

UNIVERSIDADE DA MADEIRA

SÍLVIO FILIPE VELOSA

PADRÕES DE ALEATORIEDADE
NA
MODELAÇÃO DO ACASO



Agradecimentos

Gostaria de agradecer aos meus pais Virgínia e Manuel e à minha irmã Teresa todo o apoio que me deram nas pequenas grandes necessidades do dia a dia, e a paciência que tiveram comigo, durante este trabalho.

Agradeço todo o apoio institucional e manifesto o meu apreço pelo Departamento de Matemática e Engenharias da Universidade da Madeira, e gostaria de destacar, pelo apoio sempre presente, os Professores Rita Vasconcelos, Sandra Mendonça, José Carmo e Nuno Nunes. Este destaque não põe em dúvida o meu apreço por todos os meus colegas, e obriga-me a recordar o Professor Orlando Freitas, que, como aqueles que os deuses amam, morreu tão cedo.

Gostaria também de agradecer aos colegas do meu grupo de trabalho Sandra Mendonça, Tiago Marques, Rita Vasconcelos e Fernando Sequeira a sua disponibilidade, ajuda e inspiração constantes.

Ao meu excepcional orientador de doutoramento Prof. Dinis Pestana, um agradecimento especial pela disponibilidade, pelo estímulo, e pelas lições.

Conteúdo

1	Introdução	11
1.1	Comentários e Complementos	51
2	Classes de Panjer e Extensões	53
2.1	Introdução	53
2.2	Comentários e Complementos	74
2.2.1	Relação com as somas aleatórias de subordinadora geométrica ou de Poisson	74
2.2.2	Interpretação das classes \mathcal{C}_γ	77
2.2.3	Generalização dos resultados obtidos a $\gamma < 0$ e relação com a \mathcal{N} -divisibilidade infinita	79
2.2.4	Distribuições de Panjer generalizadas com suporte finito	80
2.2.5	Expressão das variáveis $N_{\alpha, \beta, \gamma}$ como somas aleatórias	82
3	Modelos de Contagem	91
3.1	<i>Così Fan Tutte</i> — algumas reflexões sobre aleatoriedade discreta	91
3.2	Comentários e Complementos	125
4	Transformadas Integrais	139
4.1	Introdução	139
4.2	Funções geradoras de probabilidade, transformada de Laplace e extensões	140

Resumo

A soma de variáveis aleatórias com número de parcelas é *aleatório*, para além do evidente interesse conceptual e teórico, tem larga ressonância na investigação do processo de risco e em processos de ramificação.

Reformulamos a teoria de Panjer (1981), que permite o cálculo iterativo do risco agregado, com o recurso a valores médios de uniformes, descrevendo uma extensão da classe de Panjer, e estudando em detalhe a equação funcional que a caracteriza.

Aplicamos essas ideias na caracterização de aleatoriedade discreta, exemplificando com o comportamento das fêmeas de pássaros que investem na promiscuidade de parceiros para garantir a diversidade genética da prole, tendo no entanto o cuidado de manter as aparências de fidelidade, para garantir a cooperação do parceiro no sucesso da ninhada.

Apresentamos as transformadas de Laplace e funções geradoras numa perspectiva que leva a uma introdução natural de transformadas de Pareto, cuja relevância exemplificamos.

Abstract

Randomly stopped sums, aside from their conceptual interest, are of the utmost importance on practical issues, such as the risk process and branching processes.

We extend Panjer's (1981) theory on the class of subordinators allowing an algorithmic iterative approach to compute the densities of aggregated claims, by investigating a general iterative expression using moments of uniform random variables, and the functional equation it implies for the corresponding probability generating functions.

Discrete randomness patterns are then elaborated, and we exemplify their fitness in the investigation of recently discovered female bird behaviour — sexual promiscuity contributing to genetic diversity of future generations, but disguised social behaviour, apparently with a single mate, who collaborates in raising what he believes to be his nestlings — showing that equilibrium patterns resulting from generalised discrete hyperbolic laws make sense.

We also introduce Pareto transforms as natural extensions of the Laplace transform, and show their relevance.

Capítulo 1

Introdução

Na presente dissertação exponho resultados publicados com outros membros da mesma equipa de investigação, Dinis Pestana, Sandra Mendonça e Tiago Marques, complementados por alguns resultados posteriores e comentários, quando adequado.

Escolhi portanto um formato menos tradicional para a dissertação, mas actualmente muito adoptado: a transcrição de artigos publicados. Em vez de uma introdução procurando cerzir os resultados, à guisa de introdução transcrevo dois artigos por nós publicados em 2003, no início da execução do plano doutoral, “Somas e Máximos de Variáveis Aleatórias Independentes — Classes de Leis Limites” e “Classes de Leis \mathcal{N} -Infinitamente Divisíveis” (este último com Pestana). De facto, eles expõem não só a conexão entre teoria das somas e teoria dos extremos de variáveis aleatórias, e as extensões recentes de “*Poisson stopped sums*” para “*geometric stopped sums*”, como questões mais genéricas sobre classes de leis limites em esquemas aleatórios.

Esse caminho conduziu-nos ao estudo de somas aleatórias (*stopped sums*), ao processo do risco e processos de ramificação. Os ricos resultados de Panjer (1981) — que foram um salto qualitativo importante no estudo do processo do risco — culminaram na publicação na *RevStat* (com Pestana) de “Extensions of Katz-Panjer Families of Discrete Distributions”, que se transcreve como parte integrante do Capítulo 2, em cuja secção de complementos e comentários é feito o estudo exaustivo das condições de existência de soluções para a equação às diferenças em que centrámos a nossa atenção.

Este estudo levou-me a reflectir a que ponto conhecia mal as leis discretas. A leitura de um trabalho de Neuhäuser *et al.*, cujas conclusões

algo abusivas face à análise de dados (também esta com imperfeições), nos levou a uma investigação sistemática de formas de aleatoriedade discretas, chegando a conclusões cuja interpretação biológica parece mais razoável. Esse trabalho conjunto com Marques e Pestana, publicado na revista *Biometrical Letters*, é transcrito como parte integrante do Capítulo 3. Um aspecto que nos merece particular realce é que, ao transformar variáveis aleatórias restringindo-lhes o suporte, se obtêm famílias de variáveis modificadas ou truncadas com um espaço de parâmetros que pode ser bem mais amplo que o original.

É assim que chegamos a classes de “variáveis potência”, cujo papel em fractalidade é cada vez mais relevante, e, de um modo natural, estendemos a clássica lei de Zipf a classes de leis “hiperbólicas” discretas — que, na perspectiva de Mandelbrot, têm para fenómenos complexos a mesma importância que a lei gaussiana nas áreas mais tradicionais da Ciência.

Finalmente, no Capítulo 4, apresentam-se alguns resultados parcelares relacionados com trabalho conjunto com Mendonça e Pestana (2003). Neste capítulo a minha opção foi não integrar esse trabalho conjunto, por considerar ter sido modesta a minha contribuição.

Parte do trabalho realizado durante o período de investigação conducente ao pedido de provas de doutoramento não foi incluído, por se afastar da linha principal do trabalho aqui apresentado.⁽¹⁾

Passo então à transcrição dos dois artigos que são parte integrante desta introdução.

⁽¹⁾ Por exemplo:

- Marques, T., Pestana, D. D., Vasconcelos, R., Papoila, A. L., Velosa, S. (2003). “Meta-Análise, a Síntese de Evidência Estatística”, *Notas e Comunicações do CE-AUL*, Lisboa.
- Pestana, D. D., Sequeira, F., Velosa, S. F. (2004). “Relação de Parseval, densidades definidas positivas e estabilidade”, *Estatística com Acaso e Necessidade — Actas do XI Congresso Anual da Sociedade Portuguesa de Estatística*, 615-627.
- Gaspar, H., Almeida, M. J., Freitas, D., Velosa, S., Maia, J. (2004). “Growth, physical fitness and specific motor skills in Portuguese soccer players”, *9th Annual Congress of the ECSS*, Clermont-Ferrand.

Somas e máximos de variáveis aleatórias independentes — Classes de leis limites

Sílvia Filipe Velosa ⁽¹⁾

*Departamento de Matemática, Universidade da Madeira, e
CEAUL — Centro de Estatística e Aplicações da Universidade de Lisboa
svelosa@math.uma.pt*

Resumo: Depois de uma panorâmica do desenvolvimento clássico de leis estáveis e seus domínios de atracção, e de leis infinitamente divisíveis, na perspectiva de convergências de classes, apresentam-se alguns resultados sobre refinamentos de classes de leis limites de funções de variáveis independentes construindo M_∞ , a menor classe fechada para produtos e limites de funções de distribuição que contém todas as distribuições max-estáveis. Aborda-se ainda a problemática de somas e de máximos de um número aleatório de variáveis aleatórias i.i.d., quando a variável de contagem não é Poisson, resolvendo o problema de Zolotarev sobre auto-decomponibilidade aleatória no contexto de máximos e caracterizando as leis limites estáveis de máximos $X_{N:N} = \max \{X_1, \dots, X_N\}$ de variáveis aleatórias X_k i.i.d., com $N \sim \text{Geométrica}(\theta)$ independente dos termos X_k , e seus domínios de atracção.

Palavras-chave: estabilidade, divisibilidade infinita, auto-decomponibilidade, somas, somas aleatórias, máximos.

Abstract: After a brief review of the classical extensions of the central limit theorem and of the extremal limit theorem, namely of infinite divisibility and of stable laws and their domains of attraction, we focus on some new refinements and extensions. Starting from Mejsler's M class of limit distributions of maxima of independent random variables, we refine the concept of self-decomposability for maxima, and construct and characterize M_∞ , the smallest class containing all max-stable distributions that is closed under products and limits. In what concerns an important extension of the classical results — stopped sums and stopped maxima, with non-Poisson subordinator — we study the analogue for maxima of Zolotarev's question on random self-decomposability, and characterize the max-stable laws when the subordinator is geometric.

Keywords: stability, infinite divisibility, self-decomposability, sums, random sums, maxima.

MSC2000: 60F05; 60E07.

⁽¹⁾Investigação parcialmente financiada por FCT/POCTI/FEDER (Projecto VEXTRA).

1 Teorema limite central e suas extensões

A convergência em distribuição de somas parciais de variáveis aleatórias independentes e identicamente distribuídas (i.i.d.), convenientemente centradas e reduzidas, para uma gaussiana padrão, tornou-se um resultado de tal importância que Pólya (1920) o apodou de Teorema Limite Central, um nome que caiu em graça e é desde então usado. O teorema clássico foi sujeito a inúmeras extensões, nomeadamente o Teorema Limite Central com as condições de Lindeberg-Feller, estabelecendo resultado análogo para a sucessão de somas parciais de variáveis aleatórias independentes, desde que se imponha uma condição de limitação uniforme no crescimento da soma das variâncias truncadas (uma condição que, em última análise, assegura que as caudas das parcelas são suficientemente leves, pelo que a contribuição de cada uma das parcelas nunca se torna preponderante — para mais comentários, cf. Pestana e Velosa (2002, Cap. 8), Laha e Rohatgi (1979) ou Galambos (1995)).

Lévy (1925, 1937) desenvolveu duas extensões de grande importância ao teorema clássico de de Moivre (1738) e Laplace (1812): a teoria geral das leis estáveis e seus domínios de atracção, e a teoria ainda mais geral das leis infinitamente divisíveis.

A teoria das leis estáveis generaliza o teorema limite central clássico a situações em que o peso das caudas torna preponderante o papel de alguma(s) das parcelas: Seja $\{X_k\}_{k \geq 1}$ uma sucessão de variáveis aleatórias i.i.d., $S_n = \sum_{k=1}^n X_k$ e suponha-se que existem *constantes de atracção* $a_n > 0$ (destinadas a estabilizar a dispersão) e $b_n \in \mathbb{R}$ (destinadas a estabilizar a localização) tais que

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

A variável aleatória limite Y é *estável* para somas, no sentido em que se “replica” por somas: se tomarmos a sucessão de somas parciais de réplicas Y_k independentes de Y , existem *constantes normalizadoras* $A_n > 0$ e $B_n \in \mathbb{R}$, para $n = 1, 2, \dots$, tais que é verificada a *equação de estabilidade*

$$\frac{\sum_{k=1}^n Y_k - B_n}{A_n} \stackrel{d}{=} Y.$$

Nestas circunstâncias diz-se que $X \stackrel{d}{=} X_k$ está no domínio de atracção da variável aleatória estável Y , $X \in \mathcal{D}(Y)$. É óbvio que estamos a falar de *convergências de tipos*, no sentido de Khinchine, pelo que em geral concentramos a discussão em variáveis padrão, representantes particularmente simples das classes que tipificam.

No teorema limite central clássico, demonstra-se convergência para a Gaussiana no caso de X ter variância (uma condição suficiente, mas não necessária:

basta que uma função simples da variância truncada seja de variação lenta). Neste caso servem as constantes de atracção $a_n = \sigma\sqrt{n}$ e $b_n = n\mu$, notações tão usuais que nos dispensamos de explicações supérfluas, e, na equação de estabilidade, neste caso $A_n = n^{\frac{1}{2}}$. Prova-se, mais geralmente, que na equação de estabilidade $A_n = n^{\frac{1}{\alpha}}$ para algum $\alpha \in (0, 2]$, a que se chama *expoente característico* da variável aleatória estável. Assim, a gaussiana é a estável de expoente característico $\alpha = 2$ (e a Cauchy (1853) é uma estável com expoente característico $\alpha = 1$, um facto já relatado em Laplace, 1812).

A invenção das estáveis era bastante promissora: muitos dos fenómenos podem ser modelados invocando algum mecanismo aditivo, e que modelo poderia ser mais simples do que aquele em que a adjunção de nova informação apenas iria alterar parâmetros de localização e escala? É esta uma das razões do sucesso do modelo “normal” clássico, e seria decerto de grande utilidade dispor de modelos análogos com caudas mais pesadas.

Infelizmente a equação de estabilidade é fácil de resolver apenas no espaço transformado das funções características, mostrando-se que as soluções definidas positivas (condição necessária e suficiente de Bochner para uma função φ com $\varphi(0) = 1$ ser função característica de uma variável aleatória honesta) da equação de estabilidade

$$\varphi(t) = e^{-it \frac{B_n}{A_n}} \left[\varphi\left(\frac{t}{A_n}\right) \right]^n$$

são da forma

$$\varphi(t) = \exp \left\{ iat - c|t|^\alpha \left[1 - i\beta \frac{t}{|t|} \omega(t, \alpha) \right] \right\},$$

onde $a \in \mathbb{R}$ é um parâmetro de localização, $c > 0$ é um parâmetro de escala, $\beta \in [-1, 1]$ é um parâmetro de assimetria, α é o expoente característico ou *índice*, e

$$\omega(t, \alpha) = \begin{cases} \operatorname{tg} \frac{\pi\alpha}{2} & \alpha \in (0, 1) \cup (1, 2] \\ -\frac{2}{\pi} \ln |t| & \alpha = 1 \end{cases}.$$

(Zolotarev (1986) usa representações alternativas, que para determinados objectivos são mais cómodas). E a expressão acima não é fácil de inverter analiticamente, apenas sendo conhecidas as funções densidade de probabilidade estáveis para os já referidos casos $\alpha = 2$ (gaussiana) e $\alpha = 1, \beta = 0$ (Cauchy), e para $\alpha = \frac{1}{2}, \beta = 1$ (Lévy). Assim, o que parecia tão promissor é decepcionante, pois em geral apenas está acessível o tratamento computacional destes modelos.

Anote-se desde já que o papel desempenhado pelas funções características no esquema de somas de variáveis aleatórias independentes (a função característica da soma de independentes é o produto das funções características das parcelas) é desempenhado pelas funções de distribuição no esquema de máximos de variáveis aleatórias i.i.d.

O problema da caracterização das funções de distribuição estáveis para máximos é assim o da resolução da *equação de estabilidade*

$$F^n(A_n x + B_n) = F(x)$$

que Fréchet (1927), um discípulo de Lévy, importou da teoria das somas para a teoria de valores extremos, inaugurando os estudos conducentes ao *teorema limite extremal* que paraleliza o *teorema limite central*. A solução da equação funcional de estabilidade para extremos deve-se a Fisher e Tippett (1928), creditando-se a Gnedenko (1943) a demonstração de que as soluções por eles avançadas são as únicas possíveis. Mais geralmente, é possível estudar os elementos estáveis de convoluções generalizadas (Bingham, 1971, na sequência de um trabalho em que Kingman (1963), para estudar passeios aleatórios em esferas, tinha introduzido uma operação integral partilhando de muitas das propriedades da convolução) e seus domínios de atracção. Este desenvolvimento foi possível depois de Feller (1967) ter mostrado que a caracterização dos domínios de atracção de leis estáveis para somas (Doebelin, 1940; Gnedenko, 1940) e de leis estáveis para máximos (Gnedenko, 1943, com excepção do caso do domínio de atracção da Gumbel, que seria resolvido por de Haan em 1970) se faria com maior eficiência recorrendo à teoria das funções de variação lenta e de variação regular de Karamata (1930). Para uma excelente resenha, consulte-se Bingham *et al.* (1987).

É fácil notar que se X é estável, então é também *infinitamente divisível*, no sentido em que X pode ser decomposta como soma de tantas parcelas i.i.d. quantas desejarmos, isto é

$$X \stackrel{d}{=} Y_1 + \dots + Y_n,$$

qualquer que seja n , com parcelas Y_k i.i.d. Por outras palavras, para todo natural $n \geq 1$, existe uma função característica ψ_n tal que a função característica φ de X verifica $\varphi(t) = \psi_n^n(t)$, ou seja $\varphi^{1/n}$ é uma função característica.

É de facto imediato estabelecer que, no caso de X ser estável, a raiz de índice n , qualquer que seja n , da sua função característica φ , continua a ser uma função característica: da equação de estabilidade decorre imediatamente que

$$\varphi^{1/n}(t) = \exp\left(-it \frac{B_n}{nA_n}\right) \varphi\left(\frac{t}{A_n}\right).$$

(É fácil provar — cf. Lukacs (1970) — que uma função característica infinitamente divisível não tem zeros reais, pelo que se pode definir sem ambiguidade o seu logaritmo, e consequentemente qualquer potência positiva).

A noção de divisibilidade infinita surgiu nos anos 30, tendo de Finetti (1930) demonstrado um resultado crucial: as leis infinitamente divisíveis podem ser identificadas com Poissons compostas ou seus limites. Kolmogorov (1932) deduziu uma representação integral para as funções características infinitamente

este pode ser aproximado por $\exp \left[- \sum \left(\varphi_{x_{nk}}(t) - 1 \right) \right]$. Afinal, as representações canónicas não são mais do que uma elaboração rigorosa desta constatação. Veja-se também Burrill (1977, pp. 337-341), onde se procede a uma fascinante preparação intuitiva da representação de Lévy-Khinchine, que é depois rigorosamente demonstrada.

O estudo actual das leis estáveis é usualmente feito a partir da representação integral das funções características infinitamente divisíveis, usando a equação de estabilidade, veja-se por exemplo Lukacs (1970). As descobertas de Lévy e de Khinchine permitiram o desenvolvimento de uma “aritmética das leis de probabilidade”, inspirada na aritmética do semigrupo multiplicativo dos inteiros, mas com patologias estranhas em que se enredou muita investigação. Estes problemas são hoje enquadrados na investigação de semi-grupos topológicos (que Kendall (1967) apodou de semi-grupos délficos, por os ter apresentado num congresso realizado em Loutraki, nas imediações de Delfos), veja-se também Kendall e Harding (1973).

Mais uma vez, o que se estabelece para o esquema de somas tem paralelo no esquema de máximos: num caso usa-se produtos de funções características, no outro produtos de funções de distribuição. Durante muito tempo não se deu qualquer relevo à noção de divisibilidade infinita para máximos, pela simples razão que qualquer função de distribuição univariada é infinitamente divisível, o que retira qualquer interesse à questão. No entanto, Balkema e Resnick (1977) mostraram que no caso multivariado deixa de se ter esta situação trivial (e as funções de distribuição infinitamente divisíveis para máximos surgem naturalmente associadas aos processos extremais, tal como as distribuições infinitamente divisíveis para somas estão naturalmente associadas aos processos com incrementos independentes); e adiante, ao generalizar a classe de Meijer, demonstraremos que o conceito tem, mesmo no caso univariado, importância.

O contexto de estabilidade é demasiado estrito em modelação, o de divisibilidade infinita demasiado vago. Tal como o Teorema Limite Central com as condições de Lindeberg-Feller, no caso de parcelas com caudas moderadas (ou seja, em que nenhuma toma valor preponderante), é o verdadeiro contexto que justifica a importância do modelo gaussiano, como limite de somas de parcelas independentes (é aliás extensível para formas brandas de dependência), há que relaxar o espartilho de identidade distribucional na situação de caudas pesadas.

O impulso inicial nesta direcção deve-se a Khinchine, que colocou o problema, devendo-se a solução geral a Lévy (1937). A classe de leis limites não degeneradas de somas de variáveis aleatórias independentes convenientemente normalizadas $\sum_{k=1}^n \frac{X_k - b_n}{a_n}$, com a “condição de densidade” $a_n \xrightarrow{n \rightarrow \infty} \infty$ e $\frac{a_{n+1}}{a_n} \xrightarrow{n \rightarrow \infty} 1^{(4)}$, denomina-se classe L de Khinchine; o problema análogo

(4) As generalizações $\frac{a_{n+1}}{a_n} \xrightarrow{n \rightarrow \infty} r > 1$ e $\frac{a_{n+1}}{a_n} \xrightarrow{n \rightarrow \infty} r \in (0, 1)$ foram estudadas por Sreehari (1970) e por Kruglov (1972), respectivamente.

para máximos foi estudado por um discípulo de Gnedenko, Mejlzer (1956), pelo que denotamos classe M a classe das possíveis leis limites não degeneradas de máximos convenientemente normados de variáveis independentes.

Prova-se que uma função característica φ é da classe L se e só se verificar a equação funcional

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

onde ψ_a é uma função característica. Prova-se, além disso, que ψ_a é infinitamente divisível.

Por outras palavras, a correspondente variável aleatória X admite, para todo $a \in (0, 1)$, a decomposição em parcelas independentes $X \stackrel{d}{=} aX + Y_a$, onde Y_a é uma variável aleatória infinitamente divisível. Por esta razão, autores há que preferem chamar *classe das auto-decomponíveis* à classe L de Khinchine. Obviamente

$$\{\text{estáveis}\} \subset \{\text{auto-decomponíveis}\} \subset \{\text{infinitamente divisíveis}\}.$$

Poderia parecer, à primeira vista, que estava encontrado o contexto adequado para a modelação de fenómenos sujeitos a um mecanismo aditivo. Duas objecções, no entanto, têm que ser feitas:

- O número de observações é sempre finito, pelo que, mesmo que cada qual provenha de um modelo parente diferente (como é viável neste esquema de parentes independentes mas não necessariamente identicamente distribuídas), nunca estará em causa toda a complexidade conceptual admissível no problema de Khinchine. Quando estão em causa parcelas de diversos tipos, intuitivamente esperar-se-ia obter como lei limite uma mistura de estáveis (na Secção 2, ao comentarmos a *conjectura de Gnedenko*, veremos que a verdade é, neste caso particular, contra-intuitiva). Não se sabe a que ponto a classe L é mais vasta do que todas as possíveis misturas de estáveis.
- Não há uma caracterização das situações que correspondem a cada uma das possíveis leis limites (enquanto para estáveis há o estudo completo dos domínios de atracção, em termos de variação regular das caudas; e no Teorema Limite Central de Lindberg-Feller, há também um esclarecimento das condições a exigir, em termos da sucessão de variâncias truncadas, para se ter convergência para a normal).

Por estas razões, podemos considerar que os trabalhos clássicos de Lévy, Khinchine e seus discípulos (com relevo para Doeblin, Gnedenko e o seu estudante Mejlzer) deram resposta a problemas importantes, mas sobretudo suscitaram um amplo leque de novas perguntas e campos de investigação, que abordamos na Secção que segue.

2 Extensões do Esquema Clássico

As descobertas de Lévy sobre leis estáveis e leis infinitamente divisíveis (e, mais geralmente, sobre processos com incrementos independentes), e extensões para situações de dependência fraca (por exemplo martingalas, uma outra importante criação de Lévy) estão na génese da moderna Teoria da Probabilidade, cuja história na primeira metade do século XX pode, em grande medida, ser feita acompanhando os desenvolvimentos daqueles campos.

Não é por isso de estranhar que os conceitos de estabilidade e de divisibilidade infinita tenham invadido outros campos. Só a título de exemplo, refira-se o estudo de POT (*Peaks Over Thresholds*), cf.

Reiss e Thomas (1991), cujos elementos estáveis são as Paretos generalizadas, ou o recente conceito de planos experimentais infinitamente divisíveis.

A teoria clássica situa-se no campo da convergência de tipos, assumindo transformações lineares. Mesmo estas podem ser mais sofisticadas, usando transformações lineares diferentes para cada uma das parcelas, isto é procurando possíveis limites fracos não degenerados de

$$\sum_{k=1}^n \frac{X_k - b_{k,n}}{a_{k,n}}.$$

Tanto quando sabemos, a questão foi apenas parcialmente tratada por Schreiber (1977), e Graça Martins e Pestana (1988) trataram um problema idêntico no esquema de máximos. Os resultados são evidentemente mais ricos do que os clássicos: no caso de parcelas i.i.d., obtém-se como classe de leis limites um conjunto convexo que, no caso de a distribuição parente estar no domínio de atracção de uma estável, contém, naturalmente, essa lei estável. No caso de parcelas independentes, obtém-se uma classe mais vasta do que a classe L , que será porventura a classe das infinitamente divisíveis, mas que, por enquanto, não foi caracterizada.

A tentação de usar transformações não lineares é natural. Afinal, se conhecêssemos a lei limite, podíamos chegar a ela em dois passos, usando o teorema da transformação uniformizante! Mas, referindo questões menos platónicas, e exemplificando agora com máximos: Desde Fisher e Tippett (1928) que se sabe que máximos normalizados de gaussianas convergem muito lentamente para a “*ultimate limit law*” Gumbel, e que para todo n existe uma “*penultimate approximation*” Weibull mais próxima de $F^n(a_n x + b_n)$. Este assunto, que fez correr rios de tinta até ser finalmente explicado por Gomes (1984) e por Gomes e Pestana (1987) — que mostraram a importância de particionar o domínio de atracção em *domínio de atracção normal* e *domínio de atracção não normal* consoante a função de variação regular envolvida na escolha das constantes de atracção converge ou não para uma constante não nula; veja-se também Iglésias Pereira *et al.* (1996), no que respeita resultados pré-assintóticos em somas, que mostram que se trata de um falso problema, pois aproximações pré-assintóticas são a regra, e não a excepção que o trabalho de Fisher e Tippett parecia apontar —,

trouxe em particular a constatação (Haldane e Jayakar, 1963) de que máximos normalizados de quadrados de Gaussianas convergem mais rapidamente para o limite Gumbel do que os máximos normalizados de Gaussianas.

Mais recentemente houve abordagens considerando normalização pela aplicação de operadores (“operator stable laws”), ou “auto-normalizações”. Veja-se o interessante trabalho de Logan *et al.* (1973), no que refere somas, e os resultados de Mendonça (2000) em máximos. Também Hall (1998), por trabalhar com variáveis aleatórias discretas, preferiu usar *rarefação* (“thinning”) — que abre a porta das somas aleatórias, que abordamos na Secção 3 — em vez de normalização linear.

Todas estas abordagens são interessantes, abrem novos caminhos e propiciam uma visão mais abrangente da problemática de teoremas limites fracos em Teoria da Probabilidade. No entanto o esquema clássico admite extensões não triviais, que nos ocupam, e de que passamos a expor alguns resultados.

No intuito de tornar o esquema clássico mais interessante do ponto de vista de modelação, Gnedenko conjecturou que se usássemos parcelas de diferentes tipos, e em diferentes domínios de atracção se obteriam classes de leis limites que corresponderiam a misturas das estáveis que surgiriam como limite de cada um dos tipos de parcelas consideradas.

Por exemplo: sabe-se que se $X \sim \text{Gaussiana}$, $Y = \frac{1}{X} \in \mathcal{D}(\text{Cauchy})$, cf. Gomes e Pestana (1981). Então, se considerarmos uma sucessão $\{W_k\}_{k \geq 1}$, em que para cada n se tem n_1 réplicas $W_k \stackrel{d}{=} \frac{1}{X} = Y$ e $n_2 = n - n_1$ réplicas $W_k \stackrel{d}{=} X$, então a soma

$$\frac{\sum_{k=1}^n W_k}{\sqrt{n}} = \frac{\sum_{k=1}^{n_1} Y_k}{n_1} \frac{n_1}{\sqrt{n}} + \frac{\sum_{k=1}^{n_2} X_k}{\sqrt{n_2}} \sqrt{\frac{n_2}{n}}$$

no caso de $\frac{n_1}{\sqrt{n}} \xrightarrow{n \rightarrow \infty} \alpha \in (0, 1)$ converge para uma mistura de Gaussiana e Cauchy.

Zolotarev e Koroljuk (1961) mostraram que a conjectura de Gnedenko é válida para $n = 2$, mas Zinger (1965) mostrou que para $n \geq 3$ é falsa. Assim, o caminho de ir ampliando o esquema de estáveis relaxando progressivamente a hipótese de identidade distribucional não parece o adequado. Veja-se Graça Martins e Pestana (1987, 1988) no que refere problema idêntico para extremos.

Urbanik (1973) tomou o caminho oposto, partindo da classe das infinitamente divisíveis para refinamentos da classe L de Khinchine.

Denote-se a classe das infinitamente divisíveis L_0 , e L_1 a classe das auto-decomponíveis de Khinchine. Vimos já a caracterização das funções características φ auto-decomponíveis:

$$\varphi \in L_1 \text{ se e só se } \forall a \in (0, 1), \varphi(t) = \varphi(at) \psi_a(t), \text{ com } \psi_a \in L_0.$$

Podemos então definir iterativamente as subclasses L_k , $k = 2, 3, \dots$ usando

$$\varphi \in L_k \text{ se e só se } \forall a \in (0, 1), \varphi(t) = \varphi(at) \psi_a(t), \text{ com } \psi_a \in L_{k-1}.$$

É óbvio que $L_k \subset L_{k-1}$; Urbanik tem implícita uma interpretação em termos de somas de variáveis aleatórias, construindo o que denomina de disposições triangulares de variação lenta.

Vejamos a construção análoga que se pode fazer em termos de máximos de variáveis aleatórias independentes.

Definam-se, para cada variável aleatória X , os *pontos terminais* esquerdo e direito, $\alpha = \alpha_{F_X} = \inf \{x : F_X(x) > 0\}$, $\omega = \omega_{F_X} = \sup \{x : F_X(x) < 1\}$, respectivamente (“*left endpoint*” e “*right endpoint*”), e a cada função de distribuição F_X associem-se

$$\text{Caso 1. } F_X^{(1)}(x) = F_X(x)$$

$$\text{Caso 2. } F_X^{(2)}(x) = F_X(\omega - e^{-x}) = F_{-\ln(\omega-X)}(x) \quad \text{se } \omega < \infty$$

$$\text{Caso 3. } F_X^{(3)}(x) = F_X(\alpha + e^x) = F_{\ln(X-\alpha)}(x) \quad \text{se } \alpha > -\infty$$

Denote-se M_0 a classe de todas as funções de distribuição (vimos já que, no caso univariado, são infinitamente divisíveis para máximos), e defina-se M_1 da seguinte forma:

$$F_X \in M_1 \text{ se e só se } \forall a > 0 \text{ existe } Y_a \in M_0 \text{ tal que } G_X^{(i)}(x) = G_X^{(i)}(x+a) G_{Y_a}^{(i)}(x),$$

para o $i \in \{1, 2, 3\}$ que seja apropriado aos pontos terminais de F_X (escrevemos indiferentemente $X \in M_k$ ou $F_X \in M_k$; por outro lado, quando tal for conveniente distinguimos as subclasses $M_k^{(i)}$, $i = 1, 2, 3$, sendo $M_k = \bigcup_{i=1}^3 M_k^{(i)}$).

A classe M_1 , descoberta por Mejsler (1956), é, no esquema de máximos de variáveis independentes, o símile da classe L de Khinchine, no sentido em que é a classe das leis limites de máximos adequadamente normalizados de variáveis independentes, sob uma “condição de uniformidade para o máximo” (que é o símile, para máximos, da condição de negligibilidade assintótica para somas)

$$\max_{1 \leq k \leq n} \mathbb{P} \left[\frac{X_k - c_n}{d_n} > x \right] = \max_{1 \leq k \leq n} \left\{ 1 - F_{X_k}(c_n + d_n x) \right\} \xrightarrow{n \rightarrow \infty} 0$$

e, para todo $t \in (0, 1)$, fixo, existe

$$\lim_{n \rightarrow \infty} \sum_{k=1}^{I(nt)} \mathbb{P} \left[\frac{X_k - c_n}{d_n} > x \right] = \lim_{n \rightarrow \infty} \sum_{k=1}^{I(nt)} \left[1 - F_{X_k}(c_n + d_n x) \right] = z(t, x),$$

onde $I(x)$ denota o maior inteiro não superior a x , finito no caso de $z(1, x)$ o ser (admite-se que $x > \alpha_{F_X}$ em que F_X é a distribuição limite dos máximos normalizados).

A equação funcional acima pode ser reformulada em termos de log-concavidade de $G_X^{(i)}$, no suporte de X , e para o(s) i apropriado(s), cf. Galambos (1987, pp. 216–220).

Comente-se a equação $G_X^{(i)}(x) = G_X^{(i)}(x+a) G_{Y_a}^{(i)}(x)$, onde $Y_a \in M_a$, $\forall a > 0$. Para o caso $i = 1$, por exemplo, é $F_X(x) = F_X(x+a) F_{Y_a}(x)$, $\forall a > 0$, o que significa que existe, para todo $a > 0$, uma variável aleatória Y_a independente de X e tal que

$$X \stackrel{d}{=} \max\{X - a, Y_a\}$$

justificando-se assim que também no esquema de máximos falemos de auto-decomponibilidade. Para mais detalhes veja-se também Hall (1998).

Passemos agora a uma construção de refinamentos da classe M , por iteração de aplicação daquela equação funcional, inspirando-nos em Urbanik (1973) e em Graça Martins e Pestana (1988), mas levando mais longe a construção e a interpretação em termos de máximos. Referiremos também os resultados de Pestana e Mendonça (2001) sobre monotonia generalizada de ordem não inteira para definir classes M_α , $\alpha > 0$. No que diz respeito a somas, veja-se também Kumar e Schreiber (1978), que procedem a uma análise e representação integral de Choquet identificando os pontos extremos destes conjuntos convexos, e Bai e Yin (1984).

As observações acima, capitalizando na similitude formal entre funções de distribuição e funções características no tratamento de máximos e de somas, respectivamente, de variáveis independentes, não são totalmente satisfatórios, por não ser explícito o sentido probabilista profundo daquela *auto-decomponibilidade* formal.

Detalhemos por isso os pormenores da construção das classes M_k , em que será patente o que é a variável Y_a na decomposição $X \stackrel{d}{=} \max\{X - a, Y_a\}$. Seguindo os passos da construção feita por Urbanik no contexto de somas, definimos as subclasses M_j da classe M de Meizler com base num critério de estacionaridade, que depois mostramos ser equivalente à condição de auto-decomponibilidade. Adaptamos o conceito de *slowly varying triangular arrays* de Urbanik (1973) ao contexto de máximos:

Seja $X_1, X_2, \dots, X_n, \dots$ uma sucessão de variáveis aleatórias independentes. Dizemos que $\{X_k\} \in S_1$ se e só se existem sucessões de constantes c_n e $d_n > 0$ tais que os máximos normalizados $\frac{X_{n:n} - c_n}{d_n}$ têm uma distribuição limite não degenerada F e $\{\hat{X}_k\}$ verifica, para qualquer $x > \alpha_F$, as condições de uniformidade para o máximo.

Dizemos que $\{X_k\} \in S_j$ ($j=2, 3, \dots$) se e só se $\{X_k\} \in S_1$ e para todo o $c > 0$ a disposição triangular $X_{nk} = X_{I(cn)+k}$ ($k = 1, 2, \dots, n$; $n = 1, 2, \dots$) é equivalente a um elemento de S_{j-1} — isto é, se a sucessão dos máximos $\max_{1 \leq k \leq n} X_{nk}$, devidamente normalizada, tiver uma distribuição limite não degenerada, e co-

num a alguma sucessão de S_{j-1} . Definimos também $S_\infty = \bigcap_{j=1}^{\infty} S_j$.

Designaremos por M_j , $j = 1, 2, \dots, \infty$, a classe das possíveis distribuições limites de sucessões de S_j , aqui representadas pelas respectivas funções de distribuição. M_1 é a classe de Meijler das funções de distribuição log-côncavas. Convencionaremos designar por M_0 a classe de todas as funções de distribuição.

Começamos por notar que uma função de distribuição log-côncava é necessariamente contínua — se o não fosse seria ilimitada em qualquer intervalo, cf. Hardy *et al.*, 1978, p. 96, o que é manifestamente impossível —, e honesta (sem átomos em $+\infty$ e $-\infty$). Como apenas vamos considerar sucessões de variáveis aleatórias honestas, também a distribuição limite F não tem átomos de probabilidade nesses pontos.

A condição “ $\ln F$ côncava” equivale a

$$\forall a > 0 \quad \frac{F(x)}{F(x+a)} = G_a(x), \quad x \geq \alpha,$$

onde G_a é uma função não decrescente, e $\alpha = \alpha_F$. Nestas condições, podemos tomar para G_a uma função de distribuição.

De facto, definindo $F_a(x) = G_a(x)$ se $x \geq \alpha$ e $F_a(x) = 0$ caso contrário, temos que F_a é uma função contínua, dado que F é contínua. Pela mesma razão, e por F_a ser não decrescente,

$$\sup_{x \geq \alpha} F_a(x) = \frac{\lim_{x \rightarrow \omega} F(x)}{\lim_{x \rightarrow \omega} F(x+a)},$$

com $\omega = \omega_F$. Quer o numerador quer o denominador são não nulos. Logo, F_a é majorada por $1 = \sup_{x \geq \alpha} F_a(x)$, e portanto $F_a(x) = \eta + (1 - \eta)G(x)$, onde G é uma função de distribuição. Por outro lado,

$$\eta = \inf_{x \geq \alpha} \frac{F(x)}{F(x+a)} = \frac{\lim_{x \rightarrow \alpha} F(x)}{\lim_{x \rightarrow \alpha} F(x+a)} = 0.$$

Por definição de α , o denominador é positivo. Segue-se então que $G = F_a$. Notamos também que F e F_a têm o mesmo suporte. Assim, a condição de log-concavidade $F(x) = F(x+a)F_a(x)$ equivale a $X \stackrel{d}{=} \max\{X-a, Y_a\}$, com Y_a uma variável aleatória independente de X .

Definam-se as variáveis aleatórias $X^{(i)}$, $i = 1, 2, 3$, tais que:

1. $X^{(1)} = X$,
2. $X^{(2)} = \ln(X - \alpha)$, se $\alpha > -\infty$,

$$3. \quad X^{(3)} = -\ln(\omega - X), \text{ se } \omega < +\infty.$$

Então a classe M_1 de Meizler é constituída pelas funções de distribuição das variáveis aleatórias X tais que para algum $i = 1, 2, 3$ se verifica a condição

$$\forall a > 0 \quad X^{(i)} \stackrel{d}{=} \max\{X^{(i)} - a, Y_a^{(i)}\},$$

onde Y_a é uma variável aleatória independente de X , com o mesmo suporte. Note-se também que $X = \alpha + \exp(X^{(2)})$ e $X = \omega - \exp(-X^{(3)})$.

Teorema:

A função de distribuição F de uma variável aleatória X pertence à classe M_j ($j = 1, 2, \dots$) se e só se para algum $i = 1, 2, 3$ se verifica a condição

$$\forall a > 0, \quad F^{(i)}(x) = F^{(i)}(x + a) G_a^{(i)}(x), \text{ com } G_a \in M_{j-1}.$$

Em termos de variáveis aleatórias, $X^{(i)} \stackrel{d}{=} \max\{X^{(i)} - a, Y_a^{(i)}\}$, com Y_a independente de X .

Demonstração:

Dada a extensão da prova vamos chamar a atenção, em itálico, para os diversos passos construtivos da demonstração, como guia de leitura.

Vamos agora demonstrar a condição necessária, isto é que se F for distribuição limite dos máximos normalizados de uma sucessão $\{X_k\} \in S_j$ então

$$\forall a > 0 \quad F^{(i)}(x) = F^{(i)}(x + a) G_a^{(i)}(x), \text{ com } G_a \in M_{j-1}.$$

Começemos por explicitar uma *decomposição* útil no que segue:

Seja $\{X_k\} \in S_1$ uma sucessão arbitrária de variáveis aleatórias independentes cujos máximos normalizados $\frac{X_{n:n} - c_n}{d_n}$ converjam para uma variável aleatória não degenerada X com função de distribuição F . Para todo $t \in (0, 1]$, definimos

$$\begin{aligned} \bullet \quad X_{nt:n} &= \max_{1 \leq k \leq I(nt)} X_k; \\ \bullet \quad X'_{nt:n} &= \max_{k > I(nt)} X_k = \max_{I(nt)+1 \leq k \leq n} X_k. \end{aligned}$$

Então,

$$\frac{X_{n:n} - c_n}{d_n} = \max \left\{ \frac{X_{nt:n} - c_n}{d_n}, \frac{X'_{nt:n} - c_n}{d_n} \right\}.$$

A decomposição acima corresponde a uma factorização das respectivas funções de distribuição em $F_n(x) = F_{nt}(x) H_{nt}(x)$.

Para o primeiro termo da decomposição, verifica-se:

$$\frac{X_{nt:n} - c_n}{d_n} \xrightarrow[n \rightarrow \infty]{d} \frac{X - A_t}{B_t}.$$

Assim, a distribuição limite do máximo retardado $X_{nt:n}$ é afectada apenas na localização e na escala⁽⁵⁾.

Há três possibilidades para A_t e B_t :

- I. $B_t = 1$ e $A_t = r \ln t$;
- II. $B_t = t^m$, com $m < 0$, e $A_t = -\alpha(t^m - 1)$;
- III. $B_t = t^m$, com $m > 0$, e $A_t = \omega(1 - t^m)$.

Caso I

Se $B_t = 1$ e $A_t = r \ln t$, então $r < 0$, visto que $X_{nt:n} \leq X_{n:n}$ (a distribuição limite de $X_{nt:n}$ vem antecipada relativamente à de $X_{n:n}$). Escolhendo $t = e^{\frac{a}{r}}$, tem-se

$$\frac{X_{nt:n} - c_n}{d_n} \xrightarrow[n \rightarrow \infty]{d} X - a$$

Relativamente ao segundo termo da decomposição efectuada, acabamos de ver que $F_n(x) \xrightarrow[n \rightarrow \infty]{} F(x)$ e $F_{nt}(x) \xrightarrow[n \rightarrow \infty]{} F(x + a)$, do que resulta

$$H_{nt}(x) \xrightarrow[n \rightarrow \infty]{} G_a(x) = \frac{F(x)}{F(x + a)}, \quad x \geq \alpha_F.$$

Daqui conclui-se imediatamente que a distribuição limite G_a é não degenerada, contínua e honesta, e verifica $F(x) = F(x + a) G_a(x)$ para todo o $a > 0$.

Resta apenas provar que $G_a \in M_{j-1}$:

Por hipótese, qualquer disposição triangular $X_{nk} = X_{I(cn)+k}$ é equivalente a um elemento de S_{j-1} . Ora,

$$X'_{nt:n} = \max_{I(nt)+1 \leq k \leq n} X_k = \max_{1 \leq k \leq n-I(nt)} X_{k+I(nt)}.$$

Por outro lado, $k_n = n - I(nt) \xrightarrow[n \rightarrow \infty]{} +\infty$. Logo, $X'_{nt:n}$ é equivalente a uma sucessão de S_{j-1} , e portanto $G_a \in M_{j-1}$.

Caso II

Se $B_t = t^m$, com $m < 0$, então a distribuição de X tem suporte limitado inferiormente e $A_t = -\alpha(t^m - 1)$. Nestas condições, por um raciocínio análogo ao que foi feito para o caso anterior, vê-se que as distribuições limites dos termos da decomposição verificam para $x > \alpha$ a igualdade

$$F(x) = F(A_t + B_t x) G_a(x),$$

⁽⁵⁾ Decorre de Galambos (1987, pp. 218–219), tal como as expressões para os parâmetros de localização e escala.

onde $G_a \in M_{j-1}$. Daqui vem que

$$F(\alpha + e^x) = F\left(A_t + B_t(\alpha + e^x)\right) G_a(\alpha + e^x).$$

Tomando $t = e^{\frac{a}{m}}$, fica $F\left(A_t + B_t(\alpha + e^x)\right) = F(\alpha + e^{x+a})$, ou seja

$$F^{(2)}(x) = F^{(2)}(x+a) G_a^{(2)}(x),$$

para todo o x real e $a > 0$, como queríamos demonstrar.

Caso III

Se $B_t = t^m$, com $m > 0$, então a distribuição de X tem suporte limitado superiormente e $A_t = \omega(1 - t^m)$. De modo análogo aos casos anteriores, e tomando $t = e^{-\frac{a}{m}}$, obtemos

$$F^{(3)}(x) = F(\omega - e^{-x}) = F(\omega - e^{-x-a}) G_a(\omega - e^{-x}) = F^{(3)}(x+a) G_a^{(3)}(x)$$

para todo o a positivo, com $G_a \in M_{j-1}$.

Passemos agora à condição suficiente: Admita-se que

$$\forall a > 0, F^{(i)}(x) = F^{(i)}(x+a) G_a^{(i)}(x), G_a \in M_{j-1}.$$

Pretendemos mostrar que F é distribuição limite dos máximos normalizados de uma sucessão $\{X_k\} \in S_j$.

Caso $i = 1$

Começemos por *construir uma sucessão* apropriada, cujos máximos normalizados *têm distribuição limite* F , da seguinte forma:

Atendendo à relação $F(x) = F(x+a) G_a(x)$, vemos que G_a é uma função de distribuição não degenerada, para todo o $a > 0$. Considere-se então uma sucessão de variáveis aleatórias independentes $\{X_k\}$ com funções de distribuição

$$F_1(x) = F(x) \quad \text{e} \quad F_k(x) = \frac{F(x - \ln k)}{F(x - \ln(k-1))}, \quad k = 2, 3, \dots$$

Os máximos desta sucessão têm a função de distribuição $F(x - \ln n)$. Tomando as constantes normalizadoras $c_n = \ln n$ e $d_n = 1$, tem-se $F_{X_{n:n}}(c_n + d_n x) = F(x)$. Logo, F é distribuição limite dos máximos normalizados.

No que respeita as *condições de uniformidade para o máximo*:

Facilmente se verifica que $F_k(c_n + d_n x) \xrightarrow{n \rightarrow \infty} 1$, $k = 1, 2, \dots$. Logo,

$$\max_{1 \leq k \leq n} \mathbb{P} \left[\frac{X_k - c_n}{d_n} > x \right] = \max_{1 \leq k \leq n} [1 - F_k(c_n + d_n x)] \xrightarrow{n \rightarrow \infty} 0.$$

Nestas condições, para qualquer $0 < t \leq 1$ a expressão

$$\sum_{k=1}^{I(nt)} [1 - F_k(c_n + d_n x)]$$

é (cf. Galambos, 1987, p. 127) assintoticamente equivalente ($n \rightarrow \infty$) a

$$\begin{aligned} -\ln \prod_{k=1}^{I(nt)} F_k(c_n + d_n x) &= -\sum_{k=1}^{I(nt)} \ln F_k(c_n + d_n x) \\ &= -\ln F\left(x - \ln \frac{I(nt)}{n}\right) \xrightarrow{n \rightarrow \infty} -\ln F(x - \ln t), \end{aligned}$$

que é finito para todo $t \in (0, 1]$ sempre que o seja para $t = 1$. (Se for finito para $t = 1$ tem-se $x > \alpha_F$, pelo que $x - \ln t > \alpha_F$ quando $0 < t < 1$.)

Quanto à condição de estacionaridade de Urbanik:

Dado um número $c > 0$, seja $X_{nk} = X_{I(cn)+1}$ ($k = 1, 2, \dots; n = 1, 2, \dots$). Para $n \geq 2$, a função de distribuição de $\max_{1 \leq k \leq n} X_{nk}$ é

$$\frac{F(x - \ln(I(cn) + n))}{F(x - \ln(I(cn) + 1))}$$

e a função de distribuição de $\max_{1 \leq k \leq n} X_{nk} - \ln n$ é

$$\frac{F\left(x - \ln \frac{I(cn)+n}{n}\right)}{F\left(x - \ln \frac{I(cn)+1}{n}\right)} \xrightarrow{n \rightarrow \infty} \frac{F(x - \ln(c+1))}{F(x - \ln c)} = G_{\ln(c+1) - \ln c}(x).$$

Como esta distribuição pertence a M_{j-1} , $\forall c > 0$, é por definição distribuição limite dos máximos normalizados de uma sucessão de S_{j-1} . Logo, $F \in M_j$, ficando assim estabelecida a condição suficiente no caso $i = 1$.

Caso $i = 2$

Construção e convergência da sucessão:

Para $i = 2$, consideramos a sucessão de variáveis aleatórias independentes $\{X_k^{(2)}\}$ com funções de distribuição

$$F_1^{(2)}(x) = F^{(2)}(x) \quad \text{e} \quad F_k^{(2)}(x) = \frac{F^{(2)}(x - \ln k)}{F^{(2)}(x - \ln(k-1))}, \quad k = 2, 3, \dots, n, \dots$$

Por razões análogas às que foram apresentadas no caso anterior, $X_{n:n}^{(2)} - \ln n \xrightarrow{d} X^{(2)}$, donde resulta, para a sucessão $Z_k = \exp(X_k^{(2)})$, que

$$\frac{Z_{n:n}}{n} + \alpha \xrightarrow{d} X.$$

Condições de uniformidade para o máximo:

As constantes normalizadoras para $Z_{n:n}$ são $c_n = -n\alpha$ e $d_n = n$, e é fácil verificar para $x > \alpha$ as condições

$$\max_{1 \leq k \leq n} \mathbb{P} \left[\frac{Z_k - c_n}{d_n} > x \right] = \max_{1 \leq k \leq n} \mathbb{P} \left[X_k^{(2)} - \ln n > x \right] \xrightarrow{n \rightarrow \infty} 0$$

e

$$\sum_{k=1}^{I(nt)} [1 - F_k(c_n + d_n x)] \xrightarrow{n \rightarrow \infty} -\ln F^{(2)}(x - \ln t) = -\ln F \left(\alpha + \frac{e^x}{t} \right),$$

que é um valor finito para qualquer $t \in (0, 1]$.

