FAMILIES OF DENSITIES WITH NON-CONSTANT CARRIERS WHICH HAVE FINITE DIMENSIONAL SUFFICIENT STATISTICS

By J. K. GHOSH and K. K. ROY

Indian Statistical Institute

SUMMARY. Let $X_1, X_2, \ldots, X_n$ be independently and identically distributed random variables with common density function $f_\theta(x)$. Let the carrier $X_\theta = \{x : f_\theta(x) > 0\}$ of the density $f_\theta(x)$ be an open interval $(a(\theta), b(\theta))$. If $a(\theta)$, $b(\theta)$ are continuous and monotonic in opposite sense and if there is a continuous real-valued statistic which is sufficient for the given family of densities for a sample of size $n \geqslant 2$, then we prove $f_\theta(x)$ has the following form $f_\theta(x) = g(\theta).h(x)$.

In our second theorem we have the same conclusion where $a(\theta)$, $b(\theta)$ are continuous and monotonic in the same sense and there is a continuous statistic taking values in R^2, which is sufficient for a sample of size $n \geqslant 3$. In the last section we investigate cases when $a(\theta)$, $b(\theta)$ are monotonic and the sufficient statistic is k-dimensional, $k \geqslant 2$.

1. Introduction

Let $X_1, X_2, \ldots, X_n$ be independently and identically distributed random variables with common density function $f_\theta(x)$. Let the carrier

$$X_\theta = \{x : f_\theta(x) > 0\}$$

of the density $f_\theta(x)$ be an open interval $(a(\theta), b(\theta))$. In Section 3 we prove Theorem 1 which states that if $a(\theta)$, $b(\theta)$ are continuous and monotonic in opposite sense and if there is a continuous real-valued statistic which is sufficient for the given family of densities for a sample of size $n \geqslant 2$, then $f_\theta(x)$ has the following form

$$f_\theta(x) = g(\theta) \cdot h(x).$$

In Theorem 2 we have the same conclusion where $a(\theta)$, $b(\theta)$ are continuous and monotonic in the same sense and there is a continuous statistic, taking values in R^2, which is sufficient for a sample of size $n \geqslant 3$. In the last section we investigate cases when $a(\theta)$, $b(\theta)$ are monotonic and there is a statistic T which is regular and k-dimensional at each point x (this essentially means that the rank of the matrix $\left(\left(\frac{\partial T_i}{\partial x_j}\right)\right)$ is k where $T = (T_1, T_2, \ldots, T_k)$) of a dense subset of the sample space and is sufficient for the given family of densities. More specifically in Theorem 3 we prove that if T is a minimal sufficient statistic for the family of densities Γ satisfying certain regularity conditions, for a sample of size $(k+1)$, $k \geqslant 1$, and if T is regular and k-dimensional at each point of a dense subset of the sample space then the family of densities Γ have the $(k-1)$-dimensional exponential form. (In these theorems sufficiency is used in the stronger sense of Dynkin (1951); for a definition see Section 2).

Results like Theorem 3 have received wide attention in the case where $a(\theta)$ and $b(\theta)$ are independent of θ. One may see, for example, Pitman (1936), Koopman (1936), Darmois (1935), Dynkin (1951), Brown (1964, 1970), Barankin and Maitra (1963), for extensions in many directions under the basic hypothesis of constant carriers for the family of densities. Pitman (1936) had considered, in addition to the preceding case, the problems that arise when the carriers are intervals of the sort considered in the previous paragraph. He had pointed out the conclusions of our Theorems 1, 2 and 3, but did not state the conditions under which they are valid. As may be expected, his arguments are not rigorous. Huzurbazar (1965) has shed some light on these problems but his treatment remains non-rigorous. (The somewhat easier problem of finding minimal sufficient statistic and their dimension for given families with non-constant carriers has been taken up successfully in a number of papers, Fraser (1966), Barankin (1966)). As the subsequent sections will show, a rigorous formulation of Pitman's results and arguments turns out to be far from simple.

2. General definitions and preliminaries

Let $\mathcal{P} = \{P_\theta : \theta \varepsilon I\}$ be a family of probability measures defined on a Borel subset X of the real line. We shall assume that each P_θ in $\mathcal{P}$ is dominated by a single σ-finite measure μ on X. We shall denote by $\Gamma(\mathcal{P})$ (or simply Γ) the family of probability densities $\{p_\theta(\cdot) : \theta \varepsilon I\}$ where $p_\theta(\cdot)$ is a fixed version of the Radon-Nikodym derivative of P_θ with respect to μ.

Definition 1 : The set

$$X_\theta = \{x \varepsilon X : p_\theta(x) > 0\}$$

is called the carrier of the density p_θ.

Definition 2 : The family of densities Γ is said to have non-constant carriers if all the X_θ's are not the same.

Due to a theorem of Halmos and Savage (for a proof one may see Lehmann (1959) p. 354, Theorem 2), it follows that there exist countably many θ_n's such that

$$P_\theta\left(X_\theta - \bigcup_1^\infty X_{\theta_n}\right) = 0 \;\forall\; \theta.$$

Hence, without much loss of generality, we may assume that the fixed version of the densities $p_\theta(\cdot)$ are so chosen that $X_\theta \subset \bigcup_1^\infty X_{\theta_n}$. Also, we shall take

$$X = \bigcup X_\theta = \bigcup_1^\infty X_{\theta_n}.$$

We shall consider n independent observations from the distribution P_θ and in what follows by *a sample of size n* we shall mean n independent observations from P_θ on X.

Definition 3 : A *statistic* for a sample of size n is any Borel-measurable function on X^n into a measurable space.

Definition 4 : A statistic T for a sample of size n is said to be *functionally sufficient* (in short f-sufficient) for a sample of size n, for the family of densities Γ if there are two functions g and h on $I \times T(X^n)$ and X^n respectively, such that

$$\prod_{i=1}^{n} p_\theta(x_i) = g(\theta, T(x_1, x_2, \ldots, x_n))\, h(x_1, x_2, \ldots, x_n) \ \forall\ (x_1, x_2, \ldots, x_n) \text{ in } X^n \text{ and } \theta \text{ in } I. \quad \ldots \quad (2.1)$$

Dynkin (1951), calls a statistic sufficient if it satisfies the above definition. We use the term functionally sufficient to differentiate this concept with the usual concept of a sufficient statistic. If the equality in (2.1) is satisfied almost everywhere with respect to the dominating measure μ, then we call T sufficient for the family Γ for a sample of size n. In proving our main results we assume the slightly stronger criterion of functional sufficiency of a statistic T than the usual sufficiency. We follow Dynkin in this respect.

Definition 5 : A statistic T for a sample of size n is said to be a *minimal f-sufficient statistic* for the family Γ for a sample of size n if, for any f-sufficient statistic S for Γ for a sample of size n, $S(x_1, x_2, \ldots, x_n) = S(x_1', \ldots, x_n')$ implies that $T(x_1, x_2, \ldots, x_n) = T(x_1', \ldots, x_n')$.

In other words T is minimal f-sufficient for Γ for a sample of size n if and only if T is a function of every other f-sufficient statistic for Γ for a sample of the same size.

Definition 6 : Following Dynkin (1951), we call a function f defined on a non-empty open subset U of R^n, *trivial* if there exists a non-empty open subset V of U such that T on V is one-one. f is *non-trivial* on U if f is not trivial on U.

The following definition of the rank of a family of densities is also due to Dynkin (1951).

Definition 7 : The *rank* $r(\Gamma)$ of the family of densities Γ is defined as follows :

$r(\Gamma) = 0$ if for a sample of size 1 there is a non-trivial f-sufficient statistic for Γ

otherwise,

$$r(\Gamma) = \sup\{1 \leqslant r \leqslant \infty : \text{ for any finite } n \leqslant r \text{ there is no non-trivial f-sufficient statistic for } \Gamma \text{ for a sample of size } n\}.$$

Remark : If X is an interval and $r(\Gamma) = 0$ and the non-trivial f-sufficient statistic T, for a sample of size one, is once continuously differentiable on X, then T is a constant, i.e., $\mathscr{P}$ collapses to a singleton.

Definition 8 : A function $f = (f_1, f_2, \ldots, f_k)$ defined on a non-empty open subset U of R^n taking values in R^k, $1 \leqslant k \leqslant n$, is said to be *at most k-dimensional* if f is continuous on U.

We may say that such a function is of dimension exactly k at $x^0 \epsilon U$ if given any open set $V \ni x^0$ $\exists$ an open set $V' \ni x^0$ such that $f(V')$ is a k-dimensional set as defined in Hurewicz and Wallman (1948). f is k-dimensional if at least one such x^0 exists. However we shall not need the concept of a k-diemensional statistic in this generality. We define below a statistic which is k-dimensional at all points of a dense open subset V of U and satisfies some additional conditions.

