Minimal sufficiency of order statistics in convex models

Lutz Mattner
Universität Hamburg

17 March 2000

This is a preprint version of an article which appeared in *Proceedings of the American Mathematical Society* **129**, 3401-3411 (2001). Except for this paragraph, the present version is unchanged from 17 March 2000.

Abstract

Let $\mathcal{P}$ be a convex and dominated statistical model on the measurable space $(\mathcal{X}, \mathcal{A})$, with $\mathcal{A}$ minimal sufficient, and let $n \in \mathbb{N}$. Then $\mathcal{A}^{\otimes n}_{\mathrm{sym}}$, the σ-algebra of all permutation invariant sets belonging to the n-fold product σ-algebra $\mathcal{A}^{\otimes n}$, is shown to be minimal sufficient for the corresponding model for n independent observations, $\mathcal{P}^n = \{P^{\otimes n} : P \in \mathcal{P}\}$.

The main technical tool provided and used is a functional analogue of a theorem of Grzegorek (1982) concerning generators of $\mathcal{A}^{\otimes n}_{\mathrm{sym}}$.

1 Introduction and main results

1.1 Aim. Perhaps the most natural first step in the analysis of a statistical model consists in determining a minimal sufficient σ-algebra. Since this is not always easy, it appears to be worthwhile to supply theorems yielding minimal sufficient σ-algebras for statistically interesting classes of models. Restricting attention to models for independent and identically distributed observations, we note that the case of exponential families is well understood (see, for example, Theorem 1.6.9 in Pfanzagl (1994) and, concerning a possible misinterpretation, Theorem 2.3 of Mattner (1999a)) and that the case of location-scale parameter models on the real line has been treated in Mattner (1999b).

The aim of the present paper is to prove, roughly speaking, that the order statistic is always minimal sufficient in the case of independent and identically distributed observations from convex models (Theorem 1.5). Previously, this was known for boundedly complete convex models only (see Remarks 1.7 b) and c)).

[0] *1991 Mathematics Subject Classification:* 62B05, 62G30, 28A35.
[0] *Keywords and phrases:* Comparison of σ-algebras, nonparametric models, permutation invariance, symmetric sets.

1.2 Guide. The main result of the present paper is the "if" statement in Theorem 1.5. The corresponding "only if" statement does not depend on convexity and is hence stated separately as Theorem 1.4. The latter Theorem seems intuitively obvious but is apparently nontrivial to prove. Theorem 1.5 is illustrated by proving Theorem 1.6. Some remarks and references to the literature are collected in 1.7.

Proofs of Theorems 1.4, 1.5 and 1.6 are supplied in Section 3. The necessary auxiliary facts are presented in Section 2. Of these, the result of 2.5 a), on the generation of the σ-algebra of all permutation invariant sets, is of central importance for the present paper and appears to be of independent interest.

1.3 Notation. In this paper, a statistical model (or experiment) $\mathcal{P}$ on a measurable space $(\mathcal{X}, \mathcal{A})$ is a subset of $\mathrm{Prob}(\mathcal{X}, \mathcal{A})$, the set of all probability measures on $(\mathcal{X}, \mathcal{A})$. (This deviation from the more usual convention, of requiring $\mathcal{P}$ to be an indexed family of probability measures, ensures that the phrase "$\mathcal{P}$ convex" makes sense.) We assume as known, for example from Chapter 1 of Torgersen (1991), the usual definitions and basic facts concerning dominatedness of $\mathcal{P}$, and concerning sufficiency, minimal sufficiency, and bounded completeness of sub-σ-algebras of $\mathcal{A}$. For sets $C, D \in \mathcal{A}$, we write $C = D\,[\mathcal{P}]$ or $C =_{\mathcal{P}} D$ if the symmetric difference $C \vartriangle D$ has P-measure zero for every $P \in \mathcal{P}$. Similarly, for sub-σ-algebras $\mathcal{C}, \mathcal{D}$ of $\mathcal{A}$, we write $\mathcal{C} \subset \mathcal{D}\,[\mathcal{P}]$ or $\mathcal{C} \subset_{\mathcal{P}} \mathcal{D}$ if for every $C \in \mathcal{C}$ there is a $D \in \mathcal{D}$ with $C = D\,[\mathcal{P}]$. We write $\mathcal{P} \ll \mu$ if $\mathcal{P}$ is dominated by μ, which means that μ is a σ-finite measure on $(\mathcal{X}, \mathcal{A})$ and that every $P \in \mathcal{P}$ has a density with respect to μ.

Given the model $\mathcal{P}$ on $(\mathcal{X}, \mathcal{A})$ (corresponding to sample size 1), we write $\mathcal{P}^n$ for the model $\{P^{\otimes n} : P \in \mathcal{P}\}$ of n-fold product measures on the n-fold product measurable space $(\mathcal{X}^n, \mathcal{A}^{\otimes n})$ (corresponding to arbitrary sample size $n \in \mathbb{N}$). If now $\mathcal{C}$ is a sub-σ-algebra of $\mathcal{A}$, then $\mathcal{C}^{\otimes n}_{\mathrm{sym}}$ denotes the σ-algebra of all permutation invariant sets belonging to the n-fold product (or power) σ-algebra $\mathcal{C}^{\otimes n}$. More explicitly, let $\mathfrak{S}_n$ denote the group of all permutations of the set $\{1, \ldots, n\}$, and let us put $x_\pi := (x_{\pi(1)}, \ldots, x_{\pi(n)})$ for $x \in \mathcal{X}^n$ and $\pi \in \mathfrak{S}_n$. Then

$$\mathcal{C}^{\otimes n}_{\mathrm{sym}} \;:=\; \big\{C \in \mathcal{C}^{\otimes n} : x \in C,\, \pi \in \mathfrak{S}_n \Rightarrow x_\pi \in C\big\}.$$

In Sections 2 and 3 below we also use the notation $(\bigotimes_{i=1}^n f_i)(x) := \prod_{i=1}^n f_i(x_i)$ for $x \in \times_{i=1}^n \mathcal{X}_i$ and functions $f_i : \mathcal{X}_i \to \mathbb{R}$, and accordingly $f^{\otimes n} := \bigotimes_{i=1}^n f$.

