JOURNAL OF MULTIVARIATE ANALYSIS 54, 1–17 (1995)

Multivariate Liouville Distributions, IV

RAMESHWAR D. GUPTA*

University of New Brunswick, St. John, Canada

AND

DONALD ST. P. RICHARDS[†]

University of Virginia

DEDICATED TO INGRAM OLKIN ON THE OCCASION OF HIS SEVENTIETH BIRTHDAY

We define a class of distributions, containing the classical Dirichlet and Liouville distributions, in which the random variables take values in a locally compact Abelian group or semigroup. These generalizations retain many properties of the Dirichlet and Liouville distributions, including properties of the marginal and conditional distributions and regression functions. We present a number of examples illustrating the general theory. © 1995 Academic Press, Inc.

1. INTRODUCTION

Let us recall [15–18] that an absolutely continuous random vector $(X_1, ..., X_n) \in \mathbb{R}_+^n$ has a multivariate Liouville distribution if its joint density function is of the form

$$f\left(\sum_{i=1}^{n} x_i\right) \prod_{i=1}^{n} x_i^{a_i - 1}, \tag{1.1}$$

Received July 15, 1993; revised November 1994.

AMS 1990 subject classifications: primary 43A05, 60B05, 62H05; secondary 33D80, 43A85, 43A90, 60B15, 62H10.

Key words and phrases: convolution family, convolution semigroup, Dirichlet distribution, Gauss distribution, hypergeometric function on symmetric cones, hypergeometric function of matrix argument, Liouville–Dirichlet integral, locally compact Abelian group, mixtures of exponential distributions, multinomial distribution, multivariate negative binomial distribution, Poisson distribution, symmetric cone, Weyl fractional derivative.

* Research supported in part by the Natural Sciences and Engineering Research Council Grant OGP0004850.

[†] Research supported in part by National Science Foundation Grant DMS-9101740.

0047-259X/95 $12.00

where $a_i > 0$, $i = 1, ..., n$, and $f: \mathbb{R}_+ \to \mathbb{R}_+$ is such that (1.1) is a density function. In [15–18] we have treated various properties and applications of the Liouville distributions, in statistical and probabilistic settings. In this paper we introduce some broad generalizations of the densities (1.1), to be also referred to as Liouville distributions. These generalizations retain many characteristics of the Dirichlet and the Liouville distributions and subsume virtually all classical examples of those distributions.

Initially, we were motivated to consider these generalizations because of a basic observation about the proofs of many properties of the Dirichlet or Liouville distributions. Let us illustrate this observation with an example. In (1.1), choose $n = 3$ and $f(t) = c(1-t)^{a_4 - 1}$, $0 < t < 1$, where c is a constant. Then $(X_1, X_2, X_3) \sim D(a_1, a_2, a_3; a_4)$, a Dirichlet distribution with parameters a_1, a_2, a_3, a_4. Then it is well known, and not difficult to prove, that the marginal distribution of (X_1, X_2) is also Dirichlet, $(X_1, X_2) \sim D(a_1, a_2; a_3 + a_4)$. By writing out the joint density function of (X_1, X_2, X_3), integrating over X_3, and analyzing carefully the evaluation of the resulting integral, the reader will discover that at the heart of the proof lies the well-known identity

$$\frac{x^{a_3 - 1}}{\Gamma(a_3)} * \frac{x^{a_4 - 1}}{\Gamma(a_4)} = \frac{x^{a_3 + a_4 - 1}}{\Gamma(a_3 + a_4)}, \tag{1.2}$$

$x > 0$, where $*$ denotes convolution. Similar remarks apply when the problem is to prove that the joint distribution of $(X_1, X_2 + X_3)$ is $D(a_1, a_2 + a_3; a_4)$. Further, it can be checked that the derivation of many other properties of the Dirichlet distributions depend in a fundamental way on the identity (1.2); these properties include the calculation of moments and even the evaluation of the normalization constant for the Dirichlet densities.

This observation raises the possibility of replacing each monomial, $x_i^{a_i - 1}$, in (1.1) by $\phi_{a_i}(x_i)$, where the set of functions $\{\phi_a : a > 0\}$ satisfies the convolution property

$$\phi_{a_1} * \phi_{a_2} = \phi_{a_1 + a_2} \tag{1.3}$$

for all $a_1, a_2 > 0$. To avoid some technical difficulties, we shall choose the ϕ_a to be probability densities; in the case of the densities (1.1), this amounts to choosing $\phi_a(x) = x^{a-1} e^{-x} / \Gamma(a)$, $x > 0$, and rewriting (1.1) in the form

$$\tilde{f}\left(\sum_{i=1}^{n} x_i\right) \prod_{i=1}^{n} \frac{x_i^{a_i - 1} e^{-x_i}}{\Gamma(a_i)}, \tag{1.4}$$

where $\tilde{f}(t) = \left(\prod_{i=1}^{n} \Gamma(a_i)\right) e^t f(t)$.

We define the new class of multivariate Liouville distributions in exactly this manner, so that a Liouville density function will be one of the form

$$f\left(\sum_{i=1}^{n} x_i\right) \prod_{i=1}^{n} \phi_{a_i}(x_i),$$

where the densities ϕ_a satisfy (1.3). In fact, we will work with a set $\mathscr{F}$ of densities ϕ_a (or probability measures μ_a), where the indices a belong to an abstract Abelian semigroup I, instead of only the positive real line $\mathbb{R}_+$, and the elements of $\mathscr{F}$ satisfy the convolution property (1.3). This construction produces a large class of distributions because of the great lattitude we enjoy in the choice of the densities ϕ_a or measures μ_a, and the semigroup I. Moreover, we may as well attain maximum generality by allowing the random variables $X_1, ..., X_n$ to take values in a locally compact Abelian group or semigroup. It is remarkable that, despite this rather abstract setting, the resulting class of distributions retain many of the classical features of the Dirichlet distributions. As we can expect from the above discussion, marginal and conditional distributional properties are retained fully.