Definition 9 : A function $f = (f_1, f_2, ..., f_k)$ defined on a non-empty open subset U of R^n taking values in R^k, $1 \leqslant k \leqslant n$ is said to be a *regular k-dimensional* function if f satisfies the following conditions :

(a) f is continuous on U,

(b) there is an open subset $V \subset U$, V is dense in U and V satisfies the following :

(i) all the partial derivatives $\dfrac{\partial f_i}{\partial x_j}$ $i = 1, 2, ..., k$; $j = 1, 2, ... n$, exist at each point $x = (x_1, x_2, ..., x_n)$ in V, they are continuous functions of x on V,

(ii) rank of the matrix

$$\left(\left(\frac{\partial f_i}{\partial x_j}\right)\right)_{i=1,2,...,k;\ j=1,2,...,n;}$$

is exactly k for every point x in V.

We shall frequently write the set V of this definition as $V(f)$.

Definition 10 : We call a function $f = (f_1, f_2, ..., f_k)$ on a subset U of R^n into R^k, *continuously differentiable* on $A \subset U$, if all the partial derivatives

$$\frac{\partial f_i}{\partial x_j} \quad i = 1, 2, ..., k ; \quad j = 1, 2, ..., n ;$$

exist and are continuous on A.

Definition 11 : The family of densities Γ is said to be a *k-dimensional exponential family* of densities for $k \geqslant 0$, if there are $(k+1)$ real-valued functions $\phi_0, \phi_1, ..., \phi_k$ on $X = \bigcup_{\theta \epsilon I} X_\theta$ and $(k+1)$ real-valued functions $c_0, c_1, ..., c_k$ on the parameter space I such that

(i) $1, \phi_1, ..., \phi_k$ are linearly independent,

(ii) $1, c_1, c_2, ..., c_k$ are linearly independent,

and (iii) the densities $\{p_\theta(\cdot) ; \theta \epsilon I\}$ have the following form :

$$p_\theta(x) = \begin{cases} \exp\left[c_0(\theta)+\phi_0(x)+\sum_{i=1}^{k} c_i(\theta)\,\phi_i(x)\right] & \text{for } x \text{ in } X_\theta \\ 0 & \text{for } x \text{ in } X - X_\theta \end{cases}$$

$\forall\ \theta$ in I.

In the following sections, we shall consider families of densities Γ having non-constant carriers, each carrier being an open interval and we shall prove the exponential nature of the families when an f-sufficient statistic, (or a minimal f-sufficient statistic) satisfying certain properties, exists.

In the remainder of this section we present a few auxiliary lemmas which we need for proving our main results. Lemmas 2.1, 2.2 are easy consequences of the inverse function and the implicit function theorems (see, for example, Apostol (1965)). Lemma 2.3 follows easily from Lemma 2.2, chain-rule and the rank theorem for products of matrices. Lemma 2.1 can alternatively be proved by using a well-known result in dimension theory, which says that there is no one-one continuous function defined on a non-empty open subset of R^n into R^k where $n > k \geqslant 1$ (see p. 5 of Hurewicz and Wallman (1948)). Lemma 2.2 in a slightly different form may be found in Barankin and Katz (1959). We shall make use of Lemmas 2.1, 2.2 and 2.3 in Section 5 and Lemmas 2.4 and 2.5 in Section 3 and Section 4.

Lemma 2.1 : *Let U be a non-empty open subset of R^n, $n > 1$. Let $f = (f_1, f_2, \ldots, f_k)$ defined on U, be a regular k-dimensional function* (Definition 9) *on U, where $1 \leqslant k < n$. Then f is non-trivial* (Definition 6) *on U.*

Lemma 2.2 : *Let $f = (f_1, f_2, \ldots, f_k)$ be a continuously differentiable function defined on an open neighbourhood U of a point $x^0 = (x_1^0, \ldots x_n^0)$ in R^n.*

Let ϕ be a real-valued continuously differentiable function on U such that

$$\phi(x_1, x_2, \ldots, x_n) = g(f(x_1, x_2, \ldots, x_n)) \ \forall \ (x_1, x_2, \ldots, x_n) \text{ in } U,$$

where g is a real-valued function defined on $f(U)$. Let the rank of the matrix

$$M_f(x^0) = \left(\left(\frac{\partial f_i}{\partial x_j}\right)\right)_{i = 1, 2, \ldots, k, \, j = 1, 2, \ldots, n, \, x = x^0}$$

be k. Then there is a neighbourhood N of x^0 such that $f(N)$ is a neighbourhood of $f(x^0)$ and g is continuously differentiable on $f(N)$.

Lemma 2.3 : *Let $T = (T_1, T_2, \ldots, T_k)$ and $\phi = (\phi_1, \phi_2, \ldots, \phi_{k'})$ be continuously differentiable functions defined on a non-empty open subset U of R^n into R^k and $R^{k'}$ respectively, such that the ranks of the matrices*

$$M_T(x) = \left(\left(\frac{\partial T_i}{\partial x_j}\right)\right)_{i = 1, 2, \ldots, k; \ j = 1, 2, \ldots, n}$$

and

$$M_\phi(x) = \left(\left(\frac{\partial \phi_i}{\partial x_j}\right)\right)_{i = 1, 2, \ldots, k'; \ j = 1, 2, \ldots, n}$$

are r and s, respectively, for all x in U, with $s < r$.

Then there cannot exist a function g on $\phi(U)$ such that

$$T(x) = g(\phi(x)) \quad \forall \ x \text{ in } U.$$

Lemma 2.4 : *Let $A_1, A_2, \ldots, A_n$ be non-degenerate bounded closed intervals in R. Let $T = (T_1, T_2, \ldots, T_k)$ be a continuous function defined on $A_1 \times A_2 \times \ldots \times A_n$ into R^k, $1 \leqslant k < n$, such that there are real-valued functions $\phi_1, \phi_2, \ldots, \phi_k$ on $T(A_1 \times \ldots \times A_n)$ satisfying the following relations :*

$$x_i = \phi_i(T(x_1, x_2, \ldots, x_n)) \ \forall (x_1, x_2, \ldots, x_n) \text{ in } A_1 \times A_2 \times \ldots \times A_n \quad \text{and} \quad i = 1, 2, \ldots, k.$$

Then for each fixed point $(x^0_{k+1}, ..., x^0_n)$ in the interior of $A_{k+1}\times ...\times A_n$ we have open intervals $B_{k+1}, ..., B_n$ with $x^0_j \in B_j \subset A_j$, $\forall j = k+1, ..., n$ and a point $(x^0_1, ..., x^0_k)$ in $A_1\times ...\times A_k$ such that $T(x^0_1, x^0_2, ..., x^0_k, x_{k+1}, ..., x_n)$ is constant over $B_{k+1}\times ...\times B_n$.

Proof: Let $(x^0_{k+1}, ..., x^0_n)$ be in the interior of $A_{k+1}\times ...\times A_n$.

Let

$$T^*(x_1, x_2, ..., x_k) = T(x_1, x_2, ..., x_k, x^0_{k+1}, ..., x_n) \;\forall\; (x_1, x_2, ..., x_k) \text{ in } A_1\times ...\times A_k.$$

Then T^* is $1-1$ and continuous on $A_1\times ...\times A_k$ and hence is a homeomorphism (since $A_1\times ...\times A_k$ is compact) of $A_1\times ...\times A_k$ and $T^*(A_1\times ...\times A_k)$. It follows that the subset $T^*(A_1\times A_2\times ...\times A_k)$ of R^n is k-dimensional, and hence by a well-known theorem of dimension theory (Hurewicz and Wallman (1948) p. 44) contains an open sphere, say,

$$S(t^0, \epsilon) = \{t \epsilon R^k : \|t-t_0\| < \epsilon\}$$

where $\|\cdot\|$ denotes the usual Euclidean norm in R^k.

Let

$$t^0 = (t^0_1, ..., t^0_k) = T(x^0_1, x^0_2, ..., x^0_k, x^0_{k+1}, ..., x^0_n)$$

where $(x^0_1, x^0_2, ..., x^0_k) \epsilon A_1\times ... \times A_k$.