1.4 Theorem ($\mathcal{P}$ arbitrary). *Let $\mathcal{P}$ be a dominated model on the measurable space $(\mathcal{X}, \mathcal{A})$, let $\mathcal{C} \subset \mathcal{A}$ be a σ-algebra, and let $n \in \mathbb{N}$. If $\mathcal{C}^{\otimes n}_{\mathrm{sym}}$ is minimal sufficient for $\mathcal{P}^n$, then $\mathcal{C}$ is minimal sufficient for $\mathcal{P}$.*

1.5 Theorem ($\mathcal{P}$ convex). *Let $\mathcal{P}$ be a convex and dominated model on the measurable space $(\mathcal{X}, \mathcal{A})$, let $\mathcal{C} \subset \mathcal{A}$ be a σ-algebra, and let $n \in \mathbb{N}$. Then $\mathcal{C}^{\otimes n}_{\mathrm{sym}}$ is minimal sufficient for $\mathcal{P}^n$ iff $\mathcal{C}$ is minimal sufficient for $\mathcal{P}$.*

1.6 Theorem ($\mathcal{P}$ defined by moment conditions). *Let μ be an atomless and σ-finite measure on the measurable space $(\mathcal{X}, \mathcal{A})$, let $\mathcal{U}$ be a finite set of $\mathcal{A}$-measurable $\mathbb{R}$-valued functions on $\mathcal{X}$, let $(c_u : u \in \mathcal{U})$ be a family of real numbers, and assume that*

$$(1) \quad \mathcal{P} \;:=\; \left\{P \in \mathrm{Prob}(\mathcal{X}, \mathcal{A}) : P \ll \mu,\, \int |u|\, dP < \infty,\, \int u\, dP = c_u \quad (u \in \mathcal{U})\right\}$$

is nonempty. Then, for every $n \in \mathbb{N}$, $\mathcal{A}_{\mathrm{sym}}^{\otimes n}$ is minimal sufficient for $\mathcal{P}^n$.

1.7 Remarks. **a)** Do Theorems 1.4 and 1.5 remain true if the assumption of dominatedness is omitted? The following result of Landers (1972), although inconclusive in this respect, suggests that the answer could be in the negative: For a nondominated model $\mathcal{P}$, the existence of a minimal sufficient σ-algebra neither implies nor is implied by the existence of a minimal sufficient σ-algebra for $\mathcal{P}^2$. See Examples 2 and 4 of Landers (1972).

b) Theorem 1.5 is analogous to the known theorem which results if "minimal sufficient" is replaced by the stronger "boundedly complete and sufficient" in both cases, and if the condition of dominatedness is omitted. See Pfanzagl (1994), pages 19-21, and, for extensions and further references, Mandelbaum & Rüschendorf (1987) and Mattner (1996). It is somewhat surprising, in view of the literature just indicated, that the question leading to Theorem 1.5 has apparently not been treated before.

c) A very simple example to which Theorem 1.5 applies, via its corollary Theorem 1.6, while the theorem mentioned in b) does not, is given by $\mathcal{P} :=$ set of all probability measures on $\mathbb{R}$ with Lebesgue-density and with median zero: Here 1.6 immediately yields minimal sufficiency of the order statistic (which is not boundedly complete), for every sample size $n \in \mathbb{N}$.

d) Some scholars seem to be inclined to think that, under suitable "regularity conditions" on $\mathcal{P}$, the two properties

(2) $\mathcal{A}_{\mathrm{sym}}^{\otimes n}$ is not minimal sufficient for $\mathcal{P}^n \quad (n \geq n_0(\mathcal{P}))$,

(3) $\mathcal{P}$ is an exponential family

should be equivalent, and that hence a result like the "if" statement in Theorem 1.5 could perhaps be proved easily, subject to the "regularity conditions", by proving nonexponentiality of $\mathcal{P}$. Indeed, for many familiar examples of models $\mathcal{P}$, either both or neither of (2) and (3) hold. Nevertheless the two properties are known to be independent:

To prove (2) $\not\Rightarrow$ (3), we may use the following beautiful example, indicated on page 18 of Torgersen (1965) and explained in more detail in paragraphs 1.12 b) and c) of Mattner (1999b): Let f be a probability density with respect to Lebesgue measure $\boldsymbol{\lambda}$ on $\mathbb{R}$, such that $f = f_1 f_2$ with f_1 a normal density and f_2 periodic, and let $\mathcal{P}$ be the corresponding location parameter model, $\mathcal{P} = \{f(\cdot - \vartheta)\boldsymbol{\lambda} : \vartheta \in \mathbb{R}\}$. Then, for every $n \geq 2$, the order statistic is not minimal sufficient for $\mathcal{P}^n$ but, except for very special and explicitly known f_2, the model $\mathcal{P}$ is not an exponential family.

To prove (3) $\not\Rightarrow$ (2), the most elementary example is the family of Bernoulli distributions $\mathcal{P} = \{B_p = (1-p)\delta_0 + p\delta_1 : p \in]0,1[\}$, with $\mathcal{X} = \{0,1\}$ and $\mathcal{A} =$ power set of $\mathcal{X}$. This $\mathcal{P}$ is an exponential family but, for example by convexity of $\mathcal{P}$ and by applying Theorem 1.5, has $\mathcal{A}_{\mathrm{sym}}^{\otimes n}$ minimal sufficient for every $n \in \mathbb{N}$.

A more sophisticated example is supplied by Theorem 2.3 of Mattner (1999a). It shows that the implication (3) $\Rightarrow$ (2) can fail to hold even for a smooth one-parameter model on $\mathbb{R}$ with continuous Lebesgue densities.

e) It is perhaps appropriate to recall the standard practical interpretation of minimal sufficiency of the order statistic for a model $\mathcal{P}^n$: If one uses observations $x_1, \ldots, x_n \in \mathcal{X}$ modelled by $\mathcal{P}^n$ *only* to perform a statistical procedure (such as an estimator, or a confidence interval, or a test) from which the observations can not be recovered up to their order (and this is the case for many standard statistical procedures), then one loses information contained in the observations. For example, just a knowledge of some good estimator of one quantity of interest together with the estimators value attained at the observations may be useless for computing the value of any good estimator of another quantity of interest.

