The Annals of Statistics
1987, Vol. 15, No. 3, 1229–1244

COMPLETE AND SYMMETRICALLY COMPLETE FAMILIES OF DISTRIBUTIONS

By Avi Mandelbaum[*] and Ludger Rüschendorf

Stanford University and Münster University

Polarization is used to unify and extend classical results concerning the completeness of generalized order statistics with dominated and non-dominated distributions.

1. Introduction to completeness.

1.0. The subject of the present paper is completeness of families of probability measures and statistics. The development is based mainly on the idea of polarization. The first section provides a unified introduction to completeness and prepares grounds for the new results developed in later sections.

1.1. The basic ingredient of a statistical model is a triplet $(X, \mathscr{B}, \mathscr{P})$ where $(X, \mathscr{B})$, the *sample space*, is a measurable space and $\mathscr{P} = \{P\}$ is a family of probability measures on it. We adopt the standard convention that statements with respect to $\mathscr{P}$ should be interpreted as simultaneous statements about *all* P in $\mathscr{P}$. For example, "a.s. $\mathscr{P}$" stands for "a.s. P for all $P \in \mathscr{P}$"; $L^p(X, \mathscr{B}, \mathscr{P})$ is the intersection of $L^p(X, \mathscr{B}, P)$ over $P \in \mathscr{P}$, and so on. Let $\mathscr{T} = \{f\}$ be a family of $\mathscr{B}$-measurable real-valued functions that includes the zero function. The family $\mathscr{P}$ is *complete with respect to* $\mathscr{T}$ ($\mathscr{T}$-complete) if any f in $\mathscr{T} \cap L^1(X, \mathscr{B}, \mathscr{P})$ that satisfies

$$\int_X f(x)\, dP = 0, \quad \forall\, P \in \mathscr{P} \tag{1a}$$

must also satisfy

$$f(x) = 0 \quad \text{a.s. } \mathscr{P}. \tag{1b}$$

Informally, completeness says that among the elements in $\mathscr{T}$, the only unbiased estimator of 0 is the zero function. The concept of completeness was anticipated by Halmos ([9], page 35) who looked for a proper way to describe families of measures that are "sufficiently large." It was later formalized by Lehmann and Scheffé ([16]–[18]) and has been widely applied since.

The term "complete" was apparently borrowed from functional analysis, as indicated in [7] (see also Section 3.2). Surprisingly, only a few works have used the function analytic relation explicitly. Examples are Farrell [4] who considered completeness in a Banach space setup and Siebert [24].

Received April 1984; revised December 1986.

AMS 1980 *subject classifications*. Primary 62G05, 62E10.

Key words and phrases. Completeness, symmetric completeness, polarization, unbiased estimation, tensor products of Hilbert spaces.

1.2. The notion of completeness is often associated with the size of the family $\mathscr{P}$. However, it is easy to construct examples of pairs $\mathscr{P}_1 \subset \mathscr{P}_2$ such that $\mathscr{P}_1$ is complete while $\mathscr{P}_2$ is not (indeed, let $\mathscr{P}_1$ be a complete family of measures supported on a subset of X; construct $\mathscr{P}_2$ by adding to $\mathscr{P}_1$ an incomplete family of measures which are supported on the complement of that subset). It is also important to note that completeness is not only a consequence of the fact that the family $\mathscr{P}$ is "sufficiently rich," but also a consequence of the fact that the set $\mathscr{T}$ is "sufficiently small." Indeed, all families of distributions are complete with respect to the singleton set $\{0\}$. Thus, different classes of test functions give rise to different (nonequivalent) notions of completeness. For a text treatment of completeness the reader is referred to Lehmann's books, [14] and [15].

1.3. In statistical application, restrictions on the test functions $\mathscr{T}$ are of two categories: quantitative and qualitative. Quantitative restrictions are technical and constrain the growth of the test functions while qualitative restrictions constrain their form.

An example of a quantitative restriction is to fix a number p, $1 \leq p \leq \infty$, and let

$$\mathscr{T} = L^p(X, \mathscr{B}, \mathscr{P}). \tag{2}$$

A family that is complete with respect to $\mathscr{T}$ in (2) is called *L^p-complete*. A family that is L^1-complete will simply be called *complete*. The L^∞-complete families are the *boundedly complete* families that arise in statistical testing.

1.4. Qualitative restrictions typically constrain the test functions in $\mathscr{T}$ to depend on a *statistic* T which is of interest. Formally, $\mathscr{T}$ consists of all functions which are measurable with respect to the σ-field generated by the statistic T. When $\mathscr{P}$ is complete w.r.t. such a family, then the statistic T is called a *complete statistic*. Completeness of a statistic amounts to completeness of the family of its distributions which, in turn, can be verified by (1) with test functions f of the form

$$f(x) = g(T(x)), \tag{3}$$

for some real-valued measurable functions g.