Let $\delta > 0$ be so chosen that

$$(x^0_j-\delta, x^0_j+\delta) \subset A_j \qquad \forall j = k+1, ..., n$$

and

$$|x_j-x^0_j| < \delta \;\forall j = k+1, k+2, ..., n \Longrightarrow \|T(x^0_1, x^0_2, ..., x^0_k, x_{k+1}, ..., x_n) - T(x^0_1, ..., x^0_n)\| < \epsilon.$$

Let $\qquad B_j = (x^0_j-\delta, x^0_j+\delta) \quad \forall j = k+1, ..., n$

and

$$W_{x^0_1, ..., x^0_k} = \{T(x^0_1, ..., x^0_k, x_{k+1}, ..., x_n) : (x_{k+1}, ..., x_n) \epsilon B_{k+1}\times ...\times B_n\}.$$

Now

$$t \epsilon W_{x^0_1, ..., x^0_k} \cap S(t^0, \epsilon)$$

$$\Longrightarrow t = T(x^0_1, ..., x^0_k, x'_{k+1}, ..., x'_n) \text{ for some } (x'_{k+1}, ..., x'_n) \text{ in } B_{k+1}\times ...\times B_n$$

and

$t = T(x'_1, ..., x'_k, x^0_{k+1}, ..., x^0_n)$ for some $(x'_1, ..., x'_k)$ in $A_1\times ...\times A_k \Longrightarrow$ $(x'_1, x'_2, ..., x'_k) = (x^0_1, x^0_2, ..., x^0_k)$ by the assumption in the lemma $\Longrightarrow t = t^0$.

Hence $W_{x^0_1, ..., x^0_k} \cap S(t^0, \epsilon) = \{t^0\}$ is a singleton. But because of choice of $\delta > 0$

$$W_{x^0_1, ..., x^0_k} \subset S(t^0, \epsilon).$$

It follows that $W_{x^0_1, ..., x^0_k}$ is a singleton. Hence $(x^0_1, ..., x^0_k)$, $B_{k+1}, ..., B_n$ satisfy the requirements of the lemma. [Q.E.D.]

Lemma 2.4 can further be strengthened as follows, in the case when $n = 2$ and $k = 1$.

Lemma 2.5. *Let A_1, A_2 be two non-degenerate bounded closed intervals in R. Let T be a continuous real-valued function defined on $A_1 \times A_2$ such that there is a real-valued function ϕ defined on $T(A_1 \times A_2)$, which satisfies*

$$x_1 = \phi(T(x_1, x_2)) \ \forall \ (x_1, x_2) \text{ in } A_1 \times A_2.$$

Then there is a real-valued function g defined on A_1 such that

$$T(x_1, x_2) = g(x_1) \ \forall \ (x_1, x_2) \text{ in } A_1 \times A_2.$$

Proof: Let us define for each x in A_1

$$N_x = \{T(x, y) : y \in A_2\}.$$

N_x is a connected subset of R and hence is a degenerate or non-degenerate interval.

By the assumption in the lemma we have,

$$\text{for} \quad x \neq x', \ N_x \cap N_{x'} = \phi.$$

It follows that the set

$D = \{x \epsilon A_1 : N_x$ is a non-degenerate interval$\}$ is a countable subset of A_1. The lemma will follow as soon as we show that

$$D = \phi.$$

Let $x_1 \epsilon D$ and y_1, y_2 be in A_2.

Let $\epsilon > 0$ be arbitrary; choose $\delta > 0$ such that

$$|x_1 - x_1'| < \delta \Longrightarrow |T(x_1, y) - T(x_1', y)| < \delta \ \forall \ y \text{ in } A_2.$$

Since D is countable, we can choose $x_1' \epsilon A_1 - D$ such that $|x_1 - x_1'| < \delta$.

Then

$$|T(x_1, y_1) - T(x_1, y_2)| \leqslant |T(x_1, y_1) - T(x_1', y_1)| + |T(x_1', y_1) - T(x_1', y_2)|$$
$$+ |T(x_1', y_2 - T(x_1, y_2)| < 2\,\epsilon.$$

Hence $T(x_1, y_1) = T(x_1, y_2)$. Since y_1, y_2 are arbitrary, N_{x_1} is also a singleton, which contradicts that x_1 is in D. [Q.E.D.]

Remark: To prove our results in the later sections we do not actually need this stronger version of Lemma 2.4 given by Lemma 2.5.

3. Carrier intervals with end-points monotonic in opposite sense and a one-dimensional f-sufficient statistic exists

Let Γ be a family of densities $\{p_\theta, \theta \text{ in } I\}$ such that the carrier X_θ of each p_θ is an open interval $(a(\theta), b(\theta))$, where $a(\theta), b(\theta)$ satisfy the following conditions:

(A-3.1) The parameter space I is an interval of either of the two following forms:

(i) $I = [\lambda_0, \lambda_1), \quad -\infty < \lambda_0 < \lambda_1 \leqslant \infty$

(ii) $I = [\lambda_0, \lambda_1], \quad -\infty < \lambda_0 < \lambda_1 < \infty$

(A-3.2) $a(\theta)$ is a strictly increasing and $b(\theta)$ is a strictly decreasing function of θ in I.

(A-3.3) $a(\theta)$, $b(\theta)$ are continuous functions of θ in I.

Without loss of generality we take

$$\lambda_0 = 0, \quad \lambda_1 = 1.$$

Let
$$a(\theta) \uparrow c_1 \quad \text{as} \quad \theta \uparrow \lambda_1 = 1$$
and
$$b(\theta) \downarrow c_2 \quad \text{as} \quad \theta \uparrow \lambda_1 = 1.$$

Then $c_1 \leqslant c_2$ (of course if $c_1 = c_2$, then $\lambda_1 \notin I$, i.e., the interval I is of the type (i) of (A-3.1)).

Here we also have

$$X = \bigcup_{\theta \in I} X_\theta = X_0 = (a(0),\ b(0)).$$

Let

$$q_\theta(x) = p_\theta(x)/p_0(x) \text{ for } x \text{ in } X_\theta \text{ and } \theta \text{ in } I.$$

Let us define

$$\alpha(x) = \begin{cases} a^{-1}(x) & \text{for } a(0) < x < c_1 \\ 1 & \text{for } c_1 \leqslant x \leqslant c_2 \\ b^{-1}(x) & \text{for } c_2 < x < b(0). \end{cases}$$

(We observe that since $a(\theta)$, $b(\theta)$ are strictly montonic functions, their inverses exist.)

$$\beta(x, y) = \min\{\alpha(x), \alpha(y)\}.$$

Theorem 1 : *If for any $n \geqslant 2$, there is a statistic of dimension at most one, i.e., a continuous real-valued statistic T defined on X^n, which is f-sufficient for the family of densities $\Gamma = \{p_\theta(\cdot);\ \theta \epsilon I\}$ satisfying the conditions* (A-3.1), (A-3.2), (A-3.3), *then Γ is a zero-dimensional exponential family, i.e., the densities p_θ, θ in I have the following form :*

$$p_\theta(x) = \begin{cases} g(\theta) \cdot h(x) & \text{for } x \text{ in } X_\theta \\ 0 & \text{for } x \text{ in } X - X_\theta \end{cases}$$

for all θ in I.

Proof : Without loss of generality we can and shall assume that T is a continuous f-sufficient statistic for Γ, for a sample of size two. In fact, if T is a continuous f-sufficient statistic for Γ for a sample of size $n > 2$, we can consider the statistic T^* defined by

$$T^*(x_1, x_2) = T(x_1, x_2, c, c, \ldots, c)$$

where $c_1 \leqslant c \leqslant c_2$ is a fixed point.

The statistic T^* is f-sufficient for Γ for a sample of size two and is also continuous.

Fix a θ_0 in I. We shall first show that for every point x_2^0 in X_{θ_0}, there is an open interval $N_{x_2^0}$ containing x_2^0, $N_{x_2^0} \subset X_{\theta_0}$ such that $q_{\theta_0}(x)$ is constant over $N_{x_2^0}$. This, in turn, will show that on every compact sub-interval of X_{θ_0}, $q_{\theta_0}(x)$ is constant. From this it easily follows that $q_{\theta_0}(x)$ is constant over X_{θ_0}, which allows us to conclude that the theorem is true.

It remains to prove that for every fixed point x_2^0 in X_{θ_0}, there is an open interval $N_{x_2^0}$ containing x_2^0 and $N_{x_2^0} \subset X_{\theta_0}$ such that $q_{\theta_0}(x)$ is constant over $N_{x_2^0}$. Fix x_2^0 in X_{θ_0}. Let us choose a non-degenerate bounded closed interval $A_2 \subset X_{\theta_0}$, containing x_2^0 as an interior point and another non-degenerate bounded closed interval $A_1 \subset X_{\theta_0}$ such that

$$(1) \quad \beta(x_1, x_2) = \min\{\alpha(x_1), \alpha(x_2)\} = \alpha(x_1) \quad \forall (x_1, x_2) \text{ in } A_1 \times A_2$$

and $$(2) \quad \alpha \text{ is one-one on } A_1 \text{ i.e., } A_1 \subset (a(\theta_0), c_1) \cup (c_2, b(\theta_0)).$$

Such a choice of A_1 and A_2 is possible by our assumptions (A-3.1), (A-3.2) and (A-3.3).

