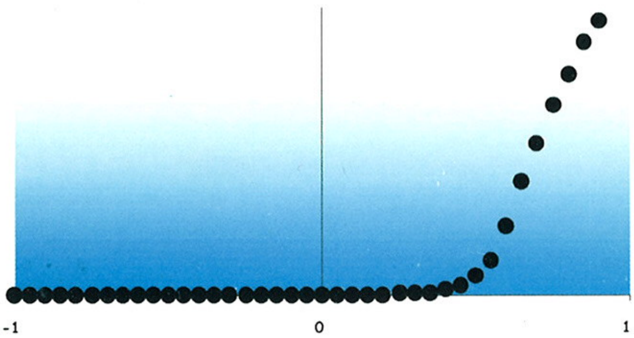
UNIVERSIDADE DA MADEIRA

Tópicos sobre a Convergência Fraca de Sucessões de Variáveis Aleatórias

Sandra Mendonça



Dezembro de 2000

Tópicos sobre a Convergência Fraca de Sucessões de Variáveis Aleatórias

Sandra Mendonça

Universidade da Madeira
Dezembro de 2000

Investigação parcialmente suportada pela FCT / PRAXIS XXI / FEDER
e POCTI - Plano Operacional de Ciência, Tecnologia e Inovação.

Tópicos sobre a Convergência Fraca de Sucessões de Variáveis Aleatórias

de

Sandra Mendonça

é uma monografia escrita com o objectivo de obter o grau de Doutor em Matemática, na especialidade de Probabilidades e Estatística. A elaboração da monografia teve orientação do Professor Dinis Pestana, da Faculdade de Ciências da Universidade de Lisboa. A apresentação pública terá lugar na Universidade da Madeira.

Capa: Distribuição empírica da variável $\dfrac{\max_{1\leq i\leq 6} X_i}{\left(\sum_{i=1}^{6}|X_i|^2\right)^{1/2}}$, com base em uma amostra, criada pelo programa Microsoft Excel, de 600 vectores pseudo-aleatórios cujas entradas têm a distribuição de uma Gumbel padrão.

Agradecimento

Às pessoas
que me formaram
e às pessoas
com quem partilho os meus dias

Índice

Notação

Ao longo deste trabalho faremos uso da seguinte notação:

O símbolo/ /A expressão	designa/abrevia/substitui
v.a.	variável aleatória ("v.a.'s" no plural)
f.d.	função de distribuição de probabilidade
f.d.p.	função de densidade de probabilidade
i.i.d.	independentes e identicamente distribuídas
$\xrightarrow{d}$	converge em distribuição
$\xrightarrow{p}$	converge em probabilidade
$X \frown Dist$	X com distribuição $Dist$
$X \stackrel{d}{=} Y$	X e Y têm a mesma função de distribuição
F_X	função de distribuição de probabilidade da v.a. X
f_X	função de densidade de probabilidade da v.a. X
μ_F	medida de Lebesgue-Stieltjes induzida pela f.d. F
$\alpha(F)$	$\inf\left\{x \in \overline{\mathbf{R}} : F(x) > 0\right\}$
$\omega(F)$	$\sup\left\{x \in \overline{\mathbf{R}} : F(x) < 1\right\}$
$\mathbf{N}_1$	números naturais positivos
$\{X_i\}_{i\in\mathbf{N}_1}$	sucessão de variáveis aleatórias independentes
S_n	$\sum\limits_{i=1}^{n} X_i$
$F * G\,(x)$	$\int_{\mathbf{R}} F(x-y)\,dG(y) := \int_{\mathbf{R}} F(x-y)\,d\mu_G(y)$
$X_{i:n}$	i-ésima estatística ordinal de um vector aleatório $(X_1, ..., X_n)$
$\mathbf{I}_A$	função indicadora do conjunto A
$[x]\,,\, x \in \mathbf{R}$	o maior inteiro inferior ou igual a x
$\mathrm{Be}(\alpha, \beta)$	distribuição Beta de parâmetros $\alpha, \beta > 0$, de densidade $f_X(x) = \frac{x^{\alpha-1}(1-x)^{\beta-1}}{B(p,q)}\mathbf{I}_{[0,1]}(x)$
$\mathrm{Gau}(\lambda, \sigma)$	distribuição Gaussiana de valor médio λ e desvio-padrão σ, de densidade $f_X(x) = \frac{1}{\sqrt{2\pi}\sigma}\exp\left[-\frac{(x-\mu)^2}{2\sigma^2}\right]$

1 Sobre o assunto desta tese

I hear, I know. I see, I remember. I do, I understand.

Confucius

A presente tese foi preparada e escrita durante os anos lectivos 1998-1999, 1999-2000 e os três primeiros meses do presente ano lectivo. O apoio do Departamento de Matemática da Universidade da Madeira que me concedeu a dispensa de serviço durante todo este tempo, da própria Universidade, do Ministério da Educação (via uma bolsa do Programa de Desenvolvimento Educativo para Portugal), do Centro de Estatística e Aplicações da Universidade de Lisboa, e, claro está, o apoio dos meus familiares, amigos e colegas, em particular do Professor Dinis Pestana, orientador da presente tese, tornaram possível a elaboração deste trabalho. Esta oportunidade, conjuntamente com a ânsia e o prazer pelo saber, conduziram a muitas leituras – das quais a bibliografia serve de prova – e à abordagem de assuntos diversos (a abrangência do título assim o indica). Existe no entanto, e como não poderia deixar de ser, um eixo condutor em torno do qual os diferentes temas se associam. Esse eixo é a convergência fraca de sucessões de variáveis aleatórias, com especial relevância para sucessões de funções de somas ou de máximos de variáveis aleatórias independentes.

Depois de uma breve resenha histórica sobre a Teoria da Probabilidade e da Estatística (Capítulo 2), vamos proceder a normalizações aleatórias, construídas a partir das próprias variáveis aleatórias em jogo (Secção 3.1).

Logan *et al.* [41] tinham já abordado a questão no contexto de somas

$$Y_n(p) = \sum_{i=1}^{n} X_i \Big/ \left(\sum_{i=1}^{n} |X_i|^p \right)^{1/p} \quad (p > 0)$$

(que, no caso $p \longrightarrow +\infty$, resulta em $Y_n(+\infty) = \sum_{i=1}^{n} X_i \Big/ \max_{1 \le i \le n} X_i$, sucessão definida e estudada por Darling [7]). Os resultados são fascinantes, e pareceu interessante investigar no caso dos máximos, as variáveis aleatórias

$$W_n(p) = \frac{\max_{1 \le i \le n} X_i}{\left(\sum_{i=1}^{n} |X_i|^p \right)^{1/p}} \quad \text{e} \quad W_n(+\infty) = \frac{\max_{1 \le i \le n} X_i}{\max_{1 \le i \le n} |X_i|};$$

sobre o assunto apresentamos resultados exactos e assintóticos, e estudamos $P[W_n(+\infty) = 1]$.

Na nossa abordagem ao estudo das classes de Mejzler começamos por considerar algumas restrições no que se refere à proporção do número de réplicas de cada um dos tipos considerados. Alargamos depois os nossos resultados à situação em que o número de réplicas de cada tipo é aleatório (Parágrafo 3.2.1). Como generalização do esquema clássico de máximos, abordamos ainda a teoria dos domínios de atracção parcial (Parágrafo 3.2.2).

Procedemos depois a investigações de simetria usando funções de estatísticas ordinais,

$$Q_{1n} = \frac{X_{n:n} - X_{n-1:n}}{X_{2:n} - X_{1:n}} \quad \text{e} \quad Q_{2n} = \frac{X_{n:n} - X_{2:n}}{X_{n-1:n} - X_{1:n}},$$

que avaliam o equilíbrio entre as caudas direita e esquerda da distribuição parente (Secção 3.3).

Recordamos que k é completamente monótona se, e só se, $(-1)^n k^{(n)}(x) \geq 0$, $n = 0, 1, \ldots$, e que pode ser representada como transformada de Laplace de uma função monótona não decrescente (resultado de Bernstein, cf. Teorema 4). A representação integral de funções monótonas de ordem r, que surgiram no trabalho de Graça Martins e Pestana [23], aquando a construção do *simile* das classes de Khintchine no esquema de máximos, foi estudada por Pestana e Mendonça [56], que apresentam fórmulas de inversão da transformada integral, e relacionam o seu estudo com unimodalidade generalizada, estatísticas ordinais e transformadas beta (Secção 3.4).

2 O passado — Um pouco de história

History is the witness that testifies to the passing of time; it illumines reality, vitalizes memory, provides guidance in daily life and brings us tidings of antiquity.

Marcus Tulius Cicero

É usual considerar como data do nascimento da Teoria da Probabilidade e da Estatística os anos situados nos meados do século XVIII, pois foi nesta altura que primeiro se deu a transposição de métodos e resultados originados em problemas particulares (surgidos essencialmente nos jogos e na astronomia) para outras áreas, distintas das originais. Este corte umbilical — a representação abstracta e geral de resultados concretos e particulares — constitui a base de qualquer desenvolvimento da grande malha que é a ciência.

Nesta monografia vamos trabalhar um pouco na parte desta malha que trata da convergência fraca de sucessões de variáveis aleatórias, em particular as associadas à teoria das somas e à teoria dos extremos de variáveis aleatórias independentes. Embora historicamente estas teorias se tenham desenvolvido separadamente, é uma separação que aponta mais para um paralelismo entre as duas teorias. Enquanto nas somas se procurou relaxar a hipótese de identidade distribucional, no caso dos extremos procurou-se relaxar a hipótese de independência. Para compreender melhor a evolução destas duas sub-teorias da Teoria da Probabilidade e da Estatítisca, segue-se uma compilação de alguns factos históricos (cf. Stigler [71], Le Cam [37] e Gnedenko e Kolmogorov [17]). Fazer história, o que, absolutamente, não é o caso, ou mesmo tentar apresentar os acontecimentos mais importantes de algum período histórico, é sempre um pouco arriscado, porque sempre fica muito por dizer, e o que é dito deixa para segundo plano tudo o resto. É importante não esquecer que a realidade vai sempre além de tudo o que se possa dizer.

2.1 As primeiras leis limite

The epistemological value of the theory of probability is revealed only by limit theorems.

Gnedenko e Kolmogorov

Desde muito cedo na história da Teoria da Probabilidade e da Estatística a soma de variáveis aleatórias teve um papel fulcral no seu desenvolvimento, papel este que, obviamente, foi evoluindo ao longo do tempo. Inicialmente a soma de variáveis aleatórias surgiu na *combinação das observações*, na sua maioria observações ligadas à astronomia. Embora os astrónomos do século XVIII se mostrassem reticentes em usar uma estatística, hoje tão comum, como a média, pois receavam que ao fazê-lo estariam a propagar os erros, e não a compensá-los, foi nesta área que os métodos estatísticos tiveram o seu primeiro impulso; atenda-se por exemplo à criação do que hoje designamos como método dos mínimos quadrados por Adrien Marie Legendre, em 1805, quando este tentava encontrar a forma da terra com base nas medições de um arco de meridiano.

Mas, citando uma frase presente no prefácio do notável livro de Gnedenko e Kolmogorov [17], *o valor epistemológico da Teoria da Probabilidade é baseado no facto de que os fenómenos aleatórios em larga escala, na sua acção colectiva produzem efeitos precisos, não aleatórios, e é apenas revelado através dos teoremas limite.* Os dois primeiros teoremas limite — a lei dos grandes números e o teorema limite central — surgiram de uma forma simplificada e a versão actual é fruto da complexificação e da generalização dos primeiros resultados.

A lei dos grandes números começou por surgir no caso particular da distribuição binomial pela mente criativa de um de entre os doze membros da família Bernoulli — esta é a estimativa que Stigler [71] apresenta do número de elementos da família Bernoulli que, de alguma forma, contribuíram de uma forma significativa para a evolução da Matemática e da

Física. Falamos de Jacob Bernoulli, o *pai da quantificação da incerteza.* Alguns dos seus trabalhos não publicados em vida foram compilados pelo seu sobrinho, Nicholas Bernoulli, num livro intitulado *Ars Conjectandi,* datado de 1713, e que inclui a primeira versão da lei dos grandes números.

O trabalho de Jacob Bernoulli teve continuação com o seu sobrinho Nicolas Bernoulli e com o francês Abraham De Moivre. De Moivre no seu livro *The Doctrine of Chances* (de 1738) mostra, usando o que viria a ser a fórmula de Stirling, que as distribuições binomiais podem ser aproximadas pelas gaussianas. Este embrião do **teorema limite central,** título que Polyá deu ao resultado (cf. Le Cam [37]), para assim salientar a sua importância, desenvolveu-se com a ajuda de Pierre Simon de Laplace e de Siméon Dennis Poisson: Laplace em 1812 generalizou o resultado de De Moivre a variáveis independentes, limitadas e suportadas por reticulados e, por sua vez, Poisson, em 1824, estendeu o resultado de Laplace a somas de variáveis independentes e limitadas, não necessariamente idênticas. Gauss foi mais longe e estendeu o resultado à soma de um grande número de parcelas negligíveis, mas sem grandes formalismos. Foi Chebyshev quem, em 1887, de facto formalizou os problemas e quem delineou uma versão final do teorema.

Actualmente existem várias versões do teorema limite central, com condições mais ou menos fortes, consoante o caso, algumas até relativas a variáveis com valores em espaços que não necessariamente $\mathbf{R}^n$, (cf. e.g. Araújo e Giné [2]), outras relativas a sucessões de variáveis dependentes, como os *martingales* e as cadeias de Markov. No entanto, a versão mais conhecida do teorema limite central será talvez a de Lindeberg (1922)-Feller (1935) que passamos a enunciar: dada uma sucessão de variáveis aleatórias (v.a.'s) independentes, $\{X_i\}_{i\in\mathbf{N}_1}$, onde cada variável aleatória X_i tem distribuição F_i, valor médio $\mu_i = E(X_i)$ e variância $\sigma_i^2 = \mathbf{E}\left[(X_i-\mu_i)^2\right]$ finitos, temos que, considerando $s_n = \left(\sum\limits_{i=1}^{n}\sigma_i^2\right)^{\frac{1}{2}}$,

$$\frac{\sum_{i=1}^{n} X_i - \sum_{i=1}^{n}\mu_i}{s_n} \underset{n\to+\infty}{\overset{d}{\longrightarrow}} Z \frown \mathrm{Gau}\,(0,1)$$

se, para todo o real positivo t, a condição de Lindeberg,

$$\frac{1}{s_n^2}\sum_{i=1}^{n}\int_{|x|>ts_n} x^2 dF_i(x) \underset{n\to+\infty}{\longrightarrow} 0,$$

for satisfeita.

Relativamente à **lei dos grandes números**, Chebyshev verificou que dada uma sucessão de variáveis aleatórias independentes, $\{X_i\}_{i\in\mathbf{N}_1}$, com valores médios μ_i finitos e variâncias σ_i^2 finitas $(i=1,\ldots,n)$, a condição $\frac{s_n}{n} \underset{n\to+\infty}{\longrightarrow} 0$, era suficiente para garantir a convergência em probabilidade de $\frac{S_n-\sum_{i=1}^{n}\mu_i}{n}$ para zero, onde $S_n=\sum\limits_{i=1}^{n}X_i$. Para resolver o problema relativamente à convergência em distribuição de uma soma de variáveis normalizadas para a distribuição gaussiana, Chebyshev criou o método dos momentos. Treze anos depois, Ljapunov, seguindo uma outra via - o método das funções características - encontrou também condições suficientes para a referida convergência.

2.2 Variáveis estáveis, a classe L, as leis infinitamente divisíveis e as classes L_m de Urbanik

Another characteristic of mathematical thought is that it can have no success where it cannot generalize.

C. S. Pierce

Na Matemática é quase sempre possível *subir* na escada da generalização, no sentido em que cada conceito pode ser visto como um caso particular de um outro conceito mais geral. A distribuição gaussiana pode ser obtida como o limite, em distribuição, de somas normalizadas de variáveis aleatórias independentes e identicamente distribuídas (i.i.d.) e, como tal, é um caso particular de uma v.a. estável, noção que passamos a recordar:

Definição 1

*Dizemos que a função de distribuição (f.d.) F de uma v.a. X, ou mesmo a v.a. X, é **estável,** se para todo o natural n positivo existirem $b_n > 0$, $a_n \in \mathbf{R}$ tais que*

$$F_{\sum_{i=1}^n X_i}(x) = F_X^{*n}(x) = \underbrace{F_X * \ldots * F_X}_{n}(x) = F_{\frac{X-a_n}{b_n}}(x),$$

onde as variáveis X_i, $i = 1, \ldots, n$, são independentes e idênticas a X.

Por outras palavras, e aproveitando para introduzir um novo conceito — os tipos de Khintchine, uma v.a. X é estável se o seu **tipo,** i.e., o conjunto das variáveis aleatórias obtidas de X por deslizamentos e mudanças de escala, for fechado para somas.

É de salientar que sempre que falamos em classes limite, estamos a falar de convergências de tipos.

A classe das funções de distribuição estáveis surgiu no segundo quartel do nosso século e o seu estudo envolve nomes como Paul Lévy, Khintchine, Gnedenko, Doeblin e Feller. Lévy foi o pioneiro no estudo das variáveis estáveis. Em 1917 foram encomendadas a Lévy três conferências sobre o teorema limite central. Lévy foi além do teorema limite central, e resolveu investigar mais geralmente em que condições somas de variáveis aleatórias i.i.d, convenientemente normalizadas (centradas e com dispersão estabilizada) convergem para uma v.a. Y não degenerada; por outras palavras, em que condições existem constantes de atracção $b_n > 0$ e $a_n \in \mathbf{R}$ tais que

$$\frac{S_n - a_n}{b_n} \xrightarrow[n \to +\infty]{d} Y, \text{ não degenerada.} \tag{1}$$

Lévy em 1925 e Khintchine em 1938 mostraram a coincidência da classe das distribuições limite destas somas normalizadas com a classe das distribuições estáveis. Em 1936, Lévy e Khintchine, mostraram que as distribuições estáveis se caracterizam por ter uma função característica, φ digamos, que satisfaz a seguinte condição

$$\ln \varphi(t) = iat - c\,|t|^{\alpha}\,[1 + i\beta\,\mathrm{sign}\,t\;\omega(t,\alpha)],$$

onde $0 < \alpha \leq 2$, é o **expoente característico** da distribuição estável, β, γ e c são constantes reais, sendo $c \geq 0$ (parâmetro de escala), $|\beta| \leq 1$

(parâmetro de assimetria) e

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

Prova-se ainda que as constantes de normalização em (1) são da forma $b_n = bn^{1/\alpha}$. A distribuição gaussiana, já o referimos, é uma distribuição estável e corresponde a ter $\alpha = 2$. Poucos mais exemplos têm função de distribuição exprimível em forma analítica explícita. Temos ainda as distribuições de Cauchy ($\alpha = 1$; se $\beta = 0$ obtemos a distribuição padrão de Cauchy) e a distribuição de Lévy ($\alpha = \frac{1}{2}$ e $\beta = 1$). Zolotarev [75] apresenta alguns exemplos de distribuições estáveis expressas através de funções especiais como os integrais de Fresnel ($\alpha = \frac{1}{2}$ e $\beta = 0$), a função de MacDonald de ordem 1/3 ($\alpha = \frac{1}{3}$ e $\beta = 1$), as funções de Whittaker ($\alpha = \frac{2}{3}$ e $\beta = 0$, $\alpha = \frac{2}{3}$ e $\beta = 1$, $\alpha = \frac{3}{2}$ e $\beta = \frac{1}{3}$, $\alpha = \frac{3}{2}$ e $\beta = -\frac{1}{3}$), entre outras.

O **domínio de atracção** das distribuições estáveis, i.e., o conjunto das funções de distribuição que, devidamente normalizadas, convergem para uma determinada distribuição estável, foi também completamente caracterizado. Esta caracterização advém da junção de resultados obtidos por Lévy (1935), Khintchine (1935), Feller (1935), Gnedenko (1939) e Doeblin (1940):

Teorema 1

Uma função de distribuição F pertence ao domínio de atracção de uma *distribuição estável com índice α, $0 < \alpha < 2$, se e só se,*[1]

$$\lim_{x \to +\infty} \frac{F(-x)}{1 - F(x)} \in [0, +\infty] \qquad e \qquad \frac{1 - F(x) + F(-x)}{1 - F(kx) + F(-kx)} \underset{x \to +\infty}{\longrightarrow} k^{\alpha}.$$

Uma função de distribuição F pertence ao domínio de atracção de uma distribuição normal se e só se

$$\frac{z^2 \int_{|x| \geq z} dF(x)}{\int_{|x|<z} x^2 dF(x)} \underset{z \to +\infty}{\longrightarrow} 0.$$

Mas a distribuição gaussiana pode também ser obtida como o limite em distribuição de somas normalizadas de variáveis aleatórias independentes, não necessariamente idênticas (tal como se constata no teorema de Lindeberg-Feller); a gaussiana é também um caso particular de uma v.a. da **classe** L que passamos a definir:

Definição 2

A classe L é formada pelas distribuições limite de somas de variáveis aleatórias independentes, $S_n = \sum_{i=1}^{n} X_i$, normalizadas por sucessões de constantes $\{a_n\}_{n \in \mathbf{N}_1}$ e $\{b_n\}_{n \in \mathbf{N}_1}$ ($\forall n$, $b_n > 0$), i.e., de $\frac{\sum_{i=1}^{n} X_i - a_n}{b_n}$, com a condição de que $\frac{X_i}{b_n} \underset{n \to +\infty}{\overset{p}{\longrightarrow}} 0$.

As distribuições da classe L foram caracterizadas por Lévy em 1937: F pertence à classe L se, e só se, $\forall a \in (0, 1)$, $F(x)$ é a convolução de

[1] Recorde-se que a condição $\frac{1-F(x)+F(-x)}{1-F(kx)+F(-kx)} \underset{x \to +\infty}{\longrightarrow} k^{\alpha}$ poderia ser traduzida pela expressão "a função $1 - F(x) + F(-x)$ é de **variação regular** de índice $-\alpha$ em infinito", assunto que será adiante discutido com mais pormenor.

$F\left(\frac{x}{a}\right)$ com uma outra distribuição $F_a(x)$ referente a uma v.a. infinitamente divisível, conceito que definiremos já de seguida (note-se que, por esta razão, as funções da classe L são também designadas por funções de distribuição **auto-decomponíveis**: a v.a. X associada a F pode ser decomposta na soma de duas v.a.'s independentes, $X = aX + Y_a$, em que uma das parcelas é apenas uma redução da v.a. original e a outra, como já referimos, é infinitamente divisível).