1.5. Complete statistics are most useful when they are also sufficient. A fundamental example of a sufficient statistic which is often complete is the order statistic of k independent observations from a common distribution. The statistical model is the triplet $(X^k, \mathscr{B}^k, \mathscr{P}^k)$, where the superscript k denotes the usual product and $\mathscr{P}^k = \{P^k\}$; $\mathscr{P}$ is the set of possible common distributions. The (generalized) order statistic T is

$$T(x_1, \dots, x_k) = \{x_1, \dots, x_k\},$$

which assigns to the vector sample $(x_1, \dots, x_k)$ the *set* of points $\{x_1, \dots, x_k\}$. The set of test functions $\mathscr{T}$ which satisfy the qualitative restriction (3) with respect to the order statistic T are the $\mathscr{B}^k$-measurable symmetric functions. It follows that the order statistic is complete if and only if any integrable *symmetric* function f that satisfies

(4a) $$\int_{X^k} f(x_1, \dots, x_k)\, dP^k = 0, \quad \forall\, P \in \mathscr{P},$$

must also satisfy

(4b) $$f = 0 \quad \text{a.s. } \mathscr{P}^k.$$

A family $\mathscr{P}$ for which (4) holds is called *symmetrically complete of order* k. Thus, $\mathscr{P}$ is symmetrically complete of order k if and only if the k-order statistic is complete. Symmetric completeness was first treated by Halmos [9] who discussed discrete measures. Fraser [5] studied symmetric completeness of families dominated by nonatomic measures. Bell, Blackwell and Breiman [1] coined the term "symmetrically complete" and used the results in [5] to establish the symmetric completeness of a wide range of nonparametric families.

1.6. A family $\mathscr{P}$ that is symmetrically complete of some order $k \geq 2$ must be complete as well (see also Remark 6 in Section 3.2). Indeed, if f satisfies (1a), then

$$\int_{X^k} \sum_1^k f(x_i)\, dP^k = k \int_X f(x)\, dP = 0, \quad \forall\, P \in \mathscr{P}.$$

Since $\Sigma_1^k f(x_i)$ is symmetric,

$$\sum_1^k f(x_i) = 0 \quad \text{a.s. } \mathscr{P}^k,$$

implying that for any measurable set $A \in \mathscr{B}$,

$$\int_A f(x)\, dP = \int_{A^k} \sum_1^k f(x_i)\, dP^k / \left(kP(A)^{k-1}\right) = 0.$$

Hence, f must also satisfy (1b). However, completeness does not imply symmetric completeness and examples could be constructed on a sample space X with only two points. A key question is, therefore,

1.6A. What conditions should be added to "$\mathscr{P}$ is complete" in order to guarantee also that "$\mathscr{P}$ is symmetrically complete"?

1.7. When symmetric completeness holds, symmetric statistics are uniformly minimum variance unbiased estimators of their expectation. For square integrable symmetric statistics, the theory of symmetric tensor products of Hilbert spaces is a natural framework [3]. Thus, it is not surprising that when the qualitative constraint of symmetry and the quantitative constraint of square integrability both apply, that the same Hilbert space theory is again natural and useful. Hilbert space ideas, together with the polarization techniques (see Section 2) allow one to provide partial answers to 1.6A and unify and extend completeness results that have been proved in [9], [1], [5] and [14]. For an introduction to tensor products of Hilbert spaces, the reader is referred to [23]. A comprehensive treatment of the subject is developed in Neveu [19].

1.8. The structure of a statistical model is greatly simplified if its family $\mathscr{P} = \{P\}$ is dominated by a σ-finite measure μ. It is then possible to parametrize $\mathscr{P}$ by the densities of its members with respect to μ, say $D = \{h\}$, and deal with functions (the densities h) rather than measures (the probability distributions P). For dominated models, it is natural to replace $\mathscr{T}$ in (2) by

$$\mathscr{T} = L^p(X, \mathscr{B}, \mu)$$

and consider completeness of the family of densities D. This approach is taken in Section 3 where the concepts of L^2-complete and symmetrically complete families of densities are identified with the standard concept of a "complete set in a Hilbert space."

1.9. The rest of the present work is structured as follows. In Section 2 we state the polarization results (which are proved later in the appendices). The Hilbert space relation to completeness is developed in Section 3 and the main results are given in Section 4.

2. Polarization.

2.0. In this section we state two identities and a lemma that form the basis of the polarization technique. As before, $(X, \mathscr{B})$ is a measurable space: All functions considered are either $\mathscr{B}$-measurable or $\mathscr{B}^k$-measurable, where $\mathscr{B}^k$ is the k product σ-field in X^k. Functions that are used as integrands will always be assumed integrable.

2.1. For a function $f = f(x_1, \ldots, x_k)$, let Sf denote the symmetric function

$$Sf(x_1, \ldots, x_k) = \sum_{\sigma} f(x_{\sigma_1}, \ldots, x_{\sigma_k}),$$

where the sum is over all $k!$ permutations $\sigma = (\sigma_1, \ldots, \sigma_k)$ of $(1, 2, \ldots, k)$. Given k functions $h_1(x), \ldots, h_k(x)$, define the function of k variables $\otimes_1^k h_i$ by

$$\left(\bigotimes_1^k h_i \right)(x_1, \ldots, x_k) = h_1(x_1) \cdots h_k(x_k). \tag{5}$$

Similarly, if $P_1, \ldots, P_k$ are measures on $(X, \mathscr{B})$, denote by $\otimes_1^k P_i$ the usual product measure on $(X^k, \mathscr{B}^k)$.

2.2. The polarization identities, which we are about to state, have been used, no doubt, numerous times implicitly, but we are not aware of any explicit statement of them [an example is [3], page 743, the derivation of (2.6)].

Polarization identity for functions.

$$S \bigotimes_{1}^{k} h_i = \sum_{R} (-1)^{k-|R|} \bigotimes_{1}^{k} h_R,$$

where the sum is over all subsets R of $\{1, \dots, k\}$ and

$$h_R(x) = \sum_{j \in R} h_j(x), \qquad h_\phi(x) = 0.$$

Polarization identity for measures.

$$\sum_{\sigma} \bigotimes_{1}^{k} P_{\sigma_i} = \sum_{R} (-1)^{k-|R|} \bigotimes_{1}^{k} P_R,$$

where $P_R = \sum_{j \in R} P_j$ and $P_\phi = 0$.