The fact that

$$\beta(X_1, X_2) = \sup\{\theta : p_\theta(x_1)\, p_\theta(x_2) > 0\}$$

implies that $\beta(.,.)$ defined on $X_0 \times X_0$ is a function of every f-sufficient statistic for Γ.

Hence on $X_0 \times X_0$, $$\beta(x_1, x_2) = \psi(T(x_1, x_2)).$$

Now on $A_1 \times A_2$ $$x_1 = a(\alpha(x_1)) = a(\beta(x_1, x_2))$$
$$= a \circ \psi(T(x_1, x_2)).$$

By applying Lemma 2.4, we have an open interval $N_{x_2^0}$ containing x_2^0, $N_{x_2^0} \subset A_2$ and a point x_1^0 in A_1, such that $T(x_1^0, x_2)$ is constant over $N_{x_2^0}$.

Now

$q_{\theta_0}(x_1^0) \cdot q_{\theta_0}(x_2) = g(T(x_1^0, x_2))\ \forall\ x_2$ in $N_{x_2^0}$ for some function g (by f-sufficiency of T). Since $x_1^0 \epsilon A_1 \subset X_{\theta_0}$, $q_{\theta_0}(x_1^0) \neq 0$; It follows that $q_{\theta_0}(x_2)$ is constant over $N_{x_2^0}$.

[Q.E.D.]

Remarks (0) It is clear that T in the above theorem must be one-dimensional in the sense of the paragraph following Definition 8.

(1) If the sample space $X = X_{\lambda_1}$ is of the form $(a(\lambda_1), b(\lambda_1))$, $-\infty \leqslant a(\lambda_1) < b(\lambda_1) \leqslant \infty$ and the parameter space is either $[\lambda_0, \lambda_1]$ or $(\lambda_0, \lambda_1]$, the conclusion of Theorem 1 remains true provided $a(\theta)$, $b(\theta)$ satisfy (A-3.3) along with (A-3.2*) mentioned below :

(A-3.2*) $a(\theta)$ is a strictly decreasing function of θ in I and $b(\theta)$ is a strictly increasing function of θ in I.

The proof is exactly analogous.

(2) We may also modify the assumptions (A-3.2) and (A-3.2*) by assuming that either $a(\theta)$ or $b(\theta)$ [but not both] is constant in θ. In this case also the same proof yields the conclusion of Theorem-1.

(3) The following converse of Theorem 1 is also true and is easy to prove :

If Γ satisfies the conditions (A-3.1), (A-3.2) and (A-3.3) and if the densities in Γ have the form

$$p_\theta(x) = \begin{cases} g(\theta)\ h(x), & \text{for } x \text{ in } X_\theta \\ 0 & \text{otherwise} \end{cases}$$

for each θ, then $\beta(x_1, x_2)$ is a continuous f-sufficient statistic for Γ for a sample of size 2.

4. Carrier intervals with end-points monotonic in the same sense and a 2-dimensional f-sufficient statistic exists

In this section we consider a family $\Gamma = \{p_\theta(\cdot) : \theta \epsilon I\}$ of probability densities on the real line satisfying the following conditions :

(A-4.1) The parameter-space I is an open interval (λ_0, λ_1) where $-\infty \leqslant \lambda_0 < \lambda_1 \leqslant \infty$.

(A-4.2) The carrier of $p_\theta(\cdot)$ is an open interval $(a(\theta), b(\theta))$, for each θ in I.

(A-4.3) Both $a(\theta)$ and $b(\theta)$ are strictly increasing functions of θ in I.

(A-4.4) Both $a(\theta)$, $b(\theta)$ are continuous functions of θ in I.

Let us take $\lambda_0 = 0$, and $\lambda_1 = 1$.

Let $$a_0 = \lim_{\theta \downarrow 0} a(\theta), \qquad b_0 = \lim_{\theta \downarrow 0} b(\theta)$$

and $$a_1 = \lim_{\theta \uparrow 1} a(\theta), \qquad b_1 = \lim_{\theta \uparrow 1} b(\theta)$$

(a_0, b_0 may be $-\infty$ and a_1, b_1 may be $+\infty$).

Then $$X = \bigcup_{\theta \epsilon I} X_\theta = (a_0, b_1).$$

Let us define

$$\alpha_1(x) = \begin{cases} 0 & \text{if } a_0 < x \leq b_0 \\ b^{-1}(x) & \text{if } b_0 < x < b_1 \end{cases}$$

and
$$\alpha_2(x) = \begin{cases} a^{-1}(x) & \text{if } a_0 < x < a_1 \\ 1 & \text{if } a_1 \leq x < b_1 \end{cases}$$

We see that for each x in $X = (a_0, b_1)$

$$\{\theta : p_\theta(x) > 0\} = \{\theta : \alpha_1(x) < \theta < \alpha_2(x)\}.$$

Let

$$\beta_1(x_1, x_2, \ldots, x_n) = \max \{\alpha_1(x_1), \alpha_1(x_2), \ldots, \alpha_1(x_n)\}$$

$$\beta_2(x_1, x_2, \ldots, x_n) = \min \{\alpha_2(x_1), \ldots, \alpha_2(x_n)\}.$$

With these definitions we have $p_\theta(x_1) \ldots p_\theta(x_n) > 0$ if and only if

$$\beta_1(x_1, x_2, \ldots, x_n) < \theta < \beta_2(x_1, x_2, \ldots, x_n).$$

Hence if T is any f-sufficient statistic for the family Γ for a sample of size n then

$$\beta(x_1, x_2, \ldots, x_n) = (\beta_1(x_1, x_2, \ldots, x_n),\ \beta_2(x_1, x_2, \ldots, x_n))$$

can be expressed as a function of T.

Theorem 2 : *If there is a statistic T of dimension at most two (i.e., a continuous statistic on X^n into R^2) defined on X^n ($n \geq 3$) such that T is f-sufficient for the family of densities Γ satisfying the conditions (A-4.1) through (A-4.4) for a sample of size n, then the densities $p_\theta(\cdot)$, θ in I have the following form :*

$$p_\theta(x) = \begin{cases} g(\theta) \,.\, h(x), & \text{for } x \text{ in } X_\theta \\ 0, & \text{for } x \text{ in } X - X_\theta \end{cases}$$

$\forall$ θ in I where $h(x)$ is a strictly positive function throughout X.

Proof: (1) We know that β_1, β_2 can be expressed as follows :

$\beta_i(x_1, x_2, \ldots, x_n) = \psi_i(T(x_1, \ldots, x_n))$, $\quad i = 1, 2 \quad \forall (x_1, x_2, \ldots, x_n)$ in X^n.

(2) We shall prove that if θ_1, θ_2 are in I such that

$$X_{\theta_1} \cap X_{\theta_2} \neq \phi$$

then on $X_{\theta_1} \cap X_{\theta_2}$, $\dfrac{p_{\theta_1}(x)}{p_{\theta_2}(x)}$ is a constant, which we shall denote by $g(\theta_1, \theta_2)$.

Let θ_1, θ_2 in I be fixed such that $X_{\theta_1} \cap X_{\theta_2} \neq \phi$ and let $\theta_1 > \theta_2$. Let x_3^0 be a fixed point in $X_{\theta_1} \cap X_{\theta_2}$. It is easy to see that we can choose three closed intervals

$$A_1 = [u, v]$$
$$A_2 = [p, q]$$
$$A_3 = [r, s]$$

such that

(1) $\max\{a(\theta_1), a(\theta_2)\} = a(\theta_1) < p < q < r < x_3^0 < s < u < v$
$< \min\{b(\theta_1), b(\theta_2)\} = b(\theta_2)$,

(2) $q < a_1$ and $u > b_0$.

Let $A_j = [r, s]$ for $3 < j \leqslant n$.

In view of the assumptions (A-4.1) to (A-4.4) α_1 is one-one on A_1 and α_2 is one-one on A_2 and

$$\beta_1(x_1, x_2, \ldots, x_n) = \alpha_1(x_1)$$

$$\beta_2(x_1, x_2, \ldots, x_n) = \alpha_2(x_2) \;\; \forall (x_1, x_2, \ldots, x_n) \text{ in } A_1 \times \ldots \times A_n.$$

So we have on $A_1 \times A_2 \times \ldots \times A_n$

$$\begin{aligned} x_1 = b(\alpha_1(x_1)) &= b(\beta_1(x_1, \ldots, x_n)) \\ &= b \circ \psi_1 \circ T(x_1, x_2, \ldots, x_n), \\ x_2 = a(\alpha_2(x_2)) &= a(\beta_2(x_1, x_2, \ldots, x_n)) \\ &= a \circ \psi_2 \circ T(x_1, x_2, \ldots, x_n). \end{aligned}$$

By taking $x_j^0 = x_3^0$ for $3 \leqslant j \leqslant n$, we may apply Lemma 2.4 and get a point (x_1^0, x_2^0) in $A_1 \times A_2$ and an open interval N containing x_3^0 such that $T(x_1^0, x_2^0, x_3, x_4, \ldots, x_n)$ is constant for all $(x_3, x_4, \ldots, x_n)$ in N^{n-2}.

