Chapter 4

Liouville distributions and their generalization

Joseph Liouville (1809-82) generalized the Dirichlet integral to the Liouville multiple integral over the positive orthant $\Re_+^n$ (e.g., see Edwards, 1922, p. 160; Whittaker and Watson, 1952; Fichtenholz, 1967, p. 391). Marshall and Olkin (1979) use this integral to define the so-called "Liouville-Dirichlet" distributions. Sivazlian (1981a) focused on deriving the Dirichlet and beta distributions from the generalized Liouville family. Sivazlian (1981b) introduced two kind of Liouville distributions and show that these new classes could be used to derive some well-known statistical distributions. Anderson and Fang (1982, 1987) study a subclass of Liouville distributions, and Gupta and Richards (1987, 1990, 1991, 1992, 1995) give a more comprehensive treatment of the multivariate Liouville distributions and also extend some results to their matrix analogues. Rayens and Srinivasan (1994) give some results on the generalized Liouville distributions on the unit simplex. Smith (1994), starting from generalized Liouville distributions on the positive orthant, introduces the so-called conditional generalized Liouville distributions and gives some of their properties. Gupta and Song (1996) study some properties of the generalized Liouville distribution, give a one-to-one correspondence between two kinds of generalized Liouville distributions, and derive the stochastic representations and marginal and conditional distributions of this family. All of the authors mentioned above use the so-called "Liouville integral" (Fang et al., 1990) or its extension (Sivazlian, 1981a) in their approach.

Fang et al. (1990) devote a chapter to the standard Liouville distributions, using an alternative approach that we follow in section 4.1. This approach makes dealing with these distributions much easier. Also, this approach is a helpful tool for investigating the covariance structures of Liouville random vectors. This tool is unlike the one used by Gupta and Richards (1987) for this purpose, which Rayens (1993) recently showed to be of limited use on a simplex sample space. In section 4.2, we will define the generalized Liouville distribution on the simplex (Rayens and Srinivasan, 1994). In section 4.3 we will use the approach of Fang et al. to introduce generalized Liouville distributions on the positive orthant as well as conditional generalized Liouville distributions. Our conclusions will be given in section 4.4. In all of these sections, we will adopt the notation used by Fang et al. (1990).

4.1 Liouville distributions

Liouville distributions can be viewed in some sense as extensions of the Dirichlet distribution to the positive orthant. In particular,

Definition 4.1 A random vector $\mathbf{x}$ in $\Re_{+}^{n}$ is said to have a ***Liouville distribution*** if $\mathbf{x} \overset{d}{=} \mathrm{r}.\mathbf{y}$, where $\mathbf{y} = (y_1, y_2, \ldots, y_n) \sim D_n(\alpha)$ on H_n and r is an independent r.v. with p.d.f. f ; in symbols, $\mathbf{x} \sim L_n(\alpha; f)$. We shall call $\mathbf{y}$ the *Dirichlet base,* α the *Dirichlet parameter,* r the *generating variate*, and f the *generating density*.

It should be noticed immediately that when r=1 with probability one, the Liouville distribution reduces to the Dirichlet distribution $D_n(\alpha)$ on H_n. Later in this dissertation we will consider an extension to the Liouville distribution by allowing for non-Dirichlet base (e.g., adaptive Dirichlet base).

Theorem 2.1 stated that the Dirichlet distribution can result from an independent basis in which each element is gamma distributed and all of these distributions have the same scale parameter. In fact, from definition 4.1, the Dirichlet distribution can also result from many bases that do not have independent elements. This is easily seen by noting that, since $\mathbf{y} \in H_n$, we have $y_1 + y_2 + ... + y_n = 1$. Also, it is clear from the definition above (as noted by Fang et al, 1990) that $\mathbf{x} \sim L_n(\alpha; f)$ if and only if $(x_1 / \sum x_i, ..., x_n / \sum x_i) \sim D_n(\alpha)$ on H_n and is independent of the total size $\sum x_i$ (equivalent to the r.v. in the definition of the Liouville distribution).

The following theorem gives the density function of a Liouville distribution with generating density function f; the proof can be found in Fang et al. (1990).