2 Auxiliary facts

2.1 A functional Sierpiński lemma with possibly unbounded functions in the generator. *Let $\mathcal{X}$ be a set and let $\mathcal{F}$ and $\mathcal{G}$ be sets of $\mathbb{R}$-valued functions on $\mathcal{X}$. Assume that $\mathcal{F} \subset \mathcal{G}$, $\mathcal{F}$ is stable under multiplication, $\mathcal{G}$ is a vector space with $1 \in \mathcal{G}$, and $\mathcal{G}$ is closed with respect to pointwise sequential convergence. Then every $\mathcal{F}$-measurable $\mathbb{R}$-valued function belongs to $\mathcal{G}$.*

Remark. This must be well known, but I am not aware of any reference.

Sketch of proof. One may proceed in 6 steps: 1. One may assume that $\mathcal{F}$ is countable, $\mathcal{F} = \{f_k : k \in \mathbb{N}\}$. (Compare Halmos (1950), page 24, Theorem D.) 2. For every polynomial function $p : \mathbb{R}^n \to \mathbb{R}$, we have $p(f_1, \ldots, f_n) \in \mathcal{G}$. 3. For every continuous function $F : \mathbb{R}^n \to \mathbb{R}$, we have $F(f_1, \ldots, f_n) \in \mathcal{G}$. 4. For every continuous function $F : \mathbb{R}^{\mathbb{N}} \to \mathbb{R}$, we have $F(f_k : k \in \mathbb{N}) \in \mathcal{G}$. 5. For every measurable function $F : \mathbb{R}^{\mathbb{N}} \to \mathbb{R}$, we have $F(f_k : k \in \mathbb{N}) \in \mathcal{G}$. (By Doob (1994), page 59, Theorem.) 6. By the measurable factorization theorem (Dudley (1989), Theorem 4.2.8), every $\mathcal{F}$-measurable function belongs to $\mathcal{G}$. ∎

2.2 Generation of symmetric σ-algebras. *Let $\mathcal{X}$ be a set, let $\mathcal{F}$ be a set of $\mathbb{R}$-valued functions on $\mathcal{X}$, and let $\mathcal{A} := \sigma(\mathcal{F})$ be the σ-algebra generated by $\mathcal{F}$. Let further Γ be a finite group of $\mathcal{A}$-measurable transformations operating on $\mathcal{X}$, and put $\mathcal{A}_\Gamma := \{A \in \mathcal{A} : A = \gamma A \ (\gamma \in \Gamma)\}$. If $\mathcal{F}$ is stable under multiplication, then*

$$\mathcal{A}_\Gamma \quad = \quad \sigma(\left\{ \sum_{\gamma \in \Gamma} f \circ \gamma \ : \ f \in \mathcal{F} \right\}). \tag{4}$$

Remark. This would be wrong if the assumption "$\mathcal{F}$ stable under multiplication" were omitted, as is obvious from the example $\mathcal{X} = \mathbb{R}$, $\mathcal{F} = \{\text{the identiy}\}$, $\Gamma =$ the reflection group.

Proof. Let us denote the right hand side of (4) by $\mathcal{B}$. Clearly, $\mathcal{B} \subset \mathcal{A}_\Gamma$, so it remains to be shown that $\mathcal{A}_\Gamma \subset \mathcal{B}$. Let us put

$$\mathcal{G} \quad := \quad \left\{ g \in \mathbb{R}^{\mathcal{X}} \ : \ \sum_{\gamma \in \Gamma} g \circ \gamma \ \mathcal{B}\text{-measurable} \right\}.$$

Then $\mathcal{X}$, $\mathcal{F}$ and $\mathcal{G}$ fulfill the assumptions of 2.1. Hence every $\mathcal{A}$-measurable $\mathbb{R}$-valued function belongs to $\mathcal{G}$. Let now g be an $\mathcal{A}_\Gamma$-measurable $\mathbb{R}$-valued function. Then g is $\mathcal{A}$-measurable and Γ-invariant. Hence $g \in \mathcal{G}$ and thus $g = \sum_{\gamma \in \Gamma} g \circ \gamma$ is $\mathcal{B}$-measurable. ∎

2.3 Generation of product σ-algebras. *Let $n \in \mathbb{N}$ and let, for each $i \in \{1, \ldots, n\}$, $(\mathcal{X}_i, \mathcal{A}_i)$ be a measurable space and $\mathcal{F}_i$ be a set of $\mathbb{R}$-valued functions with $\mathcal{A}_i = \sigma(\mathcal{F}_i)$ and with $1 \in \mathcal{F}_i$. Then $\otimes_{i=1}^n \mathcal{A}_i = \sigma(\otimes_{i=1}^n f_i \ : \ f_i \in \mathcal{F}_i$ for $i = 1, \ldots, n)$.*

Proof. One easily shows that the right hand side contains every set of the form $\mathcal{X}_1 \times \ldots \times \mathcal{X}_{i-1} \times A_i \times \mathcal{X}_{i+1} \times \ldots \times \mathcal{X}_n$ with $A_i \in \mathcal{A}_i$. ∎

2.4 The "first main theorem" on symmetric polynomials in vector variables. *Let K be a field of characteristic zero, let $m, n \in \mathbb{N}$, and let $p \in K[y_{i,j} : 1 \le i \le m,\, 1 \le j \le n]$ be a polynomial. Then p has the invariance property*

$$(5)\quad p(y_{i,\pi(j)} : 1 \le i \le m,\, 1 \le j \le n) \;=\; p(y_{i,j} : 1 \le i \le m,\, 1 \le j \le n) \qquad (\pi \in \mathfrak{S}_n)$$

iff p can be written as a polynomial in the polynomials

$$\varphi_{\nu,\alpha_1,\ldots,\alpha_\nu}(y_{i,j} : 1 \le i \le m,\, 1 \le j \le n) \;=\; \sum_{\substack{\beta_1,\ldots,\beta_\nu \in \{1,\ldots,n\} \\ \text{pairwise different}}} \prod_{l=1}^{\nu} y_{\alpha_l,\beta_l}$$

with $\nu \in \{1, \ldots, n\}$ and $\alpha_1, \ldots \alpha_\nu \in \{1, \ldots, m\}$.

