somente se for uma distribuição estável.*

Teorema 1.8. *Para* $0 < \alpha < 2$*, seja* G_α *uma distribuição* α*-estável e assuma que* $G \in \mathcal{D}(G_\alpha)$*. Então*

$$G(-x) \sim x^{-\alpha}\ell(x) \ \ \text{para} \ \ x \to \infty \tag{1.17}$$

e

$$\bar{G}(x) \sim x^{-\alpha}\ell(x) \ \ \text{para} \ \ x \to \infty, \tag{1.18}$$

em que $\sim$ *está no sentido de (1.11) e* $\ell(.)$ *é uma função lentamente variante. Além disso, se* $X \stackrel{d}{=} G$ *então*

$$\begin{cases} E|X|^{\alpha'} < \infty & \text{se } \alpha' < \alpha \\ E|X|^{\alpha'} = \infty & \text{se } \alpha' > \alpha. \end{cases}$$

Em particular, temos

$$\begin{cases} VarX = \infty & \text{se } \alpha < 2 \\ E|X| < \infty & \text{se } \alpha > 1 \\ E|X| = \infty & \text{se } \alpha < 1. \end{cases}$$

1.6 Distribuições Extremais

A Teoria dos Valores Extremos nos fornece uma classe importante de distribuições de probabilidade, a saber, as distribuições assintóticas das variáveis

$$X_{(n,n)} := \max\{X_1, ..., X_n\} \ \ \text{e} \ \ X_{(1,n)} := \min\{X_1, ..., X_n\},$$

onde $\{X_1, ..., X_n\}$ é uma amostra da distribuição base F. Para o máximo, o estudo analítico da expressão $(1 - \bar{F})^n$ nos permite concluir que existem apenas três tipos de distribuições maximais. Os estudos pioneiros deste tema pertencem ao matemático Mauricé Fréchet(1878 - 1973) e aos estatísticos Ronald Fischer(1890 - 1962) e Leonard

Tippet, seguidos pelas contribuições devidas a Boris Gnedenko(1912 - 1995) e a consolidação teórica por parte de Emil Gumbel(1891 - 1966) e podem ser encontrados em [25], [27] e [30].

Definição 1.10. *Sejam $X_1, ..., X_n$ v.a.'s i.i.d com distribuição comum F e assuma que existam sequências de constantes $a_n > 0$ e b_n tais que*

$$\lim_{n\to\infty} P\left(\frac{\max\{X_1, ..., X_n\} - b_n}{a_n} \leq x\right) = \lim_{n\to\infty} F^n(a_n x + b_n) = H(x), \tag{1.19}$$

onde H é uma distribuição não degenerada. Então H é chamada de distribuição maximal e dizemos que F pertence ao domínio de atração maximal da distribuição H (notação: $F \in \mathcal{D}_{\max}(H)$). Similarmente, se temos

$$\lim_{n\to\infty} P\left(\frac{\min\{X_1, ..., X_n\} - b_n}{a_n} \leq x\right) = \lim_{n\to\infty} \{1 - [1 - F(a_n x + b_n)]^n\} = H(x),$$

então H é chamada de distribuição minimal e dizemos que F pertence ao domínio de atração minimal da distribuição H (notação: $F \in \mathcal{D}_{\min}(H)$).

Como o tratamento do mínimo é simétrico ao do máximo, enunciamos os próximos resultados apenas para as extremais maximais.

Teorema 1.9. *(Tippet-Fischer) Existem apenas três classes de distribuições maximais: sejam $X_1, ..., X_n$ v.a.'s i.i.d e assuma que $H(x)$ satisfaça (1.19). Então $H(x)$ é uma das seguintes distribuições:*

Tipo Gumbel:

$$\lambda(x) = e^{-e^{-x}} \qquad -\infty < x < \infty;$$

Tipo Fréchet:

$$\Phi_\alpha(x) = \begin{cases} e^{-x^{-\alpha}} & se \;\; x \geq 0 \\ 0 & se \;\; x < 0; \end{cases}$$

Tipo Weibull:

$$\psi_\alpha(x) = \begin{cases} e^{-(-x)^{\alpha}} & se \;\; x \leq 0 \\ 1 & se \;\; x > 0; \end{cases}$$

onde α é uma constante positiva.

O próximo teorema caracteriza o domínio de atração maximal de Fréchet quanto à variação regular e essa caracterização nos permitirá, no Capítulo 2, estabelecer a conexão com as distribuições estáveis.

Teorema 1.10. *Seja Φ_α a distribuição maximal de Fréchet e seja $\mathcal{D}_{\max}(\Phi_\alpha)$ seu domínio de atração maximal. Então $F \in \mathcal{D}_{\max}(\Phi_\alpha)$ se e somente se $1 - F \in RV_{-\alpha}$, isto é, se e somente se a cauda à direita de $\bar{F}$ é regularmente variante com índice $-\alpha$,*

$$\lim_{x\to\infty} \frac{\bar{F}(xt)}{\bar{F}(x)} = t^{-\alpha},\ t > 0.$$

Além disso, se

$$\frac{\max\{X_1, ..., X_n\}}{a_n} \xrightarrow{d} X, \tag{1.20}$$

onde X tem distribuição Fréchet

$$\Phi_\alpha(x) = \begin{cases} e^{-x^{-\alpha}} & se\ \ x \geq 0 \\ 0 & se\ \ x < 0, \end{cases}$$

podemos escolher $a_n = F^{-1}\left(1 - \dfrac{1}{n}\right)$ e a_n satisfaz

$$\lim_{n\to\infty} n\bar{F}(a_n x) = x^{-\alpha}, \forall x > 0. \tag{1.21}$$

Capítulo 2

Testes de Ajuste de Bondade

2.1 Introdução

Considere a classe das distribuições α-estáveis com variância infinita $\{S_\alpha(\sigma,0,\mu),\ \mu \in \mathbb{R},\ \sigma > 0,\ 0 < \alpha < 2\}$ e a classe das distribuições extremais de Fréchet, $\{\Phi_\alpha,\ \alpha > 0\}$. Buscamos formular testes de ajuste de bondade para avaliar a proximidade de uma dada distribuição F relativa a essas classes. Veremos que, para dado índice α e fazendo-se uso da distância de Mallows-Wasserstein, podemos avaliar esta proximidade com relação à família de escala-locação gerada por $G_\alpha = S_\alpha(\sigma,0,\mu)$ ou Φ_α dada por:

$$\mathcal{G}_{G_\alpha} = \left\{ G:\ G(x) = G_\alpha\left(\frac{x-\mu}{\sigma}\right),\ \mu \in \mathbb{R},\ \sigma > 0 \right\}. \tag{2.1}$$

Para o caso $\alpha = 2$, isto é $S_2(1,0,0) = G_2 = \Phi$, foi proposto em [11] um teste de ajuste de bondade para avaliar a normalidade dos dados empíricos e a distância-2 de M-W foi utilizada com sucesso para derivar funcionais de processos quantis empíricos e determinar as correspondentes distribuições assintóticas. A relação fundamental

$$d_2^2(F,\mathcal{G}_\Phi) = \left(\inf\{d_2(F,G) | G \in \mathcal{G}_\Phi\}\right)^2 = \sigma_F^2 - \left(\int_0^1 F^{-1}(t)\Phi^{-1}(t)dt\right)^2 \tag{2.2}$$

permitiu a escolha da estatística adequada,

$$R_n := 1 - \frac{\left(\int_0^1 F_n^{-1}(t)\Phi^{-1}(t)dt\right)^2}{S_n^2}. \tag{2.3}$$

Aqui F_n^{-1} é a inversa generalizada da função de distribuição empírica, Φ é a função de distribuição da normal padrão, S_n^2 é a variância amostral. A estatística R_n mede o desvio da normalidade e pode ser interpretada como uma medida de correlação. Em

[11], mostra-se que existe uma sequência de constantes $\{a_n\}$ tal que, assintoticamente, $nR_n - a_n$ comporta-se como

$$\int_0^1 \frac{B^2(t) - E(B^2(t))}{\phi^2(\Phi^{-1}(t))} dt - \left(\int_0^1 \frac{B(t)dt}{\phi(\Phi^{-1}(t))} \right)^2 - \left(\int_0^1 \frac{B(t)\Phi^{-1}(t)dt}{\phi(\Phi^{-1}(t))} \right)^2, \tag{2.4}$$

onde $\{B(t)|\ 0 \leq t \leq 1\}$ é uma ponte Browniana e ϕ é a densidade da distribuição normal padrão.

A estatística resultante da substituição em (2.3) da distribuição Gaussiana Φ por uma distribuição $G_0 \in \mathcal{L}_2$, isto é, com segundo momento finito, foi apresentada em [12] como um teste de correlação com propriedades assintóticas idênticas ao do caso Gaussiano. De Wet, em uma discussão intitulada "*Testing Weibull as a case of correlation test*", introduziu a noção de distância de M-W ponderada e mostrou a aplicabilidade de testes de correlação ponderados para a família Weibull de escala (vide [17]). Em [6], [8] e, mais recentemente, em 2009, em ([9]), onstatou-se que estes testes ponderados poderiam ser utilizados para outras famílias de distribuição que incluíam as famílias Gamma, Lognormal e Gumbel de escala com a função de Wet como peso. Para estender a aplicabilidade do teste à família Pareto, outras escolhas de função peso foram necessárias, vide [6]. Outras discussões relacionadas a esses testes podem ser encontradas em [32] e [34].

Os métodos propostos são muito restritivos e seu uso não se estende a importantes classes de distribuição de cauda pesada, tal como as estáveis e as demais distribuições no domínio de atração para máximos da distribuição Fréchet. Vale ressaltar que as distribuições α-estáveis, $S_\alpha(\sigma, \beta, \mu)$, devido a propriedade de infinita divisibilidade, têm um papel fundamental no estudo do comportamento assintótico de somas parciais estabilizadas. Para contornar as dificuldades encontradas, proporemos o uso da distância de M-W entre distribuições ponderadas ao invés da distância de M-W ponderada usada por Csorgo e de Wet. O nosso Lema 2.2 exibe a equivalência desses procedimentos e permite a extensão dessas estatísticas para distribuições estáveis e extremais de Fréchet.

A Seção 2.2 é dedicada à citação de resultados anteriores, que motivaram a nossa abordagem. Para as distribuições α-estáveis, a alternativa de se fazer uso da distância-α de M-W apresentou dificuldades que não pudemos contornar, visto que ainda não dispomos de estudos sobre "pontes de Lévy", correspondentes às pontes Brownianas. Exploramos na Seção 2.3 as propriedades das distribuições ponderadas e determinamos o comportamento caudal (variação regular) das funções de distribuições ponderadas associadas a distribuições α-estáveis e extremais de Fréchet. Nosso Lema 2.1 ilustra o efeito da função peso sobre o comportamento caudal da distribuição. Apresentamos na

Condição 2.2 uma classe geral de funções peso. Para essa classe, o Lema 2.2 mostra a conexão entre distribuições ponderadas e a distância de M-W ponderada. O Lema 2.3 mostra que a relação em (2.2) pode ser obtida para distribuições com segundo momento ponderado finito. Na Seçao 2.4, apresentamos alguns exemplos ilustrativos dos nossos procedimentos.

Na Seção 2.5, propomos testes de ajuste de bondade, similares a (2.3), para as distribuições α-estáveis e extremais de Fréchet. Os principais resultados são os Teoremas 2.3 e 2.4, que estabelecem resultados assintóticos e constituem uma extensão do uso da metodologia apresentada na Seção 2.2. Ainda nessa seção, apresentamos vários resultados preliminares, de interesse matemático, que são necessários para a prova dos teoremas, vide Lema 2.4 a Lema 2.8.

2.2 Preliminares e Motivação

A estatística R_n, definida em (2.3), foi introduzida em [11] com a finalidade de testar $F \in \mathcal{G}_\Phi$, onde Φ é a função de distribuição da normal padrão e

$$\mathcal{G}_\Phi = \left\{ G : G(x) = \Phi\left(\frac{x-\mu}{\sigma}\right), \mu \in \mathbb{R},\ \sigma > 0 \right\}.$$

Como medida de desvio da normalidade foi proposto o uso da distância de Wasserstein (equivalentemente distância de M-W de ordem 2),

$$d_2(F, \mathcal{G}_\Phi) = \inf\{d_2(F, G) :\ G \in \mathcal{G}_\Phi\}.$$

Mostra-se que o ínfimo é atingido para G com média $\mu_G = \mu_F$ e variância $\sigma_G^2 = \left(\int_0^1 F^{-1}(t)\Phi^{-1}(t)dt\right)^2$. Nesse caso, temos

$$d_2^2(F, \mathcal{G}_\Phi) = \sigma_F^2 - \left(\int_0^1 F^{-1}(t)\Phi^{-1}(t)dt\right)^2,$$

o que justifica o uso de (2.3) como a estatística de teste. Os mesmos autores estenderam esses resultados para a família de escala-locação, $\mathcal{G}_{G_0}$, gerada por um distribuição G_0 de média zero e variância unitária, que enunciaremos a seguir. Para tanto considere a seguinte condição:

Condição 2.1. *Assuma que a f.d. F satisfaça a Condição 1.1 e que*

$$\int_0^1 \frac{t(1-t)}{f^2(F^{-1}(t))}dt < \infty. \tag{2.5}$$

Além disso, assuma que as distribuições empíricas associadas F_n satisfazem

$$n\int_{1-\frac{1}{n}}^{1}\left(F_n^{-1}(t)) - F^{-1}(t)\right)^2 dt \xrightarrow{p} 0 \tag{2.6}$$

e

$$n\int_{0}^{\frac{1}{n}}\left(F_n^{-1}(t) - F^{-1}(t)\right)^2 dt \xrightarrow{p} 0. \tag{2.7}$$

Teorema 2.1. *(Cuesta et al, (2000))([12]) Seja G_0 uma f.d. com média zero, variância unitária e função densidade g_0 positiva em (a_{G_0}, b_{G_0}), a_{G_0} e b_{G_0}, definidos por (1.2). Seja F uma f.d, satisfazendo a Condição 2.1, e seja R_n definido por*

$$R_n = 1 - \frac{\left(\int_0^1 F_n^{-1}(t)G_0^{-1}(t)dt\right)^2}{S_n^2}.$$

Então, se $F \in \mathcal{G}_{G_0}$, a distribuição limite de nR_n é dada por

$$\int_0^1 \frac{B^2(t)}{f^2(F^{-1}(t))}dt - \left(\int_0^1 \frac{B(t)dt}{f(F^{-1}(t))}\right)^2 - \left(\int_0^1 \frac{B(t)F^{-1}(t)}{f(F^{-1}(t))}dt\right)^2, \tag{2.8}$$

onde $\{B(t):\ 0 \le t \le 1\}$ é uma ponte Browniana e f é a função densidade de F.

Claramente, a restrição $G_0 \in \mathcal{L}_2$ reduz a aplicação desses testes ao caso F *versus* distribuições de cauda leve. Para contornar essa restrição e estender essas técnicas para distribuições de cauda pesada, em [16], de Wet propôs o uso da seguinte função peso:

Definição 2.1. *Seja F uma f.d. com densidade $f(x) > 0$ em (a_F, b_F), assuma que F é duas vezes diferenciável e que*

$$\lim_{x\downarrow a_F} x^2 f'(x) = 0 = \lim_{x\uparrow b_F} x^2 f'(x)$$

e

$$I = \int_{-\infty}^{\infty} L^2(x)f(x)dx < \infty \ \ \textit{onde} \ \ L(x) = -1 - \frac{xf'(x)}{f(x)}.$$

Nesse caso, definimos a função peso de de Wet associada a F como sendo $w_F : (0,1) \to [0,\infty)$ satisfazendo

$$w_F(t) := \frac{1}{I}\frac{L^{'}(F^{-1}(t))}{F^{-1}(t)}, \ \ 0 < t < 1. \tag{2.9}$$

Nos trabalhos [6], [7], [8], [9], [16] e [17], a função peso w_F foi usada com sucesso para as famílias Weibull, Gamma, Lognormal e Gumbel. Para a Pareto, a escolha da função de de Wet não pode ser feita, mas são exibidas em [6] duas funções peso para a família de escala Pareto. Os casos Gamma e Lognormal são tratados em [8], e, em [9], são apresentados os casos relativos às famílias Gumbel, Weibull e Pareto. Em [7], estes resultados foram estendidos para a família escala-locação. Nessas abordagens e em [10], para $w : (0,1) \longrightarrow [0,\infty)$ uma função Borel mensurável com $\int_0^1 w(t)dt = 1$, temos que a distância-2 de M-W ponderada foi utilizada como a métrica para o desvio e foram formuladas versões ponderadas das condições em (2.5), (2.6) e (2.7) conforme segue

$$\int_0^1 \frac{t(1-t)}{f^2(F^{-1}(t))} w(t)dt < \infty, \tag{2.10}$$

$$n \int_0^{\frac{1}{n}} \left(F_n^{-1}(t) - F^{-1}(t)\right)^2 w(t)dt \xrightarrow{p} 0 \tag{2.11}$$

e

$$n \int_{1-\frac{1}{n}}^1 \left(F_n^{-1}(t) - F^{-1}(t)\right)^2 w(t)dt \xrightarrow{p} 0. \tag{2.12}$$

Enunciamos abaixo resultados quanto à convergência das estatísticas ponderadas e recortadas associadas a R_n, adaptados a nossa necessidade e que serão utilizados adiante. Para dada distribuição F, denotamos por $\mu_{2,F}(w)$ o segundo momento ponderado pela função peso w para R_n

$$\mu_{2,F}(w) := \int_{-\infty}^{\infty} x^2 w(F(x))dF(x) = \int_0^1 (F^{-1}(t))^2 w(t)dt. \tag{2.13}$$

Em relação à distribuição geradora G_0, considere as estatísticas ponderadas $R_n^l(w)$, $R_n^r(w)$ e $R_n^{l,r}(w)$ aparada à esquerda, à direita e em ambos os lados, com os segundos momentos aparados ponderados por w sendo denotados respectivamente por $\mu_{2,F}^{n,l}(w)$, $\mu_{2,F}^{n,r}(w)$ e $\mu_{2,F}^{n,l,r}(w)$. Temos:

$$R_n^l(w) = 1 - \frac{\left(\int_{\frac{1}{n+1}}^1 F_n^{-1}(t) G_0^{-1}(t) w(t)dt\right)^2}{\mu_{2,F}^{n,l}(w)\mu_{2,G_0}^{n,l}(w)},$$

$$R_n^r(w) = 1 - \frac{\left(\int_0^{\frac{n}{n+1}} F_n^{-1}(t) G_0^{-1}(t) w(t)dt\right)^2}{\mu_{2,F}^{n,r}(w)\mu_{2,G_0}^{n,r}(w)}$$

e

$$R_n^{l,r}(w) = 1 - \frac{\left(\int_{\frac{1}{n+1}}^{\frac{n}{n+1}} F_n^{-1}(t) G_0^{-1}(t) w(t) dt\right)^2}{\mu_{2,F}^{n,l,r}(w)\mu_{2,G_0}^{n,l,r}(w)}.$$

Teorema 2.2. *(Csorgo, (2002, 2003)([6], [7])) Seja G_0 uma f.d. com média zero e assuma que temos função densidade g_0 positiva em (a_{G_0}, b_{G_0}), a_{G_0} e b_{G_0}, definidos por (1.2). Seja F uma f.d que satisfaça a Condição 1.1 e as condições em (2.10), (2.11) e (2.12). Então, se $F \in \mathcal{G}_{G_0}$, as estatísticas $nR_n^l(w)$, $nR_n^r(w)$ e $nR_n^{l,r}(w)$ convergem em distribuição para*

$$\frac{1}{\mu_{2,F}(w)} \int_0^1 \frac{B^2(t)}{f^2(F^{-1}(t))} w(t) dt - \left(\frac{1}{\mu_{2,F}(w)} \int_0^1 \frac{B(t)F^{-1}(t)}{f(F^{-1}(t))} w(t) dt\right)^2, \tag{2.14}$$

onde $\{B(t)|0 \leq t \leq 1\}$ é uma ponte Browniana e f é a densidade de F.