Theorem 4.1 *The density function of a Liouville distribution with generating density function* f *is given by*

$$g(x_1, ..., x_n) = \prod_{i=1}^{n} \frac{x_i^{\alpha_i - 1}}{\Gamma(\alpha_i)} \frac{\Gamma(\alpha^*)}{(\sum_{i=1}^{n} x_i)^{\alpha^* - 1}} f(\sum_{i=1}^{n} x_i), \text{ where } \alpha^* = \sum_{i=1}^{n} \alpha_i \tag{4.1}$$

This density is defined on the a-simplex $S_a^n = \{(x_1, ..., x_n) | 0 < \sum_{i=1}^{n} x_i < a\}$ if and only if f is defined on the interval (0,*a*). Fang et al. noted that if the generating density f is defined only on the interval (0,1), then equation (4.1) above provides an alternative to the logistic normal class of distributions for compositional data.

Note that equation (4.1) can also be written in the form

$$g(x_1, ..., x_n) = h(\sum_{i=1}^{n} x_i) \prod_{i=1}^{n} x_i^{\alpha_i - 1}, \tag{4.2}$$

where

$$h(t) = \frac{\Gamma(\alpha^*)}{\prod_{i=1}^n \Gamma(\alpha_i)} \frac{1}{t^{\alpha^*-1}} f(t), \text{ for all } t > 0 \tag{4.3}$$

Here, $h(.)$ is called the *density generator* of the Liouville distribution, to identify it from the generating density function f. Throughout this proposal, we adopt the notation used by Fang et al. (1990), who use $L_n(\alpha; f)$ for a Liouville distribution with generating density function f, and $L_n(h;\alpha)$ for a Liouville distribution with density generator h.

Remark 4.1

It is fairly easy to show that

i) If $\mathbf{x} \sim L_n(\alpha; f)$, then $f \sim Be(\alpha^*, b)$ if and only if $\mathbf{x} \sim D_n(\alpha_1, \ldots, \alpha_n; b)$ on S_1^n.

ii) If $\mathbf{x} \sim L_n(\alpha; f)$ and $\alpha_j - \alpha_{j+1} - \ldots - \alpha_{n-1} - \alpha_n > 0$ for all $j = 2, \ldots, n-1$, then

$f \sim Be(a,b)$ if and only if $(1 - \sum_{i=1}^n x_i, x_1, \ldots, x_{n-1}) \sim CM_n(\hat{\alpha}, \hat{\lambda})$ on S_1^n, where

$\hat{\alpha} = (b, \alpha_1, \ldots, \alpha_{n-1})$ and $\hat{\lambda} = (a, \alpha_2 - \alpha_3 - \ldots - \alpha_n, \alpha_3 - \alpha_4 - \ldots - \alpha_n, \ldots, \alpha_{n-1} - \alpha_n, \alpha_n)$.

iii) If $x \sim L_1(\alpha; f)$, then by equation (4.1), f is the density function of x.

Sivazlian (1981b) refers to Liouville distributions on the a-simplex S_a^n as *Liouville distributions of the first kind* when $a = \infty$, and *Liouville distributions of the second kind* when $a = 1$. However, there is room for confusion, since Fang et al. (1990) use the term *Liouville distributions of the second kind* for distributions on S_a^n for any $a < \infty$. Therefore, in this thesis, we will refer to Liouville distributions on S_a^n as *Liouville distributions on the positive orthant* when $a = \infty$, and *Liouville distributions on the unit simplex* when $a = 1$, and will write $L_n^{(\infty)}(\alpha; f)$ or $L_n^{(\infty)}(h;\alpha)$ for Liouville distributions on the positive orthant and

$L_n^{(1)}(\alpha; f)$ or $L_n^{(1)}(h;\alpha)$ for Liouville distributions on the unit simplex when there is a need to distinguish between them.

Gupta and Richards (1991) consider a Liouville distribution with density of the form

$$\mathrm{g}(\mathbf{x}) = A x_1^{\alpha_1 - 1} \cdots x_n^{\alpha_n - 1} f\left(\frac{x_1 + \cdots + x_n}{\theta}\right)$$

for $x_i > 0$. They discuss applications of this distribution to reliability theory, as well as estimation of both θ and α_i when $\alpha_1 = \cdots = \alpha_n$.

Liouville distributions on the unit simplex provide an infinite class of distributions for modeling compositional data. Fang et al. (1990) show that Liouville distributions have the following nice mathematical properties:

Theorem 4.2 *Let* $\mathbf{x} \sim L_n(\alpha; f)$, *such that* $\mathbf{x}$ *has zero probability at the origin and* f *is positive on (0,a), where* $0 < a < \infty$. *Then all the marginal distributions of* $\mathbf{x}$ *are Liou-ville distributions. In particular, let* $1 \le m \le n$, $\alpha^* = \sum_{i=1}^n \alpha_i$, $\alpha_1^* = \sum_{i=1}^m \alpha_i$, *and* $\alpha_2^* = \sum_{i=m+1}^n \alpha_i$. *Then the generating density function of* $(x_1, \cdots, x_m) \sim L_m(\alpha_1, \cdots, \alpha_m; f_m)$ *is*

$$f_m(t) = \frac{\Gamma(\alpha^*)}{\Gamma(\alpha_1^*)\Gamma(\alpha_2^*)} t^{\alpha_1^* - 1} \int_t^\infty \frac{(r - t)^{\alpha_2^* - 1}}{r^{\alpha^* - 1}} f(r) dr \tag{4.4}$$

for all $0 < t < a$.