As distribuições da classe L são **infinitamente divisíveis**, i.e., para todo o natural positivo n, podem ser escritas como a n-ésima convolução de uma distribuição, e têm função característica φ tal que

$$\ln \varphi(t) = i\gamma t + \int_{-\infty}^{+\infty} \left(\exp(itx) - 1 - \frac{itx}{1+x^2}\right) \frac{1+x^2}{x^2} dG(x), \quad (2)$$

onde $\gamma \in \mathbf{R}$ e G é monótona não decrescente e de variação limitada (em zero a função integranda é definida por continuidade). Esta é a chamada **representação de Lévy-Khintchine** e a função G é conhecida como a **função espectral de Lévy-Khintchine** de φ. Mas nem todas as distribuições infinitamente divisíveis são estáveis; basta, por exemplo, pensar na distribuição de Poisson; no entanto, toda a f.d. infinitamente divisível pode também ser escrita como o limite de funções de distribuição de somas de variáveis aleatórias. De facto, dado um sistema triangular infinitesimal, i.e., uma "matriz" de v.a.'s $\{X_{nk}\}_{n \in \mathbf{N}_1,\, 1 \leq k \leq k_n}$, onde $k_n \underset{n \to +\infty}{\longrightarrow} +\infty$, para cada n, $X_{n1}, \ldots , X_{nk_n}$ são independentes, e

$$\forall \varepsilon > 0, \quad \max_{1 \leq k \leq k_n} \mathbf{P}\left(|X_{nk}| > \varepsilon\right) \underset{n \to +\infty}{\longrightarrow} 0,$$

prova-se que F é infinitamente divisível se, e só se, F é o limite (nos pontos de continuidade de F) de $F_n = F_{X_{n1}+\ldots+X_{nk_n}}$.

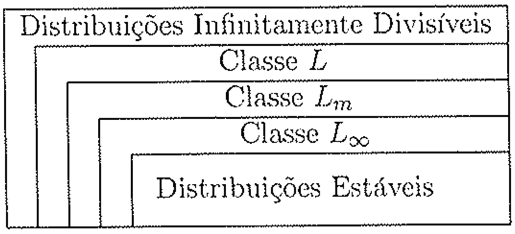
A classe das estáveis é demasiado estreita para modelação, a classe das auto-decomponíveis demasiado larga (admite que as parcelas podem ser de qualquer tipo). Numa tentativa de equilíbrio, Gnedenko conjecturou que se nas parcelas houvesse k tipos de v.a.'s (no sentido de Khintchine) os possíveis limites seriam combinações convexas de, no máximo, k estáveis. Koroljuk e Zolotarev [34] estabeleceram o resultado para $k = 2$, mas Zinger [74] demonstrou que para $k > 2$ a conjectura era falsa.

Uma abordagem, porventura mais interessante ainda, deve-se a Urbanik [72], que usou aquilo que apelidou de "esquemas triangulares de variação lenta". A construção de Urbanik leva a considerar a equação funcional para funções características

$$\varphi(t) = \varphi(at)\, \varphi_a(t), \ \forall a \in (0, 1),$$

onde $\varphi_a \in L_m$; L_0 é a classe das infinitamente divisíveis, resultando deste modo que L_1 é a classe L. Constrói-se assim iterativamente L_m, $m = 1, 2, \ldots$, e $L_\infty = \bigcap_{k=0}^{+\infty} L_m$ parece ser uma classe particularmente tentadora para modelação de fenómenos em que haja a considerar qualquer mecanismo aditivo.

Resumindo, sob a condição $\sup_k \mathbf{P}\left(\left|\frac{X_k}{b_n}\right| > \varepsilon\right) \underset{n \to +\infty}{\longrightarrow} 0$, $\forall \varepsilon > 0$, temos que



Como já referimos os teoremas limite que primeiro emergiram foram versões particulares da lei dos grandes números e do teorema do limite central. Em qualquer um dos casos, mesmo depois da generalização dos teoremas, a finitude da variância é exigida, tornando assim desprezável a contribuição dos extremos para a soma. Esta é talvez a razão do estudo mais tardio da teoria dos extremos.

2.3 Variáveis max-estáveis, a classe M e as classes M_m de Mejzler

Live the Golden Mean (avoid extremes).

Aristóteles

Esta seria a resposta de Aristóteles à pergunta "Como devemos viver?". Durante muito tempo a atitude de Aristóteles aplicou-se de certa forma também à Estatística, já que a preocupação principal dos estatísticos encontrava-se no comportamento central dos dados, comportamento mais facilmente adivinhável.

Quando falamos em extremos de variáveis aleatórias, de imediato pensamos na distribuição limite do máximo de variáveis aleatórias. Poderíamos ampliar os nosso problemas, pensando no segundo máximo, no terceiro máximo, e assim sucessivamente até chegar ao mínimo, mas verifica-se que as distribuições assintóticas destes, quer conjuntamente, quer marginalmente, são funções da distribuição limite do máximo pelo que a nossa atenção debruçar-se-á apenas sobre o estudo de máximos.

Sabemos que, quando as variáveis são independentes, a função característica da soma é o produto das funções características de cada variável. Com a substituição adequada dos objectos matemáticos, algo de semelhante acontece na teoria dos valores extremos: a f.d. do máximo de variáveis aleatórias independentes é o produto das respectivas funções de distribuição. É então natural estabelecer um paralelo entre a teoria das somas e a teoria dos extremos.

Comecemos por considerar uma sucessão de variáveis aleatórias i.i.d. $\{X_i\}_{i\in\mathbf{N}_1}$. O interesse por extremos surgiu já, muito timidamente, nos trabalhos de Bernoulli mas só cerca de duzentos anos mais tarde foram encontrados os três tipos limite possíveis e a caracterização dos domínios de atracção. Passemos aos pormenores. A equação funcional

$$F^{*n}(b_n x + a_n) = F(x) \qquad (b_n > 0),$$

cujas soluções são as distribuições estáveis, foi transportada por Fréchet [13], em 1927, para a teoria dos máximos na forma

$$F^{n}(b_n x + a_n) = F(x) \qquad (b_n > 0).$$

As soluções desta equação, usualmente apelidadas de **max-estáveis**, são precisamente as distribuições que surgem como limite de máximos normalizados: $\frac{X_{n:n}-a_n}{b_n}$ $(b_n > 0)$. Se supusermos, por exemplo, que

$$\lim_{x \to +\infty} x^{\alpha} [1 - F(x)] = b > 0,$$

para algum α positivo então, para $x > 0$,

$$\begin{aligned}
\lim_{n \to +\infty} P\left(X_{n:n} \leq x (bn)^{1/\alpha}\right) &= \lim_{n \to +\infty} \left[F\left(x (bn)^{1/\alpha}\right)\right]^n \\
&= \lim_{n \to +\infty} \left[1 - \tfrac{1}{x^{\alpha} bn} x^{\alpha} bn \left(1 - F\left(x (bn)^{1/\alpha}\right)\right)\right]^n \\
&= \lim_{n \to +\infty} \left[1 - \frac{1}{x^{\alpha} n}\right]^n = \exp\left(-x^{-\alpha}\right).
\end{aligned}$$

Para $x < 0$ temos $\lim\limits_{n \to +\infty} P\left(X_{n:n} \leq x (bn)^{1/\alpha}\right) = 0$. Esta foi a primeira solução a ser encontrada e foi Fréchet [13] quem o fez; esta família de distribuições é hoje designada pelo seu nome, a família de distribuições de Fréchet:

$$\Phi\gamma(x) = \exp\left(-x^{-\gamma}\right) \mathbf{I}_{[0,+\infty)}(x), \quad (\gamma > 0).$$

Em 1928, Fisher e Tippet [12] encontraram as outras duas possíveis soluções: a família das distribuições Weibull

$$\Psi_{\gamma}(x) = [1 - \exp(-(-x)^{\gamma})] \mathbf{I}_{(-\infty,0)}(x) + \mathbf{I}_{[0,+\infty)}(x), \quad (\gamma > 0),$$

e a distribuição Gumbel

$$\Lambda(x) = \exp(-\exp(-x)).$$

Nos pontos interiores aos respectivos suportes, estas famílias de distribuições podem ser escritas conjuntamente através da **representação de Von-Mises–Jenkinson**

$$\Gamma_\tau(x) = \exp\left(-(1+\tau x)^{-1/\tau}\right),\ 1+\tau x > 0$$

(sendo $\Gamma_0(x) := \lim_{\tau\to 0}\Gamma_\tau(x)$). Algumas funções de distribuição absolutamente contínuas foram devidamente enquadradas nos respectivos domínios de atracção de cada uma destas distribuições por von Mises em 1936 [49] (*algumas*, pois as condições de von Mises não eram necessárias, apenas suficientes). Seis anos mais tarde, em 1943, Gnedenko [15] caracterizou completamente os domínios de atracção da família de distribuições Fréchet e da família de distribuições Weibull. Gnedenko encontrou também uma caracterização para o domínio de atracção de Gumbel, mas foi Haan [26] quem caracterizou completamente este domínio.

Sumariando, os domínios de atracção das max-estáveis caracterizam-se do seguinte modo (supõe-se que a v.a X tem a si associada a f.d. F):

- A f.d. F pertence ao domínio de atracção da família de distribuições Fréchet se, e só se, $1-F$ é de variação regular de índice $-\gamma$ (cf. a nota de rodapé 1) e se o seu suporte for ilimitado superiormente; simbolicamente,

$$\exists\,\{b_n\}_{n\in\mathbf{N}_1},\{a_n\}_{n\in\mathbf{N}_1},\gamma>0:$$
$$(\forall n\in\mathbf{N}_1, b_n>0)\wedge\left(\lim_{n\to+\infty}P\left[X_{n:n}\le b_nx+a_n\right]=\Phi\gamma(x)\right)\Leftrightarrow$$

$$\Leftrightarrow \omega(F) = +\infty \wedge \left(\exists\gamma > 0, \forall x > 0, \lim_{t\to+\infty} \frac{1 - F(tx)}{1 - F(x)} = x^{-\gamma}\right); \tag{3}$$

- O domínio de atracção da família de distribuições Weibull define-se de forma análoga ao domíno de atracção da família de distribuições Fréchet, considerando que neste caso F tem obrigatoriamente suporte superiormente limitado: uma f.d. F pertence ao domínio de atracção da família de distribuições Weibull se o seu suporte for limitado superiormente e se a função definida em x por $1 - F(\omega(F) - x)$ for de variação regular de índice $-\gamma$. Simbolicamente,

$$\begin{aligned}
&\exists \{b_n\}_{n\in\mathbf{N}_1}, \{a_n\}_{n\in\mathbf{N}_1}, \gamma > 0: \\
&(\forall n \in \mathbf{N}_1, b_n > 0) \wedge \left(\lim_{n\to+\infty} P[X_{n:n} \leq b_n x + a_n] = \Psi_\gamma(x)\right) \\
&\Leftrightarrow \\
&(\omega(F) < +\infty) \wedge \left(\exists\gamma > 0, \forall x > 0, \lim_{t\to+\infty} \frac{1-F\left(\omega(F)-\frac{1}{tx}\right)}{1-F\left(\omega(F)-\frac{1}{x}\right)} = x^{-\gamma}\right);
\end{aligned} \tag{4}$$

- Finalmente, relativamente ao domínio de atracção da distribuição Gumbel, para que a f.d. F pertença a este domínio é necessária e suficiente a existência de uma função h que satisfaça a condição dada em (5), i.e.,

$$\begin{aligned}
&\exists \{b_n\}_{n\in\mathbf{N}_1}, \{a_n\}_{n\in\mathbf{N}_1}: \\
&(\forall n \in \mathbf{N}_1, b_n > 0) \wedge \left(\lim_{n\to+\infty} P[X_{n:n} \leq b_n x + a_n] = \Lambda(x)\right) \\
&\Leftrightarrow \exists h, \text{ função positva}: \forall x, \lim_{t\to\omega(F)} \frac{1 - F(t + xh(t))}{1 - F(t)} = \exp(-x).
\end{aligned} \tag{5}$$

Gnedenko explicitou também uma possível solução para as constantes de atracção em cada caso:

- no caso da Fréchet: $a_n = 0$ e $b_n = \inf\left\{x : 1 - F(x) \leq \frac{1}{n}\right\}$;
- no caso da Weibull: $a_n = \omega(F)$ e $b_n = \omega(F) - \inf\left\{x : 1 - F(x) \leq \frac{1}{n}\right\}$;
- e no caso da Gumbel: $a_n = \inf\left\{x : 1 - F(x) \leq \frac{1}{n}\right\}$ e $b_n = \frac{\int_{a_n}^{\omega(F)}(1-F(t))dt}{1-F(a_n)}$.

Relativamente à misteriosa função h que surge no terceiro caso, foi Haan [26] quem, em 1970, deu uma expressão explícita para esta,

$$h(t) = \frac{\int_t^{\omega(F)} (1 - F(y))\, dy}{1 - F(t)}, \tag{6}$$

definida para $\alpha(F) < t < \omega(F)$.

Passemos agora ao que se designaria por classe das auto-decomponíveis no esquema de máximos. Um discípulo de Gnedenko, Mejzler [43] estudou a classe das leis limite de máximos de v.a.'s independentes, mas não identicamente distribuídas, classe semelhante à de auto-decomponíveis L de Khintchine; em honra a Mejzler, Pestana [54] denotou-a **classe** M. Graça Martins e Pestana [22], inspirando-se no trabalho de Urbanik e no facto de que toda a f.d. F univariada é infinitamente divisível, no sentido de que, para todo o natural n, $F^{1/n}$ ser ainda uma função de distribuição, usaram a equação funcional $F(x) = F(ax)\, F_a(x)$ para definir iterativamente as classes M_1 ($\equiv M$ de Mejzler), $M_2, \ldots$, e caracterizar $M_\infty = \bigcap_{k=0}^{+\infty} M_m$. Curiosamente, são funções de distribuição cujo logaritmo é completamente monótono, i.e., uma transformada de Laplace. Tal como a classe L_∞ de Urbanik parece ser ideal para modelar somas, a classe M_∞ parece ideal para modelar máximos.

2.4 O quociente entre somas e extremos

Vimos no parágrafo anterior que o desenvolvimento da teoria relacionada com o máximo de variáveis aleatórias independentes e da teoria das somas de variáveis aleatórias independentes têm dado alguns passos paralelos. Além desta dialéctica entre teorias, alguns autores têm tentado proceder a uma comparação mais directa usando o quociente entre $M_n = \max\limits_{1\leq i\leq n} X_i$ e $S_n = \sum\limits_{i=1}^{n} X_i$, revelando desta forma, a susceptibilidade da soma de ser influenciada pela parcela máxima. Entre estes autores encontramos, por exemplo, Darling, O'Brien, Maller, Resnick, Teugels e Bingham.

Antes de continuarmos, convém recordar que uma função f, positiva, cujo domínio inclui um intervalo do tipo $(a, +\infty)$ diz-se de **variação lenta** se

$$\forall\lambda > 0, \quad \frac{f(\lambda x)}{f(x)} \underset{x\to+\infty}{\longrightarrow} 1. \tag{7}$$

Um conceito fortemente associado à variação lenta é o de variação regular: dizemos que uma função f, positiva, cujo domínio inclui um intervalo do tipo $(a, +\infty)$ é de **variação regular** de índice ρ no infinito se

$$\forall\lambda > 0, \quad \frac{f(\lambda x)}{f(x)} \underset{x\to+\infty}{\longrightarrow} \lambda^\rho. \tag{8}$$

Uma vez mais, considere-se $\{X_i\}_{i\in\mathbf{N}_1}$ uma sucessão de variáveis aleatórias independentes, identicamente distribuídas a uma variável aleatória X, F a respectiva função de distribuição e $\omega(F)$ o supremo do suporte de X, i.e., $\omega(F) = sup\{x \in \mathbf{R} : F(x) < 1\}$. Suponhamos que $X \geq 0$. Bingham *et al.* [4] coleccionaram alguns resultados relativos a este assunto que passamos a enunciar:

- $M_n/S_n \xrightarrow{p} 0$ é equivalente à variação lenta de $\int_0^x y dF(y)$;
- $M_n/S_n \xrightarrow{p} 1$ é equivalente à variação lenta de $1 - F$;
- A convergência de M_n/S_n para um limite não degenerado é equivalente a F ser atraída para uma lei estável de índice $\alpha \in (0, 1)$, o que por sua vez é equivalente a $E(S_n/M_n)$ ter limite finito superior a um;
- Se X tem valor médio finito μ então: a convergência de $(S_n - n\mu)/M_n$ para um limite não degenerado é equivalente a F ser atraída para uma lei estável de índice $\alpha \in (1, 2)$, o que, por sua vez, é equivalente a que $E((S_n - n\mu)/M_n) \to c \in (1, +\infty)$, e, neste caso, $\alpha = (1 + c)/c$.

3 O presente — Novos desenvolvimentos

There are lies, damned lies and statistics.

Mark Twain

De facto, se olharmos com atenção para a Estatística, verificamos que a sua grande missão é a construção de modelos *adequados* à realidade, através da leitura/interpretação dos sinais que esta nos oferece. Mas um modelo — a base de toda a ciência — será sempre um modelo: uma mentira.

Um modelo pode despertar a nossa admiração pelo seu detalhe e precisão - um exemplo notável é o passeio aleatório, notável também pelas muitas generalizações a que se presta (nascimentos e morte, difusão, etc.) Pode ferir a nossa imaginação por, através de caracterizações locais, pode ser globalizado com hipóteses de *estacionaridade* - como é o caso do processo de Poisson, ou o estudo das probabilidades de absorção em barreiras no passeio aleatório ou outras cadeias de Markov. Pode ser útil, por substituir cálculos deprimentes.

Nestes próximos parágrafos vamos centrar a nossa atenção em modelos de somas generalizadas, médias ponderadas, médias-p. E desde já notamos que qualquer estatística ordinal $X_{k:n}$ pode ser considerada uma média ponderada, e que $(\sum_{i=1}^{n}|X_i|^p)^{1/p} \underset{p\to+\infty}{\longrightarrow} X_{n:n}$. Estas médias generalizadas podem ser encaradas como *sinais* de localização, e diferenças entre estatísticas ordinais são *leituras* de escala, que nos convêm por vezes comparar (por exemplo $\frac{X_{n:n}-X_{2:n}}{X_{n-1:n}-X_{1:n}}$) a fim de avaliar pesos de caudas, uma das características mais relevantes e determinantes. Usaremos conceitos de variação regular de Karamata (cf. Bingham *et al.* [4]) para precisar a noção de peso de cauda, e atendendo às ideias de Resnick sobre *tail equivalence* (cf. Resnick [60]) damos importância de relevo ao modelo de Pareto generalizado, que sabemos também serem os elementos estáveis na abordagem

POT (*Peaks Over Thresholds*).

3.1 Auto-normalização

> *Limit theorems are useless in statistical contexts if they depend on parameters which must be computed from unknown distributions. A way around this is to replace such parameters by functions of the observations.*
>
> Sidney Resnick

Resnick [62] justificou assim, muito cristalinamente, a necessidade do estudo da auto-normalização.

Existem já alguns resultados respeitantes a este assunto, essencialmente no que se refere à soma de variáveis aleatórias i.i.d.. Inspirados por Logan *et al.* [41] que estudaram a estatística

$$\sum_{i=1}^{n} X_i \bigg/ \left(\sum_{i=1}^{n} |X_i|^p \right)^{1/p}$$

($p > 0$) e usando alguns dos resultados que Darling [7] apresenta no seu estudo relativamente ao caso limite $\sum_{i=1}^{n} X_i \Big/ \max_{1 \leq i \leq n} X_i$, vamos tentar encontrar o paralelo na teoria dos extremos considerando

$$W_n(p) = \max_{1 \leq i \leq n} X_i \bigg/ \left(\sum_{i=1}^{n} |X_i|^p \right)^{1/p} \text{ e } W_n(+\infty) = \max_{1 \leq i \leq n} X_i \Big/ \max_{1 \leq i \leq n} |X_i| .$$

Note-se desde já que ambas as estatísticas têm suporte contido no intervalo $[-1,1]$.

Se $X \geq 0$, então $W_n(+\infty) = 1$, facto que retira o interesse desta estatística para variáveis não negativas. Algo análogo se passa no comportamento assintótico de $W_n(+\infty)$ quando o suporte da variável X (supp$X$) é limitado e onde obtemos que

$$W_n(+\infty) \underset{x\to+\infty}{\overset{d}{\longrightarrow}} \frac{\sup \operatorname{supp} X}{\max\left[\left|\inf \operatorname{supp} X\right|, \left|\sup \operatorname{supp} X\right|\right]}.$$

No entanto, $W_n(+\infty)$ revelou-se potencialmente útil no estudo da (as)simetria da variável X.

Proposição 1

Seja X uma v.a. absolutamente contínua. Temos então que

(a) $F_{W_n(+\infty)}(x) = n \int\limits_{-\infty}^{0} \left[F(-xt) - F(t)\right]^{n-1} f(t)\, dt \mathbf{I}_{[-1,1)}(x) + \mathbf{I}_{[1,+\infty)}(x).$

(b) Se $\omega(F) > 0$ então $\mathbf{P}\left[W_n(+\infty) = 1\right] \neq 0$.

(c) Se X é simétrica então $\mathbf{P}\left[W_n(+\infty) = 1\right] = \frac{1}{2}$.