Com base nas técnicas acima, introduzimos na próxima seção uma família de funções peso sem a dependência intrínseca que a função w_F de de Wet tem da distribuição F, que é justamente o elemento desconhecido. Isso permitirá a extensão dos resultados de [11] para as distribuições α-estáveis e extremais de Fréchet.

2.3 Distribuições Ponderadas

Com o objetivo de desenvolver estatísticas análogas a

$$R_n = 1 - \frac{\left(\int_0^1 F_n^{-1}(t)\Phi^{-1}(t) dt\right)^2}{S_n^2}$$

para distribuições α-estáveis, a alternativa de se fazer uso da distância-α de M-W apresentou dificuldades que não pudemos contornar, pois temos resultados equivalentes ao do Teorema do Limite Central para distribuições estáveis ([1] e [33]) que precisariam ser estendidos para incluir resultados quanto à convergência ao movimento de Lévy e que seriam semelhantes ao Teorema de Donsker para o movimento Browniano. Além disso, seriam necessários resultados assintóticos semelhantes aos das pontes Brownianas.

Diferentemente dos trabalhos [6], [8] e [9], porém inspiradas neles, optamos por trabalhar com distribuições ponderadas.

Definição 2.2. *Dizemos que* $w : (0,1) \longmapsto [0,\infty)$ *é uma função peso se* $w(.) \geq 0$ *e* $\displaystyle\int_0^1 w(t)dt = 1$*, isto é,* w *é uma função densidade em (0,1). E dada uma função de distribuição* F *definimos a sua distribuição ponderada associada por*

$$F_w(x) = \int_{-\infty}^{x} w(F(y))dF(y). \tag{2.15}$$

Considere o conjunto das distribuições com segundo momento ponderado finito,

$$\mathcal{L}_2(w) = \left\{ F : \int_{-\infty}^{\infty} x^2 w(F(x))dF(x) < \infty \right\}. \tag{2.16}$$

Naturalmente, se $F \in \mathcal{L}_2(w)$ então $F_w \in \mathcal{L}_2$ (espaço de distribuições com segundo momento finito). Além disso, para a média ponderada e a variância ponderada temos

$$\mu_F(w) = \int_{-\infty}^{\infty} xw(F(x))dF(x) = \int_{-\infty}^{\infty} xdF_w(x) = \mu_{F_w}$$

e

$$\sigma_F^2(w) = \int_{-\infty}^{\infty} x^2 w(F(x))dF(x) - \mu_F^2(w) = \int_{-\infty}^{\infty} x^2 dF_w(x) - \mu_{F_w}^2 = \sigma_{F_w}^2.$$

Mais ainda, tomando $u = F(y)$ em (2.15) podemos provar a proposição:

Proposição 2.1. *Seja* $F \in \mathcal{L}_2(w)$ *então* $F_w \in \mathcal{L}_2$*,*

$$F_w(x) = W(F(x)) \;\; e \;\; F_w^{-1}(u) = F^{-1}(W^{-1}(u))$$

onde

$$W(v) = \int_0^v w(u)du. \tag{2.17}$$

Iremos considerar f.d.'s F, que possuem variância infinita, e procurar condições sobre a função peso w que assegurem que $F \in \mathcal{L}_2(w)$. O seguinte exemplo ilustra o nosso procedimento.

Exemplo 2.1. *Para* $0 < \alpha < 2$ *considere a f.d.*

$$F(x) = \begin{cases} \dfrac{(-x)^{-\alpha}}{2} & se \;\; x < -1 \\ 0 & se \;\; -1 < x \geq 1. \\ 1 - \dfrac{x^{-\alpha}}{2} & se \;\; x \geq 1. \end{cases}$$

Um cálculo direto nos mostra que $\int |x|dF(x) = \infty$. *No entanto, para* $\beta < 1 - \frac{2}{\alpha}$, *podemos definir a função peso* $w_\beta(.)$:

$$w_\beta(u) = \begin{cases} (1-\beta)2^{-\beta}u^{-\beta} & se \quad 0 < u < \frac{1}{2} \\ (1-\beta)2^{-\beta}(1-u)^{-\beta} & se \quad \frac{1}{2} \leq u < 1. \end{cases}$$

Temos que $\int_0^1 w_\beta(u)du = 1$ *e portanto* w_β *é uma função peso. Além disso,*

$$\frac{dF(x)}{dx} = \frac{\alpha|x|^{-\alpha-1}}{2} I_{(|x|\geq 1)}$$

e

$$w(F(x)) = (1-\beta)|x|^{\alpha\beta} I_{(|x|\geq 1)}.$$

E é fácil verificar que

$$\int_{-\infty}^{-1} x^2 w(F(x))dF(x) = \int_1^\infty x^2 w(F(x))dF(x)$$

e

$$(1-\beta)\alpha \int_1^\infty x^{1+\alpha(\beta-1)}dx < \infty$$

desde que $\beta < \frac{-2}{\alpha}$. *De onde segue que* $F \in \mathcal{L}_2(w_\beta)$.

O exemplo acima mostra que, com a devida ponderação, distribuições com caudas regularmente variantes podem ter segundo momento finito. Por outro lado, as funções de variação regular caracterizam as caudas das distribuições α-estáveis (Proposição 1.6). Assim, motivados pelo Exemplo 2.1, vamos estabelecer a seguinte condição sobre a função peso.

Condição 2.2. *Para algum* $\beta < -\frac{2}{\alpha}$, $0 < \alpha < 2$, *a função peso* w *satisfaz,*
(a) para distribuições com caudas pesadas do mesmo tipo em ambos os lados:

$$w(u) = \begin{cases} k_* u^{-\beta} & se \quad 0 < u < u_* \\ k_*(1-u)^{-\beta} & se \quad u_* \leq u < 1 \end{cases} \tag{2.18}$$

onde $0 < u_* < 1$ *e* $k_* = \dfrac{1-\beta}{u_*^{1-\beta} + (1-u_*)^{1-\beta}}$*;*

(b) para distribuições com cauda pesada apenas a direita:

$$w(u) = \begin{cases} 1 & se \quad 0 < u < u_* \\ k_*(1-u)^{-\beta} & se \quad u_* \leq u < 1, \end{cases} \tag{2.19}$$

onde $0 < u_* < 1$ *e* $k_* = (1-\beta)(1-u_*)^{\beta}$*;*

(c) para distribuições com cauda pesada apenas a esquerda:

$$w(u) = \begin{cases} k_* u^{-\beta} & se \quad 0 < u < u_* \\ 1 & se \quad u_* \leq u < 1 \end{cases} \tag{2.20}$$

onde $0 < u_* < 1$ *e* $k_* = (1-\beta)u_*^{\beta}$.

Observação 2.1. *Note que a nossa Condição 2.2 é muito menos restritiva do que a proposta por de Wet (Definição 2.1), que foi definida em termos da distribuição desconhecida* F*. De fato, para* F *conhecida e pela prova do Lema 2.1 abaixo, poderíamos ter simplificado a nossa condição tomando* w *como sendo qualquer função satisfazendo* $w(F(x)) \sim |x|^{\gamma}$ *para* $\gamma \leq -3$ *. Naturalmente, se a distribuição em consideração possuir o segundo momento finito, isto é, caudas leves, a função peso será a constante* 1.

Lema 2.1. *Seja* F *uma f.d. e assuma que as caudas* $\bar{F}(x) = 1 - F(x), F(-x) \in RV_{-\alpha}$*. Considere* w *uma função peso satisfazendo a Condição 2.2. Então,*

(i) $\bar{F}_w(x), F_w(-x) \in RV_{\alpha(\beta-1)}$.

(ii) $F \in \mathcal{L}_2(w)$ *e* $F_w \in \mathcal{L}_2$.

Demonstração. (i) Para $x \geq 1/u_*$, defina

$$V_-(x) = W\left(\frac{1}{x}\right) = \int_0^{1/x} w(u)du.$$

Da Proposição 2.1, segue que $F_w(x) = W(F(x))$ e consequentemente temos

$$F_w(-x) = V_-\left(\frac{1}{F(-x)}\right). \tag{2.21}$$

Fazendo $u = \dfrac{1}{y}$ em (2.17), temos

$$V_-(x) = W\left(\frac{1}{x}\right) = \int_x^{\infty} v_-(y)dy, \tag{2.22}$$

onde $v_-(y) = w\left(\frac{1}{y}\right) y^{-2}$.

Como $x \geq 1/u_*$ segue da Condição 2.2 que $w\left(\frac{1}{x}\right) = k_* x^\beta$ e $v_-(y) \in RV_{\beta-2}$ e, portanto, já que $\beta - 2 < -3 < -1$, usando o Teorema de Karamata (Teorema 1.4) aplicado a (2.22) temos que $V_-(x) \in RV_{\beta-1}$.

Com esta informação e aplicando em (2.21) a Proposição 1.1 (ii) e (iii) e a Proposição 1.2 (v), segue que $F_w(-x) \in RV_{\alpha(\beta-1)}$.

Para a cauda a direita de F_w aplicando-se a Proposição 2.1 temos que

$$\bar{F_w}(x) = \int_{F(x)}^{1} w(y)dy.$$

Logo,

$$\bar{F_w}(x) = V_+\left(\frac{1}{\bar{F}(x)}\right), \tag{2.23}$$

onde

$$V_+(x) = \int_x^\infty v_+(y)dy \tag{2.24}$$

com

$$v_+(y) = w\left(1 - \frac{1}{y}\right) y^{-2}.$$

Pela Condição 2.2, temos que $w\left(1 - \frac{1}{x}\right) = k_* x^\beta$. Este fato e o uso do Teorema de Karamata em (2.24) assegura que $V_+(x) \in RV_{\beta-1}$. Aplicando-se as Proposições 1.1 e 1.2 em (2.23), temos $\bar{F_w}(x) \in RV_{\alpha(\beta-1)}$.

(ii) Dos argumentos acima e do Teorema 1.5 e novamente do Teorema de Karamata segue que para x^* fixado, temos

$$\int_{x^*}^\infty x^2 dF_w(-x) < \infty$$

e

$$\int_{x^*}^\infty x^2 d\bar{F}_w(x) < \infty.$$

Logo, $F_w \in \mathcal{L}_2$ e $F \in \mathcal{L}_2(w)$.

□

O próximo resultado visa comparar a distância de M-W ponderada

$$d_{2,w}^2(F,G) = \int_0^1 (F^{-1}(t) - G^{-1}(t))^2 w(t)dt \tag{2.25}$$

com a correspondente distância-2 de M-W das distribuições ponderadas,

$$d_2^2(F_w, G_w) = \int_0^1 (F_w^{-1}(t) - G_w^{-1}(t))^2 dt.$$

Também, para a família de escala-locação gerada por G_0

$$\mathcal{G}_{G_0} := \left\{ G : G(x) = G_0\left(\frac{x-\mu}{\sigma}\right), \mu \in \mathbb{R},\ \sigma > 0 \right\}$$

comparamos as distâncias

$$d_{2,w}(F, \mathcal{G}_{G_0}) = \inf\{d_{2,w}(F,G) : G \in \mathcal{G}_{G_0}\}$$

e

$$d_2(F_w, \mathcal{G}_{G_{0,w}}) = \inf\{d_2(F_w, G) : G \in \mathcal{G}_{G_{0,w}}\},$$

onde $G_{0,w}$ é a distribuição ponderada de G_0,

$$G_{0,w}(x) = \int_{-\infty}^{x} w(G_0(y))dG_0(y). \tag{2.26}$$

Lema 2.2. *Sejam F e G distribuições com segundo momento ponderado finito, isto é, $F \in \mathcal{L}_2(w)$ e $G \in \mathcal{L}_2(w)$. Assuma que a função peso w satisfaça a Condição 2.2, então*

(i) $d_{2,w}(F,G) = d_2(F_w, G_w)$*;*

(ii) $\mathcal{G}_{G_w} \equiv \mathcal{G}_{G_{w,0}} \equiv \mathcal{G}_{G_{0,w}}$ *onde $G_{0,w}$ é a distribuição ponderada por w de G_0 e $G_0(x) = G(\sigma_G(w)x + \mu_G(w))$ com $\sigma_G(w)$ e $\mu_G(w)$ sendo, respectivamente, o desvio padrão ponderado e a média ponderada de G;*

(iii) $d_{2,w}(F, \mathcal{G}_G) = d_2(F_w, \mathcal{G}_{G_w})$.

Demonstração. (i) Temos pela Proposição 2.1

$$\begin{aligned}
d_{2,w}^2(F,G) &= \int_0^1 (F^{-1}(t) - G^{-1}(t))^2 w(t)dt \\
&= \sigma_F^2(w) + \mu_F^2(w) + \sigma_G^2(w) + \mu_G^2(w) - 2\int_0^1 F^{-1}(t)G^{-1}(t)w(t)dt \\
&= \sigma_{F_w}^2 + \mu_{F_w}^2 + \sigma_{G_w}^2 + \mu_{G_w}^2 - 2\int_0^1 F^{-1}(t)G^{-1}(t)w(t)dt
\end{aligned}$$

e

$$\begin{aligned} d_2^2(F_w, G_w) &= \int_0^1 (F_w^{-1}(t) - G_w^{-1}(t))^2 dt \\ &= \sigma^2_{F_w} + \mu^2_{F_w} + \sigma^2_{G_w} + \mu^2_{G_w} - 2\int_0^1 F_w^{-1}(t) G_w^{-1}(t) dt. \end{aligned}$$

Por outro lado, para

$$W(v) = \int_0^v w(u)du, \quad 0 < v < 1$$

e usando novamente a Proposição 2.1 temos que

$$\begin{aligned} \int_0^1 F_w^{-1}(t) G_w^{-1}(t) dt &= \int_0^1 F^{-1}(W^{-1}(t)) G^{-1}(W^{-1}(t)) dt \\ &= \int_0^1 F^{-1}(t) G^{-1}(t) dW(t) \\ &= \int_0^1 F^{-1}(t) G^{-1}(t) w(t) dt. \end{aligned}$$

E o item (i) está provado.

(ii) Se $G_0 \in \mathcal{G}_G$ então G_0 é uma versão escalada e locada de G, e, portanto, temos que $\mathcal{G}_G = \mathcal{G}_{G_0}$ ou equivalentemente

$$G_0(x) = G(\sigma x + \mu) \tag{2.27}$$

para algum $\sigma > 0$ e $\mu \in \mathbb{R}$. Sem perda de generalidade, como $G \in \mathcal{L}_2(w)$, suponha G_0 com $\mu_{G_0}(w) = 0$ e $\sigma^2_{G_0}(w) = 1$. Nesse caso, teremos $\mu_G(w) = \mu$ e $\sigma^2_G(w) = \sigma^2$.

Assim pela Proposição 2.1, podemos reescrever (2.27) como:

$$G_0(x) = G\left(x\sigma_G(w) + \mu_G(w)\right) = G(\sigma_{G_w} x + \mu_{G_w}). \tag{2.28}$$

Similarmente, considere $G_{w,0} \in \mathcal{G}_{G_w}$, da mesma forma, tomando $\mu_{G_{w,0}} = 0$ e $\sigma^2_{G_{w,0}} = 1$, obtemos

$$G_{w,0}(x) = G_w(\sigma_{G_w} x + \mu_{G_w}). \tag{2.29}$$

Considere a função de distribuição ponderada $G_{0,w}$ de G_0. Substituindo (2.28) em (2.26), obteremos da definição de função de distribuição ponderada e da relação em (2.29) que $G_{w,0} = G_{0,w}$ e com este argumento está demonstrado que o gerador $G_{w,0}$ da família $\mathcal{G}_{G_w}$ é a versão ponderada $G_{0,w}$ do gerador G_0 da família $\mathcal{G}_G$ e, portanto, temos

$$\mathcal{G}_{G_w} = \mathcal{G}_{G_{w,0}} = \mathcal{G}_{G_{0,w}}. \tag{2.30}$$

(iii) Por (2.30) se $H \in \mathcal{G}_{G_w}$ então $H \in \mathcal{G}_{G_{0,w}}$. Então para algum μ e σ, temos

$$H(x) = \int_{-\infty}^{\frac{x-\mu}{\sigma}} dG_{0w} = \int_{-\infty}^{\frac{x-\mu}{\sigma}} w(G_0(y))dG_0(y). \tag{2.31}$$

Por outro lado, se $\tilde{H} \in \mathcal{G}_{G_0}$ é definido por $\tilde{H}(x) = G_0\left(\frac{x-\mu}{\sigma}\right)$, então por (2.31)

$$\tilde{H}_w(x) = \int_{-\infty}^{\frac{x-\mu}{\sigma}} w(G_0(y))dG_0(y) = H(x).$$

Segue que

$$d_2(F_w, \mathcal{G}_{G_w}) = \inf_{H\in\mathcal{G}_{G_0}} \{d_2(F_w, H_w)\} = \inf_{H\in\mathcal{G}_{G_0}} \{d_{2,w}(F, H)\}. \tag{2.32}$$

A última igualdade segue de (i), de onde temos que $d_2(F_w, H_w) = d_{2,w}(F, H)$, concluindo a prova do item (iii).