Remark. To avoid repetitions, one could impose the condition $\alpha_1 \le \ldots \le \alpha_\nu$.

Proof. This is stated and proved on pages 37-39 of Weyl (1946). ■

2.5 Generation of the σ-algebra of all permutation invariant sets in a power σ-algebra. *Let $n \in \mathbb{N}$, let $(\mathcal{X}, \mathcal{A})$ be a measurable space, and let $\mathcal{F}$ be a set of $\mathbb{R}$-valued functions on $\mathcal{X}$ with $\mathcal{A} = \sigma(\mathcal{F})$ and $1 \in \mathcal{F}$.*

a) *Then*

$$(6)\qquad \mathcal{A}^{\otimes n}_{\mathrm{sym}} \;=\; \sigma\Big(\mathcal{X}^n \ni x \mapsto \sum_{\pi \in \mathfrak{S}_n} \Big(\bigotimes_{i=1}^{n} f_i\Big)(x_\pi) \;:\; f_1, \ldots, f_n \in \mathcal{F}\Big).$$

b) *If $\mathcal{F}$ is convex, then*

$$(7)\qquad \mathcal{A}^{\otimes n}_{\mathrm{sym}} \;=\; \sigma(f^{\otimes n} : f \in \mathcal{F}).$$

Remark. The result (6) of part a) is analogous to Theorem 1 of Grzegorek (1982), which states that

$$(8)\qquad \mathcal{A}^{\otimes n}_{\mathrm{sym}} \;=\; \sigma\Big(\Big\{ \bigcup_{\pi \in \mathfrak{S}_n} \mathop{\times}_{i=1}^{n} A_{\pi_i} \;:\; A_1, \ldots, A_n \in \mathcal{A} \Big\}\Big).$$

It seems that neither of (6) and (8) is easily deducible from the other. In this paper, (6) is used twice, once in the proof of 2.5 b), which is the main ingredient in the proof 3.2 of Theorem 1.5, and once in the proof of 2.7. In the latter case, but apparently not in the former, we could have used Grzegorek's theorem instead.

Proof. a) Let us put $m_1 := n$. We first apply 2.2, with the present $(\mathcal{X}^n, \mathcal{A}^{\otimes n})$ in place of $(\mathcal{X}, \mathcal{A})$, and with $\Gamma := \mathfrak{S}_n$, operating on $\mathcal{X}^n$ via $(\pi, x) \mapsto x_\pi$. The role of the generator of $\mathcal{A}^{\otimes n}$ is played by

$$\Big\{ \prod_{i_2=1}^{m_2} \bigotimes_{i_1=1}^{m_1} f_{i_1,i_2} \;:\; m_2 \in \mathbb{N},\, f_{i_1,i_2} \in \mathcal{F} \Big\},$$

which is trivially stable under multiplication and indeed generates $\mathcal{A}^{\otimes n}$ by 2.3. Thus 2.2 yields

$$(9)\qquad \mathcal{A}^{\otimes n}_{\mathrm{sym}} \;=\; \sigma\Big(\mathcal{X}^n \ni x \mapsto \sum_{\pi \in \mathfrak{S}_n} \prod_{i_2=1}^{m_2} \Big(\bigotimes_{i_1=1}^{m_1} f_{i_1,i_2}\Big)(x_\pi) \;:\; f_{i_1,i_2} \in \mathcal{F}\Big).$$

Let us fix $m_2 \in \mathbb{N}$ and put

$$M := \{1,\ldots,m_1\} \times \{1,\ldots,m_2\}, \quad N := \{1,\ldots,n\}.$$

Let us further fix $f_{i_1,i_2} \in \mathcal{F}$ for $(i_1,i_2) \in M$. Then

$$\sum_{\pi\in\mathfrak{S}_n} \prod_{i_2=1}^{m_2} \left(\bigotimes_{i_1=1}^{m_1} f_{i_1,i_2} \right)(x_\pi) \;=\; p(f_{i_1,i_2}(x_j) : (i_1,i_2) \in M,\, j \in N) \tag{10}$$

where p is a polynomial in $m_1 m_2 n$ variables, namely

$$p(y_{i_1,i_2,j}) \;=\; \sum_{\pi\in\mathfrak{S}_n} \prod_{(i_1,i_2)\in M} y_{i_1,i_2,\pi(i_1)},$$

which satisfies (5) for $m := m_1 m_2$, except for the irrelevant difference that the indexing $i \in \{1,\ldots,m\}$ in 2.4 is here replaced by $(i_1,i_2) \in M$. Thus an application of 2.4 shows that the left hand side of (10) can be written as a polynomial in the functions

$$\sum_{\substack{\beta_1,\ldots,\beta_\nu\in\{1,\ldots,n\} \\ \text{pairwise different}}} \prod_{l=1}^{\nu} f_{\alpha_{1,l},\alpha_{2,l}}(x_{\beta_l}) \qquad (\nu \in N,\, (\alpha_{1,l},\alpha_{2,l}) \in M). \tag{11}$$

Since $1 \in \mathcal{F}$, the sum in (11) can be written as

$$\frac{1}{(n-\nu)!} \sum_{\pi\in\mathfrak{S}_n} \prod_{i=1}^{n} f_i(x_{\pi_i})$$

with $f_1 := f_{\alpha_{1,1},\alpha_{2,1}}, \ldots, f_\nu := f_{\alpha_{1,\nu},\alpha_{2,\nu}}$ and $f_{\nu+1} := \ldots := f_n := 1$ all belonging to $\mathcal{F}$. Hence the left hand side of (10) can be written as a polynomial function of functions occurring in the generator of the right hand side of (6). It follows that the left hand side of (10) is measurable with respect to the right hand side of (6). In view of (9), this proves (6).

