On a class of distributions on the simplex

(Article begins on next page)

On a class of distributions on the simplex

Stefano Favaro[a], Georgia Hadjicharalambous[a], Igor Prünster[a,1,*]

[a]*Department of Statistics and Applied Mathematics & Collegio Carlo Alberto, University of Torino, Corso Unione Sovietica 218/bis, 10134 Torino, Italy*

Abstract

In the present paper we define and investigate a novel class of distributions on the simplex, termed normalized infinitely divisible distributions, which includes the Dirichlet distribution. Distributional properties and general moment formulae are derived. Particular attention is devoted to special cases of normalized infinitely divisible distributions which lead to explicit expressions. As a by–product also new distributions over the unit interval and a generalization of the Bessel function distribution are obtained.

Keywords: Bayes estimator, Dirichlet distribution, distribution on the simplex, generalized Bessel function distribution, infinitely divisible distribution, inverse Gaussian distribution, Lévy measure, positive stable distribution

1. Introduction

The Dirichlet distribution on the simplex is motivated by the well–known Dirichlet integral and arises in a variety of probabilistic and statistical contexts. These include Bayesian analysis, order statistics, limit laws, statistical genetics, Pearson systems of curves, nonparametric inference, distribution–free tolerance intervals, reliability theory, probability inequalities, multivariate analysis, characterization problems, stochastic processes and other areas. In order to define the Dirichlet distribution, let us consider a collection of random variables (r.v.) $X_1, \ldots, X_n$ which are assumed to be independent and distributed according to a gamma distribution with parameters

---

*Corresponding author

Email addresses: `stefano.favaro@unito.it` (Stefano Favaro),
`georgia.hadjicharalambous@unito.it` (Georgia Hadjicharalambous),
`igor@econ.unito.it` (Igor Prünster)

[1]Also affiliated to ICER

$(\alpha_1, 1), \ldots, (\alpha_n, 1)$, respectively, being $\alpha_i > 0$ the shape parameters for $i = 1, \ldots, n$. Then, by defining the r.v. $W := X_1 + \cdots + X_n$, the Dirichlet distribution with parameter $(\alpha_1, \ldots, \alpha_n)$ is defined as the distribution of the random vector $(P_1, \ldots, P_n) := (X_1/W, \ldots, X_n/W)$ on the $(n-1)$–dimensional simplex $\Delta^{(n-1)} := \{(p_1, \ldots, p_{n-1}) : p_i > 0, i = 1, \ldots, n-1, \sum_{i=1}^{n-1} p_i \leq 1\}$. In other terms, the Dirichlet distribution is definable via a "normalization" operation as the joint distribution of a set of independent r.v. distributed according to a gamma distribution divided by their sum.

The Dirichlet distribution represents one of the most widely studied distributions over the last century. The reason for its popularity can be mainly traced back to its mathematical tractability: however, this also implies some constraints in terms of flexibility, which constitute a limitation in some of its numerous applications. Therefore, various approaches for constructing distributions on the simplex have been undertaken with the aim of overcoming some of the drawbacks of the Dirichlet distribution. For instance, in the context of the analysis of compositional data, which are subject to nonnegativity and constant-sum constraints, and in Bayesian analysis we mention the scaled Dirichlet distribution [2], the generalized logistic–normal distribution [3], the Liouville distribution and its various generalizations (see [30] and references therein), the generalized Dirichlet distribution [8, 17, 18], the grouped Dirichlet distribution [26], the nested Dirichlet distribution [27, 33] and the Dirichlet tree distribution usually also known as the hyper–Dirichlet distribution [9, 10]. In the present paper we propose an alternative approach to the problem of defining distributions on the simplex by resorting to the above mentioned "normalization" approach. Such an approach has been fruitfully introduced in [31] for the definition of random probability measures for Bayesian nonparametric inference (see, e.g., [28, 19, 22, 14, 20, 15] for following developments).

Let us consider a r.v. X distributed according to a gamma distribution with parameter $(\alpha, 1)$; it is well–known that the gamma distribution is infinitely divisible (ID), i.e for any $n \in \mathbb{N}$ there exists a collection of independent and identically distributed r.v. $Y_1, \ldots, Y_n$ such that $X \stackrel{\mathrm{d}}{=} Y_1 + \cdots + Y_n$ or, alternatively, the r.v. X can be "separated" into the sum of an arbitrary number of independent and identically distributed r.v. In particular, according to the Lévy–Khintchine representation theorem for ID distributions, the Laplace transform of the gamma distribution with parameter $(\alpha, 1)$ coincides with

$$\mathbb{E}\left[\mathrm{e}^{-u\,X}\right] = \exp\left\{-\alpha \int_0^\infty (1 - \mathrm{e}^{-ux}) \frac{\mathrm{e}^{-x}}{x} dx\right\} = (1 + u)^{-\alpha} \qquad u \geq 0,$$

2

where $\alpha x^{-1}\mathrm{e}^{-x}dx$ is the so–called Lévy measure and it completely characterizes the gamma distribution. See the comprehensive and stimulating monograph [32] and references therein. At this stage we are in a position to describe the main purpose of the present paper, that is, the generalization of the "normalization" approach applied for defining the Dirichlet distribution to cases in which the gamma r.v. are replaced by arbitrary positive ID r.v. In other terms, we consider a collection of positive r.v. $X_1, \ldots, X_n$ which are assumed to be independent and distributed according to, not necessarily coinciding, ID distributions. The Laplace transform of each X_i can be represented canonically via the Lévy–Khintchine representation as

$$\mathbb{E}[\mathrm{e}^{-uX_i}] = \exp\left\{-\Psi_i(u)\right\} = \exp\left\{-\int_0^\infty (1 - \mathrm{e}^{-ux})\nu_i(dx)\right\} \qquad u \geq 0,$$

where Ψ_i is typically referred to as Laplace exponent and the Lévy measure ν_i is any nonnegative Borel measure satisfying the condition

$$\int_0^\infty \min(1, x)\nu_i(dx) < \infty$$

which completely characterizes the distribution of the r.v. X_i, for each $i = 1, \ldots, n$. Thus, by defining the r.v. $W := X_1 + \cdots + X_n$, the "normalization" approach yields a wide class of distributions over $\Delta^{(n-1)}$ for the random vector $(P_1, \ldots, P_n) := (X_1/W, \ldots, X_n/W)$. In particular, each of these distributions is completely characterized by the corresponding collection of Lévy measures $\{\nu_1, \ldots, \nu_n\}$. We term this class of distributions the normalized ID (NID) distributions.

The class of NID distributions represents a natural extension of the Dirichlet distribution, which can be recovered as special case of NID distributions by assuming the collection of Lévy measures to be $\nu_i(dx) = \alpha_i x^{-1}\mathrm{e}^{-x}dx$ for $i = 1, \ldots, n$. Further important examples of NID distributions, which have found important applications in Bayesian theory, are the normalized inverse–Gaussian distribution [22], the normalized $1/2$–stable distributions [5] and the normalized tempered stable distribution [21, 24]. Our main aim is to investigate the class of NID distributions in terms of their distributional properties and to provide general moment formulae. Specifically, the key tools exploited for deriving the results are given by Gurland's inversion formula for characteristic functions [16] and the Faà di Bruno formula which allows to obtain expressions in terms of partial Bell polynomials [7]. As an application of the general results, some interesting examples of NID distributions are considered by combining three well–known ID

3

distributions: the gamma distribution, the positive $1/2$–stable distribution and the inverse–Gaussian distribution. These lead to relatively simple and tractable expressions, which are readily exploitable in applications. Following these guidelines, in Section 2 we state the definition of NID distribution, investigate properties of the corresponding distribution and provide general moment formulae. In Section 3 interesting special cases of NID distributions are analyzed in detail. As a by–product also new distributions over the unit interval and a generalization of the Bessel function distribution are obtained.

2. Normalized infinitely divisible distributions

Relying on the "normalization" concept briefly described in Section 1, we start by stating the formal definition of NID distribution on the simplex.

Definition 1. Let $X_1, \ldots, X_n$ be a collection of independent strictly positive ID r.v. and $W := X_1 + \cdots + X_n$. A normalized infinitely divisible (NID) distribution is defined as the distribution of the random vector $(P_1, \ldots, P_n) := (X_1/W, \ldots, X_n/W)$ on the simplex $\Delta^{(n-1)} := \{(p_1, \ldots, p_{n-1}) : p_i > 0, i = 1, \ldots, n-1, \sum_{i=1}^{n-1} p_i \leq 1\}$.

It is important to note that since the r.v. $X_1, \ldots, X_n$ are uniquely identified by their Lévy measures $\nu_1, \ldots, \nu_n$, also a NID distribution is characterized by the collection of Lévy measures $\nu_1, \ldots, \nu_n$, although it is clearly not ID anymore. Nonetheless, the relation to the underlying ID r.v. allows to derive properties of NID distributions directly in terms of the underlying Lévy measures $\nu_1, \ldots, \nu_n$. The assumption of strict positivity of the r.v. $X_1, \ldots, X_n$ restated in terms of the corresponding Lévy measures boils down to requiring $\int_{(0,\infty)} \nu_i(dx) = \infty$ (see, e.g., [31]). In fact, for the normalization operation to be well defined, at least one of the X_i's needs to be strictly positive. For simplicity, we assume all X_i's to be strictly positive.

2.1. Distributions

In the sequel we restrict our attention to NID distributions based on r.v., which are absolutely continuous w.r.t. the Lebesgue measure on $\mathbb{R}$. In this way, any NID distribution admits a probability density function (p.d.f.) on the $(n-1)$–dimensional simplex $\Delta^{(n-1)}$ which is absolutely continuous w.r.t. the Lebesgue measure on $\mathbb{R}^{n-1}$. In order to investigate the distribution of a NID r.v. $(P_1, \ldots, P_n)$ we can distinguish two different scenarios with respect to the existence of a closed–form expression for the p.d.f. of the

4

corresponding ID r.v. $X_1, \ldots, X_n$. If the p.d.f. of the r.v. X_i, denoted by f_i, admits a closed–form expression for every $i = 1, \ldots, n$ then, the p.d.f. of the corresponding NID random vector $(P_1, \ldots, P_n)$ can be obtained by a simple transformation procedure. In particular, if h denotes the p.d.f. of $(P_1, \ldots, P_{n-1}, W)$, one has

$$g_{(P_1,\ldots,P_n)}(p_1, \ldots, p_{n-1}) = \int_0^\infty h(p_1, \ldots, p_{n-1}, w) dw$$

$$= \int_0^\infty \prod_{i=1}^{n-1} f_i(p_i w) \, f_n \left(w - \sum_{i=1}^{n-1} p_i w \right) w^{n-1} \, dw. \quad (1)$$

Example. As an illustrative example let us consider the case of $X_1, \ldots, X_n$ having positive stable distribution. By $X \sim \mathrm{St}(\gamma, c)$ we denote a r.v. distributed according to a γ–stable distribution with parameter (γ, c), where $\gamma \in (0, 1)$ and $c > 0$ is the shape parameter. In general, the p.d.f. of a γ–stable r.v. does not admit a closed–form expression for its p.d.f. A notable exception is represented by the case $\gamma = 1/2$ for which the p.d.f. is given by

$$f_X(x) = \frac{c}{(2\pi)^{1/2}} \mathrm{e}^{-\frac{c^2}{2x}} x^{-\frac{3}{2}} \mathbb{1}_{\mathbb{R}^+}(x).$$

For later use, observe that the $1/2$–stable distribution coincides with the inverted gamma distribution with parameter $(1/2, c)$. Now, consider the NID distribution, termed normalized $1/2$-stable distribution, arising by considering independent r.v. $X_1, \ldots, X_n$ such that $X_i \sim \mathrm{St}(\gamma, c_i)$ for $i = 1, .., n$. A straightforward application of (1) yields

$$g_{(P_1,\ldots,P_n)}(p_1, \ldots, p_{n-1}) = \frac{\Gamma(\frac{n}{2}) \prod_{i=1}^n c_i}{\pi^{\frac{n}{2}}} \frac{\prod_{i=1}^{n-1} p_i^{-\frac{3}{2}} \left(1 - \sum_{i=1}^{n-1} p_i\right)^{-\frac{3}{2}}}{\left(\sum_{i=1}^{n-1} \frac{c_i^2}{p_i} + \frac{c_n^2}{1-\sum_{i=1}^{n-1} p_i}\right)^{\frac{n}{2}}}, \quad (2)$$

a distribution first obtained in [5], which represents an interesting special case of NID. □

The second more complicated scenario is represented by the case of X_i not admitting a closed–form p.d.f. for some (or every) $i = 1, \ldots, n$. In such a case, based on the only knowledge of the Lévy measures ν_i for $i = 1, \ldots, n$, one can still derive an expression for the marginal density of P_i or of a linear functional of P_i, i.e. $\sum_{i=1}^n f(y_i) P_i$, where $y_1, \ldots, y_n$ can be interpreted as the "locations" on which the P_i's sit with f being a measurable real–valued function. In particular, by adapting arguments of [31] we have the following proposition which yields the cumulative distribution function (c.d.f.) of P_j expressed in terms of the underlying Lévy measures.