$\square$

A seguir, mostraremos a validade da relação fundamental em (2.2) para distribuições com segundo momento ponderado finito. Vale ressaltar que a nossa demonstração é muito mais simples e direta que a apresentada em [6] para a distância ponderada $d_{2,w}(.,.)$.

Lema 2.3. *Assuma que as hipóteses do Lema 2.2 sejam satisfeitas e seja*

$$G_0(x) = G(\sigma_G(w)x + \mu_G(w)).$$

Então

$$\begin{aligned}
\frac{d_2(F_w, \mathcal{G}_{G_w})}{\sigma^2_{F_w}} &= \frac{d_{2,w}(F, \mathcal{G}_G)}{\sigma^2_F(w)} \\
&= 1 - \frac{\left(\int_0^1 F^{-1}(t)G_0^{-1}(t)w(t)dt\right)^2}{\sigma^2_F(w)} \\
&= 1 - \frac{\left(\int_0^1 F^{-1}(t)(G^{-1}(t) - \mu_G(w))w(t)dt\right)^2}{\sigma^2_F(w)\sigma^2_G(w)}.
\end{aligned} \tag{2.33}$$

Demonstração. Seja $H(x) = G_0\left(\dfrac{x-\mu}{\sigma}\right)$. Então $H^{-1}(t) = \sigma G_0^{-1}(t) + \mu$ para algum $\mu \in \mathbb{R}$ e $\sigma > 0$. Seja $\mu_F(w)$ e $\sigma_F^2(w)$ a média ponderada e variância ponderada de F.

$$\mu_H(w) = \int_0^1 H^{-1}(t)w(t)dt = \int_0^1 (\sigma G_0^{-1}(t) + \mu)w(t)dt = \sigma\mu_{G_0}(w) + \mu = \mu. \tag{2.34}$$

$$\mu_{2,H}(w) = \int_0^1 (H^{-1}(t))^2 w(t)dt = \int_0^1 (\sigma G_0^{-1}(t) + \mu)^2 w(t)dt = \sigma^2 + \mu^2. \tag{2.35}$$

Segue de (2.34) e (2.35) que $\sigma_F^2(w) = \sigma_H^2(w)$. Temos também,

$$d_{2,w}^2(F,H) \;=\; \sigma_F^2(w) + \mu_F^2(w) + \sigma^2 + \mu^2 - 2\int_0^1 F^{-1}(t)(\sigma G_0^{-1}(t) + \mu)w(t)dt,$$

onde

$$2\int_0^1 F^{-1}(t)(\sigma G_0^{-1}(t) + \mu)w(t)dt = 2\mu\mu_F(w) + 2\sigma\int_0^1 F^{-1}(t)G_0^{-1}(t)w(t)dt.$$

Mais ainda,

$$\begin{aligned} d_{2,w}^2(F,H) &= \sigma_F^2(w) + \left(\sigma - \int_0^1 F^{-1}(t)G_0^{-1}(t)w(t)dt\right)^2 + (\mu_F(w) - \mu)^2 \\ &- \left(\int_0^1 F^{-1}(t)G_0^{-1}(t)w(t)dt\right)^2. \end{aligned}$$

Consequentemente, para minimizar $d_{2,w}^2(F,H)$ basta tomar $\mu = \mu_F(w)$ e

$$\sigma = \int_0^1 F^{-1}(t)G_0^{-1}(t)w(t)dt.$$

Minimizando com respeito a μ e a σ, temos:

$$d_{2,w}^2(F,\mathcal{G}_G) = \sigma_F^2(w) - \left(\int_0^1 F^{-1}(t)G_0^{-1}(t)w(t)dt\right)^2.$$

Porém, como $\mu_{G_0}(w) = 0$, temos que

$$\begin{aligned} \int_0^1 F^{-1}(t)G_0^{-1}(t)w(t)dt &= \int_0^1 (F^{-1}(t) - \mu_F(w))G_0^{-1}(t)w(t)dt \\ &= \int_0^1 (F^{-1}(t) - \mu))G_0^{-1}(t)w(t)dt, \end{aligned}$$

que é invariante com respeito a μ e a $\mu_F(w)$. E assim

$$\begin{aligned}
\frac{d^2_{2,w}(F,\mathcal{G}_G)}{\sigma^2_F(w)} &= 1 - \frac{\left(\int_0^1 F^{-1}(t)G_0^{-1}(t)w(t)dt\right)^2}{\sigma^2_F(w)} \\
&= 1 - \frac{\left(\int_0^1 (G^{-1}(t)-\mu_G(w))F^{-1}(t)w(t)dt\right)^2}{\sigma^2_F(w)\sigma^2_G(w)},
\end{aligned}$$

onde a última igualdade segue do fato da relação em (2.28) ter como consequência que

$$\begin{aligned}
G_0^{-1}(t) &= \inf\{s|G(s) \geq \sigma_G(w)t + \mu_G(w)\} \\
&= \frac{G^{-1}(t)-\mu_G(w)}{\sigma_G(w)}.
\end{aligned}$$

□

2.4 Exemplos Ilustrativos

Nos exemplos abaixo, apresentaremos o uso da função peso w da Condição 2.2 para diversas distribuições de cauda pesada. Propomos também uma adaptação de w a ser utilizada para distribuições F com segundo momento finito (cauda leve) que satisfazem

$$\sup_{0\leq t\leq 1} \frac{t(1-t)|f'(F^{-1}(t))|}{f^2(F^{-1}(t))} \leq \gamma < \infty. \tag{2.36}$$

No entanto, não vale a condição

$$\int_0^1 \frac{t(1-t)}{f^2(F^{-1}(t))}dt < \infty. \tag{2.37}$$

A condição em (2.36), que está incluída na Condição 1.1, é nada mais do que a condição de Von Mises para processos empíricos. E a condição em (2.37) impõe certa uniformidade no controle das variâncias do processo limite, pois sabemos que a distribuição empírica F_n satisfaz

$$\sqrt{n}(F_n^{-1}(t) - F^{-1}(t)) \xrightarrow{d} \frac{-B(t)}{f(F^{-1}(t))},$$

onde $\{B(t) : 0 \leq t \leq 1\}$ é uma ponte Browniana e f é a densidade de F. E a variável limite $\dfrac{B(t)}{f(F^{-1}(t))}$ possui variância $\dfrac{t(1-t)}{f^2(F^{-1}(t))}$.

Veremos na próxima seção que, para a obtenção do limite assintótico de $R_n(w)$, será também necessário um controle mais refinado do comportamento caudal (2.6) e (2.7).

Estas condições serão verificadas para a classe de distribuições estáveis e Fréchet, nas provas dos Teoremas 2.3 e 2.4.

Nos Exemplos 2.5 e 2.6, discutimos o uso da função peso de de Wet para as distribuições Weibull ([6]) e Gamma ([8]) respectivamente, e apresentamos uma função peso alternativa e mais simples.

Exemplo 2.2. *Considere a distribuição estável $S_1(0,0,1)$, isto é, a Cauchy padrão com densidade dada por:*

$$g(x) = \frac{1}{\pi(1+x^2)}, \ x \in \mathbb{R}.$$

Temos $\mu_G = \infty$, no entanto, a Condição 1.1 é satisfeita, pois, $a_G = \infty$, $b_G = \infty$, $g(x) > 0$ em (a_G, b_G), $g^{'}$ existe e

$$\begin{aligned} \sup_{0\leq t\leq 1} \frac{t(1-t)|g^{'}(G^{-1}(t))|}{g^2(G^{-1}(t))} &= \sup_{x\in\mathbb{R}} \frac{G(x)\bar{G}(x)|g^{'}(x)|}{g^2(x)} \\ &= \sup_{x\in\mathbb{R}} 2\pi x G(x)\bar{G}(x). \end{aligned}$$

Como o supremo ocorre quando $x \to \pm\infty$ e $g^{'}(x) = \dfrac{-2x}{\pi(1+x^2)^2}$, podemos aplicar a regra de l'Hôpital para

$$\lim_{x\to\infty} x\bar{G}(x) \ e \ \lim_{x\to-\infty} xG(x). \tag{2.38}$$

Para o primeiro limite em (2.38) temos

$$\lim_{x\to\infty} \frac{\bar{G}(x)}{\dfrac{1}{x}} = \lim_{x\to\infty} \frac{x^2}{\pi(1+x^2)} = 1.$$

Para o segundo limite, usando argumentos similares obtemos que

$$\lim_{x\to-\infty} xG(x) = -1.$$

Quanto a (2.5), fazendo-se $t = G(x)$ temos $x = G^{-1}(t)$, pois a distribuição G é estritamente crescente

$$\begin{aligned} \int_0^1 \frac{t(1-t)}{g^2(G^{-1}(t))} dt &= \int_{-\infty}^{\infty} \frac{G(x)\bar{G}(x)dG(x)}{g^2(x)} \\ &= \int_0^1 \frac{G(x)\bar{G}(x)dx}{g(x)} = \infty. \end{aligned}$$

Pois, por l'Hôpital temos

$$\lim_{x\to\infty}\frac{\bar{G}(x)}{g(x)}=\lim_{x\to\infty}\frac{-g(x)}{g'(x)}=\lim_{x\to\infty}\frac{1+x^2}{2x}=\infty.$$

Esta dificuldade pode ser contornada pela função peso

$$w(t)=\begin{cases}32t^3 & se \quad 0<t<\dfrac{1}{2}\\[2ex] 32(1-t)^3 & se \quad \dfrac{1}{2}\leq t<1.\end{cases}$$

Segue que

$$\begin{aligned}\int_0^1\frac{t(1-t)w(t)}{g^2(G^{-1}(t))}dt &= 32\int_{-\infty}^{F^{-1}(\frac{1}{2})}\frac{G^4(x)\bar{G}(x)dx}{g(x)}+32\int_{F^{-1}(\frac{1}{2})}^{\infty}\frac{G(x)\bar{G}^4(x)dx}{g(x)}\\ &= A+B.\end{aligned}$$

Como as provas são análogas, mostraremos que $B<\infty$*. Pela Proposição 1.6, existe uma constante c tal que* $\bar{G}(x)\sim cx^{-1}$ *e, portanto,* $\bar{G}^4(x)\sim c^*x^{-4}$*. Então, para y grande*

$$\int_y^{\infty}\frac{G(x)\bar{G}^4(x)dx}{g(x)}\leq\int_y^{\infty}\frac{\bar{G}^4(x)dx}{g(x)}\sim c^*\int_y^{\infty}\frac{\pi(1+x^2)dx}{x^4}<\infty.$$

E, portanto, $B<\infty$*, pois,* $\displaystyle\int_{F^{-1}(\frac{1}{2})}^{y}\frac{G(x)\bar{G}^4(x)dx}{g(x)}<\infty$ *pela Proposição 1.5, item (iii). E assim a condição em (2.5) é satisfeita.*

Para testar $F\in\mathcal{G}$ *onde*

$$\mathcal{G}=\left\{H:H(x)=\frac{\left(1+\left(\frac{x-\mu}{\sigma}\right)^2\right)^{-1}}{\pi},\mu\in\mathbb{R},\sigma>0\right\},$$

usamos a estatística ponderada

$$R_n(w)=1-\frac{\left(\displaystyle\int_0^1 F_n^{-1}(t)G^{-1}(t)w(t)dt\right)^2}{S^2_{F_n}(w)}.$$

Nesse caso, o Teorema 2.3, que apresentaremos adiante, determina a distribuição assintótica de $nR_n(w)$*, que será dada por (2.48).*

Exemplo 2.3. *Considere a distribuição maximal da Fréchet. Para $\alpha > 0$,*

$$\Phi_\alpha(x) = \begin{cases} e^{-x^{-\alpha}} & se \;\; x \geq 0 \\ 0 & se \;\; x < 0. \end{cases}$$

Nesse caso, temos $a_{\Phi_\alpha} = 0$, $b_{\Phi_\alpha} = \infty$ e a densidade é dada por

$$\phi_\alpha(x) = \alpha x^{-(\alpha+1)} e^{-x^{-\alpha}}, x > 0.$$

Note que temos $\phi_\alpha(x) \sim \alpha x^{-(\alpha+1)}$, pois, $\lim_{x\to\infty} \dfrac{\phi_\alpha(x)}{\alpha x^{-(\alpha+1)}} = 1$. Assim, temos para $y \to \infty$ que

$$\int_y^\infty x^2 \phi_\alpha(x) dx \sim \int_y^\infty x^2 \alpha x^{-(\alpha+1)} dx,$$

se uma das integrais for finita, e isto ocorre se $2 - (\alpha + 1) < -1$ ou $\alpha > 2$. E para $0 < \alpha < 2$, temos o segundo momento infinito (cauda pesada).

Assim, como no exemplo anterior, verifica-se a Condição 1.1. Mostraremos que, para todo $\alpha > 0$, a condição em (2.5) não é satisfeita.

Observe que

$$\frac{\Phi_\alpha(x)\bar{\Phi}_\alpha(x)}{\phi_\alpha(x)} = \frac{\bar{\Phi}_\alpha(x)}{\alpha x^{-(\alpha+1)}} = \frac{x^{\alpha+1}\bar{\Phi}_\alpha(x)}{\alpha} \sim \frac{x}{\alpha},$$

pois

$$\lim_{x\to\infty} \frac{\bar{\Phi}_\alpha(x) x^{\alpha+1}}{x} = \lim_{x\to\infty} \frac{\bar{\Phi}_\alpha(x)}{x^{-\alpha}} = \lim_{x\to\infty} \frac{\phi_\alpha(x)}{\alpha x^{-(\alpha+1)}} = 1.$$

Segue que

$$\int_0^1 \frac{t(1-t)dt}{\phi_\alpha^2(\Phi_\alpha^{-1}(t))} = \int_0^\infty \frac{\Phi_\alpha(x)\bar{\Phi}_\alpha(x)dx}{\phi_\alpha(x)} = \infty, \forall \alpha > 0.$$

Para a situação de cauda pesada, $0 < \alpha < 2$, podemos usar a função da Condição 2.2 e obter para $\beta < \dfrac{-2}{\alpha}$

$$w(t) = \begin{cases} 1 & se \quad 0 < t < u_* \\ k_*(1-t)^{-\beta} & se \quad u_* \leq t < 1. \end{cases}$$

Um Cálculo direto mostrará que temos

$$\int_0^1 \frac{t(1-t)}{\phi_\alpha^2(\Phi_\alpha^{-1}(t))} w(t)dt < \infty.$$

E o Teorema 2.4, que apresentaremos adiante, assegurará a convergência em distribuição de $nR_n(w)$ para esse caso.

Observação 2.1. *Muito embora a Condição 2.2 proponha função peso w para distribuições de cauda pesada, podemos formular o equivalente para distribuições com segundo momento finito (cauda leve), conforme veremos nos Exemplos 2.4, 2.5 e 2.6. Para tanto, basta tomar $\beta < -1$ na Condição 2.2.*

Exemplo 2.4. *Considere no Exemplo 2.3 a situação de cauda leve, $\Phi_\alpha \in L_2$ com $\alpha > 2$. Vimos que a condição em (2.5) não é satisfeita. Definindo a função peso de acordo com a Observação 2.1, por exemplo, seja $\beta = -2$*

$$w(t) = \begin{cases} 1 & se \quad 0 < t < u_* \\ k_*(1-t)^2 & se \quad u_* \leq t < 1. \end{cases}$$

Seja $x_* = \Phi_\alpha^{-1}(u_*)$. Temos

$$\begin{aligned} \int_0^1 \frac{t(1-t)w(t)dt}{\phi_\alpha^2(\Phi_\alpha^{-1}(t))} &= \int_0^{x_*} \frac{\Phi_\alpha(x)\bar{\Phi}_\alpha(x)dx}{\phi_\alpha(x)} + \int_{x^*}^\infty k^* \frac{\Phi_\alpha(x)\bar{\Phi}_\alpha^3(x)dx}{\phi_\alpha(x)} \\ &= A + B. \end{aligned}$$

A primeira integral é trivialmente finita pois

$$A \leq \int_0^{x_*} \frac{\Phi_\alpha(x)dx}{\phi_\alpha(x)} = \int_0^{x_*} \frac{x^{\alpha+1}dx}{\alpha} < \infty.$$

Para mostrar que $B < \infty$, note que, do Exemplo 2.3, temos $\phi_\alpha(x) \sim \alpha x^{-(\alpha+1)}$ e $\bar{\Phi}_\alpha(x) \sim x^{-\alpha}$. Segue que, para y grande,

$$\frac{\bar{\Phi}_\alpha^3(y)}{\phi_\alpha(y)} \sim \frac{y^{-2\alpha+1}}{\alpha}.$$

Consequentemente,

$$\frac{B}{k_*} \leq \int_{x_*}^y \frac{\bar{\Phi}_\alpha^3(x)dx}{\phi_\alpha(x)} + \int_y^\infty \frac{\bar{\Phi}_\alpha^3(x)dx}{\phi_\alpha(x)} \sim \int_{x_*}^y \frac{\bar{\Phi}_\alpha^3(x)dx}{\phi_\alpha(x)} + \int_y^\infty \frac{x^{-2\alpha+1}dx}{\alpha} < \infty,$$

pois $-2\alpha + 1 < 0$ já que $\alpha > 2$ e $\dfrac{\bar{\Phi}_\alpha^3(y)}{\phi_\alpha(y)}$ é contínua. Novamente, o Teorema 2.4 determina a distribuição limite de $nR_n(w)$ para esse caso.