Demonstração

(a) Comecemos por provar a primeira afirmação. Consideremos $x \in [-1, 1)$,

$$F_{W_n(+\infty)}(x) = \mathbf{P}\left[\frac{\max\limits_{1\le i\le n} X_i}{\max\limits_{1\le i\le n}|X_i|} \le x\right]$$

$$= \sum_{k=1}^{n} \mathbf{P}\left[\frac{\max\limits_{1\le i\le n} X_i}{\max\limits_{1\le i\le n}|X_i|} \le x,\ \max_{1\le i\le n}|X_i| = |X_k|, X_k > 0\right]$$
$$+ \sum_{k=1}^{n} \mathbf{P}\left[\frac{\max\limits_{1\le i\le n} X_i}{\max\limits_{1\le i\le n}|X_i|} \le x,\ \max_{1\le i\le n}|X_i| = |X_k|, X_k \le 0\right]$$

$$= \sum_{k=1}^{n} \mathbf{P}\left[\max_{1\le i\le n} X_i \le xX_k,\ \max_{1\le i\le n}|X_i| = X_k, X_k > 0\right]$$
$$+ \sum_{k=1}^{n} \mathbf{P}\left[\max_{1\le i\le n} X_i \le -xX_k,\ \max_{1\le i\le n}|X_i| = -X_k, X_k \le 0\right]$$

$$= \sum_{k=1}^{n} \mathbf{P}\left[\max_{1\le i\le n} X_i \le -xX_k,\ \max_{1\le i\le n}|X_i| = -X_k, X_k \le 0\right],$$

pois

$$\mathbf{P}\left[\max_{1\le i\le n} X_i \le xX_k,\ \max_{1\le i\le n}|X_i| = X_k, X_k > 0\right] \le \mathbf{P}\left[X_k \le xX_k\right] = 0$$

(recorde-se que $x \in [-1, 1)$). Temos assim que

$$F_{W_n(+\infty)}(x) = \sum_{k=1}^{n} \mathbf{P}\left[\max_{1\le i\le n} X_i \le -xX_k,\ \max_{1\le i\le n}|X_i| = -X_k, X_k \le 0\right]$$

$$= \sum_{k=1}^{n} \mathbf{P}\left[X_1 \le -xX_k, \ldots (k) \ldots, X_n \le -xX_k,\right.$$
$$\left.|X_1| \le -X_k, \ldots (k) \ldots, |X_n| \le -X_k, X_k \le 0\right].$$

("(k)" indica a ausência do termo de ordem k). Como as variáveis são i.i.d. e $X_k \leq -xX_k$, para $x \in (-1,0]$ e $-xX_k \leq -X_k$, para $x \in [-1,1)$ resulta que

$$
\begin{aligned}
F_{W_n(+\infty)}(x) &= n \int_{(-\infty,0]} P^{n-1}\left[t \leq X_1 \leq -xt\right] dF(t) \\
&= n \int_{(-\infty,0]} \left[F(-xt) - F\left(t^-\right)\right]^{n-1} dF(t) \\
&= n \int_{-\infty}^{0} \left[F(-xt) - F(t)\right]^{n-1} f(t)\, dt.
\end{aligned}
$$

(b) Se $x = 1$, sabemos já que $F_{W_n(+\infty)}(x) = 1$ donde resulta que

$$
\begin{aligned}
\mathbf{P}[W_n(+\infty) = 1] &= F_{W_n(+\infty)}(1) - F_{W_n(+\infty)}\left(1^-\right) \\
&= 1 - n \int_{-\infty}^{0} \left[F(-t) - F(t)\right]^{n-1} f(t)\, dt \\
&= 1 + \int_{-\infty}^{0} n\left[F(-t) - F(t)\right]^{n-1}\left[-f(-t) - f(t)\right] dt \\
&\quad + \int_{-\infty}^{0} n\left[F(-t) - F(t)\right]^{n-1} f(-t)\, dt \\
&= \int_{-\infty}^{0} n\left[F(-t) - F(t)\right]^{n-1} f(-t)\, dt \\
&= \int_{0}^{+\infty} n\left[F(t) - F(-t)\right]^{n-1} f(t)\, dt,
\end{aligned}
$$

o que implica que

$$
\mathbf{P}\left[W_n(+\infty) = 1\right] = 0 \Leftrightarrow \left[F(t) - F(-t)\right]^{n-1} f(t) = 0,
$$

quase certamente em $(0,+\infty)$. Ora, se o suporte da variável X for um subconjunto de $\mathbf{R}^-$ a condição é obviamente verificada. Caso contrário, teremos sempre $P\left[W_n(+\infty)=1\right]>0$ (note-se que estamos a trabalhar com variáveis honestas, i.e., com suporte contido em $\mathbf{R}$).

(c) Relativamente à última afirmação basta observar que

$$\begin{aligned} F_{W_n(+\infty)}\left(1^-\right) &= n\int_{-\infty}^{0}\left[F(-t)-F\left(t^-\right)\right]^{n-1}dF(t) \\ &= n\int_{-\infty}^{0}\left[1-2F(t)\right]^{n-1}dF(t)=\frac{1}{2}. \end{aligned}$$

■

Observe-se que no caso discreto esta asserção não é em geral verdadeira. Como contra-exemplo basta considerar uma variável aleatória, X, suportada pelos inteiros, tal que

$$p_k=\mathbf{P}\left[X=k\right]=\begin{cases}\frac{2^{-k}}{3}, & \text{se } k\geq 1\\ \frac{2^{k}}{3}, & \text{se } k\leq 0\end{cases}.$$

Não é difícil mostrar que

$$F_{W_n(+\infty)}\left(1^-\right)=\frac{n}{3}\sum_{m=0}^{n-1}\binom{n-1}{m}\frac{(-1)^{n-1-m}\,2^{-2(n-m)}}{1-2^{-n+m}}+\frac{n}{3^n}+\frac{n}{3}\frac{1}{2^n}.$$

Exemplo 1

As caudas paretianas são caudas de variação regular. O índice de variação regular destas caudas percorre todos os índices possíveis (para funções *de distribuição). Dado que estamos a considerar apenas distribuições com suporte ilimitado superiormente e inferiormente, vamos considerar uma generalização da variável cuja densidade tem a seguinte forma:*

$$g(x) = \beta\alpha^{\beta}x^{-\beta-1}\mathbf{I}_{[\alpha,+\infty)}(x).$$

O coeficiente de variação regular desta variável é $-\beta$ pois

$$\lim_{x\longrightarrow+\infty}\frac{1-F(\lambda x)}{1-F(x)} = \lim_{x\longrightarrow+\infty}\frac{\lambda f(\lambda x)}{f(x)} = \lambda^{-\beta} \quad (\lambda > 0).$$

Neste exemplo vamos considerar a densidade

$$f(x) = \frac{1}{2}\beta_2\alpha_2^{\beta_2}(-x)^{-\beta_2-1}\mathbf{I}_{(-\infty,-\alpha_2]}(x) + \frac{1}{2}\beta_1\alpha_1^{\beta_1}x^{-\beta_1-1}\mathbf{I}_{[\alpha_1,+\infty)}(x)$$

e correspondente função de distribuição

$$F(x) = \tfrac{1}{2}\left(\tfrac{\alpha_2}{-x}\right)^{\beta_2}\mathbf{I}_{(-\infty,-\alpha_2]}(x) + \tfrac{1}{2}\mathbf{I}_{(-\alpha_2,\alpha_1)}(x) + \left[1-\tfrac{1}{2}\left(\tfrac{\alpha_1}{x}\right)^{\beta_1}\right]\mathbf{I}_{[\alpha_1,+\infty)}(x).$$

Por simplicidade, e tendo em conta que os parâmetros α_1 e α_2 não interferem com o coeficiente de variação regular, tomaremos $\alpha_1 = \alpha_2 = \alpha$. No intervalo $(-\alpha,\alpha)$ o integral

$$n\int_{-\infty}^{0}[F(-xt)-F(t)]^{n-1}f(t)\,dt$$

é nulo. Quando $-1 < x \leq 0$, obtemos

$$F_{Wn(+\infty)}(x) = \frac{\left[1-(-x)^{\beta_2}\right]^{n-1}}{2^n}$$

que converge para zero quando n cresce para infinito; para $0 < x < 1$

$$F_{Wn(+\infty)}(x) = \frac{n\beta_2}{2}\int_0^x\left[1-\frac{1}{2}\left(\frac{y}{x}\right)^{\beta_1}-\frac{1}{2}y^{\beta_2}\right]^{n-1}y^{\beta_2-1}dy + \frac{\left[1-x^{\beta_2}\right]^n}{2^n}.$$

Se $\beta_1 = \beta_2 = \beta$ *então*

$$\begin{aligned} F_{Wn(+\infty)}(x) &= \frac{n\beta}{2}\int_0^x \left[1 - y^\beta\left(\frac{x^{-\beta}+1}{2}\right)\right]^{n-1} y^{\beta-1}dy + \frac{\left(1-x^\beta\right)^n}{2^n} \\ &= \frac{1}{1+x^\beta}\left[x^\beta + \left(\frac{1-x^\beta}{2}\right)^n\right] \end{aligned}$$

que converge para $\dfrac{x^\beta}{1+x^\beta}$, *quando* n *cresce para infinito.*

Relativamente a $W_n(p)$:

Proposição 2

Suponhamos que $\alpha(F_X) \geq 0$ *e que* $\omega(F_X) = +\infty$. *Se, para* $a > 0$,

$$\lim_{x\to+\infty}\frac{1-F_X(ax)}{1-F_X(x)} = 1$$

então

$$\lim_{n\to+\infty} F_{W_n(p)}(x) = F_1(x).$$

Demonstração

Se as variáveis são não negativas

$$F_{W_n(p)}(x) = P\left[\frac{\max\limits_{1\leq i\leq n} X_i}{\left(\sum\limits_{i=1}^n |X_i|^p\right)^{1/p}} \leq x\right] = 1 - P\left[\frac{\sum\limits_{i=1}^n Y_i}{\max\limits_{1\leq i\leq n} Y_i} \leq y\right]$$

onde $Y_i = X_i^p$ são i.i.d. e $y = \frac{1}{x^p} > 0$. Usando os resultados de Darling [7], podemos afirmar que, quando as variáveis da sucessão $\{X_n\}$ são de variação lenta no infinito,

$$\lim_{n \to +\infty} F_{W_n(p)}(x) = F_1(x),$$

i.e., $W_n(p)$ converge fracamente para a constante 1. ∎

De novo, do trabalho de Darling [7], é possível concluir que

Proposição 3

Se $X_i \geq 0$ e se $\sum_{i=0}^n X_i^p$, devidamente normalizada tiver distribuição limite estável com expoente característico $0 < \alpha < 1$, então

$$\lim_{n \to +\infty} F_{W_n(p)}(x) = 1 - G\left(\frac{1}{x^p}\right),$$

onde G é uma função tal que

$$\int_0^{+\infty} \exp(ity)\, dG(y) = \frac{\exp(it)}{1 - \alpha \int_0^1 \frac{(\exp(itu-1))}{u^{\alpha+1}} du}.$$

Ainda sobre as propriedades de $W_n(p)$, observe-se que

$$\frac{\max\limits_{1 \leq i \leq n} X_i}{\left(\sum\limits_{i=1}^n |X_i|^p\right)^{1/p}} \leq -\frac{1}{n^{1/p}} \Leftrightarrow -n^{1/p} \max_{1 \leq i \leq n} X_i \geq \left(\sum_{i=1}^n |X_i|^p\right)^{1/p}$$

$$\Leftrightarrow n\left(-\max_{1 \leq i \leq n} X_i\right)^p \geq \sum_{i=1}^n |X_i|^p \qquad (\text{tomando } \max_{1 \leq i \leq n} X_i = X_{i_0})$$

$$\Leftrightarrow (n-1)|X_{i_0}|^p \geq \sum_{\substack{i=1 \\ i \neq i_0}}^n |X_i|^p \geq (n-1)|X_{i_0}|^p$$

que tem probabilidade nula. Conclui-se assim a

Proposição 4

Se $x \in \left[-1, -1/n^{1/p}\right]$, *então* $P\left[W_n\left(p\right) \leq x\right] = 0$.

3.2 Domínios de atracção não clássicos

Quando nos afastamos dos requisitos da independência e/ou da identidade distribucional é natural que surjam novas distribuições limite. De facto, a menos que se imponham algumas restrições, este novo problema chega mesmo a perder o interesse, pois existem sucessões de v.a.'s, não independentes e/ou não idênticas em distribuição, para as quais qualquer distribuição pode ser resultado do limite em distribuição.

Infelizmente a situação de dependência é uma situação muito vaga, e só em casos muitos particulares se conhecem resultados de relevo, como, e.g., no casos das sucessões de variáveis aleatórias permutáveis e das sucessões de variáveis aleatórias estacionárias, (cf. referências de Balakrishnan *et al.* [3] e de Galambos [14]).

Quando se mantém a independência, e se relaxa a condição de identidade distribucional, os resultados surgem com mais facilidade; devemos o seguinte resultado a Mejzler (cf. Galambos [14]):

Teorema 2

Seja $\{F_i\}_{i\in\mathbf{N}_1}$ *uma sucessão de funções de distribuição para a qual existem sucessões de números reais* $\{c_i\}_{i\in\mathbf{N}_1}$ *e* $\{d_i\}_{i\in\mathbf{N}_1}$ *tais que*

$\forall i \in \mathbf{N}_1,\ d_i > 0,$

$$\lim_{n\to+\infty} \max\{1 - F_i(c_n + d_n x) : 1 \le i \le n\} = 0 \tag{9}$$

e, para todo o número $0 < t \le 1$,

$$\lim_{n\to+\infty} \sum_{i=1}^{[nt]} [1 - F_i(c_n + d_n x)] = w(t, x)$$

existe, e é finito, para todo $0 < t \le 1$, *sempre que for finito para* $t = 1$. *Quando esta condição é satisfeita dizemos que as sucessões referidas satisfazem a* ***condição de uniformidade para o máximo***. *Nesta situação, uma função de distribuição não degenerada* H *é a distribuição limite de* $\left(\max\limits_{1\le i\le n} X_i - a_n\right)/b_n$ *se, e só se, uma das seguintes condições é verificada:*

- $\log H(x)$ *é uma função côncava;*
- $\omega(H) < +\infty$ *e a função definida por* $\log[\omega(H) - \exp(-x)]$ *é côncava, para* $x > 0$;
- $\alpha(H)$ *é finito e a função definida por* $\log H[\alpha(H) + \exp(x)]$, $x > 0$, *é côncava.*

Voltaremos a falar deste resultado mais adiante, no Parágrafo 3.4.

3.2.1 Normalização linear não clássica – algumas considerações

Como habitualmente, consideremos $\{X_i\}_{i \in \mathbf{N}_1}$ uma sucessão de variáveis aleatórias independentes (X_i com f.d. F_i). Seja

$$Y_n = \max_{1 \leq i \leq n} \frac{X_i - a_{in}}{b_{in}},$$

onde a_{in} e $b_{in} > 0$, $n \in \mathbf{N}_1$, $1 \leq i \leq n$, são constantes reais. Note-se que a normalização apresentada, poderá ser diferente para cada variável aleatória. O nosso objectivo é, neste parágrafo, tentar conhecer e caracterizar o conjunto das distribuições limite, não degeneradas, de Y_n.

Comecemos por observar que, tal como no caso em que a normalização é feita de igual modo para todos os elementos de um vector aleatório de dimensão n (i.e., em que a normalização só depende do valor n), são necessárias algumas restrições - pois, caso contrário, a classe obtida é formada por todas as distribuições[2] - temos, também aqui, de proceder a algumas restrições. Na situação clássica (para máximos) a única restrição - que origina a conhecida classe de Mejzler - é a condição de uniformidade para o máximo dada por (9). Observemos que:

1. Na sucessão de funções de distribuição $\{F_i\}_{i \in \mathbf{N}_1}$ não devem existir

[2] Galambos [14] apresenta um exemplo ilustrativo deste facto: tome-se F uma função de distribuição arbitrária e $\{p_i\}_{i \in \mathbf{N}_1}$ a função massa de probabilidade de uma variável discreta, suportada pelos naturais positivos, tal que $p_i \neq 0$, $\forall i$. Considere-se uma sucessão de variáveis aleatórias independentes $\{X_i\}_{i \in \mathbf{N}}$ onde cada variável X_i tem distribuição F^{p_i}. Temos que

$$P\left[\max_{1 \leq i \leq n} X_i \leq x\right] = \prod_{i=1}^{n} F^{p_i}(x) = F^{p_1 + \ldots + p_n}(x) \longrightarrow F(x),\ n \longrightarrow +\infty.$$

distribuições que surjam na sucessão um número finito de vezes, caso contrário o estudo desta classe perde interesse.

2. O caso mais simples será quando $F_i = F$, $\forall i$, e, neste caso, quando se toma uma normalização constante para o parâmetro i, cai-se no problema clássico. É necessário pois tomar uma normalização diferente para cada variável X_i, $\frac{X_i - a_{in}}{b_{in}}$.

3. Atendendo ao primeiro ponto todas as funções de distribuição que surgem na sucessão $\{F_i\}_{i\in\mathbf{N}_1}$, fazem-no um número infinito (numerável) de vezes. Temos duas situações possíveis:

   (a) O número de distribuições distintas que surge em $\{F_i\}_{i\in\mathbf{N}_1}$ é finito. Suponhamos que temos k distribuições distintas e que para distribuições iguais temos a mesma normalização (se optarmos por normalizações distintas caímos no ponto 2). Neste caso

$$\mathbf{P}\left[\max_{1\le i\le n} \frac{X_i - a_{in}}{b_{in}} \le x\right] = \prod_{i=1}^{k} \left[F_i\left(a_{in} + b_{in}x\right)\right]^{c_i(n)}$$

   onde $c_1(n) + \ldots + c_k(n) = n$ e $c_i(n)$ crescem para $+\infty$, quando n cresce para $+\infty$. Note-se que, dado que $F_i\left(a_{in} + b_{in}x\right) \in [0, 1]$, para que o limite seja não nulo é necessário que, para todo i, $F_i\left(a_{in} + b_{in}x\right)$ convirja para 1, quando n converge para $+\infty$. Temos então um produto de factores do tipo $F^{c(n)}\left(a_n + b_n x\right)$. O comportamento assintótico de cada um destes factores é o tema de estudo dos domínios de atracção parcial, que abordaremos no parágrafo seguinte. Temos, no entanto, uma sucessão expoente particular pois $0 \leq c(n) \leq 1$.

Quando F_i pertence ao domínio de atracção de uma max-estável é necessário, de modo a que F_i tenha alguma influência no resultado final, que $\frac{c_i(n)}{n} \to 0$. Neste caso, as constantes de normalização são já conhecidas e o resultado final, supondo que todas as funções de distribuição estão no domínio de alguma max-estável, depois de aplicar o limite, é o produto de funções de extremos que são elementos da classe de Mejzler, que sabemos ser um elemento de Mejzler.

(b) O número de distribuições distintas que surgem em $\{F_i\}_{i \in \mathbf{N}_1}$ é infinito numerável. Neste caso, e procedendo como no modo anterior, obteremos como resultado também um elemento da classe de Mejzler, pois a soma infinita numerável de funções convexas, se existir, é ainda uma função convexa.

Concluímos assim que para obter elementos externos à classe de Mejzler é necessário tomar normalizações distintas para uma mesma distribuição ou considerar domínios de atracção parcial.

Uma outra generalização possível é obtida com a aleatorização do número de variáveis de cada tipo. Por exemplo, se consideramos que numa sucessão de v.a.'s independentes, $\{Y_i\}_{i \in \mathbf{N}_1}$, temos apenas dois tipos de v.a.'s, X_1 e X_2, com funções de distribuição F_1 e F_2, respectivamente, e que o número de cópias de X_1 no vector aleatório $(Y_1, \ldots, Y_n)$ é dado por

N, uma binomial de parâmetros $n \in \mathbf{N}_1$ e $p \in (0,1)$, temos que

$$\begin{aligned}\mathbf{P}\left[\max_{1\leq k\leq n} Y_i \leq x\right] &= \sum_{m=0}^{n} \mathbf{P}\left[\max_{1\leq k\leq n} Y_i \leq x \middle| N=m\right] \mathbf{P}[N=m] \\ &= \sum_{m=0}^{n} F_1^m(x) F_2^{n-m}(x) \binom{n}{m} p^m (1-p)^{n-m} \\ &= [pF_1(x) + (1-p) F_2(x)]^n . \end{aligned} \quad (10)$$

É certo que é uma situação que não traz novidades relativamente a distribuições limite, mas sugere um problema interessante. Facilmente se observa que se $\omega(F_1) > \omega(F_2)$ ou se $\omega(F_1) < \omega(F_2)$, então a combinação linear convexa de F_1 e de F_2 pertence ao mesmo domínio de atracção que a distribuição cujo suporte tem limite superior. No caso $\omega(F_1) = \omega(F_2)$, é o peso das caudas que determina qual a distribuição que será dominante na combinação convexa.

O *peso das caudas de uma f.d.* é um conceito talvez pouco preciso, dada a não unicidade da definição, mas importantíssimo pois é este que determina os comportamentos assintóticos, como os resultados limite e as velocidades de convergência (cf. Gomes e Pestana [20] e Gomes e Pestana [21] sobre comportamentos *penultimate*). Também na análise exploratória de localização e escala é o peso das caudas que determina quais os melhores estimadores, a nível de robustez e de resistência (cf. Mendonça [44]).

A definição mais apropriada será talvez o conceito de variação regular de Karamata (cf. Bingham *et al.* [4]), que por várias vezes foi referido ao longo dos parágrafos anteriores (na caracterização dos domínios da atracção das estáveis e das max-estáveis). Presentemente a variação lenta e a variação regular são áreas de investigação intensa e de aplicabilidade muito vasta, (cf. as monografias de Seneta [69] e de Bingham *et al.* [4]).

Voltemos à expressão (10) e suponhamos que F_1 e F_2 pertencem ao domínio de atracção de uma f.d. max-estável: Fréchet (Φ), Weibull (Ψ) ou Gumbel (Λ). Designando por F a combinação $pF_1 + (1-p)\,F_2$, denotando por $F \in \Gamma$ o facto de F pertencer ao domínio de atracção de uma max-estável Γ, e excluindo as situações triviais onde $\omega(F_1) \neq \omega(F_2)$, temos o seguinte lema:

Lema 1

1. *Se* $F_1 \in \Phi_{\alpha_1}$ *e* $F_2 \in \Phi_{\alpha_2}$ *então* $F \in \Phi_{\min(\alpha_1,\alpha_2)}$.
2. *Se* $F_1 \in \Psi_{\alpha_1}$ *e* $F_2 \in \Psi_{\alpha_2}$ *então* $F \in \Psi_{\min(\alpha_1,\alpha_2)}$.
3. *Se* $F_1 \in \Phi_{\alpha}$ *e* $F_2 \in \Lambda$ *então* $F \in \Phi_{\alpha}$.
4. *Se* $F_1 \in \Psi_{\alpha}$ *e* $F_2 \in \Lambda$ *então* $F \in \Psi_{\alpha}$.
5. *Se* $F_1, F_2 \in \Lambda$ *e se existir o* $\lim\limits_{x\to\omega(F)} \frac{1-F_1(x)}{1-F_2(x)}$, *finito ou não, então* $F_2 \in \Lambda$.

Demonstração

1. Se $F_1 \in \Phi_{\alpha_1}, F_2 \in \Phi_{\alpha_2}$, então $1-F_1$ e $1-F_2$ são de variação regular no infinito, com índices $-\alpha_1$ e $-\alpha_2$, respectivamente (cf. (3)). Existem

então funções de variação lenta (cf. (7)) L_1 e L_2 tais que, para todo número real x, $1 - F_i(x) = x^{-\alpha_i} L_i(x)$, $i = 1, 2$, donde resulta que

$$\begin{aligned}
1 - F(x) &= 1 - [pF_1(x) + (1-p) F_2(x)] \\
&= p - pF_1(x) + 1 - p - (1-p) F_2(x) \\
&= p(1 - F_1(x)) + (1-p)(1 - F_2(x)) \\
&= px^{-\alpha_1} L_1(x) + (1-p) x^{-\alpha_2} L_2(x),
\end{aligned}$$

i.e., F escreve-se como a soma de duas funções de variação regular, uma de índice $-\alpha_1$ outra de índice $-\alpha_2$, resultando assim que $1 - F$ é uma função de variação regular de índice $-\min(\alpha_1, \alpha_2)$ (cf. Bingham *et al.* [4]), e, consequentemente, que $F \in \Phi \min(\alpha_1, \alpha_2)$.