b) Let $\mathcal{E}$ denote the right hand side of (7). Clearly, $\mathcal{A}^{\otimes n}_{\mathrm{sym}} \supset \mathcal{E}$. To prove the converse inclusion, let $f_1,\ldots,f_n \in \mathcal{F}$. Then, for every $\lambda = (\lambda_1,\ldots,\lambda_n) \in [0,\infty[^n$ with $\sum \lambda_j = 1$, the function $(\sum_{j=1}^n \lambda_j f_j)^{\otimes n}$ is $\mathcal{E}$-measurable. The latter clearly remains true if the condition $\sum \lambda_j = 1$ is omitted. Expanding in powers of λ, that is

$$(\sum_{j=1}^n \lambda_j f_j)^{\otimes n}(x) \;=\; \sum_{\substack{\alpha\in\{0,\ldots,n\}^n \\ \alpha_1+\ldots+\alpha_n=n}} \lambda^\alpha \cdot \sum_{\pi\in\mathfrak{S}_n} f_1(x_{\pi_1})\cdot\ldots\cdot f_1(x_{\pi_{\alpha_1}})\cdot\ldots\cdot f_n(x_{\pi_{n-\alpha_n+1}})\cdot\ldots\cdot f_n(x_{\pi_n}),$$

and differencing once with respect to each λ_i and then setting $\lambda = (0.\ldots,0)$ yields the $\mathcal{E}$-measurability of the coefficient of $\lambda_1\cdot\ldots\cdot\lambda_n$ in $(\sum_{j=1}^n \lambda_j f_j)^{\otimes n}$. This coefficient is the function $\mathcal{X}^n \ni x \mapsto \sum_{\pi\in\mathfrak{S}_n} \left(\bigotimes_{i=1}^n f_i\right)(x_\pi)$. Using (6) from part a), we conclude that $\mathcal{A}^{\otimes n}_{\mathrm{sym}} \subset \mathcal{E}$. ∎

2.6 Comparison of σ-algebras modulo dominated families of measures via comparison modulo pairs. *Let $\mathcal{P}$ be a dominated model on $(\mathcal{X}, \mathcal{A})$, and let $\mathcal{C}, \mathcal{D} \subset \mathcal{A}$ be σ–algebras. Then $\mathcal{C} \subset \mathcal{D}\,[\mathcal{P}]$ iff, for every choice of $P, Q \in \mathcal{P}$, we have $\mathcal{C} \subset \mathcal{D}\,[\{P, Q\}]$.*

Remark. This follows from the proof of Proposition 1 in Le Bihan, Littaye–Petit & Petit (1970). The validity of that proof is not affected by the counterexamples of Luschgy (1978) and Siebert (1979) to the Théorème in Le Bihan, Littaye–Petit & Petit (1970). For the convenience of the reader, we nevertheless provide a shorter and more elementary version of the cited proof.

Proof. The "only if" part is trivial. So let us assume that $\mathcal{C} \subset \mathcal{D}\,[\mathcal{Q}]$ for every $\mathcal{Q} \subset \mathcal{P}$ with cardinality $|\mathcal{Q}| \le 2$. By dominatedness, there exists a countable set $\mathcal{P}_0 \subset \mathcal{P}$ having the same nullsets as $\mathcal{P}$. So we have to show that

$$\mathcal{C} \;\subset\; \mathcal{D}\,[\mathcal{P}_0]. \tag{12}$$

Let $C \in \mathcal{C}$. For every $\mathcal{Q} \subset \mathcal{P}_0$ with $|\mathcal{Q}| \le 2$, we may choose a set $D_\mathcal{Q} \in \mathcal{D}$ with $C = D_\mathcal{Q}\,[\mathcal{Q}]$, and we may put

$$\begin{aligned} D_P &:= \bigcap \{D_\mathcal{Q} : \mathcal{Q} \subset \mathcal{P}_0,\, |\mathcal{Q}| \le 2,\, P \in \mathcal{Q}\} \qquad (P \in \mathcal{P}_0), \\ D &:= \bigcup_{P \in \mathcal{P}_0} D_P. \end{aligned}$$

(Caution: D_P and $D_{\{P\}}$ are defined entirely differently.) Then $D \in \mathcal{D}$. For every $P \in \mathcal{P}_0$, we have $C =_P D_P \subset D$. Hence $C \subset D\,[\mathcal{P}_0]$. For every $\{P_0, P\} \subset \mathcal{P}_0$, we have $D_P \subset D_{\{P_0,P\}} =_{P_0} C$, and thus $D \subset C\,[P_0]$. Hence also $D \subset C\,[\mathcal{P}_0]$. Thus (12) is proved. ∎

2.7 Comparison of power σ-algebras versus comparison of their symmetrizations. *Let $\mathcal{P}$ be a dominated model on the measurable space $(\mathcal{X}, \mathcal{A})$, let $\mathcal{C}, \mathcal{D} \subset \mathcal{A}$ be σ–algebras , and let $n \in \mathbb{N}$. Then $\mathcal{C}^{\otimes n}_{\mathrm{sym}} \subset \mathcal{D}^{\otimes n}_{\mathrm{sym}}\,[\mathcal{P}^n]$ iff $\mathcal{C} \subset \mathcal{D}\,[\mathcal{P}]$.*

Remark. I do not know whether the assumption of dominatedness, used below in the proof of the "only if" part, may be omitted without substitute.

Proof. "if": An application (6) from 2.5 a), with $\mathcal{F} = \{1_C : C \in \mathcal{C}\}$, yields

$$\mathcal{C}^{\otimes n}_{\mathrm{sym}} \;=\; \sigma\Big(\mathcal{X}^n \ni x \mapsto \sum_{\pi \in \mathfrak{S}_n} \Big(\bigotimes_{i=1}^{n} 1_{C_i}\Big)(x_\pi) \;:\; C_1, \ldots, C_n \in \mathcal{C}\Big)$$

and a similar representation of $\mathcal{D}^n_{\mathrm{sym}}$. Since, as is well known, inclusion modulo nullsets of σ-algebras follows from inclusion modulo nullset of any generating set of functions, the assumption $\mathcal{C} \subset \mathcal{D}\,[\mathcal{P}]$ easily yields $\mathcal{C}^{\otimes n}_{\mathrm{sym}} \subset \mathcal{D}^{\otimes n}_{\mathrm{sym}}\,[\mathcal{P}^n]$.