Therefore, f-sufficiency of T shows that

$$\frac{p_{\theta_1}(x_1^0) p_{\theta_1}(x_2^0) p_{\theta_1}(x_3) \ldots p_{\theta_1}(x_n)}{p_{\theta_2}(x_1^0) p_{\theta_2}(x_2^0) p_{\theta_2}(x_3) \ldots p_{\theta_2}(x_n)}$$

is a constant as $(x_3, \ldots, x_n)$ varies over N^{n-2} and hence

$$\frac{p_{\theta_1}(x_3) \ldots p_{\theta_1}(x_n)}{p_{\theta_2}(x_3) \ldots p_{\theta_2}(x_n)}$$

is a constant as $(x_3, \ldots, x_n)$ vary over N^{n-2}. In particular $\left[\dfrac{p_{\theta_1}(x_3)}{p_{\theta_2}(x_3)}\right]^{n-2}$ and hence $\dfrac{p_{\theta_1}(x_3)}{p_{\theta_2}(x_3)}$ is constant over N.

It follows, as in the proof of Theorem 1, that $\dfrac{p_{\theta_1}(x)}{p_{\theta_2}(x)}$ is constant over $X_{\theta_1} \cap X_{\theta_2}$.

(3) Let N denote all integers positive, negative and zero. Let $\{\theta_n, n \in N\}$ be a set of points in (0, 1) such that

(i) $\theta_n < \theta_{n+1} \;\forall\; n$ in N,

(ii) $\lim\limits_{n \to -\infty} \theta_n = 0$ and $\lim\limits_{n \to +\infty} \theta_n = 1$,

(iii) $X_{\theta_n} \cap X_{\theta_{n+1}} \neq \phi$.

It is not difficult to check that such a sequence $\{\theta_n\}$ exists. Let us define $h_n(x)$ on X_{θ_n}, n in N as follows :

$$h_0(x) = p_{\theta_0}(x) \text{ for } x \text{ in } X_{\theta_0}.$$

For x in X_{θ_n}

$$h_n(x) = \begin{cases} p_{\theta_n}(x)\, g(\theta_{n-1}, \theta_n)\, g(\theta_{n-2}, \theta_{n-1}) \ldots g(\theta_0, \theta_1) & \text{if } n > 0 \\ p_{\theta_n}(x)\, g(\theta_{n+1}, \theta_n)\, g(\theta_{n+2}, \theta_{n+1}) \ldots g(\theta_0, \theta_{-1}) & \text{if } n < 0 \end{cases}$$

It is easy to check that on $X_{\theta_n} \cap X_{\theta_{n+1}}$

$$h_n(x) = h_{n+1}(x)$$

and hence it follows that the function

$$h(x) = h_n(x) \text{ on } X_{\theta_n} \text{ for } n \text{ in } N$$

is unambiguously defined for all x in $\bigcup_{n \in N} X_{\theta_n} = X$.

An easy verification now shows that for any θ in (0, 1), $p_\theta(x)/h(x)$ is constant $\psi(\theta)$, for all x in X_θ. [Q.E.D.]

Remarks (1) Instead of assuming that $a(\theta)$, $b(\theta)$ are increasing functions of θ we could have assumed that they both are strictly decreasing functions of θ. The same result holds and the proof is identical.

(2) The following converse of this theorem is also true and is easy to prove :

If the family of densities $\Gamma = \{p_\theta(\cdot), \theta \in I\}$ satisfy the conditions (A-4.1) through (A-4.4) and if further, $p_\theta(\cdot)$ has the following form for each θ in I

$$p_\theta(x) = \begin{cases} \psi(\theta) \cdot h(x) & \text{for } x \text{ in } X_\theta \\ 0 & \text{otherwise} ; \end{cases}$$

then there is a continuous statistic T defined on X^n into R^2, such that T is f-sufficient for Γ for a sample of size n. To be specific, $T = \beta(x_1, x_2, \ldots, x_n)$ is a continuous f-sufficient statistic in this case.

5. Carrier intervals with end-points monotonic in opposite sense and a regular k-dimensional f-sufficient statistic exists

In this section, as before, we consider a family $\Gamma = \{p_\theta(\cdot), \theta \epsilon I\}$ of probability densities on the real line R, such that the carrier X_θ of $p_\theta(\cdot)$ is an open interval $(a(\theta), b(\theta))$, for each θ in I. We use the terms defined in Section 2. We investigate the form of the density $p_\theta(\cdot)$, θ in I, when a regular k-dimensional minimal f-sufficient statistic exists for Γ for a sample of size $(k+1)$.

We make the following assumptions about the density functions $p_\theta(\cdot)$, θ in I.

(A-5.1) The parameter space I is an interval of the type $[\lambda_0, \lambda_1)$ where $-\infty < \lambda_0 < \lambda \leqslant \infty$.

(A-5.2) $p_\theta(x)$ is a once continuously differentiable function of x in X_θ for each fixed θ in I.

(A-5.3) The carrier X_θ of $p_\theta(\cdot)$ is an open interval $(a(\theta), b(\theta))$ for each θ in I.

(A-5.4) $a(\theta)$ is a once continuously differentiable function of θ with strictly positive first derivative for each θ in I. $b(\theta)$ is a once continuously differentiable function of θ with strictly negative first derivative for each θ in I.

(A-5.5) $a(\theta)$ increases to a finite constant c as θ increases to λ_1 and $b(\theta)$ decreases to the same constant c as θ increases to λ_1.

As before, we take, without loss of generality $\lambda_0 = 0$ and $\lambda_1 = 1$ and

$$X = \bigcup_{\theta \epsilon I} X_\theta = X_0.$$

For θ in $I = [0, 1)$ we define

$$r_\theta(x) = \log p_\theta(x) - \log p_0(x) \text{ for all } x \text{ in } X_\theta.$$

For any function f defined on a set A and a subset $B \subset A$, let $f|B$ denote the restriction of f to B. For each θ in I and $E \subset X_\theta$ let $\mathcal{L}_\theta(E)$ denote the smallest linear space of functions on E containing the constant functions and the family of functions :

$$\{r_{\theta'}|E : 0 \leqslant \theta' \leqslant \theta\}.$$

Let

$$\mathcal{L}_\theta = \mathcal{L}_\theta(X_\theta).$$

Also let $\mathcal{L}^*_\theta(E)$ denote the smallest linear space of functions on E containing the constant functions and the family of functions :

$$\{r_{\theta'}|E : 0 \leqslant \theta' < \theta\}, \text{ where } \theta \text{ is in } I \text{ and } E \subset X_\theta.$$

Lemma 5.1 : *Let $T = (T_1, T_2, \ldots, T_k)$ defined on $X^{(k+1)}$ be a regular k-dimensional f-sufficient statistic for the family $\{p_\theta(x),\ x \epsilon X : \theta \text{ in } I\}$ based on a sample of size $(k+1)$, $k \geqslant 1$. Then for any fixed $0 \leqslant \theta_0 < 1$, $\{p_\theta(x),\ x \text{ in } X_{\theta_0} : 0 \leqslant \theta \leqslant \theta_0\}$ has the following form :*

$$p_\theta(x) = exp\,[c_0(\theta) + \sum_1^k c_i(\theta)\phi_i(x) + \phi_0(x)] \ \forall\ x \text{ in } X_{\theta_0},\ 0 \leqslant \theta \leqslant \theta_0 \quad \ldots \quad (5.1)$$

where $\phi_0, \phi_1, \ldots, \phi_k$ are continuously differentiable functions on X_{θ_0}.

Proof: By Lemma 2.1 T is non-trivial on $X_{\theta_0}^{(k+1)}$. Hence the rank of the family of densities (they are probability densities except for a normalising constant factor) :

$$\{p_\theta(x), x \text{ in } X_{\theta_0} : 0 \leqslant \theta \leqslant \theta_0\} \text{ on } X_{\theta_0}$$

is less than or equal to k. It follows from a theorem of Dynkin (1951, Theorem 3, one should see Brown 1964; 1970 for a correction of Dynkin's theorems) that $p_\theta(x)$ has the form stated in (5.1). [Q.E.D.]