Exemplo 2.5. *Considere a classe de distribuições Weibull dada pela densidade*

$$g(x) = \alpha e^{-x^\alpha} x^{\alpha-1}, x > 0, \alpha > 0$$

e a correspondente função de distribuição $G(x) = 1 - e^{-x^\alpha}$. *Temos que* $G \in L_2$, $\mu_G = \Gamma\left(\frac{1}{\alpha}+1\right)$ *e* $\sigma_G^2 = \Gamma\left(\frac{1}{\alpha}+1\right) - \left(\Gamma\left(\frac{2}{\alpha}+1\right)\right)^2$. *Trata-se da classe das distribuições minimais de Weibull. Nesse exemplo, vamos analisar o caso* $\alpha = 1$, *isto é, a* $\exp(1)$. *A Condição 1.1 é trivialmente satisfeita (vide [12], pg 60 para detalhes) e quanto à condição em (2.5), temos*

$$\begin{aligned}\int_0^1 \frac{t(1-t)dt}{g^2(G^{-1}(t))} &= \int_0^\infty \frac{G(x)\bar{G}(x)dx}{g(x)} \\ &= \int_0^\infty (1-e^{-x})dx = \infty.\end{aligned}$$

Usando a Observação 2.1, podemos tomar a mesma função peso w *do Exemplo 2.4 e teremos para* $x_* = G^{-1}(u_*)$

$$\begin{aligned}\int_{u_*}^1 \frac{t(1-t)w(t)dt}{g^2(G^{-1}(t))} &= \int_{x_*}^\infty \frac{G(x)\bar{G}^3(x)dx}{g(x)} \\ &\leq \int_{x_*}^\infty k_* e^{-2x} dx < \infty.\end{aligned}$$

Nesse caso, sem a verificação das condições refinadas sobre as caudas, (2.6) e (2.7), podemos determinar o limite assintótico para a variante recortada $\tilde{R}_n^{r,l}(w_G)$ *de* $R_n(w)$

$$\tilde{R}_n^{r,l}(w_G) := 1 - \frac{\left(\int_{\frac{1}{n+1}}^{\frac{n}{n+1}} F_n^{-1}(t)G^{-1}(t)w_G(t)dt\right)^2}{\mu_{2,F_n}^{n,l,r}(w_G)\mu_{2,G}^{n,l,r}(w_G)}, \tag{2.39}$$

através do Teorema 2.2. Onde $\mu_{2,F_n}^{n,r,l}(w)$ *e* $\mu_{2,G}^{n,l,r}(w)$ *são, respectivamente, o segundo momento recortado ponderado da medida empírica* F_n *e da* G.

Exemplo 2.6. *Considere a distribuição Gamma*

$$G(x) = \frac{\int_0^x e^{-y} y^{\alpha-1} dy}{\Gamma(\alpha)}, \Gamma(\alpha) = \int_0^\infty e^{-y} y^{\alpha-1} dy$$

onde $\alpha > 0$. *Como* $\mu_G = \sigma_G^2 = \alpha$ *temos o segundo momento finito (cauda leve). Novamente a condição em (2.5) não é satisfeita.*

Em [8], faz-se o uso da função de Wet dada pela Definição 2.1:

$$w_G(t) := \frac{1}{I_G}\frac{L_G^{'}(G^{-1}(t))}{G^{-1}(t)}, \quad 0 < t < 1$$

com

$$I_G = \int_{-\infty}^{\infty} L^2(x)g(x)dx < \infty \ \ onde \ \ L_G(x) = -1 - \frac{xg'(x)}{g(x)},$$

desde que

$$\lim_{x \downarrow a_G} x^2 g'(x) = 0 = \lim_{x \uparrow b_G} x^2 g'(x)dx. \tag{2.40}$$

Para a distribuição Gamma, temos $(a_G, b_G) = (0, \infty)$ *e cálculo direto mostra (2.40). Temos* $L_G(x) = x - \alpha$ *e* $I_G = \alpha$*. Segue que* $w_G(t) = \dfrac{1}{\alpha G^{-1}(t)}$*. Com o uso dessa função peso, em [8], mostrou-se que, ainda assim,*

$$\int_0^1 \frac{t(1-t)w_G(t)dt}{g^2(G^{-1}(t))} = \infty.$$

No entanto, temos

$$\int_0^1 \int_0^1 \frac{(s \wedge t - st)^2}{g^2((G^{-1}(s))g^2((G^{-1}(t))} w_G(s)w_G(t)dsdt < \infty, \tag{2.41}$$

onde $s \wedge t := min\{s,t\}$ *e para* $\delta \in \left(\frac{1}{2}, 1\right)$,

$$\lim_{n\to\infty} c_n(w) := \lim_{n\to\infty} \frac{1}{n^{1-\delta}} \int_{\frac{1}{n+1}}^{\frac{n}{n+1}} \frac{(t(1-t))^{\delta} w_G(t)dt}{g^2(G^{-1}(t))} = 0. \tag{2.42}$$

As verificações de (2.41) e (2.42) podem ser encontradas em [8], páginas: 344-347. E isso possibilitou a determinação do limite assintótico para a variante recortada $\tilde{R}_n^{r,l}(w_G)$ *de* $R_n(w)$ *onde* $\tilde{R}_n^{r,l}(w_G)$ *é a estatística recortada de ambos os lados*

$$\tilde{R}_n^{r,l}(w_G) := 1 - \frac{\left(\int_{\frac{1}{n+1}}^{\frac{n}{n+1}} F_n^{-1}(t)G^{-1}(t)w_G(t)dt\right)^2}{\mu_{2,F_n}^{n,l,r}(w_G)\mu_{2,G}^{n,l,r}(w_G)}, \tag{2.43}$$

onde $\mu_{2,F_n}^{n,r,l}(w)$ *e* $\mu_{2,G}^{n,l,r}(w)$ *são, respectivamente, o segundo momento recortado ponderado da medida empírica* F_n *e da* G*.*

Alternativamente, em função da Observação 2.1, podemos tomar a função peso

$$w(u) = \begin{cases} 1 & se \quad 0 < u < \dfrac{1}{2} \\ k_*(1-u)^{-\beta} & se \quad \dfrac{1}{2} \leq u < 1, \end{cases} \tag{2.44}$$

onde $\beta < -1$ *e* k_* *uma constante apropriada.*

Nesse caso,

$$\begin{aligned}\int_{\frac{1}{2}}^{1} \frac{t(1-t)w(t)dt}{g^2(G^{-1}(t))} &= \int_{G^{-1}(\frac{1}{2})}^{\infty} \frac{G(x)\bar{G}(x)w(G(x))dx}{g(x)} \\ &\leq \int_{G^{-1}\left(\frac{1}{2}\right)}^{\infty} k_* \frac{(\bar{G}(x))^2 dx}{g(x)},\end{aligned}$$

pois $(\bar{G}(x))^{-\beta-1} \leq 1$. *Segue que para* $y \to \infty$, $y > G^{-1}\left(\frac{1}{2}\right)$ *e também*

$$\int_{y}^{\infty} k_* \frac{(\bar{G}(x))^2 dx}{g(x)} < \infty,$$

pois, por integração por partes, temos para y *grande*

$$\int_{y}^{\infty} e^{-z} z^{\alpha-1} dz \sim y^{\alpha-1} e^{-y} \tag{2.45}$$

e, consequentemente,

$$\int_{y}^{\infty} \frac{(\bar{G}(x))^2 dx}{g(x)} \sim \frac{1}{\Gamma(\alpha)} \int_{y}^{\infty} x^{\alpha-1} e^{-x} dx \sim \frac{y^{\alpha-1} e^{-y}}{\Gamma(\alpha)}.$$

Nesse caso, sem a verificação das condições refinadas sobre as caudas, (2.6) e (2.7), podemos também concluir a convergência de $\tilde{R}_n^{r,l}(w)$ *através do Teorema 2.2. Vale ressaltar que o uso de função peso* w*, satisfazendo a Condição 2.2 para* $\beta < -1$ *ao invés da função de de Wet, simplifica bastante as verificações das hipóteses necessárias à existência de distribuição assintótica. Fato similar obtivemos para a Weibull no Exemplo 2.5 quando substituímos a função peso de de Wet também pela função peso satisfazendo a Condição 2.2 para* $\beta = -2$.

2.5 Testes de Ajuste de Bondade

Baseado na amostra $\{X_1, \ldots, X_n\}$ da distribuição F, estamos interessadas em testar a hipótese nula $F \in \mathcal{G}_G$, onde a família escala-locação $\mathcal{G}_G$ é gerada ou pelas distribuições estáveis $G = G_\alpha = S_\alpha(\sigma, 0, \mu)$ com $0 < \alpha < 2$ ou pelas distribuições extremais de Fréchet $G = \Phi_\alpha$ com $\alpha > 0$. Por simplicidade de notação, usaremos Φ_α para denotar tanto a distribuição maximal

$$\Phi_\alpha(x) = \begin{cases} e^{-x^{-\alpha}} & \text{se } x \geq 0 \\ 0 & \text{se } x < 0, \end{cases} \tag{2.46}$$

quanto a distribuição minimal, caso em que temos $\Phi_\alpha(x) = e^{-(-x)^{-\alpha}}$ para $x \leq 0$. Assim,

$$\mathcal{G}_{G_\alpha} = \left\{ G : \; G(x) = G_\alpha \left(\frac{x - \mu'}{\sigma'} \right), \; \mu' \in \mathbb{R}, \; \sigma' > 0 \right\}$$

e, similarmente, $\mathcal{G}_{\Phi_\alpha}$.

Vimos pelo Lema 2.2 que $d_{2,w}(F, \mathcal{G}_G) = d_2(F_w, \mathcal{G}_{G_w})$. Mais ainda, o Lema 2.3 e (2.33) nos levam a considerar

$$\begin{aligned} R(w) &= \frac{d_2(F_w, \mathcal{G}_{G_w})}{\sigma^2_{F_w}} \\ &= 1 - \frac{\left(\int_0^1 F^{-1}(t) G_0^{-1}(t) w(t) dt \right)^2}{\sigma_F^2(w)} \\ &= 1 - \frac{\left(\int_0^1 F^{-1}(t) (G^{-1}(t) - \mu_G(w)) w(t) dt \right)^2}{\sigma_F^2(w) \sigma_G^2(w)}, \end{aligned}$$

onde $G_0(x) = G(\sigma_G(w)x + \mu_G(w))$ com $\mu_{G_0}(w) = 0$ e $\sigma^2_{G_0}(w) = 1$. Usaremos como estatística de teste a correspondente versão empírica

$$\begin{aligned} R_n(w) &= 1 - \frac{\left(\int_0^1 F_n^{-1}(t) G_0^{-1}(t) w(t) dt \right)^2}{S^2_{F_n}(w)} \qquad (2.47) \\ &= 1 - \frac{\left(\int_0^1 F_n^{-1}(t) (G^{-1}(t) - \mu_G(w)) w(t) dt \right)^2}{S^2_{F_n}(w) \sigma_G^2(w)}, \end{aligned}$$

onde F_n é a distribuição empírica e $S^2_{F_n}(w)$ a variância amostral ponderada,

$$S^2_{F_n}(w) = \int_0^1 (F_n^{-1}(t))^2 w(t) dt - \left(\int_0^1 F_n^{-1}(t) w(t) dt \right)^2.$$

Teorema 2.3. *Seja a distribuição estável $G_\alpha = S_\alpha(\sigma, 0, \mu)$ para algum $0 < \alpha \leq 2$ e assuma que a função peso w satisfaça a Condição 2.2. Então, sob a hipótese nula, $F \in \mathcal{G}_{G_\alpha}$, temos*

$$\begin{aligned} nR_n(w) &\xrightarrow{d} \frac{1}{\sigma^2_{G_\alpha}(w)} \int_0^1 \frac{B^2(t) w(t) dt}{g_\alpha^2(G_\alpha^{-1}(t))} - \left(\frac{1}{\sigma^2_{G_\alpha}(w)} \int_0^1 \frac{B(t) w(t) dt}{g_\alpha(G_\alpha^{-1}(t))} \right)^2 \qquad (2.48) \\ &- \left(\frac{1}{\sigma^2_{G_\alpha}(w)} \int_0^1 \frac{B(t) G_\alpha^{-1}(t) w(t) dt}{g_\alpha(G_\alpha^{-1}(t))} \right)^2, \end{aligned}$$

onde g_α é a função densidade de G_α e $\{B(t) : 0 \leq t \leq 1\}$ é a ponte Browniana.

Teorema 2.4. *Para $\alpha > 0$ seja Φ_α a distribuição extremal de Fréchet e assuma que a função peso w satisfaça a Condição 2.2. Então, sob a hipótese nula, $F \in \mathcal{G}_{\Phi_\alpha}$, temos (2.48) substituindo G_α por Φ_α e a densidade g_α pela densidade ϕ_α de Φ_α.*

Observação 2.2. *(a) As demonstrações dos teoremas acima serão feitas por meio de vários resultados complementares, do Lema 2.4 ao Lema 2.8. Provaremos, por exemplo, que as hipóteses do Teorema 1.1 estão satisfeitas, o que permitirá a aproximação dos processos quantis empíricos por sequências de pontes Brownianas. Verificaremos também que temos o comportamento caudal em (2.6) e em (2.7) se estas integrais forem ponderadas por w satisfazendo a Condição 2.2 e, aplicando-se o Teorema 2.2, concluiremos as demonstrações dos nossos teoremas. Para o Teorema 2.3, excluímos, nos lemas abaixo, as discussões sobre o caso $\alpha = 2$, pois é justamente o Teorema 2 de [11].*

(b) As nossas provas sugerem a possibilidade de formular um teorema mais geral baseado no comportamento caudal das distribuições. Grande parte dos argumentos utilizados diz respeito às caudas regularmente variantes das distribuições estáveis e distribuições no domínio de atração de extremais do tipo Fréchet. Essa possibilidade será explorada em um trabalho futuro.

Lema 2.4. *Assuma que w satisfaça a Condição 2.1.*

(i) Seja $F \in \mathcal{D}(G_\alpha)$. Então, $\sigma^2_{G_\alpha}(w) < \infty$ e $\sigma^2_F(w) < \infty$.

(ii) $\sigma^2_{\Phi_\alpha}(w) < \infty$.

Demonstração. (i) Seja $F \in \mathcal{D}(G_\alpha)$, então, pelo Teorema 1.8, temos $\bar{F}(x) \in RV_{-\alpha}$ e $F(-x) \in RV_{-\alpha}$. Pelo Lema 2.1 (i), temos $\bar{F}_w(x) \in RV_{\alpha(\beta-1)}$ e $F_w(-x) \in RV_{\alpha(\beta-1)}$. Como $\bar{F}_w$ e F_w são funções positivas, limitadas e $\alpha(\beta-1) < -2$, podemos utilizar a Proposição 1.4 com $\gamma = 2$

$$
\begin{aligned}
\int_x^\infty y^2 dF_w(y) &= -x^2 \bar{F}_w(x) \left(\frac{\displaystyle\int_x^\infty y^2 d\bar{F}_w(y)}{x^2 \bar{F}_w(x)} \right) \\
&\sim \frac{\alpha(1-\beta)}{2+\alpha(\beta-1)} x^{\alpha(\beta-1)+2}, \text{ quando } x \to \infty.
\end{aligned}
$$

Segue que $\int_x^\infty y^2 dF_w(y) < \infty$. Similarmente, mostramos que $\int_{-\infty}^x y^2 dF_w(y) < \infty$ usando $F_w(-x)$. Portanto, $F_w \in \mathcal{L}_2$ e $\sigma_F^2(w) < \infty$. Como $G_\alpha \in \mathcal{D}(G_\alpha)$, temos também $\sigma^2_{G_\alpha}(w) < \infty$.

(ii) Uma das formas de argumentar que $\bar{\Phi} \in RV_{-\alpha}$ é notar que $\Phi \in \mathcal{D}_{max}(\Phi_\alpha)$, pois seguirá do Teorema 1.10 (outra alternativa, a expansão de Taylor de $\bar{\Phi}_\alpha$). Os mesmos argumentos que (i) mostram que $\int_x^\infty y^2 \phi_\alpha(y) w(y) dy < \infty$. Como $\Phi_\alpha(x) = 0$ para $x < 0$ segue que $\sigma^2_{\Phi_\alpha}(w) < \infty$.

$\square$

O próximo lema visa mostrar que as hipóteses do Teorema 1.1 estão satisfeitas. Para tal precisamos mostrar que para algum $\gamma > 0$ temos

$$\sup_{0\leq t\leq 1} \frac{t(1-t)|f'(F^{-1}(t))|}{f^2(F^{-1}(t))} \leq \gamma. \tag{2.49}$$

Lema 2.5. *Assuma que $F \in \mathcal{G}_{G_\alpha}$ (ou $F \in \mathcal{G}_{\Phi_\alpha}$) então F satisfaça a Condição 1.1.*

Demonstração. (i) Seja $F \in \mathcal{G}_{G_\alpha}$. Observe que para algum $\mu \in \mathbb{R}$ e algum $\sigma > 0$ temos $F = S_\alpha(\sigma, 0, \mu)$. Como pela Proposição 1.5 as distribuições α-estáveis são infinitamente diferenciáveis (vide [37] para detalhes), temos a existência da densidade f e da sua derivada f'. Alem disso, $f(x) > 0, \forall x \in \mathbb{R}$. Como consequência desses fatos, resta mostrar (2.49). Como F é contínua, fazendo $t = F(x)$, podemos reescrever (2.49) como

$$\sup_{x\in\mathbb{R}} \frac{F(x)(1-F(x))|f'(x)|}{f^2(x)} \leq \gamma.$$

Assim, basta mostrar que

$$\lim_{x\to\infty} \frac{xf'(x)}{f(x)} = \lim_{x\to-\infty} \frac{xf'(x)}{f(x)} = -(\alpha+1), \tag{2.50}$$

$$\lim_{x\to\infty} \frac{\bar{F}(x)}{xf(x)} = \frac{1}{\alpha} \quad \text{e} \quad \lim_{x\to-\infty} \frac{F(x)}{xf(x)} = -\frac{1}{\alpha}. \tag{2.51}$$

Segue da Proposição 1.6 que $\bar{F}(x) \in RV_{-\alpha}$ e $F(-x) \in RV_{-\alpha}$. Como f e f' são derivadas de uma distribuição α-estável satisfazem as condições de monotonicidade do Teorema 1.5, segue que $f \in RV_{-(\alpha+1)}$ e $f' \in RV_{-(\alpha+2)}$. Como $-(\alpha+2) < -1$, podemos aplicar o Teorema de Karamata (Teorema 1.4) para $f'(x)$ e $f'(-y)$ e temos

$$\lim_{x\to\infty} \frac{xf'(x)}{f(x)} = \lim_{x\to\infty} \frac{xf'(x)}{-\int_x^\infty f'(t)dt} = -\alpha - 1$$

e

$$\lim_{x\to-\infty}\frac{xf'(x)}{f(x)}=\lim_{y\to\infty}\frac{-yf'(-y)}{\displaystyle\int_y^{\infty}f'(-z)dz}=-\alpha-1.$$

Temos, então, (2.50) provado.