The polarization identity for measures is verified by integrating the identity for functions with respect to the product measure μ^k where μ dominates P_i and h_i is the density of P_i w.r.t. μ. The polarization identity for functions is essentially an identity for permanents of matrices. It is discussed in Appendix 1.

2.3. For the polarization lemma we need the following

DEFINITION. A subset C of a linear space is called *weakly convex* if for every $c_1, c_2 \in C$ there *exists* a number α *strictly* between 0 and 1 such that

$$\alpha c_1 + (1 - \alpha) c_2 \in C.$$

POLARIZATION LEMMA. *Let $\mathscr{P}$ be a weakly convex set of probability measures on $(X, \mathscr{B})$. Suppose that $f(x_1, \dots, x_k)$ is a symmetric function satisfying*

$$\int_{X^k} f \, dP^k = 0, \quad \forall P \in \mathscr{P}. \tag{6a}$$

Then

$$\int_{X^k} f \, d \bigotimes_{i=1}^{k} P_i = 0, \quad \forall P_1, \dots, P_k \in \mathscr{P}. \tag{6b}$$

The proof of the polarization lemma reduces to a statement about the completeness of the multinomial distribution with a weakly convex parameter set. The lemma appears in [17], page 152, as Exercise 12 and is discussed in Appendix 2.

3. Statistical completeness in the Hilbert space framework.

3.0. The present section provides a Hilbert space framework for statistical completeness. We start with some necessary notations and proceed with the relevant Hilbert space theory.

3.1. Fix a σ-finite measure μ on $(X, \mathscr{B})$. Let $L^2(\mu) = L^2(X, \mathscr{B}, \mu)$ be the usual Hilbert space with inner product $(f, g) = \int_X f(x)g(x)\,d\mu$.

For a subset $D \subset L^2(\mu)$, let $\otimes_1^k D$ be the subset of $L^2(\mu^k) = L^2(X^k, \mathscr{B}^k, \mu^k)$ defined by

$$\bigotimes_1^k D = \left\{ \bigotimes_1^k h_i \colon h_1, \dots, h_k \in D \right\},$$

where $\otimes_1^k h_i$ is defined in (5).

Let S^k denote the sub-σ-field of symmetric sets in $\mathscr{B}^k$ and let $\odot_1^k D$ be the subset of $L_s^2(\mu^k) = L^2(X^k, S^k, \mu^k)$ defined by

$$\bigodot_1^k D = \left\{ \bigotimes_1^k h \colon h \in D \right\}.$$

The space $L_s^2(\mu^k)$ is nothing but the symmetric functions (or μ^k-equivalent classes) which are square integrable with respect to the product measure μ^k.

The inner product in $L^2(\mu^k)$ and $L_s^2(\mu^k)$ is denoted by $(f, g)_k$; the expression $(f, \otimes_1^k h_i)_k$ will be abbreviated to $(f, \otimes h_i)_k$.

3.2. It would be very surprising if, when coining the term "complete," Lehmann and Scheffé did not have the following definition in mind:

DEFINITION 1. A subset C of Hilbert space H is called *complete in H* (or just *complete*) if the only element in H which is orthogonal to C is the element 0, i.e.,

$$(f, h) = 0, \quad \forall\, h \in C \Rightarrow f = 0.$$

A standard exercise in Hilbert space theory is

PROPOSITION 2. *A subset $C \subset H$ is complete if and only if* $\operatorname{span}(C)$ *is dense in H, i.e.,* $\overline{\operatorname{span}}(C) = H$.

[Here span (C) is the subspace generated by C and $\overline{\operatorname{span}}(C)$ is its closure.]

Now let D be a subset of $L^2(\mu)$ consisting of *probability* densities with respect to μ. Elements in D are nonnegative and integrate to 1 with respect to μ.

DEFINITION 3. The family D is called *L^2-complete* (or just *complete*) if for any $f \in L^2(\mu)$,

$$(f, h) = \int_X f(x)h(x)\,d\mu = 0, \quad \forall\, h \in D \Rightarrow f = 0.$$

DEFINITION 4. The family D is called *symmetrically complete of order k* (with respect to μ) if for any $f \in L_s^2(\mu^k)$,

$$(f, \otimes h)_k = \int_{X^k} f(x_1, \dots, x_k) h(x_1) \cdots h(x_k)\, d\mu^k = 0, \quad \forall\, h \in D \Rightarrow f = 0.$$

From Proposition 2 one concludes

3.2A. The set D is symmetrically complete of order k, i.e., $\odot_1^k D$ is complete in $L_s^2(\mu^k)$, if and only if span $\odot_1^k D$ is dense in $L_s^2(\mu^k)$.

REMARKS.

1. The name "L^2-complete" distinguishes it from "boundedly complete" and "complete" used in the statistical literature. (We actually abuse the notation slightly, since in the literature completeness refers to *distributions*, while we use it for *densities*.) However, Definitions 1 and 3 coincide by identifying $C = D$ and $H = L^2(\mu)$. Hence, it is reasonable to omit the "L^2" part in Definition 3 and call D just "complete."
2. Definitions 1 and 4 coincide by identifying $H = L_s^2(\mu^k)$ and $C = \odot_1^k D$.
3. Since elements in L^2 spaces are equivalence classes, we have omitted the "almost everywhere" qualification in Definitions 3 and 4 from the conclusion "$f = 0$."
4. Definitions 3 and 4 are what properties (1) and (4), respectively, reduce to, when the measure μ is actually *equivalent* to $\mathscr{P}$ and $\mathscr{T} = L^2(\mu)$ resp. $\mathscr{T} = L_s^2(\mu^k)$.
5. It was shown in Section 1.6 that symmetric completeness of order k implies completeness of D. The same idea can be used to prove (by induction) that symmetric completeness of some order k implies symmetric completeness of *all* orders less than k.