Proposition 1 *If $(P_1, \ldots, P_n)$ has NID distribution characterized by the Lévy measures $\nu_1, \ldots, \nu_n$, then*

$$\mathbb{P}(P_j \leq y) = \frac{1}{2} - \frac{1}{\pi} \lim_{T \uparrow \infty} \int_0^T \frac{1}{t} \exp\left\{ \sum_{1 \leq i \neq j \leq n} \int_0^\infty \cos(-tvy - 1)\nu_i(dv) \right\}$$

$$\times \exp\left\{ \int_0^\infty \cos(tv(1-y) - 1)\nu_j(dv) \right\} \tag{3}$$

$$\times \sin\left(\sum_{1 \leq i \neq j \leq n} \int_0^\infty \sin(-tvy)\nu_i(dv) + \int_0^\infty \sin(tv(1-y))\nu_j(dv) \right) dt.$$

Example (continued). Consider again the case of the NID distribution based on stable r.v. but, instead of fixing $\gamma = 1/2$, we face the general case of $\gamma \in (0,1)$. In such a case the p.d.f. of the X_i is not available in closed form. However, the Laplace transform still has a particularly simple expression. In fact, if $X \sim \mathrm{St}(\gamma, c)$, then, using a slightly different parametrization than the usual for coherence with following developments, the corresponding Lévy measure is given by $\nu(dv) = \frac{c}{\sqrt{2\pi} v^{1+\gamma}} \, dv$ and the Laplace transform has exponent given by

$$\Psi_{\mathrm{St}(\gamma,c)}(u) = \frac{c\,\Gamma(1-\gamma)}{\sqrt{2\pi}\,\gamma} \, u^\gamma. \tag{4}$$

We can now derive the c.d.f. of P_j for a NID distribution based on n independent r.v. $X_i \sim \mathrm{St}(\gamma, c_i)$. By applying Proposition 1, noting the absolute integrability in $[M, \infty]$ of the integrand in (3) and some tedious algebra one obtains that

$$\mathbb{P}(P_j \leq y) = \frac{1}{2} - \frac{1}{\pi\gamma} \arctan\left(\frac{(1-y)^\gamma c_j - y^\gamma \sum_{i \neq j} c_i}{(1-y)^\gamma c_j + y^\gamma \sum_{i \neq j} c_i} \tan\frac{\pi\gamma}{2} \right), \tag{5}$$

which represents a surprisingly simple expression. Expressions for other examples of NID distributions can be obtained by simple plug–in of a set a Lévy measures into (3). □

2.2. Moments

We now turn our attention to the derivation of moment formulae for general NID distributions: in deriving such results we rely only on the knowledge of the Lévy measures characterizing a NID distribution and, therefore, they can be applied to both scenarios highlighted above. Let us start from the first moment.

Proposition 2 *If $(P_1, \ldots, P_n)$ has NID distribution characterized by the Lévy measures $\nu_1, \ldots, \nu_n$, then*

$$\mathbb{E}[P_j] = \int_0^\infty \left(\frac{d}{du} \Psi_j(u) \right) e^{-\sum_{i=1}^n \Psi_i(u)} du,$$

where $\Psi_i(u) := \int_0^\infty (1 - e^{-uv}) \nu_i(dv)$ is the Laplace exponent of the r.v. X_i.

In order to derive the r–th moment of the r.v. P_j, for any $r \in \mathbb{N}$, one can exploit arguments similar to those used in Proposition 2, which lead to

$$\mathbb{E}[P_j^r] = \frac{(-1)^r}{\Gamma(r)} \int_0^\infty u^{r-1} \left(\frac{d^r}{du^r} e^{-\Psi_j(u)} \right) e^{-\sum_{1 \leq i \neq j \leq n} \Psi_i(u)} du. \qquad (6)$$

Moreover, the r–th centered moment of P_j is

$$\mathbb{E}[P_j - \mathbb{E}[P_j]]^r \qquad (7)$$

$$= \sum_{l=0}^r \binom{r}{l} (-1)^l \left(\int_0^\infty \left(\frac{d}{du} \Psi_j(u) du \right) e^{-\sum_{i=1}^n \Psi_i(u)} \right)^l$$

$$\times \frac{(-1)^{r-l}}{\Gamma(r-l)} \int_0^\infty u^{r-l-1} \left(\frac{d^{r-l}}{du^{r-l}} e^{-\Psi_j(u)} \right) e^{-\sum_{1 \leq i \neq j \leq n} \Psi_i(u)} du.$$

Significant simplifications of the expressions for the moment formulae (6) and (7) can be obtained by using the Faá di Bruno formula (see [7, Section 3.4]). Indeed, by the Faá di Bruno formula, one has

$$\frac{d^r}{du^r} e^{-\Psi_j(u)} = e^{-\Psi_j(u)} \sum_{s=1}^r \frac{1}{s!} \sum_{(r_1, \ldots, r_s) \in \mathcal{D}_{r,s}} \frac{r!}{r_1! \cdots r_s!} \prod_{i=1}^s \frac{d^{r_i}}{du^{r_i}} (-\Psi_j(u)) \qquad (8)$$

with $\mathcal{D}_{r,s}$ being the set of indexes $(r_1, \ldots, r_s) \in \{1, \ldots, r\}^s$ such that $1 \leq r_i \leq s$ and $\sum_{i=1}^s r_i = r$. Moreover, note that

$$B_{r,s}(w_\bullet(u)) := \frac{1}{s!} \sum_{(r_1, \ldots, r_s) \in \mathcal{D}_{r,s}} \frac{r!}{r_1! \cdots r_s!} \prod_{i=1}^s \frac{d^{r_i}}{du^{r_i}} (-\Psi_j(u))$$

is the (r,s)–th partial Bell polynomial (see [7]) with weight sequence $w_\bullet(u) := \{w_i(u), i \geq 1\}$ such that $w_i(u) := d^i/du^i(-\Psi_j(u))$. Therefore, by applying (8) to the moment formulae (6) and (7) we obtain the following alternative result.

7

Proposition 3 *If $(P_1, \ldots, P_n)$ has NID distribution characterized by the Lévy measures $\nu_1, \ldots, \nu_n$, then*

$$\mathbb{E}[P_j^r] = \frac{(-1)^r}{\Gamma(r)} \sum_{s=1}^{r} \int_0^\infty u^{r-1} e^{-\sum_{i=1}^n \Psi_i(u)} B_{r,s}(w_\bullet(u)) du \qquad (9)$$

and

$$\mathbb{E}[P_j - \mathbb{E}[P_j]]^r \qquad (10)$$

$$= \sum_{l=0}^{r} \binom{r}{l} (-1)^l \left(\int_0^\infty \left(\frac{d}{du} \Psi_j(u) \right) e^{-\sum_{i=1}^n \Psi_i(u)} du \right)^l$$

$$\times \frac{(-1)^{r-l}}{\Gamma(r-l)} \sum_{s=1}^{r-l} \int_0^\infty u^{r-l-1} e^{-\sum_{i=1}^n \Psi_i(u)} B_{r-l,s}(w_\bullet(u)) du.$$

For ease of notation, in the following we indicate the r–th and r–th centered moments of the r.v. P_j by $\mu_{j,r}$ and $\bar{\mu}_{j,r}$, respectively. Using (9) and (10) one can easily obtain general formulae for the mean, the variance or any other quantity related to the moments of P_j. An interesting subclass of NID distributions corresponds to the case of independent r.v. $X_1, \ldots, X_n$ with the same ID distribution and, moreover, differing solely in their shape parameter. For instance, the γ–stable NID considered in the previous example belongs to this class. Such a subclass essentially corresponds to the framework of [19]: therefore, as a particular case of Proposition 3, we obtain, on the one hand, the formulae for the mean, variance and skewness of [19, Proposition 1] and, on the other hand, new formulae for all higher order moments. This special class of NID distributions is further discussed in Section 3.6 in relation to Bayesian nonparametric inference.

Example (continued). Considering again the case of the normalized γ–stable distribution, i.e. the NID distribution based on $X_i \sim \text{St}(\gamma, c_i)$ for $i = 1, \ldots, n$, one obtains the simple expressions

$$\mathbb{E}[P_j] = \frac{c_j}{c}$$

$$\text{Var}(P_j) = \frac{c_j(c - c_j)}{c^2} (1 - \gamma)$$

$$\gamma_1(P_j) = \frac{c - 2 c_j}{2\sqrt{c_j(c - c_j)}} \frac{2 - \gamma}{\sqrt{(1 - \gamma)}}$$

having set $c = \sum_{i=1}^n c_i$ and denoted by γ_1 the skewness. Moreover, one can determine also quantities related to higher order moments. For instance,

if one is interested in the kurtosis $\kappa(P_j) = \bar{\mu}_{j,4}/(\mathrm{Var}(P_j))^2$, some simple algebra leads to

$$\kappa(P_j) = \gamma + \frac{3}{\gamma - 1} + \frac{c^2(\gamma - 2)(\gamma - 3)}{6c_j(c - c_j)(1 - \gamma)}.$$

$\square$

3. Applications to special cases

In this section we consider some interesting special cases of NID distributions and show how the results of Section 2 can be applied in order to derive explicit distributions and moment formulae. We focus on three ID distributions: the gamma distribution, the 1/2–stable distribution and the inverse–Gaussian distribution.

Before proceeding we recall their main features also in order to fix the notation. We denote by $X \sim \mathrm{Ga}(\alpha, \gamma)$ a r.v. distributed according to a gamma distribution with parameter (α, γ), with α and γ denoting the shape and scale parameters, respectively. The corresponding p.d.f. is

$$f_X(x) = \frac{\gamma^\alpha}{\Gamma(\alpha)} \mathrm{e}^{-x\gamma} x^{\alpha-1} \mathbb{1}_{\mathbb{R}^+}(x).$$

Moreover, recall that its Lévy measure is $\nu(dv) = \alpha v^{-1} \mathrm{e}^{-v} dv$ and denote by $\Psi_{\mathrm{Ga}(\alpha,\gamma)}$ its Laplace exponent in (1) which is given by

$$\Psi_{\mathrm{Ga}(\alpha,\gamma)}(u) = \alpha \log\left(1 + \frac{u}{\gamma}\right). \tag{11}$$

Recall, from the example in Section 1, that $X \sim \mathrm{St}(\gamma, c)$ denotes a γ–stable r.v. with parameter (γ, c), which for $\gamma = 1/2$ admits a closed form p.d.f. given by (2.1). The corresponding Laplace exponent is obtained from (4) by setting $\gamma = 1/2$ yielding

$$\Psi_{\mathrm{St}(1/2,c)}(u) = c(2u)^{\frac{1}{2}}. \tag{12}$$

Finally, we denote by $X \sim \mathrm{IG}(\lambda, \theta)$ a r.v. distributed according to an inverse–Gaussian distribution with parameter (λ, θ), with λ and θ denoting the shape and scale parameters, respectively. The corresponding p.d.f. is given by

$$f_X(x) = \frac{\lambda}{(2\pi)^{1/2}} \mathrm{e}^{\theta\lambda - \frac{1}{2}\left(\frac{\lambda^2}{x} + \theta^2 x\right)} x^{-\frac{3}{2}} \mathbb{1}_{\mathbb{R}^+}(x).$$

The Lévy measure of $X \sim \text{IG}(\lambda, \theta)$ is $\nu(dv) = \lambda(2\pi)^{-1/2} v^{-1-\gamma} e^{-\frac{1}{2}\theta^2 v} dv$ and the corresponding Laplace exponent coincides with

$$\Psi_{\text{IG}(\lambda,\theta)}(u) = \lambda(2u + \theta^2)^{\frac{1}{2}} - \theta\lambda. \tag{13}$$

If $\theta = 0$, we recover the γ–stable distribution with parameter $(1/2, \lambda)$.

