Ann. Inst. Statist. Math.
35 (1983), Part A, 355-363

COMPLETENESS AND SELF-DECOMPOSABILITY OF MIXTURES

M. H. ALAMATSAZ

(Received July 19, 1982; revised Dec. 16, 1982)

Summary

In this article, we first introduce the concept of strong completeness and then show that the mixture of every strongly complete distribution is complete if the mixing distribution is complete. This, in effect, reveals the completeness of several well-known mixtures. For instance, Xekalaki (1983, *Ann. Inst. Statist. Math.*, to appear) showed that the Univariate Generalized Waring Distribution is boundedly complete only relative to one of its three parameters. Now, as a consequence of our result, it follows that this distribution is actually complete relative to any of its parameters.

Self-decomposability of mixtures is also discussed here. It is shown that a mixture of self-decomposable distributions is not necessarily self-decomposable when the mixing distribution is self-decomposable. For a special case of Poisson mixture, however, the result is valid when the mixing distribution is continuous self-decomposable, a result due to Forst (1979, *Zeit. Wahrscheinlichkeitsth.*, **49**, 349–352).

1. Completeness of mixtures

It can be seen that mixtures have many uses in practical situations. Since the completeness is of importance in the theory of estimation and testing hypotheses, it appears to be worthwhile to know whether or not a mixture family is complete. In this section, the completeness property of mixtures, in general, will be investigated. It turns out that a large number of mixtures including those of exponential families are actually complete.

First we introduce a new concept, extending the usual notion of completeness. Suppose $\{F(\boldsymbol{y}|\boldsymbol{x}, \boldsymbol{\beta})\}$ is a multivariate family of distributions with vector parameters $\boldsymbol{x}$ and $\boldsymbol{\beta}$ having real components. We give the new concept by the following:

DEFINITION 1 (Strong completeness). A family $F=\{F(\boldsymbol{y}|\boldsymbol{x}, \boldsymbol{\beta}): \boldsymbol{x} \in \underline{\boldsymbol{X}}\}$, of distributions of n-component random vectors, is said to be strongly complete, relative to $\boldsymbol{x}$, if for every arbitrary function g: $R^n \to R$ satisfying

$$\mathrm{E}_{F(\cdot\,|\,\boldsymbol{x},\boldsymbol{\beta})}\{g(\boldsymbol{Y})\}=0 \qquad \text{for every } \boldsymbol{x} \text{ in some dense subset of } \underline{X},$$

we have

$$\mathrm{P}_{F(\cdot\,|\,\boldsymbol{x},\boldsymbol{\beta})}\{g(\boldsymbol{Y})=0\}=1 \qquad \text{for every } \boldsymbol{x}\in\underline{X}.$$

Obviously, strong completeness is a stronger property than completeness. Actually, every strongly complete family is complete as every set is dense in itself.

Using the completeness, under some mild assumptions, of exponential families (see, e.g., Lehmann [7]), we shall show in the following theorem that under appropriate assumptions exponential families are indeed strongly complete. The proof is given for the univariate families. The multivariate case follows in the same way.

THEOREM 1. *The exponential families (of distinct members for distinct θ) given by*

$$dF_\theta(x)=a(\theta)b(x)\exp\{c(\theta)d(x)\}\,d\mu(x) \tag{1}$$

with $c(\theta)$ as a continuous function of θ and the parameter space corresponding to θ as an open interval Θ are strongly complete.

PROOF. Distributions of the form (1) are known to be complete (see Lehmann [7]). Thus, it is sufficient to show that for an arbitrary function $g(\cdot)$ independent of θ,

$$\mathrm{E}_\theta(g(X))=0 \qquad \text{for every } \theta\in B \tag{2}$$

with B as a dense subset of Θ implies

$$\mathrm{E}_\theta(g(X))=0 \qquad \text{for all } \theta\in\Theta. \tag{3}$$

From (2) we have $\mathrm{E}_\theta(g^+(X))=\mathrm{E}_\theta(g^-(X))<\infty$, i.e.,

