JOURNAL OF MULTIVARIATE ANALYSIS 39, 106–116 (1991)

Some Parametric Models on the Simplex

O. E. BARNDORFF-NIELSEN

Aarhus University, Aarhus, Denmark

AND

B. JØRGENSEN

Instituto de Matemática Pura e Aplicada, Rio de Janeiro, Brazil

A class of new parametric models on the unit simplex in R^m is introduced, the distributions in question being obtained as conditional distributions of m independent generalized inverse Gaussian random variables given their sum. The Dirichlet model occurs as a special case. Two other special cases, corresponding respectively to the inverse Gaussian model and the reciprocal inverse Gaussian model, are studied in some detail. In particular, several exact chi-squared decompositions are found. © 1991 Academic Press, Inc.

1. INTRODUCTION

The purpose of this note is to introduce a class of new parametric models on the unit simplex in R^m and to list some of its properties. The models may be useful in studies of compositional data, as encountered, for instance, in geology and biology. For a comprehensive account of the statistical methodology for such data see Aitchison [3].

Referring to the "absence of satisfactory parametric classes of distributions on [the simplex]," Aitchison [1] developed the logistic normal distribution on the simplex and its statistical analysis. Aitchison [2] introduced a generalization of the logistic normal distribution on the simplex that includes the Dirichlet distribution as a special case. See also Aitchison [3]. Other generalizations of the Dirichlet distribution have been proposed by Connor and Mosimann [5] and Gupta and Richards [6].

Received August 1, 1989; revised September 12, 1990.

AMS 1980 subject classifications: 62E15; 62F99; 62H99.

Key words and phrases: chi-squared decompositions, compositional data, Dirichlet distribution, exponential models, generalized inverse Gaussian distribution, inverse Gaussian distribution, reciprocal inverse Gaussian distribution.

0047-259X/91 $3.00

Copyright © 1991 by Academic Press, Inc.
All rights of reproduction in any form reserved.

The distributions which constitute the new models are derived as the conditional distributions of m independent random variables given their sum, the distributional laws of the independent random variables being generalized inverse Gaussian. The class of Dirichlet distributions is obtained as a special case, due to the fact that the gamma distributions constitute a subclass of the generalized inverse Gaussian distributions. All the models are exponential models.

By definition, the models possess basis independence, in the sense of Aitchison [1]. In Section 5 we investigate some of the other concepts of independence studies by Aitchison [1, 3].

Two of the models are of particular interest. They correspond to the cases where the random variables, on whose sum we condition, are identically distributed according to an inverse Gaussian distribution or a reciprocal inverse Gaussian distribution, respectively. We refer to these models as S_{m-1}^{-} and S_{m-1}^{+} or, briefly, as S^{-} and S^{+}. Both models are proper dispersion models in the sense of Barndorff-Nielsen and Jørgensen [4].

The model S_{m-1}^{-} has the properties that its marginal and conditional distributions are again of the same type. In the case of S_{m-1}^{+} only the conditional distributions are of the type S^{+}.

Various general statistical properties of S^{-} and S^{+} follow from the properties of dispersion models [9, 10] and proper dispersion models [4]. The particular properties of S^{-} and S^{+} derive from those of the inverse Gaussian model and the reciprocal inverse Gaussian model.

In certain cases, the inference in the S_{m-1}^{-} model leads to tests based on independent χ^2-variables, in analogy with the linear normal model and with certain models for the inverse Gaussian and hyperboloid distributions [12, 7]. These and other results, derived below, show that the statistical properties of the new models have a structure that one would not expect to find in any other distributions on the simplex derived as the conditional distributions of m independent variables given their sum.