2. Análoga à anterior.

3. Se $F_1 \in \Phi_\alpha, F_2 \in \Lambda$ é esperado que $F \in \Phi_\alpha$. De facto, dado que (designando por L a função de variação lenta tal que $1 - F_1(x) = x^{-\alpha} L(x)$)

$$\begin{aligned}
1 - F(x) &= px^{-\alpha} L(x) + (1-p)(1 - F_2(x)) \\
&= x^{-\alpha} [pL(x) + (1-p) x^{\alpha} (1 - F_2(x))],
\end{aligned}$$

basta provar que $pL(x) + (1-p) x^{\alpha} (1 - F_2(x))$ é uma função de variação lenta. Seja $\lambda > 0$; temos que,

$$\lim_{x \to +\infty} \frac{pL(\lambda x) + (1-p) \lambda^{\alpha} x^{\alpha} (1 - F_2(\lambda x))}{pL(x) + (1-p) x^{\alpha} (1 - F_2(x))}$$

$$= \lim_{x \to +\infty} \frac{\frac{L(\lambda x)}{L(x)} + \frac{(1-p)(1-F_2(\lambda x))}{p(\lambda x)^{-\alpha} L(\lambda x)} \frac{L(\lambda x)}{L(x)}}{1 + \frac{(1-p)(1-F_2(x))}{px^{-\alpha} L(x)}}$$

$$= \lim_{x\to+\infty} \frac{\frac{L(\lambda x)}{L(x)} + \frac{(1-p)}{p}\frac{1-F_2(\lambda x)}{1-F_1(\lambda x)}\frac{L(\lambda x)}{L(x)}}{1+\frac{(1-p)}{p}\frac{1-F_2(x)}{1-F_1(x)}}.$$

Tendo em conta que, por pertencer ao domínio de atracção da Gumbel e dado que $\omega(F_2) = +\infty$, $1-F_2$ é de variação rápida (cf. Embrechts [10]), donde resulta que existem funções c e δ tais que $c(x) \to c \in (0,+\infty)$, $\delta(x) \to -\infty$ quando $x \to +\infty$, e existe $z_1 > 0$ tal que

$$1 - F_2(x) = c(x) \exp\left[\int_{z_1}^{x} \frac{\delta(t)}{t} dt\right], \ z_1 < x;$$

por outro lado, $1 - F_1$ é de variação regular de índice $-\alpha$, resultando assim que

$$1 - F_1(x) = d(x) \exp\left[-\int_{z_2}^{x} \frac{h(t)}{t} dt\right], \ z_2 < x$$

para algum z_2, $d(x) \to d > 0$ e $h(x) \to -\alpha$. Destas duas representações, temos que, para $z = \max(z_1, z_2, 1)$,

$$\frac{1-F_2(x)}{1-F_1(x)} = \frac{c(x)}{d(x)} \exp\left[-\int_{z}^{x} \frac{h(t)}{t} - \frac{\delta(t)}{t} dt\right] \underset{x\to+\infty}{\longrightarrow} 0$$

e, consequentemente,

$$\lim_{x\to+\infty} \frac{pL(\lambda x) + (1-p)\lambda^\alpha x^\alpha (1-F_2(\lambda x))}{pL(x) + (1-p) x^\alpha (1-F_2(x))} = 1.$$

4. Análoga à anterior.

5. Suponhamos que existe o limite $a = \lim_{x\to\omega(F)} \frac{1-F_2(x)}{1-F_1(x)}$. Para que F pertença ao domínio da distribuição Gumbel é necessário e suficiente que (cf. Haan [26]) o integral duplo

$$\int_{x}^{\omega(F)} \int_{y}^{\omega(F)} (1 - F(t))\, dtdy$$

seja finito e que

$$\lim_{x\to\omega(F)} \frac{(1-F(x))\int_x^{\omega(F)}\int_y^{\omega(F)}(1-F(t))\,dtdy}{\left[\int_x^{\omega(F)}(1-F(t))\,dt\right]^2} = 1.$$

Ora, tomando F_1 e F_2 pertencentes ao domínio da distribuição Gumbel, com $\omega(F_1) = \omega(F_2)$, temos que

$$\int_x^{\omega(F)}\int_y^{\omega(F)}(1-F(t))\,dtdy$$
$$= p\int_x^{\omega(F)}\int_y^{\omega(F)}(1-F_1(t))\,dtdy + (1-p)\int_x^{\omega(F)}\int_y^{\omega(F)}(1-F_2(t))\,dtdy$$

é finito e que, tomando por comodidade $p = \frac{m_1}{m}$ e $m = m_1 + m_2$,

$$\frac{(1-F(x))\int_x^{\omega(F)}\int_y^{\omega(F)}(1-F(t))\,dtdy}{\left[\int_x^{\omega(F)}(1-F(t))\,dt\right]^2}$$

$$= \frac{(p(1-F_1(x))+(1-p)(1-F_2(x)))\int_x^{\omega(F)}\int_y^{\omega(F)}(p(1-F_1(t))+(1-p)(1-F_2(t)))dtdy}{\left[\int_x^{\omega(F)}(p(1-F_1(t))+(1-p)(1-F_2(t)))dt\right]^2}$$

$$= \frac{m_1^2(1-F_1(x))\int_x^{\omega(F)}\int_y^{\omega(F)}(1-F_1(t))\,dtdy}{\left[m_1\int_x^{\omega(F)}(1-F_1(t))\,dt + m_2\int_x^{\omega(F)}(1-F_2(t))\,dt\right]^2} +$$
$$\frac{m_1m_2(1-F_1(x))\int_x^{\omega(F)}\int_y^{\omega(F)}(1-F_2(t))\,dtdy}{\left[m_1\int_x^{\omega(F)}(1-F_1(t))\,dt + m_2\int_x^{\omega(F)}(1-F_2(t))\,dt\right]^2} +$$

$$\frac{m_1 m_2 (1-F_2(x)) \int_x^{\omega(F)} \int_y^{\omega(F)} (1-F_1(t))\, dtdy}{\left[m_1 \int_x^{\omega(F)} (1-F_1(t))\, dt + m_2 \int_x^{\omega(F)} (1-F_2(t))\, dt\right]^2} + \frac{m_2^2 (1-F_2(x)) \int_x^{\omega(F)} \int_y^{\omega(F)} (1-F_2(t))\, dtdy}{\left[m_1 \int_x^{\omega(F)} (1-F_1(t))\, dt + m_2 \int_x^{\omega(F)} (1-F_2(t))\, dt\right]^2}$$

$$= \frac{m_1 \frac{[m_1(1-F_1(x))+m_2(1-F_2(x))]}{(1-F_1(x))} (1-F_1(x)) \frac{\int_x^{\omega(F)} \int_y^{\omega(F)} (1-F_1(t))dtdy}{\left[\int_x^{\omega(F)} (1-F_1(t))dt\right]^2}}{\left[m_1 + m_2 \frac{\int_x^{\omega(F)}(1-F_2(t))dt}{\int_x^{\omega(F)}(1-F_1(t))dt}\right]^2} + \frac{\frac{m_2[m_1(1-F_1(x))+m_2(1-F_2(x))]}{1-F_2(x)} (1-F_2(x)) \frac{\int_x^{\omega(F)} \int_y^{\omega(F)} (1-F_2(t))dtdy}{\left[\int_x^{\omega(F)} (1-F_2(t))dt\right]^2}}{\left[m_1 \frac{\int_x^{\omega(F)}(1-F_1(t))dt}{\int_x^{\omega(F)}(1-F_2(t))dt} + m_2\right]^2}$$

$$\underset{x\to\omega(F)}{\approx} \frac{m_1\left[m_1 + m_2 \frac{1-F_2(x)}{1-F_1(x)}\right]}{\left[m_1 + m_2 \frac{\int_x^{\omega(F)}(1-F_2(t))dt}{\int_x^{\omega(F)}(1-F_1(t))dt}\right]^2} + \frac{m_2\left[m_1 \frac{1-F_1(x)}{1-F_2(x)} + m_2\right]}{\left[m_1 \frac{\int_x^{\omega(F)}(1-F_1(t))dt}{\int_x^{\omega(F)}(1-F_2(t))dt} + m_2\right]^2}$$

$$\underset{x\to\omega(F)}{\longrightarrow} \frac{m_1(m_1+m_2 a)}{(m_1+m_2 a)^2} + \frac{m_2\left(m_1 \frac{1}{a} + m_2\right)}{\left(m_1 \frac{1}{a} + m_2\right)^2}$$

$$= \frac{m_1(m_1+m_2a)+m_2a(m_1+m_2a)}{(m_1+m_2a)^2} = \frac{(m_1+m_2a)(m_1+m_2a)}{(m_1+m_2a)^2} = 1,$$

onde $a = \lim\limits_{x\to\omega(F)} \frac{1-F_2(x)}{1-F_1(x)}$. Este resultado é válido também quando $a = 0$, $a = +\infty$.

■

3.2.2 Domínios de atracção parcial para distribuições bivariadas

Comecemos por definir o conceito "domínio de atracção parcial" (já mencionado anteriormente):

Definição 3

Dada uma função de distribuição F e uma sucessão monótona, não decrescente, de números naturais, $\{k_n\}_{n\in\mathbf{N}_1}$, dizemos que F pertence ao ***domínio de atracção parcial*** *de uma função de distribuição não degenerada G se existirem sucessões $\{a_n\}_{n\in\mathbf{N}_1}$ e $\{b_n\}_{n\in\mathbf{N}_1}$, $b_n > 0, \forall n$, tais que*

$$F^{k_n}(a_n + b_n x) \longrightarrow G(x),$$

para todos os pontos x de continuidade de G.

Richard Green [24] provou que todas as funções de distribuição têm um domínio de atracção parcial não vazio.

O nosso objectivo é agora o estudo dos domínios de atracção no caso bivariado, i.e., dada uma f.d. G encontrar funções de distribuição F e sucessões $\{a_n^1\}, \{b_n^1\}, \{a_n^2\}$ e $\{b_n^2\}$ $(b_n^1, b_n^2 > 0, \forall n)$ tais que

$$F^{k_n}(a_{1n} + b_{1n}x_1, a_{2n} + b_{2n}x_2) \longrightarrow G(x_1, x_2).$$

De forma a poder utilizar os factos conhecidos da situação univariada, introduziremos o conceito de função de dependência (Galambos [14]):

Definição 4

Seja F uma f.d. de variável vectorial de dimensão m, com funções de distribuição marginais F_i, $1 \leq i \leq m$. Seja D_F uma função definida no cubo de dimensão m, crescente em cada uma das dimensões tal que

$$F(x_1, x_2, \ldots, x_m) = D_F [F_1(x_1), F_2(x_2), \ldots, F_m(x_m)].$$

A função D_F será designada ***função de dependência****.*

Não é difícil verificar que a função de dependência D_F satisfaz a seguinte igualdade:

$$D_{F^n}(y_1, y_2, \ldots, y_m) = D_F^n\left(y_1^{1/n}, y_2^{1/n}, \ldots, y_m^{1/n}\right).$$

Lema 2

Sejam G uma f.d. bivariada, absolutamente contínua, G_1, G_2 as respectivas funções de distribuição marginais, D_G a respectiva função de dependência. Se existirem funções de distribuição F_{1n} e F_{2n} tais que $F_{in} \longrightarrow G_i$, quando $n \to +\infty$ $(i = 1, 2)$, então

$$D_G(F_{1n}(x), F_{2n}(y)) \underset{n \longrightarrow +\infty}{\longrightarrow} G(x, y).$$

Demonstração

De $G(x, y) = D_G(G_1(x), G_2(y))$ sabemos que

$$D_G(x, y) = G\left(G_1^{-1}(x), G_2^{-1}(y)\right).$$

Logo,

$$D_G\left(F_{1n}(x), F_{2n}(y)\right) = G\left(G_1^{-1}\left(F_{1n}(x)\right), G_2^{-1}\left(F_{2n}(y)\right)\right) \underset{n \longrightarrow +\infty}{\longrightarrow} G(x,y).$$

■

Tendo este lema em mente, bem como os resultados no artigo de Green [24] sabemos que existem sucessões $\{k_n\}, \{a_{in}\}, \{b_{in}\}, b_{in} > 0,\ i = 1,2$ e uma f.d. F, tais que, para qualquer f.d. G bivariada, absolutamente contínua,

$$D_G\left(F_1^{k_n}(a_{1n} + b_{1n}x), F_2^{k_n}(a_{2n} + b_{2n}y)\right) \underset{n \longrightarrow +\infty}{\longrightarrow} G(x,y).$$

Sejam

$$\begin{aligned}
u_{in} &= G_i^{-1}\left(\left(1 - 10^{-2^n}\right)^{k_n}\right), \\
v_{in} &= G_i^{-1}\left(\left(1 - 10^{-2^{n-1}}\right)^{k_n}\right), \\
\alpha_{in} &= u_{in} - (v_{in} - u_{in})\left(1 - 2^{-k}\right)2^{k+1}, \\
\beta_{in} &= (v_{in} - u_{in})\,2^{k+1}, \\
k_n &= 10^{2^n + 2^{n-1}}, a_{in} = -\frac{\alpha_{in}}{\beta_{in}} \text{ e } b_{in} = -\frac{1}{\beta_{in}}
\end{aligned}$$

Sabemos que

$$G(x,y) = D_G\left(G_1(x), G_2(y)\right)$$

e uqe

$$G^n(x,y) = D_G^n\left(G_1(x), G_2(y)\right) = D_{G^n}\left(G_1^n(x), G_2^n(y)\right).$$

Por um lado

$$\begin{aligned} & D_G\left(F_1^{k_n}(a_{1n}+b_{1n}x), F_2^{k_n}(a_{2n}+b_{2n}y)\right) \\ & = \left[D_G^{1/k_n}\left(F_1^{k_n}(a_{1n}+b_{1n}x), F_2^{k_n}(a_{2n}+b_{2n}y)\right)\right]^{k_n}, \end{aligned}$$

e por outro

$$\begin{aligned} & D_G\left(F_1^{k_n}(a_{1n}+b_{1n}x), F_2^{k_n}(a_{2n}+b_{2n}y)\right) \\ & = D_{\left(G^{1/k_n}\right)^{k_n}}\left(F_1^{k_n}(a_{1n}+b_{1n}x), F_2^{k_n}(a_{2n}+b_{2n}y)\right) \\ & = \left[D_{G^{1/k_n}}\left(F_1(a_{1n}+b_{1n}x), F_2(a_{2n}+b_{2n}y)\right)\right]^{k_n}. \end{aligned}$$

Logo

$$\begin{aligned} & F(a_{1n}+b_{1n}x, a_{2n}+b_{2n}y) \\ & = D_G^{1/k_n}\left(F_1^{k_n}(a_{1n}+b_{1n}x), F_2^{k_n}(a_{2n}+b_{2n}y)\right) \\ & = D_{G^{1/k_n}}\left(F_1(a_{1n}+b_{1n}x), F_2(a_{2n}+b_{2n}y)\right). \end{aligned}$$

Existem também alguns estudos relativos a um outro tipo de atracção, que poderíamos apelidar de "atracção potência", que em vez de tomar a normalização linear $a_n + b_n x$, considera como normalização $a_n |x|^{b_n} \operatorname{sign} x$ (onde tanto a_n como b_n são positivos). Mohan *et al.* [50] encontraram condições necessárias e suficientes para que uma f.d. F (uni- ou multivariada) pertença ao domínio de atracção potência de uma f.d. G, i.e., que asseguram a existência de sucessões $\{a_n\}$ e $\{b_n\}$, $a_n, b_n > 0, \forall n$, tais que

$$F^n\left(a_n |x|^{b_n} \operatorname{sign} x\right) \longrightarrow G(x),$$

para todos os pontos de continuidade de G ($x \in C(G)$). A atracção parcial neste caso encontra-se ainda por estudar.

3.3 *Spacings*

Order is the shape upon which beauty depends.

Pearl Buck

A simetria de variáveis aleatórias em torno de a, $F(a-x) = 1-F(a+x)$, é uma propriedade "geométrica" com larga repercussão em probabilidade (Jurek [31] aliás, anota que é a nossa familiaridade com a geometria de $\mathbf{R}$ e de $\mathbf{R}^2$ que nos ofusca e nos leva a não notar que muitos dos grandes teoremas de Probabilidade dão simultaneamente resultados sobre Probabilidades e Geometria). Efron [9], por exemplo, demonstrou que simetria radial bastava para que a estatística estudentizada $T_n = \sqrt{n\,(n-1)}\ \overline{X}/\left[\sum\left(X_i - \overline{X}\right)^2\right]^{1/2}$ tivesse a distribuição clássica de t de Student, sem necessidade de pressupor parente gaussiana; e é bem conhecida a relevância de simetria na velocidade de convergência, quer no caso do teorema limite central quer em outros resultados de convergência fraca. No presente parágrafo investigamos, usando spacings extremos - no sentido de realçarem a cauda direita e a cauda esquerda da distribuição parente - as indicações que algumas estatísticas nos podem fornecer sobre a assimetria.

Considere-se então uma sucessão de variáveis aleatórias, $\{X_n\}_{n\in\mathbf{N}_1}$, absolutamente contínuas, independentes e identicamente distribuídas a uma variável X. Considere-se ainda, para $n > 1$, as estatísticas

$$Q_{1n} = \frac{X_{n:n} - X_{n-1:n}}{X_{2:n} - X_{1:n}} \text{ e } Q_{2n} = \frac{X_{n:n} - X_{2:n}}{X_{n-1:n} - X_{1:n}}.$$

Podemos desde já observar que Q_{1n} e Q_{2n} têm como suporte o intervalo $[0, +\infty)$ e que para $n = 2$, $Q_{1n} = 1$. Mais, Q_{1n} e Q_{2n} são variáveis que comparam os pesos das caudas direita e esquerda da distribuição de X, e, consequentemente, avaliam a (as)simetria desta distribuição. É então interessante estudar o valor de $P\left[Q_{1n} \leq 1\right]$, ou, equivalentemente, de $P\left[Q_{2n} \leq 1\right]$, já que, no caso de simetria das variáveis iniciais este valor é, como iremos confirmar, 0.5. Imediato é também o facto que se o suporte da variável X for limitado à direita e ilimitado à esquerda, então Q_{1n} converge, em lei, para zero; analogamente, se o suporte de X for ilimitado à direita e limitado à esquerda, então Q_{1n} converge, em lei, para $+\infty$.

3.3.1 A distribuição de Q_{1n}

Sejam $x \geq 0$, $Y_{1n} = X_{2:n} - X_{1:n}$, $Y_{2n} = X_{n:n} - X_{n-1:n}$ e $n > 3$. Temos que

$$P\left[Q_{1n} \leq x\right] = P[Y_{2n} \leq Y_{1n}x] = \int_0^{+\infty} dy_1 \int_0^{xy_1} dy_2 f_{(Y_{1n},Y_{2n})}\left(y_1, y_2\right)$$

onde $f_{(Y_{1n},Y_{2n})}$ é a função densidade de probabilidade conjunta de (Y_{1n}, Y_{2n}). Ora,

$$f_{(Y_{1n},Y_{2n})}\left(y_1, y_2\right) = \frac{\partial^2}{\partial y_1 \partial y_2} F_{(Y_{1n},Y_{2n})}\left(y_1, y_2\right)$$

e

$$F_{(Y_{1n},Y_{2n})}\left(y_1, y_2\right) = P\left[Y_{1n} \leq y_1, Y_{2n} \leq y_2\right]$$

$$= P\left[X_{2:n} \leq X_{1:n} + y_1, X_{n:n} \leq X_{n-1:n} + y_2\right]$$

$$= \int_{-\infty}^{+\infty} dx_1 \int_{x_1}^{x_1+y_1} dx_2 \int_{x_2}^{+\infty} dx_{n-1} \int_{x_{n-1}}^{x_{n-1}+y_2} dx_n$$
$$f_{(X_{1:n},X_{2:n},X_{n-1:n},X_{n:n})}\left(x_1, x_2, x_{n-1}, x_n\right)$$

$$= \int_{-\infty}^{+\infty} dx_1 \int_{x_1}^{x_1+y_1} dx_2 \int_{x_2}^{+\infty} dx_{n-1} \int_{x_{n-1}}^{x_{n-1}+y_2} dx_n$$
$$\frac{n!}{(n-4)!} \left[F(x_{n-1}) - F(x_2)\right]^{n-4} f(x_1) f(x_2) f(x_{n-1}) f(x_n)$$

Derivando em ordem a y_1 e a y_2 (note-se o teorema da convergência de Lebesgue garante a legitimidade deste passo) conclui-se que

$$P\left[Q_{1n} \leq x\right] = \int_0^{+\infty} dy_1 \int_0^{xy_1} dy_2 \int_{-\infty}^{+\infty} dx_1 \int_{x_1+y_1}^{+\infty} dx_{n-1}$$
$$\frac{n!}{(n-4)!} \left[F(x_{n-1}) - F(x_1+y_1)\right]^{n-4}$$
$$f(x_1) f(x_1+y_1) f(x_{n-1}) f(x_{n-1}+y_2)$$

$$= \frac{n!}{(n-4)!} \int_0^{+\infty} dy_1 \int_{-\infty}^{+\infty} dx_1 \int_{x_1+y_1}^{+\infty} dx_{n-1}$$
$$\left[F(x_{n-1}) - F(x_1+y_1)\right]^{n-4} f(x_1) f(x_1+y_1) f(x_{n-1})$$
$$\left[F(x_{n-1}+xy_1) - F(x_{n-1})\right]$$

$$= \frac{n!}{(n-4)!} (A - B)$$

onde

$$A = \int_0^{+\infty} dy_1 \int_{-\infty}^{+\infty} dx_1 \int_{x_1+y_1}^{+\infty} dx_{n-1}$$
$$\left[F(x_{n-1}) - F(x_1+y_1)\right]^{n-4} f(x_1) f(x_1+y_1) f(x_{n-1}) F(x_{n-1}+xy_1)$$

$$= \int_{-\infty}^{+\infty} dx_1 \int_{x_1}^{+\infty} dx_{n-1} \int_{0}^{x_{n-1}-x_1} dy_1$$
$$\left[F\left(x_{n-1}\right) - F\left(x_1 + y_1\right)\right]^{n-4} f\left(x_1\right) f\left(x_1 + y_1\right) f\left(x_{n-1}\right) F\left(x_{n-1} + xy_1\right)$$

$$= \int_{-\infty}^{+\infty} dx_1 \int_{x_1}^{+\infty} dx_{n-1} \frac{\left[F\left(x_{n-1}\right) - F\left(x_1\right)\right]^{n-3}}{n-3} f\left(x_1\right) f\left(x_{n-1}\right) F\left(x_{n-1}\right)$$
$$+ \int_{0}^{+\infty} dy_1 \int_{-\infty}^{+\infty} dx_1 \int_{x_1+y_1}^{+\infty} dx_{n-1}$$
$$\frac{\left[F\left(x_{n-1}\right) - F\left(x_1 + y_1\right)\right]^{n-3}}{n-3} f\left(x_1\right) f\left(x_{n-1}\right) f\left(x_{n-1} + xy_1\right) x$$

e

$$B = \int_{0}^{+\infty} dy_1 \int_{-\infty}^{+\infty} dx_1 \int_{x_1+y_1}^{+\infty} dx_{n-1}$$
$$\left[F\left(x_{n-1}\right) - F\left(x_1 + y_1\right)\right]^{n-4} f\left(x_1\right) f\left(x_1 + y_1\right) f\left(x_{n-1}\right) F\left(x_{n-1}\right)$$

$$= \int_{-\infty}^{+\infty} dx_1 \int_{x_1}^{+\infty} dx_{n-1} \int_{0}^{x_{n-1}-x_1} dy_1$$
$$\left[F\left(x_{n-1}\right) - F\left(x_1 + y_1\right)\right]^{n-4} f\left(x_1\right) f\left(x_1 + y_1\right) f\left(x_{n-1}\right) F\left(x_{n-1}\right)$$

$$= \int_{-\infty}^{+\infty} dx_1 \int_{x_1}^{+\infty} dx_{n-1} \frac{\left[F(x_{n-1}) - F(x_1)\right]^{n-3}}{n-3} f\left(x_1\right) f\left(x_{n-1}\right) F\left(x_{n-1}\right).$$

Somando as parcelas obtemos o seguinte resultado:

Lema 3

Dada uma sucessão de variáveis aleatórias i.i.d., absolutamente con-

tínuas, a estatística $Q_{1n} = \frac{X_{n:n} - X_{n-1:n}}{X_{2:n} - X_{1:n}}$ *tem a seguinte distribuição*

$$F_{Q_{1n}}(x) = \frac{n!}{(n-3)!} \int_0^{+\infty} dy_1 \int_{-\infty}^{+\infty} dx_1 \int_{x_1+y_1}^{+\infty} dx_{n-1}$$
$$[F(x_{n-1}) - F(x_1+y_1)]^{n-3} f(x_1) f(x_{n-1}) f(x_{n-1} + xy_1) x.$$

3.3.2 O valor de $P[Q_{1n} \leq 1]$ quando X é simétrica

Consideremos novamente $Y_{1n} = X_{2:n} - X_{1:n}$, $Y_{2n} = X_{n:n} - X_{n-1:n}$ e a densidade conjunta de Y_{1n} e Y_{2n}, $f_{(Y_{1n},Y_{2n})}$.