$$a(\theta)\int|g(x)|b(x)\exp\{c(\theta)d(x)\}\,d\mu(x)<\infty \tag{4}$$

for every θ in B. Suppose $A=\{x\in R\colon d(x)>0\}$ and let $\theta^*\in\Theta$ be arbitrary. Choose a sequence $\{\theta_n\}$ in B tending to θ^*. Since B is a dense set, it is possible to choose θ' and θ'' in B such that $c(\theta_n)<c(\theta')$ and $c(\theta_n)>c(\theta'')$ for all n. In view of (4), then we have

$$\int_A|g(x)|b(x)\exp\{c(\theta')d(x)\}\,d\mu(x)<\infty$$

and

$$\int_{A^c} |g(x)| b(x) \exp\{c(\theta'')d(x)\} d\mu(x) < \infty .$$

Since $\exp\{c(\theta_n)d(x)\} < \exp\{c(\theta')d(x)\}$ for every x in A and $\exp\{c(\theta_n)\cdot d(x)\} > \exp\{c(\theta'')d(x)\}$ for every $x \in A^c$, for all n, in view of the continuity of $c(\theta)$, it follows by the dominated convergence theorem that

$$\lim_{n\to\infty} \mathrm{E}_{\theta_n}(g(X)) = \mathrm{E}_{\theta^*}(g(X)) = 0 .$$

Since θ^* was arbitrary this completes the proof.

As a result of this theorem it follows that the normal, binomial, negative binomial (and hence geometric), Poisson, beta, gamma (and hence exponential) and several other exponential families are in fact strongly complete.

Suppose $\{F(\boldsymbol{y} | \boldsymbol{x}, \boldsymbol{\beta})\}$ is a multivariate family of distributions with the vector parameter $\boldsymbol{x}$ such that it itself is a random vector with distribution $G(\boldsymbol{x} | \boldsymbol{\theta})$, where $\boldsymbol{\beta}$ and $\boldsymbol{\theta}$ are some vector parameters with real components. As in Bayesian inference, the mixture

$$H(\boldsymbol{y} | \boldsymbol{\theta}, \boldsymbol{\beta}) = \int_{\underline{X}} F(\boldsymbol{y} | \boldsymbol{x}, \boldsymbol{\beta}) dG(\boldsymbol{x} | \boldsymbol{\theta}) \tag{5}$$

is a family of multivariate distributions. $G(\boldsymbol{x} | \boldsymbol{\theta})$ in (5) is said to be the mixing distribution. Here, the sample space $\underline{X}$ is assumed to be such that every $\boldsymbol{x} \in \underline{X}$ is a support point of some $G(\cdot | \boldsymbol{\theta})$. Following Johnson and Kotz [6], we shall denote the mixture (1) by

$$F_Y(\cdot | \boldsymbol{x}, \boldsymbol{\beta}) \underset{\boldsymbol{x}}{\wedge} G_X(\cdot | \boldsymbol{\theta}) .$$

We shall now give the following theorem:

THEOREM 2. *The mixture of a strongly complete family is complete when the mixing distribution is complete.*

PROOF. Let $F = \{F(\boldsymbol{y} | \boldsymbol{x}, \boldsymbol{\beta}) : \boldsymbol{x} \in \underline{X}\}$ be a strongly complete family and $G = \{G(\boldsymbol{x} | \boldsymbol{\theta}) : \boldsymbol{\theta} \in \underline{\Theta}\}$ be complete. Consider $g : R^n \to R$ to be arbitrary function such that it is integrable with respect to the measure induced by $H(\boldsymbol{y} | \boldsymbol{\theta}, \boldsymbol{\beta})$ given in (5). We shall show that if

$$\mathrm{E}_{H(\cdot | \boldsymbol{\theta}, \boldsymbol{\beta})}\{g(\boldsymbol{Y})\} = 0 \qquad \text{for all } \boldsymbol{\theta} \in \underline{\Theta} , \tag{6a}$$

then

$$\mathrm{P}_{H(\cdot | \boldsymbol{\theta}, \boldsymbol{\beta})}\{g(\boldsymbol{Y}) = 0\} = 1 \qquad \text{for all } \boldsymbol{\theta} \in \underline{\Theta} . \tag{6b}$$

Let (6a) hold, i.e.

$$\int_{\underline{Y}} g(\boldsymbol{y}) \int_{\underline{X}} dF(\boldsymbol{y} | \boldsymbol{x}, \boldsymbol{\beta}) dG(\boldsymbol{x} | \boldsymbol{\theta}) = 0 \qquad \text{for all } \boldsymbol{\theta} \in \underline{\Theta} .$$