The discussion of the inferential properties of the models introduced in this paper is fairly brief. We hope to take up a more detailed study, including analysis of various data sets and a comparison to the alternative logistic normal models proposed and studied by Aitchison [3], elsewhere. However, it is worth pointing out that the models introduced in the present paper do not yet provide a full alternative to the logistic normal distributions as far as statistical analysis of compositional data is concerned, principally because the latter family of distributions has $m(m-1)/2$ covariance parameters, whereas our distributions have only one "variance," or rather "precision," parameter, namely, the parameter denoted λ below.

2. Definition and Derivation from the Generalized Inverse Gaussian Model

Let ∇^{m-1} denote the unit simplex in R^m, i.e.,

$$\nabla^{m-1} = \{(y_1, ..., y_m) : y_1 + \cdots + y_m = 1, y_1 > 0, ..., y_m > 0\}.$$

For any $\delta > 0$ we define a parametric model on $\delta\nabla^{m-1} = \{\delta y : y \in \nabla^{m-1}\}$ by its probability density function, with respect to Lebesgue measure on $\delta\nabla^{m-1}$,

$$p(y; \alpha, \mu, \gamma, \delta) = A(\alpha, \mu, \gamma, \delta) \Pi y_i^{\alpha_i - 1} e^{-\gamma Q(y; \mu, \delta)/2}, \tag{1}$$

where $\alpha = (\alpha_1, ..., \alpha_m) \in R^m$, $\mu = (\mu_1, ..., \mu_m) \in \delta\nabla^{m-1}$, and $\gamma > 0$ are parameters; $A(\alpha, \mu, \gamma, \delta)$ is a norming constant; and

$$\begin{aligned} Q(y; \mu, \delta) &= \Sigma(y_i - \mu_i)^2/y_i \\ &= \Sigma\mu_i^2/y_i - \delta. \end{aligned} \tag{2}$$

It is of interest to rewrite (1) slightly as

$$p(y; \alpha, \mu, \lambda, \delta) = a(\alpha, \mu, \lambda, \delta) \Pi y_i^{\alpha_i - 1} e^{\lambda t(y; \mu, \delta)}, \tag{3}$$

where a is another norming constant and

$$\begin{aligned} \lambda &= \gamma \Pi \mu_i^{-(2\alpha_i - 1)/(m-1)}, \\ t(y; \mu, \delta) &= -\tfrac{1}{2} \Pi \mu_i^{(2\alpha_i - 1)/(m-1)} Q(y; \mu, \delta). \end{aligned} \tag{4}$$

We shall denote the distribution (3) by $S_{m-1}(\alpha, \mu, \lambda, \delta)$ and when $\delta = 1$ we write $S_{m-1}(\alpha, \mu, \lambda)$ for $S_{m-1}(\alpha, \mu, \lambda, 1)$ and $Q(y; \mu)$ for $Q(y; \mu, 1)$.

If $y \sim S_{m-1}(\alpha, \mu, \lambda, \delta)$ then

$$\delta^{-1} y \sim S_{m-1}(\alpha, \delta^{-1}\mu, \delta^{(2\alpha_+ - 1)/(m-1)}\lambda), \tag{5}$$

where $\alpha_+ = \alpha_1 + \cdots + \alpha_m$.

Further, it may be noted that in case $m = 2$ and $\delta = 1$ the form $Q(y; \mu)$ can be reexpressed as

$$Q(y; \mu) = \frac{(y - \mu)^2}{y(1 - y)},$$

where we have written y for y_1 and μ for μ_1.

One way of arriving at the distribution $S_{m-1}(\alpha, \mu, \lambda, \delta)$ is by conditioning on the sum of m independent generalized inverse Gaussian random variables, in generalization of a standard derivation of the

Dirichlet distribution as the conditional distribution given the sum of m independent gamma random variates with a common scale parameter. (For a comprehensive account of the properties of the generalized inverse Gaussian distributions see [8].) Specifically, let $N^{-}(\alpha, \chi, \psi)$ denote the generalized inverse Gaussian distribution whose probability density function is

$$p(x; \alpha, \chi, \psi) = \frac{(\psi/\chi)^{\alpha/2}}{2K_\alpha(\sqrt{(\chi\psi)})} x^{\alpha-1} e^{-(\chi x^{-1} + \psi x)/2} \tag{6}$$

