Ann Inst Stat Math (2011) 63:291–303
DOI 10.1007/s10463-009-0221-9

Mixtures of power series distributions: identifiability via uniqueness in problems of moments

Jordan Stoyanov · Gwo Dong Lin

Received: 2 April 2008 / Revised: 8 September 2008 / Published online: 10 March 2009

Abstract We treat the identifiability problem for mixtures involving power series distributions. Applying an idea of Sapatinas (Ann Inst Stat Math 47:447–459, 1995) we prove and elaborate that a mixture distribution is identifiable if a certain Stieltjes problem of moments has a unique solution while a non-uniqueness leads to a non-identifiable mixture. We describe explicitly models of identifiable mixtures and models of non-identifiable mixtures. Illustrative examples and comments on related questions are also given.

Keywords Power series distributions · Mixtures of distributions · Identifiability · Non-identifiability · Stieltjes problem of moments · Uniqueness · Non-uniqueness

1 Introduction

The problem of identifiability of mixture distributions has received a considerable attention from both theorists and practitioners over the last decades. Among the literature in this area we mention here the books by Everitt and Hand (1981), Titterington et al. (1985), Prakasa Rao (1992), Lindsay (1995) and McLachlan and Peel (2000).

In this paper we study mixtures of power series distributions. We are looking for conditions under which such a mixture is identifiable or non-identifiable. One specific feature of mixtures of power series distributions is that the identifiability problem is

J. Stoyanov (✉)
School of Mathematics and Statistics, Newcastle University, Newcastle upon Tyne NE1 7RU, UK
e-mail: jordan.stoyanov@ncl.ac.uk

G. D. Lin
Institute of Statistical Science, Academia Sinica, Taipei 11529, Taiwan, ROC
e-mail: gdlin@stat.sinica.edu.tw

associated with the uniqueness in an appropriate Stieltjes problem of moments while non-identifiability is associated with such a problem when the solution is non-unique. We use some recent developments in the classical problem of moments. In particular we exploit ideas and techniques from Slud (1993), Lin (1992, 1997), Stoyanov (1997, 2000), Pakes (2001), Pakes et al. (2001) and Gut (2002).

There are different approaches when treating identifiability problems, see for example Barndorff-Nielsen (1965), Lüxmann-Ellinghaus (1987), Prakasa Rao (1992), Sapatinas (1995), Wesolowski (1995), Papageorgiou and Wesolowski (1997), Gupta et al. (2001), Holzmann et al. (2004) and Böhning and Patilea (2005). Each approach has its own merits thus allowing to study the identifiability of mixtures from a different point of view.

The present paper is organized as follows. In Sect. 2 we recall some basic notions and formulate the identifiability problem. The main results and their proofs are given in Sects. 3 and 4. Illustrative examples are exhibited in Sect. 5 while brief comments on related topics are outlined in Sect. 6.

2 Basic notions and formulation of the problem

Suppose (X, θ) is a two-dimensional random vector defined on some underlying probability space $(\Omega, \mathscr{F}, P)$. We assume that the random variable X takes values in the set $\mathbb{N}_0 := \{0, 1, \dots, k, \dots\}$ of all non-negative integers or in some its finite or infinite subset $\mathbb{N}_0^*$. The set of values of the random variable θ is denoted by T and $T \subset [0, \infty)$.

Let $F_t = \mathscr{L}(X|\theta = t)$ be the conditional distribution of X given that $\{\theta = t\}$, where for each $t \in T$, F_t is a *power series distribution* defined by

$$F_t = \{f(k|t),\ k = 0, 1, 2, \dots\} \quad \text{and} \quad f(k|t) := P[X = k|\theta = t] = \frac{a_k\, t^k}{A(t)}. \qquad (1)$$

Here a_k, $k = 0, 1, \dots$, are non-negative real numbers and

$$A(t) = \sum_{n=0}^{\infty} a_n\, t^n$$

is the *power series function* on $[0, \infty)$ corresponding to the distribution F_t. One of our assumptions is that the infinite series $A(t)$ has a strictly positive radius of convergence ρ_A, $0 < \rho_A \leq \infty$. The case $\rho_A = \infty$ provides most possibilities, while the case $\rho_A = 0$ is trivial and of no interest.

Among the power series distributions are the Binomial, Poisson, Negative-binomial, Logarithmic and their several modifications. Later on, in Sect. 5, we will use some of these distributions to illustrate our results.

We assume that θ has a distribution function $G(t)$, $t \in T = [0, \rho_A)$. In what follows G can be arbitrary, discrete or absolutely continuous, with bounded or unbounded support.

Now in a standard way we write down the unconditional distribution $\mathscr{L}(X)$ of the random variable X and say that we "mix" the distributions of the family

$F = \{F_t,\ t \in T\}$ with respect to the *mixing distribution* G. Thus we obtain the *mixture distribution* $H = \mathscr{L}(X)$, where

$$H = \{h_k,\ k = 0, 1, 2, \ldots\} \quad \text{with} \quad h_k = P[X = k] = \int_T f(k|t)\,\mathrm{d}G(t). \tag{2}$$

Suppose that the family $F = \{F_t,\ t \in T\}$ of conditional distributions, with F_t as defined by (1), is fixed. Relation (2) leads to the question: how are the mixing distribution G and the mixture H related to each other?

If, given F, there is only one mixing distribution G on T, producing H, we say that the mixture distribution H is *identifiable*.