Similarmente, aplicando-se o Teorema de Karamata (Teorema 1.4) para $f(x)$ e $f(-y)$, temos:

$$\lim_{x\to\infty}\frac{\bar{F}(x)}{xf(x)}=\lim_{x\to\infty}\frac{\displaystyle\int_x^{\infty}f(t)dt}{xf(x)}=\frac{1}{\alpha}$$

e

$$\lim_{x\to-\infty}\frac{F(x)}{xf(x)}=\lim_{y\to\infty}\frac{\displaystyle\int_y^{\infty}f(-z)dz}{-yf(-y)}=-\frac{1}{\alpha}.$$

E (2.51) segue.

(ii) Seja $F\in\mathcal{G}_{\Phi_\alpha}$, então, temos $F(x)=\Phi_\alpha\left(\frac{x-\mu}{\sigma}\right)$ para algum $\mu\in\mathbb{R}$ e algum $\sigma>0$. Assumindo que Φ_α é distribuição maximal, o seu suporte é dado por $(a_F,b_F)=(\mu,\infty)$. Nesse caso, basta mostrar

$$\sup_{x>0}\frac{F(x)(1-F(x))|f'(x)|}{f^2(x)}\leq\gamma.$$

Ou, equivalentemente,

$$\lim_{x\to\infty}\frac{xf'(x)}{f(x)}<\infty \;\text{ e }\; \lim_{x\to\infty}\frac{\bar{F}(x)}{xf(x)}<\infty.$$

Por (2.46), temos que $f(x)>0,\forall x\in\mathbb{R}$ e ambos f e f' monótonas. Os mesmos tipos de argumentos de (i) completam a prova.

$\square$

Como vimos na Seção 1.3, a distância-2 de M-W entre a distribuição empírica e a distribuição base pode ser expressa em termos do processo quantil empírico ajustado $\rho_{n,F}(t)=\sqrt{n}(F_n^{-1}(t)-F^{-1}(t))f(F^{-1}(t))$ e satisfaz a relação

$$\sqrt{n}d_2(F_n,F)=\left(\int_0^1\frac{\rho_{n,F}^2(t)}{f^2(F^{-1}(t))}dt\right)^{\frac{1}{2}}.$$

Precisamos assegurar que a distribuição assintótica de $\sqrt{n}d_2(F_n, F)$ seja

$$\left(\int_0^1 \frac{B^2(t)}{f^2(F^{-1}(t))}dt\right)^{\frac{1}{2}}.$$

Como $E\{B^2(t)\} = t(1-t)$, uma das condições necessárias é a finitude da integral ponderada em ((2.52) ou (2.5)).

Lema 2.6. *Seja $F \in \mathcal{G}_{G_\alpha}$ (ou $F \in \mathcal{G}_{\Phi_\alpha}$) e assuma que a função peso w satisfaça a Condição 2.2. Então F possui densidade f e*

$$\int_0^1 \frac{t(1-t)}{f^2(F^{-1}(t))}w(t)dt < \infty. \tag{2.52}$$

Demonstração. (i) Segue da Proposição 1.5 que F é contínua e possui densidade contínua f. Fazendo $t = F(x)$ temos $F^{-1}(t) = x$ e podemos reescrever (2.52) como

$$\begin{aligned}\int_0^1 \frac{t(1-t)w(t)}{f^2(F^{-1}(t))}dt &= \int_{-\infty}^{\infty} \frac{F(x)[1-F(x)]w(F(x))}{f^2(x)}dF(x) \\ &= \int_{-\infty}^{\infty} \frac{F(x)\bar{F}(x)w(F(x))}{f(x)}dx.\end{aligned}$$

Para u_* e k_* definidos na Condição 2.2, tomamos $x_* = F^{-1}(u_*)$ e temos

$$w(F(x)) = \begin{cases} k_*(F(x))^{-\beta} & \text{se } x \le x_* \\ k_*(\bar{F}(x))^{-\beta} & \text{se } x > x_*. \end{cases}$$

Então para provar (2.52) é suficiente mostrar que

$$A = k_* \int_{-\infty}^{x^*} \frac{(F(x))^{-\beta+1}\bar{F}(x)}{f(x)}dx < \infty \tag{2.53}$$

e

$$B = k_* \int_{x^*}^{\infty} \frac{F(x)(\bar{F}(x))^{-\beta+1}}{f(x)}dx < \infty. \tag{2.54}$$

Como as provas são análogas, mostraremos que $B < \infty$.

(ii) Assuma que $F \in \mathcal{G}_{G_\alpha}$. Pela Proposição 1.6 e do fato da densidade f da distribuição α-estável ser monótona, temos as hipóteses do Teorema 1.5 satisfeitas, logo além de $\bar{F}(x) \in RV_{-\alpha}$, temos também $f \in RV_{-(\alpha+1)}$. Segue que existem constantes C_1 e C_2, tais que

$$\bar{F}(x) \sim C_1 x^{-\alpha} \quad \text{e} \quad f(x) \sim C_2 x^{-(\alpha+1)} \quad \text{quando } x \to \infty.$$

Assim, usando a Proposição 1.1 (i), temos $(\bar{F}(x))^{-\beta+1} \in RV_{\alpha(\beta-1)}$. Pelo fato de $0 \leq F(x) \leq 1$, temos para alguma constante C que para $y \to \infty$, temos

$$B \quad \sim \quad k_* \int_{x^*}^{y} \frac{F(x)(\bar{F}(x))^{-\beta+1}}{f(x)} dx + k_* \int_{y}^{\infty} \frac{x^{\alpha(\beta-1)}}{x^{-(\alpha+1)}} dx.$$

A finitude de $\displaystyle\int_{y}^{\infty} x^{\alpha\beta+1} dx$ é consequência do Teorema 1.4, que pode ser aplicado pelo fato de $\alpha\beta + 1 < -1$, já que de acordo com a Condição 2.2, temos $\beta < -2/\alpha$.

Enquanto que a finitude de $k_* \displaystyle\int_{x^*}^{y} \frac{F(x)(\bar{F}(x))^{-\beta+1}}{f(x)} dx$ é consequência da Proposição 1.5 (iii).

(iii) Assuma que $F \in \mathcal{G}_{\Phi_\alpha}$. Com os mesmos argumentos do passo (ii) do Lema 2.4 concluímos que $\bar{F}(x) \in RV_{-\alpha}$. Como Φ_α possui densidade ϕ_α monótona, segue pelo Teorema 1.5 que $\phi_\alpha \in RV_{-(\alpha+1)}$. O restante da demonstração é análoga ao caso (ii).

$\square$

O próximo lema estabelece um resultado auxiliar sobre funções de distribuição com cauda ρ-variantes para algum $\rho < 0$, que será necessário para as nossas demonstrações adiante.

Lema 2.7. *Seja F uma f.d. com $F(x) < 1$, $\forall x \in \mathbb{R}$. Assuma que para algum $\rho < 0$ temos $\bar{F} \in RV_\rho$. Defina*

$$a_n = \inf \left\{ y : y > 0 \ , \ \bar{F}(y) \leq \frac{1}{n} \right\}.$$

Então, se r e γ são tais que $\rho + \gamma < 0$ e $\gamma + r < 0$, temos

$$\lim_{n\to\infty} n a_n^r \int_{a_n}^{\infty} x^\gamma d\bar{F}(x) = 0. \tag{2.55}$$

Demonstração. Seja $V(x) = a_x$, $x > 0$. Como $\bar{F}$ é não crescente e $\rho < 0$, temos, pela Proposição 1.2 (iv), $V \in RV_{-1/\rho}$. Segue da Proposição 1.2 (v) que

$$\lim_{n\to\infty} V(n) = a_n = \infty. \tag{2.56}$$

Como $\gamma + r < 0$, temos

$$\lim_{n\to\infty} a_n^{\gamma+r} = 0. \tag{2.57}$$

Por outro lado, como $\bar{F} \in RV_\rho$ temos também, pelo Teorema 1.10, $F \in \mathcal{D}_{\max}(\Phi_{-\rho})$. Como

$$a_n = \inf\left\{y : y > 0\ ,\ \bar{F}(y) \leq \frac{1}{n}\right\} = F^{-1}\left(1 - \frac{1}{n}\right)$$

segue da relação (1.21) também do Teorema 1.10 que

$$\lim_{n\to\infty} n\bar{F}(a_n) = 1. \tag{2.58}$$

Note que $\bar{F}$ é positiva e limitada e temos (2.56), então, pela Proposição 1.4 segue que

$$\lim_{n\to\infty} \frac{\displaystyle\int_{a_n}^{\infty} x^\gamma d\bar{F}(x)}{a_n^\gamma \bar{F}(a_n)} = -\frac{\rho}{\rho+\gamma}, \tag{2.59}$$

pois $\rho + \gamma < 0$.

Rescrevendo (2.55), temos

$$\lim_{n\to\infty} n a_n^r \int_{a_n}^{\infty} x^\gamma d\bar{F}(x) = \lim_{n\to\infty} n\bar{F}(a_n) a_n^{\gamma+r} \frac{\displaystyle\int_{a_n}^{\infty} x^\gamma d\bar{F}(x)}{a_n^\gamma \bar{F}(a_n)}.$$

Finalmente, de (2.57), (2.58) e (2.59) temos (2.55).

□

Sabemos que

$$d_{2,w}^2(F_n, F) = \int_0^1 (F_n^{-1}(t) - F^{-1}(t))^2 w(t)dt.$$

Os resultados do lema abaixo mostram que podemos controlar o comportamento das caudas dessa distância. Mostraremos que

$$A_n = n\int_{1-\frac{1}{n}}^1 (F_n^{-1}(t) - F^{-1}(t))^2 w(t)dt \xrightarrow{p} 0 \tag{2.60}$$

e

$$B_n = n\int_0^{\frac{1}{n}} (F_n^{-1}(t) - F^{-1}(t))^2 w(t)dt \xrightarrow{p} 0. \tag{2.61}$$

Lema 2.8. *Assuma que $F \in \mathcal{G}_{G_\alpha}$ (ou $F \in \mathcal{G}_{\Phi_\alpha}$) e que w satisfaça a Condição 2.2. Então temos (2.60) e (2.61).*

Demonstração. (i) Para $F \in \mathcal{G}_{G_\alpha}$ ou $F \in \mathcal{G}_{\Phi_\alpha}$, provaremos (2.60) por meio dos passos de (ii) a (v). Quanto a (2.61), provaremos a sua validade para $F \in \mathcal{G}_{G_\alpha}$ através dos passos (ii) e de (vi) a (viii). Para o caso $F \in \mathcal{G}_{\Phi_\alpha}$, a prova de (2.61) é bem mais simples. Note que, para algum $\mu \in \mathbb{R}$ e algum $\sigma > 0$, temos $F(x) = \Phi_\alpha\left(\dfrac{x-\mu}{\sigma}\right)$. Assim, por (2.46), temos $a_F = \inf\{x : F(x) > 0\} = \mu$ e para $0 \leq t \leq 1/n$ com n suficientemente grande, com probabilidade 1 (q.c.),

$$|F_n^{-1}t) - F^{-1}(t)| < \delta < \infty \ \text{ e } \ B_n \leq n\delta^2 \int_0^{\frac{1}{n}} w(t)dt.$$

Também para n grande, da Condição 2.2 sobre a função peso w, temos

$$\begin{aligned} \int_0^{\frac{1}{n}} w(t)dt &= k_* \int_0^{\frac{1}{n}} u^{-\beta} du \\ &= \frac{k_*}{1-\beta}\left(\frac{1}{n}\right)^{1-\beta}. \end{aligned}$$

Como $\beta < -2/\alpha$ temos segue que

$$\lim_{n\to\infty} \frac{n\delta^2 k_*}{1-\beta}\left(\frac{1}{n}\right)^{1-\beta} = 0.$$

E resulta (2.61).

(ii) Usaremos a notação

$$M_n = \max\{X_1, \cdots, X_n\} \quad \text{e} \quad m_n = \min\{X_1, \cdots, X_n\}.$$

Então temos

$$F_n^{-1}(t) = \begin{cases} m_n & \text{se } 0 \leq t < \frac{1}{n} \\ M_n & \text{se } 1 - \frac{1}{n} < t \leq 1. \end{cases}$$

Assim, para

$$a_n = F^{-1}\left(1 - \frac{1}{n}\right) \quad \text{e} \quad b_n = -F^{-1}\left(\frac{1}{n}\right)$$

podemos reescrever (2.60) e (2.61) como segue

$$\begin{aligned} A_n &= n \int_{1-\frac{1}{n}}^{1} \left(M_n - F^{-1}(t)\right)^2 w(t)dt \\ &= n \int_{1-\frac{1}{n}}^{1} \left(M_n - a_n + a_n - F^{-1}(t)\right)^2 w(t)dt \\ &= A_n^1 + A_n^2 + A_n^3. \end{aligned}$$

E para provar (2.60), é suficiente mostrar que

$$A_n^1 = n\int_{1-\frac{1}{n}}^{1}(M_n - a_n)^2 w(t)dt \xrightarrow{p} 0 \tag{2.62}$$

$$A_n^2 = n\int_{1-\frac{1}{n}}^{1}(a_n - F^{-1}(t))^2 w(t)dt \longrightarrow 0 \tag{2.63}$$

$$A_n^3 = 2n\int_{1-\frac{1}{n}}^{1}(M_n - a_n)(a_n - F^{-1}(t))w(t)dt \xrightarrow{p} 0. \tag{2.64}$$

Similarmente, temos

$$\begin{aligned} B_n &= n\int_0^{\frac{1}{n}}\left(m_n - F^{-1}(t)\right)^2 w(t)dt \\ &= n\int_0^{\frac{1}{n}}\left(m_n - b_n + b_n - F^{-1}(t)\right)^2 w(t)dt \\ &= B_n^1 + B_n^2 + B_n^3. \end{aligned}$$

E para provar (2.61), é suficiente mostrar que

$$B_n^1 = n\int_0^{\frac{1}{n}}(m_n - b_n)^2 w(t)dt \xrightarrow{p} 0 \tag{2.65}$$

$$B_n^2 = n\int_0^{\frac{1}{n}}(b_n - F^{-1}(t))^2 w(t)dt\ 0 \tag{2.66}$$

$$B_n^3 = 2n\int_0^{\frac{1}{n}}(m_n - b_n)(b_n - F^{-1}(t))w(t)dt \xrightarrow{p} 0. \tag{2.67}$$

(iii) Para provar (2.62), note que temos

$$A_n^1 = na_n^2\left(\frac{M_n}{a_n} - 1\right)^2\int_{a_n}^{\infty} w(F(x))dF(x).$$

Se $F \in \mathcal{G}_{G_\alpha}$, vimos no passo (ii) da prova do Lema 2.6 que $\bar{F}(x) \in RV_{-\alpha}$. Naturalmente, se $F \in \mathcal{G}_{\Phi_\alpha}$, temos também $\bar{F}(x) \in RV_{-\alpha}$; e, pelo Teorema 1.10, temos que $F, \Phi_\alpha \in \mathcal{D}_{max}(\Phi_\alpha)$ e assim $\left(\frac{M_n}{a_n} - 1\right)$ converge para uma distribuição não degenerada.

Como $F(x) < 1$ e $\bar{F}(x) \in RV_{-\alpha}$ com $\alpha > 0$, temos satisfeitas as hipóteses do Lema 2.7, onde tomaremos $\gamma = \alpha\beta$, $r = 2$, $\rho = -\alpha$, $\rho+\gamma = \alpha(\beta-1) < 0$ e $\gamma+r = \alpha\beta+2 < 0$. Como $\bar{F}(x) \in RV_{-\alpha}$ para alguma constante c temos $\bar{F}(x) \sim cx^{-\alpha}$. Pela Condição 2.2, temos para n grande $a_n > F^{-1}(u_*)$ e também

$$\begin{aligned}\int_{a_n}^{\infty} w(F(x))dF(x) &= \int_{a_n}^{\infty} k_*(\bar{F}(x))^{-\beta}dF(x) \\ &\sim \int_{a_n}^{\infty} x^{\alpha\beta}d\bar{F}(x).\end{aligned} \tag{2.68}$$

De onde segue que

$$na_n^2\int_{a_n}^{\infty} w(F(x))dF(x) \sim na_n^r\int_{a_n}^{\infty} x^{\gamma}d\bar{F}(x) \to 0, \text{ quando } n \to \infty,$$

devido ao lema mencionado acima. E isso completa a prova de (2.62).

(iv)Para provar (2.63), note que temos

$$\begin{aligned}A_{n,2} &= na_n^2\int_{a_n}^{\infty} w(F(x))dF(x) - 2na_n\int_{a_n}^{\infty} xw(F(x))dF(x) \\ &+ n\int_{a_n}^{\infty} x^2w(F(x))dx.\end{aligned}$$

De (2.68) obtida no passo (iii), segue que para $n \to \infty$, obtemos

$$\begin{aligned}A_{n,2} &\sim na_n^2\int_{a_n}^{\infty} x^{\alpha\beta}d\bar{F}(x) + 2n\int_{a_n}^{\infty} x^{\alpha\beta+1}d\bar{F}(x) \\ &- n\int_{a_n}^{\infty} x^{\alpha\beta+2}d\bar{F}(x).\end{aligned}$$

Pelos mesmos argumentos do passo (iii), temos satisfeitas as hipóteses do Lema 2.7 para cada um dos três termos acima com $\rho = -\alpha$. Mas agora tomando no primeiro termo $\gamma = \alpha\beta$, $r = 2$, depois considerando $\gamma = \alpha\beta + 1$, $r = 1$ no segundo termo e, por último, fazendo $\gamma = \alpha\beta + 2$, $r = 0$. Note que usando o fato de $\beta < -2/\alpha$, temos $\rho + \gamma < 0$ e $\gamma + r < 0$ para qualquer uma das escolhas de γ e r feitas acima. E isso completa a prova de (2.63).