(where K_α is a Bessel function and where $x > 0$, $\chi, \psi > 0$, $\alpha \in R$). Furthermore, let $y_1, ..., y_m$ be independent random variables such that the distribution of y_i $(i = 1, ..., m)$ is $N^{-}(\alpha_i, \chi_i, \psi)$, i.e., the y_i have a common ψ parameter, and let $y_+ = y_1 + \cdots + y_m$. The conditional distribution of $y = (y_1, ..., y_m)$ given $y_+ = \delta$ is the same as (1) with

$$\gamma\mu_i^2 = \chi_i, \tag{7}$$

the parameter ψ disappearing by the conditioning due to the sufficiency for ψ of y_+. Note that

$$\gamma = \delta^{-2}(\Sigma\sqrt{\chi_i})^2. \tag{8}$$

3. The Models S^- and S^+

The two submodules corresponding to $\alpha_i = -1/2$ $(i = 1, ..., m)$ and to $\alpha_i = \frac{1}{2}$ $(i = 1, ..., m)$, respectively, are of special interest.

When the α_i equal $-1/2$ the distribution (6) is the inverse Gaussian distribution and in this case the simplex distribution (3) may be given the more explicit form

$$p(y; \mu, \lambda, \delta) = \delta^{1/2}\{\lambda/(2\pi)\}^{(m-1)/2}(\Pi y_i)^{-3/2} e^{\lambda t(y; \mu, \delta)}, \tag{9}$$

where

$$t(y; \mu, \delta) = -\tfrac{1}{2}(\Pi\mu_i)^{-2/(m-1)} Q(y; \mu, \delta). \tag{10}$$

This follows from the well-known result that the distribution of the sum y_+ is again an inverse Gaussian distribution,

$$y_+ \sim N^{-}(-\tfrac{1}{2}, (\sqrt{\chi_1} + \cdots + \sqrt{\chi_m})^2, \psi). \tag{11}$$

The distribution (9) will be denoted by $S_{m-1}^{-}(\mu, \lambda, \delta)$, and by $S_{m-1}^{-}(\mu, \lambda)$ when $\delta = 1$.

In case $\alpha_i = \frac{1}{2}$ $(i = 1, ..., m)$ the probability density (6) is that of the reciprocal inverse Gaussian distribution (in other words, x^{-1} follows the inverse Gaussian distribution). The corresponding version of (3), which we shall denote by $S^+_{m-1}(\mu, \lambda, \delta)$ and by $S^+_{m-1}(\mu, \lambda)$ when $\delta = 1$, is

$$p(y; \mu, \lambda, \delta) = a^+(\lambda, \delta)(\Pi y_i)^{-1/2} e^{\lambda t(y; \mu, \delta)}, \tag{12}$$

where

$$\begin{aligned} a^+(\lambda, \delta) &= (2\pi)^{-(m-1)/2} 2^{m/2} \Gamma(m/2) \delta^{-(m-1)/2} \lambda^{-1/2} / I_{m-1}(\delta\lambda), \\ I_{m-1}(\lambda) &= \int_0^\infty (1+z)^{-(m-1)/2} z^{(m-2)/2} e^{-\lambda z/2}\, dz \end{aligned} \tag{13}$$

and

$$t(y; \mu, \delta) = -\tfrac{1}{2} Q(y; \mu, \delta).$$

The only difficulty in establishing (12) consists in proving that the norming constant is as given by (13). To show this we invoke the result, due to P. Blæsild, that the reciprocal Gaussian distribution can be represented as the convolution of an inverse Gaussian distribution and a gamma distribution with one degree of freedom; specifically,

$$N^-(\tfrac{1}{2}, \chi, \psi) = N^-(-\tfrac{1}{2}, \chi\, \psi) * N^-(\tfrac{1}{2}, 0, \psi).$$

