The Annals of Probability
1989, Vol. 17, No. 4, 1664–1690

FISHER INFORMATION AND DICHOTOMIES IN EQUIVALENCE / CONTIGUITY[1]

By Brian J. Thelen

University of Michigan

A contiguity dichotomy for two sequences of product measures is proved under the assumption of component measures belonging to a dominated experiment which is differentiable. This generalizes Eagleson's (1981) result for Gaussian measures. The dichotomy result is then used to generalize and clarify the results of Shepp (1965) and Steele (1986) with regards to finite Fisher information and equivalence dichotomies between two product measures, one with a fixed component measure and the second with rigidly perturbed component measures.

1. Introduction. There has been much interest from both probability and statistics in equivalence/singularity dichotomies within a family of probability measures on a measurable space. Specifically, if $(\Omega, \mathscr{F}, \mathscr{P})$ is an experiment [i.e., $(\Omega, \mathscr{F})$ is a measurable space and $\mathscr{P}$ is a set of probability measures on $(\Omega, \mathscr{F})$], we say that an equivalence/singularity dichotomy holds if $P, \tilde{P} \in \mathscr{P}$ implies $P \equiv \tilde{P}$ or $P \perp \tilde{P}$.

One of the first interesting dichotomies is due to Kakutani (1948), and is as follows. Let $(\Omega, \mathscr{F}) = (\prod_1^\infty \Omega_i, \sigma(\prod_1^\infty \mathscr{F}_i))$, where $\{(\Omega_i, \mathscr{F}_i)\}_1^\infty$ is a sequence of measurable spaces satisfying conditions needed for the Kolmogorov consistency theorem. Let $\{Q_i\}_1^\infty$ be a sequence of measures, where Q_i is on $(\Omega_i, \mathscr{F}_i)$ for all i. Then letting $\mathscr{P} = \{\prod_1^\infty P_i\colon\ P_i \equiv Q_i\ \forall\ i\}$, Kakutani showed that if $P = \prod_1^\infty P_i$, $\tilde{P} = \prod_1^\infty \tilde{P}_i$ are members of $\mathscr{P}$, then $P \equiv \tilde{P}$ or $P \perp \tilde{P}$, with the former being true if and only if

(1) $$\sum_1^\infty H^2(P_i, \tilde{P}_i) < \infty.$$

In (1), $H(P_i, \tilde{P}_i)$ is the Hellinger distance between P_i and $\tilde{P}_i$ and is defined by

(2) $$2H^2(P_i, \tilde{P}_i) = \int \left(\sqrt{f_i} - \sqrt{\tilde{f}_i}\right)^2 d\nu_i,$$

where $\nu_i = P_i + \tilde{P}_i$, $f_i \in dP_i/d\nu_i$ and $\tilde{f}_i \in d\tilde{P}_i/d\nu_i$.

Another interesting dichotomy holds in the case where $\Omega = \mathbb{R}^\infty$, $\mathscr{F} = \mathscr{B}(\mathbb{R}^\infty)$ (i.e., the Borel σ-field), and $\mathscr{P}$ is the set of all Gaussian probability measures. This was proved by Feldman (1958) and Hájek (1958). In the special case where $P = \prod_1^\infty \mathscr{N}(0,1)$ and $\tilde{P} = \prod_1^\infty \mathscr{N}(\mu_i, \sigma_i^2)$, they showed that $P \equiv \tilde{P}$ [it is an easy

Received May 1986; revised June 1988.

[1]Research supported by Office of Naval Research Contract N00014-85-K-0150.

AMS 1980 *subject classifications*. Primary 60G30; secondary 62F03.

Key words and phrases. Asymptotic separability, contiguity, dichotomy, equivalence, Fisher information, singularity.

exercise to verify this directly from (1)] if and only if

$$\sum_{1}^{\infty}\left(\mu_i^2 + (1 - \sigma_i)^2\right) < \infty. \tag{3}$$

On investigating (3) in the special case of $\sigma_i = 1$ for all i, we see that $P \equiv \tilde{P}$ if and only if $\{\mu_i\} \in l^2$. Note that $\mathscr{N}(\mu_i, 1)$ is just a translate of $\mathscr{N}(0, 1)$ by μ_i and an interesting question is: What other probability measures on $(\mathbb{R}, \mathscr{B}(\mathbb{R}))$, besides $\mathscr{N}(0, 1)$, satisfy this property? This was answered by Shepp (1965) who showed that if P is a probability measure on $(\mathbb{R}, \mathscr{B}(\mathbb{R}))$ and P_t denotes the translate of P by t, then $\prod_1^\infty P \equiv \prod_1^\infty P_{t_i}$ for all $\{t_i\} \in l^2$ if and only if $P \equiv \lambda$ (λ is Lebesgue measure) and P has finite Fisher information, that is, there exists a density f such that f is locally absolutely continuous and

$$\int \frac{(f')^2}{f}\, d\lambda < \infty. \tag{4}$$

He also showed $\prod_1^\infty P \perp \prod_1^\infty P_{t_i}$ for all $\{t_i\} \notin l^2$. Thus a necessary and sufficient condition for an l^2-type of dichotomy to hold in a translation experiment is that of finite Fisher information as defined in (4).