Applying Fubini's Theorem, we have

$$\int_{\underline{X}}\left\{\int_{\underline{Y}} g(\boldsymbol{y})dF(\boldsymbol{y}\,|\,\boldsymbol{x},\,\boldsymbol{\beta})\right\}dG(\boldsymbol{x}\,|\,\boldsymbol{\theta})=0 \qquad \text{for all } \boldsymbol{\theta}\in\underline{\Theta}\,.$$

Since G is a complete family, this implies that

$$\int_{\underline{Y}} g(\boldsymbol{y})dF(\boldsymbol{y}\,|\,\boldsymbol{x},\,\boldsymbol{\beta})=0 \tag{7}$$

almost surely for $\boldsymbol{x}$ with respect to $G(\cdot\,|\,\boldsymbol{\theta})$. It is easy to see that there exists then a dense subset B in $\underline{X}$ such that (7) holds for all $\boldsymbol{x}$ in B. Consequently, by the strong completeness of F, it follows in view of (7) that

$$\mathrm{P}_{F(\cdot\,|\,\boldsymbol{x},\boldsymbol{\beta})}\{g(\boldsymbol{Y})=0\}=1 \qquad \text{for all } \boldsymbol{x}\in\underline{X}\,. \tag{8}$$

It is now immediate that (8) implies (6b) and hence the proof is complete.

As we have seen, the exponential families are (under some mild assumptions) strongly complete and hence the above theorem directly implies

COROLLARY 1. *All mixtures of an exponential family are complete if the mixing distribution is complete.*

It is immediate then that distributions like: Pólya-Egenberger [Binomial $(N,p)\bigwedge_{p}$Beta], Poisson-Normal [Poisson $(\theta)\bigwedge_{\theta}$Truncated Normal (μ,σ)], the discrete Lognormal [Poisson $(\theta)\bigwedge_{\theta}$Lognormal (μ,σ,α)], Plank [mixtures of Gamma], etc. are all complete.

Recently, Xekalaki [11] was concerned with the completeness of the Univariate Generalized Waring Distribution (UGWD). She showed that a UGWD(a,k,p); $a>0$, $k>0$, $p>0$; is boundedly complete relative to p. Dr. D. N. Shanbhag has pointed out (in a private communication) that her method could be improved slightly to deal with the completeness of the family relative to p. However, as a result of the corollary above, it is evident that this distribution is actually complete relative to any of the parameters a, k, p. This is so because a UGWD (a,k,p) can be expressed as a mixture of the Poisson distribution with mean λ, where λ given m is a random variable (r.v.) with distribution $\gamma(k,m^{-1})$ for which m is also a r.v. with distribution $\beta(a,p)$ of type II (Irwin [5]). Thus by the application of Corollary 1, twice, we have that UGWD(a,k,p) is complete when either a or p is taken as the varying parameter. On the other hand, a r.v. Z corresponding to a UGWD(a,k,p) can be expressed as $Z\overset{d}{=}X_1X_2X_3^{-1}$ with X_i's independent and distributed as $\gamma(k, 1)$, $\gamma(a, 1)$ and $\gamma(p, 1)$ respectively (Sibuya [9]). In view of the symmetry in X_1 and X_2, it follows that completeness with respect to a implies that corresponding to k. Thus we have:

COROLLARY 2. *The family* $W=\{\mathrm{UGWD}(a, k, p)\colon a>0, k>0, p>0\}$ *is complete relative to any of the parameters* a, k, p.

Xekalaki has also been concerned with unimodality and self-decomposability of this distribution in her paper. It is hence of relevance here if we also mention that this distribution is a non-strongly unimodal discrete distribution. This follows since, as it can be trivially seen, the distribution is not logconcave (a lattice distribution $\{p_n\}$ is strongly unimodal if and only if it is logconcave, i.e., $p_n^2 \geqq p_{n+1}p_{n-1}$ for all n).