Se a variável parente X for simétrica, então $f_{(Y_{1n},Y_{2n})}$ é uma função simétrica relativamente às suas variáveis-argumento, i.e.,

$$f_{(Y_{1n},Y_{2n})}(y_1, y_2) = f_{(Y_{1n},Y_{2n})}(y_2, y_1).$$

Demonstração

Sejam y_1, $y_2 \geq 0$ e $n > 3$. Temos então que

$$f_{(Y_{1n},Y_{2n})}(y_1, y_2)$$
$$= \int_{-\infty}^{+\infty} dx_1 \int_{x_1+y_1}^{+\infty} dx_{n-1} \frac{n!}{(n-4)!} [F(x_{n-1}) - F(x_1+y_1)]^{n-4}$$
$$f(x_1) f(x_1+y_1) f(x_{n-1}) f(x_{n-1}+y_2)$$

$$= \int_{-\infty}^{+\infty} dx_{n-1} \int_{-\infty}^{x_{n-1}-y_1} dx_1 \frac{n!}{(n-4)!} [F(x_{n-1}) - F(x_1+y_1)]^{n-4}$$
$$f(x_1) f(x_1+y_1) f(x_{n-1}) f(x_{n-1}+y_2)$$

$(x_{n-1} = -x - y_2)$

$$= \int_{-\infty}^{+\infty} dx \int_{-\infty}^{-x-y_2-y_1} dx_1 \frac{n!}{(n-4)!} \left[F(-x-y_2) - F(x_1+y_1)\right]^{n-4}$$
$$f(x_1) f(x_1+y_1) f(-x-y_2) f(-x)$$

$(x_1 = -z - y_1)$

$$= \int_{-\infty}^{+\infty} dx \int_{x+y_2}^{+\infty} dz \frac{n!}{(n-4)!} \left[F(-x-y_2) - F(-z)\right]^{n-4}$$
$$f(-z-y_1) f(-z) f(-x-y_2) f(-x)$$

$$= \int_{-\infty}^{+\infty} dx \int_{x+y_2}^{+\infty} dz \frac{n!}{(n-4)!} \left[F(z) - F(x+y_2)\right]^{n-4}$$
$$f(z+y_1) f(z) f(x+y_2) f(x)$$

$$= f_{(Y_{1n}, Y_{2n})}(y_2, y_1).$$

Para $y_1 < 0$ ou $y_2 < 0$, $f_{(Y_{1n}, Y_{2n})}(y_1, y_2) = f_{(Y_{1n}, Y_{2n})}(y_2, y_1) = 0.$ ∎

Lema 4

Sejam Y_1, Y_2 duas variáveis aleatórias com densidade de probabilidade conjunta $f_{(Y_1,Y_2)}$ simétrica relativamente às suas variáveis. Então

$$P[Y_1 \leq Y_2] = P[Y_2 \leq Y_1] = \frac{1}{2}.$$

Demonstração

Basta notar que

$$\begin{aligned}
P[Y_1 \leq Y_2] &= \int_{-\infty}^{+\infty} dy_1 \int_{y_1}^{+\infty} dy_2 f_{(Y_1,Y_2)}(y_1, y_2) \\
&= \int_{-\infty}^{+\infty} dy_2 \int_{y_2}^{+\infty} dy_1 f_{(Y_1,Y_2)}(y_2, y_1) \\
&= \int_{-\infty}^{+\infty} dy_2 \int_{y_2}^{+\infty} dy_1 f_{(Y_1,Y_2)}(y_1, y_2) \\
&= P[Y_2 \leq Y_1].
\end{aligned}$$

■

Dos lemas anteriores resulta o seguinte corolário:

Corolário 1

Se X for simétrica e absolutamente contínua então

$$P[Q_{1n} \leq 1] = 1/2.$$

Mais uma vez se confirma a importância da ordenação, e da informação que ela transporta. Não posso terminar esta secção a propósito de *spacings* e estatísticas ordinais, sem referir o recente trabalho de Brilhante [5], no qual a autora, usando funções de estatísticas ordinais, propõe e estuda alguns testes para inferência estatística em modelos não gaussianos, nomeadamente nos modelos de distribuição Pareto generalizada e de distribuição laplaciana.

3.4 Funções monótonas de ordem r

Beauty is the first test; there is no permanent place in the world for ugly mathematics.

G. H. Hardy

3.4.1 O teorema de Krein-Milman

As representações integrais das funções características das funções de distribuição infinitamente divisíveis são, sabe-se hoje, representações integrais de Choquet.

A forma integral do teorema de Krein-Milman, que pode ser vista como uma particularização do teorema de Choquet, é um instrumento da Análise Funcional deveras útil quando se pretende encontrar representações integrais para os elementos de um conjunto convexo e compacto (Choquet [6]). Este teorema afirma que:

Teorema 3

(Phelps [57]) Todo o ponto de um subconjunto C, compacto e convexo, de um espaço localmente convexo é o baricentro de uma medida de probabilidade μ em C que é suportada pelo fecho de pontos extremos de C.

Recorde-se que um conjunto C diz-se **convexo** se dados dois seus elementos, x e y, a combinação linear convexa destes dois elementos,

$$a = px + (1-p)\,y,\ p \in [0,1]$$

é ainda um elemento de C. Recorde-se ainda que, dado um conjunto convexo C, dizemos que $a \in C$ é um **ponto extremo** se, e só se, $C - \{a\}$ é ainda convexo. Por outras palavras, se a é um ponto extremo de C e $a = px + (1-p)y$ com $x,\ y \in C$, $p \in [0,1]$, ou $x = y = a$, ou p é igual a zero ou um. Relativamente à medida μ, quando dizemos que esta é suportada pelo conjunto dos pontos extremos de C, que vamos denotar por $\varepsilon(C)$, queremos com isso dizer que $\mu(C \backslash \varepsilon(C)) = 0$. Falta apenas esclarecer o significado de um ponto x ser "o baricentro da medida μ".

Definição 5

Dados um conjunto C não vazio, compacto, subconjunto de um espaço localmente convexo E, e μ uma medida de probabilidade em C, dizemos que um ponto $x \in E$ é o baricentro da medida μ, ou que é representado por μ se para todo o funcional linear contínuo de E, f,

$$f(x) = \int_C f d\mu. \tag{11}$$

O que o teorema anterior nos diz é que podemos em vez de C em (11) usar $\varepsilon(C)$:

$$f(x) = \int_{\varepsilon(C)} f d\mu,$$

e, no caso em que $f(x) = x$ obtemos a representação integral de x:

$$x = \int_{\varepsilon(C)} y d\mu(y).$$

Regressando às distribuições infinitamente divisíveis, foi Johansen [29] quem estabeleceu que o logarítmo de uma função característica infinitamente divisível é definida positiva, e com base nesse resultado conseguiu identificar os pontos extremos do conjunto convexo das funções características infinitamente divisíveis; são os tipos

i) degenerado;

ii) gaussiano (o que estabelece o teorema de Lévy-Cramér);

iii) Poisson generalizado com suporte em reticulados em $\mathbf{R}^+$ ou $\mathbf{R}^-$ (o que estabelece, como corolário, o teorema de Raikov).

Como curiosidade, é de mencionar que o teorema de Lévy-Cramér — que estabelece que uma variável gaussiana pode ser decomposta apenas como soma de variáveis gaussianas — foi utilizado por Lévy, o qual dedicou grande parte de um seu trabalho às consequências deste resultado, mesmo sem o ter demonstrado, tão convicto que estava da sua veracidade. Foi Cramér quem, em 1936, usando funções características, demonstrou o teorema com uma prova que Lévy classificou de excessivamente técnica, para um resultado que considerava intituitivo (cf. Loève [40]). O teorema de Raikov veio na sequência do anterior e estabelece que uma v.a. Poisson pode ser decomposta apenas como soma de Poissons.

Uma outra aplicação notável do teorema de Krein-Milman é o teorema de Bernstein para as funções completamente monótonas, cuja prova pode ser encontrada, e.g., em Phelps [57]. Recorde-se que uma função f definida em $(0, +\infty)$ diz-se **completamente monótona** se possuir derivadas $f^{(n)}$ de todas as ordens e se $(-1)^n f^{(n)}(x) \geq 0$, para todo o real positivo x.

Teorema 4 *(de Bernstein)*

Uma função f é completamente monótona sse se puder escrever da forma

$$f(t) = \int_0^{+\infty} \exp(-tx)\, dF(x) \tag{12}$$

onde F é uma medida em $[0, +\infty)$ (não necessariamente finita), i.e., sse for a transformada de Laplace de uma medida definida em $[0, +\infty)$.

Entre os exemplos de funções completamente monótonas podemos encontrar as restrições a $(0, +\infty)$ do modelo extremal Fréchet, das funções $g(x) = \frac{1}{x}$ e $g(x) = \exp(-x)$.

3.4.2 Representação integral de funções monótonas de ordem r

O conceito, um pouco vago, de "smoothness" pode ser expresso de várias formas. A monotonicidade de uma função f, a existência de derivadas monótonas, unimodalidade, são talvez as "condições regulares" usualmente necessárias. Existem muitos resultados no que diz respeito a funções monótonas, funções com derivadas monótonas, e funções convexas (um conceito facilmente expresso em termos da segunda derivada, quando $f \in C^2$); por outro lado, funções $f \in C^\infty$ com derivadas monótonas têm propriedades interessantes — por exemplo, a classe de Bernstein já referida, das funções completamente monótonas.

Vamos explorar a situação intermédia, quando $(-1)^k f^{(k)}(x) \geq 0$ para $k = 0, 1, 2, \ldots, n$ e $x \geq 0$, e obter uma representação integral de Choquet semelhante a (12); mostraremos ainda que a representação de Bernstein das funções completamente monótonas é obtida como o limite, quando n converge para $+\infty$, da representação geral que estabelecemos.

A fórmula de inversão da transformada integral é generalizável, usando o "cálculo fraccionário", ou "cálculo dif-integral" — por outras palavras, integrais e derivadas de ordem fraccionária, uma ideia inspirada no integral de Cauchy iterado (para uma apresentação bem documentada, dos desenvolvimentos iniciais da ideia de derivadas e integrais fraccionários, e uma discussão e conciliação das definições de Riemann e de Liouville, cf. Oldham e Spanier [51]).

Representações integrais de funções com derivadas monótonas, e teorema de inversão foram estudados por Williamson [73] e Lévy [38], que caracterizou funções de distribuição e funções características com derivadas monótonas, e observou que o resultado generaliza o teorema de Khintchine sobre unimodalidade e a condição de Pólya sobre funções características. Estes trabalhos resolveram os problemas principais de caracterização de funções com derivadas monótonas. De facto, a monotonicidade tem também um papel importante no estudo de funções características. As caracterizações de Bochner e de Cramér para funções características são virtualmente impossíveis de serem usadas e, em termos práticos, apenas a condição suficiente de Pólya é aproveitável. Mostramos, novamente, que esta condição é, também ela, uma representação integral de Choquet, e um simples corolário dos resultados sobre monotonicidade generalizada.

A abordagem apresentada, usando o teorema de Choquet, permite uma visão profunda da estrutura da monotonicidade generalizada, e tem uma interpretação probabilística, em termos de produtos de variáveis aleatórias independentes, sendo um dos factores uma v.a. de distribuição beta. Mostramos assim que o teorema de Khintchine sobre funções de distribuição unimodais, pode ser visto como uma representação integral de Choquet, e, consequentemente, generalizamos o conceito de unimodalidade, discutimos a transformação beta, e terminamos com uma referência à classe de Mejzler.

Algumas definições

Ao longo deste parágrafo as funções de distribuição são definidas em $[-\infty, +\infty]$. O suporte de uma função de distribuição F com átomo em $+\infty$ de peso $b \in (0, 1]$ é definido como sendo o suporte da função de distribuição $\frac{F}{b}$. Se a f.d. tiver um átomo em $-\infty$ de valor $a \in (0, 1]$, o seu suporte é, por definição, o suporte de $\frac{F-a}{1-a}$. Obviamente, quando F tem simultaneamente átomos de probabilidade em $-\infty$ e $+\infty$, por exemplo, com os valores acima referidos, o seu suporte será definido pelo suporte de $\frac{F-a}{b-a}$.

Definição 6

Seja f uma função definida em $[-\infty, +\infty]$, não negativa, limitada, monótona não crescente em $[0, +\infty]$, tal que $f(x) = 0$ para todo $x < 0$.

Dizemos então que f é ***monótona de ordem zero.*** *Suponhamos agora que f é também convexa em $A = \{x : f(x) > 0\}$. Dizemos que f é* ***monótona de ordem n em*** *$[0, +\infty]$, $n = 1, 2, \ldots$, sse*

1. *$f \in C^n(int(A))$, i.e., f possui derivada de ordem n contínua no interior do conjunto A, $int(A)$.*
2. *Para $k = 1, \ldots, n$, $(-1)^k f^{(k)}$, é não negativa, monótona não crescente em $int(A)$.*

Denotaremos por M_n^*, $n = 0, 1, \ldots$, a classe de funções monótonas de ordem n em $[0, +\infty]$. Observe-se que M_n^* é um cone convexo bicudo (*pointed*). Não é difícil provar que $M_n^* \supseteq M_{n+1}^*$ e que se $f \in M_n^*$ então $(-1)^k f^{(k)}$ é convexa para $k = 0, 1, \ldots, n-1$.

Concluímos assim que $M_\infty^* = \bigcap\limits_{n=1}^{\infty} M_n^*$ é a classe das funções completamente monótonas.

Definição 7

Dizemos que uma função f ***é monótona de ordem n em*** *$(0, +\infty]$ sse para todo real $a > 0$, a função $f_a(x) = f(a + x)$ é monótona de ordem n em $[0, +\infty]$.*

Denotamos por M_n a classe das funções monótonas de ordem n em $(0, +\infty]$.

Definição 8

Sejam F uma f.d. e $A = \left(\alpha\left(\hat{F}\right), \omega\left(\hat{F}\right)\right)$, o interior do suporte da f.d. definida por

$$\hat{F}(x) = \frac{F(x) - \lim\limits_{x \to -\infty} F(x)}{\lim\limits_{x \to +\infty} F(x) - \lim\limits_{x \to -\infty} F(x)}.$$

Suponhamos que F é diferenciável em $A - \{0\}$ e consideremos $f(x) = \frac{d}{dx}F(x)$ se $x \in A - \{0\}$, $f(x) = 0$ se $x \notin A$. Dizemos que a f.d. F é um elemento de $\mathbf{D_{n+1}}$, $n = 0, 1, \ldots$, sse as funções

$$f_1(x) = f(x)\,\mathbf{I}_{(0,+\infty)}(x) \;\; e \;\; f_2(x) = f(-x)\,\mathbf{I}_{(0,+\infty)}(x)$$

são monótonas de ordem n em $(0, +\infty]$.

Alternativamente, relativamente a f_2, podemos pôr como condição que a restrição $f\mathbf{I}_{(-\infty,0)}$ seja absolutamente monótona de ordem n, i.e., $f^{(k)} \geq 0$, $k = 1, ..., n$ para $x < 0$. O trabalho original de Bernstein é sobre funções absolutamente monótonas, i.e., tais que $f^{(k)} \geq 0$, $k = 1, 2, ...$ Foi o facto de $f(-x)$ ser então a transformada de Laplace de uma função não decrescente que trouxe as funções completamente monótonas para o primeiro plano.

A classe D^{*}_{n+1} define-se de forma análoga, com as transformações óbvias.

Não é difícil provar que (cf. Williamson [73]):

Lema 5

Se a função f é monótona de ordem n, a função $x^{k-1}f^{(k)}(x)$ é integrável em $(0,+\infty)$ e $\lim_{x\to+\infty} x^k f^{(k)}(x) = 0$ para $k = 1, 2, \ldots, n$ e $n = 1, 2, \ldots$.

Pontos extremos de D_n e representações integrais

Vamos agora identificar os pontos extremos de D_n, que é obviamente um conjunto convexo. A próxima proposição é uma consequência imediata dos seguintes factos:

1. $F \in D_n$ sse $F\mathbf{I}_{(0,+\infty]}$ e $F\mathbf{I}_{[-\infty,0)} + \mathbf{I}_{[0,+\infty]}$ pertencem a D_n.
2. Se F é a f.d. da uma v.a. X,

$$F \in D_n \text{ se, e só se, } \tilde{F} \in D_n,$$

onde $\tilde{F}$ é a f.d. da v.a. $-X$.

Proposição 5

Se F é um ponto extremo do conjunto D_n, $n = 1, 2, \ldots$, então $F(0) = 0$ ou $F(0) = 1$. As funções de distribuição degeneradas em $-\infty$, em 0 e em $+\infty$, que denotaremos por $F_{n,-\infty}$, $F_{n,0}$ and $F_{n,+\infty}$, são os únicos pontos extremos degenerados de D_n, $n = 1, 2, \ldots$.

No que diz respeito aos pontos extremos não degenerados de D_n, $n = 1, 2, \ldots$, temos que:

Teorema 5

As funções de distribuição $F_{n,a}$ definidas por

$$\frac{d}{dx}F_{n,a}(x) = \frac{n}{|a|^n}\left(|a| - x\,\mathrm{sign}(x)\right)^{n-1}\mathbf{I}_{(\min(0,a),\max(0,a))}(x), \tag{13}$$

onde $a \in \mathbf{R} - \{0\}$, são pontos extremos de D_n, $n = 1, 2, \ldots$.

Demonstração

É imediato que se $F_{n-1,a}$ é ponto extremo de D_{n-1}, então $F_{n,a}$ é um ponto extremo de D_n, $n = 2, 3, \ldots$. Tendo em conta as observações anteriores, é suficiente provar que $F_{1,a}$ é um ponto extremo de D_1 para todo $a > 0$. Observemos que, neste caso $\frac{d}{dx}F_{1,a} = \frac{1}{a}\mathbf{I}_{(0,a)}$, i.e., $F_{1,a}$ é a f.d. de uma variável uniforme em $(0, a)$. Sejam $G_1, G_2 \in D_1, \lambda \in [0, 1]$ tais que

$$F_{1,a} = \lambda G_1 + (1 - \lambda) G_2.$$

Se $G_1(a) = 0$ (resp. $G_2(a) = 0$), então $\lambda = 0$ e $G_2(a) = 1$ (resp. $\lambda = 1$ e $G_1(a) = 1$), e, consequentemente, $F_{1,a} \equiv G_2$ (resp. $F_{1,a} \equiv G_1$).

Suponhamos agora que $0 < G_1(a) < 1$ (que, dado que $F_{1,a}(a) = 1$ implica que $G_2(a) > 0$). Temos que

$$\frac{1}{a}\mathbf{I}_{(0,a)}(x) = \lambda\frac{d}{dx}G_1(x) + (1 - \lambda)\frac{d}{dx}G_2(x),$$

onde, como consequência de $G_i \in D_1$, $\frac{d}{dx}G_i$ é não negativo e monótona não crescente em $(0, a)$, para $i = 1, 2$. Resulta assim que $\frac{d}{dx}G_i$ é constante em $(0, a)$, i.e., $G_i = \mu_i F_{1,a}$, $i = 1, 2$, onde $\mu_i \in (0, 1)$ e tais que

$$1 = F_{1,a}(a) = \lambda\mu_1 + (1 - \lambda)\mu_2;$$

$F_{1,a}$ é um ponto extremo de D_1. ■

Estabelecemos agora o resultado recíproco:

Teorema 6

Se F é um ponto extremo não degenerado de D_n, $n = 1, 2, \ldots$, então $F = F_{n,a}$ para algum $a \in \mathbf{R} - \{0\}$, onde $F_{n,a}$ é a função definida por (13).

Com o intuito de demonstrar o teorema anterior vamos definir um subconjunto E_n, de D_n, definido da seguinte forma: $F \in E_n$ se, e só se, F for uma f.d. honesta e $\forall G \in D_n$

$$(G \neq 0 \text{ , } G \neq F \text{ e } (F - G) \in D_n)$$
$$\Rightarrow (\text{supp}G = \text{supp}F \text{ e } \exists \lambda \in (0,1) \ \forall x \in \text{supp}F : G(x) = \lambda F(x)) .$$

Começamos por enunciar e provar alguns lemas.

Lema 6

Se F é ponto extremo de D_n então $F(0) = 0$ ou $F(0) = 1$.

Demonstração De facto, se tivéssemos $F(0) = a \in (0,1)$ poderíamos decompor F na soma convexa $F = aF_1 + (1 - a) F_2$ com

$$F_1(x) = \frac{F(x)}{a} \mathbf{I}_{(-\infty,0)}(x) + \mathbf{I}_{[0,+\infty)}(x) \text{ e } F_2(x) = \frac{F(x) - a}{1 - a} \mathbf{I}_{[0,+\infty)}(x) ,$$

e F não seria ponto extremo. ■

Lema 7

A distribuição F_{na} definida em $\mathbf{R}$ *por*

$$F_{na}(x) = \left[1 - \left(1 - \frac{x}{a}\right)^n\right] \mathbf{1}_{(0,a)}(x) + \mathbf{1}_{[a,+\infty)}(x).$$

é um elemento de E_n.

Demonstração

Seja $G \in D_n$, com $G \neq 0$ e $G \neq F$, tal que $F_{na} - G \in D_n$.