The above was extended to the case of $(\mathbb{R}^k, \mathscr{B}(\mathbb{R}^k))$ in Le Cam (1970), Proposition 2. Specifically, if P is a probability measure on $(\mathbb{R}^k, \mathscr{B}(\mathbb{R}^k))$, then it was shown that $\prod_1^\infty P \equiv \prod_1^\infty P_{t_i}$ for all $\{t_i\} \in l^2$ if and only if $P \equiv \lambda$ and the map

$$t \in \mathbb{R}^k \to \sqrt{f(\cdot + t)} \in L^2(\lambda) \text{ is differentiable,} \tag{5}$$

where $f \in dP/d\lambda$. Differentiability here means Frechet differentiability as a function from the Hilbert space $\mathbb{R}^k$ to the Hilbert space $L^2(\lambda)$. It is interesting to note that in the case of $k = 1$ and $P \equiv \lambda$, this implies that (4) and (5) are equivalent by comparing this with Shepp's result.

Shepp's result also was extended to $(\mathbb{R}^k, \mathscr{B}(\mathbb{R}^k))$ by Steele (1986). The generalization included not only translations but rigid motions as well, that is, translations composed with rotations. However Steele's definition of finite Fisher information is quite nonstandard and is not directly comparable with the conditions given in (4) and (5). We will show that it is actually equivalent to a differentiability condition similar to that given in (5).

All of the above results are related to equivalence/singularity dichotomies between two measures. However in the asymptotic theory of statistics there are useful generalizations of equivalence and singularity which are applicable to two sequences of measures. These are contiguity and asymptotic separation. Eagleson (1981) proved that a contiguity/asymptotic separation dichotomy holds between any two sequences of finite-dimensional Gaussian probability measures. He also gave necessary and sufficient conditions for the sequences to be contiguous.

The main two results of this paper are generalizations and clarifications of the previous results of Eagleson (1981) and Steele (1986). The first results (Theorems 3.2 and 3.3) give sufficient conditions for a contiguity/asymptotic separation dichotomy in the case of two sequences of product measures, where we assume that the component probability measures are from a dominated experiment

$E = (\Omega, \mathscr{F}, \{P_\theta: \theta \in \Theta\})$ with Θ in some finite-dimensional Euclidean space. The sufficient conditions are essentially L^2-differentiability of the map which takes $\theta \in \Theta$ to the square root of the density of P_θ with respect to the dominating measure. This generalizes Eagleson in that the Gaussian probability measures are a dominated experiment which satisfy this differentiability condition. We also show that in the case of a differentiable experiment under natural assumptions the two sequences of product measures are contiguous if and only if the l^2 distance between the component parameters remains asymptotically bounded.

The second result (Theorem 5.2) is a generalization and clarification of Steele's (1986) result and is a sort of converse to the previous result. Here we are concerned with a special experiment which is generated by a probability measure P and all rigid motion perturbations of P. We parameterize this experiment by (t, R), where t is the translation vector and R is the rotation matrix. By the previous result, L^2-differentiability implies a contiguity/asymptotic separation dichotomy, with necessary and sufficient conditions for contiguity being asymptotic boundedness of the l^2 distance between the sequences of component parameters. The second result of this paper is a converse to the above result. In particular, we show that if one has a contiguity/asymptotic separation dichotomy with contiguity holding if and only if the l^2 distance is bounded, then the experiment generated by rigid motion perturbations of P is L^2-differentiable. This clarifies Steele's result for it shows that his definition of finite Fisher information is equivalent to L^2-differentiability. It also extends the dichotomy to the more general contiguity/asymptotic separation framework and does not require (as Steele's dichotomy result did) that the component rigid motions converge to the identity.

2. Notation and preliminaries. We first give the basic definitions of some concepts discussed in the Introduction. An experiment E is a measurable space $(\Omega, \mathscr{F})$ along with a class of probability measures $\mathscr{P}$ and we write $E = (\Omega, \mathscr{F}, \mathscr{P})$.

Let $(\Omega, \mathscr{F}, \nu)$ be a σ-finite measure space and $\{f_\gamma: \gamma \in \Gamma\} \subset L^1(\nu)$. Then $\{f_\gamma: \gamma \in \Gamma\}$ is uniformly integrable (u.i.) if for all $\varepsilon > 0$ there exist an $h \in L^1(\nu)$ such that

$$\int (|f_\gamma| - h)_+ \, d\nu < \varepsilon \quad \text{for all } \gamma \in \Gamma, \tag{6}$$

where $(x)_+ = \max\{0, x\}$ for all $x \in \mathbb{R}$. For more details regarding uniform integrability in this general framework, see Fabian and Hannan (1985), Section 4.8, or Bauer (1981), Section 2.12. A well-known result is that since ν is σ-finite, $\{f_\gamma: \gamma \in \Gamma\}$ is u.i. if and only if for every sequence $\{\gamma_n\} \subset \Gamma$ there exists a subsequence $\{\gamma_{n'}\}$ such that $\{f_{\gamma_{n'}}\}$ is u.i.

