Ann. Inst. Statist. Math.
Vol. 47, No. 3, 447–459 (1995)

IDENTIFIABILITY OF MIXTURES OF POWER-SERIES DISTRIBUTIONS AND RELATED CHARACTERIZATIONS

THEOFANIS SAPATINAS

Department of Mathematical Statistics and Operational Research, Exeter University, Exeter EX4-4QE, U.K.

(Received January 28, 1994; revised October 24, 1994)

Abstract. The concept of the identifiability of mixtures of distributions is discussed and a sufficient condition for the identifiability of the mixture of a large class of discrete distributions, namely that of the power-series distributions, is given. Specifically, by using probabilistic arguments, an elementary and shorter proof of the Lüxmann-Ellinghaus's (1987, *Statist. Probab. Lett.*, **5**, 375–378) result is obtained. Moreover, it is shown that this result is a special case of a stronger result connected with the Stieltjes moment problem. Some recent observations due to Singh and Vasudeva (1984, *J. Indian Statist. Assoc.*, **22**, 93–96) and Johnson and Kotz (1989, *Ann. Inst. Statist. Math.*, **41**, 13–17) concerning characterizations based on conditional distributions are also revealed as special cases of this latter result. Exploiting the notion of the identifiability of power-series mixtures, characterizations based on regression functions (posterior expectations) are obtained. Finally, multivariate generalizations of the preceding results have also been addressed.

Key words and phrases: Univariate and multivariate power-series distributions, mixtures of distributions, the moment problem, infinite divisibility, posterior expectations.

1. Introduction

Mixtures of distributions are used in building probability models quite frequently in biological and physical sciences. For instance, in order to study certain characteristics in natural populations of fish, a random sample might be taken and the characteristic measured for each member of the sample; since the characteristic varies with the age of the fish, the distribution of the characteristic in the total population will be a mixture of the distribution at different ages. In order to analyse the qualitative character of inheritance, a geneticist might observe a phenotypic value that has a mixture distribution because each genotype might produce phenotypic values over an interval. For a comprehensive account on mixtures of distributions as they occur in diverse fields, one is referred to Titterington *et al.* (1985) and Prakasa Rao (1992) among others.

In mathematical terms, mixtures of distributions may be described by

$$H(x) = \int_{\Omega} F(x \mid \theta) dG(\theta), \tag{1.1}$$

where $F(\cdot \mid \theta)$ is a distribution function for all $\theta \in \Omega$, $F(x \mid \cdot)$ is a Borel measurable function for each x and G is a distribution function defined on Ω. It is easy to see that H is a distribution function and it is called a mixture. The family $F(x \mid \theta)$, $\theta \in \Omega$ is referred to as the kernel of the mixture and G as the mixing distribution function. In order to devise statistical procedures for inferential purposes, an important requirement is the identifiability of the mixing distribution. Unless this is so, it is not meaningful to estimate the distribution either non-parametrically or parametrically, especially in the Bayesian or empirical Bayesian analysis. The mixture H defined by (1.1) is said to be identifiable if there exists a unique G yielding H, or equivalently, if the relationship

$$H(x) = \int_{\Omega} F(x \mid \theta) dG_1(\theta) = \int_{\Omega} F(x \mid \theta) dG_2(\theta)$$

implies $G_1(\theta) = G_2(\theta)$ for all $\theta \in \Omega$. Lack of identifiability is not uncommon and not artificial, as it is shown by considering the binomial distribution $p(x \mid \theta) = \binom{n}{x}\theta^x(1-\theta)^{n-x}$, $x = 0, 1, \ldots, n$, $\theta \in (0, 1)$. We see that for every x, $p(x)$ is a linear function of the first n moments $\mu_r = \int_0^1 \theta^r dG(\theta)$, $r = 1, 2, \ldots, n$ of $G(\theta)$. Consequently, any other $G^*(\theta)$ with the same first n moments will yield the same mixed distribution $p(x)$.

The identifiability problems concerning finite and countable mixtures (i.e. when the support of G in (1.1) is finite and countable respectively) have been investigated by Teicher (1963), Patil and Bildikar (1966), Yakowitz and Spragins (1968), Tallis (1969), Fraser *et al.* (1981), Tallis and Chesson (1982) and Kent (1983). In the case of a general mixture (i.e. when G is arbitrary), a totally different picture emerges. For instance, the class of arbitrary mixtures of normal distribution $\{N(\mu, \sigma^2), -\infty < \mu < \infty, 0 < \sigma^2 < \infty\}$ is not identifiable (see Teicher (1960)), whereas the class of finite mixtures of the same family is identifiable (see Teicher (1963)). So, we need to obtain some sufficient conditions for identifiability of arbitrary mixtures. Teicher (1961) studied the identifiability of mixtures of additively closed families, while Barndörff-Nielsen (1965) discussed the identifiability of mixtures of some restricted multivariate exponential families. More recently, Lüxmann-Ellinghaus (1987) has given a sufficient condition for the identifiability of a large class of discrete distributions, namely that of the power-series distributions. Using topological arguments, he has shown that mixtures of this family are identifiable provided that the family in question is infinitely divisible.