Further motivation for studying the new family of distributions is provided by the recent work of Barndorff-Nielsen and Jørgensen [1] (cf. Seshadri [28]). These authors introduce new classes of parametric models on the unit simplex by conditioning independent generalized inverse Gaussian random variables on their sum. We shall show that the models in [1] can be derived from ours by choosing the densities ϕ_a from a convolution family of generalized inverse Gaussian densities and then conditioning on their sum. In general, the process of conditioning a Liouville vector on the sum of its components leads to distributions which can be treated in depth using the general techniques presented here.

2. Preliminaries

Let G be a locally compact Abelian (LCA) group, with addition as its binary operation and its identity element denoted by 0. We assume throughout that G is equipped with a Hausdorff topology and has a countable basis of open sets. All background material on LCA groups needed here can be located in Berg and Forst [3] or Heyer [20].

In some instances, we will take G to be a LCA semigroup instead of a group. In these situations G will be a subset of a LCA group H and will be endowed with the subspace topology. An example of this situation is the case in which G is the space of real symmetric positive-definite $r \times r$ matrices and H is the group of all real symmetric $r \times r$ matrices.

With t denoting a generic element of G we denote by dt the Haar measure on G. We will assume that a particular normalization of the Haar measure has been specified. Then for $p \geqslant 1$, we denote by $L^p(G)$ the space of functions $g\colon G \to \mathbb{C}$ for which

$$\|g\|_p := \left(\int_G |g(t)|^p \, dt \right)^{1/p} < \infty.$$

We denote by $C_b(G)$ the set of continuous functions $g\colon G \to \mathbb{C}$ which vanish at infinity; endowed with the usual supremum norm, $C_b(G)$ is a Banach space. The dual of $C_b(G)$ is the Banach space $(M_b(G), \|\cdot\|)$ of all bounded Radon measures on G, where

$$\|\mu\| := \int_G |d\mu(t)|, \qquad \mu \in M_b(G).$$

For $\mu_1, \mu_2 \in M_b(G)$, the *convolution* $\mu_1 * \mu_2$ is defined by

$$\mu_1 * \mu_2(g) := \int_G \int_G g(x+y) \, d\mu_1(x) \, d\mu_2(y), \qquad g \in C_b(G). \tag{2.1}$$

Clearly the convolution operation is commutative and $\|\mu_1 * \mu_2\| \leqslant \|\mu_1\| \cdot \|\mu_2\|$. The convolution operator is also a bounded operator on $L^1(G)$: For $g \in L^1(G)$ and $\mu \in M_b(G)$,

$$g * \mu(x) := \int_G g(x-y) \, d\mu(y), \qquad x \in G, \tag{2.2}$$

converges absolutely almost everywhere on G, $g * \mu \in L^1(G)$, and $\|g * \mu\|_1 \leqslant \|g\|_1 \cdot \|\mu\|$. In the case in which μ is absolutely continuous with respect to the Haar measure, say, $d\mu(t) = g_2(t) \, dt$, then (2.2) reduces to the convolution of two functions,

$$g_1 * g_2(x) = \int_G g_1(x-t) \, g_2(t) \, dt, \tag{2.3}$$

$x \in G$. In particular, the integrability properties of $g_1 * g_2$ follow immediately from the remarks above, viz., if $g_1, g_2 \in L^1(G)$ then also $g_1 * g_2 \in L^1(G)$.

Let I be an abstract Abelian semigroup with addition as its binary operation; but not necessarily containing an identity element, 0.

Throughout, we let $\mathscr{F} = \{\mu_a : a \in I\}$ be a *convolution family* of probability measures on G. That is, the elements of $\mathscr{F}$ are probability measures indexed by I and $\mu_{a_1} * \mu_{a_2} = \mu_{a_1 + a_2}$ for all $a_1, a_2 \in I$. If a measure $\mu_a \in \mathscr{F}$ is absolutely continuous (with respect to the Haar measure dt) we denote the corresponding density function by ϕ_a. A broad class of examples arises from the situation wherein $\mathscr{F}$ is a *convolution semigroup* of probability measures or densities (cf. [2, p. 100; 3, p. 48; 12]). In these cases the index set is $I = \mathbb{R}_+$; $\mu_0 = \varepsilon_0$, the Dirac measure at the identity element, 0, in G; and the map $a \mapsto \mu_a$ is weakly continuous. This type of situation arises in (1.4), where ϕ_a is the density of the gamma distribution with shape parameter a.

Our first result is a generalization of the famous Liouville–Dirichlet integral (cf. [7, p. 160; 9; 15; 24; 25]) to the setting of LCA groups.

2.1. THEOREM. *Suppose that $\mu_{a_i} \in \mathscr{F}$, $i = 1, ..., n$, and $g: G \to \mathbb{C}$ satisfies*

$$\int_G |g(t)| \, d\mu_a(t) < \infty, \tag{2.4}$$

where $a = a_1 + \cdots + a_n$. Then

$$\int_G \cdots \int_G g\left(\sum_{i=1}^n x_i\right) \prod_{i=1}^n d\mu_{a_i}(x_i) = \int_G g(t) \, d\mu_a(t). \tag{2.5}$$

Proof. Replacing x_i by t_i, $i = 2, ..., n$, and $x_1 + \cdots + x_n$ by t, then the left-hand side of (2.5) becomes

$$\int_G \cdots \int_G g(t) \, d\mu_{a_1}(t - t_2 - t_3 - \cdots - t_n) \prod_{i=2}^n d\mu_{a_i}(t_i). \tag{2.6}$$

From the definition of the convolution integral (2.3), it follows that the integral with respect to $t_2, ..., t_n$ in (2.6) is the $(n-1)$-fold convolution, $\mu_{a_1} * \mu_{a_2} * \cdots * \mu_{a_n}$. Then the conclusion follows from the semigroup property, $\mu_{a_1} * \mu_{a_2} = \mu_{a_1 + a_2}$, $a_1, a_2 \in I$. Finally, it follows from Fubini's theorem that the integral (2.5) exists under the stated condition (2.4). ▮

We will also need a generalization of the classical Weyl fractional derivative.