We shall show, in what follows, that if T defined on $X^{(k+1)}$ is a regular k-dimensional and a minimal f-sufficient statistic for a sample of size $(k+1)$, $k \geqslant 2$, then the functions $1, \phi_1, \phi_2, \ldots, \phi_k$ given by (5.1) of Lemma 5.1 are linearly dependent and the vector-space $\mathcal{L}_{\theta_0}$ has dimension k.

Let us define, as in Section 3,

$$\alpha(x) = \begin{cases} a^{-1}(x) & \text{if } a(0) < x < c \\ 1 & \text{if } x = c \\ b^{-1}(x) & \text{if } c < x < b(0) \end{cases}$$

and $\beta(x_1, \ldots, x_{k+1}) = \min\{\alpha(x_1), \alpha(x_2), \ldots, \alpha(x_{k+1})\}$. It is easy to see, as in Section 3 for $(x_1 \ldots x_{k+1})$ in X^{k+1}, that β can be written as a function of any f-sufficient statistic T for Γ based on a sample of size $(k+1)$. It is also easy to see that β is a regular one-dimensional statistic on $X^{(k+1)}$.

Lemma 5.2 : *If $T = (T_1, \ldots, T_k)$ is a regular k-dimensional statistic defined on $X^{(k+1)}$, $k \geqslant 2$, such that T is f-sufficient for Γ for a sample of size $(k+1)$, then for any fixed θ_0 in $[0, 1)$ the functions $1, \phi_1, \phi_2, \ldots, \phi_k$ on X_{θ_0} given by (5.1) of Lemma 5.1 are linearly dependent.*

Proof: Let $0 \leqslant \theta_0 < 1$ be fixed. If possible let the continuously differentiable functions $1, \phi_1, \phi_2, \ldots, \phi_k$, given by (5.1) of Lemma 5.1, be linearly independent.

Let $V(T)$ and $V(\beta)$ be the open dense subsets of $X^{(k+1)}$, which satisfies the regularity condition (b) of Definition 9 for T and β respectively.

3

Let $$W = V(T) \cap V(\beta) \cap X_{\theta_0}^{(k+1)}.$$

W is open and dense in $X_{\theta_0}^{(k+1)}$. From linear independence of $1, \phi_1, \phi_2, \ldots, \phi_k$, it follows by an argument used by Dynkin (part 2 of Theorem 2) that there exists a point

$$(x_1^0, x_2^0, \ldots, x_{k+1}^0) \text{ in } W \cap \{(x_1, x_2, \ldots, x_{k+1}) : a(x_{k+1}) < a(x_i) \;\forall\; i = 1, 2, \ldots, k\}$$

such that

$$\det((\phi_i'(x_j^0))) \neq 0.$$

Let

$$\psi_i(x_1, x_2, \ldots, x_{k+1}) = \sum_{j=1}^{k+1} \phi_i(x_j), \quad i = 1, 2, \ldots, k$$

and $\psi_{k+1}(x_1, x_2, \ldots, x_{k+1}) = \beta(x_1, x_2, \ldots, x_{k+1})$ for $(x_1, x_2, \ldots, x_{k+1})$ in $X_{\theta_0}^{k+1}$.

Then

$$\left.\frac{\partial(\psi_1, \ldots, \psi_{k+1})}{\partial(x_1, x_2, \ldots, x_{k+1})}\right|_{x^0 = (x_1^0, \ldots, x_{k+1}^0)} = \alpha'(x_{k+1}^0) \times \det((\phi_i'(x_j^0))) \neq 0.$$

It follows that there is a neighbourhood N, $N \subset V(T) \cap X_{\theta_0}^{k+1}$ of $x^0 = (x_1^0, \ldots, x_{k+1}^0)$ in $X_{\theta_0}^{k+1}$, such that $\psi = (\psi_1, \ldots, \psi_{k+1})$ is one-one on N. It is easy to check that ψ is minimal f-sufficient for $\{p_\theta(x), x \text{ in } X_{\theta_0} : 0 < \theta < \theta_0\}$. Hence there exists a function g on $T(N)$ such that

$$\psi(x_1, x_2, \ldots, x_{k+1}) = g \circ T(x_1, x_2, \ldots, x_{k+1})$$

which is a contradiction to the fact that T is not one-one on N (Lemma 2.1).

[Q.E.D.]

Lemma 5.3 : *Let T be a statistic defined on $X^{(k+1)}$ such that T is minimal f-sufficient for Γ for a sample of size $(k+1)$. If $E \subset X^{k+1}$ and S is any statistic such that $p_\theta(x_1).\ p_\theta(x_2) \ldots \ldots p_\theta(x_{k+1}) = g(\theta, S(x_1, x_2, \ldots, x_{k+1}))h(x_1, x_2, \ldots, x_{k+1})$ for all $(x_1, x_2, \ldots, x_{k+1})$ in E and θ in I for some functions g and h, then T can be expressed as a function of S on E.*

The above lemma is self-evident and is true for any family of densities.

Proposition 5.1 : Let $T = (T_1, T_2, \ldots, T_k)$ be a regular k-dimensional statistic defined on X^{k+1}, such that T is minimal f-sufficient for Γ, for a sample of size $(k+1)$. Let θ_0 in (0, 1) be arbitrary and let U be an open interval of the form $(a(\theta_0), b)$, $a(\theta_0) < b < b(\theta_0)$. Then the dimension of $\mathcal{L}_{\theta_0}^*(U)$ is k.

Proof : Since $\mathcal{L}_{\theta_0}$ has dimension less than or equal to k and since $U \subset X_{\theta_0}$ and $\mathcal{L}_{\theta_0}^*(U) \subset \mathcal{L}_{\theta_0}(U)$, we have that the dimension of $\mathcal{L}_{\theta_0}^*(U)$ is less than or equal to

k. If possible, let the dimension of $\mathcal{L}^{*}_{\theta_0}(U)$ be strictly less than k. Then we can choose points $\theta_1, \theta_2, \ldots, \theta_s$ in $(0, \theta_0)$ such that

(i) $0 < \theta_j < \theta_0$.

(ii) $s < k-1$,

(iii) the functions $1, r_{\theta_1}|U, r_{\theta_2}|U, \ldots, r_{\theta_s}|U$ form a basis for $\mathcal{L}^{*}_{\theta_0}(U)$.

Let $\max_{1 \leqslant j \leqslant s} \theta_j = \theta^*$. Let x_0 be any point in the interval $(a(\theta^*), a(\theta_0))$ such that

$$[U^k \times \{x_0\}] \cap V(T) \neq \phi$$

where $V(T)$ is the open dense subset of X^{k+1} which satisfies the condition (b) of Definition 9. Let $\alpha(x_0) = \theta'$. Then $\theta^* < \theta' < \theta_0$. Let $\theta_{s+1}, \theta_{s+2}, \ldots, \theta_{k-1}$ be chosen from the interval $(0, \theta']$ such that the functions $1, r_{\theta_1}|X_{\theta'}, r_{\theta_2}|X_{\theta'}, \ldots, r_{\theta_{k-1}}|X_{\theta'}$ generate $\mathcal{L}_{\theta'} = \mathcal{L}_{\theta'}(X_{\theta'})$.

Let
$$B = U^k \times \{x_0\}$$
for $(x_1, x_2, \ldots, x_k)$ in U^k let us define

$$f_j(x_1, x_2, \ldots, x_k) = \sum_{i=1}^{k} r_{\theta_j}(x_i), \qquad j = 1, 2, \ldots, s. \qquad \ldots \quad (5.2)$$

Now for any $0 \leqslant \theta < \theta_0$ there exist constants $c_0(\theta), c_1(\theta), \ldots, c_s(\theta)$ such that

$$r_\theta(x) = \sum_{j=1}^{s} c_j(\theta) r_{\theta_j}(x) + c_0(\theta) \text{ for all } x \text{ in } U. \qquad \ldots \quad (5.3)$$

Also for any $0 \leqslant \theta \leqslant \theta'$, there exist constants $d_0(\theta), d_1(\theta), \ldots, d_{k-1}(\theta)$ such that

$$r_\theta(x) = \sum_{j=1}^{(k-1)} d_j(\theta)\, r_{\theta_j}(x) + d_0(\theta)$$

$$\forall\, x \text{ in } \overline{(a(\theta'), b(\theta'))} = [a(\theta'), b(\theta')] \qquad \ldots \quad (5.4)$$

(by continuity of r_θ's).

Now if $(x_1, x_2, \ldots, x_{k+1})$ is in B we have for $\theta \geqslant \theta'$,

$$p_\theta(x_{k+1}) = p_\theta(x_0) = 0 \qquad \ldots \quad (5.5)$$

and for $\theta < \theta'$

$$r_\theta(x_1) + r_\theta(x_2) + \ldots + r_\theta(x_{k+1}) = \sum_{j=1}^{s} c_j(\theta) f_j(x_1, x_2, \ldots, x_k) + kc_0(\theta) + \sum_{j=1}^{k-1} d_j(\theta)\, r_{\theta_j}(x_0) + d_0(\theta) \qquad \ldots \quad (5.6)$$

[by (5.2), (5.3) and (5.4)].