(v) Para provar (2.64), note que temos

$$\begin{aligned}A_{n,3} &= 2na_n^2\left(\frac{M_n}{a_n} - 1\right)\int_{a_n}^{\infty} w(F(x))dF(x) \\ &- 2na_n\left(\frac{M_n}{a_n} - 1\right)\int_{a_n}^{\infty} xw(F(x))dF(x).\end{aligned}$$

Note que no passo (iv) os limites para as integrais

$$na_n^2 \int_{a_n}^{\infty} w(F(x))dF(x) \text{ e } na_n \int_{a_n}^{\infty} xw(F(x))dF(x)$$

foram calculados. Além disso, vimos no passo (iii) que para $F \in \mathcal{G}_{G_\alpha}$ ou $F \in \mathcal{G}_{\Phi_\alpha}$ temos $\left(\dfrac{M_n}{a_n} - 1\right)$ convergindo para uma distribuição não degenerada. Assim, temos (2.64).

(vi) Para provar (2.65), note que temos

$$B_n^1 = nb_n^2 \left(\frac{m_n}{b_n} - 1\right)^2 \int_{b_n}^{\infty} w(F(-x))dF(-x).$$

Logo, os mesmos argumentos do passo (iii) são válidos se substituirmos $\bar{F}$ e a_n por $F(-x)$ e b_n, respectivamente. Usando a Condição 2.2, concluímos a prova de (2.64).

(vii) Para provar (2.66), note que

$$\begin{aligned} B_{n,2} &= nb_n^2 \int_{b_n}^{\infty} w(F(x))dF(x) - 2nb_n \int_{b_n}^{\infty} xw(F(x))dF(x) \\ &+ n \int_{b_n}^{\infty} x^2 w(F(x))dF(x). \end{aligned}$$

Logo, os mesmos argumentos do passo (iv) são válidos se substituirmos $\bar{F}$ e a_n por $F(-x)$ e b_n, respectivamente. Usando a Condição 2.2, concluímos a prova de (2.66).

(viii) Para provar (2.67), note que

$$\begin{aligned} B_{n,3} &= nb_n^2 \left(\frac{m_n}{b_n} - 1\right) \int_{b_n}^{\infty} w(F(-x))dF(-x) \\ &- nb_n^2 \left(\frac{m_n}{b_n} - 1\right) \int_{b_n}^{\infty} w(F(-x))dF(-x). \end{aligned}$$

E a demonstração pode ser feita exatamente com os mesmos argumentos do passo (v) se substituirmos $\bar{F}$ e a_n por $F(-x)$ e b_n. Usando a Condição 2.2, concluímos a prova de (2.67).

$\square$

Demonstração do Teorema 2.3. O Lema 2.5 mostra que temos a Condição 1.1 satisfeita; o Lema 2.4 mostra que $F \in \mathcal{L}_2(w)$; o Lema 2.6 mostra que temos (2.10);

o Lema 2.8 assegura que temos (2.11) e (2.12). Segue que temos a Condição 2.1 e o Teorema 2.2, que completam a demonstração.

□

Similarmente,

Demonstração do Teorema 2.4. O Lema 2.5 mostra que temos a Condição 1.1 satisfeita; o Lema 2.4 mostra que $\Phi_\alpha \in \mathcal{L}_2(w)$; o Lema 2.6 mostra que temos (2.10); o Lema 2.8 assegura que temos (2.11) e (2.12). Segue que temos a Condição 2.1 e o Teorema 2.2, que completam a demonstração.

□

Capítulo 3

Testes de ε-Similaridade

3.1 Introdução

No Capítulo 2, apresentamos estatísticas para testar $F \in \mathcal{G}_G$ e, para incluir as distribuições estáveis e as extremais de Fréchet, introduzimos uma função peso w na métrica de M-W. Neste capítulo, o nosso interesse é testar a similaridade entre duas distribuições F e G, que possuem cauda(s) pesada(s). Uma das métricas intuitivas para avaliar a similaridade é a da variação total

$$d_{TV}(P_F, P_G) = \sup_{B \in \mathcal{B}(\mathbb{R})} |P_F(B) - P_G(B)|,$$

onde P_F e P_G são, respectivamente, as medidas de probabilidade na reta associadas às distribuições F e G. Na busca de similaridade, podemos utilizar procedimentos e testes que possam avaliar para dado $0 \leq \varepsilon < 1$, a condição

$$d_{TV}(P_F, P_G) \leq \varepsilon.$$

A nossa abordagem tem como base os trabalhos [12], [14], [15] e [19], em que, motivadas por aplicações em que não eram necessárias que as distribuições a serem avaliadas coincidissem, foram propostas versões robustas para os testes de ajuste/correlação tratados no Capítulo 2. Para tanto, vários conceitos de similaridade e ajustes parciais de medidas de probabilidade foram introduzidos para possibilitar testar hipóteses do tipo

$$H_0 : \tau_\gamma(F, G) > \delta > 0 \quad \text{vs} \quad H_1 : \tau_\gamma(F, G) \leq \delta,$$

onde $0 \leq \gamma < 1$ e, para $h_0(.)$ convenientemente escolhido,

$$\tau_\gamma(F, G) = \left(\int_0^1 (F^{-1}(t) - G^{-1}(t))^2 h_0(t) dt \right)^{\frac{1}{2}}.$$

Em [14], sob condições que incluem o quarto momento finito, foi obtida no Teorema 2.2 a distribuição assintótica de $\sqrt{n}(\tau_\gamma(F_n, G) - \tau_\gamma(F, G))$.

A hipótese do quarto momento finito inviabiliza o uso de $\tau_\gamma(F_n, G)$ para distribuições α-estáveis, pois se $X \stackrel{d}{=} S_\alpha(\sigma, \beta, \mu)$, temos $E|X|^{\alpha'} = \infty$ para todo $\alpha' > \alpha$. Neste capítulo, temos como objetivo obter um resultado similar ao Teorema 2.2 de [14] que inclua as distribuições α-estáveis de cauda pesada e extremais de Fréchet. Na Seção 3.2, apresentaremos vários conceitos e resultados relacionados à similaridade de distribuições que motivaram a escolha do teste que vai desempenhar o papel da estatística $H_0 : \tau_\gamma(F, G) > \delta > 0 \quad \text{vs} \quad H_1 : \tau_\gamma(F, G) \leq \delta$, para distribuições com caudas regularmente variantes.

Veremos na Seção 3.3 que a distribuição ponderada F_w é uma distribuição ajustada de F e que w, apropriadamente escolhida, poderá desempenhar o papel da função de ajuste. Esse fato e o nosso Lema 2.1 indicam que a estatística $d_{2,w}(F, \mathcal{G}_G)$ dos testes de ajuste/correlação, obtida para as estáveis e extremais no Capítulo 2, é nada mais que uma distância entre distribuições ajustadas. Ainda nesta seção, apresentaremos a classe adequada de funções peso w, Condição 3.2, que garantirá a distribuição assintótica de

$$\sqrt{n}(d_{2,w}(F_n, G) - d_{2,w}(F, G))$$

Mais precisamente, mostraremos no Teorema 3.2 que para distribuições com caudas regularmente variantes e pesadas temos

$$\sqrt{n}(d_{2,w}(F_n, G) - d_{2,w}(F, G)) \xrightarrow{d} 2 \int_0^1 \frac{B(t)(F^{-1}(t) - G^{-1}(t))}{f(F^{-1}(t))} w(t) dt.$$

Podemos utilizar essa convergência acima na construção de testes para distribuições de caudas regularmente variantes e pesadas, sob a hipótese nula $d_{2,w}(F, G) > \varepsilon$ vs $d_{2,w}(F, G) \leq \varepsilon$, para os quais usaremos a terminologia testes de ε-similaridade, $0 \leq \epsilon < 1$.

3.2 Motivação

A melhor maneira de tornar um modelo robusto é o ajuste. Procedimentos de ajuste são de uso frequente na estatística robusta como um caminho para diminuir a influência de dados contaminados nas inferências. A maneira usual de ajustar amostras é de forma simétrica e restrita aos dados pertencentes às caudas. Pressupõe-se, como condição implícita, que no ajuste simétrico a contaminação seja somente por "outliers"(observações não desprezíveis de uma variável aleatória, com valores bastante distantes da média). Esse pressuposto levou à introdução de versões de ajuste de dados dependentes que permitem sanar algumas limitações de versões anteriores que simplesmente removem observações extremas nas caudas. Várias alternativas ao ajuste simétrico têm sido propostas na literatura estatística, dentre elas, o ajuste imparcial. A consequência do uso do método do ajuste imparcial (ajuste dirigido dos dados) no lugar de procedimentos de ajuste clássicos é maximizar a similaridade entre as distribuições retirando os dados que mais contribuem para a dissimilaridade estando nas caudas ou não. Por um ponto de vista descritivo, o ajuste imparcial é uma ferramenta interessante para a comparação de dados amostrais, pois além da robustez que tal ajuste dá ao modelo, ele permite uma análise descritiva pela determinação dos subconjuntos dos dados que melhoram o modelo. Reside nisso a razão pelo qual o ajuste imparcial é preterido em detrimento de outros modelos de ajuste clássicos.

Ao remover ou diminuir, até um certo grau, a importância de alguns dados ao assinalar pesos a uma probabilidade original $\mathbb{P}$, obtemos novas distribuições chamadas medidas ajustadas de $\mathbb{P}$ ou γ-ajustes de $\mathbb{P}$. Como exemplo, veremos quais são as medidas ajustadas da medida empírica em duas situações, no caso em que dados são removidos e no caso em que se muda o peso atribuído aos dados. Esse exemplo servirá de motivação para a definição formal de medidas ajustadas. Suponha que seja necessário retirar k dentre os dados $\{x_1, ..., x_n\}$ de uma amostra para eliminar a contaminação. Quando é atribuído peso zero aos pontos que mais contribuem para a dissimilaridade e é feita a redistribuição de massa aos restantes, uma nova distribuição para os dados que permanecem na amostra é obtida e, nesse caso, a medida empírica $\sum_{i=1}^{n} \delta_{x_i}/n$ é substituída pela sua medida ajustada $\sum_{i=1}^{n} \frac{b_i \delta_{x_i}}{n}$, onde $b_i = 0$, para as observações no conjunto dos k dados retirados, e $\frac{b_i}{n} = \frac{1}{n-k}$ para as observações remanescentes. Note que quando a fração de dados a ser retirada da amostra é limitada por $0 \leq \gamma < 1$, surgem limitantes tanto para o número de dados que podem ser retirados como para os pesos que podem ser atribuídos aos dados remanescentes, pois os k dados retirados da amostra correspondem a uma fração $x \leq \gamma$. Logo, temos que $k \leq n\gamma$, e , portanto, $\frac{1}{n-k} \leq \frac{1}{n(1-\gamma)}$. Se ao invés de retirar dados for atribuído um peso $\frac{b_i}{n}$ para cada um dos n dados, onde $b_1 + ... + b_n = n$, a medida empírica $\sum_{i=1}^{n} \delta_{x_i}/n$ será substituída após o ajuste por

$$\sum_{i=1}^{n} \frac{b_i \delta_{x_i}}{n}, \; 0 \leq b_i \leq \frac{1}{1-\gamma}.$$

A distribuição resultante do ajuste na medida empírica motiva os vários conceitos, introduzidos em [14] e em [15], de γ-ajuste de uma medida de probabilidade $\mathbb{P}$. Sumarizamos alguns desses conceitos e resultados que estabelecem as várias equivalências.

Definição 3.1. *(a) Seja $\mathbb{P}$ uma medida de probabilidade e seja $0 \leq \gamma < 1$. Dizemos que uma medida de probabilidade $\mathbb{P}^*$ é uma γ-ajuste de $\mathbb{P}$ se $\mathbb{P}^*$ é absolutamente contínua com relação a $\mathbb{P}(\mathbb{P}^* \prec \mathbb{P}^*)$ e a derivada Radon-Nikodym satisfaz $\dfrac{d\mathbb{P}^*}{d\mathbb{P}} \leq \dfrac{1}{1-\gamma}$.*

(b) Dizemos que duas medidas de probabilidade $\mathbb{P}_1$ e $\mathbb{P}_2$ sobre o mesmo espaço amostral são γ-similares se existirem medidas de probabilidade $\mathbb{P}_0$, ${\mathbb{P}'}_1$, $\mathbb{P}'_2$, tais que

$$\begin{aligned} \mathbb{P}_1 &= (1-\varepsilon_1)\mathbb{P}_0 + \varepsilon_1 {\mathbb{P}'}_1 \\ \mathbb{P}_2 &= (1-\varepsilon_2)\mathbb{P}_0 + \varepsilon_1 {\mathbb{P}'}_2, \end{aligned}$$

onde $0 \leq \varepsilon_i \leq \gamma < 1$, $i = 1, 2$.

O teorema abaixo apresenta a relação entre os conceitos acima e a norma da variação total. Seja $\mathcal{T}^\gamma(\mathbb{P})$ o conjunto das γ-ajustes de uma medida $\mathbb{P}$, isto é, para $0 \leq \gamma < 1$

$$\mathcal{T}^\gamma(\mathbb{P}) = \left\{ \mathbb{P}^* \in \mathcal{P} : \mathbb{P}^* \prec \mathbb{P} \;\text{ e }\; \frac{d\mathbb{P}^*}{d\mathbb{P}} \leq \frac{1}{1-\gamma} \right\}. \tag{3.1}$$

Teorema 3.1. *(Cuesta et al (2009)([15]) Seja $0 \leq \gamma < 1$ e sejam $\mathbb{P}_1$ e $\mathbb{P}_2$ medidas de probabilidade. Então são equivalentes:*

(i) $\mathbb{P}_1$ e $\mathbb{P}_2$ são γ-similares;

(ii) $\mathcal{T}^\gamma(\mathbb{P}_1) \cap \mathcal{T}^\gamma(\mathbb{P}_2) \neq \emptyset$;

(iii) $d_{TV}(\mathbb{P}_1, \mathbb{P}_2) = \sup_{B \in \mathcal{B}(\mathbb{R})} |\mathbb{P}_1(B) - \mathbb{P}_2(B)| \leq \gamma$.

Mais ainda, se $\mathbb{P}_1 \in \mathcal{L}_2$ e $\mathbb{P}_2 \in \mathcal{L}_2$, então, (i) e (ii) são equivalentes a

$$d_2(\mathcal{T}^\gamma(\mathbb{P}_1), \mathcal{T}^\gamma(\mathbb{P}_2)) = 0,$$

onde d_2 é a distância M-W de ordem 2 e $\mathcal{L}_2$ é o espaço definido em (1.5).

Note que a equivalência entre os itens (i) e (iii) do teorema acima nos diz que $d_{TV}(\mathbb{P}_1, \mathbb{P}_2)$ é a menor fração a ser retirada da amostra para obter similaridade entre as distribuições.

As versões robustas propostas em [14] para os testes de correlação de [12] foram utilizadas em testes de bioequivalência e apresentaram melhores resultados do que os já obtidos em [19]. Nesse trabalho, os autores propuseram uma medida de dissimilaridade entre medidas de probabilidade $\mathbb{P}$ e $\mathbb{Q} \in \mathcal{L}_2$, baseada na menor distância entre suas medidas ajustadas

$$\tau_\gamma(\mathbb{P}, \mathbb{Q}) = \inf_{\mathbb{P}^* \in \mathcal{T}^\gamma(\mathbb{P}), \mathbb{Q}^* \in \mathcal{T}^\gamma(\mathbb{Q})} d_2(\mathbb{P}^*, \mathbb{Q}^*),$$

onde d_2 é a distância-2 de M-W e $\gamma \in [0, 1)$ é o nível máximo de ajuste.

Para o problema de duas amostras, dada uma medida de probabilidade $\mathbb{P}$ desconhecida e uma medida de probabilidade $\mathbb{Q}$ conhecida, assuma $X_1, ..., X_n$ i.i.d's $\mathbb{P}$, o interesse em [14] é o seguinte teste que assume dissimilaridade sob a hipótese nula

$$H_0 : \tau_\gamma(\mathbb{P}, \mathbb{Q}) > \delta \text{ versus } H_1 : \tau_\gamma(\mathbb{P}, \mathbb{Q}) \leq \delta,$$

onde $\delta > 0$. Para tal, o comportamento assintótico da variável $\sqrt{n}(\tau_\gamma(\mathbb{P}_n, \mathbb{Q}) - \tau_\gamma(\mathbb{P}, \mathbb{Q}))$ foi determinado sob a hipótese da existência do quarto momento no Teorema 2.2 do referido artigo. E isso inviabiliza o uso dessa medida de dissimilaridade para distribuições de cauda pesada. Na próxima seção veremos o uso alternativo da distância de M-W para testar ε-similaridade entre distribuições de cauda pesada.

3.3 ε-Similaridade para Cauda Pesada

Note que se w for uma função peso satisfazendo $w(t) \leq \dfrac{1}{1-\varepsilon}$, então a distribuição ponderada

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

é nada mais que a distribuição associada à ε-ajuste da probabilidade associada à distribuição F. Mais especificamente, se P_{F_w} é a probabilidade associada a F_w, então

$$P_{F_w} \in \mathcal{T}^\varepsilon(P_F) = \left\{ P^* : P^* \prec P_F \;\text{ e }\; \frac{dP^*}{dP_F} \leq \frac{1}{1-\varepsilon} \right\},$$

onde P_F é a probabilidade associada a F e $\mathcal{T}^\varepsilon(P_F)$ é definido por (3.1).

As discussões da seção anterior e o fato de $d_{2,w}(F, G) = d_2(F_w, G_w)$, garantido pelo Lema 2.1, justificam a importância de se analisar o comportamento assintótico da sequência $\sqrt{n}(d_{2,w}(F_n, G) - d_{2,w}(F, G))$ para testes de ε-similaridade.