In the present paper, a sufficient condition for the identifiability of arbitrary power-series mixtures is given, and it is used to obtain some characterization results. Specifically, in Section 2, we have given an elementary and shorter proof for the Lüxmann-Ellinghaus's (1987) result by using probabilistic arguments. Moreover, we have shown that this result is a special case of a stronger result connected with the Stieltjes moment problem. Some recent observations due to Singh and

Vasudeva (1984) and Johnson and Kotz (1989) concerning characterizations based on conditional distributions are also revealed as special cases of our results. In Section 3, exploiting the notion of the identifiability of power-series mixtures, characterizations based on regression functions (posterior expectations) are obtained. Finally, in Section 4, multivariate generalizations of the preceding results have also been addressed.

2. Identifiability of arbitrary power-series mixtures

Theorem 2.1 given below shows that the concept of identifiability of arbitrary mixtures of power-series distributions is linked with the Stieltjes moment problem.

THEOREM 2.1. *Let X be a random variable with the power-series mixture distribution defined by*

$$p(x) = \int_I \frac{a_x \theta^x}{A(\theta)} dG(\theta), \qquad x = 0, 1, \dots, \tag{2.1}$$

where $a_x > 0$ for all $x = 0, 1, \dots$, $I = [0, \rho)$, $0 < \rho \leq \infty$, ρ is the radius of convergence of $A(\theta)$, and G is a distribution function defined on I. Assume that

$$\sum_{x=1}^{\infty} \left(\frac{p(x)}{a_x} \right)^{-1/2x} = \infty. \tag{2.2}$$

Then the mixture (2.1) *is identifiable.*

PROOF. It is easily seen that $\mu_x^* = \frac{p(x)a_0}{a_x p(0)}$, $x = 0, 1, \dots$ is the moment sequence of the distribution function G^*, where

$$dG^*(\theta) = \frac{a_0}{p(0)} \frac{1}{A(\theta)} dG(\theta), \qquad \theta \in I. \tag{2.3}$$

Moreover,

$$\sum_{x=1}^{\infty} \mu_x^{*-1/2x} = \infty$$

if and only if (2.2) is valid. (Incidentally, when $\rho < \infty$ the distribution function G has a bounded support and hence the above condition is always satisfied.) Hence, in view of Carleman's condition for the Stieltjes moment problem (see Shohat and Tamarkin (1963), pp. 19–20), the distribution function G^* is uniquely determined by its moments which entails that G is uniquely determined by the sequence $\{p(x) : x = 0, 1, \dots\}$ given the power-series distribution. This implies that the mixture (2.1) is identifiable and thus the proof of the theorem is completed.

Remark 1. (i) If (2.2) is replaced by

$$\sum_{x=1}^{\infty} \gamma_x^{-1} = \infty,$$

where

$$\gamma_x = \inf_{k \geq x} \left(\frac{p(k)}{a_k} \right)^{1/2k},$$

then the conclusion of Theorem 2.1 remains valid.

(ii) If the restriction $a_x > 0$ for all $x = 0, 1, \dots$ is replaced by $\{x : a_x > 0\} = \{\beta + \gamma x : x = 0, 1, \dots, \beta \in N_0, \gamma \in N\}$, then Theorem 2.1 still holds after having (2.2) replaced with

$$\sum_{x=1}^{\infty} \left(\frac{p(\beta + \gamma x)}{a_{\beta+\gamma x}} \right)^{-1/2x} = \infty.$$

(iii) The conclusion of Theorem 2.1 remains valid if $\frac{a_x \theta^x}{A(\theta)}$ in (2.1) is replaced by $\frac{a_x (h(\theta))^x}{A(h(\theta))}$, where h, h^{-1} are measurable functions and h is a one-to-one correspondence.

Note 1. It is a simple exercise to check that in the case of Poisson, negative binomial and logarithmic series mixtures the condition (2.2) is obviously met and we arrive at the result that the mixtures in question are identifiable. On the other hand, as pointed out in the introduction, for the binomial mixtures, there is no identifiability; this implies that the condition (2.2) is not met in the present case and it is easy to verify this. Obviously, Theorem 2.1 subsumes the cases of finite and countable mixtures. Thus, some known results such as the identifiability of finite mixtures of negative binomial distributions (see Yakowitz and Spragins (1968)) are included in Theorem 2.1.

In the sequel, we give an elementary and shorter proof for Lüxmann-Ellinghaus's (1987) result by using probabilistic arguments, and reveal that this result is also a corollary of Theorem 2.1.

