A Compendium of Variance Functions for Real Natural Exponential Families

Alain C. Vandal
Department of Mathematics and Statistics
McGill University
October 1994

A Thesis submitted to the Faculty of Graduate Studies and Research in partial fulfillment of the requirements of the degree of Master of Arts

Copy no. 1

National Library of Canada
Bibliothèque nationale du Canada

Acquisitions and Bibliographic Services Branch
Direction des acquisitions et des services bibliographiques

395 Wellington Street
Ottawa, Ontario
K1A 0N4

395, rue Wellington
Ottawa (Ontario)
K1A 0N4

Your file Votre référence

Our file Notre référence

ISBN 0-612-05432-2

Canada

Preface

Rapid advances in the classification of Natural Exponential families using their Variance Functions make the present work timely, both from a theoretical point of view, where a handy reference can always be helpful, and from an applied perspective, where Quasi-likelihood estimation algorithms and Generalized Linear Models can easily be customized to accomodate new and potentially useful models using only variance functions or variance and link functions.

This work was undertaken at the suggestion of Professor V. Seshadri, who as thesis supervisor, has from the onset expressed great enthusiasm for the project and displayed more patience than the author probably deserved. The author's heartfelt gratitude goes to him, as well as to Professors George P.H. Styan, David B. Wolfson and Keith J. Worsley from the Department of Mathematics and Statistics of McGill University for their unfailing support and open friendliness. The author also wishes to thank Professor Gérard Letac of the Université Paul Sabatier in Toulouse, France for an enlightening conversation concerning probability measure inversion, as well as for suggesting the inclusion of a table of Canonical Caste Members for the variance functions.

On a more personal note, the author also wishes to express his thanks and affection to Nina Gilbert, whose supportiveness was one of the few consistent aspects of the time spent in the preparation of this thesis.

Algebraic results were verified using Maple V release 2: Maple is a trademark of Waterloo Maple Software. Some numerical checks were performed with S-plus version 3.2; S-plus is a trademark of Statsci. This work was typeset using LaTeX version 2.09 by the author.

This research was supported in part by NSERC grant 283-34 and NATO collaborative research grant 921347.

Abstract

Natural Exponential Families (NEF) belonging to the Grand-Babel class have variance functions (VF) of the form $\mathbf{V}_F(m) = P(m)\Delta(m) + Q(m)\sqrt{\Delta(m)}$, with P, Q and Δ polynomials with $\deg P \geq 3$, $\deg Q \geq 2$ and $\deg \Delta \geq 2$. Although the members of this class have not as yet all been enumerated, several useful sub-classes have been fully described, namely the Morris class, with at most quadratic polynomial VFs; the Mora class, with cubic polynomial Vfs; the Babel class, with $\deg P = 0$, $\deg Q \leq 1$ and $\deg \Delta \leq 2$; and the Seshadri class, with $\deg P = \deg Q = \deg \Delta = 1$. In order to motivate a uniform presentation of each member of these classes in compendium form, the basic properties of NEFs are surveyed, with special insistance on extension models such as convolution families, exponential dispersion models and affinities of NEFs. The Grand-Babel NEFs are presented with both a canonical parametrization which emphasizes the link to their basis measure and a more familiar or utilitarian parametrization. Expressions for the variance function, the cumulant transform, the mean-domain mapping, the density (when available), the Legendre transform and some asymptotics are given for each NEF, thus providing links to the theories of Likelihood and Quasi-likelihood, Gemeralized Linear Models, Saddlepoint approximation, Large deviations, Distributions and Asymptotic approximation. The notion of Canonical Caste Member (CCM), an easily identifiable representative of the equivalence class of all affinities of a NEF, is introduced; correspondingly, a table of variance functions for the CCMs of the currently classified Grand-Babel NEFs is provided.

Résumé

Les Familles Exponentielles Naturelles (FEN) appartenant à la classe de Grand-Babel sont caractérisées par une fonction variance (FV) de la forme $\mathbf{V}_F(m) = P(m)\Delta(m) + Q(m)\sqrt{\Delta(m)}$, où P, Q et Δ sont des polynômes avec $\deg P \geq 3$, $\deg Q \geq 2$ et $\deg \Delta \geq 2$. Bien que les membres de cette classe n'aient pas tous été recensés, certaines sous-classes fort utiles sont déjà exhaustivement décrites, nommément la classe de Morris, possédant des FV polynômiales de degré inférieur ou égal à deux; la classe de Mora, aux FV polynômiales cubiques; la classe de Babel, où $\deg P = 0$, $\deg Q \leq 1$ et $\deg \Delta \leq 2$, et la classe de Seshadri, où $\deg P = \deg Q = \deg \Delta = 1$. Dans le but de motiver une présentation uniformisée de chacun des membres de ces classes sous forme de compendium, on effectue un survol des propriétés élémentaires des FEN, en insistant plus spécialement sur leurs extensions, telles que les familles de convolution, les modèles exponentiels de dispersion et les affinités de FEN. On présente les FEN de Grand-Babel à la fois sous une paramétrisation canonique qui souligne le lien qu'elles possèdent avec leur mesure génératrice et sous une paramétrisation plus familière ou utilitaire. Chaque FEN est décrite à l'aide de sa fonction variance, de sa tranformée des cumulants, de sa fonction de moyenne, de sa densité (lorsque celle-ci est disponible), de sa transformée de Legendre et de certains résultats asymptotiques, établissant ainsi des liens avec les théories de vraisemblance et de quasi-vraisemblance, des modèles linéaires généralisés, de l'approximation de point de selle, des grandes déviations, des lois de probabilités et des approximations asymptotiques. On introduit la notion de Membre Canonique de la Caste (MCC), un représentant facilement identifiable de la classe d'équivalence formée par les affinités d'une FEN; y faisant suite, on produit une table des FV de chaque MCC pour les FEN de Grand-Babel actuellement classifiées.

Contents

List of Tables

Introduction

In 1978, Ole Barndorff-Nielsen published *Information and Exponential Families in Statistical Theory*, a treatise which exposed the first firmly measure theoretic constructive approach to Exponential Families and inference theory. Four years later, the publication of Carl N. Morris' seminal "Natural exponential families with quadratic variance functions" spurred a flurry of research in the theory of Natural Exponential Families (NEF) and the classification of their Variance Functions (VF), notably though not exclusively by a team of students and researchers at the Université Paul-Sabatier in Toulouse led by Gérard Letac. For these reasons, what we have seen fit to dub "constructive NEF theory" has been variously dubbed the "Danish" and the "Toulousain" approach.

Research in the classification of NEFs using their variance functions has unearthed not only many useful techniques in distribution theory but also a good number of theoretical gems, particularly as concerns relationships between NEFs. The concept of inversion of a probability measure, for example (§2.4, ch. 1), was pioneered by M.C.K. Tweedie (1945) and later formalized by Letac (1992); it boasts of an elegant formulation as well as a deep association with Lévy processes, and still harbours several open problems, even in the one-dimensional case. Likewise, the reduction of NEFs to equivalence classes of generating measures (§2.1) has illuminated previously unsuspected kinships between NEFs, which in turn have opened up new avenues for research.

Efforts at classification of NEFs began with *power variance functions (PVFs)*, VFs of the form $\mathbf{V}_F(m) = cm^\alpha$ for $m \in \mathbb{R}^+$ and $c, \alpha \in \mathbb{R}$, introduced by Tweedie (1984). PVFs were later explored and their classification completed by Jorgensen (1986 and 1987), Bar-Lev and Enis (1986) and Bar-Lev and Stramer (1987). Another important class of NEFs has been dubbed the Grand-Babel class (see §2.6, chapter 1) by Letac (1992), a whimsical acronym based on the names of its originators, namely Bar-Lev, Bshouty, Enis (1991) and Letac (who had broached the subject in the discussion following Jorgensen's [1987] paper). The classification of Grand-Babel NEFs has been progressing steadily.

At this stage in NEF VF research, the present work intends to provide a compact handbook on Grand-Babel class natural exponential families and their variance functions, thus partially fulfilling a need expressed in the research community for a dictionary or atlas of classified variance functions. Space constraints have limited the scope of the present work to classified Grand-Babel class NEFs; PVFs are thus excluded from it, except insofar as they also belong to Grand-Babel subclasses. The PVFs incidentally covered by this work are the Normal (p. 20), the Poisson (p. 21), the Gamma (p. 26) and the Inverse Gaussian (p. 37) families. The Gamma-Poisson Mixture families are also mentioned in appendix section 1. Stable laws and a discussion of PVF classification are thus set outside the scope of our study. Recent work by Malouche (1994) on the partial classification of a Grand-Babel subclass was completed too recently to be incorporated here.

Chapter 1 is devoted to a brief, almost purely terminological, exposition of the mathematical gears and levers of constructive NEF theory in general, and includes a description of the role of the variance function in the related fields of Saddlepoint approximation, Quasi-likelihood theory and Generalized Linear Models. A note on an original extension of probability measure inversion appears in §2.4. The Grand-Babel class is presented in §2.6, along with a discussion of Canonical Caste Members, a concept suggested by G. Letac as a lexical device useful for the recognition of variance functions of *affine NEFs* (see §2.2).

Chapter 2 forms the core of this work, presenting in compendium form the classified Grand-Babel NEFs and VFs, including the Morris (p. 19), Mora (p. 29), Babel (p. 39) and Seshadri (p. 78) classes. To each *convolution family* (see §2.3) belonging to these classes corresponds a standardized entry; the originality and usefulness of the compendium lay in this standardization: variance functions, generating measure and Laplace transform related expressions are parametrized uniformly throughout using a form of extended canonical parametrization, while expressions for the density, when available, include an explicit reparametrization in more familiar, meaningful or convenient terms. On this topic, we note that several new explicit expressions for densities are stated in chapter 2 and derived in the appendix: these are the densities for the Trinomial families, the Pascal Sum, Poisson-Pascal Sum, Binomial Sum and Gamma Sum families, the S-Abel and S-Takács families, and the Reciprocal Inverse Gaussian families with arbitrary dispersion parameter.

Chapter 3 is a simple table of Canonical Caste Member variance functions for the families described in chapter 2. The inclusion of this section was suggested by G. Letac as a valuable addition to a compendium of variance functions. Appendix B contains important proofs and derivations and some relevant mathematical background.

Chapter 1

Measure and Natural Exponential Families

We define here the basic building blocks of constructive Natural Exponential Family (NEF) theory on $\mathbb{R}$, beginning with the positive generating measures and their Laplace transforms, and ending with statistical properties and concepts related to NEFs or useful for their description. Our discussion will be limited wholly to the real line; however, success has been met in structuring higher-dimensional NEFs, notably by Letac (1988,1992 ch. 6), Casalis (1992, 1994) and Hassairi (1993). The one-dimensional case evidently embodies many of the fundamental concepts of NEF theory and is also (and naturally) more fully explored than its higher-dimensional relatives.

The reader is assumed familiar with the fundamental definitions of measure theory and of statistical estimation and inference. The notation used throughout this chapter will be used uniformly throughout the compendium, and is summarized in Appendix C.

1 Basis Measures and the Laplace Transform

In the constructive, measure theoretical approach to Natural Exponential Family theory, the most basic object is the generating measure, also called the *basis* of the family. As we will see in section 2, this measure is determined up to an equivalence class. In general, however, most NEFs can be generated by simple transformations of simple measures: the transformations involved not only often have a concrete probabilistic meaning, but are also conveniently manipulated using the Laplace transform.

We will first define some operations on measures and then establish a proper panoply of generating and transformable measures.

1.1 The Laplace Transform and Operations on Measure

Let $\mathcal{M}_+$ be the set of positive measures on $\mathbb{R}$ and consider the Laplace transform of $\mu \in \mathcal{M}_+$ given by

$$\mathrm{L}_\mu(\theta) = \int_{\mathbf{R}} e^{\theta x}\mu(dx)$$

defined for $\theta \in D(\mu) = \{\theta \in \mathbb{R} \mid \mathrm{L}_\mu(\theta) < +\infty\}$. We will be concerned only with the interior of $D(\mu)$ as the domain of variation for θ, which we denote $\Theta(\mu) = \mathrm{int}D(\mu)$.

A fundamental result associates bijectively any measure with its Laplace transform, through an inversion formula (see for instance Feller [1971], XIII.4). The inversion of a Laplace transform is often a difficult task, and in some cases an explicit expression for the measure to which it corresponds remains unavailable. Nevertheless the Laplace transform enables us to define some useful and natural operations on measures.

The set of basis measures for real NEFs will be

$$\mathcal{M} = \{\mu \in \mathcal{M}_+ \mid \Theta(\mu) \neq \emptyset \;\&\; \nexists a \in \mathbb{R} : \mu\{a\} = \mu(\mathbb{R})\}$$

and our initial objective will be to construct measures belonging to this set. The main operations reviewed, apart from scaling ($\mu' = a\mu, a \in \mathbb{R}^+$, with $\mathrm{L}_{a\mu}(\theta) = a\mathrm{L}_\mu(\theta)$) and addition ($\mu' = \mu_1 + \mu_2$, with $\mathrm{L}_{\mu_1+\mu_2}(\theta) = \mathrm{L}_{\mu_1}(\theta) + \mathrm{L}_{\mu_2}(\theta)$) are imaging under affinity, convolution, power, exponentiation and geometric expansion. A discussion of measure inversion is postponed until §2.4. The operation of mixing is described in §3.1. Most of the following statements are proved or otherwise discussed in Letac (1992, chapter 1).

The Dirac measure at 0 δ_0 is defined in §1.2.

Affinity Imaging

Let $\mathcal{F} = \sigma(\mathbb{R})$ be the σ-field of interest. Consider the affine transformation

$$A : \mathbb{R} \to \mathbb{R}, t \mapsto at + b, \; a, b \in \mathbb{R}, \; a \neq 0$$

We call $A * \mu$ the *image measure of* μ *under* A, denoted and defined by $A * \mu(B) = \mu(A^{-1}(B))$, $\forall B \in A(\mathcal{F})$. With this definition (see Theorem 2.1 below) if $\mu \in \mathcal{M}$ (respectively, $\mathcal{M}_+$) then $A * \mu \in \mathcal{M}$ (respectively, $\mathcal{M}_+$) as well and $\Theta(A * \mu) = \Theta(\mu)$. Moreover $\mathrm{L}_{A*\mu}(\theta) = e^{b\theta}\mathrm{L}_\mu(a\theta)$.

Convolution

The convolution of two measures μ and ν in $\mathcal{M}_+$ is denoted $\mu * \nu$ and defined by

$$(\mu * \nu)(A) = \int_{\mathbb{R}} \nu(A - x)\mu(dx), \; A \in \mathcal{F}$$

The Laplace transform of a convolution measure is given by $\mathrm{L}_{\mu*\nu}(\theta) = \mathrm{L}_\mu(\theta)\mathrm{L}_\nu(\theta)$, with $\Theta(\mu_{\mu*\nu}) = \Theta(\mu) \cap \Theta(\nu)$ (see Letac [1992]). We abbreviate repeated convolution of the same measure by $\mu^{*n} = \mu * \mu * \ldots * \mu$ for $n \in \mathbb{N}$, with the convention that $\mu^{*0} = \delta_0$.

Power

The simplest method by which a well-defined notion of power for measures can be established is the consideration of their *Jørgensen set*. The Jørgensen set of $\mu \in \mathcal{M}$ is described by

$$\Lambda(\mu) = \left\{\lambda > 0 \mid \exists \mu_\lambda \in \mathcal{M} : \mathrm{L}_{\mu_\lambda}(\theta) = (\mathrm{L}_\mu(\theta))^\lambda, \theta \in \Theta(\mu)\right\}$$

By analogy with $\mathrm{L}_{\mu^{*n}}(\theta) = (\mathrm{L}_\mu(\theta))^n$ for $n \in \mathbb{N}$ as seen above, we define $\nu = \mu^{*\lambda}$ for $\lambda \in \Lambda(\mu)$ if $\mathrm{L}_\nu(\theta) = (\mathrm{L}_\mu(\theta))^\lambda$ for all $\theta \in \Theta(\mu)$.

A measure μ will be said to be infinitely divisible if $\Lambda(\mu) = \mathbb{R}^+$; this definition is equivalent to the usual definition of infinite divisibility as provided, for instance, by Feller (1971). We can note immediately that $\mathbb{N}^* \subset \Lambda(\mu)$ for all $\mu \in \mathcal{M}_+$ (see §2.3).

Exponentiation

If $\mu \in \mathcal{M}_+$, define

$$\exp \mu = \delta_0 + \sum_{j=1}^{\infty} \frac{\mu^{*j}}{j!}$$

Then $\exp \mu \in \mathcal{M}$ if $\mu \neq \delta_0$, $\Theta(\exp \mu) = \Theta(\mu)$ and $\mathrm{L}_{\exp \mu}(\theta) = \exp \mathrm{L}_\mu(\theta)$ for $\theta \in \Theta(\mu)$.

Geometric Expansion

Let $\mu \in \mathcal{M}_+$ be such that $\Theta^*(\mu) = \{\theta \in \Theta(\mu) \mid \mathrm{L}_\mu(\theta) < 1\} \neq \emptyset$. Define

$$\nu_r = (\delta_0 - \mu)^{*(-r)} = \delta_0 + \sum_{j=1}^{\infty} \frac{\Gamma(r+j)}{j!\Gamma(r)} \mu^{*j}$$

Then $\nu_r \in \mathcal{M}$ if ν_r is not a multiple of δ_0, $\Theta(\nu_r) = \Theta^*(\mu)$ and $\mathrm{L}_{\nu_r}(\theta) = (1 - \mathrm{L}_\mu(\theta))^{-r}$ for $\theta \in \Theta(\nu_r)$.

1.2 Examples of Measures and Transformations

We list here a certain number of important measures on $\mathbb{R}$. Not all of these may form the basis for a NEF, but all form a starting point for the construction of such a basis using the above operations.

Table 1.1: Some important positive measures on $\mathbb{R}$

Name of measure	*Definition*
Dirac	δ_x, where $\delta_x(A) = 1$ if $x \in A$ and $\delta_x(A) = 0$ otherwise, for $A \subset \mathbb{R}$
Counting Measure	$\sum_{j=0}^{\infty} \delta_j$
Abel	$\sum_{j=0}^{\infty} \frac{(j+1)^{j-1}}{j!} \delta_j$
Takács(a)	$\sum_{j=0}^{\infty} \frac{\Gamma((a+1)j+1)}{\Gamma(j+1)\Gamma(aj+2)} \delta_j$ for $a > 0$
Lebesgue on A	$\mathbb{1}_A(x)$ where $A \subset \mathbb{R}$
Kendall-Ressel	$\frac{x^x}{\Gamma(x+2)} \mathbb{1}_{\mathbf{R}^+}(x)$
Standard Normal	$\exp\left(-\frac{x^2}{2}\right) \mathbb{1}_{\mathbf{R}}(x)$
Hyperbolic Secant	$\mu = \operatorname{sech}(x) \mathbb{1}_{\mathbf{R}}(x)$

Most of the NEFs included in this compendium are rooted in one or more of the above measures. By suitably applying the various operations on measures mentioned above, we can create a rich diversity of bases for NEFs. Some fundamental examples are included here.

Example 1.2.1 Binomial Measure
Let $\mu = \delta_0 + \delta_1$. *Then for* $n \in \mathbb{N}^*$, $\mu^{*n} = \sum_{j=0}^{n} \binom{n}{j} \delta_j$, *called the Binomial measure.*

Example 1.2.2 Gamma Measure
Let $\mu = \mathbb{1}_{\mathbf{R}^+}(t)$, *the Lebesgue measure on* $\mathbb{R}^+$. *Then* $\mu^{*r} = \frac{t^{r-1}}{\Gamma(r)} \mathbb{1}_{\mathbf{R}^+}(t)$, *for* $r \in \mathbb{R}^+$, *is called the Gamma measure with scale parameter* r.

Example 1.2.3 Poisson Measure
Let $\mu = \exp(\delta_1)$. *Then* $\mu = \sum_{j=0}^{\infty} \frac{\delta_j}{j!}$, *called the Poisson measure.*

Example 1.2.4 Pascal and Negative Binomial Measures
Let $\mu = (\delta_0 - \delta_1)^{*(-r)}$. *Then* $\mu = \sum_{j=0}^{\infty} \binom{r+j-1}{r-1} \delta_j$, *the Pascal measure for* $r \in \mathbb{R}^+$. *If* $r \in \mathbb{N}^*$, *the image measure* μ' *of* μ *by the affinity* $A(t) = t + r$ *is given by* $\mu' = \sum_{j=r}^{\infty} \binom{j-1}{r-1} \delta_j$, *the Negative Binomial measure.*

2 Natural Exponential Families and Extensions

Natural exponential families are constructed here, as well as the objects which most naturally characterize them, such as cumulant transform, mean domain and, of course, variance function. The aforementioned measure operations of affinity imaging and power are recast in terms of extension models of NEFs, and probability measure inversion is introduced. We quote some basic results in the asymptotic theory of variance functions. Finally we make a note on castes, a type of lexical construct derived from variance functions, and introduce an equivalence class across affinities of NEFs with the notion of Canonical Caste Member (CCM).

2.1 Natural Exponential Families

We need a few more definitions before moving on to the construction of NEFs.

We first define the *cumulant transform of* μ as

$$k_\mu(\theta) = \log \mathrm{L}_\mu(\theta)\,,\ \theta \in \Theta\,(\mu).$$

Hölder's inequality can be used to show that $k_\mu(\theta)$ is strictly convex and analytic on $\Theta\,(\mu)$. Hence for $\mu \in \mathcal{M}$ and $\theta \in \Theta\,(\mu)$, define the probability measure

$$P(\theta, \mu)(dx) = \exp\left(\theta x - k_\mu(\theta)\right)\mu(dx).$$

We call θ the *canonical parameter* of the probability measure $P(\theta, \mu)$ and the parametrization in terms of $\theta \in \Theta\,(\mu)$ the *canonical form* of the NEF. It is a simple matter to show that the moment generating function $M(s)$ for this probability measure exists in a neighborhood of 0 when $\theta \in \Theta\,(\mu)$ and is given by

$$M(s) = \exp\left(k_\mu(\theta + s) - k_\mu(\theta)\right).$$

The basis measure generating $P(\theta, \mu)$ is not unique. If $\mu_1 = \exp(ax + b)\mu$, with $a, b \in \mathrm{I\!R}$, then $k_{\mu_1}(\theta) = k_\mu(\theta + a) + b\,,\ \theta \in \Theta\,(\mu_1) = \Theta\,(\mu) - a$. But then $P(\theta, \mu_1) = P(\theta + a, \mu)\,,\ \theta + a \in \Theta\,(\mu_1) + a = \Theta\,(\mu)$. Hence defining

$$F(\mu) = \left\{P(\theta, \mu)(dx) \mid \theta \in \Theta\,(\mu)\right\},$$

we see that since $F(\mu) = F(\mu_1)$, $F(\mu)$ depends not strictly on μ but rather on the equivalence class $\{\mu_1 \in \mathcal{M} \mid \mu_1 = \exp(ax + b)\mu\,,\ a, b \in \mathrm{I\!R}\}$; we can thus write $F = F(\mu)$, called the *natural exponential family (NEF) generated by* μ.

In the sequel, we assume as a matter of notational convenience that $X \sim P(\theta, \mu)(dx)$. Since k_μ is strictly convex on $\Theta\,(\mu)$, the mapping $\tau_\mu(\theta) = k'_\mu(\theta)$ is injective. Define

$$M_F = \tau_\mu(\Theta\,(\mu))$$

called the *mean domain* of F and easily seen to depend on F rather than μ. The name *mean domain* comes from the property

$$k'_\mu(\theta) = \int_{-\infty}^{+\infty} xP(\theta,\mu)(dx) = \mathbb{E}[X],$$

so that τ_μ maps $\theta \in \Theta(\mu)$ bijectively to the mean associated with the probability measure $P(\theta,\mu)$. We denote $\phi : M_F \rightarrow \Theta(\mu), m \mapsto \tau_\mu^{-1}(m)$. The bijectivity of ϕ allows a reparametrization of $P(\theta,\mu), \theta \in \Theta(\mu)$ as $P(m,F), m \in M_F$.

Denote now $\mathbf{V}_F(m) = k''_\mu(\phi(m))$. It is easy to show that

$$\mathbf{V}_F(m) = \int_{-\infty}^{+\infty} (x-m)^2 P(m,F)dx = \mathrm{Var}[X],$$

whence we call $(\mathbf{V}_F, M_F)$ the *variance function* associated with F. An alternative expression for the variance function is

$$\mathbf{V}_F(m) = \frac{1}{\phi'_\mu(m)}, m \in M_F,$$

sometimes useful in computation. It is a well-known and easily shown result that the variance function $(\mathbf{V}_F, M_F)$ characterizes the NEF among all NEF's (see for instance Morris [1982], Letac and Mora [1990]), though it does not characterize any particular member of the family.

The *Legendre transform* of measure μ is given by

$$\begin{aligned} \mathrm{k}^*_\mu(x) &= \sup_{\theta\in\Theta(\mu)} x\theta - \mathrm{k}_\mu(\theta) \\ &= x\phi_\mu(x) - k_\mu(\phi_\mu(x)), \, x \in M_F. \end{aligned}$$

The Legendre transform, like the cumulant transform, is a property of the measure μ and not of the family $F(\mu)$. Letac (1992) provides details about the use of the Legendre transform in large deviation theory. Further motivation is provided for the inclusion of the Legendre transform in the compendium throughout section 3.

2.2 Affine Natural Exponential Families

Consider the real affinity $A(t) = at + b, a, b \in \mathbb{R}, a \neq 0$. The following theorem summarizes facts concerning the NEF generated by $A * \mu$ as it relates to the NEF generated by μ.

Theorem 2.1 *(Letac and Mora [1990]) Let A be an affine transformation as above, and $F = F(\mu)$ be a NEF. Then*

(i) $A * \mu \in \mathcal{M}$ *and* $\Theta(A * \mu) = \Theta(\mu)$

(ii) $k_{A*\mu}(\theta) = k_\mu(a\theta) + b\theta, \theta \in \Theta(\mu)$

(iii) *If for some* $\mu' \in \mathcal{M}$, $F(\mu) = F(\mu')$, *then* $F(A*\mu) = F(A*\mu')$

(iv) $M_{A*\mu} = A(M_F)$, *and* $A * P(m, F(\mu)) = P(A(m), F(A*\mu))$

(v) $\mathbf{V}_{F(A*\mu)} = a^2 \mathbf{V}_{F(\mu)}(A^{-1}(m)), m \in M_F$

The following corollary, though trivial in nature, spells out the computational method and probabilistic interpretation associated with actual members of affine NEFs, by extending part (iv) of Theorem 2.1.

Corollary 2.2 *With A, μ and F as above,*

(vi) $A * P(\theta, \mu) = P(\theta/a, A * \mu)$.

(vii) *If* $X \sim P(\theta, \mu)(dx)$, *then* $A(X) \sim [A * P(\theta, \mu)]\,(dx)$.

Part (vii), in turn, allows us to introduce a natural notation. We will write

$$[aP(\theta,\mu)+b] \quad \text{for} \quad A * P(\theta,\mu),$$
$$[aF(\mu)+b] \quad \text{for} \quad F(A*\mu),$$

and so on, whenever convenient.

2.3 Convolution Families and Exponential Dispersion Models

Let $F = F(\mu)$ be a NEF. The Jørgensen sets of all measures in the class which generates F must be identical, and thus we can define the Jørgensen set of a NEF as $\Lambda(F) = \Lambda(\mu)$.

For $\lambda \in \Lambda(F)$, let $\mu_\lambda \in \mathcal{M}$ be such that $\mathrm{k}_{\mu_\lambda}(\theta) = \lambda \mathrm{k}_\mu(\theta)$, and denote

$$F^\lambda(\mu) = F(\mu_\lambda)$$

The notation is not abusive, since for $\mu'_\lambda = \exp(ax+b)\mu_\lambda$, $a, b \in \mathbb{R}$, we have $F^\lambda(\mu) = F(\mu_\lambda) = F(\mu'_\lambda) = F^\lambda(\mu')$, where $\mu' = \exp(ax + b/\lambda)\mu(dx)$ also generates F. We then call *convolution family (CF)* (or *additive model*) the set of probability measures

$$\mathrm{CF}(F) = \bigcup_{\lambda\in\Lambda(F)} F^\lambda$$

and *exponential dispersion model (EDM)* the set

$$\mathrm{EDM}(F) = \bigcup_{\lambda\in\Lambda(F)} \left[\frac{F^\lambda}{\lambda}\right].$$

We call λ the *dispersion parameter* of the CF or EDM. EDMs were first introduced by Jørgensen (1986 and 1987). From the above it is clear that every NEF generates a CF (respectively, EDM), and that the Jørgensen set induces and indexes a partition of the CF (respectively, EDM) into NEFs. The actual forms of the densities are

$$\mathrm{ED}^*(\theta,\lambda) = \exp\left(\theta x - \lambda k_\mu(\theta)\right)\mu^{*\lambda}(dx)$$

for CFs and

$$\mathrm{ED}(\theta,\lambda) = \exp\left[\lambda\left(\theta y - k_\mu(\theta)\right)\right](A * \mu^{*\lambda})(dy)$$

for EDMs, where the affinity A is given by $A(t) = t/\lambda$, $t \in \mathbb{R}$; in both cases the range of the parameters (θ, λ) is $\Theta(\mu) \times \Lambda(\mu)$. EDMs are sometimes reparametrized in terms of the mean m and $\sigma^2 = 1/\lambda$, yielding the density

$$\mathrm{ED}(m, \sigma^2) = \exp\left(\frac{1}{\sigma^2}\left[x\phi_\mu(m) - k_\mu(\phi_\mu(m))\right]\right)(A * \mu)(dx).$$

We note that $\Lambda(F)$ is an additive semigroup in $\mathbb{R}^+$. Let $p, q \in \Lambda(F)$. Since for $\mu, \nu \in \mathcal{M}$, $k_{\mu*\nu}(\theta) = \mathrm{k}_\mu(\theta) + k_\nu(\theta)$ we have that $k_{\mu_p * \mu_q}(\theta) = \mathrm{k}_{\mu_p}(\theta) + \mathrm{k}_{\mu_q}(\theta) = (p+q)\mathrm{k}_\mu(\theta)$. Hence there exists $\mu_{p+q} = \mu_p * \mu_q$ such that $\mathrm{k}_{\mu_{p+q}}(\theta) = (p+q)\mathrm{k}_{\mu_1}(\theta)$, whence $p + q \in \Lambda(F)$. In particular, since $1 \in \Lambda(F)$, we get $\mathbb{N}^* \subset \Lambda(F)$.

The following theorem (proposition 2.5 in Letac and Mora [1990]) summarizes the above results, and introduces others which will enable us to determine the interplay between variance function and CFs and EDMs.

Theorem 2.3 *(Letac and Mora [1990]) Let $\mu_1, \mu_1' \in \mathcal{M}$ and their λth power $\mu_\lambda = \mu_1^{*\lambda}$ and $\mu_\lambda' = \mu_1'^{*\lambda}$ with λ in $\Lambda(\mu_1)$ and $\Lambda(\mu_1')$ respectively. Assume that $F(\mu_1) = F(\mu_1')$. Then*

(i) $\Lambda(\mu_1) = \Lambda(\mu_1')$ (denoted by Λ) and $F(\mu_\lambda) = F(\mu_\lambda')$ (denoted by F^λ) for $\lambda \in \Lambda$;

(ii) For $\theta \in \Theta(\mu_1)$, $P(\theta, \mu_1)^{*\lambda} = P(\theta, \mu_\lambda)$;

(iii) For all $\lambda \in \Lambda$, $M_{F^\lambda} = \lambda M_{F^1}$;

(iv) For all $\lambda \in \Lambda$ and all $m \in M_{F^\lambda}$,

$$\mathbf{V}_{F^\lambda}(m) = \lambda \mathbf{V}_{F^1}\left(\frac{m}{\lambda}\right).$$

The proofs are all straightforward. Part (iv) above allows us, given a variance function, not only to identify the particular NEF but also the CF and EDM associated with it. (The relationship in part (iv) implies a mapping $F \to F^\lambda$ for $\lambda \in \Lambda(F)$, the so-called *Jørgensen transformation.*) For this reason, the variance functions in this compendium all represent a generic member of their NEF F^λ as a subset of the CF generated by F. The generating measure is thus always represented as the power of some fundamental measure, and the mean domain, in general, will be found to depend upon this power.

A pair of simple propositions, useful in determining the form of an affine CF or EDM from its variance function, close off this section.

Proposition 2.4 *Let F and G be NEF's, and let $\lambda \in \Lambda(G)$.*

(i) Suppose that H is a NEF and that and $a, b \in \mathbb{R}$, $a \neq 0$. Then $F = G^\lambda$ and $G = G^1 = [aH + b]$ implies that $F = [aH^\lambda + \lambda b]$.

(ii) If $m \in M_F \cap M_G$, then

$$\mathbf{V}_F(m) = \frac{1}{\lambda}\mathbf{V}_G(m) \text{ if and only if } F = \left[\frac{G^\lambda}{\lambda}\right] \text{ for } \lambda \in \Lambda(F).$$

2.4 Inverse Natural Exponential Families

Measure inversion is a remarkable operation which can best be defined in the context of probability measures. It was formalized by Letac (1992) and generalizes the concept of inverse distribution introduced by Tweedie (1945). A good treatment is found in Seshadri (1993, ch. 5). Inversion is defined as follows.

Definition 2.5 *(Letac and Mora [1990]) Let* $\mu, \mu_1 \in \mathcal{M}$. *Denote* $\Theta^+(\mu) = \Theta(\mu) \cap (0, +\infty)$. (μ, μ_1) *is an inverse pair (similarly,* μ_1 *is the inverse of* μ *or vice versa) if*

(i) $\Theta^+(\mu) \neq \emptyset \neq \Theta^+(\mu_1)$

(ii) $-k_\mu|_{\Theta^+(\mu)} : \Theta^+(\mu) \to \Theta^+(\mu_1)$ *is a bijection, and* $-k_\mu|_{\Theta^+(\mu)}\left(-k_{\mu_1}|_{\Theta^+(\mu_1)}(\theta)\right) = \theta$ *for* $\theta \in \Theta^+(\mu_1)$.

If (μ, μ_1) is an inverse pair as defined above, it can be shown that $(e^{ax+b}\mu, e^{bx+a}\mu_1)$ is also an inverse pair; hence inverse pairs of NEFs such as $(F = F(\mu), F' = F(\mu'))$ are well-defined.

Inversion sometimes admits of a probabilistic interpretation, essentially relating random processes and their hitting time distribution. The reader may wish to consult Letac and Mora (1990) for more details and references; we will simply indicate that among the most famous interpretable inverse pairs of NEFs are:

(i) the Binomial and Negative Binomial families (pages 23 and 24);
(ii) the Poisson and Gamma families (pages 21 and 26);
(iii) the Normal and Inverse Gaussian families (pages 20 and 37).

In most cases, such as the Generalized Hyperbolic Secant and Arcsine inverse pair (pages 27 and 34), such an interpretation is still lacking. Moreover a given measure need not have an inverse in $\mathcal{M}$. The Hermite families (page 42), for instance, do not have inverses, as is shown in appendix (§B.5).

