## IDENTIFIABILITY OF MIXTURES OF PRODUCT MEASURES[1]

By Henry Teicher

*Purdue University*

For any integers $n, m \geqq 1$, let $\mathfrak{F}_{n,m} = \{F(x; \alpha): \alpha \varepsilon R_1^m, x \varepsilon R^n\}$ constitute a family of $n$-dimensional cdf's indexed by a point $\alpha$ in a Borel subset $R_1^m$ of Euclidean $m$-space $R^m$ such that $F(x; \alpha)$ is measurable in $R^n \times R_1^m$. Let $\mathcal{G}$ denote the class of $m$-dimensional cdf's $G$ whose induced measures $\mu_G$ assign measure one to $R_1^m$ and define $\mathcal{H} = \{H: H(x) = \int_{R_1^m} F(x; \alpha)\, dG(\alpha), G \varepsilon \mathcal{G}\}$. Then the class $\mathcal{H}$ of mixtures of $\mathfrak{F}$ is called identifiable if the mapping of $\mathcal{G}$ onto $\mathcal{H}$ is one-to-one. Any $H \varepsilon \mathcal{H}$ is termed a finite mixture if the entire mass of the corresponding $\mu_G$ is confined to a finite set of points of $R_1^m$.

Take $n = 1$ in the preceding, in which case we write $\mathfrak{F}_{1,m} = \mathfrak{F}$. Define for each $n \geqq 1$

$$\mathfrak{F}^*_{n,mn} = \{F^*(x; \alpha): F^*(x; \alpha) = \prod_{i=1}^n F(x_i, \alpha_i), F(x_i, \alpha_i) \varepsilon \mathfrak{F}, 1 \leqq i \leqq n\}$$

so that if $X_1, X_2, \cdots, X_n$ are independent random variables each of whose distributions is in $\mathfrak{F}$, their joint distribution is an element of $\mathfrak{F}^*_{n,mn}$. Since $\mathfrak{F}^*_{n,m\cdot n}$ is a particular version of $\mathfrak{F}_{n,m\cdot n}$ (with $R_1^{m\cdot n} = R_1^m \times R_1^m \times \cdots \times R_1^m$ and distributions corresponding to product measures) the class of all mixtures (or of all finite mixtures) of $\mathfrak{F}^*_{n,mn}$ is well defined as is the notion of identifiability of this class.

Theorem 1. *If the class of all mixtures of $\mathfrak{F}_{1,m}$ is identifiable, then for every $n > 1$, the class of mixtures of $\mathfrak{F}^*_{n,mn}$ is likewise identifiable. Conversely, if for some $n > 1$, the class of all mixtures of $\mathfrak{F}^*_{n,mn}$ is identifiable, the same is true for $\mathfrak{F}_{1,m}$.*

Proof. The converse is trivial since taking $F(x; \alpha) \varepsilon \mathfrak{F}_{1,m}$, if

$$\int F(x; \alpha)\, dG(\alpha) = \int F(x; \alpha)\, d\hat{G}(\alpha)$$

then multiplying both sides by $\prod_{i=1}^{n-1} F(x_i, \alpha_0)$ where $\alpha_0 \varepsilon R_1^m$, necessarily $I_{\alpha_0} \cdots I_{\alpha_0} \cdot G = I_{\alpha_0} \cdots I_{\alpha_0} \cdot \hat{G}$ and therefore $G = \hat{G}$. (As usual $I_{\alpha_0}$ is the one dimensional distribution with unit mass at $\alpha_0$).

To prove the first part of the theorem it suffices to show for fixed but arbitrary $n$ that if the class of all mixtures of $\mathfrak{F}^*_{n,m\cdot n}$ is identifiable, the same is true for $\mathfrak{F}^*_{n+1,m(n+1)}$.

Suppose then that for $F^* \varepsilon \mathfrak{F}^*_{n,mn}$, $F \varepsilon \mathfrak{F}_{1,m}$,

$$\int F^*(x; \alpha) F(y; \beta)\, dG(\alpha, \beta) \equiv_{x,y} \int F^*(x; \alpha) F(y; \beta)\, d\hat{G}(\alpha, \beta). \tag{1}$$

Let $G_2(\beta)$, $\hat{G}_2(\beta)$ denote the marginal distributions of $\beta$ corresponding to $G$ and $\hat{G}$; let $G(\alpha \mid \beta)$, $\hat{G}(\alpha \mid \beta)$ denote versions of the conditional probabilities such that for each $\beta$, $G(\alpha \mid \beta)$ and $\hat{G}(\alpha \mid \beta)$ are distribution functions in the variable $\alpha$ and

Received 17 November 1967.