"only if": By dominatedness and by 2.6, we may assume that $\mathcal{P} = \{P_1, P_2\}$. Let $C_0 \in \mathcal{C}$. We have to construct a $D_0 \in \mathcal{D}$ with $C_0 = D_0\,[\mathcal{P}]$. Put $C := \times_{i=1}^{d} C_0$. By assumption, there is a $D \in \mathcal{D}^n$ with $D = C\,[\mathcal{P}^n]$. For $y \in \times_{i=2}^{n} \mathcal{X}$, let $D^y := \{x \in \mathcal{X} : (x, y) \in D\}$.

If $P \in \mathcal{P}$ is fixed, then, by Fubini, for P^{n-1}–almost every $y \in C_0^{n-1}$

$$C_0 = D^y\,[P]. \tag{13}$$

This would easily imply the claim if $\mathcal{P}$ were dominated by one of its elements. The general case seems to necessitate a more complicated argument, such as the following.

Put $\mu := P_1 + P_2$ and choose $\mathcal{C}$–measurable densities g_i and $\mathcal{D}$–measurable densities h_i such that, for $i = 1$ and 2,

$$P_i|\mathcal{C} = g_i \cdot \mu|\mathcal{C}, \qquad P_i|\mathcal{D} = h_i \cdot \mu|\mathcal{D}. \tag{14}$$

There is no loss in generality in assuming that one of the following three inclusions holds:

$$C_0 \subset \{g_1 = 0, g_2 = 0\}, \tag{15}$$
$$C_0 \subset \{g_1 > 0, g_2 > 0\}, \tag{16}$$
$$C_0 \subset \{g_1 > 0, g_2 = 0\}. \tag{17}$$

(Otherwise cut C_0 down to each of the above three sets and to a fourth set similar to the third, construct sets D_0 separately in each case, and join.)

Also, we may assume that $\mu(C_0) > 0$, since otherwise we may take $D_0 = \emptyset$.

In case of (15), (14) shows that $P_1(C_0) = P_2(C_0) = 0$, a case just excluded.

Now assume (16) and $\mu(C_0) > 0$. Then the statement involving (13) yields a $y \in \mathcal{X}^{n-1}$ such that, simultaneously, $C_0 = D^y\,[P_1]$ and $C_0 = D^y\,[P_2]$, so that we may take $D_0 = D^y$.

Finally assume (17) and $\mu(C_0) > 0$. Now the statement involving (13) only yields a $y_1 \in \mathcal{X}^{n-1}$ such that $C_0 = D^{y_1}\,[P_1]$. We put

$$D_0 := D^{y_1} \cap \{h_2 = 0\}.$$

Then

$$C_0 \cap \{h_2 = 0\} = D_0\,[P_1] \tag{18}$$

and, in view of $P_2(C_0) = \int_{C_0} g_2\,d\mu = 0 = \int_{D_0} h_2\,d\mu = P_2(D_0)$,

$$C_0 = D_0\,[P_2]. \tag{19}$$

By $P_2(C_0) = 0$ and by the choice of D, we have $P_2^n(D) = 0$, hence

$$\mu^n(D \cap \{h_2 > 0\}^n) = 0.$$

This implies

$$P_1^n(C \cap \{h_2 > 0\}^n) = 0,$$

so that

$$C_0 \cap \{h_2 > 0\} = \emptyset\,[P_1]. \tag{20}$$

Taken together, (18), (20) and (19) yield $C_0 = D_0\,[\mathcal{P}]$. ∎

2.8 Moment conditions, products, and measurability. Let $(\mathcal{X}, \mathcal{A})$ be a measurable space. Below, $\mathcal{L}^\infty(\mathcal{X}, \mathcal{A})$ denotes the set of all $\mathbb{R}$-valued and bounded measurable functions on $\mathcal{X}$. If μ is a measure on $(\mathcal{X}, \mathcal{A})$, then $\mathcal{L}^1(\mathcal{X}, \mathcal{A}, \mu)$ denotes the set of all $\mathbb{R}$-valued and μ-integrable functions on $\mathcal{X}$.

Lemma. *Let μ be an atomless measure on the measurable space $(\mathcal{X}, \mathcal{A})$, let $n \in \mathbb{N}$, let $u_1, \ldots, u_n \in \mathcal{L}^1(\mathcal{X}, \mathcal{A}, \mu)$, and let*

$$\mathcal{F} \;:=\; \left\{ f \in \mathcal{L}^\infty(\mathcal{X}, \mathcal{A}) : \int f u_i\,d\mu = 0 \quad (1 \le i \le n) \right\}.$$

a) *Every $f \in \mathcal{L}^\infty(\mathcal{X},\mathcal{A})$ can be written as a product $f = f_1 f_2$ with $f_1, f_2 \in \mathcal{F}$.*
b) *The σ-algebra generated by $\mathcal{F}$ is $\mathcal{A}$.*

Proof. a) Let $f \in \mathcal{L}^\infty(\mathcal{X},\mathcal{A})$. Choose $g_1, g_2 \in \mathcal{L}^\infty(\mathcal{X},\mathcal{A})$ with $f = g_1 g_2$, for example $g_1 = 1$ and $g_2 = f$. Then

$$M(A) \;:=\; \Big(\int_A g_j u_i \, d\mu \;:\; i \in \{1,\ldots n\},\, j \in \{1,2\}\Big) \qquad (A \in \mathcal{A})$$

defines an $\mathbb{R}^{2n}$-valued atomless measure M on $(\mathcal{X},\mathcal{A})$. The Liapounoff convexity theorem (see Rudin (1991), Theorem 5.5) yields an $A \in \mathcal{A}$ with $M(A) = (1/2)(M(\emptyset) + M(\mathcal{X}))$. For this A, put $f_j := (2\cdot 1_A - 1)g_j$ for $j \in \{1,2\}$. Then $(\int f_j u_i \, d\mu) = 2M(A) - M(\mathcal{X}) = 0$, hence $f_1, f_2 \in \mathcal{F}$, and $f_1 f_2 = f$.

b) If $A \in \mathcal{A}$, then, by a), 1_A is a product of two functions belonging to $\mathcal{F}$, and hence is $\mathcal{F}$-measurable. ■

2.9 Sufficiency and convex closure. Let be $\mathcal{P}$ a model on the measurable space $(\mathcal{X},\mathcal{A})$. The convex closure, with respect to setwise convergence, of $\mathcal{P}$ is, by definition, the smallest set $\tilde{\mathcal{P}}$ of probability measures on $(\mathcal{X},\mathcal{A})$ such that $\tilde{\mathcal{P}}$ is convex and closed with respect to setwise convergence (that is: $(P_j : j \in J)$ a net in $\tilde{\mathcal{P}}$, $P \in \mathrm{Prob}(\mathcal{X},\mathcal{A})$, $\lim_j P_j(A) = P(A)$ $(A \in \mathcal{A}) \Rightarrow P \in \tilde{\mathcal{P}}$). It can obviously obtained from $\mathcal{P}$ by first taking the convex hull $\mathrm{conv}\,\mathcal{P}$ of $\mathcal{P}$, and then taking the limits of all nets in $\mathrm{conv}\,\mathcal{P}$.