2.2. DEFINITION. For $a \in I$ and $g: G \to \mathbb{C}$, define

$$W^a g(x) := \int_G g(t + x) \, d\mu_a(t), \qquad x \in G, \tag{2.7}$$

as the Weyl fractional derivative of order a of g, whenever the integral exists.

Note that W^a is a convolution: $W^a g = \check{\mu}_a * g$, where $\check{\mu}_a(B) = \mu_a(-B)$ for any μ_a-measurable set $B \subseteq G$. Then it follows from earlier remarks on the integrability properties of the convolution operator that W^a is a bounded linear operator on $L^1(G)$. We also note that from the Liouville–Dirichlet integral (2.5) follows the semigroup property $W^{a_1} W^{a_2} = W^{a_1 + a_2}$, $a_1, a_2 \in I$; in fact, this property is equivalent to (2.5).

In the case in which $G = \mathbb{R}_+$, $I = \mathbb{R}_+$, and μ_a is absolutely continuous with density function $\phi_a(x) = x^{a-1} e^{-x} / \Gamma(a)$, $x > 0$, $a > 0$, the operator W^a is essentially the Weyl fractional derivative; hence it is injective. The same holds when μ_a corresponds to the Gaussian distribution on $\mathbb{R}$ with mean 0 and variance $a > 0$, in which case W^a is the well-known Gauss–Weierstrass transform. To study the injectivity of W^a in the setting of LCA groups we need a few more preliminaries.

Let $\hat{G}$ denote the *dual* (or *character*) group of G, consisting of all continuous homomorphisms γ of G into the torus $\mathbb{T} = \{\exp(\sqrt{-1}\,\theta): 0 \leqslant \theta < 2\pi\}$. Recall that, by Pontryagin duality, G may be regarded as the dual group of $\hat{G}$; so it is customary to use the notation (t, γ), instead of $\gamma(t)$, for $t \in G$, $\gamma \in \hat{G}$. For $\gamma \in \hat{G}$ and a probability measure μ on G, the *Fourier transform* of μ is

$$\hat{\mu}(\gamma) = \int_G \overline{(t, \gamma)}\, d\mu(t).$$

It is well known that the Fourier transform, $\mu \to \hat{\mu}$, is injective and that $\|\hat{\mu}\|_\infty \leqslant 1$.

We now provide a sufficient condition on μ_a for the operator W^a to be injective. This condition is satisfied, e.g., by any infinitely divisible distribution on $\mathbb{R}^n$ and by the Wishart distribution.

2.3. PROPOSITION. *Suppose that $\mu_a \in \mathscr{F}$ satisfies $\hat{\mu}_a(\gamma) \neq 0$ for all $\gamma \in \hat{G}$. Then W^a is injective on $L^1(G)$.*

Proof. For $g \in L^1(G)$, $\gamma \in \hat{G}$, and $a \in I$, it follows by applying the Fourier transform to the identity $W^a g = \check{\mu}_a * g$ that $\widehat{W^a g}(\gamma) = \widehat{\check{\mu}_a}(\gamma)\, \hat{g}(\gamma)$. Since $\widehat{\check{\mu}_a}(\gamma) = \hat{\mu}_a(-\gamma) \neq 0$ for all $\gamma \in \hat{G}$ then $\widehat{W^a g}(\gamma) = 0$ implies that $\hat{g}(\gamma) = 0$ or $g = 0$, a.e. Therefore W^a is injective. ∎

3. GENERAL PROPERTIES OF THE LIOUVILLE DISTRIBUTIONS

3.1. DEFINITION. Let $X_1, \ldots, X_n$ be random variables taking values in G, and $\mathscr{F} = \{\mu_a : a \in I\}$ be a convolution family of probability measures on G.

Then $(X_1, ..., X_n)$ is said to have a Liouville distribution if its joint probability measure is of the form

$$f\left(\sum_{i=1}^{n} x_i\right) \prod_{i=1}^{n} d\mu_{a_i}(x_i), \tag{3.1}$$

$x_1, ..., x_n \in G$, where $f: G \to \mathbb{R}_+$, and $\mu_{a_i} \in \mathscr{F}$, $i = 1, ..., n$.

When (3.1) holds, we write $(X_1, ..., X_n) \sim L_n[f; a_1, ..., a_n]$. Using the Liouville–Dirichlet integral (2.5), the following properties of the Liouville distributions are obtained immediately. These results generalize some well-known properties [15, Sections 4, 8] of the classical Liouville distributions, and the proofs are standard.

3.2. PROPOSITION. *Suppose that* $(X_1, ..., X_n) \sim L_n[f; a_1, ..., a_n]$, $1 \leqslant r < n$, *and* $a = a_{r+1} + \cdots + a_n$. *Then*

(i) $(X_1, ..., X_r) \sim L_r[f_r; a_1, ..., a_r]$, *where* $f_r(t) = W^a f(t)$, $t \in G$.

(ii) *The conditional distribution of* $(X_{r+1}, ..., X_n)$ *given* $\{X_1 = x_1, ..., X_r = x_r\}$ *is* $L_{n-r}[g_r; a_{r+1}, ..., a_n]$, *where*

$$g_r(t) = f\left(t + \sum_{i=1}^{r} x_i\right) \Big/ f_r\left(\sum_{i=1}^{r} x_i\right). \tag{3.2}$$

In particular, conditioning on $\{X_1 = x_1, ..., X_r = x_r\}$ *is equivalent to conditioning on* $\{\sum_{i=1}^{r} X_i = \sum_{i=1}^{r} x_i\}$.

(iii) *For* $h: G \to \mathbb{R}$,

$$E\left(h\left(\sum_{i=r+1}^{n} X_i\right) \Bigg| \sum_{i=1}^{r} X_i = t\right) f_r(t) = \int_G \phi_a(y)\, h(y)\, f(t+y)\, dy, \tag{3.3}$$

whenever the expectation exists.