Theorem 4.3 *Let* $\mathbf{x} \sim L_n(\alpha; f)$. *Then the mixed moments of* $\mathbf{x}$ *are given by*

$$E\left(\prod_{i=1}^n x_i^{m_i}\right) = \frac{\mu_m}{(\alpha^*)^{[m]}} \prod_{i=1}^n \alpha_i^{[m_i]} \tag{4.5}$$

where $m = \sum_{i=1}^{n} m_i$, $\alpha^* = \sum_{i=1}^{n} \alpha_i$, μ_m *is the mth raw moment of the generating variate r, and the superscript* $[\cdot]$ *denotes the ascending factorial; i.e.,* $\alpha^{[m]} = \alpha(\alpha+1) \cdots (\alpha+m-1)$.

In particular, we have

$$E(x_i) = \mu_i \frac{\alpha_i}{\alpha^*}, \qquad \mathrm{var}(x_i) = \frac{\alpha_i}{\alpha^*}\left(\mu_2 \frac{\alpha_i+1}{\alpha^*+1} - \mu_1^2 \frac{\alpha_i}{\alpha^*}\right), \tag{4.6}$$

and

$$\mathrm{cov}(x_i, x_j) = \frac{\alpha_i \alpha_j}{\alpha^*}\left(\frac{\mu_2}{\alpha^*+1} - \frac{\mu_1^2}{\alpha^*}\right), \; 1 \le i < j \tag{4.7}$$

Theorem 4.4 *Let* $\mathbf{x} \sim L_n(\alpha; f)$. *Then any amalgamation* $(w_1, \cdots, w_s)$ *of* $\mathbf{x}$ *is distributed* $L_s(\alpha_1^{'}, \cdots, \alpha_s^{'}; f)$, *where* $(\alpha_1^{'}, \cdots, \alpha_s^{'})$ *is the corresponding amalgamation of the Dirichlet parameters* $\boldsymbol{\alpha} = (\alpha_1, \alpha_2, \cdots, \alpha_n)$.

Remark 4.2

From equation (4.7), we have the following.

1) When r=1 with probability one, the Liouville distribution reduces to the Dirichlet distribution $D_n(\boldsymbol{\alpha})$ on H_n, and equation (4.7) reduces to the one corresponding to the Dirichlet distribution.

2) The sign of $\mathrm{cov}(x_i, x_j)$ for $1 \le i < j$ depends on the generating variate r through its first and second moments, but not on *i* and *j*. Thus, the upper triangular covariance matrix is either completely positive or completely negative unless $\mathbf{x} \in S^n$. This means that the covariance matrix of the Liouville distribution on the positive orthant has a very restrictive form. In chapter 6, we will extend the Liouville distribution in a way that relaxes somehow this restrictive form.

3) If $\text{cov}(x_i, x_j)$ <0 (>0) for some $i \neq j$, then $\text{cov}(x_r, x_s)$ <0 (>0) $\forall\, r \neq s$. Also, if $\mathbf{x} \in S^n$ and $\text{cov}(x_i, x_j)$ >0 for some $i \neq j$, then we must have $\text{cov}(x_k, x_{n+1})$ <0 $\forall\, k = 1,2,\cdots,n$, where $x_{n+1} = 1 - \sum_{i=1}^{n} x_i$. Thus, the covariance matrix of $x_1,\cdots,x_n,x_{n+1}$ must have one of the following two forms:

$$\begin{bmatrix} + & - & - & - & - & - & - & s_1 \\ - & + & - & - & - & - & - & s_2 \\ - & - & + & - & - & - & - & s_3 \\ - & - & - & + & - & - & - & s_4 \\ - & - & - & - & + & - & - & s_5 \\ - & - & - & - & - & + & - & s_6 \\ - & - & - & - & - & - & + & s_7 \\ s_1 & s_2 & s_3 & s_4 & s_5 & s_6 & s_7 & + \end{bmatrix} \quad OR \quad \begin{bmatrix} + & + & + & + & + & + & + & - \\ + & + & + & + & + & + & + & - \\ + & + & + & + & + & + & + & - \\ + & + & + & + & + & + & + & - \\ + & + & + & + & + & + & + & - \\ + & + & + & + & + & + & + & - \\ + & + & + & + & + & + & + & - \\ - & - & - & - & - & - & - & + \end{bmatrix},$$

where $s_i = sign[\text{cov}(x_i, x_{n+1})]$.