Condição de estacionaridade:

Falta provar que para todo o $c > 0$ a disposição triangular $Z_{nk} = Z_{I(cn)+k}$ ($k = 1, 2, \dots, n$; $n = 1, 2, \dots$) é equivalente a um elemento de S_{j-1} . Facilmente se vê que

$$\max_{1 \leq k \leq n} X_{nk}^{(2)} - \ln n \xrightarrow{d} Y^{(2)},$$

em que a função de distribuição de $Y^{(2)}$ é $G_{\ln(c+1)-\ln c}^{(2)}$. Isso implica que

$$\alpha + \frac{1}{n} \max_{1 \leq k \leq n} Z_{nk} \xrightarrow{d} Y = \alpha + \exp(Y^{(2)}),$$

cujas função de distribuição é $G_{\ln(c+1)-\ln c} \in M_{j-1}$. Logo, $F \in M_j$.

Caso $i = 3$

Construção e convergência da sucessão:

Consideremos agora a sucessão de variáveis aleatórias independentes $\{X_k^{(3)}\}$ com funções de distribuição $F_k^{(3)}$ dadas por

$$F_1^{(3)}(x) = F^{(3)}(x) \quad \text{e} \quad F_k^{(3)}(x) = \frac{F^{(3)}(x - \ln k)}{F^{(3)}(x - \ln(k-1))}, \quad k = 2, 3, \dots, n, \dots$$

Tem-se $X_{n:n}^{(3)} - \ln n \xrightarrow{d} X^{(3)}$, donde $n W_{n:n} + \omega \xrightarrow{d} X$, com $W_k = -\exp(-X_k^{(3)})$.

Condições de uniformidade para o máximo:

Deste modo, para $c_n = -\frac{\omega}{n}$, $d_n = \frac{1}{n}$ e $x > \alpha$ temos

$$\max_{1 \leq k \leq n} \mathbb{P} \left[\frac{W_k - c_n}{d_n} > x \right] = \max_{1 \leq k \leq n} \mathbb{P} \left[X_k^{(3)} - \ln n > x \right] \xrightarrow{n \rightarrow \infty} 0;$$

$$\sum_{k=1}^{I(nt)} [1 - F_k(c_n + d_n x)] \xrightarrow{n \rightarrow \infty} -\ln F^{(3)}(x - \ln t) = -\ln F(\omega - t e^{-x}).$$

Se a última expressão for finita para $t = 1$, então $\omega - t e^{-x} > \omega - e^{-x} > \alpha$, pelo que é também finita para $t \in (0, 1]$.

Condição de estacionaridade:

Seja $c > 0$, e $W_{nk} = W_{I(cn)+k}$, com $k = 1, 2, \dots, n$; $n = 1, 2, \dots$. Tal como no caso anterior verifica-se $\max_{1 \leq k \leq n} X_{nk}^{(3)} - \ln n \xrightarrow{d} Y^{(3)}$ donde resulta que

$$\omega + n \max_{1 \leq k \leq n} W_{nk} \xrightarrow{d} Y = \omega - \exp(-Y^{(3)}),$$

cujas função de distribuição é $G_{\ln(c+1) - \ln c} \in M_{j-1}$, por hipótese.

Fica assim concluída a demonstração da condição suficiente. \square

Corolário:

Uma função de distribuição F pertence à classe M_∞ se e só se para algum $i = 1, 2, 3$ se verifica a condição

$$\forall a > 0 \quad F^{(i)}(x) = F^{(i)}(x + a) G_a^{(i)}(x), \quad \text{com } G_a \in M_\infty.$$

Atendendo ao conceito de monotonia de ordem superior caracterizado por Pestana e Mendonça (2001) usando uma análise de pontos extremos,

$$F_x \in M_k^{(1)} \text{ se e só se } F_x(x) = e^{-K(x)},$$

com as restrições de $K(|x|)$ a $x > 0$ monótonas de ordem $k + 1$, podendo por isso usar-se a representação integral obtida no referido trabalho. Mais geralmente, é possível, usando a abordagem de Pestana e Mendonça, definir classes $M_\alpha^{(1)}$, $\alpha > 0$, mas até agora não obtivemos qualquer interpretação probabilística interessante, e é assunto que permanece nos limbos das curiosidades analíticas.

$$M_\infty^{(1)} = \bigcap_{k=0}^{\infty} M_k^{(1)}$$

contém as distribuições log-completamente monótonas. É sobejamente conhecido que uma função completamente monótona pode ser representada como transformada de Laplace de uma função não decrescente.

No que refere $M_k^{(2)}$ e $M_k^{(3)}$, podemos estabelecer resultados análogos. Repare-se que neste último caso ω é o supremo dos pontos terminais direitos das diversas variáveis independentes consideradas no esquema de convergência. Como o

ponto terminal direito é finito, em amostras crescentes daquela população limite há uma certa tendência para as estatísticas ordinais de topo acumularem próximo desse ponto, e neste caso é necessário um factor multiplicativo, em vez de uma mera translação, na equação funcional. Apenas os detalhes da análise são diferentes, como se pode ver por exemplo na exposição de Galambos (1987) no que respeita a análise pioneira de Meizler (1956) dos limites de máximos de variáveis independentes, classe que em sua honra denotámos M_1 .

Anote-se ainda que M_∞ é a menor classe de funções de distribuição fechada para produtos e limites pontuais que contém todas as distribuições estáveis para máximos, razão pela qual nos parece que constitui a classe natural para modelação de máximos.

3 Somas e Máximos — Esquemas Aleatórios

O esquema clássico e suas extensões podem ser consideravelmente ampliados admitindo diversas formas de aleatorização.

Por exemplo, em lugar de normalização linear, usando auto-normalização aleatória, $\frac{\sum_{k=1}^n X_k}{\left(\sum_{k=1}^n |X|^p\right)^{1/p}}$ (Logan *et al.*, 1973) ou $\frac{\max_{1 \leq k \leq n} X_k}{\left(\sum_{k=1}^n |X|^p\right)^{1/p}}$ (Mendonça,

2000). Ou, alternativamente, admitindo “falhas” aleatórias no processo de observação/registo, como Hall (1998), que usou consistentemente a ideia de rarefação (*thinning*). Ou ainda, como iremos fazer, considerando esquemas em que o número de termos é aleatório.

A própria ideia de divisibilidade infinita, que apresentámos associada ao conceito de esquemas triangulares assintoticamente negligíveis, começou por surgir, no trabalho seminal de de Finetti (1930), associada a limites de somas aleatórias de variáveis aleatórias, com a restrição de o número N de parcelas ter distribuição de Poisson — uma forma eficaz de manter a soma finita, pois o número de parcelas tipicamente não se afasta excessivamente de $\mathbb{E}(N) \pm \sqrt{\mathbb{E}(N)}$. Assim, a classe das leis infinitamente divisíveis é a classe das Poissons compostas e seus limites pontuais.

Este resultado, que parte do facto de as funções características infinitamente divisíveis não terem zeros reais, pode ser importado para o esquema de máximos, desde que haja o cuidado de considerar a restrição da função de distribuição ao suporte da variável aleatória: qualquer função de distribuição univariada (que, recordamos, é trivialmente infinitamente divisível para máximos) pode ser escrita sob a forma de “Poissons max-compostas”

$$F(x) = \lim_{n \rightarrow \infty} \exp[-\lambda_n (1 - F_n(x))], \quad x \in \mathcal{S}_x = \{x : dF_x(x) > 0\}$$

($\lambda_n = n$ e $F_n = F^{\frac{1}{n}}$ são escolhas óbvias),

uma expressão que nunca vimos explorada, porventura por o conceito de divisibilidade infinita para máximos de variáveis aleatórias ter parecido irrelevante.

É evidentemente possível estudar situações mais gerais, em que o processo subordinador não é Poisson. Rachev e os seus colaboradores têm dedicado algum esforço no que refere a um número geométrico de variáveis aleatórias, veja-se Mittnik e Rachev (1993) e Kozubowski (1994), ampliando consideravelmente, embora pareçam não se dar conta disso, os resultados obtidos por Rényi (1956) e Kovalenko (1965) no caso de parcelas positivas, e a classe das variáveis aleatórias de Linnik (1953), veja-se Pestana e Velosa (2003) e as referências que apontam. A representação de Kozubowski implica que as leis estáveis para somas de um número aleatório geométrico de variáveis aleatórias i.i.d. podem ser representadas como produto de uma estável para somas (no esquema clássico) por uma exponencial padrão dela independente. Face aos resultados de Steutel (1970), as geo-estáveis são, consequentemente, infinitamente divisíveis para somas, no sentido clássico.

Considere-se o problema, originalmente proposto por V. M. Zolotarev e investigado por Klebanov *et al.* (1984) no âmbito de somas, da descrição da classe das variáveis aleatórias X aleatoriamente max-auto-decomponíveis no sentido em que

$$\forall \theta \in (0, 1), \exists W_\theta : X \stackrel{d}{=} \max\{W_\theta, Z_\theta X\},$$

com X , W_θ e $Z_\theta \sim \text{Bernoulli}(1 - \theta)$ independentes. Denotamos F_θ a função de distribuição de W_θ . Como a função de distribuição de $Z_\theta X$ é

$$F_{Z_\theta X}(x) = \theta I_{[0, \infty)} + (1 - \theta) F_X(x)$$

a equação de Zolotarev só tem solução se as variáveis envolvidas tiverem suporte positivo. Para x positivo tem-se então

$$F_X(x) = \frac{\theta F_\theta(x)}{1 - (1 - \theta) F_\theta(x)} = \sum_{k=1}^{\infty} F_\theta^k(x) \theta (1 - \theta)^{k-1} = F_{X_{N:N}}(x),$$

a distribuição do máximo de um número aleatório de réplicas independentes de W_θ , em que o número de termos $N \sim \text{Geométrica}(\theta)$ é independente deles.

A relevância das geo(max)compostas — ou máximos aleatórios geométricos (*geometric stopped maxima*), como a discussão nos capítulos 8 e 9 de Johnson *et al.* (1992) sugere — torna-se mais expressiva quando se procura generalizar a noção de estabilidade, do contexto determinista da equação de estabilidade de Fréchet (1927) para este contexto aleatório:

Uma variável aleatória é geo(max)estável se e só se para todo $\theta \in (0, 1)$, $\exists a_\theta > 0$, $b_\theta \in \mathbb{R}$ tais que

$$F_X(a_\theta x + b_\theta) = F(a_\theta x + b_\theta) = \frac{\theta F(x)}{1 - (1 - \theta) F(x)}$$

(isto é, não só é max-aleatoriamente auto-decomponível no sentido de Zolotarev, como é obtida como máximo de um número aleatório geométrico de termos do seu tipo). No caso $b_\theta = 0$, dizemos que é estritamente geo(max)estável. Consideramos primeiro este caso:

Defina-se $G(x) = e^{1-\frac{1}{F(x)}}$, $x > \alpha_F$. A equação acima transforma-se em $G(a_\theta x) = G^{\frac{1}{\theta}}(x)$, a equação de estabilidade que foi originalmente estudada por Lévy no âmbito de somas, que o seu discípulo Fréchet adaptou a máximos, e que foi finalmente resolvida, no âmbito de máximos, por Fisher e Tippet (estabelecendo mais tarde Gnedenko que as soluções por eles encontradas, e geralmente referidas como Fréchet- α , Gumbel e Weibull- α são as únicas distribuições que a verificam).

Mas então, como $F(x) = \frac{1}{1-\ln G(x)}$, $x > \alpha_F$, da solução conhecida do problema de max-estabilidade clássico, deduz-se que as distribuições geo(max)estáveis são de um dos seguintes tipos:

1. $F(x) = \frac{1}{1+x^{-\alpha}} I_{(0,\infty)}$, $\alpha > 0$, que podemos apelidar de log-logística, que resulta da max-estável Fréchet- α .
2. $F(x) = \frac{1}{1+e^{-x}} I_{\mathbb{R}}$, a distribuição logística, associada à max-estável Gumbel.
3. $F(x) = \frac{1}{1+(-x)^{-\alpha}} I_{(-\infty,0)} + I_{[0,\infty)}$, $\alpha > 0$, simétrica da log-logística, da max-estável Weibull- α .

Ou, partindo da forma geral de von Mises-Jenkinson da distribuição de valores extremos $G(x) = \exp \left[- (1 + \gamma x)^{-\frac{1}{\gamma}} \right] I_{\{x: 1+\gamma x > 0\}}$, $\gamma \in \mathbb{R}$, a expressão sintética

$$F(x) = \frac{1}{1 + (1 + \gamma x)^{-\frac{1}{\gamma}}} I_{\{x: 1+\gamma x > 0\}}, \quad \gamma \in \mathbb{R}.$$

Rachev e Resnick (1991), obtiveram expressões análogas, num contexto independente do problema de Zolotarev.

Dos resultados sobre equivalência de caudas de Resnick (1972) e Cline (1986), vê-se que a caracterização dos domínios de atracção das log-logísticas é semelhante à caracterização dos domínios de atracção das Fréchet e das Weibull, respectivamente, no esquema clássico. Estamos a investigar a caracterização do domínio de atracção da logística, e outras situações de max-estabilidade aleatória, veja-se Pestana e Velosa (2003).

Agradecimentos:

A minha formação em Probabilidade deve-se à oportunidade que o Professor Dinis Pestana me ofereceu de trabalhar sob sua orientação. Em particular, o seu conhecimento profundo dos desenvolvimentos do teorema limite central e do teorema limite extremal entusiasmarão-me por este tópico, e proporcionaram-me um acesso simples a muita da bibliografia relevante.

Agradeço aos meus colegas do Departamento de Matemática da Universidade da Madeira, nomeadamente à Sandra Mendonça, sempre pronta a discutir pontos mais finos de Matemática, e aos Professores Rita Vasconcelos e José Molarinho Carmo, que como presidentes do DM sempre me encorajaram a investir na investigação.

Agradeço também ao *referee* os seus comentários estimulantes, que permitiram melhorar a apresentação dos resultados.

Bibliografia

- [1] Bai, Z. D. e Yin, Y. Q. (1984). Distributions of class L_α , *J. Mult. Anal.* **14**, 309–319.
- [2] Balkema, A. A. e Resnick, S. I. (1977). Max-infinite divisibility, *J. Appl. Prob.*, **14**, 309–319.
- [3] Bingham, N. H. (1971). Factorization theory and domains of attraction for generalized convolution algebras, *Proc. London Math. Soc.* (3), **23**, 16–30.
- [4] Bingham, N. H., Goldie, C. M. e Teugels, J. L. (1987). *Regular Variation*, Cambridge Univ. Press, Cambridge.
- [5] Burrill, W. C. (1977). *Measure, Integration, Probability*, McGraw-Hill, New York.
- [6] Cauchy, A. (1853). Sur les résultats moyens d'observations de même nature, et sur les résultats les plus probables, *C. R. Acad. Sci. Paris* **37**, 198–206.
- [7] Cline, D. (1986). Convolution tails, product tails and domains of attraction, *Probab. Theor.* **72**, 529–557.
- [8] Cramér, H. (1946, 1991). *Mathematical Methods of Statistics*, Princeton Univ. Press, Princeton.
- [9] de Moivre, A. (1738). *The Doctrine of Chances*, 2ª ed. (Existe edição moderna acessível, Chelsea, New York, 1967, da terceira edição, de 1756, revista e aumentada.)
- [10] Doeblin, W. (1940). Sur l'ensemble des puissances d'une loi de probabilité, *Studia Math.*, **9**, 71–96.
- [11] Feller, W. (1967). On regular variation and local limit theorems, *Proc. 5th Berkeley Symposium on Mathematical Statistics and Probability*, II, Univ. California Press, Los Angeles, 373–378.
- [12] Feller, W. (1971). *An Introduction to Probability Theory and Some of its Applications*, vol. II, Wiley, New York.
- [13] Finetti, B. de (1930). Le funzioni caratteristiche di legge istantanea, *Rend. Ac. Lincei*, (6) **12**, 278–282.

- [14] Fisher, R. A. e Tippett, L. H. C. (1928). Limiting form of the frequency distribution of the largest or smallest member of a sample, *Proc. Cambridge Philos. Soc.* **24**, 180–190.
- [15] Fréchet, M. (1927). Sur la loi de probabilité de l'écart maximum, *Ann. Soc. Polonaise Math.*, **6**, 93–116.
- [16] Galambos, J. (1987). *The Asymptotic Theory of Extreme Order Statistics*, Krieger, Florida.
- [17] Galambos, J. (1995). *Advanced Probability Theory*, 2nd ed., M. Dekker, New York.
- [18] Gnedenko, B. V. (1940). On the theory of domains of attraction of stable laws, *Uchenye Zapiski, Moskov Gos. Univ.*, **30**, 61–72.
- [19] Gnedenko, B. V. (1943) Sur la distribution limite du terme maximum d'une série aléatoire, *Ann. Math.* **44**, 423–453.
- [20] Gnedenko, B. V. e Kolmogorov, A. N. (1954). *Limit Distributions for Sums of Independent Random Variables*, Addison-Wesley, Reading, Mass.
- [21] Gomes, M. I. (1984). Penultimate limiting forms in extreme value theory, *Ann. Instit. Statist. Math.* **36A**, 71–85.
- [22] Gomes, M. I. e Pestana, D. D. (1981). On the domain of attraction of stable and of extreme value distributions, *Bull. Greek Math. Soc.* **22**, 105–120, 1981.
- [23] Gomes, M. I. e Pestana, D. D. (1987). Non-Standard Domains of Attraction and Rates of Convergence, *New Perspectives in Theoretical and Applied Statistics*, Wiley, New York, 467–477.
- [24] Graça Martins, E. e Pestana, D. D. (1987). Nonstable Limit Laws in Extreme Value Theory, *New Perspectives in Theoretical and Applied Statistics*, Wiley, New York, 449–457.
- [25] Graça Martins, E. and Pestana, D. D. (1988). The extremal limit problem — extensions, *Probability and Mathematical Statistics with Applications*, W. Grossman, J. Mogyoródi, I. Vincze and W. Wertz (eds.), Reidel, Dordrecht, 143–153.
- [26] Haan, L. de (1970). *On Regular Variation and its Application to the Weak Convergence of Sample Extremes*, Math. Centre Tracts **32**, Amsterdam.
- [27] Haldane, J. B. and Jayakar, S. D. (1963). The distribution of extremal and nearly extremal values in samples from a normal population, *Biometrika* **50**, 89–94.
- [28] Hall, A. (1998). *Extremos de Sucessões de Contagem — Do Outro Lado do Espelho*, F.C.U.L., Lisboa, tese de doutoramento.
- [29] Hardy, G. H., Littlewood, J. E., Pólya, G. (1978). *Inequalities*, Cambridge University Press, Cambridge.
- [30] Iglésias Pereira, H., Oliveira, O. e Pestana, D. D. (1996). Limites estáveis e comportamentos pré-assintóticos, *A Estatística a Decifrar o Mundo*, 109–116, Salamandra, Lisboa.
- [31] Johansen, S. (1966). An application of extreme point methods to the representation of infinitely divisible distributions, *Z. Wahrscheinlichkeitstheorie verw. Geb.* **5**, 304–316.
- [32] Johnson, N. L., Kotz, S. and Kemp, A. W. (1992). *Univariate Discrete Distributions*, Wiley, New York.

- [33] Karamata, J. (1930), Sur un mode de croissance régulière des fonctions, *Mathematica (Cluj)*, 4, 38–53.
- [34] Kendall, D. G. (1963). Extreme point methods in stochastic analysis, *Z. Wahrscheinlichkeitstheorie verw. Geb.*, 1, 295–300.
- [35] Kendall, D. G. (1967). Renewal sequences and their arithmetic, *Lecture Notes in Mathematics* 31, Springer, Berlin. (Reeditado em *Stochastic Analysis*, Kendall and Harding, eds., Wiley, New York.)
- [36] Kendall, D. G. and Harding, E. F. (1973). *Stochastic Analysis*, Wiley, London.
- [37] Kingman, J. F. C. (1963). Random walks with spherical symmetry, *Acta Mathematica*, 109, 11–53.
- [38] Klebanov, L. B., Maniya, G. M. and Melamed, I. A. (1984) A problem of Zolotarev and analogs of infinite divisibility in a scheme for summing a random number of random variables, *Theor. Probab. Appl.* 29, 791–794.
- [39] Kolmogorov, A. N. (1932) Sulla forma generale di un processo stocastico omogeneo, *Atti Acad. Naz. Lincei* (6) 15, 805–808 e 866–869.
- [40] Kovalenko, I. N. (1965). On a class of limit distributions for rarefied flows of homogeneous events, *Lit. Mat. Sbornik* 5, 569–573. (On the class of limit distributions for thinning streams of homogeneous events, *Selected Transl. Math. Statist. and Prob.* 9, Providence, Rhode Island, 1971, 75–81.)
- [41] Kozubowski, T. J. (1994). Representation and properties of geometric stable laws, *Approximation, Probability, and Related Fields*, Plenum, New York, 321–337.
- [42] Kumar, A. and Schreiber, B. M. (1978). Characterization of subclasses of class L probability distributions, *Ann. Probab.* 6, 279–293.
- [43] Kruglov, V. M. (1972). On an extension of the class of stable distributions, *Theor. Probab. Appl.*, 17, 723–732.
- [44] Laha, R. G. and Rohatgi, V. K. (1979). *Probability Theory*, Wiley, New York.
- [45] Laplace, P. H. S. de (1812). *Théorie Analytique des Probabilités*, Mme Veuve Courcier, Paris.
- [46] Lévy, P. (1925). *Calcul des Probabilités*, Gauthier-Villars, Paris.
- [47] Lévy, P. (1937). *Théorie de l'Addition des Variables Aléatoires*, Gauthier-Villars, Paris.
- [48] Linnik, Yu. V. (1953). Linear forms and statistical criteria, II, *Ukrain. Mat. Z.* 5, 247–290. (Trad. em *Selected Translations In Mathematical Statistics* 3 (1962), Amer. Math. Soc., Providence, Rhode Island.)
- [49] Logan, B. F., Mallows, C. L., Rice, S. O., and Shepp, L. A. (1973). Limit distributions of self-normalized sums, *Ann. Probab.* 1, 788–809.
- [50] Lukacs, E. (1970). *Characteristic Functions*, 2nd ed., Griffin, London.
- [51] Macis, Ju. Ju. (1971). Limit theorems in a non-classical formulation, *Theor. Probab. Appl.* 16 175–182.
- [52] Malosevskii, S. G. and Nikitin, Ja. Ju. (1977). Limit theorems without the asymptotic negligibility condition, chapt. IX in Linnik, Ju. V. and Ostrovskii, I. V. *Decomposition of Random Variables and Vectors*, Mer. Math. Soc., Providence, 285–306.

- [53] Mendonça, S. (2000). *Tópicos sobre Convergência Fraca de Sucessões de Variáveis Aleatórias*, Universidade da Madeira, tese de doutoramento.
- [54] Mejlzer, D. (1956). On the problem of the limit distribution for the maximal term of a variational series, *L'vov Pol. Inst. Nauc. Zp.*, **38**, 90–109.
- [55] Mittnik, S., and Rachev, S. T. (1993). Modeling asset returns with alternative stable distributions, *Econometric Rev.* **12**, 261–330.
- [56] Pestana, D. D., and Mendonça, S. (2001). Higher-order monotone functions and Probability Theory, *Generalized Convexity and Generalized Monotonicity*, Springer, Berlin, 317–331.
- [57] Pestana, D. D., and Velosa, S. F. (2002). *Introdução à Probabilidade e à Estatística*, Vol. I, Fundação Gulbenkian, Lisboa.
- [58] Pestana, D. D., and Velosa, S. F. (2003). Classes de Leis \mathcal{N} -Infinitamente Divisíveis, *CEAUL*, Lisboa, aceite para publicação nas Actas do X Congresso da Sociedade Portuguesa de Estatística.
- [59] Pólya, G. (1920). Über den zentralen Grenzwertsatz der Wahrscheinlichkeitsrechnung und das Momentenproblem, *Math. Zeit.* **8**, 171–181.
- [60] Rachev, S. T., and Resnick, S. (1991). Max-geometric infinite divisibility and stability, *Comm. Statist. — Stochastic Models* **7**, 191–218.
- [61] Raikov, D. A. (1938). On the decomposition of Poisson laws, *Dokl. Acad. Sci. URSS* **14**, 9–11.
- [62] Reiss, R.-D. and Thomas, M. (1991). *Statistical Analysis of Extreme Values*, Birkhäuser, Basel.
- [63] Rényi, A. (1956). A characterization of the Poisson process, *MTA Mat. Kut. Int. Közl.* **1**, 519–527 (original em húngaro, trad. inglesa em *Selected Papers of Alfred Rényi*, **1**, 1948–1956, P. Turán, ed., 622–279 Akadémiai Kiadó, Budapest; com uma nota de D. Szász sobre os desenvolvimentos posteriores, até 1976).
- [64] Resnick, S. I. (1972). Products of distribution functions attracted to extreme value laws, *J. Appl. Prob.* **8**, 781–793.
- [65] Schreiber, M. (1977). Dérivée convexe d'une loi de probabilité et lois stables. Application à un problème d'isomorphisme, *Israel J. Math.* **28**, 287–312.
- [66] Sreehari, M. (1970). On a class of limit distributions for normalized sums of independent random variables, *Theor. Probab. Appl.* **15**, 258–281.
- [67] Steutel, F. W. (1970). *Preservation of Infinite Divisibility Under Mixing and Related Topics*, Math. Centre Tracts, Amsterdam.
- [68] Urbanik, K. (1973). Limit Laws for Sequences of Normed Sums Satisfying Some Stability Conditions, *Multivariate Analysis III*, Academic Press, New York, 225–237.
- [69] Velosa, S. F. (2003). Novas classes de leis infinitamente divisíveis discretas, *CEAUL*, Lisboa, aceite para publicação nas Actas do X Congresso da Sociedade Portuguesa de Estatística.
- [70] Zinger, A. A. (1965). On a class of limit distributions for normed sums of independent random variables, *Theor. Probab. Appl.* **10**, 607–620.

- [71] Zolotarev, V. M. (1967). Generalizations of the Lindeberg–Feller theorem, *Theor. Probab. Appl.* **12**, 608–618.
- [72] Zolotarev, V. M. (1970). Théorèmes limites généraux pour les sommes de variables aléatoires indépendantes, *C. R. Acad. Sci. Paris* **270**, A899–A902.
- [73] Zolotarev, V. M. (1986). *One-Dimensional Stable Distributions*, American Mathematical Society, Rhode Island.
- [74] Zolotarev, V. M. and Koroljuk, V. S. (1961). On the hypothesis proposed by B. V. Gnedenko, *Theor. Probab. Appl.*, **6**, 431–435.

Classes de leis \mathcal{N} -infinitamente divisíveis

Dinis Duarte Pestana⁽¹⁾

*Departamento de Estatística e Investigação Operacional,
Faculdade de Ciências da Universidade de Lisboa, e
CEAUL — Centro de Estatística e Aplicações da Universidade de Lisboa
dinis.pestana@fc.ul.pt*

Sílvia Filipe Velosa⁽¹⁾

*Departamento de Matemática, Universidade da Madeira, e
CEAUL — Centro de Estatística e Aplicações da Universidade de Lisboa
svelosa@math.uma.pt*

Resumo: A aleatorização de operações deterministas proporciona em geral esclarecimentos mais profundos sobre anteriores resultados. A teoria clássica da adição de variáveis aleatórias independentes pode ser generalizada de forma não trivial considerando normalizações aleatórias (sejam elas auto-normalizações ou rarefações), e/ou somas de um número aleatório de variáveis aleatórias (*stopped sums*), que nos casos mais conhecidos podem também ser encaradas como misturas. Comentamos os pontos de contacto com funções características de Linnik, introduzimos uma classe de funções características mais rica do que a das semi- α -Laplace, e procedemos a construções análogas no esquema de máximos de um número aleatório de variáveis independentes (*stopped maxima*).

Palavras-chave: somas aleatórias, divisibilidade infinita, misturas geométricas, funções geradoras, distribuições \mathcal{N} -Gaussianas e \mathcal{N} -infinitamente divisíveis, máximos aleatórios.

Abstract: Randomization of some steps in classical frameworks may shed a new light in important problems. Random sums, namely with geometric subordinator, provide interesting counterparts of stability and infinite divisibility, and have been rediscovered in different contexts after the pioneering work of Kovalenko (1965). We comment on its relations with Linnik's characteristic functions, we introduce a wider class containing Pillai's semi- α -Laplace characteristic functions, and develop some new results for random maxima.

Keywords: stopped sums, infinite divisibility, geometric mixtures, generating functions, \mathcal{N} -Gaussian laws, \mathcal{N} -infinite divisibility, stopped maxima.

MSC2000: 60F05.

⁽¹⁾ Investigação parcialmente financiada por FCT/POCTI/FEDER (Projecto VEXTRA).

1 Introdução

O desenvolvimento da Probabilidade na primeira metade do século XX foi largamente influenciado pela revolução que Paul Lévy (1925, 1937) trouxe aos estudos do teorema limite central clássico, nomeadamente com a introdução das noções de *estabilidade* e de *divisibilidade infinita* para somas, noções que vieram a revelar-se fulcrais também no estudo de estatísticas ordinais extremas (Fréchet, 1927; Fisher and Tippett, 1928; Gnedenko, 1943; Balkema and Resnick, 1977), de produtos de variáveis aleatórias independentes (Zolotarev, 1957, 1962, 1967; Athayde 1985) e, mais geralmente, de convoluções generalizadas e seus domínios de atracção (Bingham, 1971). Sendo uma abordagem analítica límpida e económica, obscureceu temporariamente criações paralelas porventura mais ricas de potencialidades, como o teorema limite central de Lindeberg (1920, 1922) ou as somas aleatórias de variáveis aleatórias — o caso particular de Poissons compostas — de de Finetti (1930). Uma exposição magistral dos êxitos da teoria clássica encontra-se em Gnedenko and Kolmogorov (1954).

Parece-nos natural que o desenvolvimento da Probabilidade leve a generalizações de teorias estabelecidas, por aleatorização de alguns dos passos anteriormente tratados deterministicamente. A consideração de normalizações aleatórias em lugar das classes de convergência de tipos de Khinchine, sejam auto-normalizações como em Logan *et al.* (1971) ou Mendonça (2000), ou rarefações elementares como em Rényi (1956) ou Kovalenko (1965), tem permitido a construção de modelos que aprofundam os dos esquemas clássicos. Veja-se também Hall (1998), que usa rarefação elementar para investigar situações assintóticas em extremos de variáveis discretas, e Temido (2000) no que refere outras extensões do esquema clássico. Por outro lado, a teoria das somas aleatórias levou a muitas generalizações não triviais da teoria clássica das somas de variáveis aleatórias de Lévy e Khinchine, nomeadamente à definição de leis \mathcal{N} -infinitamente divisíveis (Klebanov *et al.*, 1984). Sabe-se que a Gaussiana usual é a \mathcal{N} -Gaussiana associada a variáveis aleatórias degeneradas, e a distribuição de Lapace é a \mathcal{N} -Gaussiana associada a variáveis aleatórias geométricas. Não se conhecem expressões analíticas explícitas para outros exemplos, relacionados com o processo de Galton-Watson.

Na Secção 2 colecionamos o material relevante para discutir a rarefação elementar e \mathcal{N} -divisibilidade infinita, nomeadamente a solução de um problema posto por V. M. Zolotarev: descrição da classe das variáveis aleatórias X tais que

$$\forall \theta \in (0, 1), \exists W_\theta : X \stackrel{d}{=} W_\theta + Z_\theta X,$$

com X , W_θ e $Z_\theta \sim \text{Bernoulli}(1-\theta)$ independentes. Por outro lado, comentamos as ligações entre somas aleatórias geométricas, funções densidade de probabilidade definidas positivas e funções características de Linnik, uma vez que a profunda unidade destas questões parece ter passado despercebida aos diversos investigadores que se têm ocupado delas. Na Secção 3 analisamos o problema de Zolotarev no contexto de máximos aleatórios.

Incluimos com alguma extensão resultados não inéditos por duas razões. Por um lado, são resultados ainda pouco divulgados e com desenvolvimentos potenciais interessantes, que parcialmente tivemos que reconstruir por dos trabalhos originais só haver, em inglês, resumo (como é caso da dissertação de doutoramento de Umarov (1992, U. Tashkent) sobre \mathcal{N} -divisibilidade). Por outro lado, estabelecemos relações que nos parecem interessantes entre problemas que têm sido tratadas como se fossem disjuntos, sem que os autores pareçam dar-se conta de resultados alheios relevantes, que apenas usam vocabulário diverso e foram obtidos em contexto diferente. Neste aspecto, complementa também a perspectiva apresentada por Velosa (2002) sobre somas e máximos de variáveis aleatórias, centrada nos conceitos de estabilidade e de auto-decomponibilidade, quer para somas quer para máximos, dando prioridade à abordagem determinista.

2 Somas Aleatórias e Divisibilidade Infinita

Uma variável aleatória X é infinitamente divisível se for possível decompô-la na soma de tantas parcelas i.i.d. Y_k quantas se deseje,

$$\forall n \in \mathbb{N}, \quad X \stackrel{d}{=} Y_1 + \cdots + Y_n.$$

Ou, em termos da sua função característica φ_X , $\varphi_X^{\frac{1}{n}}$ continua a ser uma função característica, para todo $n \in \mathbb{N}$ — o que um modo geral não é cómodo verificar, pois as condições necessárias e suficientes de Bochner e de Cramér para uma função ser uma transformada de Fourier de uma distribuição não são manejáveis, veja-se Lukacs (1970).

A divisibilidade infinita pode ser abordada usando esquemas triangulares assintoticamente negligíveis, cf. Feller (1971) ou Gnedenko and Kolmogoroff (1954), em que se discute de forma instrutiva por que razões é necessário usar sucessões de variáveis aleatórias duplamente indexadas para se obter esta classe. A negligibilidade assintótica é um requisito natural: a sucessão de somas parciais só pode convergir se a sucessão de parcelas for evanescente.

Por outro lado o teorema de de Finetti (1930) mostra que qualquer variável infinitamente divisível é Poisson composta ou limite (em distribuição) de uma sucessão de Poissons compostas. É uma outra forma de controlar o crescimento da soma: com elevada probabilidade o número de parcelas em geral não excederá $\lambda + 3\sqrt{\lambda}$, onde λ é o número médio de parcelas, e a soma mantém-se finita.

As somas aleatórias de parcelas aleatórias têm grande importância em diversos ramos da probabilidade. Exemplifica-se amiúde usando como subordinadora uma variável aleatória $N \sim \text{Geométrica}(p)$,

$$N = \begin{cases} k & k = 1, 2, \dots \\ p_k = p q^{k-1} & (q = 1 - p) \end{cases}$$

porque esse tipo de somas surge naturalmente em muitos contextos importantes, como processos de ramificação, teoria do risco e seguros, fiabilidade de sistemas complexos ou filas de espera.

O problema de Zolotarev, que está na génese da \mathcal{N} -divisibilidade infinita (Klebanov *et al.*, 1984) permite uma abordagem formal que se presta a desenvolvimentos diversos: pretende-se descrever a classe das variáveis aleatórias X tais que

$$\forall \theta \in (0, 1), \exists W_\theta : X \stackrel{d}{=} W_\theta + Z_\theta X,$$

com X , W_θ e $Z_\theta \sim \text{Bernoulli}(1 - \theta)$ independentes.

A função característica de Z_θ é $\varphi_{Z_\theta}(t) = \theta + (1 - \theta)e^{it}$, e a função característica de $Z_\theta X$ é $\varphi_{Z_\theta X}(t) = \int_{\mathbb{R}} (\theta + (1 - \theta)e^{itz}) dF_X(x) = \theta + (1 - \theta)\varphi_X(t)$.

Consequentemente, a condição de Zolotarev pode ser reescrita sob a forma

$$\varphi_X(t) = \varphi_{W_\theta}(t) [\theta + (1 - \theta)\varphi_X(t)] \iff \varphi_X(t) = \frac{\theta \varphi_{W_\theta}(t)}{1 - (1 - \theta)\varphi_{W_\theta}(t)}, \quad t \in \mathbb{R},$$

ou seja

$$\varphi_X(t) = \sum_{n=1}^{\infty} \theta (1 - \theta)^{n-1} \varphi_{W_\theta}^n(t), \quad t \in \mathbb{R}.$$

Conclui-se assim que X está na classe de Zolotarev se for uma soma aleatória, com subordinadora geométrica, de réplicas independentes de W_θ . Tendo em vista os resultados de Steutel (1970), todas as variáveis aleatórias que são solução do problema de Zolotarev são infinitamente divisíveis.

A classe de somas aleatórias $S_N = \sum_{k=1}^N X_k$ de variáveis i.i.d. X_k , sendo o número de parcelas $N \sim \text{Geométrica}(\theta)$ independente dos termos X_k , tem sido abordada por diversos autores e em diversos contextos. Sem pretender exaustividade:

- Rényi (1956) definiu o conceito de rarefação elementar: Seja $\{X_n\}_{n \geq 1}$ uma sucessão de variáveis aleatórias i.i.d. positivas, e a sucessão de somas parciais $\left\{Y_n = \sum_{k=1}^n X_k\right\}_{n \geq 1}$. Considere-se seguidamente a sucessão $\{Y_n^*\}_{n \geq 1}$ obtida daquela sucessão de somas parciais por *rarefação elementar*, no sentido em que na nova sucessão cada Y_k é incluído ou não, com probabilidades θ e $1 - \theta$, respectivamente, independentemente de qualquer outra soma parcial.

Denotando F_θ e L_θ a função de distribuição e a correspondente transformada de Laplace de Y_1^* , é imediato que

$$F_\theta(x) = \sum_{n=1}^{\infty} F_x^{*n}(x) \theta (1-\theta)^{n-1}, \text{ e}$$

$$L_\theta(t) = \sum_{n=1}^{\infty} L_x^n(t) \theta (1-\theta)^{n-1} = \frac{\theta L_x(t)}{1 - (1-\theta) L_x(t)}$$

Por outras palavras, $Y_1^* = \sum_{k=1}^N X_k$ com $N \sim \text{Geométrica}(\theta)$ independente das variáveis X_k . Rényi demonstrou ainda que a distribuição limite de somas aleatórias apropriadamente normalizadas daquele tipo, no caso de as parcelas serem variáveis aleatórias positivas com valor médio, é necessariamente exponencial.

- Kovalenko (1965) retomou o problema de Rényi, sem a exigência de as parcelas terem valor médio. A classe das variáveis aleatórias que surgem como possíveis limites — posteriormente denotada classe \mathcal{K} por Gnedenko (1970) — é a classe das *leis estáveis para a rarefação elementar*, cujas transformadas de Laplace verificam a equação de estabilidade

$$\forall \theta \in (0, 1), \exists a_\theta > 0: \forall t, L_\theta(t) = L(a_\theta t) = \frac{\theta L(t)}{1 - (1-\theta) L(t)},$$

donde $L(t) = \frac{1}{1+ct^\delta}$, $c > 0$ e $0 < \delta \leq 1$.

Infelizmente a situação é muito semelhante à situação clássica: as transformadas de Laplace das leis estáveis têm uma expressão analítica límpida, mas que em geral não é possível inverter decentemente: as únicas densidades estáveis para a rarefação elementar com expressões analíticas conhecidas são

$$- f(x) = \left(1 - e^{-\frac{x}{c}}\right) I_{(0,\infty)}, \text{ correspondente a } L(t) = \frac{1}{1+ct};$$

$$- f(x) = \frac{1}{\sqrt{\pi x}} - \frac{2e^x}{\sqrt{x}} \int_{\sqrt{x}}^{\infty} e^{-z^2} dz, \text{ correspondente a } L(t) = \frac{1}{1+\sqrt{t}}.$$

Gnedenko and Kovalenko (1968) usaram a classe \mathcal{K} para investigar fenómenos de *delayed waiting time* em filas de espera, iluminando nesta perspectiva a fórmula de Pollaczek-Khinchine.

- Kozubowski (1994) estudou com toda a generalidade as funções características de geométricas compostas, concluindo que a forma geral das funções

características *geo(+)*estáveis — classe que parece natural continuar a denotar \mathcal{K} — é

$$\psi(t) = \frac{1}{1 + \lambda |t|^\alpha \omega(t, \alpha, \beta) - i\mu t},$$

onde $\alpha \in (0, 2]$, $\omega(t, \alpha, \beta) = \begin{cases} 1 - \frac{t}{|t|} i\beta \operatorname{tg} \frac{\pi\alpha}{2} & \alpha \neq 1 \\ 1 + \frac{t}{|t|} i\beta \frac{2}{\pi} \ln |t| & \alpha = 1 \end{cases}$, $\lambda > 0$, $\mu \in \mathbb{R}$, e $\beta \in [-1, 1]$.

Por outras palavras, definindo a escala $\lambda = \sigma^\alpha$, uma variável aleatória Y é *geo(+)*estável se e só se puder ser escrita sob a forma

$$Y \stackrel{d}{=} \begin{cases} \mu Z + \sigma X_{\alpha, \beta} Z^{\frac{1}{\alpha}} & \alpha \neq 1 \\ \mu Z + \sigma X_{\alpha, \beta} Z + \frac{2\beta\sigma}{\pi} Z \ln(\sigma Z) & \alpha = 1 \end{cases}$$

onde $X_{\alpha, \beta}$ é uma variável aleatória estável padrão (no esquema clássico de somas) com expoente característico α e parâmetro de assimetria β , independente de $Z \sim \text{Exponencial}(1)$.

Neste contexto geral, os resultados iniciais de Rényi (1956) e de Kovalenko (1965) têm a seguinte leitura: no caso de somas aleatórias geométricas de variáveis aleatórias positivas independentes entre si e independentes da subordinadora geométrica, a exponencial tem o papel que a gaussiana desempenha no esquema clássico de estabilidade de somas. Os resultados posteriores de Kozubovski (1994) mostram que no caso de parcelas sem aquela restrição a lei limite, centrada em 0, tem função característica $\psi(t) = \frac{1}{1+\lambda|t|^\alpha}$, por outras palavras é Laplace. Assim, as “gaussianas” no esquema de somas aleatórias geométricas são as variáveis aleatórias de Laplace, um comentário que será retomado adiante.

Kozubovski (1994) não refere os resultados parciais de Kovalenko (1965) (restritos a parcelas positivas, e por isso usando transformadas de Laplace, mais cómodas). E nenhum deles refere Linnik (1953), que tinha estabelecido que $\frac{1}{1+|t|^\alpha}$ é função característica de uma variável aleatória simétrica.

De facto, se Y_α for uma variável aleatória estável positiva, com função característica $\exp(-|t|^\alpha)$, independente de X_β com função densidade de probabilidade $f_\beta(x) = \frac{\exp(-x^\beta)}{\Gamma(1+\frac{1}{\beta})} I_{(0, \infty)}$, a função característica de $W = X_\beta^{\frac{\alpha}{\beta}} Y_\alpha$ é $\psi(t) = \frac{1}{(1+|t|^\alpha)^{\frac{1}{\beta}}}$. Com $\beta = 1$, i. e., $X_1 \sim \text{Exponencial}(1)$, obtém-se a família

de funções características de Linnik, que Pestana *et al.* (2001) relacionaram a classe \mathcal{K} com a investigação de pares recíprocos de variáveis aleatórias.

As funções características de Linnik fazem parte de uma classe mais vasta de funções características da forma $\psi(t) = \frac{1}{1+f(t)}$, onde $f(t)$ é uma função

contínua com $f(0) = 0$ com algumas propriedades adicionais. Pillai (1985) investigou a classe das funções características a que chamou semi- α -Laplace de ordem $b \in (0, 1)$ e expoente $\alpha \in (0, 2]$, caracterizadas por $f(t) = b^{\frac{1}{\alpha}} f(bt)$.

Estas funções características são misturas que exibem uma curiosa propriedade de auto-decomponibilidade:

$$\exists p \in (0, 1) \text{ tal que } \psi(t) = p\psi(bt) + (1-p)\psi(t)\psi(bt).$$

É fácil de ver que a função característica da exponencial, por exemplo, verifica aquela equação funcional.

Uma classe mais geral do que a de Pillai é aquela em que $f(t) = -\ln \omega(t)$, onde $\omega(t)$ é a função característica de uma variável aleatória infinitamente divisível no esquema clássico. Neste ponto, é de recordar que o grande resultado de Johansen (1966) sobre representações de funções características infinitamente divisíveis assenta na sua descoberta de que os correspondentes logaritmos são funções definidas positivas.

Como uma função característica infinitamente divisível ω não tem zeros reais (pelo que tem sentido definir o seu logaritmo), e facilmente se estabelece que $\omega^{\frac{1}{n}}(t) \xrightarrow{n \rightarrow \infty} 1$ e $n \left[\omega^{\frac{1}{n}}(t) - 1 \right] \xrightarrow{n \rightarrow \infty} \ln \omega(t)$, tem-se que $\psi(t) = \frac{1}{1 - \ln \omega(t)}$ é uma função característica.

De facto, a função característica de $S_N = \sum_{k=1}^N W_k$ onde $N \sim \text{Geométrica}(\theta)$ independente das parcelas W_k com função característica ω^θ é

$$\begin{aligned} \psi(t) &= \sum_{n=1}^{\infty} \left[\omega^\theta(t) \right]^n \theta (1-\theta)^{n-1} = \frac{\theta \omega^\theta(t)}{1 - (1-\theta) \omega^\theta(t)} = \\ &= \frac{1}{1 + \frac{1 - \omega^\theta(t)}{\theta \omega^\theta(t)}} = \frac{1}{1 - \frac{\frac{1}{\theta} [\omega^\theta(t) - 1]}{\omega^\theta(t)}} \xrightarrow{\theta \rightarrow 0} \frac{1}{1 - \ln \omega(t)}. \end{aligned}$$

Assim, tal como se pode definir geo-estabilidade, também se pode definir geo-divisibilidade infinita: Uma variável aleatória com função característica ψ é geo-infinitamente divisível se puder ser decomposta numa soma aleatória de parcelas i.i.d., com subordinador geométrico com parâmetro $\theta \in (0, 1)$. Dos resultados acima decorre que uma função característica geo-infinitamente divisível não tem zeros reais, $\forall \theta \in (0, 1)$, $\psi^\theta(t)$ é uma função característica, e ψ é uma função característica geo-infinitamente divisível se e só se for da forma $\psi(t) = \frac{1}{1 - \ln \omega(t)}$, onde ω é uma função característica infinitamente divisível no esquema clássico de somas de variáveis i.i.d. uniformemente assintoticamente negligíveis.

As funções características de Linnik e semi- α -Laplace são então casos especiais destas que agora introduzimos.

As somas aleatórias $S_N = \sum_{k=0}^N X_k$ de variáveis aleatórias i.i.d. X_k em que o

número de parcelas N é uma variável aleatória de contagem independente das parcelas têm ajudado a esclarecer muitos aspectos da teoria clássica das somas, e proporcionado extensões não triviais de grande interesse. Uma das questões que suscitaram foi a da definição geral de leis \mathcal{N} -infinitamente divisíveis, e em particular de distribuições \mathcal{N} -Gaussianas — acima comentámos, a título de exemplo, que a Laplace preenche o papel da Gaussiana no caso de somas aleatórias com subordinadora geométrica.

Por analogia com a equação de estabilidade para a Gaussiana Z , no sentido clássico, $Z \stackrel{d}{=} \sum_{k=0}^n \frac{Z_k}{\sqrt{n}}$, dizemos que X é gaussiana associada à família de variáveis aleatórias de contagem $\mathcal{N} = \{N_\theta, \theta \in \Theta\}$ — em que assumimos que existe $\mathbb{E}(N_\theta), \forall \theta \in \Theta \subseteq (0, 1)$, e que a parametrização é tal que $\mathbb{E}(N_\theta) = \frac{1}{\theta}$ — se e só se a sua função característica φ_x verificar a equação funcional

$$\varphi_x(t) = \varphi_{\sum_{k=0}^n x_k}(\sqrt{\theta}t) = \sum_{n=0}^{\infty} \varphi_x^n(\sqrt{\theta}t) \mathbb{P}[N_\theta = n] = P_{N_\theta}(\varphi_x(\sqrt{\theta}t)),$$

onde P_{N_θ} é a função geradora de probabilidades de N_θ .

Nem sempre existe distribuição \mathcal{N} -Gaussiana associada a uma família de variáveis aleatórias de contagem $\mathcal{N} = \{N_\theta\}_{\theta \in \Theta}$. A existência de distribuição \mathcal{N} -Gaussiana associada à variável N_θ depende da resolubilidade da equação funcional $\varphi_x(t) = P_{N_\theta}(\varphi_x(\sqrt{\theta}t))$, para todo $t \in \mathbb{R}$ e para todo $\theta \in \Theta$. Definindo $\phi_x(t) = \varphi_x(\sqrt{t})$, aquela equação é equivalente à equação funcional de Poincaré

$$\phi_x(t) = P_{N_\theta}(\phi_x(\theta t)), \quad \forall t \geq 0, \theta \in \Theta.$$

que é resolúvel se e só se o semigrupo das funções geradoras de probabilidades P_{N_θ} (algebrizado com a operação de composição) for comutativo.

Mais geralmente, seja $\mathcal{N} = \{N_\theta\}_{\theta \in \Theta}$ uma família de variáveis aleatórias de contagem, e admita-se que algebrizando com a operação de composição a correspondente família de funções geradoras de probabilidade

$$\mathcal{P} = \left\{ P_\theta(s) = \mathbb{E}(s^{N_\theta}) \right\}_{\theta \in \Theta}$$

se tem um semi-grupo comutativo.