THEOREM 2.2. *Let X be a random variable with the power-series mixture distribution given by* (2.1). *Assume that*

$$p(x \mid \theta) = \frac{a_x \theta^x}{A(\theta)}, \qquad x = 0, 1, \dots,$$

is infinitely divisible (possibly shifted) for $\theta = \theta_0 \in I$. Then the mixture (2.1) *is identifiable.*

PROOF. Since $p(x \mid \theta)$ is infinitely divisible for $\theta = \theta_0 \in I$, we get that

$$\frac{A(\theta_0 s)}{A(\theta_0)} = e^{-\lambda(\theta_0) + \lambda(\theta_0 s)}, \qquad s \in [0, 1], \tag{2.4}$$

where $\lambda(\theta) = \sum_{r=1}^{\infty} c_r \theta^r$ such that $c_r \geq 0$ (see Feller (1968), p. 290). Taking k to be the smallest value such that $c_k > 0$, (2.4) can be written as

$$\begin{aligned} \frac{A(\theta_0 s)}{A(\theta_0)} &= e^{-\lambda(\theta_0) + \lambda^*(\theta_0 s) + c_k \theta_0^k s^k}, \\ &= e^{-\lambda(\theta_0)} \left(\sum_{r=0}^{\infty} \frac{c_k^r \theta_0^{kr} s^{kr}}{r!} \right) e^{\lambda^*(\theta_0 s)}, \end{aligned} \tag{2.5}$$

where $\lambda^*(\theta_0 s)$ is a power series with non-negative coefficients with $\lambda^*(0) = 0$. Relation (2.5) implies that $A(\theta_0)e^{-\lambda(\theta_0)} = a_0$ on taking $s = 0$; we shall assume without loss of generality $a_0 = 1$. On using the equation obtained by equating the coefficients of $(\theta_0 s)^{kx}$ on both sides of (2.5) in conjunction with this equation, we then get

$$a_{xk} = \sum_{r=0}^{x} \frac{c_k^{x-r}}{(x-r)!}\xi_r,$$

where ξ_r is the coefficient of $(\theta_0 s)^r$ in the expansion of $e^{\lambda^*(\theta_0 s)}$. Note that $\lambda^*(0) = 0$ implies that $\xi_0 = 1$ and as $\xi_r \geq 0$ for $r \geq 1$, we can then conclude that

$$a_{xk} \geq \frac{c_k^x}{x!}. \tag{2.6}$$

Consequently, the moment generating function (m.g.f.)

$$M(t) = \sum_{x=0}^{\infty} \frac{p(xk)a_0}{a_{xk}p(0)} \frac{t^x}{x!} \leq \frac{a_0}{p(0)} \sum_{x=0}^{\infty} \left(\frac{t}{c_k}\right)^x < \infty, \qquad \text{for} \quad |t| < c_k.$$

The latter shows that the distribution function corresponding to the moment sequence $\{\mu_{xk} : x = 0, 1, \ldots\}$ is uniquely determined which implies that the distribution function G is uniquely determined by the sequence $\{p(xk) : x = 0, 1, \ldots\}$ given the power-series distribution. This implies that the mixture (2.1) is identifiable, and hence the proof of the theorem is completed.

Remark 2. (i) It is easily seen that after establishing (2.6), Theorem 2.2 could also be made to follow as a consequence of Theorem 2.1, because, on using the Stirling formula, we now get

$$\sum_{x=1}^{\infty} \left(\frac{p(xk)}{a_{xk}}\right)^{-1/2x} \geq c_k^{1/2} \sum_{x=1}^{\infty} (x!)^{-1/2x} = \infty$$

which, in view of Remark 1 (ii), implies that the mixture (2.1) is identifiable.

(ii) Among power-series distributions, the Poisson, negative binomial and logarithmic series are infinitely divisible, and so their mixtures are identifiable. The binomial distribution is a power-series distribution as well, but it is not infinitely divisible, and so its identifiability with respect to success parameter is not established; a result which is consistent with the implications of Theorem 2.1.

Remark 3. Singh and Vasudeva (1984) proved, using an extended Stone-Weierstrass theorem, the following result: let X, Y and Z be random variables such that X and Y are non-negative and

$$P(Z = x \mid X = \theta) = P(Z = x \mid Y = \theta) = e^{-\theta}(1 - e^{-\theta})^x, \qquad \theta \geq 0, \quad x = 0, 1, \ldots.$$

Then X and Y are identically distributed. (They used this result to characterize the exponential distribution.) An extension of the above result, which was used to characterize some other distributions, was obtained by Johnson and Kotz (1989) as stated below: if X, Y have the same support, and

$$P(Z = x \mid X = \theta) = P(Z = x \mid Y = \theta) = g(\theta)(h(\theta))^x, \tag{2.7}$$

$(g(\theta), h(\theta) > 0)$ for all $x = 0, 1, \ldots$ and all θ in the common support of X and Y, and $h(\theta)$ is a strictly monotonic function of θ, then X and Y have identical distributions. (Note here that (2.7) immediately implies that $h(\theta) < 1$.) As $h(\theta)$ above is a strictly monotonic function, there is no loss of generality in assuming that $h(\theta) = \theta$, and hence the above result is an obvious corollary of Theorem 2.1. Alternatively, by taking $a_x = (-1)^x \binom{-k}{x}$, $k > 0$, and $A(h(\theta)) = (1 - h(\theta))^k$ with $h(\theta) \in (0, 1)$, (2.2) is obviously met and we get an identifiability result. This implies that Johnson and Kotz's (1989) result is obviously valid if the right hand side of (2.7) is taken as $g(\theta)\binom{-k}{x}(-h(\theta))^x$ (i.e. the one corresponding to a negative binomial distribution instead of a geometric distribution).