When inversion is possible, however, the following theorem shows that there exists a natural injective correspondence not only between NEFs, but in fact between individual members of the NEFs when the mean domains are concentrated on $\mathbb{R}^+$.

Theorem 2.6 *(Letac and Mora [1990]) Let* F *and* F_1 *be NEF's in* $\mathbb{R}$. *For* G *a NEF, denote* $M_G^+ = M_G \cap (0, +\infty)$. *Then* (F, F_1) *is an inverse pair if and only if*

(i) $M_F^+ \neq \emptyset \neq M_{F_1}^+$, *and the map* $m \mapsto 1/m$ *restricted to* M_F^+ *is a bijection onto* $M_{F_1}^+$;

(ii) *For all* $m \in M_F^+$, $\mathbf{V}_F(m) = m^3 V_{F_1}\left(\frac{1}{m}\right)$

V. Seshadri and the author, in joint unpublished work based in part on Kendall (1957), Khan and Jain (1978) and Jain and Khan (1979), have generalized somewhat the notion of inversion. Two NEFs F and G are said to form an *extended inverse pair indexed by* r if $[-brF + b]$ and $[rG + b^{-1}]$ form an inverse pair for some $r \in \mathbb{R}$

and for all $b \in \Lambda_G$. For $r > 0$, F can be thought of as a family of distributions ruling an input (discrete or continuous) in an infinite storage space subject to deterministic output, while G is a family of distributions ruling the time to first emptiness of this space. A striking property of extended inversion is that it can be used to generate the Mora class convolution families from the families in the Morris class (chapter 2). See appendix (§B.2) for details on the measure transformation involved.

2.5 Asymptotic Distributions and the Variance Function

Because they are often much simpler in form than their parent density and Laplace transform, and because they characterize their NEF, variance functions can provide quick asymptotic results through the following theorem, due to Mora (1990), and here restricted to the real line.

Theorem 2.7 *Let $(F_n)_{n=0}^{\infty}$ be a sequence of NEFs on $\mathbb{R}$ with mean domains M_n and variance functions $\mathbf{V}_n$. Assume that there exists a non-empty open subset $M \subset \cap_{n=0}^{\infty} M_n$ such that $\lim_{n\to\infty} \mathbf{V}_n(m) = \mathbf{V}(m)$ exists uniformly for all $m \in M$ and $\mathbf{V}(m) > 0$ for $m \in M$.*

Then there exists a NEF F on $\mathbb{R}$ such that

$$\mathbf{V} = \mathbf{V}_F, M_F \subset M \text{ , and } P(m, F_n) \rightarrow P(m, F),$$

in the sense of tight convergence.

Jørgensen and Martínez (1991) consider convergence of EDMs to EDMs with power variance functions ($\mathbf{V}_F(m) = am^k$).

2.6 Canonical Caste Members

The so-called *Grand-Babel* class of NEF's on $\mathbb{R}$ is characterized by VF's of the form

$$\mathbf{V}_F(m) = P(m)\Delta(m) + Q(m)\sqrt{\Delta(m)},$$

where P, Q and Δ are polynomials such that $\deg P \leq 2$, $\deg Q \leq 1$ and $\deg \Delta \leq 2$. Various specializations of the Grand-Babel form of the variance function give rise to the groupings of VFs used in this compendium. A partial list of these groupings is given below.

Table 1.2: Classified sub-classes of the Grand-Babel class

Class	P	Q	Δ
Morris	$\deg P(m) = 0$	$Q(m) \equiv 0$	$\deg \Delta(m) \leq 2$
Mora	$P(m) \equiv 0$	$\deg Q(m) = 2$	$\deg \Delta(m) = 2$
Babel	$\deg P(m) = 0$	$\deg Q(m) \leq 1$	$1 \leq \deg \Delta(m) \leq 2$
Seshadri	$\deg P(m) = 1$	$\deg Q(m) = 1$	$\deg \Delta(m) = 1$

The conditions on P, Q and Δ are not by themselves sufficient to characterize a given class. The notion of *caste* must be introduced in order to do so.

Among the various groupings it is convenient for geometrical, probabilistic and lexical reasons to express the VF's of the Grand-Babel class in a canonical form based on the polynomial $\Delta(m)$. We shall say, following Letac's (1992) and Kokonendji's (1993) lead, that if there exists an affinity $K(t) = at + b$, $a, b \in \mathbb{R}$ and a constant k such that $k\Delta\left(K^{-1}(m)\right) = 1$, m, m^2, $m^2 - 1$, $1 - m^2$ or $m^2 + 1$, then F belongs to the *caste* $k\Delta\left(K^{-1}(m)\right)$. Kokonendji (1993) lists various properties associated with castes, such as closure properties under various measure operations, geometric interpretation, etc. We note that the above definition of caste together with Table 2.6 is equivalent to Kokonendji's definition.

Morris (1982) introduces the *canonical member* of the Morris class as the family with variance function $\mathbf{V}(m) = am^2 + s$, where $a \in \mathbb{R}$ and $s \in \{-1, 0, 1\}$. We extend the concept by applying it to the polynomial Δ and arranging, without loss of generality and at the cost of an affinity, for a to belong to $\{-1, 0, 1\}$. The caste member with variance function $\mathbf{V}_{KF}(m) = a^2\mathbf{V}_F\left(K^{-1}(m)\right)$ will then be called the *canonical caste member (CCM) of the affinity family* $\left\{[aF + b] \mid (a, b) \in \mathbb{R}^2\right\}$.

The usefulness of canonical caste representation as a lexical tool is obvious, as it simplifies recognition of a particular NEF given its VF by providing a uniform representation amongst all affinities of the family. The following proposition makes explicit the transformations used to produce caste canonical members (or equivalently their variance function).

Proposition 2.8 *Let* $\mathbf{V}_F(m) = P(m)\Delta(m) + Q(m)\sqrt{\Delta(m)}$, *with* $\Delta(m) = \alpha m^2 + \beta m + \gamma$, $\alpha, \beta, \gamma \in \mathbb{R}$. *Assume* $|\beta| \neq 2\sqrt{\alpha\beta}$, *i.e. that* $\Delta(m)$ *is not a perfect square, and that* $\deg \Delta \neq 0$.

Case 1: *If* $\alpha = 0$, *let* $A_1(t) = \beta^{-1}(t - \gamma)$. *If* $G_1 = A_1^{-1}F$, *then* $\mathbf{V}_{G_1}(m) = P_1(m)m + Q_1(m)\sqrt{m}$, *with*

$$\begin{aligned} P_1(m) &= \beta^2 P\left[A_1(m)\right] \\ Q_1(m) &= \beta^2 Q\left[A_1(m)\right]. \end{aligned}$$

Case 2: *If* $\alpha \neq 0$, *let* $D = \beta^2 - 4\alpha\gamma$, $\sigma = \mathrm{sgn}\alpha$ *and* $\delta = \sigma\mathrm{sgn}D$. *Let* $A_2(t) = \dfrac{\sqrt{|D|}}{2|\alpha|}t - \dfrac{\beta}{2\alpha}$. *If* $G_2 = A_2^{-1}F$, *then* $\mathbf{V}_{G_2}(m) = P_2(m)\Delta^*(m) + Q_2(m)\sqrt{\Delta^*(m)}$, *with*

$$\begin{aligned} \Delta^*(m) &= \sigma m^2 - \delta, \\ P_2(m) &= |\alpha| P\left[A_2(m)\right], \\ Q_2(m) &= \frac{2|\alpha|^{3/2}}{\sqrt{|D|}} Q\left[A_2(m)\right] \end{aligned}$$

These transformations facilitate the automatic determination of the CCM of a CF from its variance function when the CF belongs to the Grand-Babel class.

3 Statistical Aspects

We now provide motivation for the information included in the compendium by describing some relevant statistical concepts as they relate to constructive NEF theory.

3.1 Mixture Distributions

Consider the set of measures $A = \{\mu_r \in \mathcal{M}_+ \mid r \in \mathcal{I}\}$ for some index set $\mathcal{I} \subset \mathbb{R}$, and a measure ν with support $S(\nu) \subset \mathcal{I}$. (The support of ν is the smallest set A such that $\nu(\mathbb{R} \setminus A) = 0$.) Then we define the mixture of A by ν as

$$\mu_r \wedge_r \nu = \int_{S(\nu)} \mu_r \nu(dr).$$

Consider now a measure $\mu \in \mathcal{M}_+$, $\mathcal{I} = \Lambda(\mu)$ and $A = \left\{\mu^{*\lambda} \mid \lambda \in \Lambda(\mu)\right\}$. Then define for $S(\nu) \subset \Lambda(\mu)$

$$\mu \wedge \nu = \mu^{*Y} \wedge_Y \nu = \int_{S(\nu)} \mu^{*y} \nu(dy).$$

Here μ is called the *kernel measure* and ν the *mixing measure*. Henceforth, unless otherwise specified, the term "mixture" will be used to denote the above type of mixing (with respect to powers of a single measure). Then $k_{\mu\wedge\nu}(\theta) = k_\nu\left(k_\mu(\theta)\right)$ and $\Theta(\mu \wedge \nu) = \Theta(\mu) \cap k_\mu^{-1}\left(\Theta(\nu)\right)$.

The mixture of NEF members itself belongs to a NEF, since, as is shown in appendix (§B.3),

$$P(\theta_1, \mu) \wedge P(\theta_2, \nu) = P(\theta_1, \mu_1 \wedge \nu)$$

where $\mu_1 = \exp\left(\theta_2 - k_\mu(\theta_1)\right)\mu$, $\theta_1 \in \Theta(\mu_1 \wedge \nu) = \Theta(\mu) \cap k_\mu^{-1}\left(\Theta(\nu) + \theta_2 - k_\mu(\theta_1)\right)$.

As θ_1 and θ_2 vary, the mixture itself does not in general define a NEF (since $\mu_1 \wedge \nu$ and $\Theta(\mu_1 \wedge \nu)$ depend upon θ_1 and θ_2), but each mixture belongs to and induces a NEF. Mixture distributions often make for straightforward probabilistic interpretations; these have been made explicit in several cases in the compendium.

3.2 Saddlepoint Approximation

Let $\pi \in \mathcal{M}$ be a probability measure with $0 \in \Theta(\pi)$, and consider a random sample $X_1, \ldots, X_n \sim \pi(dx)$, with $\bar{X}_n = (1/n)\sum_{j=0}^{n} X_j$. Daniels (1954) has shown that if $f_n(x)$ is the density of $\bar{X}_n$ and if $g_n(x) = \sqrt{\dfrac{n}{2\pi \mathbf{V}_F(x)}} \exp\left(-k_\pi^*(x)\right)$, where $k_\pi^*(x)$ is the Legendre transform of π, then $(f_n(x)/g_n(x)) \to 1$ as $n \to +\infty$ for all $x \in M_{F(\pi)}$. The function g_n may require renormalization to be used as an approximate density, but the error involved in the approximation is $O(1/n)$.

Since for $\mu \in \mathcal{M}$ and $\mu' = \exp(ax + b)\mu$ we get $k_{\mu'}^*(x) = \mathrm{k}_\mu^*(x) - (ax + b)$, the Legendre transforms included in this compendium may be used to derive a saddlepoint approximation for the mean of $X_1, \ldots, X_n \sim \pi(dx) = P(\theta_0 \in \Theta(\mu), \mu)(dx)$, say, by taking $k_\pi^*(x) = \mathrm{k}_\mu^*(x) - (\theta_0 x - k_\mu(\theta_0))$, for $\theta_0 \in \Theta(\mu)$. Note that in this case $0 \in \Theta(\pi)$, since $\theta_0 \in \Theta(\mu) = \Theta(\pi) + \theta_0$. The Legendre transform can thus be used to approximate densities for EDMs when closed forms are unavailable or intractable.

Daniels (1980) has also shown that the only univariate distributions for which the saddlepoint approximation is exact are the Normal and Inverse Gaussian. The Gamma distribution also yields an exact saddlepoint approximation, up to a factor which depends on n.

Seshadri (1993) provides an effective summary of saddlepoint approximation in the context of NEF theory.

3.3 Generalized Linear Models

Consider a sample $X_i, i = 1, \ldots, n$ and a covariate matrix $\mathbf{Y}_{n\times q} = [y_{ij}] = [\boldsymbol{y}_1 \ldots \boldsymbol{y}_n]'$, where $\boldsymbol{y}_i \in \mathbb{R}^q$, $i = 1, \ldots, n$. The associated *generalized linear model (GLM)* is given by

$$X_i \sim \mathrm{ED}(m_i, \sigma_i^2), i = 1, \ldots, n$$

and

$$g(\mathbb{E}[X_i]) = \sum_{j=1}^{q} y_{ij}\beta_j = \boldsymbol{y}_i'\boldsymbol{\beta}$$

where $g : M_F \to \mathbb{R}$ is an injective function, called the *link function*, and $\boldsymbol{\beta} \in \mathbb{R}^q$ is a parameter vector to estimate such that $\boldsymbol{y}_i'\boldsymbol{\beta} \in g(M_F)$. Generalized linear models were initially introduced by Nelder and Wedderburn (1972).

Under a GLM, the density of X_i is

$$\mathrm{ED}(m_i, \sigma_i^2) = \exp\left(\frac{1}{\sigma_i^2}\left[x\phi_\mu(g^{-1}(\boldsymbol{y}_i'\boldsymbol{\beta})) - k_\mu(\phi_\mu(g^{-1}(\boldsymbol{y}_i'\boldsymbol{\beta})))\right]\right)\mu(dx)$$

from which an expression for the likelihood can easily be derived. We call $g = \phi_\mu$ the *canonical link function*. Under a canonical link, the density becomes

$$\mathrm{ED}(m_i, \sigma_i^2) = \exp\left(\frac{1}{\sigma_i^2}\left[x\boldsymbol{y}_i'\boldsymbol{\beta} - k_\mu(\boldsymbol{y}_i'\boldsymbol{\beta})\right]\right)\mu(dx).$$

This form may be computationally convenient in some ways, because the expression for the likelihood function is simplified. The canonical link function, however, may in practice make unreasonable restrictions on $\boldsymbol{\beta}$ and lack a meaningful interpretation.

The *deviance* can be used to perform tests of hypotheses in the GLM setting. Assuming equal weights for the observations x_i, and denoting the estimated mean vector for the observations by $\hat{\boldsymbol{m}}$, the deviance may be written as

$$\mathrm{D}(\hat{\boldsymbol{m}}, \boldsymbol{x}) = 2\lambda \sum_{i=1}^{n} \int_{g(\hat{m}_i)}^{x_i} \frac{x_i - m}{\mathrm{V}_F(m)} dm.$$

The deviance may be construed as a measure of goodness of fit of the GLM, and is asymptotically distributed as χ^2_{n-q} as $n \to +\infty$.

3.4 Quasi-Likelihood and the Variance Function

For a general family of real non-degenerate probability measures F, define the mean domain of F by $M_F = \{m \in \mathbb{R} \mid m = \mathbb{E}_{\mu \in F}[X]\}$. Then define the variance function $\mathbf{V}_F : M_F \to \mathbb{R}^+, m \mapsto \mathrm{Var}_{\mu \in F}[X]$. For $\mu \in F$, $X \sim \mu$ and $m = \mathbb{E}[X]$, call *quasi-likelihood function* the map K defined by the relation

$$\frac{\partial K(X,m)}{\partial m} = \frac{X - m}{\mathbf{V}_F(m)}.$$

Wedderburn (1974) shows how estimation techniques using quasi-likelihood parallel those which use the log likelihood, and how, in fact, quasi- and log-likelihood are equal when F is a NEF. The properties of quasi-likelihood depend on the fact that estimation under μ is done using a NEF probability measure π which closely approximates μ in a neighborhood of the maximum likelihood estimate under π.

Chapter 2

Classified Real Grand-Babel Natural Exponential Families

All currently classified Grand-Babel variance functions are listed in the present chapter along with other basic information. The entries from each class are preceded by a synoptic listing of the class accompanied by general information, including the main references concerning the class.

The convolution families are all generated from a basis measure chosen to be as convenient as possible: expression and recognizability of the measure and of its Laplace transform, shape of the support and mean domain, amenability to interpretation, clarity of the relationship to the generated convolution family were all concerns in the choice of the basis measure, as well as a desire to simplify the form of the variance function and of the density (when available).

We briefly describe the structure of each entry.

Introduction Section

Various details concerning the families are listed before each entry, including notational information, historical details, probabilistic interpretations, occurrence in statistical literature, etc.

Variance Function Section

The variance function is expressed in terms of no more than four terms:

- m: the variance function argument in M_F;
- λ: the dispersion parameter from $\Lambda(F)$, identifying the CF;
- q: in most cases, a secondary dispersion parameter ratio;
- a: a possible index to the generating measure.

The mean domain M_F and Jørgensen set $\Lambda(F)$ are specified when known, as well as the domains of variation of q and a when appropriate. The dispersion ratio q corresponds most often to the ratio of the powers of two convoluted measures, viz. the generating measure. $\mu = \nu_1^{*\lambda} * \nu_2^{*\lambda q}$ Negative values for q, however, may sometimes occur when the relevant Laplace transform corresponds to a non-negative measure; probabilistic interpretations are quite difficult to come by in such cases.

Basis Measure Section

The basis measure is specified in terms of the simple measures of table 1.1 and operations of §1.1 in chapter 1 whenever possible. The operation of inversion is not used explicitly to derive measures in this section.

Cumulant Transform and Mean Domain Mapping Section

The mappings $k_\mu(\theta)$ and $\phi_\mu(m)$ are specified, as well as the canonical parameter space $\Theta(\mu)$.

Density Section

The density is specified in explicit form whenever such a form is known; otherwise it is expressed in terms of a Laplace or Mellin transform on the support of the basis measure. A reparametrization is effected to bring the canonical and dispersion (and possibly other) parameters in line with the usual forms of the distribution, or to simplify the expression for the density when the distribution has had little exposure in the literature. The mean is expressed using the new parametrization.

Legendre Transform Section

The Legendre transform of the basis measure is indicated under this heading. The transform is often expressed of simpler quantities. The information in §3.2, chapter 1, should be consulted in order to use the Legendre transform in the context of saddlepoint approximation.

Asymptotics Section

In accordance with Theorem 2.7 in chapter 1, asymptotic results hold true for fixed m. In some cases the asymptotics are stated using the parametrization supplied in the Density section.

Notes Section

The Notes section contains other relevant information such as the inverse distribution when it exists and belongs to a classified NEF, special forms for the density or the variance function, and other potentially useful facts.

Other References Section

References not indicated elsewhere are included in this section, along with a motivation.

1 Morris Class

The Morris class contains all convolution families with at most quadratic polynomial variances. It was shown to contain exactly the six following families by Morris (1982):

Table 2.1: Morris class convolution families

Normal:	$\left\{N(\xi,\sigma^2) \mid (\xi,\sigma^2) \in \mathbb{R} \times \mathbb{R}^+\right\}$, supported on $\mathbb{R}$
Poisson:	$\left\{Po(\xi) \mid \xi \in \mathbb{R}^+\right\}$, supported on $\mathbb{N}$
Binomial:	$\{Bin(n,p) \mid (p,n) \in (0,1) \times \mathbb{N}^*\}$, supported on $\{0,1,\ldots,n\}$
Pascal (Negative Binomial):	$\left\{NB(r,p) \mid (p,r) \in (0,1) \times \mathbb{R}^+\right\}$, supported on $\mathbb{N}$
Gamma:	$\left\{\Gamma(\alpha,\beta) \mid (\beta,\alpha) \in \mathbb{R}^+ \times \mathbb{R}^+\right\}$, supported on $\mathbb{R}^+$
Generalized Hyperbolic Secant:	$\left\{GHS(r,\beta) \mid (\beta,r) \in (-\pi/2,+\pi/2) \times \mathbb{R}^+\right\}$, supported on $\mathbb{R}^+$

Letac (1992, ch. 3) quotes and reworks characterizations of the Morris class by Meixner (1934) and Feinsilver (1986), the latter being based essentially on the Cramér transform, $f_\mu(x,m) = \exp(-k_\mu^*(x))$. Shanbhag (1972, 1976) produces another characterization of the Morris class based on the diagonality of their 3×3 Bhattacharya matrix. Laha and Lukacs (1960) characterized the Morris class as those families for which a quadratic form in the r.v. has quadratic regression on a linear form in the r.v.

The Morris class families are among the most useful in applications, a fact which convincingly illustrates Letac's principle (1992, ch. 1): *the simpler* $\mathbf{V}_F$, *the more useful is* F. The Morris and Mora classes together form a kernel of such useful F's, which may be transformed and combined to form most families of the Babel class and of the Seshadri class.

1.1 Normal Families

The Normal or Gaussian distribution is denoted $\mathrm{N}(\xi, \sigma^2)$, with parameters $\xi \in \mathbb{R}$ and $\sigma^2 \in \mathbb{R}^+$ denoting the mean and variance of the distribution respectively.

One of the earliest mentions of a Normal-like distribution occurred when De Moivre attempted to derive an approximation for the Binomial distribution (see for instance Stigler [1986]). The Normal in fact occurs as the limiting distribution of standardized sums of independent random variables with common mean, provided their second moments are suitably small (see, for instance, Billingsley [1979], sec. 27). Closely associated with this property is the fact that $\mathrm{N}(0, \sigma^2)$ is a stable distribution (see Feller [1971], ch. VI). Perhaps due to these properties, it is found to reflect or closely approximate the actual distribution of many naturally occurring random variates.

Variance Function:

$$\mathbf{V}_F(m) = \lambda$$
$$M_F = \mathbb{R},\ \Lambda(F) = \mathbb{R}^+$$

Basis Measure:

$$\mu(dx) = \frac{1}{\sqrt{2\pi\lambda}} \exp\left(-\frac{x^2}{2\lambda}\right) \mathbb{1}_{\mathbb{R}}(x)(dx)$$

Cumulant Transform and Mean Domain Mapping:

$$\mathrm{k}_\mu(\theta) = \lambda\theta^2/2,\ \Theta(\mu) = \mathbb{R}$$
$$\phi_\mu(m) = m/\lambda,\ m \in M_F$$

Density:

$$f_X(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(x-\xi)^2}{2\sigma^2}\right),\ x \in \mathbb{R}$$
$$\theta = \frac{\xi}{\sigma^2},\ \xi \in \mathbb{R}\ ;\ \lambda = \sigma^2,\ \sigma^2 \in \mathbb{R}^+\ ;$$
then $m = \xi$.

Legendre Transform:

$$\mathrm{k}^*_\mu(x) = \frac{x^2}{2\lambda},\ x \in M_F$$

Notes:

- $\mathrm{IG}(1/\sigma^2, \xi^2/\sigma^2)$ is the inverse distribution of $\mathrm{N}(\xi, \sigma^2)$ for $\xi > 0$.
- The Normal and the Inverse Gaussian families are the only univariate NEF's for which the saddlepoint approximation is exact.

1.2 Poisson Families

The Poisson distribution is denoted Po(ξ), where ξ is often called the intensity parameter, and equals both mean and variance for the distribution.

Historically, it was derived as the limiting distribution for a Binomial distribution Bin(n, p) as $n \to \infty$ with $np = \xi$ constant (Stigler [1982]). More generally, Jørgensen (1986) showed that every univariate discrete positive exponential dispersion model converges to Po(m) as $\lambda \to \infty$ while m remains fixed.

The Poisson distribution is also defined as the distribution of the number of events within a fixed time interval which occur in a Poisson process.

Variance Function:

$$\mathbf{V}_F(m) = m$$
$$M_F = \mathbb{R}^+ , \Lambda(F) = \mathbb{R}^+$$

Basis Measure:

$$\mu = \exp(\lambda\delta_1)$$

Cumulant Transform and Mean Domain Mapping:

$$\mathrm{k}_\mu(\theta) = \lambda e^\theta , \Theta(\mu) = \mathbb{R}$$
$$\phi_\mu(m) = \log\frac{m}{\lambda} , m \in M_F$$

Density:

$$f_X(x) = \frac{\xi^x}{x!}\exp(-\xi) , x \in \mathbb{N}$$
$$\theta = \log\left(\frac{\xi}{\lambda}\right) , \xi \in \mathbb{R}^+ ;$$
then $m = \xi$

Legendre Transform:

$$\mathrm{k}^*_\mu(x) = x(\log(x) - 1) , x \in M_F$$

Notes:

- $\Gamma(1, 1/\xi)$ is the inverse distribution of $\mathrm{Po}(\xi)$.
- λ can always be taken to be 1 by adjusting the canonical parameter θ appropriately. However, the convolution family or exponential dispersion model form may make the Laplace transforms of convolutions involving Poisson random variables easier to manage. We may examine, for instance, the closure under convolution of the Poisson distribution: since in general $\mathrm{ED}(\lambda_1, \theta) * \mathrm{ED}(\lambda_2, \theta) = \mathrm{ED}(\lambda_1 + \lambda_2, \theta)$, we get $\mathrm{Po}(\xi_1) * \mathrm{Po}(\xi_2) = \mathrm{Po}(\xi_1 + \xi_2)$ by taking $\mathrm{Po}^* \equiv \mathrm{ED}$, $\theta = \log(\xi_1 + \xi_2)$, $\lambda_1 = 1 + \xi_2/\xi_1$, and $\lambda_2 = 1 + \xi_1/\xi_2$ (Jørgensen [1987]). See also, for instance, the NB + P and P − NB families in the Babel class.

Other References:

- Billingsley (1979) for a discussion of central limit theorems applied to triangular arrays of random variables.

1.3 Binomial Families

The Binomial distribution $\mathrm{Bin}(n, p)$ occurs as the number of successes in a series of $n \in \mathbb{N}^*$ independent trials, each with a probability $p \in (0, 1)$ of success. The Binomial distribution arose very early in the history of probability, ostensibly in the course of mathematical descriptions of games of chance (see for instance Stigler [1986], ch. 2).

$\mathrm{Bin}(1, p)$ is often called the Bernoulli distribution.

Variance Function:

$$\mathrm{V}_F(m) = \frac{m}{\lambda}(\lambda - m)$$
$$M_F = (0, \lambda),\ \Lambda(F) = \mathbb{N}^*$$

Basis Measure:

$$\mu = (\delta_0 + \delta_1)^{*\lambda}$$

Cumulant Transform and Mean Domain Mapping:

$$\mathrm{k}_\mu(\theta) = \lambda \log(1 + e^\theta),\ \Theta(\mu) = \mathbb{R}$$

$$\phi_\mu(m) = \log\left(\frac{m}{\lambda - m}\right),\ m \in M_F$$

Density:

$$f_X(x) = \binom{n}{x} p^x (1-p)^{n-x},\ x \in \mathbb{N}$$

$$\theta = \log\left(\frac{p}{1-p}\right),\ p \in (0,1)\ ;\ \lambda = n,\ n \in \mathbb{N}^*\ ;\ \text{mm}$$

then $m = np$

Legendre Transform:

$$\mathrm{k}^*_\mu(x) = \log\left(\frac{x^x(\lambda - x)^{(\lambda - x)}}{\lambda^\lambda}\right),\ x \in M_F$$

Asymptotics:

$\mathrm{Bin}(n, p) \xrightarrow{\mathcal{D}} \mathrm{Po}(\xi)$ if $n \to \infty$ while $np = \xi$ remains fixed.

Notes:

- $\mathrm{NB}(1/n, p)$ is the inverse distribution of $\mathrm{Bin}(n, p)$.

1.4 Pascal (Negative Binomial) Families

The Pascal or Negative Binomial-1 distribution $\mathrm{NB}(r,p)$ with $r \in \mathbb{N}^*$ occurs as the number of failures required to obtain r successes in a series of independent trials when the probability of success for each trial is p. Often in the literature, the names "Negative Binomial" or "Negative Binomial-2" will refer to the $[\mathrm{NB}(r,p)+r]$ distribution while $\mathrm{NB}(r,p)$ is called the Pascal distribution for $r \in \mathbb{N}^*$. $\mathrm{NB}(1,p)$ is often called the Geometric distribution.

The choice of r may be extended to the positive real line. A simple interpretation for this case results from the fact that $\mathrm{NB}(r,p) = \mathrm{Po}(\xi) \wedge_\xi \Gamma(r, p/(1-p))$. In this compendium, we abusively retain the name "Pascal distribution" for general $r \in \mathbb{R}^+$ instead of the more acceptable but less wieldy name of "Negative Binomial-1".

Variance Function:

$$\mathbf{V}_F(m) = \frac{m}{\lambda}(m+\lambda)$$
$$M_F = (0,+\infty)\,,\ \Lambda(F) = \mathbb{R}^+$$

Basis Measure:

$$\mu = (\delta_0 - \delta_1)^{*-\lambda}$$

Cumulant Transform and Mean Domain Mapping:

$$\mathrm{k}_\mu(\theta) = -\lambda \log\left(1-e^\theta\right)\,,\ \Theta(\mu) = \mathbb{R}^-$$
$$\phi_\mu(m) = \log\left(\frac{m}{m+\lambda}\right)\,,\ m \in M_F$$

Density:

$$f_X(x) = \frac{\Gamma(x+r)}{x!\Gamma(r)} p^r (1-p)^x\,,\ x \in \mathbb{N}$$
$$\theta = \log(1-p)\,,\ p \in (0,1)\ ;\ \lambda = r \in \mathbb{R}^+\ ;$$
$$\text{then } m = r\frac{1-p}{p}.$$

Legendre Transform:

$$\mathrm{k}^*_\mu(x) = \log\left(\frac{x^x \lambda^\lambda}{(x+\lambda)^{x+\lambda}}\right)\,,\ x \in M_F$$

Asymptotics:

$\mathrm{NB}(r,p) \xrightarrow{\mathcal{D}} \mathrm{Po}(m)$ if $r \to \infty$ while m as in the Density section above remains constant.

Notes:

- If $1/r \in \mathbb{N}^*$, then $\mathrm{Bin}(1/r, p)$ is the inverse distribution of $[\mathrm{NB}(r,p) + r]$.
- $\mathrm{NB}\left(1, \dfrac{r(1-p)}{r(1-p)+p^2}\right)$ is the inverse distribution of $\mathrm{NB}(r,p)$.

1.5 Gamma Families

The Gamma distribution with parameters $\alpha, \beta \in \mathbb{R}^+$ is denoted $\Gamma(\alpha, \beta)$, and is a left-skewed distribution with applications in modeling failure-time distributions, among other phenomena. It arises as a result of convolution and/or division of the simpler $\text{EXP}(\beta)$ distribution (see below), which has the useful memorylessness property (i.e. if $X \sim \text{EXP}(\cdot)$, $P[t+h > X \geq t | X \geq t] = P[h > X]$).

The case of $\Gamma(1, \beta)$ is the exponential distribution with parameter β, denoted $\text{EXP}(\beta)$. The case of $\Gamma(\nu/2, 2)$ is the chi-squared distribution with ν degrees of freedom, denoted χ^2_ν. In particular, this latter distribution rules sums of ν squared standard normal random variables when $\nu \in \mathbb{N}^*$.

Variance Function:

$$\mathbf{V}_F(m) = \frac{m^2}{\lambda}$$
$$M_F = \mathbb{R}^+ ,\ \Lambda(F) = \mathbb{R}^+$$

Basis Measure:

$$\mu = \gamma^{*\lambda}$$
where $\gamma = 1\!\!1_{\mathbb{R}^+}(x)$

Cumulant Transform and Mean Domain Mapping:

$$\text{k}_\mu(\theta) = -\lambda \log(-\theta),\ \Theta(\mu) = \mathbb{R}^-$$
$$\phi_\mu(m) = -\frac{\lambda}{m},\ m \in M_F$$

Density:

$$f_X(x) = \frac{1}{\Gamma(\alpha)} \frac{1}{\beta^\alpha} x^{\alpha-1} \exp\left(-\frac{x}{\beta}\right),\ x \in \mathbb{R}^+$$
$$\theta = -\frac{1}{\beta},\ \beta \in \mathbb{R}^+ \ ;\ \lambda = \alpha \in \mathbb{R}^+ \ ;\ \text{then } m = \alpha\beta$$

Legendre Transform:

$$\text{k}^*_\mu(x) = -\lambda \left[\log\left(\frac{x}{\lambda}\right) + 1\right],\ x \in M_F$$

Notes:

- $[(1/\alpha)\text{Po}(1/\beta)]$ is the inverse distribution of $\Gamma(\alpha, \beta)$.
- The saddlepoint approximation for the Gamma distribution is exact, up to a constant factor.

1.6 Generalized Hyperbolic Secant Families

The Generalized Hyperbolic Secant distribution, denoted GHS(r, β), is the distribution of the area comprised between an arc in a planar Brownian curve and the chord subtending it (Talacko [1956]). Morris (1982) notes that the GHS distribution is the natural observation of the Beta distribution. Harkness and Harkness (1965) characterize GHS($n, 0$), $n \in \mathbb{N}^*$, as the distribution of the logarithm of the geometric mean of independent Cauchy random variables. Shanbhag (1976) arrives at a full form of the GHS distribution through a Bhattacharya matrix characterization.

Variance Function:

$$\mathbf{V}_F(m) = \frac{m^2}{\lambda} + \lambda\,,\ \lambda \in \mathbb{R}^+$$
$$M_F = \mathbb{R}\,,\ \Lambda(F) = \mathbb{R}^+$$

Basis Measure:

$$\mu = \eta^{*\lambda}$$
$$\text{where } \eta = \frac{\text{sech}(\pi x/2)}{2}\mathbb{1}_{\mathbb{R}^+}(x)$$

Cumulant Transform and Mean Domain Mapping:

$$\mathrm{k}_\mu(\theta) = \lambda \log(\sec\theta)\,,\ \Theta(\mu) = \left(-\frac{\pi}{2}, \frac{\pi}{2}\right)$$
$$\phi_\mu(m) = \arctan\left(\frac{m}{\lambda}\right),\ m \in M_F$$

Density:

$$f_X(x) = 2^{r-2}\frac{|\Gamma(r/2 + ix)|^2}{\Gamma(r)\Gamma^2(r/2)}\cos^r(\beta)e^{\beta x}\,,\ x \in \mathbb{R};$$
$$\theta = \beta\,,\ \beta \in \left(-\frac{\pi}{2}, \frac{\pi}{2}\right)\ ;\ \lambda = r\,,\ r \in \mathbb{R}^+\ ;$$
$$\text{then } m = r\tan\beta$$

Legendre Transform:

$$\mathrm{k}^*_\mu(x) = x\arctan\left(\frac{x}{\lambda}\right) - \frac{\lambda}{2}\log\left(\frac{x^2}{\lambda^2} + 1\right),\ x \in M_F$$

Notes:

- $[(1/r)\text{Arc}(1, 0, \pi/2 - \beta)]$ is the inverse distribution of $\text{GHS}(r, \beta)$.

- Correspondingly to the form of the generating measure with $\lambda = 1$, the density for $r = 1$ can be written as

$$f_X(x) = \frac{\cos(\beta)}{2\cosh\left(\frac{\pi x}{2}\right)} e^{\beta x}, \; x \in \mathbb{R}$$

Other References:

- Johnson and Kotz (1970) for general properties of the distribution.

- Abramowitz and Stegun (1964) for expansions of the Gamma function in the complex plane.