This, in combination with the convolution properties of the inverse Gaussian distribution (cf. formula (11) above) and of the gamma distribution, implies that the sum $y_+ = y_1 + \cdots + y_m$ (where $y_i \sim N^-(\frac{1}{2}, \chi_i, \psi)$) is distributed as

$$N^-\left(-\frac{1}{2}, \delta^2\gamma, \psi\right) * N^-\left(\frac{n}{2}, 0, \psi\right).$$

(Recall that χ, γ, μ and the given δ are related by (7), (8), and $\mu_+ = \delta$.) Consequently, the density function of y_+ may be expressed as

$$\delta\{\lambda/(2\pi)\}^{1/2} 2^{-m/2} \Gamma(m/2)^{-1} \psi^{m/2} e^{\delta\sqrt{(\lambda\psi)}} y_+^{(m-3)/2} e^{-(\lambda\delta^2/y_+ + \psi y_+)/2} I_{m-1}(\lambda\delta^2/y_+),$$

and (13) follows.

Henceforth we assume that y is distributed according to either (9) or (12) with $\delta = 1$, i.e., either

$$\begin{aligned} y &\sim S^-_{m-1}(\mu, \lambda) \\ &\sim \{\lambda/(2\pi)\}^{(m-1)/2} (\Pi y_i)^{-3/2} e^{-\lambda(\Pi\mu_i)^{-2/(m-1)} \Sigma(y_i - \mu_i)^2/2y_i}, \end{aligned} \tag{14}$$

or

$$y \sim S^+_{m-1}(\mu, \lambda)$$
$$\sim (2\pi)^{-(m-1)/2} 2^{m/2} \Gamma(m/2) \lambda^{-1/2} I_{m-1}(\lambda)^{-1} (\Pi y_i)^{-1/2} e^{-\lambda \Sigma (y_i - \mu_i)^2/2y_i}. \quad (15)$$

We shall refer to the parameters μ and λ of (14) and (15) as the *position* and the *precision*, respectively.

4. Relations to Some General Model Types

The models $S^-_{m-1}(\mu, \lambda)$ and $S^+_{m-1}(\mu, \lambda)$ are proper dispersion models in the sense of Barndorff-Nielsen and Jørgensen [4]; i.e., (14) as well as (15) are of the form

$$a(\lambda) b(y) e^{\lambda t(y; \mu)},$$

for certain functions a, b, and t.

It is, however, very noteworthy that neither $S^-_{m-1}(\mu, \lambda)$ nor $S^+_{m-1}(\mu, \lambda)$ is a transformation model, not even for λ fixed. Also, they are not exponential dispersion models (in the sense of Jørgensen [11]).

On the other hand, as is apparent from (1) and (2), both models are exponential models with canonical statistic $(y_1^{-1}, ..., y_m^{-1})$, the canonical parameter of (14) being

$$\theta = -\tfrac{1}{2}\lambda (\Pi \mu_i)^{-2/(m-1)} (\mu_1^2, ..., \mu_m^2).$$

while that of (15) is

$$\theta = -\tfrac{1}{2}\lambda (\mu_1^2, ..., \mu_m^2).$$

Under both models the parameter μ may be expressed in terms of the coordinates of θ by

$$\mu_i = s(\theta)^{-1} \sqrt{(-2\theta_i)},$$

while for S^-_{m-1},

$$\lambda = \{s(\theta)^{-1} \Pi \sqrt{(-2\theta_i)}\}^{2/(m-1)}$$

and for S^+_{m-1}

$$\lambda = s(\theta)^2,$$

where

$$s(\theta) = \Sigma \sqrt{(-2\theta_i)}.$$

The cumulant function of the exponential model (14) is, consequently,

$$\kappa(\theta) = -\Sigma \log \sqrt{(-2\theta_i)} + \log s(\theta) - \tfrac{1}{2}s(\theta)^2,$$

while that of (15) may be written

$$\kappa(\theta) = \log I_{m-1}(s(\theta)^2) + \log s(\theta) - \tfrac{1}{2}s(\theta)^2.$$