Let $g(\boldsymbol{x})$ be a one-to-one correspondence of $\boldsymbol{x}$ such that it is homeomorphism with values in R^m. It is clear that the distribution $F(\boldsymbol{y}|\boldsymbol{x}, \boldsymbol{\beta})$ is strongly complete relative to $\boldsymbol{x}$ if and only if the corresponding transformed distribution $F^*(\boldsymbol{y}|g(\boldsymbol{x}), \boldsymbol{\beta})$ is strongly complete relative to $g(\boldsymbol{x})$. Thus if $g(x)$ has a complete distribution $G(\cdot|\boldsymbol{\theta})$, it follows, from Theorem 1, that the mixture

$$F_{\underline{Y}}(\cdot|\boldsymbol{x}, \boldsymbol{\beta}) \underset{g(\boldsymbol{x})}{\wedge} G_{g(\boldsymbol{X})}(\cdot|\boldsymbol{\theta})$$

is also complete, relative to $\boldsymbol{\theta}$.

In view of this fact, it can be seen for example that, the Pareto distribution and the Contagious or Neyman's type A distribution are also complete. The Pareto distribution has the form: $\exp(\lambda) \underset{\lambda^{-1}}{\wedge} \text{Gamma}$, and the Contagious distribution is a mixture of the form: $\text{Poisson}(\theta) \underset{\theta/\phi}{\wedge} \text{Poisson}(\lambda)$ with $\phi>0$ as a constant.

Assuming that the measure induced by F is continuous in $\boldsymbol{x}$ for each $\boldsymbol{\beta}$, it is not difficult to see that the following theorem involving the notion of bounded completeness remains valid.

THEOREM 3. *The family H, as in Theorem 2, is boundedly complete if F and G are boundedly complete families. Moreover, if H is a boundedly complete family, then so is the family F.*

Note that in the second part of Theorem 3, there is no condition on G.

2. Self-decomposability of mixtures

A distribution function (d.f.) is said to be self-decomposable (s.d.) or of class L if its characteristic function (ch.f.) $\varphi(t)$ satisfies $\varphi(t)=\varphi(\alpha t)\cdot\varphi_\alpha(t)$ for all $\alpha \in (0, 1)$, where $\varphi_\alpha(t)$ is a ch.f. In the following we state, some well-known results in the literature.

LEMMA 1. *All non-degenerate s.d. distributions are absolutely continuous (see, e.g., Fisz and Varadarajan* [3]).

LEMMA 2. *A non-degenerate bounded r.v. can not be s.d. (see, e.g., Baxter and Shapiro* [1], *Chatterjee and Pakshirajan* [2] *and Ruegg* [8]).

LEMMA 3. *The convolution of two s.d. distributions is s.d.*

Steutel and van Harn [10] defined the analogue of self-decomposability for lattice distributions. They define a non-degenerate lattice distribution $\{p_n: n=0, 1, \cdots; p_0 \neq 0\}$ to be s.d. if its probability generating function (p.g.f.) $P(z)$ satisfies $P(z)=P(1-\alpha+\alpha z)P_\alpha(z)$; $|z|\leqq 1$; for all $\alpha \in (0, 1)$ with $P_\alpha(z)$ as a p.g.f. They have proved that this is equivalent to the assertion that $P(z)$ has the following form

$$(9) \qquad P(z)=\exp\left\{-\lambda\int_z^1 \frac{1-G(u)}{1-u}du\right\}; \qquad \lambda>0,\ |z|\leqq 1$$

with $G(u)$ as a p.g.f. such that $G(0)=0$.

As it is implicit from their proof of expression (9), this criterion can be restated in a simpler form as given in the following lemma:

LEMMA 4. *A lattice distribution is s.d. if and only if its p.g.f., $P(z)$ has the form*

$$(10) \qquad P(z)=\exp\left\{-\int_z^1 R(u)du\right\}; \qquad |z|\leqq 1$$

with $R(u)$ as the g.f. of a non-increasing sequence of non-negative real numbers.

Now let us go back to the mixtures. Consider, in general, $f(\cdot\,|x)$ to be a density function* of a continuous or a non-negative integer-valued lattice r.v. that is s.d. Note that by Lemma 1 any continuous s.d. distribution is absolutely continuous. Let then $h(\cdot)$ be the mixture of the s.d. densities $f(\cdot\,|x)$ defined as

$$(11) \qquad h(y)=\int_0^\infty f(y\,|x)dG(x)\,.$$

The question is whether or not self-decomposability of $G(x)$, the mixing distribution, implies that of $h(\cdot)$, the mixture. Forst [4] showed that if $f(\cdot\,|x)$ is a Poisson distribution, then the answer to this question is in the affirmative provided $G(x)$ is the d.f. of a continuous r.v.