2 Mora Class

The Mora class in its restricted form contains all convolution families with strictly cubic variance functions. In the literature, "Mora class" and "Morris-Mora class" are also used to denote the class of convolution families with at most cubic polynomial variance functions. In her thesis, Mora (1986) showed that the following five families, up to affinity, account for all cubic variance functions:

Table 2.2: Mora class convolution families

Generalized Poisson (Abel):	$\left\{\mathrm{GP}(\kappa_1, \kappa_2) \mid (\kappa_2, \kappa_1) \in (0,1) \times \mathbb{R}^+\right\}$, supported on $\mathbb{N}$
Generalized Negative Binomial (Takács):	$\left\{\mathrm{GNB}(r_1, r_2, p) \mid (p, r_1) \in (r_2/(r_2+1), 1) \times \mathbb{R}^+\right\}$, for $r_2 \in \mathbb{R}^+$, supported on $\mathbb{N}$
Arcsine:	$\left\{\mathrm{Arc}(r, a, \xi) \mid (\xi, r) \in (0, \arctan a^{-1}) \times \mathbb{R}^+\right\}$, for $a \in \mathbb{R}^+ \cup \{0\}$, supported on $\mathbb{N}$
Kendall-Ressel:	$\left\{\mathrm{KR}(r, \xi) \mid (\xi, r) \in (1, +\infty) \times \mathbb{R}^+\right\}$, supported on $\mathbb{R}^+$
Inverse Gaussian:	$\left\{\mathrm{IG}(\chi, \psi) \mid (\psi, \chi) \in \mathbb{R}^+ \times \mathbb{R}^+\right\}$, supported on $\mathbb{R}^+$

Although we retain Mora's original classification concerning the Arcsine type, Letac and Mora (1990) in an important paper surveying cubic variance exponential families see fit to distinguish between the Strict Arcsine and the Large Arcsine families; eliminating the distinction provides us with a uniform way of treating this family in the context of convolution families.

An immediate application of Theorem 2.6 shows that the Morris-Mora families are closed under inversion. It is likewise straightforward to show that these families are closed under weak convergence.

2.1 Generalized Poisson (Abel) Families

The Generalized Poisson distribution, denoted $GP(\kappa_1, \kappa_2)$, $\kappa_1 \in \mathbb{R}^+$, was introduced by Consul and Jain (1973) to provide a Poisson-like model which admits of unequal mean and variance. It also arises as a limiting distribution for the Generalized Negative Binomial distribution. $[GP(\kappa_1, \kappa_1) + \kappa_1\kappa_2]$ is usually called the Borel-Tanner distribution, introduced by Borel (1942) for $\kappa_1\kappa_2 = 1$ and generalized by Tanner (1953) for $\kappa_1\kappa_2 \in \mathbb{R}^+$.

The Generalized Poisson NEFs are called *Abel type* by Mora (1986) and Letac and Mora (1990).

Variance Function:

$$\mathbf{V}_F(m) = \frac{m}{\lambda^2}(m+\lambda)^2$$
$$M_F = \mathbb{R}^+ \,,\, \Lambda(F) = \mathbb{R}^+$$

Basis Measure:

$$\mu = \nu^{*\lambda}$$
$$\text{where } \nu = \sum_{j=0}^{\infty} \frac{(j+1)^{j-1}}{j!}\delta_j$$

Cumulant Transform and Mean Domain Mapping:

$$\mathrm{k}_\mu(\theta) = \lambda f^{-1}(e^\theta)\,,\, \Theta(\mu) = (-\infty, -1)$$
$$\text{where } f^{-1} \text{ is the reciprocal of } f : (0,1) \to (0, e^{-1})\,,\, t \mapsto \frac{t}{e^t}$$
$$\phi_\mu(m) = \log\left(\frac{m}{m+\lambda}\right) - \frac{m}{m+\lambda}\,,\, m \in M_F$$

Density:

$$f_X(x) = \frac{\kappa_1}{x!}(\kappa_1 + \kappa_2 x)^{x-1} e^{-(\kappa_1+\kappa_2 x)}\,,\, x \in \mathbb{N}$$
$$\theta = \log(\kappa_2) - \kappa_2\,,\, \kappa_2 \in (0,1) \;;\; \lambda = \frac{\kappa_1}{\kappa_2}\,,\, \kappa_1 \in \mathbb{R}^+ \;;\text{ then } m = \frac{\kappa_1}{1-\kappa_2}.$$

Legendre Transform:

$$\mathrm{k}_\mu^*(x) = x\left[\log\left(\frac{x}{x+\lambda}\right) - 1\right]\,,\, x \in M_F$$

Notes:

- $[(\kappa_2/\kappa_1)\Gamma(1, 1/\kappa_1) - (\kappa_2/\kappa_1)]$ is the inverse distribution of $\text{GP}(\kappa_1, \kappa_2)$.
- The form of the density for the Borel-Tanner distribution is

$$\text{BT}(r, \kappa) = \frac{r x^{x-r-1}}{\Gamma(x - r + 1)} \kappa^{x-r} e^{-\kappa x}$$

for $r \in \mathbb{R}^+$, $\kappa \in (0, 1)$, $x \in \{r, r+1, \ldots\}$. (The form relates to the GP distribution through $\text{BT}(r, \kappa) = [\text{GP}(r/\kappa, \kappa) + r]$.) Note the curious formal relationship between the densities of the Borel-Tanner and the Kendall-Ressel distributions.

Other References:

- Consul (1989) devotes a complete monograph to the Generalized Poisson distribution.
- Haight and Breuer (1960) provide a survey of the Borel-Tanner distribution, including tables of probability.

2.2 Generalized Negative Binomial (Takács) Families

The Generalized Negative Binomial distribution, denoted $\mathrm{GNB}(r_1, r_2, p)$, was introduced in restricted form by Takács (1962), Mohanty (1966), and generalized to its present form by Jain and Consul (1973), as a probability distribution involved in queueing theory and, more generally, as a discrete distribution the variance of which increases with the mean at a rate which may differ from that of the Pascal.

The Generalized Negative Binomial family is also called *Fluctuation type*, after Feller, by Mora (1986) and *Takács type* by Letac and Mora (1990).

Variance Function:

$$\mathbf{V}_F(m) = \frac{m}{\lambda^2}[(a+1)m+\lambda][am+\lambda]$$
where $a > 0$
$M_F = \mathbb{R}^+$, $\Lambda(F) = \mathbb{R}^+$

Basis Measure:

$$\mu(dx) = \sum_{j=0}^{\infty} \frac{\lambda}{j!} \frac{\Gamma(\lambda+(a+1)j)}{\Gamma(\lambda+aj+1)} \delta_j(dx)$$

Cumulant Transform and Mean Domain Mapping:

$$\mathrm{k}_\mu(\theta) = \lambda \log(1+f^{-1}(e^\theta)),\ \Theta(\mu) = \left(-\infty, \log\left[\frac{a^a}{(a+1)^{a+1}}\right]\right)$$
where f^{-1} is the reciprocal of $f:(0,1/a) \to \left(0, a^a/(a+1)^{a+1}\right)$, $t \mapsto \dfrac{t}{(1+t)^{a+1}}$
$$\phi_\mu(m) = \log\left(\frac{m[am+\lambda]^a}{[(a+1)m+\lambda]^{a+1}}\right),\ m \in M_F$$

Density:

$$f_X(x) = \frac{r_1\Gamma((r_2+1)x+r_1)}{x!\Gamma(r_2x+r_1+1)} p^x(1-p)^{r_1+r_2x},\ x \in \mathbb{N}\ ;$$
$\theta = \log(p(1-p)^{r_2})$, $p \in (0, 1/(r_2+1))$; $\lambda = r_1 \in \mathbb{R}^+$; $a = r_2$, $r_2 \in (0,+\infty)$;
then $m = r_1 \dfrac{p}{1-(r_2+1)p}$

Legendre Transform:

$$\mathrm{k}^*_\mu(x) = \log\left(\frac{x^x[(a-1)x+\lambda]^{(a-1)x+\lambda}}{[ax+\lambda]^{ax+\lambda}}\right),\ x \in M_F$$

Asymptotics:

$$\mathrm{GNB}(r_1, r_2, p) \xrightarrow{\mathcal{D}} \mathrm{NB}(r_1, p) \text{ as } r_2 \to 0.$$

Notes:

- $\left[\frac{1}{r_1}\mathrm{NB}\left(1, \frac{p}{1+p}\right) + \frac{r_2}{r_1}\right]$ is the inverse distribution of $\mathrm{GNB}(r_1, r_2, p)$.

2.3 Arcsine Families

The Arcsine type, which we denote $\mathrm{Arc}(r, a, \xi), a \geq 0$, was introduced by Mora (1986) as one the five NEF's with cubic variances. Seshadri (private communication) has discovered that the Arcsine families represent the distribution of time to vanishing of a queue with binomial input and deterministic output.

The Arcsine type was partitioned into two distinct classes by Letac and Mora (1990), corresponding to $\mathrm{Arc}(\lambda, 0, \xi)$ ("Strict Arcsine") and $\mathrm{Arc}(\lambda, a > 0, \xi)$ ("Extended" or "Large Arcsine"). The reparametrization with respect to $\lambda \in \Lambda(F)$ allows us to construct the two types simultaneously with a single generating measure.

Variance Function:

$$\mathbf{V}_F(m) = \frac{m}{\lambda^2}\left[(am+\lambda)^2 + m^2\right]$$
where $a \in \mathbb{R}^+ \cup \{0\}$
$M_F = \mathbb{R}^+ \,,\, \Lambda(F) = \mathbb{R}^+$

Basis Measure:

$\mu = \nu^{*\lambda}$
where $\nu = \sum_{j=0}^{\infty} \frac{p_j(aj+1)}{aj+1}\frac{\delta_j}{j!}$;
with $p_{2j}(t) = \prod_{i=0}^{j-1}\left(t^2+4i^2\right)$ and $p_{2j+1}(t) = t\prod_{i=0}^{j-1}\left(t^2+(2i+1)^2\right)$

Cumulant Transform and Mean Domain Mapping:

$\mathrm{k}_\mu(\theta) = \lambda \arcsin(f^{-1}(e^\theta)) \,,\, \Theta(\mu) = \left(-\infty, -\log\sqrt{1+a^2} - a\arctan a^{-1}\right)$
where f^{-1} is the reciprocal of
$f : \left(0, \sqrt{1+a^2}\right) \to \left(0, \sqrt{1+a^2}\exp(a\arctan a^{-1})\right) \,,\, t \mapsto \frac{t}{\exp(a\arcsin t)}$
$\phi_\mu(m) = \log\left(\sin\arctan\left[\frac{m}{am+\lambda}\right]\right) - a\arctan\left[\frac{m}{am+\lambda}\right] \,,\, m \in M_F$

Density:

$f_X(x) = \frac{r}{x!}\frac{p_x(ax+r)}{ax+r}\frac{\sin^x\xi}{\xi^r}e^{-a\xi x} \,,\, x \in \mathbb{N}$;
where p_x is as in the Measure section above;
$\theta = \log(\sin\xi) - a\xi \,,\, \xi \in (0, \arctan a^{-1})$; $\lambda = r \,,\, r \in \mathbb{R}^+ \,,\, a \in \mathbb{R}^+ \cup \{0\}$;
then $m = \frac{r}{\cot\xi - a}$

Legendre Transform:

$$k_\mu^*(x) = x \log\left(\sin \arctan \frac{x}{ax+\lambda}\right) - (ax+\lambda) \arctan \frac{x}{ax+\lambda}, \; x \in M_F$$

Asymptotics:

$\mathrm{Arc}(r, a, \xi) \xrightarrow{\mathcal{D}} \mathrm{Po}(m)$ as $\lambda \to +\infty$ while m as in the Density section above remains constant.

Notes:

- $[(1/r)\mathrm{GHS}(1, \frac{\pi}{2} - \xi) - a/r]$ is the inverse distribution of $\mathrm{Arc}(r, a, \xi)$.

2.4 Kendall-Ressel Families

The Kendall-Ressel distribution, which we denote $\mathrm{KR}(r,\xi)$, arises naturally as the first-passage time distribution of a Gamma process (Letac and Mora [1990]). Kendall (1957) observed this and derived the density function of this hitting time in the context of storage theory. Ressel, in an unpublished result reported by Mora (1986), demonstrates that the exponentiated reciprocal of $f : \mathbb{R}^+ \to \mathbb{R}^+ , t \mapsto e^t - t - 1$ is the Laplace transform of a positive measure by showing its complete monotonicity (see Feller [1971], XIII.4).

Variance Function:

$$\mathbf{V}_F(m) = \frac{m^2}{\lambda^2}(m+\lambda),$$
$$M_F = \mathbb{R}^+ , \Lambda(F) = \mathbb{R}^+$$

Basis Measure:

$$\mu = \nu^{*\lambda}$$
$$\text{where } \nu = \frac{x^x e^{-x}}{\Gamma(x+2)} \mathbb{1}_{\mathbb{R}^+}(x)$$

Cumulant Transform and Mean Domain Mapping:

$$\mathrm{k}_\mu(\theta) = -\lambda f^{-1}(-\theta), \Theta(\mu) = \mathbb{R}^-$$
where f^{-1} is the reciprocal function of $f : \mathbb{R}^+ \to \mathbb{R}^+ , t \mapsto e^t - t - 1$
$$\phi_\mu(m) = 1 + \log\left(\frac{m+p}{m}\right) - \left(\frac{m+p}{m}\right), m \in M_F$$

Density:

$$f_X(x) = \frac{r x^{x+r-1}}{\Gamma(x+r+1)} \xi^{x+r} e^{-\xi x}, x \in \mathbb{R}^+$$
$$\theta = 1 + \log\xi - \xi , \xi \in (1, +\infty) \; ; \; \lambda = r , r \in \mathbb{R}^+ \; ; \text{ then } m = \frac{r}{\xi - 1}$$

Legendre Transform:

$$\mathrm{k}^*_\mu(x) = (x+\lambda)\log\left(\frac{x+\lambda}{x}\right) - \lambda, x \in M_F$$

Notes:

- $[(1/r)\mathrm{Po}(\xi) - 1/r]$ is the inverse distribution of $\mathrm{KR}(r,\xi)$.

2.5 Inverse Gaussian Families

The Inverse Gaussian distribution, denoted IG(χ, ψ), is the first-passage time distribution of a Wiener process. Its many interesting properties and relationship with the Normal distribution have made it a choice object of study in the 20th century. Tweedie (1956, 1957) provided the name for this distribution, even though it had been discovered on several occasions since the early 20th century. Seshadri (1993) provides a detailed monograph on the Inverse Gaussian families in the context of NEF theory, including a historical survey.

The IG(0, 1) distribution is a stable distribution. In general, the Inverse Gaussian also acts as a limiting distribution for all members of the Mora class of NEFs, among others.

Variance Function:

$$V_F(m) = \frac{m^3}{\lambda^2}$$
$$M_F = \mathbb{R}^+ ,\ \Lambda(F) = \mathbb{R}^+$$

Basis Measure:

$$\mu = \nu^{*\lambda}$$
$$\text{where } \nu = \frac{1}{\sqrt{2\pi x^3}} \exp\left(-\frac{1}{2x}\right) \mathbb{1}_{\mathbb{R}^+}(x)dx$$

Cumulant Transform and Mean Domain Mapping:

$$k_\mu(\theta) = -\lambda\sqrt{-2\theta} ,\ \Theta(\mu) = \mathbb{R}^-$$
$$\phi_\mu(m) = -\frac{\lambda^2}{2m^2} ,\ m \in M_F$$

Density:

$$f(x) = \frac{e^{\sqrt{\chi\psi}}}{\sqrt{2\pi x^3}}\sqrt{\chi}\, e^{-\frac{1}{2}(\chi x^{-1}+\psi x)} ,\ x \in \mathbb{R}^+$$
$$\theta = -\frac{\psi}{2} ,\ \psi \in \mathbb{R}^+ \ ;\ \lambda = \sqrt{\chi} ,\ \chi \in \mathbb{R}^+ \ ;$$
$$\text{then } m = \frac{\sqrt{\chi}}{\sqrt{\psi}}$$

Legendre Transform:

$$k^*_\mu(x) = \frac{\lambda^2}{x} ,\ x \in M_F$$

Notes:

- $N(\sqrt{\psi/\chi}, 1/\chi)$ is the inverse distribution of $IG(\chi, \psi)$.
- The Inverse Gaussian and the Normal distributions are the only univariate distributions for which the saddlepoint approximation is exact.

Other References:

- Chhikara and Folks (1989) for a monograph on the statistical properties of the Inverse Gaussian distribution.

3 Babel Class

The Babel class is defined as that class of NEFs F for which the variance function has the form

$$\mathbf{V}_F(m) = P\Delta(m) + Q(m)\sqrt{\Delta(m)}$$

with $\deg P = 0$, $\deg Q \le 1$ and $1 \le \deg \Delta \le 2$. If a family sporting a variance function of this form belongs to caste 1 or m^2, it will be seen immediately to belongs to the Morris class. Only the so-called non-degenerate castes m, $m^2 - 1$, $1 - m^2$ and $m^2 + 1$ will therefore be admitted within the Babel class. (See ch. 1, §2.6 for more details concerning castes). The Babel class was exhaustively classified by Letac (1992); convolution families are listed below by caste, up to affinity.

Caste m contains three sets of families.

Table 2.3: Babel caste m convolution families

Hermite:	$\left\{\text{Hermite}(r,\xi) \mid (\xi,r) \in \mathbb{R}^+ \times \mathbb{R}^+\right\}$, supported on $\mathbb{N}$.
Laguerre:	$\left\{\text{Laguerre}(r_1,r_2,\xi) \mid (\xi,r_1) \in \mathbb{R}^+ \times \mathbb{R}^+\right\}$ for $r_2 \in (-r_1,+\infty)$, supported on $\mathbb{N}$.
Non-Central Chi-Squared:	$\left\{\Gamma'(\alpha,\beta,\delta) \mid (\beta,\delta) \in \mathbb{R}^+ \times \mathbb{R}^+\right\}$ for $\alpha \in \mathbb{R}^+ \cup \{0\}$, supported on $\mathbb{R}^+$.

Babel caste m distributions are fairly well-documented in the statistical literature. Exponentiation is involved in the expression of the basis measure for all three types.

Caste m^2-1 contains two complex sets of convolution families indexed by two parameters.

Table 2.4: Babel caste $m^2 - 1$ convolution families

Mixed Geometric:	$\left\{\text{MG}(r_1,r_2,a,\xi) \mid (\xi,r_1) \in (0,1) \times \mathbb{R}^+\right\}$, for $(r_2,a) \in [0,+\infty) \times [-1,1) \setminus \{(0,-1)\}$, supported on $\mathbb{N}$.
Mixed Exponential:	$\left\{\text{ME}(r_1,r_2,b,\beta) \mid (\beta,r_1) \in \mathbb{R}^+ \times \mathbb{R}^+\right\}$, for $(r_2,b) \in \mathbb{R}^+ \times \mathbb{R}^+$, supported on $\mathbb{R}^+$.

These distributions are probably common in the literature, but can also be difficult to identify at a glance (see the examples for the Mixed Geometric families).

Caste 1-m^2 contains a single set of families, indexed by one parameter.

Table 2.5: Babel caste $1 - m^2$ convolution family

Trinomial:	$\{\text{Trin}(n, a, \beta) \mid (\beta, n) \in \mathbb{R} \times \mathbb{N}^*\}$, for $a \in (0, 1)$, supported on $\{-n, -n+1, \ldots, n\}$.

Caste m^2+1 is the most populous of the Babel class with 12 sets of families.

Table 2.6: Babel caste $m^2 + 1$ convolution families

Pascal Sum:	$\left\{\text{NB} + \text{NB}(r_1, r_2, p_1, p_2) \mid (p_1, r_1) \in (0,1) \times \mathbb{R}^+\right\}$, for $(p_2, r_2) \in (0,1) \times \mathbb{R}^+$, supported on $\mathbb{N}$
Pascal Difference:	$\left\{\text{NB} - \text{NB}(r_1, r_2, p_1, p_2) \mid (p_1, r_1) \in (0,1) \times \mathbb{R}^+\right\}$, for $(p_2, r_2) \in (0,1) \times \mathbb{R}^+$, supported on Z
Pascal-Binomial Sum:	$\{\text{B} + \text{NB}(n, r, p_1, p_2) \mid (p_1, n) \in (0,1) \times A\}$, for $(p_2, r) \in (0,1) \times \mathbb{R}^+$, supported on $\mathbb{N}$. A so far is still unknown.
Poisson-Pascal Sum:	$\left\{\text{P} + \text{NB}(\xi, r, p) \mid (p, r) \in (0,1) \times \mathbb{R}^+\right\}$, for $\xi \in \mathbb{R}^+$, supported on $\mathbb{N}$.
Poisson-Pascal Difference:	$\left\{\text{P} - \text{NB}(\xi, r, p) \mid (p, r) \in (0,1) \times \mathbb{R}^+\right\}$, for $\xi \in \mathbb{R}^+$, supported on Z.
Binomial Sum:	$\{\text{B} + \text{B}(n_1, n_2, p_1, p_2) \mid (p_1, n_1) \in (0,1) \times \mathbb{N}^*\}$, for $(p_2, n_2) \in (0,1) \times \mathbb{N}^*$, supported on $\mathbb{N}$.
Poisson-Binomial Sum:	$\{\text{P} + \text{B}(\xi, n, p) \mid (p, n) \in (0,1) \times \mathbb{N}^*\}$, for $\xi \in \mathbb{R}^+$, supported on $\mathbb{N}$.
Poisson Difference:	$\left\{\text{P} + \text{P}(\xi_1, \xi_2) \mid (\xi_1, \xi_2) \in \mathbb{R}^+ \times \mathbb{R}^+\right\}$, supported on Z.
Gamma Sum:	$\left\{\text{G} + \text{G}(\alpha_1, \alpha_2, \beta_1, \beta_2) \mid (\beta_1, \alpha_1) \in \mathbb{R}^+ \times \mathbb{R}^+\right\}$, for $(\beta_2, \alpha_2) \in \mathbb{R}^+ \times \mathbb{R}^+$, supported on $\mathbb{R}^+$.
Gamma Difference:	$\left\{\text{G} - \text{G}(\alpha_1, \alpha_2, \beta_1, \beta_2) \mid (\beta_1, \alpha_1) \in \mathbb{R}^+ \times \mathbb{R}^+\right\}$, for $(\beta_2, \alpha_2) \in \mathbb{R}^+ \times \mathbb{R}^+$, supported on $\mathbb{R}$.
Normal-Gamma Sum:	$\left\{\text{N} + \text{G}(\sigma^2, \alpha, \xi, \beta) \mid (\beta, \alpha) \in \mathbb{R}^+ \times \mathbb{R}^+\right\}$, for $(\xi, \sigma^2) \in \mathbb{R} \times \mathbb{R}^+$, supported on $\mathbb{R}$.
Hyperbolic Secant Sum:	$\left\{\text{H} + \text{H}(r_1, r_2, \beta_1, \beta_2) \mid (\beta_1, r_1) \in (-\pi/2, \pi/2) \times \mathbb{R}^+\right\}$, for $(\beta_2, r_2) \in (-\pi/2, \pi/2) \times \mathbb{R}^+$, supported on $\mathbb{R}$.

These families consist of distributions of sums or differences of discrete families in the Morris class, of Normal and Gamma families, and of two Generalized Hyperbolic Secant families. Since Poisson random variables are reproductive under convolution, their sum belongs to the Morris class. As well, an operation of difference can be replaced by a sum with a translated family member or with another member of the family in the following cases:

$$
\begin{aligned}
[-\mathrm{Bin}(n,p)] &\overset{\mathcal{D}}{=} [\mathrm{Bin}(n,1-p)-n] \\
[-\mathrm{N}(\xi,\sigma^2)] &\overset{\mathcal{D}}{=} \mathrm{N}(-\xi,\sigma^2) \\
[-\mathrm{GHS}(r,\beta)] &\overset{\mathcal{D}}{=} \mathrm{GHS}(r,-\beta)
\end{aligned}
$$

The operation of difference does not create new families in these situations.

3.1 Caste m — Hermite Families

The Hermite distribution, which we denote Hermite(r,ξ), arises as a generalization of the Poisson distribution. Kemp and Kemp (1965) propose several probabilistic interpretations of the Hermite random variable:

(1) As a special case of the Poisson Binomial $\text{Po}(J\xi)\wedge_J \text{Bin}(2,r)$.

(2) As the distribution of the sum of two correlated Poisson random variables (being an application of case (3) below, in accordance with Ahmed [1961], section 2).

(3) As the sum of a Poisson random variable and a so-called Poisson doublet: if $X\sim\text{Po}(\xi_1)$ and $Y\sim\text{Po}(\xi_2)$ with X and Y independent, then

$$X+2Y\sim\text{Hermite}\left(\frac{1}{2}(\xi_1^2/\xi_2^2),\xi_2\right).$$

Variance Function:

$$\mathbf{V}_F(m)=\frac{1}{2\lambda}\left(\Delta(m)-\lambda\sqrt{\Delta(m)}\right)$$
$$\Delta(m)=4\lambda m+\lambda^2$$
$$M_F=\mathbb{R}^+,\ \Lambda(F)=\mathbb{R}^+$$

Basis Measure:

$$\mu=\exp(\lambda\nu)$$
where $\nu=\delta_1+\delta_2/2$

Cumulant Transform and Mean Domain Mapping:

$$\mathrm{k}_\mu(\theta)=\lambda\left(e^\theta+\frac{e^{2\theta}}{2}\right),\ \Theta(\mu)=\mathbb{R}$$
$$\phi_\mu(m)=\log\left(\frac{\sqrt{\Delta(m)}-\lambda}{2\lambda}\right),\ m\in M_F$$

Density:

$$f_X(x)=\frac{(r\xi)^x}{x!}\exp\left(-r^2\xi[\xi+2]\right)H_x^*(r),\ x\in\mathbb{N}$$
where $H_j^*(x)=i^{-j}H_j(ix)$, H_j the j^{th} Hermite polynomial.
$\theta=\log\xi,\ \xi\in\mathbb{R}^+$; $\lambda=2r^2,\ p\in\mathbb{R}^+$; then $m=2r^2\xi(1+\xi)$.

Legendre Transform:

$$\mathrm{k}_\mu^*(x)=x\log\left(\frac{\sqrt{\Delta(x)}-\lambda}{2\lambda}\right)-\frac{1}{4}\left(2x-\lambda+\sqrt{\Delta(x)}\right)$$

Asymptotics:

$$\text{Hermite}(r,\xi)\xrightarrow{\mathcal{D}}[2\text{Po}(m)] \text{ as } r \to 0 \text{ while } m = 2r^2\xi(1+\xi) \text{ remains constant.}$$

Notes:

- Unlike Kemp and Kemp (1965) or Kendall and Stuart (1953), we are using the classical definition of the Hermite polynomial (see, for instance, Szegö [1939]), so that
$$\sum_{j=0}^{\infty} H_j(a)\frac{z^j}{j!} = \exp\left(2az - z^2\right)$$
- The polynomials $H_j^*(z)$, $j \in \mathbb{N}$ above are simply polynomials with coefficients equal to the absolute values of the coefficients of the Hermite polynomials $H_j(z)$.
- Letac (1992) starts with the basis measure $\mu' = \lambda^{-\frac{x}{2}}\mu$ so that $k_{\mu'}(\theta) = \sqrt{\lambda}e^{\theta} + e^{2\theta}/2$. In general, if $k_\nu(\theta) = \exp(a_1 e^\theta + a_2 e^{2\theta}/2)$, $\theta \in \mathbb{R}$, then we get $F(\nu) = F(\mu')$, taking $\lambda = a_1^2/a_2$.
- Letac (1992) shows how the Hermite type can be thought of as a limiting case of the Laguerre type.
- See appendix (§B.5) for a derivation of the basis measure for arbitrary power λ.

Other References:

- McKendrick (1926) for an early application of the Hermite distribution.
- Jain (1983) for a general discussion of the Hermite distribution.
- Watson (1988) for another approach to the derivation.

3.2 Caste m — Laguerre Families

The Laguerre distribution, denoted Laguerre(r_1, r_2, ξ), arises in practice in quantum theory. Gurland et al. (1983) describe the Laguerre distribution, which Letac (1992) generalizes by allowing $r_2 \geq -r_1$ instead of $r_2 > 0$.

For $a \geq 0$ and $\xi > 0$, if $X \sim \mathrm{NB}(ra, (1+\xi)^{-1})$ (assuming a degenerate distribution at 0 if $a = 0$) and $S \sim [\mathrm{NB}(N, (1+\xi)^{-1}) + N] \wedge_N \mathrm{Po}(r\xi^{-1})$, then $(X + S) \sim \mathrm{Laguerre}(r, a, \xi)$. It is easy with this interpretation to see that the Laguerre type is a generalization of the so-called *Poisson-Pascal* distribution (given by $\mathrm{NB}(N, p) \wedge_N \mathrm{Po}(\cdot)$) (see, for instance, Johnson and Kotz [1969]).

Variance Function:

$$\mathbf{V}_F(m) = \frac{1}{2\lambda}\left(-q\Delta(m) + [2m + \lambda q(1+q)]\sqrt{\Delta(m)}\right)$$
$$\Delta(m) = 4\lambda m + \lambda^2(q+1)^2$$
where $q \geq -1$
$$M_F = \mathbb{R}^+,\ \Lambda(F) = \mathbb{R}^+$$

Basis Measure:

$$\mu = (\delta_0 - \delta_1)^{*(-\lambda q)} * \exp\left(\lambda\left[\delta_1 * (\delta_0 - \delta_1)^{*(-1)}\right]\right)$$

Cumulant Transform and Mean Domain Mapping:

$$\mathrm{k}_\mu(\theta) = -\lambda q \log(1 - e^\theta) + \frac{\lambda e^\theta}{1 - e^\theta},\ \Theta(\mu) = \mathbb{R}^-$$
$$\phi_\mu(m) = \log\left(\frac{2m + \lambda(1+q) - \sqrt{\Delta(m)}}{2(m + \lambda q)}\right),\ m \in M_F$$

Density:

$$f_X(x) = \mathrm{L}_x^{r_2-1}(-r_1)\, e^{-r_1\xi} \frac{\xi^x}{(1+\xi)^{x+r_2}},\ x \in \mathbb{N}$$
where $\mathrm{L}_j^\alpha(x),\ \alpha - x \geq 0,\ j \in \mathbb{N}$,
is the j^{th} Generalized Laguerre polynomial in $x \in \mathbb{R}$ of order α.
$$\theta = \log\left(\frac{\xi}{1+\xi}\right),\ \xi \in \mathbb{R}^+\ ;\ \lambda = r_1,\ r_1 \in \mathbb{R}^+\ ;\ q = r_2/r_1\ ;\ r_2 \in (-r_1, +\infty)\ ;$$
then $m = r_1\xi(1+\xi) + r_2\xi$

Legendre Transform:

$$\mathrm{k}_{\mu}^{*}(x) = \frac{1}{2}\left(\lambda(q+1) - \sqrt{\Delta(x)}\right) + \log\left(\frac{\left[2m + \lambda(q+1) - \sqrt{\Delta(x)}\right]^{x}\left[\lambda(q-1) + \sqrt{\Delta(x)}\right]^{\lambda q}}{\left[2(x+\lambda q)\right]^{x+\lambda q}}\right), x \in M_F$$

Other References:

- Helstrom (1976) for an application to quantum theory.
- Watson (1988) for another derivation of the distribution.
- Szegö (1939) for background material concerning generalized Laguerre polynomials.

3.3 Caste m — Non-central Chi-Squared Families

The Non-central Chi-squared distribution, which we will denote $\Gamma'(\alpha,\beta,\delta)$, is in fact a generalization of the distribution of the same name introduced by Fisher (1928), the properties of which are summarized in Johnson and Kotz (1970). The less general version of the distribution is denoted $\chi'^2_\nu(\delta)$ in the latter work, corresponding to $\Gamma'(\nu/2, 2, \delta/2)$ in our notation.

For $\alpha,\beta,\delta > 0$, if $X\sim\Gamma(\alpha,\beta)$ and $S\sim\Gamma(J,\beta)\wedge_J \mathrm{Po}(\delta)$, then $X+S\sim\Gamma'(\alpha,\beta,\delta)$.

A typical application arises if $X_i\sim\mathrm{N}(\xi_i,1)$, $i=1,\ldots,n$ and $S=\frac{1}{n}\sum_{i=1}^n(X_i-\bar{X})^2$, for then $S\sim\chi'^2_\nu\left(\sum_{i=1}^n(\xi_i-\bar{\xi})^2\right)$, where $\bar{X}=n^{-1}\sum_{i=0}^n X_i$ and $\bar{\xi}=n^{-1}\sum_{i=0}^n\xi_i$.

Variance Function:

$$\mathbf{V}_F(m)=\frac{1}{2\lambda}\left(-q\Delta(m)+[2m+\lambda q^2]\sqrt{\Delta(m)}\right)N$$
$$\Delta(m)=4\lambda m+\lambda^2q^2\,,\; q\geq 0$$
$$M_F=\mathbb{R}^+\,,\;\Lambda(F)=\mathbb{R}^+$$

Basis Measure:

$$\mu=\gamma^{*\lambda q}*\exp\lambda\gamma$$
where $\gamma=\mathbb{1}_{\mathbb{R}^+}(x)$

Cumulant Transform and Mean Domain Mapping:

$$\mathrm{k}_\mu(\theta)=\frac{\lambda}{-\theta}-\lambda q\log(-\theta)\,,\;\Theta(\mu)=\mathbb{R}^-$$
$$\phi_\mu(m)=\frac{-\lambda q-\sqrt{\Delta(m)}}{2m}\,,\; m\in M_F$$

Density:

$$f_X(x)=x^{\alpha-1}e^{-(\delta\beta+x/\beta)}\frac{1}{\beta^\alpha}\sum_{j=0}^{\infty}\frac{(\delta x)^j}{\Gamma(j+1)\Gamma(\alpha+j)}\,,\; x\in\mathbb{R}^+$$
$$\theta=-\frac{1}{\beta}\,,\;\beta\in\mathbb{R}^+\;;\;\lambda=\delta\,,\;\delta\in\mathbb{R}^+\;;\;q=\frac{\alpha}{\delta}\,,\;\alpha\in\mathbb{R}^+\cup\{0\}$$
then $m=\beta(\alpha+\beta\delta)$.

Legendre Transform:

$$\mathrm{k}^*_\mu(x)=\lambda q\log\left(\frac{\lambda q+\sqrt{\Delta(x)}}{2x}\right)-\sqrt{\Delta(x)}\,,\; x\in M_F$$

Notes:

- The density of the Non-central Chi-squared distribution $\chi'^2_\nu(\delta) = \Gamma'(\nu/2, 2, \delta/2)$ can be expressed in terms of elementary functions – see Johnson and Kotz (1970) for results and references.

- $\Gamma'(0, \beta, \delta) = \Gamma(J, \beta) \wedge_J \mathrm{Po}(\delta)$ is a member of the Tweedie scale.

3.4 Caste $m^2 - 1$ — Mixed Geometric Families

The Mixed Geometric type, denoted $\mathrm{MG}(r_1, r_2, a, \xi)$, with fixed $(a, r_2) \in [-1,1)\times[0,+\infty)\setminus\{-1,0\}$, and parameter $\xi \in (0,1)$, is a rich and complex type which allows for several different interpretations. Among these are:

- (Letac [1992]) The distributions with Möbius probability generating functions appearing in Harris (1960) belong to the NEF $F\left((b+1)\delta_0 + \sum_{n=1}^{\infty} \delta_n\right)$, which can be shown to correspond to the $[(b+1)\mathrm{MG}(b/(b+1), 1/b, b/(b+1), \xi)]$, $\xi \in (0,1)$ family.

