

Problems Relating to the Existence of Maximal and
Minimal Elements in Some Families of Statistics (Sub-fields)

by

D. Basu

University of North Carolina
and
Indian Statistical Institute

Institute of Statistics Mimeo Series No. 435

June 1965

This research was supported by the Air Force Office of
Scientific Research Contract No. AF-AFOSR-760-65.

DEPARTMENT OF STATISTICS
UNIVERSITY OF NORTH CAROLINA
Chapel Hill, N. C.

Problems Relating to the Existence of Maximal and
Minimal Elements in Some Families of Statistics (Sub-fields)

by

D. Basu

(University of North Carolina
and
Indian Statistical Institute)

1. SUMMARY:

In statistical theory one comes across various families of statistics (sub-fields). For each such family, it is of some interest to ask oneself as to whether the family has maximal and/or minimal elements. The author proves here the existence of such elements in a number of cases and leaves the question unsolved in a number of other cases. A number of problems of an allied nature are also discussed.

2. INTRODUCTION:

Let (X, G, P) be a given probability structure (or statistical model). A statistic is a measurable transformation of (X, G) to some other measurable space. Each such statistic induces, in a natural manner, a sub-field* of G and is, indeed, identifiable with the induced sub-field.

Between sub-fields of G there exists the following natural partial order of inclusion relationship.

Definition 1: The sub-field G_1 is said to be larger than the sub-field G_2 if every member of G_2 is also a member of G_1 .

* abbreviation for sub- σ -field

A slightly weaker version of the above partial order is the following:

Definition 2: The sub-field G_1 is said to be essentially larger than the sub-field G_2 if every member of G_2 is \mathcal{P} -equivalent to some member of G_1 .

[Two measurable sets A and B are said to be \mathcal{P} -equivalent if their symmetric difference $A \Delta B$ is \mathcal{P} -null for each $P \in \mathcal{P}$.]

Given a family \mathfrak{F} of sub-field (statistics), one naturally enquires as to whether \mathfrak{F} has a largest and/or least element in the sense of definition 1. In the absence of such elements in \mathfrak{F} , one may enquire about the possible existence of maximal and/or minimal elements. [An element G_0 of \mathfrak{F} is a maximal (minimal) element of \mathfrak{F} , if, there exists no other element G_1 in \mathfrak{F} such that G_1 is larger (smaller) than G_0 .] In the absence of maximal (minimal) elements in \mathfrak{F} , one may look for elements that are essentially largest (least) or are essentially maximal (minimal) in the sense of the weaker partial order of definition 2.

The particular case where \mathfrak{F} is the family of all sufficient sub-fields has received considerable attention. The largest element of \mathfrak{F} is clearly the total sub-field G itself. If \mathcal{P} be a dominated family of measures then it is well-known that \mathfrak{F} has an essentially least element in terms of the weaker partial order of definition 2. In general, \mathfrak{F} does not have even essentially minimal elements. If, however, an essentially minimal element exists then it must be essentially unique and, thus, be the essentially least element of \mathfrak{F} (see Cor. 3 to Theorem 4 in [3]).

In [1] the author considered the family \mathfrak{F} of ancillary sub-fields. [A sub-field G_0 is said to be ancillary if the restriction of the class \mathcal{P} of probability measures to G_0 shrinks the class down to a single probability measure.] The least ancillary sub-field is clearly the trivial sub-field

consisting of only the empty set ϕ and the whole space \mathcal{X} . Existence of maximal elements in the family of ancillary sub-fields was demonstrated in [1]. In general there exist a multiplicity of maximal ancillary sub-fields. In sections 3 to 6 we list four problems that are similar to the problem of ancillary sub-fields. In section 7 we develop a general method to demonstrate the existence of maximal elements in these four cases. In section 8 we discuss some related questions. In section 9 we list a number of other problems.

3. \mathcal{B} -independent sub-fields (\mathcal{F}_1)

Let \mathcal{B} be a fixed sub-field. A sub-field \mathcal{C} is said to be \mathcal{B} -independent (independent of \mathcal{B}) if

$$P(BC) \equiv P(B)P(C)$$

for all $B \in \mathcal{B}$, $C \in \mathcal{C}$ and $P \in \mathcal{P}$.