Remark 1. All three distributions above are special cases of the generalized inverse–Gaussian distribution with parameter $(\alpha, \lambda, \theta)$, which is an absolutely continuous r.v. with p.d.f. given by

$$f_X(x) = \left(\frac{\theta}{\lambda}\right)^\alpha \frac{1}{2K_\alpha(\lambda\theta)} x^{\alpha-1} e^{-\frac{1}{2}(\lambda^2 x^{-1} + \theta^2 x)} \mathbb{1}_{\mathbb{R}^+}(x)$$

where K_r is the modified Bessel function of the third kind with index $r \in \mathbb{R}$ and $\alpha \in \mathbb{R}$, $\lambda, \theta \in \mathbb{R}^-$ and such that $(\lambda, \theta) \neq (0, 0)$. The Laplace exponent of a r.v. $X \sim \text{GIG}(\alpha, \lambda, \theta)$ is given by

$$\Psi_{\text{GIG}_{\alpha,\lambda,\theta}}(u) = \log\left(\frac{K_\alpha(\lambda\theta)}{K_\alpha(\lambda(2u+\theta^2)^{1/2})}\left(\frac{\theta}{\lambda}\right)^{-\alpha}\left(\frac{2u+\theta^2}{\lambda^2}\right)^{\frac{\alpha}{2}}\right).$$

The gamma distribution with parameter $(\alpha, \theta^2/2)$ corresponds to a $\text{GIG}(\alpha, 0, \theta)$. Moreover, the inverse–Gaussian distribution with parameter (λ, θ) can be recovered as $X \sim \text{GIG}(\alpha, \lambda, \theta)$ by setting $\alpha = -1/2$ and the $1/2$–stable distribution by further setting $\theta = 0$. □

3.1. Classification of NID distributions

Definition 1 leaves quite some freedom in the construction of a NID distribution and, hence, it is useful to identify subclasses of NID distributions according to the structure of the r.v. $X_1, \ldots, X_n$ in terms of which a NID distribution is defined. For our purposes, it is useful to distinguish two main setups: (i) for some fixed $k \in \{1, \ldots, n\}$, the r.v. $X_1, \ldots, X_k$ follow the same positive ID distribution, while the r.v. $X_{k+1}, \ldots, X_n$ follow another positive ID distribution; if $k = n$, then the second set of r.v. vanishes by convention; (ii) the r.v. $X_1, \ldots, X_n$ follow the same positive ID distribution therefore representing the special case of (i) corresponding to $k = n$. In both cases all parameters (not only the shape parameters) of the involved classes of ID distributions are allowed to differ.

In particular, let us first focus on setup (i): let F_1 and F_2 be two different positive ID distributions with vectors of parameters c_1 and c_2, respectively, chosen among the three notable above mentioned distributions.

Hence, $X_1, \ldots, X_n$ is a collection of independent r.v. such that, for some fixed $k \in \{1, \ldots, n\}$, X_i has distribution F_1 with parameters $c_{1,i}$, for $i = 1, \ldots, k$ and X_i has distribution F_2 with parameters $c_{2,i}$, for $i = k+1, \ldots, n$. Exploiting collections of r.v. $X_1, \ldots, X_n$ structured in this way, from Definition 1 we obtain three new interesting NID, namely:

(A) NID distributions based on 1/2–stable and inverse–Gaussian r.v.;

(B) NID distributions based on gamma and inverse–Gaussian r.v.;

(C) NID distributions based on gamma and 1/2–stable r.v.

For each of the three NID distributions we derive the p.d.f. and some moment formulae. Moreover, as an illustration, we consider the problem of estimating the success probability in a binary experiment in Bayes framework and derive the corresponding explicit Bayes estimators. These represent alternatives to the usually exploited estimator corresponding to the conjugate beta prior choice. Specifically, let $\{Y_n, n \geq 1\}$ be an exchangeable sequence taking value in $\{0, 1\}^\infty$: by de Finetti's representation theorem, there exists a r.v. p taking values in $(0, 1)$ such that conditionally on p, the r.v. Y_i are independent and identically distributed according to a Bernoulli distribution with parameter p. In a Bayesian setup, the distribution of p is commonly referred to as a "prior" distribution and, given a sample $Y_1, \ldots, Y_n$, interest relies in determining a Bayes estimator for p, which, for a quadratic loss function, coincides with $\mathbb{E}[p|Y_1, \ldots, Y_n]$. We provide such an estimator assuming p is distributed according to one of the three NID distributions listed above.

We then focus on Case (ii) and, specifically, when the r.v. $X_1, \ldots, X_n$ are gamma distributed but with different both scale and shape parameters. Finally, some known NID distributions are recalled: these correspond to the very special case of $X_1, \ldots, X_n$ having the same ID distribution and, moreover, differing only in their shape parameters.

3.2. Normalized 1/2–stable and inverse–Gaussian distributions

For any $n \in \mathbb{N}$ and fixed $k \in \{1, \ldots, n\}$, let us consider n independent r.v. $X_1, \ldots, X_n$ such that $X_i \sim \mathrm{St}(1/2, c_i)$ for $i = 1, \ldots, k$ and $X_i \sim \mathrm{IG}(\lambda_i, \theta_i)$ for $i = k+1, \ldots, n$. Since the p.d.f. of the r.v. $X_1, \ldots, X_n$ are known, the p.d.f. of the NID random vector $(P_1, \ldots, P_n)$ is obtained by applying (1). In fact, some algebra and formula 3.471.9 in [13] lead to

$$g_{(P_1,\ldots,P_n)}(p_1, \ldots, p_{n-1}) = 2 \prod_{i=1}^{k} \frac{c_i}{(2\pi)^{1/2}} \prod_{i=k+1}^{n} \frac{\lambda_i}{(2\pi)^{1/2}} e^{\theta_i \lambda_i} \tag{14}$$

11

$$\times \prod_{i=1}^{n-1} p_i^{-\frac{3}{2}} \left(1 - |\mathbf{p}|\right)^{-\frac{3}{2}} \left(\frac{C_1(p_1, \ldots, p_{n-1})}{L_1(p_1, \ldots, p_{n-1})}\right)^{-\frac{n}{4}}$$

$$\times K_{-n/2} \left(C_1^{\frac{1}{2}}(p_1, \ldots, p_{n-1}) L_1^{\frac{1}{2}}(p_1, \ldots, p_{n-1})\right) \mathbb{1}_{\Delta^{(n-1)}}(p_1, \ldots, p_{n-1}),$$

where $|\mathbf{p}| := \sum_{i=1}^{n-1} p_i$, K_ν denotes the modified Bessel function of the third kind with index $\nu \in \mathbb{R}$ and, moreover, we set

$$C_1(p_1, \ldots, p_{n-1}) := \sum_{i=1}^{k} \frac{c_i^2}{p_i} + \sum_{i=k+1}^{n-1} \frac{\lambda_i^2}{p_i} + \frac{\lambda_n^2}{1 - |\mathbf{p}|}$$

$$L_1(p_1, \ldots, p_{n-1}) := \sum_{i=k+1}^{n-1} \theta_i^2 p_i + \theta_n^2 \left(1 - |\mathbf{p}|\right).$$

The same idea can be used for constructing new distributions over the unit interval, which obviously correspond to NID distributions with $n = 2$. Hence, let X_1 and X_2 be two independent r.v. such that $X_1 \sim \mathrm{St}(1/2, c)$ and $X_2 \sim \mathrm{IG}(\lambda, \theta)$. The distribution of the NID r.v. $P := X_1/(X_1 + X_2)$ is then given by

$$g_P(p) = \frac{c \lambda e^{\theta \lambda}}{\pi} p^{-\frac{3}{2}} (1-p)^{-\frac{3}{2}} \left(\frac{C_1(p)}{L_1(p)}\right)^{-\frac{1}{2}} K_{-1} \left(C_1^{\frac{1}{2}}(p) L_1^{\frac{1}{2}}(p)\right) \mathbb{1}_{(0,1)}(p), \quad (15)$$

with $C_1(p) := c^2 p^{-1} + \lambda^2 (1-p)^{-1}$ and $L_1(p) := \theta^2(1-p)$. Moment formulae for a NID random vector, whose distribution admits density (14), can be obtained by applying Proposition 3. In particular, define

$$S_1(u) := (2u)^{\frac{1}{2}} \sum_{i=1}^{k} c_i + \sum_{i=k+1}^{n} \lambda_i ((2u + \theta_i^2)^{\frac{1}{2}} - \theta_i).$$

The next proposition summarizes these moment formulae.

Proposition 4 *If $(P_1, \ldots, P_n)$ has NID distribution admitting density (14), then, for any $r \in \mathbb{N}$, for $1 \leq j \leq k$*

$$\mu_{j,r} = 2^{-2r} \sum_{s=1}^{r} \frac{2^{2s} c_j^s}{\Gamma(s)} \left(\frac{1}{2}\right)^{\frac{s}{2}} \binom{2r-s-1}{r-1} \int_0^\infty u^{\frac{s}{2}-1} e^{-S_1(u)} du \qquad (16)$$

and for $k + 1 \leq j \leq n$

$$\mu_{j,r} = 2^{-r} \sum_{s=1}^{r} \frac{(2\lambda_j)^s}{\Gamma(s)} \binom{2r-s-1}{r-1} \int_0^\infty u^{r-1}(2u + \theta_j^2)^{\frac{s}{2}-r} e^{-S_1(u)} du. \quad (17)$$

As special case of Proposition 4 we recover the r–th moment for a NID r.v. P over $(0,1)$ distributed according to a distribution characterized by (15), i.e.

$$\mu_r = \sum_{s=1}^{r} \frac{(2c)^s}{\Gamma(s)} \left(\frac{1}{2}\right)^{\frac{s}{2}} 2^{s-2r} \binom{2r-s-1}{r-1} \int_0^{+\infty} u^{\frac{s}{2}-1} \mathrm{e}^{-S_1(u)} du. \qquad (18)$$

Remark 2. Note that expressions (16) and (17) can be further simplified by solving the sum over the index s. In particular, for $1 \le j \le k$ we can write (16) as

$$\mu_{j,r} = 2^{-2r} \int_0^{\infty} \frac{\mathrm{e}^{-S_1(u)}}{u} \sum_{s=1}^{r} \frac{1}{\Gamma(s)} \binom{2r-s-1}{r-1} (2^{3/2} c_j u^{1/2})^s du$$

$$= \frac{(-1)^{r+1} \pi^{1/2}}{\Gamma(r)} 2^{-r/2-1/4} c_j^{1/2+r}$$

$$\times \int_0^{+\infty} \frac{\mathrm{e}^{-S_1(u)+2^{1/2} c_j u^{1/2}}}{u^{3/4-r/2}} I_{1/2-r}(2^{1/2} c_j u^{1/2}) du$$

and for $k+1 \le j \le n$ we can write (17) as

$$\mu_{j,r} = 2^{-r} \int_0^{+\infty} \frac{u^{r-1} \mathrm{e}^{-S_1(u)}}{(2u+\theta_j^2)^r} \sum_{s=1}^{r} \frac{1}{\Gamma(s)} \binom{2r-s-1}{r-1} (2\lambda_j (2u+\theta_j^2)^{1/2})^s du$$

$$= \frac{(-1/2)^{r+1} \pi^{1/2}}{\Gamma(r)} (2\lambda_j)^{1/2+r}$$

$$\times \int_0^{+\infty} \frac{u^{r-1} \mathrm{e}^{-S_1(u)+\lambda_j (\lambda_j (2u+\theta_j^2))^{1/2}}}{(2u+\theta_j^2)^{r/2-1/4}} I_{1/2-r}(\lambda_j (2u+\theta_j^2)^{1/2}) du,$$

where I_ν denotes the modified Bessel function of the first kind with index $\nu \in \mathbb{R}$. Accordingly, we can simplify expression (18). $\qquad \square$

Finally, as far as the central moments are concerned we have the following result.