- (Letac [1992]) For fixed $a \in (0,1)$, the $[\mathrm{Bin}\,(N, \kappa_1(\xi)) + N]\wedge_N \mathrm{NB}\,(1, 1-\kappa_2(\xi))$, $\xi \in (0,1)$ family, with $\kappa_1(\xi) = a\xi/(a\xi + 1 - a)$ and $\kappa_2(\xi) = \xi(a\xi + 1 - a)$ can be shown to correspond to the family $\mathrm{MG}(a, (1-a)/a, -a, \xi)$.

- (Letac [1992]) $[2\mathrm{NB}(\lambda_1, 1-\xi)] * [\mathrm{NB}(\lambda_1\lambda_2, 1-\xi) + \lambda_1] = \mathrm{MG}(\lambda_1, \lambda_2, -1, \xi)$. This distribution is analogous to the Hermite distribution in its interpretation as the sum of a Poisson and a Poisson doublet.

- (Seshadri [1991]) The mixture distribution

  $$\frac{p\omega}{p\omega + 1}\mathrm{NB}(r,p) + \frac{1}{p\omega+1}\mathrm{NB}^*(r,p)$$

  where $\mathrm{NB}^*(r,p) = [\mathrm{NB}(r+1,p)+1]$ is the so-called length-biased distribution of $\mathrm{NB}(r,p)$, belongs to the Mixed Geometric type if $\omega > 1$. The distribution is then given by $\mathrm{MG}\,(2(p+1)/(k+1), k, 1/(1-p),)$, where k is a rather complicated expression in p and ω.

In the following we use the notation $(1-0z)^{b/0}$ to mean $\lim_{a\to 0}(1-az)^{b/a} = \exp(-bz)$, for $b \in \mathbb{R}$.

Variance Function:

$$\begin{aligned}
&\mathbf{V}_F(m) = P\Delta(m) + Q(m)\sqrt{\Delta(m)} \\
&P = \frac{1-a(1+q)}{2\lambda(1+q)}\,,\; Q(m) = \frac{(1-a)[1+a(1+q)]m - \lambda q[1-a(1+q)]}{2\lambda(1-a)(1+q)}\,, \\
&\Delta(m) = m^2 + 2\lambda\frac{q+2}{1-a}m + \left(\frac{\lambda q}{1-a}\right)^2\;; \\
&a \in (-1,1)\,, q \geq 0\,,\;\text{ or } a = -1\,, q > 0\,, \\
&M_F = \mathbb{R}^+\,, \Lambda(F) = \mathbb{R}^+
\end{aligned}$$

Basis Measure:

$$\mu = \nu_a^{*\lambda} * (\delta_0 - \delta_1)^{*(-\lambda q)}$$

where $\nu_a = \sum_{j=0}^{\infty} C_a^1(j)\delta_j$ with $C_a^1(j)$ such that

$$\sum_{j=0}^{\infty} C_a^1(j)z^j = \frac{(1-az)^{1/a}}{(1-z)}, \; a \in [-1,1)$$

Cumulant Transform and Mean Domain Mapping:

$$\mathrm{k}_\mu(\theta) = \lambda \log\left(1 - ae^\theta\right)^{1/a} - \lambda(q+1)\log\left(1-e^\theta\right), \; \Theta(\mu) = (-\infty, 0)$$

$$\phi_\mu(m) = \log\frac{1}{2}\left(\frac{\sqrt{\Delta(m)} - (1+a)m - \lambda q}{\lambda[1-a(1+q)] - am}\right), \; m \in M_F$$

Density:

$$f_X(x) = \frac{(1-\xi)^{r_1+r_2}}{(1-a\xi)^{r_1/a}}\xi^x C_a^{r_1,r_2}(x), \; x \in \mathbb{N},$$

where $\sum_{j=0}^{\infty} C_a^{b,c}(j)z^j = \frac{(1-bz)^{c/b}}{(1-z)^{c+d}}$ for $b \in [-1,1)$, $c \in \mathbb{R}^+$, $d \in \mathbb{R}^+ \cup \{0\}$, $z \in (0,1)$

$\theta = \log\xi$, $\xi \in (0,1)$; $\lambda = r_1$, $r_1 \in \mathbb{R}^+$; $q = r_2/r_1$, $r_2 \in \mathbb{R}^+ \cup \{0\}$;

$a \in [-1,1)$ such that $(a, r_2) \neq (-1, 0)$; then

$$m = r_1\xi\left(\frac{r_2}{1-\xi} - \frac{1}{1-a\xi}\right)$$

Legendre Transform:

$$\mathrm{k}_\mu^*(x) = \log\left(\frac{A(x)^x\,[1-A(x)]^{\lambda(q+1)}}{[1-aA(x)]^{\lambda/a}}\right)$$

where $A(x) = \dfrac{(1+a)x + \lambda q - \sqrt{\Delta(x)}}{2\,(ax + \lambda[a(q+1)-1])}$.

Asymptotics:

$\mathrm{MG}(r_1, r_2, a, \xi) \xrightarrow{\mathcal{D}} \mathrm{Po}(m)$ as $r_1 \to +\infty$ or $r_2 \to +\infty$ while m as in the Density section above remains fixed.

$\mathrm{MG}(r_1, r_2, -1, \xi) \xrightarrow{\mathcal{D}} [2\mathrm{NB}(r_1, 1-\xi)]$ as $r_2 \to 0$.

$\mathrm{MG}(r_1, r_2, a, \xi) \xrightarrow{\mathcal{D}} \mathrm{NB}(r_1 q, 1-\xi)$ as $a \to 1$.

Notes:

- The combination $(a, r_2) = (-1, 0)$ is disallowed simply because in such a case the basis measure $\mu = \delta_0$.
- If $\frac{r_1}{r_1 + r_2} < a < 1$, we can get an explicit form for the density:

$$f_X(x) = F(-x, -r_1/a; 1 - r_1 - r_2 - x, a) \binom{r_1 + r_2}{x} \frac{(1-\xi)^{r_1+r_2}}{(1-a\xi)^{r_1/a}} \xi^x , \, x \in \mathbb{N}.$$

where F is the hypergeometric function. See appendix (§B.8) for a derivation.

3.5 Caste $m^2 - 1$ — Mixed Exponential Families

The Mixed Exponential type, which we denote $ME(\kappa, \alpha_1, \alpha_2, \beta)$ is another rich class of NEF's. It arises naturally (Seshadri [1991]) as the mixture distribution

$$\frac{\omega}{\omega + \alpha\beta}\Gamma(\alpha, \beta) + \frac{\alpha\beta}{\omega + \alpha\beta}\Gamma^*(\alpha, \beta),$$

where $\Gamma^*(\alpha, \beta) = \Gamma(\alpha + 1, \beta)$ is the length-biased distribution of $\Gamma(\alpha, \beta)$; this distribution belongs to the Mixed Exponential family $ME[\omega/\alpha, 1, \alpha, \beta]$, $\beta \in \mathbb{R}^+$.

Variance Function:

$$\mathbf{V}_F(m) = \frac{1}{2\lambda(q+1)}\left(-q\Delta(m) + \left[(q+2)m + \lambda q^2 a\right]\sqrt{\Delta(m)}\right)$$
$$\Delta(m) = m^2 + 2\lambda a(q+2)m + \lambda^2 a^2 q^2 \ ;$$
where $a, q \in \mathbb{R}^+$
$$M_F = \mathbb{R}^+ \, , \, \Lambda(F) = \mathbb{R}^+$$

Basis Measure:

$$\mu = (a\delta_0 + \gamma)^{*\lambda} * \gamma_{\lambda q}$$
where $\gamma = \mathbb{1}_{\mathbf{R}^+}(x)$

Cumulant Transform and Mean Domain Mapping:

$$\mathrm{k}_\mu(\theta) = \lambda \log\left(a + \frac{1}{-\theta}\right) - \lambda q \log(-\theta) \, , \, \Theta(\mu) = \mathbb{R}^-$$
$$\phi_\mu(m) = -\frac{1}{2am}\sqrt{\Delta(m)} - m + \lambda q a \, , \, m \in M_F$$

Density:

$$fx(x) = \frac{\beta^{\alpha_1+\alpha_2}}{(\beta+\kappa)^{\alpha_1}}\frac{x^{\alpha_2-1}}{\Gamma(r_2)}e^{-(\beta+\kappa)x}{}_1F_1[\alpha_1 + \alpha_2; \alpha_2; \kappa x] \, , \, x \in \mathbb{R}^+,$$
where ${}_1F_1$ is the confluent hypergeometric function;
$$\theta = -\frac{1}{\beta}\, , \, \beta \in \mathbb{R}^+ \ ; \ \lambda = \alpha_1 \, , \, \alpha_1 \in \mathbb{R}^+ \ ; \ q = \alpha_2/\alpha_1 \, , \, \alpha_2 \in \mathbb{R}^+ \ ;$$
$$a = \kappa^{-1} \, , \, \kappa \in \mathbb{R}^+ \ ;$$
then $m = \beta\left(\frac{\alpha_1\kappa\beta}{1+\kappa\xi} + \frac{\alpha_2}{\alpha_1}\right)$

Legendre Transform:

$$\mathrm{k}_{\mu}^{*}(x) = \lambda \log \left(\frac{\left[\lambda q a - x + \sqrt{\Delta(x)}\right]^{\lambda(q+1)}}{[2ax]^{\lambda q}\left[\lambda q a^2 + ax + a\sqrt{\Delta(x)}\right]^{\lambda}} \right) - \frac{\Delta(x)}{2a}, \ x \in M_F$$

Asymptotics:

$$\mathrm{ME}(\kappa, \alpha_1, \alpha_2, \beta) \xrightarrow{\mathcal{D}} \Gamma\left(\alpha_1 + \frac{\alpha_2}{\alpha_1}, \beta\right) \text{ as } \kappa \to +\infty.$$

$$\mathrm{ME}(\kappa, \alpha_1, \alpha_2, \beta) \xrightarrow{\mathcal{D}} \Gamma\left(\frac{\alpha_2}{\alpha_1}, \beta\right) \text{ as } \kappa \to 0.$$

Notes:

- See appendix (§B.9) for a derivation of the Mixed Exponential density.

3.6 Caste $1 - m^2$ — Trinomial Families

The Trinomial type, which we denote Trin(n, a, β), is a generalization of the Binomial type, to which it converges, up to an affinity, as a approaches the boundaries of $(0, 1)$. The Trinomial type is not infinitely divisible.

The Trinomial type was apparently introduced initially by Jørgensen, Letac and Seshadri (1989) as a counter-example to a property of quadratic variance function NEF's. It is the only Babel type in caste $1 - m^2$.

Variance Function:

$$\mathbf{V}_F(m) = \frac{1}{\lambda(1-a^2)}\left(\Delta(m) - a\sqrt{\Delta(m)}\right)$$
$$\Delta(m) = \lambda^2 - (1-a^2)m^2,$$
where $a \in (0,1)$
$$M_F = (-\lambda, \lambda), \Lambda(F) = \mathbb{N}^*$$

Basis Measure:

$$\mu = (\delta_{-1} + 2a\delta_0 + \delta_1)^{*\lambda}$$

Cumulant Transform and Mean Domain Mapping:

$$\mathrm{k}_\mu(\theta) = \lambda \log(2\cosh(\theta) + 2a), \Theta(\mu) = \mathbb{R}$$
$$\phi_\mu(m) = \log\left(\frac{am + \sqrt{\lambda^2 - (1-a^2)m^2}}{\lambda - m}\right), m \in M_F$$

Density:

$$f_X(x) = \left(\frac{a}{a + \cosh\beta}\right)^n \left[\sum_{i=|x|,2}^{n} \binom{n}{i}\binom{i}{\frac{x+i}{2}}(2a)^{-i}\right] e^{\beta x}, x \in \{-n, -n+1, \ldots, n\}$$

$\theta = \beta, \beta \in \mathbb{R}$; $\lambda = n, n \in \mathbb{N}^*$; $a \in (0,1)$;

then $m = n\dfrac{\sinh\beta}{\cosh\beta + a}$

Legendre Transform:

$$k^*_\mu(x) = \log\left(\frac{A(x)^{x+\lambda}}{[A(x)^2 + 2aA(x) + 1]^\lambda}\right)$$

$$\text{where } A(x) = \frac{ax + \Delta(x)}{\lambda - x},\ x \in M_F$$

Asymptotics:

$\text{Trin}(n, a, \beta) \xrightarrow{\mathcal{D}} [2\text{Bin}\left(n, e^\beta/(1+e^\beta)\right) - n]$ as $a \to 0$.

$\text{Trin}(n, a, \beta) \xrightarrow{\mathcal{D}} [\text{Bin}\left(2n, e^\beta/(1+e^\beta)\right) - n]$ as $a \to 1$.

$[\text{Trin}(n, a, \beta) + n] \xrightarrow{\mathcal{D}} \text{Po}(e^\beta)$ as $n \to +\infty$.

Notes:

- See appendix (§B.10) for a derivation of the Trinomial density.

Other References:

- Jørgensen (1986) for a discussion of convergence of univariate discrete distributions to the Poisson distribution.

3.7 Caste $1+m^2$ — Pascal Sum Families

The $\mathrm{NB}+\mathrm{NB}(r_1,r_2,p_1,p_2)$ rules the sum of two Pascal random variables, viz.

$$\mathrm{NB}+\mathrm{NB}(r_1,r_2,p_1,p_2)=\mathrm{NB}(r_1,p_1)*\mathrm{NB}(r_2,p_2)\ .$$

Variance Function:

$\mathbf{V}_F(m)=P\Delta(m)+Q(m)\sqrt{\Delta(m)}$, where

$P=\dfrac{q+1}{2\lambda q}\,,\,Q(m)=\dfrac{1}{2q}\left(\dfrac{q-1}{\lambda}m-[q+1]\dfrac{a^2+q}{a^2-1}\right)$

$\Delta(m)=m^2+2\lambda\dfrac{a^2-q}{a^2-1}m+\lambda^2\dfrac{(a^2+q)^2}{(a^2-1)^2}$;

where $q\in\mathbb{R}^+$, $a\in\mathbb{R}^+$

$M_F=\mathbb{R}^+$, $\Lambda(F)=\mathbb{R}^+$

Basis Measure:

$$\mu=a^x(\delta_0-\delta_1)^{*(-\lambda)}*a^{-x}(\delta_0-\delta_1)^{*(-\lambda q)}$$

Cumulant Transform and Mean Domain Mapping:

$\mathrm{k}_\mu(\theta)=-\lambda\log(1-ae^\theta)-\lambda q\log(1-e^\theta/a)\,,\,\Theta(\mu)=(-\infty,-\log a)$

$\phi_\mu(m)=\log\left(\dfrac{1}{2a}\left[a^2+1-\dfrac{(a^2-1)\sqrt{\Delta(m)}+\lambda(a^2q+1)}{m+\lambda(q+1)}\right]\right)$, $m\in M_F$

Density:

$$f_X(x)=p_1^{r_1}p_2^{r_2}\binom{r_1+r_2+x-1}{x}(1-p_2)^xF\left(-x;r_1;r_1+r_2;\frac{p_1-p_2}{1-p_2}\right)\ ,x\in\mathbb{N}$$

where F is the hypergeometric function;

$\theta=\dfrac{1}{2}\log[(1-p_1)(1-p_2)]$; $a=\sqrt{(1-p_1)/(1-p_2)}\,,p_1,p_2\in(0,1)$;

$\lambda=r_1\,,r_1\in\mathbb{R}^+$; $q=r_2/r_1\,,r_2\in\mathbb{R}^+$;

then $m=r_1\dfrac{1-p_1}{p_1}+r_2\dfrac{1-p_2}{p_2}$.

Legendre Transform:

$$\mathrm{k}_\mu^*(x) = \log\left(\frac{[(a^2+1)B(x)-C(x)]^x}{[2aB(x)]^{B(x)}\,a^{\lambda(q-1)}}\right) + \log\left(\left[(1-a^2)B(x)+C(x)\right]^\lambda\left[(a^2-1)B(x)+C(x)\right]^{\lambda q}\right),$$

where $B(x) = x + \lambda(q+1)$ and $C(x) = \lambda(aq^2+1) + (a^2-1)\sqrt{\Delta(x)}$

Asymptotics:

$\mathrm{NB}+\mathrm{NB}(p_1,p_2,r_1,r_2) \xrightarrow{\mathcal{D}} \mathrm{NB}(r_1,p_1)$ as $p_2 \to 1$ or $r_2 \to 0$.

$\mathrm{NB}+\mathrm{NB}(p_1,p_2,r_1,r_2) \xrightarrow{\mathcal{D}} \mathrm{NB}(r_2,p_2)$ as $p_1 \to 1$ or $r_1 \to 0$.

$\mathrm{NB}+\mathrm{NB}(p_1,p_2,r_1,r_2) \xrightarrow{\mathcal{D}} \mathrm{Po}(m)$ as r_1 or $r_2 \to +\infty$ while m as in the Density section above remains constant.

Notes:

- Watson (1988) produces distributions based on the Legendre and on the Gegenbauer polynomials. The Gegenbauer polynomials $C_j^r(k)$ have generating function

$$\sum_{j=0}^{\infty} C_j^r(k) z^j = (1-2kz+z^2)^{-r}$$

  Letting $r_1 = r_2 = r$ in our parametrization of the NB+NB distribution thus yields the Gegenbauer distribution. In this case, putting $q_1 = 1-p_1$, $q_2 = 1-p_2$, $\bar{q}_A = (q_1+q_2)/2$ and $\bar{q}_G = \sqrt{q_1 q_2}$, the density can be written as

$$f_X(x) = (p_1 p_2)^r (q_1 q_2)^{x/2} C_x^r\left(\frac{\bar{q}_A}{\bar{q}_G}\right)$$

  for $x \in \mathbb{N}$. Since $C_j^1(k)$ is the Legendre polynomial, the so-called Legendre distribution occurs for $r = 1$ (Watson [1988]).

- See appendix (sec. B.11) for a derivation of the Pascal Sum density.

Other References:

- Meixner (1938 and 1941) for a definition of Meixner polynomials, on which the NB + NB density is based.
- Erdélyi (1953, sec. 10.24) for properties of the Meixner polynomials.
- Abramowitz and Stegun (1970) for properties of the Legendre and Gegenbauer polynomials.

3.8 Caste $1 + m^2$ — Pascal Difference Families

The $\mathrm{NB}-\mathrm{NB}(r_1, r_2, p_1, p_2)$ rules the difference of two Pascal random variables, viz.

$$\mathrm{NB}-\mathrm{NB}\,(r_1, r_2, p_1, p_2) = \mathrm{NB}(r_1, p_1) * [-\mathrm{NB}(r_2, p_2)] \, .$$

Variance Function:

$$\mathbf{V}_F\,(m) = P\Delta\,(m) + Q(m)\sqrt{\Delta\,(m)} \ \text{, where}$$

$$P = \frac{q+1}{2\lambda q}\,,\, Q(m) = \frac{1}{2q}\left(\frac{q-1}{\lambda}m - \frac{a^2+a^2q^2+2q}{a^2-1}\right)$$

$$\Delta\,(m) = m^2 + 2\lambda\left(\frac{a^2-q}{a^2-1} - q\right)m + \lambda^2 a^2 \frac{4q+a^2(q-1)^2}{(a^2-1)^2} \ ;$$

where $q \in \mathrm{I\!R}^+$, $a \in (0, 1)$

$M_F = \mathrm{I\!R}\,,\, \Lambda\,(F) = \mathrm{I\!R}^+$

Basis Measure:

$$\mu = a^x(\delta_0 - \delta_1)^{*(-\lambda)} * a^{-x}(\delta_0 - \delta_{-1})^{*(-\lambda q)}$$

Cumulant Transform and Mean Domain Mapping:

$$\mathrm{k}_\mu\,(\theta) = -\lambda\log(1-ae^\theta) - \lambda q\log(1-ae^{-\theta})\,,\, \Theta\,(\mu) = (\log a, -\log a)$$

$$\phi_\mu\,(m) = \log\left(\frac{(a^2+1)m - \lambda a^2(q-1) - (a^2-1)\sqrt{\Delta\,(m)}}{2a(m+\lambda)}\right) \ ,\ m \in M_F$$

Density:

$$f_X(x) = p_1^{r_1}p_2^{r_2}\ {}_2A^{r_1,r_2}_{a(p_1,p_2)}(x)\left(\frac{1-p_1}{1-p_2}\right)^{x/2} \ ,\ x \in \mathrm{Z}$$

where $a(p_1, p_2) = \sqrt{(1-p_1)(1-p_2)}$ and

where $\sum_{j=-\infty}^{\infty} {}_2A^{b,c}_k(j)z^j = (1-kz)^{-b}\left(1-\frac{k}{z}\right)^{-c}$, $b, c \in \mathrm{I\!R}^+$, $0 < |z| < k < 1$

$\theta = \frac{1}{2}\log\left(\frac{1-p_1}{1-p_2}\right)$, $p_1, p_2 \in (0,1)$; $a = a(p_1, p_2)$;

$\lambda = r_1$, $r_1 \in \mathrm{I\!R}^+$; $q = r_2/r_1$, $r_2 \in \mathrm{I\!R}^+$;

then $m = r_1\frac{1-p_1}{p_1} - r_2\frac{1-p_2}{p_2}$

Legendre Transform:

$$\mathrm{k}_{\mu}^{*}(x) = \log\left(\frac{[2\lambda + B(x) + C(x)]^{\lambda}[C(x) - B(x)]^{\lambda q}}{a^{x}[2(x+\lambda)]^{\lambda+x}[2x - B(x) - C(x)]^{\lambda q - x}}\right),\ x \in M_F$$

where $B(x) = a^2\lambda q + (a^2 - 1)\sqrt{\Delta(x)}$; $C(x) = x - a^2(x + \lambda)$

Asymptotics:

With m as in the Density section above remaining constant:

$$\mathrm{NB} - \mathrm{NB}(r_1, r_2, p_1, p_2) \xrightarrow{\mathcal{D}} \mathrm{NB}\left(r_1, \frac{r_1}{r_1 + m}\right) \text{ if } m > 0 \text{ as } (1 - p_1)(1 - p_2) \to 0.$$

$$\mathrm{NB} - \mathrm{NB}(r_1, r_2, p_1, p_2) \xrightarrow{\mathcal{D}} \left[-\mathrm{NB}\left(r_2, \frac{r_2}{r_2 - m}\right)\right] \text{ if } m < 0 \text{ as } (1 - p_1)(1 - p_2) \to 0.$$

$$\mathrm{NB} - \mathrm{NB}(r_1, r_2, p_1, p_2) \xrightarrow{\mathcal{D}} [-\mathrm{Po}(-m) + r_1] \text{ as } r_2 \to +\infty.$$

$$\mathrm{NB} - \mathrm{NB}(r_1, r_2, p_1, p_2) \xrightarrow{\mathcal{D}} [\mathrm{Po}(m) - r_2] \text{ as } r_1 \to +\infty.$$

3.9 Caste $1 + m^2$ — Pascal-Binomial Sum Families

The B + NB distributions rule the sum of a Binomial and a Pascal random variables, viz.

$$\mathrm{B+NB}(n, r, p_1, p_2) = \mathrm{Bin}(n, p_1) * \mathrm{NB}(r, p_2)$$

Variance Function:

$$\mathbf{V}_F(m) = P\Delta(m) + Q(m)\sqrt{\Delta(m)} \text{ , where}$$
$$P = \frac{1-q}{2\lambda q}, \ Q(m) = \frac{1}{2q}\left(\frac{q+1}{\lambda}m + \frac{(q-1)(a^2+q)}{a^2+1}\right)$$
$$\Delta(m) = m^2 - 2\lambda\frac{a^2-q}{a^2+1}m + \lambda^2\frac{(a^2+q)^2}{(a^2+1)^2} \ ;$$
where $q \in \mathbb{R}^+$, $a \in \mathbb{R}^+$
$M_F = \mathbb{R}^+$, $\Lambda(F)$ as yet unknown

Basis Measure:

$$\mu = a^x(\delta_0 + \delta_1)^{*\lambda} * a^{-x}(\delta_0 - \delta_1)^{*-(\lambda q)}$$

Cumulant Transform and Mean Domain Mapping:

$$\mathrm{k}_\mu(\theta) = \lambda \log(1 + ae^\theta) - \lambda q \log(1 - e^\theta/a), \ \Theta(\mu) = (-\infty, \log a)$$
$$\phi_\mu(m) = \log\left(\frac{(a^2-1)m - \lambda(a^2+q) + (a^2+1)\sqrt{\Delta(m)}}{2a\left[m + \lambda(q-1)\right]}\right)$$

Density:

$$f_X(x) = (1-p_1)^n p_2^r \, {}_3A^{n,r}_{a(p_1,p_2)}(x)\left(p_1\frac{1-p_2}{1-p_1}\right)^{x/2}, \ x \in \mathbb{N}$$
where $a(p_1, p_2) = \sqrt{(1-p_2)(1-p_1)/p_1}$ and
where $\sum_{j=0}^{\infty} {}_3A_k^{b,c}(j)z^j = \dfrac{(1+kz)^b}{(1-z/k)^c}$ and k, b, c are such that ${}_3A_k^{b,c}(j) \geq 0$, $j \in \mathbb{N}$.
$\theta = \dfrac{1}{2}\log\left(p_1\dfrac{1-p_2}{1-p_1}\right)$, $p_1, p_2 \in (0,1)$; $a = a(p_1, p_2)$;
$\lambda = n$, n with unknown range ; $q = \dfrac{r}{n}$, $r \in \mathbb{R}^+$;
then $m = np_1 + r\dfrac{1-p_2}{p_2}$

Legendre Transform:

$$k_\mu^*(x) = \log\left(\frac{[(a^2-1)B(x)-C(x)]^x[(a^2+1)B(x)+C(x)]^{\lambda q}}{a^{\lambda(q+1)}[2aB(x)]^{B(x)}[(a^2+1)B(x)-C(x)]^{\lambda}}\right),\ x \in M_F$$

where $B(x) = x + \lambda(q-1)$ and $C(x) = \lambda(a^2 q + 1) - (a^2+1)\sqrt{\Delta(x)}$

Asymptotics:

With m as in the density section above remaining constant:

$\mathrm{NB+B}(n,r,p_1,p_2) \xrightarrow{\mathcal{D}} [\mathrm{NB}(r, r/(r+m)) + n]$ as $a \to +\infty$.
$\mathrm{NB+B}(n,r,p_1,p_2) \xrightarrow{\mathcal{D}} \mathrm{NB}(r, r/(r+m))$ as $a \to 0$.
$\mathrm{NB+B}(n,r,p_1,p_2) \xrightarrow{\mathcal{D}} \mathrm{Po}(m)$ as $r_1 \to +\infty$ or $r_2 \to +\infty$.

Notes:

- Necessary and sufficient conditions on a, λ, q for $\mu\{n\} \geq 0$, $n \in \mathbb{N}$, are not yet known. Letac (1992, ch. 5) points out that the basis measure remains positive for $a = 1$ and $q = 1$, letting $\lambda \in \mathbb{R}^+$.

3.10 Caste $1 + m^2$ — Poisson-Pascal Sum Families

The $\mathrm{P} + \mathrm{NB}(\xi, r, p)$ rules the sum of a Poisson and a Pascal random variables, viz.

$$\mathrm{P} + \mathrm{NB}(\xi, r, p) = \mathrm{Po}(r\xi) * \mathrm{NB}(r, p).$$

Variance Function:

$$\mathbf{V}_F(m) = \frac{1}{2\lambda}\left(\Delta(m) + [m - \lambda(a^2+1)]\sqrt{\Delta(m)}\right)$$
$$\Delta(m) = m^2 - 2\lambda(a^2-1)m + \lambda^2(a^2+1)$$
where $a \in \mathbb{R}^+$
$$M_F = \mathbb{R}^+ \,,\, \Lambda(F) = \mathbb{R}^+$$

Basis Measure:

$$\mu = a^x \exp(\lambda\delta_1) * a^{-x}(\delta_0 - \delta_1)^{*(-\lambda)}$$

Cumulant Transform and Mean Domain Mapping:

$$\mathrm{k}_\mu(\theta) = \lambda a e^\theta - \lambda \log\left(1 - e^\theta/a\right) \; \Theta(\mu) = (-\infty, \log a)$$
$$\phi_\mu(m) = \log\left(\frac{m + \lambda(a^2-1) - \Delta(m)}{2\lambda a}\right) \,,\, m \in M_F$$

Density:

$$f_X(x) = (pe^{-\xi})^r \xi^x \left[(-1)^x L_x^{-r-x}\left(\frac{r\xi}{1-p}\right)\right] \,,\, x \in \mathbb{N}$$
where $a(\xi, p) = \sqrt{\xi/(1-p)}$
$\theta = \frac{1}{2}\log[\xi(1-p)] \,,\, \xi \in \mathbb{R}^+, p \in (0,1)$; $a = a(\xi, p)$; $\lambda = r \,,\, r \in \mathbb{R}^+$;
then $m = r\left(\xi + \frac{1-p}{p}\right)$

Legendre Transform:

$$\mathrm{k}^*_\mu(x) = \log\left(\frac{B(x)^x\,[2\lambda a^2 - B(x)]^\lambda}{(2\lambda a)^{x+\lambda} a^\lambda}\right) - \frac{1}{2}B(x) \,,\, x \in M_F$$
where $B(x) = x + \lambda(a^2+1) - \sqrt{\Delta(x)}$

Asymptotics:

With m as in the Density section above remaining constant:

$$\text{NB} + \text{P}(r, \xi, p) \xrightarrow{\mathcal{D}} \text{Po}(m) \text{ as } r \to +\infty \text{ or as } a \to +\infty.$$
$$\text{NB} + \text{P}(r, \xi, p) \xrightarrow{\mathcal{D}} \left[\text{NB}\left(r, \frac{r}{r+m}\right) - r\right] \text{ as } a(\xi, p) \to 0.$$

Notes:

- See appendix (§B.12) for a derivation of the density.

3.11 Caste $1 + m^2$ — Poisson-Pascal Difference Families

The P − NB distribution rules the difference of a Poisson and a Pascal random variables, viz.

$$\mathrm{P-NB}(\xi, r, p) = \mathrm{Po}(r\xi) * [-\mathrm{NB}(r, p)]$$

Variance Function:

$$\mathbf{V}_F(m) = \frac{1}{2\lambda}\left(\Delta(m) + [\lambda(a^2+2) - m]\sqrt{\Delta(m)}\right)$$
$$\Delta(m) = m^2 - 2\lambda a^2 m + \lambda^2 a^2(a^2+4)$$
where $a \in \mathbb{R}^+$
$$M_F = \mathbb{R}, \Lambda(F) = \mathbb{R}^+$$

Basis Measure:

$$\mu = a^x \exp(\lambda\delta_1) * a^{-x}(\delta_0 - \delta_{-1})^{*(-\lambda)}$$

Cumulant Transform and Mean Domain Mapping:

$$\mathrm{k}_\mu(\theta) = \lambda a e^\theta - \lambda\log(1 - ae^{-\theta}),\ \Theta(\mu) = (\log a, +\infty)$$
$$\phi_\mu(m) = \log\left(\frac{m + \lambda a^2 + \sqrt{\Delta(m)}}{2\lambda a}\right),\ m \in M_F$$

Density:

$$f_X(x) = (pe^{-\xi})^r{}_5A^r_{a(\xi,p)}(x)\left(\frac{\xi}{1-p}\right)^{x/2},\ x \in \mathrm{Z}$$
where $a(\xi,p) = \sqrt{\xi(1-p)}$ and
where $\sum_{j=-\infty}^{\infty} {}_5A^b_k(j)z^j = \left(e^{kz}(1-k/z)\right)^b$, $b, k \in \mathbb{R}^+$, $z > k$
$$\theta = \frac{1}{2}\log\left(\frac{\xi}{1-p}\right),\ \xi = \mathbb{R}^+,\ p \in (0,1)\ ;\ a = a(\xi,p)\ ;\ \lambda = r,\ r \in \mathbb{R}^+\ ;$$
then $m = r\left(\xi - \frac{1-p}{p}\right)$

Legendre Transform:

$$\mathrm{k}^*_\mu(x) = \log\left(\frac{B(x)^{x-\lambda}[B(x) - 2\lambda a^2]^\lambda}{[2\lambda a]^x}\right) - \frac{1}{2}B(x),\ x \in M_F$$
where $B(x) = x + \lambda a^2 + \sqrt{\Delta(x)}$

Asymptotics:

$\mathrm{P-NB}(\xi, r, p) \xrightarrow{\mathcal{D}} \mathrm{Po}(m)$ as $r \to +\infty$ while m remains constant.
$\mathrm{P-NB}(\xi, r, p) \xrightarrow{\mathcal{D}} \mathrm{Po}(m)$ as $\xi(1-p) \to 0$ while m remains constant.

3.12 Caste $1 + m^2$ — Binomial Sum Families

The $\mathrm{B+B}(n_1, n_2, p_1, p_2)$ distribution rules the sum of two Binomial random variables, viz.

$$\mathrm{B} + \mathrm{B}(n_1, n_2, p_1, p_2) = \mathrm{Bin}(n_1, p_1) * \mathrm{Bin}(n_2, p_2)$$

Variance Function:

$$\mathbf{V}_F(m) = \frac{1}{2\lambda q}\left(-[q+1]\Delta(m) + \left[(q-1)m + \lambda(q+1)\frac{a^2+q}{a^2-1}\right]\sqrt{\Delta(m)}\right)$$

$$\Delta(m) = m^2 - 2\lambda\frac{a^2-q}{a^2-1}m + \lambda^2\frac{(a^2+q)^2}{(a^2-1)^2}$$

where $q \in \{j/\lambda;\ j \in \mathbb{N}^*\}$, $a \in \mathbb{R}^+$

$M_F = (0, \lambda(q+1))$, $\Lambda(F) = \mathbb{N}^*$

Basis Measure:

$$\mu = a^x(\delta_0 + \delta_1)^{*\lambda} * a^{-x}(\delta_0 + \delta_1)^{*\lambda q}$$

Cumulant Transform and Mean Domain Mapping:

$$\mathrm{k}_\mu(\theta) = \lambda\log(1 + ae^\theta) + \lambda q\log(1 + e^\theta/a),\ \Theta(\mu) = \mathbb{R}$$

$$\phi_\mu(m) = \log\left(\frac{(a^2+1)m - \lambda(a^2+q) + (a^2-1)\sqrt{\Delta(m)}}{a[\lambda(q+1) - m]}\right),\ m \in M_F$$

Density:

$$f_X(x) = q_1^{r_1-x}q_2^{r_2-x}(p_2-p_1)^x P_x^{(n_1-x, n_2-x)}\left(\frac{p_1q_2 + p_2q_1}{p1 - p2}\right),\ x \in \mathbb{N}$$

where P is the Jacobi polynomial,

$a(p_1, p_2) = \sqrt{\frac{p_1q-2}{p_2q_1}}$ and

$\theta = \frac{1}{2}\log\left(\frac{p_1p_2}{q_1q_2}\right)$, $a = \sqrt{\frac{p_1q_2}{p_2q_1}}$, $p_1, p_2 \in (0,1)$, $q_i = 1 - p_i, i = 1, 2$;

$\lambda = n_1$, $q = n_2/n_1$, $n_1, n_2 \in \mathbb{N}^*$;

then $m = n_1p_1 + n_2p_2$

Legendre Transform:

$$k^*_\mu(x) = \log\left(\frac{[B(x) - C(x) + D(x)]^r [2aC(x)]^{C(x)} a^{\lambda(q-1)}}{[B(x) + C(x) + D(x)]^\lambda}\right)$$

$$- \log\left([2\lambda(a^2 q + 1) - B(x) - C(x) + D(x)]^{\lambda(q+1)}\right) \ , x \in M_F$$

where $B(x) = a^2 x - \lambda(a^2 - 1)$; $C(x) = \lambda(q + 1) - x$; $D(x) = (a^2 - 1)\Delta(x)$

Asymptotics:

With m as in the Density section above remaining constant:

$B + B(n_1, n_2, p_1, p_2) \xrightarrow{\mathcal{D}} Po(m)$ as $n_1 \to +\infty$ or $n_2 \to +\infty$.