Let \mathcal{F}_1 be the family of all \mathcal{B} -independent sub-fields. Clearly, the least element of \mathcal{F}_1 is the trivial sub-field. Even in very simple situations, \mathcal{F}_1 has no largest (or essentially largest) element. In section 7 we shall show that \mathcal{F}_1 always has maximal elements. Consider the two examples.

Example 1(a): Let \mathcal{X} consist of the four points a, b, c and d and let \mathcal{P} consist of only one probability measure--the one that allots equal probabilities to the four points. Let \mathcal{B} consist of the four sets

$$\phi, \mathcal{X}, [a,b] \text{ and } [c,d].$$

Then the two sub-fields \mathcal{C}_1 and \mathcal{C}_2 , consisting, respectively, of

$$\mathcal{C}_1: \phi, \mathcal{X}, [a,c] \text{ and } [b,d],$$

$$\mathcal{C}_2: \phi, \mathcal{X}, [a,d] \text{ and } [b,c],$$

are each a maximal \mathfrak{B} -independent sub-field. Incidentally, in this case C_1 and C_2 happen to be independent of each other.

Example 1(b): Let x_1, x_2, \dots, x_n be n independent normal variables with equal unknown means φ and equal unknown standard deviations θ . Let \mathfrak{B} be the sub-field induced by the statistic

$$\bar{x} = (x_1 + x_2 + \dots + x_n)/n$$

and let C be induced by the set of differences

$$D = (x_1 - x_n, x_2 - x_n, \dots, x_{n-1} - x_n).$$

Here C is \mathfrak{B} -independent but is not the largest \mathfrak{B} -independent sub-field. Indeed, in this situation, there are infinitely many maximal elements in \mathfrak{F}_1 (see example 1 in [1]). It is, however, possible to show that C is an essentially maximal element in \mathfrak{F}_1 . In the above example, one may reverse the role of \bar{x} and D and ask oneself as to whether \bar{x} is a maximal statistic that is D -independent. It is of some interest to speculate about the truth or falsity of the following general

Proposition 1: If C be a maximal (or essentially maximal) \mathfrak{B} -independent sub-field, then \mathfrak{B} is a maximal (or essentially maximal) C -independent sub-field.

4. φ -free sub-fields (\mathfrak{F}_2):

Let us suppose that the members of the class \mathfrak{P} are indexed by two independent parameters θ and φ , i.e.,

$$\mathfrak{P} = \{ P_{\theta, \varphi} \mid \theta \in \Theta, \varphi \in \Phi \},$$

the parameter space being the Cartesian-product $\Theta \times \Phi$.

A sub-field C is called φ -free if the restriction of \mathcal{P} to C leads to a class of probability measures that may be indexed by θ alone, i.e. for all $C \in \mathcal{C}$ the probability $P_{\theta, \varphi}(C)$ is a function of θ only. Let \mathcal{F}_2 be the family of all φ -free sub-fields. Evidently, the notion of φ -free sub-fields is a direct generalisation of the notion of ancillary sub-fields.

The trivial sub-field is again the least element of \mathcal{F}_2 . That \mathcal{F}_2 always has maximal elements will be demonstrated later. In general, \mathcal{F}_2 has a plurality of maximal elements.

Example 2(a): Let \mathcal{X} consist of the five points a, b, c, d and e and let $\mathcal{P} = \{P_{\theta, \varphi}\}$ consist of the probability measures

x	:	a	b	c	d	e
$P_{\theta, \varphi}(x)$:	$1-\theta$	$\theta\varphi$	$\theta\varphi$	$\theta(\frac{1}{2} - \varphi)$	$\theta(\frac{1}{2} - \varphi)$

where $0 < \theta < 1$ and $0 < \varphi < \frac{1}{2}$.

There are exactly 12 sub-sets of \mathcal{X} whose probability measure is φ -free, and they are $\mathcal{X}, [a], [b,d], [b,e], [c,d], [c,e]$ and their complements. As these 12 sets do not constitute a sub-field it is clear that there cannot exist a largest element in \mathcal{F}_2 . The two sub-fields C_1 and C_2 consisting respectively of

C_1 : $\mathcal{X}, [a], [b,d], [c,e]$ and their complements

C_2 : $\mathcal{X}, [a], [b,e], [c,d]$ and their complements,

are the two maximal elements of \mathcal{F}_2 .