Diz-se que função de distribuição F é \mathcal{N} -infinitamente divisível se para todo $\theta \in \Theta$ existir uma função de distribuição F_θ tal que

$$F(x) = \sum_{n=1}^{\infty} F_\theta^{*n}(x) \mathbb{P}(N_\theta = n), \quad x \in \mathbb{R}.$$

Isto constata-se facilmente para duas famílias $\mathcal{N} = \{N_\theta\}$:

1. Família das N_θ degeneradas em $\frac{1}{\theta}$, $\theta = \frac{1}{n}$, $n = 1, 2, \dots$ (a parametrização é feita de forma a $\mathbb{E}(N_\theta) = \frac{1}{\theta}$).

De facto, $P_{N_\theta}(s) = P_\theta(s) = s^{\frac{1}{\theta}}$, e consequentemente $P_{\theta_1}\left(P_{N_{\theta_0}}(s)\right) = \left(s^{\frac{1}{\theta_0}}\right)^{\frac{1}{\theta_1}} = P_{\theta_0}\left(P_{N_{\theta_1}}(s)\right)$.

Esta família viabiliza encarar o esquema clássico como um caso especial deste esquema aleatório mais vasto da \mathcal{N} -divisibilidade infinita.

2. Família das $N_\theta \curvearrowright \text{Geométrica}(\theta)$. Como $P_{N_\theta}(s) = P_\theta(s) = \frac{\theta s}{1 - (1 - \theta)s}$,

$$P_{\theta_1}\left(P_{N_{\theta_0}}(s)\right) = \frac{\theta_1 \frac{\theta_0 s}{1 - (1 - \theta_0)s}}{1 - (1 - \theta_1) \frac{\theta_0 s}{1 - (1 - \theta_0)s}} = \frac{\theta_0 \theta_1 s}{1 - (1 - \theta_0 \theta_1)s} = P_{\theta_0}\left(P_{N_{\theta_1}}(s)\right).$$

Atrás, usando argumentos mais simples, construímos as classes de \mathcal{N} -estáveis e de \mathcal{N} -infinitamente divisíveis desta família.

Esperávamos que a família de variáveis aleatórias

$$N_\theta = \begin{cases} 2k + 1 & k = 0, 1, \dots \\ \frac{2\theta}{1+\theta} \left(\frac{1-\theta}{1+\theta}\right)^k & \end{cases},$$

que surge como caso limite em Velosa (2002a), fornecesse outro exemplo. Mas o semi-grupo das suas funções geradoras de probabilidades $P_\theta(s) = \frac{2\theta s}{1+\theta-(1-\theta)s^2}$, algebrizado com a composição de funções, não é comutativo.

3 Máximos aleatórios

Hall (1998) mostrou a relevância de, no estudo de leis limites de extremos de sucessões de variáveis aleatórias discretas, usar a rarefação elementar em vez de estabilização de localização e escala por transformações lineares — é de notar que assim se recorre a uma aleatorização em vez de uma transformação determinística.

Velosa (2002), transpondo o problema de Zolotarev para máximos seguindo um trajecto diverso de Rachev and Resnick (1991), deduziu a forma geral das funções de distribuição max(geo)estáveis,

$$F(x) = \frac{1}{1 + (1 + \gamma x)^{-\frac{1}{\gamma}}} \mathbf{I}_{\{x: 1 + \gamma x > 0\}}, \quad \gamma \in \mathbb{R},$$

(tendo como caso limite a distribuição logística, $F(x) = \frac{1}{1+e^{-x}} \mathbb{I}_{\mathbb{R}}$, quando $\gamma \rightarrow 0$), caracterizando domínios de atracção quando $\gamma \neq 0$, à custa de equivalência de caudas.

De forma análoga ao atrás exposto, pode definir-se \mathcal{N} -max-divisibilidade infinita. No caso de

$$\mathcal{N} = \{N_\theta, \theta \in \Theta\}$$

ser a família das $N_\theta \sim \text{Geométrica}(\theta)$, $\theta \in (0, 1)$, a construção é em tudo análoga ao que atrás fizemos no esquema de somas aleatórias geométricas, apenas há que ter o cuidado de trabalhar no suporte das variáveis aleatórias, e notar que qualquer potência positiva de uma função de distribuição é uma função de distribuição, veja-se em Velosa (2002) os comentários sobre max-estabilidade univariada:

Seja X uma variável aleatória com função de distribuição F e suporte S ; $\forall x \in S$, $n \left[F^{\frac{1}{n}}(x) - 1 \right] \xrightarrow{n \rightarrow \infty} \ln F(x)$ e $F^{\frac{1}{n}}(x) \xrightarrow{n \rightarrow \infty} 1$.

A função de distribuição $G(x) = \frac{1}{1 - \ln F(x)}$ pode então ser interpretada como distribuição limite do máximo de um número aleatório geométrico (*geometric stopped maxima*) de variáveis i.i.d.:

De facto, $X_{N:N} = \max_{1 \leq k \leq N} W_k$, onde $N \sim \text{Geométrica}(\theta)$ independente dos termos W_k com função de distribuição F^θ , tem função de distribuição

$$\begin{aligned} G(x) &= \sum_{n=1}^{\infty} \left[F^\theta(x) \right]^n \theta (1-\theta)^{n-1} = \frac{\theta F^\theta(x)}{1 - (1-\theta) F^\theta(x)} = \\ &= \frac{1}{1 + \frac{1-F^\theta(x)}{\theta F^\theta(x)}} = \frac{1}{1 - \frac{\frac{1}{\theta} [F^\theta(x)-1]}{F^\theta(x)}} \xrightarrow{\theta \rightarrow 0} \frac{1}{1 - \ln F(x)}. \end{aligned}$$

Podemos assim exprimir as funções de distribuição max-geo-infinitamente divisíveis sob a forma $G(x) = \frac{1}{1 - \ln F(x)}$, o que em si mesmo tem pouco interesse, mas que abre a perspectiva de desenvolvermos uma teoria da geo(max)autodecomponibilidade mais ampla do que a inspirada nas questões postas por Zolotarev para somas.

Bibliografia

- [1] Athayde, E. (1985). *Estudos Sobre Multiplicação de Variáveis Aleatórias Independentes*, Universidade de Lisboa.
- [2] Balkema, A. A. and Resnick, S. I. (1977). Max-infinite divisibility, *J. Appl. Prob.*, **14**, 309–319.
- [3] Bingham, N. H. (1971). Factorization theory and domains of attraction for generalized convolution algebras, *Proc. London Math. Soc.* (3), **23**, 16–30.

- [4] Feller, W. (1968). *An Introduction to Probability Theory and Some of its Applications*, vol. I, Wiley, New York.
- [5] Finetti, B. de (1930). Le funzioni caratteristiche di legge istantanea, *Rend. Ac. Lincei*, (6) **12**, 278–282.
- [6] Fisher, R. A. and Tippett, L. H. C. (1928). Limiting form of the frequency distribution of the largest or smallest member of a sample, *Proc. Cambridge Philos. Soc.* **24**, 180–190.
- [7] Fréchet, M. (1927). Sur la loi de probabilité de l'écart maximum, *Ann. Soc. Polonaise Math.*, **6**, 93–116.
- [8] Gnedenko, B. V. (1943) Sur la distribution limite du terme maximum d'une série aléatoire, *Ann. Math.* **44**, 423–453.
- [9] Gnedenko, B. V. (1970). Limit theorems for sums of a random number of positive independent random variables, *Proc. 6th Berkeley Symp. Math. Statist. Probab.*, vol. 2, 537–549, California University Press, Berkeley.
- [10] Gnedenko, B. V. and Kolmogorov, A. N. (1954). *Limit Distributions for Sums of Independent Random Variables*, Addison-Wesley, Reading, Mass.
- [11] Gnedenko, B. V. and Kovalenko, I. N. (1968). *Introduction to Queueing Theory*, Program for Scientific Translations, Jerusalem.
- [12] Hall, A. (1998). *Extremos de Sucessões de Contagem — Do Outro Lado do Espelho*, F.C.U.L., Lisboa.
- [13] Johansen, S. (1966). An application of extreme point methods to the representation of infinitely divisible distributions, *Z. Wahrscheinlichkeitstheorie verw. Geb.* **5**, 304–316.
- [14] Klebanov, L. B., Manija, G. M. and Melamed, I. A. (1984) A problem of Zolotarev and analogs of infinite divisibility in a scheme for summing a random number of random variables, *Theor. Probab. Appl.* **29**, 791–794.
- [15] Kovalenko, I. N. (1965). On a class of limit distributions for rarefied flows of homogeneous events, *Lit. Mat. Sbornik* **5**, 569–573. (*Selected Transl. Math. Statist. and Prob.* **9**, Providence, Rhode Island, 1971, 75– 81.)
- [16] Kozubowski, T. J. (1994). Representation and properties of geometric stable laws, *Approximation, Probability, and Related Fields*, Plenum, New York. 321–337.
- [17] Kozubowski, T. J. and Rachev, S. T. (1999a). Univariate geometric stable laws, *J. Comp. Anal. Appl.* **1**, 177–217.
- [18] Lévy, P. (1925). *Calcul des Probabilités*, Gauthier-Villars, Paris.
- [19] Lévy, P. (1937). *Théorie de l'Addition des Variables Aléatoires*, Gauthier-Villars, Paris.
- [20] Lindeberg, J. W. (1920). Über das Exponentialgesetz in der Wahrscheinlichkeitsrechnung, *Ann. Acad. Sci. Fenn.* **16**, 1–23.
- [21] Lindeberg, J. W. (1922). Eine neue Herleitung des Exponentialgesetzes in der Wahrscheinlichkeitsrechnung, *Math. Zeitschr.* **15**, 211–225.
- [22] Linnik, Yu. V. (1953). Linear forms and statistical criteria, II, *Ukrain. Mat. Z.* **5**, 247–290. (Trad. em *Selected Translations In Mathematical Statistics* **3** (1962), Amer. Math. Soc., Providence, Rhode Island.)

- [23] Logan, B. F., Mallows, C. L., Rice, S. O., and Shepp, L. A. (1973). Limit distributions of self-normalized sums, *Ann. Probab.* **1**, 788–809.
- [24] Mendonça, S. (2000). *Tópicos sobre Convergência Fraca de Sucessões de Variáveis Aleatórias*, Universidade da Madeira.
- [25] Pestana, D. D., Sequeira, F. and Velosa, S. F. (2001). Parseval's relation and self-reciprocal characteristic functions, *Rev. Estat./Statist. Rev.* — 23rd European Meeting of Statisticians, *Contributed Papers II*, 315–316.
- [26] Pillai, R. N. (1985). Semi- α Laplace distributions, *Communications in Statistics — Theory and Methods* **14**, 991–1000.
- [27] Rachev, S. T., and Resnick, S. (1991). Max-geometric infinite divisibility and stability, *Comm. Statist.* — *Stochastic Models* **7**, 191–218.
- [28] Rényi, A. (1956). A characterization of the Poisson process, *MTA Mat. Kut. Int. Közl.* **1**, 519–527 (*Selected Papers of Alfred Rényi*, **1**, 1948–1956, P. Turán, ed., 622–279 Akadémiai Kiadó, Budapest).
- [29] Steutel, F. W. (1970). *Preservation of Infinite Divisibility Under Mixing and Related Topics*, Math. Centre, Amsterdam.
- [30] Temido, M. G. (2000). *Classes de Leis Limites em Teoria de Valores Extremos — Estabilidade e Semiestabilidade*, DM, FCTUC, Coimbra.
- [31] Velosa, S. F. (2003). Somas e Máximos de Variáveis Aleatórias Independentes — Classes de Leis Limites, *CEAUL*, Lisboa.
- [32] Velosa, S. F. (2003a). Novas Classes de Leis Infinitamente Divisíveis Discretas, *CEAUL*, Lisboa.
- [33] Zolotarev, V. M. (1957). Mellin-Stieltjes transforms in Probability Theory, *Theor. Probab. Appl.* **2**, 433–460.
- [34] Zolotarev, V. M. (1962). On the general theory of multiplication of independent random variables, *Soviet Math. Dokl.* **3**, 166–170.
- [35] Zolotarev, V. M. (1967). On the M-divisibility of a stable law, *Theor. Probab. Appl.* **12**, 506–508.

1.1 Comentários e Complementos

A evolução deste trabalho de investigação veio a centrar-se sobre variáveis aleatórias discretas, e nesse sentido é imperativo juntar algumas indicações sobre modelos de contagem.

Johnson *et al.* (1992, p. 324–325) consideram que os três resultados mais profundos da teoria da probabilidade discreta são o teorema de Lévy sobre a identidade de Poissons compostas e infinitamente divisíveis discretas, o teorema de Maceda (1948) sobre a divisibilidade infinita de misturas discretas de Poissons se e só se a lei que governa a mistura for ela própria infinitamente divisível, e o teorema de Gurland (1957) identificando leis compostas e misturas de variáveis aleatórias, sempre que a função geradora de probabilidade \mathcal{G} da subordinadora dependa de um parâmetro θ , verificando a equação funcional $\mathcal{G}(z \mid k\theta) = [\mathcal{G}(z \mid \theta)]^k$.

O teorema de Lévy foi consideravelmente generalizado por de Finetti (1929), que mostrou que a classe das infinitamente divisíveis é constituída por Poissons compostas e limites de Poissons compostas. Por outro lado Rényi (1956) e Kovalenko (1965) iniciaram o estudo das somas aleatórias subordinadas por uma variável geométrica, o que veio a originar as novas áreas da geo-divisibilidade infinita, e depois da \mathcal{N} -divisibilidade infinita.

Creio nunca ter sido anotado que o notável trabalho de Panjer (1981) desenvolvendo expressões recursivas para o cálculo do processo de risco contém implicitamente aqueles célebres resultados, e define classes mais vastas, considerando somas aleatórias subordinadas por binomiais, por binomiais negativas, e por Poissons. No Capítulo 2, estas classes são ainda generalizadas de forma não trivial, em trabalho conjunto com Pestana. Não foi possível, no entanto, encontrar uma expressão recursiva para o cálculo das densidades que lhes conferisse a utilidade imediata da classe de Panjer. À luz do teorema de Gurland (1957), podem os referidos resultados ser perspectivados na óptica da dualidade entre misturas e somas aleatórias, esclarecendo a sua interpretação.

No Capítulo 3, o comportamento de algumas fêmeas de pássaros, que decerto cantam como a Dorabella e a Fiordiligi da ópera *Così Fan Tutte*, levou-nos a reflectir sobre o papel que a investigação das estruturas de aleatoriedade deve desempenhar na modelação de fenómenos aleatórios. Além de integrarmos nesse capítulo um trabalho conjunto com Marques e Pestana, apresentamos algumas anotações sublinhando um aspecto a que não se dá, a meu ver, a relevância devida: a eventual transfiguração radical do espaço dos parâmetros quando se faz a truncatura do suporte

de uma distribuição.

A truncatura foi usada por Sundt and Jewell (1981) e por Willmot (1987) na generalização dos resultados de Panjer (1981). Recentemente, Hess *et al.* (2002) propuseram uma classe muito interessante de distribuições a que chamaram muito apropriadamente *general claim number distributions*. De facto, correspondem às variáveis de contagem que têm importância efectiva como subordinadoras de somas aleatórias. Nos complementos ao Capítulo 3, fornecemos demonstrações completas da caracterização dos membros dessa classe.

No capítulo 4 apresentamos alguns resultados de interesse independente, sobre o que nos pareceu apropriado designar por transformadas de Pareto.

Capítulo 2

Classes de Panjer e Extensões

2.1 Introdução

As funções massa de probabilidade das variáveis de contagem mais usadas — binomiais, de Poisson, e binomiais negativas — verificam a equação

$$p_{n+1} = \left(a + \frac{b}{n+1}\right) p_n, \quad n = 0, 1, 2, \dots$$

Esta expressão recursiva foi originalmente usada por Katz (1965) para definir uma família ampla de distribuições, que Johnson, Kotz and Kemp (2003) consideram ter entre as variáveis aleatórias discretas um papel análogo ao que têm as curvas de Pearson no caso contínuo. A fórmula ganhou nova importância quando Panjer (1981) descobriu como facilitava a determinação da função de distribuição de indenizações agregadas no processo do risco, em teoria dos seguros.

Na sequência do trabalho seminal de Panjer, Sundt e Jewell (1981) provaram que binomial, Poisson e binomial negativa são os únicos modelos de contagem (não degenerados) que verificam aquela expressão recursiva — por simplicidade, passaremos a referi-las colectivamente como variáveis de Panjer.

Mais, mostra-se que a lei da soma aleatória

$$S_N = \sum_{k=1}^N X_k,$$

onde as variáveis X_k são réplicas independentes do modelo X para o valor das indenizações, e independentes do número de indenizações N com

distribuição de Panjer, pode ser calculada iterativamente. De facto, em vez de ser necessário calcular as convoluções associadas, como é habitual, basta usar um algoritmo iterativo simples.

A expressão recursiva para a função de distribuição das indemnizações agregadas pode ser considerada uma extensão da fórmula a que a função massa de probabilidade da subordinadora N obedece. Considerando que os valores X_k das indemnizações individuais são variáveis aleatórias discretas positivas com função massa de probabilidade f_j , a função de distribuição da soma aleatória S_N é dada nos inteiros positivos por

$$g_i = \sum_{j=1}^i \left(a + \frac{bj}{i} \right) f_j g_{i-j}, \quad i = 1, 2, 3, \dots$$

Existem generalizações simples para $f_0 > 0$ e para parcelas X_k contínuas (veja-se Panjer (1981), e especialmente Sundt and Jewell (1981)).

A demonstração de Panjer, quando especializada para o caso de parcelas discretas e subordinadora $N \sim \text{Poisson}(\mu)$ ou $N \sim \text{Geométrica}(p)$, respectivamente, fornece também a representação das funções geradoras de probabilidades das leis infinitamente divisíveis discretas como Poissons compostas e das geo-infinitamente divisíveis discretas como geométricas compostas, como se detalha na secção de comentários e complementos.

A classe de Panjer foi imediatamente generalizada por Sundt e Jewell (1981), que mostraram que a variável logarítmica verificava a expressão

$$p_{n+1} = \left(a + \frac{b}{n+1} \right) p_n, \quad n = 1, 2, \dots,$$

gozando, como subordinadora de somas aleatórias, de características tão boas quanto as clássicas distribuições de Panjer.

Willmot (1987) mostrou que a distribuição de Engen verifica a mesma recursão para $n = 1, 2, \dots$, e mais recentemente Hess *et al.* (2002) introduziram a noção de *basic claim number distributions* como aquelas cuja função massa de probabilidade verifica

$$p_{n+1} = \left(a + \frac{b}{n+1} \right) p_n, \quad n = k, k+1, k+2, \dots$$

para algum k inteiro não negativo. Podemos descrevê-las brevemente usando o conceito de variáveis modificadas em $\{0, 1, \dots, k-1\}$: são as

variáveis binomiais, de Poisson, binomiais negativas, logarítmicas e binomiais negativas generalizadas de Engen, em cada caso modificadas naquela secção inicial de \mathbb{N} .

A extensão da classe de Panjer de que se ocupa o trabalho que é parte integrante deste capítulo baseia-se em uma visão estrutural artificial da sua construção usando momentos de uniformes. Como na situação clássica, usamos a expressão iterativa mais geral que investigamos para estabelecer uma equação funcional que a função geradora de probabilidades deve verificar. Na secção de complementos e comentários apresentamos resultados completos sobre os ternos (α, β, γ) para os quais a solução da equação funcional é, de facto, uma função absolutamente monótona normada, isto é uma função geradora de probabilidades.

Não tivemos sucesso em um objectivo que surgiu como extensão natural dos resultados que alcançámos: seria interessante desenvolver expressões iterativas para a lei de somas aleatórias subordinadas por variáveis de contagem das classes de variáveis de contagem assim definidas, mas a sua complexidade não permitiu tal extensão.

A motivação inicial desta extensão foi a constatação de que nas classes assim definidas, se $0 < \gamma_1 < \gamma_2 \leq 1$ então $\mathcal{C}_{\gamma_1} \subset \mathcal{C}_{\gamma_2}$. Como a nossa construção permite a definição de classes \mathcal{C}_γ para $\gamma \in (-1, 1]$, esperávamos provar que aquela inclusão era válida nesse domínio alargado. Tal não é, infelizmente, verdade. Assim, um outro objectivo nosso — exhibir leis \mathcal{N} -gaussianas para além das trivialmente conhecidas — ficou também gorado. (Note-se que, para além da gaussiana clássica que está associada, no esquema de \mathcal{N} -divisibilidade, às uniformes discretas, o exemplo conhecido é a Laplace, que surge no contexto das geométricas compostas, a nossa classe \mathcal{C}_0 .)

Aprofundamos, também, nessa secção final, o estudo das classes \mathcal{C}_γ para $\gamma \in (-1, 0)$, e fazemos observações estruturais que complementam o material publicado na *RevStat*.

EXTENSIONS OF KATZ–PANJER FAMILIES OF DISCRETE DISTRIBUTIONS *

Authors: DINIS D. PESTANA

– Departamento de Estatística e Investigação Operacional,
Faculdade de Ciências da Universidade de Lisboa, Bloco C6, Piso 4,
Campo Grande, 1749-016 Lisboa, Portugal, e
CEAUL – Centro de Estatística e Aplicações da Universidade de Lisboa
(dinis.pestana@fc.ul.pt)

SÍLVIO F. VELOSA

– Departamento de Matemática e Engenharias, Universidade da Madeira,
Campus Universitário da Penteada, 9000-390 Funchal, Portugal, e
CEAUL – Centro de Estatística e Aplicações da Universidade de Lisboa
(sfilipe@uma.pt)

Received: July 2004

Revised: September 2004

Accepted: September 2004

Abstract:

- Let $N_{\alpha, \beta, \gamma}$ be a discrete random variable whose probability atoms $\{p_n\}_{n \in \mathbb{N}}$ satisfy $\frac{f(n+1)}{f(n)} = \alpha + \beta \frac{\mathbb{E}(U_\gamma^n)}{\mathbb{E}(U_\gamma^k)}$, $n=0, 1, \dots$, for some $\alpha, \beta \in \mathbb{R}$, where $U_\gamma \sim \text{Uniform}(\gamma, 1)$, $\gamma \in (-1, 1]$. When $\gamma \rightarrow 1$, $U_\gamma \rightarrow U_1$, the degenerate random variable with unit mass at 1, and the above iterative expression is $\frac{p_{n+1}}{p_n} = \alpha + \frac{\beta}{n+1}$ for $n = k, k+1, \dots$, used by Katz and by Panjer ($k = 0$), by Sundt and Jewell and by Willmot ($k = 1$) and, for general $k \in \mathbb{N}$, by Hess, Lewald and Schmidt. We investigate the case $U_\gamma \sim \text{Uniform}(\gamma, 1)$ with $\gamma \in (-1, 1)$ in detail for $\alpha = 0$. We then construct classes \mathcal{C}_γ of discrete infinitely divisible randomly stopped sums such that $N_{0, \beta, \gamma} \in \mathcal{C}_\gamma$. \mathcal{C}_0 is the class of compound geometric random variables, \mathcal{C}_1 is the class of compound Poissons, and $|\gamma_1| < \gamma_2 \leq 1$ implies $\mathcal{C}_{\gamma_1} \subset \mathcal{C}_{\gamma_2} \subseteq \mathcal{C}_1$.

Key-Words:

- *Poisson stopped sums (compound Poisson); geometric stopped sums (compound geometric); Panjer's algorithm.*

AMS Subject Classification:

- 60G50, 60E10, 91B30.

*Research partially supported by FCT/POCTI/FEDER.

1. INTRODUCTION

Let us consider the discrete random variables $N_{\alpha,\beta}$ whose probability mass functions (p.m.f.) $\{f_{N_{\alpha,\beta}}(n)\}_{n \in \mathbb{N}}$ satisfy

$$(1.1) \quad f_{N_{\alpha,\beta}}(n+1) = \left(\alpha + \frac{\beta}{n+1} \right) f_{N_{\alpha,\beta}}(n), \quad \alpha, \beta \in \mathbb{R}, \quad n = 0, 1, \dots$$

From (1.1) it follows that $f_{N_{\alpha,\beta}}(n) = f_{N_{\alpha,\beta}}(0) \prod_{k=1}^n \left(\alpha + \frac{\beta}{k} \right)$. In particular,

$$f_{N_{\alpha,0}}(n) = f_{N_{\alpha,0}}(0) \alpha^n = (1-\alpha) \alpha^n \implies N_{\alpha,0} \sim \text{Geometric}(1-\alpha),$$

and we may write

$$(1.2) \quad f_{N_{\alpha,0}}(n+1) = \alpha f_{N_{\alpha,0}}(n) = \sum_{k=0}^n f_{N_{\alpha,0}}(k) r_{n-k},$$

where $r_0 = \alpha$ is the ratio of a geometric series and $r_1 = \dots = r_n = 0$.

On the other hand,

$$f_{N_{0,\beta}}(n) = f_{N_{0,\beta}}(0) \prod_{k=1}^n \frac{\beta}{k} = f_{N_{0,\beta}}(0) \frac{\beta^n}{n!} = e^{-\beta} \frac{\beta^n}{n!} \implies N_{0,\beta} \sim \text{Poisson}(\beta),$$

and we may write

$$(1.3) \quad (n+1) f_{N_{0,\beta}}(n+1) = \beta f_{N_{0,\beta}}(n) = \sum_{k=0}^n f_{N_{0,\beta}}(k) r_{n-k},$$

where $r_0 = \beta$ and $r_1 = \dots = r_n = 0$. Note that similar expressions do not hold for randomly stopped sums $S_{N_{\alpha,\beta}} = S_{N_{\alpha,\beta}}(Y) = \sum_{k=1}^{N_{\alpha,\beta}} Y_k$, where the summands Y_k are i.i.d. and independent of the subordinator $N_{\alpha,\beta}$, with p.m.f. satisfying (1.1), whenever both $\alpha \neq 0$ and $\beta \neq 0$. However, for geometric stopped sums $\sum_{k=1}^{N_{\alpha,0}} Y_k$ and for Poisson stopped sums, $\sum_{k=1}^{N_{0,\beta}} Y_k$ (i.e., when either $\beta = 0$ or $\alpha = 0$) we get nice similar expressions, with the $r_k \geq 0$ and convergence of $\sum_{k=0}^{\infty} r_k$, in the case of geometric stopped sums, and convergence of $\sum_{k=0}^{\infty} \frac{r_k}{k+1}$, for Poisson stopped sums. In the definition of randomly stopped sums, $\mathbb{P}[S_{N_{\alpha,\beta}}=0 \mid N_{\alpha,\beta}=0] = 1$, and therefore $\mathbb{P}[S_{N_{\alpha,\beta}}=0] = \mathbb{P}[N_{\alpha,\beta}=0] = f_{N_{\alpha,\beta}}(0)$ whenever $\mathbb{P}[Y_k > 0] = 1$.

Panjer (1981) has remarked that the discrete (nondegenerate) random variables whose p.m.f.'s satisfy equation (1.1) are

- $N_{0,\beta} \sim \text{Poisson}(\beta)$, $\beta > 0$,
- $N_{\alpha,\beta} \sim \text{Binomial}\left(-1 - \frac{\beta}{\alpha}, \frac{\alpha}{\alpha-1}\right)$, in case $\alpha < 0$ and $-\frac{\beta}{\alpha} \in \mathbb{N}^+$, and
- $N_{\alpha,\beta} \sim \text{NegativeBinomial}\left(\frac{\alpha+\beta}{\alpha}, 1 - \alpha\right)$ if $\alpha \in (0, 1)$ and $\alpha + \beta > 0$.

The dispersion index $\frac{\text{var}(N_{\alpha,\beta})}{\mathbb{E}(N_{\alpha,\beta})} = \frac{1}{1-\alpha}$ is less than 1 (underdispersion) for the binomial and greater than 1 (overdispersion) for the negative binomial. On the other hand, $N_{0,\beta} \sim \text{Poisson}(\beta)$ is a yardstick, with dispersion index 1. We denote Π the class of random variables $N_{\alpha,\beta}$ described above.

These random variables play an important role as subordinators in randomly stopped sums. Compound or generalized random variables (other names traditionally given to $S_{N_{\alpha,\beta}}$, cf. the discussion on terminology in Johnson, Kotz and Kemp, 1992) are at the core of branching processes and many other subjects where the aim is to obtain the distribution of randomly stopped sums, namely in the study of aggregate claims in the risk process, see Klugman, Panjer and Willmot (1998) and Rólski, Schmidli, Schmidt and Teugels (1999).

Katz (1965) had used an iterative expression equivalent to (1.1) to organize a coordinated presentation of count distributions. Panjer's (1981) pathbreaking result has been to use the iterative expression satisfied by the p.m.f. of the subordinator $N_{\alpha,\beta}$ to get an iterative algorithm to compute the density function (probability mass function or probability density function) of $S_{N_{\alpha,\beta}}$. This is used in section 2 to establish characterization theorems for infinitely divisible and for geometric infinitely divisible generating functions.

In section 3, we investigate discrete random variables $N_{\alpha,\beta,\gamma}$ whose probability mass function (p.m.f.) $\{p_n\}_{n \in \mathbb{N}}$ satisfies the more general relation

$$(1.4) \quad \frac{f_{N_{\alpha,\beta,\gamma}}(n+1)}{f_{N_{\alpha,\beta,\gamma}}(n)} = \alpha + \beta \frac{\mathbb{E}(U_0^n)}{\mathbb{E}(U_\gamma^n)} = \alpha + \frac{\beta}{\sum_{k=0}^n \gamma^k}, \quad \alpha, \beta \in \mathbb{R}, \quad n = 0, 1, \dots$$

where $U_\gamma \sim \text{Uniform}(\gamma, 1)$, $\gamma \in (-1, 1)$. As

$$(1.5) \quad \mathbb{E}(U_\gamma^n) = \frac{1}{n+1} \frac{1 - \gamma^{n+1}}{1 - \gamma} \xrightarrow{\gamma \rightarrow 1} 1,$$

Panjer's class corresponds to the degenerate limit case, letting $\gamma \rightarrow 1$ so that $U_\gamma \rightarrow U_1$, the degenerate random variable with unit mass at 1.

When $\alpha = 0$, the iterative expression for the p.m.f. of $N_{0,\beta,\gamma}$ verifies

$$(1.6) \quad \frac{1 - \gamma^{n+1}}{1 - \gamma} f_{N_{0,\beta,\gamma}}(n+1) = \sum_{k=0}^n f_{N_{0,\beta,\gamma}}(k) r_{n-k}$$

with $r_0 = \beta$ and $r_1 = \dots = r_n = 0$, of which (1.2) and (1.3) aren't but the cases $\gamma = 0$ and $\gamma = 1$, respectively. We shall investigate the classes \mathcal{C}_γ of randomly

stopped sums $\sum_{k=0}^{N_{0,\beta,\gamma}} Y_k$, whose members satisfy (1.6) for nonnegative r_k , with $\sum_{k=0}^{\infty} r_k < \infty$.

In section 4 we show that when $|\gamma_1| < \gamma_2 \leq 1$, $\mathcal{C}_{\gamma_1} \subset \mathcal{C}_{\gamma_2}$. Also, for $\gamma \in [0, 1]$, the classes \mathcal{C}_γ form an increasing chain of classes of infinitely divisible random variables, spanning from \mathcal{C}_0 , the class of discrete geometric stopped sums, to \mathcal{C}_1 , the class of discrete Poisson stopped sums.

Many of these results rely on properties of absolutely monotone functions scattered in the literature, that we shall discuss in section 2 below in conjunction with Panjer theory. Ospina and Gerber (1987) remarked that the representation theorem for the generating functions of discrete stopped Poisson sums (discrete infinitely divisible laws) follows from Panjer's theory, and the same is true for the representation of geometric infinitely divisible generating functions, see section 2, and for wider classes of generating functions whose bearing on general p -infinite divisibility is worth noting. This will be further discussed in the concluding section.

2. BASIC RESULTS

Let $\mathcal{G}(s) = \sum_{n=0}^{\infty} f(n)s^n$, $s \in [0, r)$, be the generating function of the sequence $\{f(n)\}_{n \in \mathbb{N}}$; in other words, $f(n) = \frac{\mathcal{G}^{(n)}(0)}{n!}$, $n \in \mathbb{N}$.

If $p_n \geq 0$, $n \in \mathbb{N}$, then $\mathcal{G}^{(n)}(s) \geq 0$, $s \in [0, r)$, and we say that \mathcal{G} is absolutely monotone (abs. mon.) in $[0, r)$. If there exists $r > 0$ such that \mathcal{G} is abs. mon. in $[0, r)$, we say that the function \mathcal{G} is abs. mon. (Bernstein, 1928).

We refer to Widder (1946, chapt. IV) and to Feller (1968, chap. XI) for basic information on absolutely monotone functions and generating functions; Skellam and Shelton (1957) or Srivastava and Manocha (1984) provide a thorough discussion. It is obvious that the sum or the product of abs. mon. functions is abs. mon.; we shall need the following results:

1. \mathcal{G} is abs. mon. $\iff \mathcal{G}(0) \geq 0$ and $\frac{d\mathcal{G}}{ds}$ is abs. mon. $\iff \frac{d}{ds} [s\mathcal{G}(s)]$ is abs. mon. (since $p_n \geq 0 \iff (1+n)p_n \geq 0$).

2. Let $\gamma \in (-1, 1)$; then, \mathcal{G} abs. mon. $\iff \mathcal{G}(s) - \gamma \mathcal{G}(\gamma s)$ abs. mon. (it is sufficient to note that $p_n \geq 0 \iff p_n(1 - \gamma^{n+1}) \geq 0$).

Let $|\gamma| \leq \eta < 1$; then, \mathcal{G} abs. mon. $\implies \eta \mathcal{G}(\eta s) - \gamma \mathcal{G}(\gamma s)$ abs. mon. ($p_n \geq 0$ implies $p_n(\eta^{n+1} - \gamma^{n+1}) \geq 0$).

Note that $\eta \mathcal{G}(\eta s) - \gamma \mathcal{G}(\gamma s)$ is no longer abs. mon. if $-1 < \eta < \gamma \leq 0$.

3. If \mathcal{G}_1 is abs. mon. in $[0, r_1)$, \mathcal{G}_2 is abs. mon. in $[0, r_2)$, and $\mathcal{G}_2(s) < r_1$ for all $s \in [0, r_2)$, the compound function $\mathcal{G}_1 \circ \mathcal{G}_2 = \mathcal{G}_1(\mathcal{G}_2)$ is abs. mon. in $[0, r_2)$. In particular:

- (a) As $\mathcal{G}_1(s) = e^s$ is the generating function of $p_n = \frac{1}{n!}$, \mathcal{G}_2 abs. mon. implies that $(\mathcal{G}_1 \circ \mathcal{G}_2)(s) = e^{\mathcal{G}_2(s)}$ is abs. mon.
- (b) As $\mathcal{G}_1(s) = \frac{1}{1-s}$ is the generating function of $p_n = 1$, \mathcal{G}_2 abs. mon. in $[0, r_2)$ with $\mathcal{G}_2(s) < 1$ for $s \in [0, r_2)$ implies that $(\mathcal{G}_1 \circ \mathcal{G}_2)(s) = \frac{1}{1-\mathcal{G}_2(s)}$ is abs. mon.

Let us now consider the randomly stopped sum $S_{N_{\alpha, \beta}} = \sum_{k=1}^{N_{\alpha, \beta}} Y_k$, where $Y_k \stackrel{d}{=} Y$, $k=1, 2, \dots$, are i.i.d. counting random variables, with p.m.f. $\{f_Y(n)\}_{n \in \mathbb{N}}$, independent of the Panjer subordinator $N_{\alpha, \beta}$.

As

$$\mathbb{E} \left[\frac{k}{n+1} Y_1 \mid \sum_{i=1}^k Y_i = n+1 \right] = 1$$

and

$$\mathbb{P} \left[Y_1 = j \mid \sum_{i=1}^k Y_i = n+1 \right] = \frac{f_Y(j) f_Y^{*(k-1)}(n+1-j)}{f_Y^{*k}(n+1)}, \quad j = 0, \dots, n+1$$

(Rólski *et al.*, 1999, p. 119), where as usual f^{*k} denotes the k -fold convolution ($f^{1*} = f$, $f^{*k} = f * f^{*(k-1)}$), it follows that the probability mass function of a Poisson stopped sums $(N_{0, \beta}, \beta > 0)$ verifies

$$\begin{aligned} (n+1) f_{S_{N_{0, \alpha}}}(n+1) &= \sum_{k=0}^n f_{S_{N_{0, \alpha}}}(k) \beta (n+1-k) f_Y(n+1-k) \\ (2.1) \quad &= \sum_{k=0}^n f_{S_{N_{0, \alpha}}}(k) r_{n-k}, \end{aligned}$$

with $r_k = \beta (k+1) f_Y(k+1) \geq 0$, $k=0, 1, \dots$, and it therefore follows that the generating function $\mathcal{H}_{N_{0, \beta}}(s) = \sum_{k=0}^{\infty} r_k s^k$ of the $\{r_k\}_{k \in \mathbb{N}}$ is absolutely monotone, with

$$\sum_{k=0}^{\infty} \frac{r_k}{k+1} = \sum_{k=0}^{\infty} \beta f_Y(k+1) = \beta (1 - f_Y(0)).$$

Assuming that $f_Y(0) = 0$ (i.e., enforcing

a unique representation by fixing this free parameter), multiplying both sides of (1.6) by s^n and summing for $n = 0, 1, \dots$, we get,

$$(2.2) \quad \mathcal{G}_{S_{N_{0,\beta}}}(s) = \exp \left[\beta \left(\frac{1}{\beta} \sum_{k=0}^{\infty} \frac{r_k}{k+1} s^{k+1} - 1 \right) \right] = e^{\beta[\mathcal{P}(s)-1]},$$

where $\mathcal{P}(s) = \frac{1}{\beta} \sum_{k=0}^{\infty} \frac{r_k}{k+1} s^{k+1}$ is a (unique) p.g.f., such that $\mathcal{P}(0) = 0$.

On the other hand, for geometric stopped sums $(N_{\alpha,0}, 0 < \alpha < 1)$ we get

$$(2.3) \quad f_{S_{N_{\alpha,0}}}(n+1) = \sum_{k=0}^n f_{S_{N_{\alpha,0}}}(k) r_{n-k}$$

where $r_k = \frac{\alpha f_Y(k+1)}{1 - \alpha f_Y(0)}$. As in the treatment of Poisson stopped sums, we may get a unique representation theorem by letting the free parameter $f_Y(0) = 0$, which implies $\sum_{k=0}^{\infty} r_k = \alpha$, multiplying both sides of (2.3) by s^n and summing for $n = 0, 1, \dots$. In terms of generating functions,

$$\frac{\mathcal{G}_{S_N}(s) - f_{N_{\alpha,0}}(0)}{s} = \mathcal{G}_{S_N}(s) \mathcal{H}_{N_{\alpha,0}}(s)$$

where $\mathcal{H}_{N_{\alpha,0}}$ is the generating function of $\{r_n\}_{n \in \mathbb{N}}$, which are all nonnegative, with $\sum_{n=0}^{\infty} r_n = \alpha \in (0, 1)$, i.e. $\mathcal{H}_{N_{\alpha,0}}$ is abs. mon. From $\mathcal{G}_{S_N}(1) = \frac{p_0}{1 - \mathcal{H}_{N_{\alpha,0}}(1)} = \frac{p_0}{1 - \alpha} = 1$, it follows that

$$\mathcal{G}_{S_N}(s) = \frac{p_0}{1 - s \mathcal{H}_{N_{\alpha,0}}(s)} = \frac{1 - \alpha}{1 - s \sum_{k=0}^{\infty} r_k s^k} = \frac{1 - \alpha}{1 - \alpha \mathcal{P}(s)}$$

where $\mathcal{P}(s) = \sum_{k=0}^{\infty} \frac{r_k}{\alpha} s^{k+1}$, such that $\mathcal{P}(0) = 0$, is a p.g.f., because it is abs. mon. and $\mathcal{P}(1) = 1$. In other words, Panjer's iteration also provides a straightforward proof of the representation theorem for geometric infinitely divisible lattice distributions.

We record these representation theorems for the sake of the corollaries that we then establish, which will be instrumental in the proof of the extensions in sections 3 and 4.

Theorem 2.1. *The p.g.f. $\mathcal{G}_{S_{N_{0,\beta}}}$ of a discrete Poisson stopped sum such that $\mathbb{P}[S_{N_{0,\beta}} = 0] = f_{S_{N_{0,\beta}}}(0) > 0$ has a unique representation $\mathcal{G}_{S_{N_{0,\beta}}}(s) = e^{\beta[\mathcal{P}(s)-1]}$, where \mathcal{P} is a p.g.f. such that $\mathcal{P}(0) = 0$, and $\beta = -\ln \mathcal{G}_{S_{N_{0,\beta}}}(0)$.*

The p.g.f. $\mathcal{G}_{S_{N_{\alpha,0}}}$ of a discrete geometric stopped sum such that $\mathbb{P}[S_N = 0] = f_{S_{N_{\alpha,0}}}(0) > 0$ has a unique representation $\mathcal{G}_{S_{N_{\alpha,0}}}(s) = \frac{1-\alpha}{1-\alpha \mathcal{P}(s)}$, where \mathcal{P} is a p.g.f. such that $\mathcal{P}(0) = 0$, and $\alpha = 1 - \mathcal{G}_{S_{N_{\alpha,0}}}(0)$.

Observe also that $\exp\left(1 - \frac{1}{\mathcal{G}_{s_{N_{\alpha,0}}}}\right) = e^{\frac{\alpha}{1-\alpha}[\mathcal{P}(s)-1]} = \mathcal{G}_{s_{N_0, \frac{\alpha}{1-\alpha}}}(s)$. On the other hand, $\frac{1}{1 - \ln\left(\mathcal{G}_{s_{N_0, \beta}}(s)\right)} = \frac{1 - \frac{\beta}{\beta+1}}{1 - \frac{\beta}{\beta+1}\mathcal{P}(s)} = \mathcal{G}_{s_{N_0, \frac{\beta}{\beta+1}}}(s)$.

Corollary 2.1.1.

- (1) Let \mathcal{G} be a probability generating function such that $\mathcal{G}(0) > 0$; then, \mathcal{G} is the p.g.f. of a discrete Poisson stopped sum iff $\frac{\mathcal{G}'(s)}{\mathcal{G}(s)}$ is abs. mon.
- (2) Let \mathcal{G} be a p.g.f. such that $\mathcal{G}(0) > 0$, and $\gamma \in (-1, 1)$. If \mathcal{G} is the p.g.f. of a discrete Poisson stopped sum, then $\frac{\mathcal{G}(s)}{\mathcal{G}(\gamma s)}$ is abs. mon., and $\mathcal{G}_\gamma(s) = \frac{\mathcal{G}(\gamma)\mathcal{G}(s)}{\mathcal{G}(\gamma s)}$ is also the p.g.f. of a Poisson stopped sum.
- (3) Let \mathcal{G} be a p.g.f. such that $\mathcal{G}(0) > 0$, and $|\gamma_1| \leq \gamma_2 < 1$. If \mathcal{G} is the p.g.f. of a discrete Poisson stopped sum, then $\frac{\mathcal{G}(\gamma_2 s)}{\mathcal{G}(\gamma_1 s)}$ is abs. mon. and $\mathcal{G}_{\gamma_1, \gamma_2}(s) = \frac{\mathcal{G}(\gamma_1)\mathcal{G}(\gamma_2 s)}{\mathcal{G}(\gamma_2)\mathcal{G}(\gamma_1 s)}$ is also the p.g.f. of a Poisson stopped sum.
- (4) Any discrete geometric stopped sum such that $\mathbb{P}[S_N=0] = \tilde{p}_0 > 0$ is a Poisson stopped sum, i.e. infinitely divisible.

Proof: (1) From Theorem 2.1 we know that \mathcal{G} , with $\mathcal{G}(0) > 0$, is the p.g.f. of a Poisson stopped sum iff $\frac{\mathcal{G}'(s)}{\mathcal{G}(s)} = \mathcal{H}_{N_0, \beta}(s) = \sum_{k=0}^{\infty} r_k s^k$, where $r_k = \beta(k+1)f_Y(k+1) \geq 0$, $k = 0, 1, \dots$, and therefore its generating function $\mathcal{H}_{N_0, \beta}(s) = \sum_{k=0}^{\infty} r_k s^k$ is absolutely monotone.

(2) From formula (2.2), we see that $\mathcal{G}(s) > 0$ for all s , therefore $\frac{\mathcal{G}(s)}{\mathcal{G}(\gamma s)} \geq 1$ if $0 \leq s \leq 1$. On the other hand, $\frac{d}{ds} \left[\ln \frac{\mathcal{G}(s)}{\mathcal{G}(\gamma s)} \right] = \frac{\mathcal{G}'(s)}{\mathcal{G}(s)} - \gamma \frac{\mathcal{G}'(\gamma s)}{\mathcal{G}(\gamma s)}$ is abs. mon., by 2.1.1.(1) and property 2 of abs. mon. functions. As $\ln \frac{\mathcal{G}(s)}{\mathcal{G}(\gamma s)}$ is nonnegative for $s=0$, it is also abs. mon., by property 1 of abs. mon. functions. From property 3(a) of abs. mon. functions, it follows that $\frac{\mathcal{G}(s)}{\mathcal{G}(\gamma s)}$ is abs. mon. Since $\mathcal{G}_\gamma(0) = \mathcal{G}(\gamma) > 0$, $\mathcal{G}_\gamma(1) = 1$, and $\frac{\mathcal{G}'_\gamma(s)}{\mathcal{G}_\gamma(s)} = \frac{d}{ds} \left[\ln \frac{\mathcal{G}(s)}{\mathcal{G}(\gamma s)} \right]$ is abs. mon., we conclude that \mathcal{G}_γ is the p.g.f. of a Poisson stopped sum.

(3) By 2.1.1.(2), $\frac{\mathcal{G}(\gamma_2 s)}{\mathcal{G}(\gamma_1 s)}$ is abs. mon., and by property 2 of abs. mon. functions $\frac{\mathcal{G}'_{\gamma_1, \gamma_2}(s)}{\mathcal{G}_{\gamma_1, \gamma_2}(s)} = \gamma_2 \frac{\mathcal{G}'(\gamma_2 s)}{\mathcal{G}(\gamma_2 s)} - \gamma_1 \frac{\mathcal{G}'(\gamma_1 s)}{\mathcal{G}(\gamma_1 s)}$ is abs. mon. Since $\mathcal{G}_{\gamma_1, \gamma_2}(0) = \frac{\mathcal{G}(\gamma_1)}{\mathcal{G}(\gamma_2)} > 0$ and $\mathcal{G}_{\gamma_1, \gamma_2}(1) = 1$, it follows that $\mathcal{G}_{\gamma_1, \gamma_2}(s) = \frac{\mathcal{G}(\gamma_1)\mathcal{G}(\gamma_2 s)}{\mathcal{G}(\gamma_2)\mathcal{G}(\gamma_1 s)}$ is the p.g.f. of a Poisson stopped sum.

(4) As we have seen, \mathcal{G} with $\mathcal{G}(0) > 0$ is the p.g.f. of a discrete geometric stopped sum iff $\mathcal{G}(s) = \frac{\mathcal{G}(0)}{1 - s\mathcal{H}_{N_{\alpha,0}}(s)}$, where $\mathcal{H}_{N_{\alpha,0}}(s) < 1$ for $s \in [0, 1)$ is abs. mon.

As

$$\frac{\mathcal{G}'(s)}{\mathcal{G}(s)} = \frac{\frac{\mathcal{G}(0) \frac{d}{ds} [s\mathcal{H}_{N_{\alpha,0}}(s)]}{(1 - s\mathcal{H}_{N_{\alpha,0}}(s))^2}}{\frac{\mathcal{G}(0)}{1 - s\mathcal{H}_{N_{\alpha,0}}(s)}} = \frac{\frac{d}{ds} [s\mathcal{H}_{N_{\alpha,0}}(s)]}{1 - s\mathcal{H}_{N_{\alpha,0}}(s)},$$

and from property 1 of abs. mon. functions $\frac{d}{ds} [s\mathcal{H}_{N_{\alpha,0}}(s)]$ is abs. mon., from property 3(b) we know that $\frac{1}{1 - s\mathcal{H}_{N_{\alpha,0}}(s)}$ is abs. mon., and the product of abs. mon. functions is abs. mon., it follows that $\frac{\mathcal{G}'(s)}{\mathcal{G}(s)}$ is abs. mon. From part (1) of Corollary 2.1.1., it follows that \mathcal{G} is the p.g.f. of a Poisson stopped sum. \square

If the probability generating function \mathcal{G}_Y of $Y \frown F_Y$ depends on the parameter θ so that $\mathcal{G}_Y(s|k\theta) = [\mathcal{G}_Y(s|\theta)]^k$, then $F_X \vee F_Y = F_Y \wedge_K F_X$, where \vee denotes the stopped sum of Y independent copies of X , and \wedge_K denotes the mixture of $Y|K$, with mixing distribution F_X (Gurland, 1957). Therefore the class of discrete Poisson stopped sums coincides with the class of discrete mixtures of Poisson random variables. In what mixtures of geometric random variables and geometric stopped sums, the former is strictly included in the later.

3. EXTENSIONS

We now investigate the nondegenerate discrete random variables $N_{\alpha, \beta, \gamma}$ whose probability mass function $\{p_n\}_{n \in \mathbb{N}}$ satisfies

$$(3.1) \quad \frac{p_{n+1}}{p_n} = \alpha + \beta \frac{\mathbb{E}(U_0^n)}{\mathbb{E}(U_\gamma^n)} = \alpha + \beta \frac{1 - \gamma}{1 - \gamma^{n+1}} \quad \text{for } n = 0, 1, \dots, \quad \alpha, \beta \in \mathbb{R},$$

where $U_\gamma \frown \text{Uniform}(\gamma, 1)$, $\gamma \in (-1, 1)$, with $p_0 > 0$. If $\gamma = 0$, all possible solutions are geometric random variables, and when $\gamma \rightarrow 1$ we get Panjer's class of counting distributions.