$B + B(n_1, n_2, p_1, p_2) \xrightarrow{\mathcal{D}} Bin(n_1, m/n_1)$ as $\dfrac{p_1(1 - p_2)}{p_2(1 - p_1)} \to +\infty$ and n_2 varies appropriately.

$B + B(n_1, n_2, p_1, p_2) \xrightarrow{\mathcal{D}} Bin(n_2, m/n_2)$ as $\dfrac{p_1(1 - p_2)}{p_2(1 - p_1)} \to 0$ and n_1 varies appropriately.

Other References:

- Szegö (1939) for properties of the Jacobi polynomial.

3.13 Caste $1 + m^2$ — Poisson-Binomial Sum Families

The P + B distribution rules the sum of a Binomial and a Poisson random variables, viz.

$$\mathrm{P+B}(\xi, n, p) = \mathrm{Po}(n\xi) * \mathrm{Bin}(n, p)$$

Variance Function:

$$\mathbf{V}_F(m) = \frac{1}{2\lambda}\left(-\Delta(m) + [m + \lambda(a^2+1)]\sqrt{\Delta(m)}\right)$$
$$\Delta(m) = m^2 + 2\lambda(a^2-1)m + \lambda^2(a^2+1)^2$$
where $a \in \mathbb{R}^+$
$$M_F = \mathbb{R}^+ , \Lambda(F) = \mathbb{N}^*$$

Basis Measure:

$$\mu = a^x(\delta_0 + \delta_1)^{*\lambda} * a^{-x}\exp(\lambda\delta_1)$$

Cumulant Transform and Mean Domain Mapping:

$$\mathrm{k}_\mu(\theta) = \lambda\left(ae^\theta + \log(1 + e^\theta/a)\right) , \Theta(\mu) = \mathbb{R}$$
$$\phi_\mu(m) = \log\left(\frac{m - \lambda(a^2+1) + \sqrt{\Delta(m)}}{2\lambda a}\right) , m \in M_F$$

Density:

$$f_X(x) = \left(e^{-\xi}(1-p)\right)^n {}_7A^n_{a(\xi,p)}(x)\left(\frac{p\xi}{1-p}\right)^{x/2} , x \in \mathbb{N},$$
where $a(\xi, p) = \sqrt{(1-p)\xi/p}$ and

where $\sum_{j=0}^{b} {}_7A^b_k(j)z^j = [\exp(zk)(1 + z/k)]^b$, $b \in \mathbb{N}^*$, $k \in \mathbb{R}^+$, $z \in \mathbb{R}^+$

$$\theta = \frac{1}{2}\log\left(\frac{p\xi}{1-p}\right) , a = a(\xi, p) , \xi \in \mathbb{R}^+ , p \in (0,1) ;$$
$\lambda = n$, $n \in \mathbb{N}^*$;

then $m = n(\xi + p)$

Legendre Transform:

$$\mathrm{k}^*_\mu(x) = \log\left(\frac{B(x)^x a^\lambda [2\lambda a]^{\lambda - x}}{[B(x) + 2\lambda a^2]^\lambda}\right) - \frac{B(x)}{2} , x \in M_F$$
where $B(x) = x - \lambda(a^2+1) + \sqrt{\Delta(x)}$

Asymptotics:

With m as in the Density section above remaining constant:

$\mathrm{P+B}(\xi,n,p)\xrightarrow{\mathcal{D}}\mathrm{Po}(m)$ as $n\to+\infty$ or as $(1-p)\xi/p\to+\infty$.

$\mathrm{P+B}(\xi,n,p)\xrightarrow{\mathcal{D}}[\mathrm{Po}(m)+n]$ if $k=\xi+p>1$, as $p\to 1$ while n (and thus k) also remains fixed.

$\mathrm{P+B}(\xi,n,p)\xrightarrow{\mathcal{D}}\mathrm{Bin}(n,m/n)$ if $k=\xi+p<1$, as $\xi\to 0$ while n also remains fixed.

Notes:

- If $\xi<1$, from Erdélyi (1953) the density becomes

$$f_X(x)=e^{-n\xi}L_x^{(n-x)}\left(-n\xi\frac{1-p}{p}\right)p^x(1-p)^{n-x}\,,\ x\in\mathbb{N}$$

  where $L_j^{(\alpha)}(\cdot)$ is the jth Generalized Laguerre polynomial of order α.

- See appendix (§B.14) for a note on the asymptotics of the P + B families.

Other References:

- Szegö (1939) for properties of the Generalized Laguerre polynomial.

3.14 Caste $1 + m^2$ — Poisson Difference Families

The P – P distribution rules the difference of two Poisson random variables, viz.

$$\mathrm{P-P}(\xi_1, \xi_2) = \mathrm{Po}(\xi_1) * [-\mathrm{Po}(\xi_2)]$$

Variance Function:

$$\mathbf{V}_F(m) = \sqrt{\Delta(m)}$$
$$\Delta(m) = m^2 + 4\lambda^2$$
$$M_F = \mathbb{R}\,,\ \Lambda(F) = \mathbb{R}^+$$

Basis Measure:

$$\mu = \exp(\lambda\delta_1) * \exp(\lambda\delta_{-1})$$

Cumulant Transform and Mean Domain Mapping:

$$\mathrm{k}_\mu(\theta) = 2\lambda\cosh\theta\,,\ \Theta(\mu) = \mathbb{R}$$
$$\phi_\mu(m) = \log\left(\frac{m + \sqrt{\Delta(m)}}{2\lambda}\right),\ m \in M_F$$

Density:

$$f(x) = \exp\left[-\frac{1}{2}(\xi_1 + 1/\xi_2)\right] I_x\left(2\sqrt{\xi_1\xi_2}\right)(\xi_1/\xi_2)^{x/2}\ ,\ x \in \mathrm{Z}$$

where $I_x(\cdot)$ is the modified Bessel function of the first kind;

$$\theta = \frac{1}{2}\log\left(\frac{\xi_1}{\xi_2}\right)\ ;\ \lambda = \sqrt{\xi_1\xi_2}\,,\ \xi_1, \xi_2 \in \mathbb{R}^+$$

then $m = \xi_1 - \xi_2$

Legendre Transform:

$$\mathrm{k}^*_\mu(x) = x\log\left(\frac{x + \sqrt{\Delta(x)}}{2\lambda}\right) - \sqrt{\Delta(x)},\ x \in M_F$$

Asymptotics:

With m as in the Density section above remaining constant:

$\text{P} - \text{P}(\xi_1, \xi_2) \xrightarrow{\mathcal{D}} \text{Po}(m)$ as $\xi_2 \to 0$.
$\text{P} - \text{P}(\xi_1, \xi_2) \xrightarrow{\mathcal{D}} [-Po(-m)]$ as $\xi_1 \to 0$.

Notes:

- The Bessel distribution is the inverse of the P – P distribution.

Other References:

- Abramowitz and Stegun (1970) for a reference on modified Bessel functions.

3.15 Caste $1 + m^2$ — Gamma Sum Families

The G + G distributions rule the sum of two gamma random variables, viz.

$$\mathrm{G+G}(\alpha_1, \alpha_2, \beta_1, \beta_2) = \Gamma(\alpha_1, \beta_1) * \Gamma(\alpha_2, \beta_2)$$

Variance Function:

$$\mathbf{V}_F(m) = \frac{1}{2\lambda q}\left([q+1]\Delta(m) - \left[(q-1)m + \frac{\lambda(q+1)^2}{2a}\right]\sqrt{\Delta(m)}\right)$$

$$\Delta(m) = m^2 + \frac{\lambda(q-1)}{a}m + \left(\frac{\lambda(q+1)}{2a}\right)^2$$

where $q \in \mathbb{R}^+$, $a \in \mathbb{R}$

$M_F = \mathbb{R}^+$, $\Lambda(F) = \mathbb{R}^+$

Basis Measure:

$$\mu = \exp(-ax)\gamma^{*\lambda} * \exp(ax)\gamma^{*\lambda q}$$

where $\gamma = \mathbb{1}_{\mathbb{R}^+}(x)$

Cumulant Transform and Mean Domain Mapping:

$$\mathrm{k}_\mu(\theta) = -\lambda\log(a-\theta) - \lambda q\log(-a-\theta), \; \Theta(\mu) = (-\infty, -|a|)$$

$$\phi_\mu(m) = -\frac{2a\sqrt{\Delta(m)} + \lambda(q+1)}{2m}, \; m \in M_F$$

Density:

$$f(x) = \frac{x^{\alpha_1+\alpha_2-1}e^{-x/\beta_2}}{\Gamma(\alpha_1+\alpha_2)\beta_1^{\alpha_1}\beta_2^{\alpha_2}}\, {}_1F_1\left(\alpha_2; \alpha_1+\alpha_2; \left[\frac{1}{\beta_1} - \frac{1}{\beta_2}\right]x\right), \; x \in \mathbb{R}^+$$

where ${}_1F_1$ is the confluent hypergeometric function;

$$\theta = -\frac{1}{2}\left(\frac{1}{\beta_1} + \frac{1}{\beta_2}\right), \; a = \frac{1}{2}\left(\frac{1}{\beta_1} - \frac{1}{\beta_2}\right) \; ; \; \beta_1, \beta_2 \in \mathbb{R}^+$$

$\lambda = \alpha_1$, $q = \alpha_2/\alpha_1$; $\alpha_1, \alpha_2 \in \mathbb{R}^+$; then $m = \alpha_1\beta_1 + \alpha_2\beta_2$

Legendre Transform:

$$\mathrm{k}_\mu^*(x) = \log\left(\frac{[B(x)+2ax]^\lambda[B(x)-2ax]^{\lambda q}}{[2x]^{\lambda(q+1)}}\right) - \frac{B(x)}{2}, \; x \in M_F$$

where $B(x) = \lambda(q+1) + 2a\sqrt{\Delta(x)}$

Asymptotics:

With m as in the Density section above remaining constant:

$$\mathrm{G+G}(\alpha_1,\alpha_2,\beta_1,\beta_2)\xrightarrow{\mathcal{D}}\Gamma\left(\alpha_1,\frac{m}{\alpha_1}\right) \text{ as } \left(\frac{1}{\beta_1}-\frac{1}{\beta_2}\right)\to-\infty.$$
$$\mathrm{G+G}(\alpha_1,\alpha_2,\beta_1,\beta_2)\xrightarrow{\mathcal{D}}\Gamma\left(\alpha_2,\frac{m}{\alpha_2}\right) \text{ as } \left(\frac{1}{\beta_1}-\frac{1}{\beta_2}\right)\to+\infty.$$

Notes:

- See appendix (§B.15) for a derivation of the density of the G + G distribution.

3.16 Caste $1 + m^2$ — Gamma Difference Families

The G − G distributions rule the difference of two gamma random variables, viz.

$$\mathrm{G-G}(\alpha_1, \alpha_2, \beta_1, \beta_2) = \Gamma(\alpha_1, \beta_1) * [-\Gamma(\alpha_2, \beta_2)]$$

Variance Function:

$$\mathbf{V}_F(m) = \frac{1}{2\lambda q}\left([q+1]\Delta(m) + \left[(q-1)m + \frac{\lambda(q+1)^2}{2a}\right]\sqrt{\Delta(m)}\right)$$

$$\Delta(m) = m^2 + \frac{\lambda(q-1)}{a}m + \left(\frac{\lambda(q+1)}{2a}\right)^2$$

where $q \in \mathbb{R}^+$, $a \in \mathbb{R}^+$

$M_F = \mathbb{R}$, $\Lambda(F) = \mathbb{R}^+$

Basis Measure:

$\mu = \exp(-ax)\gamma_+^{*\lambda} * \exp(ax)\gamma_-^{*\lambda q}$

where $\gamma_+ = \mathbb{1}_{\mathbf{R}^+}(x)$ and $\gamma_- = \mathbb{1}_{\mathbf{R}^-}(x)$

Cumulant Transform and Mean Domain Mapping:

$\mathrm{k}_\mu(\theta) = -\lambda\log(a-\theta) - \lambda q\log(a+\theta)$, $\Theta(\mu) = (-a, a)$

$$\phi_\mu(m) = \frac{2a\sqrt{\Delta(m)} - \lambda(q+1)}{2m}, \; m \in M_F$$

Density:

$$f_X(x) = \frac{e^{(1/\beta_2 - 1/\beta_1)x/2}}{\beta_1^{\alpha_1}\beta_2^{\alpha_2}} {}_2B^{\alpha_1,\alpha_2}_{a(\beta_1,\beta_2)}(x), \; x \in \mathbb{R}$$

where $a(\beta_1, \beta_2) = \frac{1}{2}\left(\frac{1}{\beta_1} + \frac{1}{\beta_2}\right)$

where $\int_{-\infty}^{+\infty} e^{tx}{}_2B_k^{b,c}(x)dx = (k-t)^{-b}(k+t)^{-c}$, $b, c, k \in \mathbb{R}^+$, $t \in (-k, k)$

$\theta = \frac{1}{2}\left(\frac{1}{\beta_2} - \frac{1}{\beta_1}\right)$, $a = a(\beta_1, \beta_2)$; $\beta_1, \beta_2 \in \mathbb{R}^+$

$\lambda = \alpha_1$, $q = \alpha_2/\alpha_1$; $\alpha_1, \alpha_2 \in \mathbb{R}^+$;

then $m = \alpha_1\beta_1 - \alpha_2\beta_2$

Legendre Transform:

$$k_\mu^*(x) = \log\left(\frac{[2ax+B(x)]^{\lambda}[2ax-B(x)]^{\lambda q}}{[2x]^{\lambda(q+1)}}\right) - \frac{B(x)}{2},\ x \in M_F$$

where $B(x) = \lambda(q+1) - 2a\sqrt{\Delta(x)}$

Asymptotics:

With m as in the Density section above remaining constant:

$$\mathrm{G}-\mathrm{G}(\alpha_1,\alpha_2,\beta_1,\beta_2) \xrightarrow{\mathcal{D}} \Gamma\left(\alpha_1, \frac{m}{\alpha_1}\right) \text{ as } \alpha_2 \to 0.$$

$$\mathrm{G}-\mathrm{G}(\alpha_1,\alpha_2,\beta_1,\beta_2) \xrightarrow{\mathcal{D}} \left[-\Gamma\left(\alpha_2, -\frac{m}{\alpha_2}\right)\right] \text{ as } \alpha_1 \to 0.$$

$$\mathrm{G}-\mathrm{G}(\alpha_1,\alpha_2,\beta_1,\beta_2) \xrightarrow{\mathcal{D}} \Gamma\left(\alpha_1, \frac{m}{\alpha_1}\right) \text{ as } \left(\frac{1}{\beta_1}+\frac{1}{\beta_2}\right) \to +\infty \text{ if } \beta_1 > \beta_2.$$

$$\mathrm{G}-\mathrm{G}(\alpha_1,\alpha_2,\beta_1,\beta_2) \xrightarrow{\mathcal{D}} \left[-\Gamma\left(\alpha_2, -\frac{m}{\alpha_2}\right)\right] \text{ as } \left(\frac{1}{\beta_1}+\frac{1}{\beta_2}\right) \to +\infty \text{ if } \beta_1 < \beta_2.$$

3.17 Caste $1 + m^2$ — Normal-Gamma Sum Families

The N + G distributions rule the sum of a gamma and a normal random variables, viz.

$$\mathrm{N+G}(\sigma^2, \alpha, \xi, \beta) = \mathrm{N}(\xi, \sigma^2) * \Gamma(\alpha, \beta)$$

Variance Function:

$\mathbf{V}_F(m) = \frac{1}{2\lambda}\left(\Delta(m) + [m - 2\lambda q a]\sqrt{\Delta(m)}\right)$

$\Delta(m) = m^2 - 4\lambda q a m + 4\lambda^2 q(qa^2 + 1)$

where $q \in \mathbb{R}^+$, $a \in \mathbb{R}$

$M_F = \mathbb{R}^+$, $\Lambda(F) = \mathbb{R}^+$

Basis Measure:

$\mu = \exp(-ax)\gamma^{*\lambda} * \exp(ax)\nu^{*\lambda q}$

where $\gamma = \mathbb{1}_{\mathbb{R}^+}(x)$ and $\nu = \frac{1}{\sqrt{2\pi}}\exp(x^2/2)$

Cumulant Transform and Mean Domain Mapping:

$\mathrm{k}_\mu(\theta) = -\lambda\log(a-\theta) + \frac{\lambda q}{2}(\theta + a)^2$, $\Theta(\mu) = (-\infty, a)$

$\phi_\mu(m) = \dfrac{m + \sqrt{\Delta(m)}}{2\lambda q}$, $m \in M_F$

Density:

$f(x) = \frac{1}{\beta^\alpha}\exp\left(-\frac{\xi^2}{2\sigma^2}\right) {}_3B^{\alpha,\sigma^2/2}_{a(\xi,\sigma^2,\beta)}(x)\exp\left[\frac{x}{2}\left(\frac{\xi}{\sigma^2} - \frac{1}{\beta}\right)\right]$, $x \in \mathbb{R}$

where $a(\xi, \sigma^2, \beta) = \frac{1}{2}\left(\frac{\xi}{\sigma^2} + \frac{1}{\beta}\right)$ and

where $\int_{-\infty}^{+\infty} e^{tx}\, {}_3B^{b,c}_k(x)dx = \frac{\exp[c(k+t)^2]}{(k-t)^b}$, $b, c \in \mathbb{R}^+$, $k \in \mathbb{R}$, $t \in (-\infty, k)$

$\theta = \frac{1}{2}\left(\frac{\xi}{\sigma^2} - \frac{1}{\beta}\right)$, $a = a(\xi, \sigma^2, \beta)$; $\beta \in \mathbb{R}^+$, $\xi \in \mathbb{R}$;

$\lambda = \alpha$, $q = \sigma^2/\alpha$; $\alpha, \sigma^2 \in \mathbb{R}^+$;

then $m = \xi + \alpha\beta$

Legendre Transform:

$$k_{\mu}^{*}(x) = \lambda \log\left(\frac{B(x)+\sqrt{\Delta(x)}}{2\lambda q}\right) + \frac{\left(B(x)+\sqrt{\Delta(x)}\right)^2}{8\lambda q} - (aB(x)+\lambda),\ x \in M_F$$

where $B(x) = 2\lambda q a - x$

Asymptotics:

With m as in the Density section above remaining constant:

$\mathrm{N}+\mathrm{G}(\sigma^2,\alpha,\xi,\beta) \xrightarrow{\mathcal{D}} \mathrm{N}(m, 2\alpha)$ as $\sigma^2 \to +\infty$.

$\mathrm{N}+\mathrm{G}(\sigma^2,\alpha,\xi,\beta) \xrightarrow{\mathcal{D}} \Gamma\left(\sigma^2, \frac{m}{\sigma^2}\right)$ as $\alpha \to 0$.

$\mathrm{N}+\mathrm{G}(\sigma^2,\alpha,\xi,\beta) \xrightarrow{\mathcal{D}} \mathrm{N}(m, \alpha)$ as $\frac{\xi}{\sigma^2}+\frac{1}{\beta} \to +\infty$.

3.18 Caste $1 + m^2$ — Hyperbolic Secant Sum Families

The H + H distributions rule the sum of two Generalized Hyperbolic Secant random variables, viz.

$$\mathrm{H+H}(r_1, r_2, \xi_1, \xi_2) = \mathrm{GHS}(r_1, \beta_1) * \mathrm{GHS}(r_2, \beta_2)$$

Variance Function:

$$\mathbf{V}_F(m) = P\Delta(m) + Q(m)\sqrt{\Delta(m)}\,,$$
$$P = \frac{\sin^2 a}{2\lambda q}(1+q)\,,$$
$$Q(m) = \frac{\sin^2 a}{2\lambda q}\left((1-q)m + \lambda\csc(2a)\left((1-q)^2\cos(2a) - 4q\right)\right) \text{ and}$$
$$\Delta(m) = m^2 + 2\lambda(1-q)\cot(2a)m + \lambda^2\left[(1-q)^2\cot^2(2a) + 4q\csc(2a)\right]$$
where $q \in \mathbb{R}^+$, $a \in (-\pi/2, \pi/2)$

$M_F = \mathbb{R}$, $\Lambda(F) = \mathbb{R}^+$

Basis Measure:

$$\mu = \exp(ax)\eta^{*\lambda} * \exp(-ax)\eta^{*\lambda q}$$
$$\text{where } \eta = \frac{\operatorname{sech}(\pi x/2)}{2}\mathbb{1}_{\mathbb{R}^+}(x)$$

Cumulant Transform and Mean Domain Mapping:

$$\mathrm{k}_\mu(\theta) = \lambda\log\sec(\theta + a) + \lambda q\log\sec(\theta - a)\,,\ \Theta(\mu) = (-\pi/2 + |a|, \pi/2 - |a|)$$
$$\phi_\mu(m) = -\frac{\lambda(q+1)\csc(2a) + \sqrt{\Delta(m)}}{m\tan a + \lambda(1-q)}\,,\ m \in M_F$$

Density:

$$f(x) = \exp\left(\frac{[\beta_1 + \beta_2]^x}{2}\right){}_4B^{r_1,r_2}_{a(\beta_1,\beta_2)}(x)\cos^{r_1}\beta_1\cos^{r_2}\beta_2\,,\ x \in \mathbb{R}\,,$$

where $a(\beta_1, \beta_2) = \frac{1}{2}(\beta_1 - \beta_2)$ and

where $\int_{-\infty}^{+\infty} e^{tx}{}_4B^{b,c}_k(x)dx = \sec^b(t+k)\sec^c(t-k)$,

$k \in (-\pi/2, \pi/2)$, $b, c \in \mathbb{R}^+$, $t \in (-\pi/2 + |k|, \pi/2 - |k|)$;

$\theta = \frac{1}{2}(\beta_1 + \beta_2)$, $a = a(\beta_1, \beta_2)$, $\beta_1, \beta_2 \in (-\pi/2, \pi/2)$;

$\lambda = r_1$, $q = r_2/r_1$;

then $m = r_1\tan\beta_1 + r_2\tan\beta_2$

Legendre Transform:

$$\mathrm{k}_{\mu}^{*}(x) = \log\left(\cos^{\lambda}(B(x)-a)\cos^{\lambda q}(B(x)+a)\right) - xB(x)$$

$$\text{where } B(x) = \frac{\lambda(q+1)\csc(2a) + \sqrt{\Delta(x)}}{x\tan a + \lambda(1-q)}$$

Asymptotics:

$\left[\frac{2\sqrt{r_1 r_2(r_1+r_2)}}{|r_1^2 - r_2^2|}\mathrm{H} + \mathrm{H}(r_1, r_2, \xi_1, \xi_2)\right] \xrightarrow{\mathcal{D}} \mathrm{N}(m, 1)$ as $\beta_1 - \beta_2 \to 0$, with m as in the Density section above remaining constant.

4 Seshadri Class

The Seshadri class contains all convolution families with variance functions of the form

$$\mathbf{V}_F(m) = P(m)\Delta(m) + Q(m)\sqrt{\Delta(m)}$$

where $\deg P = \deg Q = \deg \Delta = 1$. Kokonendji (1993, ch. 2) has exhaustively classified its members, by showing that for $\deg \Delta = 1$, their variance function could be written in the form

$$\mathbf{V}_F = \sqrt{\Delta} P_*(\sqrt{\Delta})$$

where P_* is a polynomial of degree 3. P_*, in fact, ranges across the forms of all variance functions of Mora class members.

Kokonendji and Seshadri (1994) show that families generated by the Lindsay transform of measures with strictly cubic variance functions are infinitely divisible and belong to the Seshadri class. Seshadri (1991) investigates several members of the Seshadri class in the context of finite mixtures of Mora class members.

The Seshadri class contains, up to affinities, 5 convolution families.

Table 2.7: Seshadri class convolution families

S-Abel:	$\{SAb(r_1, r_2, \xi) \mid (\xi, r_1) \in (0,1) \times [\max(0^+, -r_2), +\infty)\}$, for $r_2 \in \mathbb{R}$, supported on $\mathbb{N}$
S-Takács:	$\{STa(r_1, r_2, a, \xi) \mid (\xi, r_1) \in (0, (a+1)^{-1}) \times [\max(0^+, -r_2), +\infty)\}$, for $(r_2, a) \in \mathbb{R} \times \mathbb{R}^+$, supported on $\mathbb{N}$
S-Arcsine:	$\{SArc(r_1, r_2, a, \xi) \mid (\xi, r_1) \in (0, \arctan a^{-1}) \times [\max(0^+, -ar_2), +\infty)\}$,for $(r_2, a) \in \mathbb{R} \times [0, +\infty)$, supported on $\mathbb{N}$
S-Kendall-Ressel:	$\{SKR(r_1, r_2, \xi) \mid (\xi, r_1) \in (1, +\infty) \times [\max(0^+, -r_2), +\infty)\}$, for $r_2 \in \mathbb{R}$, supported on $\mathbb{R}^+$
Reciprocal Inverse Gaussian:	$\left\{RIG(\chi, \psi, r) \mid (\chi, r) \in \mathbb{R}^+ \times \mathbb{R}^+\right\}$, for $\psi \in \mathbb{R}^+$, supported on $\mathbb{R}^+$

Kokonendji (1993, ch. 2) distinguishes between the S-Strict Arcsine ($a = 0$) and S-Large Arcsine ($a > 0$) distributions.

4.1 S-Abel Families

The S-Abel families, denoted $SAb(r_1, r_2, \xi)$, arise as generalizations of the Babel Generalized Poisson families. Members of these families were discussed in the context of random mapping theory by Berg and Nowicki (1991), while Kokonendji and Seshadri (1994) recast these distributional results within exponential family theory.

Variance Function:

$$\mathbf{V}_F(m) = \frac{1}{2\lambda^2}\left([m + \lambda q^2]\Delta(m) + \left[\lambda(1-3q)m - \lambda^2 q^2(q+1)\right]\sqrt{\Delta(m)}\right)$$
$$\Delta(m) = 4\lambda m + \lambda^2(q+1)^2\,,$$
where $q \in [-1, +\infty)$
$$M_F = \mathbb{R}^+\,,\ \Lambda(F) = \mathbb{R}^+$$

Basis Measure:

$$\mu = (\delta_0 - \nu)^{*(-\lambda)} * \nu^{*\lambda q}\,,$$
where $\nu = \delta_0 + \sum_{j=1}^{\infty} \frac{(j+1)^{j-1}}{j!}\delta_j$

is a generating measure for $q \geq 0$.

Cumulant Transform and Mean Domain Mapping:

$$\mathrm{k}_\mu(\theta) = -\lambda \log\left(1 - f^{-1}(e^\theta)\right) + \lambda q f^{-1}(e^\theta)\,,\ \Theta(\mu) = (-\infty, -1)$$
where f^{-1} is the reciprocal of $f : (0,1) \to (0, e^{-1})\,,\ t \mapsto \frac{t}{e^t}$
$$\phi_\mu(m) = \log(A(m)) - A(m)$$
where $A(m) = \frac{2m + \lambda(q+1) - \sqrt{\Delta(m)}}{2(m + \lambda q)}$

Density:

$$f_X(x) = \frac{r_1 r_2}{\Gamma(r_1+1)}\, {}_1C^{(r_1,r_2)}(x) e^{-\xi(x+r_2)}(1-\xi)^{r_1}\xi^x\,,\ x \in \mathbb{N}\,,$$
where ${}_1C^{(b,c)}(0) = 1$ and ${}_1C^{(b,c)}(z) = \frac{c}{z!}\left[(b+z) + (c+1)^{\sim}\right]^{z-1}$,

for $b \in \mathbb{R}^+\,,\ c \in (-b, +\infty)\,,\ z \in \mathbb{N}^*$, and

where $[s + t^{\sim}]^n = \sum_{j=0}^{n} \binom{n}{j} s^j (t)^{(n-j)}$

$\theta = \log\xi - \xi\,,\ \xi \in (0,1)$; $\lambda = r_1\,,\ r_1 \in \mathbb{R}^+$; $q = r_2/r_1\,,\ r_2 \in [-r_1, +\infty)$

then $m = \frac{\xi}{(1-\xi)^2}[r_1 + r_2(1-\xi)]$

Legendre Transform:

$$\mathrm{k}_\mu^*(x) = \frac{1}{2}\log\left(\frac{\left[2x+\lambda(q+1)-\sqrt{\Delta(x)}\right]^x\left[\lambda(q-1)+\sqrt{\Delta(x)}\right]^\lambda}{\left[2(x+\lambda q)\right]^{m+\lambda}}\right)$$
$$-\frac{1}{2}\left(2x+\lambda(q+1)-\sqrt{\Delta(x)}\right),\ x\in M_F$$

Asymptotics:

With m as in the Density section above remaining constant:

$\mathrm{SAb}(r_1,r_2,\xi)\xrightarrow{\mathcal{D}}\mathrm{Po}(m)$ as $r_1\to+\infty$ or $r_2\to+\infty$.
$\mathrm{SAb}(r_1,r_2,\xi)\xrightarrow{\mathcal{D}}\mathrm{GP}\left(\frac{mr_2}{m+r_2},\frac{m}{m+r_2}\right)$ as $r_1\to 0$ for $r_2>0$.

Notes:

- The variance function may also be written out as

$$\mathbf{V}_F(m) = \frac{\sqrt{\Delta(m)}}{8\lambda^3}\left(\sqrt{\Delta(m)}-\lambda(q+1)\right)\left(\sqrt{\Delta(m)}+\lambda(1-q)\right)^2$$

  a representation which emphasizes the relationship between the S-Abel families and the Mora Abel families.

- When $r_2=-r_1$ the variance function takes the following simple form:

$$\mathbf{V}_F(m) = \frac{2m}{\lambda}\left(\sqrt{m}+\sqrt{\lambda}\right)^2$$

- See appendix (§B.17) for a derivation of the density.

4.2 S-Takács Families

The S-Takács families, denoted $\mathrm{STa}(r_1, r_2, a, \xi)$, arise as further generalizations of the Babel Generalized Negative Binomial families.

Variance Function:

$$
\begin{aligned}
&\mathbf{V}_F(m) = P(m)\Delta(m) + Q(m)\sqrt{\Delta(m)} \\
&P(m) = \frac{1}{\lambda^2 a^2 (a+1)^2}\left[a(a+1)m + \lambda q(q-1)\right], \\
&Q(m) = \frac{1}{\lambda^2 a^2 (a+1)^2}\left[\lambda a(a+1)(a+2-3q)m - \lambda^2 q(q-1)(a+q)\right] \\
&\Delta(m) = 4\lambda a(a+1)m + \lambda^2 (a+q)^2, \\
&\text{where } a \in \mathbb{R}^+ \text{ and } q \in [-a, +\infty) \\
&M_F = \mathbb{R}^+,\ \Lambda(F) = \mathbb{R}^+
\end{aligned}
$$

Basis Measure:

$$
\begin{aligned}
&\mu = (\delta_0 - a\nu)^{*(-\lambda)} * \nu^{*\lambda q}, \\
&\text{where } \nu = \sum_{j=0}^{\infty} \frac{1}{j!}\frac{\Gamma(aj+j+1)}{\Gamma(aj+2)}\delta_j \\
&\text{is a generating measure for } q \geq 0.
\end{aligned}
$$

Cumulant Transform and Mean Domain Mapping:

$$
\begin{aligned}
&\mathrm{k}_\mu(\theta) = -\lambda \log\left(1 - a f^{-1}(e^\theta)\right) + \lambda q \log\left(1 + f^{-1}(e^\theta)\right), \\
&\Theta(\mu) = \left(-\infty, \log(a^a) - \log\left((a+1)^{a+1}\right)\right) \\
&\text{where } f^{-1} \text{ is the reciprocal of} \\
&f : (0, 1/a) \to \left(0, a^a/(a+1)^{a+1}\right),\ t \mapsto t/(1+t)^{a+1} \\
&\phi_\mu(m) = \log\left(\frac{\left[2(a+1)m + p(a+q) - \sqrt{\Delta(m)}\right]}{\left[2(a+1)\left((a+1)m + pq\right)\right]^{a+1}}\right) \\
&\qquad + a\log\left(2a(a+1)m + p(q + 2aq - a) + \sqrt{\Delta(m)}\right),\ m \in M_F
\end{aligned}
$$

Density:

$$f_X(x) = (1-(a+1)\xi)^{r_1}\,\xi^x(1-\xi)^{r_2-ax}{}_2C_a^{(r_1,r_2)}(x)\,,\ x\in\mathbb{N}\,,\text{ where}$$
$${}_2C_k-1^{(b,c)}(0)=1\ \text{ and}$$
$${}_2C_k^{(b,c)}(x)=\frac{1}{x!}\left(c[(c-1+kx+x)_\sim+kb^\sim]^{x-1}+kb[(r_2+kx+x)_\sim+k(b-1)^\sim]^{x-1}\right)$$
$$\text{where } (s_\sim+tu^\sim)^n=\sum_{j=0}^{n}\binom{n}{j}(s)_{(j)}t^{n-j}(u)^{(n-j)}\ ,\text{ and}$$
$$\text{where } (b)_c=b(b+1)\dots(b+c-1)\,,\ c\in\mathbb{N}\,,\ x\in\mathbb{N}^*$$
$$\theta=\log\xi-a\log(1-\xi)\,,\ \xi\in(0,1/(a+1))\ ;\ a=a\,,\ a\in\mathbb{R}^+$$
$$\lambda=r_1\,,\ r_1\in\mathbb{R}^+\ ;\ q=r_2/r_1\,,\ r2\in[-ar_1,+\infty)\ ;$$
$$\text{then } m=\xi\frac{ar_1+(1-(a+1)\xi)r_2}{(1-(a+1)\xi)^2}$$

Legendre Transform:

$$\mathrm{k}^*_\mu(x)=\log\left(\frac{[2B(x)+C-\sqrt{\Delta(x)}]^x[2aB(x)-C+\sqrt{\Delta(x)}]^{ax}}{[2B(x)-2(x+\lambda)]^{\lambda(1-q)}[2(a+1)B(x)]^{B(x)-\lambda q}}\right)$$
$$+\log\left(\frac{[\sqrt{\Delta(x)}-C-2\lambda]^\lambda a^{\lambda q}}{[2aB(x)-C-sqrt\Delta(x)]^{\lambda q}}\right)\ ,\ x\in M_F$$
$$\text{where } B(x)=(a+1)x+\lambda q\,,\ C=\lambda(a-q)$$

Asymptotics:

$\mathrm{STa}(r_1,r_2,a,\xi)\xrightarrow{\mathcal{D}}\mathrm{Po}(m)$ as $r_1\to+\infty$ or $r_2\to+\infty$ for m as in the Density section above remaining constant.