Caso 1, $x < 0$: Neste caso temos que

$$(F_{na} - G)(x) = -G(x).$$

Como $F_{na} - G$ é uma função de distribuição resulta que, para $x < 0$, $G(x) = 0 = F(x) = \lambda F(x)$, $\forall \lambda \in [0, 1]$.

Caso 2, $x \in [0, a)$: Temos que

$$(F_{na} - G)(x) = \left[1 - \left(1 - \frac{x}{a}\right)^n\right] - G(x) \geq 0$$

e, para $1 \leq k \leq n$, (a partir desta ordem não temos a garantia de existência de derivadas)

$$(-1)^{k-1}(F_{na} - G)^{(k)}(x) = \frac{n!}{(n-k)!}\frac{1}{a^k}\left(1 - \frac{x}{a}\right)^{n-k} - (-1)^{k-1}G^{(k)}(x) \tag{14}$$

e

$$(-1)^{k-1}G^{(k)}(x) \tag{15}$$

são não negativas, convexas, monótonas não crescentes em $(0,a)$, e, consequentemente (por serem **convexas**, cf. Roberts e Varberg [64]), existem funções h_{1k}, h_{2k}, monótonas não decrescentes, e elementos c_{1k}, $c_{2k} \in (0,a)$, tais que, para $x \in (0,a)$

$$(14) = \frac{n!}{(n-k)!}\frac{1}{a^k}\left(1-\frac{c_{2k}}{a}\right)^{n-k} - (-1)^{k-1} G^{(k)}(c_{2k}) + \int_{c_{2k}}^{x} h_{2k}(t)\,dt, \tag{16}$$

$$(15) = (-1)^{k-1} G^{(k)}(x) = (-1)^{k-1} G^{(k)}(c_{1k}) + \int_{c_{1k}}^{x} h_{1k}(t)\,dt$$

e, substituindo em (16) a expressão obtida para (15),

$$\begin{aligned} &\frac{n!}{(n-k)!}\frac{1}{a^k}\left(1-\frac{x}{a}\right)^{n-k} - (-1)^{k-1} G^{(k)}(c_{1k}) - \int_{c_{1k}}^{x} h_{1k}(t)\,dt \\ &= \frac{n!}{(n-k)!}\frac{1}{a^k}\left(1-\frac{c_{2k}}{a}\right)^{n-k} - (-1)^{k-1} G^{(k)}(c_{2k}) + \int_{c_{2k}}^{x} h_{2k}(t)\,dt. \end{aligned}$$

Em particular, para $k=n$,

$$\begin{aligned} &\frac{n!}{a^n} - (-1)^{n-1} G^{(n)}(c_{1n}) - \int_{c_{1n}}^{x} h_{1n}(t)\,dt \\ &= \frac{n!}{a^n} - (-1)^{n-1} G^{(n)}(c_{2n}) + \int_{c_{2n}}^{x} h_{2n}(t)\,dt \Leftrightarrow \end{aligned}$$

$$\Leftrightarrow (-1)^{n-1}\left[G^{(n)}(c_{2n}) - G^{(n)}(c_{1n})\right] = \int_{c_{1k}}^{x} h_{1n}(t)\,dt + \int_{c_{2k}}^{x} h_{2n}(t)\,dt$$

$$\Leftrightarrow (-1)^{n-1}\left[G^{(n)}(c_{2n}) - G^{(n)}(c_{1n})\right] - \int_{c_{2k}}^{x} h_{2n}(t)\,dt = \int_{c_{1k}}^{x} h_{1n}(t)\,dt$$

donde resulta que

$$(-1)^{n-1}\left[G^{(n)}(c_{2n}) - G^{(n)}(c_{1n})\right] - \int_{c_{1k}}^{x} h_{1n}(t)\,dt \text{ e } \int_{c_{2k}}^{x} h_{2n}(t)\,dt$$

são funções afins em $(0,a)$ (apenas neste caso se pode dar a igualdade entre uma função côncava e uma função convexa). Mais, como $\int_{c_{1k}}^{x} h_{1n}(t)\,dt$ e $\int_{c_{2k}}^{x} h_{2n}(t)\,dt$ são ambas funções monótonas não crescentes, resulta que, sendo a sua soma constante relativamente a x, $\int_{c_{1k}}^{x} h_{1n}(t)\,dt$ e $\int_{c_{2k}}^{x} h_{2n}(t)\,dt$ são funções constantes. Existe então uma constante b tal que

$$\int_{c_{1k}}^{x} h_{1n}(t)\,dt = b$$

e

$$(15) = (-1)^{n-1} G^{(n)}(x) = (-1)^{n-1} G^{(n)}(c_{1n}) + b \Leftrightarrow$$

$$G^{(n)}(x) = \left[G^{(n)}(c_{1n}) + (-1)^{n-1} b\right],$$

ou seja, G é uma função polinomial de grau n em $(0,a)$. Podemos estender este domínio a $[0,a)$ pois G é uma função contínua à direita. Temos então dois polinómios de grau n, F_{na} e G, elementos de D_n. De modo a simplificar a apresentação suporemos que estes se escrevem, para $x \in [0,a)$ na forma

$$F(x) = \sum_{i=0}^{n} f_i \left(\frac{x}{a}\right)^i \text{ e } G(x) = \sum_{i=0}^{n} g_i \left(\frac{x}{a}\right)^i.$$

com $F(0) = G(0) = 0\,(\Rightarrow f_0 = g_0 = 0)$. A função F é uma função contínua em $\mathbf{R}$ tal que $F(a) = \sum_{i=0}^{n} f_i = 1$ e $F^{(k)}(a) = 0$ $\left(\Rightarrow G^{(k)}(a) = 0\right)$, para $k = 1, ..., n-1$; mais, $F^{(k)}$, $k = 1, ..., n-1$, é contínua em a. Temos também que, para $k = 1, ..., n$ e $x \in [0,a)$,

$$(-1)^{k-1} F^{(k)}(x) = (-1)^{k-1} \sum_{i=k}^{n} \frac{f_i}{a^k} \frac{i!}{(i-k)!} \left(\frac{x}{a}\right)^{i-k} \geq 0,$$

$$(-1)^{k-1} G^{(k)}(x) = (-1)^{k-1} \sum_{i=k}^{n} \frac{g_i}{a^k} \frac{i!}{(i-k)!} \left(\frac{x}{a}\right)^{i-k} \geq 0$$

e

$$(-1)^{k-1} \left[F^{(k)}(x) - G^{(k)}(x)\right] = (-1)^{k-1} \sum_{i=k}^{n} \frac{f_i - g_i}{a^k} \frac{i!}{(i-k)!} \left(\frac{x}{a}\right)^{i-k} \geq 0.$$

Para $x = 0$ (e $k = 1, ..., n$) temos que

$$(-1)^{k-1} F^{(k)}(0) = (-1)^{k-1} \frac{f_k}{a^k} k! \geq 0 \Rightarrow (-1)^{k-1} f_k \geq 0,$$

$$(-1)^{k-1} G^{(k)}(0) = (-1)^{k-1} \frac{g_k}{a^k} k! \geq 0 \Rightarrow (-1)^{k-1} g_k \geq 0$$

e

$$(-1)^{k-1} \left[F^{(k)}(0) - G^{(k)}(0)\right] = (-1)^{k-1} \frac{f_k - g_k}{a^k} k! \geq 0$$

$$\Rightarrow (-1)^{k-1} f_k \geq (-1)^{k-1} g_k \geq 0$$

$$\Rightarrow \exists \lambda_k \in [0,1] : (-1)^{k-1} g_k = (-1)^{k-1} \lambda_k f_k$$

$$\Rightarrow \exists \lambda_k \in [0,1] : g_k = \lambda_k f_k.$$

Pretende-se provar que $\lambda_1 = ... = \lambda_n$. Para $x = a$ e $k < n$

$$(-1)^{k-1} F^{(k)}(a) = (-1)^{k-1} \sum_{i=k}^{n} \frac{f_i}{a^k} \frac{i!}{(i-k)!} = 0 \tag{17}$$

$$\Rightarrow \sum_{i=k}^{n} f_i \frac{i!}{(i-k)!} = 0, \tag{18}$$

o que implica que

$$(-1)^{k-1} G^{(k)}\left(a^-\right) = \sum_{i=k}^{n} \lambda_i f_i \frac{i!}{(i-k)!} = 0. \tag{19}$$

Fazendo $k = n-1$ em (18) e (19) vem que

$$\left\{ \begin{array}{l} \sum_{i=n-1}^{n} f_i \frac{i!}{(i-k)!} = 0 \\ \sum_{i=n-1}^{n} \lambda_i f_i \frac{i!}{(i-k)!} = 0 \end{array} \right. \Leftrightarrow \left\{ \begin{array}{l} f_{n-1}(n-1)! + f_n n! = 0 \\ \lambda_{n-1} f_{n-1}(n-1)! + \lambda_n f_n n! = 0 \end{array} \right.$$

$$\Leftrightarrow \left\{ \begin{array}{l} \lambda_n f_{n-1}(n-1)! + \lambda_n f_n n! = 0 \\ \lambda_{n-1} f_{n-1}(n-1)! + \lambda_n f_n n! = 0 \end{array} \right.$$

$$\Rightarrow \lambda_n = \lambda_{n-1},$$

pois $f_i \neq 0$, $i = 1, ..., n$; para $k = n-2$ vem que

$$\left\{ \begin{array}{l} f_{n-2}(n-2)! + f_{n-1}(n-1)! + f_n \frac{n!}{2!} = 0 \\ \lambda_{n-2} f_{n-2}(n-2)! + \lambda_{n-1} f_{n-1}(n-1)! + \lambda_n f_n \frac{n!}{2} = 0 \end{array} \right.$$

$$\Leftrightarrow \left\{ \begin{array}{l} \lambda_n f_{n-2}(n-2)! + \lambda_n f_{n-1}(n-1)! + \lambda_n f_n \frac{n!}{2!} = 0 \\ \lambda_{n-2} f_{n-2}(n-2)! + \lambda_n f_{n-1}(n-1)! + \lambda_n f_n \frac{n!}{2} = 0 \end{array} \right.$$

$$\Rightarrow \lambda_n = \lambda_{n-2};$$

e assim sucessivamente. Suponhamos que $\lambda_n = \lambda_{n-1} = ... = \lambda_{n-m}$, com $m \leq n-2$. Então

$$\left\{ \begin{array}{l} \sum_{i=n-m-1}^{n} f_i \frac{i!}{(i-k)!} = 0 \\ \sum_{i=n-m-1}^{n} \lambda_i f_i \frac{i!}{(i-k)!} = 0 \end{array} \right.$$

$$\Leftrightarrow \begin{cases} f_{n-m-1}(n-m-1)! + \sum_{i=n-m}^{n} f_i \frac{i!}{(i-k)!} = 0 \\ \lambda_{n-m-1} f_{n-m-1}(n-m-1)! + \lambda_n \sum_{i=n-m}^{n} f_i \frac{i!}{(i-k)!} = 0 \end{cases}$$

$$\Rightarrow \lambda_n = \lambda_{n-m-1}.$$

Concluímos assim que no intervalo $[0, a)$: $G = \lambda_n F$.

Caso 3, $x \geq a$:

$$(F_a - G)(x) = 1 - G(x)$$

e

$$(F_a - G)^{(1)}(x) = -G^{(1)}(x).$$

Como $G \in D_n$, $-G^{(1)}(x) \geq 0$ e $G^{(1)}(x) \geq 0$, donde resulta que $G(x)$ é constante para $x \geq a$. Como, ainda, $1 - G(x) \geq 0$ resulta que $G = \lambda F$ em $[a, +\infty]$, com $\lambda = G(a) \in [0, 1]$. ∎

Lema 8

A distribuição $1 - F_{na}(-x)$ *é também um elemento de* E_n

$$1 - F_{na}(-x) = \left(1 - \left[1 - \left(1 + \frac{x}{a}\right)^n\right]\right) \mathbf{1}_{(0,a)}(-x) + \mathbf{1}_{(-\infty,0]}(-x)$$

$$= \left(1 + \frac{x}{a}\right)^n \mathbf{1}_{(-a,0)}(x) + \mathbf{1}_{[0,+\infty)}(x).$$

Demonstração

Análoga à prova do Lema 7. ∎

Lema 9

A função de distribuição definida por $F(x) = \mathbf{1}_{[0,+\infty)}(x)$ é um elemento de E_n.

Demonstração

Análoga aos casos 1 e 3 do Lema 7. ■

O seguinte lema dispensa demonstração.

Lema 10

As funções de distribuição não honestas, F_1 e F_2, definidas em $\mathbf{R}$ por $F_1(x) = 1$ e $F_2(x) = 0$ são pontos extremos de D_n.

Lema 11

Os elementos de E_n são pontos extremos de D_n.

Demonstração

De facto, se assim não fosse, poderíamos escrever um elemento arbitrário F de E_n como a combinação linear convexa não trivial de dois elementos, F_1 e F_2, de D_n, distintos de F. Ora, de $F = \lambda F_1 + (1-\lambda) F_2$ resultaria, por definição de E_n, a existência de elementos $\alpha_1, \alpha_2 \in (0,1)$, tais que

$\lambda F_1 = \alpha_1 F$ e $(1-\lambda) F_2 = \alpha_2 F$. Substituindo,

$$F = \lambda F_1 + (1-\lambda) F_2 \Leftrightarrow F = \alpha_1 F + \alpha_2 F \Leftrightarrow \alpha_1 + \alpha_2 = 1,$$

obteríamos que

$$\lambda F_1 = \alpha_1 F \text{ e } (1-\lambda) F_2 = (1-\alpha_1) F$$

e, consequentemente, que $F_1 = \frac{\alpha_1}{\lambda} F$ e $F_2 = \frac{1-\alpha_1}{1-\lambda} F$. Recorde-se que F é uma função de distribuição honesta, donde, obrigatoriamente, teríamos de ter $\frac{\alpha_1}{\lambda} \leq 1$ e $\frac{1-\alpha_1}{1-\lambda} \leq 1$, i.e., $\alpha_1 = \lambda$, o que contrariaria a hipótese de F_1 e F_2 serem funções de distribuição distintas de F. ■

Podemos, deste último lema e do Lema 6, concluir que se $F \in E_n$ então $F(0) = 0$ ou $F(0) = 1$.

Lema 12

Os pontos extremos não degenerados de D_n são funções de distribuição honestas, e contínuas em $\mathbf{R}$.

Demonstração

Suponhamos, por exemplo, que F é uma função de distribuição da classe D_n tal que $F(-\infty) = a \in (0,1)$. Podemos decompor F na soma

$$F(x) = aF_{-\infty}(x) + (1-a)\frac{F(x) - a}{1-a}.$$

De modo análogo, se $F(+\infty) = b \in (0,1)$,

$$F(x) = b\frac{F(x) - (1-b) F_{+\infty}(x)}{b} + (1-b) F_{+\infty}(x).$$

Quanto à continuidade de F, basta fazer a análise em zero. Se F tivesse um átomo de probabilidade em zero, digamos de valor $a \in (0,1)$, poderíamos escrever $F = aF_0 + (1-a)\,\frac{F-aF_0}{1-a}$. ■

Lema 13

Os pontos extremos não degenerados de D_n são elementos de E_n.

Demonstração

Sejam F um ponto extremo de D_n e G um elemento de D_n distinto de F, não nulo, tal que $F - G \in D_n$. Podemos então escrever F como a soma de dois elementos de D_n: $F = G + (F - G)$. Sejam $\alpha = \lim_{x \to +\infty} F(x) = 1$ e $\lambda = \lim_{x \to +\infty} [F(x) - G(x)]$. Dado que $G \neq 0$ e $G \neq F$ temos $\lambda \neq 1$ e $\lambda \neq 0$. As funções $\frac{G}{1-\lambda}$ e $\frac{F-G}{\lambda}$ são elementos de D_n e

$$F = (1-\lambda)\,\frac{G}{1-\lambda} + \lambda\frac{F-G}{\lambda},$$

donde resulta, por F ser um ponto extremo de D_n, que

$$F = \frac{G}{1-\lambda} \Leftrightarrow G = (1-\lambda)\,F,$$

i.e., F é um elemento de E_n. ■

Lema 14

Os pontos extremos não degenerados de D_n obedecem à condição definida por (13).

Demonstração

Suponhamos que $n > 1$. Seja $F \in \mathbf{D}_n$ não degenerada, ponto extremo de $\mathbf{D}_n$ (logo, honesta) e suponhamos, sem perda de generalidade, que $F(0) = 1$. Note-se que neste caso temos que

$$F^{(k)}(x) \geq 0, \ k = 1, ..., n.$$

Seja supp$F$ o suporte de F e $\beta \in$ supp$F$, tal que $0 < \int_{-\infty}^{\beta} F^{(1)}(x)\, dx < 1$. Defina-se

$$F_1(x) = \begin{cases} F(x) & \text{se} \quad x < \beta \\ \sum_{i=0}^{n} F^{(n-i)}(\beta) \frac{(x-\beta)^{n-i}}{(n-i)!} & \text{se} \quad \beta \leq x \leq 0 \\ F_1(0) = \sum_{i=1}^{n} F^{(n-i)}(\beta) \frac{(-\beta)^{n-i}}{(n-i)!} & \text{se} \quad 0 < x \end{cases}. \quad (20)$$

Provemos que F_1 e $F - F_1$ são elementos de $\mathbf{D}_n$. Ora,

$$F_1^{(k)}(x) = \begin{cases} F^{(k)}(x) & \text{se} \quad x < \beta \\ \sum_{i=0}^{n-k} F^{(n-i)}(\beta) \frac{(x-\beta)^{n-k-i}}{(n-k-i)!} & \text{se} \quad \beta \leq x < 0 \\ 0 & \text{se} \quad 0 < x \end{cases}, \quad (21)$$

que é sem dúvida, para $k = 1, ..., n$, uma função não negativa, e, para $\beta < x < 0$,

$$F^{(k)}(x) - F_1^{(k)}(x) = F^{(k)}(x) - \sum_{i=1}^{n-k} F^{(n-i)}(\beta) \frac{(x-\beta)^{n-k-i}}{(n-k-i)!}. \quad (22)$$

Para mostrar a não negatividade desta última função precisamos de recorrer a fórmula de Taylor que nos diz que uma função $(n+1)$ vezes continuamente diferenciável num intervalo fechado $[\beta, x]$ poder ser escrita na seguinte forma:

$$f(x) = \sum_{i=0}^{n} f^{(i)}(\beta)\frac{(x-\beta)^i}{i!} + \int_{\beta}^{x} f^{(n+1)}(t)\frac{(x-t)^n}{n!}dt$$

$$= \sum_{i=0}^{n} f^{(i)}(\beta)\frac{(x-\beta)^i}{i!} + f^{(n+1)}(\xi)\frac{(x-\beta)^{n+1}}{(n+1)!}$$

para algum $\xi \in [\beta, x]$. No nosso caso temos uma função F, n vezes continuamente diferenciável, pelo que $F^{(k)}$ é $(n-k)$ vezes continuamente diferenciável e

$$F^{(k)}(x) = \sum_{i=0}^{n-k-1} F^{(k+i)}(\beta)\frac{(x-\beta)^i}{i!} + F^{(n-k)}(\xi)\frac{(x-\beta)^{n-k}}{(n-k)!}$$

para algum $\xi \in [\beta, x]$ e $x \in [\beta, 0)$. Substituindo em (22) vem que

$$F^{(k)}(x) - F_1^{(k)}(x)$$

$$= F^{(k)}(x) - \sum_{i=1}^{n-k} F^{(n-i)}(\beta)\frac{(x-\beta)^{n-k-i}}{(n-k-i)!}$$

$$= \sum_{i=0}^{n-k-1} F^{(k+i)}(\beta)\frac{(x-\beta)^i}{i!} + F^{(n-k)}(\xi)\frac{(x-\beta)^{n-k}}{(n-k)!} - \sum_{i=1}^{n-k} F^{(n-i)}(\beta)\frac{(x-\beta)^{n-k-i}}{(n-k-i)!}$$

$$= \sum_{i=0}^{n-k-1} F^{(n-1-i)}(\beta) \frac{(x-\beta)^{n-k-1-i}}{(n-k-1-i)!} +$$

$$+ F^{(n-k)}(\xi) \frac{(x-\beta)^{n-k}}{(n-k)!} - \sum_{i=1}^{n-k} F^{(n-i)}(\beta) \frac{(x-\beta)^{n-k-i}}{(n-k-i)!}$$

$$= \sum_{i=1}^{n-k} \frac{F^{(n-i)}(\beta)(x-\beta)^{n-k-i}}{(n-k-i)!} + \frac{F^{(n-k)}(\xi)(x-\beta)^{n-k}}{(n-k)!} - \sum_{i=1}^{n-k} \frac{F^{(n-i)}(\beta)(x-\beta)^{n-k-i}}{(n-k-i)!}$$

$$= F^{(n-k)}(\xi) \frac{(x-\beta)^{n-k}}{(n-k)!} \geq 0,$$

como se pretendia.

Podemos então escrever F como a soma de dois elementos de D_n,

$$F = F_1 + (F - F_1).$$

Atendendo a que F é ponto extremo resulta que $F_1 = \lambda F$ para algum $\lambda \in [0,1]$, e que o valor de λ só poderá ser 1, pois, para $x \leq \beta$, $F = F_1$. A função de distribuição F é então, para $x \in [\beta, 0)$, um polinómio de grau n. Dada a arbitrariedade de β resulta que a expressão de F é dada por uma função polinomial de grau n, e o suporte de F é obrigatoriamente limitado, digamos por a. Podemos pois escrever

$$F(x) = \sum_{i=0}^{n} \left(\frac{x}{a}\right)^i f_i$$

para certos valores reais f_i, e $x \in [a, 0]$. De F ser extremo resulta que, para $k = 0, ..., n-1$, $F^k(a) = 0$, e, consequentemente, $f_i = \frac{n!}{i!(n-i)!}(-1)^i$, i.e.,

$$F(x) = \sum_{i=0}^{n} \left(-\frac{x}{a}\right)^i \frac{n!}{i!(n-i)!} = \left(1 - \frac{x}{a}\right)^n,$$

como se pretendia. ■

Vamos agora dar uma representação integral dos elementos de M_n^*, i.e., das funções monótonas de ordem n em $[0, +\infty]$.

Teorema 7

Uma função f definida em $[-\infty, +\infty]$ é monótona de ordem n em $[0, +\infty]$ sse $f(x) = 0$ para $x < 0$ e

$$f(x) = \int_0^{1/x} (1 - xt)^n \, dG(t), \quad x \geq 0$$

onde G é uma função limitada, contínua à direita, monótona, não decrescente e não negativa.