If, however, there are at least two distributions on T, say G_1 and G_2, such that

$$h_k = \int_T f(k|t)\,\mathrm{d}G_1(t) = \int_T f(k|t)\,\mathrm{d}G_2(t) \quad \text{for all} \quad k = 0, 1, \ldots, \quad \text{but} \quad G_1 \neq G_2,$$

then the mixture distribution $H = \{h_k\}$ is *non-identifiable*.

The meaning of the above is as usual: if the mixture H is identifiable, we can recover *uniquely* the mixing distribution G, while G cannot be recovered uniquely if H is non-identifiable.

If the random vector (X, θ) is considered as a stochastic model of some phenomenon, it is clear then from the above setting that the model is described by the triplet (F, G, H), where $F = \{F_t,\ t \in T\}$ is the set of conditional distributions $F_t = \mathscr{L}(X|\theta = t)$ (this is equivalent to knowing the power series function A), $G = \mathscr{L}(\theta)$ is the mixing distribution and $H = \mathscr{L}(X)$ is the resulting mixture distribution. Since a part of the available information is expressed in terms of mixture distributions, for the pair (X, θ) one can use the name *mixture model*. Thus, if the mixture H is identifiable, we say also that the mixture model (X, θ) is identifiable. Otherwise, if H is non-identifiable, the model is non-identifiable.

We are looking for conditions, hopefully easy to check, which guarantee identifiability or non-identifiability of mixture distributions.

3 Main results and their proofs

First we describe a step which is crucial for what follows next. One of our assumptions is that the numbers a_k in (1), and hence the coefficients of the series $A(t) = \sum_{k=0}^{\infty} a_k t^k$, are non-negative and that $A(t)$ has a radius of convergence $\rho_A \in (0, \infty]$. If an index $k \in \mathbb{N}_0$ is such that $a_k > 0$, this means that the number k is an atom of the distribution F_t of the random variable X conditionally on $\{\theta = t\}$ for any t and also unconditionally. We assume always that $a_0 > 0$, but in general, some of the remaining coefficients a_k may be zero; we analyze such a case in Theorem 1 below. For now, unless specified otherwise, we assume that all coefficients of $A(t)$ are strictly positive, i.e., $a_0 > 0, a_1 > 0, \ldots, a_k > 0, \ldots$. This property of a_k's together with relation (2) immediately implies that h_k are also strictly positive and this is true for any mixing distribution G. As in Sapatinas (1995), let us define a "new" function, say $\tilde{G}$, as follows:

$$\mathrm{d}\tilde{G}(t) := \frac{a_0}{h_0}\frac{1}{A(t)}\,\mathrm{d}G(t), \quad t \in T = [0, \rho_A). \tag{3}$$

It is easy to check that $\tilde{G} = (\tilde{G}(t),\ t \in T)$ is a distribution function. Hence, there is a random variable, say $\tilde{\theta}$, whose distribution is $\tilde{G}$, i.e., $\tilde{\theta} \sim \tilde{G}$. Denoting by $\tilde{m}_k$ the kth order moment of $\tilde{\theta}$, and also of $\tilde{G}$, we find that

$$\tilde{m}_k := E[\tilde{\theta}^k] = \int_T u^k\,\mathrm{d}\tilde{G}(u) = \frac{a_0}{h_0}\frac{h_k}{a_k} \quad \text{for } k = 1, 2, \dots. \tag{4}$$

Therefore, $\tilde{\theta}$ has finite moments $\tilde{m}_k$ of all positive orders and, moreover, relation (4) tells us how the moment sequence $\{\tilde{m}_k\}$ is explicitly expressed in terms of $\{a_k\}$ and $\{h_k\}$.

In the sequel we use the term "M-determinate" or "M-indeterminate" to indicate that a distribution is, or is not, determined uniquely by its moments.

It is clear from the above setting that the problem of identifiability is directly related to a certain Stieltjes problem of moments. We formulate this as follows.

Proposition 1 (a) *Given the power series function $A(t)$, $t \in [0, \rho_A)$, the distribution function $\tilde{G}$ determines uniquely the mixing distribution G, and vice versa.*

(b) *The identifiability of the mixture H, and hence of the model (X, θ), is equivalent to the moment determinacy of the distribution $\tilde{G}$.*

Let us start with a result which is based on classical arguments. It helps us in describing conditions under which a mixture distribution is identifiable or non-identifiable.

Theorem 1 (a) (Non-identifiable mixtures) *Assume that for any $t \in [0, \infty)$ the conditional distribution $F_t = \mathscr{L}(X|\theta = t)$ is a* ***finite*** *power series distribution which is defined by the power series function $A(t) = a_0 + a_1 t + \cdots + a_k t^k + \cdots + a_{n^*} t^{n^*}$. Here n^* is a fixed positive integer and the coefficients $a_0, a_1, \dots, a_{n^*}$ are positive; all other coefficients in (1) are equal to zero. Then for any mixing distribution G, the mixture H is non-identifiable.*

(b) (Identifiable mixtures) *Suppose that $F_t = \mathscr{L}(X|\theta = t),\ t \in [0, \infty)$, is an infinite power series distribution based on the power series function $A(t)$ whose radius of convergence is ρ_A, $0 < \rho_A \le \infty$. Let the mixing distribution G have a* ***bounded*** *support,* $\mathrm{supp}(G) \subset [0, \rho_A)$. *Then the mixture distribution H is identifiable.*