Thus, unlike Dirichlet distributions, Liouville-distributed random variables can have positive as well as negative covariance. For example, let $\mathrm{r} \sim Be(0.1,1)$ and α^* >1. Then it is easy to show that $\text{cov}(x_i, x_j)$ >0 for all $i \neq j$, $i, j \neq n+1$. Also, Rayens and Srinivasan (1994) considered the Liouville distribution on the simplex with n=2, $f(u) = \exp(\beta e^{-u})$, $\alpha_1 = 3$, $\alpha_2 = 7$, and values of β in the range from -10 to 100. They showed that near $\beta = 28$, the correlation coefficient of x_1 and x_2 is approximately zero, positive for $\beta > 28$, and negative for $\beta < 28$.

Smith (1994) noted that "important work is left in trying to compare different choices of f to find the 'best' Liouville distribution. One criteria [sic] that could be set forth is the ability to model correlation close to 1." In addition to that, however, I would note that the covariance matrix of the Liouville distribution has a very restrictive form regardless of the choice of the generating density f . This violates one of the criteria for successful simplex distributions set forth in chapter 1.

Because Liouville distributions on the simplex have the ability to model positive covariance, it should be clear that they are an improvement over the Dirichlet, (although not necessarily over the adaptive Dirichlet distributions). Liouville distributions on the simplex, also, do not satisfy complete right neutrality except in the special case where they reduce to the Dirichlet distribution (Fang et al., 1990). However, all Liouville distributions on the simplex possess complete n-subcompositional independence (in our terminology), as shown by Rayens and Srinivasan (1994). Thus, while Liouville distributions on the simplex have a less severe independence structure than the Dirichlet (which is one of the criteria proposed earlier in this thesis), they may not be flexible enough in this regard for some applications. In addition, there are some severe restrictions on the structure of the covariance matrix.

4.2 Generalized Liouville distributions on the simplex

Smith (1994) devotes a complete chapter to generalized Liouville distributions on the unit simplex, which he calls generalized Liouville distributions "of the second kind." In this section, some of the material in that chapter will be summarized. The interested reader should note that the domain of the generalized Liouville distribution of the second kind in Smith (1994) is the unit simplex, while the domain of the so-called generalized Liouville distribution of the second kind in Gupta and Song (1996) is not.

Definition 4.2 A random vector $\mathbf{x} = (x_1, \cdots, x_{n+1})$ is said to have a ***generalized Liouville distribution on the simplex*** $GL_n^{(1)}(h; \alpha_1, \cdots, \alpha_{n+1}, \beta_1, \cdots, \beta_{n+1}, q_1, \cdots, q_{n+1})$ if $\mathbf{x}$ has density function

$$g(\mathbf{x}) = A x_1^{\alpha_1 - 1} \cdots x_{n+1}^{\alpha_{n+1} - 1} h\Big((\frac{x_1}{q_1})^{\beta_1} + \cdots + (\frac{x_{n+1}}{q_{n+1}})^{\beta_{n+1}} \Big) \tag{4.8}$$

when $\mathbf{x} \in S^{n+1}$ and 0 otherwise, where $\alpha_i > 0$, $q_i > 0$, and $\beta_i > 0$ for all *i*, h is positive and continuous, and

$$A = \frac{1}{\iint \cdots \int\limits_{S^{n+1}} x_1^{\alpha_1 - 1} \cdots x_{n+1}^{\alpha_{n+1} - 1} h\Big((\frac{x_1}{q_1})^{\beta_1} + \cdots + (\frac{x_{n+1}}{q_{n+1}})^{\beta_{n+1}} \Big) dx_{n+1} \cdots dx_1}$$

Notes:

1. *A* must be non-zero for the density to exist.

2. The generalized Liouville distribution on the simplex reduces to the standard Liouville distribution when $\beta_i = 1$ and $q_i = 1$ for all *i*, and $\mathbf{x} \in \Re_+^n$.