3.3. Denote by 1_A the indicator function of the set A and let

$$\mathscr{B}^* = \{A \in \mathscr{B}: 0 < \mu(A) < \infty\}. \tag{7}$$

Given a set $A \in \mathscr{B}^*$, define the uniform density u_A (with respect to μ) by $u_A(x) = (1/\mu(A))1_A(x)$, $x \in X$. The set $D = \{\otimes_1^k u_{A_i};\ A_i \in \mathscr{B}^*\}$ is complete in $L^2(\mu^k)$. Hence, $\otimes_1^k L^2(\mu^k)$ is complete in $L^2(\mu^k)$. In fact,

PROPOSITION 5. *D is complete in $L^2(\mu)$ if and only if $\otimes_1^k D$ is complete in $L^2(\mu^k)$.*

PROOF. First assume that D is complete. The proof that $\otimes_1^k D$ is complete is based on the two relations

$$\bigotimes_1^k \operatorname{span}(D) \subset \operatorname{span}\left(\bigotimes_1^k D\right) \tag{8}$$

and

$$\bigotimes_1^k \overline{\operatorname{span}}(D) \subset \overline{\bigotimes_1^k \operatorname{span}(D)}. \tag{9}$$

Then

$$L^2(\mu^k) = \overline{\operatorname{span}\left\{\bigotimes_1^k \overline{\operatorname{span}}(D)\right\}} \underset{(9)}{\subset} \overline{\overline{\operatorname{span}\left\{\bigotimes_1^k \operatorname{span}(D)\right\}}} \underset{(8)}{\subset} \overline{\operatorname{span} \bigotimes_1^k D},$$

which indeed verifies that span $(\otimes_1^k D)$ is dense in $L^2(\mu^k)$. Now, (8) is a consequence of simple manipulations and (9) follows from the fact that if, for $i = 1, \dots, k$, $\psi_i^n \to h_i$ in $L^2(\mu)$ as $n \to \infty$, then also $\otimes_1^k \psi_i^n \to \otimes_1^k h_i$ in $L^2(\mu^k)$. Now suppose that $\otimes_1^k D$ is complete. To see that D must be complete, just note that

$$(f, h_1) = \int_{X^k} f(x_1) h_1(x_1) \cdots h_k(x_k)\, d\mu^k, \quad \forall\, h_1, \dots, h_k \in D,$$

because h_i integrate to 1 and μ^k is a product measure. □

3.4. The definition of symmetric completeness applies to sets of functions that are not necessarily densities. For example, using observation 3.2A, the following proposition shows that $D = L^2(\mu)$ is symmetrically complete for all orders. This is a fundamental and typical application of the polarization identity in that it reduces "symmetric completeness" statements in $L_s^2(\mu^k)$ to "completeness" statements in $L^2(\mu^k)$.

Proposition 6. *The set* $\odot_1^k L^2(\mu)$ *is complete in* $L_s^2(\mu^k)$.

Proof. Let f be a symmetric function in $L_s^2(\mu^k)$ for which $(f, \otimes h)_k = 0$, $\forall\, h \in L^2(\mu)$. Then for $h_1, \dots, h_k \in L^2(\mu)$,

$$(f, \otimes h_i)_k = \frac{1}{k!}(f, S \otimes h_i)_k = \frac{1}{k!} \sum_R (-1)^{k-|R|} (f, \otimes h_R)_k = 0.$$

Since $\otimes_1^k L^2(\mu)$ is complete, we conclude that $f = 0$. □

4. Symmetric completeness—main results.

4.0. We now proceed with the set up of Section 3 and use the polarization results from Section 2 to partially answer question 1.6A.

4.1. As a motivation, we first show

4.1A. If $D \subset L^2(\mu)$ is a set of probability densities which is complete and *convex*, then D is symmetrically complete for all orders.

Indeed, suppose that $f \in L_s^2(\mu^k)$ satisfies $(f, \otimes h)_k = 0$, $\forall\ h \in D$. Let $h_1, \dots, h_k \in D$. Then $(1/|R|) h_R \in D$ by convexity, and as in the proof of Proposition 6, we get from the polarization identity,

$$(f, \otimes h_i)_k = \frac{1}{k!} \sum_{R \neq \phi} (-1)^{k-|R|} |R|^k \left(f, \otimes \frac{h_R}{|R|} \right) = 0.$$

To conclude that $f = 0$, it suffices to show that $\otimes_1^k D$ is complete. This follows from Proposition 5.

An immediate consequence of 4.1A is

4.1B. Let U be the set of uniform densities with respect to μ over $(X, \mathscr{B})$, i.e.,

$$U = \left\{ u_A(x) = \frac{1}{\mu(A)} 1_A(x),\ A \in \mathscr{B}^* \right\}, \tag{10}$$

where $\mathscr{B}^*$ is defined in (7). Then the *convex set generated by U* is symmetrically complete of all orders (cf. Lehmann [14], page 153, problem 13).

4.2. The statement 4.1A can be extended in at least two directions. In one direction, the polarization lemma combined with Proposition 5 yields

THEOREM 7. *Let $D \subset L^2(\mu)$ be a set of probability densities with respect to μ. If D is complete and weakly convex then D is symmetrically complete of all orders.*

The second direction, which will now be pursued, is suggested by 4.1B.