There are many other properties that follow this result. As an example, and for future reference, suppose that the measures μ_{a_i} are all absolutely continuous with density ϕ_{a_i}, respectively, $i = 1, ..., n$. Then it follows from Proposition 3.2(i), (ii), that for $1 \leqslant r \leqslant n-1$, the conditional density function of $(X_1, ..., X_r)$ given $X_1 + \cdots + X_n = t$ is

$$\frac{\prod_{i=1}^{r} \phi_{a_i}(x_i)}{\phi_{a_1 + \cdots + a_n}(t)} \phi_{a_{r+1} + \cdots + a_n}(t - x_1 - x_2 - \cdots - x_r). \tag{3.4}$$

We now obtain some characterizations of the Liouville distributions. The proof of the following result is similar to the proof of [15, Proposition 9.1].

3.3. PROPOSITION. *Let* $(X_1, ..., X_n) \sim L_n[f; a_1, ..., a_n]$. *Then the following are equivalent*:

(i) *The function* f *is multiplicative*; *i.e.*, $f(x+y) = f(x)\, f(y)$, $x, y \in G$.

(ii) $X_1, ..., X_n$ *are mutually independent.*

(iii) *There exists* i, j *with* X_i *and* X_j *mutually independent.*

In the following result, we assume that the Liouville distributions are *identifiable*; that is, if the distributions $L_n[f_1; a_1, ..., a_n]$ and $L_n[f_2; a'_1, ..., a'_n]$ represent the same probability measure then $f_1 \equiv f_2$ and $a_i = a'_i$ for all $i = 1, ..., n$.

3.4. PROPOSITION. *Let* $(X_1, ..., X_n) \sim L_n[f; a_1, ..., a_n]$ *and suppose that* $(X_1, ..., X_r) \sim L_r[f; a_1, ..., a_r]$ *for some* $1 \leqslant r < n$. *Then for each* $x \in G$,

$$f(x+t) = f(x) \tag{3.5}$$

for almost all (*with respect to* μ_a) $t \in G$. *If* f *is continuous then, for each* $x \in G$, (3.5) *holds for each* $t \in \operatorname{supp}[\mu_a]$, *the support of* μ_a.

Proof. Since $(X_1, ..., X_n) \sim L_n[f; a_1, ..., a_n]$ then, by Proposition 3.2(i), $(X_1, ..., X_r) \sim L_r[W^a f; a_1, ..., a_r]$, where $a = a_{r+1} + \cdots + a_n$. Since we are also given that $(X_1, ..., X_r) \sim L_r[f; a_1, ..., a_r]$ then, by identifiability, $f(t) = W^a f(t)$ for all $t \in G$; that is,

$$f(t) = \int_G f(x+t)\, d\mu_a(t), \qquad x \in G. \tag{3.6}$$

The integral equation (3.6) has been studied extensively [27] and its solution is known as the Choquet–Deny theorem. Applying the theorem of [29] we obtain (3.5). ∎

Note that the solution (3.5) of (3.6) requires only the minimal assumptions that G is an Abelian semigroup and $f: G \to \mathbb{R}_+$ is a bounded Borel measurable function. If additional assumptions are made on G then an explicit integral representation may be obtained for f in terms of the multiplicative functions on G (cf. [22; 27]); we stress that these additional assumptions are satisfied for all the semigroups appearing in the examples of Section 4.

4. LIOUVILLE DISTRIBUTIONS ON LCA GROUPS AND SEMIGROUPS

In this section, we consider various examples of semigroups I and $\mathscr{F}$ and groups or semigroups G, and we describe the related Liouville distributions.

4.1. *Poisson Distributions*

Let G be a LCA group. There are several definitions of a Poisson measure on G, resulting in numerous formulations of the corresponding Liouville distributions [20, Chap. III]. For simplicity, we will work only with the simplest notion of a Poisson measure on G.

Let x_0 be a fixed element of G and $a>0$. For any $x \in G$, let δ_x denote the measure assigning unit mass to x. The *Poisson measure associated with* x_0 *and* a is the measure

$$\mu_a = e^{-a}\left(\delta_0 + \sum_{j=1}^{\infty} \frac{\delta_{x_0}^{*j} a^j}{j!}\right). \tag{4.1.1}$$

Then it follows from (4.1.1) that

$$\hat{\mu}_a(\gamma) = \exp(a(e^{i\gamma(x_0)} - 1)) \tag{4.1.2}$$

for all $\gamma \in \hat{G}$. Each measure μ_a, $a>0$, is an example of a *Poisson measure* on G. It also follows from (4.1.2) that the family $\{\mu_a : a>0\}$ is a convolution semigroup of probability measures on G, called a *Poisson semigroup*.

Note that the Poisson measure μ_a is concentrated on elements of G of the form $0, x_0, 2x_0, 3x_0, ...$, when x_0 is of infinite order, and then $\mu_a\{jx_0\} = e^{-a}a^j/j!$ for $j = 0, 1, 2, 3, ...$ If x_0 is of finite order p, then μ_a is concentrated on elements of G of the form $0, x_0, 2x_0, ..., (p-1)x_0$, and then

$$\mu_a\{jx_0\} = e^{-a} \sum_{l \equiv j \,(\mathrm{mod}\, p)} \frac{a^l}{l!}, \qquad j = 0, 1, ..., p-1.$$

Thus, for Poisson measures μ_{a_j}, $j = 1, ..., n$, the measure

$$f(x_1 + \cdots + x_n) \prod_{j=1}^{n} d\mu_{a_j}(x_j) \tag{4.1.3}$$

is an example of a Liouville distribution on $G \times \cdots \times G$.