(c) (More identifiable mixtures) *Consider the model (X^*, θ), where $\mathscr{L}(X^*|\theta = t)$ is an infinite power series distribution based on the power series function $A^*(t)$ which is defined as follows. Assume that in (1) strictly positive are infinitely many of the coefficients, but not necessarily all: $a_k > 0$ only for indices $k \in \mathbb{N}_0^*$, where $\mathbb{N}_0^*$ is an infinite set of integers:*

$$\mathbb{N}_0^* := \{0 = k_0 < k_1 < k_2 < \cdots\} \subset \mathbb{N}_0 \quad \textit{with} \quad \sum_{j=1}^{\infty} \frac{1}{k_j} = \infty.$$

Thus, the power series function is $A^*(t) := \sum_{k\in\mathbb{N}_0^*} a_k\, t^k$ *and let its radius of convergence be* ρ_{A^*}, $0 < \rho_{A^*} \le \infty$. *Let finally* G *be an arbitrary mixing distribution with a bounded support,* $\mathrm{supp}(G) \subset [0, \rho_{A^*})$. *Then the mixture* H *is identifiable.*

Proof (a) Since the values of the random variable X are in the set $\mathbb{N}_0^* = \{0, 1, \ldots, n^*\}$ which is finite, the power series function $A(t) = \sum_{k=0}^{n^*} a_k\, t^k$ is well-defined for any real t, so its radius of convergence is $\rho_A = \infty$. Hence we can take any mixing distribution G with bounded or unbounded support and use relation (2) to obtain the mixture H. Thus we have in our disposal two finite sets of numbers, $\{a_0, a_1, \ldots, a_{n^*}\}$ and $\{h_0, h_1, \ldots, h_{n^*}\}$.

We look now at the distribution function $\tilde{G}$, see (3). The random variable $\tilde{\theta} \sim \tilde{G}$ has moments $\tilde{m}_k$ up to order n^*, which are expressed in terms of a_k's and h_k's by (4). Hence, all we know about $\tilde{G}$ are the first n^* moments $\{\tilde{m}_1, \tilde{m}_2, \ldots, \tilde{m}_{n^*}\}$. In such a case, however, there are infinitely many distributions, all different from $\tilde{G}$, and all having the same moments $\tilde{m}_k,\ k = 1, \ldots, n^*$. Therefore the mixture distribution H is non-identifiable.

(b) Since $F_t = \mathscr{L}(X|\theta = t)$ is an *infinite* power series distribution, we may think that X takes values in the set $\mathbb{N}_0$ and let $A(t)$, the power series function, have a radius of convergence ρ_A, $0 < \rho_A \le \infty$. One of the assumptions is that the mixing distribution $G = \mathscr{L}(\theta)$ has a bounded support, $\mathrm{supp}(G) \subset [0, \rho_A)$. Hence the distribution $\tilde{G}$ has also a bounded support $\mathrm{supp}(\tilde{G})$. In this case $\tilde{\theta}$, and also $\tilde{G}$, has finite all moments $\tilde{m}_k = (a_0 h_k)/(h_0 a_k)$, $k = 1, 2, \ldots$. Moreover, the function $\tilde{M}(u) = E[\exp(u\tilde{\theta})]$ is well-defined for any real u, hence $\tilde{\theta}$ obeys a moment generating function. This, see Feller (1971), implies that $\tilde{G}$ is M-determinate, i.e., $\tilde{G}$ is the only distribution with the moment sequence $\{\tilde{m}_k\}$. Hence, because of (3), the mixing distribution G itself is uniquely determined which means, by definition, that the mixture H is identifiable.

(c) We need now the “new” distribution function $\tilde{G}^*$ defined by

$$\mathrm{d}\tilde{G}^*(t) := \frac{a_0}{h_0}\, \frac{1}{A^*(t)}\, \mathrm{d}G(t), \quad t \in [0, \rho_{A^*}).$$

Note that the support of $\tilde{G}^*$ is also bounded because the function A^* is bounded away from zero. Let us point out that the conclusion about the moment determinacy of the distribution $\tilde{G}^*$ follows from Lemma 1 below, which is based on the well-known Müntz–Szász theorem, or its refinement, see Lin (1992, 1993) and the references therein. □

Remark 1 The requirement $\sum_{j=1}^{\infty} (k_j)^{-1} = \infty$ in Theorem 1(c) is essential. We can take $\mathbb{N}_0^* = \mathbb{N}_0$, the case treated first, or $\mathbb{N}_0^*$ to be the set of all odd integer numbers, or the set of all even numbers, or even the set of all prime numbers. However, we cannot take $\mathbb{N}_0^*$ to consist of all perfect squares.

Remark 2 Let us note that in the literature there are several and different proofs of the following classical result: a mixture of the Binomial distribution (appropriately parameterized) is non-identifiable, see Tallis (1969) or Titterington et al. (1985). The above arguments, which confirm the same statement, are based on the fact that a distribution cannot be characterized by knowing only a finite number of its moments.

Lemma 1 *Let Y be a random variable obeying a distribution G on the bounded interval $[a, b]$ with $a \geq 0$. Assume that $\{t_n\}_{n=1}^{\infty}$ is a sequence of positive, distinct real numbers satisfying the condition $\sum_{n=1}^{\infty}(t_n)^{-1} = \infty$ and is bounded away from zero. Then the distribution G is characterized by the sequence of moments $\{E(Y^{t_n})\}_{n=1}^{\infty}$.*

The cases treated in Theorem 1 show that interesting to study are mixture models involving *infinite* power series distributions when the mixing distributions have an *unbounded* support, say $[0, \infty)$.