Proposition 5 *If $(P_1,\ldots,P_n)$ has NID distribution admitting density (14), then, for any $r \in \mathbb{N}$, for $1 \le j \le k$*

$$\bar{\mu}_{j,r} = \sum_{l=0}^{r} \binom{r}{l} (-1)^l \left(\int_0^{+\infty} e^{-S_1(u)} c_j (2u)^{-\frac{1}{2}} du\right)^l$$

$$\times \sum_{s=1}^{r-l} \frac{(2c_j)^s}{\Gamma(s)} \left(\frac{1}{2}\right)^{\frac{s}{2}} 2^{s-2(r-l)} \binom{2(r-l)-s-1}{r-l-1} \int_0^{+\infty} u^{\frac{s}{2}-1} e^{-S_1(u)} du$$

and for $k+1 \le j \le n$

$$\bar{\mu}_{j,r} = \sum_{l=0}^{r} \binom{r}{l} (-1)^l \left(\int_0^{+\infty} e^{-S_1(u)} \lambda_j (2u + \theta_j^2)^{-\frac{1}{2}} du\right)^l$$

$$\times 2^{l-r} \sum_{s=1}^{r-l} \frac{(2\lambda_j)^s}{\Gamma(s)} \binom{2(r-l)-s-1}{r-l-1} \int_0^{+\infty} u^{r-l-1} (2u + \theta_j^2)^{\frac{s}{2}-r+l} e^{-S_1(u)} du.$$

Again, as a special case of Proposition 5 we recover the r–th centered moment for a NID r.v. P on (0,1) distributed according to a distribution characterized by (15), which is given by

$$\bar{\mu}_r = \sum_{l=0}^{r} \binom{r}{l} (-1)^l e^{l\theta\lambda} (\theta K_1(\theta\lambda) c)^l$$

$$\times \sum_{s=1}^{r-l} \frac{(2c)^s}{\Gamma(s)} \left(\frac{1}{2}\right)^{\frac{s}{2}} 2^{s-2(r-l)} \binom{2(r-l)-s-1}{r-l-1} \int_0^{+\infty} u^{\frac{s}{2}-1} \mathrm{e}^{-S_1(u)} du$$

We complete this subsection by providing the Bayes estimator for the success probability in a binary experiment under the assumption that the "prior" distribution for p is characterized by (15). The Bayes estimate for the parameter p is given in terms of the Meijer G–function $G_{p,q}^{l,m}$. Refer to [12, Section 5] for a thorough discussion of the Meijer G–functions: here we just note that they are very general functions whose special cases cover most of the mathematical functions such as the trigonometric functions, Bessel functions and generalized hypergeometric functions.

Proposition 6 *Let $p \sim g_P(p)$ be defined by (15). If $Y_1, \ldots, Y_n | p$ are independent and identically distributed according to a Bernoulli distribution with parameter p, then*

$$\mathbb{E}[p|Y_1, \ldots, Y_n] = \left(\frac{\theta c}{2}\right)^2 \frac{G_{1,3}^{3,0}\left(\frac{c^2\theta^2}{4} \middle| \begin{matrix} 0 \\ -n+m-1/2, -n-1, -n-2 \end{matrix}\right)}{G_{1,3}^{3,0}\left(\frac{c^2\theta^2}{4} \middle| \begin{matrix} 0 \\ -n+m-1/2, -n, -n-1 \end{matrix}\right)}.$$

where $m := \sum_{i=1}^n Y_i$ is the number of successes.

14

3.3. Normalized gamma and inverse–Gaussian distributions

For any $n \in \mathbb{N}$ and fixed $k \in \{1, \ldots, n\}$, let us consider n independent r.v. $X_1, \ldots, X_n$ such that $X_i \sim \mathrm{Ga}(\alpha_i, \gamma_i)$ for $i = 1, \ldots, k$ and $X_i \sim \mathrm{IG}(\lambda_i, \theta_i)$ for $i = k+1, \ldots, n$. Since the p.d.f. of the r.v. $X_1, \ldots, X_n$ are known, the p.d.f. of the NID random vector $(P_1, \ldots, P_n)$ is obtained by applying (1). Some algebra and formula 3.471.9 in [13] lead to

$$g_{(P_1,\ldots,P_n)}(p_1, \ldots, p_{n-1}) = 2 \prod_{i=1}^{k} \frac{\gamma_i^{\alpha_i}}{\Gamma(\alpha_i)} \prod_{i=k+1}^{n} \frac{\lambda_i}{(2\pi)^{1/2}} e^{\theta_i \lambda_i} \tag{19}$$

$$\times \prod_{i=1}^{k} p_i^{\alpha_i - 1} \prod_{i=k+1}^{n-1} p_i^{-\frac{3}{2}} (1 - |\mathbf{p}|)^{-\frac{3}{2}} \left(\frac{C_2(p_1, \ldots, p_{n-1})}{L_2(p_1, \ldots, p_{n-1})} \right)^{\frac{a}{2} - \frac{(n-k)}{4}}$$

$$\times K_{a-(n-k)/2} \left(2 C_2^{\frac{1}{2}}(p_1, \ldots, p_{n-1}) L_2^{\frac{1}{2}}(p_1, \ldots, p_{n-1}) \right) \mathbb{1}_{\Delta^{(n-1)}}(p_1, \ldots, p_{n-1}),$$

where $|\mathbf{p}| := \sum_{i=1}^{n-1} p_i$, $a := \sum_{i=1}^{n} \alpha_i$, K_ν denotes the modified Bessel function of the third kind with index $\nu \in \mathbb{R}$ and where we defined

$$C_2(p_1, \ldots, p_{n-1}) := \frac{1}{2} \sum_{i=k+1}^{n-1} \frac{\lambda_i^2}{p_i} + \frac{1}{2} \frac{\lambda_n^2}{1 - |\mathbf{p}|}$$

and

$$L_2(p_1, \ldots, p_{n-1}) := \sum_{i=1}^{k} p_i \gamma_i + \frac{1}{2} \sum_{i=k+1}^{n-1} \theta_i^2 p_i + \frac{1}{2} \theta_n^2 (1 - |\mathbf{p}|).$$

Let us know consider the particular case of $n = 2$, which yields a new distribution over the unit interval. Let X_1 and X_2 be two independent r.v. such that $X_1 \sim \mathrm{Ga}(\alpha, \gamma)$ and $X_2 \sim \mathrm{IG}(\lambda, \theta)$. The distribution of the NID r.v. $P := X_1/(X_1 + X_2)$ has p.d.f. given by

$$g_P(p) = \frac{2 \gamma^\alpha \lambda e^{\theta \lambda}}{\Gamma(\alpha)(2\pi)^{1/2}} p^{\alpha - 1}(1 - p)^{-\frac{3}{2}} \left(\frac{L_2(p)}{C_2(p)} \right)^{-\frac{\alpha}{2} + \frac{1}{4}} \tag{20}$$

$$\times K_{\alpha - 1/2} \left(2 C_2^{\frac{1}{2}}(p) L_2^{\frac{1}{2}}(p) \right) \mathbb{1}_{(0,1)}(p),$$

where we defined $C_2(p) := \lambda^2 [2(1-p)]^{-1}$ and $L_2(p) := p\gamma + \frac{1}{2}\theta^2(1-p)$. After some algebra and setting $\beta := \lambda\sqrt{2\gamma}$ and $\delta := \frac{\theta\beta}{\sqrt{2\gamma}}$, (20) simplifies to

$$g_P(p) = \frac{2^{-\alpha + 1} \beta^{\alpha + \frac{1}{2}} e^\delta}{\Gamma(\alpha)\sqrt{2\pi}} \frac{p^{\alpha - 1}(1 - p)^{-\frac{\alpha}{2} - \frac{5}{4}}}{(p + \delta^2(1-p))^{\frac{\alpha}{2} - \frac{1}{4}}} K_{\alpha - \frac{1}{2}} \left(\beta \frac{\sqrt{p + \delta^2(1-p)}}{\sqrt{1-p}} \right). \tag{21}$$

Figure 1 depicts the p.d.f. of this NID distribution for some particular choice of the parameters both on the simplex and on the unit interval. The flexibility of such a distribution is apparent and, in particular, it allows also for proper bimodal distributions on $(0, 1)$.

<center>FIGURE 1 AROUND HERE</center>

Moment formulae for a NID random vector $(P_1, \ldots, P_n)$, whose distribution admits density (19), can be obtained by applying Proposition 3. In particular, define

$$S_2(u) := \sum_{i=1}^{k} \alpha_i \log\left(1 + \frac{u}{\gamma_i}\right) + \sum_{i=k+1}^{n} \lambda_i((2u + \theta_i^2)^{\frac{1}{2}} - \theta_i)$$

The next proposition summarizes these moment formulae.

Proposition 7 *If $(P_1, \ldots, P_n)$ has NID distribution admitting density (19), then, for any $r \in \mathbb{N}$, for $1 \le j \le k$*

$$\mu_{j,r} = \frac{1}{\Gamma(r)}(\alpha_j)_r \int_0^\infty u^{r-1}(\gamma_j + u)^{-r} e^{-S_2(u)} du \tag{22}$$

and for $k + 1 \le j \le n$

$$\mu_{j,r} = 2^{-r} \sum_{s=1}^{r} \frac{(2\lambda_j)^s}{\Gamma(s)} \binom{2r - s - 1}{r - 1} \int_0^\infty u^{r-1}(2u + \theta_j^2)^{\frac{s}{2} - r} e^{-S_2(u)} du. \tag{23}$$

As special case of Proposition 7 we recover the r–th moment for a NID r.v. P over $(0, 1)$ distributed according to a distribution characterized by (21). Indeed, we have

$$\mu_r = \frac{1}{\Gamma(r)}(\alpha)_r \int_0^\infty u^{r-1}(\gamma + u)^{-r} e^{-\alpha \log(1 + u/\gamma) - \lambda((2u + \theta^2)^{\frac{1}{2}} - \theta)} du. \tag{24}$$

Remark 3. Note that expression (23) can be further simplified by solving the sum over the index s. In particular, for $k + 1 \le j \le n$ we can write (23) as

$$\mu_{j,r} = 2^{-r} \int_0^{+\infty} \frac{u^{r-1} e^{-S_1(u)}}{(2u + \theta_j^2)^r} \sum_{s=1}^{r} \frac{1}{\Gamma(s)} \binom{2r - s - 1}{r - 1} (2\lambda_j(2u + \theta_j^2)^{1/2})^s du$$

$$= \frac{(-1/2)^{r+1} \pi^{1/2}}{\Gamma(r)}(2\lambda_j)^{1/2 + r}$$

<center>16</center>

$$\times \int_0^{+\infty} \frac{u^{r-1}\mathrm{e}^{-S_2(u)+\lambda_j(\lambda_j(2u+\theta_j^2))^{1/2}}}{(2u+\theta_j^2)^{r/2-1/4}}I_{1/2-r}(\lambda_j(2u+\theta_j^2)^{1/2})du,$$

where I_ν denotes the modified Bessel function of the first kind with index $\nu \in \mathbb{R}$. $\qquad\square$

Moreover, the central moments are provided in the following proposition.

Proposition 8 *If $(P_1,\ldots,P_n)$ has NID distribution admitting density* (19), *then, for any $r \in \mathbb{N}$, for $1 \le j \le k$*

$$\bar{\mu}_{j,r} = \sum_{l=0}^r \binom{r}{l}(-1)^l \left(\int_0^\infty e^{-S_2(u)}\alpha_j(\gamma_j+u)^{-1}du\right)^l$$

$$\times \frac{(\alpha_j)_{(r-l)}}{\Gamma(r-l)}\int_0^\infty u^{r-l-1}(\gamma_j+u)^{-r+l}e^{-S_2(u)}du$$

and for $k+1 \le j \le n$

$$\bar{\mu}_{j,r} = \sum_{l=0}^r \binom{r}{l}(-1)^l \left(\int_0^{+\infty} e^{-S_2(u)}\lambda_j(2u+\theta_j^2)^{-\frac{1}{2}}du\right)^l$$

$$\times 2^{l-r}\sum_{s=1}^{r-l}\frac{(2\lambda_j)^s}{\Gamma(s)}\binom{2(r-l)-s-1}{r-l-1}\int_0^{+\infty}u^{r-l-1}(2u+\theta_j^2)^{\frac{s}{2}-r+l}e^{-S_2(u)}du.$$

As special case of Proposition 8 we recover the r–th centered moment for a NID r.v. P on (0,1) distributed according to a distribution characterized by (21). We have

$$\bar{\mu}_r = \sum_{l=0}^r \binom{r}{l}(-1)^l\mathrm{e}^{l\theta\lambda}\alpha^l \left(\int_0^\infty \mathrm{e}^{-\lambda(2u+\theta^2)^{\frac{1}{2}}}(\gamma+u)^{-1}du\right)^l$$

$$\times \frac{(\alpha)_{(r-l)}}{\Gamma(r-l)}\int_0^\infty u^{r-l-1}(\gamma+u)^{-r+l}\mathrm{e}^{-\alpha\log(1+u/\gamma)-\lambda((2u+\theta^2)^{\frac{1}{2}}-\theta)}du.$$

A final aspect we consider is the derivation of the Bayes estimator for the problem of estimating the success probability in a binary experiment under the assumption that the "prior" distribution for p is given by (21).