Many classical discrete distributions given in [21] are special cases of the Liouville distributions (4.1.3). As examples, if we choose μ_a as the standard Poisson distribution on $\mathbb{Z}_+$ with mean parameter a and we denote $a_1 + \cdots + a_n$ by α, then (4.1.3) is (i) the joint distribution of a set of multinomial random variables in the case in which $\alpha = 1$ and $f(t) = t!\, e$, $t \in \mathbb{Z}_+$; (ii) the multivariate negative binomial distribution when $f(t) = e^{\alpha}(1+\alpha)^{-N-t}\,\Gamma(N+t)/\Gamma(N)$, $t \in \mathbb{Z}_+$, where $N>0$; (iii) the multivariate logarithmic series distribution in the case in which $f(t) = e^{\alpha}(1+\alpha)^{-t}\,\Gamma(t)/\log(1+\alpha)$, $t \in \mathbb{N}$.

Another approach to defining Liouville distributions using Poisson-type distributions is as follows. Let ψ be a fixed real number. For $a > 0$, let μ_a denote the generalized Poisson distribution with parameter (a, ψ). The distribution μ_a is supported on $\mathbb{Z}_+$, the semigroup of nonnegative integers and has probability function

$$\mu_a\{x\} = \begin{cases} a(a+\psi x)^{x-1} e^{-a-\psi x}/x! & \text{for } x = 0, 1, 2, \ldots, \\ 0 & \text{for } x > m \text{ when } \psi < 0, \end{cases} \tag{4.1.4}$$

and $\mu_a\{x\} = 0$ otherwise, where $\max(-1, -a/m) < \psi \leqslant 1$ and m ($\geqslant 4$) is the largest positive integer such that $a + m\psi > 0$ in the case in which $\psi < 0$. By [6, Theorem 1.4.1, p. 15], the convolution property $\mu_{a_1} * \mu_{a_2} = \mu_{a_1+a_2}$ holds for all $a_1, a_2 > 0$. Then, proceeding as in (4.1.3), we obtain a Liouville distribution on $\mathbb{Z}_+^n$.

4.2. *Generalized Convolution of Mixtures of Exponential Distributions*

Let $\mathbf{Q}$ denote the family of all nonnegative measures Q on $\mathbb{R}_+$ such that

$$\int_0^1 t^{-1}\, dQ(t) < \infty, \qquad \int_1^\infty t^{-2}\, dQ(t) < \infty. \tag{4.2.1}$$

Let $I = [0, \infty) \times \mathbf{Q}$; equipped with the binary operation, $(a_1, Q_1) + (a_2, Q_2) = (a_1 + a_2, Q_1 + Q_2)$, I becomes an Abelian semigroup.

For $(a, Q) \in I$, a probability measure $\mu_{a,Q}$ on $\mathbb{R}_+$ is a *generalized convolution of mixtures of exponential distributions* (gcmed) if the moment generating function of $\mu_{a,Q}$ exists and is given by

$$\int_{\mathbb{R}_+} e^{\lambda x}\, d\mu_{a,Q}(x) = \exp\left[\lambda a + \int_{\mathbb{R}_+} \left(\frac{1}{t-\lambda} - \frac{1}{t}\right) dQ(t)\right] \tag{4.2.2}$$

for $\lambda \leqslant 0$. A general description of the class of gcmed's is provided by Bondesson [4]. In particular, we note that the class of gcmed's is closed under convolutions, a result which follows immediately from (4.2.1) and (4.2.2); and each $\mu_{a,Q}$ is infinitely divisible. Moreover, for the gcmed distribution $\mu_{a,Q}$ to be a mixture of exponential distributions it is necessary and sufficient that $a = 0$; and that the measure $dQ(t)$ be absolutely continuous with respect to Lebesgue measure, with a density function $q(t)$ satisfying $q(t) \leqslant 1$.

Then, for gcmed distributions μ_{a_j, Q_j}, $j = 1, \ldots, n$, the probability measure

$$f(x_1 + \cdots + x_n) \prod_{j=1}^n d\mu_{a_j, Q_j}(x_j) \tag{4.2.3}$$

is an example of a Liouville distribution on $\mathbb{R}_+^n$.

An important class of examples of gcmed's are the generalized inverse Gaussian distributions, denoted $N^{-}(\beta, \chi, \psi)$, with probability densities of the form

$$p(x; \beta, \chi, \psi) = \frac{(\psi/\chi)^{\beta/2}}{2K_\beta(\sqrt{\chi\psi})} x^{\beta-1} e^{-(\psi x + \chi x^{-1})/2} \tag{4.2.4}$$

for $x > 0$, where K_β is a modified Bessel function of order β, $\psi > 0$, $\chi > 0$, and $\beta \in \mathbb{R}$. (To realize the distribution $N^{-}(\beta, \chi, \psi)$ as a gcmed requires the result [4, Remark 3.2] that a gcmed uniquely determines its "parameter" (a, Q). It follows from [4, Eq. (3.4)] that $a = 0$, but more work is required to determine explicitly the measure Q [4, p. 53].)

Choosing the gcmed's in (4.2.3) from the class of generalized inverse Gaussian distributions (4.2.4), we obtain a Liouville distribution on $\mathbb{R}_{+}^{n}$. Next, we may construct a probability distribution on the unit simplex in $\mathbb{R}^n$, $\mathscr{S} = \{(y_1, ..., y_n) : y_i > 0,\ i = 1, ..., n;\ y_1 + \cdots + y_n = 1\}$, by conditioning the distribution of $(X_1, ..., X_n)$ on the sum $X_1 + \cdots + X_n$, as in (3.4); that is,

$$(Y_1, ..., Y_n) = (X_1, ..., X_n) \mid X_1 + \cdots + X_n = 1. \tag{4.2.5}$$

In [15], the marginal distribution of $(Y_1, ..., Y_{n-1})$ would be called a Liouville distribution *of type II*; the resulting distributions are themselves Liouville distributions as defined within our abstract framework, and the general theory applies. In particular, by Proposition 3.2, the marginal distributions of subvectors of $(Y_1, ..., Y_{n-1})$ belong to the same class.

When the X_i in (4.2.5) are such that $X_i \sim N^{-}(\beta_i, \chi_i, \psi)$, that is, the X_i have a common ψ parameter, then the resulting distributions on the simplex $\mathscr{S}$ coincide with the distributions studied in [1, 28].