Example 2(b): Let x_1, x_2, \dots, x_n be n independent and identically distributed variables with a cumulative distribution function (cdf) of the type

$$F\left(\frac{x - \varphi}{\theta}\right), \quad -\infty < \varphi < \infty, \quad 0 < \theta < \infty,$$

where the function F is known and φ, θ are the so-called location and scale parameters.

The sub-field \mathcal{C} generated by the $n-1$ dimensional statistic

$$D = (x_1 - x_n, x_2 - x_n, \dots, x_{n-1} - x_n)$$

is φ -free in the sense defined before. In general, it is not true that \mathcal{C} is the largest element of the family \mathfrak{F}_2 of φ -free sub-fields. In the particular case, where F is the cdf of a normal variable, the sub-field \mathcal{C} may be shown to be an essentially maximal element of \mathfrak{F}_2 . Let us observe that, in this particular case, \mathfrak{F}_2 is the same as \mathfrak{F}_1 of example 1(b). The following proposition may well be true.

Proposition 2: Whatever may be F , the sub-field \mathcal{C} (as defined above) is an essentially maximal element of the family \mathfrak{F}_2 of φ -free (φ being the location parameter) sub-fields.

Suppose in example 2(b) we reverse the role of φ and θ and concern ourselves with the family \mathfrak{F}_2^* of θ -free sub-fields, i.e. with sub-fields every member of which has a probability measure that does not involve the scale parameter θ . The author believes that the following proposition is generally true.

Proposition 3: Every θ -free sub-field is also φ -free, i.e. $\mathfrak{F}_2^* \subset \mathfrak{F}_2$. (In the particular case where F is the cdf of a normal variable, the truth of Proposition 3 has been established in [4].)

5. \mathcal{G} -similar sub-fields (\mathfrak{F}_3):

Let $\mathcal{G} = \{g\}$ be an arbitrary (but fixed) class of measurable transformations of (X, \mathcal{G}) into itself. For each $P \in \mathcal{P}$, the transformation $g \in \mathcal{G}$ induces a probability measure Pg^{-1} on (X, \mathcal{G}) . A sub-field \mathcal{C} will be called

\mathcal{G} -similar if, for each $g \in \mathcal{G}$ and $P \in \mathcal{P}$, the restriction of the two measures P and Pg^{-1} to \mathcal{C} are identical. In other words, \mathcal{C} is \mathcal{G} -similar if for all $C \in \mathcal{C}$

$$Pg^{-1}(C) \equiv P(C) \quad \text{for all } P \in \mathcal{P} \text{ and } g \in \mathcal{G} .$$

Let \mathfrak{F}_3 be the family of all \mathcal{G} -similar sets. One may look upon \mathfrak{F}_3 as the family of sub-fields that are induced by statistics $T(x)$ such that $T(x)$ and $T(gx)$ are identically distributed for each $P \in \mathcal{P}$ and $g \in \mathcal{G}$. The least element of \mathfrak{F}_3 is, of course, the trivial sub-field. As we shall see later, \mathfrak{F}_3 always has maximal elements and, in general, a plurality of them.

Example 3(a):

Let \mathcal{X} be the real line and $\mathcal{P} = \{P_\theta \mid -\infty < \theta < \infty\}$, where P_θ is the uniform distribution over the interval $(\theta, \theta + 1)$. Let \mathcal{G} consist of the single transformation g defined as

$$gx = \text{the fractional part of } x .$$

It is easy to check that for all θ in $(-\infty, \infty)$

$$P_\theta g^{-1} = P_\theta .$$

In this example, the sub-field \mathcal{C} is \mathcal{G} -similar if and only if each member of \mathcal{C} has a probability that is θ -free. Thus, the family \mathfrak{F}_3 of \mathcal{G} -similar sub-fields is the same as the family of ancillary sub-fields. Here, \mathfrak{F}_3 has a largest element and that is the sub-field of all Borel-sets A such that the two sets A and $A+1$ are essentially equal with respect to the Lebesgue measure.

Example 3(b): Let $(\mathcal{X}, \mathcal{G}, \mathcal{P})$ be as in example 2(b) where F is known and φ, θ are the location and scale parameters. Define the shift transformation g_a as

$$g_a(x_1, x_2, \dots, x_n) = (x_1 + a, x_2 + a, \dots, x_n + a)$$

where a is a fixed real number. Let $\mathcal{G} = \{g_a \mid -\infty < a < \infty\}$ be the class of all shift transformations.