Also for $(x_1, x_2, \ldots, x_{k+1})$ in B

$$\beta(x_1, x_2, \ldots, x_{k+1}) = \alpha(x_0). \qquad \ldots \quad (5.7)$$

By (5.5), (5.6) and (5.7), for $(x_1, x_2, \ldots, x_{k+1})$ in B and for all θ in $[0, 1)$ we can write $p_\theta(x_1) \ldots p_\theta(x_{k+1})$ in the following form :

$$p_\theta(x_1)\ldots p_\theta(x_{k+1}) = \psi(\theta, x_0, f_1(x_1, x_2, \ldots, x_k), f_2(x_1, \ldots, x_k), \ldots, f_s(x_1, \ldots, x_k)) \times h(x_1, x_2, \ldots, x_{k+1})$$

where $h(x_1, x_2, \ldots, x_{k+1}) = p_0(x_1) \ldots p_0(x_{k+1})$.

By the minimal f-sufficiency of T and Lemma 5.3 we have

$$T(x_1, x_2, \ldots, x_k, x_0) = \phi(f_1(x_1, x_2, \ldots, x_k), f_2(x_1, x_2, \ldots, x_k), \ldots, f_s(x_1, x_2, \ldots, x_k)) \;\forall\; (x_1, x_2, \ldots, x_k) \text{ in } U^k. \quad \ldots \quad (5.8)$$

Now since $U^k \times \{x_0\} \cap V(T)$ is not empty, there is a point $(x_1, x_2, \ldots, x_k)$ in $(a(\theta_0), b)^k$ where $T(x_1, x_2, \ldots, x_k, x_0)$ is continuously differentiable in all the variables $x_1, x_2, \ldots, x_k$ and the rank of the matrix

$$\left(\left(\frac{\partial T_i}{\partial x_j}\right)\right)_{\substack{i = 1, 2, \ldots, k \\ j = 1, 2, \ldots, k \\ x = (x_1, x_2, \ldots, x_k, x_0)}}$$

is at least $(k-1)$. But then the relation (5.8) contradicts Lemma 2.3 and the proposition is proved. ... (Q.E.D.)

Corollary : *Under the conditions of Proposition* 5.1 *if* $0 < \theta_0 < 1$ *and if* $U = (a(\theta_0), b)$ *where* $a(\theta_0) < b \leqslant b(\theta_0)$, *then the dimension* $\mathcal{L}_{\theta_0}(U)$ *is* k *and we can choose* $\theta_1, \theta_2, \ldots, \theta_{k-1}$ *in the open interval* $(0, \theta_0)$ *such that the functions* $1, r_{\theta_1} | U, r_{\theta_2} | U, \ldots, r_{\theta_{k-1}} | U$ *form a basis for* $\mathcal{L}_{\theta_0}(U)$.

Lemma 5.4 : *Let* $\{f_1, f_2, \ldots, f_n\}$ *be a finite family of real-valued functions on an arbitrary set* A. $f_1, f_2, \ldots, f_n$ *are linearly independent if and only if there exist* n *points* $x_1, x_2, \ldots, x_n$ *in* A *such that the determinant of the matrix*

$$((f_i(x_j))) \;\; i = 1, 2, \ldots, n; \; j = 1, 2, \ldots, n$$

is non-zero.

The proof is well-known for the case when A is a finite set. The proof for infinite A is similar.

Proposition 5.2 : Let $T = (T_1, T_2, \ldots, T_k)$ be a regular k-dimensional statistic defined on X^{k+1} such that T is minimal f-sufficient for the family $\Gamma = \{p_\theta(\,.\,) : \theta \epsilon I\}$, for a sample of size $(k+1)$ and let c be the point in X, given by assumption (A-5.5) on the densities. Then for each $0 < \theta_0 < 1$ and each neighbourhood U of c, $U \subset X_{\theta_0}$, $\mathcal{L}_{\theta_0}(U)$ has dimension k.

Proof : Let, if possible, $U = (a, b)$ be an open interval containing c, $U \subset X_{\theta_0}$ such that $\mathcal{L}_{\theta_0}(U)$ has dimension strictly less than k; a fortiori we have that the dimension of $\mathcal{L}_{\theta_0}((a, c])$ is less than k. Let

$$x_0 = \inf\{x : a(\theta_0) < x < c, \dim[\mathcal{L}_{\theta_0}((x, c])] < k\}.$$

Lemma 5.4 shows that if $\mathcal{L}_{\theta_0}((x, c])$ has dimension ν then there is an x' with $x < x' < c$ such that $\mathcal{L}_{\theta_0}((x', c])$ also has dimension ν. It follows that the infimum x_0 is attained. Thus $\mathcal{L}_{\theta_0}((x_0, c])$ has dimension $(s+1) < k$ and further Proposition 5.1 shows that $a(\theta_0) < x_0 \leqslant a < c$. By the definition of x_0 we further have that if $a(\theta_0) < x < x_0$ then

$$\dim [\mathcal{L}_{\theta_0}((x, c])] = k.$$

Let $\theta^* = \alpha(x_0)$. Then $\theta_0 < \theta^* < 1$ and $a(\theta^*) = x_0$. By Corollary to Proposition 5.1 let us choose, $\theta_1, \theta_2, \ldots, \theta_{k-1}$ in $(0, \theta_0)$ such that

(i) the functions $1, r_{\theta_1} | (x_0, c], \ldots, r_{\theta_s} | (x_0, c]$ form a basis for $\mathcal{L}_{\theta_0}[(x_0, c]]$, $s < k-1$.

(ii) the functions $1, r_{\theta_1} | X_{\theta_0}, \ldots, r_{\theta_{k-1}} | X_{\theta_0}$ form a basis for $\mathcal{L}_{\theta_0}$.

Now for $s+1 \leqslant j \leqslant k-1$, for x in $(x_0, c]$, $r_{\theta_j}(\,.\,)$ can be expressed as a linear combination of $1, r_{\theta_1}, \ldots, r_{\theta_s}$. Let $0 < \theta < \theta^*$ be arbitrary. If $0 < \theta \leqslant \theta_0$, then by (ii) $r_\theta(\cdot)$ on $(x_0, c]$ can be expressed as a linear combination of $1, r_{\theta_1}, \ldots, r_{\theta_s}$.

If $\theta_0 < \theta < \theta^*$, then there is an $x < x_0$ such that $(x, c] \subset X_\theta \subset X_{\theta_0}$. Therefore $1, r_{\theta_1}, r_{\theta_2}, \ldots, r_{\theta_{k-1}}$, when restricted to $(x, c]$ generate $\mathcal{L}_{\theta_0}((x, c])$ which has dimension k because $x < x_0$. It follows that $1, r_{\theta_1}, \ldots, r_{\theta_{k-1}}$ are linearly independent on $(x, c]$ and hence when restricted to $(x, c]$ they generate $\mathcal{L}_\theta(x, c]$. Thus on $(x, c]$ and a fortiori on $(x_0, c]$, r_θ is a linear combination of $1, r_{\theta_1}, \ldots, r_{\theta_{k-1}}$. But on $(x_0, c]$, $r_{\theta_{s+1}}, \ldots, r_{\theta_{k-1}}$ are linear combinations of $1, r_{\theta_1}, \ldots, r_{\theta_s}$. It follows that $1, r_{\theta_1}, \ldots, r_{\theta_s}$ when restricted to $(x_0, c]$ generate $\mathcal{L}^*_{\theta^*}((a(\theta^*), c])$. This contradicts Proposition 5.1. (Q.E.D.)

Remark: It is easily seen that the arguments used in the above Proposition yield a stronger conclusion viz. that under the conditions of Proposition 5.2, for each $0 < \theta_0 < 1$ and U a non-empty open subset of X_{θ_0}, $\mathcal{L}_{\theta_0}(U)$ has dimension k.

We shall now state our main result of this section.