5. Marginal and Conditional Distributions, and Amalgamation

It follows from the interpretation of $S_{m-1}^{-}(\mu, \lambda, \delta)$ as a conditional distribution under an inverse Gaussian specification that both the conditional distributions and the marginal distributions of $S_{m-1}^{-}(\mu, \lambda, \delta)$ are of the type S^{-}.

To be more specific, with y distributed according to (14), let $\tilde{y} = (y_1, ..., y_k)$ and $\acute{y} = (y_{k+1}, ..., y_m)$, for some k with $1 \leqslant k < m$, and let $\tilde{\mu}$ and $\acute{\mu}$ be the defined similarly. Then the conditional distribution of $\acute{y}$ given $\tilde{y}$ is

$$\begin{aligned} \acute{y} \mid \tilde{y} \sim S_{m-k-1}^{-}&((1-\tilde{y}_+)(1-\tilde{\mu}_+)^{-1}\acute{\mu}, (1-\tilde{y}_+)^{-2}(1-\tilde{\mu}_+)^2 \\ &\times \gamma[\Pi_i\{(1-\tilde{y}_+)(1-\tilde{\mu}_+)^{-1}\acute{\mu}_i\}]^{2/(m-k-1)}, 1-\tilde{y}_+), \end{aligned} \tag{16}$$

where $\tilde{y}_+ = y_1 + \cdots + y_k$, $\tilde{\mu}_+ = \mu_1 + \cdots + \mu_k$. This appears from the sequence of distributional equivalences,

$$\begin{aligned} y \mid \tilde{y} &\sim y \mid \tilde{y}, y_+ = 1 \\ &\sim y \mid \tilde{y}, \acute{y}_+ = 1 - \tilde{y}_+ \\ &\sim \acute{y} \mid \acute{y}_+ = 1 - \tilde{y}_+. \end{aligned} \tag{17}$$

Furthermore, by division of the conditional density of $\acute{y}$ given $\tilde{y}$ in (14) we obtain that the marginal distribution of $\tilde{y}^+ = (\tilde{y}, 1-\tilde{y}_+) = (y_1, ..., y_k, 1-\tilde{y}_+)$ is given by

$$\tilde{y}^+ \sim S_k^{-}(\tilde{\mu}^+, \lambda(\Pi_i \mu_i)^{-2/(m-1)}\{(\Pi_i \tilde{\mu}_i)(1-\tilde{\mu}_+)\}^{2/k}), \tag{18}$$

where $\tilde{\mu}^+ = (\tilde{\mu}, 1-\tilde{\mu}_+) = (\mu_1, ..., \mu_k, 1-\tilde{\mu}_k)$. Note that, by (4), the γ parameter corresponding to the distribution (18) equals the γ parameter for y. In other words, the γ parameter is invariant under marginalization (whereas the λ parameter is not).

Repeated application of (18) shows that if $y = (y_1, ..., y_m) \sim S_{m-1}^{-}(\mu, \lambda)$ and if $z_1 = y_1 + \cdots + y_{i_1}$, $z_2 = y_{i_1+1} + \cdots + y_{i_2}$, ..., $z_k = y_{i_{k-1}+1} + \cdots + y_m$, where $1 \leqslant i_1 < i_2 < \cdots < i_{k-1} \leqslant m-1$, then $z = (z_1, ..., z_k)$ follows again an