Denoting the joint distribution of (x_1, x_2, \dots, x_n) by $P_{\varphi, \theta}$, we note at once that

$$P_{\varphi, \theta} g_a^{-1} = P_{\varphi+a, \theta} .$$

In this example, the family \mathfrak{F}_3 of \mathcal{G} -similar sub-fields is the same as the family \mathfrak{F}_2 of φ -free sub-fields.

Let us call the set A \mathcal{G} -invariant if $A \in \mathcal{G}$ and

$$g^{-1}A = A \quad \text{for all } g \in \mathcal{G} .$$

Likewise, let us call A almost \mathcal{G} -invariant if the two sets $g^{-1}A$ and A are \mathcal{P} -equivalent for all $g \in \mathcal{G}$. Let \mathcal{B}_i and \mathcal{B}_a be, respectively, the class of \mathcal{G} -invariant and almost \mathcal{G} -invariant sets. It is easy to check that \mathcal{B}_i and \mathcal{B}_a are members of the family \mathfrak{F}_3 of \mathcal{G} -similar sub-fields. The following proposition should be provable under some conditions.

Proposition 4: The sub-field \mathcal{B}_a of almost \mathcal{G} -invariant sets is a maximal \mathcal{G} -similar sub-field.

Under some general conditions it should also be true that the sub-field \mathcal{B}_i of \mathcal{G} -invariant sets is an essentially maximal element of \mathfrak{F}_3 . This is so in the case of example 3(b) where F is the cdf of a normal variable.

6. \mathcal{B} -linked sub-fields (\mathfrak{F}_4):

Let \mathcal{B} be a fixed sub-field of \mathcal{G} . A sub-field \mathcal{C} will be called \mathcal{B} -linked if \mathcal{B} be sufficient for $(\mathcal{C}, \mathcal{P})$, i.e., for every $C \in \mathcal{C}$, there exists a \mathcal{B} -measurable mapping $Q(C, \cdot)$ of \mathcal{X} into the unit interval such that, for all $B \in \mathcal{B}$ and $P \in \mathcal{P}$,

$$P(BC) = \int_B Q(C, \cdot) d P(\cdot) .$$

Let \mathfrak{F}_4 be the family of all \mathfrak{B} -linked sub-fields. The trivial sub-field is again the least element of \mathfrak{F}_4 . We shall presently see that \mathfrak{F}_4 always has maximal elements.

Examples 4(a): (i) Let \mathfrak{B} be the trivial sub-field. It is easy to see that, in this instance, \mathfrak{F}_4 is the same as the family of all ancillary sub-fields.

(ii) Let us suppose that \mathfrak{P} is indexed by the parameters φ and θ . Let \mathfrak{B} be a fixed φ -free sub-field, i.e., a member of \mathfrak{F}_2 as defined in section 4. In this instance, every \mathfrak{B} -linked sub-field is also φ -free.

(iii) Let \mathfrak{B} be a sufficient sub-field. In this case \mathfrak{F}_4 is the family of all sub-fields.

Example 4(b): Let $(\mathfrak{X}, \mathfrak{G}, \mathfrak{P})$ be as in example 1(b) and let \mathfrak{B}_0 be the sub-field induced by the sample variance $\Sigma(x_i - \bar{x})^2/n$. If \mathfrak{C} be the sub-field induced by

$$D = (x_1 - x_n, x_2 - x_n, \dots, x_{n-1} - x_n),$$

then it is easy to check that \mathfrak{C} is \mathfrak{B}_0 -linked. Since \mathfrak{B}_0 is φ -free it follows that every \mathfrak{B}_0 -linked sub-field is also φ -free. It is possible to show that \mathfrak{C} is an essentially maximal \mathfrak{B}_0 -linked sub-field. The truth of the following proposition is worth investigating.

Proposition 5: If \mathfrak{B}_0 and \mathfrak{C} be as in example 4(b), then \mathfrak{C} is an essentially largest element of the family \mathfrak{F}_4 of the \mathfrak{B}_0 -linked sub-fields.