4.3. Fraser [5] proved

4.3A. If the dominating measure μ is *nonatomic*, the set U in (10) is itself symmetrically complete for all orders.

Indeed, suppose $f \in L_s^2(\mu^k)$ is such that

$$(f, \otimes u_A)_k = 0, \quad \text{for all } A \in \mathscr{B}^*.$$

By the polarization identity, for $A_1, \dots, A_k \in \mathscr{B}^*$ *disjoint* we get

$$(f, \otimes 1_{A_i})_k = \frac{1}{k!} \sum_R (-1)^{k-|R|} (f, \otimes u_{A_R})_k \mu(A_R)^k = 0,$$

where $A_R = \bigcup_{j \in R} A_j$. The fact that $f = 0$ is a consequence of

LEMMA 8. *If μ is nonatomic, then $\{ \otimes_1^k u_{A_i}: A_i \in \mathscr{B}^*, A_i \cap A_j = \varnothing, i \neq j\}$ is complete in $L^2(\mu^k)$.*

Lemma 8 was proved independently in [5] and by Itô [10] who used it to define multiple Wiener integrals with respect to Gaussian random measures.

The reason for taking disjoint sets in proving 4.3A was to be able to express 1_{A_R} as $\sum_{j \in R} 1_{A_j}$. This suggests the following generalized form of 4.1A which is deduced from unifying the proofs of 4.1A and 4.3A.

THEOREM 9. *Suppose that* $\otimes_1^k D$ *contains a subset C which is complete in* $L^2(\mu^k)$ *and such that for any* $\otimes_1^k h_i \in C$ *there exist* $\alpha_1, \dots, \alpha_k > 0$ *for which*

$$\sum_{j \in R} \alpha_j h_j \Big/ \sum_{j \in R} \alpha_j \in D, \quad \text{for all } R \subset \{1, \dots, k\}.$$

Then D is symmetrically complete of order k.

REMARKS.

1. If D is convex and complete, taking $C = \otimes_1^k D$ and $\alpha_1 = \cdots = \alpha_k = 1$ yields 4.1A.
2. If μ is nonatomic, taking C to be the set from Lemma 8 yields 4.3A.

4.4. An obvious question is whether the nonatomic nature of μ is necessary for 4.3A to hold. The answer is

PROPOSITION 10. (a) *The set U is always symmetrically complete of order* 2.

(b) *If* $|\mathscr{B}| \geq 3$, *then there exists a discrete measure* μ *for which U is not symmetrically complete of any order* $k \geq 3$.

PROOF. (a) Suppose that $(f, \otimes u_A)_2 = 0$ for some $f \in L_s^2(\mu^2)$ and all $A \in \mathscr{B}^*$. Now, the symmetric σ-field S^2 is generated by $\{A_1 \times A_2 \cup A_2 \times A_1; A_1, A_2 \in \mathscr{B}\}$ or, equivalently, by $\{A_1 \times A_2 \cup A_2 \times A_1; A_1, A_2 \in \mathscr{B}, A_1 \cap A_2 = \varnothing \text{ or } A_1 = A_2\}$.

For $A_1 \cap A_2 = \varnothing$ we have

$$A_1 \times A_2 \cup A_2 \times A_1 = (A_1 \cup A_2)^2 - (A_1^2 \cup A_2^2),$$

hence, $\{A^2; A \in \mathscr{B}\}$ generates S^2. We conclude that $\int_C f \, d\mu^2 = 0$ for all $C \in S^2$. Taking C equal first to $\{f \geq 0\}$ and then $\{f \leq 0\}$ implies that $f = 0$.

(b) First, let $k = 3$, $A_0, A_1 \in \mathscr{B}$, $A_0 \cap A_1 = \varnothing$, $A_i \neq \varnothing$, $i = 0, 1$, and $x_i \in A_i$, $i = 0, 1$. Define

$$S_1 = A_0 \times A_0 \times A_1 \cup A_0 \times A_1 \times A_0 \cup A_1 \times A_0 \times A_0,$$

$$S_2 = A_1 \times A_1 \times A_0 \cup A_1 \times A_0 \times A_1 \cup A_0 \times A_1 \times A_1$$

and $\mu = \frac{1}{2}\varepsilon_{x_0} + \frac{1}{2}\varepsilon_{x_1}$, ε_{x_i} denoting the one point measures supported on $\{x_i\}$. For $h = 1_{S_1} - 1_{S_2}$, $\int_{A^3} h \, d\mu^3 = 0$ holds for all $A \in \mathscr{B}^*$, i.e.,

$$(h, \otimes u_A)_3 = 0, \quad \text{for all } A \in \mathscr{B}^*,$$

while, clearly, h is not equal to zero a.s. μ^3.

In the general case, $k \geq 3$, consider a similar construction where S_1 is the symmetrized version of $A_0 \times A_0 \times A_1 \times X^{k-3}$, etc. □

REMARKS.

1. The statements 4.1B, 4.3A, Lemma 8 and Proposition 10 all still hold if $\mathscr{B}^*$ is replaced by a subset of $\mathscr{B}^*$, as long as that subset is a semiring that generates $\mathscr{B}$ (cf. also Fraser [5] concerning $\mathscr{U}$).
2. Proposition 10 is related to the results of Rao [22] and Grzegorek [8] stating that S^k is generated by A^k, $A \in \mathscr{B}$ for $k = 2$, while for $k \geq 3$ equality does not hold except in the trivial case $\mathscr{B} = \{\varnothing, X\}$.

4.5. Several of the results for dominated families discussed in the previous sections have analogs for nondominated families. Plachky [20] and Landers and Rogge [13] proved the following analogue of Proposition 5 (Plachky considered the case of bounded completeness).