Proposition 9 *Let $p \sim g_p(p)$ be defined by* (21). *If $Y_1,\ldots,Y_n|p$ are independent and identically distributed according to a Bernoulli distribution with parameter p, then*

$$\mathbb{E}[p|Y_1,\ldots,Y_n] = (\alpha+m)\left(\frac{\lambda\theta}{2}\right)^2$$

$$\times \frac{\mathrm{G}_{1,3}^{3,0}\left(\frac{\lambda^2\theta^2}{4}\middle|\begin{array}{c}0\\-m-\alpha-1,-n-1,-n-1/2-\alpha\end{array}\right)}{\mathrm{G}_{1,3}^{3,0}\left(\frac{\lambda^2\theta^2}{4}\middle|\begin{array}{c}0\\-m-\alpha,-n,-n+1/2-\alpha\end{array}\right)},$$

where $m := \sum_{i=1}^{n} Y_i$ is the number of successes.

3.4. Normalized gamma and 1/2–stable distributions

For any $n \in \mathbb{N}$ and fixed $k \in \{1,\dots,n\}$, let us consider n independent r.v. $X_1,\dots,X_n$ such that $X_i \sim \mathrm{Ga}(\alpha_i,\gamma_i)$ for $i = 1,\dots,k$ and $X_i \sim \mathrm{St}(1/2,c_i)$ for $i = k+1,\dots,n$. Recall that the 1/2–stable distribution with parameter c can be recovered from the inverse–Gaussian distribution with parameter (c,θ) by setting $\theta = 0$. Therefore, we only highlight the main findings, all other properties are easily obtained from those in Section 3.3. First, since equation 3.471.9 in [13] can be still applied when $\theta = 0$, the NID distribution obtained by combining gamma distributions and 1/2–stable distributions can be recovered from (19) by setting $\theta_i = 0$ and $\lambda_i = c_i$ for $i = k+1,\dots,n$. In particular, we obtain

$$g_{(P_1,\dots,P_n)}(p_1,\dots,p_{n-1}) = 2\prod_{i=1}^{k}\frac{\gamma_i^{\alpha_i}}{\Gamma(\alpha_i)}\prod_{i=k+1}^{n}\frac{c_i}{(2\pi)^{1/2}} \tag{25}$$

$$\times \prod_{i=1}^{k}p_i^{\alpha_i-1}\prod_{k+1}^{n-1}p_i^{-\frac{3}{2}}(1-|\mathbf{p}|)^{-\frac{3}{2}}\left(\frac{C_2(p_1,\dots,p_{n-1})}{L_2(p_1,\dots,p_{n-1})}\right)^{\frac{a}{2}-\frac{(n-k)}{4}}$$

$$\times K_{a-(n-k)/2}\left(2C_2^{\frac{1}{2}}(p_1,\dots,p_{n-1})L_2^{\frac{1}{2}}(p_1,\dots,p_{n-1})\right)\mathbb{1}_{\Delta^{(n-1)}}(p_1,\dots,p_{n-1}),$$

where $a := \sum_{i=1}^{k}\alpha_i$, K_ν denotes the modified Bessel function of the third kind with index $\nu \in \mathbb{R}$,

$$C_2(p_1,\dots,p_{n-1}) := \frac{1}{2}\sum_{i=k+1}^{n-1}\frac{c_i^2}{p_i} + \frac{c_n^2}{1-|\mathbf{p}|}$$

and $L_2(p_1,\dots,p_{n-1}) := \sum_{i=1}^{k}p_i\gamma_i$. Focusing on the special case of NID r.v. P on the unit interval, let X_1 and X_2 be two independent r.v. such that $X_1 \sim \mathrm{Ga}(\alpha,\gamma)$ and $X_2 \sim \mathrm{St}(1/2,c)$. The distribution of the NID r.v. $P := X_1/(X_1+X_2)$ has p.d.f. coinciding with

$$g_P(p) = \frac{2\gamma^\alpha c}{\Gamma(\alpha)(2\pi)^{1/2}}p^{\alpha-1}(1-p)^{-\frac{3}{2}}\left(\frac{L_2(p)}{C_2(p)}\right)^{-\frac{\alpha}{2}+\frac{1}{4}} \tag{26}$$

18

$$\times K_{\alpha-1/2}\left(2C_2^{\frac{1}{2}}(p)L_2^{\frac{1}{2}}(p)\right)\mathbb{1}_{(0,1)}(p),$$

where $C_2(p) := c^2[2(1-p)]^{-1}$ and $L_2(p) := p\gamma$. In (26) one can set without loss of generality $\beta := c\sqrt{2\gamma}$, which yields the simple expression

$$g_P(p) = \frac{2^{-\alpha+1}\,\beta^{\alpha+\frac{1}{2}}}{\Gamma(\alpha)\sqrt{2\pi}}p^{\frac{\alpha}{2}-\frac{3}{4}}(1-p)^{-\frac{\alpha}{2}-\frac{5}{4}}K_{\alpha-\frac{1}{2}}\left(\frac{\beta\sqrt{p}}{\sqrt{1-p}}\right). \qquad (27)$$

Moment formulae for a NID random vector $(P_1,\ldots,P_n)$, whose distribution admits density (25), can be easily recovered from Proposition 7 and Proposition 8 by setting $\theta_i = 0$ and $\lambda_i = c_i$ for $i = k+1,\ldots,n$.

3.4.1. A generalized Bessel function distribution

The distributions in Sections 3.3 and 3.4 are intimately connected with the quite popular Bessel function distribution first introduced in [25]. In fact, the distribution on the unit interval introduced in (21) allows to derive a generalized Bessel function distribution with an additional parameter determining the support of the distribution.

Among several procedures known in the literature for constructing the Bessel function distribution, for our purposes it is useful to recall that a Bessel function distribution with parameter (α, β) can be derived as the distribution of the r.v. $Y = (X_1/X_2)^{1/2}$ where X_1 and X_2 are two independent r.v. such $X_1 \sim \text{Ga}(\alpha, \gamma)$ and $X_2 \sim \text{St}(1/2, c)$. Therefore, a Bessel function r.v. Y admits p.d.f. with respect to the Lebesgue measure on $\mathbb{R}$ given by

$$f_Y(y) = \frac{2^{-\alpha+2}\beta^{\alpha+1/2}}{\Gamma(\alpha)\sqrt{2\pi}}y^{\alpha-\frac{1}{2}}K_{\alpha-1/2}(\beta y)\mathbb{1}_{(0,\infty)}(y), \qquad (28)$$

having set $\beta = c\sqrt{2\gamma}$. This is the one-sided Bessel function distribution, the two-sided is readily obtained by replacing y with $|y|$ and dividing by 2. Now, if P is the r.v. with p.d.f. (27), then one has

$$\left(\frac{X_1}{X_2}\right)^{\frac{1}{2}} \overset{\text{d}}{=} \left(\frac{P}{1-P}\right)^{\frac{1}{2}}$$

and the Bessel function distribution can be seen as a transformation of the r.v. P with p.d.f. (27).

At this point, it seems natural to introduce a generalization of the Bessel function distribution by replacing the NID distribution (27) with the more

general NID distribution based on the inverse Gaussian r.v. which admits density (21). Specifically, for P with p.d.f. (21) define the r.v.

$$Y = \left(\frac{P + \delta(1 - P)}{1 - P} \right)^{\frac{1}{2}}, \tag{29}$$

which takes values in (δ, ∞) and has parameters (α, β, δ). By a simple change of variable from (21) one obtains the p.d.f. of Y, which is given by

$$f_Y(y) = \frac{2^{-\alpha+2}\beta^{\alpha+1/2}e^{\delta}}{\Gamma(\alpha)\sqrt{2\pi}} \left(y^2 - \delta^2 \right)^{\alpha-1} y^{\frac{3}{2}-\alpha} K_{\alpha-1/2}\left(\beta y \right) \mathbb{1}_{(\delta,\infty)}(y), \tag{30}$$

which clearly reduces to the Bessel function distribution (28) if $\delta = 0$. Hence, Y defined in (29) represents a generalized Bessel function distribution. Its two–sided analog with support $(-\infty, \delta) \cup (\delta, +\infty)$.

For the special case of α being an integer one can derive a mixture representation of (30) which is given in the next proposition.

Proposition 10 *The generalized Bessel function distribution defined in (30) with $\alpha = n \in \mathbb{N}$, admits the following mixture representation*

$$f_Y(y) = \sum_{i=0}^{n-1} f_i(y)p_i \tag{31}$$

where

$$f_i(y) = \frac{\beta^{n-i}y^{-n-i+1}(y^2 - \delta^2)^{n-1}e^{-\beta y}}{\sum_{k=0}^{n-1} \binom{n-1}{k}(-1)^k(\beta\delta)^{2k}\Gamma(n - i - 2k, \delta)}$$

$$p_i = \frac{2^{-n-i+1}\Gamma(n+i)e^{\delta} \sum_{k=0}^{n-1} \binom{n-1}{k}(-1)^k(\beta\delta)^{2k}\Gamma(n - i - 2k, \delta)}{\Gamma(n)\Gamma(i+1)\Gamma(n-i)}$$

and $\Gamma(\cdot, \cdot)$ stands for the incomplete gamma function.

Note that this result allows to recover the representation of the Bessel function distribution as mixture of gamma distributions as a special case (see [11]). In fact, by setting $\delta = 0$, $f_i(y)$ becomes the p.d.f. of gamma r.v. with shape parameter $(n - i)$ and scale parameter β and p_i coincides with $2^{-n-i+1}\Gamma(n+i)[\Gamma(n)\Gamma(i+1)]^{-1}$, for $i = 0, \ldots, n - 1$.

We now consider a distribution belonging to Case (ii) according to the classification introduced at the beginning of Section 3, namely the case of $X_1, \ldots, X_n$ having the same ID distribution but different parameters. In particular, we consider the NID distribution based on $X_1, \ldots, X_n$ such that $X_i \sim \text{Ga}(\alpha_i, \gamma_i)$ for $i = 1, \ldots, n$. This represents possibly the simplest generalization of the Dirichlet distribution which arises if $\gamma_i = \gamma$ for $i = 1, \ldots, n$. However, to the authors knowledge, no results are present in the literature. The p.d.f. of this NID random vector $(P_1, \ldots, P_n)$ is easily obtained by applying (1), some algebra and formula 3.381.4 in [13] leading to

$$g_{(P_1,\ldots,P_n)}(p_1, \ldots, p_{n-1}) = \Gamma(a) \prod_{i=1}^{n} \frac{\gamma_i^{\alpha_i}}{\Gamma(\alpha_i)} \prod_{i=1}^{n-1} p_i^{\alpha_i - 1} (1 - |\mathbf{p}|)^{\alpha_n - 1} \qquad (32)$$

$$\times \left(\sum_{i=1}^{n-1} \gamma_i p_i + \gamma_n (1 - |\mathbf{p}|) \right)^{-a} \mathbb{1}_{\Delta^{(n-1)}}(p_1, \ldots, p_{n-1}),$$

where $|\mathbf{p}| := \sum_{i=1}^{n-1} p_i$.

In order to determine the corresponding distribution on the unit interval, let X_1 and X_2 be two independent r.v. such that $X_1 \sim \text{Ga}(\alpha_1, \gamma_1)$ and $X_2 \sim \text{Ga}(\alpha_2, \gamma_2)$. The distribution of the NID r.v. $P := X_1/(X_1 + X_2)$ has p.d.f given by

$$g_P(p) = \frac{\Gamma(a)}{\Gamma(\alpha_1)\Gamma(\alpha_2)} \gamma_1^{\alpha_1} \gamma_2^{\alpha_2} p^{\alpha_1 - 1} (1 - p)^{\alpha_2 - 1} (\gamma_1 p + \gamma_2 (1 - p))^{-a} \mathbb{1}_{(0,1)}(p).$$

$$(33)$$

Figure 2 depicts the p.d.f. of this NID distribution for various choices of the parameters both on the bi–dimensional simplex and on the unit interval.

<center>FIGURE 2 AROUND HERE</center>

Moment formulae for such a NID random vector $(P_1, \ldots, P_n)$ can be obtained as particular cases of Proposition 7 and Proposition 8 by assuming $k = n$. The resulting moment formulae are summarized in the next propositions.