7. Existence of maximal elements:

In this section we develop some general methods to prove the existence of maximal elements in the families $\mathfrak{F}_1, \mathfrak{F}_2, \mathfrak{F}_3$ and \mathfrak{F}_4 . Let us first note a common feature of the four families of sub-fields. Each \mathfrak{F}_i ($i = 1, 2, 3, 4$) is

the totality of all sub-fields that can be embedded in a certain class \mathcal{E}_i of measurable sets. This will be clear once we define the four classes $\mathcal{E}_1, \mathcal{E}_2, \mathcal{E}_3$ and \mathcal{E}_4 of measurable sets.

Definitions:

(i) Let \mathcal{E}_1 be the class of all \mathcal{G} -independent (see section 3) sets, i.e. $\mathcal{E}_1 = \{A \mid P(AB) = P(A)P(B), \text{ for all } P \in \mathcal{P}, B \in \mathcal{G}\}$

(ii) Let \mathcal{E}_2 be the class of all φ -free (see section 4) sets, i.e. $\mathcal{E}_2 = \{A \mid P_{\varphi, \theta}(A) \text{ does not involve } \varphi\}$

(iii) Let \mathcal{E}_3 be the class of all \mathcal{Q} -similar (see section 5) sets, i.e. $\mathcal{E}_3 = \{A \mid P(g^{-1}A) = P(A) \text{ for all } P \in \mathcal{P}, g \in \mathcal{Q}\}$.

(iv) Let \mathcal{E}_4 be the class of all \mathcal{G} -linked (see section 6) sets, i.e. $A \in \mathcal{E}_4$ if and only if there exists a \mathcal{G} -measurable mapping $Q(A, \cdot)$ of \mathcal{X} into the unit interval such that

$$P(AB) = \int_B Q(A, \cdot) dP(\cdot) \text{ for all } P \in \mathcal{P} \text{ and } B \in \mathcal{G}.$$

It is now clear that, for $i = 1, 2, 3, 4$.

$$\mathcal{F}_i = \{C \mid C \text{ is a subfield and } C \subset \mathcal{E}_i\},$$

i.e., \mathcal{F}_i is the family of all sub-fields that can be embedded in the class \mathcal{E}_i of measurable sets.

Our first general result is the following:

Theorem 1: Each \mathcal{E}_i ($i = 1, 2, 3, 4$) has the following properties

- (a) $\phi \in \mathcal{E}_i, \mathcal{X} \in \mathcal{E}_i$
- (b) $A \in \mathcal{E}_i, B \in \mathcal{E}_i, A \subset B \implies B - A \in \mathcal{E}_i$.
- (c) \mathcal{E}_i is closed for countable disjoint unions.

The proof of Theorem 1 is routine and hence omitted. An immediate consequence of Theorem 1 is the

Corollary: Each \mathcal{E}_i ($i = 1, 2, 3, 4$) is a monotone class of sets.

The following is our fundamental existence theorems.

Theorem 2: If \mathcal{E} be a given monotone class of sets and \mathcal{F} be the family of all Borel-fields that could be embedded in \mathcal{E} , then corresponding to each element \mathcal{C} of \mathcal{F} there exists a maximal element $\tilde{\mathcal{C}}$ of \mathcal{F} such that $\mathcal{C} \subset \tilde{\mathcal{C}}$.

Proof: Let $\{\mathcal{C}_t \mid t \in T\}$ be an arbitrary linearly ordered (wrt the partial order of inclusion relationship) sub-family of \mathcal{F} and let

$$\mathcal{C}_0 = \bigcup_{t \in T} \mathcal{C}_t$$

Since $\{\mathcal{C}_t\}$ is linearly ordered it follows that \mathcal{C}_0 is a field of sets. The monotone extension of \mathcal{C}_0 is then the same as the Borel-extension \mathcal{C}_1 of \mathcal{C}_0 . Since \mathcal{E} is monotone and $\mathcal{C}_0 \subset \mathcal{E}$, it follows that $\mathcal{C}_1 \subset \mathcal{E}$ and hence $\mathcal{C}_1 \in \mathcal{F}$. Thus, every linearly ordered sub-family of \mathcal{F} has an upper bound in \mathcal{F} . Theorem 2 is then a consequence of Zorn's Lemma.

An immediate consequence of Theorem 1 and 2 is

Theorem 3: For each $\mathcal{C} \in \mathcal{F}_i$ there exists a maximal element $\tilde{\mathcal{C}}$ in \mathcal{F}_i such that $\mathcal{C} \subset \tilde{\mathcal{C}}$, ($i = 1, 2, 3, 4$).