Lemma. *Let $\mathcal{P}$ and $\mathcal{Q}$ be models on $(\mathcal{X},\mathcal{A})$ with identical convex closures with respect to setwise convergence. Then the partial orders $\subset_{\mathcal{P}}$ and $\subset_{\mathcal{Q}}$ on the sub-σ-algebras of $\mathcal{A}$ are identical, and $\mathcal{P}$ and $\mathcal{Q}$ have the same sufficient σ-algebras and the same, if any, minimal sufficient σ-algebras.*

Proof. Follows easily from the definitions. For the statement concerning "sufficient", one uses the fact that setwise convergence of a net (P_j) to P implies $\lim_j P_j \varphi = P\varphi$ for every bounded measurable function φ. ■

3 Proofs of the main results

3.1 Proof of Theorem 1.4. Suppose that $\mathcal{A}$ is not minimal sufficient for $\mathcal{P}$. Then, by definition, there is a σ–algebra $\mathcal{C} \subset \mathcal{A}$, which is sufficient for $\mathcal{P}$ and does not satisfy $\mathcal{C} = \mathcal{A}\,[\mathcal{P}]$. It easily follows that $\mathcal{C}^{\otimes n}_{\mathrm{sym}} \subset \mathcal{A}^{\otimes n}_{\mathrm{sym}}$ is sufficient for $\mathcal{P}^n$ (one may use Lemma 1 of Landers (1972) here). Also, by Lemma 2.7 applied to the pair $(\mathcal{A},\mathcal{C})$ in place of $(\mathcal{C},\mathcal{D})$, it is not true that $\mathcal{C}^{\otimes n}_{\mathrm{sym}} = \mathcal{A}^{\otimes n}_{\mathrm{sym}}\,[\mathcal{P}^n]$. Hence $\mathcal{A}^{\otimes n}_{\mathrm{sym}}$ is not minimal sufficient for $\mathcal{P}^n$. ■

3.2 Proof of Theorem 1.5. "only if": See Theorem 1.4.

"if": We proceed in two steps.

STEP 1: *Without loss of generality, we may assume that there is a probability measure $\mu \in \mathcal{P}$ which dominates $\mathcal{P}$.*

PROOF. Let $(P_k : k \in \mathbb{N})$ be a sequence in $\mathcal{P}$ such that $\mu := \sum_{k=1}^\infty 2^{-k} P_k$ dominates $\mathcal{P}$, and let $\mathcal{Q}$ be the convex hull of $\mathcal{P} \cup \{\mu\}$. Then $\mathcal{Q}$ is a model satisfying the assumptions of Theorem 1.5, and $\mu \in \mathcal{Q}$ dominates $\mathcal{Q}$. Since $\mathcal{Q} \supset \mathcal{P}$, the claim now follows via the lemma from 2.9, if we show for $m = 1$ and for $m = n$ that $\mathcal{Q}^m$ is contained in the convex closure, with respect to setwise convergence, of $\mathcal{P}^m$.

So let $m \in \mathbb{N}$ and $Q \in \mathcal{Q}$. Then there is $P \in \mathcal{P}$ and $\lambda \in [0,1]$ such that $Q = \lambda P + (1-\lambda)\mu$. (Here convexity of $\mathcal{P}$ is used, but only for convenience.) It follows that

$$
\begin{aligned}
Q^{\otimes m} &= (\lambda P + (1-\lambda)\mu)^{\otimes m} \\
&= (\lambda P + (1-\lambda)\lim_{K\to\infty}(1-2^{-K})\sum_{k=1}^{K} 2^{-k}P_k)^{\otimes m} \\
&= \lim_{K\to\infty}(\lambda P + (1-\lambda)(1-2^{-K})\sum_{k=1}^{K} 2^{-k}P_k)^{\otimes m}, \qquad (21)
\end{aligned}
$$

where the indicated limit relations even hold in total variation, and not merely setwise. Making now essential use of the convexity of $\mathcal{P}$, we observe that the probability measures on the right of "lim" in (21) belong to $\mathcal{P}^m$. Hence $Q^{\otimes m}$ indeed belongs to the convex closure with respect to setwise convergence of $\mathcal{P}^m$.

Step 2: Let $\mathcal{C}$ be minimal sufficient for $\mathcal{P}$. Let μ be as in Step 1. For each $P \in \mathcal{P}$, let f_P be a μ-density of P, with $f_P = 1$ in case $P = \mu$, and let $\mathcal{F} := \{f_P : P \in \mathcal{P}\}$. Then, by Bahadur's version of a theorem of Lehmann & Scheffé [see Torgersen (1991), p. 69], the σ-algebra $\sigma(\mathcal{F})$ is minimal sufficient for $\mathcal{P}$. Hence $\mathcal{C} = \sigma(\mathcal{F})\,[\mathcal{P}]$. By Lemma 2.7, we can assume that

$$\mathcal{C} \;=\; \sigma(\mathcal{F}).$$

Since $\mu^{\otimes n} \in \mathcal{P}^n$, the σ-algebra

$$\mathcal{E}_0 \;:=\; \sigma(f^{\otimes n} : f \in \mathcal{F})$$

is minimal sufficient for $\mathcal{P}^n$. By the convexity of $\mathcal{P}$, we have $\mathcal{E}_0 = \mathcal{E}\,[\mathcal{P}]$ with

$$\mathcal{E} \;:=\; \sigma((\sum_{j=1}^{k}\lambda_j f_j)^{\otimes n} \;:\; k \in \mathbb{N},\, f_j \in \mathcal{F},\, \lambda_j \in [0,1],\, \sum \lambda_j = 1).$$

By 2.5 b) applied to the convex hull $\operatorname{conv}\mathcal{F}$ of $\mathcal{F}$, we have $\mathcal{E} = (\sigma(\operatorname{conv}\mathcal{F}))^{\otimes n}_{\mathrm{sym}} = (\sigma(\mathcal{F}))^{\otimes n}_{\mathrm{sym}} = \mathcal{C}^{\otimes n}_{\mathrm{sym}}$. ■

3.3 Proof of Theorem 1.6. Since $\mathcal{P}$ from (1) is convex and dominated, Theorem 1.5 will yield the stated conclusion if we prove that $\mathcal{A}$ is minimal sufficient for $\mathcal{P}$. This will be done in three steps.