Theorem 2 *Suppose (X, θ) is a mixture model, where $F_t = \mathscr{L}(X|\theta = t)$ is an infinite power series distribution defined by (1) with all a_k's strictly positive and such that the power series function $A(t) = \sum_{k=0}^{\infty} a_k\, t^k$ has a radius of convergence $\rho_A = \infty$. If $G(t)$, $t \in T = [0, \infty)$, is the mixing distribution, we obtain the mixture $H = \{h_k\}$ by relation (2).*

(a) *Suppose that the following condition is satisfied:*

$$\sum_{k=1}^{\infty}\left(\frac{a_k}{h_k}\right)^{1/2k} = \infty. \tag{5}$$

Then the mixture distribution H is identifiable.

(b) *Suppose that the following condition is satisfied:*

$$\sum_{k=1}^{\infty}\left(\frac{a_k}{h_k}\right)^{1/2k} < \infty. \tag{6}$$

Assume further that G is absolutely continuous with density g and that the ratio-function $g(t)/A(t)$, $t \geq 0$, is ultimately log-concave. Under these conditions the mixture distribution H is non-identifiable.

Proof Part (a) is one of the results in Sapatinas (1995). The idea is to examine the distribution $\tilde{G}$ and the *Carleman quantity* $\tilde{\mathrm{C}}$ calculated for the moment sequence $\{\tilde{m}_k\}$: $\tilde{\mathrm{C}} = \mathrm{C}[\{\tilde{m}_k\}] := \sum_{k=1}^{\infty}(\tilde{m}_k)^{-1/2k}$. Condition (5) easily implies that $\tilde{\mathrm{C}} = \infty$. Hence, see Feller (1971) or Stoyanov (1997), the distribution $\tilde{G}$ is M-determinate. Since $\tilde{G}$ determines uniquely the mixing distribution G, then in view of Proposition 1, the mixture H is identifiable.

For part (b), we see that the distribution function $\tilde{G}$ is absolutely continuous with density $\tilde{g}(t) = \tilde{c}\, g(t)/A(t), t \geq 0$, where $\tilde{c}$ is a normalizing constant. We now observe that even for arbitrary G, condition (6) leads to a finite Carleman quantity calculated for the moments $\{\tilde{m}_k\}$ of $\tilde{G}$, i.e., $\mathrm{C}[\{\tilde{m}_k\}] < \infty$. Moreover, the condition on the ratio $g(t)/A(t)$ yields that the density $\tilde{g}$ is ultimately log-concave. Thus we are in a position to apply Theorem 1 of Pakes (2001) and conclude that the distribution $\tilde{G}$ is M-indeterminate. As a consequence, there are "many" (even infinitely many) choices for G and they all produce the same mixture H, so H is non-identifiable. □

The next two results cover mixture models in which the mixing distribution G is assumed to be absolutely continuous. We need some notations.

For any (strictly) positive density $p(x)$, $x \in (0, \infty)$, we define the following normalized logarithmic integral which is called *Krein quantity*:

$$\mathrm{K}[p] := \int_0^\infty \frac{-\ln p(x^2)}{1+x^2}\,\mathrm{d}x.$$

It can be shown that always $\mathrm{K}[p] \geq -\frac{1}{2}$ and for the "value" of K there are two possibilities, namely

$$\mathrm{K}[p] = \infty \quad \text{or} \quad \mathrm{K}[p] < \infty.$$

The "value" of the Krein quantity $\mathrm{K}[p]$ is directly related to the uniqueness and non-uniqueness of a distribution with density p in terms of the moments, see Slud (1993), Lin (1997), Stoyanov (1997, 2000), Pakes (2001), Pakes et al. (2001) and Gut (2002).

For a smooth and positive function $b(x)$, $x \in (0, \infty)$, we define the following function which is called *Lin function*:

$$\mathrm{L}[b(x)] := \frac{-x\, b'(x)}{b(x)}, \quad x \in (0, \infty).$$

Let us present two results which involve the Krein quantity K and the Lin function L and elaborate the use of Proposition 1.

Theorem 3 *Assume that the mixing distribution G is absolutely continuous and has density g, which is positive. Let further g and the power series function A be such that the Krein quantity, calculated for the ratio g/A, is finite, that is,*

$$\mathrm{K}[g/A] = \int_0^\infty \frac{-\ln\left[g(x^2)/A(x^2)\right]}{1+x^2}\,\mathrm{d}x < \infty. \tag{7}$$

Then the mixture distribution H is non-identifiable.

Proof Since $\tilde{g}(t) = (a_0/h_0)\, g(t)/A(t)$ for $t > 0$, the above condition (7) and elementary properties of integrals together imply that the Krein quantity $\mathrm{K}[\tilde{g}]$ is finite: $\mathrm{K}[\tilde{g}] < \infty$. Hence, see Slud (1993) or Lin (1997), the distribution $\tilde{G}$ is M-indeterminate, i.e., $\tilde{G}$ is not determined uniquely by its moment sequence. Therefore, by Proposition 1(b), the mixture distribution H is non-identifiable. □

Theorem 4 *Let (X, θ) be the mixture model as described above and let g be the density of G. Assume further that the ratio g/A has a positive right-hand limit at* 0 *and that g and A are such that the Krein quantity* K *calculated for the ratio g/A is divergent:*

$$\mathrm{K}[g/A] = \int_0^\infty \frac{-\ln\left[g(x^2)/A(x^2)\right]}{1+x^2}\,\mathrm{d}x = \infty \tag{8}$$

and for some fixed $x_0 > 0$,

$$\mathrm{L}[g(x)/A(x)] \nearrow \infty \quad \textit{as} \quad x_0 \leq x \to \infty.$$

Under these conditions the mixture distribution H is identifiable.