PROPOSITION 5′. *A family $\mathscr{P}$ is complete if and only if*

$$\bigotimes_1^k \mathscr{P} = \left\{ \bigotimes_1^k P_i;\ P_i \in \mathscr{P}, 1 \leq i \leq k \right\} \quad \textit{is complete}.$$

Using Proposition 5′ and the polarization lemma we get

THEOREM 7′. *A family $\mathscr{P}$ which is complete and weakly convex is symmetrically complete of all orders.*

Moreover, the polarization identity for measures can be used to verify

THEOREM 9′. *Suppose that $\otimes_1^k \mathscr{P}$ contains a complete subset C such that for any $\otimes_1^k P_i \in C$ there exist $\alpha_1, \ldots, \alpha_k > 0$ for which*

$$\sum_{j \in R} \alpha_j P_j \Big/ \sum_{j \in R} \alpha_j \in \mathscr{P}, \quad \textit{for all } R \subset \{1, \ldots, k\}.$$

Suppose also that C dominates $\mathscr{P}$ in the sense that $f \in L^1(X, \mathscr{B}, P)$ which satisfies $\int f\, dP = 0$, $\forall\, P \in C$ must also satisfy $f = 0$ a.s. P, $\forall\, P \in \mathscr{P}$. Then $\mathscr{P}$ is symmetrically complete of order k.

4.6. We end this section with a generalization of 4.1B which allows one to deduce a completeness result for U-statistics and also generalize Halmos' treatment of discrete distributions. Given a σ-finite measure μ, fix $k = 1, 2, \ldots$ and define a set of densities V_k by

$$V_k = \left\{ \frac{1}{\sum_{i=1}^m r_i \mu(A_i)} \sum_{i=1}^m r_i 1_{A_i}:\ A_1, \ldots, A_m \text{ disjoint sets in } \mathscr{B}^*; \right.$$
$$\left. \sum_{i=1}^m r_i \leq k,\ r_1, \ldots, r_m = 1, 2, 3, \ldots \right\}. \tag{11}$$

Let $\mathscr{P}_{\mu, k}$ be the set of all probability measure densities with respect to μ in V_k.

THEOREM 11. *For* $k = 1, 2, \ldots,$

(a) V_k *is symmetrically complete of order* k;
(b) $\mathscr{P}_{\mu,k}$ *is symmetrically complete of order* k.

PROOF. We only prove part (b) since the proof of part (a) is similar. For $A \in \mathscr{B}^*$, let μ_A be the μ-uniform distribution on A, i.e., the distribution on $\mathscr{B}$ with density $(1/\mu(A))\, 1_A$ w.r.t. μ. Define

$$C^* = \left\{ \bigotimes_{i=1}^{k} \mu_{A_i};\ A_i \in \mathscr{B}^*,\ A_i = A_j \text{ or } A_i \cap A_j = \varnothing \text{ for } i \neq j \right\}.$$

If $\otimes_{i=1}^{k} \mu_{A_i} \in C^*$, then

$$\frac{1}{\sum_{j \in R} \mu(A_j)} \sum_{j \in R} \mu(A_j) \mu_{A_j} \in \mathscr{P}_{\mu,k}, \quad \text{for } R \subset \{1, \ldots, k\}.$$

The symmetric completeness of order k of $\mathscr{P}_{\mu,k}$ now follows from Theorem 9′ and the following lemma. □

LEMMA 12. C^* *is complete*.

PROOF. If $f \in L^1(C^*)$ and

$$\int f\, d \bigotimes_{i=1}^{k} \mu_{A_i} = \frac{1}{\prod_{i=1}^{k} \mu(A_i)} \int_{A_1 \times \cdots \times A_k} f\, d\mu^k = 0,$$

for all $A_i \in \mathscr{B}^*$ such that $A_i \cap A_j = \varnothing$ or $A_i = A_j$ for $i \neq j$, then $\int_{A_1 \times A_k} f\, d\mu^k = 0$, for all $A_i \in \mathscr{B}^*$, since each product $A_1 \times \cdots A_k$ can be written as a finite sum of products $B_1 \times \cdots \times B_k$, $B_i \in \mathscr{B}$ with $B_i \cap B_j = \varnothing$ or $B_i = B_j$ for $i \neq j$. This implies that $f = 0$ a.s. μ^k, i.e., C^* is complete. □

REMARKS.

1. As in the remark following Proposition 10, the collection of sets $\mathscr{B}^*$ in Theorem 11 can be replaced by a semiring generating $\mathscr{B}$.
2. Theorem 11 includes, as a special case, the result of Halmos [9], i.e., if $\mathscr{P}$ contains all finite discrete measures on $(X, \mathscr{B})$, then $\mathscr{P}$ is symmetrically complete of any order. For the proof note that for any discrete measure μ on $(X, \mathscr{B})$, $\mathscr{P}_{\mu,k} \subset \mathscr{P}$.
3. By reasoning analogously to Remark (2), Theorem 11 implies that, given an arbitrary σ-finite measure μ on $(X, \mathscr{B})$, the set of all probability measures which are absolutely continuous with respect to μ is symmetrically complete of all orders. Fraser [6] proved this result when μ is the Lebesgue measure on $\mathbb{R}^1$ and Bell, Blackwell and Breiman [1] extended it to nonatomic μ. Similarly, one concludes that the set of all nonatomic measures on $(X, \mathscr{B})$ is symmetrically complete of all orders. This result is proven in Lehmann [14] for $X = \mathbb{R}^1$ and in [1] for a general X.