Notes:

- As a factored polynomial of $\sqrt{\Delta}$, the variance function is

$$\mathrm{V}_F(m)=\frac{\sqrt{\Delta}}{8\lambda^3a^2(a+1)^2}\left(\sqrt{\Delta}-\lambda(a+q)\right)\left(\sqrt{\Delta}+\lambda(a-q)\right)\left(\sqrt{\Delta}+\lambda(a+2-q)\right)$$

4.3 S-Arcsine Families

The S-Arcsine families, denoted $\text{SArc}(r_1, r_2, a, \xi)$, first arose in the course of the classification of the members of the Seshadri class by Kokonendji (1993).

No probabilistic interpretation has been proposed for these families thus far.

Variance Function:

$$\begin{aligned}
&\mathbf{V}_F(m) = P(m)\Delta(m) + Q(m)\sqrt{\Delta(m)} \\
&P(m) = \frac{1}{2\lambda^2(a^2+1)^2}\left[(a^2+1)m + \lambda(q^2+1)\right], \\
&Q(m) = \frac{1}{2\lambda^2(a^2+1)^2}\left[\lambda(a^2+1)(a-3q)m - \lambda^2(a+q)(q^2+1)\right] \\
&\Delta(m) = 4\lambda(a^2+1)m + 4\lambda^2(q+a)^2, \\
&\text{where } a \in \mathbb{R}^+ \cup \{0\}, \; q \in [-a, +\infty) \\
&M_F = \mathbb{R}^+, \; \Lambda(F) = \mathbb{R}^+
\end{aligned}$$

Basis Measure:

$$\begin{aligned}
&\mu = \nu_1^{*\lambda} * \nu_2^{*\lambda q}, \\
&\text{where } \nu_1 = \sum_{j=0}^{\infty} \nu_{1,j}\delta_j \text{ and } \nu_2 = \sum_{j=0}^{\infty} \nu_{2,j}\delta_j, \\
&\text{with } \sum_{j=0}^{\infty} \nu_{1,j} z^j = \left(\sqrt{1 - f^{-1}(z)^2} - a f^{-1}(z)\right)^{-1}, \; \nu_{2,j} = \frac{1}{j!}\frac{p_j(aj+1)}{aj+1} \\
&\text{where } p_{2j}(t) = \prod_{i=0}^{j-1}(t^2 + 4i^2) \text{ and } p_{2j+1}(t) = t\prod_{i=0}^{j-1}\left(t^2 + (2i+1)^2\right) \\
&\text{and } f^{-1}(\cdot) \text{ as below,} \\
&\text{is a generating measure for } q \geq 0.
\end{aligned}$$

Cumulant Transform and Mean Domain Mapping:

$$\begin{aligned}
&\mathrm{k}_\mu(\theta) = -\lambda \log\left(\sqrt{1 - f^{-1}(e^\theta)} - ae^\theta\right) + \lambda q \arcsin(f^{-1}(e^\theta)), \\
&\Theta(\mu) = \left(-\infty, -\log\sqrt{1+a^2} - a\arctan a^{-1}\right) \\
&\text{where } f^{-1} \text{ is the reciprocal of} \\
&f : \left(0, \sqrt{1+a^2}\right) \to \left(0, \sqrt{1+a^2}\exp\left(a \arctan a^{-1}\right)\right), \; t \mapsto \frac{t}{\exp(a \arcsin t)}; \\
&\phi_\mu(m) = \log(\sin\arctan A(m)) - a\arctan A(m), \\
&\text{where } A(m) = \frac{2am + \lambda(a+q) - \sqrt{\Delta(m)}}{2(a^2 m + \lambda(aq-1))}
\end{aligned}$$

Density:

$$f_X(x) = (\cos\xi - a\sin\xi)^{r_1}\,{}_3C_a^{r_1,r_2}(x)\sin^x\xi e^{-\xi(ax+r_2)}\,,$$

$$\text{where } \sum_{j=0}^{\infty} {}_3C_k^{b,c}(j)z^j = \frac{\exp\left(c\arcsin f^{-1}(z)\right)}{\left(\sqrt{1-f^{-1}(z)^2} - kf^{-1}(z)\right)^b}$$

$$\text{for } b \in \mathbb{R}^+\,,\ k \in \mathbb{R}^+ \cup \{0\}\,,\ c \in [-ar_1, +\infty)\,,\ z \in \left(0, \frac{\exp(\arctan a^{-1})}{1+a^2}\right)$$

$$\theta = \log\sin\xi - a\xi\,,\ \xi \in (0, \arctan a^{-1})\ ;\ \lambda = r_1\,,\ r_1 \in \mathbb{R}^+\ ;$$

$$a = a\,,\ a \in \mathbb{R}^+\ ;\ q = r_2/r_1\,,\ r_2 \in (-ar_1, +\infty)$$

$$\text{then } m = \frac{\xi}{(1-a\xi)^2}\left[(ar_1 + r_2) + \xi(r_1 - ar_2)\right)$$

Legendre Transform:

$$\mathrm{k}^*_\mu(x) = \frac{1}{2}\log\left(\frac{\left[B(x) - \sqrt{\Delta(x)}\right]^{2x}\left[a\sqrt{\Delta(x)} - D\right]^{2\lambda}}{2^{x+\lambda}\left[2C(x)(C(x)+x) + \Delta(x) - B(x)\sqrt{\Delta(x)}\right]^{x+\lambda}}\right)\,,\ x \in M_F$$

$$\text{where } B(x) = 2ax + \lambda(a+q)\,,\ C(x) = a^2x + \lambda(aq-1)\,,\ D = \lambda[a(a+q) - 2)]$$

Asymptotics:

$\mathrm{SArc}(r_1, r_2, a, \xi) \xrightarrow{\mathcal{D}} \mathrm{Po}(m)$ as $r_1 \to +\infty$ or $r_2 \to +\infty$ while m as in the Density section above remains fixed.

Notes:

- Kokonendji (1993) treats as separate the cases where $a = 0$ and $a > 0$; in the first case the distribution is denoted SSA (Seshadri Strict Arcsine) and in the second SLA (Seshadri Large Arcsine).

- The variance function of the SArc families may also be written out as

$$\mathrm{V}_F(m) = \frac{\sqrt{\Delta(m)}}{8\lambda^3(a^2+1)}\left(\sqrt{\Delta(m)} - \lambda(a+q)\right)\left(\left[\sqrt{\Delta(m)} + \lambda(a-q)\right]^2 + 4\lambda^2\right)$$

  a representation which emphasizes the relationship between the Seshadri Arcsine families and the Mora Arcsine families.

- The generating measure becomes more tractable when $a = 0$. In this case we get

$$\mathrm{SArc}(r_1, r_2, 0, \xi) = [2\mathrm{NB}(r_1/2, p)] * \mathrm{Arc}\left(r_2, 0, \xi/\sqrt{1+\xi^2}\right)$$

  where $p = \log\left(1 - \xi/\sqrt{1+\xi^2}\right)$.

4.4 S-Kendall-Ressel Families

The S-Kendall-Ressel families, denoted SKR(r_1, r_2, ξ), arise as generalizations of the Babel Kendall-Ressel families. Kokonendji (1993) gives a straightforward probabilistic interpretation of the SKR families for $r_1 > 0$ and $r_2 \geq 0$: let $X \sim \text{KR}(R, \xi) \wedge_R \text{NB}(r_1, 1 - 1/\xi)$ and $Y \sim \text{KR}(r_2, \xi)$; then $X + Y \sim \text{SKR}(r_1, r_2, \xi)$.

Variance Function:

$$\mathbf{V}_F(m) = \frac{1}{2\lambda^2}\left(\left[m + \lambda q(q-1)\right]\Delta(m) - \left[\lambda(3q-2)m + \lambda^2 q^2(q-1)\right]\sqrt{\Delta(m)}\right)$$
$$\Delta(m) = 4\lambda m + \lambda^2 q^2,$$
where $q \in [-1, +\infty)$
$$M_F = \mathbb{R}^+, \; \Lambda(F) = \mathbb{R}^+$$

Basis Measure:

$$\mu = (\delta_0 - \nu)^{*\lambda} * \nu^{*\lambda q},$$
where $\nu = \dfrac{x^x e^{-x}}{\Gamma(x+2)} \mathbb{1}_{\mathbb{R}^+}(x)$
is a generating measure for $q \geq 0$.

Cumulant Transform and Mean Domain Mapping:

$$\mathrm{k}_\mu(\theta) = -\lambda \log\left(1 - f^{-1}(-\theta)\right) - \lambda q f^{-1}(-\theta), \; \Theta(\mu) = \mathbb{R}^-$$
where f^{-1} is the reciprocal of $f : \mathbb{R}^+ \to \mathbb{R}^+, \; t \mapsto e^t - t - 1$
$$\phi_\mu(m) = \log\left(\frac{2m + \lambda q + \sqrt{\Delta(m)}}{2m}\right) - \frac{\lambda q + \sqrt{\Delta(m)}}{2m}, \; m \in M_F$$

Density:

$$f_X(x) = \frac{\xi^x x^{x+r_2-1}}{\Gamma(x+r_2)}\left(1 - e^{-\xi}\right)^{r_1} e^{-\xi(x+r_2)} {}_1F_1(r_1 - 1; x + r_2; x), \; x \in \mathbb{R}^+$$
where ${}_1F_1$ is the confluent hypergeometric function
and $\theta = \log\xi + 1 - \xi, \; \xi \in (1, +\infty)$; $\lambda = r_1, \; r_1 \in \mathbb{R}^+$; $q = r_2/r_1, \; r_2 \in [-1, +\infty)$
then $m = \dfrac{r_1 + r_2(e^\xi - 1)}{(e^\xi - 1)^2}$

gendre Transform:

$$k_\mu^*(x) = \log\left(\frac{[2x + B(x)]^{x+\lambda(q-1)}[B(x)]^{\lambda}}{[2m]^{x+\lambda q}}\right) - \frac{B(x)}{2}, x \in M_F$$

where $B(x) = \lambda q + \sqrt{\Delta(m)}$

Notes:

- An alternative expression for the variance function is given by Kokonenjdji (1993) and clarifies the relationship between the variance functions of the Kendall-Ressel and the S-Kendall-Ressel families:

$$\mathbf{V}_F(m) = \frac{\sqrt{\Delta(m)}}{8\lambda^3}\left(\sqrt{\Delta(m)} - \lambda q\right)^2\left(\sqrt{\Delta(m)} + \lambda(2-q)\right)$$

- The form of the density is due to Kokonendji (1993).

4.5 Reciprocal Inverse Gaussian Families

The Reciprocal Inverse Gaussian families, denoted RIG(ψ, χ, r), consist of the powers of the reciprocal of an IG(ψ, χ) random variable, namely $X \sim \text{RIG}(\psi, \chi, 1/2) \Rightarrow X^{-1} \sim \text{IG}(\phi, \chi)$.

Since the Inverse Gaussian has a natural interpretation as a first-passage time in Brownian motion, the Reciprocal Inverse Gaussian r.v. is proportional to the speed with which the process approaches a bound. RIG$(\cdot, \cdot, 1/2)$ and IG are both special cases of the Generalized Inverse Gaussian distribution. An important feature of the RIG families in application is that they consist of convolutions of Gamma and Inverse Gaussian random variables.

The RIG families correspond to Kokonendji's (1993) SIG type.

Variance Function:

$$\mathbf{V}_F(m) = \frac{1}{2\lambda^2}\left([m + \lambda q^2]\Delta(m) + \lambda q[3m + \lambda q^2]\sqrt{\Delta(m)}\right)$$
$$\Delta(m) = 4\lambda m + \lambda^2 q^2$$
where $q \in \mathbb{R}^+$
$$M_F = \mathbb{R}^+,\ \Lambda(F) = \mathbb{R}^+$$

Basis Measure:

$$\mu = (2\gamma)^{*\lambda/2} * \nu^{\lambda q}$$
where $\nu = \dfrac{1}{\sqrt{2\pi x^3}} \exp\left(-\dfrac{1}{2x}\right) \mathbb{1}_{\mathbb{R}^+}(x)$

Cumulant Transform and Mean Domain Mapping:

$$\mathrm{k}_\mu(\theta) = -\frac{\lambda}{2}\log(-2\theta) - \lambda q\sqrt{-2\theta},\ \Theta(\mu) = \mathbb{R}^-$$
$$\phi_\mu(m) = -\frac{2\lambda m + \lambda^2 q^2 + \lambda q\sqrt{\Delta(m)}}{4m^2},\ m \in M_F$$

Density:

$$f(x) = \frac{\chi^r e^{\sqrt{\chi\psi}}}{2^r\Gamma(r)} x^{r-1} e^{-1/2(\chi x + \psi/x)} \left[{}_1F_1\left(r - \frac{1}{2}; \frac{1}{2}; \frac{\psi}{2x}\right) + (1 - 2r)\sqrt{\frac{\psi}{2x}} \frac{\Gamma(r)}{\Gamma(r + 1/2)} {}_1F_1\left(r; \frac{3}{2}; \frac{\psi}{2x}\right)\right]$$

where ${}_1F_1$ is the confluent hypergeometric function;

$\theta = -\dfrac{\chi}{2},\ \chi \in \mathbb{R}^+,\ \lambda = 2r,\ q = \dfrac{\sqrt{\psi}}{2r},\ r \in \mathbb{R}^+ \psi \in \mathbb{R}^+$;

then $m = \dfrac{\sqrt{\chi\psi} + 2r}{\chi}$

Legendre Transform:

$$k_{\mu}^{*}(x) = \frac{\lambda}{2}\log\left(\frac{A(x)}{x}\right) + \lambda q\sqrt{\frac{A(x)}{x}} - \frac{A(x)}{2}$$

where $A(x) = 2\lambda x + \lambda^2 q^2 + \lambda q\sqrt{\Delta(x)}$

Asymptotics:

With m as in the Density section above remaining constant:

$\mathrm{RIG}(\psi, \chi, r) \xrightarrow{\mathcal{D}} \mathrm{IG}(m/\sqrt{\psi}, \psi)$ as $r \to 0$ while m, ψ remain fixed.

$\mathrm{RIG}(\psi, \chi, r) \xrightarrow{\mathcal{D}} \Gamma(r, m/r)$ as $\psi \to 0$ while m, r remain fixed.

Notes:

- $\mathrm{RIG}(\chi, \psi, r) \xrightarrow{\mathcal{D}} \Gamma\left(\frac{1}{2}, \frac{2}{\chi}\right)$ as $\psi \to 0$.
- See appendix (§B.16) for a derivation of the density.

Other References:

- Jørgensen (1982) for a complete survey of the Generalized Inverse Gaussian.
- Seshadri (1993) for a complete survey of the Inverse Gaussian and related distributions in the context of exponential family theory.

Chapter 3

Canonical Caste Members

We provide in this tabular chapter a synoptic table of the variance functions of the Canonical Caste Members for all classified convolution families in the compendium. The Grand-Babel form of the variance function $V_F(m) = P(m)\Delta(m) + Q(m)\sqrt{\Delta(m)}$ is assumed; accordingly, the variance of the CCMs is expressed in terms of P and Q, with Δ corresponding to the caste. The background material is presented in §2.6, chapter 1.

Family	Caste $\Delta(m)$	$K(t)$ / $V_{K*F}(m)$
Morris Class		
N	1	t $P \equiv \lambda$, $Q \equiv 0$
Po	m	t $P \equiv 1$, $Q \equiv 0$
Bin	$1 - m^2$	$2t/\lambda - 1$ $P \equiv 1/\lambda$, $Q \equiv 0$
NB	$m^2 - 1$	$2t/\lambda + 1$ $P \equiv 1/\lambda$, $Q \equiv 0$
Γ	m^2	t $P \equiv 1/\lambda$, $Q \equiv 0$
GHS	$m^2 + 1$	t/λ $P \equiv 1/\lambda$, $Q \equiv 0$

Family	Caste $\Delta(m)$	$K(t)$ / $V_{K*F}(m)$
Mora Class		
GP	m^2	t $P \equiv 0,$ $Q(m) = (m/\lambda + 1)^2$
GNB	m^2	t $P \equiv 0,$ $Q(m) = [(a+1)m/\lambda + 1][am/\lambda + 1]$ $a > 0$
Arc	m^2	t $P \equiv 0,$ $Q(m) = (am/\lambda + 1)^2 + (m/\lambda)^2$ $a \geq 0$
KR	m^2	$t + \lambda$ $P \equiv 0,$ $Q(m) = (m/\lambda - 1)^2$
IG	m^2	t $P \equiv 0,$ $Q(m) = (m/\lambda)^2$
Babel Class		
Hermite	m	$4\lambda t + \lambda^2$ $P = 8\lambda,$ $Q(m) = -8\lambda^2$
Laguerre	m	$4\lambda t + \lambda^2(q+1)^2$ $P = -8pq,$ $Q(m) = 4[m + \lambda^2(q^2 - 1)]$ $q \geq -1$
Γ'	m	$4\lambda t + \lambda^2 q^2$ $P = -8pq,$ $Q(m) = 4(m + \lambda^2 q^2)$ $q \geq 0$

Family	Caste $\Delta(m)$	$K(t)$ / $\mathbf{V}_{K*F}(m)$
Babel Class (continued)		
MG	m^2-1	$\frac{a-1}{2\lambda\sqrt{q+1}}t-\frac{q+2}{\sqrt{q+1}^2}$ $P=\frac{1-a(1+q)}{2\lambda(q+1)}$, $Q(m)=-\frac{1}{2\lambda(q+1)}\left([1+a(q+1)]m+\frac{a(q+1)(q+4)+3q+4}{2\sqrt{q+1}}\right)$ $(a,q)\in[-1,1)\times[0,+\infty)\setminus\{(-1,0)\}$
ME	m^2-1	$\frac{t}{2\lambda a\sqrt{q+1}}+\frac{q+2}{2\sqrt{q+1}}$ $P=-\frac{q}{2\lambda(q+1)}$, $Q(m)=\frac{1}{2\lambda(q+1)}(q+2)m-\frac{q^2+8q+8}{2\sqrt{q+1}}$
Trin	$1-m^2$	$\frac{\sqrt{1-a^2}}{\lambda}t$ $P=1/\lambda$, $Q(m)=a/\lambda^2$ $a\in(0,1)$
NB + NB	m^2+1	$\frac{a^2-1}{2\lambda a\sqrt{q}}t+\frac{a^2-q}{2a\sqrt{q}}$ $P=\frac{q+1}{2\lambda q}$, $Q(m)=\frac{q-1}{2\lambda q}m-\frac{3a^2q-q^2-a^2+3q}{\lambda aq\sqrt{q}}$ $(a,q)\in\mathbb{R}^+\times\mathbb{R}^+$
NB − NB	m^2+1	$\frac{a^2-1}{2\lambda a\sqrt{q}}t-\frac{a(q-1)}{2\sqrt{q}}$ $P=\frac{q+1}{2\lambda q}$, $Q(m)=\frac{q-1}{2\lambda q}m-\frac{a^2(q^2+1)+2q}{2(a^2-1)q}$ $(a,q)\in(0,1)\times\mathbb{R}^+$

Babel Class
(continued)

NB + B	m^2+1	$\frac{a^2+1}{2\lambda a\sqrt{q}}t - \frac{a^2-q}{2a\sqrt{q}}$ $P = \frac{1-q}{2\lambda q}$, $Q(m) = \frac{q+1}{2\lambda q}m + \frac{3a^2q+a^2-3q-q^2}{4\lambda aq\sqrt{q}}$ $(a,q) \in \mathbb{R}^+ \times \mathbb{R}^+$
P + NB	m^2+1	$\frac{1}{2\lambda a}t - \frac{a^2-1}{2a}$ $P = \frac{1}{2\lambda}$, $Q(m) = \frac{m}{2\lambda}m + \frac{a^2-3}{4\lambda a}$ $a \in \mathbb{R}^+$
P − NB	m^2+1	$\frac{1}{2\lambda a}t - \frac{a}{2}$ $P = \frac{1}{2\lambda}$, $Q(m) = -\frac{m}{2\lambda} - \frac{a^2-2}{4\lambda a}$ $a \in \mathbb{R}^+$
B + B	m^2+1	$\frac{a^2-1}{2\lambda a\sqrt{q}}t - \frac{a^2-q}{2a\sqrt{q}}$ $P = -\frac{q+1}{2\lambda q}$, $Q(m) = \frac{1}{2\lambda q}\left((q-1)m - \frac{a^2-3a^2q-3q+q^2}{a\sqrt{q}}\right)$ $(a,q) \in \mathbb{R}^+ \times \{j/\lambda \mid j \in \mathbb{N}^*\}$
P + B	m^2+1	$\frac{t}{2\lambda a} + \frac{a^2-1p}{2a}$ $P = -\frac{1}{2\lambda}$, $Q(m) = \frac{m}{2\lambda} - \frac{a^2-3}{4\lambda a}$ $a \in \mathbb{R}^+$

Babel Class
(continued)

P − P	m^2+1	$\frac{t}{2\lambda}$ $P=0,$ $Q=\frac{1}{2\lambda}$
G + G	m^2+1	$\frac{at}{\lambda\sqrt{q}}+\frac{q-1}{2\sqrt{q}}$ $P=\frac{q+1}{2\lambda q},$ $Q(m)=-\frac{q-1}{2\lambda q}m+\frac{q^2-6q+1}{4\lambda q\sqrt{q}}$ $(a,q)\in\mathbb{R}\times\mathbb{R}^+$
G − G	m^2+1	$\frac{at}{\lambda\sqrt{q}}+\frac{q-1}{2\sqrt{q}}$ $P=\frac{q+1}{2\lambda q},$ $Q(m)=\frac{1}{2\lambda q}\left((q-1)m-\frac{q^2-6q+1}{2\sqrt{q}}\right)$ $(a,q)\in\mathbb{R}^+\times\mathbb{R}^+$
N + G	m^2+1	$\frac{t}{2\lambda\sqrt{q}}-a\sqrt{q}$ $P=\frac{1}{2\lambda},$ $Q(m)=\frac{m}{2\lambda}+\frac{a\sqrt{q}}{2\lambda}$ $(a,q)\in\mathbb{R}\times\mathbb{R}^+$
H + H	m^2+1	$\frac{\sqrt{\lvert\sin(2a)\rvert}}{2\sqrt{q}}\left(\frac{t}{\lambda}+\tan(2a)(q-1)\right)$ $P=\frac{q+1}{2\lambda q}\sin^2(a),$ $Q(m)=\frac{\sin^2 a}{2\lambda q}\left((q-1)m+\frac{\sqrt{\lvert\sin(2a)\rvert}}{\sqrt{q}}\left[\tan(2a)(q-1)^2+4q\csc(2a)\right]\right)$ $(a,q)\in(-\pi/2,\pi/2)\times\mathbb{R}^+$

Family	Caste $\boldsymbol{\Delta}(m)$	$\boldsymbol{K}(t)$ / $V_{K*F}(m)$
Seshadri Class		
SAb	m	$4\lambda t + \lambda^2(q+1)^2$ $P(m) = 2m/\lambda + 2\lambda(3q+1)(q-1)$, $Q(m) = (2-6q)m - 2\lambda^2(q+1)(q-1)^2$ $q \in [-1, +\infty)$
STa	m	$4\lambda a(a+1)t + \lambda^2(a+q)^2$ $P(m) = 2m/\lambda - 2\lambda(a^2+2aq+4q-3q^2)$, $Q(m) = 2(a+2-3q)m - 2\lambda^2(a+2-q)(a^2-q^2)$ $a \in \mathbb{R}^+$, $q \in [-a, +\infty)$
SArc	m	$4\lambda(a^2+1)m + 4\lambda^2(q+a)^2$ $P(m) = 2m/\lambda - 8\lambda(a^2+2aq-1)$ $Q(m) = 2(a-3q)m - 8\lambda^2(a+q)(a^2-2aq-2q^2+1)$ $a \in [0, +\infty)$, $q \in [-a, +\infty)$
SKR	m	$4\lambda t + \lambda^2 q^2$ $P(m) = 2m/\lambda + 2\lambda q(3q-4)$ $Q(m) = -2(3q-2)m - 2\lambda^2 q^2(q-2)$ $q \in [-1, +\infty)$
RIG	m	$4\lambda m + \lambda^2 q^2$ $P(m) = 2m/\lambda + 6\lambda q^2$, $Q(m) = -6qm - 2\lambda^2 q^3$ $q \in \mathbb{R}^+$

Appendix A

Examples of Other Variance Functions

We include here a few examples of variance functions which do not belong to the fully classified sets of families. Our purpose here is simply to provide an immediate reference to two sets of convolution families to which chapter 2 refers.

1 Gamma-Poisson Mixture Families

The gamma-Poisson mixture distribution, denoted $\Gamma \wedge \mathrm{Po}(\beta, p)$ has the interesting property of being a self-inverse distribution. The full form of the mixture, $\Gamma(J, \beta) \wedge_J \mathrm{Po}(\lambda)$, clarifies its probabilistic interpretation.

Variance Function:

$$\mathbf{V}_F(m) = \frac{2m^{\frac{3}{2}}}{\sqrt{\lambda}}$$
$$M_F = \mathbb{R}^+ , \Lambda(F) = \mathbb{R}^+$$

Basis Measure:

$$\mu(dx) = \exp \lambda \gamma_1$$
$$\text{where } \gamma_1(dx) = \sum_{n=0}^{\infty} \delta_n(dx) \ .$$

Cumulant Transform and Mean Domain Mapping:

$$\mathrm{k}_\mu(\theta) = \frac{\lambda}{-\theta} , \Theta(\mu) = \mathbb{R}^-$$
$$\phi_\mu(m) = -\sqrt{\frac{\lambda}{m}} , m \in M_F$$

Density:

$$f_X(x) = \exp\left(-p\frac{\theta+1}{\theta}\right) x^{-1} e^{-\beta x} \sum_{j=0}^{\infty} \frac{(px)^j}{(j+1)(j!)^2} , x \in \mathbb{R}^+$$
$$\theta = -\beta , \beta \in \mathbb{R}^+ \ ; \ \lambda = p , p \in \mathbb{R}^+ \ ; \text{then } m = \frac{p}{\beta^2} \ .$$

Legendre Transform:

$$\mathrm{k}^*_\mu(x) = -2\sqrt{\lambda x} , x \in M_F$$

Notes:

- The $\Gamma \wedge \mathrm{Po}(\beta, p)$ distribution is its own inverse.

Other References:

- Tweedie (1984) in the context of a discussion of Tweedie scale families (families F for which $\mathbf{V}_F(m) = am^b$ for some $(a, b) \in \mathbb{R} \times \mathbb{R}$).

2 Bessel Families

Variance Function:

$$V_F(m) = \frac{m^2}{\lambda^2}\sqrt{m^2 + \lambda^2}$$
$$M_F = \mathbb{R}^+ , \Lambda(F) = \mathbb{R}^+$$

Basis Measure:

$$\mu(dx) = \lambda I_\lambda(x) x^{-1} 1\!\!1_{\mathbb{R}^+}(x) dx ,$$

where $I_\lambda(\cdot)$ is the modified Bessel function of the first kind with index λ .

Cumulant Transform and Mean Domain Mapping:

$$k_\mu(\theta) = \lambda \log\left(-\theta - \sqrt{\theta^2 - 1}\right) , \Theta(\mu) = (-\infty, -1)$$
$$\phi_\mu(m) = -\frac{\sqrt{m^2 + \lambda^2}}{m} , m \in M_F$$

Density:

$$f_X(x) = \frac{p\xi^p}{x} I_p(x) \exp\left[-\frac{1}{2}\left(\frac{1}{\xi} + \xi\right) x\right] , x \in \mathbb{R}^+$$
$$\theta = -\frac{1}{2}\left(\frac{1}{\xi} + \xi\right) , \xi \in (1, +\infty) \; ; \; \lambda = p , p \in \mathbb{R}^+ \; ; \text{ then } m = p\frac{2\xi}{\xi^2 + 1}$$

Legendre Transform:

$$k_\mu^*(x) = -\sqrt{m^2 + \lambda^2} - \log\left(\frac{\sqrt{m^2 + \lambda^2} - \lambda}{m}\right) , x \in M_F$$

Notes:

- $[(1/p)\mathrm{P} - \mathrm{P}(\lambda = 1, \cdot)]$, the Poisson Difference type, is the inverse distribution of Bessel(p, ξ).
- The density $g(x;\lambda)dx = e^{-x}\mu(dx) , \lambda \in \mathbb{N}^*$ was determined by Feller to be the density of the first-passage time through λ for a randomized random walk.
- Bessel$(p, \xi) \xrightarrow{\mathcal{D}}$ IG$(p^2, 4\xi^2/(\xi^2 + 1)^2)$ as $p \to 0$.

Other References:

- Feller (1971) p.437
- Jørgensen (1987) presents this type as an example of EDM generation from a NEF.

Appendix B

Proofs and Derivations

B.1 Preliminaries

Measures Concentrated on $\mathbb{N}$

We reproduce here part of Proposition 4.4 from Letac and Mora (1990), which will help us to determine whether inverse families exist in certain cases.

Theorem B.1 *Let F be a NEF concentrated on $\mathbb{R}$ with variance function $\mathbf{V}$ defined on M_F. Then $F = F(\mu)$ is concentrated on $\mathbb{N}$ such that $\mu\{0\}, \mu\{1\} > 0$ if and only if*

1. *$M_F = (0, b)$ for some $0 < b \geq +\infty$;*
2. *There exists an open subset A of the complex plane containing $[0, b)$ and an analytic function ϕ' on A such that $\phi'(m) = m/\mathbf{V}(m)$ if $0 < m < b$ and such that $\phi'(0) = 1$.*

Lagrange Expansion

Let g be a mapping analytic at 0 and such that $g(0) \neq 0$. Consider the transformation $u = \dfrac{t}{g(t)}$. If f is analytic in a neighborhood of 0, then

$$f(t(u)) = f(0) + \sum_{j=1}^{\infty} \frac{u^j}{j!} \left[\frac{d^{j-1}}{dt^{j-1}} f'(t) g(t)^j \right] \Bigg|_{t=0}, \tag{B.1}$$

called the *Lagrange expansion* (sometimes *Lagrange expansion of the first kind*) of f in terms of u. See for instance Consul and Shenton (1972) for an application to the generation of probability distributions.

(Mora [1986, ch. II, Prop. 4.1] provides a streamlined version of the Lagrange expansion formula, directly applicable to the computation of variance functions for the discrete Mora class families, among others.)

Meixner Polynomials

From Meixner (1941) we get that with $k < 0$, $\alpha + \beta \in \mathbb{R}$ $\alpha\beta \geq 0$, if

$$\sum_{j=0}^{\infty} P_j(x) \frac{z^j}{j!} = \frac{(1 - \beta z)^{\frac{\beta x + k}{\beta(\alpha - \beta)}}}{(1 - \alpha z)^{\frac{\alpha x + k}{\alpha(\alpha - \beta)}}} \tag{B.2}$$

then

$$P_j(x) = \binom{k/\alpha\beta}{j}(-\beta)^j F\left(-j; \frac{x+k/\alpha}{\alpha-\beta}; -k/\alpha\beta; 1-\alpha/\beta\right) \quad (B.3)$$

where $F(\ldots)$ is the hypergeometric function. We call $P_j(x)$ the Meixner polynomial.

Erdélyi (1953, §10.24) defines the Meixner polynomial as

$$\begin{aligned} P_j^{\beta,c}(x) &= (\beta+x)_j F\left(-j,-x;1-\beta-j-x;c^{-1}\right) \\ &= (\beta)_j F(-j,-x;\beta;1-c^{-1}) \end{aligned} \quad (B.4)$$

for $\beta > 0$ and $0 < c < 1$. In this case, it has generating function

$$\sum_{j=0}^{\infty} P_j^{\beta,c}(x)\frac{z^j}{j!} = (1-z/c)^x(1-z)^{-x-\beta} \quad (B.5)$$

for $|z| < \min(1,|c|)$. We shall use identities B.3 and B.4 associated respectively with generating functions B.2 and B.5 according to convenience.

Jacobi Polynomials

Erdélyi (1953, §10.8) gives the following explicit expression for the Jacobi polynomial:

$$P_j^{(\alpha,\beta)}(x) = 2^{-j}\sum_{k=0}^{j}\binom{j+\alpha}{k}\binom{j+\beta}{j-k}(x-1)^{j-k}(x+1)^k \quad (B.6)$$

Erdélyi's Generating Function for Laguerre Polynomials

Erdélyi (1953, §10.12) produces the following generating function for Laguerre polynomials:

$$\sum_{j=0}^{\infty} L_j^{\alpha-j} z^j = \exp(-xz)(1+z)^{\alpha} \quad (B.7)$$

for $|z| < 1$.

Meijer's G-functions

Meijer's G-functions turn out to be quite useful in the determination of closed-form densities, mainly because in spite of their apparent unwieldiness, they can easily be manipulated into inverse Laplace transforms and can represent a wide array of higher functions. We list here some key identities and properties taken or adapted from Mathai (1993, ch. 2 and 3).

Identity 1 *Rational Function*

$$\frac{z^{\beta}}{(1+az^{\alpha})^{\gamma}} = \frac{a^{-\beta/\alpha}}{\Gamma(\gamma)} G_{1,1}^{1,1}\left(az^{\alpha}\Big|_{\beta/\alpha}^{1-\gamma+\beta/\alpha}\right) \quad (B.8)$$

for $|az^{\alpha}| < 1$.