Proof Condition (8) implies that the density $\tilde{g}$ has a divergent Krein quantity: $\mathrm{K}[\tilde{g}] = \infty$. We combine this with the Lin condition and refer to results from Lin (1997) and Pakes (2001) to conclude that the distribution $\tilde{G}$ with density $\tilde{g}$ is M-determinate. This, according to Proposition 1, guarantees the identifiability of the mixture H. □

4 Mixture models of an exponential type

Let us analyze a mixture model (X, θ) assuming that the mixing distribution G and the power series function A have a special form, in particular, that they both are of an exponential type.

We use the standard notation $RV(\rho)$, where $\rho \in (-\infty, \infty)$, for the class of measurable and positive functions $u(t), t \in [0, \infty)$, which are regularly varying at infinity with exponent ρ. Recall that $u \in RV(\rho)$ means that $u(tx)/u(t) \to x^\rho$ as $t \to \infty$ for all $x > 0$. If the exponent is $\rho = 0$, then $RV(0)$, usually denoted by SV, is called the class of slowly varying functions. For details see Bingham et al. (1989), Feller (1971) or Gut (2002).

Assumption 1 Let $v_1 = (v_1(t),\ t \geq 0)$ be a positive function such that $v_1 \in RV(\rho_1)$ for some $\rho_1 > 0$. Define the power series function $A = (A(t),\ t \geq 0)$ to be $A(t) = \exp[v_1(t)], t \geq 0$. We also assume that $A(t)$ is an infinitely differentiable function with the Taylor/Maclaurin expansion $A(t) = \sum_{k=0}^{\infty} a_k\, t^k$ converging for any $t \geq 0$; hence the radius of convergence of $A(t)$ is $\rho_A = \infty$. Finally, we assume that all coefficients $a_k, k = 0, 1, 2, \ldots$, are strictly positive.

Assumption 2 Suppose that $G = (G(t),\ t \geq 0)$ is an absolutely continuous distribution function with density g of the following exponential type:

$$g(t) = c\, u(t) \exp[-v(t)], \quad t > 0.$$

Here u and v are positive functions such that $u \in RV(\rho_0)$ and $v \in RV(\rho_2)$ for some $\rho_0 > -1$ and $\rho_2 > 0$; $c > 0$ is a normalizing constant.

We use the functions A and g as described in Assumptions 1 and 2 to define the mixture model (X, θ), equivalently the triplet (F, G, H). As in Sect. 2, $F_t = \mathscr{L}(X|\theta = t)$, the infinite power series distribution, is defined by (1). With $G = \mathscr{L}(\theta)$ as a mixing distribution, we use relation (2) and obtain the mixture H.

Theorem 5 *Suppose that (X, θ) is a mixture model such that F, G and H are defined as above. Let ρ_1 and ρ_2 be the exponents introduced in Assumptions 1 and 2, respectively. Then we have the following:*

(a) *if* $\max\{\rho_1, \rho_2\} > \frac{1}{2}$, *the mixture distribution H is identifiable;*
(b) *if* $\max\{\rho_1, \rho_2\} < \frac{1}{2}$, *the mixture distribution H is non-identifiable.*

Proof With A, ρ_1, g and ρ_2 as in Assumptions 1 and 2, we consider the density function $\tilde{g}(t) := (a_0/h_0)\, g(t)/A(t)$, $t > 0$. Let $\tilde{G}$ be the corresponding distribution function and $\tilde{\theta}$ a random variable such that $\tilde{\theta} \sim \tilde{G}$. It is seen that $\tilde{g}(t) = \tilde{c}\, u(t)\, \exp[-\tilde{v}(t)]$, $t > 0$, where $\tilde{v}(t) = v(t) + v_1(t)$ and $\tilde{c}$ is the normalizing constant. Moreover, $\tilde{v} \in RV(\tilde{\rho})$ with

$$\tilde{\rho} = \max\{\rho_1, \rho_2\}.$$

Since $\rho_1 > 0$ and $\rho_2 > 0$, the random variable $\tilde{\theta}$ has all moments finite. If $\tilde{\rho} \geq 1$, $\tilde{\theta}$ has a moment generating function and hence the distribution $\tilde{G}$ is M-determinate. In this case the conclusion is clear; Proposition 1(b) tells us that the mixture H is identifiable. If $\tilde{\rho} < 1$, there is no moment generating function, so $\tilde{G}$ has a heavy tail (as $t \to \infty$), but it is still possible $\tilde{G}$ to be M-determinate, see, e.g., Stoyanov (1997). We need a result, see Lemma 2 below, which is essentially due to Pakes et al. (2001) and Gut (2002). These authors give quite general conditions under which a distribution is M-determinate or is M-indeterminate. In both cases, however, the value $\tilde{\rho} = \frac{1}{2}$ is critical. More precisely, if $\tilde{\rho} > \frac{1}{2}$, the distribution $\tilde{G}$ is M-determinate, while for $\tilde{\rho} < \frac{1}{2}$, $\tilde{G}$ is M-indeterminate. Therefore, by Proposition 1, we conclude that the mixture distribution H is identifiable for $\tilde{\rho} > \frac{1}{2}$ and non-identifiable for $\tilde{\rho} < \frac{1}{2}$. □