Demonstração

Não é difícil verificar que M_n^*, $n = 0, 1, \ldots$, é um cone convexo. Seja E o espaço de todas as funções reais com suporte em $[0, +\infty]$ cuja restrição a $(0, +\infty)$ é derivável n vezes, e considere-se a topologia induzida pela família contável de semi-normas

$$p_{m,k}(f) = \sup\left\{ \left| f^{(i)}(x) \right| : m^{-1} \leq x \leq m, 0 \leq i \leq k \right\}$$

$(k = 1, 2, \ldots, n; m = 1, 2, \ldots)$. Definamos o sub-conjunto de E,

$$k_n = \left\{ f \in M_n^* : \int_0^{+\infty} f(x) \, dx = 1 \right\}.$$

O conjunto k_n é uma base convexa e compacta para o cone M_n^* (cf. Pestana [53]).

Tendo em conta o que foi referido anteriormente, segue facilmente que os pontos extremos de k_n, $n = 0, 1, \ldots$, são

$$\begin{aligned} f_{n,0}(x) &= \delta_0(x)\,; \\ f_{n,a}(x) &= \frac{n}{a^n}\left(1 - \frac{x}{a}\right)^n \mathbf{I}_{(0,a)}(x)\,; \\ f_{n,+\infty}(x) &= \delta_\infty(x)\,. \end{aligned}$$

Consideremos a aplicação

$$\begin{array}{rccc} \varphi: & [0,+\infty] & \longrightarrow & k_n \\ & a & \longmapsto & \varphi(a) = f_{n,a} \end{array}.$$

A aplicação φ é obviamente contínua, quando consideramos a topologia induzida e, consequentemente o conjunto dos pontos extremos de k_n é $\text{ext}(k_n) = \varphi([0,+\infty])$ que, sendo a imagem de um conjunto compacto, é também compacto. Logo podemos fazer corresponder a cada $f \in k_n$ uma medida de Borel normada, P, suportada pelo conjunto dos pontos extremos de k_n (cf. Phelps [57]) tal que

$$L(f) = \int_{\text{ext}(k_n)} L\,dP$$

para cada funcional linear contínuo L definido em E. No caso particular para o qual $L_x(f) = f(x)$, $f \in E$, obtemos

$$f(x) = L_x(f) = \int_{\text{ext}(k_n)} L_x\,dP, \quad x \geq 0.$$

Defina-se μ em cada subconjunto boreliano B de $[0,+\infty]$ por

$$\mu(B) = P[\varphi(B)]\,.$$

Dado que $L_x[\varphi(a)] = f_{n,a}(x)$ temos que

$$\begin{aligned} f(x) &= \int_{[0,+\infty]} f_{n,a}(x)\, d\mu(a) \\ &= \int_{(0,+\infty)} f_{n,a}(x)\, d\mu(a) + \int_{\{0\}} f_{n,a}(x)\, d\mu(a) \\ &= \int_{(0,+\infty)} \left(1 - \frac{x}{a}\right)^n \mathrm{I}_{[0,a)}(x)\, d\mu(a) + f_{n,0}(x)\, \mu(\{0\}) \\ &= \int_x^{+\infty} \left(1 - \frac{x}{a}\right)^n d\mu(a) + f_{n,0}(x)\, \mu(\{0\}), \quad x \geq 0. \end{aligned}$$

■

Uma função monótona de ordem n em $(0, +\infty]$ pode ser expressa da forma seguinte

$$f(x) = \int_0^{1/x} (1 - xt)^n \, dG(t),$$

para valores de x reais positivos, onde G é uma medida.

Existe também uma representação integral dos elementos de $M_\infty^* = \bigcap_{n=1}^{\infty} M_n^*$, a classe das funções completamente monótonas.

Teorema 8

Uma condição necessária e suficiente para que f seja completamente monótona em $[0, +\infty)$ é que

$$f(x) = \int_0^{+\infty} \exp(-xt)\, dG(t),$$

onde G é limitada e monótona não decrescente, e o integral converge para todo o real x não negativo, i.e., f é a transformada de Laplace de G (cf. Feller [11]).

Demonstração

Este resultado torna-se de fácil demonstração se observarmos que $\exp(-kx)-(1-x)^{k-1}$ aproxima-se de zero uniformemente em $[0,1]$ quando $k \longrightarrow +\infty$, e se tivermos em mente o resultado do Lema 5. ■

O teorema de Bernstein referido (cf. Feller [11]) sobre a representação integral de funções completamente monótonas foi estabelecido seguindo estas linhas por Mattner [42] e por Pollard [58].

Funções de distribuição da classe D_n e funções características da classe C_n

Com base nos resultados da secção anterior temos a seguinte caracterização dos elementos da classe D_n.

Teorema 9

Uma f.d. F é um elemento de D_n sse F é a f.d. de uma variável aleatória $Y \stackrel{d}{=} \left(1-U^{1/n}\right) Y_n$, onde U denota a distribuição uniforme em $[0,1]$ e Y_n é uma v.a. independente de U.

Demonstração

Observemos que $U^{1/n}$ é uma variável beta de parâmetros n e 1; logo $\left(1-U^{1/n}\right)$ é uma variável Be$(1,n)$. Denotemos por $X_{1,n}$ a variável $\left(1-U^{1/n}\right)$. Ora,

$$f_{X_{1,n}Y}(x) = \int_{\mathbf{R}} f_{X_{1,n}}\left(\frac{x}{y}\right)\frac{1}{|y|}f_Y(y)\,dy$$

$$= \mathbf{I}_{\mathbf{R}_0^+}(x)\int_x^{+\infty} n\left(1-\tfrac{x}{y}\right)^{n-1}\tfrac{f_Y(y)}{y}dy + \mathbf{I}_{\mathbf{R}^-}(x)\int_{-\infty}^{x} n\left(1-\tfrac{x}{y}\right)^{n-1}\tfrac{f_Y(y)}{|y|}dy$$

$\left(t=\frac{|x|}{y}\right)$

$$= \mathbf{I}_{\mathbf{R}_0^+}(x)\int_0^1 \tfrac{n(1-t)^{n-1}}{t}f_Y\left(\tfrac{|x|}{t}\right)dt + \mathbf{I}_{\mathbf{R}^-}(x)\int_{-1}^{0} \tfrac{n(1+t)^{n-1}}{|t|}f_Y\left(\tfrac{|x|}{t}\right)dt$$

$$= \mathbf{I}_{\mathbf{R}_0^+}(x)\int_0^1 \tfrac{n(1-t)^{n-1}}{t}f_Y\left(\tfrac{|x|}{t}\right)dt + \mathbf{I}_{\mathbf{R}^-}(x)\int_{-1}^{0} \tfrac{n(1-|t|)^{n-1}}{|t|}f_Y\left(\tfrac{x}{|t|}\right)dt$$

$$= \int_0^1 \frac{n(1-t)^{n-1}}{t}f_Y\left(\frac{x}{t}\right)dt \overset{y=t/x}{=} \int_0^{1/x}\frac{n(1-xy)^{n-1}}{y}f_Y\left(\frac{1}{y}\right)dy$$

$$= \int_0^{1/x}(1-xy)^{n-1}\,dG(y),$$

onde $dG(y)=\frac{1}{y}f_Y\left(\frac{1}{y}\right)dy$. ■

Deste último teorema e do facto de D_1 ser formado por todas as funções de distribuição unimodais com vértice 0 (cf. Definição 12), temos o seguinte resultado obtido por Khintchine em 1938:

Corolário 2

Uma f.d. F é unimodal com vértice 0 sse a função característica correspondente φ puder ser representada na seguinte forma

$$\varphi(t) = \frac{1}{t}\int_0^t \theta(x)\,dx, \quad t \in \mathbf{R}, \tag{23}$$

onde θ é uma função característica.

Para proceder à prova do corolário basta recordar que a função característica do produto de duas variáveis aleatórias independentes, digamos X e Y, é dada por

$$\varphi_{XY}(t) = \int_{\mathbf{R}} \varphi_X(yx)\,dF_Y(x)\,.$$

Recordando ainda que uma variável aleatória de distribuição $B(p,q)$ tem função característica $\varphi(t) = M(p, p+q; it)$, onde $M(\alpha;\beta;z)$ denota a função hipergeométrica confluente de parâmetros α e β, e argumento z,

$$M(\alpha,\beta;z) = \sum_{k=0}^{\infty} \frac{\Gamma(\alpha+k)}{\Gamma(\alpha)}\frac{\Gamma(\beta)}{\Gamma(\beta+k)}\frac{z^k}{k!},$$

podemos reescrever o resultado de Khintchine na forma

$$\varphi(t) = \int_{\mathbf{R}} M(1, 1+1; itx)\,dG(x)\,, \quad t \in \mathbf{R}, \tag{24}$$

onde G é uma função de distribuição. Mais geralmente:

Corolário 3

Uma f.d. F é um elemento de D_n sse a função caracterísica correspondente φ é da forma

$$\varphi(t) = \int_{\mathbf{R}} M(1; 1+n; itx)\, dG(x), \quad t \in \mathbf{R},$$

onde G é uma função de distribuição.

Como resultado particular, para $n = 2$, temos o teorema de Sakovič's sobre funções características de funções de densidade convexas

$$\varphi(t) = \frac{2}{t^2} \int_{\mathbf{R}} \frac{1 + itx - \exp(itx)}{x^2} dG(x), \quad t \in \mathbf{R},$$

onde G é uma função de distribuição. É interessante apontar as semelhanças com a representação das variáveis aleatórias infinitamente divisíveis, cf. Steutel [70]. Observe-se também que, Johansen [29], a representação de Lévy das funções características infinitamente divisíveis não é mais do que uma representação integral de Choquet.

Definimos agora a classe C_n e caracterizamos os seus elementos usando uma representação integral.

Definição 9

Dizemos que uma função característica φ pertence à classe C_n, $n = 0, 1, \ldots$, sse φ corresponde a uma v.a. simétrica e $\varphi \mathbf{I}_{[0,+\infty)}$ é monótona de ordem n em $[0, +\infty]$.

Observe-se que

1. Se $\varphi \in C_n$, $n = 0, 1, \ldots$, φ é uma função real e par.
2. Para $n \geq 1$, $C_n \subseteq C_{n-1}$.
3. A função característica φ pertence à classe C_n, $n = 0, 1, \ldots$, se, e só se, puder ser representada para todo o real t positivo na forma
$$\varphi(t) = \varphi(-t) = \int_0^{1/t} (1 - xt)^n \, dG(x),$$
onde G é uma f.d. tal que $G(0) = 0$.

É sempre complicado decidir se uma dada função φ é ou não uma função característica. As caracterizações de Bochner e de Cramér não são de fácil manuseamento. No teorema de Bochner exige-se, além de que $\varphi(0) = 1$, (i.e. que a v.a. em questão seja uma variável honesta), que a função seja definida não negativa. Quanto ao critério de Cramér, o integral
$$\int_0^{\alpha} \int_0^{\alpha} \varphi(t - u) \exp[ix(t - u)] \, dt du$$
tem de ser real e não negativo, para todo o número real x e para todo o real positivo α. A condição de Pólya dá-nos condições suficientes muito simples para que φ, real e par, seja uma função característica, nomeadamente a convexidade de $\varphi(t)$, $t > 0$.

Se na definição acima exposta tomarmos $n = 1$ obtemos as distribuições de Pólya e a conhecida representação integral para as funções características de Pólya surgem de imediato: φ é uma função característica de Pólya se, e só se,

$$\varphi(t) = \int_0^{1/|t|} (1 - x|t|)\, dG\,(x)\,,$$

onde G é uma função distribuição tal que $G(0) = 0$.

Infelizmente a utilidade destes modelos é limitada, pois as funções características da classe de Pólya não têm valor médio finito. No entanto existem alguns resultados recentes, Gneiting [18] e Gneiting [19], dando critérios para funções características associadas a variáveis aleatórias com valor médio e variância finitos.

Transformada inversa

Tendo em mente os resultados do Lema 5 e usando a expansão em série de Taylor dos elementos de M_n^* não é difícil provar o seguinte resultado.

Teorema 10

Se $f\,(x) = \int_0^{1/x} (1 - tx)^n\, dG\,(t)\,, x \geq 0$, *com* G *limitada e monótona não decrescente,* $G(0) = 0$, *então a função* G *é unicamente determinada nos seus pontos de continuidade, e nestes pontos*

$$G\,(t) = \sum_{k=0}^{n} \frac{(-1)^k}{k!} \left(\frac{1}{t}\right)^k f^{(k)}\left(\frac{1}{t}\right), \quad t > 0.$$

Caracterização alternativa das funções monótonas de ordem n, e extensões a ordem fraccionária

Dos parágrafos anteriores sabemos já que uma função monótona de ordem n em $(0, +\infty]$ pode ser representada da seguinte forma

$$f(x) = \int_0^{1/x} (1 - xt)^n \, dG(t), \quad x > 0,$$

onde G é não negativa e monótona não decrescente. A aplicação G, conhecido o valor de $G(0)$, pode ser facilmente obtida como uma transformação de f pois

$$x^n f\left(\frac{1}{x}\right) = \int_0^x (x - t)^n \, dG(t)$$

e

$$\frac{d^n}{dx^n}\left[x^n f\left(\frac{1}{x}\right)\right] = n!\,[G(x) - G(0)].$$

Por outro lado, se assumirmos que f é tal que $\frac{d^n}{dx^n}\left[x^n f\left(\frac{1}{x}\right)\right]$ é não negativa e monótona não decrescente, e definirmos G^* em $[0, +\infty]$ tal que

$$G^*(x) = \frac{1}{n!}\frac{d^n}{dx^n}\left[x^n f\left(\frac{1}{x}\right)\right], \quad x > 0$$

e $0 \leq G^*(0) \leq \lim\limits_{x \to 0^+} G^*(x)$, então

$$h(x) = \int_0^{1/x} (1 - xt)^n \, dG^*(t), \quad x > 0$$

é uma função monótona de ordem n e, assumindo que

$$\lim_{x \to +\infty} f(x) \text{ e } \lim_{x \to +\infty} h(x)$$

são finitos, $f = h$, desde que $G^*(0) = G(0)$.

Podemos então afirmar que uma função f é monótona de ordem n em $(0, +\infty]$ sse $0 \leq \lim_{x \to +\infty} f(x) < +\infty$, $\frac{d^n}{dx^n}\left[x^n f\left(\frac{1}{x}\right)\right]$ é não negativa e monótona não decrescente em $(0, +\infty]$. Esta definição alternativa tem a vantagem de nos permitir a extensão do conceito através do uso de derivadas fraccionárias. Procedamos então à substituição do parâmetro n por um parâmetro não negativo ν (em vez de M_n e M_n^* teremos M_ν e M_ν^*). Este operador de derivação é definido do seguinte modo (cf. Oldham e Spanier [51]):

$$D^\nu [f(x)] = I^{-\nu} [f(x)] = \frac{d^n}{dx^n} I^{n-\nu} [f(x)],$$

onde n é um inteiro maior do que ν, e, para $\beta > 0$

$$I^\beta [f(x)] = \frac{x^\beta}{\Gamma[\beta]} \int_0^1 (1-t)^{\beta-1} f(tx)\, dt.$$

Definição 10

Uma função f diz-se monótona de ordem ν $(\nu \geq 0)$ em $(0, +\infty]$ se, e só se,

1. $0 \leq \lim_{x \to +\infty} f(x) < +\infty$;
2. $D^\nu \left[x^\nu f\left(\frac{1}{x}\right)\right]$ *é não negativa e monótona não decrescente.*

A classe das funções monótonas de ordem ν $(\nu \geq 0)$ em $(0, +\infty]$ será denotada por M_ν.

A representação integral dos elementos de M_ν pode ser facilmente deduzida:

Teorema 11

Uma função f é monótona de ordem ν ($\nu \geq 0$) em $(0,+\infty]$ se, e só se,

$$f(x) = \int_0^{1/x} (1 - xt)^{\nu} \, dG(t), \quad x > 0,$$

onde G é uma função não negativa e monótona não decrescente em $(0,+\infty]$.

Também aqui podemos definir uma classe de distribuições especiais:

Definição 11

Sejam F uma f.d. e $A = \left(\alpha\left(\hat{F}\right), \omega\left(\hat{F}\right)\right)$, o interior do suporte da f.d. definida por

$$\hat{F}(x) = \frac{F(x) - \lim_{x \to -\infty} F(x)}{\lim_{x \to +\infty} F(x) - \lim_{x \to -\infty} F(x)}.$$

Suponhamos que F é diferenciável em $A - \{0\}$ e consideremos $f(x) = \frac{d}{dx}F(x)$ se $x \in A - \{0\}$, $f(x) = 0$ se $x \notin A$. Dizemos que a f.d. F é um elemento de $\mathbf{D}_{\nu}$, $\nu \geq 1$, se e só se as funções definidas por

$$f_1(x) = f(x)\, \mathbf{I}_{(0,+\infty)}(x) \text{ e } f_2(x) = f(-x) \mathbf{I}_{(0,+\infty)}(x)$$

são monótonas de ordem $\nu - 1$ em $(0,+\infty]$.

Como consequência do Teorema 11 temos o seguinte corolário:

Corolário 4

Uma função de distribuição F pertence à classe D_ν, $\nu \geq 1$, se, e só se, for a f.d. da variável aleatória

$$X \stackrel{d}{=} \left(1 - U^{1/\nu}\right) Y_\nu,$$

onde U denota a distribuição uniforme em $[0,1]$ e Y_ν é independente de U.

E entre as classes temos a seguinte relação:

Teorema 12

Se $0 < \mu < \nu$, então $D_{\nu+1} \subsetneq D_{\mu+1}$.

Demonstração

Usaremos a transformada de Mellin nesta prova. A transformada de Mellin de uma v.a. X é definida como sendo

$$\mathcal{M}_X(s) = \int_0^{+\infty} x^s dF_X(x)$$

onde F_X é a f.d. associada a X. Se X é uma v.a. beta de parâmetros p e q não é difícil mostrar que

$$\mathcal{M}_X(s) = \mathcal{M}_{p,q}(s) = \frac{\mathrm{B}(p+s,q)}{\mathrm{B}(p,q)}, \ \mathrm{Re}(s) > -p.$$

Desta igualdade e do facto da transformada de Mellin do produto de variáveis aleatórias, não negativas e independentes, ser o produto da transformada de Mellin das variáveis envolvidas, a seguinte identidade é válida

$$\mathcal{M}_{1,\nu+1}(s) = \mathcal{M}_{1,\mu+1}(s)\,\mathcal{M}_{\mu+2,\nu-\mu}(s), \quad \mathrm{Re}(s) > -1. \tag{25}$$

Do corolário anterior sabemos que $F \in D_{\nu+1}$ se, e só se, a f.d. de

$$X \stackrel{d}{=} \left(1 - U^{\frac{1}{\nu+1}}\right) Y_{\nu+1},$$

com $Y_{\nu+1}$ independente de U, a distribuição uniforme em $[0,1]$. Como mencionamos anteriormente, a variável $1 - U^{\frac{1}{\nu+1}}$ é uma variável beta de parâmetros 1 e $\nu + 1$. Usando (25) obtemos que F é a f.d. da variável aleatória

$$X \stackrel{d}{=} \left(1 - U^{\frac{1}{\mu+1}}\right) Y_{\mu+1},$$

onde $Y_{\mu+1} = Z_{\mu+2,\nu-\mu} Y_{\nu+1}$, donde resulta que $F \in D_{\mu+1}$. ■

Corolário 5

Se $0 < \mu < \nu$, *então* $M_{\nu+1} \subseteq M_{\mu+1}$.

3.4.3 A classe de Mejzler

A monotonicidade generalizada tem também um papel importante nas conhecidas classes de Mejzler e de Khintchine, abordadas nas secções 2.2 e 2.3. Recordemos o caso particular da classe de Mejzler, definida pelos limites não degenerados de sucessões do tipo

$$\frac{X_{n:n} - b_n}{a_n}, \ (a_n > 0)$$

obtidas pela normalização de $X_{n:n} = \max_{1 \leq k \leq n} X_k$, onde $\{X_k\}_{k \in \mathbf{N}_1}$ é uma sucessão de variáveis aleatórias independentes (não necessariamente idênticas). Esta classe, designada por classe M, foi caracterizada por Mejzler (cf. Teorema 2): dada uma f.d. F e a função

$$G(x) = \begin{cases} G_1(x) = F(\alpha(F) + \exp(x)), & \text{se } \alpha(F) > -\infty \\ G_2(x) = F(\omega(F) - \exp(-x)), & \text{se } \omega(F) < +\infty \\ G_3(x) = F(x), & \text{caso contrário} \end{cases},$$

sabemos que F pertence a M sse G_i for log-côncava para algum $i = 1, 2, 3$.

Note-se ainda que $F \in M$ se, e só se,

$$G_i(x) = G_i(a + x) F_a(x) \tag{26}$$

para algum $i = 1, 2, 3$, $\forall a < 0$ se $i = 1, \forall a > 0$ caso contrário, sendo F_a uma função não negativa e monótona não decrescente.

A classe de Mejzler foi "refinada" no seguinte sentido: designando por $\mathcal{M}_{-1}$ o conjunto de todas as funções de distribuição e por $\mathcal{M}_0$ a classe M, definiu-se iterativamente as classes $\mathcal{M}_r, r = 1, 2, \ldots$, do seguinte modo: $F \in \mathcal{M}_r$ se, e só se, G_i, para algum $i = 1, 2, 3$, satisfaz (26) com F_a proporcional a uma f.d. de $\mathcal{M}_{r-1}$. À intersecção destas classes deu-se a designação de $\mathcal{M}_\infty$ $\left(\mathcal{M}_\infty = \bigcap_{r=1}^{\infty} \mathcal{M}_r\right)$. As funções de distribuição da classe $\mathcal{M}_r$ têm propriedades de monotonicidade surpreendentes: $F \in \mathcal{M}_r$ se, e só se,

$$F(x) = \exp[-k(x)],$$

com k uma função monótona de ordem $r + 1$; em particular

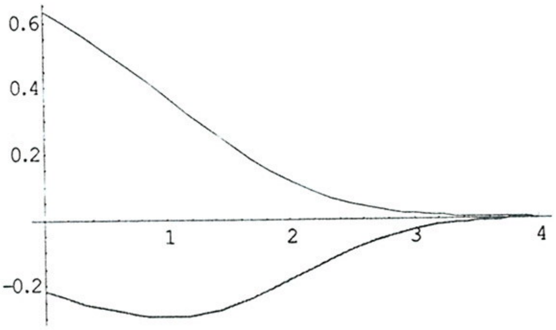
$$F \in \mathcal{M}_{+\infty} \text{ sse } F(x) = \exp[-k(x)]$$

com k uma função completamente monótona (cf. Pestana [54]).