Identity 2 *Confluent Hypergeometric Function*

$$G_{1,2}^{1,1}\left(z\middle|_{\beta,\gamma}^{\alpha}\right) = \frac{\Gamma(1-\alpha+\beta)}{\Gamma(1-\gamma+\beta)} z^{\beta}{}_1F_1\left(1-\alpha+\beta; 1+\beta-\gamma; -z\right) \quad \text{(B.9)}$$

Identity 3 *Exponential Function*

$$G_{0,1}^{1,0}\left(pz^{\alpha}\middle|_{\beta/\alpha}\right) = p^{\beta/\alpha} z^{\beta} \exp\left(pz^{\alpha}\right) \quad \text{(B.10)}$$

Identity 4 *Whittaker Function*

$$G_{1,2}^{2,0}\left(z\middle|_{\beta,\gamma}^{\alpha}\right) = z^{\frac{\beta+\gamma-1}{2}} e^{-z/2} W_{\frac{\beta+\gamma-2\alpha+1}{2},\frac{\beta-\gamma}{2}}(z), \quad \text{(B.11)}$$

where $W_{\cdot,\cdot}(\cdot)$ *is the Whittaker function (see Abramowitz and Stegun [1964]).*

Property 5 *Analytic Continuation:*

$$G_{p,q}^{m,n}\left(z\middle|_{b_1,\ldots,b_q}^{a_1,\ldots,a_p}\right) = G_{q,p}^{n,m}\left(\frac{1}{z}\middle|_{1-b_1,\ldots,1-b_q}^{1-a_1,\ldots,1-a_p}\right) \quad \text{(B.12)}$$

Property 6 *Order increase/reduction:*

$$\begin{aligned} G_{p,q}^{m,n}\left(z\middle|_{b_1,\ldots,b_q}^{a_1,\ldots,a_p}\right) &= G_{p+1,q+1}^{m,n+1}\left(z\middle|_{b_1,\ldots,b_q,k}^{k,a_1,\ldots,a_p}\right) \text{ for } p,q,n \geq 0; \quad \text{(B.13)} \\ &= G_{p+1,q+1}^{m+1,n}\left(z\middle|_{k,b_1,\ldots,b_q}^{a_1,\ldots,a_p,k}\right) \text{ for } p,q,m \geq 0; \end{aligned}$$

Property 7 *Additivity of indices:*

$$z^{\alpha} G_{p,q}^{m,n}\left(z\middle|_{b_1,\ldots,b_q}^{a_1,\ldots,a_p}\right) = G_{p,q}^{m,n}\left(z\middle|_{b_1+\alpha,\ldots,b_q+\alpha}^{a_1+\alpha,\ldots,a_p+\alpha}\right) \quad \text{(B.14)}$$

Property 8 *Multiplicativity of indices:*

$$G_{p,q}^{m,n}\left(z\middle|_{b_1,\ldots,b_q}^{a_1,\ldots,a_p}\right) = (2\pi)^{(1-r)\delta} r^{u} G_{rp,rq}^{rm,rn}\left(z^{r} r^{r(p-q)}\middle|_{\Delta(r,b_1),\ldots,\Delta(r,b_q)}^{\Delta(r,a_1),\ldots,\Delta(r,a_p)}\right) \quad \text{(B.15)}$$

$$\begin{aligned} \text{where } \delta &= m+n-\frac{p+q}{2}, \\ u &= \sum_{j=1}^{q} b_j - \sum_{i=1}^{p} a_i + \frac{p-q}{2} + 1, \\ \Delta(r,a) &= \left(\frac{a}{r}, \frac{a+1}{r}, \ldots, \frac{a+r-1}{r}\right). \end{aligned}$$

Property 9 *Laplace Transform:*

$$u^{a-1} G_{p+1,q}^{m,n+1}\left(\frac{k}{u}\middle|_{b_1,\ldots,b_q}^{a,a_1,\ldots,a_p}\right) = \int_0^{\infty} e^{-ux} x^{-a} G_{p,q}^{m,n}\left(kx\middle|_{b_1,\ldots,b_q}^{a_1,\ldots,a_p}\right) dx \quad \text{(B.16)}$$

for $\Re(k) > 0$, $\delta = m+n-\frac{1}{2}(p+q) > 0$, $|\arg k| < \delta\pi$, $\Re(\min_j b_j - a) > -1$.

B.2 Extended Inversion

We make explicit the basis measure transformation involved in extended inversion (§2.4, ch. 1). The following is a reworking and extension of the work of Khan and Jain (1978), itself largely based on Kendall (1957).

Consider a measure $\mu \in \mathcal{M}$, and a measure ν defined by its Laplace transform

$$\mathrm{L}_\nu(\theta) = \mathrm{L}_\mu\left(g^{-1}(e^\theta)\right)^u,$$

where

$u \in \Lambda_{F(\mu)}$, g^{-1} is the reciprocal of $g : t \mapsto \dfrac{e^t}{\mathrm{L}_\mu(t)^r}$,
$r \in \mathbb{R}$ is such that $\nu \in \mathcal{M}$, and
$\theta \in \{t \in \mathbb{R} \mid t - rk_\mu(t) \in \Theta(\mu)\}$.

Then, writing $F = F(\mu)$ and $G = G(\nu)$, and applying the definitions of section 2 in chapter 1, it is a simple matter to show that

$$\mathbf{V}_G(m) = (rm+u)^3 \mathbf{V}_F\left(\frac{m}{rm+u}\right)$$

whence

$$r^2 \mathbf{V}_G\left(\frac{m-u}{r}\right) = r^2 m^3 \mathbf{V}_F\left(-\frac{1}{r}\left[\frac{u}{m}-1\right]\right)$$

so that, using Theorem 2.1,

$$\mathbf{V}_{[rG+u]}(m) = m^3 \mathbf{V}_{[-(r/u)F+(1/u)]}\left(\frac{1}{m}\right).$$

Letting $1/u = b \in \Lambda_G$ yields the parametrization of §2.4, ch. 1.

B.3 Mixture Measures

Under the setup described in §3.1, the Laplace transform of $\mu \wedge \nu$ is given by

$$\begin{aligned}
\mathrm{L}_{\mu\wedge\nu}(\theta) &= \int_{S(\mu)} \exp(\theta x)\left(\int_{S(\nu)} \mu^{*\lambda}\nu(d\lambda)\right)(dx) \\
&= \int_{S(\nu)}\left(\int_{S(\mu)} \exp(\theta x)\mu^{*\lambda}(dx)\right)\nu(d\lambda) \\
&= \int_{S(\nu)} \exp(\lambda \mathrm{k}_\mu(\theta))\nu(d\lambda) \text{ , for } \theta \in \Theta(\mu) \\
&= L_\nu(k_\mu(\theta)) \text{ , for } k_\mu(\theta) \in \Theta(\nu).
\end{aligned}$$

Thus $k_{\mu\wedge\nu}(\theta) = k_\nu(k_\mu(\theta))$ and $\Theta(\mu\wedge\nu) = \Theta(\mu) \cap k_\mu^{-1}(\Theta(\nu))$.

Now consider the NEF generated by $\mu \wedge \nu$, with $\theta \in \Theta_{\mu\wedge\nu}$ and $S(\nu) \in \Lambda(\mu)$.

$$\begin{aligned}
P(\theta, \mu \wedge \nu) &= [\exp(\theta x) - k_\nu (k_\mu(\theta))] \int_{S(\nu)} \mu^{*\lambda} \nu(d\lambda) \\
&= \int_{S(\nu)} \exp(\theta x - \lambda \mathrm{k}_\mu(\theta)) \mu^{*\lambda} \exp(\lambda \mathrm{k}_\mu(\theta)) - k_n u(\mathrm{k}_\mu(\theta))) \nu(d\lambda) \\
&= \int_{S(\nu)} P(\theta, \mu^{*\lambda}) P(\mathrm{k}_\mu(\theta), \nu)(d\lambda) \\
&= P(\theta, \mu) \wedge P(\mathrm{k}_\mu(\theta), \nu)
\end{aligned}$$

Since $P(\theta, \mu) = P(\theta, e^b \mu)$ for all $b \in \mathbb{R}$, it is a simple matter to see that

$$P(\theta_1, \mu) \wedge P(\theta_2, \nu) = P(\theta_1, \mu_1 \wedge \nu)$$

where $\mu_1 = \exp(\theta_2 - k_\mu(\theta_1))\mu$. Then we need $\theta \in \Theta(\mu_1 \wedge \nu)$. By putting $z = k_{\mu_1}(t) = k_\mu(t) + \theta_2 - k_\mu(\theta_1)$, we get that $k_{\mu_1}^{-1}(z) = k_\mu^{-1}(z + k_\mu(\theta_1) - \theta_2)$; since also $\Theta(\mu_1) = \Theta(\mu)$, we see that $\Theta(\mu_1 \wedge \nu) = \Theta(\mu) \cap k_\mu^{-1}(\Theta(\nu) + k_\mu(\theta_1) - \theta_2)$.

For $\mu \in \mathcal{M}$, we establish the relationship between $P(\theta, \exp(r\mu))$ (and, respectively, $P\left(\theta, (\delta_0 - \mu)^{*(-\lambda)}\right)$) and the mixture distribution $P(\theta, \mu^{*J}) \wedge_J \mathrm{Po}(\xi)$ (respectively, $P(\theta, \mu^{*J}) \wedge_J \mathrm{NB}(r, p)$) directly.

In the Poisson mixing case,

$$\begin{aligned}
P(\theta, \exp(r\mu)) &= \exp\left(\theta x - k_{\exp(r\mu)}(\theta)\right) \exp(r\mu) \\
&= \exp(\theta x - r\mathrm{L}_\mu(\theta)) \sum_{j=0}^{\infty} \frac{r^j}{j!} \mu^{*j} \\
&= \sum_{j=0}^{\infty} \exp(-r\mathrm{L}_\mu(\theta)) \frac{\left(re^{\mathrm{k}_\mu(\theta)}\right)^j}{j!} \exp(\theta x - j\mathrm{k}_\mu(\theta)) \mu^{*j} \\
&= \sum_{j=0}^{\infty} \exp(-r\mathrm{L}_\mu(\theta)) \frac{(r\mathrm{L}_\mu(\theta))^j}{j!} P(\theta, \mu^{*j}) \\
&= P(\theta, \mu^{*J}) \wedge_J \mathrm{Po}(r\mathrm{L}_\mu(\theta))
\end{aligned}$$

To determine the NEF in canonical form produced by mixing $P(\theta, \mu^{*j})$ with $\mathrm{Po}(\xi)$, $\xi > 0$ arbitrary, the parameter r may be adjusted appropriately.

To mix with a Pascal probability measure, consider

$$\begin{aligned}
P\left(\theta, (\delta_0 - \mu)^{*(-r)}\right) &= \exp\left(\theta x - k_{(\delta_0 - \mu)^{*(-r)}}(\theta)\right) (\delta_0 - \mu)^{*(-r)} \\
&= \exp(\theta x + r \log(1 - \mathrm{L}_\mu(\theta))) \sum_{j=0}^{\infty} \binom{r+j-1}{r-1} \mu^{*j} \\
&= \sum_{j=0}^{\infty} \binom{r+j-1}{r-1} \exp(r \log[1 - \mathrm{L}_\mu(\theta)]) \\
&\qquad \exp(j\mathrm{k}_\mu(\theta)) \exp(\theta x - j\mathrm{k}_\mu(\theta)) \mu^{*j} \\
&= \sum_{j=0}^{\infty} \binom{r+j-1}{r-1} [1 - \mathrm{L}_\mu(\theta)]^r \mathrm{L}_\mu(\theta)^j P(\theta, \mu^{*j}) \\
&= P(\theta, \mu^{*J}) \wedge_J \mathrm{NB}(r, 1 - \mathrm{L}_\mu(\theta))
\end{aligned}$$

To determine the NEF in canonical form produced by mixing $P(\theta, \mu^{*j})$ with $\mathrm{NB}(r,p)$, $p \in (0,1)$ arbitrary, the basis measure must be adjusted appropriately, viz.

$$P(\theta, \mu^{*J}) \wedge_J \mathrm{NB}(r,p) = P\left(\theta, (\delta_0 - \mu_1)^{*(-r)}\right)$$

where $\mu_1 = \left(\dfrac{1-p}{\mathrm{L}_\mu(\theta)}\right)\mu$.

B.4 Basis Measure for the Hermite Families

Using Szegö's (1939) definition of the Hermite polynomials,

$$\sum_{j=0}^{\infty} H_j(x)\frac{z^j}{j!} = \exp(2xz - z^2)$$

so that if we let $H_j^*(x) = i^{-n}H_j(ix)$ (so that $H_j^*(x)$ has coefficients equal to the coefficients of $H_j(x)$ in absolute value),

$$\sum_{j=0}^{\infty} H_j^*(x)\frac{z^j}{j!} = \exp(2xz + z^2).$$

If we now let $z = \sqrt{\lambda/2}e^\theta$ and $x = \sqrt{\lambda/2}$, we get

$$\sum_{j=0}^{\infty} H_j^*(\sqrt{\lambda/2})\left(\sqrt{\lambda/2}e^\theta\right)^j = \exp\left(\lambda e^\theta + \lambda\frac{e^{2\theta}}{2}\right).$$

The explicit form of the basis measure for the Hermite distribution with power λ is thus

$$\mu = \sum_{j=0}^{\infty}\left(\frac{\lambda}{2}\right)^{j/2} H_j^*\left(\sqrt{\frac{\lambda}{2}}\right).$$

B.5 Inversion and the Hermite Families

Proposition B.1 *The Hermite families have no inverse.*

Proof. If F is the inverse family, then $\mathbf{V}_F(m) = \lambda^{-2}\mathbf{V}_G(\lambda m)$, where $\mathbf{V}_G(m) = m(m^2+4m) - m^2\sqrt{m^2+4m}$. Thus we can take $\phi_G(m) = \frac{1}{4}\log(m^2 + m\sqrt{m^2+4m})$ and $k_G(\theta) = \frac{1}{2}\sqrt{1+\exp(4\theta)/2}$. Up to an affinity, we can therefore take the Laplace transform of the inverse family to be $\mathrm{L}_{F_1}(\theta) = \exp\left(\dfrac{1}{2}\sqrt{1+e^\theta/2}\right)$, which in turn corresponds to a variance function $\mathbf{V}_{F_1}(m) = m\Delta(m) - 2m^2\sqrt{\Delta(m)}$, with $\Delta(m) = 4m^2+1$.

Now $\mathbb{R}^+ \supset M_{F_1} = M_G$, since $M_F = \mathbb{R}^+$ and the map $M_F \to M_G, m \mapsto 1/m$ is bijective. Also, $V_{F_1}(m) < 0$ for $m < 0$, so that $M_{F_1} = \mathbb{R}^+$. Hence a straightforward application of theorem B.1 shows that F_1 is concentrated on $\mathbb{N}$ if F_1 is a NEF.

Computation of the first terms of the entire series expansion

$$\sum_{j=0}^{\infty} a_j z^j = \frac{1}{2}\exp\left(\frac{1}{2}\sqrt{1+z/2}\right)$$

shows that $a_0 = e^{1/2}$, $a_1 = e^{1/2}/8$ as required by the theorem, but $a_2 = -e^{1/2}/128$. Hence F_1 is not a NEF, and the Hermite type does not have an inverse. □

B.6 Letac's Second Mixed Geometric Example

Let $P[X_i = 1] = 1 - \kappa_1$ and $P[X_i = 2] = \kappa_1$, $\kappa_1 \in (0,1)$, $i = 1, 2, \ldots$. Then if $S_n = \sum_{i=1}^{n} X_i$, $S_n = T_n + n$, where $T_n \sim \text{Bin}(n, \kappa_1)$. Hence $M_{S_n}(t) = e^{tn} M_{T_n}(t)$ where $M_{T_n}(t) = (1 - \kappa_1 + \kappa_1 e^t)^n$. Now let $P[N = n] = (1-\kappa_2)\kappa_2^n$, $n = 0, 1, \ldots$, so that $N \sim [\text{NB}(1, 1-\kappa_2) - 1]$. Then

$$\begin{aligned} M_{S_N}(t) &= (1-\kappa_2)\sum_{n=0}^{\infty} M_{S_n}(t)\kappa_2^n \\ &= (1-\kappa_2)\sum_{n=0}^{\infty}\left[\kappa_2 e^t(1-\kappa_1+\kappa_1 e^t)\right]^n \\ &= \frac{(1-\kappa_2)}{1-\kappa_2(1-\kappa_1)e^t - \kappa_2\kappa_1 e^{2t}} \end{aligned}$$

Now setting $\kappa_2(1-\kappa_1) = ae^\theta$ and $\kappa_2\kappa_1 = ae^{2\theta}$ and solving, we get from $\log M_{S_N}(t) = k_\mu(\theta + t) - k_\mu(\theta)$ that

$$\begin{aligned} L_\mu(\theta) &= \frac{1}{1-(1-a)e^\theta - ae^{2\theta}} \\ &= L_{e^{*a}_{-a} * g_{1-a}}(\theta). \end{aligned}$$

This completes the proof.

B.7 Length-Biased Distribution of a Pascal Random Variable and Seshadri's Mixture NEFs

If $X \sim P(\theta, \mu)$, then the length-biased distribution of X is given by $P^*(\theta, \mu) = \dfrac{xP(\theta,\mu)}{m}$, where m is the expected value of X. Thus if

$$P(\theta(p), \mu_r) = \begin{pmatrix} x+r-1 \\ r-1 \end{pmatrix} p^r(1-p)^x \delta_x = [\text{NB}(r, p)],$$

we get

$$P^*(\theta(p), \mu_r) = \begin{pmatrix} x+r-1 \\ r \end{pmatrix} p^{r+1}(1-p)^{x-1}\delta_x$$

for $x \in \mathbb{N}$, since $\mathbb{E}[X] = r(1-p)/p$. Thus $P^*(\theta(p), \mu_r) = [\text{NB}(r+1, p) + 1]$, since if $Y \sim P^*(\theta, \mu)dx$, then $P[Y = x] = P[X = x-1] = P[X+1 = x]$ so that $Y \stackrel{\mathcal{D}}{=} (X+1) \sim [\text{NB}(r+1,p)+1]$.

In the case mentioned under the Mixed Geometric type, we have

$$\begin{aligned} P_\omega(\theta, \nu) &= \frac{\omega'}{\omega'+m} P(\theta, \mu) + \frac{m}{\omega'+m} P^*(\theta, \mu) \\ &= \frac{\omega p}{\omega p + 1 - p}[\text{NB}(r,p) - r] + \frac{1-p}{\omega p + 1 - p}[\text{NB}(r+1,p) - r] \end{aligned}$$

where $\omega' = r\omega$, yielding a Mixed Geometric type, a Pascal type or a NB+NB type as $\omega > 1$, $\omega = 1$ or $\omega < 1$ respectively (from the form of the mixture variance function; see Seshadri [1991]). Looking more closely at the case $\omega = 1$, we get that in fact

$$\begin{aligned} P_1(\theta(p), \nu_r) &= p[\text{NB}(r,p)] + (1-p)[\text{NB}(r+1,p)+1] \\ &= \binom{x+r-1}{r-1} p^{r+1}(1-p)^x + \binom{x+r-1}{r} p^{r+1}(1-p)^x \\ &= \binom{x+r}{r} p^{r+1}(1-p)^x \\ &= [\text{NB}(r+1,p)] \end{aligned}$$

as required by the variance function.

B.8 Density for a Special Case of the Mixed Geometric Type

Assume that $\frac{r_1}{r_1+r_2} < a < 1$. We wish to find $C_a^{r_1,r_2}(j)$, $j \in \mathbb{N}$, such that

$$\sum_{j=0}^{\infty} C_a^{r_1,r_2}(j) z^j = \frac{(1-az)^{r_1/a}}{(1-z)^{r_1+r_2}}.$$

Let $c = 1/a$, $x = r_1/a$ and $\beta = r_1 + r_2 - r_1/a$, , we see that $0 < c < 1$ and $\beta > 0$. We can therefore use equations B.5 and B.4 to show that

$$C_a^{r_1,r_2}(j) = \binom{r_1+r_2}{j} F(-j, -r_1/a; 1 - r_1 - r_2 - j, a)$$

The final form of the density is thus

$$\begin{aligned} f_X(x) &= \frac{(1-\xi)^{r_1+r_2}}{(1-a\xi)^{r_1/a}} \xi^x C_a^{r_1,r_2}(x) \\ &= F(-x, -r_1/a; 1 - r_1 - r_2 - x, a) \binom{r_1+r_2}{x} \frac{(1-\xi)^{r_1+r_2}}{(1-a\xi)^{r_1/a}} \xi^x \end{aligned}$$

where F is the hypergeometric function.

B.9 Density for the Mixed Exponential Type

We need to find $\mu(dx) = E_{a,\lambda,q}(x)(dx)$ such that

$$\mathrm{L}(t) = \int_0^\infty e^{-tx} E_{a,\lambda,q}(x)dx = \frac{(1+at)^\lambda}{t^{\lambda(q+1)}}\,,\; t \in \mathbb{R}^+ \,.$$

Now with $x = ay$ and

$$E_{a,\lambda,q}(x) = a^{\lambda(q+1)-1} E^*_{a,\lambda,q}(x/a)\,, \tag{B.17}$$

we have

$$\int_0^\infty e^{-ty} E^*_{a,\lambda,q}(y)dx = \frac{(1+t)^\lambda}{t^{\lambda(q+1)}}\,.$$

We let $u = u(t) = (1+t)^{-1}$, so that $0 < u < 1$ and $t = t(u) = (1-u)/u$. We denote $\mathrm{L}^*(u) = \mathrm{L}\,(t(u))$; then

$$\begin{aligned}\mathrm{L}^*(u) &= \frac{u^{\lambda q}}{(1-u)^{\lambda(q+1)}} \\ &= u\frac{u^{\lambda q-1}}{(1-u)^{\lambda(q+1)}}\end{aligned}$$

From equation B.8,

$$\frac{u^{\lambda q-1}}{(1-u)^{\lambda(q+1)}} = \frac{(-1)^{\lambda q-1}}{\Gamma[\lambda(q+1)]} G^{1,1}_{1,1}\left(-u\big|^{1-\lambda(q+1)+\lambda q-1}_{\lambda q-1}\right)$$

where G is Meijer's G-function, whence we get

$$\mathrm{L}^*(u) = \frac{(-1)^{\lambda q-1}}{\Gamma[\lambda(q+1)]} u G^{1,1}_{1,1}\left(-u\big|^{-\lambda}_{\lambda q-1}\right)\,.$$

From equation B.13,

$$G^{1,1}_{1,1}\left(-u\big|^{-\lambda}_{\lambda q-1}\right) = G^{1,2}_{2,2}\left(-u\big|^{0,-\lambda}_{\lambda q-1,0}\right)$$

and thus

$$\mathrm{L}^*(u) = \frac{(-1)^{\lambda q-1}}{\Gamma[\lambda(q+1)]}\left\{ u G^{1,2}_{2,2}\left(\frac{-1}{u^{-1}}\big|^{0,-\lambda}_{\lambda q-1,0}\right)\right\}\,.$$

We now use equation B.16, which governs the Laplace transform of G-functions. Taking

$$\begin{gathered}\omega = u^{-1}\;,\text{ so } \Re(\omega) > 0\,, \\ m = n = p = 1\,, p = 2\;,\text{ so } \delta = 1/2 > 0\,, \\ \eta = -1\;,\text{ so } |\arg\eta| = 1 < \delta\pi\;,\text{ and} \\ \sigma = 1\;,\text{ so } \Re(\sigma + \min_j b_j) = 1 > 0\end{gathered}$$

we get

$$\begin{aligned}(u^{-1})^{-1} G^{1,2}_{2,2}\left(\frac{-1}{u^{-1}}\big|^{0,-\lambda}_{\lambda q-1,0}\right) &= \int_0^\infty e^{-u^{-1}y} G^{1,1}_{1,2}\left(-y\big|^{-\lambda}_{\lambda q-1,0}\right)dx \\ &= \int_0^\infty e^{-ty}e^{-y} G^{1,1}_{1,2}\left(-y\big|^{-\lambda}_{\lambda q-1,0}\right)dx\end{aligned}$$

so that we can put, by uniqueness of the Laplace transform,

$$E^*_{a,\lambda,q}(x) = \frac{(-1)^{\lambda q-1}}{\Gamma[\lambda(q+1)]} e^{-y} G^{1,1}_{1,2}\left(-y\big|^{-\lambda}_{\lambda q-1,0}\right) \tag{B.18}$$

From equation B.14,

$$G^{1,1}_{1,2}\left(-y\big|^{-\lambda}_{\lambda q-1,0}\right) = (-y)^{\lambda q-1} G^{1,1}_{1,2}\left(-y\big|^{1-\lambda(q+1)}_{0,1-\lambda q}\right) \tag{B.19}$$

and from equation B.9,

$$G^{1,1}_{1,2}\left(-y\big|^{1-\lambda(q+1)}_{0,1-\lambda q}\right) = \frac{\Gamma[\lambda(q+1)]}{\Gamma[\lambda q]} {}_1F_1[\lambda(q+1);\lambda q; y], \tag{B.20}$$

where ${}_1F_1$ is the confluent hypergeometric function. Combining equations B.19 and B.20 with equation B.18, we get

$$E^*_{a,\lambda,q}(y) = \frac{y^{\lambda q-1} e^{-y}}{\Gamma(\lambda q)} {}_1F_1[\lambda(q+1);\lambda q; y]$$

and finally, from B.17

$$E_{a,\lambda,q}(x) = a^{\lambda} \frac{x^{\lambda q-1} e^{x/a}}{\Gamma(\lambda q)} {}_1F_1[\lambda(q+1);\lambda q; x/a] \ .$$

Thus we can write the density f_X of a Mixed Exponential random variable as

$$\begin{aligned} f_X(x) &= \left(\frac{a\xi^{q+1}}{1+a\xi}\right)^{\lambda} \frac{x^{\lambda q-1}}{\Gamma(\lambda q)} e^{-(\xi+1/a)x} {}_1F_1[\lambda(q+1);\lambda q; x/a] \\ &= \frac{\xi^{r_1+r_2}}{(\alpha+\xi)^{r_1}} \frac{x^{r_2-1}}{\Gamma(r_2)} e^{-(\xi+\alpha)x} {}_1F_1[r_1+r_2; r_2; \alpha x] \end{aligned}$$

where $r_1 = \lambda$, $r_2 = \lambda q$, $\alpha = 1/a$ and $\xi = -\theta$.

B.10 Density for the Trinomial Families

We wish to find a_k such that $\sum_{k=-n}^{n} a_k z^k = (z + 1/z + 2a)^n$, for $n \in \mathbb{N}$. We define the symbols $\mathcal{E}p$ and $\mathcal{O}p$ as the nearest even, respectively odd, integer not exceeding p.

$$\begin{aligned} (z+1/z+2a)^n &= \sum_{i=0}^{n} \binom{n}{i} \left(z+\frac{1}{z}\right)^i (2a)^{n-i} \\ &= (2a)^n \sum_{i=0}^{n} \binom{n}{i} (2a)^{-i} \sum_{j=0}^{i} \binom{i}{j} z^{2j-i} \end{aligned}$$

Now let $k = 2j - i$, so that $k \in \{-i, -i+2, \ldots, i\}$. Then

$$\begin{aligned}
(z + 1/z + 2a)^n &= (2a)^n \sum_{i=0}^{n} \binom{n}{i} (2a)^{-i} \sum_{k=-i,2}^{i} \binom{i}{\frac{k+i}{2}} z^k \\
&= (2a)^n \left[\sum_{i=0}^{\frac{\mathcal{E}n}{2}} \binom{n}{2i} (2a)^{-2i} \sum_{k=-2i,2}^{2i} \binom{2i}{\frac{k+2i}{2}} z^k + \right. \\
&\quad \left. \sum_{i=0}^{\frac{\mathcal{O}n-1}{2}} \binom{n}{2i+1} (2a)^{-2i-1} \sum_{k=-2i-1,2}^{2i+1} \binom{2i+1}{\frac{k+2i+1}{2}} z^k \right] \\
&= (2a)^n \left[\sum_{k=-\mathcal{E}n,2}^{\mathcal{E}n} \sum_{i=\frac{|k|}{2}}^{\frac{\mathcal{E}n}{2}} \binom{n}{2i} (2a)^{-2i} \binom{2i}{\frac{k+2i}{2}} z^k + \right. \\
&\quad \left. \sum_{k=-\mathcal{O}n,2}^{\mathcal{O}n} \sum_{i=\frac{|k-1|}{2}}^{\frac{\mathcal{O}n-1}{2}} \binom{n}{2i+1} (2a)^{-2i-1} \binom{2i+1}{\frac{k+2i+1}{2}} z^k \right]
\end{aligned}$$

Thus the coefficients are given by

$$a_k = (2a)^n \sum_{i=|k|,2}^{n} \binom{n}{i} \binom{i}{\frac{k+i}{2}} (2a)^{-i}.$$

B.11 Density for the Pascal Sum Families

The Mellin transform of the basis measure of a Pascal sum family is given by

$$\sum_{j=0}^{\infty} {}_1A^{r_1,r_2}(j) z^j = \mathrm{M}(z) = \frac{(1 - z/a)^{-\lambda q}}{(1 - az)^{\lambda}}.$$

From equation B.3 above, putting $\alpha = a$, $\beta = 1/a$, $x = \lambda(q + a^2)/a$ and $k = -\lambda(q+1)$, we have that if

$$P_n(\lambda, q, a) = \binom{-\lambda(q+1)}{n} (-a)^{-n} F(-n; \lambda; \lambda(q+1); 1 - a^2)$$

then

$$L(z) = \sum_{n=0}^{\infty} P_n(\lambda, q, a) z^n$$

Thus we obtain

$$\begin{aligned}
{}_1A_a^{r_1,r_2}(x) &= \binom{-(r_1 + r_2)}{x} (-a)^{-x} F(-x; r_1; r_1 + r_2; 1 - a^2) \\
&= \binom{r_1 + r_2 + x - 1}{x} a^{-x} F(-x; r_1; r_1 + r_2; 1 - a^2)
\end{aligned}$$

The density is then given by

$$\begin{aligned} f_X(x) &= p_1^{r_1} p_2^{r_2} \, {}_1A^{r_1,r_2}_{a(p_1,p_2)}(x)[(1-p_1)(1-p_2)]^{x/2} \\ &= p_1^{r_1} p_2^{r_2} \binom{r_1+r_2+x-1}{x} \left(\frac{1-p_1}{1-p_2}\right)^{-x/2} \\ &\quad F\left(-x; r_1; r_1+r_2; \frac{p_1-p_2}{1-p_2}\right) [(1-p_1)(1-p_2)]^{x/2} \\ &= p_1^{r_1} p_2^{r_2} \binom{r_1+r_2+x-1}{x} (1-p_2)^x F\left(-x; r_1; r_1+r_2; \frac{p_1-p_2}{1-p_2}\right) \end{aligned}$$

The density is of course symmetrical in subscripts 1 and 2, as can easily be shown by starting with $M(z) = \dfrac{(1-az)^{-\lambda}}{(1-z/a)^{\lambda q}}$.

B.12 Density for the Poisson-Pascal Sum Families

We wish to find ${}_4A^r_a(j)$, $j \in \mathbb{N}$, such that

$$\sum_{j=0}^{n_1+n_2} {}_4A^r_a(j) z^j = \frac{\exp(-raz)}{(1-z/a)^r}.$$

Using equation B.7, replacing z with $-z/a$ and putting $x = ra^2$ and $\alpha = -r$, we get

$$\sum_{j=0}^{\infty} L_j^{-r-j}(ra^2)(-z/a)^j = \frac{\exp(-raz)}{(1-z/a)^r},$$

hence ${}_4A^r_a(j) = (-a)^{-j} L_j^{-r-j}(ra^2)$. Since $a = \sqrt{\xi(1-p)}$, the density is given by

$$\begin{aligned} f_X(x) &= (pe^{-xi})^r \left[\xi(1-p)\right]^{x/2} {}_4A^r_a(x) \\ &= (pe^{-\xi})^r \xi^x \left[(-1)^x L_x^{-r-x}\left(\frac{r\xi}{1-p}\right)\right] \end{aligned}$$

for $x \in \mathbb{N}$.

B.13 Density for the Binomial Sum Families

We wish to find ${}_6A^{n_1,n_2}_a(j)$, $j \in \mathbb{N}$, such that

$$\sum_{j=0}^{n_1+n_2} {}_6A^{n_1,n_2}_a(j) z^j = (1+az)^{n_1}(1+z/a)^{n_2}$$

Expanding the right-hand side yields

$$\begin{aligned}
(1+az)^{n_1}(1+z/a)^{n_2} &= \left[\sum_{k=0}^{n_1}\binom{n_1}{j}(az)^j\right]\left[\sum_{k=0}^{n_2}\binom{n_2}{j}\left(\frac{z}{a}\right)^j\right] \\
&= \sum_{j=0}^{n_1+n_2}\sum_{k=0}^{j}\binom{n_1}{k}(az)^k\binom{n_2}{j-k}\left(\frac{z}{a}\right)^{j-k} \\
&= \sum_{j=0}^{n_1+n_2}\left[\sum_{k=0}^{j}\binom{n_1}{k}\binom{n_2}{j-k}a^{2k}\right]\left(\frac{z}{a}\right)^j
\end{aligned}$$

From equation B.6 with $\alpha = n_1 - j$, $\beta = n_2 - j$, $x = \dfrac{a^2+1}{a^2-1}$ and P the Jacobi polynomial we get

$$\begin{aligned}
P_j^{(n_1-j,n_2-j)}\left(\frac{a^2+1}{a^2-1}\right) &= 2^{-j}\sum_{k=0}^{j}\binom{n_1}{k}\binom{n_2}{j-k}\left(\frac{a^2+1}{a^2-1}-1\right)^{j-k} \\
&\qquad \left(\frac{a^2+1}{a^2-1}+1\right)^k \\
&= (a^2-1)^{-j}\sum_{k=0}^{j}\binom{n_1}{k}\binom{n_2}{j-k}a^{2k} \\
&= \left(\frac{a}{a^2-1}\right)^j {}_6A_a^{n_1,n_2}(j)
\end{aligned}$$

Since $a = \sqrt{\dfrac{p_1(1-p_2)}{p_2(1-p_1)}}$, the density becomes (putting $q_i = 1 - p_i, i = 1, 2$)

$$\begin{aligned}
fx(x) &= q_1^{r_1}q_2^{r_2}\left(\frac{p_1}{q_1}+\frac{p_2}{q_2}\right)^{x/2} {}_6A_a^{n_1,n_2}(x) \\
&= q_1^{r_1-x}q_2^{r_2-x}(p_2-p_1)^x P_x^{(n_1-x,n_2-x)}\left(\frac{p_1q_2+p_2q_1}{p1-p2}\right), \; x \in \mathbb{N}
\end{aligned}$$

The apparent asymmetry between subscripts 1 and 2 disappears in light of the relationship

$$P_j^{(\alpha,\beta)}(k) = (-1)^j P_j^{(\beta,\alpha)}(-k).$$

B.14 Asymptotics for the Poisson-Binomial Sum Families

Let $a = a(\xi, p) \stackrel{\text{def}}{=} \sqrt{\dfrac{(1-p)\xi}{p}}$. We then get

$$\mathbf{V}_F(m) \to \frac{1}{2n}\left((m+n)|m-n| - (m-n)^2\right) \text{ as } a \to 0$$

for fixed m and n.

Now if $m > n$, or equivalently since $m = n(\xi + p)$ if $k = \xi + p > 1$, then $\mathbf{V}_F(m) \to m - n$, so that $F(m) \to [\text{Po}(m) + n]$. Also k must remain fixed, since both m and n must themselves remain fixed. But then $a \to 0 \Leftrightarrow \frac{(1-p)\xi}{p} \to 0 \Leftrightarrow (1-p)(k-p) \to 0 \Leftrightarrow p \to 1$, this last condition being equivalent to $\xi \to k-1$ under k fixed.

If $m < n$, or equivalently if $k = \xi + p < 1$, then $\mathbf{V}_F(m) \to m - m^2/n$, so that $F(m) \to \text{Bin}(n, m/n)$. As before, with $k < 1$ fixed, $a \to 0 \Leftrightarrow (1-k+\xi)\xi \to 0 \Leftrightarrow \xi \to 0$, this last condition being equivalent to $p \to k$ under k fixed.

B.15 Density for the Gamma Sum Families

The Laplace transform of the measure $\exp(-ax)\gamma^{*r_1} * \exp(ax)\gamma^{*r_2}$, which generates the G + G distributions, is given by

$$L_{\mathrm{G+G}}(\theta) = (a-\theta)^{-r_1}(-a-\theta)^{-r_2} = a^{-r_1-r_2}(1-\theta/a)^{-r_1}(1+\theta/a)^{-r_2}\,,\ a > 0$$

with $\theta \in (-\infty, -|a|)$.