Lemma 2 *Let $G = (G(t),\ t \geq 0)$ be an absolutely continuous distribution function whose density g is as in Assumption 2. Then:*

(a) *G is M-determinate if $\rho_2 > \frac{1}{2}$;*
(b) *G is M-indeterminate if $\rho_2 < \frac{1}{2}$.*

Proof It is well-known that decisive for the moment determinacy of a distribution, in our case G, is the behavior at infinity of its tail $1 - G(t)$. This is why we involve the behavior of the density $g(t)$ for large t. Clearly, $g(t)$ may depend on both regularly varying functions from some class $RV(\rho)$ and slowly varying functions from the class $SV = RV(0)$. Recall that if $\ell \in SV$ and $\varepsilon > 0$, then we have $t^{-\varepsilon}\, \ell(t) \to 0$ and $t^{\varepsilon}\, \ell(t) \to \infty$ as $t \to \infty$ [see page 16 in Bingham et al. (1989)]. Combining this fact with our assumptions and referring to the discussion in Pakes et al. (2001), see Theorem 1, page 108, we arrive at claim (a). Finally, claim (b) is one of the results in Gut (2002), namely Theorem 6.1(a). □

5 Illustrative examples

Example 1 Consider a mixture model (X, θ), where $F_t = \mathscr{L}(X|\theta = t)$ is the Poisson distribution with parameter $t > 0$. This is an infinite power series distribution with $a_k = 1/k!$, the power series function is $A(t) = e^t$ and its radius of convergence is $\rho_A = \infty$. Now assume that the random variable θ has a gamma distribution, $\theta \sim \gamma(a, b)$, for some $a > 0$, $b > 0$, so this will be the mixing distribution G. It is well-known that the unconditional distribution of X, i.e., the mixture $H = \{h_k\}$, is the negative binomial distribution.

Thus we ask the converse question: if the result of mixing a Poisson distribution is the negative binomial distribution, is it true that the gamma distribution is the only mixing distribution? The answer is "yes" and it follows from Theorem 2.

Indeed, first we use formula (2) and find explicitly the mixture H:

$$H = \{h_k,\ k = 0, 1, \ldots\}, \quad \text{where} \quad h_k = \frac{b^a\,\Gamma(k+a)}{k!(1+b)^{k+a}\Gamma(a)}.$$

The next step is to write the ratio a_k/h_k and analyze the behavior of $(a_k/h_k)^{1/2k}$. By using properties of the gamma function and the Stirling formula we arrive at the divergence of the series in (5). Hence the model (X, θ) is identifiable. In other words, gamma distribution is the only mixing distribution in this case.

Let us mention that Lüxmann-Ellinghaus (1987) has given another proof of the identifiability in this case by using the infinite divisibility property of the Poisson distribution. And yet another proof of the same statement is given by Sapatinas (1995).

Example 2 Following Assumptions 1 and 2, let us take the power series function $A = (A(t), t \geq 0)$ and the mixing density $g = (g(t), t > 0)$ to be of the form:

$$A(t) = c_1 \exp[t^{\rho_1}], \quad t \geq 0 \quad \text{and} \quad g(t) = c_2\, t^{\rho_0} \exp[-t^{\rho_2}], \quad t > 0,$$

for some $\rho_1 > 0$, $\rho_0 > -1$ and $\rho_2 > 0$. We follow the same procedure as described before and define the mixture model (X, θ) and its triplet (F, G, H). The identifiability of the model depends on the values of ρ_1 and ρ_2.

The identifiability of gamma-mixture of the Poisson distribution, see Example 1, is not surprising in view of Theorem 5. Simply, in this case $A(t) = e^t$, so $\rho_1 = 1$. Moreover, $\rho_0 = a - 1$ for some $a > 0$, $\rho_2 = 1$ and hence $\tilde{\rho} \geq 1$.

On the other hand, for a fixed power series function $A(t) = c_1 \exp[t^{\rho_1}]$ with exponent $\rho_1 \in \left(0, \frac{1}{2}\right)$, the mixture distribution H is identifiable or non-identifiable according to whether the exponent (of the mixing density g) is $\rho_2 > \frac{1}{2}$ or $\rho_2 < \frac{1}{2}$.

Example 3 How to deal with mixture models (X, θ) involving the exponents $\rho_1 = \frac{1}{2}$ and $\rho_2 = \frac{1}{2}$? Clearly this boundary case is not covered by Theorem 5.