* Observe that in the continuous case the density is with respect to Lebesgue measure and in the lattice case the density is with respect to counting measure.

In fact, he proved that a continuous distribution $G(x)$ is s.d. if and only if $\mathrm{P}_n=\int_0^\infty e^{-xt}[(xt)^n/n!]dG(x)$ is a lattice s.d. distribution for every $t>0$.

In what follows, we give three counterexamples in this connection illustrating that the result is not valid in general. Note that the normal and Gamma distributions are known to be s.d. and Poisson distributions are lattice s.d.

Example 1. (Illustrating that if $f(\cdot\,|x)$ and $G(\cdot)$ are both lattice s.d. then $h(\cdot)$ is not necessarily s.d.). Take $f(\cdot\,|x)$ to be Poisson (x) and $G(\cdot)$ to have the p.g.f. $Q(z)=\exp\{-\lambda(1+\beta-z-\beta z^2)\}$; $|z|\leqq 1$, $\lambda>0$ and $0<\beta\leqq 1/2$. Since $Q(z)$ can be written as in (10), with $R(u)$ being the g.f. of the sequence $\{\lambda, 2\beta\lambda\}$, by Lemma 4, $G(\cdot)$ is s.d. It is easy to see that the p.g.f. of $h(\cdot)$, $\mathrm{P}(z)$, has the form

$$\mathrm{P}(z)=\exp\{-\lambda(1+\beta-e^{-(1-z)}-\beta e^{-2(1-z)})\}; \qquad |z|\leqq 1,\ \lambda>0 \text{ and } 0<\beta\leqq 1/2\,.$$

By Lemma 4, $h(\cdot)$ is s.d. if and only if $R(u)$ satisfying

$$\int_z^1 R(u)du=\lambda\{1+\beta-e^{-(1-z)}-\beta e^{-2(1-z)}\} \tag{12}$$

is the g.f. of a non-increasing sequence of non-negative real numbers, $\{r_n\}_0^\infty$ say. From (12), it follows that

$$R(u)=\lambda\{e^{-(1-u)}+2\beta e^{-2(1-u)}\}\,, \qquad \lambda>0\,,$$

which is the g.f. of the sequence $\{r_n\}$ with

$$r_0=\lambda(e^{-1}+2\beta e^{-2}) \quad\text{and}\quad r_1=\lambda(e^{-1}+4\beta e^{-2})\,.$$

It is clear that $r_1>r_0$ and hence $h(\cdot)$ is not s.d.

Remark 1. It is easily seen that the result of Forst remains valid when "for every $t>0$" is replaced by "for a sequence $\{t_n;\ n=0, 1, 2, \cdots\}$ of positive numbers converging to infinity". By the virtue of this fact it follows that, generally, if $G(\cdot)$ is not a continuous s.d. distribution, there exists some $t_0>0$ such that for every $t\in[t_0, \infty)$, $p_n=\int_0^\infty e^{-xt}[(xt)^n/n!]dG(x)$; $t>0$, is a non-self-decomposable lattice distribution. In Example 1, for instance, if $f(\cdot\,|x)$ is taken to be Poisson (tx); $t>0$, t_0 can be taken to be any value greater than or equal to 1.

Example 2. (Illustrating that if $f(\cdot\,|x)$ is a s.d. density corresponding to a continuous r.v. and $G(\cdot)$ is lattice s.d., $h(\cdot)$ need not, in general, be s.d.). Take $f(\cdot\,|x)$ to be Normal $(0, x)$ and $G(\cdot)$ to be Poisson (λ). (Note that here the degenerate distribution for $x=0$ is taken as

Normal (0, 0)). Thus the ch.f. of $h(\cdot)$ will be

$$\begin{aligned}E(e^{itY}) &= \sum_{x=0}^{\infty} e^{-(xt^2)/2} e^{-\lambda} \lambda^x / x! \\ &= \exp\{-\lambda(1-e^{-t^2/2})\}\ ; \qquad -\infty < t < \infty,\ \lambda > 0\ .\end{aligned}$$

The distribution corresponding to this ch.f. is well-known to have a discontinuity at the origin and hence by Lemma 1 it can not be s.d., as required.