8. Some general results:

Let \mathcal{E} be a class of measurable sets having the same characteristics as those of the classes \mathcal{E}_i in Theorem 1. That is,

- a) $\phi \in \mathcal{E}$, $\mathcal{X} \in \mathcal{E}$
- b) $A \in \mathcal{E}$, $B \in \mathcal{E}$, $A \subset B \implies B - A \in \mathcal{E}$.
- c) \mathcal{E} is closed for countable disjoint unions.

Let \mathcal{F} be the family of all the sub-fields that may be embedded in \mathcal{E} and let \mathcal{F}_0 be the sub-family of all the maximal elements in \mathcal{F} . That \mathcal{F}_0 is not vacuous, has been established in Theorem 2.

Two members A and B of \mathcal{E} are said to 'conform' if $AB \in \mathcal{E}$. The set $A \in \mathcal{E}$ is said to be 'conforming' if $AB \in \mathcal{E}$ for all $B \in \mathcal{E}$. If every member of \mathcal{E} be conforming, then \mathcal{E} must itself be a Borel-field; and hence there is no problem as \mathfrak{F}_0 consists of a single member namely \mathcal{E} itself. A sub-field is 'conforming' if every one of its members is so.

Let \mathfrak{D} be the class of all the conforming sets in \mathcal{E} , i.e.

$$A \in \mathfrak{D} \implies A \in \mathcal{E} \text{ and } AB \in \mathcal{E} \text{ for all } B \in \mathcal{E}.$$

Let \mathcal{M} stand for a typical element of \mathfrak{F}_0 , i.e. \mathcal{M} is a maximal element of \mathfrak{F} . We prove the following

Theorem 4

i) \mathcal{M} is a maximal element of \mathfrak{F} , if and only if, $A \in \mathcal{E} - \mathcal{M}$ implies that A does not conform to at least one member of \mathcal{M} .

ii) \mathfrak{D} is a sub-field and is equal to the intersection of all the maximal elements in \mathfrak{F} . It is the largest conforming sub-field.

iii) \mathcal{C} is a conforming sub-field, if and only if, for all $\mathcal{B} \in \mathfrak{F}$ it is true that $\mathcal{C} \vee \mathcal{B} \in \mathfrak{F}$. [$\mathcal{C} \vee \mathcal{B}$ stands for the least sub-field containing both \mathcal{C} and \mathcal{B} .]

Proof: Let $\mathcal{M} \in \mathfrak{F}_0$ and let A be a fixed member of $\mathcal{E} - \mathcal{M}$. If possible, let A conform to all the members of \mathcal{M} . Consider the class \mathcal{M}^* of sets of the type

$$AM_1 \cup A'M_2, \quad ,$$

where M_1 and M_2 are arbitrary members of \mathcal{M} . It is easy to check that

$\mathcal{M}^* \in \mathfrak{F}$ and that $A \in \mathcal{M}^*$ and $\mathcal{M} \subset \mathcal{M}^*$. This violates the supposition that \mathcal{M} is a maximal element of \mathfrak{F} . Thus, the 'only if' part of (i) is proved.

To prove the 'if' part we have only to observe that if \mathcal{M} is not maximal then there exists a larger sub-field $\mathcal{M}^* \subset \mathcal{E}$ and this implies the existence of an $A \in \mathcal{E} - \mathcal{M}$ that conforms to every member of \mathcal{M} .

Since, every member of \mathcal{D} conforms (by definition) to every member of \mathcal{E} , it is an immediate consequence of (i) that $\mathcal{D} \subset \mathcal{M}$ for each $\mathcal{M} \in \mathcal{F}_0$, i.e. $\mathcal{D} \subset \bigcap \mathcal{M}$.

Now, let M and E be typical members of $\bigcap \mathcal{M}$ and \mathcal{E} respectively. From Theorem 2, there exists a maximal element \mathcal{M}_0 in \mathcal{F} which contains the sub-field consisting of ϕ , E , E' and \mathcal{X} . Thus, M and E are together in the sub-field \mathcal{M}_0 and hence they must conform. Since E is arbitrary, it follows that $M \in \mathcal{D}$. We have thus proved the equality of \mathcal{D} and $\bigcap \mathcal{M}$ and has incidentally proved the equality of \mathcal{E} and $\bigcup \mathcal{M}$. Since each \mathcal{M} is a sub-field it is now clear that $\mathcal{D} = \bigcap \mathcal{M}$ is also a sub-field. That it is the largest conforming sub-field, follows from its definition.