For a nondegenerate solution of (3.1) with infinite support to exist, we must have

$$\alpha + \beta \frac{\mathbb{E}(U_0^n)}{\mathbb{E}(U_\gamma^n)} = \alpha + \beta \frac{1 - \gamma}{1 - \gamma^{n+1}} > 0$$

for every integer n . According to the signs of β and γ , the infimum of this factor is either $\alpha + \beta$ (for $n = 0$), $\alpha + \frac{\beta}{1 + \gamma}$ (for $n = 1$), or $\alpha + \beta(1 - \gamma)$ (when $n \rightarrow \infty$), so we must have $\alpha + \beta > 0$, $\alpha + \frac{\beta}{1 + \gamma} > 0$, and $\alpha + \beta(1 - \gamma) \geq 0$. Then, applying

the ratio test to the sum

$$\sum_{k \geq 0} p_k = p_0 \sum_{k=0}^{\infty} \prod_{n=0}^{k-1} \left(\alpha + \beta \frac{1-\gamma}{1-\gamma^{n+1}} \right)$$

we see that it converges iff $0 \leq \alpha + \beta(1-\gamma) < 1$. Thus a necessary and sufficient condition for a solution of (3.1) with infinite support (random variable with finite support cannot be infinitely divisible) to exist is that

$$\min \left\{ \alpha + \beta, \alpha + \frac{\beta}{1+\gamma} \right\} > 0 \quad \text{and} \quad 0 \leq \alpha + \beta(1-\gamma) < 1 .$$

Rewriting (3.1) as

$$(3.2) \quad (1-\gamma^{n+1}) f_{N_{\alpha, \beta, \gamma}}(n+1) = [\alpha + \beta(1-\gamma)] f_{N_{\alpha, \beta, \gamma}}(n) - \alpha \gamma^{n+1} f_{N_{\alpha, \beta, \gamma}}(n) ,$$

for $\gamma \in (-1, 1)$, $n = 0, 1, \dots$, multiplying both sides by s^{n+1} and summing we get

$$(3.3) \quad [1 - (\alpha + \beta(1-\gamma)) s] \mathcal{G}_{\alpha, \beta, \gamma}(s) = (1 - \alpha \gamma s) \mathcal{G}_{\alpha, \beta, \gamma}(\gamma s) ,$$

where $\mathcal{G}_{\alpha, \beta, \gamma}(s) = \sum_{n=0}^{\infty} f_{N_{\alpha, \beta, \gamma}}(n) s^n$ denotes the probability generating function of the probability mass function $\{f_{N_{\alpha, \beta, \gamma}}(n)\}_{n=0}^{\infty}$, and from that

$$(3.4) \quad \mathcal{G}_{\alpha, \beta, \gamma}(s) = \mathcal{G}_{\alpha, \beta, \gamma}(\gamma^{n+1} s) \prod_{k=0}^n \frac{1 - \alpha \gamma^{k+1} s}{1 - [\alpha + \beta(1-\gamma)] \gamma^k s} .$$

Observing that

$$(3.5) \quad \frac{\mathcal{G}_{\alpha, \beta, \gamma}(s)}{\mathcal{G}_{\alpha, \beta, \gamma}(1)} = \frac{\mathcal{G}_{\alpha, \beta, \gamma}(\gamma^{n+1} s)}{\mathcal{G}_{\alpha, \beta, \gamma}(\gamma^{n+1})} \prod_{k=0}^n \frac{\frac{1 - \alpha \gamma^{k+1} s}{1 - [\alpha + \beta(1-\gamma)] \gamma^k s}}{\frac{1 - \alpha \gamma^{k+1}}{1 - [\alpha + \beta(1-\gamma)] \gamma^k}}$$

and letting $n \rightarrow \infty$,

$$(3.6) \quad \mathcal{G}_{\alpha, \beta, \gamma}(s) = \prod_{k=0}^{\infty} \frac{1 - \alpha \gamma^{k+1} s}{1 - \alpha \gamma^{k+1}} \frac{1 - [\alpha + \beta(1-\gamma)] \gamma^k}{1 - [\alpha + \beta(1-\gamma)] \gamma^k s} .$$

If $\gamma \in [0, 1)$, $\alpha < 0$ and $\beta \in (-\frac{\alpha}{1-\gamma}, \frac{1-\alpha}{1-\gamma})$, we recognize in

$$\mathcal{G}_{\alpha, \beta, \gamma}(s) = \prod_{k=0}^{\infty} \frac{1 - \alpha \gamma^{k+1} s}{1 - \alpha \gamma^{k+1}} \frac{1 - [\alpha + \beta(1-\gamma)] \gamma^k}{1 - [\alpha + \beta(1-\gamma)] \gamma^k s} ,$$

the probability generating function of an infinite sum of independent random variables, the k -th summand being the result of randomly adding 1, with probability $\frac{\alpha \gamma^{k+1}}{\alpha \gamma^{k+1} - 1}$, to an independent *Geometric* $(1 - [\alpha + \beta(1-\gamma)] \gamma^k)$ random variable.

The limiting case $\gamma = 1$ may be approached as follows: rewriting (3.3) as

$$\frac{\mathcal{G}_{\alpha,\beta,\gamma}(s) - \mathcal{G}_{\alpha,\beta,\gamma}(\gamma s)}{\alpha s [\mathcal{G}_{\alpha,\beta,\gamma}(s) - \mathcal{G}_{\alpha,\beta,\gamma}(\gamma s)] + (1-\gamma)s [\beta \mathcal{G}_{\alpha,\beta,\gamma}(s) + \alpha \mathcal{G}_{\alpha,\beta,\gamma}(\gamma s)]} = 1 ,$$

dividing the numerator and the denominator by $(1-\gamma)s$ and letting $\gamma \rightarrow 1$, we get

$$\frac{\mathcal{G}'_{\alpha,\beta,1}(s)}{\alpha s \mathcal{G}'_{\alpha,\beta,1}(s) + \beta \mathcal{G}_{\alpha,\beta,1}(s) + \alpha \mathcal{G}_{\alpha,\beta,1}(s)} = 1 \iff \frac{\mathcal{G}'_{\alpha,\beta,1}(s)}{\mathcal{G}_{\alpha,\beta,1}(s)} = \frac{\alpha + \beta}{1 - \alpha s} ,$$

the expression we obtain working out the probability generating function in Panjer's iterative expression $p_{\alpha,\beta}(n+1) = (\alpha + \frac{\beta}{n+1})p_{\alpha,\beta}(n)$, $\alpha, \beta \in \mathbb{R}$, $n = 0, 1, \dots$.

We now focus on the case $\alpha = 0$, for which $\beta \in (0, \frac{1}{1-\gamma})$, and

$$(3.7) \quad \mathcal{G}_{0,\beta,\gamma}(s) = \prod_{k=0}^{\infty} \frac{1 - \beta(1-\gamma)\gamma^k}{1 - \beta(1-\gamma)\gamma^k s} = \prod_{k=0}^{\infty} \frac{1 - w_k}{1 - w_k s} ,$$

where $w_k = \beta(1-\gamma)\gamma^k$. If $\gamma \in [0, 1)$, we get that have $N_{0,\beta,\gamma} = \sum_{k=0}^{\infty} W_k$, with $W_k \sim \text{Geometric}(1 - \beta(1-\gamma)\gamma^k)$ independent summands. If $\gamma = 0$, the above expression simplifies to $\mathcal{G}_{0,\beta,0}(s) = \frac{1-\beta}{1-\beta s}$. Therefore we conclude that $N_{0,\beta,0} = N_{\beta,0} \sim \text{Geometric}(1 - \beta)$, $\beta \in (0, 1)$.

Let us point out that the probability mass function of a random variable $N_{0,\beta,\gamma}$, $\gamma \in (-1, 1)$, trivially satisfies

$$\frac{1 - \gamma^{n+1}}{1 - \gamma} p_{n+1} = \sum_{k=0}^n p_k r_{n-k} ,$$

with $r_0 = \beta$ and $r_1 = r_2 = \dots = r_n = 0$, provided that

$$0 < \beta = \sum_{n=0}^{\infty} r_n s^n = \mathcal{H}(s) < \frac{1}{1-\gamma} ,$$

a point which will be of relevance in the following section.

4. DISCRETE INFINITELY DIVISIBLE DISTRIBUTIONS AND \mathcal{C}_γ CLASSES

In what follows we investigate the classes \mathcal{C}_γ , $\gamma \in (-1, 1)$, of nondegenerate counting random variables (distributions, p.g.f.) whose probability mass function satisfies $\tilde{p}_0 > 0$ and the general recursive relation

$$(4.1) \quad \frac{1 - \gamma^{n+1}}{1 - \gamma} \tilde{p}_{n+1} = \sum_{k=0}^n \tilde{p}_k r_{n-k} , \quad n = 0, 1, \dots ,$$

with $r_k \geq 0$, which extends Panjer's recursive expression for the probability mass function of the classes of Poisson stopped sums (\mathcal{C}_1) and of geometric stopped sums (\mathcal{C}_0). It is well known that any geometric infinitely divisible lattice distribution is infinitely divisible in the classical sense, a result that follows from the fact that $\frac{1-p}{1-ps} = \exp\{\ln(1-p)[\mathcal{P}(s) - 1]\}$, where $\mathcal{P}(s) = -\frac{1}{\ln(1-p)} \sum_{k=0}^{\infty} \frac{(ps)^k}{k}$ is the p.g.f. of a logarithmic random variable.

As before, multiplying both members of (4.1) by s^{n+1} and summing for $n \geq 0$, we obtain

$$(4.2) \quad \frac{\mathcal{G}(s) - \mathcal{G}(\gamma s)}{1 - \gamma} = s \mathcal{G}(s) \mathcal{H}_\gamma(s),$$

where $\mathcal{G}(s) = \sum_{n=0}^{\infty} \tilde{p}_n s^n$ and $\mathcal{H}_\gamma(s) = \sum_{n=0}^{\infty} r_n s^n$ converges at least for $|s| \leq 1$. Thus \mathcal{H}_γ is by definition abs.mon. Since we have excluded degenerate solutions to (4.1), we must have $\mathcal{H}_\gamma(0) = r_0 = \frac{\tilde{p}_1}{\tilde{p}_0} > 0$.

If $\gamma \in [0, 1)$, we have

$$\begin{aligned} 1 &\geq \sum_{n=0}^{\infty} \tilde{p}_{n+1} = \sum_{n=0}^{\infty} \frac{1-\gamma}{1-\gamma^{n+1}} \sum_{k=0}^n \tilde{p}_k r_{n-k} \\ &= \sum_{k=0}^{\infty} \tilde{p}_k \sum_{n=0}^{\infty} \frac{(1-\gamma) r_n}{1-\gamma^{n+k+1}} \\ &> \sum_{k=0}^{\infty} \tilde{p}_k \sum_{n=0}^{\infty} (1-\gamma) r_n = (1-\gamma) \sum_{n=0}^{\infty} r_n, \end{aligned}$$

and therefore $|\mathcal{H}_\gamma(s)| \leq \mathcal{H}_\gamma(1) = \sum_{n=0}^{\infty} r_n < \frac{1}{1-\gamma}$ for $|s| \leq 1$.

If $\gamma \in (-1, 0)$, then $\frac{1-\gamma^{i+1}}{1-\gamma} \leq 1$ for $i = 0, 1, \dots$, and by a similar reasoning we conclude that in this case $|\mathcal{H}_\gamma(s)| < 1$ for $|s| \leq 1$.

As was seen in the previous section, the p.m.f. of $N_{0,\beta,\gamma}$ verifies recursion (4.1) with $r_0 = \beta$ and $r_1 = r_2 = \dots = 0$, with $0 < \beta < \frac{1}{1-\gamma}$.

We have the following result:

Theorem 4.1. *Let W be a random variable with p.g.f. \mathcal{G} , and $\gamma \in (-1, 1)$.*

$$W \in \mathcal{C}_\gamma \quad \text{iff} \quad \mathcal{G}(s) = \prod_{k=0}^{\infty} \frac{1 - (1-\gamma)\gamma^k \mathcal{H}_\gamma(\gamma^k)}{1 - (1-\gamma)\gamma^k s \mathcal{H}_\gamma(\gamma^k s)},$$

where \mathcal{H}_γ is a unique abs.mon. function such that $\mathcal{H}_\gamma(0) > 0$ and $\mathcal{H}_\gamma(1) < \max\{1, \frac{1}{1-\gamma}\}$.

Thus, if $\gamma \in [0, 1)$ the elements of \mathcal{C}_γ are infinite sums $W = \sum_{k=0}^{\infty} X_k$ of independent geometric stopped sums $X_k = \sum_{i=1}^{N_k} Y_{ki}$, whose subordinators are $N_k \sim \text{Geometric}(1 - (1-\gamma)\gamma^k \mathcal{H}_\gamma(\gamma^k))$ random variables, and whose i.i.d. summands $Y_{ki} \stackrel{d}{=} Y_k$ have the p.g.f. $\mathcal{P}_k(s) = \frac{s \mathcal{H}_\gamma(\gamma^k s)}{\mathcal{H}_\gamma(\gamma^k)}$.

Proof: We have established that

$$\frac{\mathcal{G}(s) - \mathcal{G}(\gamma s)}{1 - \gamma} = s \mathcal{G}(s) \mathcal{H}_\gamma(s) \iff \frac{\mathcal{G}(s)}{\mathcal{G}(\gamma s)} = \frac{1}{1 - (1-\gamma)s \mathcal{H}_\gamma(s)}.$$

Iterating the above expression, similarly to what we have done to obtain (3.6), we finally get

$$(4.3) \quad \mathcal{G}(s) = \prod_{k=0}^{\infty} \frac{1 - (1-\gamma)\gamma^k \mathcal{H}_\gamma(\gamma^k)}{1 - (1-\gamma)\gamma^k s \mathcal{H}_\gamma(\gamma^k s)}.$$

If $\gamma \in [0, 1)$, we further have

$$\mathcal{G}(s) = \prod_{k=0}^{\infty} \frac{1 - w_k}{1 - w_k \frac{s \mathcal{H}_\gamma(\gamma^k s)}{\mathcal{H}_\gamma(\gamma^k)}} = \prod_{k=0}^{\infty} \frac{1 - w_k}{1 - w_k \mathcal{P}_k(s)}$$

where $w_k = (1-\gamma)\gamma^k \mathcal{H}_\gamma(\gamma^k)$, and the $\mathcal{P}_k(s) = \frac{s \mathcal{H}_\gamma(\gamma^k s)}{\mathcal{H}_\gamma(\gamma^k)}$ are (unique) probability generating functions such that $\mathcal{P}_k(0) = 0$. \square

Theorem 4.2. Let W be a counting random variable with p.g.f. \mathcal{G} , and $\gamma \in (-1, 1)$. $W \in \mathcal{C}_\gamma$ iff $\mathcal{H}_\gamma(s) = \frac{\mathcal{G}(s) - \mathcal{G}(\gamma s)}{(1-\gamma)s \mathcal{G}(s)}$ is abs. mon.

We can use this result to show that the geometric distribution verifies (4.1) for nonnegative γ . In fact, if $X_\theta \sim \text{Geometric}(1 - \theta)$, with $0 < \theta < 1$, we have $r_k = \gamma^k \theta^{k+1} \geq 0$ and $\mathcal{H}_\gamma(1) = \frac{\theta}{1-\gamma\theta} < \frac{1}{1-\gamma}$. Given the uniqueness of the coefficients of \mathcal{H}_γ , we may also conclude that the geometric distribution does not belong to \mathcal{C}_γ when $\gamma \in (-1, 0)$.

The truncated geometric distribution with support on the even integers, Y_θ , given by the p.m.f.

$$p_n = \begin{cases} (1 - \theta^2) \theta^n & \text{if } n = 2k \text{ even} \\ 0 & \text{if } n = 2k + 1 \text{ odd} \end{cases}, \quad 0 < \theta < 1,$$

is an element of \mathcal{C}_γ for all $\gamma \in (-1, 1]$, since it verifies (4.1) with $r_{2k} = 0$, $r_{2k+1} = (1+\gamma)\gamma^{2k}\theta^{2k+2}$, and $\mathcal{H}_\gamma(1) = \frac{(1+\gamma)\theta^2}{1-(\gamma\theta)^2} < \frac{1}{1-\gamma}$.

It's interesting to note that the p.g.f. of Y_θ is $\mathcal{G}_{Y_\theta}(s) = \frac{1-\theta^2}{1-\theta^2 s^2} = \mathcal{G}_{X_{\theta^2}}(s^2)$. It is not difficult to show that if $X \in \mathcal{C}_0$ has the p.g.f. \mathcal{G} , then $\mathcal{G}(s^2)$ is the p.g.f. of an element of \mathcal{C}_γ , for every $\gamma \in (-1, 1]$.

Corollary 4.2.1. Let W be a counting random variable with p.g.f. \mathcal{G} , and $\gamma \in (-1, 1)$. If $W \in \mathcal{C}_\gamma$, $\frac{\mathcal{G}(s)}{\mathcal{G}(\gamma s)}$ is absolutely monotone.

Proof: From the proof of Theorem 4.1, $\frac{\mathcal{G}(s)}{\mathcal{G}(\gamma s)} = \frac{1}{1-(1-\gamma)s\mathcal{H}_\gamma(s)}$. If $\gamma \in [0, 1)$, we have $s\mathcal{H}_\gamma(s) \leq \mathcal{H}_\gamma(1) < \frac{1}{1-\gamma}$ for $0 \leq s \leq 1$; on the other hand, if $\gamma \in (-1, 0)$ we have $(1-\gamma)s\mathcal{H}_\gamma(s) \leq (1-\gamma)\mathcal{H}_\gamma(1) < 1$ for $0 \leq s \leq \frac{1}{1-\gamma}$. Thus, it follows from property 3(b) of abs.mon. functions that $\frac{\mathcal{G}(s)}{\mathcal{G}(\gamma s)}$ is abs.mon. (in $[0, 1]$ for nonnegative γ , and in $[0, \frac{1}{1-\gamma}]$ for negative γ). \square

Corollary 4.2.2. For $\gamma \in (-1, 1)$, $\mathcal{C}_\gamma \subset \mathcal{C}_1$.

Proof: Taking derivatives on both sides of $1 - \frac{\mathcal{G}(\gamma s)}{\mathcal{G}(s)} = (1-\gamma)s\mathcal{H}_\gamma(s)$, we obtain

$$\frac{\mathcal{G}'(s)\mathcal{G}(\gamma s) - \gamma\mathcal{G}'(\gamma s)\mathcal{G}(s)}{\mathcal{G}^2(s)} = (1-\gamma) \frac{d}{ds} [s\mathcal{H}_\gamma(s)] ,$$

equivalent to

$$(4.4) \quad \frac{\mathcal{G}'(s)}{\mathcal{G}(s)} - \gamma \frac{\mathcal{G}'(\gamma s)}{\mathcal{G}(\gamma s)} = (1-\gamma) \frac{\mathcal{G}(s)}{\mathcal{G}(\gamma s)} \frac{d}{ds} [s\mathcal{H}_\gamma(s)] .$$

Therefore, in view of Corollary 4.2.1 and of property 1 of abs.mon. functions $\frac{\mathcal{G}'(s)}{\mathcal{G}(s)} - \gamma \frac{\mathcal{G}'(\gamma s)}{\mathcal{G}(\gamma s)}$ is abs.mon. which in turn (property 2 of abs.mon. functions) implies that $\frac{\mathcal{G}'(s)}{\mathcal{G}(s)}$ is abs.mon.

The result follows from Corollary 2.1.1. \square

The inclusion is strict: the $Poisson(\mu)$ distribution belongs to \mathcal{C}_1 for all $\mu > 0$, but does not belong to \mathcal{C}_γ when $\gamma \in (-1, 1)$, since from Theorem 4.2 we have $r_k = (-1)^k (1-\gamma)^k \frac{\mu^{k+1}}{(k+1)!}$, so that \mathcal{H}_γ is not abs.mon.

Corollary 4.2.3. For $|\gamma_1| \leq \gamma_2 < 1$, $\mathcal{C}_{\gamma_1} \subset \mathcal{C}_{\gamma_2}$.

Proof: Let \mathcal{G} be the p.g.f. of a random variable $W \in \mathcal{C}_{\gamma_1} \subset \mathcal{C}_1$.

$$\frac{\mathcal{H}_{\gamma_2}(s) - \gamma_1 \mathcal{H}_{\gamma_2}(\gamma_1 s)}{\mathcal{H}_{\gamma_1}(s) - \gamma_2 \mathcal{H}_{\gamma_1}(\gamma_2 s)} = \frac{1 - \gamma_1 \frac{\mathcal{G}(\gamma_1 \gamma_2 s)}{\mathcal{G}(\gamma_1 s)} - \frac{\mathcal{G}(\gamma_2 s)}{\mathcal{G}(s)}}{1 - \gamma_2 \frac{\mathcal{G}(\gamma_1 \gamma_2 s)}{\mathcal{G}(\gamma_2 s)} - \frac{\mathcal{G}(\gamma_1 s)}{\mathcal{G}(s)}} = \frac{1 - \gamma_1}{1 - \gamma_2} \frac{\mathcal{G}(\gamma_2 s)}{\mathcal{G}(\gamma_1 s)} .$$

From Corollary 2.1.1.(3), $\frac{\mathcal{G}(\gamma_2 s)}{\mathcal{G}(\gamma_1 s)}$ is abs. mon., and from property 2 of abs. mon. functions $\mathcal{H}_{\gamma_1}(s) - \gamma_2 \mathcal{H}_{\gamma_1}(\gamma_2 s)$ is abs. mon. Then $\mathcal{H}_{\gamma_2}(s) - \gamma_1 \mathcal{H}_{\gamma_2}(\gamma_1 s)$, and therefore \mathcal{H}_{γ_2} , are also abs. mon., which proves that $W \in \mathcal{C}_{\gamma_2}$. \square

We can see that the inclusion is strict directly from (4.1). Suppose that $-1 < \gamma < \eta < 1$ and $0 < \beta < \frac{1}{1-\eta}$. We know that $N_{0,\beta,\eta} \in \mathcal{C}_\eta$, since its p.m.f. satisfies $\frac{1-\eta}{1-\eta} p_{n+1} = \beta p_n$. Assume that $N_{0,\beta,\eta} \in \mathcal{C}_\gamma$, that is, $\frac{1-\gamma}{1-\gamma} p_{n+1} = \sum_{k=0}^n p_k r_{n-k}$. Then $p_1 = p_0 r_0 = \beta p_0$ implies $r_0 = \beta$, and

$$(4.5) \quad (1 + \gamma) p_2 = \frac{\beta}{1 + \eta} p_1 (1 + \eta + \gamma - \eta) = p_1 r_0 + p_0 r_1$$

implies $r_1 = -\frac{\eta-\gamma}{1+\eta} \beta^2$. But this is negative, therefore $N_{0,\beta,\eta} \notin \mathcal{C}_\gamma$.

Corollary 4.2.4. Let W be a counting random variable with p.g.f. \mathcal{G} , W_γ the random variable with p.g.f. $\mathcal{G}_\gamma(s) = \frac{\mathcal{G}(\gamma) \mathcal{G}(s)}{\mathcal{G}(\gamma s)}$, and $\gamma \in (-1, 1)$. $W \in \mathcal{C}_\gamma$ iff $W_\gamma \in \mathcal{C}_0$.

Proof: As $W \in \mathcal{C}_\gamma \implies W \in \mathcal{C}_1$, from Corollary 2.1.1 we know that $\mathcal{G}_\gamma(s) = \frac{\mathcal{G}(\gamma) \mathcal{G}(s)}{\mathcal{G}(\gamma s)}$ is a p.g.f.

From the proof of Theorem 2.2 (or simply by taking $\gamma = 0$ in Theorem 4.1), in what concerns this p.g.f. \mathcal{G}_γ we obtain, with self-explaining notations,

$$(4.6) \quad \mathcal{H}_0^{(g_\gamma)}(s) = \frac{\mathcal{G}_\gamma(s) - \mathcal{G}_\gamma(0)}{s \mathcal{G}_\gamma(s)} = \frac{\mathcal{G}(s) - \mathcal{G}(\gamma s)}{s \mathcal{G}(s)} = (1 - \gamma) \mathcal{H}_\gamma^{(g)}(s)$$

and therefore $\mathcal{H}_0^{(g_\gamma)}$ is abs. mon. iff $\mathcal{H}_\gamma^{(g)}$ is abs. mon. \square

5. FURTHER COMMENTS

1. Geometric infinite divisibility arose from Kovalenko's (1965) extensions of Rényi's (1956) work on random rarefaction, with the general characterization of geometric stable laws given in Kozubowski (1994). This led to a general definition of \mathcal{N} -summation schemes, the classical summation scheme being the special case $N_p = \frac{1}{p}$ (degenerate random variables, and therefore a non-random sum of random variables). It is well known that for some families $\mathcal{N} = \{N_p, p \in (0, 1), \mathbb{E}(N_p) = \frac{1}{p}\}$ there exists \mathcal{N} -Gaussian laws (for instance for $N_p \prec \text{Geometric}(p)$, the corresponding \mathcal{N} -Gaussian random variables being the Laplace random variables), while other N_p , for instance $N_p \prec \text{Poisson}(\frac{1}{p})$,

do not admit \mathcal{N} -Gaussian laws. Although it is easy to prove that in more general branching settings \mathcal{N} -Gaussian laws do exist, only the usual Gaussian law and the Laplace geometric–Gaussian law are explicitly exhibited in the references we know.

This research arose from the observation that $\mathcal{C}_0 \subset \mathcal{C}_1$ and that, more generally, $0 < \gamma_1 < \gamma_2 < 1 \implies \mathcal{C}_0 \subset \mathcal{C}_{\gamma_1} \subset \mathcal{C}_{\gamma_2} \subset \mathcal{C}_1$.

Our aim was either to prove that there exist $\gamma \in (0, 1)$ such that for $\gamma_1 \leq \gamma$ we could exhibit a \mathcal{N} -Gaussian law in \mathcal{C}_{γ_1} — which we couldn't — or else to extend \mathcal{C}_γ classes for $\gamma < 0$ — which we did — and show that for those it was possible to construct \mathcal{N} -Gaussian random variables. Unfortunately for $-1 < \gamma_1 < \gamma_2 < 0$ the chain of inclusions $\mathcal{C}_{\gamma_1} \subset \mathcal{C}_{\gamma_2} \subset \mathcal{C}_0$ is no longer valid.

2. The extension of Katz–Panjer's iterative relation

$$(5.1) \quad \frac{f(n+1)}{f(n)} = \alpha + \frac{\beta}{n+1} = \alpha + \beta \mathbb{E}(U_0^n), \quad n = 0, 1, \dots, \quad \alpha, \beta \in \mathbb{R},$$

by

$$(5.2) \quad \frac{f(n+1)}{f(n)} = \alpha + \beta \frac{\mathbb{E}(U_0^n)}{\mathbb{E}(U_\gamma^n)}, \quad n = 0, 1, \dots, \quad \alpha, \beta \in \mathbb{R},$$

where $U_\gamma \sim \text{Uniform}(\gamma, 1)$, $\gamma \in (-1, 1]$ may seem arbitrary at this stage, unless it is considered as a first step in extending (5.1) by using more general *Beta*, of which the *Uniform* in (5.2) isn't but a special case, or even more general random variables. Naturally $\{f(n)\}_{n \in \mathbb{N}}$ is not a p.m.f. unless the restrictions in the parameters are very strong.

3. Panjer's class $\Pi = \Pi^{(0)}$ has been generalized by Sundt and Jewell (1981), who considered the class $\Pi^{(1)}$ of discrete random variables whose probability mass function satisfies

$$(5.3) \quad f_{\alpha, \beta}(n+1) = \left(\alpha + \frac{\beta}{n+1} \right) f_{\alpha, \beta}(n), \quad \alpha, \beta \in \mathbb{R}, \quad n = 1, 2, \dots$$

Willmot (1987) published the definitive characterization of $\Pi^{(1)}$: the probability mass function of a discrete random variable N , with support $\mathcal{S} = \{1, 2, \dots\}$, satisfies the above expression if N is either a zero-truncated Binomial, Poisson or Negative Binomial random variable, or a Logarithmic (when $\alpha \in (0, 1)$ and the index $\frac{\alpha}{\alpha+\beta} \rightarrow 0$) or an Engen (1974) Extended Negative Binomial random variable (index $\frac{\alpha}{\alpha+\beta} \in (-1, 0)$, where $\alpha \in (0, 1]$), and general solutions N^* , with support $\mathcal{S} = \{0, 1, 2, \dots\}$, arise from a *hurdle process* (Cameron and Trivedi, 1998, pp.123–125) $N^* = \begin{cases} 0 & N \\ p_0 & 1 - p_0 \end{cases}$, where N is one of the above variables. Klugman *et al.* (1998) describe the solutions as *zero modified N variables*.

Hess, Lewald and Schmidt (2002) considered the even more general setting $\Pi^{(k)}$, $k = 0, 1, \dots$, in which the probability mass functions satisfy

$$(5.4) \quad f_{\alpha, \beta}(n+1) = \left(\alpha + \frac{\beta}{n+1} \right) f_{\alpha, \beta}(n), \quad \alpha, \beta \in \mathbb{R}, \quad n = k, k+1, \dots,$$

giving a complete description of $\Pi^{(k)}$, $k = 0, 1, \dots$ in terms of $\{0, 1, \dots, k-1\}$ modified basic claim number distributions, i.e., the left k -truncated binomial, Poisson, and negative binomial distributions, the other basic claim number distributions are the left truncated *Logarithmic*(k, θ) distribution, and the left truncated *Engen*(k, β, θ) distribution. The extension

$$(5.5) \quad \frac{f_{\alpha, \beta}(n+1)}{f_{\alpha, \beta}(n)} = \alpha + \beta \frac{1 - \gamma}{1 - \gamma^{n+1}}, \quad \alpha, \beta \in \mathbb{R}, \quad n = k, k+1, \dots,$$

of (3.1) may be investigated along similar lines, but with very cumbersome results.

REFERENCES

- [1] BERNSTEIN, S. (1928). Sur les fonctions absolument monotones, *Acta Mathematica*, **51**, 1–66.
- [2] CAMERON, A.C. and TRIVEDI, P.K. (1998). *Regression Analysis of Count Data*, Cambridge University Press, Cambridge.
- [3] DHAENE, J. and SUNDT, B. (1998). On approximating distributions by approximating their De Pril transforms, *Scand. Actuar. J.*, 1–23.
- [4] ENGEN, S. (1974). On species frequency models, *Biometrika*, **61**, 263–270.
- [5] FELLER, W. (1968). *An Introduction to Probability Theory and its Applications*, vol. I, Wiley, New York.
- [6] GURLAND, J. (1957). Some interrelations among compound and generalized distributions, *Biometrika*, **44**, 265–268.
- [7] HESS, K.TH.; LEWALD, A. and SCHMIDT, K.D. (2002). An extension of Panjer's recursion, *ASTIN Bulletin*, **32**, 283–297.
- [8] JOHNSON, N.L.; KOTZ, S. and KEMP, A.W. (1992). *Univariate Discrete Distributions*, Wiley, New York.
- [9] KATZ, L. (1965). *Unified treatment of a broad class of discrete probability distributions*. In "Classical and Contagious Discrete Distributions", Pergamon Press, Oxford, 175–182.
- [10] KLUGMAN, S.A.; PANJER, H.H. and WILLMOT, G.E. (1998). *Loss Models: From Data to Decisions*, Wiley, New York.
- [11] KOVALENKO, I.N. (1965). On a class of limit distributions for rarefied flows of homogeneous events, *Lit. Mat. Sbornik*, **5**, 569–573. (*Selected Transl. Math. Statist. and Prob.* **9**, Providence, Rhode Island, 1971, 75–81.)

- [12] KOZUBOWSKI, T.J. (1994). Representation and properties of geometric stable laws, *Approximation, Probability, and Related Fields*, Plenum, New York, 321–337.
- [13] OSPINA, A.V. and GERBER, H.U. (1987). A simple proof of Feller's characterization of the compound Poisson distribution, *Insurance: Mathematics and Economics*, **6**, 63–64.
- [14] PANJER, H.H. (1981). Recursive evaluation of a family of compound distributions, *ASTIN Bulletin*, **12**, 22–26.
- [15] RÉNYI, A. (1956). A characterization of the Poisson process, *MTA Mat. Kut. Int. Közl.*, **1**, 519–527. (English translation: *Selected Papers of Alfred Rényi*, **1**, 1948–1956, P. Turán, ed., 622–279, Akadémiai Kiadó, Budapest, with a note by D. Szász on ulterior developments up to 1976).
- [16] RÓLSKI, T.; SCHMIDLI, H.; SCHMIDT, V. and TEUGELS, J. (1999). *Stochastic Processes for Insurance and Finance*, Wiley, New York.
- [17] SKELLAM, J.G. and SHENTON, L.R. (1957). Distributions associated with random walks and recurrent events (with discussion), *J. Roy. Statist. Soc.*, **B 19**, 64–118.
- [18] SRIVASTAVA, H.M. and MANOCHA, H.L. (1984). *A Treatise on Generating Functions*, Horwood, Chichester.
- [19] SUNDT, B. and JEWELL, W.S. (1981). Further results on recursive evaluation of compound distributions, *ASTIN Bulletin*, **12**, 27–39.
- [20] WIDDER, D.V. (1946). *The Laplace Transform*, Princeton Univ. Press, Princeton.
- [21] WILLMOT, G.E. (1987). Sundt and Jewell's family of discrete distributions, *ASTIN Bulletin*, **18**, 17–29.

2.2 Comentários e Complementos

2.2.1 Relação com as somas aleatórias de subordinadora geométrica ou de Poisson

A generalização da relação iterativa de Katz-Panjer

$$\frac{f(n+1)}{f(n)} = \alpha + \frac{\beta}{n+1} = \alpha + \beta \mathbb{E}(U_0^n), \quad n = 0, 1, \dots; \quad \alpha, \beta \in \mathbb{R} \quad (2.1)$$

através de

$$\frac{f(n+1)}{f(n)} = \alpha + \beta \frac{\mathbb{E}(U_0^n)}{\mathbb{E}(U_\gamma^n)}, \quad n = 0, 1, \dots; \quad \alpha, \beta \in \mathbb{R} \quad (2.2)$$

onde $U_\gamma \sim \text{Uniforme}(\gamma, 1)$, $\gamma \in (-1, 1]$, pode parecer arbitrária neste ponto, a menos que seja encarada como um primeiro passo para generalizar (2.2) usando uma variável beta geral, de que a uniforme em (2.2) é um caso especial, ou mesmo variáveis aleatórias mais gerais.

Por outro lado, as variáveis aleatórias $N_{0,\beta,\gamma}$ pertencem às classes \mathcal{C}_γ de variáveis aleatórias discretas cuja função massa de probabilidade verifica a equação recursiva

$$\frac{1 - \gamma^{n+1}}{1 - \gamma} f_{N_{\alpha,\beta,\gamma}}(n+1) = \sum_{k=0}^n f_{N_{\alpha,\beta,\gamma}}(k) r_{n-k}, \quad r_n \geq 0, \quad n = 0, 1, \dots$$

as quais interpolam classes importantes na teoria da \mathcal{N} -divisibilidade infinita. De facto, \mathcal{C}_0 é a classe das geo-infinitamente divisíveis discretas (geométricas compostas) e \mathcal{C}_1 é a classe das infinitamente divisíveis discretas (Poisson compostas), e, para $0 \leq \gamma_1 < \gamma_2 \leq 1$, verifica-se a inclusão $\mathcal{C}_0 \subseteq \mathcal{C}_{\gamma_1} \subset \mathcal{C}_{\gamma_2} \subseteq \mathcal{C}_1$.

Somas aleatórias com subordinadora geométrica

Consideremos uma soma aleatória $S_{N_{\alpha,\beta}} = \sum_{k=1}^{N_{\alpha,\beta}} Y_k$, em que $Y_k \stackrel{d}{=} Y$, $k = 1, 2, \dots$, são variáveis de contagem i.i.d., com f.m.p. $\{q_n\}_{n \in \mathbb{N}}$, independente da subordinadora de Panjer $N_{\alpha,\beta}$, e representemos por $\{\tilde{p}_n\}_{n \in \mathbb{N}}$

a f.m.p. de $S_{N_{\alpha, \beta}}$. As distribuições de somas aleatórias geométricas (ou geométricas compostas) têm um papel importante nos esquemas modernos de somas, por intermédio da rarefação elementar, ou emagrecimento.

Notemos que um ponto fulcral na demonstração dos teoremas de representação das somas aleatórias com subordinadora geométrica $N_{\alpha, 0}$ (geométrica) ou $N_{0, \beta}$ (Poisson), via teoria de Panjer, é que $\mathbb{E} \left[\frac{k}{n+1} Y_1 \mid \sum_{i=1}^k Y_i = n+1 \right] = 1$ e $\mathbb{P} \left[Y_1 = j \mid \sum_{i=1}^k Y_i = n+1 \right] = \frac{q_j q_{n+1-j}^{*(k-1)}}{q_{n+1}^{*k}}$, $j = 0, \dots, n+1$ (Rólski *et al.*, 1999, p. 119), onde como é habitual q_n^{*k} representa a convolução de ordem k , e portanto $\sum_{j=0}^{n+1} \frac{q_j q_{n+1-j}^{*(k-1)}}{q_{n+1}^{*k}} = 1$.

$$\begin{aligned}
 \tilde{p}_{n+1} &= \mathbb{P}[S_{N_{\alpha, 0}} = n+1] = \sum_{k=0}^{\infty} \mathbb{P}[S_{N_{\alpha, 0}} = n+1 \mid N_{\alpha, 0} = k] \mathbb{P}[N_{\alpha, 0} = k] = \\
 &= \sum_{k=1}^{\infty} \mathbb{P}[S_k = n+1] \alpha p_{k-1} = \quad \quad \quad (\text{dado que } \mathbb{P}(S_0 \geq 1) = 0) \\
 &= \sum_{k=1}^{\infty} q_{n+1}^{*k} \alpha p_{k-1} \sum_{j=0}^{n+1} \mathbb{P} \left(Y_1 = j \mid \sum_{i=1}^k Y_i = n+1 \right) = \\
 &= \sum_{k=1}^{\infty} q_{n+1}^{*k} \alpha p_{k-1} \sum_{j=0}^{n+1} \frac{q_j q_{n+1-j}^{*(k-1)}}{q_{n+1}^{*k}} = \\
 &= \sum_{j=0}^{n+1} \alpha q_j \sum_{k=1}^{\infty} p_{k-1} q_{n+1-j}^{*(k-1)} = \sum_{j=0}^{n+1} \tilde{p}_{n+1-j} \alpha q_j = \\
 &= \alpha q_0 \tilde{p}_{n+1} + \sum_{k=0}^n \tilde{p}_k \alpha q_{n+1-k}.
 \end{aligned}$$

Portanto

$$\tilde{p}_{n+1} = \sum_{k=0}^n \tilde{p}_k r_{n-k} \quad (2.3)$$

onde $r_k = \frac{\alpha q_{k+1}}{1 - \alpha q_0}$.

Somas aleatórias com subordinadora de Poisson

$$\begin{aligned}
\tilde{p}_{n+1} &= \mathbb{P}[S_{N_{0,\beta}} = n+1] = \sum_{k=0}^{\infty} \mathbb{P}[S_{N_{0,\beta}} = n+1 \mid N_{0,\beta} = k] \mathbb{P}[N_{0,\beta} = k] = \\
&= \sum_{k=1}^{\infty} \mathbb{P}[S_k = n+1] \frac{\beta}{k} p_{k-1} = \quad (\text{dado que } \mathbb{P}(S_0 \geq 1) = 0) \\
&= \sum_{k=1}^{\infty} q_{n+1}^{*k} \frac{\beta}{k} p_{k-1} \mathbb{E} \left(\frac{k}{n+1} Y_1 \mid \sum_{i=1}^k Y_i = n+1 \right) = \\
&= \sum_{k=1}^{\infty} q_{n+1}^{*k} \frac{\beta}{k} p_{k-1} \sum_{j=0}^{n+1} \frac{jk}{n+1} \frac{q_j q_{n+1-j}^{*(k-1)}}{q_{n+1}^{*k}} = \\
&= \sum_{j=0}^{n+1} \beta \frac{jq_j}{n+1} \sum_{k=1}^{\infty} p_{k-1} q_{n+1-j}^{*(k-1)} = \sum_{j=1}^{n+1} \tilde{p}_{n+1-j} \frac{\beta j}{n+1} q_j.
\end{aligned}$$

Portanto

$$(n+1) \tilde{p}_{n+1} = \sum_{k=0}^n \tilde{p}_k \beta (n+1-k) q_{n+1-k} = \sum_{k=0}^n \tilde{p}_k r_{n-k}, \quad (2.4)$$

com $r_k = \beta(k+1)q_{k+1} \geq 0$, $k = 0, 1, \dots$, e conclui-se que a função geradora $\mathcal{H}_{N_{0,\beta}}(s) = \sum_{k=0}^{\infty} r_k s^k$ of the $\{r_k\}_{k \in \mathbb{N}}$ é absolutamente monótona.

Por outro lado,

$$\sum_{k=0}^{\infty} \frac{r_k}{k+1} = \sum_{k=0}^{\infty} \beta q_{k+1} = \beta(1 - q_0). \quad (2.5)$$

Se nos restringirmos ao caso em que $q_0 = 0$ (i.e., impusermos uma representação única fixando este parâmetro livre), obtemos uma nova demonstração da conhecida caracterização das distribuições de contagem infinitamente divisíveis, i.e. somas aleatórias discretas com subordinadora de Poisson (Poisson compostas), com $\tilde{p}_0 > 0$ (cf. Feller, 1968, cap. XI), baseada na teoria de Panjer, como foi visto acima.

Seria interessante estender estes resultados a uma variável aleatória $N_{\alpha,\beta,\gamma}$ qualquer. A estrutura das duas demonstrações anteriores sugere que para isso bastaria achar uma função $f : \{(y, s) : 0 < y < s\} \rightarrow \mathbb{R}$ tal que a expressão para $m_n = \mathbb{E}[f(Y_1, s) \mid \sum_{j=1}^n Y_j = s]$ não dependesse de s , e simplificasse convenientemente factor $\frac{1}{1-\gamma^k}$ (onde os $Y_j \stackrel{d}{=} Y$, $j = 1, 2, \dots$ são variáveis de contagem i.i.d.) Infelizmente, não foi possível descobrir tal função.

2.2.2 Interpretação das classes \mathcal{C}_γ

A segunda parte de cada demonstração consiste em aproveitar a relação recursiva entre \tilde{p}_{n+1} e a convolução $\sum_{k=0}^n \tilde{p}_k r_{n-k}$ para encontrar uma equação para a função geradora de probabilidade da soma aleatória. Mais pormenorizadamente, multiplicando ambos os membros das relações recursivas mencionadas anteriormente por s^n e somando para $n = 0, 1, \dots$, tem-se o seguinte.

Somas aleatórias com subordinadora geométrica

$$\sum_{n=0}^{\infty} \tilde{p}_{n+1} s^n = \sum_{n=0}^{\infty} \underbrace{\left(\sum_{k=0}^n \tilde{p}_k r_{n-k} \right)}_{(\{\tilde{p}_n\} \star \{r_n\})} s^n$$

ou

$$\frac{\mathcal{G}_{s_N}(s) - p_0}{s} = \mathcal{G}_{s_N}(s) \mathcal{H}_{N_{\alpha,0}}(s)$$

onde $\mathcal{H}_{N_{\alpha,0}}$ é a função geradora dos $\{r_n\}_{n \in \mathbb{N}}$, que são todos não negativos,

com $\sum_{n=0}^{\infty} r_n = \alpha \in (0, 1)$, i.e. $\mathcal{H}_{N_{\alpha,0}}$ é abs. mon.

Como $1 = \mathcal{G}_{s_N}(1) = \frac{p_0}{1 - \mathcal{H}_{N_{\alpha,0}}(1)} = \frac{p_0}{1 - \alpha}$, resulta que

$$\mathcal{G}_{s_N}(s) = \frac{p_0}{1 - s \mathcal{H}_{N_{\alpha,0}}(s)} = \frac{1 - \alpha}{1 - s \sum_{k=0}^{\infty} r_k s^k} = \frac{1 - \alpha}{1 - \alpha \mathcal{P}(s)}$$

onde $\mathcal{P}(s) = \sum_{k=0}^{\infty} \frac{r_k}{\alpha} s^{k+1}$, tal que $\mathcal{P}(0) = 0$, é uma f.g.p., pois é abs. mon. e $\mathcal{P}(1) = 1$.

Somas aleatórias com subordinadora de Poisson

$$\sum_{n=0}^{\infty} (n+1) \tilde{p}_{n+1} s^n = \sum_{n=0}^{\infty} \underbrace{\left(\sum_{k=0}^n \tilde{p}_k r_{n-k} \right)}_{(\{\tilde{p}_n\} \star \{r_n\})} s^n$$

ou

$$\begin{aligned} \mathcal{G}'_{s_{N_0, \beta}}(s) = \mathcal{G}_{s_{N_0, \beta}}(s) \mathcal{H}_{N_0, \beta}(s) &\Longleftrightarrow \frac{\mathcal{G}'_{s_{N_0, \beta}}(s)}{\mathcal{G}_{s_{N_0, \beta}}(s)} = \mathcal{H}_{N_0, \beta}(s) = \sum_{k=0}^{\infty} r_k s^k \Longleftrightarrow \\ &\Longleftrightarrow \mathcal{G}_{s_{N_0, \beta}}(s) = \exp \left(\sum_{k=0}^{\infty} r_k \frac{s^{k+1}}{k+1} + C \right). \end{aligned}$$

Como $1 = \mathcal{G}_{s_{N_0, \beta}}(1) = \exp \left(\sum_{k=0}^{\infty} \frac{r_k}{k+1} + C \right) = e^{\beta+C}$, resulta que $C = -\beta$, e podemos reescrever

$$\mathcal{G}_{s_{N_0, \beta}}(s) = \exp \left[\beta \left(\frac{1}{\beta} \sum_{k=0}^{\infty} \frac{r_k}{k+1} s^{k+1} - 1 \right) \right] = e^{\beta[\mathcal{P}(s)-1]}, \quad (2.6)$$

onde $\mathcal{P}(s) = \frac{1}{\beta} \sum_{k=0}^{\infty} \frac{r_k}{k+1} s^{k+1}$ é uma f.g.p. (única), tal que $\mathcal{P}(0) = 0$.

Tal como no tratamento das somas aleatórias com subordinadora geométrica, conseguimos um teorema de representação única pondo o parâmetro livre $q_0 = 0$, o que implica $\sum_{k=0}^{\infty} r_k = \alpha$:

A representação para a função geradora de probabilidade dos elementos das classes \mathcal{C}_γ :

$$\mathcal{G}(s) = \prod_{k=0}^{\infty} \frac{1 - (1-\gamma) \gamma^k \mathcal{H}_\gamma(\gamma^k)}{1 - (1-\gamma) \gamma^k s \mathcal{H}_\gamma(\gamma^k s)} = \prod_{k=0}^{\infty} \frac{1 - w_k}{1 - w_k \mathcal{P}_k(s)}, \quad (2.7)$$

onde $w_k = (1-\gamma) \gamma^k \in [0, 1]$ e os $\mathcal{P}_k(s) = \frac{s \mathcal{H}_\gamma(\gamma^k s)}{\mathcal{H}_\gamma(\gamma^k)}$ são funções geradoras de probabilidade (únicas) tais que $\mathcal{P}_k(0) = 0$, permite interpretar $W \in \mathcal{C}_\gamma$ como somas de geométricas compostas, quando $\gamma \in [0, 1]$. No caso de $\gamma \in (-1, 0)$, deixa de ser válida esta interpretação directa, visto que γ^k toma valores alternadamente positivos e negativos.

Recordando que os elementos das classes \mathcal{C}_γ foram definidas como variáveis aleatórias de contagem cuja função massa de probabilidade satisfaz a relação recursiva geral

$$\frac{1 - \gamma^{n+1}}{1 - \gamma} \tilde{p}_{n+1} = \sum_{k=0}^n \tilde{p}_k r_{n-k}, \quad r_n \geq 0, \quad \tilde{p}_0 > 0, \quad (2.8)$$

que generaliza as classes de somas aleatórias com subordinadora Poisson (\mathcal{C}_1) e de somas aleatórias com subordinadora geométrica (\mathcal{C}_0), seria também interessante relacionar $W \in \mathcal{C}_\gamma$ (γ negativo ou não) com as somas aleatórias de subordinadora $N_{0,\beta,\gamma}$, ou seja as variáveis aleatórias discretas com função geradora de probabilidade da forma:

$$\mathcal{Q}(s) = \prod_{k=0}^{\infty} \frac{1 - \beta(1 - \gamma)}{1 - \beta(1 - \gamma) \mathcal{P}(s)}, \quad (2.9)$$

onde $\mathcal{P}(s)$ é uma função geradora de probabilidade tal que $\mathcal{P}_k(0) = 0$, mas não pôde ser alcançado esse objectivo do presente trabalho.

2.2.3 Generalização dos resultados obtidos a $\gamma < 0$ e relação com a \mathcal{N} -divisibilidade infinita

A divisibilidade infinita geométrica surgiu das extensões feitas por Kovalenko (1965) ao trabalho de Rényi sobre rarefacção aleatória (1956), tendo a caracterização geral das leis geométricas estáveis sido feita por Kozubowski (1994). Isso levou à definição geral de esquemas de somas de classe \mathcal{N} , em que o esquema aditivo clássico é o caso especial $N_p = \frac{1}{p}$ (variáveis aleatórias degeneradas, e portanto uma soma não aleatória de variáveis aleatórias). É bem sabido que para algumas famílias $\mathcal{N} = \{N_p, p \in (0, 1), \mathbb{E}(N_p) = \frac{1}{p}\}$ existem leis \mathcal{N} -gaussianas (por exemplo, para $N_p \curvearrowright \text{Geométrica}(p)$, as variáveis aleatórias \mathcal{N} -gaussianas correspondentes são variáveis aleatórias de Laplace), enquanto outras famílias de variáveis N_p , tais como $N_p \curvearrowright \text{Poisson}\left(\frac{1}{p}\right)$, não admitem leis \mathcal{N} -gaussianas. Ainda que seja fácil provar que existem leis \mathcal{N} -gaussianas em leis de processos de ramificação mais gerais, apenas a lei gaussiana usual e a distribuição de Laplace, lei geométrica-gaussiana são indentificadas explicitamente nas referências que conhecemos.