Let $u = 1 + \theta/a$, so that $u \in (0, 2)$. Then

$$\begin{aligned}
L_{\mathrm{G+G}}(\theta(u)) &= \frac{1}{a^{r_1+r_2}} \frac{u^{-r_2}}{(1-u+1)^{r_1}} \\
&= \frac{1}{a^{r_1+r_2}} \frac{1}{2^{r_1}} \frac{u^{-r_2}}{(1-u/2)^{r_1}} \\
&= \frac{1}{(2a)^{r_1+r_2}} \frac{1}{\Gamma(r_1)} G^{1,1}_{1,1}\left(-\frac{u}{2}\Big|^{1-r_1-r_2}_{-r_2}\right) && \text{from eq. B.8} \\
&= \frac{1}{(2a)^{r_1+r_2}} \frac{1}{\Gamma(r_1)} G^{1,1}_{1,1}\left(-\frac{2}{u}\Big|^{1+r_2}_{r_1+r_2}\right) && \text{from eq. B.12} \\
&= \frac{1}{(2a)^{r_1+r_2}} \frac{1}{\Gamma(r_1)} G^{1,2}_{2,2}\left(-\frac{2}{u}\Big|^{1,1+r_2}_{r_1+r_2,1}\right) && \text{from eq. B.13} \\
&= \frac{1}{(2a)^{r_1+r_2}} \frac{1}{\Gamma(r_1)} \int_0^\infty \frac{e^{-ux}}{x} G^{1,1}_{1,2}\left(-2x\Big|^{1+r_2}_{r_1+r_2,1}\right) dx && \text{from eq. B.16} \\
&= \frac{1}{(2a)^{r_1+r_2}} \frac{1}{\Gamma(r_1)} \int_0^\infty e^{-\theta(x/a)} \frac{e^{-x}}{x} G^{1,1}_{1,2}\left(-2x\Big|^{1+r_2}_{r_1+r_2,1}\right) dx \\
&= \frac{1}{(2a)^{r_1+r_2}} \frac{1}{\Gamma(r_1)} \int_0^\infty e^{-\theta y} \frac{e^{-ay}}{y} G^{1,1}_{1,2}\left(-2ay\Big|^{1+r_2}_{r_1+r_2,1}\right) dy && (y = ax)
\end{aligned}$$

Thus the generating measure μ for a G + G family can be given by

$$\begin{aligned}
\mu(dx) &= \frac{1}{(2a)^{r_1+r_2}} \frac{1}{\Gamma(r_1)} \frac{e^{-ax}}{x} G^{1,1}_{1,2}\left(-2ax\Big|^{1+r_2}_{r_1+r_2,1}\right) \\
&= \frac{e^{-ax} x^{r_1+r_2-1}}{(2a)^{r_1+r_2}\Gamma(r_1+r_2)} {}_1F_1(r_1, r_1+r_2; 2ax) && \text{from eq. B.9}
\end{aligned}$$

Reparametrizing as in 3.15 yields the final form of the density. Note that the familiar reproductivity property of the Gamma distribution in its dispersion parameter α for fixed β is obvious from the form of the density when $\beta_1 = \beta_2$.

B.16 Density for the Reciprocal Inverse Gaussian Families

We wish to find $\mu(dx)$ such that

$$\mathrm{L}_{\mathrm{RIG}}(\theta) = (-2\theta)^{-\lambda/2}\exp(-\lambda q\sqrt{-2\theta})$$

i.e. we need $E_{\lambda,q}(x)$ such that

$$\int_0^\infty e^{-zx}E_{\lambda,q}(x)dx = z^{-\lambda/2}\exp\left(-\lambda q z^{1/2}\right)$$

where $z = -2\theta > 0$. Now

$$\begin{aligned}
z^{-\lambda/2}\exp\left(-\lambda q z^{1/2}\right) &= z^{-\lambda/2}G^{1,0}_{0,1}\left(\lambda q z^{1/2}\middle|_{0}\right) && \text{from eq. B.10}\\
&= \frac{z^{-\lambda/2}}{\sqrt{\pi}}G^{2,0}_{0,2}\left(\frac{\lambda^2q^2}{4}z\middle|_{0,1/2}\right) && \text{from eq. B.15}\\
&= \frac{z^{-\lambda/2}}{\sqrt{\pi}}G^{0,2}_{2,0}\left(\frac{4}{\lambda^2q^2z}\middle|^{1,1/2}\right) && \text{from eq. B.12}\\
&= \frac{z^{-\lambda/2}}{\sqrt{\pi}}G^{0,3}_{3,1}\left(\frac{4}{\lambda^2q^2z}\middle|^{1-\lambda/2,1,1/2}_{1-\lambda/2}\right) && \text{from eq. B.14}\\
&= \frac{1}{\sqrt{\pi}}\int_0^\infty e^{-zx}x^{\lambda/2-1}G^{0,2}_{2,1}\left(\frac{4x}{\lambda^2q^2}\middle|^{1,1/2}_{1-\lambda/2}\right)dx && \text{from eq. B.16}\\
&= \frac{1}{\sqrt{2^\lambda\pi}}\int_0^\infty e^{\theta y}y^{\lambda/2-1}G^{0,2}_{2,1}\left(\frac{2y}{\lambda^2q^2}\middle|^{1,1/2}_{1-\lambda/2}\right)dx && (y = 2x)
\end{aligned}$$

Hence

$$\begin{aligned}
\mu(dx) &= \frac{1}{\sqrt{2^\lambda\pi}}x^{\lambda/2-1}G^{2,0}_{1,2}\left(\frac{\lambda^2q^2}{2x}\middle|^{\lambda/2}_{0,1/2}\right)dx && \text{from eq. B.12}\\
&= \frac{1}{\sqrt{2^\lambda\pi}}x^{\lambda/2-1}\left(\frac{\lambda^2q^2}{2x}\right)^{-1/4}\exp\left(-\frac{\lambda^2q^2}{4x}\right)W_{\frac{3-2\lambda}{4},-\frac{1}{4}}\left(\frac{\lambda^2q^2}{2x}\right)dx && \text{from eq. B.11}
\end{aligned}$$

with W the Whittaker function.

From Abramowitz and Stegun (1964)

$$
\begin{aligned}
W_{\frac{3-2\lambda}{4},-\frac{1}{4}}(z) &= \frac{\Gamma(1/2)}{\Gamma(\lambda/2)} z^{1/4}\exp(-z/2)\,{}_1F_1\left(\frac{\lambda-1}{2};\frac{1}{2};z\right) \\
&\quad + \frac{\Gamma(-1/2)}{\Gamma(\lambda/2-1/2)} z^{3/4}\exp(-z/2){}_1F_1\left(\frac{\lambda}{2};\frac{3}{2};z\right) \\
&= \frac{\Gamma(1/2)}{\Gamma(\lambda/2)} z^{1/4}\exp(-z/2)\left[{}_1F_1\left(\frac{\lambda-1}{2};\frac{1}{2};z\right)+\right. \\
&\quad \left.(1-\lambda)\frac{\Gamma(\lambda/2)}{\Gamma(\lambda/2+1/2)} z^{1/2}{}_1F_1\left(\frac{\lambda}{2};\frac{3}{2};z\right)\right]
\end{aligned}
$$

Simplifying, we get

$$
\begin{aligned}
\mu(dx) &= \frac{x^{\lambda/2-1}}{2^{\lambda/2}\Gamma(\lambda/2)}\exp\left(-\frac{\lambda^2q^2}{2x}\right)\left[{}_1F_1\left(\frac{\lambda-1}{2};\frac{1}{2};\frac{\lambda^2q^2}{2x}\right)+\right. \\
&\quad \left.(1-\lambda)\frac{\lambda q}{\sqrt{2x}}\frac{\Gamma(\lambda/2)}{\Gamma(\lambda/2+1/2)}{}_1F_1\left(\frac{\lambda}{2};\frac{3}{2};\frac{\lambda^2q^2}{2x}\right)\right]
\end{aligned}
$$

B.17 Explicit Measure for the S-Abel Families

We determine an explicit expansion for the measure μ described in section 4.1 of chapter 2. The Laplace transform of this measure is given by

$$L_\mu(\theta) = \frac{\exp(r_2 f^{-1}(\theta))}{(1-f^{-1}(e^\theta))^{r_1}}$$

where $r_1 = \lambda > 0$, $r_2 = \lambda q \geq -r_1$, $0 < t < 1$ and

$$u = f(t) = \frac{t}{e^t}$$

If we write $g(t) = e^{r_2 t}(1-t)^{-r_1}$, then the Lagrange expansion (eq. B.1) of g in terms of u yields

$$
\begin{aligned}
g(t(u)) &= g(0) + \sum_{j=1}^{\infty}\left[\frac{d^{j-1}}{dt^{j-1}}e^{jt}g'(t)\right]_{t=0}\frac{u^j}{j!} \\
&= g(0) + \sum_{j=1}^{\infty}\left[\frac{d^{j-1}}{dt^{j-1}}\frac{r_1 e^{(j+r_2)t}}{(1-t)^{r_1+1}}\right]_{t=0}\frac{u^j}{j!}
\end{aligned}
$$

It is easily shown by induction that

$$\frac{d^i}{dt^i}\frac{e^{rt}}{(1-t)^s} = \sum_{k=0}^{i}\binom{i}{k} r^k s^{(i-k)}\frac{e^{rt}}{(1-t)^{s+i-k}}$$

where $a^{(n)} = \dfrac{\Gamma(a+n)}{\Gamma(a)}$. Hence

$$g(t(u)) = 1 + \sum_{j=1}^{\infty}\left[r_1\sum_{k=0}^{j-1}\binom{j-1}{k}(j+r_2)^k(r_1+1)^{(j-k-1)}\right]\frac{u^j}{j!}$$

$$= 1 + \sum_{j=1}^{\infty} \left[r_1 \sum_{k=0}^{j-1} \frac{(r_1 + j - k - 1)!}{r_1!(j-k-1)!} \frac{j^k}{k!} \right] \frac{u^j}{j}$$

$$= 1 + \sum_{j=1}^{\infty} \frac{r_1}{j!} [(j + r_2) + (r_1 + 1)^{\sim}]^{j-1} u^j$$

where we introduce the notation

$$[a + b^{\sim}]^n = \sum_{k=0}^{n} a^k (b)^{(n-i)}. \tag{B.21}$$

Note that in particular, for $r_2 = 0$ and $r_1 = 1$ we get

$$\begin{aligned}
\frac{1}{1 - t(u)} &= 1 + \sum_{j=1}^{\infty} \left[\sum_{k=0}^{j-1} \frac{(j-k)j^{k-1}}{k!} \right] u^j \\
&= 1 + \sum_{j=1}^{\infty} \left[\sum_{k=0}^{j-1} \frac{j^k}{k!} - \sum_{k=1}^{j-1} \frac{j^{k-1}}{(k-1)!} \right] u^j \\
&= 1 + \sum_{j=1}^{\infty} \left[\sum_{k=0}^{j-1} \frac{j^k}{k!} - \sum_{k=0}^{j-2} \frac{j^k}{k!} \right] u^j \\
&= 1 + \sum_{j=1}^{\infty} \frac{j^{j-1}}{(j-1)!} u^j \\
&= 1 + \sum_{j=1}^{\infty} \frac{j^j}{j!} u^j
\end{aligned}$$

B.18 Explicit Measure for the S-Takács Families

Here we derive an explicit measure for μ as defined in section 4.2 in chapter 2. Consider first $f_{b,c,d}(t) = (1+t)^b (1-ct)^{-d}$, with $b > -d$ and $d > 0$. It is easy to show by induction that

$$\frac{d^i}{dt^i} f(t) = \sum_{k=0}^{i} \binom{i}{k} b_{(k)} c^{i-k} d^{(i-k)} k_{b-k,c,d+i-k}(t)$$

so that

$$\frac{d^i}{dt^i} f(0) = [b_{\sim} + cd^{\sim}]^i ,$$

introducing an obvious notation analogous to that of eq. B.21, where $a^{(n)} = \dfrac{\Gamma(a+n)}{\Gamma(a)}$ and $a_{(n)} = \dfrac{\Gamma(a+1)}{\Gamma(a-n+1)}$.

Now if we let $u = \dfrac{t}{(1+t)^{a+1}}$ and $g(t) = \dfrac{(1+t)^{r_2}}{(1-at)^{r_1}} = f_{r_2,a,r_1}(t)$, with $r_2 \geq -r_1$, the Lagrange expansion (B.1) for g in terms of u becomes immediately available since

$$g(t(u)) = g(0) + \sum_{j=1}^{\infty} \frac{u^j}{j!} \left[\frac{d^{j-1}}{dt^{j-1}} f'_{r_2,a,r_1}(t)(1+t)^{(a+1)j} \right]_{t=0}$$

$$= 1 + \sum_{j=1}^{\infty} \frac{u^j}{j!} \left[\frac{d^{j-1}}{dt^{j-1}} r_2 f_{r_2-1+(a+1)j,a,r_1}(t) + \frac{d^{j-1}}{dt^{j-1}} a r_1 f_{r_2+(a+1)j,a,r_1-1}(t) \right]_{t=0}$$

$$= 1 + \sum_{j=1}^{\infty} \left(\frac{1}{j!} r_2 \left[(r_2 - 1 + (a+1)j)_{\sim} + a r_1^{\sim} \right]^{j-1} + a r_1 \left(r_2 + (a+1)j \right)_{\sim} + a (r_1 - 1)^{\sim}]^{j-1} \right) u^j$$

from which an explicit form for the measure may be determined to produce a density.

Appendix C

Notation

Notation used in the compendium

$\mathbb{1}_A(x)$	indicator function: $\mathbb{1}_A(x) = \begin{cases} 1 & \text{if } x \in A \\ 0 & \text{if } x \notin A \end{cases}$
$(a)^{(k)}$	$\dfrac{\Gamma(a+k)}{\Gamma(a)}$
$(a)_{(k)}$	$\dfrac{\Gamma(a+1)}{\Gamma(a-k+1)}$
$A \xrightarrow{\mathcal{D}} B$	family, distribution, random variable, A converges weakly to B
$\bar{a}, \bar{a}_n, \bar{a}_A$	arithmetic mean of a collection $a_1, a_2, \ldots, a_n$
$\bar{a}_G$	geometric mean of a collection $a_1, a_2, \ldots, a_n$
$[aA(\cdot)+b]$	affine density, distribution or family of probability measures; if $X \sim [aA(\cdot)+b]$, $(X-b)/a \sim A(\cdot)$ (ch. 1, §2.1)
$A \stackrel{\mathcal{D}}{=} B$	A and B belong to the same distribution
$\begin{pmatrix} a \\ b \end{pmatrix}$	$\dfrac{\Gamma(a+1)}{\Gamma(b+1)\Gamma(a-b+1)}$
$\boldsymbol{A}', \boldsymbol{a}'$	transpose of the matrix $\boldsymbol{A}$, the vector $\boldsymbol{a}$
$A * \mu$	image of a measure μ under an affinity A (ch.1, §1.1)
$\mu * \nu$	convolution of two measures μ and ν (ch.1, §1.1)
$\mu^{*\lambda}$	λth power of a measure μ (ch.1, §1.1)
$(\delta_0 - \mu)^{*(-\lambda)}$	λth power of the geometric expansion of μ (ch.1, §1.1)
$\sum_{k=j,i}^{n} a_k$	$a_j + a_{j+i} + a_{j+2i} + \ldots + a_{n-(n-j)(\text{mod } i)}$
$X \sim A$	random variable X has distribution, probability measure family or density A, depending on the context
$(ab_\sim + cd^\sim)^n$	notation for $\sum_{j=0}^{n} \begin{pmatrix} n \\ j \end{pmatrix} a^j b_{(j)} c^{n-j} d^{(n-j)}$
$\mu_r \wedge_r \nu$	mixture measure of the measure set $\{\mu_r \mid r \in \mathcal{I})\}$ by the measure ν defined on $\mathcal{I}$ (ch.1, §3.1)
$\mathrm{Arc}(\cdot,\cdot,\cdot)$	Arcsine distribution (ch.2, §2.3)
$\mathrm{B}+\mathrm{B}(\cdot,\cdot,\cdot,\cdot)$	Binomial Sum distribution (ch.2, §3.12)
$\mathrm{B}+\mathrm{NB}(\cdot,\cdot,\cdot,\cdot)$	Pascal-Binomial Sum distribution (ch.2, §3.9)

Notation used in the compendium (continued)

$\text{Bessel}(\cdot,\cdot)$	Bessel distribution (ch.2, §2)
$\text{Bin}(\cdot,\cdot)$	Binomial distribution (ch.2, §1.3)
CCM	canonical caste member (ch.1, §2.6)
$\text{CF}, \text{CF}(F)$	convolution family generated by F (ch.1, §2.3)
$D(\mu)$	for $\mu \in \mathcal{M}_+$, $D(\mu) = \{\theta \in \mathbb{R} \mid \text{L}_\mu(\theta) < +\infty\}$ (ch.1, §1.1)
$\text{D}(\hat{m}, x)$	deviance associated with a GLM and an observation x from it (ch.1, §3.3)
$\Delta(\cdot)$	polynomial function of degree ≤ 2
δ_a	Dirac measure at a (ch.1, §1.2)
$\mathbb{E}[\cdot], \mathbb{E}_\pi[\cdot]$	expected value, expected value under probability measure π
$\mathcal{E}a$	nearest even integer not exceeding a
$\text{ED}^*_\mu(\theta,\lambda), \text{ED}^*(\theta,\lambda)$	convolution family density (ch.1, §2.3)
$\text{ED}_\mu(\theta,\lambda), \text{ED}(\theta,\lambda)$	exponential dispersion model density (ch.1, §2.3)
$\text{EDM}(F)$	exponential dispersion model generated by F (ch.1, §2.3)
$\exp(\mu)$	exponentiation of a measure μ (ch.1, §1.1)
$F, F(\mu)$	NEF generated by the measure $\mu \in \mathcal{M}$ (ch.1, §2.1)
$G^{m,n}_{p,q}\left(z\middle\vert\begin{matrix}a_1,\dots,a_p\\ b_1,\dots,b_q\end{matrix}\right)$	Meijer's G-function (see appendix, §B.1
$\text{G}-\text{G}(\cdot,\cdot,\cdot,\cdot)$	Gamma Difference distribution (ch.2, §3.16)
$\text{G}+\text{G}(\cdot,\cdot,\cdot,\cdot)$	Gamma Sum distribution (ch.2, §3.15)
$\text{GHS}(\cdot,\cdot)$	Generalized Hyperbolic Secant distribution (ch.2, §1.6)
GLM	Generalized Linear Model (ch.1, §3.3)
$\text{GNB}(\cdot,\cdot,\cdot)$	Takács (Generalized Negative Binomial) distribution (ch.2, §2.2)
$\text{GP}(\cdot,\cdot)$	Abel (Generalized Poisson) distribution (ch.2, §2.1)
$\Gamma(x)$	gamma function $\int_0^\infty t^{x-1}e^{-t}dt$
$\Gamma'(\cdot,\cdot,\cdot)$	Non-central Chi-squared distribution (ch.2, §3.3)
$\Gamma(\cdot,\cdot)$	Gamma distribution (ch.2, §1.5)
$\text{Hermite}(\cdot,\cdot)$	Hermite distribution (ch.2, §3.1)
$\text{H}+\text{H}(\cdot,\cdot,\cdot,\cdot)$	Hyperbolic Secant Sum distribution (ch.2, §3.18)
$\text{KR}(\cdot,\cdot)$	Kendall-Ressel distribution (ch.2, §2.4)
$\text{IG}(\cdot,\cdot)$	Inverse Gaussian distribution (ch.2, §2.5)
$k_\mu(\theta)$	cumulant transform of μ: $\log \text{L}_\mu(\theta)$ (ch.1, §2.1)
$\text{k}^*_\mu(x)$	Legendre transform of μ with argument $x \in M_F$, $F = F(\mu)$ (ch.1, §2.1)
$\mathcal{L}(X)$	probability distribution of the r.v. X
$\text{L}_\mu(\theta)$	Laplace transform of μ with argument θ (ch.1, §1.1)

Notation used in the compendium (continued)

$L_j^{\alpha}(\cdot)$	Generalized Laguerre polynomial
Laguerre$(\cdot,\cdot,\cdot)$	Laguerre distribution (ch.2, §3.2)
$\Lambda(\mu), \Lambda(F)$	the Jørgensen set associated with measure μ or family F (ch.1, §2.3)
$\mathcal{M}_+$	set of positive measures on $\mathbb{R}$ (ch.1, §1.1)
$\mathcal{M}$	subset of $\mathcal{M}_+$ not concentrated on a single point (ch.1, §1.1)
M_F	mean domain of the NEF F (ch.1, §2.1)
ME$(\cdot,\cdot,\cdot)$	Mixed Exponential distribution (ch.2, §3.5)
MG$(\cdot,\cdot,\cdot)$	Mixed Geometric distribution (ch.2, §3.4)
$\mathbb{N}, \mathbb{N}^*$	$\{0,1,2,\ldots\}$, $\{1,2,3,\ldots\}$
N$(\cdot,\cdot)$	Normal distribution (ch.2, §1.1)
N + G$(\cdot,\cdot,\cdot)$	Normal-Gamma Sum distribution (ch.2, §3.17)
NB$(\cdot,\cdot)$	Pascal (Negative Binomial) distribution (ch.2, §1.4)
NB − NB$(\cdot,\cdot,\cdot,\cdot)$	Pascal Difference distribution (ch.2, §3.8)
NB + NB$(\cdot,\cdot,\cdot,\cdot)$	Pascal Sum distribution (ch.2, §3.7)
NEF	Natural Exponential Family (ch.1, §2)
$\mathcal{O}a$	nearest odd integer not exceeding a
$P(m,F)$	reparametrization of $P(\theta,\mu)$ in terms of its mean m and its NEF $F = F(\mu)$ (ch.1, §2.1)
$P(\cdot), Q(\cdot)$	polynomials of degrees ≤ 2 and ≤ 1 respectively
$P(\theta,\mu)$	NEF probability measure generated by μ (ch.1, §2.1)
P − NB$(\cdot,\cdot,\cdot)$	Poisson-Pascal Difference distribution (ch.2, §3.11)
P + NB$(\cdot,\cdot,\cdot)$	Poisson-Binomial Sum distribution (ch.2, §3.13)
P + NB$(\cdot,\cdot,\cdot)$	Poisson-Pascal Sum distribution (ch.2, §3.10)
$P_j^{(\alpha,\beta)}(\cdot)$	Jacobi polynomial
P − P$(\cdot,\cdot)$	P-P distribution (ch.2, §3.14)
$\phi_\mu(\cdot)$	mean domain mapping $M_F \to \Theta(\mu)$, $m \to \theta$, m the mean of $P(\theta,\mu)$ (ch.1, §2.1)
Po$(\cdot)$	Poisson distribution (ch.2, §1.2)
$\mathbb{R}^+, \mathbb{R}^-$	$(0,+\infty)$, $(-\infty,0)$
RIG$(\cdot,\cdot,\cdot)$	Reciprocal Inverse Gaussian distribution (ch.2, §4.5)
r.v.	random variable
SAb$(\cdot,\cdot,\cdot)$	S-Abel distribution (ch.2, §4.1)
SArc$(\cdot,\cdot,\cdot,\cdot)$	S-Arcsine distribution (ch.2, §4.3)
sgn$a$	signum function: $\text{sgn}a = \begin{cases} -1 & \text{if } a < 0 \\ 0 & \text{if } a = 0 \\ 1 & \text{if } a > 0 \end{cases}$

Notation used in the compendium (continued)

$\mathrm{STa}(\cdot,\cdot,\cdot,\cdot)$	S-Takács distribution (ch.2, §4.2)
$\mathrm{SKR}(\cdot,\cdot,\cdot)$	S-Kendall-Ressel distribution (ch.2, §4.4)
$\mathrm{Trin}(\cdot,\cdot,\cdot)$	Trinomial distribution (ch.2, §3.6)
$\tau_\mu(\cdot)$	the mapping $\Theta(\mu)\to M_F$, $\theta \mapsto m$, m the mean of $P(\theta,\mu)$ (ch.1, §2.1)
$\Theta(\mu)$	canonical parameter space of μ for $\mu \in \mathcal{M}_+$, $\Theta(\mu) = \mathrm{int}D(\mu)$ (ch.1, §1.1)
$\mathbf{V}_F(m)$	variance function $M_F \to \mathbb{R}^+$, $\mathbb{E}[X] \mapsto \mathrm{Var}[X]$ if $X \sim P(m,F)$ (ch.1, §2.1)
Var, Var_π	variance, variance under probability measure π
$W_{\alpha,\beta}$	Whittaker function (ch.2, §B.1)

Bibliography

Abramowitz, M. and Stegun, I.A. (1964). *Handbook of Mathematical Functions*, National Bureau of Standards Applied Mathematics Series. US Government Printing Office, Washington D.C.

Ahmed, M.S. (1961). On a locally most powerful boundary randomized similar test for the independence of two Poisson random variables. *Ann. Math. Statist.*, 32, 809-827.

Bar-Lev, S.K. and Enis, P. (1986). Reproducibility and natural exponential families with power variance functions. *Ann. Statist.*, 14, 1507-1522.

Bar-Lev, S.K. and Stramer, O. (1987). Characterizations of natural exponential families with power variance functions by zero regression properties. *Prob. Th. Rel. Fields.*, 76, 509-522.

Berg, S. and Nowicki, K. (1991). Statistical inference for a class of modified power series distributions with applications to random mappings theory. *J. Statist. Plann. Inference.*, 28, 247-261.

Billingsley, P. (1979). *Probability and Measure.* John Wiley and Sons, Inc., New York.

Borel, E. (1942). Sur l'emploi du théorème de Bernoulli pour faciliter le calcul d'une infinité de coefficients. Application au problème de l'attente à un guichet. *C. R. Acad. Sc. Paris série I*, 214, 452-456.

Casalis, M. (1992). Les familles exponentielles sur $\mathbb{R}^2$ à fonction variance $\mathbf{V}(m) = am \otimes m + B(m) + C$. *C. R. Acad. Sc. Paris série I*, 314, 635-638.

Casalis, M. (1994). Les $2n + 4$ types de familles exponentielles naturelles sur $\mathbb{R}^n$ à fonction variance quadratique simple. *C. R. Acad. Sc. Paris série I*, 318, 261-264.

Chhikara, R.S. and Folks, J.L. (1989). *The Inverse Gaussian Distribution: Theory, Methodology and Applications.* Marcel Dekker Inc., New York.

Consul, P.C. (1989). *Generalized Poisson Distributions: Properties and Applications*. Marcel Drekker Inc., New York.

Consul, P.C. and Jain, G.C. (1973). A generalization of the Poisson distribution. *Technometrics*, 15, 791-799.

Consul, P.C. and Shenton, L.R. (1972). Use of Lagrange expansion for generating discrete generalized probability distributions. *SIAM J. Appl. Math.*, 23, 239-249.

Daniels, H.E. (1954). Saddlepoint approximations in statistics. *Ann. Math. Stat.*, 25, 631-650.

Daniels, H.E. (1980). Exact saddlepoint approximations. *Biometrika*, 67, 59-63.

Erdélyi, A. (1953). *Higher Transcendental Functions*, vol. 1 & 2. Robert E. Krieger Publishing, Malabar.

Feinsilver, P. (1986). Some classes of orthogonal polynomials associated with martingales. *Proceedings of the AMS*, 98, 298-302.

Feller, W. (1971). *An Introduction to Probability Theory and its Applications*. John Wiley and Sons, New York.

Fisher, R.A. (1928). The general sampling distribution of the multiple correlation coefficient. *Proceedings of the Royal Society of London*, 121A, 654-673.

Fisher, R.A. (1951). Introduction. *Mathematical Tables*. British Mathematical Association.

Gould, H.W. (1972). *Combinatorial Identities*. Published by the author, Morgantown.

Gurland, J., Chen, E.E., and Hernandez, F.M. (1983). A new discrete distribution involving Laguerre polynomials. *Comm. Statist. – Theory and Methods*, 12, 1987-2004.

Haight, F.A. and Breuer, M.A. (1960). The Borel-Tanner distribution. *Biometrika*, 40, 143-150.

Harkness, W.L. and Harkness, M.L. (1968). Generalized hyperbolic secant distributions. *J. Amer. Statist. Assoc.*, 63, 329-337.

Harris, T. (1960). *Branching Processes*. Springer, Berlin.

Hassairi, A. (1993). Les $d+3$ G-orbites de la classe de Morris-Mora des familles exponentielles sur $\mathbb{R}^d$. *C. R. Acad. Sc. Paris série I*, 317, 887-890.

Helstrom, C.W. (1976). *Quantum Detection and Estimation Theory.* Academic Press, New York.

Jain, G.C. (1983). Hermite distributions. *Encyclopaedia of Statistical Sciences* (S. Kotz ed.), John Wiley and Sons, New York.

Jain, G.C. and Consul, P.C. (1971). A generalized negative binomial distribution. *SIAM J. Appl. Math.*, 21, 501-513.

Jain, G.C. and Khan, M.S.H. (1979). On an exponential family. *Math. Operationsforsch. Statist. Ser. Statistics*, 10(1), 153-168.

Janardan, K.G. and Rao, B.R. (1983). Lagrange distributions of the second kind and weighted distributions. *SIAM J. Appl. Math.*, 43, 302-313.

Johnson, N.L. and Kotz, S. (1969). *Distributions in Statistics: Discrete Distributions.* Houghton-Mifflin, Boston.

Johnson, N.L. and Kotz, S. (1970). *Distributions in Statistics: Continuous Univariate Distributions*, vol. 1 & 2. John Wiley and Sons, New York.

Jorgensen, B. (1982). *Statistical Properties of the Generalized Inverse Gaussian Distribution*, Lecture Notes in Statistics. Springer-Verlag, New York.

Jorgensen, B. (1986). Some properties of exponential dispersion models. *Scan. J. Statist.*, 13, 187-198.

Jorgensen, B. (1987). Exponential dispersion models. *J. Roy. Statis. Soc. B.*, 49, 127-162.

Jorgensen, B., Letac, G., and Seshadri, V. (1989). Sur une propriété des familles exponentielles naturelles de variance quadratique. *Can. J. Statist.*, 17(1), 1-8.

Jorgensen, B. and Martinez, J.R. (1991). On the asymptotic behaviour of the variance function. *Informes de Matemática.* IMPA, Rio de Janeiro.

Kemp, C.D. and Kemp, A.W. (1965). Some properties of the Hermite distribution. *Biometrika*, 52, 187-198.

Kendall, D.G. (1957). Some problems in the theory of dams. *J. Roy. Statist. Soc. B.*, 19, 207-212.

Kendall, M.G. and Stuart, A. (1963). *The Advanced Theory of Statistics*, vol. 1 & 2. Charles Griffin and Co., London.

Kokonendji, C.C. (1993). *Familles exponentielles naturelles de fonction variance en* $R\Delta + Q\sqrt{\Delta}$. PhD thesis, Université Paul Sabatier, Toulouse.

Kokonendji, C.C. and Seshadri, V. (1992). La méthode de Lindsay appliquée à la construction de familles exponentielles de fonction variance de degré 4 en $\sqrt{m}$. *C. R. Acad. Sc. Paris série I*, 314, 305-308.

Kokonendji, C.C. and Seshadri, V. (1994). The Lindsay transform of natural exponential families. *Can. J. Statist.*, 22(2), 259-272.

Laha, R.G. and Lukács, E. (1960). On a problem connected with quadratic regression. *Biometrika*, 47, 335-343.

Letac, G. (1988). Le problème de la classification des familles exponentielles naturelles sur $\mathbb{R}^d$ ayant une fonction variance quadratique. *Probability Measures on Groups*, Lecture Notes in Mathematical Statistics, vol. IX, Springer, New York.

Letac, G. (1992). *Lectures on Natural Exponential Families and their Variance Functions*, RIMPA Monograph no. 50. IMPA, Rio de Janeiro.

Letac, G. and Mora, M. (1990). Natural real exponential families with cubic variance functions. *Ann. Statis.*, 18, 1-37.

Malouche, D. (1994). Les fonctions variances dans $\mathbb{R}$ en $Q\sqrt{\Delta}$ où Q et Δ sont de degré inférieur ou égal à deux. Master's thesis, Université Paul Sabatier, Toulouse.

Mathai, A.M. (1993). *A Handbook of Generalized Special Functions for Statistical and Physical Sciences*. Oxford University Press, Oxford.

McKendrick, A.G. (1926). The application of mathematics to medical problems. *Proc. Edin. Math. Soc.*, 44, 98-130.

Meixner, J. (1934). Orthogonale Polynomsysteme mit einer besonderen Gestalt der erzeugenden Funktion. *J. Lond. M. Soc.*, 9, 6-13.

Meixner, J. (1941). Umformung gewisser Reihen, deren Glieder Produkte hypergeometrischer Funktionen sind. *Deutsche Math.*, 6, 341-349.

Mohanty, S.G. (1966). On a generalized two-coin tossing problem. *Biometrische Z.*, 8, 266-272.

Mora, M. (1990). La convergence des fonctions variance des familles exponentielles naturelles. *Ann. Fac. Sci. Toulouse Math.*, 5.XI(2), 105-120.

Morris, C.N. (1982). Natural exponential families with quadratic variance functions. *Ann. Statis.*, 10, 65-80.

Nelder, J.A. and Wedderburn, R.W.M. (1972). Generalized Linear Models. *J. Roy. Statis. Soc. A.*, 135(3), 370-384.

Seshadri, V. (1991). Finite mixtures of natural exponential families. *Can. J. Statist.*, 19(4), 437-445.

Seshadri, V. (1993). *The Inverse Gaussian Distribution: A Case Study in Exponential Families*, Oxford Science Publications, Oxford.

Shanbhag, D.N. (1972). Some characterizations based on the Bhattacharya matrix. *J. Appl. Prob.*, 9, 580-587.

Shanbhag, D.N. (1979). Diagonality of the Bhattacharya matrix as a characterization. *Teor. Veroyanost. i Primenen.*, 24, 424-427.

Stigler, S.M. (1982). Poisson on the Poisson distribution. *Stat. Prob. Lett.*, 1, 33-35.

Stigler, S.M. (1986). *The History of Statistics: the Measurement of Uncertainty before 1900.* Belknap Press, Cambridge.

Szegö, G. (1939). *Orthogonal Polynomials*, AMS Colloquium Publications. American Mathematical Society, Providence.

Takács, L. (1962). A generalization of the ballot problem and its application in the theory of queues. *J. Amer. Statist. Assoc.*, 57, 327-337.

Talacko, J. (1957). Perk's distributions and their role in the theory of Wiener's stochastic variables. *Trabajos Estadistica*, 7, 159-174.

Tanner, J.C. (1953). A problem of interference between two queues. *Biometrika*, 40, 58-69.

Tweedie, M.C.K. (1946). The regression of the sample variance on the sample mean. *J. London Math. Soc.*, 21, 21-28.

Tweedie, M.C.K. (1957). Statistical properties of Inverse Gaussian distributions – I. *Ann. Math. Statist.*, 28, 362-377.

Tweedie, M.C.K. (1984). An index which distinguishes between some important exponential families. *Statistics: Applications and New Directions. Proceedings of the Indian Statistical Institute Golden Jubilee International Conference* (J.K. Ghosh and J. Roy, eds.) Indian Statistical Institute, Calcutta.

Wedderburn, R.W.M. (1974). Quasi-likelihood functions, generalized linear models and the Gauss-Newton method. *Biometrika*, 61, 439-447.