Condição 3.1. *Seja F uma distribuição com densidade estritamente positiva e contínua no seu suporte (a_F, b_F). Assuma que, para $1 < \alpha < 2$, as caudas de F são regularmente variantes satisfazendo uma das condições:*

(a) $\bar{F}(x) \in RV_{-\alpha}$ e $F(-x) \in RV_{-\alpha}$;

(b) $\bar{F}(x) \in RV_{-\alpha}$ e $F(-x)$ é lentamente variante;

(c) $\bar{F}(x)$ é lentamente variante e $F(-x) \in RV_{-\alpha}$.

Motivadas pelo Teorema 3.1, propomos uma classe de funções peso que, além de controlar as caudas pesadas, satisfazem $w(\cdot) \leq \dfrac{1}{1-\varepsilon}$.

Condição 3.2. *Seja $\beta < -\dfrac{2}{\alpha} - 1$. Dado $0 \leq \varepsilon < 1$, assuma que a função peso w satisfaça uma das condições abaixo, onde k_*, k_*^- k_*^+, u_*, u_*^- e u_*^+ são constantes e escolhidas de tal forma a assegurar $w(t) \leq \dfrac{1}{1-\varepsilon}$.*

(a) Para distribuições com caudas pesadas do mesmo tipo em ambos os lados:

$$w(u) = \begin{cases} k_*^- u^{-\beta} & se \quad 0 < u < u_*^- \\ k_* & se \quad u_*^- \leq u < u_*^+ \\ k_*^+ (1-u)^{-\beta} & se \quad u_*^+ \leq u < 1, \end{cases} \tag{3.2}$$

onde $0 < u_*^- < u_*^+ < 1$;

(b) Para distribuições com cauda pesada apenas a direita:

$$w(u) = \begin{cases} k_*^- & se \quad 0 < u < u_* \\ k_*^+ (1-u)^{-\beta} & se \quad u_* \leq u < 1, \end{cases} \tag{3.3}$$

onde $0 < u_* < 1$;

(c) Para distribuições com cauda pesada apenas a esquerda:

$$w(u) = \begin{cases} k_*^- u^{-\beta} & se \quad 0 < u < u_* \\ k_*^+ & se \quad u_* \leq u < 1, \end{cases} \tag{3.4}$$

onde $0 < u_* < 1$.

Observação 3.1. *Esta nova restrição sobre β na Condição 3.2 em relação a Condição 2.2 se deve ao Lema 3.3.*

Exemplo 3.1. *Considere a distribuição α-estável com $0 < \alpha < 2$ e $\beta < \frac{-2}{\alpha} - 1$. Dado $0 \leq \varepsilon < 1$, para $m \in \mathbb{N}$, tome*

$$u_*^- = \frac{1}{m},\ u_*^+ = 1 - \frac{1}{m} \text{ e } k_*^- = k_*^+ = m^{-\beta}.$$

Com estas escolhas, teremos uma função peso w satisfazendo a Condição 3.2. Como para as distribuições α-estáveis, ambas as caudas são regularmente variantes, w satisfaz (3.2). Note que

$$\int_0^{u_*^-} w(t)dt + \int_{u_*^+}^1 w(t)dt = 2\int_0^{\frac{1}{m}} m^{-\beta}t^{-\beta}dt = \frac{2}{1-\beta}\frac{1}{m}. \tag{3.5}$$

Defina $k_* = 1 - \frac{2}{1-\beta}\frac{1}{m}$ então

$$k_*(u_*^+ - u_*^-) = 1 - \frac{2}{1-\beta}\frac{1}{m}. \tag{3.6}$$

Como $\frac{2}{1-\beta}\frac{1}{m} < \frac{1}{m}$ segue que $0 < 1 - \frac{2}{1-\beta}\frac{1}{m} < 1$. De (3.5) e (3.6), concluímos que $\int_0^1 w(t)dt = 1$. Podemos verificar que $w(u) \leq \frac{1}{1-\varepsilon}$. Se $0 < u < u_*^-$ temos

$$w(u) = k_*^- u^{-\beta} \leq k_*^- \left(\frac{1}{m}\right)^{-\beta} = 1 \leq \frac{1}{1-\varepsilon}.$$

Por argumentos similares, para $u_*^+ \leq u < 1$ temos $w(u) \leq \frac{1}{1-\varepsilon}$. E para $u_*^- \leq u < u_*^+$ note que

$$w(u) = k_* = 1 - \frac{2}{1-\beta}\frac{1}{m} < 1 \leq \frac{1}{1-\varepsilon}.$$

O Teorema 3.2 propõe para distribuições com caudas regularmente variantes um resultado similar ao Teorema 2.2 de [14] válido para distribuições com quarto momento finito que foi mencionado na Seção 3.2. A convergência em (3.7) constitui um resultado auxiliar para testes de ε-similaridade que envolvem distribuições de cauda pesada.

Teorema 3.2. *Sejam F e G funções de distribuição com caudas do mesmo tipo em que ambas satisfazem a Condição 3.1 para índices de variação regular α e γ, respectivamente. Considere ainda que $G^{-1}(F)f \in RV_\rho$, para algum $\rho < -1$. Dado $0 \leq \varepsilon < 1$, assuma que w satisfaça a Condição 3.2. Então,*

$$\sqrt{n}(d_{2,w}(F_n, G) - d_{2,w}(F, G)) \xrightarrow{d} 2\int_0^1 \frac{B(t)(F^{-1}(t) - G^{-1}(t))w(t)dt}{f(F^{-1}(t))}, \tag{3.7}$$

em que f é a densidade de F e $\{B(t) : 0 \leq t \leq 1\}$ é a ponte Browniana.

A demonstração do teorema acima será feita por meio de vários resultados auxiliares, alguns obtidos no Capítulo 2 (Lema 2.4 ao Lema 2.8), e os demais apresentaremos adiante. Os lemas abaixo têm caráter um pouco mais geral, mas incluem os casos de G ser uma distribuição α-estável de cauda pesada, bem com as extremais de Fréchet com $0 < \alpha < 2$. Vale ressaltar que as distribuições estáveis e as extremais de Fréchet satisfazem as hipóteses do teorema acima referido, já que tal teorema decorre dos lemas mencionados.

Observação 3.2. *O Lema 3.1 é válido para funções de distribuição com caudas regularmente variantes de índice $0 < \alpha < 2$, pois é corolário da Prova do Lema 2.8. No entanto, optamos por enunciá-lo conforme o seguinte: $1 < \alpha < 2$. Isso se justifica porque esse lema tem como função subsidiar a prova do Teorema 3.2, para o qual essa restrição, sobre o expoente de variação regular, é necessária.*

Lema 3.1. *Seja F_n a distribuição empírica de F em que F satisfaz a Condição 3.1. Dado $0 \leq \varepsilon < 1$, assuma que a função peso w satisfaça a Condição 3.2, então*

(i)

$$\sqrt{n}\int_0^{\frac{1}{n}}(F_n^{-1}(t))^2 w(t)dt \xrightarrow{p} 0 \;\; e \;\; \sqrt{n}\int_{1-\frac{1}{n}}^{1}(F_n^{-1}(t))^2 w(t)dt \xrightarrow{p} 0.$$

(ii)

$$\lim_{n\to\infty}\sqrt{n}\int_0^{\frac{1}{n}}(F^{-1}(t))^2 w(t)dt = 0 \;\; e \;\; \lim_{n\to\infty}\sqrt{n}\int_{1-\frac{1}{n}}^{1}(F^{-1}(t))^2 w(t)dt = 0,$$

(iii)

$$\sqrt{n}\int_{1-\frac{1}{n}}^{1}F_n^{-1}(t)F^{-1}(t)w(t)dt \xrightarrow{p} 0 \;\; e \;\; \sqrt{n}\int_0^{\frac{1}{n}}F_n^{-1}(t)F^{-1}(t)w(t)dt \xrightarrow{p} 0.$$

Demonstração.

A prova do Lema 3.1 para o caso em que a função peso w é definida por (3.2) é análoga a prova feita para o Lema 2.8 no caso em que $F \in G_\alpha = S_\alpha(\sigma, 0, \mu)$. Mas se ao invés de considerarmos escolhas de u_*^+, u_*^-, k_*^-, k_* e k_*^+ como no exemplo 3.2, assumimos que w é definida por (3.3), a prova do Lema 3.1 é análoga a feita para o Lema 2.8 no caso em que $F \in G_{\Phi_\alpha}$. Além disso, os mesmos argumentos aplicados para a Fréchet no Lema 2.8 podem ser aplicados quando a cauda regularmente variante é a esquerda.

Então, utilizando-se dos mesmos argumentos do Lema 2.8, com a devida modificação na função peso w, a conclusão da prova é imediata.

$\square$

Lema 3.2. *Sob as hipóteses do Teorema 3.2 temos*

$$\int_0^1 \frac{\sqrt{t(1-t)}(F^{-1}(t) - G^{-1}(t))}{f(F^{-1}(t))} w(t)dt < \infty. \tag{3.8}$$

Demonstração.
Caso 1): suponha que w satisfaça (3.2) na Condição 3.2.
Para provar (3.8) para o caso 1 é suficiente mostrar que

$$\int_0^1 \frac{\sqrt{t(1-t)}F^{-1}(t)}{f(F^{-1}(t))} w(t)dt < \infty. \tag{3.9}$$

$$\int_0^1 \frac{\sqrt{t(1-t)}G^{-1}(t)}{f(F^{-1}(t))} w(t)dt < \infty. \tag{3.10}$$

Para obter (3.9), basta repetir os passos da prova do Lema 2.6 que a demonstração é análoga.

Considere $x_*^+ = F^{-1}(u_*^+)$, u_*^+ definido na Condição 3.2. Para obter (3.10), basta verificar que

$$\int_{u_*^+}^1 \frac{\sqrt{t(1-t)}G^{-1}(t)}{f(F^{-1}(t))} w(t)dt = k_*^+ \int_{x_*^+}^\infty (\bar{F}(x))^{-\beta} \sqrt{\bar{F}(x)F(-x)}G^{-1}(F(x))dx < \infty, \tag{3.11}$$

pois a prova de que $\displaystyle\int_0^{u_*^-} \frac{\sqrt{t(1-t)}G^{-1}(t)}{f(F^{-1}(t))} w(t)dt < \infty$ é análoga. Desse fato e da variação regular da cauda $\bar{F}$, como F é limitada, concluímos que é suficiente mostrar que para y suficientemente grande

$$k_*^+ \int_y^\infty x^{\alpha\beta} \sqrt{\bar{F}(x)}G^{-1}(F(x))dx < \infty, \tag{3.12}$$

pois a finitude da integral $k_*^+ \displaystyle\int_{x_*^+}^y (\bar{F}(x))^{-\beta} \sqrt{\bar{F}(x)}G^{-1}(F(x))dx$ é garantida pela continuidade de F, G e f.

Para obter (3.12), note que da integração por partes obtemos

$$\begin{aligned}\int_y^\infty x^{\alpha\beta}\sqrt{\bar{F}(x)}G^{-1}(F(x))dx &= \left.\frac{\sqrt{\bar{F}(x)}G^{-1}(F(x))x^{\alpha\beta+1}}{(\beta\alpha+1)}\right|_y^\infty \\ &+ \int_y^\infty \frac{x^{\beta\alpha+1}f(x)G^{-1}(F(x))}{2\sqrt{\bar{F}(x)}(\alpha\beta+1)}dx \\ &- \int_y^\infty \frac{x^{\alpha\beta+1}f(x)\sqrt{\bar{F}(x)}}{(\alpha\beta+1)g(G^{-1}(F(x)))}dx.\end{aligned}$$

Segue da Proposição 1.1 que $\sqrt{\bar{F}(x)}$ é $RV_{-\frac{\alpha}{2}}$. Desse fato e das hipóteses de $G^{-1}(F)f \in RV_\rho$ e $\bar{F} \in RV_{-\alpha}$, obtemos

$$\frac{\sqrt{\bar{F}(x)}}{f(x)}G^{-1}(F(x))f(x)\frac{x^{\alpha\beta+1}}{(\alpha\beta+1)} \sim x^{\rho+\alpha\beta+2+\frac{\alpha}{2}},\ x \to \infty$$

e, portanto, como $\rho + \alpha\beta + 2 + \frac{\alpha}{2} < 0$, temos que

$$\lim_{x\to\infty} \frac{\sqrt{\bar{F}(x)}G^{-1}(F(x))x^{\alpha\beta+1}}{(\alpha\beta+1)} = 0. \tag{3.13}$$

Recorrendo novamente aos comportamento assintóticos polinomiais conhecidos, segue do fato de $G^{-1}(F)f \in RV_\rho$, $\bar{F} \in RV_{-\alpha}$ e $\sqrt{\bar{F}(x)} \in RV_{-\frac{\alpha}{2}}$, que

$$\frac{G^{-1}(F(x))f(x)x^{\alpha\beta+1}}{2(\alpha\beta+1)\sqrt{\bar{F}(x)}} \sim x^{\rho+\alpha\beta+1+\frac{\alpha}{2}},\ x \to \infty,$$

onde $\rho + \alpha\beta + 1 + \dfrac{\alpha}{2} < -1$. Logo, segue do Teorema de Karamata (Teorema 1.4), que

$$\int_y^\infty \frac{x^{\alpha\beta+1}f(x)G^{-1}(F(x))}{2\sqrt{\bar{F}(x)}(\alpha\beta+1)}dx < \infty. \tag{3.14}$$

Do Teorema 1.4, temos que

$$g^{-1}(G^{-1}(F(x)))G^{-1}(F(x)) \sim \bar{G}(G^{-1}(F(x))) = \bar{F}(x).$$

Desse fato e das hipóteses $G^{-1}(F)f \in RV_\rho$ e $\bar{F} \in RV_{-\alpha}$, temos

$$\frac{x^{\alpha\beta+1}f(x)G^{-1}(F(x))\sqrt{\bar{F}(x)}}{(\alpha\beta+1)g^{-1}(G^{-1}(F(x)))G^{-1}(F(x))} \sim \frac{\alpha x^{\alpha\beta+1}f(x)G^{-1}(F(x))}{(\alpha\beta+1)\sqrt{\bar{F}(x)}} \sim x^{\alpha\beta+\rho+\frac{\alpha}{2}},$$

onde $\alpha\beta + \rho + \frac{\alpha}{2} < -2$.

Então pelo Teorema de Karamata (Teorema 1.4), concluímos que

$$\int_y^\infty \frac{x^{\alpha\beta+1} f(x)\sqrt{\bar{F}(x)}}{(\alpha\beta+1)g^{-1}(G(F(x)))} dx < \infty. \tag{3.15}$$

De (3.13), (3.14) e (3.15), obtemos (3.12). Isso completa a prova de (3.10) para o caso 1. As demonstrações para os demais casos seguem de argumentos similares.

□

Lema 3.3. *Sob as hipóteses do Teorema 3.2, temos*

(i)

$$\lim_{n\to\infty} \sqrt{n} \int_{1-\frac{1}{n}}^1 F^{-1}(t)G^{-1}(t)w(t)dt = 0 \ , \ \lim_{n\to\infty} \sqrt{n} \int_0^{\frac{1}{n}} F^{-1}(t)G^{-1}(t)w(t)dt = 0,$$

(ii)

$$\sqrt{n} \int_{1-\frac{1}{n}}^1 F_n^{-1}(t)G^{-1}(t)w(t)dt \xrightarrow{p} 0 \ \ e \ \ \sqrt{n} \int_0^{\frac{1}{n}} F_n^{-1}(t)G^{-1}(t)w(t)dt \xrightarrow{p} 0.$$

Demonstração. (*i*)

Caso 1): suponha que w satisfaça (3.2) na Condição 3.2.

É suficiente mostrar que

$$\lim_{n\to\infty} \sqrt{n} \int_{1-\frac{1}{n}}^1 F^{-1}(t)G^{-1}(t)w(t)dt = 0, \tag{3.16}$$

pois a verificação de que $\lim_{n\to\infty} \sqrt{n} \int_0^{\frac{1}{n}} F^{-1}(t)G^{-1}(t)w(t)dt = 0$ pode ser feita de maneira análoga. Vimos na prova do Lema 2.7 que

$$a_{n,F} = \inf\left\{y : y > 0 \ , \ \bar{F}(y) \le \frac{1}{n}\right\} = F^{-1}\left(1 - \frac{1}{n}\right)$$

e

$$a_{n,G} = \inf\left\{y : y > 0 \ , \ \bar{G}(y) \le \frac{1}{n}\right\} = G^{-1}\left(1 - \frac{1}{n}\right).$$

E que temos $\lim_{n\to\infty} a_{n,F} = \infty$ e $\lim_{n\to\infty} a_{n,G} = \infty$.