S^- distribution, with the same γ as that for y. This is a useful amalgamation property. However, the S^- distributions and the Dirichlet distributions differ in that, for the latter, the $1+n$ vectors $(z_1, ..., z_k)$ and $(y_{i_{v-1}+1} z_v^{-1}, ..., y_{i_v} z_v^{-1})$, $v = 1, ..., k$, $i_0 = 0$, $i_k = m$, are stochastically independent. This independence property makes the Dirichlet distribution unrealistic for most problems of compositional analysis, as emphasized by Aitchison [3]. Like the S^- distributions, the logistic normal distributions do not have such an independence property. On the other hand, the logistic normal distributions differ from the S^- distributions by not constituting an invariant system under amalgamation [3, Section 6.6]. Thus, in this respect, the S^- distributions occupy a middle ground between the Dirichlet distributions and the logistic normal distributions.

The reasoning expressed in the sequence of relations (17) applies equally to the model S_{m-1}^+; i.e., the conditional distributions of S_{m-1}^+ are again of the type S^+. However, the marginal distributions of S_{m-1}^+ are not S^+ distributions as they involve ratios of functions of the form (15).

6. Some Inferential Properties

Again, let $y \sim S_{m-1}^-(\mu, \lambda)$. From (10) and (14) it follows that the Laplace transform of $^-2t(y; \mu)$ is given by

$$\{1 - 2u/\lambda\}^{-(m-1)/2};$$

i.e., we have

$$(\Pi \mu_i)^{-2/(m-1)} Q(y; \mu) \sim \lambda^{-1} \chi^2(m-1)$$

or, equivalently,

$$Q(y; \mu) = \gamma^{-1} \chi^2(m-1), \tag{19}$$

where $\chi^2(m-1)$ indicates the chi-squared distribution on $m-1$ degrees of freedom.

Further, let $(y_{1*}, ..., y_{k*})$, where $y_{j*} = (y_{j1}, ..., y_{jn_j})$ for $j = 1, ..., k$, be a partition of y; let $(\mu_{1*}, ..., \mu_{k*})$ be the corresponding partition of μ; and let $y_{j+} = y_{j1} + \cdots + y_{jn_j}$ and $\mu_{j+} = \mu_{j1} + \cdots + \mu_{jn_j}$. Because of (16) and (5) we find that

$$y_{j+}^{-1} y_{j*} \mid y_{j+} \sim S_{n_j-1}^-(\mu_{j+}^{-1} \mu_{j*}, y_{j+}^{-1} \lambda_j(\mu)),$$

where

$$\lambda_j(\mu) = \mu_{j+}^{-2/(n_j-1)} \left[\prod_{v=1}^{n_j} \mu_{jv} \right]^{2/(n_j-1)} \left[\prod_{j=1}^{k} \prod_{v=1}^{n_j} \mu_{jv} \right]^{-2(m-1)} \lambda,$$

and combining this with (19) shows that, conditionally on $y_{*+} = (y_{1+}, ..., y_{k+})$,

$$y_{j+}\mu_{j+}^{-2} \sum_{v=1}^{n_j} \frac{\mu_{jv}^2}{y_{jv}} - 1 \sim y_{j+}\mu_{j+}^{-2}\gamma^{-1}\chi^2(n_j - 1)$$

or, equivalently, that

$$Q(y_{j*}; \mu_{j*}, y_{j+}^{-1}\mu_{j+}^2) = \sum \frac{\mu_{jv}^2}{y_{jv}} - \frac{\mu_{j+}^2}{y_{j+}} \sim \gamma^{-1}\chi^2(n_j - 1). \tag{20}$$

Since the right-hand side of (20) does not depend on y_{*+} we have that (20) holds marginally, as well as conditionally given y_{*+}, that

$$Q(y_{1*}; \mu_{1*}, y_{1+}^{-1}\mu_{1+}^2), ..., Q(y_{k*}; \mu_{k*}, y_{k+}^{-1}\mu_{k+}^2) \quad \text{and} \quad y_{*+}$$

are mutually independent and that

$$Q(y_{1*}; \mu_{1*}, y_{1+}^{-1}\mu_{1+}^2), ..., Q(y_{k*}; \mu_{k*}, y_{k+}^{-1}\mu_{k+}^2) \quad \text{and} \quad Q(y_{*+}; \mu_{*+})$$

constitute an exact χ^2-decomposition of $Q(y; \mu)$ into $k+1$ independent components.