Now, let \mathcal{C} be an arbitrary conforming sub-field i.e. let \mathcal{C} be a sub-field of \mathcal{D} . For each $\mathcal{B} \in \mathcal{F}$, there exists (Theorem 2) a maximal element \mathcal{M} of \mathcal{F} such that $\mathcal{B} \subset \mathcal{M}$. But $\mathcal{C} \subset \mathcal{D} \subset \mathcal{M}$. Therefore, $\mathcal{C} \vee \mathcal{B} \subset \mathcal{M} \subset \mathcal{E}$, i.e. $\mathcal{C} \vee \mathcal{B} \in \mathcal{F}$. This proves the 'only if' part of (iii). The 'if' part is trivial.

For example, let \mathcal{E} be the class of all \mathcal{B} -linked sets (see Sections 6 and 7) in the probability structure $(\mathcal{X}, \mathcal{G}, \mathcal{P})$ where \mathcal{B} is a fixed sub-field of \mathcal{G} . If the set A be \mathcal{B} -measurable function $Q(A, \cdot)$ satisfying definition (iv) of section 7, then it is easily seen that AB is \mathcal{B} -linked for every $B \in \mathcal{B}$. We have only to define $Q(AB, \cdot)$ as

$$Q(A, \cdot) I(B, \cdot)$$

where $I(B, \cdot)$ is the indicator of B .

In this case, \mathcal{B} is a conforming sub-field. Theorem 4(iii) then asserts that for every \mathcal{B} -linked sub-field \mathcal{C} the sub-field $\mathcal{B} \vee \mathcal{C}$ is also \mathcal{B} -linked. It will be of some interest to find out conditions under which \mathcal{B} is the largest

conforming sub-field, i.e. $\mathcal{G} = \emptyset$.

9. Some further problems

In this section we list four problems that are mostly unsolved.

(A) Separating sub-fields: Let \mathcal{P} be a class of 'distinct' probability measures on a measurable space (X, \mathcal{G}) . That is, for each pair P_1, P_2 of members of \mathcal{P} there exists a measurable set $A \in \mathcal{G}$ such that $P_1(A) \neq P_2(A)$. A sub-field \mathcal{G} will be called 'separating' if the restriction of \mathcal{P} to \mathcal{G} gives rise to a class of distinct measures. For example, every sufficient sub-field is separating. No ancillary or ϕ -free (see Section 4) sub-field is separating. Let \mathcal{F}_5 be the family of all separating sub-fields. By definition \mathcal{G} is the largest element of \mathcal{F}_5 . What can we say about the existence of minimal elements in \mathcal{F}_5 ? A variant of this problem has recently received some attention in USSR ([6] and [8]). A partition Π of X into a class of disjoint measurable sets $\{A_t\}$ will be called 'separating' if for each pair P_1, P_2 of member of \mathcal{P} there exists a member A_t of the partition Π such that $P_1(A_t) \neq P_2(A_t)$. A separating partition is called minimal if there exists no other separating partition with a smaller number of parts. Let $\nu(\mathcal{P})$ stand for the number (possibly infinite) of parts in a minimal separating partition. What can we say about $\nu(\mathcal{P})$?

Example 5(a): Consider the class \mathcal{P} of all normal distributions on the real line with unit variances. Here $\nu(\mathcal{P}) = 2$. Any partition of the real line into two half lines is clearly separating and, of course, minimal. The corresponding sub-field is a minimal element of \mathcal{F}_5 .

Example 5(b): Let \mathcal{P} be the family of uniform distributions on $[0, \theta]$, $0 < \theta < 1$. In this case $\nu(\mathcal{P}) = 3$ (see [6]).

Example 5(c): If \mathcal{P} consists of a finite number of measures P_1, P_2, \dots, P_n then $v(\mathcal{P}) \leq n$. If \mathcal{P} consists of a countable number of continuous measures then $v(\mathcal{P}) = 2$. (see [6] and [8]).