Sejam k_*^+ e u_*^+ definidos na Condição 3.2. Para obter (3.16), note que temos para n grande $a_{n,F} > F^{-1}(u_*^+)$ e também

$$\begin{aligned}\lim_{n\to\infty} \sqrt{n} \int_{1-\frac{1}{n}}^{1} F^{-1}(t)G^{-1}(t)w(t)dt &= \sqrt{n} \int_{a_{n,F}}^{\infty} xG^{-1}(F(x))w(F(x))f(x)dx \\ &= k_*^+ \sqrt{n} \int_{a_{n,F}}^{\infty} x(\bar{F}(x))^{-\beta} dU(x),\end{aligned}$$

onde

$$U(x) = \int_{x}^{\infty} G^{-1}(F(y))f(y)dy. \tag{3.17}$$

Como $\bar{F} \in RV_{-\alpha}$, para alguma constante c temos $\bar{F}(x) \sim cx^{-\alpha}$. Então

$$\begin{aligned}k_*^+ \sqrt{n} \int_{a_{n,F}}^{\infty} x(\bar{F}(x))^{-\beta} dU(x) &\sim k_*^+ \sqrt{n} \int_{a_{n,F}}^{\infty} x^{\alpha\beta+1} dU(x) \\ &= k_*^+ \sqrt{n} U(a_{n,F}) a_{n,F}^{\alpha\beta+1} \frac{\displaystyle\int_{a_{n,F}}^{\infty} x^{\alpha\beta+1} dU(x)}{a_{n,F}^{\alpha\beta+1} U(a_{n,F})}.\end{aligned}$$

Por hipótese, $G^{-1}(F)f \in RV_\rho$ com $\rho < -1$, segue do Teorema de Karamata (Teorema 1.4) que $U \in RV_{\rho+1}$ e que U é limitada, assim, as hipóteses da Proposição 1.4 estão satisfeitas, pois temos também $\alpha\beta + 1 + \rho < 0$. Desse modo temos

$$\lim_{n\to\infty} \frac{\displaystyle\int_{a_{n,F}}^{\infty} x^{\alpha\beta+1} dU(x)}{a_{n,F}^{\alpha\beta+1} U(a_{n,F})} = \frac{-\rho}{\alpha\beta + 1 + \rho}. \tag{3.18}$$

Ademais, note que fazendo $x = a_{n,F}$ em (3.17) obtemos

$$U(a_{n,F}) = \int_{a_{n,G}}^{\infty} z d\bar{G}(z),$$

e, portanto,

$$\begin{aligned}\lim_{n\to\infty} k_*^+ \sqrt{n} U(a_{n,F}) a_{n,F}^{\alpha\beta+1} \frac{\displaystyle\int_{a_{n,F}}^{\infty} x^{\alpha\beta+1} dU(x)}{a_{n,F}^{\alpha\beta+1} U(a_{n,F})} &= k_*^+ \left(\frac{-\rho}{\alpha\beta + 1 + \rho} \right) \lim_{n\to\infty} \sqrt{n} \int_{a_{n,\bar{G}}}^{\infty} z d\bar{G}(z) a_{n,F}^{\alpha\beta+1} \\ &= c_* \lim_{n\to\infty} \sqrt{n} \bar{G}(a_{n,G}) a_{n,F}^{\alpha\beta+1} a_{n,G} \frac{\displaystyle\int_{a_{n,G}}^{\infty} z d\bar{G}(z)}{a_{n,G} \bar{G}(a_{n,G})},\end{aligned}$$

onde $c_* = k_* \left(\dfrac{-\rho}{\alpha\beta + \rho + 1} \right)$. Como $-\gamma + 1 < 0$, usando a Proposição 1.4 segue que

$$\lim_{n\to\infty} \frac{\displaystyle\int_{a_{n,G}}^{\infty} z d\bar{G}(z)}{a_{n,G}\bar{G}(a_{n,G})} = \frac{\gamma}{-\gamma + 1}.$$

E para completar a prova, note que $\sqrt{n}\bar{G}(a_{n,G}) = \dfrac{1}{\sqrt{n}}$. Além disso, vimos na prova do Lema 2.7 que $a_n \in RV_{\frac{1}{\alpha}}$ e $a_{n,\bar{G}} \in RV_{\frac{1}{\gamma}}$, consequentemente, temos que

$$\lim_{n\to\infty} \frac{a_n^{\alpha\beta+1} a_{n,\bar{G}}}{\sqrt{n}} = 0,$$

pois $\beta + \dfrac{1}{\gamma} + \dfrac{1}{\alpha} - \dfrac{1}{2} < 0$, já que $\beta < \dfrac{-2}{\alpha} - 1$ e $1 < \alpha, \gamma < 2$. E (3.16) está provado para o caso 1. Para os demais casos, a prova do item (i) é análoga.

(ii)

Caso 1): suponha que w satisfaça (3.2) na Condição 3.2.

Seja $\{X_1, ..., X_n\}$ uma amostra de F. Note que $F \in D_{max}(\Phi_\alpha)$, logo $\dfrac{\max\{X_1, ..., X_n\}}{a_{n,F}}$ tem distribuição limite não degenerada pelo Teorema 1.10. Além disto,

$$F_n^{-1}(t) = \max\{X_1, ..., X_n\}, \text{ para } 1 - \frac{1}{n} < t \leq 1.$$

E, portanto, teremos que

$$\lim_{n\to\infty} \sqrt{n} \int_{1-\frac{1}{n}}^{1} F_n^{-1}(t) G^{-1}(t) w(t) dt = \sqrt{n} \max\{X_1, ..., X_n\} \int_{a_{n,F}}^{\infty} G^{-1}(F(x)) w(F(x)) f(x) dx.$$

A continuação da prova pode ser omitida. Isto se justifica, pois a demonstração é, essencialmente, similar a do item (i) feita para o caso 1. A demonstração dos demais casos segue ideias similares. □

Dispomos de todos os resultados auxiliares necessários à prova do Teorema 3.2, então, daremos continuidade com a prova do teorema.

Demonstração do Teorema 3.2. Note que temos

$$\sqrt{n}(d_{2,w}(F_n, G) - d_{2,w}(F, G)) \;=$$

$$
\begin{aligned}
&= \sqrt{n}\int_0^{\frac{1}{n}} \left(F_n^{-1}(t) - F^{-1}(t)\right)^2 w(t)dt + \\
&+ \sqrt{n}\int_{1-\frac{1}{n}}^{1} \left(F_n^{-1}(t) - F^{-1}(t)\right)^2 w(t)dt + \\
&+ 2\sqrt{n}\int_0^{\frac{1}{n}} (F_n^{-1}(t) - F^{-1}(t))(F^{-1}(t) - G^{-1}(t))w(t)dt + \\
&+ 2\sqrt{n}\int_{1-\frac{1}{n}}^{1} (F_n^{-1}(t) - F^{-1}(t))(F^{-1}(t) - G^{-1}(t))w(t)dt + \\
&+ \sqrt{n}\int_{\frac{1}{n}}^{1-\frac{1}{n}} \left(F_n^{-1}(t) - F^{-1}(t)\right)^2 w(t)dt + \\
&+ 2\sqrt{n}\int_{\frac{1}{n}}^{1-\frac{1}{n}} (F_n^{-1}(t) - F^{-1}(t))(F^{-1}(t) - G^{-1}(t))w(t)dt = \\
&= A_{n,1} + A_{n,2} + A_{n,3} + A_{n,4} + A_{n,5} + A_{n,6}.
\end{aligned}
$$

Pelo Lema 3.1, $A_{n,1} \xrightarrow{p} 0$ e $A_{n,2} \xrightarrow{p} 0$. Usando o Lema 3.3 e novamente o Lema 3.1, temos que $A_{n,3} \xrightarrow{p} 0$ e $A_{n,4} \xrightarrow{p} 0$. Consequentemente, temos

$$
\lim_{n\to\infty} \sqrt{n}(d_{2,w}(F_n, G) - d_{2,w}(F, G)) =
$$

$$
\begin{aligned}
&= \lim_{n\to\infty} \sqrt{n}\int_{\frac{1}{n}}^{1-\frac{1}{n}} \left(F_n^{-1}(t) - F^{-1}(t)\right)^2 w(t)dt \\
&+ 2\lim_{n\to\infty} \sqrt{n}\int_{\frac{1}{n}}^{1-\frac{1}{n}} (F_n^{-1}(t) - F^{-1}(t))(F^{-1}(t) - G^{-1}(t))w(t)dt \\
&= \frac{1}{\sqrt{n}}\int_{\frac{1}{n}}^{1-\frac{1}{n}} \frac{\rho_{n,F}^2(t)w(t)}{f^2(F^{-1}(t))}dt + 2\lim_{n\to\infty}\int_{\frac{1}{n}}^{1-\frac{1}{n}} \frac{\rho_{n,F}(t)(F^{-1}(t) - G^{-1}(t))}{f(F^{-1}(t))}w(t)dt,
\end{aligned}
$$

em que $\rho_{n,F}(t) := \sqrt{n}(F_n^{-1}(t) - F^{-1}(t))f(F^{-1}(t))$.

O Teorema 1.1 nos garante a existência de uma sequência de pontes Brownianas

$\{B_n(t)\}_n$, tal que

$$n^{\frac{1}{2}-\nu} \sup_{\frac{1}{n}\le t\le 1-\frac{1}{n}} \frac{|\rho_{n,F}(t)-B_n(t)|}{(t(1-t))^\nu} = \begin{cases} O(\log n), & \text{se } \nu = 0; \\ O(1), & \text{se } 0<\nu\le \frac{1}{2}. \end{cases}$$

Segue que

$$\begin{aligned}
\sqrt{n}(d_{2,w}(F_n,G)-d_{2,w}(F,G)) &= \frac{1}{\sqrt{n}}\int_{\frac{1}{n}}^{1-\frac{1}{n}} \frac{(\rho_{n,F}(t)-B_n(t))^2 w(t)}{f^2(F^{-1}(t))}dt \\
&+ \frac{1}{\sqrt{n}}\int_{\frac{1}{n}}^{1-\frac{1}{n}} \frac{B_n^2(t)w(t)}{f^2(F^{-1}(t))}dt \\
&+ \frac{2}{\sqrt{n}}\int_{\frac{1}{n}}^{1-\frac{1}{n}} \frac{B_n(t)(\rho_{n,F}(t)-B_n(t))w(t)}{f^2(F^{-1}(t))}dt \\
&+ 2\int_{\frac{1}{n}}^{1-\frac{1}{n}} \frac{(\rho_{n,F}(t)-B_n(t))(F^{-1}(t)-G^{-1}(t))w(t)}{f(F^{-1}(t))}dt \\
&+ 2\int_{\frac{1}{n}}^{1-\frac{1}{n}} \frac{B_n(t)(F^{-1}(t)-G^{-1}(t))w(t)}{f(F^{-1}(t))}dt \\
&= D_{n,1}+D_{n,2}+D_{n,3}+D_{n,4}+D_{n,5}.
\end{aligned}$$

Tomando-se $\nu = \dfrac{1}{2}$ no Teorema 1.1, obtemos

$$D_{n,1} \le \frac{O(1)}{\sqrt{n}} \int_0^1 \frac{t(1-t)w(t)}{f^2(F^{-1}(t))}dt.$$

Provamos no Lema 2.6 que $\displaystyle\int_0^1 \frac{t(1-t)w(t)}{f^2(F^{-1}(t))}dt < \infty$ e isso completa a prova de que o $\lim\limits_{n\to\infty} D_{n,1} = 0$. Por argumentos similares, temos que o $\lim\limits_{n\to\infty} D_{n,2} = 0$. Novamente do Teorema 1.1 para $\nu = \dfrac{1}{2}$, obtemos que o $\lim\limits_{n\to\infty} D_{n,3} = 0$ pelo Lema 2.6.

Pelo Lema 3.2, temos que

$$\int_{\frac{1}{n}}^{1-\frac{1}{n}} \frac{\sqrt{t(1-t)}(F^{-1}(t)-G^{-1}(t))}{f(F^{-1}(t))} w(t)dt < \infty.$$

Desse fato e do Teorema 1.1 para $\nu \in \left(0, \dfrac{1}{2}\right)$, temos que

$$D_{n,4} \le \frac{2}{n^{\frac{1}{2}-\nu}} \int_{\frac{1}{n}}^{1-\frac{1}{n}} \frac{(t(1-t))^\nu(F^{-1}(t)-G^{-1}(t))}{f(F^{-1}(t))} w(t)dt \to 0,$$

o que completa a prova.

$\square$

Conseguimos concluir, dessa forma, os objetivos propostos nesta tese. Apresentamos testes de similaridade para as distribuições estáveis. Esses testes de correlação para estáveis obtidos estendem o uso da metodologia de testes de correlação na distância Mallows-Wasserstein para distribuições contínuas regularmente variantes com cauda pesada. O nosso Teorema 2.3 surgiu da necessidade de obter uma versão do Teorema 2.1, apresentado em [11] para a distribuição gaussiana, que fosse válido para essas distribuições. Já o nosso Teorema 3.2, por sua vez, foi a alternativa encontrada para obter um resultado análogo ao que foi apresentado em [14] no Teorema 2.2, que possui uma exigência muito forte sobre os momentos.

Referências Bibliográficas

[1] Barbosa, E.G., Dorea, C.C.Y. *A note on the Lindeberg condition for convergence to stable laws in Mallows distance*, Bernoulli, Vol. 15, 922-924, 2009.

[2] Del Barrio, E., Giné, E., Utzet, F., *Asymptotic for L_2 Functionals of the Empirical Quantile Process, with Applications to Tests of Fit Based on Weighted Wasserstein Distance*. Bernoulli **11**, 131-189, 2005.

[3] Bickel, P.J., Freedman, D.A., *Some asymptotic theory for the Bootstrap*. The Annals of Statistics, Vol.9, 1196-1217, 1981.

[4] Bingham, N., Goldie, C., Teugels, J., *Regular Variation*. Cambridge University Press, 1989.

[5] Csorgo, M., *Quantile Processes with Statistical Application*. Society for Industrial and Applied Mathematics, 1983.

[6] Csorgo, S., *Weighted Correlation Tests for Scale Families*. Sociedad de Estadística e Investigación Operativa Vol. 11, Nº 1, 219-248, 2002.

[7] Csorgo, S., *Weighted Correlation Tests for Location-Scale Families*. Mathematical and Computer Modelling **38**, 753-762, 2003.

[8] Csorgo, S., Szabó, T., *Weighted Correlation Tests for Gamma and Lognormal Families*. Tatra Mountains Mathematical Publications **26**, 337-356, 2003.

[9] Csorgo, S., Szabó, T., *Weighted Quantile Correlation Tests for Gumbel, Weibull and Pareto Families*. Probability and Mathematical Statistics Vol. 29, Fasc2, 227-250, 2009.

[10] Csorgo, M., Csorgo, S., Horvath, L., Mason, D., *Weighted Empirical and Quantile Processes*. The Annals of Probability Vol. 14, Nº 1, 31-85, 2010.

[11] Cuesta, A., del Barrio, E., Matrán, C., Rodriguez, R.,*Tests of Goodness of Fit Based on the L_2-Wasserstein Distance*. The Annals of Statistics, 1230-1239, 1999.

[12] Cuesta, A., del Barrio, E., Matrán, C., *Contributions of Empirical and Quantile Processes to the Asymptotic Theory of Goodness-of-Fit Tests*. Sociedad de Estadística e Investigación Operativa Vol. 0, N^o 1, 1-96, 2000.

[13] Cuesta, A., del Barrio, E., Esteban, P., Matrán, C., *Similarity of probability measures through trimming*. Annals of Probability, 2008.

[14] Cuesta, A., del Barrio, E., Esteban, P., Matrán, C., *Trimmed Comparison of Distributions*. Journal of the American Statistical Association Vol. **103**, N^o 482, 697-704, 2008.

[15] Cuesta, A., Del Barrio, E., Esteban, P., Matrán, C., *Similarity of Samples and Trimming*. Preprint submitted to the Annals of Statistics, 2009.

[16] 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, N^o 1, 89-107, 2002.

[17] De Wet, T., *Discussion of del Barrio, Cuesta e Matrán([12])*. Test **9**(2000), 74-79.

[18] Czado, C., Freitag, G., Munk, Axel., *A Nonparametric Test for Similarity of Marginals - with Applications to the Assessment of Population Bioequivalence*. Journal of Statistical Planning and Inference **137**, 697-711, 2007.

[19] Czado, C., Munk, A., *Nonparametric Validation of Similar Distributions and Assessment of Goodness of Fit*. Journal Royal Statistical Society B **60**, Part I, 223-241, 1998.

[20] Dorea, C.C.Y., Ferreira D.B. *Conditions for Equivalence Between Mallows Distance and Convergence to Stable Laws*. Acta Math. Hungar., Vol. 134, 1-11, 2012.

[21] Embrechts, P., Kluppelberg, C., Mikosch, T., *Modelling Extremal Events for Insurance and Finance*. Springer, 1997.

[22] Fama, E., *The behavior of Stock Market Prices*. The Journal of business Vol. 38, N^o 1, 34-105, 1965.

[23] Fama, E., Rool, R., *Some Properties of Symmetric Stable Distributions*. Americal Statistical Association Journal, 817-836, 1968.

[24] Feller, W., *An Introduction to Probability theory and its Applications II.* Wiley, New York, 1971.

[25] Fisher, R. A., Tippet, L. H. C., *Limiting Forms of the Frequency Distributions of the Largest or Smallest Member of a Sample.* Proc. Cambridge Philos. Soc. **24**, 180-190, 1928.

[26] Fofack, H., Nolan, J., *Tail Behavior, Modes and Others Characteristics of Stable Distributions.* Extremes 2:1, 39-58, 1999.

[27] Fréchet, M., *Sur la loi de probabilité de l'écart maximum.* Ann. de la Soc. polanaise de Math. **6**, 93, 1927, Cracow.

[28] Galambos, J., *The Asymptotic Theory of Extreme Order Statistics.* Jonh Wiley, New York, 1978.

[29] Galambos, J., *The Development of the Mathematical Theory of Extremes in the Past Half Century.* Theory Probab. Appl. **39**, 234-248, 1994.

[30] Gnedenko, B. V., *Sur la distribution limite du terme maximum d´une série aleatóire.* Ann. Math. **44**, 423-453, 1963.

[31] Gordaliza, A., *Best approximations to random variables based on trimming procedures.* Journal Approx Theor., **64**, 162-180, 1991.

[32] Jonhson, O., Samworth, R., *Convergence of the Empirical Process in Mallows Distance, with an Application to Goodness-of-Fit Tests.* Preprint, 2005.

[33] Jonhson, O., Samworth, R., *Central Limit Theorem and Convergence to Stable Laws in Mallows distance.* Bernoulli **11**, N° 5, 829-845, 2005.

[34] Jonhson, O., Samworth, R., *The Empirical Process in Mallows Distance, with Application to Bootstrap Performance.* Preprint, 2008.

[35] Haan, L., *On Regular Variation and its Application to the Weak Convergence of Sample Extremes.* Mathematical Centre, 1970.

[36] Nolan, J., *Parametrizations and Modes of Stable Distributions.* Statistical and Probability Letters **38**, 187-195, 1998.

[37] Nolan, J., *Stable distributions models for heavy tailed data.* Birkhauser, 187-195, 2007. **Nota:** Em progresso, Capítulo 1 online em http://academic2.american.edu/jp-nolan/stable/cap1.pdf.

[38] Resnick, S., *Tail Equivalence and its Applications.* Journal Applied Probability **8**, 135-156, 1971.

[39] Resnick, I., *Extreme Values, Regular Variation and Point Process.* Springer-Verlag, Berlim, 1987.

[40] Samorodnistky, G., Taqqu, M., *Stable non Gaussian Random Variable Processes.* Chapman Hall, 1994.

[41] Shorack, G., Wellner, J., *Empirical processes with applications to statistics.* Wiley, New York, 1986.