Step 1: There is no loss in assuming that $c_u = 0$ $(u \in \mathcal{U})$.

Proof. Replace each u by $u - c_u$.

Step 2: There is no loss in assuming that $\mu \in \mathcal{P}$.

Proof. Since $\mathcal{P}$ is nonempty and dominated, there is a sequence (P_n) in $\mathcal{P}$ such that $\{P_n : n \in \mathbb{N}\}$ and $\mathcal{P}$ have the same nullsets. Let $b_n := \int \sum_{u\in\mathcal{U}} |u(x)|\, dP_n(x)$ and choose a sequence (ε_n) in $]0,\infty[$ with $\sum_1^\infty \varepsilon_n = 1$ and $\sum_1^\infty \varepsilon_n b_n < \infty$. Then $\sum_1^\infty \varepsilon_n P_n$ belongs to $\mathcal{P}$ and dominates $\mathcal{P}$.

Step 3: We assume that $c_u = 0$ $(u \in \mathcal{U})$ and that $\mu \in \mathcal{P}$. Let us put

$$\mathcal{F} \;:=\; \Big\{ f \in \mathcal{L}^1(\mathcal{X},\mathcal{A},\mu) \;:\; f \geq 0,\, \int f\, d\mu = 1,\, \int f u\, d\mu = 0 \;\; (u \in \mathcal{U}) \Big\}.$$

(The notation $\mathcal{L}^1$ and $\mathcal{L}^\infty$ here and below is as in 2.8.) Using Bahadur's theorem again (compare Step 2 in 3.2), we observe that $\mathcal{C} := \sigma(\mathcal{F})$ is minimal sufficient for $\mathcal{P}$. Put

$$\mathcal{G} \;:=\; \left\{ g \in \mathcal{L}^\infty(\mathcal{X}, \mathcal{A}) \,:\, \int gu\,d\mu = 0 \;\; (u \in \{1\} \cup \mathcal{U}) \right\}.$$

Since $1 \in \mathcal{F}$, we have $\mathcal{F} \supset \{1 + g \,:\, g \in \mathcal{G}, g \geq -1\}$ and hence

$$\mathcal{C} \;\supset\; \sigma(g \,:\, g \in \mathcal{G},\, g \geq -1) \;=\; \sigma(\mathcal{G}).$$

By the lemma from 2.8, we have $\sigma(\mathcal{G}) = \mathcal{A}$. Hence $\mathcal{C} = \mathcal{A}$. ∎

Acknowledgement. The present work was supported by a Heisenberg grant of the Deutsche Forschungsgemeinschaft. I thank O. Riemenschneider for a helpful discussion while writing this paper, and a referee for useful comments.

References

DOOB, J.L. (1994). *Measure Theory.* Springer, New York.

DUDLEY, R.M. (1989). *Real Analysis and Probability.* Wadsworth, Pacific Grove, California.

GRZEGOREK, E. (1982). Symmetric σ–fields and universal null sets. In: *Measure Theory. Oberwolfach 1981.* Lecture Notes in Mathematics 945, Springer, pp. 101-109.

HALMOS, P.R. (1950). *Measure Theory.* Van Nostrand. Reprinted 1974 by Springer, New York.

LE BIHAN, M.-F., LITTAYE–PETIT, M. & PETIT, J.-L. (1970). Exhaustivité par paire. *C. R. Acad. Sci. Paris* **270**, Sér. A, 1753-1756.

LANDERS, D. (1972). Sufficient and minimal sufficient σ-fields. *Z. Wahrscheinlichkeitstheorie verw. Gebiete* **23**, 197-207.

LUSCHGY, H. (1978). Sur l'existence d'une plus petite sous-tribu exhaustive par paire. *Ann. Inst. Henri Poincaré*, Section B: *Calcul des Probabilités et Statistique*, **14**, 391-398.

MANDELBAUM, A. & RÜSCHENDORF, L. (1987). Complete and symmetrically complete families of distributions. *Ann. Statist.* **15**, 1229-1244.

MATTNER, L. (1996). Complete order statistics in parametric models. *Ann. Statist.* **24**, 1265-1282.

MATTNER, L. (1999 a). Sufficiency, exponential families, and algebraically independent numbers. *Math. Meth. Statist.* **8**, 397-406.

MATTNER, L. (1999 b). Minimal sufficienct statistics in location-scale parameter models. Preprint.

PFANZAGL, J. (1994). *Parametric Statistical Theory.* de Gruyter, Berlin.

RUDIN, W. (1991). *Functional Analysis.* 2nd. ed. McGraw–Hill, New York.

SIEBERT, E. (1979). Pairwise sufficiency. *Z. Wahrscheinlichkeitstheorie verw. Gebiete* **46**, 237-246.

TORGERSEN, E. (1965). Minimal sufficiency of order statistics in the case of translation- and scale parameters. *Skand. Aktuarietidsskrift* **48**, 16-21.

TORGERSEN, E. (1991). *Comparison of Statistical Experiments.* Cambridge University Press, Cambridge.

WEYL, H. (1946). *The Classical Groups. Their Invariants and their Representations.* 2nd. Ed. Princeton University Press, Princeton.

UNIVERSITÄT HAMBURG, INSTITUT FÜR MATHEMATISCHE STOCHASTIK, BUNDESSTR. 55, 20146 HAMBURG, GERMANY

E–mail address: `mattner@math.uni-hamburg.de`