Let us consider a power series distribution $F_t = \{f(k|t)\}$ defined as follows:

$$f(k|t) = P[X = k|\theta = t] = \frac{1}{(2k)!}\,\frac{t^k}{A(t)}, \quad k = 0, 1, 2, \ldots.$$

The random variable X takes values in the set $\mathbb{N}_0$ and we find explicitly the following compact expression for the function $A(t)$:

$$A(t) = \frac{1}{2}\left(e^{\sqrt{t}} + e^{-\sqrt{t}}\right) = \frac{1}{2}\,\frac{e^t + 1}{e^{\sqrt{t}}}, \quad t \geq 0.$$

Let us take the mixing distribution G with the density g given by

$$g(t) = \frac{1}{2}\, e^{-\sqrt{t}}, \quad t > 0.$$

It is clear from the expressions for the functions A and g that indeed $\rho_1 = \frac{1}{2}$, $\rho_2 = \frac{1}{2}$. Theorem 5 cannot characterize the identifiability of the mixture H. Our conclusion will be based on the properties of the density $\tilde{g}(t)$. We use the Krein–Lin techniques. First, we easily check that the Krein quantity for $\tilde{g}$ is divergent: $\mathrm{K}[\tilde{g}] = \infty$. Thus we need the behavior as $t \to \infty$ of the Lin function $\mathrm{L}[\tilde{g}(t)]$. Since here $\mathrm{L}[\tilde{g}(t)] = (\ln 2)^{-1} t\, e^t/(e^t + 1)$, we easily see that this function is ultimately monotone and diverging to infinity, i.e., the Lin condition is satisfied. Hence the distribution $\tilde{G}$ is M-determinate, and by Proposition 1, the mixture H is identifiable.

Example 4 If we want to consider a mixture model with $\rho_1 = \frac{1}{4}$ and $\rho_2 = \frac{1}{3}$, we can take the power series distribution $F_t = \mathscr{L}(X|\theta = t) = \{f(k|t)\}$ to be given by

$$f(k|t) = P[X = k|\theta = t] = \frac{1}{(4k)!}\,\frac{t^k}{A(t)}, \quad k = 0, 1, 2, \ldots,$$

where

$$A(t) = \frac{1}{4}\left(e^{t^{1/4}} + e^{-t^{1/4}}\right) + \frac{1}{2}\cos\left(t^{1/4}\right), \quad t > 0.$$

This explicit expression for the power series function $A(t)$ was found by using Formula 5.2.7.10 from Prudnikov et al. (1992).

Suppose now that the mixing distribution G has the following density g:

$$g(t) = c\, e^{-t^{1/3}}, \quad t > 0 \quad (c \text{ is a normalizing constant}).$$

From the expressions for A and g we have indeed that $\rho_1 = \frac{1}{4}$, $\rho_2 = \frac{1}{3}$.

The next step is to see that the density $\tilde{g}$ of the distribution $\tilde{G}$ has the following tail behavior:

$$\tilde{g}(t) \sim \tilde{c}\left(e^{-t^{1/3}} + e^{-t^{1/4}}\right) \quad \text{for large } t.$$

There are different ways to find the answer to the question of whether the mixture H obtained from these A and G is identifiable or non-identifiable.

We are in a good position to apply Theorem 5(b). Crucial is that in this case $\tilde{\rho} = \max\{\rho_1, \rho_2\} = \frac{1}{3} < \frac{1}{2}$. Even without referring to this Theorem, we see that a random variable with density $\tilde{g}(t) = \tilde{c}\, g(t)/A(t)$, where g and A are given as above, has all moments finite. In order to establish the moment indeterminacy of the distribution $\tilde{G}$ whose density is $\tilde{g}$, we examine the Krein quantity $\mathrm{K}[\tilde{g}]$. We easily find that $\mathrm{K}[\tilde{g}] < \infty$. Therefore, see Slud (1993), Lin (1997) or Stoyanov (1997, 2000), the distribution $\tilde{G}$ is M-indeterminate, and in view of Proposition 1, the mixture distribution H, based on A and G in this case, is non-identifiable.

Another approach is to try to apply Theorem 2(b). We know the coefficients a_k and use formula (2) to calculate h_k. Even though we cannot find explicit expressions for h_k, we can use some properties of the gamma function and the Stirling formula and show the convergence of the series $\sum_{k=1}^{\infty} (a_k/h_k)^{1/2k}$ and hence the finiteness of the

Carleman quantity. Then we have to check that the density $\tilde{g}(t) = \tilde{c}g(t)/A(t)$, $t \geq 0$, is ultimately log-concave. Thus the non-identifiability of the mixture H will follow from Theorem 2(b). The details are left to the reader.

Example 5 Here is a case which is "close" to Example 4. We consider the same power series distribution F_t as given above, and instead of the mixing density $g(t) = c\, e^{-t^{1/3}}$, now we take the following one:

$$g(t) = \frac{1}{4}\, t^{-3/4}\, e^{-t^{1/4}}, \quad t > 0.$$

The mixture model (X, θ) is well defined and if we repeat the same reasoning as above, we arrive at the conclusion that in this case the mixture distribution H is non-identifiable. Perhaps the reader can recognize that $g(t) = \frac{1}{4}\, t^{-3/4}\, e^{-t^{1/4}}$, $t > 0$, is exactly the density of the random variable ξ^4, where $\xi \sim \mathrm{Exp}(1)$. The distribution with density g is M-indeterminate, see, e.g., Stoyanov (1997, 2000). The non-identifiability of the mixture H is a consequence of the "slow" exponential growing rate of the power series function $A(t)$ as $t \to \infty$. Here the rate is equal to $1/4$. If the rate were 1, as for the Poisson distribution, under the same mixing density g (the density of ξ^4), the mixture distribution H will be identifiable.

6 Comments on related topics

(a) When studying mixture models with non-identifiable mixture distributions H, we can make further one essential step. Namely, we are in a position to write explicitly "many" (in fact infinitely many) mixing distributions, a whole family G_ε, $\varepsilon \in [-1, 1]$ such that any G_ε produces the same mixture $H = \{h_k\}$. For this purpose we have to construct a Stieltjes class for the M-indeterminate distribution $\tilde{G}$, see Stoyanov (2004) and Stoyanov and Tolmatz (2005). We do not provide details here.