4. In the theory of unbiased estimation one often considers the sets of distributions on $(X, \mathscr{B})$

$$\mathscr{P}_f = \left\{P: \int_{X^k} f^2(x_1, \dots, x_k)\, dP^k < \infty\right\},$$

$$\mathscr{P}_{f,\mu} = \left\{P \in \mathscr{P}_f: P \text{ absolutely continuous w.r.t. } \mu\right\},$$

for f a fixed function in $L_s^2(P^k)$. Theorem 11 implies that both $\mathscr{P}_f$ and $P_{f,\mu}$ are symmetrically complete of all orders. In particular, U-statistics are complete and hence uniformly minimum variance unbiased estimators of their means.

APPENDIX 1

The polarization identity. The polarization identity is an immediate consequence of the following representation for the permanent of a $k \times k$ matrix $A = [a_{ij}]$:

Permanent identity.

$$\operatorname{per}(A) = \sum_R (-1)^{k-|R|} \prod_{i=1}^{k} \left(\sum_{j \in R} a_{ij} \right),$$

where

$$\operatorname{per}(A) = \sum_\sigma \prod_{i=1}^{k} a_{i\sigma_i}.$$

The permanent identity reduces to the polarization identity by letting $a_{ij} = h_j(x_i)$. A proof of the permanent identity using the inclusion–exclusion principle can be found in Jacobs ([11], Satz 3.13, page 31).

REMARK. The polarization identity can be derived from a relation which is stated on page 52 of [2] and attributed to P. Cartier. It is also a consequence of the Möbius inversion formula ([21], page 2).

APPENDIX 2

Completeness of multinomial families. We were not able to follow the outline suggested in Lehmann [14] to prove the polarization lemma. [To see the difficulty just try to prove that $\Sigma_{\alpha(1,3)} = 0$. Note that Lehmann's equation (1) does *not* guarantee

$$(\alpha_1 P_1 + \alpha_3 P_3)/(\alpha_1 + \alpha_3) \in \mathscr{P}, \quad \text{only that} \quad (\alpha_1 P_1 + \alpha_2 P_2)/(\alpha_1 + \alpha_2) \in \mathscr{P}.]$$

Recently, Kallenberg et al. [12] provided a proof that follows Lehmann's suggestions. Here is an alternative proof which was derived independently. It may be of independent interest since it also provides a sufficient condition for completeness of the multinomial distribution.

DEFINITION. Call a set $S \subset R^k$ a *weak simplex* if it is a *subset* of the $(k-1)$-simplex

$$\left\{\alpha = (\alpha_1, \dots, \alpha_k) \colon \alpha_1 \geq 0,\ \sum_{i=1}^{k} \alpha_i = 1\right\}$$

and it contains the k *unit vectors* and is *weakly convex.*

PROOF OF THE POLARIZATION LEMMA. Let f be a symmetric function. Then

$$\int f\, d\left(\sum_{i=1}^{k} \alpha_i P_i\right)^k = \sum \binom{k}{x_1, \dots, x_k} \alpha_1^{x_1} \cdots \alpha_k^{x_k} \int f\, dP_1^{x_1} \cdots dP_k^{x_k},$$

where the sum is over $\sum_{i=1}^{k} x_i = k$ and $\binom{k}{x_1, \dots, x_k} = k!/x_1! \cdots x_k!$.

Let us fix $P_1, \dots, P_k \in \mathscr{P}$. Define a set S by

$$S = \left\{\alpha = (\alpha_1, \dots, \alpha_k) \colon \alpha_1 \geq 0,\ \sum_{i=1}^{k} \alpha_i P_i \in \mathscr{P}\right\}.$$

Then S is a weak simplex and the lemma will follow if we prove that the multinomial distribution, *indexed by a weak simplex,* is a complete family. More precisely,

PROPOSITION. *Let S be a weak simplex in R^k. Then for all n, the family of multinomial distributions $\{B(n; \alpha = (\alpha_1, \dots, \alpha_k)), \alpha \in S\}$ is complete.*

REMARK. The polarization lemma follows by taking $k = n$.

PROOF. We must show that if

$$E_\alpha f(X_1, \dots, X_k) = \sum \binom{n}{x_1, \dots, x_k} \alpha_1^{x_1} \cdots \alpha_k^{x_k} f(x_1, \dots, x_k) = 0, \quad \forall\, \alpha \in S,$$

then $f = 0$. The proof is by induction on k. Start with the binomial case, $k = 2$. Let $X \sim B(n, \alpha)$ and suppose that

$$E_\alpha f(x) = \sum_{x=0}^{n} \binom{n}{x} \alpha^x (1-\alpha)^{n-x} f(x) = 0, \quad \forall\, \alpha \text{ such that } (\alpha, 1-\alpha) \in S.$$

The set $\{\alpha \colon (\alpha, 1-\alpha) \in S\}$ is a weak simplex in R^1. Hence, it must contain n distinct $\alpha_1, \dots, \alpha_n$. The $n \times n$ matrix $[\alpha_j^x (1-\alpha_j)^{n-x}]$ is nonsingular, being equivalent (up to nonsingular matrices) to a Vandermonde matrix. Hence, $f(x) = 0$, $x = 0, 1, \dots, n$.