Unfortunately, Smith was not able to provide many results about the covariance structure of the generalized Liouville distribution. Rayens and Zhang (1993) show that there do exist generalized Liouville distributions with positive covariance. Since the class of generalized Liouville distributions contains the Liouville distribution as a special case, they clearly must be at least as flexible as the Liouville in modeling positive covariance, but it is not clear whether they are sufficiently general to satisfy our criteria. In addition, this class of distributions is intractable, particularly with respect to its moments as well as its normalizing constant. However, Rayens and Zhang (1993) have developed software that enables one to calculate the normalizing constant *A*.

Smith concludes that generalized Liouville distributions on the simplex are an improvement over Liouville distributions on the simplex because they do have less severe independence structures. However, they cannot model some independence concepts such as right neutrality for a partition of order 1. In addition, they are not invariant with respect to the choice of which variable is the fill-up value.

4.3 Generalized Liouville distributions on the positive orthant and conditional generalized Liouville distributions

In the last section, we summarized some results in Smith (1994) regarding generalized Liouville distributions on the simplex. In this section, we summarize some results in Smith (1994) regarding generalized Liouville distributions on the positive orthant and conditional generalized Liouville distributions. However, these distributions will be introduced following the approach used in section 4.1.

Definition 4.3 A random vector $\mathbf{z} = (z_1, \cdots, z_{n+1}) \in \Re_+^{n+1}$ is said to have a ***generalized Liouville distribution***, $GL_{n+1}^{\infty}(\alpha_1, \cdots, \alpha_{n+1}, \beta_1, \cdots, \beta_{n+1}, q_1, \cdots, q_{n+1}; f)$, if $\tilde{\mathbf{z}} \overset{d}{=} \mathrm{r}.\mathbf{y}$, where $\tilde{\mathbf{z}}_i = (z_i / q_i)^{\beta_i}$, $\alpha_i > 0$, $\beta_i > 0$, $q_i > 0$ for $i = 1, \cdots, n+1$, $\mathbf{y} = (y_1, y_2, \ldots, y_{n+1}) \sim D_{n+1}(\alpha_1/\beta_1, \cdots, \alpha_{n+1}/\beta_{n+1})$ on H_{n+1} and r is an independent r.v. with p.d.f. f.

Analogous to the standard Liouville distribution, we will call $\mathbf{y}$ the *Dirichlet base*, $(\alpha_1/\beta_1, \cdots, \alpha_{n+1}/\beta_{n+1})$ the *Dirichlet parameter*, r the *generating variate*, and f the *generating density*.

It should be noticed immediately that $\tilde{\mathbf{z}}$ follows the standard Liouville distribution and when $\beta_1 = \cdots = \beta_{n+1} = q_1 = \cdots = q_{n+1} = 1$, the generalized Liouville distribution reduces to the standard Liouville distribution on $\Re_+^{n+1}$. Also, $\mathbf{z} \sim GL_{n+1}^{\infty}(\alpha, \beta, \boldsymbol{q}; f)$ iff $(\tilde{z}_1 / \sum_{j=1}^{n+1}\tilde{z}_j, \cdots, \tilde{z}_{n+1} / \sum_{j=1}^{n+1}\tilde{z}_j) \sim D_{n+1}(\alpha_1/\beta_1, \ldots, \alpha_{n+1}/\beta_{n+1})$ on H_{n+1} and is independent of the total size $\sum_{j=1}^{n+1}\tilde{z}_j$, which plays the role of r. For a proof of this, see Sivazlian (1981a) and Fang et al. (1990).

The following theorem gives the density function of a generalized Liouville distribution on the positive orthant with generating density function f.

Theorem 4.5 *The density function of a generalized Liouville distribution with generating density function* f *is given by*

$$g(\mathbf{z}) = A \prod_{i=1}^{n+1} \frac{z_i^{\alpha_i - 1}}{[\sum_{j=1}^{n+1}(z_j / q_j)^{\beta_j}]^{\theta - 1}} f\left((\frac{z_1}{q_1})^{\beta_1} + \cdots + (\frac{z_{n+1}}{q_{n+1}})^{\beta_{n+1}} \right), \tag{4.9}$$

where $A = \dfrac{\beta_1 \cdots \beta_{n+1} \Gamma(\theta)}{\prod_{i=1}^{n+1} q_i^{\alpha_i} \Gamma(\alpha_i / \beta_i)}$ and $\theta = \sum_{i=1}^{n+1} \alpha_i / \beta_i$.

This density is defined on the set $Z_a^{n+1} = \{(z_1, \ldots, z_{n+1}) | 0 < \sum_{i=1}^{n+1}(z_i / q_i)^{\beta_i} < a\}$ if and only if f is defined on the interval $(0, a)$.