Proposition 11 *If $(P_1, \ldots, P_n)$ has NID distribution admitting density (32), then, for any $r \in \mathbb{N}$, for $1 \le j \le n$*

$$\mu_{j,r} = \frac{\gamma_j^{-r} (\alpha_j)_r}{\Gamma(r)} \int_0^{\infty} u^{r-1} \left(1 + \frac{u}{\gamma_j} \right)^{-r} \prod_{i=1}^{n} \left(1 + \frac{u}{\gamma_i} \right)^{-\alpha_i} du$$

<center>21</center>

As a special case of Proposition 11 we recover the r–th moment for a NID r.v. P over $(0,1)$ with p.d.f. (33)

$$\mu_{j,r} = \frac{(\alpha_1)_r}{(\alpha_1 + \alpha_2)_r} \left(\frac{\gamma_2}{\gamma_1}\right)^r {}_2F_1\left(\alpha_1 + r, r; \alpha_1 + \alpha_2 + r; 1 - \frac{\gamma_2}{\gamma_1}\right)$$

where ${}_2F_1(\cdot, \cdot; \cdot; \cdot)$ is the Gauss hypergeometric function (see [4]).

Proposition 12 *If $(P_1, \ldots, P_n)$ has NID distribution admitting density (32), then, for any $r \in \mathbb{N}$, for $1 \leq j \leq n$*

$$\bar{\mu}_{j,r} = \sum_{l=0}^{r} \binom{r}{l}(-1)^l \left(\int_0^\infty e^{-\sum_{i=1}^n \alpha_i \log(1+u/\gamma_i)} \frac{\alpha_j}{\gamma_j + u} du\right)^l$$

$$\times \frac{\gamma_j^{-r+l}(\alpha_j)_{(r-l)}}{\Gamma(r-l)} \int_0^\infty u^{r-l-1}\left(1 + \frac{u}{\gamma_j}\right)^{-r+l} \prod_{i=1}^n \left(1 + \frac{u}{\gamma_i}\right)^{-\alpha_i} du.$$

From Proposition 12 one can also deduce the r–th centered moment for the corresponding NID r.v. P over $(0,1)$

$$\bar{\mu}_r = \left(\frac{\gamma_2}{\gamma_1}\right)^r \sum_{l=0}^{r} \binom{r}{l}(-\alpha_1)^l \left(\frac{1}{\alpha_2}{}_2F_1\left(1, 1; 1 + \alpha_2; 1 - \frac{\gamma_2}{\gamma_1}\right)\right)^l$$

$$\times \frac{(\alpha_1)_{(r-l)}}{(\alpha_1 + \alpha_2)_{(r-l)}}{}_2F_1\left(\alpha_1 + r - l, r - l; \alpha_1 + \alpha_2 + r - l; 1 - \frac{\gamma_2}{\gamma_1}\right).$$

3.6. Some known NID distributions

In the literature one can find various NID distributions corresponding to the very special case of independent r.v. $X_1, \ldots, X_n$ with the same ID distribution and, moreover, differing solely in their shape parameter. Such a case is important for Bayesian nonparametric applications since the resulting NID distributions are additive and, therefore, they can be used for constructing random probability measures. In fact, one can show that there exists a random probability measure which admits any such NID distribution as underlying family of finite-dimensional distributions and the additivity property guarantees that the necessary consistency conditions are satisfied. In addition to the Dirichlet distribution and the normalized stable distribution considered in Section 2, other such NID distributions are the normalized inverse Gaussian distribution [22], the generalized Dirichlet distribution [23] and the normalized tempered stable distribution [21, 24], which is often also termed normalized generalized gamma distribution. Apart from the Dirichlet case, only the first few moments of these distributions are known. Hence, by applying the general moment formulae derived in Section 2, one can easily derive both their raw and central moments of any order.

4. Discussion and further developments

The simplex plays an important role as sample space in many practical situations where compositional data, in the form of proportions of some whole, are to be analyzed. Historically, due to its simplicity in terms of both mathematics and interpretation, the Dirichlet distribution has represented the first tool for modeling such data. However, as remarked by [2], the Dirichlet distribution is over–structured for practical applications because of its many strong independence properties. Such properties make it indisputably the reference model for expressing the non trivial idea of substantial independence for compositions but at the same time heavily restrict its potential for applications. In the light of such inadequacies a powerful methodology based on log ratio transformations of the original variables has been proposed in [1, 2]. In such an approach parametric models are built in the unconstrained transformed sample space. It is still an open problem to find a tractable parametric class of distributions which contains the Dirichlet distribution as well as other models able to exhibit significant departures from its strong independence properties.

In the present paper we proposed a simple alternative approach, based on the idea of "normalization", to the problem of defining distributions on the simplex: we introduced the class of NID distributions which represents a natural extension of the Dirichlet distribution. Distributional properties and some general moment formulae for the class of NID distributions have been derived and interesting and tractable special cases have been identified by combining three well-known ID distributions: the gamma distribution, the $1/2$–stable distribution and the inverse–Gaussian distribution.

In light of the results obtained for the NID distributions, we believe that the "normalization" approach could represent a useful way to introduce new distributions on the simplex handling various dependence structures, exhibiting a richer parametrization, containing the Dirichlet distribution as a special case and preserving some of its good mathematical properties. Future work on NID distributions will consist of a thorough study of the theoretical properties most relevant for the analysis of compositional data.

Acknowledgements

Appendix

PROOF OF PROPOSITION 1. For the linear functional $\sum_{i=1}^{n} f(y_i)P_i$ one can write

$$\mathbb{P}\left(\sum_{i=1}^{n} f(y_i)P_i \le y\right) = \mathbb{P}\left(\sum_{i=1}^{n} (f(y_i) - y)X_i \le 0\right),$$

where $\sum_{i=1}^{n} X_i(f(y_i) - y)$ is a linear functional of the r.v. $X_1, \ldots, X_n$. Now, it is known that the characteristic function of a linear functional of an ID r.v. is given by

$$\mathbb{E}[e^{it\sum_{i=1}^{n} f(y_i)X_i}] = \exp\left\{-\sum_{i=1}^{n} \int_0^\infty (1 - e^{itvf(y_i)})\nu_i(dv)\right\}$$

for every $t \in \mathbb{R}$. Hence, by resorting to Gurland's inversion formula for characteristic functions ([16]), we obtain

$$\mathbb{P}\left(\sum_{i=1}^{n} (f(y_i) - y)X_i \le 0\right)$$

$$= \frac{1}{2} - \frac{1}{\pi} \lim_{\substack{\varepsilon \downarrow 0 \\ T \uparrow \infty}} \int_\varepsilon^T \frac{1}{t} \text{Im}\left[\exp\left\{-\sum_{i=1}^{n} \int_0^\infty (1 - e^{itv(f(y_i)-y)})\nu_i(dv)\right\}\right] dt$$

$$= \frac{1}{2} - \frac{1}{\pi} \lim_{T \uparrow \infty} \int_0^T \frac{1}{t} \exp\left\{\sum_{i=1}^{n} \int_0^\infty (\cos(tv(f(y_i) - y)) - 1)\nu_i(dv)\right\}$$

$$\times \sin\left(\sum_{i=1}^{n} \int_0^\infty \sin(tv(f(y_i) - y))\nu_i(dv)\right) dt,$$

where $\text{Im}[z]$ stands for the imaginary part of $z \in \mathbb{C}$ and the absolute integrability in the origin follows by minimal modifications of the arguments exploited [31, Proposition 2]. In particular, if we set the function f equal to the indicator function $\mathbb{1}_{\{y_j\}}$ we obtain the distribution of the r.v. P_j. $\qquad\square$

PROOF OF PROPOSITION 2. From Fubini's theorem and by using the independence of the X_i's we obtain

$$\mathbb{E}[P_j] = \int_0^\infty \mathbb{E}[X_j e^{-uX_j}]\mathbb{E}[e^{-u\sum_{1 \le i \ne j \le n} X_i}]du$$

$$= \int_0^\infty \mathbb{E}\left[-\frac{d}{du}e^{-uX_j}\right]\mathbb{E}[e^{-u\sum_{1 \le i \ne j \le n} X_i}]du$$

24

$$= \int_0^\infty \left(-\frac{d}{du} \mathbb{E}[\mathrm{e}^{-uX_j}] \right) \mathbb{E}[\mathrm{e}^{-u \sum_{1 \leq i \neq j \leq n} X_i}] du$$

$$= \int_0^\infty \left(\frac{d}{du} \Psi_j(u) \right) \mathrm{e}^{-\sum_{i=1}^n \Psi_i(u)} du,$$

where the last equality follows by the Lévy–Khintchine representation of an ID distributions (1). □

PROOF OF PROPOSITION 4. We start by introducing some useful tools and notations to the used also in the sequel. For any non-negative integer N, $(a)_N = \Gamma(a+N)/\Gamma(a)$ is the N–th ascending factorial of $a \in \mathbb{R}$. Moreover, for any $n, k \in \mathbb{N}$ with $k \leq n$ and for any $\alpha \in \mathbb{R}$ the symbol $\mathscr{C}(n, k, \alpha)$ stands for the coefficient of the kth order factorial of t in the expansion of the nth generalized factorial of t with scale parameter α, i.e. $(\alpha t)_n = \sum_{k=0}^n \mathscr{C}(n, k, \alpha)(t)_k$. See [6]. Let us now consider the moment formula (9). For $j = 1, \ldots, k$ we have the following identity for the Lapace exponent of a r.v. $X_j \sim \mathrm{St}(1/2, c_j)$ (see (12))

$$-c_j \frac{d^r}{du^r} (2u)^{\frac{1}{2}} = (-1)^r c_j u^{-r} \left(\frac{u}{2} \right)^{\frac{1}{2}} \left(\frac{1}{2} \right)_{(r-1)}.$$

Then, for $j = 1, \ldots, k$, we can write

$$\mu_{j,r} = \frac{1}{\Gamma(r)} \int_0^\infty u^{-1} \mathrm{e}^{-S_1(u)} \sum_{s=1}^r c_j^s \left(\frac{u}{2} \right)^{\frac{s}{2}}$$

$$\times \frac{1}{s!} \sum_{(r_1, \ldots, r_s) \in \mathcal{D}_{r,s}} \frac{r!}{r_1! \cdots r_s!} \prod_{i=1}^s \left(\frac{1}{2} \right)_{(r_i-1)} du.$$

By combining expression (97), (98) and (88) in [29] one finds the following identity

$$\frac{1}{s!} \sum_{(r_1, \ldots, r_s) \in \mathcal{D}_{r,s}} \frac{r!}{r_1! \cdots r_s!} \prod_{i=1}^s \left(\frac{1}{2} \right)_{(r_i-1)} = 2^s \mathscr{C} \left(r, s, \frac{1}{2} \right). \qquad (\text{A.1})$$

Then,

$$\mu_{j,r} = \frac{1}{\Gamma(r)} \sum_{s=1}^r (2c_j)^s \left(\frac{1}{2} \right)^{\frac{s}{2}} \mathscr{C} \left(r, s, \frac{1}{2} \right) \int_0^\infty u^{\frac{s}{2}-1} \mathrm{e}^{-S_1(u)} du$$

(by expression (86) in [29])

$$= 2^{-2r} \sum_{s=1}^{r} \frac{2^{2s} c_j^s}{\Gamma(s)} \left(\frac{1}{2}\right)^{\frac{s}{2}} \binom{2r-s-1}{r-1} \int_0^\infty u^{\frac{s}{2}-1} \mathrm{e}^{-S_1(u)} du.$$

For $j = k+1, \ldots, n$ we can proceed using similar arguments and have the following identity for the Laplace exponent of a r.v. $X_j \sim \mathrm{IG}(\lambda_j, \theta_j)$ (see (13))

$$-\lambda_j \frac{d^r}{du^r} (2u + \theta_j^2)^{\frac{1}{2}} = (-1)^r \lambda_j 2^{r-1} (2u + \theta_j^2)^{\frac{1}{2}-r} \left(\frac{1}{2}\right)_{(r-1)}.$$

Then, for $j = k+1, \ldots, n$, we have

$$\mu_{j,r} = \frac{1}{\Gamma(r)} \int_0^\infty u^{r-1} \mathrm{e}^{-S_1(u)} \sum_{s=1}^{r} \lambda_j^s 2^{r-s} (2u + \theta_j^2)^{\frac{s}{2}-r}$$

$$\times \frac{1}{s!} \sum_{(r_1,\ldots,r_s) \in \mathcal{D}_{r,s}} \frac{r!}{r_1! \cdots r_s!} \prod_{i=1}^{s} \left(\frac{1}{2}\right)_{(r_i-1)} du$$

and by using identity (A.1)

$$\mu_{j,r} = \frac{1}{\Gamma(r)} 2^r \sum_{s=1}^{r} \lambda_j^s \mathscr{C}\left(r, s, \frac{1}{2}\right) \int_0^\infty u^{r-1} (2u + \theta_j^2)^{\frac{s}{2}-r} \mathrm{e}^{-S_1(u)} du$$