4.3. *Gauss Distributions*

Let $\mathscr{M}^1(G)$ denote the family of probability measures on the LCA group G. Let $\mu \in \mathscr{M}^1(G)$ be a *Gauss measure* (*in the sense of Parthasarathy*) [20, p. 349 ff]. Further, let $\mathbf{Q}_{+}(\hat{G})$ denote the set of all positive *quadratic forms* on $\hat{G}$. It is well known [20, Theorem 5.2.7, p. 353] that μ is a Gauss measure if and only if there exists a unique element $(t, Q) \in G \times \mathbf{Q}_{+}(\hat{G})$ such that $\hat{\mu}$, the Fourier transform of μ, is of the form

$$\hat{\mu}(\gamma) = (t, \gamma)\, e^{-Q(\gamma)}, \qquad \gamma \in \hat{G}. \tag{4.3.1}$$

From (4.3.1) it follows readily that the set of Gauss measures on G is closed under convolutions. If we write $\mu \sim \mathscr{G}(t, Q)$ whenever (4.3.1) holds, then if μ_1 and μ_2 are independent (in the sense that the underlying random variables are independent) and $\mu_j \sim \mathscr{G}(t_j, Q_j)$ for $j = 1, 2$, we have

$\mu_1 * \mu_2 \sim \mathscr{G}(t_1 + t_2, Q_1 + Q_2)$. Now let $I = G \times \mathbf{Q}_+(\hat{G})$, which is an Abelian semigroup under addition. Then, for independent Gauss measures $\mu_j \sim \mathscr{G}(t_j, Q_j)$, $j = 1, ..., n$, the probability measure

$$f(x_1 + \cdots + x_n) \prod_{j=1}^{n} d\mu_j(x_j) \tag{4.3.2}$$

is an example of a Liouville distribution on $G \times \cdots \times G$.

In the classical case, where $G = \hat{G} = \mathbb{R}^r$, the parameters t and Q play the role of the mean vector and covariance matrix, respectively, and $\mathbf{Q}_+(\hat{G})$ may be identified with the space of real symmetric positive-semidefinite $r \times r$ matrices. Thus, (4.3.2) is absolutely continuous if the "matrices" Q_j are all positive-definite, and then the corresponding density function can be written in terms of the multivariate normal densities.

Other examples may be constructed starting with measures on G which are *Gaussian in the sense of Bernstein* [20, p. 362]; or we can even restrict G to be the torus $\mathbb{T}$ and choose the Gauss measures μ_j as the *wrapped normal distributions* [20, p. 350].

4.4. *Liouville Distributions on the Space of Positive-Definite Symmetric Matrices*

Let S_r be the space of all $r \times r$ real symmetric matrices, and Ω be the cone of $r \times r$ real symmetric positive-definite matrices. For $x \in \Omega$ and $j = 1, ..., r$, let $\Delta_j(x)$ denote the jth principal minor of x. Let $\rho = (r+1)/2$, and set $d_* x = (\det x)^{-\rho}\, dx$, where dx denotes the Haar measure on the LCA group S_r.

For $\alpha = (\alpha_1, ..., \alpha_r) \in \mathbb{C}^r$, define the *generalized power function* $\Delta_\alpha : \Omega \to \mathbb{C}$ by

$$\Delta_\alpha(x) := (\det x)^{\alpha_r} \prod_{j=1}^{r-1} \Delta_j(x)^{\alpha_j - \alpha_{j+1}}. \tag{4.4.1}$$

The *gamma function for the cone* Ω is defined [11, 13] by

$$\Gamma_\Omega(\alpha) := \int_\Omega e^{-\operatorname{tr} x} \Delta_\alpha(x)\, d_* x, \tag{4.4.2}$$

whenever the integral converges absolutely. By [11, VII.1.1; 13, Theorem 2.1, p. 22], the integral (4.4.2) converges absolutely if and only if α satisfies

$$\operatorname{Re} \alpha_j > (j-1)/2, \qquad j = 1, ..., r. \tag{4.4.3}$$

Moreover, in this range, Γ_Ω is evaluated in terms of the classical gamma function as

$$\Gamma_\Omega(\alpha) = (2\pi)^{r(r-1)/2} \prod_{j=1}^{r} \Gamma(\alpha_j - \tfrac{1}{2}(j-1)). \tag{4.4.4}$$

Define the convolution of two functions g_1, g_2 on Ω by

$$g_1 * g_2(y) := \int_{(0,y)} g_1(x)\, g_2(y-x)\, dx, \tag{4.4.5}$$

where the "interval" $(0, y)$ consists of all $x \in \Omega$ such that $y - x \in \Omega$. Then it follows from the results of Gindikin [13, Proposition 2.4, p. 24] that, for $a \in \mathbb{R}^r$ satisfying (4.4.3), the set of density functions

$$\phi_a(x) = \frac{\Delta_a(x)\, e^{-\operatorname{tr} x}}{\Gamma_\Omega(a)} (\det x)^{-\rho}, \tag{4.4.6}$$

where $x \in \Omega$, is a convolution family with respect to (4.4.5). Note that the densities ϕ_a are generalizations of the Wishart distributions.

Thus, we obtain Liouville random variables $(X_1, ..., X_n)$ with joint density function

$$f(x_1 + \cdots + x_n) \prod_{j=1}^{n} \phi_{a_j}(x_j), \tag{4.4.7}$$

where $x_1, ..., x_n \in \Omega$ and the vectors $a_1, ..., a_n \in \mathbb{R}^r$ satisfy (4.4.3).

The following special cases of (4.4.7) have appeared in the literature:

(i) Choose each a_j as a vector of the form $a_j = b_j(1, ..., 1)$, where $b_j \in \mathbb{R}$, $j = 1, ..., n$; and

$$f(x) = \begin{cases} \det(\varepsilon - x)^{b_{n+1} - \rho}, & x \in (0, \varepsilon), \\ 0, & \text{otherwise}, \end{cases}$$

where $b_{n+1} > (r-1)/2$ and ε is the $r \times r$ identity matrix. In this case, $(X_1, ..., X_n)$ follows a matrix Dirichlet distribution [15, 24, 25]. As a further special case, when $n = 1$ we recover the multivariate beta distribution.