3. Characterizations based on posterior expectations

In this section, characterizations based on posterior expectations are obtained. Johnson (1957, 1967), in the case of a Poisson (θ) mixture, has shown that the relation

$$E(\Theta \mid X = x) = ax + b, \qquad x = 0, 1, \ldots, \tag{3.1}$$

where a, b are real, implies a unique form of prior distribution for Θ. Ericson (1969) has shown that if for certain problems (3.1) holds, then the first two factorial moments of the prior distribution are uniquely determined by a, b and the $\mathrm{Var}(X)$, whilst Diaconis and Ylvisaker (1979), in the case of an exponential family, characterized conjugate prior measures through the property

$$E\{E(X \mid \Theta) \mid X = x\} = ax + b, \qquad x = 0, 1, \ldots.$$

A natural question that arises now is the following: under what conditions, in the case of a power-series mixture, can a given regression function (posterior expectation) of Θ on X be used to determine the distribution of X and Θ uniquely. An answer to that question is given in Theorem 3.1 below, on exploiting the notion of identifiability of mixtures of power-series distributions with respect to a scalar parameter θ as stated in Theorem 2.1. The result is applied to some univariate distributions.

THEOREM 3.1. *Let X be a random variable with the power-series mixture distribution given by* (2.1). *Also, let the regression function* (*posterior expectation*)

$$m(x) = E(\Theta \mid X = x), \qquad x = 0, 1, \ldots$$

be such that

$$\sum_{x=1}^{\infty}\left(\prod_{i=0}^{x-1} m(i)\right)^{-1/2x} = \infty. \tag{3.2}$$

Then, the distribution of Θ *is uniquely determined by* m.

PROOF. The regression function $m(x)$ can be expressed as

$$m(x) = E(\Theta \mid X = x) = \frac{\int_I \frac{\theta^{x+1}}{A(\theta)} dG(\theta)}{\int_I \frac{\theta^{x}}{A(\theta)} dG(\theta)}, \qquad x = 0, 1, \ldots,$$

which entails that

$$m(x) = \frac{p(x+1)a_x}{p(x)a_{x+1}}, \qquad x = 0, 1, \ldots, \tag{3.3}$$

where $p(x)$ is as defined in (2.1). Relation (3.3) implies now that

$$\frac{p(x)}{a_x} = \frac{p(0)}{a_0}\prod_{i=0}^{x-1} m(i), \qquad x = 1, 2, \ldots,$$

which yields, by making use of condition (3.2) and Theorem 2.1, that the distribution of Θ is uniquely determined by m. (This, in turn, implies that the joint distribution of X and Θ is uniquely determined by m.) This completes the proof of the theorem.

Remark 4. It is a simple exercise to check that (3.2) is met if $\limsup_{x\to\infty} \frac{m(x)}{x} < \infty$; consequently, we have that Theorem 3.1 holds under this stronger condition in place of (3.2). This is so, because, for sufficiently large x, we have $\frac{m(x)}{x} \le k(<\infty)$. Assume that this is so for $x \ge x_0$. Then

$$\sum_{x=1}^{\infty}\left(\prod_{i=0}^{x-1} m(i)\right)^{-1/2x} \ge \sum_{x=x_0+1}^{\infty}\left[\left(\prod_{i=0}^{x_0} m(i)\right)^{-1/2x} (kx)^{-(x-x_0-1)/2x}\right] = \infty.$$

The following corollaries of Theorem 3.1 are some interesting applications.

COROLLARY 3.1. *Let* $X \mid \Theta = \theta \sim$ *Poisson* $(\lambda\theta)$. *Then for some* $\alpha, \beta > 0$

$$m(x) = \frac{\alpha + x}{\beta + \lambda}, \qquad x = 0, 1, \ldots$$

if and only if $X \sim$ *Negative Binomial* $(\alpha, \beta/(\beta+\lambda))$; *then* $\Theta \sim$ *Gamma* (α, β).