$\mathcal{M}_{+\infty}$ é a mais pequena classe fechada para produtos e limites pontuais, contendo todas as funções de distribuição de valores extremos. Foram estas propriedades que levaram Graça Martins e Pestana [23] a conjecturar que $\mathcal{M}_{\infty}$ é a classe apropriada para modelação genérica de máximos, e que as funções de distribuição usadas por utilizadores fora do esquema clássico (como por exemplo as log-gamas e gaussianas usadas por hidrologistas e climatologistas) seriam elementos de $\mathcal{M}_{\infty}$. Sobre a segunda conjectura, infelizmente, no que refere a f.d. gaussiana, que é de facto um elemento da classe de Mejzler, esta distribuição não é um elemento de $\mathcal{M}_{\infty}$. De facto, se designarmos por $F_{\mu,\sigma}$ a distribuição gaussiana de parâmetros $\mu \in \mathbf{R}$ (valor médio) e $\sigma > 0$ (desvio padrão) temos que (para $t \in \mathbf{R}$)

$$\begin{aligned} F_{\mu,\sigma}(t) &= \int_{-\infty}^{t} \frac{1}{\sqrt{2\pi}\sigma} \exp\left[-\frac{(x-\mu)^2}{2\sigma^2}\right] dx \\ &= \exp\left[-k(t)\right] \end{aligned}$$

com $k_{\mu,\sigma}(t) = \ln\left(\int_{-\infty}^{t} \frac{1}{\sqrt{2\pi}\sigma} \exp\left[-\frac{(x-\mu)^2}{2\sigma^2}\right] dx\right)^{-1}$ que é convexa. A função $k_{\mu,\sigma}$ não é completamente monótona. Anote-se, por exemplo, a representação gráfica das derivadas de $k_{0,1}$, $k_{0,1}^{(2)}$ e $k_{0,1}^{(3)}$:



$k_{0,1}^{(2)}$(a preto) e $k_{0,1}^{(3)}$ (a cinzento)

3.5 Uma nota sobre unimodalidade

O conceito intuitivo de v.a. unimodal foi já referido sem que se tenha procedido a sua definição formal.

Definição 12

Diz-se que uma variável aleatória real X, ou a sua distribuição F, é unimodal em torno de uma moda (ou vértice) v se F é convexa em $(-\infty, v)$ e côncava em $(v, +\infty)$ (cf. Dharmadhikari e Joag-Dev [8]).

A unimodalidade foi caracterizada por Khintchine, que identificou cada variável aleatória unimodal em zero ao produto de duas variáveis aleatórias independentes, sendo uma delas uma variável aleatória uniforme em $[0, 1]$.

Sem perda de generalidade tomaremos $v = 0$, e, por comodidade usaremos a expressão "unimodal" para abreviar "unimodal em torno de zero" e "unimodal em ν" para abreviar "unimodal em torno de ν".

No parágrafo anterior constatamos que a classe D_1 é formada por todas as funções de distribuição unimodais com vértice 0, pelo que a unimodalidade pode ser vista como o caso particular mais simples da monotonia generalizada. A convolução de distribuições é uma das ferramentas essenciais em Probabilidade e Estatística. A convolução de distribuições unimodais foi abordada por Gnedenko e Kolmogorov [17] na sua notável obra sobre as distribuições limite de somas de v.a. independents. Gnedenko e Kolmogorov afirmam nesta mesma obra que a convolução de f.d.'s unimodais continua a ser uma f.d. unimodal, lapso que Chung, o tradutor da obra, corrigiu. Sato [67] construíu mesmo duas distribuições unimodais cuja convolução pode ter o número que se queira de modas. Temos assim, como consequência, que a classe D_1,anteriormente definida, não é fechada para o operador convolução. Sabe-se apenas que

Proposição 6

Se as distribuições originais forem, além de unimodais, simétricas, a convolução é também unimodal (Dharmadhikari et al. [8]).

Quando a convolução de uma distribuição F com qualquer uma outra distribuição unimodal (num dado ponto ν) é unimodal (não necessariamente em torno do mesmo ponto) dizemos que F é **fortemente unimodal**. O seguinte teorema apresenta-nos uma caracterização destas distribuições:

Teorema 13

(caracterização de Ibragimov) Uma distribução G é fortemente unimodal se, e só se, G é contínua e a sua densidade g é log-côncava.

Esta caracterização permite-nos identificar algumas distribuições fortemente unimodais entre as mais conhecidas:

- a normal, $N(\mu, \sigma)$,

- a uniforme em (a, b), $U_{(a,b)}$,
- a gama com parâmetro de forma $p \geq 1$, $G(p, \beta)$,
- e a beta de parâmetros (p, q), com $p \geq 1$ e $q \geq 1$, $B(p, q)$

são fortemente unimodais.

Recorde-se que esta última família de distribuições, mais precisamente, distribuições das variáveis do tipo $aB(1, n)$, constituem os pontos extremos não degenerados das classes D_n (Teorema 6).

3.5.1 A unimodalidade forte e a transformada beta

Considere-se uma variável aleatória $X_{1,n} \frown B(1, n)$, $(n \geq 1)$. Dada uma outra v.a. Y, absolutamente contínua, definimos a **transformada beta** $(1, n)$ **de Y** como sendo a v.a. $X_{1,n}Y$, cuja função densidade de probabilidade é

$$f_{X_{1,n}Y}(x) = \int_0^{1/x} (1 - xy)^{n-1}\, dG(y)$$

onde $dG(y) = \frac{1}{y} f_Y\left(\frac{1}{y}\right) dy$.

A transformada beta de uma v.a. é uma variável unimodal, pois estas transformadas formam precisamente a classe D_n e, tendo em conta que $D_n \subseteq D_1$, resulta imediatamente que $X_{1,n}Y$ é unimodal (independentemente da unimodalidade de Y).

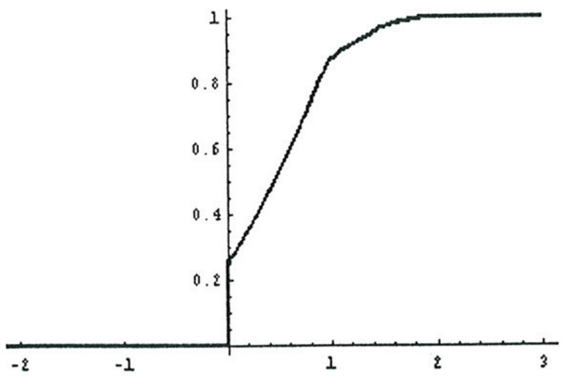
Quanto à conjectura da unimodalidade forte de $X_{1,n}Y$ esta não é verificada. Basta pensar na soma de duas variáveis uniformes em $(0,1)$: são variáveis unimodais em 0 e a variável soma é unimodal em 1, e não em 0. Se atendermos ao exemplo de Dharmadhikari e Joag-Dev [8]: a função de distribuição F definida por

$$F(x) = \begin{cases} 0, & \text{se } x < 0 \\ \frac{x+1}{2}, & \text{se } 0 \leq x < 1 \\ 1, & \text{se } x \geq 1 \end{cases}.$$

pertence à classe D_n, mas

$$F * F(x) = \begin{cases} 0, & \text{se } x < 0 \\ \frac{x^2+4x+2}{8}, & \text{se } 0 \leq x < 1 \\ \frac{4-x^2+4x}{8}, & \text{se } 1 \leq x < 2 \\ 1, & \text{se } x \geq 2. \end{cases}$$

não é unimodal:



$F_X * F_X$

Atendendo a Proposição 6, poderíamos ser tentados a substituir na transformada, $B(1,n)$ por uma variável beta simétrica, transladada de modo a que a respectiva função densidade seja uma função par. Será esta um boa candidata a distribuição transformadora fortemente unimodal? A resposta, também aqui é negativa. De facto se tomarmos para Y a variável degenerada, por exemplo, em $a > 0$, e para variável transformadora uma variável $X \stackrel{d}{=} X_{2,2} - 1/2 \frown B(2,2) - 1/2$, obtemos uma variável $Z = aX$ de distribuição $F_Z(x) = F_X(x/a)$ onde X é uma v.a. com função densidade de probabilidade

$$\begin{aligned} f_X(x) &= f_{X_{2,2}}\left(x+\frac{1}{2}\right) = 6\left(x+\frac{1}{2}\right)\left(\frac{1}{2}-x\right)\mathbf{I}_{[0,1]}\left(x+\frac{1}{2}\right) \\ &= 6\left(\frac{1}{4}-x^2\right)\mathbf{I}_{\left[-\frac{1}{2},\frac{1}{2}\right]}(x). \end{aligned}$$

As betas são fortemente unimodais, mas o resultado não é necessariamente uma f.d. unimodal em zero. Consideremos a f.d. dada por

$$F_W(x) = [1-\exp(-x-1)]\,\mathbf{I}_{[0,+\infty)}(x).$$

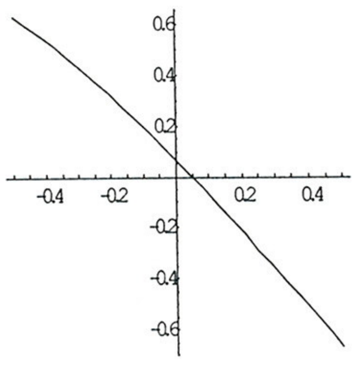
Temos que

$$\begin{aligned} F_W * F_X(x) &= \int_{\mathbf{R}} F_W(x-y)\, f_X(y)\, dy \\ &= 6\int_{-1/2}^{1/2} F_W(x-y)\left(\frac{1}{4}-y^2\right) dy \\ &= 6\left(\int_{-1/2}^{1/2} [1-\exp(y-x-1)]\left(\frac{1}{4}-y^2\right) dy\right)\mathbf{I}_{\left(\frac{1}{2},+\infty\right)}(x) \\ &\quad + 6\left(\int_{-1/2}^{x} [1-\exp(y-x-1)]\left(\frac{1}{4}-y^2\right) dy\right)\mathbf{I}_{\left[-\frac{1}{2},\frac{1}{2}\right)}(x). \end{aligned}$$

Após alguns cálculos resulta que

$$
\begin{aligned}
&F_W * F_X\left(x\right) \\
&= 6\left(\frac{1}{6} + e^{-\frac{1}{2}-x} - 3e^{-\frac{3}{2}-x}\right)\mathbf{I}_{\left(\frac{1}{2},+\infty\right)}\left(x\right) \\
&\quad + 6\left(\frac{x}{4} + \frac{1}{12} - 2\frac{x}{e} + \frac{x^2}{e} + \frac{7}{4e} - \frac{x^3}{3} - \frac{3}{e^{\frac{3}{2}+x}}\right)\mathbf{I}_{\left[-\frac{1}{2},\frac{1}{2}\right)}\left(x\right)
\end{aligned}
$$

que não é unimodal em zero como o demonstra o gráfico da segunda derivada de $F_W * F_X$ no intervalo $\left(-\frac{1}{2}, \frac{1}{2}\right)$, definida pela expressão $2e^{-1} - 2x - 3e^{-\frac{3}{2}-x}$:



A segunda derivada de $F_W * F_X$ em $\left(-\frac{1}{2}, \frac{1}{2}\right)$.

4 O futuro

I keep the subject before me and wait till the first dawnings open slowly, by little and little, into a full and clear light.

Isaac Newton

É possível apontar variados rumos que este trabalho pode vir a prosseguir. As variáveis discretas e mistas não foram abordadas. Os possíveis testes de simetria merecem mais algum estudo, bem como as extensões da classe de Mejzler; a atracção parcial no caso da normalização potência está ainda por desbravar, tal como, em alguns assuntos, a extensão à situação multivariada.

Não é pois um trabalho acabado – existirá, em ciência, um trabalho acabado? As nossas responsabilidades para com as instituições e para com os nossos colegas levaram-nos a interromper, neste âmbito de doutoramento, esta investigação.

Bibliografia

[1] Abramowitz, M. e Stegun, I. (1974), *Handbook of Mathematical Functions,* Dover.

[2] Araújo, A. e Giné, E. (1980), *The Central Limit Theorem for Real and Banach Valued Random Variables*, John Wiley & Sons, Inc.

[3] Balakrishnan, N. e Rao, C. (eds.) (1998), *Order Statistics: Theory & Methods* (Handbook of Statistics **16**), North Holland.

[4] Bingham, N., Goldie, C. e Teugels, J. (1989), *Regular Variation (Encyclopedia of Mathematics and Its Applications* **27***)*, Cambridge University Press.

[5] Brilhante, M. F. A. (1999), *Inferência Estatística em Modelos Não Gaussianos com Recurso a Spacings e Outras Funções de Estatísticas Ordinais*, Tese de Doutoramento, Universidade dos Açores.

[6] Choquet, G. (1969), *Lectures on Analysis, II – Representation Theory*, W. A. Benjamin, Inc.

[7] Darling, D. A. (1952), The Influence of the Maximum Term in the Addition of Independent Random Variables, *Transactions of American Mathematical Society* **73**, 95-107.

[8] Dharmadhikari, S. e Joag-Dev (1988), *Unimodality, Convexity, and Applications*, Academic Press.

[9] Efron (1969), Student's t-Test Under Symmetry Conditions, *JASA* **64**, 1278-1302.

[10] Embrechts, P., Klüppelberg, e Mikosch, T. (1999), *Modelling Extremal Events for Insurance and Finance* (Applications of Mathematics. Stochastic Modelling and Applied Probability **33**), Springer.

[11] Feller, W. (1971), *An Introduction to Probability Theory and Its Applications, II*, John Wiley & Sons.

[12] Fisher, R. A e Tippett, L. (1928), Limiting forms of the frequency distributions of the largest or smallest member of a sample, *Proc. Cabridge Philos. Soc.* **24**, 180-190.

[13] Fréchet, M. (1927), Sur la Loi de Probabilité de l'Écart Maximum, *Ann. de la Soc. Polonaise de Math. (Cracow)* **6**, 93-116.

[14] Galambos, J. (1987), *The Asymptotic Theory of Extreme Order Statistics*, 2^aEd., Robert E Krieger.

[15] Gnedenko, B. V. (1943) Sur la distribution limite du terme maximum d'une série aléatoire, *Ann. Math.* **44**, 423-453.

[16] Gnedenko, B. V. (1989), *The Theory of Probability and the Elements of Statistics*, Chelsea Publishing Company.

[17] Gnedenko, B. V. e Kolmogorov, A. (1954), *Limit Distributions for Sums of Independent Random Variables*, Addison-Wesley.

[18] Gneiting, T. (1999), Radial Positive Definite Functions Generated by Euclid's Hat, *Journal of Multivariate Analysis* **69**, 88-119.

[19] Gneiting, T. (2000), Kuttner's Problem and a Pólya Type Criterion for Characteristic Functions, *Proceedings of the American Mathematical Society* **128**, 1721-1728.

[20] Gomes, M. I. e Pestana, D. D. (1985), Domains of Attraction and Penultimate Behavior, *Abs. 16th European Meeting of Statisticians,* 202-203.

[21] Gomes, M. I. e Pestana, D. D. (1987), Non-Standard Domains of Attraction and Rates of Convergence, *in New Perspectives in Theoretical and Applied Statistics* (Puri, M. et al (eds.)), John Wiley & Sons, 467-477.

[22] Graça Martins, M. E. e Pestana, D. D. (1985), The Extremal Limit Theorem – Extensions, *V Pannonian Symp. Math. Statist.*, Viségrad.

[23] Graça Martins, M. E. e Pestana, D. D. (1987), Nonstable limit laws in extreme value theory, *in New Perspectives in Theoretical and Applied Statistics* (Puri, M. et al (eds.)), John Wiley & Sons, 449-458.

[24] Green, R. (1976), Partial Attraction of Maxima, *J. Appl. Prob.* **13**, 159-163.

[25] Grinevich, I. V. (1993), Domains of Attraction of the max-semistable laws under linear and power normalizations, *Theory Probab. Appl.* **38**, 640-650.

[26] Haan, L. de (1970), *On regular variation and its application to the weak convergence of sample extremes,* Tese de Doutoramento, Universidade de Amsterdão/Mathematical Centre tract **32**.

[27] Hall, P. e Heyde, C. C. (1980), *Martingale Limit Theory and Its Application,* Academic Press.

[28] Hirschman, I. I. and Widder, D. V. (1955), *The Convolution Transform,* Princeton University Press, Princeton.

[29] Johansen, S. (1966), An Application of Extreme Point Methods to the Representation of Infinitely Divisible Distributions, *Z. Wahrscheinlichkeitstheorie Verw. Geb.* **5**, 304-316.

[30] Johnson, N. e Kotz, S. (1969), *Distributions in Statistics: Continuous Univariate Distributions,* Vols. I e II, Houghton Mifflin Company.

[31] Jurek, Z. J. (1981), The Problem of B. V. Gnedenko on Banach Spaces, *Bull. Acad. Polonaise Sc., Ser. Sc. Math.* **29**, 399-407.

[32] Keilson, J. e Steutel, W. (1974), Mixtures of Distributions, Moment Inequalities and Measures of Exponentiality and Normality, *The Annals of Probability* **2**, 112-130.

[33] Kendall, D. H. e Harding, E. F. (1973), *Stochastic Analysis,* Wiley.

[34] Koroljuk, V. S. e Zolotarev, V. M. (1961), On a hypothesis proposed by B. V. Gnedenko, *Teor. Verojatnost. i Primenen* **6**, 469-474.

[35] Laha, R. e Rohatgi, V. (1979), *Probability Theory,* Wiley.

[36] Lamperti, J. (1996), *Probability: A Survey of the Mathematical Theory,* John Wiley & Sons.

[37] Le Cam, L. (1986), The Central Limit around 1935, *Statistical Science* **1**, 78-96.

[38] Lévy, P. (1962) Extensions d'un Théorème de D. Dugué et M. Girault, *Z. für Wahrscheinlichkeitstheorie Verw. Geb.* **1**, 159-173.

[39] Linnik, Ju. V. e Ostrovskii (1977), *Decomposition of Random Variables and Vectors*, Translations of Mathematical Monographs, American Mathematical Society.

[40] Loéve, M. (1973), Paul Lévy 1886-1971, *The Annals of Probability* **1**, 1-18.

[41] Logan, B. F., Mallows, C. L., Rice, S. O. e Shepp, L. A. (1973), Limit Distributions of Self-Normalized Sums, *The Annals of Probability* **1**, 788-809.

[42] Mattner, L. (1993), Bernstein's Theorem, Inversion Formula of Post and Widder, and the Uniqueness Theorem for Laplace Transforms, *Expositiones Mathematicae* **11**, 137-140.

[43] Mejzler, D. G. (1956), On the problem of the limit distribution for the maximal term of a variational series, *L'vov Pol. Inst. Naucn Zp.* **38**, 90-109.

[44] Mendonça, S. (1997), *Fugindo à Gaussiana: Modelos Lineares Generalizados & Estimação Robusta de Localização e de Escala*, Tese de Mestrado, Universidade de Lisboa.

[45] Mendonça, S. (1998), A note on random normalization, *Notas e Comunicações do CEAUL* **23/98**, Universidade de Lisboa.

[46] Mendonça, S. (1999a), Convergence of Classes and Limit Distributions, *Extreme Values and Additive Laws*, 64-67, Lisboa.

[47] Mendonça, S. (1999b), A Note on Random Normalizations, *Proceedings of the 11th European Young Statistitians Meeting,* Marly-le-Roi.

[48] Mendonça, S. (2000), On Sums and Extremes of Random Variables, Proceedings of the WDS2000, *WDS'00 Proceedings of Contributed Papers, Part I - Mathematics,* Charles University, Praga.

[49] Mises, R. de [von Mises, R.] (1936), La distribution de la plus grande de n valeurs, Reprinted in *Selected Papers II* . Amer. Math. Soc., Providence, R. I., 1954, 271-294.

[50] Mohan, N. R. e Ravi, S. (1992), Max Domains of Attraction of Univariate and Multivariate p-Max Stable Laws, *Theory Probab. Appl.* **37**, 632-643.

[51] Oldham, K. B. e Spanier, J. (1974), *The Fractional Calculus*, Academic Press.

[52] Olshen, R. A. e Savage, L. J. (1970), A Generalized Unimodality, *J. Appl. Probab.* **7**, 21-34.

[53] Pestana, D. (1978), *Some contributions to unimodality, infinite divisibility, and related topics*, Tese de Doutoramento, Universidade de Sheffield.

[54] Pestana, D. (1984), Urbanik's L_r Classes, Mejzler's M_r Classes, and Higher Order Monotone Functions, *Actas III, Colóquio de Estatística e Investigação Operacional,* Lagos.

[55] Pestana, D. e Mendonça, S. (1999), Higher Order Monotone Functions and Probability Theory, *Notas e Comunicações do CEAUL* **13/98**, Universidade de Lisboa.

[56] Pestana, D. e Mendonça, S. (2000), A Probabilistic Approach to Higher Order Convexity, aceite para publicação em *Lecture Notes in Economics and Mathematical Systems*, Springer Verlag, 2000.

[57] Phelps, R. R. (1966) *Lectures on Choquet's Theorem*, van Nostrand, New Jersey.

[58] Pollard, H. (1944), The Bernstein-Widder Theorem on Completely Monotonic Functions, *Duke Mathematical Journal* **11**, 427-430.

[59] Pyke, R. (1965), Spacings, *J. R. Statist. Soc. B* **27**, 395-436, Discussion: 437-49.

[60] Resnick, S. (1971), Tail Equivalence and its Applications, *Journal of Applied Probability* **8**, 135-136.

[61] Resnick, S. (1972), Products of Distribution Functions Attracted to Extreme Value Laws, *J. Appl. Prob.* **8**, 781-793.

[62] Resnick, S. (1986), Point Processes. Regular Variation and Weak Convergence, *Adv. Appl. Prob.* **18**, 66-138.

[63] Resnick, S. (1987), *Extreme Values, Regular Variation, and Point Processes* (Applied Probability **4**), Springer-Verlag.

[64] Roberts, A. W. e Varberg, D. E. (1973), *Convex Functions*, Academic Press.

[65] Rohatgi, V. K. (1976), *An Introduction to Probability Theory and Mathematical Statistics*, John Wiley & Sons.

[66] Sakovič, D. N. (1972), On the Characteristic Functions of Concave Distributions, *Transl. in Theor. Prob. and Math. Statist.* **6** (1975), 103-108.

[67] Sato, K.-I. (1993), Convolution of Unimodal Distributions Can Produce Any Number of Modes, *The Annals of Probability* **21**, 1543-1549.

[68] Schoenberg, I. J. (1941), On Integral Representations of Completely Monotone and Related Functions (abstract), *Bulletin of the American Mathematical Society* **47**, 208.

[69] Seneta, E. (1976), *Regularly Varying Functions* (Lecture Notes in Mathematics **508**), Springer-Verlag.

[70] Steutel, F. W. (1968), A Class of Infinitely Divisible Mixtures, *Ann. Math. Statist.* **39**, 1153-1157.

[71] Stigler, S. M. (1998), *The History of Statistics: The Measurement of Uncertainty before 1900*, Belknap-Harvard.

[72] Urbanik, K. (1973), Limit Laws for Sequences of Normed Sums Satisfying Some Stability Conditions, in Krishnaiah (ed.), *Multivariate Analysis III,* Academic Press, 225-237.

[73] Williamson, R. E. (1956), Multiply Monotone Functions and their Laplace Transforms, *Duke Mathematical Journal* **23**, 189-207.

[74] Zinger, A. A. (1965), On a class of limit distributions for normed sums of independent random variables, *Theor. Probab. Appl.* **10**, 607-620.

[75] Zolotarev, V. M. (1986), *One-Dimensional Stable Distributions* (Translations of Mathematical Monographs **65**), American Mathematical Society.