Esta investigação partiu da observação de que $\mathcal{C}_0 \subset \mathcal{C}_1$ e, mais geralmente, $0 < \gamma_1 < \gamma_2 < 1 \implies \mathcal{C}_0 \subset \mathcal{C}_{\gamma_1} \subset \mathcal{C}_{\gamma_2} \subset \mathcal{C}_1$. O nosso objectivo era provar que existia um $\gamma \in (0, 1)$ tal que para $\gamma_1 \leq \gamma$ se conseguia indicar a lei \mathcal{N} -gaussiana em \mathcal{C}_{γ_1} — o que não lográmos — ou então prolongar as classes \mathcal{C}_γ a $\gamma < 0$ — o que fizemos — e mostrar que para essa extensão era possível construir variáveis aleatórias \mathcal{N} -gaussianas. Infelizmente, para $-1 < \gamma_1 < \gamma_2 < 0$ a cadeia de inclusões $\mathcal{C}_{\gamma_1} \subset \mathcal{C}_{\gamma_2} \subset \mathcal{C}_0$ deixa de ser válida.

2.2.4 Distribuições de Panjer generalizadas com suporte finito

No artigo, apresentou-se uma generalização da classe de Panjer, composta pelas variáveis aleatórias discretas $N_{\alpha,\beta}$, cuja função massa de probabilidade (f.m.p.) $\left\{f_{N_{\alpha,\beta}}(n)\right\}_{n \in \mathbb{N}}$ satisfaz

$$f_{N_{\alpha,\beta}}(n+1) = \left(\alpha + \frac{\beta}{n+1}\right) f_{N_{\alpha,\beta}}(n), \quad n = 0, 1, \dots$$

Para tal, considerámos as variáveis aleatórias discretas $N_{\alpha,\beta,\gamma}$, com $\alpha, \beta \in \mathbb{R}$, cuja função massa de probabilidade $\{p_n\}_{n \in \mathbb{N}} = \left\{f_{N_{\alpha,\beta,\gamma}}(n)\right\}_{n \in \mathbb{N}}$ satisfaz a relação mais geral

$$\frac{f_{N_{\alpha,\beta,\gamma}}(n+1)}{f_{N_{\alpha,\beta,\gamma}}(n)} = \alpha + \beta \frac{\mathbb{E}(U_0^n)}{\mathbb{E}(U_\gamma^n)} = \alpha + \frac{\beta}{\sum_{k=0}^n \gamma^k}, \quad n = 0, 1, \dots \quad (2.10)$$

onde $U_\gamma \sim \text{Uniforme}(\gamma, 1)$, $\gamma \in (-1, 1]$. Como

$$\mathbb{E}(U_\gamma^n) = \frac{1}{n+1} \frac{1 - \gamma^{n+1}}{1 - \gamma} \xrightarrow{\gamma \rightarrow 1} 1,$$

a classe de Panjer corresponde ao caso limite degenerado que se obtém fazendo $\gamma \rightarrow 1$, de forma que $U_\gamma \rightarrow U_1$, a variável aleatória degenerada com massa unitária no 1.

A investigação focou-se nas soluções com suporte infinito, que eram as mais interessantes do ponto de vista da investigação que nos propusemos fazer, já que só estas podem ser infinitamente divisíveis. No entanto, estudámos também o caso de funções massa de probabilidade com suporte finito, confirmando assim que a extensão que definimos abrangia a classe de Panjer.

Teorema 2.2.1: *Seja $\gamma \in (-1, 1)$. As soluções $N_{\alpha,\beta,\gamma}$ de (2.10) com suporte finito são:*

- (a) a degenerada em 0, se $\alpha + \beta = 0$;
- (b) a Bernoulli $\left(\frac{-\alpha\gamma}{1-\alpha\gamma}\right)$, se $\alpha\gamma < 0$ e $\alpha + \frac{\beta}{1+\gamma} = 0$;

(c) as variáveis aleatórias com a função massa de probabilidade dada por

$$p_n = \frac{1}{C(\gamma, \alpha)} (-\alpha)^n \prod_{k=1}^n \frac{\gamma^k - \gamma^{N+1}}{1 - \gamma^k}, \quad n = 0, 1, \dots, N,$$

$$\text{onde } C(\gamma, \alpha) = \sum_{n=0}^N (-\alpha)^n \prod_{k=1}^n \frac{\gamma^k - \gamma^{N+1}}{1 - \gamma^k},$$

se $\gamma \in (0, 1)$, $\alpha < 0$ e $N \geq 2$ é um inteiro tal que $\alpha + \beta \frac{1-\gamma}{1-\gamma^{N+1}} = 0$.

Demonstração: (a) Observe-se que $\alpha + \beta \frac{\mathbb{E}(U_0^n)}{\mathbb{E}(U_\gamma^n)} = \alpha + \beta \frac{1-\gamma}{1-\gamma^{n+1}}$ quando $\gamma \neq 1$.

Para que haja solução, temos de ter $p_1 = (\alpha + \beta) p_0 \geq 0$, donde $\alpha + \beta \geq 0$. Se $\alpha + \beta = 0$, a solução é degenerada em 0.

(b) Admita-se $\alpha + \beta > 0$. O suporte de uma solução não degenerada de (2.10) é finito se e só se existe um inteiro $N \geq 1$ para o qual $\alpha + \beta \frac{1-\gamma}{1-\gamma^{N+1}} = 0$. Se $N = 1$, acha-se $\beta = -\alpha(1 + \gamma)$, donde $p_1 = -\alpha\gamma p_0$, e é necessário que $\alpha\gamma < 0$. Nestas condições, $p_0 = \frac{1}{1-\alpha\gamma}$, $p_1 = \frac{-\alpha\gamma}{1-\alpha\gamma}$, e a solução é uma variável aleatória de Bernoulli.

(c) Suponha-se agora $\alpha + \beta > 0$ e $\alpha + \beta \frac{\beta}{1+\gamma} \neq 0$, mas $\alpha + \beta \frac{1-\gamma}{1-\gamma^{N+1}} = 0$ para algum inteiro $N \geq 2$. Então, $\gamma \neq 0$. Tomando $\alpha < 0$ arbitrário, temos de ter $\beta > 0$ para que $\alpha + \beta > 0$. Se $\gamma \in (0, 1)$, a sucessão $\frac{1-\gamma}{1-\gamma^{n+1}}$ é estritamente decrescente. Nestas condições, $0 = \alpha + \beta \frac{1-\gamma}{1-\gamma^{N+1}} < \alpha + \beta \frac{1-\gamma}{1-\gamma^{n+1}}$ para $n = 0, 1, \dots, N-1$, e portanto a iteração (2.10) define a função massa de probabilidade de uma variável aleatória com suporte finito, que explicitamos adiante.

Mostremos que não há mais soluções com suporte finito. Quando $\gamma \in (0, 1)$, a sucessão estritamente decrescente $\frac{1-\gamma}{1-\gamma^{n+1}}$ tem máximo 1 ($n = 0$) e ínfimo $1 - \gamma$ ($n \rightarrow \infty$). Se $\alpha \geq 0$ e $\beta > 0$, temos $0 < \alpha + \beta(1 - \gamma) < \alpha + \beta \frac{1-\gamma}{1-\gamma^{n+1}}$. O factor $\alpha + \beta \frac{1-\gamma}{1-\gamma^{n+1}}$ não se anula para n finito.

Admitindo $\gamma \in (0, 1)$, $\alpha + \beta > 0$, $\alpha \geq 0$ e $\beta \leq 0$, vem $0 < \alpha + \beta \leq \alpha + \beta \frac{1-\gamma}{1-\gamma^{n+1}}$, para todo o n .

Consideremos agora $\gamma \in (-1, 0)$, $\alpha + \beta > 0$. Então, a sucessão $\frac{1-\gamma}{1-\gamma^{n+1}}$ é constituída por duas subsucessões monótonas, uma crescente, para n par, e a outra decrescente, para n ímpar. O seu mínimo é 1 ($n = 0$), e o máximo $\frac{1}{1+\gamma}$ ($n = 1$). Se $\beta > 0$, então $0 < \alpha + \beta \leq \alpha + \beta \frac{1-\gamma}{1-\gamma^{n+1}}$ para

todo o n , e $\{p_n\}_{n \in \mathbb{N}}$ não pode ter suporte finito.

Se $\beta \leq 0$, então $\alpha + \frac{\beta}{1+\gamma} \leq \alpha + \beta \frac{1-\gamma}{1-\gamma^{n+1}}$. Não podemos ter $\alpha + \frac{\beta}{1+\gamma} < 0$, pois isso implica $p_1 < 0$. Então só pode ser $0 < \alpha + \frac{\beta}{1+\gamma} \leq \alpha + \beta \frac{1-\gamma}{1-\gamma^{n+1}}$, e de novo não há solução com suporte finito.

Passemos à descrição da solução, quando existe.

A condição $\alpha + \beta \frac{1-\gamma}{1-\gamma^{N+1}} = 0$ equivale a $\frac{1-\gamma^{N+1}}{1-\gamma} = 1 + \gamma + \dots + \gamma^N = -\frac{\beta}{\alpha}$, donde vem:

$$p_1 = \alpha \left(1 + \frac{\beta}{\alpha}\right) p_0 = -\alpha (\gamma + \dots + \gamma^N) p_0;$$

$$p_2 = \alpha \left(1 + \frac{\beta}{\alpha} \frac{1}{1+\gamma}\right) p_1 = \alpha \left(1 - \frac{1+\gamma+\dots+\gamma^N}{1+\gamma}\right) p_1 = (-\alpha)^2 \frac{\gamma^2+\dots+\gamma^N}{1+\gamma} \frac{\gamma+\dots+\gamma^N}{1} p_0;$$

e, em geral,

$$p_n = p_0 (-\alpha)^n \prod_{k=1}^n \frac{\sum_{j=k}^N \gamma^j}{\sum_{j=0}^{k-1} \gamma^j} = p_0 (-\alpha)^n \prod_{k=1}^n \frac{\gamma^k - \gamma^{N+1}}{1 - \gamma^k},$$

para $n = 1, 2, \dots, N$. O valor de p_0 fica determinado pela condição normalizadora

$$\sum_{n=0}^N p_n = 1 \iff p_0^{-1} = 1 + \sum_{n=1}^N \prod_{k=1}^n \frac{\gamma^k - \gamma^{N+1}}{1 - \gamma^k} (> 1).$$

É fácil ver então que $p_n = p_n(\gamma) \rightarrow \binom{N}{n} \left(\frac{-\alpha}{1-\alpha}\right)^n \left(\frac{1}{1-\alpha}\right)^{N-n}$ quando $\gamma \rightarrow 1$, a função massa de probabilidade de $N_{\alpha, \beta} \curvearrowright \text{Binomial}(N, \frac{-\alpha}{1-\alpha})$.

□

2.2.5 Expressão das variáveis $N_{\alpha, \beta, \gamma}$ como somas aleatórias

Naturalmente, $\left\{f_{N_{\alpha, \beta, \gamma}}(n)\right\}_{n \in \mathbb{N}}$ apenas poderá ser uma função massa de probabilidade se forem impostas restrições bastante fortes aos seus parâmetros. Nesta secção, caracterizamos com mais pormenor as

variáveis aleatórias discretas $N_{\alpha, \beta, \gamma}$ com suporte finito, cuja f.m.p. $\{p_n\}_{n \in \mathbb{N}} = \left\{ f_{N_{\alpha, \beta, \gamma}}(n) \right\}_{n \in \mathbb{N}}$ satisfaz a relação (2.10).

Lema 2.2.1: *Seja $\gamma \in (-1, 1)$. Para que exista $N_{\alpha, \beta, \gamma}$ com suporte infinito, é necessário e suficiente que se verifique uma das seguintes condições mutuamente exclusivas:*

- (a) $\gamma\beta = 0$ e $0 < \alpha + \beta < 1$;
- (b) $\gamma \in (0, 1)$, $\alpha < 1$, $\max \left\{ 0, \frac{-\alpha}{1-\gamma} \right\} \leq \beta < \frac{1-\alpha}{1-\gamma}$ e $\beta \neq 0$;
- (c) $\gamma \in (0, 1)$, $0 < \alpha < \frac{1}{\gamma}$ e $-\alpha < \beta < \min \left\{ 0, \frac{1-\alpha}{1-\gamma} \right\}$;
- (d) $\gamma \in (-1, 0)$, $\frac{1}{\gamma} < \alpha < 1$ e $\max \{0, -\alpha\} < \beta < \frac{1-\alpha}{1-\gamma}$;
- (e) $\gamma \in (-1, 0)$, $0 < \alpha < \frac{1}{\gamma^2}$ e $-\alpha(1+\gamma) < \beta < \min \left\{ 0, \frac{1-\alpha}{1-\gamma} \right\}$.

Demonstração: Começemos pelo caso (a). Se $\gamma = 0$ ou $\beta = 0$, então a progressão geométrica $\prod_{k=0}^{n-1} \left(\alpha + \beta \frac{\mathbb{E}(U_0^k)}{\mathbb{E}(U_\gamma^k)} \right) = \prod_{k=0}^{n-1} \left(\alpha + \beta \frac{1-\gamma}{1-\gamma^{k+1}} \right) = (\alpha + \beta)^n$, $n = 1, 2, \dots$, é somável e não negativa se e só se $0 < \alpha + \beta < 1$. Como é estritamente decrescente, $\{p_n\}_{n \in \mathbb{N}}$ tem suporte infinito.

Para $\gamma \neq 0$ e $\beta \neq 0$, há quatro casos principais a distinguir: $\gamma \in (0, 1)$ e $\beta > 0$ (b); $\gamma \in (0, 1)$ e $\beta < 0$ (c); $\gamma \in (-1, 0)$ e $\beta > 0$ (d); $\gamma \in (-1, 0)$ e $\beta < 0$ (e).

Conforme os sinais de γ e β , o ínfimo de $\alpha + \beta \frac{1-\gamma}{1-\gamma^{n+1}}$ é $\alpha + \beta$ ($n = 0$), $\alpha + \frac{\beta}{1+\gamma}$ ($n = 1$), ou $\alpha + \beta(1-\gamma)$ ($n \rightarrow \infty$). Assim, para garantir que a sucessão $\{p_n\}_{n \in \mathbb{N}}$ definida por (2.10) é positiva, basta exigir

$$\alpha + \beta > 0, \quad \alpha + \frac{\beta}{1+\gamma} > 0 \quad \text{e} \quad \alpha + \beta(1-\gamma) \geq 0,$$

donde se tiram as minorações para β . Por outro lado, é necessário que convirja a soma

$$\sum_{n=0}^{\infty} p_n = p_0 \left[1 + \sum_{n=1}^{\infty} \prod_{k=0}^{n-1} \left(\alpha + \beta \frac{\mathbb{E}(U_0^k)}{\mathbb{E}(U_\gamma^k)} \right) \right].$$

Nesse caso, a condição normalizadora $\sum p_n = 1$ define p_0 , e verifica-se $0 < p_0 < 1$.

Aplicando o teste da razão ao termo geral do somatório, conclui-se que há convergência para $\alpha + \beta(1 - \gamma) < 1$ e divergência para $\alpha + \beta(1 - \gamma) > 1$. O caso $\alpha + \beta(1 - \gamma) = 1$ será estudado mais adiante.

Conjugando a majoração $\alpha + \beta(1 - \gamma) < 1$ com as minorações acima referidas, obtemos as duplas desigualdades para β . Por exemplo, consideremos o caso $\gamma \in (0, 1)$ e $\beta > 0$. Então tem-se $\min\{\alpha + \beta, \alpha + \frac{\beta}{1+\gamma}, \alpha + \beta(1 - \gamma)\} = \alpha + \beta(1 - \gamma)$, pelo que existe uma solução $p_n > 0$, para qualquer n inteiro quando $\alpha + \beta(1 - \gamma) \geq 0$. Resolvendo $0 \leq \alpha + \beta(1 - \gamma) < 1$ em ordem a β , tira-se $\max\{0, \frac{-\alpha}{1-\gamma}\} \leq \beta < \frac{1-\alpha}{1-\gamma}$ e $\beta \neq 0$, pois β é por hipótese positivo. Mas esta condição só faz sentido se $\frac{-\alpha}{1-\gamma} < 0 < \frac{1-\alpha}{1-\gamma}$ ou se $0 \leq \frac{-\alpha}{1-\gamma} < \frac{1-\alpha}{1-\gamma}$. Deste par de conjunções resulta $\alpha < 1$.

No caso $\gamma \in (-1, 0)$ e $\beta > 0$, temos $\min\{\alpha + \beta, \alpha + \frac{\beta}{1+\gamma}, \alpha + \beta(1 - \gamma)\} = \alpha + \beta$. Exigindo $\alpha + \beta > 0$ e $\alpha + \beta(1 - \gamma) < 1$, tira-se a desigualdade $\max\{0, -\alpha\} < \beta < \frac{1-\alpha}{1-\gamma}$. Para que exista um β nestas condições, é preciso ter $-\alpha < 0 < \frac{1-\alpha}{1-\gamma}$ ou $0 \leq -\alpha < \frac{1-\alpha}{1-\gamma}$, donde se tira $\frac{1}{\gamma} < \alpha < 1$.

De forma análoga se faz o estudo dos casos (c) e (e), quando $\alpha + \beta(1 - \gamma) < 1$.

O caso $\alpha + \beta(1 - \gamma) = 1$, no qual $\alpha + \beta \frac{\mathbb{E}(U_0^k)}{\mathbb{E}(U_\gamma^k)} = \alpha - \frac{\alpha-1}{1-\gamma^{k+1}} = \frac{1-\alpha\gamma^{k+1}}{1-\gamma^{k+1}}$, tem de ser estudado separadamente. Note-se que $\beta \neq 0 \iff \alpha \neq 1$.

- Se $\gamma \in (0, 1)$ e $\alpha < 1$, esta sucessão é estritamente decrescente, com ínfimo 1, o que não pode ser, porque implica $\sum p_n > \sum p_0 = \infty$.
- Para $\gamma \in (-1, 0)$ e $\alpha < 1$, tem-se $\beta > 0$ e $1 < \alpha + \frac{\beta}{1+|\gamma|+|\gamma|^2+\dots+|\gamma|^k} < \alpha + \frac{\beta}{1+\gamma+\gamma^2+\dots+\gamma^k}$. Pelo que foi visto acima, não há solução.
- Se $\gamma \in (0, 1)$ e $\alpha > 1$, então a sucessão é estritamente crescente e tem mínimo $\frac{1-\alpha\gamma}{1-\gamma}$ ($k = 0$). Este valor é positivo quando $\alpha\gamma < 1$. Mas então $0 < \frac{1-\gamma^k}{1-\gamma^{k+1}} < \frac{1-\alpha\gamma^{k+1}}{1-\gamma^{k+1}}$ para $k \geq 1$, donde

$$\prod_{k=0}^n \left(\alpha + \beta \frac{\mathbb{E}(U_0^k)}{\mathbb{E}(U_\gamma^k)} \right) \geq \frac{1-\alpha\gamma}{1-\gamma} \prod_{k=1}^n \frac{1-\gamma^k}{1-\gamma^{k+1}} = \frac{1-\alpha\gamma}{1-\gamma^{n+1}} \xrightarrow{n \rightarrow \infty} 1 - \alpha\gamma > 0,$$

e portanto a série $\sum p_n$ diverge.

- Para $\gamma \in (-1, 0)$ e $\alpha > 1$, a subsucessão $\alpha - \frac{\alpha-1}{1+|\gamma|^{2j+1}}$ é decrescente e tem por ínfimo 1. A subsucessão $\alpha - \frac{\alpha-1}{1-|\gamma|^{2j+2}}$ é crescente e tem mínimo $\frac{1-\alpha\gamma^2}{1-\gamma^2}$ ($j = 0$), que é positivo quando $\alpha\gamma^2 < 1$. Mas então $0 < \frac{1-\gamma^{2j}}{1-\gamma^{2j+2}} < \frac{1-\alpha\gamma^{2j+2}}{1-\gamma^{2j+2}}$ para $j \geq 1$. Assim,

$$\begin{aligned} \prod_{k=0}^{2n+1} \frac{1-\alpha\gamma^{k+1}}{1-\gamma^{k+1}} &> \prod_{j=0}^n \frac{1-\alpha\gamma^{2j+2}}{1-\gamma^{2j+2}} \geq \frac{1-\alpha\gamma^2}{1-\gamma^2} \prod_{j=1}^n \frac{1-\gamma^{2j}}{1-\gamma^{2j+2}} = \\ &= \frac{1-\alpha\gamma^2}{1-\gamma^{2n+2}} \xrightarrow{n \rightarrow \infty} 1-\alpha\gamma^2 > 0, \end{aligned}$$

e também neste caso não há solução, pois $\sum p_n$ diverge.

□

Lema 2.2.2: *Suponhamos que existe $N_{\alpha, \beta, \gamma} \in \Pi_{\alpha, \beta, \gamma}$ com suporte infinito. Então:*

1. $\alpha + \beta > 0$ e $0 \leq \alpha + \beta(1 - \gamma) < 1$, com desigualdade estrita se $\gamma \in (-1, 0)$ ou $\alpha \geq 0$;
2. $1 - \alpha\gamma^{k+1} > 0$, para $k \geq 0$;
3. Se $\gamma \in (-1, 0)$ e $\beta \geq 0$ ou $-\alpha\frac{1+\gamma}{1-\gamma^3} < \beta < 0$, tem-se $\alpha(1 + \gamma) + \beta(1 - \gamma^{k+1}) > 0$, para $k \geq 1$.

Demonstração: (1) Imediata.

(2) O resultado é trivialmente válido se $\gamma = 0$, ou $\gamma \in (0, 1)$ e $\alpha \leq 0$. Então, consideremos $\gamma \in (0, 1)$ e $\alpha > 0$. Quer no caso (b) do lema 2.2.1, quer no caso (c), ficamos com $0 < \alpha < \frac{1}{\gamma}$, donde vem $\alpha\gamma^{k+1} \leq \alpha\gamma < 1$, e o resultado segue. Suponhamos agora que γ é negativo; então, $\frac{1}{\gamma} < \alpha < \frac{1}{\gamma^2}$. Se $\alpha \leq 0$, vem $0 \leq \alpha\gamma < 1$, donde $\alpha\gamma^{k+1} \leq \alpha\gamma < 1$, e o resultado segue. Se $\alpha > 0$, vem $\alpha\gamma < 1$ e $0 < \alpha\gamma^2 < 1$, donde $\alpha\gamma^{k+1} < \alpha\gamma^2 < 1$ para $k \geq 1$.

(3) De facto, $1 - \gamma^{k+1} = (1 - \gamma)(1 + \gamma + \dots + \gamma^k)$ tem mínimo $(1 - \gamma)(1 + \gamma)$ (para $k = 1$) e máximo $(1 - \gamma)(1 + \gamma + \gamma^2) = 1 - \gamma^3$ (para $k = 2$). Portanto,

se $\beta \geq 0$ tem-se

$$\alpha(1+\gamma) + \beta(1-\gamma^{k+1}) \geq \alpha(1+\gamma) + \beta(1-\gamma)(1+\gamma) = [\alpha + \beta(1-\gamma)](1+\gamma),$$

que é positivo.

Se $\beta < 0$, tem-se $\alpha > 0$ e $\alpha(1+\gamma) + \beta(1-\gamma^{k+1}) \geq \alpha(1+\gamma) + \beta(1-\gamma^3)$. Como $\beta > -\alpha \frac{1+\gamma}{1-\gamma^3}$ equivale a $\alpha(1+\gamma) + \beta(1-\gamma^3) > 0$, fica estabelecido o resultado.

Para caracterizar as variáveis $N_{\alpha, \beta, \gamma} \in \Pi_{\alpha, \beta, \gamma}$, bem como para generalizações posteriores, é necessário usar o facto de que o produtório

$$\prod_{k=0}^{\infty} \frac{1 - \alpha\gamma^{k+1}s}{1 - \alpha\gamma^{k+1}} \frac{1 - [\alpha + \beta(1-\gamma)]\gamma^k}{1 - [\alpha + \beta(1-\gamma)]\gamma^k s} \quad (2.11)$$

converge absolutamente para todos os valores de α, β e $|s| \leq 1$, quando $|\gamma| < 1$.

De facto, escrevendo

$$\prod_{k=0}^{\infty} \frac{1 - [\alpha + \beta(1-\gamma)]\gamma^k}{1 - [\alpha + \beta(1-\gamma)]\gamma^k s} = \prod_{k=0}^{\infty} \left(1 + \frac{[\alpha + \beta(1-\gamma)]\gamma^k(s-1)}{1 - [\alpha + \beta(1-\gamma)]\gamma^k s} \right) = \prod_{k=0}^{\infty} (1 + a_k) \quad (2.12)$$

e fazendo o teste da razão,

$$\left| \frac{a_{k+1}}{a_k} \right| = |\gamma| \left| \frac{1 - [\alpha + \beta(1-\gamma)]\gamma^k s}{1 - [\alpha + \beta(1-\gamma)]\gamma^{k+1} s} \right| \xrightarrow{k \rightarrow \infty} |\gamma|,$$

vê-se que este produtório converge absolutamente para $|s| \leq 1$, quaisquer que sejam α e β (ver, por exemplo, Rohatgi, 1976, p. 8). De modo análogo se justifica que

$$\prod_{k=0}^{\infty} \frac{1 - \alpha\gamma^{k+1}s}{1 - \alpha\gamma^{k+1}},$$

converge absolutamente para qualquer α , donde vem o resultado. Um desenvolvimento análogo mostra que o produtório (2.11) diverge se $|\gamma| > 1$.

Teorema 2.2.2: *Seja $\gamma \in (-1, 1)$. Se existe $N_{\alpha, \beta, \gamma} \in \Pi_{\alpha, \beta, \gamma}$ com suporte infinito, então a correspondente função geradora de probabilidade é dada, para $|s| \leq 1$, por:*

$$\mathcal{G}_{\alpha, \beta, \gamma}(s) = \prod_{k=0}^{\infty} \frac{1 - \alpha\gamma^{k+1}s}{1 - \alpha\gamma^{k+1}} \frac{1 - [\alpha + \beta(1-\gamma)]\gamma^k}{1 - [\alpha + \beta(1-\gamma)]\gamma^k s}.$$

A demonstração deste resultado encontra-se no artigo.

O caso limite $\gamma \rightarrow 1$ pode ser estabelecido directamente, em vez de passar por uma equação diferencial como no artigo, embora seja mais fastidioso. Por exemplo, no caso especialmente simples $\alpha = 0$, fazendo $\gamma = 1 - \frac{1}{n} \xrightarrow{n \rightarrow \infty} 1$,

$$\begin{aligned} \mathcal{G}_{0,\beta,\gamma}(s) &= \prod_{k=0}^{\infty} \frac{1 - \beta(1-\gamma)\gamma^k}{1 - \beta(1-\gamma)\gamma^k s} = \lim_{n \rightarrow \infty} \prod_{k=0}^{n-1} \frac{1 - \beta(1-\gamma)\gamma^k}{1 - \beta(1-\gamma)\gamma^k s} = \\ &= \lim_{n \rightarrow \infty} \left[1 + \frac{\beta(s-1)}{n} \right]^n \prod_{k=0}^{n-1} \left(1 + \frac{\beta(s-1)}{n} \frac{\gamma^k - 1 + \frac{\beta}{n} \gamma^k s}{\left[1 - \frac{\beta}{n}(1-\gamma)\gamma^k s \right] \left[1 + \frac{\beta}{n}(s-1) \right]} \right) = \\ &= \lim_{n \rightarrow \infty} \left[1 + \frac{\beta(s-1)}{n} \right]^n \prod_{k=0}^{n-1} \left[1 + \frac{C\beta^2}{n^2} \left(1 + o\left(\frac{1}{n}\right) \right) \right] \xrightarrow{n \rightarrow \infty} e^{\beta(s-1)} = \mathcal{G}_{0,\beta,1}(s). \end{aligned}$$

Salientemos outros casos notáveis:

- Se $\gamma = 0$ ou $\beta = 0$, a fórmula anterior reduz-se evidentemente à função geradora de probabilidade de $X \curvearrowright Geométrica(1 - (\alpha + \beta))$,

$$\mathcal{G}_{\alpha,\beta,\gamma}(s) = \frac{1 - (\alpha + \beta)}{1 - (\alpha + \beta)s};$$

- Se $\gamma \in (0, 1)$ e $\alpha = 0$, então $0 < \beta < \frac{1}{1-\gamma}$ e a fórmula reduz-se a

$$\mathcal{G}_{\alpha,\beta,\gamma}(s) = \prod_{k=0}^{\infty} \frac{1 - \beta(1-\gamma)\gamma^k}{1 - \beta(1-\gamma)\gamma^k s},$$

e a solução é $Y \stackrel{d}{=} \sum_{k=0}^{\infty} W_k$, com os $W_k \curvearrowright Geométrica(1 - \beta(1-\gamma)\gamma^k)$ variáveis aleatórias independentes;

- Se $\gamma \in (0, 1)$ e $\alpha + \beta(1-\gamma) = 0$, então $\alpha < 0$ e fica

$$\mathcal{G}_{\alpha,\beta,\gamma}(s) = \prod_{k=0}^{\infty} \frac{1 - \alpha\gamma^{k+1}s}{1 - \alpha\gamma^{k+1}}.$$

A solução é da forma $Z \stackrel{d}{=} \sum_{k=0}^{\infty} V_k$, onde as variáveis independentes

V_k são *Bernoulli* $\left(\frac{\alpha\gamma^{k+1}}{\alpha\gamma^{k+1}-1}\right)$.

Para $\gamma \in (0, 1)$ e $\alpha < 0$ em geral, as soluções são misturas das anteriores, da forma $Y + Z$, com Y e Z independentes. Quando $\gamma \in (0, 1)$ e $\alpha > 0$, obtêm-se somas da forma $\sum_{k=0}^{\infty} S_k$, onde S_k são geométricas modificadas em zero independentes, como se estabelece no seguinte resultado.

Corolário 2.2.1: *Seja $\gamma \in (-1, 1)$. As soluções $N_{\alpha, \beta, \gamma}$ da iteração (2.10) com suporte infinito, onde α e β verificam as condições necessárias dadas no lema 2.2.1, são da forma:*

(a) $N_{\alpha, \beta, \gamma} \stackrel{d}{=} \sum_{k=0}^{\infty} (V_k + W_k)$, se $\gamma \in [0, 1)$ e $\alpha \leq 0$, onde as variáveis $V_k \sim \text{Bernoulli}\left(\frac{\alpha\gamma^{k+1}}{\alpha\gamma^{k+1}-1}\right)$ e $W_k \sim \text{Geométrica}(1 - [\alpha + \beta(1 - \gamma)]\gamma^k)$ são independentes;

(b) $N_{\alpha, \beta, \gamma} \stackrel{d}{=} \sum_{k=0}^{\infty} S_k$ se $\gamma \in [0, 1)$ e $\alpha > 0$, onde os S_k são variáveis aleatórias geométricas modificadas independentes, com função massa de probabilidade dada por:

$$\begin{cases} p_{0,k} = \frac{1 - [\alpha + \beta(1 - \gamma)]\gamma^k}{1 - \alpha\gamma^{k+1}} \\ p_{n,k} = p_{0,k} (\alpha + \beta)(1 - \gamma)\gamma^k ([\alpha + \beta(1 - \gamma)]\gamma^k)^{n-1}, \quad n \geq 1; \end{cases}$$

(c) $N_{\alpha, \beta, \gamma} \stackrel{d}{=} \sum_{k=0}^{\infty} T_k$ se $\gamma \in (-1, 0)$ e $\beta \geq 0$ ou $-\alpha\frac{1+\gamma}{1-\gamma^3} < \beta < 0$, onde os T_k são variáveis aleatórias independentes com função massa de probabilidade dada por:

$$\begin{cases} q_{0,k} = \frac{1 - [\alpha + \beta(1 - \gamma)]\gamma^{2k}}{1 - \alpha\gamma^{2k+1}} \frac{1 - [\alpha + \beta(1 - \gamma)]\gamma^{2k+1}}{1 - \alpha\gamma^{2k+2}} \\ q_{n,k} = q_{0,k} \frac{(\alpha + \beta)(1 - \gamma)}{\alpha + \beta(1 - \gamma)} ([\alpha + \beta(1 - \gamma)]\gamma^{2k})^{n-1} [\alpha(1 + \gamma) + \beta(1 - \gamma^{n+1})], \quad n \geq 1, \end{cases}$$

Demonstração: Quando $\gamma \in [0, 1)$ e $\alpha \leq 0$, tem-se $0 \leq \alpha + \beta(1 - \gamma) < 1$ em virtude do lema 2.2.2, e o termo geral do produtório

$$\mathcal{G}_{\alpha, \beta, \gamma}(s) = \prod_{k=0}^{\infty} \frac{1 - \alpha\gamma^{k+1}}{1 - \alpha\gamma^{k+1}} s \frac{1 - [\alpha + \beta(1 - \gamma)]\gamma^k}{1 - [\alpha + \beta(1 - \gamma)]\gamma^k s} \quad (2.13)$$

decompõe-se nas funções geradoras de $V_k \curvearrowright Bernoulli\left(\frac{\alpha\gamma^{k+1}}{\alpha\gamma^{k+1}-1}\right)$ e de $W_k \curvearrowright Geométrica(1 - [\alpha + \beta(1 - \gamma)]\gamma^k)$. Logo, cada termo do produtório representa uma variável aleatória $X_k = V_k + W_k$, com V_k e W_k independentes. Como o produtório converge para $|s| \leq 1$, o teorema da continuidade para funções geradoras de probabilidade (Feller, 1968, p. 280) permite concluir que é a função geradora de probabilidade da soma infinita de variáveis aleatórias independentes $\sum_{k=0}^{\infty} X_k = \sum_{k=0}^{\infty} (V_k + W_k)$.

Por outras palavras, temos uma soma infinita de variáveis aleatórias independentes, em que a parcela de ordem k é o resultado de adicionar aleatoriamente 1, com probabilidade $\frac{\alpha\gamma^{k+1}}{\alpha\gamma^{k+1}-1}$, a uma variável aleatória $Geométrica(1 - [\alpha + \beta(1 - \gamma)]\gamma^k)$.

Se $\gamma \in [0, 1)$ e $\alpha > 0$, os dois factores do termo geral do produtório deixam de ser funções geradoras de probabilidade, mas efectuando a divisão inteira vê-se que

$$\begin{aligned} \mathcal{P}_k(s) &= \frac{1 - \alpha\gamma^{k+1}s}{1 - \alpha\gamma^{k+1}} \frac{1 - [\alpha + \beta(1 - \gamma)]\gamma^k}{1 - [\alpha + \beta(1 - \gamma)]\gamma^k s} = \\ &= \frac{1 - [\alpha + \beta(1 - \gamma)]\gamma^k}{1 - \alpha\gamma^{k+1}} \frac{1}{\alpha + \beta(1 - \gamma)} \left(\alpha\gamma + \frac{(\alpha + \beta)(1 - \gamma)}{1 - [\alpha + \beta(1 - \gamma)]\gamma^k s} \right). \end{aligned}$$

Temos $0 < \alpha + \beta(1 - \gamma) < 1$, pelo que do desenvolvimento

$$\frac{1}{1 - [\alpha + \beta(1 - \gamma)]\gamma^k s} = \sum_{n \geq 0} ([\alpha + \beta(1 - \gamma)]\gamma^k)^n s^n \quad (2.14)$$

extraímos os coeficientes $\{p_{n,k}\}_{n \in \mathbb{N}}$ definidos no enunciado do teorema. Levando em conta o lema 2.2.2, conclui-se que $p_{0,k} = \mathcal{P}_k(0) = \frac{1 - [\alpha + \beta(1 - \gamma)]\gamma^k}{1 - \alpha\gamma^{k+1}} > 0$ e $p_{n,k} > 0$ para $n \geq 1$. Também é fácil ver que $\sum_{n \geq 0} p_{n,k} = \mathcal{P}_k(1) = 1$. Logo, os $\{p_{n,k}\}_{n \in \mathbb{N}}$, $k = 0, 1, \dots$, são funções massa de probabilidade de variáveis aleatórias discretas S_k , e a solução é uma soma infinita de variáveis aleatórias independentes, da forma $N_{\alpha, \beta, \gamma} \stackrel{d}{=} \sum_{k=0}^{\infty} S_k$. Como as parcelas S_k têm suporte \mathbb{N} , também a sua soma tem suporte \mathbb{N} .

Se $\gamma \in (-1, 0)$, os dois factores no termo geral do produtório (2.13) não são, de novo, funções geradoras de probabilidade. O lema 2.2.2

garante que $|[\alpha + \beta(1 - \gamma)]\gamma^k| < 1$, pelo que o desenvolvimento (2.14) continua a ser válido, mas agora γ^k toma valores alternadamente positivos e negativos. Neste caso associam-se os termos consecutivos dois a dois,

$$\mathcal{Q}_k(s) = \frac{1 - \alpha\gamma^{2k+1}s}{1 - \alpha\gamma^{2k+1}} \frac{1 - [\alpha + \beta(1 - \gamma)]\gamma^{2k}}{1 - [\alpha + \beta(1 - \gamma)]\gamma^{2k}s} \frac{1 - \alpha\gamma^{2k+2}s}{1 - \alpha\gamma^{2k+2}} \frac{1 - [\alpha + \beta(1 - \gamma)]\gamma^{2k+1}}{1 - [\alpha + \beta(1 - \gamma)]\gamma^{2k+1}s},$$

e considera-se a decomposição

$$\begin{aligned} & \left(\alpha\gamma + \frac{(\alpha + \beta)(1 - \gamma)}{1 - [\alpha + \beta(1 - \gamma)]\gamma^{2k}s} \right) \left(\alpha\gamma + \frac{(\alpha + \beta)(1 - \gamma)}{1 - [\alpha + \beta(1 - \gamma)]\gamma^{2k+1}s} \right) = \\ & = (\alpha\gamma)^2 + (\alpha + \beta)(1 - \gamma) \left(\frac{\alpha(1 + \gamma) + \beta}{1 - [\alpha + \beta(1 - \gamma)]\gamma^{2k}s} - \frac{\beta\gamma}{1 - [\alpha + \beta(1 - \gamma)]\gamma^{2k+1}s} \right). \end{aligned}$$

Expandindo em série as fracções simples, obtêm-se os coeficientes $\{q_{n,k}\}_{n \in \mathbb{N}}$. Do lema 2.2.2 resulta $q_{n,k} > 0$ para todo o $n \geq 0$, e facilmente se verifica que $\sum_{n \geq 0} q_{n,k} = \mathcal{Q}_k(1) = 1$.

□

Assim, as variáveis $N_{\alpha, \beta, \gamma}$, com $|\gamma| < 1$, são variáveis aleatórias geométricas (de suporte nos inteiros não negativos), ou somas de variáveis aleatórias independentes, com parcelas:

- geométricas e/ou de Bernoulli;
- geométricas modificadas em zero;
- geométricas modificadas em zero e nos inteiros pares,

desde que se verifique $\gamma \in [0, 1)$, ou $\beta \geq 0$, ou $\gamma \in (-1, 0)$ e $-\alpha \frac{1+\gamma}{1-\gamma^3} < \beta < 0$. No caso em que $\gamma \in (-1, 0)$ e $\beta \leq -\alpha \frac{1+\gamma}{1-\gamma^3} < 0$, não existe uma decomposição da função geradora de probabilidade que permita dar uma interpretação simples à variável $N_{\alpha, \beta, \gamma}$.

Capítulo 3

Modelos de Contagem

3.1 *Così Fan Tutte* — algumas reflexões sobre aleatoriedade discreta

Os modelos elementares de contagem — hipergeométrico, binomial, Poisson e binomial negativo — podem ser apresentados de uma forma muito simples. Prestam-se, por outro lado, a modelar muitos fenómenos reais. Hipergeométrica e binomial (e as suas extensões multivariadas) aparecem naturalmente no contexto de contagens em amostragem simples, sem e com reposição, como tradução dos conceitos de permutabilidade e de independência, respectivamente. O modelo de Poisson, abordado como limite de binomiais, ilustra as potencialidades de contagens capitalizadas na noção de estabilidade “em média”.

A modificação daqueles modelos por truncatura torna-os mais aptos a descrever contagens em que, por exemplo, o valor zero seja inobservável. Por outro lado, constatações simples como $\forall \theta \in (0, 1), \sum_{k=0}^{\infty} \theta^k = \frac{1}{1-\theta} \implies \sum_{k=0}^{\infty} \frac{\theta^{k+1}}{k+1} = -\ln(1-\theta)$ permitem definir novos modelos, no caso vertente a variável $X \sim \text{Logarítmica}(\theta)$

$$X = \begin{cases} k & k = 1, 2, \dots \\ -\frac{1}{\ln(1-\theta)} \frac{\theta^k}{k} \end{cases}$$

Situações concretas diversas, nomeadamente na área da Biologia, e

na área emergente de Geometria Fractal, têm levada à invenção de muitos outros modelos discretos, qual deles o mais complicado. Quando se percorre o livro de Johnson, Kotz e Kemp (2004), é fácil de perceber por que são evitados nos textos “normais” de Probabilidade. De facto, usando técnicas em si mesmo simples como composição e misturas, aqueles autores fazem uma panorâmica notável de modelos discretos — mas que em geral têm de ser apresentados através de funções geradoras de tratamento complexo.

A complexidade dos modelos pode torná-los apelativos, no sentido de que eventualmente se ajustam de forma mais perfeita a dados disponíveis, mas porventura o ensinamento de Saki “*a little inaccuracy saves tons of explanation*” reflecte melhor a sagesa com que se deve abordar a área da modelação.

O trabalho que se reproduz como parte integrante desta tese, “Count Data Models in Biometry and Randomness Patterns in Birds Extra-Pair Paternity”, nasceu da tentativa de corrigir alguns erros elementares que encontrámos em Neuhauser *et al.* (2001). Aqueles autores usaram dados sobre o número de crias “bastardas” achadas em ninhos de diversas espécies consideradas até há pouco monoândricas, mas que por análise de sangue foi possível comprovar serem descendentes de macho(s) diferente do parceiro regular da progenitora.

Count Data Models in Biometry and Randomness Patterns in Birds Extra-Pair Paternity

Tiago André Marques¹, Dinis Duarte Pestana^{2*}, Sílvia Filipe Velosa³

¹Universidade de Lisboa, CEAUL, Centro de Estatística e Aplicações da Universidade de Lisboa.
 Campo Grande, 1749-016 Lisboa, Portugal, +351-21-7500000 tiagomarques@fc.ul.pt

²Universidade de Lisboa, DEIO and CEAUL, Centro de Estatística e Aplicações da Universidade de
 Lisboa, Campo Grande, 1749-016 Lisboa, Portugal, +351-21-7500040 dinis.pestana@fc.ul.pt

³Universidade da Madeira, DMCE and CEAUL, Centro de Estatística e Aplicações da Universidade
 de Lisboa, Campus da Penteada, 9000-390 Funchal, Portugal, +351-291-705177
 svelosa@math.uma.pt

SUMMARY

The number of extra pair nestlings in a brood is the basic information to investigate extra-pair fertilization in socially monogamous birds, an interesting pattern of behaviour that has been observed in some species. Under unconstrained randomness, Poisson streams of events are expected. But other patterns of randomness may arise, suggesting new research questions. Starting from a coordinated approach to count models, we discuss Zipf-Mandelbrot self-organizing scaling laws, which are typical of phenomena shaped as a result of conflicting interests, and some extensions of Mandelbrot's model. While the traditional count models (Poisson, binomial, negative binomial or hypergeometric) seem inappropriate, the logarithmic, truncated logarithmic, Zipf-Mandelbrot and discrete lognormal models consistently provide the best fit to the available data, indicating that probably some females are more prone than others to have extra pair nestlings. This suggests a delicate balance: the number of extra pair nestlings in the progeny is the result of conflicting behaviours, the search for genetic diversity and the need to ensure male cooperation in raising the brood.

* Corresponding author: Dinis Pestana, fax +351-21-7500081

Key words: Count data, extra-pair fertilization, goodness-of-fit, Poisson related and Zipf – Mandelbrot related models, randomness.

This book concentrates, not on how to do an analysis, but on how to choose the right sort of analysis, and how to make sense of the answers. [...]

Statistical methods, apparently quite unrelated to each other, are in fact different aspects of the same central theory. [...] Statistical methods constitute a tool, often useful but sometimes abused.

[...] Biologists now use computers to proliferate statistical analyses, which have become an integral part of their work. But [...] sometimes an analysis is applied to data which contradict the assumptions of that analysis!

N. Gilbert, *Biometrical Interpretation* (Preface)

1. Introduction

The reason why some females of socially monogamous bird species seek extra-pair fertilizations is not clear, and it is to some extent important to assert whether the number of extra-pair young within and among broods follows some pattern of randomness which may support one of the possible explanations put forward by other researchers (Petrie and Kempenaers, 1998).

The Poisson model may be looked at as a yardstick, compared to which other models are overdispersed or underdispersed. Moreover, the fact that the differential entropy $I(p) = \int_{-\infty}^{\infty} f_x(x) \ln[f_x(x)] dx$ attains its maximum over all f_x which are positive only for $x \geq 0$ and have finite expectation for the exponential density $f_x(x) = \frac{1}{\delta} e^{-\frac{x}{\delta}} I_{(0,\infty)}$ may be interpreted as meaning that the exponential distribution is the most uncertain among all distributions of nonnegative random variables with finite expectation. Therefore, for all $t > 0$, the Poisson process has the greatest λt -dimensional entropy in the interval $(0, t)$ among all homogeneous point processes with the same given intensity $\lambda = \frac{1}{\delta} > 0$ (Rényi, 1964). For that reason, Poisson streams of events are generally interpreted as representing unconstrained randomness, but this does not preclude the data from exhibiting other patterns of randomness.

Recently, Neuhauser *et al.* (2001), using data about the number of extra-pair nestlings in broods of yellow warblers (*Dendroica petechia*), collected by Yezerinac *et al.* (1995), of hooded warblers (*Wilsonia citrina*), published by Stutchbury *et al.* (1994), and collared flycatchers (*Ficedula albicollis*), from Sheldon and Ellegren (1999), cf. columns 1–3 in Tables 1–3, challenged previous Poisson fits, and carried out a detailed investigation into the departure from a specific randomness hypothesis, namely an exchangeability hypothesis, expressed by a multivariate hypergeometric model, a form of mild dependence, with stationary probabilities. But in fact the

observed value of the G^2 goodness-of-fit statistic and the corresponding p -values show that their multivariate hypergeometric model is in general worse than the Poisson model. On the other hand, this complex multi-hypergeometric model and the cumbersome computations associated with it are irrelevant: as we shall explain in Section 4, to evaluate expected values only univariate hypergeometric margins are needed, and hence their exceedingly complex algorithm is unjustified.

The purpose of modelling is to achieve a generality that does not exist in actual observational results, in other words, to abstract from the data a general model that encompasses all possible sampling results, to extract knowledge from information. The built in dependence assumption that comes from sampling results rather than a sensible rationale is far from convincing, and the poor fit exhibited in Neuhäuser *et al.* (2001) ought to be expected, as will be further discussed in Section 4. But one of the merits of Neuhäuser *et al.* (2001) is to show that there is evident departure from the kind of randomness expressed by the hypergeometric model. We may therefore suspect that in those species where extra-mating has been studied, there are individual differences from female to female, as regards extra-mating strategies. We cannot conclude, however, whether this means different attitudes towards this kind of sexual behaviour, or whether different females have diverse extra-mating opportunities. We also ignore the distribution of successful extra-mating.

Models that are more closely related to the Poisson — even a crude Poisson model itself, with infinite support — are more appropriate than the hypergeometric model. The Poisson model may, as a side effect, incorporate the randomness derived from extra-mating opportunities, and on the other hand, as the Bernoulli filtered Poisson process is still Poisson, it may also incorporate extra-mating success.

The fact that both the Poisson and the multivariate hypergeometric provide exceedingly bad fit, in the generality of cases, does not mean that randomness has to be rejected, since other patterns of randomness may apply. We examine several alternative count models that stem from a basic Poissonian assumption in Section 2, and in Section 3 we investigate their goodness-of-fit to the number of extra-pair nestlings in broods of size s . Although much care must be taken in drawing conclusions based on such small samples, we shall observe that Zipf's law or the more general Zipf-Mandelbrot and discrete lognormal models account fairly well for the observed data.

The presentation of discrete models in Section 2 focuses on enhancing relations between different count models and how the unconstrained randomness of the Poisson model is progressively reduced by conditioning to find other well-known models such as the binomial and the hypergeometric, or by allowing for individual variability in negative binomial models. The logarithmic model, which can be derived as the limit of zero-truncated negative binomial distributions, may then be considerably generalized — in the sense that the parameter space is much wider — by right truncation, and the resulting Zipf or extended Zipf-Mandelbrot scaling laws exhibit an interesting "manicheist" pattern of randomness, that is appropriate to account for the equilibrium between conflicting tensions observed in many social

phenomena. A likely pattern for these females' behaviour, which seems to combine the eagerness for genetic diversity in the future with the present need to ensure male cooperation in raising the brood, which is supported by an excellent fit.

Most of the models we use are from the Katz family (Katz, 1965), whose probability mass functions exhibit nice recursive relations (Panjer, 1981). We also work out a discrete lognormal distribution that seems appropriate for skewed data; moreover, it generalizes Zipf–Mandelbrot laws for self-organizing phenomena that maintain a similar structure at various scales: scale has a bearing in structural organization and, on the other hand, the limit of discrete lognormal distributions when the location parameter goes to $-\infty$ is a Zipf–Mandelbrot distribution. The fit is in most cases almost as good as with the logarithmic or truncated-logarithmic distributions, and for very small samples it is even better. In the concluding Section 4, we comment on some points in the philosophy and practice of model fitting, stating some disagreements with Neuhäuser *et al.* (2001).

2. Count models and random patterns — a coordinate presentation of useful models

The classical presentation of count models is to use $X \sim \text{Hypergeometric}(N, n, p)$ for the number of successes in random sampling without replacement from a finite population (which implies a mild form of dependence, *exchangeability*), to use $Y \sim \text{Binomial}(n, p)$ as suitable for the number of successes in Bernoulli trials, i.e. when sampling with replacement (independent trials), and to adopt $W \sim \text{Poisson}(\lambda)$ as the limit of a sequence of binomial random variables under a “stable in average” condition, $E(X_n) = np_n \xrightarrow{n \rightarrow \infty} \lambda > 0$. Multivariate extensions with these margins are easily derived along similar paths.