Theorem 3: *If for any $k > 1$ there exists a statistic T, regular and k-dimensional, defined on $X^{(k+1)}$, such that T is minimal f-sufficient for the family of densities $\Gamma = \{p_\theta(\cdot) : \theta \epsilon I\}$ satisfying the conditions* (A-5.1) *to* (A-5.5), *for a sample of size* $(k+1)$, *then the family of densities Γ form a $(k-1)$-dimensional exponential family of densities.*

Proof: Let a sequence $\{\theta_n, n = 0, 1, 2, ...\}$ of points in $I = [0, 1)$ be so chosen that

(i) $0 < \theta_{n+1} < \theta_n \ \forall \ n = 0, 1, 2, ...$

(ii) $\lim_{n \to \infty} \theta_n = 0.$

Let $1, r_1^{(n)}, r_2^{(n)}, ..., r_{k-1}^{(n)}$ be chosen to be a basis for $\mathcal{L}_{\theta_n} = \mathcal{L}_{\theta_n}(X_{\theta_n})$ for $n = 0, 1, 2, ...$ For each $n \geqslant 0$ the functions $1, r_1^{(n)}, r_2^{(n)}, ..., r_{k-1}^{(n)}$ restricted to X_{θ_0} are linearly independent (Proposition 5.2). Therefore there exist uniquely determined constants $a_{ij}^{(n)}, \ i = 0, 1, 2, ..., k-1, \ j = 1, 2, ..., k-1, \ n = 0, 1, 2, ...$ such that

$$r_j^{(0)}(x) = \sum_{i=1}^{(k-1)} a_{ij}^{(n)} r_i^{(n)}(x) + a_{0j}^{(n)} \ \forall \ x \text{ in } X_{\theta_0}, \quad j = 1, 2, ..., k-1 \quad \text{and} \quad n = 0, 1, 2, ...$$

Let

$$\phi_j^{(n)}(x) = \sum_{i=1}^{(k-1)} a_{ij}^{(n)} r_i^{(n)}(x) + a_{0j}^{(n)} \ \forall \ x \text{ in } X_{\theta_n}.$$

Then it can be easily checked, by using Proposition 5.2, that

$$\phi_j^{(n)}(x) = \phi_j^{(n+1)}(x) \text{ on } X_{\theta_n} \quad \forall \ j = 1, 2, ..., k-1, \ n = 0, 1, 2, ...$$

Let $\phi_0, \phi_1, ..., \phi_{k-1}$ on X be defined as follows :

$$\phi_0(x) = -\log p_0(x) \ \forall \ x \text{ in } X$$

and

$$\phi_j(x) = \phi_j^{(n)}(x) \ \forall \ x \text{ in } X_{\theta_n}, \ n = 0, 1, 2, ...$$

It is easily verified by using Proposition 5.2 that on any neighbourhood of c and hence on each X_θ the restrictions of the functions $1, \phi_1, \phi_2, ..., \phi_{k-1}$ are linearly independent. Hence for each θ in I there exist constants $c_0(\theta), c_1(\theta), ..., c_{k-1}(\theta)$ such that

$$r_\theta(x) = \log p_\theta(x) - \log p_0(x) = c_0(\theta) + \sum_{j=1}^{k-1} c_j(\theta)\phi_j(x) \ \forall \ x \text{ in } X_\theta.$$

Hence

$$p_\theta(x) = \begin{cases} \exp\left[c_0(\theta) + \sum_1^{k-1} c_j(\theta)\, \phi_j(x) + \phi_0(x)\right] \ \forall \ x \text{ in } X_\theta \\ 0 \qquad \text{for } x \text{ in } X - X_\theta \end{cases}$$

where $1, \phi_1, \phi_2, ..., \phi_{k-1}$ are linearly independent functions on X.

It remains to verify that $1, c_1, c_2, ..., c_{k-1}$ are linearly independent functions on I. If they are linearly dependent, then by changing the numberings, if necessary, we shall have

$$c_{k-1}(\theta) = \sum_{i=1}^{k-2} b_i c_i(\theta) + b_0 \quad \forall \ \theta \text{ in } I.$$

Hence we have, for all θ in I,

$$r_\theta(x) = \sum_{i=1}^{k-2} c_i(\theta)\phi_i^*(x)+c_0(\theta)+b_0\phi_{k+1}(x) \qquad \forall\, x \text{ in } X_\theta$$

where
$$\phi_i^*(x) = \phi_i(x)+b_i\phi_{k-1}(x).$$

For $\theta = 0$ in particular

$$0 = b_0\,\phi_{k-1}(x) \quad \forall\, x \text{ in } X_0 = X.$$

Hence
$$r_\theta(x) = \sum_{1}^{k-2} c_i(\theta)\,\phi_i^*(x)+c_0(\theta) \quad \forall\, x \text{ in } X_\theta \text{ and } \theta \text{ in } I. \qquad \dots (5.9)$$

It follows from (5.9) that the dimension of $\mathcal{L}_\theta$, for any θ, $0 < \theta < 1$ is less than or equal to $(k-1)$, which is a contradiction to Proposition 5.1. (Q.E.D)

Remarks (1) If we replace the assumptions (A-5.1), (A-5.4) and (A-5.5), by the assump tions (A*-5.1), (A*-5.4) and (A*-5.5) respectively, which are stated below, then for the family of densities Γ satisfying (A*-5.1), (A-5.2), A-5.3), (A*-5.4) and (A*-5.5) and the other hypotheses of Theorem 3, the same conclusion viz. that Γ is a $(k-1)$-dimensional exponential family holds.

(A*-5.1) The parameter-space I is an interval of the type $(\lambda_0, \lambda_1]$ where $-\infty \leqslant \lambda_0 < \lambda_1 < \infty$.

(A*-5.4) $a(\theta)$ is a once continuously differentiable function of θ with strictly negative first derivative for each θ in I. $b(\theta)$ is a once continuously differentiable function of θ with strictly positive first derivative for each θ in I.

(A*-5.5) $a(\theta)$ increases to a finite constant c as θ decreases to λ_0 and $b(\theta)$ decreases to the same constant as θ decreases to λ_0.

In this case, the proof is identical with obvious modifications,

(2) The following converse of Theorem 3 is easy to prove.

If the family of densities Γ satisfies the conditions (A-5.1) to (A-5.5) and if Γ form a $(k-1)$-dimensional exponential family then there exists a statistic T, for a sample of size $(k+1)$ which is regular k-dimensional and is minimal f-sufficient for Γ for a sample of size $(k+1)$.

(3) In Theorem 3 we may consider $X^{(n)}$, $n > k+1$, instead of $X^{(k+1)}$ provided we suitably modify the regularity assumptions on T.

(4) The conditions on T in Theorem 3 are not wholly satisfactory. It would be interesting to get similar results when the rank of the matrix

$$\left(\left(\frac{\partial T_i}{\partial x_j}\right)\right) \quad i = 1, \dots, k\,; \;\; j = 1, \dots, k+1$$

is not a constant over the dense open set $V(T)$.

(5) The case when both $a(\theta)$ and $b(\theta)$ are increasing (or decreasing) can be handled in a similar way.

ACKNOWLEDGEMENT

We should like to thank the referee for some improvements in the presentation.

REFERENCES

APOSTOL, T. M. (1965): : *Mathematical Analysis*, Addison-Wesley Publishing Company, Inc.

BARANKIN, E. W. (1960): Sufficient statistics in the case of non-constant carrier, I and II. *Sankhyā*, Series A, 28, 101-114, 115-122.

BARANKIN, E. W. and MAITRA, A. P. (1963): Generalisation of the Fisher-Darmois-Koopman-Pitman theorem on sufficient statistics. *Sankhyā*, Series A, 25, 217-244.

BARANKIN, E. W. and KATZ, M. L. JR. (1959): Sufficient statistics of minimal dimension. *Sankhyā*, 21, 217-246.

BROWN, L. D. (1964): Sufficient statistics in the case of independent random variables. *Ann. Math. Stat.*, 35, 1456-1474.

——— (1970): Sufficient statistics in the case of independent random variables, Revisited and Revised. To be published in the *Ann. Math. Stat.*

DARMOIS, G. (1935): Sur les lois de probabilite a estimation exaustive. *C. R. Acad. Sci. Paris*, 260, 1265-1266.

DYNKIN, E. B. (1951): Necessary and sufficient statistics for a family of probability distributions. *Select Transl. Math. Stat. and Prob.*, 1, 23-41.

FRASER, D. A. S. (1966): Sufficiency for selection models, *Sankhyā*, Series A, 28, 329-334.

HUREWICZ, W. and WALLMAN, H. (1948): *Dimension Theory*, Princeton University Press.

HUZURBAZAR, V. S. (1965): The general form of distributions admitting sufficient statistics for parameters in non-regular cases. Unpublished.

KOOPMAN, L. H. (1936): On distributions admitting a sufficient statistic. *Trans. Amer. Math. Soc.* 39, 399-409.

LEHMANN, E. L. (1959): *Testing Statistical Hypotheses*, John Wiley and Sons, Inc., New York.

PITMAN, E. J. G. (1936): Sufficient statistics and intrinsic accuracy. *Proc. Camb. Phil. Soc.*, 32, 567-579.

Paper received : December, 1971.
Revised : May, 1972.