(ii) Choose f to be either of the functions

$$f_1(x) = \begin{cases} \phi_\alpha(x), & x \in (0, \varepsilon), \\ 0, & \text{otherwise}, \end{cases}$$

or $f_2(x) = \phi_\alpha(\varepsilon + x)$, $x \in \Omega$, for some $\alpha \in \mathbb{R}^r$ satisfying (4.4.3). These distributions have been utilized in Bayesian statistical inference problems by Guttman and Tan [19].

(iii) Let $f(x) = e^{-\operatorname{tr} x}$, $x \in \Omega$, and each a_j be of the form listed in (i). In this case $X_1, ..., X_n$ are mutually independent, and their marginal distributions are Wishart distributions. In these examples, the parameters b_j are the "degrees of freedom" of the Wishart matrices. Note that the density function of, say, X_1, exists if and only if $b_1 > (r-1)/2$, and then the characteristic function of X_1 is

$$E(e^{i \operatorname{tr} w X_1}) = \det(\varepsilon - iw)^{-b_1}, \qquad w \in S_r. \tag{4.4.8}$$

For $b_1 \in \mathbb{R}$, it is a theorem of Gindikin (cf. [11, 13, 26]) that (4.4.8) is a characteristic function if and only if b_1 belongs to the *Wallach set* $W_\Omega = \{0, \frac{1}{2}, 1, \frac{3}{2}, ..., (r-1)/2\} \cup ((r-1)/2, \infty)$. Proceeding as before, we obtain examples of (non-absolutely continuous) Liouville random variables $(X_1, ..., X_n)$, where $\mathscr{F}$ is the set of all Wishart measures indexed by the elements of $I = W_\Omega$.

In closing this section, we use the hypergeometric functions of matrix argument (cf. Muirhead [23]) to construct new examples of Liouville distributions on Ω.

(iv) Let ${}_1F_1$ denote the confluent hypergeometric function on Ω. Let $I = \{(\alpha, \beta) \in \mathbb{R}^2 : \alpha, \beta > (r-1)/2\}$, so that I is an Abelian semigroup under addition. Further, let $\mathscr{F}$ denote the class of probability density functions $\phi_{\alpha,\beta}: \Omega \to \mathbb{R}_+$, where for any $(\alpha, \beta) \in I$,

$$\phi_{\alpha,\beta}(x) = \frac{(\det x)^{\alpha+\beta-\rho}\, e^{-2 \operatorname{tr} x}}{2^{-r\beta}\Gamma_r(\alpha+\beta)}\; {}_1F_1(\alpha; \alpha+\beta; x), \qquad x \in \Omega, \tag{4.4.9}$$

where Γ_r is the multivariate gamma function [23, p. 61]. The proof that $\mathscr{F}$ is a convolution family with respect to (4.4.5) is similar to the classical case [8, p. 271]. Indeed, using results from [10; 14], it may be shown that the Laplace transform of $\phi_{\alpha,\beta}$ is

$$\int_\Omega e^{-\operatorname{tr} xw}\phi_{\alpha,\beta}(x)\, dx = \det(\tfrac{1}{2}w + \varepsilon)^{-\beta} \det(w + \varepsilon)^{-\alpha}, \qquad w \in \Omega. \tag{4.4.10}$$

Then the conclusion follows immediately from the convolution formula for Laplace transforms. Moreover, since $\phi_{\alpha,\beta}$ is a nonnegative function and the right-hand side of (4.4.10) converges to 1 as $w \to 0$ in the cone Ω then it follows that $\phi_{\alpha,\beta}$ is a probability density.

Finally, the corresponding Liouville distributions are as in (4.4.7).

5. Concluding Remarks

There are many other examples of Liouville distributions which we have not covered here. For example, as noted earlier in Section 2, we may choose the class $\mathscr{F}$ to be any convolution semigroup on the space G; a classical example would be the semigroup of stable distributions on $\mathbb{R}^n$ (cf. [3, p. 48] for a table of convolution semigroups on $\mathbb{R}$). Other examples arise when we choose G to be a symmetric cone [5]; the resulting examples would generalize those on the cone of positive-definite symmetric matrices.

Proceeding as in Section 4.1, we can construct additional classes of discrete Liouville distributions. For example, let us choose the measure μ_{a_i}, $a_i \in \mathbb{Z}_+$, to be the binomial distribution $b(a_i, p)$, so that the class $\mathscr{F}$ consists of binomial distributions with common probability of success p. Further, choose $f(t) = p^{-t}(1-p)^{t-\alpha}/\binom{\alpha}{t}$, $t \in \mathbb{Z}_+$, where $\alpha = a_1 + \cdots + a_n$. Then the Liouville distribution (3.1) reduces to a multivariate hypergeometric distribution with parameters $a_1, ..., a_n$.

In classical contexts, one method used extensively to construct multivariate distributions is the method of compounding (or mixing) [21, Chap. 11.8]: Consider a random vector $(Y_1, ..., Y_n)$ with a given distribution, and a random vector $(Z_1, ..., Z_n)$ such that the conditional distribution of $(Z_1, ..., Z_n)$ given $(Y_1, ..., Y_n)$ is specified, then the method of compounding amounts to calculating the (unconditional) distribution of $(Z_1, ..., Z_n)$. Many of the examples in [21] deal with the situation where both the marginal distribution of $(Y_1, ..., Y_n)$ and the conditional distribution of $(Z_1, ..., Z_n)$ given $(Y_1, ..., Y_n)$ are special cases of Liouville distributions (or derived from Liouville distributions through the conditioning process (3.4)); in particular, the two examples treated in [21, Chap. 11.8] are of this form. Applied to the newly defined classes of Liouville distributions, the method of compounding produces many new families of probability distributions. As an example, suppose that the conditional distribution of $(Z_1, ..., Z_n)$ given $(Y_1, ..., Y_n)$ is $L_n[f; Y_1, ..., Y_n]$, where each μ_{Y_i} is a Poisson distribution with mean Y_i, and the distribution of $(Y_1, ..., Y_n) \sim L_n[g; a_1, ..., a_n]$, where the corresponding μ_{a_i} are gamma distributions with shape parameter a_i. Using the fractional calculus techniques to reduce the multiple integrals to a single integral, we find that the unconditional distribution of $(Z_1, ..., Z_n)$ is

$$\tilde{f}(z_1 + \cdots + z_n) \prod_{i=1}^{n} \frac{\Gamma(a_i + z_i)}{z_i! \, \Gamma(a_i)},$$

where

$$\tilde{f}(t) = \frac{f(t)}{\Gamma(t + a_1 + \cdots + a_n)} \int_0^\infty e^{-2y} y^{t + a_1 + \cdots + a_n - 1} g(y) \, dy.$$