Corollary 3.2. *Let* $X \mid \Theta = \theta \sim$ *Poisson* (θ). *Then for some* $\alpha > 0$

$$m(x) = \frac{(x+1)(x+\alpha+3)}{(\alpha+1)(\alpha+x+2)}, \qquad x = 0, 1, \ldots$$

if and only if $X \sim$ *"Poisson-Lindley" distribution (see Sankaran* (1970)) *with probability distribution (p.d.) given by*

$$p(x) = \alpha^2 \frac{(x+\alpha+3)}{(\alpha+1)^{x+3}}, \qquad x = 0, 1, \ldots;$$

then Θ *has probability density function (p.d.f.) given by*

$$g(\theta) = \frac{\alpha^2}{(\alpha+1)}(\theta+1)e^{-\alpha\theta}, \qquad \theta > 0.$$

Corollary 3.3. *Let* $X \mid \Theta = \theta \sim$ *Negative Binomial* $(k, 1-\theta)$. *Then for some* $\alpha, \beta > 0$,

$$m(x) = \frac{x+\alpha}{x+\alpha+k+\beta}, \qquad x = 0, 1, \ldots$$

if and only if $X \sim$ *Negative Binomial-Beta* $(-\alpha, \beta, -k)$; *then* $\Theta \sim$ *Beta* $(-\alpha+\beta+1, \alpha)$.

Corollary 3.4. *Let* $X \mid \Theta = \theta \sim$ *Logarithmic Series* $(1-\theta)$. *Then for some* $\lambda > 0$, $\mu \geq 0$

$$m(x) = \frac{\mu+x}{\lambda+\mu+x}, \qquad x = 1, 2, \ldots$$

if and only if $X \sim$ *Digamma* (μ, λ) *(see Sibuya* (1979)) *with p.d. given by*

$$p(x) = \frac{1}{\psi(\lambda+\mu) - \psi(\lambda)} \frac{(\mu)_x}{(\lambda+\mu)_x} \frac{1}{x}, \qquad x = 1, 2, \ldots,$$

where ψ *is the digamma function and* $(a)_x = a(a+1)\cdots(a+x-1)$, $x = 1, 2, \ldots$, $(a)_0 = 1$; *then* $\Theta \sim$ *"One-End Accented Beta"* (λ, μ) *(see Sibuya* (1979)) *with p.d.f. given by*

$$g(\theta) = \frac{1}{C(\lambda, \mu)}(-\log(1-\theta))(1-\theta)^{\lambda-1}\theta^{\mu-1}, \qquad 0 < \theta < 1,$$

where $C(\lambda, \mu)$ *is the normalized constant. (The case* $\mu = 0$ *gives a characterization for the trigamma distribution.)*

4. Identifiability of multivariate power-series mixtures and related characterizations

In this section, multivariate generalizations of the results of Sections 2 and 3 are given. Theorem 4.1 given below, provides us with a general identifiability result for arbitrary mixtures of multivariate (p-variate) power-series distributions (which essentially is a multivariate extension of Theorem 2.1) and links the concept of identifiability of all mixtures under consideration with the p-dimensional moment problem.

THEOREM 4.1. *Let* $\boldsymbol{X} = (X_1, \ldots, X_p)$ *be a* p*-dimensional random vector with the multivariate power-series mixture distribution defined by*

$$(4.1)\qquad p(x_1, \ldots, x_p) = \int_{I^p} \frac{a_{x_1,\ldots,x_p}\theta_1^{x_1}\cdots\theta_p^{x_p}}{A(\theta_1, \ldots, \theta_p)} dG(\theta_1, \ldots, \theta_p), \quad x_i = 0, 1, \ldots, \quad i = 1, \ldots, p$$

where $a_{x_1,\ldots,x_p} > 0$ *for all* $x_i = 0, 1, \ldots,$ $i = 1, \ldots, p$, I^p *is the* p*-fold cartesian product of* $I = [0, \rho)$, *and* G *is a multivariate distribution function defined on* I^p. *Assume that*

$$(4.2)\qquad \sum_{x=1}^{\infty} \left(\frac{p(x, 0, \ldots, 0)}{a_{x,0,\ldots,0}} + \frac{p(0, x, 0, \ldots, 0)}{a_{0,x,0,\ldots,0}} + \cdots + \frac{p(0, \ldots, 0, x)}{a_{0,\ldots,0,x}} \right)^{-1/2x} = \infty.$$

Then the mixture (4.1) *is identifiable.*

PROOF. It is easily seen that $\mu^*_{x_1,\ldots,x_p} = \frac{p(x_1,\ldots,x_p)a_{0,\ldots,0}}{a_{x_1,\ldots,x_p}p(0,\ldots,0)}$, $x_i = 0, 1, \ldots, i = 1, \ldots, p$ is the moment sequence of the distribution function G^*, where

$$(4.3)\qquad dG^*(\theta_1, \ldots, \theta_p) = \frac{a_{0,\ldots,0}}{p(0, \ldots, 0)} \frac{1}{A(\theta_1, \ldots, \theta_p)} dG(\theta_1, \ldots, \theta_p), \quad \boldsymbol{\theta} = (\theta_1, \ldots, \theta_p) \in I^p.$$

Moreover,

$$\sum_{x=1}^{\infty} (\mu^*_{x,0,\ldots,0} + \mu^*_{0,x,0,\ldots,0} + \mu^*_{0,\ldots,0,x})^{-1/2x} = \infty$$

if and only if (4.2) is valid. (Note that when G has a bounded support the above condition is always satisfied.) Hence, in view of the Cramer-Wold condition for the p-dimensional Stieltjes moment problem (see Akhiezer (1965), p. 227), the distribution function G^* is uniquely determined by its moments which entails that G is uniquely determined by the sequence $\{p(x_1, \ldots, x_p) : x_i = 0, 1, \ldots, i = 1, \ldots, p\}$ given the multivariate power-series distribution. This implies that the mixture (4.1) is identifiable and thus the proof of the theorem is completed.