This classical presentation has the advantage of immediate interpretation in terms of the most common sampling strategies. Moreover, when $n \ll N$ it is almost irrelevant to sample with or without replacement, and in that case $X \sim \text{Hypergeometric}(N, n, p)$ may be approximated by the simpler (two parameters instead of three) $Y \sim \text{Binomial}(n, p)$. On the other hand, for large values of n and $p \approx 0$, $W \sim \text{Poisson}(np)$ is a good approximation for $Y \sim \text{Binomial}(n, p)$, a further simplification, since we have to deal with only one parameter.

For some population studies, more sophisticated counting models are needed. In what follows, we present a coordinate description of the most important count models. Starting from the Poisson model, appropriate for unconstrained randomness, we obtain the binomial model by conditioning a Poisson summand on the observed sum, and the hypergeometric model by further conditioning a binomial summand on the observed sum, and comment on the constrained randomness brought in by

increased information (smaller variance). On the other hand, there may be room for individual variability, and mixing leads to negative binomial models.

At that stage, the concepts of underdispersion and overdispersion deserve some comments, as well as the recursive expressions used by Katz (1965) to organize discrete families of distributions, later explored by Panjer (1981) to obtain simple expressions for the density of randomly stopped sums. Binomial, Poisson and negative binomial are the non-degenerate solutions of Panjer's functional equation. If we relax the recursive relation, we obtain two other nondegenerate solutions, the Engen and the logarithmic models; the latter is appropriate when data show some tendency for clustering.

All these models can be further specialized by truncating the right tail. An important point is that the parameter space may be then considerably enlarged. The truncated logarithmic model, for instance, no longer has the restriction $\theta \in (0,1)$, and for $\theta = 1$ we have Zipf's model $P(X = k) \propto \frac{1}{k}$, $k = 1, \dots, N$, which accounts for "least effort", namely for equilibrium in accommodating conflicting needs.

Mandelbrot (1983) considerably generalized the useful part of Zipf's ideas by introducing location (λ) and shape (δ) parameters, $P(X = k) \propto \frac{1}{(k-\lambda)^{1+\delta}}$, $k = 1, \dots, N$.

Observe that a scale parameter is irrelevant in these "scaling laws", since it would be absorbed into the multiplicative normalizing constant. They are therefore useful for self-organizing phenomena, those that in some sense are scale invariant.

We introduce a non-trivial extension of Mandelbrot's class of models, by letting the shape $\theta(k)$ depend on k . The discrete lognormal random variable, with probability mass function $p_k \propto \frac{1}{k} \exp\left(-\frac{(\ln k)^2}{2}\right) = \frac{1}{k^{1+\ln k}}$, $k = 1, 2, \dots$, is an important model from that extended class. More generally, the discrete lognormal family, with probability mass function $p_k \propto \frac{1}{k} \exp\left(-\frac{1}{2}\left(\frac{\ln k - \mu}{\sigma}\right)^2\right)$, $k = 1, 2, \dots$, $\mu \in R$, $\sigma > 0$, seems a likely model for many discrete skewed data.

Alongside the description of models, we comment on parameter estimation and give appropriate references.

2a. Poisson Randomness and Count Data Models — Poisson and truncated Poisson, binomial and hypergeometric.

First order approximations are successfully used in all branches of mathematical modelling, and this is a natural explanation of the wide use of the Poisson model.

In fact, apart from stationarity and independence (in disjoint observation windows) of the increments, a third axiom rules out coincidental observations, and postulates local linearity, in the sense that the probability of a single observation is approximately proportional to the size of an infinitesimal observation window of area dA :

$$P(X_{A+dA} - X_A = 1) = \lambda dA + o(dA).$$

These three basic ideas are enough to establish that the number of occurrences in a region of size A is

$$X_A = \begin{cases} k & k = 0, 1, \dots \\ p_k = e^{-\lambda A} \frac{(\lambda A)^k}{k!} \end{cases}.$$

$E(X_A) = \lambda A$, so that the local linearity of the probability of a single occurrence in an infinitesimal window is mirrored by a global linearity for the expected values. In what follows we shall assume that the global area has unit size $A = 1$.

Poisson randomness is therefore appropriate whenever we feel that expected values may be considered “stable on average” (Recall that the Poisson arises as the limiting form of a sequence $X_n \sim \text{Binomial}(n, p_n)$ when the expected value $E(X_n) = n p_n \approx \lambda$). Examples are the number of sugar cane plants germinating per 10 m² plot, the number of nests per 100 m² plot in a pine wood and the number of butterflies caught per hour in a specific field. This is, indeed, the simpler mathematical translation of our faith in the regularity (and therefore predictability) of phenomena. It is natural to estimate λ by \bar{x} , which is in fact the maximum likelihood estimate of the rate λ .

There are, moreover, many other mathematical advantages of the Poisson model, for instance general Poisson random variables are the building blocks of infinitely divisible random variables, those that may be decomposed as sums of infinitesimal independent random summands. This is, of course, a strong modelling asset, since many observed phenomena are the result of infinitely many contributing effects. We shall pursue the matter no further here, but draw the reader’s attention to the guidance that statistical knowledge may provide on the choice of models, that on one hand they must be useful — i.e. mathematically tractable — and on the other hand they must be appropriate, in the sense that they have built in properties that reflect known properties of the phenomena at hand.

A crude Poisson model may be useful, even though it has infinite support, and the phenomenon we wish to model is clearly finite. In order to use a chi-square goodness-of-fit test, care must be taken that $\sum_k O_k = \sum_k e_k$, and a standard practice is to consider the N -th class as a composite aggregate class, corresponding to $k \geq N$.

Another possibility is to truncate the right tail of the Poisson model, considering that for physical reasons $k > N$ is unobservable. The truncated model has probabilities

$$p_k = \frac{P(X=k)}{P(X \leq N)} = \frac{1}{\sum_{j=1}^N \frac{\lambda^j}{j!}} \frac{\lambda^k}{k!}, \quad k=0,1,\dots,N.$$

In this case, in order to estimate λ (maximum likelihood) from the data, we must solve

$$\sum_{j=0}^N (\bar{x} - j) \frac{\lambda^j}{j!} = 0,$$

which is straightforward using Cohen's (1961) tables. Moore (1954) suggested the estimator

$$\tilde{\lambda} = \sum_j \frac{X_j}{m},$$

where m stands for the number of observed values that are less than $N-1$, which we shall use since it is easier to compute and is an unbiased estimator of λ .

The success of the Poisson model derives from striking "conservative" mathematical properties: if $X \sim \text{Poisson}(\lambda)$, $Y \sim \text{Poisson}(\mu)$ and X and Y are independent, then $X+Y \sim \text{Poisson}(\lambda+\mu)$, and thus incorporation of new information can be achieved, in many situations, through a simple modification of the parameter, the structure of the model not being modified. A similar conservative result holds for binomial filtering (or thinning): if $X \sim \text{Poisson}(\lambda)$ goes through a binomial filter with probability of success p , the resulting model is $X_j \sim \text{Poisson}(p\lambda)$.

From the result on the addition of independent Poisson random variables, by conditioning one of the summands on the observed value of the sum we obtain a binomial (more information yields a model with smaller variance): if $X_k \sim \text{Poisson}(\lambda_k)$, for $k=1,\dots,r$, are independent,

$$X_k \mid \sum_{j=1}^r X_j = s \sim \text{Binomial}(s, \pi_k), \quad \pi_k = \frac{\lambda_k}{\sum_{j=1}^r \lambda_j}.$$

More generally, if $\mathbf{X} = (X_1, \dots, X_n)$ is multi-Poisson, the conditional distribution of \mathbf{X} given that $X_1 + \dots + X_n = n$ is multinomial; this is the basis for the analysis of count data with general chi-square statistics comparing observed counts with the corresponding expected values, using either the classical chi-square statistic X_n^2 , the likelihood ratio statistic G^2 , or general power divergence statistics (Cressie and Read, 1984).

A similar result holds for independent binomial summands: if $X_k \sim \text{Binomial}(n_k, p)$, and $N = \sum_{j=1}^r n_j$, then $X_1 + \dots + X_n \sim \text{Binomial}(N, p)$ and $X_k | \sum_{j=1}^r X_j = s \sim \text{Hypergeometric}(N, s, \pi_k)$, where $\pi_k = \frac{n_k}{N}$. As this last relation may have

been one (unstated) reason for Neuhäuser *et al.*'s (2001) choice of hypergeometric randomness, we also investigate a binomial fit, since we feel more inclined towards an unconditional Poisson model than towards an unconditional binomial model. We use the maximum likelihood estimate of the binomial parameter p , the individual probability of success in the sequence of Bernoulli trials, which in this case coincides with the minimum chi-square and the method of moments estimates.

The above results show that the binomial and the hypergeometric (and their multivariate extensions), usually derived as count models associated with simple random sampling, with or without replacement, respectively, may also arise in connection with Poisson randomness. In the following sections, we shall show that other important discrete distributions are closely related to the Poisson law.

As we observed above, binomial filtering (or *thinning*) of the Poisson distribution results in a filtered Poisson distribution. For illustration purposes, consider the following situation. Let us assume that the number of times a female encounters another male (not her social mate) may be modelled as $X \sim \text{Poisson}(\lambda)$. Assume that the probability that copulation ensues from each encounter is α , and that the probability that a copulation will eventually lead to a (extra-pair) nestling is β , independently of what happens in any other occurrence. Thus the number of extra-pair nestlings is modelled by the filtered $X_f \sim \text{Poisson}(\lambda\alpha\beta)$.

Assume also that the number of times the female copulates with her social mate can be modelled as $Y \sim \text{Poisson}(\mu)$, and that the probability that a copulation will eventually lead to a nestling is π , so that the number of non extra-pair nestlings is $Y_f \sim \text{Poisson}(\mu\pi)$.

Accepting independence between X and Y (and, for honesty sake, this might not be true in the real world), then, if we draw a sample of broods and consider a specific brood size of s (that is, conditioning on the fact that $X_f + Y_f = s$), the model that describes the number of extra-pair nestlings will be $X_f \sim \text{Binomial}(s, \frac{\lambda\alpha\beta}{\lambda\alpha\beta + \mu\pi})$. Thus it seems worthwhile to investigate a binomial fit in this case study (with disappointing results, as we shall see).

A final observation on Poisson, binomial and hypergeometric models: we derived the binomial distribution by conditioning a Poisson summand on the observed value of the sum, and the hypergeometric law by conditioning a binomial summand on the observed value of the sum (assuming independence of summands, and equal parameter p in the latter case). Comparing the variances of $X \sim \text{Poisson}(\lambda)$, $Y \sim \text{Binomial}(n, p)$ and $W \sim \text{Hypergeometric}(N, n, p)$, with equal means, $np = \lambda$,

$$\text{var}(W) = np(1-p) \frac{N-n}{N-1} < \text{var}(Y) = np(1-p) < \text{var}(X) = \lambda = np.$$

The Poisson model is mathematically simpler — one single parameter, and the family of Poisson random variables is closed under summation and binomial filtering —, but on the other hand the hypergeometric model has more information.

As we have seen, the Poisson model accounts for unconstrained randomness. The binomial model, as an alternative to the usual presentation via counting successes in sampling with replacement, has been obtained as the posterior distribution of a Poisson summand conditioned on the value of the sum — extra information that diminishes randomness, bringing in more information, and hence a smaller variance. The hypergeometric model, aside from being the model for counting successes in sampling without replacement, has been shown to be the model for a binomial summand conditioned on the observed sum, and increased information once again decreases dispersion and further constrains randomness.

If the sampling fraction $\frac{n}{N}$ is low, i.e. $n \ll N$, $\text{var}(W) \approx \text{var}(Y)$: the probability of including the same element more than once in the sample, when sampling with replacement, is very low, and sampling with or without replacement is an idle question, since both sampling strategies give the same amount of information. On the other hand, if $p \approx 0$, $\text{var}(Y) \approx \text{var}(X)$, a clear indication that the Poisson law is a good approximation of the binomial when the variance is close to the expectation.

2b. Negative binomial as a gamma mixture of Poisson distributions; truncated negative binomial.

The geometric model — and, more generally, the negative binomial model of which the geometric is not but a special case — are interesting alternatives to the Poisson model, when we need to account for individual variability. In the case at hand, this is individual variability in the females' patterns of sexual behaviour.

Let us suppose that the appropriate model for each female is $X \sim \text{Poisson}(\lambda)$

$$X = \begin{cases} k & k = 0, 1, \dots \\ p_k = e^{-\lambda} \frac{\lambda^k}{k!} \end{cases}$$

($\text{var}(X) = E(X) = \lambda$). As regards the whole population, we may model λ as a random variable $\Lambda \sim \text{Exponential}(\delta)$, i. e. with distribution function

$$F_{\Lambda}(\lambda) = \left(1 - e^{-\frac{\lambda}{\delta}}\right) I_{(0,\infty)}(\lambda), \quad \delta > 0.$$

In this hierarchical model,

$$E(X_{ii}) = E[E(X | \Lambda)] = E(\Lambda) = \delta$$

and

$$\text{var}(X_{ii}) = E[\text{var}(X | \Lambda)] + \text{var}[E(X | \Lambda)] = E(\Lambda) + \text{var}(\Lambda) = \delta + \delta^2.$$

Thus we obtain a model with $\text{var}(X_{ii}) > \text{var}(X)$, which clearly accounts for higher diversity.

The mixture distribution can easily be derived: $f_{\Lambda}(\lambda) = \frac{1}{\delta} e^{-\frac{\lambda}{\delta}} I_{(0,\infty)}(\lambda)$, hence

$$\begin{aligned} P(X_{ii} = k) &= \int_0^{+\infty} e^{-\lambda} \frac{\lambda^k}{k!} \frac{1}{\delta} e^{-\frac{\lambda}{\delta}} d\lambda = \frac{1}{\delta k!} \int_0^{+\infty} \lambda^k e^{-(1+\frac{1}{\delta})\lambda} d\lambda = \\ &= \frac{1}{\delta k!} \int_0^{+\infty} \left(\frac{y}{1+\frac{1}{\delta}}\right)^k e^{-y} \frac{dy}{1+\frac{1}{\delta}} = \frac{1}{1+\delta} \left(\frac{\delta}{1+\delta}\right)^k, \quad k=0,1,\dots \end{aligned}$$

and thus $X_{ii} \sim \text{Geometric}\left(\frac{1}{1+\delta}\right)$. For that reason, some authors consider the geometric model (and more generally $Y \sim \text{NegativeBinomial}(\nu, \frac{1}{1+\delta})$ with index $\nu > 0$, i.e. $p_k = \binom{k+\nu-1}{\nu-1} \left(\frac{1}{1+\delta}\right)^{\nu} \left(\frac{\delta}{1+\delta}\right)^k$, $k=0,1,\dots$, which may be derived using similar arguments with the Gamma with index ν as mixing distribution) as a “more dispersed” Poisson.

In the general case, when both parameters are unknown, the easiest estimation method is the method of moments, equating sample and population means and variances, respectively. As we are working with the negative binomial located at 0, $E(Y) = \frac{\nu(1-p)}{p}$ and $\text{var}(Y) = \frac{\nu(1-p)}{p^2}$, therefore

$$\tilde{p} = \frac{\bar{x}}{s} \quad \text{and} \quad \tilde{\nu} = \frac{\bar{x}^2}{s^2 - \bar{x}}.$$

As an alternative, the *Mean-and-Zero-Frequency* method equates the observed and expected number of zero values, and the sample mean and population mean. Thus

$$f_0 = (p^*)^{\nu^*} \Rightarrow \nu^* = \frac{\ln f_0}{\ln p^*}$$

and

$$\bar{\nu} \cdot \frac{1-p^*}{p^*} = \bar{x},$$

from which we get $\frac{\ln f_0}{\ln p^*} \frac{1-p^*}{p^*} = \bar{x} \iff \frac{1-p^*}{p^* \ln p^*} = \frac{\bar{x}}{\ln f_0} \iff \frac{p^*}{\ln(1+p^*)} = -\frac{\bar{x}}{\ln f_0}$, where

$P^* = \frac{1-p^*}{p^*}$. Piegorsch (1990) recommends minimum chi-square estimation, on the grounds that it is slightly less biased than maximum likelihood or method of moments estimation, provided that the sample size is greater than 20. For a discussion of the relative merits of minimum chi-square and maximum likelihood criteria, cf. Berkson (1978) and the ensuing discussion.

Although the negative binomial distribution is infinitely divisible, and hence a non-integer index may be used, we shall limit our investigation to the classical case, and use an integer approximation for $\bar{\nu}$. Whenever $\bar{\nu}$ is too small, we use the shifted logarithmic model (Fisher *et al.*, 1943, derived the logarithmic model as the limit of zero truncated negative binomial distributions).

As in the Poisson case, the support of the negative binomial is infinite (as happens for all infinitely divisible laws), and we may derive finite support models using truncation. There are no closed forms for parameter estimation in the case of truncated models. For instance, for the geometric model truncated to the right of s ,

with $p_k = \frac{p(1-p)^k}{1-(1-p)^{s+1}}$, $k=0, \dots, s$, the maximum likelihood estimate \hat{p} satisfies the equation

$$\frac{n}{p} - \sum_{j=1}^n \frac{x_j}{1-p} - \frac{n(s+1)(1-p)^s}{1-(1-p)^{s+1}} = 0,$$

which, however, is easily handled in the example that we work out.

2c. Katz family of discrete random variables.

If $W \sim \text{Poisson}(\lambda)$, the recursive expression $p_{k+1} = \frac{\lambda}{k+1} p_k = \left(\alpha + \frac{\beta}{k+1}\right) p_k$ holds for $k=0, 1, \dots$, with $\alpha=0$ and $\beta=\lambda$.

If $Y \sim \text{Binomial}(n, p)$, we may write $p_{k+1} = \frac{n-k}{k+1} \frac{p}{1-p} p_k = \left(\alpha + \frac{\beta}{k+1}\right) p_k$, with $\alpha = \frac{p}{p-1}$ and $\beta = \frac{(n+1)p}{1-p}$, for $k=0, \dots, n-1$.

On the other hand, if $X \sim \text{NegativeBinomial}(\nu, p)$, i.e. $P(X=k) = \binom{k+\nu-1}{\nu-1} p^\nu (1-p)^k$, for $k=0,1,\dots$ we have $p_{k+1} = \frac{(\nu+k)(1-p)}{k+1} p_k = \left(\alpha + \frac{\beta}{k+1}\right) p_k$, with $\alpha = 1-p$ and $\beta = (\nu-1)(1-p)$.

It is interesting to observe that the models mentioned above are the only ones whose probability mass function satisfies the recursive relation

$$\frac{p_{n+1}}{p_n} = \alpha + \frac{\beta}{n+1} \quad n=0,1,\dots$$

(Panjer's recursive expression, which had been used by Katz (1965) in the equivalent form $\frac{p_{n+1}}{p_n} = \frac{\alpha + \beta}{n+1}$, $n=0,1,\dots$, to classify families of discrete distributions).

In fact, multiplying $p_{n+1} = \alpha + \frac{\beta}{n+1} p_n$ by s^{n+1} and adding for $n \geq 0$, we obtain the differential equation

$$\frac{P'(s)}{P(s)} = \frac{\alpha + \beta}{1 - \alpha s},$$

for the probability generating function $P_{X_{\alpha,\beta}}(s) = \sum_n p_n s^n$, whose absolutely monotone solutions are the probability generating functions

1. $P_{X_{0,0}}(s) = 1$, case $\alpha = \beta = 0$, i. e. $X_{0,0} = 0$, the degenerate random variable with unit mass at 0.
2. $P_{X_{0,\beta}}(s) = e^{\beta(s-1)}$ if $\alpha = 0$ (and necessarily $\beta > 0$), and therefore $X_{0,\beta} \sim \text{Poisson}(\beta)$.
3. $P_{X_{\alpha,\beta}}(s) = \left(\frac{1-\alpha}{1-\alpha s}\right)^{\frac{\alpha+\beta}{\alpha}}$; $X_{\alpha,\beta} \sim \text{NegativeBinomial}\left(\frac{\alpha+\beta}{\alpha}, 1-\alpha\right)$ if $\alpha \in (0,1)$ and $\alpha + \beta > 0$.
4. $P_{X_{\alpha,\beta}}(s) = \left((1-\frac{\alpha}{\alpha-1}) + \frac{\alpha}{\alpha-1}s\right)^{-\left(1+\frac{\beta}{\alpha}\right)}$; if $\alpha < 0$, $\frac{\alpha}{\alpha-1} = p > 0$ and we recognize the probability generating function of $X_{\alpha,\beta} \sim \text{Binomial}\left(-1-\frac{\beta}{\alpha}, \frac{\alpha}{\alpha-1}\right)$ with $-1-\frac{\beta}{\alpha} \in \mathbb{N}^+$.

Cf. Rolski *et al.* (1999) for an alternative proof.

An important point is that if $W \sim \text{Poisson}(\lambda)$, $\frac{\text{var}(W)}{E(W)} = \frac{\lambda}{\lambda} = 1$, a good reason to use the Poisson distribution as a yardstick in what regards dispersion.

On the other hand, if $Y \sim \text{Binomial}(n, p)$, $\frac{\text{var}(Y)}{E(Y)} = \frac{np(1-p)}{np} < 1$, Y is *underdispersed*.

We have seen that the geometric random variable, allowing for individual diversity, is more dispersed than the Poisson, and more generally if $X \sim \text{NegativeBinomial}(v, p)$, $\frac{\text{var}(X)}{E(X)} = \frac{1}{p} > 1$, X is *overdispersed*.

Panjer's family may be considerably extended by using weights,

$$\frac{p_{n+1}}{p_n} = \alpha + \frac{\beta}{(n+1)\gamma_n} \quad n = 0, 1, \dots$$

(Velosa, 2003), or if we relax the defining condition, by allowing

$$\frac{p_{n+1}}{p_n} = \alpha + \frac{\beta}{n+1} \quad n = k, k+1, \dots,$$

i.e. the relation holds only for $n \geq k > 0$ (Hess *et al.*, 2002). The case $k=1$ has been studied by Sundt and Jewell (1981) and by Willmot (1987). The results are worth recording:

If $p_0 = 0$ and $\frac{p_{n+1}}{p_n} = \alpha + \frac{\beta}{n+1}$, $n = 1, 2, \dots$, multiplying by s^{n+1} and summing for $n \geq 1$, we obtain the differential equation

$$(1 - \alpha s) P'(s) - (\alpha + \beta) P(s) = p_1,$$

whose non-degenerate absolutely monotone solutions are:

1. $P_{X_{\alpha, \beta}} = \frac{1 - (1 - \alpha s)^{\frac{\alpha + \beta}{\alpha}}}{1 - (1 - \alpha)^{\frac{\alpha + \beta}{\alpha}}}$, the probability generating function of
 - (a) the zero-truncated negative binomial distribution if $\alpha \in (0, 1)$ and $\beta > -\alpha$;
 - (b) the zero truncated Engen generalized negative binomial distribution, if $\alpha \in (0, 1]$ and $\beta \in (-2\alpha, -\alpha)$;
 - (c) *Logarithmic* (α) if $-\beta = \alpha \in (0, 1)$; the logarithmic random variable may be obtained either as the limit of zero truncated Engen's extended negative binomial random variables, or of zero truncated negative binomial random variables, with index (shape) parameter going to 0. Each of these models has a prestigious history of applications to population studies, since they were introduced by Fisher *et al.* (1943) and by Engen (1974).
2. $P_{X_{0, \beta}} = \frac{e^{-\beta}}{1 - e^{-\beta}} (e^{\beta s} - 1)$ (the zero truncated Poisson, if $\alpha = 0$ and $\beta > 0$);
3. $P_{X_{\alpha, \beta}} = \frac{(1 - \alpha s)^{\binom{1 + \beta}{\alpha} - 1}}{(1 - \alpha)^{\binom{1 + \beta}{\alpha} - 1}}$ (the zero truncated binomial, if $\alpha < 0$ and $-\frac{\beta}{\alpha} \in \mathbb{N}$).

More generally, the probability mass function of the random variable X satisfies $\frac{p_{n+1}}{p_n} = \alpha + \frac{\beta}{n+1}$, $n = 1, 2, \dots$, if the corresponding probability generating function $\tilde{P}_x(s)$ can be written

$$\tilde{P}_x(s) = \tau + (1 - \tau)P_{x_{\alpha, \beta}}(s),$$

where $P_{x_{\alpha, \beta}}$ is one of the probability generating functions in the preceding enumeration, and $\tau \in \left[\frac{P(0)}{P(0)-1}, 1 \right)$. Hence, the ratio of observed frequencies $\frac{o_{k+1}}{o_k} \approx \frac{p_{k+1}}{p_k}$ may be of some guidance on the choice of an appropriate model, by fitting $\frac{(k+1)o_{k+1}}{o_k} \approx \hat{\alpha}k + (\hat{\alpha} + \hat{\beta})$. Considerations on dispersion may give further insight on which model to prefer.

2d. Clustering and the logarithmic model; truncated logarithmic and Zipf-Mandelbrot models.

If none of the nestlings is extra-pair we cannot presume that the female did not mate with other males aside from her social partner. It may be that she has a very promiscuous behaviour, but extra-pair mating turns out to be infertile. On the other hand, the availability of extra-pair males may also have a bearing on the number of extra-pair nestlings. Promiscuous mating is certainly an opportunistic behaviour, perhaps a delicate equilibrium between compulsion to bring in genetic diversity to the progeny, and the need to maintain a monogamic social organization that favours the raising of nestlings.

The fact that we cannot exclude promiscuous mating behaviour when there are no extra-pair young in the nest is to some extent compensated by clustering when we do find "bastard" progeny. This is a common situation in biology. For instance, some female insects lay eggs on leaves appropriate for feeding their maggots. Of course the fact that we do not find any eggs on a particular leaf does not mean that none of these females landed on that particular leaf. On the other hand, we often find several eggs on the same leaf, and "blindness" in what regards 0 is compensated by some clustering. We now describe an appropriate model for this kind of phenomena: for any $\theta \in (0, 1)$, we have $\ln(1 - \theta) = -\sum_{k=1}^{\infty} \frac{\theta^k}{k}$. Therefore $p_k = -\frac{1}{\ln(1 - \theta)} \frac{\theta^k}{k}$, $k = 1, 2, \dots$, is the probability mass function of a random variable $W_\theta \sim \text{Logarithmic}(\theta)$, with support $k = 1, 2, \dots$, since it sums to 1:

$$W_\theta = \begin{cases} k & k = 1, 2, \dots \\ p_k = -\frac{1}{\ln(1-\theta)} \frac{\theta^k}{k} & (0 < \theta < 1) \end{cases}$$

As $\frac{\text{var}(W_\theta)}{E(W_\theta)} = \frac{1+\frac{\theta}{\ln(1-\theta)}}{1-\theta}$, for $\theta = 1 - \frac{1}{e}$ we have $E(W_\theta) = \text{var}(W_\theta)$ as in the Poisson model.

For $\theta \in (1 - \frac{1}{e}, 1)$ the random variable W_θ is overdispersed, and for $\theta \in (0, 1 - \frac{1}{e})$ we observe that W_θ is an interesting infinite support underdispersed model.

Fisher *et al.*'s (1943) derivation of the logarithmic random variable shows that it is the weak limit as $\nu \rightarrow 0$ of a sequence of zero-truncated negative binomial random variables with index ν , cf. Johnson *et al.* (1992, p. 286), and hence it is also related to the Poisson randomness model. They have shown that if in a batch the number of species represented by exactly one individual is n_1 then, denoting $\alpha = \frac{n_1}{\theta}$ the *index of diversity*, the number of species represented by k individuals is approximated by $\frac{\alpha \theta^k}{k}$, $k = 2, 3, \dots$ The logarithmic distribution is considered a good fit for count data whenever there is underlying clustering — number of bacteria per colony, number of inhabitants per house, or number of animals per litter, for instance.

The maximum likelihood estimate of θ is the solution $\hat{\theta}$ of

$$-\frac{\theta}{(1-\theta)\ln(1-\theta)} = \bar{x} = \frac{1}{n} \sum_{k=1}^n x_k,$$

which is easily computed using, for instance, the Newton–Raphson method.

By truncating to the right of s , we obtain the truncated logarithmic model

$$W^* = \begin{cases} k & k = 1, \dots, s \\ \frac{1}{\sum_{j=1}^s \frac{\theta^j}{j}} \frac{\theta^k}{k} & \end{cases}$$

The maximum likelihood estimator of θ is the solution of

$$\frac{\theta(1-\theta^s)}{(1-\theta) \sum_{j=1}^s \frac{\theta^j}{j}} = \bar{x},$$

which may be computed from Patil and Wani's (1965) tables. But for our purposes it is simpler to equate sample and population moments, obtaining the explicit estimate

$$\tilde{\theta} = \frac{m'_3 - (s+2)m'_2 + (s+1)m'_1}{m'_3 - sm'_2},$$

where $m'_k = \frac{1}{n} \sum_{j=1}^n x_j^k$ denotes the k -th sample moment.

We shall, naturally, shift these random variables to 0, $X = W - 1$ and $X^* = W^* - 1$; this is equivalent to taking $x_k + 1$ instead of x_k as observations.

An important remark: in this truncated case it is no longer necessary to consider that the parameter space is $\Theta = (0,1)$. In fact, $\{p_k\}_{k=1}^s$, $p_k = \frac{\theta^k}{k \sum_{j=1}^s \frac{\theta^j}{j}}$, is a probability

mass function for any $\theta > 0$, as in Table 1 ($s=4,5$), Table 2 ($s=2,3,5$) and Table 3 ($s=4,5$). In particular, if $\theta=1$ we get Zipf's law $\{p_k = \frac{c}{k}\}_{k=1}^s$, where

$\frac{1}{c} = \sum_{k=1}^s \frac{1}{k} \approx \ln s + \gamma$ ($\gamma \approx 0.577$ is Euler's constant), which Zipf claimed was tied to the

'principle of least effort' in the sense that it would model phenomena shaped by conflicting interests (Zipf's primary examples are verbal communication — the need to be socially understood constraining the use of rich personal vocabulary — and the size of cities, seen as the result of the attraction/repulsion feelings that large human settlements exert upon different individuals).

More general "scaling laws" or "power laws" $\left\{p_k = \frac{c}{(k-\lambda)^{1+\rho}}\right\}_{k=1}^s$, where

$\frac{1}{c} = \sum_{k=1}^s \frac{1}{(k-\lambda)^{1+\rho}}$, with location parameter λ and shape parameter $\rho > 0$, have been

used by Mandelbrot (1983), namely to model self-organizing phenomena. Gell-Mann (1994, p. 91–94) is an excellent naive introduction to the matter, with enlightening comments on the role of parameters. When $\rho > 0$, s may be ∞ ; in the simpler case

$\lambda = 0$, $\left\{p_k = \frac{1}{\zeta(1+\rho)} \frac{1}{k^{1+\rho}}\right\}_{k \geq 1}$, where $\zeta(\cdot)$ is Riemann's zeta function. Some authors

(Seal, 1952; Adamic, 2001) regard this *zeta distribution* as a *discrete Pareto* law. The maximum likelihood estimator $\hat{\rho}$ of the shape parameter ρ satisfies

$$\frac{\zeta'(\rho+1)}{\zeta(\rho+1)} = -\frac{1}{n} \sum_{k=1}^n \ln X_k$$

(Seal, 1952). Alternatively, equating the population and sample means we obtain the estimator $\tilde{\rho}$, that satisfies $\frac{\zeta(\rho)}{\zeta(\rho+1)} = \bar{X}$, which may be easily solved using the tables provided by Moore (1956).

Power laws play an important role in modelling dynamic phenomena where scale effects contribute to self-organization. With the sole exception of the hypergeometric distribution, all the other models related to Poisson randomness that we have described are generalized power laws, cf. Johnson *et al.* (1992, p. 81). We believe this is one more reason to prefer any of the other models (moreover, in the context of initial Poisson randomness, the hypergeometric arises from a double addition/conditioning scheme that seems farfetched in this context of modelling the number of extra-pair young in each nest). Pérez-Abreu (1991) has shown that under very mild assumptions the Poisson arises as the limit of power series distributions, further justifying for the central role we have chosen for the Poisson model:

2e. Discrete lognormal model.

Discrete probability mass functions of the form $\left\{ p_k = \frac{1}{C(\theta)} \frac{1}{k^{1+\theta(k)}} \right\}_{k \geq 1}$ provide an even wider choice of models, among which Zipf–Mandelbrot's, corresponding to $\theta(k) = \rho$, is just the easiest to deal with. The generalized case, on the other hand, may provide appropriate models where location may influence scale effects.

A particular choice $\theta(k) = \ln(\sqrt{k})$ leads to $p_k \propto \frac{1}{k^{1+\ln \sqrt{k}}}$, $k = 1, 2, \dots$, i.e. to the *discrete lognormal* random variable

$$X_{0,1} = \begin{cases} k & k = 1, 2, \dots \\ p_k = \frac{1}{C(0,1)k} \exp\left(-\frac{(\ln k)^2}{2}\right) \end{cases}$$

where $C(0,1) = \sum_{k=1}^{\infty} \frac{1}{k} \exp\left(-\frac{(\ln k)^2}{2}\right)$ is the normalizing constant.

More generally, the discrete lognormal law with parameters μ and σ

$$X_{\mu,\sigma} = \begin{cases} k & k = 1, 2, \dots \\ p_k = \frac{1}{C(\mu,\sigma)k} \exp\left[-\frac{1}{2}\left(\frac{\ln k - \mu}{\sigma}\right)^2\right] \end{cases}$$

with $C(\mu, \sigma)$ the appropriate normalizing constant, seems a likely model for discrete skew populational data. The parameters μ and σ may be estimated by numerical maximization of the likelihood function.

Further observe that when $\mu \rightarrow -\infty$, the discrete lognormal law may be considered a Zipf-Mandelbrot law with shape parameter $\rho = \frac{\mu}{\sigma^2}$. In fact, if $\ln k \ll |\mu|$,

$$p_k = \frac{1}{C(\mu, \sigma)k} \exp \left[-\frac{1}{2} \left(\frac{\ln k - \mu}{\sigma} \right)^2 \right] \\ \propto \frac{1}{k} \exp \left[-\frac{\ln k (\ln k - 2\mu)}{2\sigma^2} \right] \approx \frac{1}{C(\mu, \sigma)} \frac{1}{k^{\frac{1-\mu}{\sigma^2}}}.$$

3. Goodness-of-fit

For the data provided in Tables 1–6 — Tables 1 and 2 for yellow warblers (*Dendroica petechia*) studied by Yezerinac *et al.* (1995), Tables 3 and 4 for hooded warblers (*Wilsonia citrina*) provided by Stutchbury *et al.* (1994), Tables 5 and 6 for collared flycatchers (*Ficedula albicollis*) studied by Sheldon and Ellegren (1999) —, and using G^2 as a comparison criterion, we evaluate the goodness-of-fit for the relevant models described above: the more traditional hypergeometric, Poisson, truncated Poisson, binomial, negative binomial (in general geometric), truncated geometric, in the odd-numbered tables; and the less common logarithmic, truncated logarithmic, truncated Zipf-Mandelbrot and discrete lognormal models in the even-numbered tables. We present asymptotic theory, since this way our results may be checked easily using a spreadsheet, since they do not depart substantially from exact results using exact algorithms such as those incorporated in StaXact.

For comparison purposes, our Tables 1 to 6 exhibit some overlap — i.e., for the hypergeometric and Poisson models — with Tables 2–4 in Neuhäuser *et al.* (2001); our goal is to show that the simple univariate hypergeometric model achieves exactly the same results as their multi-hypergeometric model. We omitted brood size $s=1$ (hooded warblers) and $s=2$ (collared flycatchers), and in fact should have refrained from analyzing other cases: statistical analysis with exceedingly small sample sizes cannot be recommended. But this is a case study to assess the potential interest of these models, not a statistical data analysis aimed at drawing conclusions from the data. Brood size $s=4$ (collared flycatchers), and $s=5$ (hooded warblers) are maintained in our analysis, as paradigmatic cases, important for the discussion in Section 4 and comparison with Neuhäuser *et al.*'s (2001) views.

For the Poisson, logarithmic and geometric models, which have infinite support, we give results both using the last class as an “aggregate” class, and using the

corresponding truncated models. We use truncated negative binomial models only for estimated index $\nu = 1$ (geometric), since parameter estimation for the general truncated negative binomial is unreliable. Anyway, only in the case of yellow warblers, $s = 4$ (Table 1) do we have the unexpected index estimate $\nu^* \approx 10$; in all the other cases (we would have $\nu^* \approx 5$ for hooded warblers, $s = 2$ in Table 2; but for this case, since we need to estimate two parameters, the number of degrees of freedom would be 0), the estimate is either $\nu^* \approx 0$ — in which case the logarithmic model seems appropriate — or $\nu^* \approx 1$.

As we have seen in Section 2, the logarithmic random variable is the limit of zero-truncated negative binomial random variables with index going to zero. Hence, whenever the index estimate of the negative binomial model is near zero we proceed directly to a logarithmic model fit.

Parameter estimation has been done either by maximum likelihood estimation or the simpler — as a rule, unbiased — estimation described in the appropriate subsection of Section 2; we emphasize simple methods, and this is the reason why we described them beforehand, providing a fair choice to the reader. Computer intensive methods, bootstrap estimates, and other sophisticated methods may be preferable in many situations — but in all cases the researcher must ponder whether the sample size calls for such investment. In addition to the G^2 observed value, degrees of freedom and p -values, we present parameter estimates (though the index of the negative binomial may be fractional — $\frac{pe^{\theta}}{1-(1-p)e^{\theta}}$ is an infinitely divisible characteristic

function, and thus $\left(\frac{pe^{\theta}}{1-(1-p)e^{\theta}}\right)^{\nu}$ is a characteristic function for any $\nu > 0$ — we shall use the integer part of ν^* as the binomial index), and expected frequencies, for immediate visual comparison with the observed values.

There is no need, in the hypergeometric case, for explicit evaluation of the parameter estimate $\hat{p} = \frac{r}{N}$; but this does not preclude the fact that it is implicitly used in the computation of expected values under the null hypothesis, and that the appropriate number of degrees of freedom is $(s+1) - 2 = s - 1$.

As regards the hypergeometric, the computation is done with the univariate hypergeometric marginal model, which provides exactly the same expected values, with much less computational effort than in Neuhauser *et al.* (2001). In fact, only expected values of the univariate model are needed, and it seems farfetched to compute univariate moments using the complicated algorithm for the multi-hypergeometric model given in the appendix to Neuhauser *et al.* (2001); this matter will be further discussed in Section 4.

The hypergeometric and the binomial model give poor fits, even worse than the crude Poisson model criticized by Neuhauser *et al.* (2001), or the negative binomial that could account for some individual variability. Thus the classical count models seem useless in the present context.

On the other hand, the logarithmic and truncated logarithmic (power laws, in the strict sense) provide in general excellent fits, the more general Zipf–Mandelbrot scaling laws and the skewed discrete lognormal distribution seem interesting candidates to elicit the randomness exhibited by these data. This clear pattern, exhibited by the comparison of Tables 1 and 2, Tables 3 and 4, and Tables 5 and 6, has been confirmed using StaXact exact algorithms to compute observed values and p -values.

It therefore seems plausible to consider the working hypothesis that eagerness for genetic diversity and the need to have male cooperation in raising the brood are two polarities that generate an elaborate equilibrium in these females mating behaviour.

4. Concluding Remarks

The purpose of the present paper, aside from contributing to the understanding of some female birds' remarkable sexual behaviour, is twofold:

1. To stress that the discrete models presented in elementary and intermediate Probability and Statistics courses (hypergeometric, binomial and negative binomial, and Poisson) are insufficient to model the rich randomness patterns that arise when counting biostatistical phenomena. Johnson *et al.*'s (1992) presentation, on the other hand, which in many cases uses Gauss hypergeometric generating functions, is too specialized for most practical users. Our choice would be to invest in a coordinate approach to count data models as outlined in Section 2, stressing patterns of randomness. One point we would like to emphasize once again: truncation (and indeed other techniques aimed at restricting the support) may be compensated by an interesting stretching of the parameter space (in our examples, although the parameter space for the *Logarithmic*(θ) is $\Theta = (0, 1)$, for the right truncated logarithmic — power laws — the parameter space $\Theta = (0, \infty)$ is much wider).
2. To underline the importance of clearly understanding statistical concepts, in order to avoid misconceptions and mishandling of statistical algorithms. This is worrying mostly because errors tend to have a fertile progeny following their publication. For that reason, we make explicit some points detailed in Marques and Pestana (2002):
 - Neuhäuser *et al.* (2001) claim that the Poisson is not a plausible model, since it attributes positive probabilities to unobservable values, and since the percentage of extra-pair nestlings is high, this not a rare event. The first objection — pointing out that impossible configurations have positive probabilities in a model — is a statistical misconception (truncation of the right tail solves part of the problem). Sharing their views about modelling would prevent us from using almost all useful (i.e., mathematically tractable) statistical models, most of which have infinite support, while all

our observations are finite. "*A little inaccuracy saves tons of explanations*", wrote Saki; in our view, this is the role of models: a good model is general — since there is no Science of particular facts — free from concrete details, although its properties reflect the striking patterns of the phenomena that it represents; and this inbuilt simplification makes it mathematically tractable. Regarding rare events, notice that the same model can be generated by very diverse mechanisms, and that the point is whether or not it provides a good fit to the available data. The fact that the Poisson distribution can be derived as the limit of binomial distributions under a condition of stability in the mean that has an appealing interpretation as a limit law of rare events cannot be taken as a crude statement that it applies only to rare events; and it should be noted that "rarity" concerns the binomial parent distribution, not the limiting Poisson law: $np_n \xrightarrow{n \rightarrow \infty} \lambda \in (0, \infty)$, and hence the mean value $E(X) = \lambda$ can be very large!

Although the question addressed in Neuhäuser *et al.* (2001), as they clearly stated, was not "*to model observed distributions, but to provide a pattern of randomness with which observed patterns can be compared*", their multi-hypergeometric distribution seems quite unrealistic. It comes as no surprise that Neuhäuser *et al.* (2001) obtained very low p -values using G^2 goodness-of-fit statistics for their hypergeometric fitting. In fact they were, as a rule, lower than using the Poisson that they criticize, as can be seen in their Tables 2–4 and in Tables 1–3 below.

It should be noted that there is a systematic error in their computations: in order to use the chi-square goodness-of-fit statistic care must be taken that, under the validity of H_0 , $\sum_{k=1}^N p_k = 1$. This is an important point, namely in

deriving the asymptotic chi-square statistic — the classes considered must be a partition of the support of the distribution, $\sum_{k=0}^N e_k = \sum_{k=0}^N o_k$ is the reason why

the number of degrees of freedom in the asymptotic chi-square is always $N-1$ (minus the number of estimated parameters, if any). This accounts for the discrepant values between their Tables 2–4 and our Tables 1–3. The relative error in their Table 4, $s=4$, is quite high, 12%. We point this out because it is a frequent error that needs to be eradicated, a mishandling that can be misleading.

The computation of expected frequencies is highly simplified by using a marginal hypergeometric model, instead of their complex algorithm for computing expected values using the multivariate distribution.

The reader can verify that all calculations in Tables 2, 3 and 4 in Neuhäuser *et al.* (2001) are greatly simplified with $e_k = k p_k$, the p_k being the **univariate** hypergeometric probabilities

$$p_k = \frac{\binom{r}{k} \binom{n-r}{s-k}}{\binom{n}{s}} = \frac{\binom{s}{k} \binom{n-s}{r-k}}{\binom{n}{r}},$$

where s is the brood size, n is the number of nests with brood size s , $n = s n_s$ is the total number of young in broods of size s , and r the total number of extra-pair young (we use the notations s , n and r of Neuhäuser *et al.* (2001) for an easy comparison).

In fact, the appropriate model for the number of individuals from each of r classes C_1, \dots, C_r in a sample of size n taken randomly, without replacement, from a population of size $N = K_1 + \dots + K_r$ such that K_j individuals in the population belong to class C_j , $j = 1, \dots, r$, is

$\mathbf{X} = (X_1, \dots, X_r) \sim \text{Multihypergeometric}(N; n; p_1, \dots, p_r)$, where $p_j = \frac{K_j}{N}$:

$$X = (X_1, \dots, X_r) = \begin{cases} (k_1, \dots, k_r) & k_1 + \dots + k_r = n \\ P_{(k_1, \dots, k_r)} = \frac{\binom{K_1}{k_1} \dots \binom{K_r}{k_r}}{\binom{N}{n}} \end{cases}$$

But it is obvious that any univariate marginal random variable X_j counting the number of elements classified in class C_j has *Hypergeometric*(N, n, p_j) distribution: we need only to collect all the other classes in a residual $\bar{C}_j = \bigcup_{\substack{k=1 \\ (k \neq j)}}^r C_k$. Therefore we get

$$P(X_j = x_j) = \frac{\binom{K_j}{x_j} \binom{N-K_j}{n-x_j}}{\binom{N}{n}}, \text{ and } E_{(X_1, \dots, X_N)}(X_j^r) = E_{X_k}(X_j^r).$$

Thus a much easier univariate hypergeometric model achieves the same goals as the multi-hypergeometric put forward by Neuhäuser *et al.* (2001).

Finally, we shall comment on the most important issue, i.e. what can be grasped from quantitative data as regards the understanding of mating behaviour. As has been established by Neuhäuser *et al.* (2001), the multivariate hypergeometric model and the Poisson law must be rejected, since the associated p -values are always

exceedingly small. Our results confirm this, and further establish that there is departure from the kind of randomness defined by the classical (Poisson, binomial, negative binomial and hypergeometric) count models, which always overestimate frequencies for the intermediate cases.

But other patterns of randomness may be an excellent fit with interesting behavioral interpretations — they are indeed an excellent fit, although as a general remark we would like to point out that it is easier, of course, to obtain a better fit with a model depending on more parameters, and that Akaike's criterion ought to be used to evaluate the relative benefit from using a more complicated model.

The observed frequencies seem indicate that the most common situation is either zero or at most one extra-pair, or, more seldom, many extra-pairs. This seems to indicate that only a small proportion of these females have a very promiscuous behaviour (or successful promiscuous behaviour), hypothesis that deserves further investigation.

Also, we have to bear in mind that we are modelling the number of extra-pair young as a step in trying to ascertain why females seek extra-pair fertilization. We are therefore looking in the present for questions which arose in the past, at an evolutionary level, which are far from straightforward. On the other hand, even random search by some females of extra pair fertilizations might lead to more elaborate random models for the number of extra-pair nestlings if other self-control or self-organizing mechanisms intervene in the process (such as regulation of the probability of success of extra pair fertilizations, or the probability that a female finds an available male, based on territorial constraints and nearest neighbour models). More sophisticated models, such as Zipf's equilibrium, therefore seem appropriate for seeking a much better explanation.

Nevertheless, the importance of modelling the observed number of extra pair nestlings should not be overlooked, as close fits to specific models may lead to interesting working hypotheses. The different models presented in this work are some of the possible alternatives when dealing with count data. For the available data, which is very scarce, there is clear indication that the logarithmic or the truncated logarithmic model often provide the best fit. This agrees with the explanation that different females have different behaviour, with some of them displaying more "eagerness" for promiscuous mating than others. Power laws, Zipf's law and Pareto models are closely related (Adamic, 2001, Mandelbrot, 1983), and seem to account for sophisticated forms of equilibrium reached by natural populations. The more general Zipf-Mandelbrot scaling models, and the discrete lognormal model whose structure changes with scale, provide good fits and deserve further investigation with larger samples. Scale regulation seems, indeed, to be the governing force behind most dynamic self-organizing phenomena.

Acknowledgements: This work has been partially supported by FCT/POCTI/FEDER, VEXTRA Project. T. A. Marques' research has been sponsored by FCT, SFRH/BM/2068/2000, FSE, III Quadro Comunitário de Apoio.

REFERENCES

- Adamic L.A. (2001). Zipf, Power-Laws and Pareto — a ranking tutorial.
<http://ginger.hpl.hp.com/shl/papers/ranking/>
- Berkson J. (1980). Minimum chi-square, not maximum likelihood! (with discussion), *Annals of Statistics*, 457–487.
- Cohen A. C. (1961). Estimating the Poisson parameter from samples that are truncated to the right, *Technometrics*, 433–438.
- Cressie N., Read T. R. C. (1984). Multinomial Goodness-of-fit tests, *Journal of the Royal Statistical Society B* 46, 440–464.
- Engen S. (1974). On species frequency models, *Biometrika*, 263–270.
- Fisher R. A., Corbet A. S., Williams C. B. (1943). The relation between the number of species and the number of individuals in a random sample of an animal population, *Journal of Animal Ecology*, 42–58.
- Gell-Mann M. (1994). *The Quark and the Jaguar*, Freeman, New York.
- Gilbert, N. (1989). *Biometrical Interpretation. Making Sense of Statistics in Biology*, Oxford University Press, Oxford.
- Hess K. Th., Lewald A., Schmidt K. D. (2002). An extension of Panjer's recursion, *ASTIN Bulletin* 32, 283–297.
- Johnson N. L., Kotz S., Kemp A. W. (1992). *Univariate Discrete Distributions*, Wiley, New York.
- Katz L. (1965). Unified treatment of a broad class of discrete probability distributions, *Classical and Contagious Discrete Distributions*, G. P. Patil (ed.), Pergamon Press, Oxford, 175–182.
- Mandelbrot B. (1983). *The Fractal Geometry of Nature*, Freeman, New York.
- Marques T. A., Pestana D. D. (2002). *Extra-Pair Fertilizations and the Distribution of Extra-Pair Young Within and Among Broods: A Case Study on Count Data Models in Biometry*, CEAUL, Lisboa.
- Moore P. G. (1954). A note on truncated Poisson distribution, *Biometrics* 10, 402–406.
- Moore P. G. (1956). The geometric, logarithmic and discrete Pareto forms of series, *Journal of the Institute of Actuaries* 82, 130–136.
- Neuhäuser M., Forstmeier W., Bretz F. (2001). The distribution of extra-pair young within and among broods — a technique to calculate deviations from randomness, *Journal of Avian Biology* 32, 358–363.
- Panjer H. H. (1981). Recursive evaluation of a family of compound distributions, *ASTIN Bulletin*, 22–26.
- Patil G. P., Wani J. K. (1965). Maximum likelihood estimation for the complete and truncated logarithmic series distribution, in G. P. Patil (ed.): *Classical and Contagious Discrete Distributions*, 398–409, Pergamon Press, Oxford.
- Pérez-Abreu V. (1991). Poisson approximation to power series distributions, *American Statistician* 45, 42–45.
- Petrie M., Kempeneers B. (1998). Extra-pair paternity in birds: explaining variation between species and populations, *Trends Ecol. Evol.* 13, 52–58.
- Piegorsch W.W. (1990). Maximum likelihood estimation for the negative binomial dispersion parameter, *Biometrics* 46, 863–867.

