Ann. Inst. Statist. Math.
Vol. 40, No. 2, 261–265 (1988)

IDENTIFIABILITY OF FINITE MIXTURES USING A NEW TRANSFORM

KHALAF E. AHMAD

Mathematics Department, Faculty of Sciences, Assiut University, Assiut, Egypt

(Received March 17, 1986)

Abstract. Identifiability of finite mixtures of the following families of distributions are proved: Weibull, normal log, chi, pareto and power function.

Key words and phrases: Identifiability, mixtures, Weibull, lognormal, chi, pareto, power function.

1. Introduction

Let $f=\{F(x;\ \alpha),\ \alpha \in R_1^m\}$ constitute a family of one-dimensional cumulative distribution function (c.d.f.'s) indexed by a point α in a Borel subset R_1^m of Euclidean m-space R^m such that $F(x; \alpha)$ is measurable in $R^1 \times R_1^m$. Then, the one-dimensional c.d.f.

$$H(x) = \int_{R_1^m} F(x;\ \alpha) dG(\alpha) , \tag{1.1}$$

is the image of the above mapping, say Q, of the n-dimensional c.d.f.'s G. In the integral of (1.1) H, F and G are called mixture, kernel and mixings d.f.'s, respectively. H is called a finite mixture if its mixing distribution or rather the corresponding measure is discrete and doles out positive mass to only a finite number of points in R_1^m. Thus the set H of all finite mixtures of a class f of distributions is the convex hull of f:

$$H = \left\{ H(x)\text{: } H(x) = \sum_{i=1}^{N} C_i F(x;\ \alpha_i),\ C_i > 0, \right. \\ \left. \sum_{i=1}^{N} C_i = 1,\ F(x;\ \alpha_i) \in f,\ N = 1, 2,\ldots \right\} \tag{1.2}$$

In the context of finite mixtures, the definition of "identifiable" implies f generates identifiable finite mixtures if and only if the convex hull of f has the uniqueness of representation property:

$$\sum_{i=1}^{N} C_i F_i = \sum_{i=1}^{M} C_i' F_i' ,$$

implies $N=M$ and for each i, $1 \le i \le N$, there is some j, $1 \le j \le N$, such that $C_i = C_i'$ and $F_i = F_i'$.

Clearly, identifiability questions must be settled before one can meaningfully discuss the problem of estimating the mixing c.d.f. G on the basis of observations from the mixture H. So, estimation of the mixing distribution G, testing hypotheses about the mixing distribution G, etc., can be meaningfully discussed only if the family of mixing distributions is identifiable. Discussions of identifiability of mixtures may be found in several papers, among others, by Teicher (1961, 1963, 1967), Yakowitz and Spragins (1968), Mohanty (1972), Rennie (1972), Blum and Susarla (1977), Al-Hussaini and Ahmad (1981), Fraser *et al.* (1981), Ahmad and Al-Hussaini (1982) and Kent (1983).

Teicher (1963), Yakowitz and Spragins (1968), Mohanty (1972) and Al-Hussaini and Ahmad (1981) used moment generating function, characteristic function and Laplace transforms in Theorem (2) of Teicher (1963), which is stated with weaker assumption by Chandra (1977) as the required transforms. Sometimes, we cannot find any of these transforms for some distribution function. So, we can use the moment generating function of $\log(X)$ as the required transforms.

A graphical procedure for estimation of mixed Weibull parameters in life-testing of electron tubes is given by Kao (1959). Falls (1970) represents an attempt at estimating by the method of sample moments, the five parameters of the compound Weibull distribution. However, I know of no published work giving the discussions of identifiability of the mixture of Weibull distributions.

In this paper I show the finite mixtures of Weibull, lognormal, chi, pareto and power function distributions are identifiable using the moment generating function of $\log(X)$, as the required transforms in Theorem (2.4) of Chandra (1977).

THEOREM 1.1. (Chandra (1977)) *Let there be associated with each $P_i \in \Phi$ a transform ϕ_i having the domain of definition D_{ϕ_i}, and suppose that the mapping M: $P_i \to \phi_i$ is linear. Suppose also that there exists a total ordering $(\le)$ of Φ such that*

(i) *$P_1 \le P_2$ (P_1, $P_2 \in \Phi$) implies $D_{\phi_1} \subseteq D_{\phi_2}$*

(ii) *for each $P_1 \in \Phi$, there exists some t_1 in the closure of $T_1 = \{t: \phi_1(t) \ne 0\}$ such that $\lim_{\substack{t \to t_1 \\ t \in T_1}} (\phi_2(t)/\phi_1(t)) = 0$, for each $P_1 < P_2$ (P_1, $P_2 \in \Phi$).*

Then the class Λ of all finite mixing distributions is identifiable relative to Φ.