Remark 5. In the special case where $A(\theta_1, \ldots, \theta_p)$ is replaced by $A(k_1\theta, \ldots, k_p\theta)$ with $k_1, \ldots, k_p > 0$, (4.1) can be written as

$$(4.4)\qquad p(x_1, \ldots, x_p) = \int_I \frac{a_{x_1,\ldots,x_p}k_1^{x_1}\cdots k_p^{x_p}\theta^z}{A(k_1\theta, \ldots, k_p\theta)} dG(\theta), \quad x_i = 0, 1, \ldots, \quad i = 1, \ldots, p,$$

where $x_1 + \cdots + x_p = z$, $z = 0, 1, \ldots$.

It is easily seen that $\mu_z^* = \frac{p(x_1,\ldots,x_p)a_{0,\ldots,0}}{a_{x_1,\ldots,x_p}p(0,\ldots,0)} \frac{1}{k_1^{x_1}\cdots k_p^{x_p}}$, $z = 0, 1, \ldots$ is the moment sequence of the distribution function G^*, where

$$dG^*(\theta) = \frac{a_{0,\ldots,0}}{p(0,\ldots,0)} \frac{1}{A(k_1\theta,\ldots,k_p\theta)} dG(\theta), \qquad \theta \in I. \tag{4.5}$$

From the above we conclude that if

$$\sum_{z=1}^{\infty} \mu_z^{**-1/2z} = \infty, \tag{4.6}$$

where $\mu_z^{**} = \frac{p(z,0,\ldots,0)a_{0,0,\ldots,0}}{a_{z,0,\ldots,0}p(0,\ldots,0)} \frac{1}{k_1^z}$, $z = 0, 1, \ldots$, then the mixture (4.4) is identifiable.

Note 2. (i) It is a simple exercise, in view of Theorem 4.1 and Remark 5, to check that arbitrary mixtures of multiple (p independent laws) Poisson or negative multinomial or multivariate logarithmic series distributions are identifiable. Obviously, Theorem 4.1 subsumes the cases of finite and countable mixtures.

(ii) According to Theorem 10 of Patil (1965), the mixture (4.4) is always identifiable because, in the notation of his theorem, (4.4) is a θ-mixture of multivariate GPSD$(k_1\theta,\ldots,k_p\theta)$ with range T_n such that $n = \infty$. However, the claim needs to be justified; to have the validity of the "if" part of his theorem, one has to show that the moment sequence constructed by him determines the distribution when $n = \infty$. However, if one assumes the validity of (4.6) his result holds as seen in Remark 5.

The next theorem gives a multivariate extension of Theorem 2.2.

THEOREM 4.2. *Let* $\boldsymbol{X} = (X_1,\ldots,X_p)$ *be a p-dimensional random vector with the multivariate power-series mixture distribution defined by* (4.1). *Assume that*

$$p(x_1,\ldots,x_p \mid \theta_1,\ldots,\theta_p) = \frac{a_{x_1,\ldots,x_p}\theta_1^{x_1}\cdots\theta_p^{x_p}}{A(\theta_1,\ldots,\theta_p)}, \qquad x_i = 0,1,\ldots, \qquad i = 1,\ldots,p$$

has infinitely divisible (possibly shifted) univariate marginals for $\theta_i = \theta_{0i} \in I$, $i = 1,\ldots,p$. *Then the mixture* (4.1) *is identifiable.*

PROOF. If the marginals are infinitely divisible then, following essentially the argument in the proof of Theorem 2.2, we see that the moment sequences corresponding to the marginals of the mixing distribution have m.g.f.'s with zero as an interior point of the corresponding domains of definitions, and hence it follows that the moment sequence of the mixing distribution satisfies the Cramer-Wold condition (in obvious notation)

$$\sum_{x=1}^{\infty} (\mu_{x,0,\ldots,0} + \mu_{0,x,0,\ldots,0} + \mu_{0,\ldots,0,x})^{-1/2x} = \infty.$$

Hence it follows that this moment sequence determines the corresponding mixing distribution. This implies that the mixture (4.1) is identifiable.