Next, let y_h, $h = 1, ..., n$, denote a random sample from either (14) or (15). We introduce the notation $\overset{\#}{y}$ for the vector of the m harmonic means of the coordinates of y_h, $h = 1, ..., n$, i.e., $\overset{\#}{y} = (\overset{\#}{y}_1, ..., \overset{\#}{y}_m)$ where

$$\overset{\#}{y}_i = n\left[\sum_{h=1}^{n} y_{hi}^{-1}\right]^{-1}$$

Furthermore, we let

$$\begin{aligned} \overset{\#}{y}_+ &= \overset{\#}{y}_1 + \cdots + \overset{\#}{y}_m, \\ R(\overset{\#}{y}_+) &= 1 - \overset{\#}{y}_+, \end{aligned} \tag{21}$$

and

$$\overset{\#}{y}_. = \overset{\#}{y}_+^{-1}\overset{\#}{y}.$$

Note that $R(\overset{\#}{y}_+) \geqslant 0$, in consequence of the fact that the harmonic mean is always less than or equal to the arithmetic mean.

Suppose now that the model is given by (14), i.e., $y_h \sim S_{m-1}^{-}(\mu, \lambda)$ $(h = 1, ..., n)$.

The log likelihood function is

$$l(\mu, \lambda) = (n/2)\{(m-1)\log\lambda - \lambda(\Pi\mu_i)^{-2/(m-1)}n^{-1}\Sigma Q(y_h; \mu)\}.$$

In view of the definitions (2) and (21) we may reexpress the log likelihood function as

$$l(\mu, \lambda) = (n/2)](m-1)\log\lambda - \lambda(\Pi\mu_i)^{-2/(m-1)}\{\Sigma\mu_i^2/\overset{\#}{y}_{\cdot i} - 1\}]$$

or, equivalently, as

$$\begin{aligned} l(\mu, \lambda) = (n/2)[(m-1)\log\lambda - \lambda(\Pi\mu_i)^{-2/(m-1)}\overset{\#}{y}_{+}^{-1} \\ \times\{Q(\overset{\#}{y}_{\cdot};\mu) + R(\overset{\#}{y}_{+})\}]. \end{aligned}$$

It follows, in particular, that $(\overset{\#}{y}_{+}, \overset{\#}{y}_{\cdot})$ is a minimally sufficient statistic and that the likelihood equations for μ and λ may be transformed to

$$(m-1)\mu_i^2/\overset{\#}{y}_{\cdot i} + \{Q(\overset{\#}{y}_{\cdot};\mu) - m\overset{\#}{y}_{+} + 1\}\mu_i - Q(\overset{\#}{y}_{\cdot};\mu) - R(\overset{\#}{y}_{+}) = 0 \qquad (22)$$

and

$$\lambda^{-1} = (m-1)^{-1}(\Pi\mu_i)^{-2/(m-1)}\overset{\#}{y}_{+}^{-1}\{Q(\overset{\#}{y}_{\cdot};\mu) + R(\overset{\#}{y}_{+})\}. \qquad (23)$$

Equation (22) does not involve λ; however, to yield the maximum likelihood estimate $\hat{\mu}$ for μ it must be solved numerically. The maximum likelihood estimate for λ then emerges by inserting $\hat{\mu}$ for μ on the right-hand side of (23).

Furthermore, in extension of (19), we have

$$\overset{\#}{y}_{+}^{-1}\{Q(\overset{\#}{y}_{\cdot};\mu) + R(\overset{\#}{y}_{+})\} \sim \gamma^{-1}\chi^2(n(m-1)).$$

The proof is a direct generalization of that for (19).