Example 3. (Illustrating that if $f(\cdot|x)$ and $G(\cdot)$ are both distributions corresponding to continuous s.d. r.v.'s, $h(\cdot)$ is not necessarily s.d.). Let Y_r and Y_r' be two independent r.v.'s with a common distribution $\gamma(r, \mu^{-1}r)$; $r>0$, $\mu>0$. Suppose $X_r = Z + Y_r'$, where Z is a non-degenerate non-negative s.d. r.v. It is clear that, since Y_r and Y_r' tend to μ in distribution as r tends to infinity, $X_r \xrightarrow{d} Z+\mu$ as $r\to\infty$. Hence, the r.v. Y_r/X_r tends to $\mu/(Z+\mu)$, in distribution, as $r\to\infty$. $\mu/(Z+\mu)$, the limit r.v., is bounded (between 0 and 1) and hence, by Lemma 2, it can not be s.d. As the limit of any sequence of s.d. distributions is s.d., this implies that at least for one r, r_0 say, Y_r/X_r is not s.d.

Now take $f(y|x)$ to be the density function of the r.v. Y_{r_0}/X_{r_0} given $X_{r_0}=x$, for any fixed x, and $G(x)$ to be the d.f. of X_{r_0}. Observe that $f(y|x)$ is s.d. because Y_{r_0} is a Gamma r.v. and $G(x)$ is s.d. by Lemma 3. Obviously then $h(y)$, in (11), will be the density function of the r.v. Y_{r_0}/X_{r_0} which is not s.d., as required.

Acknowledgments

The author is greatly indebted to Dr. Shanbhag, D. N. for his constructive suggestions and comments which improved the paper substantially in its content as well as presentation. My thanks are also due to the referee for his careful reading of the paper.

UNIVERSITY OF SHEFFIELD*

REFERENCES

[1] Baxter, G. and Shapiro, J. M. (1960). On bounded infinitely divisible random variables, *Sankhyā*, **17**, 253-260.

[2] Chatterjee, S. D. and Pakshirajan, R. P. (1956). On the unboundedness of infinitely divisible laws, *Sankhyā*, **17**, 349-350.

[3] Fisz, M. and Varadarajan, V. S. (1963). A condition for the absolute continuity of infinitely divisible distributions, *Zeit. Wahrscheinlichkeitsth.*, **1**, 335-339.

[4] Forst, G. (1979). A characterization of self-decomposable probabilities in the half-

* Present address: University of Isfahan, Iran.

line, *Zeit. Wahrscheinlichkeitsth.*, **49**, 349-352.

[5] Irwin, J. O. (1968). The generalized waring distribution applied to accident theory, *J. R. Statist. Soc.*, A, **131**, 205-225.

[6] Johnson, N. I. and Kotz, S. (1970). *Continuous Univariate Distributions* 1, Houghton Mifflin, Boston.

[7] Lehmann, E. L. (1964). *Testing Statistical Hypotheses*, Wiley, New York.

[8] Ruegg, A. (1970). A characterization of certain infinitely divisible laws, *Ann. Math. Statist.*, **41**, 1354-1356.

[9] Sibuya, M. (1979). Generalized hypergeometric, di-gamma and tri-gamma distributions, *Ann. Inst. Statist. Math.*, A, **31**, 373-390.

[10] Steutel, F. W. and van Harn, K. (1979). Discrete analogues of self-decomposability and stability, *Ann. Prob.*, **7**, 894-899.

[11] Xekalaki, E. (1983). Infinite divisibility, completeness and regression properties of the Univariate Generalized Waring Distribution, *Ann. Inst. Statist. Math.*, to appear.

Added in the proof

The non-strong unimodality of the UGW distributions mentioned in the paper also follows from the fact that this distribution does not have finite moments of all order. Further, an alternative counterexample to Example 1 is the Contagious or Neyman's type A distribution. This follows because of the fact that this distribution may have several modes while a s.d. lattice (or non-lattice) distribution is unimodal. For details of these, see Alamatsaz (1983, Ph. D. thesis, University of Sheffield).