Remark 6. The assumption that a multivariate distribution has infinitely divisible univariate marginals is weaker than that the distribution itself is infinitely divisible. We support our claim with the following example: assume that a_{x_1,x_2} in the bivariate power-series distribution is given by

$$a_{x_1,x_2} = \begin{cases} e^{-\lambda}\dfrac{\lambda^{x_1}}{x_1!}e^{-\mu}\dfrac{\mu^{x_2}}{x_2!} + \alpha, & \text{if } (x_1,x_2) = (0,1) \text{ or } (1,0) \\ e^{-\lambda}\dfrac{\lambda^{x_1}}{x_1!}e^{-\mu}\dfrac{\mu^{x_2}}{x_2!} - \alpha, & \text{if } (x_1,x_2) = (0,0) \text{ or } (1,1) \\ e^{-\lambda}\dfrac{\lambda^{x_1}}{x_1!}e^{-\mu}\dfrac{\mu^{x_2}}{x_2!}, & \text{otherwise,} \end{cases}$$

where α is positive and less than $\min\{e^{-(\lambda+\mu)}, e^{-(\lambda+\mu)}\lambda\mu\}$. The bivariate p.g.f. $G(t_1,t_2)$ (at $(\theta_1,\theta_2) = (1,1)$) is not equal to $e^{-\lambda(1-t_1)-\mu(1-t_2)}$ but the limit as $t_1, t_2 \to \infty$ for the ratio of these two p.g.f.'s equals to 1. As $G(0,0) > 0$, we can then claim that the bivariate distribution that we have constructed is non-infinitely divisible. However, as the marginals in this case are Poisson, they are infinitely divisible.

We turn now to a characterization based on regression functions. Theorem 4.3 given below is a multivariate extension of Theorem 3.1. Incidentally, Cacoullos (1987) showed that when $\boldsymbol{X} = (X_1,\dots,X_p)$ is random vector from a regular p-parameter exponential family and its continuous mixture with respect to parameter vector is identifiable, the posterior expectation vector uniquely determines the distributions of $\boldsymbol{\Theta}$.

THEOREM 4.3. *Let $\boldsymbol{X} = (X_1,\dots,X_p)$ be a p-dimensional random vector with the multivariate power-series mixture distribution defined by* (4.1). *Also, let the regression functions*

$$m_i(x_{(i)}) = E(\theta_i \mid X_{(i)} = x_{(i)}), \qquad i = 1,\dots,p,$$

where $x_{(i)} = (x_1,\dots,x_i,0,\dots,0)$, $i = 1,\dots,p$, be such that

$$\sum_{x=1}^{\infty}\left(\prod_{i=0}^{x-1} m_1(i,0,\dots,0) + \prod_{i=0}^{x-1} m_2(0,i,0,\dots,0) + \cdots + \prod_{i=0}^{x-1} m_p(0,\dots,0,i)\right)^{-1/2x} = \infty. \tag{4.7}$$

Then, the distribution of $\boldsymbol{\Theta}$ is uniquely determined by m_i, $i = 1,\dots,p$.

PROOF. In view of the assumptions of the theorem, it is not difficult to see that

$$\frac{p(x_1,0,\ldots,0)}{a_{x_1,0,\ldots,0}} = \frac{p(0,\ldots,0)}{a_{0,\ldots,0}} \prod_{i_1=0}^{x_1-1} m_1(i_1,0,\ldots,0), \qquad x_1 = 1,2,\ldots$$

$$\frac{p(x_1,x_2,0,\ldots,0)}{a_{x_1,x_2,0,\ldots,0}} = \frac{p(x_1,0,\ldots,0)}{a_{x_1,0,\ldots,0}} \prod_{i_2=0}^{x_2-1} m_2(x_1,i_2,0,\ldots,0), \qquad x_2 = 1,2,\ldots$$

$$\vdots$$

$$\frac{p(x_1,\ldots,x_p)}{a_{x_1,\ldots,x_p}} = \frac{p(x_1,\ldots,x_{p-1},0)}{a_{x_1,\ldots,x_{p-1}}} \prod_{i_p=0}^{x_p-1} m_p(x_1,\ldots,x_{p-1},i_p), \qquad x_p = 1,2,\ldots,$$

where $p(x_1,\ldots,x_p)$ is as defined in (4.1). From the above, by making use of (4.2) and Theorem 4.1, we conclude that the distribution of $\boldsymbol{\Theta}$ is uniquely determined by m_i, $i = 1,2,\ldots,p$. (This, in turn, implies that the joint distribution of $\boldsymbol{X}$ and $\boldsymbol{\Theta}$ is uniquely determined by m_i, $i = 1,2,\ldots,p$.) This completes the proof of the theorem.

Remark 7. It is easy to check that (4.7) is met if

$$\limsup_{x\to\infty}\left(\frac{m_1(x,0,\ldots,0)}{x} + \frac{m_2(0,x,0,\ldots,0)}{x} + \cdots + \frac{m_p(0,\ldots,0,x)}{x}\right) < \infty;$$

consequently, we have that Theorem 4.3 holds under this stronger condition in place of (4.7).

The following corollary of Theorem 4.3 is used as an illustrative example.