(by expression (86) in [29])

$$= 2^{-r} \sum_{s=1}^{r} \frac{(2\lambda_j)^s}{\Gamma(s)} \binom{2r-s-1}{r-1} \int_0^\infty u^{r-1} (2u + \theta_j^2)^{\frac{s}{2}-r} \mathrm{e}^{-S_1(u)} du.$$

$\square$

PROOF OF PROPOSITION 5. Using the moment formula (10), the proof follows by exploiting essentially the same arguments of the proof of Proposition (4). $\square$

PROOF OF PROPOSITION 6. First, observe that the p.d.f. (15) can be written as

$$g_P(p) = \frac{c\theta\lambda\mathrm{e}^{\theta\lambda}}{\pi} p^{-1}(1-p)^{-\frac{1}{2}} (\lambda^2 p + c^2(1-p))^{-\frac{1}{2}}$$

$$\times K_{-1}\left(\theta p^{-\frac{1}{2}} (\lambda^2 p + c^2(1-p))^{\frac{1}{2}}\right) \mathbb{1}_{(0,1)}(p);$$

26

then

$$\int_0^1 p^m(1-p)^{n-m}q(p)dp = \frac{\theta c^2 e^{\theta c}}{\pi} \int_0^1 p^{m-1}(1-p)^{n-m-\frac{1}{2}}((1-p)c^2+pc^2)^{-\frac{1}{2}}$$

$$\times K_{-1}\left(\theta p^{-\frac{1}{2}}\left((1-p)c^2+pc^2\right)^{\frac{1}{2}}\right)dp$$

$$= \frac{\theta c e^{\theta c}}{\pi} \int_0^1 p^{m-1}(1-p)^{n-m-\frac{1}{2}}K_{-1}\left(\theta c p^{-\frac{1}{2}}\right)dp.$$

(by using the change of variable $1/p = x$ and formula 6.592.4 in [13])

$$= \left(\frac{\theta c}{2}\right)^{2n+2}\frac{1}{\pi}e^{\theta c}\gamma\left(n-m+\frac{1}{2}\right)$$

$$\times \mathrm{G}_{1,3}^{3,0}\left(\frac{\theta^2 c^2}{4}\,\middle|\,\begin{matrix}0\\-n+m-1/2,-n,-n-1\end{matrix}\right).$$

Therefore, an application of the Bayes theorem leads to obtain the p.d.f. of the r.v. $p|Y_1,\ldots,Y_n$, i.e.

$$f_{p|Y_1,\ldots,Y_n}(p) = \theta c \left(\frac{2}{\theta c}\right)^{2n+2}\gamma\left(n-m+\frac{1}{2}\right)p^{m-1}(1-p)^{n-m-\frac{1}{2}}$$

$$\times K_{-1}\left(\theta c p^{-\frac{1}{2}}\right)\left(\mathrm{G}_{1,3}^{3,0}\left(\frac{\theta^2 c^2}{4}\,\middle|\,\begin{matrix}0\\-n+m-1/2,-n,-n-1\end{matrix}\right)\right)^{-1}\mathbb{1}_{(0,1)}(p)$$

and

$$\mathbb{E}[p|Y_1,\ldots,Y_n] = \theta c \left(\frac{2}{\theta c}\right)^{2n+2}\gamma\left(n-m+\frac{1}{2}\right)$$

$$\times \left(\mathrm{G}_{1,3}^{3,0}\left(\frac{c^2\theta^2}{4}\,\middle|\,\begin{matrix}0\\-n+m-1/2,-n-1,-n\end{matrix}\right)\right)^{-1}$$

$$\times \int_0^1 p^m(1-p)^{n-m-\frac{1}{2}}K_{-1}\left(\theta c p^{-\frac{1}{2}}\right)dp$$

$$= \left(\frac{\theta c}{2}\right)^2\frac{\mathrm{G}_{1,3}^{3,0}\left(\frac{c^2\theta^2}{4}\,\middle|\,\begin{matrix}0\\-n+m-1/2,-n-1,-n-2\end{matrix}\right)}{\mathrm{G}_{1,3}^{3,0}\left(\frac{c^2\theta^2}{4}\,\middle|\,\begin{matrix}0\\-n+m-1/2,-n,-n-1\end{matrix}\right)},$$

where the last equality is obtained by using the change of variable $1/p = x$ and by applying formula 6.592.4 in [13]. $\qquad\square$

PROOF OF PROPOSITION 7. In the following, for any $n, k \in \mathbb{N}$ with $k \leq n$, the symbol $s(n, k)$ denotes the (n, k)th Stirling number of the first kind. See [6]. The proof follows lines similar to those of the proof of Proposition 4. Let us consider the moment formula (9). For $j = 1, \ldots, k$ we have the following identity for the Laplace exponent of a r.v. $X_j \sim \text{Ga}(\alpha_j, \gamma_j)$ (see (11))

$$-\alpha_j \frac{d^k}{du^k} \log\left(1 + \frac{u}{\gamma_j}\right) = (-1)^r \alpha_j (\gamma_j + u)^{-k} (1)_{(k-1)}.$$

Then, for $j = 1, \ldots, k$, we can write

$$\mu_{j,r} = \frac{1}{\Gamma(r)} \int_0^\infty u^{r-1} e^{-S_2(u)} \sum_{s=1}^r \alpha_j^s (\gamma_j + u)^{-r}$$

$$\times \frac{1}{s!} \sum_{(r_1, \ldots, r_s) \in \mathcal{D}_{r,s}} \frac{r!}{r_1! \cdots r_s!} \prod_{i=1}^s (1)_{(r_i-1)} du.$$

By combining expression (97), (98) and (88) in [29] one finds the following identity

$$\frac{1}{s!} \sum_{(r_1, \ldots, r_s) \in \mathcal{D}_{r,s}} \frac{r!}{r_1! \cdots r_s!} \prod_{i=1}^s \left(\frac{1}{2}\right)_{(r_i-1)} = |s(r, s)|.$$

Then,

$$\mu_{j,r} = \frac{1}{\Gamma(r)} \int_0^\infty u^{r-1} e^{-S_2(u)} \sum_{s=1}^r \alpha_j^s (\gamma_j + u)^{-r} |s(r, s)| du$$

$$= \frac{1}{\Gamma(r)} \sum_{s=1}^r \alpha_j^s |s(r, s)| \int_0^\infty u^{r-1} (\gamma_j + u)^{-r} e^{-S_2(u)} du$$

$$= \frac{1}{\Gamma(r)} (\alpha_j)_r \int_0^\infty u^{r-1} (\gamma_j + u)^{-r} e^{-S_2(u)} du.$$

For $j = k+1, \ldots, n$ we can proceed using similar arguments. For $j = k+1, \ldots, n$ we have the following identity for the Lapace exponent of a r.v. $X_j \sim \text{IG}(\lambda_j, \theta_j)$ (see (13))

$$-\lambda_j \frac{d^r}{du^r} (2u + \theta_j^2)^{\frac{1}{2}} = (-1)^r \lambda_j 2^{r-1} (2u + \theta_j^2)^{\frac{1}{2}-r} \left(\frac{1}{2}\right)_{(r-1)}.$$

Then, for $j = k+1, \ldots, n$, we have

$$\mu_{j,r} = \frac{1}{\Gamma(r)} \int_0^\infty u^{r-1} e^{-S_2(u)} \sum_{s=1}^r \lambda_j^s 2^{r-s} (2u + \theta_j^2)^{\frac{s}{2}-r}$$

28

$$\times \frac{1}{s!} \sum_{(r_1,\ldots,r_s)\in\mathcal{D}_{r,s}} \frac{r!}{r_1!\cdots r_s!} \prod_{i=1}^{s} \left(\frac{1}{2}\right)_{(r_i-1)} du$$

and by using identity (A.1)

$$\mu_{j,r} = \frac{1}{\Gamma(r)} 2^r \sum_{s=1}^{r} \lambda_j^s \mathscr{C}\left(r,s,\frac{1}{2}\right) \int_0^\infty u^{r-1}(2u+\theta_j^2)^{\frac{s}{2}-r} e^{-S_2(u)} du$$

(by expression (86) in [29])

$$= 2^{-r} \sum_{s=1}^{r} \frac{(2\lambda_j)^s}{\Gamma(s)} \binom{2r-s-1}{r-1} \int_0^\infty u^{r-1}(2u+\theta_j^2)^{\frac{s}{2}-r} e^{-S_2(u)} du.$$

$\square$

PROOF OF PROPOSITION 8. Using the moment formula (10), the proof follows by the same arguments of the proof of Proposition (7). $\square$

PROOF OF PROPOSITION 9. First, observe that the p.d.f. (21) can be written as

$$g_P(p) = \frac{2\gamma^\alpha \lambda^{\alpha+1/2} e^{\theta\lambda}}{\Gamma(\alpha)(2\pi)^{1/2}} p^{\alpha-1}(1-p)^{-\frac{\alpha}{2}-\frac{5}{4}}(2\gamma p + \theta^2(1-p))^{-\frac{\alpha}{2}+\frac{1}{4}}$$

$$\times K_{\alpha-1/2}\left(\lambda(1-p)^{-\frac{1}{2}}(2p\gamma+\theta^2(1-p))^{\frac{1}{2}}\right) \mathbb{1}_{(0,1)}(p).$$

then

$$\int_0^1 p^m(1-p)^{n-m} q(p)dp = \frac{e^{\theta\lambda}(2\lambda\theta)^{\alpha+1/2}}{\Gamma(\alpha)\pi^{1/2}}$$

$$\times \int_0^1 p^{\alpha+m-1}(1-p)^{n-\frac{\alpha}{2}-m-\frac{5}{4}} K_{\alpha-1/2}\left(\lambda\theta(1-p)^{-\frac{1}{2}}\right) dp.$$

(by using the change of variable $1/1-p = x$ and formula 6.592.4 in [13])

$$= \frac{e^{\theta\lambda} 2^{-2n}(\lambda\theta)^{2n+2\alpha}}{\Gamma(\alpha)\pi^{1/2}} \Gamma(m+\alpha)$$

$$\times G_{1,3}^{3,0}\left(\frac{\lambda^2\theta^2}{4}\,\middle|\,\begin{matrix} 0 \\ -m-\alpha, -n, -n+1/2-\alpha \end{matrix}\right).$$

Therefore, an application of the Bayes theorem leads to the p.d.f. of the r.v. $p|Y_1,\ldots,Y_n$, i.e.

$$f_{p|Y_1,\ldots,Y_n}(p) = \frac{2^{2n+\alpha+1/2}(\lambda\theta)^{1/2-2n-\alpha}}{\Gamma(m+\alpha)} p^{m+\alpha-1}(1-p)^{n-m-\frac{\alpha}{2}-\frac{5}{4}}$$

29

$$\times \left(G_{1,3}^{3,0}\left(\frac{\lambda^2\theta^2}{4} \middle| \begin{array}{c} 0 \\ -m-\alpha, -n, -n+1/2-\alpha \end{array} \right) \right)^{-1} \mathbb{1}_{(0,1)}(p)$$

and

$$\mathbb{E}[p|Y_1,\ldots,Y_n] = \frac{2^{2n+\alpha+1/2}(\lambda\theta)^{1/2-2n-\alpha}}{\Gamma(m+\alpha)}$$

$$\times \left(G_{1,3}^{3,0}\left(\frac{\lambda^2\theta^2}{4} \middle| \begin{array}{c} 0 \\ -m-\alpha, -n, -n+1/2-\alpha \end{array} \right) \right)^{-1}$$

$$\times \int_0^1 p^{\alpha+m}(1-p)^{n-m-\frac{\alpha}{2}-\frac{5}{4}} K_{\alpha-1/2}\left(\lambda\theta(1-p)^{-\frac{1}{2}} \right) dp$$

$$= (\alpha+m)\left(\frac{\lambda\theta}{2} \right)^2$$

$$\times \frac{G_{1,3}^{3,0}\left(\frac{\lambda^2\theta^2}{4} \middle| \begin{array}{c} 0 \\ -m-\alpha-1, -n-1, -n-1/2-\alpha \end{array} \right)}{G_{1,3}^{3,0}\left(\frac{\lambda^2\theta^2}{4} \middle| \begin{array}{c} 0 \\ -m-\alpha, -n, -n+1/2-\alpha \end{array} \right)},$$

where the last expression is obtained by using the change of variable $1/1-p = x$ and then by applying equation 6.592.4 in [13]. $\square$

PROOF OF PROPOSITION 10. For $\alpha = n \in \mathbb{N}$, one can exploit the finite sum representation of the Bessel function given by

$$K_{n-1/2}(z) = \sqrt{\frac{\pi}{2z}} e^{-z} \sum_{i=0}^{n-1} \frac{(n-1+i)!}{i!(n-1-i)!(2z)^i}.$$

Hence, by applying it to (30) and collecting the terms in y one obtains $f_i(y)$, for $i = 0,\ldots,n-1$, up to the constant of proportionality. The latter can be obtained by applying Newton's binomial formula to the integrand so to recognize the incomplete gamma function after switching integral and sum. By carrying out the appropriate simplifications, the result then follows. $\square$

References

[1] J. Aitchison, S.M. Shen, Logistic–Normal distributions: some properties and uses, Biometrika 67 (1980), 261–272.