By appropriate specialization of the functions f and g, we can recover many of the formulas in [21, Chap. 11].

ACKNOWLEDGMENT

We thank Kenneth Gross for introducing us to the subject of harmonic analysis on symmetric cones.

REFERENCES

[1] BARNDORF-NIELSEN, O. E., AND JØRGENSEN, B. (1991). Some parametric models on the simplex. *J. Multivariate Anal.* **39** 106–116.
[2] BERG, C., CHRISTENSEN, J. P. R., AND RESSEL, P. (1984). *Harmonic Analysis on Semigroups: Theory of Positive-Definite and Related Functions.* Springer-Verlag, New York.
[3] BERG, C., AND FORST, G. (1975). *Potential Theory on Locally Compact Abelian Groups.* Springer-Verlag, New York.
[4] BONDESSON, L. (1981). Classes of infinitely divisible distributions and densities. *Z. Wahrsch. Verw. Geb.* **57** 39–71.
[5] CASALIS, M. (1991). Les familles exponentielles à variance quadratique homogenie sont des lois de Wishart sur un cone symmétrique. *C.R. Acad. Sci. Paris Ser. I* **312** 537–540.
[6] CONSUL, P. C. (1989). *Generalized Poisson Distributions: Properties and Applications.* Dekker, New York.
[7] EDWARDS, J. (1922). *A Treatise on the Integral Calculus,* Vol. 2. Macmillan, New York.
[8] ERDÉLYI, A. (1953). *Higher Transcendental Functions,* Vol. 1. McGraw–Hill, New York.
[9] FANG, K.-T., KOTZ, S., AND NG, K. W. (1990). *Symmetric Multivariate and Related Distributions.* Chapman & Hall, London.
[10] FARAUT, J., AND KORANYI, A. (1988). Fonctions hypergéométriques associées aux cônes symmétriques. *C.R. Acad. Sci. Paris Ser. I* **307** 555–558.
[11] FARAUT, J., AND KORANYI, A. (1995). *Analysis on Symmetric Cones.* Oxford Univ. Press, New York.
[12] FELLER, W. (1967). *An Introduction to Probability Theory and its Applications,* Vol. 2, 3rd ed. Wiley, New York.
[13] GINDIKIN, S. G. (1964). Analysis in homogeneous domains. *Russian Math. Surveys* **19** 1–90.
[14] GROSS, K. I., AND RICHARDS, D. ST. P. (1987). Special functions of matrix argument. I. Algebraic induction, zonal polynomials and hypergeometric functions, *Trans. Amer. Math. Soc.* **301** 781–811.
[15] GUPTA, R. D., AND RICHARDS, D. ST. P. (1987). Multivariate Liouville distributions. *J. Multivariate Anal.* **23** 233–256.
[16] GUPTA, R. D., AND RICHARDS, D. ST. P. (1990). The Dirichlet distributions and polynomial regression. *J. Multivariate Anal.* **32** 95–102.
[17] GUPTA, R. D., AND RICHARDS, D. ST. P. (1991). Multivariate Liouville distributions, II. *Probab. Math. Statist.* **12** 291–309.
[18] GUPTA, R. D., AND RICHARDS, D. ST. P. (1992). Multivariate Liouville distributions, III. *J. Multivariate Anal.* **43** 29–57.
[19] GUTTMAN, I., AND TAN, W. Y. (1973). The use of the disguised Wishart distribution in a Bayesian approach to tolerance region construction. *Ann. Inst. Statist. Math.* **25** 549–556.
[20] HEYER, H. (1977). *Probability Measures on Locally Compact Groups.* Springer-Verlag, New York.
[21] JOHNSON, N. L., AND KOTZ, S. (1969). *Discrete Distributions.* Wiley, New York.
[22] LAU, K.-S., AND ZENG, W.-B. (1990). The convolution equation of Choquet and Deny on semigroups. *Studia Math.* **97** 115–135.

[23] MUIRHEAD, R. J. (1982). *Aspects of Multivariate Statistical Theory*. Wiley, New York.

[24] OLKIN, I. (1959). A class of integral identities with matrix argument. *Duke Math. J.* **26** 207–214.

[25] OLKIN, I. (1979). Matrix extensions of Liouville–Dirichlet-type integrals. *Linear Algebra Appl.* **28** 155–160.

[26] PEDDADA, S. D., AND RICHARDS, D. ST. P. (1991). Proof of a conjecture of M. L. Eaton on the characteristic function of the Wishart distribution. *Ann. Probab.* **19** 868–874; **20** (1992), 1107.

[27] RAO, C. R., AND SHANBHAG, D. N. (1991). An elementary proof for an extended version of the Choquet–Deny theorem. *J. Multivariate Anal.* **38** 141–148.

[28] SESHADRI, V. (1992). General exponential models on the unit simplex and related multivariate inverse Gaussian distributions. *Statist. Probab. Lett.* **14** 385–391.

[29] SZÉKELY, G. J., AND ZENG, W. (1990). The Choquet–Deny convolution equation $\mu = \mu * \sigma$ for probability measures on Abelian semigroups. *J. Theoret. Probab.* **3** 361–365.