COROLLARY 4.1. *Let* $\boldsymbol{X} = (X_1,\ldots,X_p) \mid \boldsymbol{\Theta} = (\theta_1,\ldots,\theta_p) \sim$ *Negative Multinomial* $(r,\theta_1,\ldots,\theta_p)$. *Then for some* $\alpha_i > 0$, $i = 0,1,\ldots,p$, $\alpha = \sum_{i=0}^{p}\alpha_i$,

$$m_i(x_{(i)}) = \frac{x_i+\alpha_i}{x_0+r+\alpha}, \qquad x_i = 0,1,\ldots, \qquad i = 1,\ldots,p,$$

where $x_0 = \sum_{i=1}^{p} x_i$ *if and only if* $\boldsymbol{X} \sim$ *compound Negative Multinomial* $(r;\alpha_0,\alpha_1,\ldots,\alpha_p)$ (*see Mossimann* (1963)); *then* $\boldsymbol{\Theta} \sim$ *Dirichlet* $(\alpha_0,\ldots,\alpha_p)$.

Acknowledgements

I am greatly indebted to Professor D. N. Shanbhag (Sheffield University, U.K.) for constructive suggestions and comments. I am also grateful to referees for their useful suggestions which led to a substantial improvement in the presentation of the paper.

References

Akhiezer, N. I. (1965). *The Classical Moment Problem*, Oliver & Boyd, Edinburgh.

Barndörff-Nielsen, O. (1965). Identifiability of mixtures of exponential families, *J. Math. Anal. Appl.*, **12**, 115–121.

Cacoullos, T. (1987). Characterizing priors by posterior expectations in multiparameter exponential family, *Ann. Inst. Statist. Math.*, **39**, 399–405.

Diaconis, P. and Ylvisaker, D. (1979). Conjugate priors for exponential families, *Ann. Statist.*, **7**, 269–281.

Ericson, W. A. (1969). A note on the posterior mean of a population mean, *J. Roy. Statist. Soc. Ser. B*, **31**, 332–334.

Feller, W. (1968). *An Introduction to Probability Theory and its Applications*, Vol. 1, 3rd ed., Wiley, New York.

Fraser, M. D., Hsu, Yu-Sheng and Walker, J. J. (1981). Identifiabilities of finite mixtures of Von-Mises distribution, *Ann. Statist.*, **9**, 1130–1134.

Johnson, N. L. (1957). Uniqueness of a result in the theory of accident proneness, *Biometrika*, **44**, 530–531.

Johnson, N. L. (1967). Note on a uniqueness of a result in certain accident proneness model, *J. Amer. Statist. Assoc.*, **62**, 288–289.

Johnson, N. L. and Kotz, S. (1989). Characterization based on conditional distributions, *Ann. Inst. Statist. Math.*, **41**, 13–17.

Kent, J. T. (1983). Identifiability of finite mixtures for directional data, *Ann. Statist.*, **11**, 984–988.

Lüxmann-Ellinghaus, U. (1987). On the identifiability of mixtures of infinitely divisible power-series distributions, *Statist. Probab. Lett.*, **5**, 375–378.

Mossiman, J. E. (1963). On the compound negative multinomial distribution and correlations among inversely sampled pollen counts, *Biometrika*, **50**, 47–54.

Patil, G. P. (1965). Certain characteristic properties of multivariate discrete probability distributions akin to the Bates-Neyman model in the theory of accident proneness, *Sankhyā Ser. A*, **27**, 259–270.

Patil, G. P. and Bildikar (Talwalker), S. (1966). Identifiability of countable mixtures of discrete probability distributions using methods of infinite matrices, *Proc. Camb. Phil. Soc.*, **62**, 485–494.

Prakasa Rao, B. L. S. (1992). *Identifiability in Stochastic Models: Characterization of Probability Distributions*, Academic Press, London.

Sankaran, M. (1970). The discrete Poisson-Lindley distribution, *Biometrics*, **26**, 145–149.

Shohat, J. A. and Tamarkin, J. D. (1963). *The Problem of Moments*, 2nd ed., American Mathematical Society, New York.

Sibuya, M. (1979). Generalized hypergeometric, digamma and trigamma distributions, *Ann. Inst. Statist. Math.*, **31**, 373–390.

Singh, H. and Vasudeva, H. L. (1984). A characterization of exponential distribution by Yule distribution, *J. Indian Statist. Assoc.*, **22**, 93–96.

Tallis, G. M. (1969). The identifiability of mixtures of distributions, *J. Appl. Probab.*, **6**, 389–398.

Tallis, G. M. and Chesson, P. (1982). Identifiability of mixtures, *J. Austral. Math. Soc. Ser. A*, **32**, 339–348.

Teicher, H. (1960). On the mixture of distributions, *Ann. Math. Statist.*, **31**, 55–73.

Teicher, H. (1961). Identifiability of mixtures, *Ann. Math. Statist.*, **32**, 244–248.

Teicher, H. (1963). Identifiability of finite mixtures, *Ann. Math. Statist.*, **34**, 1265–1269.

Titterington, D. M., Smith, A. F. M. and Makov, U. E. (1985). *Statistical Analysis of Finite Mixtures Distributions*, Wiley, New York.

Yakowitz, S. J. and Spragins, J. D. (1968). On the identifiability of finite mixtures, *Ann. Math. Statist.*, **39**, 209–214.