PROPOSITION 1.1. *The classes of all finite mixtures of the families of the following distributions are identifiable*:

(1) *Weibull,*
(2) *Lognormal,*
(3) *Chi,*
(4) *Pareto,*
(5) *Power function.*

PROOF. Justifications of the proposition is immediate by applying the above theorem to the kernel probability function (kpf) (1)–(5) and making use of Table 1 in the next page. The kernel probability functions are listed together with the corresponding transforms as follows:

(1) kpf:

$$f(x;\, \alpha, c) = \frac{c}{\alpha}\left(\frac{x}{\alpha}\right)^{c-1} \exp\left[-\left(\frac{x}{\alpha}\right)^{c}\right], \quad c > 0, \quad \alpha > 0, \quad x > 0.$$

Transform: Moment generating function of log (X)

$$\phi(t) = \left(\frac{1}{\alpha}\right)^{t} \Gamma\left(\frac{t}{c} + 1\right).$$

(2) kpf:

$$f(x;\, \theta, \sigma) = \frac{1}{\sigma x \sqrt{2\pi}} \exp\left[-\frac{1}{2\sigma^2}(\log x - \theta)^2\right],$$
$$x > 0, \quad -\infty < \theta < \infty, \quad \sigma > 0\,.$$

Transform: Moment generating function of log (X)

$$\phi(t) = \exp\left[t\theta + t^2\sigma^2/2\right].$$

(3) kpf:

$$f(x;\, n) = \frac{x^{n-1} e^{-x^2/2}}{2^{(n/2)-1}\Gamma\left(\frac{n}{2}\right)}, \quad x \geq 0, \quad n \geq 0\,.$$

Transform: Moment generating function of log (X)

$$\phi(t) = \Gamma\left(\frac{n+t}{2}\right) 2^{t/2} \Big/ \Gamma\left(\frac{n}{2}\right).$$

(4) kpf:

$$f(x;\,\alpha,\,\theta)=\frac{\theta}{x}\left(\frac{\alpha}{x}\right)^{\theta},\quad x\geq\alpha,\quad \theta>0\,.$$

Transform: Moment generating function of log (X)

$$\phi(\theta)=\frac{\theta\alpha^{t}}{\theta-t},\quad \theta>t\,.$$

(5) kpf:

$$f(x;\,c,\,\theta)=\frac{c}{x}\left(\frac{x}{\theta}\right)^{c},\quad 0<x<\theta\,,\quad \theta>0\,,\quad c>0\,.$$

Transform: Moment generating function of log (X)

$$\phi(t)=\theta^{t}\left(\frac{c}{c+t}\right).$$

Finally we order each of the five families lexicographically by an ordering such that the two conditions of the theorem are satisfied and thus the identifiability of the finite mixtures is established. The ordering, domains of ϕ_1 and ϕ_2, and the value of t are listed for the five kpf's in the Table 1.

Table 1.

kpf	$P_1<P_2$ implies			D_{ϕ_1}	D_{ϕ_2}	t
(1)	$\alpha_1<\alpha_2$	and	$c_1=c_2$	$(-c_1,\infty)$	$(-c_2,\infty)$	∞
(2)	$\theta_1<\theta_2$	and	$\alpha_1=\alpha_2$	$(-\infty,\infty)$	$(-\infty,\infty)$	∞
(3)	$n_1<n_2$			$(-n_1,\infty)$	$(-n_2,\infty)$	$-n_1$
(4)	$\alpha_1<\alpha_2$	and	$\theta_1=\theta_2$	$(-\infty,\theta_1)$	$(-\infty,\theta_2)$	$-\infty$
(5)	$\theta_1>\theta_2$	and	$c_1=c_2$	$(-\infty,\infty)$	$(-\infty,\infty)$	∞

REFERENCES

Ahmad, K. E. and Al-Hussaini, E. K. (1982). Remarks on the non-identifiability of mixtures of distributions, *Ann. Inst. Statist. Math.*, **34**, 543–544.

Al-Hussaini, E. K. and Ahmad, K. E. (1981). On the identifiability of finite mixtures of distributions, *IEEE Trans. Inform. Theory*, **27**, 664–668.

Blum, J. R. and Susarla, V. (1977). Estimation of a mixing distribution, *Ann. Probab.*, **5**, 201–209.

Chandra, S. (1977). On the mixtures of probability distributions, *Scand. J. Statist.*, **4**, 105–112.
Falls, L. W. (1970). Estimation of parameters in compound Weibull distributions, *Technometrics*, **12**, 399–407.
Fraser, M. D., Hsu, Y. S. and Walker, J. J. (1981). Identifiability of finite mixtures of von Mises distributions, *Ann. Statist.*, **9**, 1130–1131.
Kao, J. H. K. (1959). A graphical estimation of mixed Weibull parameters in life-testing electron tubes, *Technometrics*, **1**, 389–407.
Kent, J. T. (1983). Identifiability of finite mixtures for directional data, *Ann. Statist.*, **11**, 984–988.
Mohanty, N. C. (1972). On the identifiability of finite mixtures of Laguerre distributions, *IEEE Trans. Inform. Theory*, **18**, 514–515.
Rennie, R. R. (1972). On the interdependence of the identifiability of multivariate mixtures and the identifiability of the marginal mixtures, *Sankhyā Ser. A*, **34**, 449–452.
Teicher, H. (1961). Identifiability of mixtures, *Ann. Math. Statist.*, **33**, 244–248.
Teicher, H. (1963). Identifiability of finite mixtures, *Ann. Math. Statist.*, **34**, 1265–1269.
Teicher, H. (1967). Identifiability of product measures, *Ann. Math. Statist.*, **38**, 1300–1302.
Yakowitz, S. J. and Spragins, J. D. (1968). On the identifiability of finite mixtures, *Ann. Math. Statist.*, **39**, 209–214.