(B): Partially sufficient sub-fields:

The notion of partial sufficiency, as introduced by Fraser ([5]), is as follows.

Let $\mathcal{P} = \{P_{\varphi, \theta}\}$, $\varphi \in \Phi$, $\theta \in \Theta$, be a family of probability measures indexed by the two independent parameters φ and θ . A sub-field $\mathcal{B} \subset \mathcal{G}$ will be called θ -sufficient for \mathcal{G} (or simply θ -sufficient) if

- i) \mathcal{B} is φ -free in the sense of section 4 and
- ii) for each $A \in \mathcal{G}$, there exists a choice of the conditional probability (function) of A given \mathcal{B} that does not depend on θ , i.e., for each $\varphi \in \Phi$, there exists a \mathcal{B} -measurable function $Q_{\varphi}(A, \cdot)$ that maps \mathcal{X} to the unit interval in such a manner that

$$P_{\varphi, \theta}(AB) \equiv \int_B Q_{\varphi}(A, \cdot) d P_{\varphi, \theta}(\cdot)$$

for all $B \in \mathcal{B}$ and $\theta \in \Theta$.

Let \mathcal{F}_6 be the family of all θ -sufficient sub-fields. Under what conditions can we prove that \mathcal{F}_6 is not vacuous? What about the minimal and maximal elements in \mathcal{F}_6 ?

(C): Complete sub-fields:

Given a probability structure $(\mathcal{X}, \mathcal{G}, \mathcal{P})$, we call a sub-field \mathcal{B} 'complete' if, for a \mathcal{B} -measurable, \mathcal{P} -integrable function f , the integral

$$\int_{\mathcal{X}} f d P \equiv 0, \text{ for all } P \in \mathcal{P},$$

when and only when f is \mathcal{P} -equivalent to zero. Let \mathcal{F}_7 be the family of all complete sub-fields. What can we say about the existence of maximal and minimal elements in \mathcal{F}_7 ?

Let us terminate this list of problems with a final one.

(D): Complementary sub-field

Let (X, \mathcal{G}) be a given measurable space and let \mathcal{B} be a fixed sub-field of \mathcal{G} . A sub-field \mathcal{C} will be called a complement to \mathcal{B} if

$$\mathcal{B} \vee \mathcal{C} = \mathcal{G},$$

i.e., if \mathcal{G} is the least Borel-field that contains both \mathcal{B} and \mathcal{C} .

Let $\mathfrak{F}_{\mathcal{B}}$ be the family of all sub-fields that are complements to \mathcal{B} . For example, if \mathcal{B} is the trivial sub-field then $\mathfrak{F}_{\mathcal{B}}$ consists of a single element namely \mathcal{G} itself. If $\mathcal{B} = \mathcal{G}$, then $\mathfrak{F}_{\mathcal{B}}$ consists of all sub-fields of \mathcal{G} .

\mathcal{G} is, of course, the largest element of $\mathfrak{F}_{\mathcal{B}}$. It is easy to construct examples where $\mathfrak{F}_{\mathcal{B}}$ has a multiplicity of minimal elements. Does $\mathfrak{F}_{\mathcal{B}}$ always have a minimal element?

References

1. Basu, D. (1959): The family of ancillary statistics, Sankhya, 21, 247-256.
2. Basu, D. (1965): On maximal and minimal sub-fields of certain types, Institute of Statistics Mimeo Series No. 422.
3. Burkholder, D. L. (1961): Sufficiency in the undominated case, Ann. Math. Stat., 32, 1191-1200.
4. Dantzig, G. B. (1941): On the non-existence of tests of Students' hypothesis having power functions independent of σ , Ann. Math. Stat., 11, 186-191.
5. Fraser, D. A. S. (1956): Sufficient statistics with nuisance parameters Ann. Math. Stat., 27, 838-842.
6. Kagan, A. M. and Sudakov, V. N. (1964): Separating partitions of certain families of measures (in Russian). Vestnik Leningrad University, No. 13, 147-150.
7. Pitcher, T. S. (1957): Sets of measures not admitting necessary and sufficiency statistics or sub-fields, Ann. Math. Stat., 28, 267-268.
8. Visik, S. M., Cobrinksii, A. A. and Rosenthal, A. L. (1964): A separating partition for a finite family of measures (in Russian), Theory of Prob. and its application, 9, 165-167.