- Rényi A. (1964). On an extremal property of the Poisson process, *Annals of the Institute of Statistical Mathematics* 16, 129–133.
- Rolski T., Schmidli H., Schmidt V., Teugels J. (1999). *Stochastic Processes for Insurance and Finance*, Wiley, New York.
- Seal H. L. (1952). The maximum likelihood fitting of the discrete Pareto law, *Bulletin of the Assoc. Institute of Actuaries* 78, 115–121.
- Sheldon B. C., Ellegren H. (1999). Sexual selection resulting from extra-pair paternity in collared flycatchers, *Animal Behavior* 57, 285–298.
- Stutchbury B. J., Rhymer J. M., Morton E. S. (1994). Extrapair paternity in hooded warblers, *Behavioral Ecology* 5, 384–392.
- Sundt B. Jewell W.S. (1981). Further results on recursive evaluation of compound distributions. *Astin Bulletin* 12, 27–39.
- Velosa S. F. (2003). *Discrete infinitely divisible laws*, CEAUL, Lisboa.
- Velosa S. F., (2003). New classes of discrete infinitely divisible laws in *Literacy and Statistics*, P. Brito, A. Figueiredo, F. Sousa, P. Teles e F. Rosado, eds., Sociedade Portuguesa de Estatística, Porto, 673–684.
- Willmot G. E. (1987). Sundt and Jewell's family of discrete distributions, *ASTIN Bulletin* 18, 17–29.
- Yezerinac S. M., Weatherhead P. J., Boag P. T. (1995). Extra-pair paternity and the opportunity for sexual selection in a socially monogamous bird (*Dendroica petechia*), *Behavioral Ecology and Sociobiology* 37, 179–188.
- Zar J. H. (1999). *Biostatistical Analysis*, 4th ed., Prentice Hall, Upper Saddle River.

Table 1. Extra pair nestlings in broods of size s , yellow warblers (*Dendroica petechia*), data from Yezerinac *et al.* (1995) - traditional count models. Observed and expected frequencies, G^2 observed value, degrees of freedom (d.f.), and corresponding p -value. For the untruncated models with infinite support, the last residual class represents $P(X \geq s)$.

	k	o_k	Hyperg.	Pois.	Tr.Pois.	Bin.	Neg.Bin.	Tr.Neg.Bin.
<i>est.par.</i>				0.429	0.500	0.143	1; 0.750	0.680
$s=3$	0	10	8.707	9.120	8.506	8.816	10.500	9.618
	1	2	4.610	3.909	4.253	4.408	2.625	3.080
	2	2	0.659	0.838	1.063	0.735	0.656	0.986
	3	0	0.024	0.134	0.177	0.041	0.219	0.316
G^2			3.872	2.643	2.745	3.364	2.394	1.880
d.f.			2	2	2	2	2	1
p -value			0.144	0.267	0.254	0.186	0.122	0.170
<i>est.par.</i>				1.730	2.462	0.432	10; 0.850	
$s=4$	0	10	3.720	6.561	3.522	3.839	7.247	
	1	7	11.757	11.349	8.668	11.701	10.902	
	2	9	11.350	9.816	10.669	13.373	9.021	
	3	5	6.748	5.659	8.754	6.793	5.428	
	4	6	1.225	3.614	5.387	1.294	4.403	
G^2			21.220	4.944	10.511	20.170	3.090	
d.f.			3	3	3	3	2	
p -value			0	0.176	0.015	0	0.213	
<i>est. par.</i>			~	1.538	1.935	0.308	1; 0.458	0.299
$s=5$	0	17	6.060	8.374	5.712	6.202	17.855	13.216
	1	8	13.878	12.883	11.055	13.783	9.681	9.271
	2	3	12.406	9.910	10.698	12.252	5.249	6.504
	3	3	5.410	5.082	6.902	5.445	2.846	4.563
	4	3	1.151	1.955	3.340	10.210	1.543	3.201
	5	5	0.095	0.797	1.293	0.108	1.827	2.246
G^2			59.537	27.051	32.162	57.398	6.299	6.659
d.f.			4	4	4	4	3	3
p -value			0	0	0	0	0.098	0.084

Table 2. Extra pair nestlings in broods of size s , yellow warblers (*Dendroica petechia*), data from Yezerinac *et al.* (1995) - power, generalized Zipf-Mandelbrot and discrete lognormal models.

Observed and expected frequencies, G^2 observed value, degrees of freedom (d.f.), and corresponding p -value. Maximum p -value, indicating the best fit, in boldface. For the untruncated models with infinite support, the last residual class represents $P(X \geq s)$.

	k	o_k	$Log.$	$Tr.Log.$	$Mand.$	$D.Logn.$
<i>est. par.</i>			0.491	0.563	-3.32; 5.02	0.07; 0.62
$s=3$	0	10	10.175	9.782	9.795	9.788
	1	2	2.500	2.751	2.801	2.972
	2	2	0.819	1.032	0.993	0.829
	3	0	0.505	0.435	0.410	0.411
G^2			2.330	1.813	1.664	2.387
d.f.			2	2	1	1
p -value			0.312	0.404	0.197	0.122
<i>est. par.</i>			0.827	1.251	-4.57; 0.00	0.86; 0.59
$s=4$	0	10	17.424	11.829	9.697	8.519
	1	7	7.209	7.401	8.211	11.840
	2	9	3.977	6.174	7.135	7.570
	3	5	2.468	5.795	6.303	4.139
	4	6	5.922	5.801	5.644	4.933
G^2			10.402	1.573	0.970	2.916
d.f.			3	3	2	2
p -value			0.015	0.665	0.616	0.233
<i>est. par.</i>			0.807	1.140	0.10; 0.00	0.55; 0.84
$s=5$	0	17	19.148	12.866	16.589	15.829
	1	8	7.722	7.332	7.858	9.665
	2	3	4.152	5.571	5.148	5.282
	3	3	2.512	4.763	3.828	2.987
	4	3	1.621	4.343	3.047	1.771
	5	5	3.846	4.125	2.530	3.467
G^2			1.956	4.085	3.499	2.890
d.f.			4	4	3	3
p -value			0.744	0.395	0.321	0.409

Table 3. Extra pair nestlings in broods of size s , hooded warblers (*Wilsonia citrina*), data from Stutchbury et al. (1994) - traditional count models. Observed and expected frequencies, G^2 observed value, degrees of freedom (d.f.), and corresponding p -value. For the untruncated models with infinite support, the last residual class represents $P(X \geq s)$.

	k	o_k	Hyperg.	Pois.	Tr.Pois.	Bin.	Neg.Bin.	Tr.Neg.Bin.
<i>est. par.</i>				0.692	1.200	0.346		
$s=2$	0	15	11	13.011	8.904	11.115		
	1	4	12	9.008	10.685	11.769		
	2	7	3	3.982	6.411	3.115		
G^2			12.378	5.673	9.016	11.692		
d.f.			1	1	1	1		
p -value			0	0.017	0.003	0.001		
<i>est. par.</i>				0.841	1.194	0.280	1; 0.566	0.437
$s=3$	0	28	16.255	18.978	13.797	16.402	24.904	21.388
	1	3	19.402	15.959	16.467	19.165	10.808	12.033
	2	5	7.430	6.710	9.827	7.464	4.691	6.770
	3	8	0.913	2.353	3.910	0.969	3.598	3.809
G^2			50.026	28.388	34.119	48.589	12.300	15.594
d.f.			2	2	2	2	1	1
p -value			0	0	0	0	0	0
<i>est. par.</i>				0.944	1.172	0.236	1; 0.473	0.455
$s=4$	0	23	12.098	14.000	11.225	12.258	17.027	17.217
	1	2	15.377	13.222	13.161	15.155	8.974	9.378
	2	4	7.048	6.244	7.715	7.027	4.729	5.108
	3	4	1.379	1.966	3.015	1.448	2.492	2.782
	4	3	0.097	0.568	0.884	0.112	2.777	1.515
G^2			45.957	27.390	29.800	44.203	10.734	12.187
d.f.			3	3	3	3	2	2
p -value			0	0	0	0	0.005	0
<i>est. par.</i>				1.667	2.500	0.333	$(\hat{p} \approx 0)$	
$s=5$	0	2	0.252	0.567	0.257	0.395		
	1	0	1.049	0.944	0.643	0.988		
	2	0	1.199	0.787	0.803	0.988		
	3	0	0.450	0.437	0.669	0.494		
	4	0	0.050	0.182	0.418	0.123		
	5	1	0.001	0.083	0.209	0.012		
G^2			22.107	10.032	11.335	15.276		
d.f.			4	4	4	4		
p -value			0	0.040	0.023	0.004		

Table 4. Extra pair nestlings in broods of size s , hooded warblers (*Wilsonia citrina*), data from Stutchbury *et al.* (1994) - power, generalized Zipf-Mandelbrot and discrete lognormal models.

Observed and expected frequencies, G^2 observed value, degrees of freedom (d.f.), and corresponding p -value. Maximum p -value, indicating the best fit, in boldface. For the untruncated models with infinite support, the last residual class represents $P(X \geq s)$.

	k	o_k	<i>Log.</i>	<i>Tr.Log.</i>	<i>Mand.</i>	<i>D.Logn.</i>
<i>est. par.</i>			0.622	1.261		
$s=2$	0	15	16.616	12.035		
	1	4	5.171	7.587		
	2	7	4.213	6.378		
G^2			1.985	2.789		
d.f.			1	1		
p -value			0.159	0.095		
<i>est. par.</i>			0.673	1.137	0.54; 0.00	-0.03; 0.91
$s=3$	0	28	26.487	18.585	26.911	25.924
	1	3	8.915	10.568	8.479	9.457
	2	5	4.001	8.012	5.032	4.007
	3	8	4.597	6.834	3.578	4.612
G^2			7.668	13.200	9.051	7.310
d.f.			2	2	1	1
p -value			0.022	0.001	0.003	0.007
<i>est. par.</i>			0.702	0.968	0.63; 0.00	-0.24; 1.05
$s=4$	0	23	20.882	16.368	22.213	21.124
	1	2	7.327	7.924	5.999	7.304
	2	4	3.428	5.115	3.468	3.207
	3	4	1.804	3.715	2.439	1.634
	4	3	2.560	2.877	1.881	2.731
G^2			7.809	9.014	4.441	7.669
d.f.			3	3	2	2
p -value			0.050	0.029	0.109	0.022
<i>est. par.</i>			0.821	3.000	0.69; 0.00	-2.47; 2.01
$s=5$	0	2	1.432	0.044	1.881	1.667
	1	0	0.588	0.065	0.445	0.518
	2	0	0.322	0.131	0.252	0.247
	3	0	0.198	0.294	0.176	0.142
	4	0	0.130	0.705	0.135	0.912
	5	1	0.331	1.762	0.110	0.335
G^2			3.549	14.179	8.232	2.381
d.f.			4	4	3	3
p -value			0.470	0.007	0.041	0.497

Table 5. Extra pair nestlings in broods of size s , collared flycatchers (*Ficedula albicollis*), data from Sheldon and Ellegren (1999) - traditional count models. Observed and expected frequencies, G^2 observed value, degrees of freedom (d.f.), and corresponding p -value. For the untruncated models with infinite support, the last residual class represents $P(X \geq s)$.

	k	o_k	Hyperg.	Pois.	Tr.Pois.	Bin.	Neg.Bin.	Tr.Neg.Bin.
<i>est.par.</i>				2.500	5.000	0.625		
$s=4$	0	0		0.164	0.031	0.040		
	1	0	0.143	0.410	0.153	0.264		
	2	1	0.857	0.513	0.382	0.659		
	3	1	0.857	0.428	0.637	0.732		
	4	0		0.485	0.797	0.305		
G^2			0.617	3.034	2.823	1.456		
d.f.			3	3	3	3		
p -value			0.893	0.386	0.420	0.692		
<i>est.par.</i>				1.529	2.000	0.306	1; 0.393	0.301
$s=5$	0	9	2.595	3.683	2.339	2.739	6.684	5.795
	1	2	6.133	5.633	4.679	6.035	4.056	4.050
	2	0	5.476	4.308	4.679	5.319	2.461	2.830
	3	2	2.306	2.196	3.119	2.344	1.493	1.978
	4	2	0.457	0.840	1.560	0.516	0.906	1.382
	5	2	0.034	0.340	0.624	0.046	1.399	0.966
G^2			39.528	22.126	24.729	36.906	8.291	9.538
d.f.			4	4	4	4	3	3
p -value			0	0	0	0	0.04	0.02
<i>est.par.</i>				0.804	0.822	0.134	1; 0.425	0.546
$s=6$	0	31	19.230	20.579	20.215	19.395	19.533	25.204
	1	3	18.244	16.553	16.622	18.015	11.239	11.45
	2	6	6.987	6.657	6.833	6.973	6.466	5.201
	3	4	1.382	1.785	1.873	1.439	3.721	2.363
	4	1	0.149	0.359	0.385	0.167	2.141	1.073
	5	0	0.008	0.058	0.063	0.010	1.232	0.488
	6	1	0	0.009	0.009	0	0.669	0.221
G^2			46.464	31.893	32.149	44.73	17.846	13.596
d.f.			5	5	5	5	4	4
p -value			0	0	0	0	0.001	0.009
<i>est.par.</i>				0.250	0.250	0.036	$(\hat{v} \approx 0)$	
$s=7$	0	11	9.212	9.346	9.346	9.303		
	1	0	2.579	2.336	2.336	2.412		

	2	0	0.204	0.292	0.292	0.268
	3	1	0.004	0.024	0.024	0.017
	4	0		0.002	0.002	0.001
	5	0		0	0	0
	6	0		0	0	0
	7	0		0	0	0
G^2			14.75	11.017	11.017	11.89
d.f.			6	6	6	6
p -value			0.022	0.088	0.088	0.064

Table 6. Extra pair nestlings in broods of size s , collared flycatchers (*Ficedula albicollis*), data from Sheldon and Ellegren (1999) - power, generalized Zipf-Mandelbrot and discrete lognormal models.

Observed and expected frequencies, G^2 observed value, degrees of freedom (d.f.), and corresponding p -value. Maximum p -value, indicating the best fit, in boldface. For the untruncated models with infinite support, the last residual class represents $P(X \geq s)$.

	k	o_k	Log.	Tr.Log.	Mand.	D.Logn.
<i>est. par.</i>			0.882	1.235	-500; 0.00	1.25; 0.08
$s=4$	0	0	0.825	0.653	0.402	0
	1	0	0.364	0.403	0.401	0
	2	1	0.214	0.332	0.900	0.974
	3	1	0.142	0.308	0.904	1.026
	4	0	0.455	0.304	0.398	0.001
G^2			6.993	4.562	3.005	0.0015
d.f.			3	3	2	2
p -value			0.072	0.207	0.223	0.999
<i>est. par.</i>			0.806	1.215	0.40; 0.00	0.19; 1.07
$s=5$	0	9	8.364	4.966	8.458	8.005
	1	2	3.368	3.016	3.172	3.641
	2	0	1.809	2.442	1.952	1.890
	3	2	1.093	2.225	1.410	1.088
	4	2	0.704	2.162	1.103	0.675
	5	2	1.662	2.188	0.906	1.701
G^2			6.568	7.961	4.716	6.173
d.f.			4	4	3	3
p -value			0.160	0.093	0.194	0.103
<i>est. par.</i>			0.662	0.724	0.77; 0.00	-0.58; 1.11
$s=6$	0	31	28.088	26.419	30.736	29.231

	1	3	9.291	9.569	5.747	8.678
	2	6	4.098	4.622	3.170	3.360
	3	4	2.033	2.511	2.189	1.744
	4	1	1.076	1.455	1.671	0.958
	5	0	0.593	0.879	1.352	0.570
	6	1	0.820	0.546	1.135	1.259
G^2			9.572	10.274	6.978	9.036
d.f.			5	5	4	4
<i>p</i> -value			0.088	0.068	0.13?	0.060
<i>est. par.</i>			0.350	0.426	0.96; 0.01	-29.99; 3.61
<i>s</i> =7	0	11	9.749	9.214	10.935	10.47
	1	0	1.707	1.960	0.407	1.043
	2	0	0.398	0.556	0.206	0.266
	3	1	0.105	0.177	0.138	0.100
	4	0	0.029	0.060	0.103	0.047
	5	0	0.009	0.021	0.083	0.025
	6	0	0.003	0.008	0.069	0.015
	7	0	0.001	0.003	0.059	0.035
G^2			7.172	7.357	6.323	9.548
d.f.			6	6	5	5
<i>p</i> -value			0.305	0.289	0.276	0.089

3.2 Comentários e Complementos

A Lei de Zipf e as suas extensões hiperbólicas generalizadas discretas têm vindo a ganhar grande relevo em áreas muito diversas, nomeadamente em Ecologia, em Ciências Sociais e em Linguística. A sua ligação às variáveis de Paretos, que tem como case limite a ligação da Geométrica à Exponencial, confere-lhe uma importância acrescida, podendo conjecturar-se que o seu estudo mais aprofundado contribua para a teoria das somas e dos máximos de variáveis discretas.

Por outro lado, seria deveras interessante conseguir estabelecer resultados sobre processos pontuais de que o processo de Poisson fosse um caso extremo, em que os “tempos de espera” tivessem distribuição de Pareto, e procurar as associações estruturais entre as leis desses processos pontuais e as leis hiperbólicas.

Procurámos também estabelecer resultados sobre somas de hiperbólicas generalizadas, no intuito de depois determinar a lei de uma parcela condicional na soma (por outras palavras, estender os conhecidos resultados sobre a lei de uma parcela Poisson condicional ao valor da soma de Poissons independentes), mas as expressões combinatórias que se obtêm para a lei da soma tornam os resultados inusáveis do ponto de vista analítico.

Como se estabelece no artigo acima transcrito, ao restringir o suporte da lei logarítmica a uma secção inicial finita de \mathbb{N} passamos à logarítmica truncada à direita, que admite o espaço de parâmetros $\Theta' = (0, \infty)$, muito mais vasto que o espaço de parâmetros $\Theta = (0, 1)$ da logarítmica usual com suporte $S = \{1, 2, \dots\}$. Assim, as distribuição-potência (“*power laws*”) surgem numa iluminação diversa do habitual, que possibilitará porventura uma abordagem mais profícua ao problema.

Os modelos naturais de contagem foram recentemente expostos por Hess *et al.* (2002). Com a devida vénia ao trabalho de Abamahomed (2006), apresentamos uma demonstração completa dos resultados, que a publicação em revista não permite.

A classe de Hess *et al.* (2002) em certo sentido usa os resultados de Sundt and Jewell (1981) e de Wilmott, que tinham generalizado Panjer (1981) usando truncatura em 0 e modificação em 0, que são casos especiais da truncatura em k e da modificação em k de Hess *et al.* (2002). Nesse paradigma mais geral, as *basic claim number distributions*, ou distribuições de Panjer⁽¹⁾ são as enumeradas no quadro seguinte.

⁽¹⁾Ou distribuições de Katz, como as intitulámos no artigo, por ter sido Katz o

distribuição básica	k	dist. truncada em k
$Binomial(m, \vartheta)$	$\{0, 1, \dots, m-1\}$	$Binomial(m, \vartheta; k)$
$Poisson(\alpha)$	$\{0, 1, \dots\}$	$Poisson(\alpha; k)$
$BNegativa(\beta, \vartheta)$	$\{0, 1, \dots\}$	$BNegativa(\beta, \vartheta; k)$
$Logarítmica(\vartheta)$	$\{1, 2, \dots\}$	$Logarítmica(\vartheta; k)$
$ENB(m, \beta, \vartheta)$	$\{m, m+1, \dots\}$	$ENB(m, \beta, \vartheta; k)$
$ELOG(m, \vartheta)$	$\{m, m+1, \dots\}$	$ELOG(m, \vartheta; k)$

onde $BNegativa$ designa a binomial negativa, ENB é a binomial negativa generalizada, e $ELOG$ a logarítmica generalizada, com suporte $\{m, m+1, \dots\}$, $m \in \mathbb{N}_0$, e função massa de probabilidade $\{q_n\}$ proporcional a $\vartheta^n \binom{n}{m}^{-1}$ para $n \geq m$.

Definição (3.2.1): Para uma claim number distribution $Q = \{q_n\}$, com $n \in \mathbb{N}_0$ e $k \in \mathbb{N}_0$, tal que, $q_k > 0$ e $\sum_{n=k+1}^{\infty} q_n > 0$, define-se $Q^{(k)} = \{q_n^{(k)}\}$, com $n \in \mathbb{N}_0$, onde, $q_n^{(k)} = 0$ para $n \leq k-1$ e $q_n^{(k)} = \frac{q_n}{1 - \sum_{j=0}^{k-1} q_j}$ para $n \geq k$.

Então, $Q^{(k)}$ é uma claim number distribution não degenerada, pois

$$q_n^{(k)} = \frac{q_n}{1 - \sum_{j=0}^{k-1} q_j} = \frac{q_n}{\sum_{j=k}^{\infty} q_j} = \frac{q_n}{q_k + \sum_{j=k+1}^{\infty} q_j};$$

$$0 \leq \frac{q_n}{q_k + \sum_{j=k+1}^{\infty} q_j} < 1, \text{ para } n \geq k$$

uma vez que $q_k > 0$ e $\sum_{n=k+1}^{\infty} q_n > 0$.

Além disso, $\sum_{n=k}^{\infty} q_n^{(k)} = \sum_{n=k}^{\infty} \frac{q_n}{1 - \sum_{j=0}^{k-1} q_j} = \frac{\sum_{n=k}^{\infty} q_n}{\sum_{j=k}^{\infty} q_j} = 1$.

A sua função geradora de probabilidades é

primeiro a classificá-las sistematicamente com base na fórmula recursiva verificada pelas suas funções massa de probabilidade.

$$\begin{aligned}
m_{Q^{(k)}}(t) &= \sum_{n=0}^{\infty} q_n^{(k)} t^n = \sum_{n=k}^{\infty} \frac{q_n}{1 - \sum_{j=0}^{k-1} q_j} t^n = \frac{1}{1 - \sum_{j=0}^{k-1} q_j} \sum_{n=k}^{\infty} q_n t^n = \\
&= \frac{1}{1 - \sum_{j=0}^{k-1} q_j} (\sum_{n=0}^{\infty} q_n t^n - \sum_{j=0}^{k-1} q_j t^j) = \frac{m_Q(t) - \sum_{j=0}^{k-1} q_j t^j}{1 - \sum_{j=0}^{k-1} q_j}
\end{aligned}$$

A distribuição $Q^{(k)}$ chama-se k -truncatura de Q . Em particular, temos

$$\begin{aligned}
Binomial(m, \vartheta) &= Binomial(m, \vartheta; 0) \\
Poisson(\alpha) &= Poisson(\alpha; 0) \\
BNegativa(\beta, \vartheta) &= BNegativa(\beta, \vartheta; 0) \\
Logarítmica(\vartheta) &= Logarítmica(\vartheta; 1) \\
ENB(m, \beta, \vartheta) &= ENB(m, \beta, \vartheta; m) \\
ELOG(m, \vartheta) &= ELOG(m, \vartheta; m)
\end{aligned}$$

As extensões de Hess *et al.* resultam de exigir que a recursão de Panjer

$$q_{n+1} = \left(a + \frac{b}{n+1} \right) q_n \quad (3.1)$$

se verifique apenas a partir de certo $k \geq 0$, sendo $p_n = 0$ para $n < k$. As soluções de (3.1), que se designam por distribuições de Panjer de ordem k e se representam por $Panjer(a, b; k)$, são as truncaturas em k de *basic claim number distributions*, quando existam. É fácil ver que a truncatura em k de cada distribuição básica de Panjer verifica a recursão. Para provar o recíproco é necessário o seguinte resultado.

Lema (3.2.2): *Se $Q = Panjer(a, b; k)$, então $(k+1)a + b > 0$. Mais, $a + b \geq 0 \Rightarrow a < 1$ e $a + b < 0 \Rightarrow a \leq 1$.*

Demonstração:

$$q_{k+1} = \left(a + \frac{b}{k+1} \right) q_k \iff q_{k+1} = \frac{(k+1)a + b}{k+1} q_k$$

Uma vez que $q_k > 0$ e $k+1 > 0$ terá de ser $(k+1)a + b > 0$ para que se tenha $q_{k+1} > 0$.

Sejam $a > 0$ e $a + b \geq 0$; então para $n \geq k$,

$$\begin{aligned}
q_{n+1} = \frac{na+a+b}{n+1}q_n &\geq \frac{na}{n+1}q_n \geq \frac{na}{n+1} \frac{(n-1)a}{n}q_{n-1} \geq \\
&\geq \frac{na}{n+1} \frac{(n-1)a}{n} \frac{(n-2)a}{n-1}q_{n-2} \geq \frac{n-2}{n+1}a^3q_{n-2}
\end{aligned}$$

e genericamente,

$$q_{n+1} \geq \frac{n-(j-1)}{n+1}a^jq_{n-(j-1)};$$

fazendo $j = n - k$, resulta

$$q_{n+1} \geq \frac{k+1}{n+1}a^{n-k}q_{k+1} \quad (3.2)$$

A série $\sum_{n=k}^{\infty} \frac{k+1}{n+1}a^{n-k}$ é divergente para $a \geq 1$, pois $\lim_{n \rightarrow \infty} \frac{k+1}{n+1}a^{n-k} = \infty$, para $a > 1$. Para $a = 1$, a série $\sum_{n=k}^{\infty} \frac{k+1}{n+1}$ é divergente, pois, aplicando-lhe o critério de Raabe vem:

$$\lim_{n \rightarrow \infty} n \left(\frac{k+1}{n+1} \times \frac{n+2}{k+1} - 1 \right) = \lim_{n \rightarrow \infty} \frac{n}{n+1} = 1^-$$

Mas de (3.2) deduz-se que $\sum_{n=k}^{\infty} \frac{k+1}{n+1}a^{n-k}q_{k+1} \leq \sum_{n=k}^{\infty} q_{n+1} \leq 1$. Assim, $q_{k+1} \sum_{n=k}^{\infty} \frac{k+1}{n+1}a^{n-k} \leq 1$ que implica, usando o critério da comparação, que a série $\sum_{n=k}^{\infty} \frac{k+1}{n+1}a^{n-k}$ é convergente. Logo, não pode ser $a \geq 1$; e portanto $a < 1$.

Sejam agora $a > 0$ e $a + b < 0$. Então, para $n \geq k$,

$$q_{n+1} = \frac{(n-k)a + (k+1)a + b}{n+1}q_n > \frac{n-k}{n+1}aq_n,$$

pois $(k+1)a + b > 0$. Procedendo como anteriormente,

$$q_{n+1} \geq \frac{a^3}{\frac{n+1}{n-k} \frac{n}{n-k-1} \frac{n-1}{n-k-2}}q_{n-2}; \text{ genericamente,}$$

$$q_{n+1} \geq \frac{a^j}{\frac{(n+1)n \dots (n-(j-2))}{(n-k)(n-k-1) \dots (n-k-(j-1))}}q_{n-(j-1)}; \text{ fazendo } j = n - k,$$

$$q_{n+1} \geq \frac{a^{n-k}}{\frac{(n+1)n \dots (k+2)(k+1)!}{(n-k)(n-k-1) \dots 1(k+1)!}}q_{k+1} = \frac{a^{n-k}}{\binom{n+1}{k+1}}q_{k+1}$$

A série $\sum_{n=k}^{\infty} \frac{a^{n-k}}{\binom{n+1}{k+1}}$ é divergente para $a > 1$, pois $\lim_{n \rightarrow \infty} \frac{a^{n-k}}{\binom{n+1}{k+1}} = \infty$ nesse caso. Para $a = 1$, a série $\sum_{n=k}^{\infty} \frac{1}{\binom{n+1}{k+1}}$ é convergente, pois, aplicando o critério de Raabe,

$$\lim_{n \rightarrow \infty} n \left(\frac{\binom{n+1}{k+1}^{-1}}{\binom{n+2}{k+1}^{-1}} - 1 \right) = \lim_{n \rightarrow \infty} n \left(\frac{n+2}{n-k+1} - 1 \right) = \lim_{n \rightarrow \infty} \frac{n(k+1)}{n-k+1} = k+1$$

Mas da desigualdade anteriormente estabelecida resulta $q_{k+1} \sum_{n=k}^{\infty} \frac{a^{n-k}}{\binom{n+1}{k+1}} \leq \sum_{n=k}^{\infty} q_{n+1} \leq 1$; conclui-se, assim, usando o critério da comparação, que a série $\sum_{n=k}^{\infty} \frac{a^{n-k}}{\binom{n+1}{k+1}}$ é convergente. Portanto, não pode ser $a > 1$; logo $a \leq 1$. □

Teorema (3.2.3): *Seja Q uma claim number distribution não degenerada. Para $k \in \mathbb{N}_0$, as seguintes proposições são equivalentes:*

- (a) Q é uma distribuição de Panjer de ordem k .
- (b) Q é a k -truncatura de uma claim number distribution básica.

Demonstração:

Mostremos que (a) \implies (b). Seja $Q = \{q_n\}_{n \in \mathbb{N}_0} = \text{Panjer}(a, b; k)$. Pelo teorema 2.1 de Hess *et al.* temos

$$\frac{m_Q^{(k+1)}(t)}{m_Q^{(k)}(t)} = \frac{d}{dt} (\ln m_Q^{(k)})(t) = \frac{(k+1)a + b}{1 - at},$$

para $t \in [0, 1)$, e $m_Q^{(n)}(0) = 0$ quando $n \leq k-1$.

Para resolver a equação diferencial anterior vão-se considerar 3 casos:

- caso $a < 0$:

$$q_{n+1} = \frac{(n-k)a + (k+1)a + b}{n+1} q_n \quad \text{para } n \geq k;$$

pelo lema 3.2.3, $(k+1)a + b > 0$; $q_k > 0$ até existir um valor m tal que $q_m > 0$ e $q_{m+1} = 0 \iff m = -1 - \frac{b}{a}$.

$$\frac{d}{dt}(\ln m_Q^{(k)})(t) = \frac{(k+1)a+b}{1-at} \iff \ln m_Q^{(k)}(t) = \frac{(k+1)a+b}{-a} \ln(1-at) + \ln c$$

$$= \ln c(1-at)^{-(k+1+\frac{b}{a})},$$

onde c é uma constante. $m_Q^{(k)}(t) = c(1-at)^{m-k}$, onde $m = -1 - \frac{b}{a}$; a solução geral da equação diferencial é portanto da forma

$$m_Q(t) = \sum_{j=0}^{k-1} c_j t^j + c_k (1-at)^m,$$

com os c_j constantes. $m_Q(1) = 1$ e as condições iniciais $m_Q^{(i)}(0) = 0$ para $i \leq k-1$, implicam:

$$m_Q(1) = 1 \iff \sum_{j=0}^{k-1} c_j + c_k (1-a)^m = 1 \quad (3.3)$$

$$m'_Q(0) = 0 \iff \left[\sum_{j=0}^{k-1} j c_j t^{j-1} - a c_k m (1-at)^{m-1} \right]_{t=0} = 0 \iff$$

$$\iff c_1 - a c_k m = 0 \iff c_1 = \frac{a m c_k}{1!}$$

$$m''_Q(0) = 0 \iff \left[\sum_{j=0}^{k-1} j(j-1) c_j t^{j-2} + a^2 c_k m(m-1) (1-at)^{m-2} \right]_{t=0} = 0 \iff$$

$$\iff c_k + a^2 m(m-1) c_k = 0 \iff c_2 = \frac{-(-1)^2 a^2 m(m-1) c_k}{2!}$$

\vdots

$$c_j = \frac{-(-1)^j a^j m(m-1) \cdots (m-(j-1))}{j!} c_k \iff c_j = -(-a)^j \binom{m}{j} c_k \quad (3.4)$$

Substituindo (3.4) em (3.3), resulta

$$\sum_{j=0}^{k-1} -(-a)^j \binom{m}{j} c_k + c_k (1-a)^m = 1 \iff c_k = \frac{1}{(1-a)^m - \sum_{j=0}^{k-1} (-a)^j \binom{m}{j}}$$

Substituindo (3.4) e esta expressão de c_k na solução geral,

$$\begin{aligned} m_Q(t) = \sum_{j=0}^{k-1} c_j t^j + c_k (1-at)^m &= -c_k \sum_{j=0}^{k-1} (-a)^j \binom{m}{j} t^j + \frac{(1-at)^m}{(1-a)^m - \sum_{j=0}^{k-1} (-a)^j \binom{m}{j}} \\ &= \frac{-\sum_{j=0}^{k-1} (-a)^j \binom{m}{j} t^j}{(1-a)^m - \sum_{j=0}^{k-1} (-a)^j \binom{m}{j}} + \frac{(1-at)^m}{(1-a)^m - \sum_{j=0}^{k-1} (-a)^j \binom{m}{j}} \end{aligned}$$

Dividindo ambos os termos de ambas as frações por $(1-a)^m$ e atendendo a que $(\frac{1}{1-a})^m = (\frac{1}{1-a})^{m-j} (\frac{1}{1-a})^j$, vem

$$\begin{aligned} &\frac{-\sum_{j=0}^{k-1} \binom{m}{j} (\frac{1}{1-a})^{m-j} (\frac{-a}{1-a})^j}{1 - \sum_{j=0}^{k-1} (-a)^j \binom{m}{j} (\frac{1}{1-a})^m} + \frac{(\frac{1-at}{1-a})^m}{1 - \sum_{j=0}^{k-1} (-a)^j \binom{m}{j} (\frac{1}{1-a})^m} \\ &= \frac{(\frac{1}{1-a} - \frac{a}{1-a} t)^m - \sum_{j=0}^{k-1} \binom{m}{j} (\frac{1}{1-a})^{m-j} (\frac{-a}{1-a})^j}{1 - \sum_{j=0}^{k-1} \binom{m}{j} (\frac{1}{1-a})^{m-j} (\frac{-a}{1-a})^j} \end{aligned}$$

Esta última expressão é a da função geradora de probabilidades de uma *Binomial* $(\frac{a+b}{-a}, \frac{-a}{1-a}; k)$.

- caso $a = 0$:

Pelo lema 3.2.3, $(k+1)a + b > 0 \iff b > 0$; partindo de

$$\frac{d}{dt}(\ln m_Q^{(k)})(t) = \frac{(k+1)a + b}{1-at} = b$$

temos

$$\ln m_Q^{(k)}(t) = bt + c' \iff m_Q^{(k)}(t) = ce^{bt},$$

onde c é constante. A solução geral desta equação diferencial é da forma

$$m_Q(t) = \sum_{j=0}^{k-1} c_j t^j + c_k e^{bt}$$

$m_Q(1) = 1$ e as condições iniciais $m_Q^{(i)}(0) = 0$ para $i \leq k-1$ implicam que

$$\sum_{j=0}^{k-1} c_j + c_k e^b = 1 \tag{3.5}$$

$$\left[\sum_{j=0}^{k-1} j c_j t^{j-1} + b c_k e^{bt} \right]_{t=0} = 0 \iff c_1 + b c_k = 0 \iff c_1 = \frac{-b c_k}{1!}$$

$$\begin{aligned}
\left[\sum_{j=0}^{k-1} j(j-1)c_j t^{j-2} + b^2 c_k e^{bt} \right]_{t=0} &= 0 \iff c_2 + b^2 c_k = 0 \iff c_2 = \frac{-b^2 c_k}{2!} \\
&\vdots \\
c_j &= \frac{-b^j c_k}{j!}
\end{aligned} \tag{3.6}$$

Substituindo (3.6) em (3.5) resulta,

$$\sum_{j=0}^{k-1} \frac{-b^j c_k}{j!} + c_k e^b = 1 \iff c_k = \frac{1}{e^b - \sum_{j=0}^{k-1} \frac{b^j}{j!}}$$

Substituindo a expressão (3.6) e c_k na solução geral temos,

$$m_Q(t) = \sum_{j=0}^{k-1} c_j t^j + c_k e^{bt} = \sum_{j=0}^{k-1} \frac{-b^j c_k}{j!} t^j + \frac{1}{e^b - \sum_{j=0}^{k-1} \frac{b^j}{j!}} e^{bt}$$

Multiplicando ambos os termos da segunda fracção por e^{-b} resulta

$$-c_k \sum_{j=0}^{k-1} \frac{b^j}{j!} t^j + \frac{e^{-b(1-t)}}{1 - \sum_{j=0}^{k-1} \frac{e^{-b} b^j}{j!}} = \frac{\sum_{j=0}^{k-1} \frac{b^j}{j!} t^j}{\sum_{j=0}^{k-1} \frac{b^j}{j!} - e^b} + \frac{e^{-b(1-t)}}{1 - \sum_{j=0}^{k-1} \frac{e^{-b} b^j}{j!}}$$

Multiplicando ambos os termos da primeira fracção por $-e^{-b}$ resulta

$$-\frac{\sum_{j=0}^{k-1} \frac{e^{-b} (bt)^j}{j!}}{1 - \sum_{j=0}^{k-1} \frac{e^{-b} b^j}{j!}} + \frac{e^{-b(1-t)}}{1 - \sum_{j=0}^{k-1} \frac{e^{-b} b^j}{j!}} = \frac{e^{-b(1-t)} - \sum_{j=0}^{k-1} \frac{e^{-b} (bt)^j}{j!}}{1 - \sum_{j=0}^{k-1} \frac{e^{-b} b^j}{j!}}$$

Esta última expressão é a da função geradora de probabilidades de uma *Poisson*($b; k$).

- caso $a > 0$:

Há a distinguir cinco possibilidades para b :

- caso $b > -a$:

Pelo lema 3.2.3, $a + b \geq 0 \implies a < 1$; logo $a \in (0, 1)$. Partindo de $m_Q^{(k)}(t) = c(1-a)^{-(k+1+\frac{b}{a})}$ e fazendo $\beta = \frac{a+b}{a}$, conclui-se que $\beta \in (0, \infty)$ e que $m_Q^{(k)}(t) = c(1-a)^{-(k+\beta)}$. A solução geral desta equação diferencial é da forma

$$m_Q(t) = \sum_{j=0}^{k-1} c_j t^j + c_k (1-at)^{-\beta} \tag{3.7}$$

$m_Q(1) = 1$ e as condições iniciais $m_Q^{(i)}(0) = 0$ para $i \leq k-1$ implicam que

$$\sum_{j=0}^{k-1} c_j + c_k (1-a)^{-\beta} = 1 \tag{3.8}$$

$$\left[\sum_{j=0}^{k-1} c_j j t^{j-1} + a\beta (1-at)^{-\beta-1} c_k \right]_{t=0} = 0 \iff c_1 + a\beta c_k = 0 \iff c_1 = -a\beta c_k$$

$$\left[\sum_{j=0}^{k-1} c_j j(j-1) t^{j-2} - a^2 \beta (-\beta-1) (1-at)^{-\beta-2} c_k \right]_{t=0} = 0 \iff$$

$$\begin{aligned}
&\Longleftrightarrow 2c_2 - a^2\beta(-\beta-1)c_k = 0 \Longleftrightarrow c_2 = \frac{-a^2\beta(\beta+1)c_k}{2!} \\
&\quad \vdots \\
&c_j = \frac{-a^j\beta(\beta+1)\cdots(\beta+(j-1))}{j!}c_k = -a^j\binom{\beta+j-1}{j}c_k \quad (3.9)
\end{aligned}$$

Substituindo (3.9) em (3.8),

$$-\sum_{j=0}^{k-1}\binom{\beta+j-1}{j}a^jc_k + c_k(1-a)^{-\beta} = 1 \Longleftrightarrow c_k = \frac{1}{(1-a)^{-\beta} - \sum_{j=0}^{k-1}\binom{\beta+j-1}{j}a^j}$$

Substituindo a expressão (3.9) e c_k na solução geral,

$$\begin{aligned}
m_Q(t) &= \sum_{j=0}^{k-1} c_j t^j + c_k(1-at)^{-\beta} = \\
&= \frac{-\sum_{j=0}^{k-1}\binom{\beta+j-1}{j}a^j t^j}{(1-a)^{-\beta} - \sum_{j=0}^{k-1}\binom{\beta+j-1}{j}a^j} + \frac{(1-at)^{-\beta}}{(1-a)^{-\beta} - \sum_{j=0}^{k-1}\binom{\beta+j-1}{j}a^j}
\end{aligned}$$

Multiplicando ambos os termos das duas frações por $(1-a)^\beta$ resulta

$$\begin{aligned}
&\frac{-\sum_{j=0}^{k-1}\binom{\beta+j-1}{j}(at)^j(1-a)^\beta}{1 - \sum_{j=0}^{k-1}\binom{\beta+j-1}{j}a^j(1-a)^\beta} + \frac{\left(\frac{1-at}{1-a}\right)^{-\beta}}{1 - \sum_{j=0}^{k-1}\binom{\beta+j-1}{j}a^j(1-a)^\beta} = \\
&= \frac{\left(\frac{1-at}{1-a}\right)^{-\beta} - \sum_{j=0}^{k-1}\binom{\beta+j-1}{j}(at)^j(1-a)^\beta}{1 - \sum_{j=0}^{k-1}\binom{\beta+j-1}{j}(1-a)^\beta a^j}
\end{aligned}$$

Esta última expressão é a da função geradora de probabilidades de uma *BinomialNegativa* $\left(\frac{a+b}{a}, a; k\right)$.

o caso $b = -a$:

$b = -a \Longleftrightarrow a + b = 0$ e, pelo lema 3.2.3, $a + b = 0 \implies a < 1$. Logo $a \in (0, 1)$. Partindo de

$$\frac{d}{dt}(\ln m_Q^{(k)})(t) = \frac{(k+1)a+b}{1-at} = \frac{ka}{1-at},$$

resulta

$$\begin{aligned}
\ln m_Q^{(k)}(t) &= -k \ln(1-at) + \ln c = \ln[c(1-at)^{-k}] \Longleftrightarrow \\
&\Longleftrightarrow m_Q^{(k)}(t) = c(1-at)^{-k}
\end{aligned}$$

A solução geral desta equação diferencial é da forma

$$m_Q(t) = \sum_{j=0}^{k-1} c_j t^j + c_k \ln(1-at)$$

$m_Q(1) = 1$ e as condições iniciais $m_Q^{(i)}(0) = 0$ para $i \leq k-1$ implicam:

$$\sum_{j=0}^{k-1} c_j + c_k \ln(1-a) = 1 \quad (3.10)$$

$$\begin{aligned} m'_Q(0) = 0 &\iff \left[\sum_{j=0}^{k-1} c_j j t^{j-1} + c_k \frac{-a}{1-at} \right]_{t=0} = 0 \iff c_1 + c_k(-a) = 0 \\ &\iff c_1 = \frac{ac_k}{1} \end{aligned}$$

$$\begin{aligned} m''_Q(0) = 0 &\iff \left[\sum_{j=0}^{k-1} c_j j(j-1) t^{j-2} - c_k(-a)^2(1-at)^{-2} \right]_{t=0} = 0 \iff \\ &\iff 2c_2 - c_k(-a)^2 = 0 \iff c_2 = \frac{a^2 c_k}{2} \end{aligned}$$

$$m'''_Q(0) = 0 \iff$$

$$\iff \left[\sum_{j=0}^{k-1} c_j j(j-1)(j-2) t^{j-3} - c_k(-a)^3(-2)(1-at)^{-3} \right]_{t=0} = 0 \iff$$

$$\iff 3.2.1c_3 + 2c_k(-a)^3 = 0 \iff c_3 = \frac{-2c_k(-a)^3}{3.2} \iff c_3 = \frac{a^3 c_k}{3}$$

\vdots

$$c_j = \frac{a^j c_k}{j} \quad (3.11)$$

Substituindo (3.11) em (3.10) resulta:

$$\sum_{j=0}^{k-1} \frac{a^j c_k}{j} + c_k \ln(1-a) = 1 \iff c_k = \frac{1}{\ln(1-a) + \sum_{j=0}^{k-1} \frac{a^j}{j}}$$

Substituindo a expressão (3.11) e c_k na solução geral temos

$$m_Q(t) = \sum_{j=0}^{k-1} c_j t^j + c_k \ln(1-at) = \frac{\sum_{j=0}^{k-1} \frac{a^j t^j}{j}}{\ln(1-a) + \sum_{j=0}^{k-1} \frac{a^j}{j}} + \frac{\ln(1-at)}{\ln(1-a) + \sum_{j=0}^{k-1} \frac{a^j}{j}}$$

Dividindo ambos os termos de ambas as frações por $\ln(1-a)$,

$$\frac{\sum_{j=0}^{k-1} \frac{(at)^j}{j \ln(1-a)}}{1 + \sum_{j=0}^{k-1} \frac{a^j}{j \ln(1-a)}} + \frac{\frac{\ln(1-at)}{\ln(1-a)}}{1 + \sum_{j=0}^{k-1} \frac{a^j}{j \ln(1-a)}} = \frac{\frac{\ln(1-at)}{\ln(1-a)} - \sum_{j=0}^{k-1} \frac{(at)^j}{j |\ln(1-a)|}}{1 - \sum_{j=0}^{k-1} \frac{a^j}{j |\ln(1-a)|}}$$

Esta última expressão é a da função geradora de probabilidades de uma $LOG(a; k)$.

◦ caso $-(m+1)a < b < -ma, m \in \{1, 2, \dots, k\}$.

Para $m = 1$ resulta $-2a < b < -a \implies a + b < 0 \implies a \leq 1$, pelo lema 3.2.3; logo $a \in (0, 1]$; fazendo $\beta = \frac{a+b}{a}$, resulta $b = (\beta - 1)a$. Assim,

$$-(m+1)a < (\beta - 1)a < -ma \Leftrightarrow -(m+1) < \beta - 1 < -m \Leftrightarrow$$

$$-m < \beta < -m + 1 \Leftrightarrow \beta \in (-m, -m + 1)$$

Procedendo analogamente ao caso $b > -a$, mas tendo em conta que a pode ser igual a 1, obtém-se,

$$m_Q(t) = \frac{(1-at)^{-\beta} - \sum_{j=0}^{k-1} \binom{\beta+j-1}{j} (at)^j}{(1-a)^{-\beta} - \sum_{j=0}^{k-1} \binom{\beta+j-1}{j} (a)^j}$$

Esta última expressão é a da função geradora de probabilidades de uma $ENB(m, \frac{a+b}{a}, a; k)$

◦ caso $b = -ma, m \in \{2, 3, \dots, k\}$.

Para $m = 2$, resulta $a + b = a - 2a = -a < 0$, uma vez que $a > 0$. Pelo lema 3.2.3, $a + b < 0 \implies a \leq 1$; logo $a \in (0, 1]$.

Partindo de $\frac{d}{dt}(\ln m_Q^{(k)})(t) = \frac{(k+1)a+b}{1-at}$ e fazendo $b = -ma$ obtém-se $m_Q^{(k)}(t) = c(1-at)^{m-k-1}$.

Provemos, por indução sobre k ($k \geq 2$), que

$$m_Q(t) = \frac{\sum_{n=k}^{\infty} \binom{n}{m}^{-1} (at)^n}{\sum_{n=k}^{\infty} \binom{n}{m}^{-1} a^n},$$

a função geradora de probabilidades de uma $ELOG(m, a; k)$, é solução da equação diferencial anterior. Para $k = 2$:

$$\begin{aligned}
m_Q''(t) &= \frac{\sum_{n=2}^{\infty} \binom{n}{m}^{-1} n(n-1) a^n t^{n-2}}{\sum_{n=2}^{\infty} \binom{n}{m}^{-1} a^n} = \frac{\sum_{n=2}^{\infty} \frac{(n-m)!m!}{n!} n(n-1) a^n t^{n-2}}{\sum_{n=2}^{\infty} \binom{n}{m}^{-1} a^n} = \\
&= \frac{\sum_{n=2}^{\infty} \frac{(n-m)!m!}{(n-2)!} a^n t^{n-2}}{\sum_{n=2}^{\infty} \binom{n}{m}^{-1} a^n} = \frac{\sum_{n=0}^{\infty} \frac{(n+2-m)!m!}{n!} a^{n+2} t^n}{\sum_{n=2}^{\infty} \binom{n}{m}^{-1} a^n} = \\
&= \frac{a^2(2-m)!m! \sum_{n=0}^{\infty} \frac{(n-m+2)!}{n!(2-m)!} (at)^n}{\sum_{n=2}^{\infty} \binom{n}{m}^{-1} a^n} = \frac{a^2 m!(2-m)!}{\sum_{n=2}^{\infty} \binom{n}{m}^{-1} a^n} \sum_{n=0}^{\infty} \binom{n-m+2}{n} (at)^n \\
&= c(1-at)^{m-3} = m_Q^{(2)}(t)
\end{aligned}$$

Provemos agora a hereditariedade:

$$\begin{aligned}
m_Q^{(k+1)}(t) &= (m_Q^{(k)})'(t) = (c(1-at)^{m-k-1})' = \\
&= -ac(m-k-1)(1-at)^{m-k-2} \\
&= -ac(m-k-1)(1-at)^{m-(k+1)-1} \\
&= c(1-at)^{m-(k+1)-1}
\end{aligned}$$

Além disso,

$$m_Q(1) = \frac{\sum_{n=k}^{\infty} \binom{n}{m}^{-1} a^n}{\sum_{n=k}^{\infty} \binom{n}{m}^{-1} a^n} = 1$$

e

$$m_Q^{(j)}(0) = \left[\frac{\sum_{n=k}^{\infty} \binom{n}{m}^{-1} a^n n(n-1) \cdots (n-(j-1)) t^{n-j}}{\sum_{n=k}^{\infty} \binom{n}{m}^{-1} a^n} \right]_{t=0} = 0$$

pois $0 \leq j \leq k-1 < k \leq n$ implica $j < n \iff n-j > 0$.

o caso $b \leq -(k+1)a$.

Daqui resulta $(k+1)a+b \leq 0$, que é impossível, atendendo ao lema 3.2.3. Provou-se, assim, que (a) implica (b).