Turning now to the model S_{m-1}^{+} we find that the log likelihood function based on y_h, $h = 1, ..., n$, is

$$l(\mu, \lambda) = (n/2)[-\log\lambda - 2\log I_{m-1}(\lambda) - \lambda\overset{\#}{y}_{+}^{-1}\{Q(\overset{\#}{y}_{\cdot};\mu) + R(\overset{\#}{y}_{+})\}].$$

Thus in this case, too, $(\overset{\#}{y}_{+}, \overset{\#}{y}_{\cdot})$ is a minimally sufficient statistic. Furthermore, the maximum likelihood estimates for μ and λ are given by

$$\hat{\mu} = \overset{\#}{y}_{\cdot};$$

i.e., $\hat{\mu}$ is the vector of harmonic means, and

$$\begin{aligned} H'(\lambda) &= \overset{\#}{y}_{+}^{-1}R(\overset{\#}{y}_{+}) \\ &= \overset{\#}{y}_{+}^{-1} - 1, \end{aligned}$$

where

$$H(\lambda) = -\log\lambda - 2\log I_{m-1}(\lambda).$$

If λ is large then $H(\lambda)$ may be approximated by $(m-1)\log\lambda$ and, correspondingly, we have the distributional approximations

$$n\overset{\#}{y}_{+}^{-1}R(\overset{\#}{y}_{+}) \overset{\cdot}{\sim} \lambda^{-1}\chi^2((n-1)(m-1)) \tag{24}$$

$$n\overset{\#}{y}_{+}^{-1}Q(\hat{\mu};\mu) \overset{\cdot}{\sim} \lambda^{-1}\chi^2(m-1). \tag{25}$$

The result (24) may be used for testing homogeneity of n values of μ, and (25) provides a test for a point hypothesis concerning μ.

These results follow from the theory of small-dispersion (i.e., λ large) asymptotics discussed in Jørgensen [10]. Similar results hold for the S^- model.

Acknowledgment

We are grateful to Joe Whittaker for a useful remark concerning the relation to the inverse Gaussian distribution.

References

[1] Aitchison, J. (1982). The statistical analysis of compositional data (with Discussion). *J. Roy. Statist. Soc. B* **44** 139–177.

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

[3] Aitchison, J. (1986). *The Statistical Analysis of Compositional Data. Chapman & Hall*, London.

[4] Barndorff-Nielsen, O. E., and Jørgensen, B. (1991). Proper dispersion models. In preparation.

[5] Connor, R. J., and Mosimann, J. E. (1969). Concepts of independenc for proportions with a generalization of the Dirichlet distribution. *J. Amer. Statist. Assoc.* **64** 194–206.

[6] Gupta, R. D., and Richards, D. St. P. (1987). Multivariate Liouville distributions. *J. Multivariate Anal.* **23** 233–256.

[7] Jensen, J. L. (1981). On the hyperboloid distribution. *Scand. J. Statist.* **8** 193–206.

[8] Jørgensen, B. (1982). *Statistical Properties of the Generalized Inverse Gaussian Distribution.* Lecture Notes in Statistics, Vol. 9. Springer-Verlag, New York.

[9] Jørgensen, B. (1983). Maximum likelihood estimation and large-sample inference for generalized linear and nonlinear regression models. *Biometrika* **70** 19–28.

[10] Jørgensen, B. (1987). Small-dispersion asymptotics. *Rev. Bras. Prob. Estat.* **1** 59–90.

[11] Jørgensen, B. (1987). Exponential dispersion models (with Discussion). *J. Roy. Statist. Soc. B* **49** 127–162.

[12] Tweedie, M. C. K. (1957). Statistical properties of the inverse Gaussian distribution I. *Ann. Math. Statist.* **28** 362–377.