[2] J. Aitchison, The statistical analysis of compositional data, J. Roy. Statist. Soc. Ser. B 44 (1982), 139–177.

[3] J. Aitchison, A general class of distributions on the simplex, J. Roy. Statist. Soc. Ser. B 47 (1985), 136–146.

[4] G.E. Andrews, R. Askey, R., R. Roy, Special functions, Cambridge University Press, Cambridge, 2001.

[5] M.A. Carlton, A family of densities derived from the three-parameter Dirichlet process, Adv. Appl. Probab. 39 (2002), 764–774.

[6] C.A. Charalambides, Combinatorial methods in discrete distributions, Wiley, Hoboken, 2005.

[7] L. Comtet, Advanced Combinatorics, Reidel Publishing Company, Boston, 1974.

[8] R.J. Connnor, I.E. Mosimann, Concepts of independence for proportions with a generalization of the Dirichlet distribution. J. Amer. Statist. Assoc. 64 (1969), 194-206.

[9] S.Y. Dennis, On the Hyper-Dirichlet Type 1 and Hyper-Liouville distributions, Commun. Statist. A 20 (1991), 4069-4081.

[10] S.Y. Dennis, A Bayesian analysis of tree-structured statistical decision problems. J. Statist. Plann. Inference 53 (1992), 323-344.

[11] D. Dufresne, Algebraic Properties of Beta and Gamma Distributions, and Applications Adv. Appl. Math. 20 (1998), 285-299.

[12] A. Erdélyi, W. Magnus, F. Oberhettinger, F.G. Tricomi, Higher trascendental functions Vol. I. McGraw-Hill, New York, 1953.

[13] L.S. Gradshteyn, L.M. Ryzhik, Table of integrals, series, and products, Academic Press, New York, 2000.

[14] J.E. Griffin, The Ornstein Uhlenbeck Dirichlet Process and other time-varying processes for Bayesian nonparametric inference, CRISM Working Papers 07-03.

[15] J. Griffin, S.G. Walker, Posterior Simulation of Normalized Random Measure Mixtures, J. Comput. Graph. Statist., doi:10.1198/jcgs.2010.08176.

[16] J. Gurland, Inversion formulae for the distribution of ratios, Ann. Math. Statist. 19 (1948), 228-237.

[17] I.R. James, Products of independent beta variables with applications to Connor and Mosimann's generalized Dirichlet distribution, J. Amer. Statist. Assoc. 67 (1972), 910-912

[18] I.R. James, Multivariate distributions which have beta conditional distributions, J. Amer. Statist. Assoc. 70 (1975), 681-684.

[19] L.F. James, A. Lijoi, I. Prünster, Conjugacy as a distinctive feature of the Dirichlet process, Scand. J. Statist. 20 (2006), 1–16

[20] L.F. James, A. Lijoi, I. Prünster, Posterior analysis of normalized random measures with independent increments, Scand. J. Statist. 36 (2009), 76-97.

[21] M. Kolossiatis, J.E. Griffin, M.F.J. Steel, Modeling overdispersion with the normalized tempered stable distribution, Computational Statistics & Data Analysis, in press (2011), DOI: 10.1016/j.csda.2011.01.016.

[22] A. Lijoi, R.H. Mena, I. Prünster, Hierarchical mixture modelling with normalized inverse–Gaussian priors, J. Amer. Stat. Assoc. 100 (2005), 1278-1291.

[23] A. Lijoi, R.H. Mena, I. Prünster, Bayesian Nonparametric analysis for a generalized Dirichlet process prior, Stat. Inf. Stoc. Proc. 8 (2005), 283–309.

[24] A. Lijoi, R.H. Mena, I. Prünster, Controlling the reinforcement in Bayesian nonparametric mixture models, J. Roy. Statist. Soc. Ser. B 69 (2007), 715–740.

[25] A.T. McKay, A Bessel Function Distribution, Biometrika 24 (1932), 39-44.

[26] K.W. Ng KW, M.L. Tang, M. Tan, G.L. Tian, Grouped Dirichlet distribution: A new tool for incomplete categorical data analysis, J. Multiv. Anal. 99 (2008), 490–509.

[27] K.W. Ng, M.L. Tang, G.L. Tian, M. Tan, The nested Dirichlet distribution and incomplete categorical data analysis, Statist. Sinica 19 (2009), 251-271.

[28] L.E. Nieto-Barajas, I. Prünster, S.G. Walker, Normalized random measures driven by increasing additive processes, Ann. Statist. 32 (2004), 2343-2360.

[29] J. Pitman, Brownian motion, bridge, excursion and meander characterized by sampling at independent uniform times, Electron. J. Probab. 4 (1999), 1-33.

[30] W.S. Rayens, C. Srinivasan, Dependence properties of generalized Liouville distributions on the simplex, J. Amer. Statist. Assoc. 89 (1994), 1465-1470.

[31] E. Regazzini, A. Lijoi, I. Prünster, Distributional results for means of random measures with independent increments, Ann. Statist. 31 (2003), 560-585.

[32] W.F. Steutel, K. Van Harn, Infinite divisibility of probability distributions on the real line, Marcel Dekker, New York, 2004.

[33] G.L. Tian, M.L. Tang, K.C. Yuen, K.W. Ng, Further properties and new applications for the nested Dirichlet distribution, Comput. Statist. Data Anal. 54 (2010), 394–405.
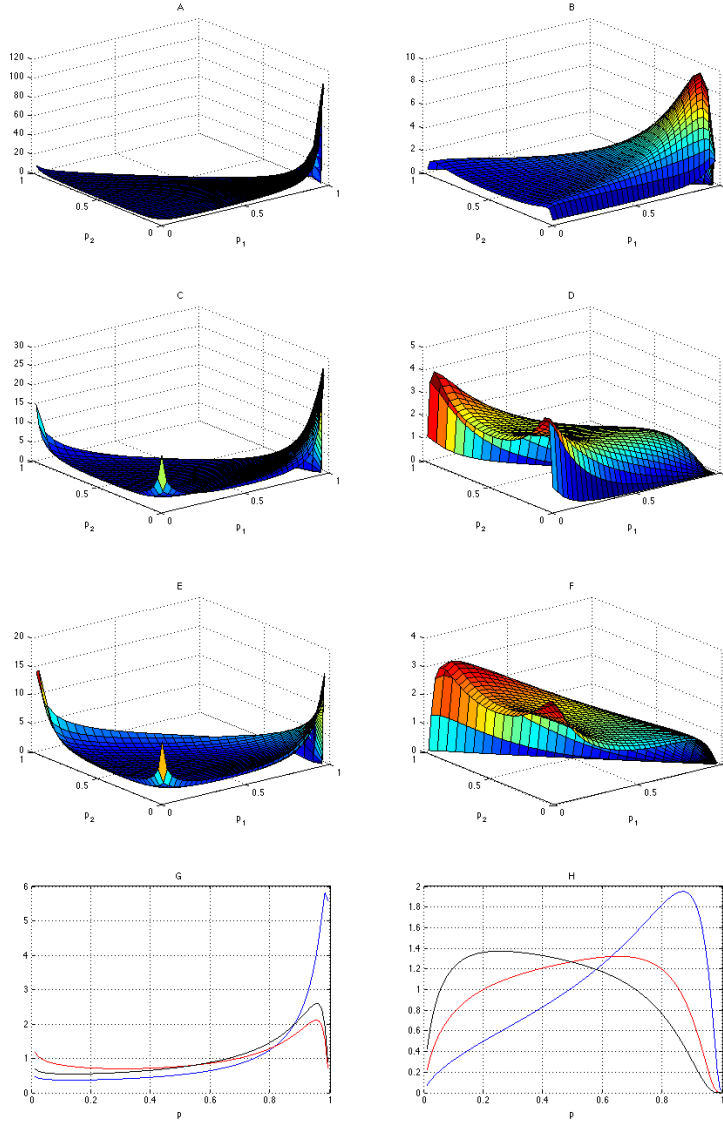
Figure .1: NID distribution based on gamma and inverse–Gaussian r.v.: A) on $\Delta^{(2)}$ with $\alpha_1 = 0.5790$, $\gamma_1 = 1$, $\lambda_i = 0.1667$, $\theta_i = 1$, $i = 1, 2$. B) on $\Delta^{(2)}$ with $\alpha_i = 1.09$, $\gamma_1 = 1$, $\lambda_i = 0.6$, $\theta_i = 1$, $i = 1, 2$. C) on $\Delta^{(2)}$ with $\alpha_1 = 0.5790$, $\gamma_1 = 3$, $\lambda_i = 0.1667$, $\theta_i = 1$, $i = 1, 2$. D) on $\Delta^{(2)}$ with $\alpha_1 = 1.09$, $\gamma_1 = 3$, $\lambda_i = 0.6$, $\theta_i = 1$, $i = 1, 2$. E) on $\Delta^{(2)}$ with $\alpha_i = 0.5790$, $\gamma_i = 3$, $i = 1, 2$, $\lambda_1 = 0.1667$, $\theta_1 = 1$. F) on $\Delta^{(2)}$ with $\alpha_i = 1.09$, $\gamma_i = 3$, $i = 1, 2$, $\lambda_1 = 0.6$, $\theta_1 = 1$. G) on $(0, 1)$ with $\alpha_1 = 0.8685$, $\lambda_1 = 0.25$, $\gamma_1 = 1$, $\theta_1 = 1$ (blue), $\gamma_1 = 3$, $\theta_1 = 1$ (red) and with $\gamma_1 = 3$, $\theta_1 = 2$ (black). H) on $(0, 1)$ with $\alpha_1 = 1.6350$, $\lambda_1 = 0.9$, $\gamma_1 = 1$, $\theta_1 = 1$ (blue), $\gamma_1 = 2$, $\theta_1 = 1$ (red) and with $\gamma_1 = 3$, $\theta_1 = 1$ (black).

33

Figure .2: NID distribution based on gamma r.v. with different scale and shape parameteres: A) on $\Delta^{(2)}$ with $\alpha_i = 0.5790$, $\gamma_i = 1$, $i = 1, 2, 3$. B) on $\Delta^{(2)}$ with $\alpha_i = 1.09$, $\gamma_i = 1$, $i = 1, 2, 3$. C) on $\Delta^{(2)}$ with $\alpha_i = 0.5790$, $i = 1, 2, 3$, $\gamma_1 = 3$, $\gamma_2 = \gamma_3 = 1$. D) on $\Delta^{(2)}$ with $\alpha_i = 1.09$, $i = 1, 2, 3$, $\gamma_1 = \gamma_2 = 1$, $\gamma_3 = 1$. E) on $\Delta^{(2)}$ with $\alpha_i = 0.5790$, $i = 1, 2, 3$, $\gamma_1 = \gamma_2 = 3$, $\gamma_3 = 2$. F) on $\Delta^{(2)}$ with $\alpha_i = 1.09$, $i = 1, 2, 3$, $\gamma_1 = 2$, $\gamma_2 = 1$, $\gamma_3 = 3$. G) on $(0, 1)$ with $\alpha_i = 0.8685$, $i = 1, 2$, $\gamma_i = 1$, $i = 1, 2$, (blue), with $\alpha_i = 0.8685$ $i = 1, 2$, $\gamma_1 = 3$, $\gamma_2 = 2$ (red) and with $\gamma_1 = 2$, $\gamma_2 = 3$ (black). H) on $(0, 1)$ with $\alpha_i = 1.6350$, $i = 1, 2$ (blue), with $\alpha_i = 1.6350$, $i = 1, 2$, $\gamma_1 = 3$, $\gamma_2 = 2$ (red) and with $\gamma_1 = 2$, $\gamma_2 = 3$ (black).

34