Provemos, agora, que (b) \Rightarrow (a). Cada uma das seis *claim number distribution* básicas é uma distribuição de Panjer, como se viu, pois elas verificam a relação $q_{n+1} = (a + \frac{b}{n+1})q_n$, com os valores de a e de b então determinados. A k -truncatura de cada uma delas também é uma distribuição de Panjer de ordem k , uma vez que:

$$q_{n+1}^{(k)} = \frac{q_{n+1}}{1 - \sum_{j=0}^{k-1} q_j} = \frac{(a + \frac{b}{n+1})q_n}{1 - \sum_{j=0}^{k-1} q_j} = (a + \frac{b}{n+1}) \frac{q_n}{1 - \sum_{j=0}^{k-1} q_j} = (a + \frac{b}{n+1}) q_n^{(k)}$$

□

As variáveis aleatórias discretas assim generalizadas verificam uma expressão recursiva do tipo de Panjer. Seja N uma variável aleatória cuja distribuição Q é uma *claim number distribution*, e seja $\{X_n\}_{n \in \mathbb{N}}$ uma sucessão de variáveis aleatórias i.i.d. independentes de N com função massa de probabilidade $\{f_n\}_{n \in \mathbb{N}_0}$. Então a função de distribuição $\{g_n\}_{n \in \mathbb{N}_0}$

da soma aleatória $S_N = \sum_{n=1}^N X_n$ é dada por $g_0 = m_Q(f_0)$ e

$$g_n = \frac{1}{1 - af_0} \left[\sum_{i=1}^n \left(a + b \frac{i}{n} \right) g_{n-i} f_i + q_k f_n^{*k} \right], \quad n \geq 1.$$

Capítulo 4

Transformadas Integrais

4.1 Introdução

Usando a inversa de exponenciais é possível estabelecer resultados surpreendentes, como se pode constatar em Shanbhag, Pestana and Sreehari (1977) e em Pestana (1978), cf. também Mendonça, Pestana e Velosa (2005).

Neste último trabalho — que, ao contrário do que se fez nos capítulos anteriores, não é transcrito como parte integrante deste capítulo por o autor considerar que a sua participação nele foi menor — observa-se que duas das mais importantes transformadas instrumentais na investigação da aritmética das leis de probabilidade, a função geradora das probabilidades e a transformada de Laplace, surgem naturalmente na comparação de variáveis aleatórias de contagem com uma variável independente, inversa de $X \sim \text{Geométrica}(1 - s)$, e de variáveis aleatórias positivas com uma variável aleatória independente, inversa de uma $Y \sim \text{Exponencial}(v)$, respectivamente.

Tal levou-nos a investigar transformadas mais gerais. Como a exponencial é uma Weibull-1, e é o limite quando $\gamma \rightarrow 0$ da rica família das Pareto generalizadas, pareceram-nos estes os casos mais promissores. A exploração da família das Weibull fora já iniciada por Shanbhag, Pestana and Sreehari (1977) e Pestana (1978), estabelecendo relações importantes com as estáveis de suporte positivo, de que decorrem resultados importantes sobre divisibilidade infinita e autodecomponibilidade de log-gamas e de log-betas. Privilegiámos por isso a exploração das transformadas Pareto, de que a transformada de Laplace será então o caso particular quando $\gamma \rightarrow 0$.

Na secção que se segue, apresentamos alguns dos resultados que complementam o estudo de Mendonça, Pestana e Velosa (2005), e exemplificam a inversão dessas transformadas usando os métodos simbólicos de Hirschman and Widder (2005).

4.2 Funções geradoras de probabilidade, transformada de Laplace e extensões

Excluindo alguns casos elementares, não é em geral fácil efectuar operações com variáveis aleatórias, mesmo na condição de que os termos a operar sejam independentes. No caso da adição, por exemplo, o tratamento analítico directo implica o cálculo de um integral de convolução. De um ponto de vista algébrico abstracto, as dificuldades radicam no facto de que as variáveis aleatórias independentes, junto com a operação de adição, não constituem um grupo. Este problema pode muitas vezes ser ultrapassado recorrendo a transformadas integrais adequadas, as quais são extremamente importantes no estudo de operações entre variáveis aleatórias.

Entre as transformadas de variáveis aleatórias mais utilizadas encontram-se a função geradora de probabilidade, para o caso discreto, e a função geradora dos momentos, ou transformada de Laplace, aplicável quer no caso discreto quer no caso contínuo. Relações notáveis envolvendo transformadas de variáveis aleatórias podem iluminar o nosso conhecimento destas.

Seja $Y \sim \text{Geométrica}(1-s)$, com suporte nos inteiros positivos, independente de uma variável aleatória discreta X , tomando valores inteiros não negativos. Então

$$\mathbb{P}\left[\frac{X}{Y} \leq 1\right] = \mathbb{P}[X \leq Y] = \sum_{k=0}^{\infty} \mathbb{P}[Y \geq X \mid X = k] \mathbb{P}[X = k] = \mathcal{G}_X(s),$$

a função geradora de probabilidade da variável X .

Analogamente, se W é uma variável aleatória exponencial com valor médio $1/v$, independente de uma variável aleatória contínua $X > 0$, então

$$\mathbb{P}\left[\frac{X}{W} \leq 1\right] = \mathbb{P}[W \geq X] = \int_0^{\infty} e^{-vx} dF_X(x) = \mathcal{L}_X(v),$$

a transformada de Laplace da variável X .

Isto inspirou-nos a considerar outras transformadas integrais potencialmente interessantes em probabilidade. Nomeadamente, recordando que a variável exponencial é o caso particular $\gamma = 0$ da rica família das variáveis de Pareto generalizadas Y_γ , $\gamma \in \mathbb{R}$, exploramos as transformadas integrais de Pareto definidas por analogia com a expressão acima. Como casos particulares, encontramos a transformada de Stieltjes e a transformada potencial, tratadas em Widder (1971), e derivadas fraccionárias.

Por outro lado, fazendo a mudança de escala $W \leftrightarrow vW$ a relação anterior pode ser expressa em termos da distribuição exponencial padrão $Y_0 = vW$ como $\mathbb{P}[X/Y_0 \leq 1/v] = \mathcal{L}_X(v)$, e esta perspectiva torna também patente que estamos perante uma fórmula que permite relacionar o comportamento de uma transformada de Laplace ($\mathcal{L}_X(v)$) em 0 e $+\infty$ com o comportamento em $+\infty$ e 0, respectivamente, de uma medida μ concentrada em $(0, +\infty)$ (a medida de probabilidade associada ao quociente X/Y_0). Estas relações tomam o nome de teoremas abelianos quando o comportamento de $\mathcal{L}_X(v)$ é deduzido a partir do de μ e de teoremas abelianos quando o oposto acontece, e são de extrema importância na teoria da probabilidade, em particular no que diz respeito à teoria das somas de variáveis aleatórias.

Seja então X uma variável aleatória não negativa, com função densidade de probabilidade $\Psi(u)$, e seja Y_γ uma variável aleatória de Pareto generalizada independente de X , com a função de distribuição $H_\gamma(y) = 1 - \tilde{H}_\gamma(y)$, cuja cauda direita é $\tilde{H}_\gamma(y) = \mathbb{P}[Y_\gamma \geq y] = (1 + \gamma y)^{-1/\gamma} \mathbf{I}_{(0, \infty)}$, com $\gamma > 0$. O quociente X/Y_γ também toma valores positivos, e verifica:

$$\mathbb{P}\left[\frac{X}{Y_\gamma} \leq \frac{1}{v}\right] = \mathbb{P}[Y_\gamma \geq vX] = \int_0^\infty \frac{\Psi(u)du}{(1 + \gamma vu)^{1/\gamma}} = \mathcal{F}_X(v),$$

para qualquer v positivo. Como veremos de imediato, $\mathcal{F}(v) \equiv \mathcal{F}_X(v)$ pode ser expressa em termos de uma transformada de convolução. Aqui e no que segue, adoptamos a notação do teorema 7.1 de Hirschmann and Widder (p. 180) — usando ψ e Ψ em vez de ϕ e Φ , para evitar confusão com a distribuição gaussiana —, que será aplicado para obter a inversão desta transformada.

Um caso particular com interesse é $\gamma = 1$, em que ficamos com

$$\mathcal{F}_x(v) = \int_0^{\infty} \frac{\Psi(u)du}{1+vu} = \frac{1}{v} \mathcal{S}_x\left(\frac{1}{v}\right),$$

onde \mathcal{S}_x representa a transformada de Stieltjes de $\Psi(u)$.

Fazendo a mudança de variáveis $u = e^t$, $v = e^{-x}$, vem $du = e^t dt$ e fica-se com:

$$\mathcal{F}(e^{-x}) = \int_{-\infty}^{+\infty} \frac{e^t \Psi(e^t) dt}{(1 + e^{-(x-t)})^{1/\gamma}}.$$

Reescrevendo $1/\gamma = \nu$ e multiplicando o numerador e o denominador por $e^{\frac{\nu}{2}(x-t)}$, vem

$$\int_{-\infty}^{+\infty} \frac{1}{2^\nu} \left(\frac{2}{e^{\frac{x-t}{2}} + e^{-\frac{x-t}{2}}} \right)^\nu e^{\frac{x\nu}{2} + t(1-\frac{\nu}{2})} \Psi(e^t) dt,$$

donde

$$e^{-\frac{x\nu}{2}} \mathcal{F}(e^{-x}) = \int_{-\infty}^{+\infty} \frac{1}{2^\nu} \operatorname{sech}^\nu \frac{x-t}{2} e^{-\frac{t}{2}(\nu-2)} \Psi(e^t) dt,$$

ou seja

$$f(x) = \int_{-\infty}^{+\infty} G(x-t) \psi(t) dt$$

com

$$f(x) = e^{-\frac{x\nu}{2}} \mathcal{F}(e^{-x});$$

$$\psi(t) = e^{-\frac{t}{2}(\nu-2)} \Psi(e^t);$$

$$G(x) = \frac{1}{2^\nu} \operatorname{sech}^\nu \frac{x}{2}.$$

Estamos nas condições do teorema 7.1 de Hirschmann and Widder, (p. 180):

- ψ é limitada e contínua
- $f = G * \psi$

Então, é possível inverter a transformada de convolução, de acordo com a fórmula $\psi(x) = E(D)f(x) = \lim_{n \rightarrow \infty} P_n(D)f(x)$, com

$$P_n(s) = A e^{(b-\epsilon_n)s} \prod_{k=1}^n \left(1 - \frac{s}{a_k}\right) e^{s/a_k}, \quad \epsilon_n \rightarrow 0,$$

onde b e os $\{a_k\}_{k=1}^{\infty}$ são os coeficientes da representação de $E(s)$ em produto infinito, A é uma constante normalizadora, e $1/E(s)$ é a transformada bilateral de Laplace de $G(x)$:

$$\frac{1}{E(s)} = \int_{-\infty}^{\infty} G(x) e^{-sx} dx = 2^{-\nu} \int_{-\infty}^{\infty} \frac{2^{\nu}}{(e^{x/2} + e^{-x/2})^{\nu}} e^{-sx} dx = \int_{-\infty}^{\infty} \frac{e^{-x(s+\frac{\nu}{2})}}{(1 + e^{-x})^{\nu}} dx.$$

Fazendo a mudança de variável $w = 1/(1 + e^{-x})$, vem $1 - w = e^{-x}/(1 + e^{-x})$, $x = \ln w - \ln(1 - w)$, $dx = \frac{1}{w(1-w)}dw$ e

$$\int_0^1 w^{\frac{\nu}{2}} (1-w)^{\frac{\nu}{2}} \left(\frac{w}{1-w}\right)^s \frac{dw}{w(1-w)} = \int_0^1 w^{\frac{\nu}{2}+s-1} (1-w)^{\frac{\nu}{2}-s-1} dw = B\left(\frac{\nu}{2} + s, \frac{\nu}{2} - s\right).$$

Preciso portanto de um desenvolvimento da função beta em produto infinito. Da conhecida fórmula para a função gama (Erdélyi *et al.*, 1955):

$$\Gamma(z) = \lim_{n \rightarrow \infty} \frac{n! n^z}{z(z+1) \dots (z+n)}$$

resulta que

$$\Gamma\left(\frac{\nu}{2} + s\right) \Gamma\left(\frac{\nu}{2} - s\right) =$$

$$\begin{aligned} & \lim_{n \rightarrow \infty} \frac{n! n^{\frac{\nu}{2}+s}}{\left(\frac{\nu}{2} + s\right) \left(\frac{\nu}{2} + s + 1\right) \dots \left(\frac{\nu}{2} + s + n\right)} \frac{n! n^{\frac{\nu}{2}-s}}{\left(\frac{\nu}{2} - s\right) \left(\frac{\nu}{2} - s + 1\right) \dots \left(\frac{\nu}{2} - s + n\right)} \\ &= \lim_{n \rightarrow \infty} \frac{(n!)^2 n^{\nu}}{\left[\left(\frac{\nu}{2}\right)^2 - s^2\right] \left[\left(\frac{\nu}{2} + 1\right)^2 - s^2\right] \dots \left[\left(\frac{\nu}{2} + n\right)^2 - s^2\right]} \\ &= \lim_{n \rightarrow \infty} \frac{(n!)^2 n^{\nu}}{\left[\frac{\nu}{2} \left(\frac{\nu}{2} + 1\right) \dots \left(\frac{\nu}{2} + n\right)\right]^2 \left[1 - \frac{s^2}{\left(\frac{\nu}{2}\right)^2}\right] \left[1 - \frac{s^2}{\left(\frac{\nu}{2} + 1\right)^2}\right] \dots \left[1 - \frac{s^2}{\left(\frac{\nu}{2} + n\right)^2}\right]} \\ &= \Gamma^2\left(\frac{\nu}{2}\right) \prod_{k=1}^{\infty} \left[1 - \frac{4s^2}{(\nu + 2k - 2)^2}\right]^{-1}. \end{aligned}$$

Assim, $E(s)$ é da forma

$$\begin{aligned} \frac{\Gamma(\nu)}{\Gamma\left(\frac{\nu}{2} + s\right) \Gamma\left(\frac{\nu}{2} - s\right)} &= \frac{\Gamma(\nu)}{\Gamma^2\left(\frac{\nu}{2}\right)} \prod_{k=1}^{\infty} \left(1 - \frac{2s}{\nu + 2k - 2}\right) \left(1 + \frac{2s}{\nu + 2k - 2}\right) = \\ &= A e^{bs} \prod_{k=1}^{\infty} \left(1 - \frac{s}{a_k}\right) e^{s/a_k}, \end{aligned}$$

com $\{a_k\} = \left\{\frac{\nu}{2}, -\frac{\nu}{2}, \frac{\nu+2}{2}, -\frac{\nu+2}{2}, \dots\right\}$, $b = -\sum \frac{1}{a_k} = 0$ e $A = \frac{\Gamma(\nu)}{\Gamma^2\left(\frac{\nu}{2}\right)}$.

Temos

$$\sum_k \frac{1}{a_k^2} = 2 \sum_{j \geq 0} \frac{2^2}{(\nu + 2j)^2} < \frac{8}{\nu} + 8 \sum_{j \geq 1} \frac{1}{(2j)^2} = \frac{8}{\nu} + 2 \sum_{j \geq 1} \frac{1}{j^2} < \infty$$

portanto é aplicável o teorema 7.1 de Hirschmann e Widder, p.180. A transformada inversa de $f(x) = e^{-\frac{x\nu}{2}} \mathcal{F}(e^{-x})$ é dada por $\psi(x) = E(D)f(x)$. É por isso necessário determinar

$$\prod_{k=1}^n \left(1 + \frac{2D}{\nu + 2k - 2}\right) [e^{-\frac{x\nu}{2}} \mathcal{F}(e^{-x})], \quad n = 1, 2, \dots$$

Para $n = 1$, fica-se com

$$\begin{aligned}
\left(1 + \frac{2D}{\nu}\right) [e^{-\frac{x\nu}{2}} \mathcal{F}(e^{-x})] &= \\
&= e^{-\frac{x\nu}{2}} \mathcal{F}(e^{-x}) + \frac{2}{\nu} \left[-\frac{\nu}{2} e^{-\frac{x\nu}{2}} \mathcal{F}(e^{-x}) - e^{-x-\frac{x\nu}{2}} \mathcal{F}'(e^{-x}) \right] \\
&= -\frac{2}{\nu} e^{-\frac{x}{2}(\nu+2)} \mathcal{F}'(e^{-x})
\end{aligned}$$

Para $n = 2$,

$$\begin{aligned}
\left(1 + \frac{2D}{\nu+2}\right) \left(1 + \frac{2D}{\nu}\right) [e^{-\frac{x\nu}{2}} \mathcal{F}(e^{-x})] &= \\
&= -\frac{2}{\nu} e^{-\frac{x}{2}(\nu+2)} \mathcal{F}'(e^{-x}) - \frac{2}{\nu+2} \left[\frac{2}{\nu} \left(-\frac{\nu+2}{2} \right) e^{-\frac{x}{2}(\nu+2)} \mathcal{F}'(e^{-x}) \right] - \\
&\quad - \frac{2}{\nu+2} \left[\frac{2}{\nu} e^{-\frac{x}{2}(\nu+2)} (-e^{-x}) \mathcal{F}''(e^{-x}) \right] = \\
&= \frac{2}{\nu} \frac{2}{\nu+2} e^{-\frac{x}{2}(\nu+4)} \mathcal{F}''(e^{-x})
\end{aligned}$$

Conjectura: o produto parcial de ordem n é

$$\frac{(-1)^n 2^n}{\nu(\nu+2) \dots (\nu+2n-2)} e^{-\frac{x}{2}(\nu+2n)} \mathcal{F}^{(n)}(e^{-x}), \quad n = 1, 2, \dots$$

Demonstração da propriedade por indução. Hipótese:

$$\begin{aligned}
\prod_{k=1}^n \left(1 + \frac{2D}{\nu+2k-2}\right) [e^{-\frac{x\nu}{2}} \mathcal{F}(e^{-x})] &= \\
&= \frac{(-1)^n 2^n}{(\nu+2n-2) \dots (\nu+2)\nu} e^{-\frac{x}{2}(\nu+2n)} \mathcal{F}^{(n)}(e^{-x})
\end{aligned}$$

Tese:

$$\begin{aligned} \prod_{k=1}^{n+1} \left(1 + \frac{2D}{\nu + 2k - 2} \right) [e^{-\frac{x\nu}{2}} \mathcal{F}(e^{-x})] &= \\ &= \frac{(-1)^{n+1} 2^{n+1}}{(\nu + 2n)(\nu + 2n - 2) \dots (\nu + 2)\nu} e^{-\frac{x}{2}(\nu + 2n + 2)} \mathcal{F}^{(n+1)}(e^{-x}) \end{aligned}$$

Ora,

$$\begin{aligned} \prod_{k=1}^{n+1} \left(1 + \frac{2D}{\nu + 2k - 2} \right) [e^{-\frac{x\nu}{2}} \mathcal{F}(e^{-x})] &= \\ &= \left(1 + \frac{2D}{\nu + 2n} \right) \prod_{k=1}^n \left(1 + \frac{2D}{\nu + 2k - 2} \right) [e^{-\frac{x\nu}{2}} \mathcal{F}(e^{-x})] \\ &= \left(1 + \frac{2D}{\nu + 2n} \right) \left[\frac{(-1)^n 2^n e^{-\frac{x}{2}(\nu + 2n)}}{(\nu + 2n - 2) \dots (\nu + 2)\nu} \mathcal{F}^{(n)}(e^{-x}) \right] = \\ &= \frac{(-1)^n 2^n e^{-\frac{x}{2}(\nu + 2n)}}{(\nu + 2n - 2) \dots (\nu + 2)\nu} \mathcal{F}^{(n)}(e^{-x}) + \frac{2}{\nu + 2n} \frac{(-1)^n 2^n}{(\nu + 2n - 2) \dots (\nu + 2)\nu} \times \\ &\quad \times \left[-\frac{\nu + 2n}{2} e^{-\frac{x}{2}(\nu + 2n)} \mathcal{F}^{(n)}(e^{-x}) - e^{-x} e^{-\frac{x}{2}(\nu + 2n)} \mathcal{F}^{(n+1)}(e^{-x}) \right] = \\ &= \frac{(-1)^{n+1} 2^{n+1} e^{-\frac{x}{2}(\nu + 2n + 2)}}{(\nu + 2n)(\nu + 2n - 2) \dots (\nu + 2)\nu} \mathcal{F}^{(n+1)}(e^{-x}). \end{aligned}$$

Passemos aos outros factores do produtório infinito; os primeiros n são

$$\prod_{k=1}^n \left(1 - \frac{2D}{\nu + 2k - 2} \right) [e^{\frac{x}{2}(\nu + 2n - 2)} R(e^{-x})], \quad n = 1, 2, \dots$$

a menos de uma constante multiplicativa, onde R é a função definida por

$$e^{-\frac{x}{2}(\nu+2n)} \mathcal{F}^{(n)}(e^{-x}) = e^{\frac{x}{2}(\nu+2n-2)} R(e^{-x}) \iff R(e^{-x}) = e^{-x(\nu+2n-1)} \mathcal{F}^{(n)}(e^{-x})$$

ou seja, $R(x) = x^{-(\nu+2n-1)} \mathcal{F}^{(n)}(x)$. Como os operadores diferenciais em cada factor do produtório têm coeficientes constantes, a ordem por que se aplicam pode ser qualquer (ver Hirschmann e Widder, p. 66). Seguindo a sequência inversa $k = n, n-1, \dots, 2, 1$, vem primeiro

$$\begin{aligned} & \left(1 - \frac{2D}{\nu-2+2n}\right) [e^{\frac{x}{2}(\nu-2+2n)} R(e^{-x})] = \\ &= e^{\frac{x}{2}(\nu-2+2n)} R(e^{-x}) - \frac{2}{\nu-2+2n} \frac{\nu-2+2n}{2} e^{\frac{x}{2}(\nu-2+2n)} R(e^{-x}) + \frac{2}{\nu-2+2n} e^{\frac{x}{2}(\nu-4+2n)} R'(e^{-x}) = \\ &= \frac{2}{\nu-2+2n} e^{\frac{x}{2}(\nu-4+2n)} R'(e^{-x}). \end{aligned}$$

Como esta fórmula vale para qualquer $\nu-2+2n \neq 0$, podemos iterá-la para obter os outros produtos. Por exemplo, multiplicando os termos correspondentes a $k = n$ e $k = n-1$, fica

$$\frac{2}{\nu-2+2n} \left(1 - \frac{2D}{\nu-4+2n}\right) [e^{\frac{x}{2}(\nu-4+2n)} R'(e^{-x})] = \frac{2}{\nu-2+2n} \frac{2}{\nu-4+2n} e^{\frac{x}{2}(\nu-6+2n)} R''(e^{-x}).$$

e, mais geralmente, para $r = n, n-1, \dots, 2, 1$,

$$\begin{aligned} & \prod_{k=1}^r \left(1 - \frac{2D}{\nu+2k-2}\right) [e^{\frac{x}{2}(\nu+2n-2)} R(e^{-x})] = \\ &= \frac{2}{\nu-2+2n} \frac{2}{\nu-4+2n} \cdots \frac{2}{\nu-2r+2n} e^{\frac{x}{2}(\nu-2-2r+2n)} R^{(r)}(e^{-x}). \end{aligned}$$

Com n termos, fica

$$\frac{2^n}{\nu(\nu+2)\cdots(\nu+2n-2)} e^{\frac{x}{2}(\nu-2)} R^{(n)}(e^{-x}), \quad n = 1, 2, \dots$$

Então

$$\begin{aligned} \prod_{k=1}^n \left(1 - \frac{2D}{\nu + 2k - 2}\right) \left(1 + \frac{2D}{\nu + 2k - 2}\right) [e^{-\frac{x\nu}{2}} \mathcal{F}(e^{-x})] = \\ = \frac{(-1)^n 2^{2n}}{[\nu(\nu+2) \dots (\nu+2n-2)]^2} e^{\frac{x}{2}(\nu-2)} \left[e^{-x(\nu+2n-1)} \mathcal{F}^{(n)}(e^{-x}) \right]^{(n)}. \end{aligned}$$

Esta expressão pode ser simplificada notando que

$$\frac{2^n}{\nu(\nu+2) \dots (\nu+2n-2)} = c_n \frac{(2n-1)(2n-3) \times \dots \times 3 \times 1}{(\nu+2n-2)(\nu+2n-4) \dots (\nu+2)\nu} = \frac{c_n}{\left(1 + \frac{\nu-1}{2n-1}\right) \left(1 + \frac{\nu-1}{2n-3}\right) \dots \left(1 + \frac{\nu-1}{3}\right) \left(1 + \frac{\nu-1}{1}\right)},$$

onde $c_n = \frac{4^n n!}{(2n)!}$. Atendendo aos resultados

$$c_n \sim \left(\frac{e}{n}\right)^n \quad (n \rightarrow \infty) \quad \text{e} \quad \frac{1}{\Gamma\left(\frac{1-s}{2}\right)} = \lim_{n \rightarrow \infty} \frac{n^{s/2}}{\sqrt{\pi}} \prod_{j=1}^n \left(1 - \frac{s}{2j-1}\right),$$

(Hirschmann and Widder, p. 73, 75), vê-se que a expressão acima é assintoticamente equivalente a

$$\frac{c_n n^{-(\nu-1)/2}}{\sqrt{\pi} \prod_{j=1}^n \left(1 + \frac{\nu-1}{2j-1}\right)} \sim \left(\frac{e}{n}\right)^n \frac{n^{-(\nu-1)/2}}{\sqrt{\pi}} \Gamma\left(\frac{\nu}{2}\right)$$

Temos portanto

$$\psi(t) = \lim_{n \rightarrow \infty} \frac{\Gamma(\nu)}{\Gamma\left(\frac{\nu}{2}\right)} \left(\frac{e}{n}\right)^{2n} \frac{n^{-(\nu-1)/2}}{\sqrt{\pi}} e^{\frac{t}{2}(\nu-2)} \left[e^{-t(\nu+2n-1)} \mathcal{F}^{(n)}(e^{-t}) \right]^{(n)} = e^{-\frac{t}{2}(\nu-2)} \Psi(e^t)$$

$$\Psi(e^t) = \lim_{n \rightarrow \infty} \frac{\Gamma(\nu)}{\Gamma\left(\frac{\nu}{2}\right)} \left(\frac{e}{n}\right)^{2n} \frac{n^{-(\nu-1)/2}}{\sqrt{\pi}} e^{t(\nu-2)} \left[e^{-t(\nu+2n-1)} \mathcal{F}^{(n)}(e^{-t}) \right]^{(n)}$$

e, fazendo de novo as mudanças de variável $\nu = 1/\gamma$ e $u = e^t$,

$$\Psi(u) = \lim_{n \rightarrow \infty} \frac{\Gamma\left(\frac{1}{\gamma}\right)}{\Gamma\left(\frac{1}{2\gamma}\right)} \left(\frac{e}{n}\right)^{2n} \frac{n^{(\gamma-1)/2\gamma}}{\sqrt{\pi}} u^{2-\frac{1}{\gamma}} \left[u^{2n-1+\frac{1}{\gamma}} \mathcal{F}^{(n)}(u) \right]^{(n)}.$$

No caso de γ ser negativo, a transformada \mathcal{F}_x pode-se exprimir em termos de integrais fraccionários da função densidade de probabilidade de X :

$$\mathcal{F}_x(v) = \int_0^{-\frac{1}{\gamma v}} (1+\gamma x)^{-1/\gamma} f_x(x) dx = (-\gamma v)^{-\frac{1}{\gamma}} \Gamma\left(1 - \frac{1}{\gamma}\right) I^{1-\frac{1}{\gamma}} f_x\left(-\frac{1}{\gamma v}\right),$$

onde $I^\nu f$ é o integral fraccionário de f de ordem ν ,

$$I^\nu f(y) = \frac{1}{\Gamma(\nu)} \int_0^y (y-x)^{\nu-1} f(x) dx.$$

Reescrevendo a relação na forma

$$G(v) \equiv \frac{(-\gamma v)^{\frac{1}{\gamma}} \mathcal{F}_x(v)}{\Gamma\left(1 - \frac{1}{\gamma}\right)} = I^{1-\frac{1}{\gamma}} f_x\left(-\frac{1}{\gamma v}\right),$$

fica claro que, dada a transformada $\mathcal{F}_x(v)$, a densidade $f_x(x)$ pode ser recuperada aplicando a conhecida fórmula de inversão do integral fraccionário a $G(v)$.

Bibliografia

- Abamahomed, I. (2006). *As Classes de Panjer e Extensões*, FCUL.
- Adamic, L. A. (2001). Zipf, Power-Laws and Pareto — a ranking tutorial. <http://ginger.hpl.hp.com/shl/papers/ranking/>
- Athayde, E. (1985). *Estudos Sobre Multiplicação de Variáveis Aleatórias Independentes*, Universidade de Lisboa.
- Bai, Z. D. and Yin, Y. Q. (1984). Distributions of class L_α , *J. Mult. Anal.* **14**, 309–319.
- Balkema, A. A. and Resnick, S. I. (1977). Max-infinite divisibility, *J. Appl. Prob.*, **14**, 309–319.
- Berkson, J. (1980). Minimum chi-square, not maximum likelihood! (com discussão), *Annals of Statistics* **8**, 457–487.
- Bernstein, S. (1928). Sur les fonctions absolument monotones, *Acta Mathematica* **51**, 1–66.
- Bingham, N. H. (1971). Factorization theory and domains of attraction for generalized convolution algebras, *Proc. London Math. Soc.* (3), **23**, 16–30.
- Bingham, N. H., Goldie, C. M. and Teugels, J. L. (1987). *Regular Variation*, Cambridge Univ. Press, Cambridge.
- Burrill, W. C. (1977). *Measure, Integration, Probability*, McGraw-Hill, New York.
- Cameron, A. C., and Trivedi, P. K. (1998). *Regression Analysis of Count Data*, Cambridge University Press, Cambridge.
- Cauchy, A. (1853). Sur les résultats moyens d’observations de même nature, et sur les résultats les plus probables, *C. R. Acad. Sci. Paris* **37**, 198–206.
- Cline, D. (1986). Convolution tails, product tails and domains of attraction, *Probab. Theor.* **72**, 529–557.

- Cohen, A. C. (1961). Estimating the Poisson parameter from samples that are truncated to the right, *Technometrics* **3**, 433–438.
- Cramér, H. (1946, 1991). *Mathematical Methods of Statistics*, Princeton Univ. Press, Princeton.
- Cressie, N., and Read, T. R. C. (1984). Multinomial Goodness-of-fit tests, *Journal of the Royal Statistical Society* **B 46**, 440–464.
- De Finetti, B. (1930). Le funzioni caratteristiche di legge istantanea, *Rend. Ac. Lincei*, (6) **12**, 278–282.
- De Haan, L. (1970). *On Regular Variation and its Application to the Weak Convergence of Sample Extremes*, Math. Centre Tracts **32**, Amsterdam.
- De Moivre, A. (1738). *The Doctrine of Chances*, 2^a ed. (Existe edição moderna acessível, Chelsea, New York, 1967, da terceira edição, de 1756, revista e aumentada.)
- Dhaene, J., and Sundt, B. (1998). On approximating distributions by approximating their De Pril transforms, *Scand. Actuar. J.*, 1–23.
- Doeblin, W. (1940). Sur l'ensemble des puissances d'une loi de probabilité, *Studia Math.*, **9**, 71–96.
- Engen, S. (1974). On species frequency models, *Biometrika* **61**, 263–270.
- Erdélyi, A., Magnus, W., Oberhettinger, F., Tricomi, F. (1955) *Higher Transcendental Functions*, Vol I, McGraw-Hill, New York.
- Feller, W. (1967). On regular variation and local limit theorems, *Proc. 5th Berkeley Symposium on Mathematical Statistics and Probability*, II, Univ. California Press, Los Angeles, 373–378.
- Feller, W. (1968). *An Introduction to Probability Theory and Some of its Applications*, vol. I, Wiley, New York.
- Feller, W. (1971). *An Introduction to Probability Theory and Some of its Applications*, vol. II, Wiley, New York.
- Fisher, R. A., Corbet, A. S., and Williams, C. B. (1943). The relation between the number of species and the number of individuals in a random sample of an animal population, *Journal of Animal Ecology* **12**, 42–58.
- Fisher, R. A. and Tippett, L. H. C. (1928). Limiting form of the frequency distribution of the largest or smallest member of a sample, *Proc. Cambridge Philos. Soc.* **24**, 180–190.

- Fréchet, M. (1927). Sur la loi de probabilité de l'écart maximum, *Ann. Soc. Polonaise Math.*, **6**, 93–116.
- Galambos, J. (1987). *The Asymptotic Theory of Extreme Order Statistics*, Krieger, Florida.
- Galambos, J. (1995). *Advanced Probability Theory*, 2nd ed., M. Dekker, New York.
- Gell-Mann, M. (1994). *The Quark and the Jaguar*, Freeman, New York.
- Gilbert, N. (1989) *Biometrical Interpretation. Making Sense of Statistics in Biology*, Oxford University Press, Oxford.
- Gnedenko, B. V. (1940). On the theory of domains of attraction of stable laws, *Uchenye Zapiski, Moskov Gos. Univ.*, **30**, 61–72.
- Gnedenko, B. V. (1943) Sur la distribution limite du terme maximum d'une série aléatoire, *Ann. Math.* **44**, 423–453.
- Gnedenko, B. V. (1970). Limit theorems for sums of a random number of positive independent random variables, *Proc. 6th Berkeley Symp. Math. Statist. Probab.*, vol. 2, 537–549, California University Press, Berkeley.
- Gnedenko, B. V. and Kolmogorov, A. N. (1954). *Limit Distributions for Sums of Independent Random Variables*, Addison-Wesley, Reading, Mass.
- Gnedenko, B. V. and Kovalenko, I. N. (1968). *Introduction to Queueing Theory*, Program for Scientific Translations, Jerusalem.
- Gomes, M. I. (1984). Penultimate limiting forms in extreme value theory, *Ann. Instit. Statist. Math.* **36A**, 71–85.
- Gomes, M. I. and Pestana, D. D. (1981). On the domain of attraction of stable and of extreme value distributions, *Bull. Greek Math. Soc.* **22**, 105–120, 1981.
- Gomes, M. I. and Pestana, D. D. (1987). Non-Standard Domains of Attraction and Rates of Convergence, *New Perspectives in Theoretical and Applied Statistics*, Wiley, New York, 467–477.
- Graça Martins, E. and Pestana, D. D. (1987). Nonstable Limit Laws in Extreme Value Theory, *New Perspectives in Theoretical and Applied Statistics*, Wiley, New York, 449–457.
- Graça Martins, E. and Pestana, D. D. (1988). The extremal limit problem — extensions, *Probability and Mathematical Statistics with Ap-*

plications, W. Grossman, J. Mogyoródi, I. Vincze and W. Wertz (eds.), Reidel, Dordrecht, 143–153.

- Gurland, J. (1957). Some interrelations among compound and generalized distributions, *Biometrika* **44**, 265–268.
- Haldane, J. B. and Jayakar, S. D. (1963). The distribution of extremal and nearly extremal values in samples from a normal population, *Biometrika* **50**, 89–94.
- Hall, A. (1998). *Extremos de Sucessões de Contagem — Do Outro Lado do Espelho*, F.C.U.L., Lisboa.
- Hardy, G. H., Littlewood, J. E., Pólya, G. (1978). *Inequalities*, Cambridge University Press, Cambridge.
- Hess, K. Th., Lewald, A., and Schmidt, K. D. (2002). An extension of Panjer's recursion, *ASTIN Bulletin* **32**, 283–297.
- Hirschman, I. I., Widder, D. V. (2005) *The Convolution Transform*, Dover Publications.
- Iglésias Pereira, H., Oliveira, O. e Pestana, D. D. (1996). Limites estáveis e comportamentos pré-assintóticos, *A Estatística a Decifrar o Mundo*, 109–116, Salamandra, Lisboa.
- Johansen, S. (1966). An application of extreme point methods to the representation of infinitely divisible distributions, *Z. Wahrscheinlichkeitstheorie verw. Geb.* **5**, 304–316.
- Johnson, N. L., Kotz, S., and Kemp, A. W. (2005). *Univariate Discrete Distributions*, Wiley, New York.
- Karamata, J. (1930), Sur un mode de croissance régulière des fonctions, *Mathematica (Cluj)*, **4**, 38–53.
- Katz, L. (1965). Unified treatment of a broad class of discrete probability distributions, *Classical and Contagious Discrete Distributions*, G. P. Patil (ed.), Pergamon Press, Oxford, 175–182.
- Kendall, D. G. (1963). Extreme point methods in stochastic analysis, *Z. Wahrscheinlichkeitstheorie verw. Geb.*, **1**, 295–300.
- Kendall, D. G. (1967). Renewal sequences and their arithmetic, *Lecture Notes in Mathematics* **31**, Springer, Berlin. (Reeditado em *Stochastic Analysis*, Kendall and Harding, eds., Wiley, New York.)
- Kendall, D. G. and Harding, E. F. (1973). *Stochastic Analysis*, Wiley, London.

- Kingman, J. F. C. (1963). Random walks with spherical symmetry, *Acta Mathematica*, **109**, 11–53.
- Klebanov, L. B., Manija, G. M. and Melamed, I. A. (1984) A problem of Zolotarev and analogs of infinite divisibility in a scheme for summing a random number of random variables, *Theor. Probab. Appl.* **29**, 791–794.
- Klugman, S. A., Panjer, H. H., and Willmot, G. E. (1998) *Loss Models: From Data to Decisions*, Wiley, New York.
- Kolmogorov, A. N. (1932) Sulla forma generale di un processo stocastico omogeneo, *Atti Acad. Naz. Lincei* (6) **15**, 805–808 e 866–869.
- Kovalenko, I. N. (1965). On a class of limit distributions for rarefied flows of homogeneous events, *Lit. Mat. Sbornik* **5**, 569–573. (*Selected Transl. Math. Statist. and Prob.* **9**, Providence, Rhode Island, 1971, 75–81.)
- Kozubowski, T. J. (1994). Representation and properties of geometric stable laws, *Approximation, Probability, and Related Fields*, Plenum, New York. 321–337.
- Kozubowski, T. J. and Rachev, S. T. (1999a). Univariate geometric stable laws, *J. Comp. Anal. Appl.* **1**, 177–217.
- Kruglov, V. M. (1972). On an extension of the class of stable distributions, *Theor. Probab. Appl.*, **17**, 723–732.
- Kumar, A. and Schreiber, B. M. (1978). Characterization of subclasses of class L probability distributions, *Ann. Probab.* **6**, 279–293.
- Laha, R. G. and Rohatgi, V. K. (1979). *Probability Theory*, Wiley, New York.
- Laplace, P. H. S. de (1812). *Théorie Analytique des Probabilités*, Mme Veuve Courcier, Paris.
- Lévy, P. (1925). *Calcul des Probabilités*, Gauthier-Villars, Paris.
- Lévy, P. (1937). *Théorie de l'Addition des Variables Aléatoires*, Gauthier-Villars, Paris.
- Lindeberg, J. W. (1920). Über das Exponentialgesetz in der Wahrscheinlichkeitsrechnung, *Ann. Acad. Sci. Fenn.* **16**, 1–23.
- Lindeberg, J. W. (1922). Eine neue Herleitung des Exponentialgesetzes in der Wahrscheinlichkeitsrechnung, *Math. Zeitschr.* **15**, 211–225.
- Linnik, Yu. V. (1953). Linear forms and statistical criteria, II, *Ukrain. Mat. Z.* **5**, 247–290. (Trad. em *Selected Translations In Mathematical Statistics* **3** (1962), Amer. Math. Soc., Providence, Rhode Island.)

- Loève, M. (1956). Ranking limit problem, *Proceedings III Berkeley Symposium on Mathematical Statistics and Probability*, Univ. California Press, Berkeley, 177–194.
- Logan, B. F., Mallows, C. L., Rice, S. O., and Shepp, L. A. (1973). Limit distributions of self-normalized sums, *Ann. Probab.* **1**, 788–809.
- Lukacs, E. (1970). *Characteristic Functions*, 2nd ed., Griffin, London.
- Maceda, E. C. (1948). On the compound and generalized Poisson distributions, *Ann. Math. Statistics* **19**, 414–416.
- Macis, Ju. Ju. (1971). Limit theorems in a non-classical formulation, *Theor. Probab. Appl.* **16** 175–182.
- Malosevskii, S. G. and Nikitin, Ja. Ju. (1977). Limit theorems without the asymptotic negligibility condition, chapt. IX in Linnik, Ju. V. and Ostrovskii, I. V. *Decomposition of Random Variables and Vectors*, Mer. Math. Soc., Providence, 285–306.
- Mandelbrot, B. (1983). *The Fractal Geometry of Nature*, Freeman, New York.
- Marques, T. A., and Pestana, D. D., Velosa, S. F. (2005). Count Data Models in Biometry and Randomness Patterns in Birds Extra-Pair Paternity, *Listy Biometryczne — Biometrical Letters*, Vol. 42, No. 2, 81–112.
- Mejzler, D. (1956). On the problem of the limit distribution for the maximal term of a variational series, *L'vov Pol. Inst. Nauc. Zp.*, **38**, 90–109.
- Mendonça, S. (2000). *Tópicos sobre Convergência Fraca de Sucessões de Variáveis Aleatórias*, tese de doutoramento, Universidade da Madeira.
- Mendonça, S., Pestana, D. D., e Velosa, S. (2005). Teoremas abelianos e tauberianos, variação regular e exponencial inversa, *Estatística Jubilar — Actas do XII Congresso Anual da SPE*.
- Mittnik, S., and Rachev, S. T. (1993). Modeling asset returns with alternative stable distributions, *Econometric Rev.* **12**, 261–330.
- Moore, P. G. (1954). A note on truncated Poisson distribution, *Biometrics* **10**, 402–406.
- Moore, P. G. (1956). The geometric, logarithmic and discrete Pareto forms of series, *Journal of the Institute of Actuaries* **82**, 130–136.
- Neuhaus, M., Forstmeier W., and Bretz, F. (2001). The distribution

of extra-pair young within and among broods — a technique to calculate deviations from randomness, *Journal of Avian Biology* **32**, 358–363.

- Ospina, A. V. and Gerber, H. U. (1987). A simple proof of Feller's characterization of the compound Poisson distribution, *Insurance: Mathematics and Economics*, **6**, 63–64.
- Panjer, H. H. (1981). Recursive evaluation of a family of compound distributions, *ASTIN Bulletin* **12**, 22–26.
- Patil, G. P., and Wani, J. K. (1965). Maximum likelihood estimation for the complete and truncated logarithmic series distribution, in G. P. Patil (ed.): *Classical and Contagious Discrete Distributions*, 398–409, Pergamon Press, Oxford.
- Pérez-Abreu, V. (1991). Poisson approximation to power series distributions, *American Statistician* **45**, 42–45.
- Pestana, D. D. (1978). *Some Contributions to Unimodality, Infinite Divisibility and Related Topics*. dissertação de doutoramento, Univ. Sheffield.
- Pestana, D. D., and Mendonça, S. (2001). Higher-order monotone functions and Probability Theory, *Generalized Convexity and Generalized Monotonicity*, Springer, Berlin, 317–331.
- Pestana, D. D., Sequeira, F. and Velosa, S. F. (2001). Parseval's relation and self-reciprocal characteristic functions, *Rev. Estat./Statist. Rev.* — 23rd European Meeting of Statisticians, *Contributed Papers II*, 315–316.
- Pestana, D. D., e Velosa, S. F. (2003). Classes de Leis \mathcal{N} -Infinitamente Divisíveis, *Literacia e Estatística — Actas do X Congresso Anual da SPE*.
- Petrie M., and Kempeneers, B. (1998). Extra-pair paternity in birds: explaining variation between species and populations, *Trends Ecol. Evol.* **13**, 285–298.
- Piegorsch, W. W. (1990). Maximum likelihood estimation for the negative binomial dispersion parameter, *Biometrics* **46**, 863–867.
- Pillai, R. N. (1985). Semi- α Laplace distributions, *Communications in Statistics — Theory and Methods* **14**, 991–1000.
- Pólya, G. (1920). Über den zentralen Grenzwertsatz der Wahrscheinlichkeitsrechnung und das Momentenproblem, *Math. Zeit.* **8**, 171–181.
- Rachev, S. T., and Resnick, S. (1991). Max-geometric infinite divisibility and stability, *Comm. Statist. — Stochastic Models* **7**, 191–218.

- Raikov, D. A. (1937). On the decomposition of Poisson laws, *Dokl. Acad. Sci. URSS* **14**, 9–11.
- Reiss, R.-D. and Thomas, M. (1991). *Statistical Analysis of Extreme Values*, Birkhäuser, Basel.
- Rényi, A. (1956). A characterization of the Poisson process, *MTA Mat. Kut. Int. Közl.* **1**, 519–527 (original em húngaro, trad. inglesa em *Selected Papers of Alfred Rényi*, **1**, 1948–1956, P. Turán, ed., 622–279 Akadémiai Kiadó, Budapest; com uma nota de D. Szász sobre os desenvolvimentos posteriores, até 1976).
- Rényi, A. (1964). On an extremal property of the Poisson process, *Annals of the Institute of Statistical Mathematics* **16**, 129–133.
- Resnick, S. I. (1972). Products of distribution functions attracted to extreme value laws, *J. Appl. Prob.* **8**, 781–793.
- Rohatgi, V. K. (1976) *An Introduction to Probability Theory and Mathematical Statistics*, John Wiley & Sons Inc.
- Rólski, T., Schmidli, H., Schmidt, V., and Teugels, J. (1999). *Stochastic Processes for Insurance and Finance*, Wiley, New York.
- Schreiber, M. (1977). Dérivé convexe d’une loi de probabilité et lois stables. Application à un problème d’isomorphisme, *Israel J. Math.* **28**, 287–312.
- Seal, H. L. (1952). The maximum likelihood fitting of the discrete Pareto law, *Bulletin of the Assoc. Institute of Actuaries* **78**, 115–121.
- Shanbhag, D. N., Pestana, D. D., and Sreehari, M. (1977) Some further results on infinite divisibility, *Math. Proc. Camb. Phil. Soc.*, **82**, 289
- Sheldon, B. C., and Ellegren, H. (1999). Sexual selection resulting from extrapair paternity in collared flycatchers, *Animal Behavior* **57**, 285–298.
- Skellam, J. G., and Shenton, L. R. (1957). Distributions associated with random walks and recurrent events (com discussão), *J. Roy. Statist. Soc. B* **19**, 64–118.
- Sreehari, M. (1970). On a class of limit distributions for normalized sums of independent random variables, *Theor. Probab. Appl.* **15**, 258–281.
- Srivastava, H. M. and Manocha, H. L. (1984). *A Treatise on Generating Functions*, Horwood, Chichester.
- Steutel, F. W. (1970). *Preservation of Infinite Divisibility Under Mixing and Related Topics*, Math. Centre, Amsterdam.

- Stutchbury, B. J., Rhymer, J. M., and Morton, E. S. (1994). Extrapair paternity in hooded warblers, *Behavioral Ecology* **5**, 384–392.
- Sundt, B. (1992). On some extensions of Panjer’s class of counting distributions, *ASTIN Bulletin* **22**, 61–80.
- Sundt, B. and Jewell, W. S. (1981). Further results on recursive evaluation of compound distributions, *ASTIN Bulletin* **12**, 27–39.
- Temido, M. G. (2000). *Classes de Leis Limites em Teoria de Valores Extremos — Estabilidade e Semi estabilidade*, DM, FCTUC, Coimbra.
- Urbanik, K. (1973). Limit Laws for Sequences of Normed Sums Satisfying Some Stability Conditions, *Multivariate Analysis III*, Academic Press, New York, 225–237.
- Velosa, S. F. (2003). Somas e Máximos de Variáveis Aleatórias Independentes — Classes de Leis Limites, *Literacia e Estatística — Actas do X Congresso Anual da SPE*.
- Velosa, S. F. (2003). New classes of discrete infinitely divisible laws in *Literacy and Statistics*, P. Brito, A. Figueiredo, F. Sousa, P. Teles e F. Rosado, eds., Sociedade Portuguesa de Estatística, Porto, 673–684.
- Velosa, S. F. (2003). Novas Classes de Leis Infinitamente Divisíveis Discretas, *Literacia e Estatística — Actas do X Congresso Anual da SPE*.
- Velosa, S. F. (2004). Modelos de Katz-Panjer e somas aleatórias, *Estatística com Acaso e Necessidade — Actas do XI Congresso Anual da SPE*.
- Widder, D. V. (1941). *The Laplace Transform*, Princeton Univ. Press, Princeton.
- Widder, D. V. (1971). *An Introduction to Transform Theory (Pure & Applied Mathematics)*, Academic Press, New York.
- Willmot, G. E. (1987). Sundt and Jewell’s family of discrete distributions, *ASTIN Bulletin* **18**, 17–29.
- Yezerinac, S. M., Weatherhead, P. J., and Boag, P. T. (1995). Extra-pair paternity and the opportunity for sexual selection in a socially monogamous bird (*Dendroica petechia*), *Behavioral Ecology and Sociobiology* **37**, 179–188.
- Zar, J. H. (1999). *Biostatistical Analysis*, 4th ed., Prentice Hall, Upper Saddle River.
- Zinger, A. A. (1965). On a class of limit distributions for normed sums of independent random variables, *Theor. Probab. Appl.* **10**, 607–620.

- Zolotarev, V. M. (1957). Mellin-Stieltjes transforms in Probability Theory, *Theor. Probab. Appl.* **2**, 433–460.
- Zolotarev, V. M. (1962). On the general theory of multiplication of independent random variables, *Soviet Math. Dokl.* **3**, 166–170.
- Zolotarev, V. M. (1967). Generalizations of the Lindeberg–Feller theorem, *Theor. Probab. Appl.* **12**, 608–618.
- Zolotarev, V. M. (1967). On the M-divisibility of a stable law, *Theor. Probab. Appl.* **12**, 506–508.
- Zolotarev, V. M. (1970). Théorèmes limites généraux pour les sommes de variables aléatoires indépendantes, C. R. Acad. Sci. Paris **270**, A899–A902.
- Zolotarev, V. M. (1986). *One-Dimensional Stable Distributions*, American Mathematical Society, Rhode Island.
- Zolotarev, V. M. and Koroljuk, V. S. (1961). On the hypothesis proposed by B. V. Gnedenko, *Theor. Probab. Appl.*, **6**, 431–435.