(b) In our recent paper, see Lin and Stoyanov (2002), we have considered a model of the type $(\tilde{S}, N)$, where N is an integer-valued positive random variable and $\tilde{S}$ is the sum of N independent copies of a given random variable X. Our goal in that paper was to study the moment determinacy of the compound sum $\tilde{S}$, which is different from the questions discussed in the present paper. These two questions, the moment determinacy and the identifiability of mixtures, can be studied simultaneously.

Acknowledgments This study was started some time ago and completed when one of the authors (JS) was visiting the Institute of Statistical Science, Academia Sinica, Taipei. The support for the visit and the hospitality of the colleagues at Academia Sinica are greatly appreciated. We are grateful to Prof. Samuel Kotz for the interest in this study and for his useful comments. We would like to thank the two anonymous referees and the editors for their useful constructive suggestions which were taken into account when revising the paper.

References

Barndorff-Nielsen, O. (1965). Identifiability of mixtures of exponential families. *Journal of Mathematical Analysis and Applications*, *12*, 115–121.

Bingham, N., Goldie, C., Teugels, J. (1989). *Regular variation*. Cambridge: Cambridge University Press.

Böhning, D., Patilea, V. (2005). Asymptotic normality in mixtures of power series distributions. *Scandinavian Journal of Statistics*, *32*, 115–131.

Everitt, B. S., Hand, D. J. (1981). *Finite mixture distributions*. London: Chapman & Hall.

Feller, F. (1971). *An introduction to probability theory and its applications* (Vol. 2, 2nd ed.). New York: Wiley.

Gupta, A. K., Nguyen, T. T., Wang, Y., Wesolowski, J. (2001). Identifiability of modified power series mixtures via posterior means. *Journal of Multivariate Analysis*, *77*, 163–174. Erratum, ibid, 80(2002), 189.

Gut, A. (2002). On the moment problem. *Bernoulli*, *8*, 407–421.

Holzmann, H., Munk, A., Stratmann, B. (2004). Identifiability of finite mixtures—with applications to circular distributions. *Sankhya*, *66*, 440–449.

Lin, G. D. (1992). Characterizations of distributions via moments. *Sankhya*, *A54*, 128–132.

Lin, G. D. (1993). Characterizations of the exponential distribution via the blocking time in a queueing system. *Statistica Sinica*, *3*, 577–581.

Lin, G. D. (1997). On the moment problems. *Statistics and Probability Letters*, *35*, 85–90. Erratum, ibid, 50 (2000), 205.

Lin, G. D., Stoyanov, J. (2002). On the moment determinacy of the distributions of compound geometric sums. *Journal of Applied Probability*, *39*, 545–554.

Lindsay, B. G. (1995). *Mixture models: theory, geometry and applications*. Hayward: Institute of Mathematical Statistics.

Lüxmann-Ellinghaus, U. (1987). On the identifiability of mixtures of infinitely divisible power series distributions. *Statistics and Probability Letters*, *5*, 375–378.

McLachlan, G., Peel, D. (2000). *Finite mixture models*. New York: Wiley.

Pakes, A. (2001). Remarks on converse Carleman and Krein criteria for the classical moment problem. *Journal of the Australian Mathematical Society*, *71*, 81–104.

Pakes, A., Hung, W.-L., Wu, J.-W. (2001). Criteria for the unique determination of probability distributions by moments. *Australian and New Zealand Journal of Statistics*, *43*, 101–111.

Papageorgiou, H., Wesolowski, J. (1997). Posterior mean identifies the prior distribution in NB and related models. *Statistics and Probability Letters*, *36*, 127–134.

Prakasa Rao, B. L. S. (1992). *Identifiability in stochastic models: characterization of probability distributions*. London: Academic Press.

Prudnikov, A. P., Brychkov, Yu. A., Marichev, O. I. (1992). *Integrals and series* (Vol. 1). New York: Gordon & Breach.

Sapatinas, T. (1995). Identifiability of mixtures of power-series distributions and related characterizations. *Annals of the Institute of Statistical Mathematics*, *47*, 447–459.

Slud, E. V. (1993). The moment problem for polynomial forms in normal random variables. *Annals of Probability*, *21*, 2200–2214.

Stoyanov, J. (1997). *Counterexamples in probability* (2nd ed.). Chichester: Wiley.

Stoyanov, J. (2000). Krein condition in probabilistic moment problems. *Bernoulli*, *6*, 939–949.

Stoyanov, J. (2004). Stieltjes classes for moment-indeterminate probability distributions. *Journal of Applied Probability*, *41A*, 281–294.

Stoyanov, J., Tolmatz, L. (2005). Method for constructing Stieltjes classes for M-indeterminate probability distributions. *Applied Mathematics and Computation*, *165*, 669–685.

Tallis, G. M. (1969). The identifiability of mixtures of distributions. *Journal of Applied Probability*, *6*, 389–398.

Titterington, D. M., Smith, A. F. M., Makov, U. E. (1985). *Statistical analysis of finite mixtures distributions*. New York: Wiley.

Wesolowski, J. (1995). Bivariate discrete measures via a power series conditional distributions and a regression function. *Journal of Multivariate Analysis*, *55*, 219–229.