Now assume that the preceding proposition holds for $k-1$. We shall use the fact that if $(X_1, \dots, X_k) \sim B(n; (\alpha_1, \dots, \alpha_k))$, then

$$X_1 \sim B(n, \alpha_1),$$

$$X_2, \dots, X_k | X_1 = y \sim B\left(n - y; \frac{\alpha_2}{1-\alpha_1}, \dots, \frac{\alpha_k}{1-\alpha_1}\right).$$

For $\alpha = \varepsilon(1, 0, \dots, 0) + (1 - \varepsilon)(0, \alpha_2, \dots, \alpha_k) = (\varepsilon, (1 - \varepsilon)\alpha_2, \dots, (1 - \varepsilon)\alpha_k)$, we have

$$0 = E_\alpha f(X_1, \dots, X_k) = \sum_{y=0}^{n} \binom{n}{y} \varepsilon^y (1 - \varepsilon)^{n-y} E_\alpha\big[f(X_1, \dots, X_k) | X_1 = y\big]$$

$$= \sum_{y=0}^{n} \binom{n}{y} \varepsilon^y (1 - \varepsilon)^{n-y} E_{(\alpha_2, \dots, \alpha_k)} f(y, X_2, \dots, X_k).$$

First, note that the set $S' = \{(\alpha_2, \dots, \alpha_k): (0, \alpha_2, \dots, \alpha_k) \in S\}$ is a weak simplex in R^{k-1}. Now fix $(\alpha_2, \dots, \alpha_k) \in S'$. The set

$$S'' = \{\varepsilon: (\varepsilon, (1 - \varepsilon)\alpha_2, \dots, (1 - \varepsilon)\alpha_k) \in S\}$$

is a weak simplex in R^1. By the binomial case

$$E_{(\alpha_2, \dots, \alpha_k)} f(y, X_2, \dots, X_k) = 0, \qquad y = 0, 1, \dots, n.$$

The last argument can be repeated for every $(\alpha_2, \dots, \alpha_k) \in S'$. Using the induction hypotheses we conclude that $f = 0$. □

Acknowledgments. The authors are grateful to Professor Charles Stein and Professor Detlef Plachky for their help. References [12] and [24] were brought to our attention by a referee who also read the manuscript carefully and proposed changes that have enhanced the presentation.

REFERENCES

[1] Bell, C. B., Blackwell, D. and Breiman, L. (1960). On the completeness of order statistics. *Ann. Math. Statist.* **31** 794–797.

[2] Dellacherie, C. and Meyer, P.-A. (1980). *Probabilités et Potentiel. Theorie des Martingales.* Hermann, Paris.

[3] Dynkin, E. B. and Mandelbaum, A. (1983). Symmetric statistics, Poisson point processes and multiple Wiener integrals. *Ann. Statist.* **11** 739–745.

[4] Farrell, R. H. (1962). Dense algebras of functions in L^p. *Proc. Amer. Math. Soc.* **13** 324–328.

[5] Fraser, D. A. S. (1954). Completeness of order statistics. *Canad. J. Math.* **31** 703–709.

[6] Fraser, D. A. S. (1957). *Nonparametric Methods in Statistics.* Wiley, New York.

[7] Gordeesch, J. (1972). Completeness and unbiased estimation. *Amer. Statist.* **26** 45–46.

[8] Grzegorek, E. (1981). Symmetric σ-fields of sets and universal null sets. *Lecture Notes in Math.* **945** 101–109. Springer, Berlin.

[9] Halmos, P. (1946). The theory of unbiased estimation. *Ann. Math. Statist.* **17** 34–43.

[10] Itô, K. (1951). Multiple Wiener integrals. *J. Math. Soc. Japan.* **3** 157–169.

[11] Jacobs, K. (1983). *Einführung in die Kombinatorik.* Walter de Gruyter, Berlin.

[12] Kallenberg, W. C. M. et al. (1984). *Testing Statistical Hypotheses: Worked Solutions.* Mathematical Centre, Amsterdam.

[13] Landers, D. and Rogge, L. (1976). A note on completeness. *Scand. J. Statist.* **3** 139.

[14] Lehmann, E. L. (1959). *Testing Statistical Hypothesis.* Wiley, New York.

[15] Lehmann, E. L. (1983). *Theory of Point Estimation.* Wiley, New York.

[16] Lehmann, E. L. and Scheffé, H. (1947). On the problem of similar regions. *Proc. Nat. Acad. Sci. U.S.A.* **33** 382–386.

[17] Lehmann, E. L. and Scheffé, H. (1950). Completeness, similar regions and unbiased estimation. *Sankhyā* **10** 305–340.

[18] LEHMANN, E. L. and SCHEFFÉ, H. (1955). Completeness, similar regions and unbiased estimation. *Sankhyā* **15** 219–236.

[19] NEVEU, J. (1968). *Processus Aleatoires Gaussiens*. Les Presses de l'Université de Montréal.

[20] PLACHKY, D. (1974). A characterization of bounded completeness in the undominated case. In *Transactions Seventh Prague Conference*. 477–480. Reidel, Dordrecht.

[21] PRESTON, C. J. (1974). *Gibbs States on Countable Sets*. Cambridge Tracts **68**. Cambridge Univ. Press, London.

[22] RAO, B. V. (1970). On Borel structures. *Colloq. Math.* **21** 199–204.

[23] REED, M. and SIMON, B. (1980). *Methods of Modern Mathematical Physics: Functional Analysis*. Academic, New York.

[24] SIEBERT, E. (1979). Statistical experiments and their conical measures. *Z. Wahrsch. verw. Gebiete* **46** 247–258.

GRADUATE SCHOOL OF BUSINESS
STANFORD UNIVERSITY
STANFORD, CALIFORNIA 94305-5015

INSTITUT FÜR MATHEMATISCHE STATISTIK
WESTFÄLISCHE WILHELMS UNIVERSITÄT
EINSTEINSTRASSE 62
4400 MÜNSTER
WEST GERMANY