[1] Research under National Science Foundation Contract GP-6073.

for each $\alpha$, $G(\alpha \mid \beta)$ and $\hat{G}(\alpha \mid \beta)$ are equal almost everywhere to measurable functions of $\beta$. Then (1) may be rewritten as

$$\int F(y;\beta)H(x;\beta)\,dG_2(\beta) \equiv_{x,y} \int F(y;\beta)\hat{H}(x;\beta)\,d\hat{G}_2(\beta) \tag{2}$$

where

$$\begin{aligned} H(x;\beta) &= \int F^*(x;\alpha)\,d_\alpha G(\alpha \mid \beta), \\ \hat{H}(x;\beta) &= \int F^*(x;\alpha)\,d_\alpha \hat{G}(\alpha \mid \beta). \end{aligned} \tag{3}$$

In turn, (2) may be expressed as

$$\int F(y;\beta)\,dJ_x(\beta) \equiv_{x,y} \int F(y;\beta)\,d\hat{J}_x(\beta) \tag{4}$$

where for each $\beta \varepsilon R^m$

$$\begin{aligned} J_x(\beta) &= \int_{-\infty}^{\beta} H(x;\gamma)\,dG_2(\gamma) \leqq G_2(\beta), \\ \hat{J}_x(\beta) &= \int_{-\infty}^{\beta} \hat{H}(x;\gamma)\,d\hat{G}_2(\gamma) \leqq \hat{G}_2(\beta). \end{aligned} \tag{5}$$

Dominated convergence applied to (4) insures that for each $x$, $J_x(\infty) = \hat{J}_x(\infty)$ since this common value is finite by (5).

Thus, from (4) and the part of the theorem already proved, $J_x \equiv_x \hat{J}_x$ or equivalently from (5), for all $\beta \varepsilon R^m$,

$$\int_{-\infty}^{\beta} H(x;\gamma)\,dG_2(\gamma) \equiv_x \int_{-\infty}^{\beta} \hat{H}(x;\gamma)\,d\hat{G}_2(\gamma). \tag{6}$$

On the other hand, letting $x \to \infty$ in (3) and then in (2) yields

$$\int F(y;\beta)\,dG_2(\beta) \equiv_y \int F(y;\beta)\,d\hat{G}_2(\beta) \tag{7}$$

implying as above that

$$G_2 = \hat{G}_2. \tag{8}$$

However, (8) in conjunction with (6) necessitates $H(x;\beta) = \hat{H}(x;\beta)$ for almost all $\beta$. The latter, together with (3) and the fact that the class of all mixtures of $\mathfrak{F}^*_{n,m\cdot n}$ is identifiable by hypothesis, entails

$$G(\cdot \mid \beta) = \hat{G}(\cdot \mid \beta), \qquad \text{almost all} \quad \beta. \tag{9}$$

Finally, combining (8) and (9), $G(\cdot,\cdot) = \hat{G}(\cdot,\cdot)$ so that $\mathfrak{F}^*_{n+1,m(n+1)}$ is identifiable.

Since the class of all mixtures of one-dimensional normal distributions

$$(\mathfrak{F} = \mathfrak{F}_{1,2} = \{F(x;\theta,\sigma^2) : F(x;\theta,\sigma^2) = (2\pi\sigma^2)^{-\frac{1}{2}} \int_{-\infty}^{x} \exp[-\tfrac{1}{2}((u-\theta)/\sigma)^2]\,du\})$$

is not identifiable [2], it follows that neither is the class of all mixtures of $\mathfrak{F}^*_{n,2n}$ (of products of $n$ normal distributions), $n > 1$. This may be contrasted with Corollary 4 of [1] which states that the class of all mixtures of products of $n$ *identical* normal distributions is identifiable for $n > 1$.

Since the argument of Theorem 1 (or a much simpler version of it) applies when all mixtures under consideration are finite, we have

THEOREM 2. *If the class of all finite mixtures of* $\mathfrak{F}_{1,m}$ *is identifiable, then for every* $n > 1$, *the class of finite mixtures of* $\mathfrak{F}^*_{n,mn}$ *is likewise identifiable. Conversely, if, for some* $n > 1$, *the class of all finite mixtures of* $\mathfrak{F}^*_{n,mn}$ *is identifiable, the same is true for* $\mathfrak{F}_{1,m}$.

The first part of Theorem 2 in the special case of exponential distributions is proved in [5].

Clearly, analogous results hold with $\mathfrak{F}_{1,m}$ and $\mathfrak{F}^*_{n,mn}$ replaced by $\mathfrak{F}_{k,m}$ and $\mathfrak{F}^*_{kn,m\cdot n}$, $k > 1$.

## REFERENCES

[1] BARNDORFF-NIELSEN, O. (1965). Identifiability of mixtures of exponential families. *J. Math. Anal. Appl.* **12** 115–121.

[2] TEICHER, HENRY. (1960). On the mixture of distributions. *Ann. Math. Statist.* **31** pp. 55–73.

[3] TEICHER, HENRY. (1961). Identifiability of Mixtures. *Ann. Math. Statist.* **32** 244–248.

[4] TEICHER, HENRY. (1963). Identifiability of finite mixtures. *Ann. Math. Statist.* **34** 1265–1269.

[5] YAKOWITZ, S. and Spragins, A. A characterization theorem for the identifiability of finite mixtures.[2]

[2] The author thanks the editors for calling his attention to this article which is to appear.