Let $(\Omega_n, \mathscr{F}_n, \{P^n, \tilde{P}^n\})$ be a sequence of experiments. The sequence $\{\tilde{P}^n\}$ is contiguous to the sequence $\{P^n\}$ ($\tilde{P}^n \lhd P^n$) if for each sequence $\{B_n\}$ such that $B_n \in \mathscr{F}_n$ for each n and $P^n(B_n) \to 0$ as $n \to \infty$ implies that $\tilde{P}^n(B_n) \to 0$ as $n \to \infty$. The sequences $\{P^n\}$ and $\{\tilde{P}^n\}$ are mutually contiguous ($P^n \lhd \rhd \tilde{P}^n$) if $\{P^n\}$ is contiguous to $\{\tilde{P}^n\}$ and vice versa. The sequence $\{P^n\}$ is asymptotically separated from $\{\tilde{P}^n\}$ $(\tilde{P}^n \vartriangle P^n)$ if there exists a subsequence $\{n'\}$ and a corresponding subsequence of subsets $\{B_{n'}\}$ such that $B_{n'} \in \mathscr{F}_{n'}$, $P^{n'}(B_{n'}) \to 1$ and $\tilde{P}^{n'}(B_{n'}) \to 0$. Note that it is possible for $\{P^n\}$ and $\{\tilde{P}^n\}$ to be asymptotically separated and there exist a subsequence $\{n'\}$ of $\{n\}$ such that $\tilde{P}^{n'} \lhd P^{n'}$. It is easy to see in the special case of $(\Omega_n, \mathscr{F}_n) = (\Omega, \mathscr{F})$, $P^n = P$ and $\tilde{P}^n = \tilde{P}$ for all n that $\tilde{P}^n \lhd \rhd P^n$ if and only if $\tilde{P} \equiv P$, and $\tilde{P}^n \vartriangle P^n$ if and only if $\tilde{P} \perp P$. It is in this sense that contiguity and asymptotic separation are generalizations of equivalence and singularity.

Let $(\Omega, \mathscr{F}, \{P_\theta: \theta \in \Theta\})$ be an experiment with $\Theta \subset \mathbb{R}^d$. The experiment is dominated if there exists a σ-finite measure ν on $(\Omega, \mathscr{F})$ such that $P_\theta \ll \nu$ for all $\theta \in \Theta$. In the case of a dominated experiment there is a notion of differentiability which is defined as follows. Let $f_\theta \in dP_\theta/d\nu$ and $h_\theta = \sqrt{f_\theta} \in L^2(\nu)$ for all $\theta \in \Theta$. The experiment is differentiable at $\theta = \theta_0$ if the mapping $\theta \in \Theta \to h_\theta$ is differentiable as a mapping from Θ to the Hilbert space $L^2(\nu)$ at $\theta = \theta_0$, that is, there exists $\nabla h_{\theta_0} \in \prod_1^d L^2(\nu)$ such that

$$\lim_{\theta \to \theta_0} \frac{\left\| h_\theta - h_{\theta_0} - (\theta - \theta_0)^t \cdot \nabla h_{\theta_0} \right\|}{|\theta - \theta_0|} = 0, \tag{7}$$

where the limit is through $\theta \in \Theta$, $(\theta - \theta_0)^t$ is the transpose of $(\theta - \theta_0)$, and the norm on $\mathbb{R}^d$, $|\cdot|$, is the usual one. Note that we do not assume that Θ is open. The experiment E is differentiable if it is differentiable at all points in Θ and E is regular if it is continuously differentiable. It is easy to see that differentiability is not dependent on the dominating measure. Throughout this paper all parameter spaces are assumed to be in finite-dimensional Euclidean space, and in any Euclidean space we use the usual norm which we always denote by $|\cdot|$.

Suppose $(\Omega, \mathscr{F}, \{P, \tilde{P}\})$ is an experiment. The Hellinger distance, $H(P, \tilde{P})$, between P and $\tilde{P}$ is defined by

$$2H^2(P, \tilde{P}) = \int (h - \tilde{h})^2 \, d\nu, \tag{8}$$

where $\nu = P + \tilde{P}$, $h \in \sqrt{dP/d\nu}$ and $\tilde{h} \in \sqrt{d\tilde{P}/d\nu}$. It is easy to see that ν could be replaced by any σ-finite measure which dominates P and $\tilde{P}$ (when of course h and $\tilde{h}$ are replaced by the obvious functions), and that $0 \le H(P, \tilde{P}) \le 1$. Also it is obvious that $H(P, \tilde{P}) = 0$ if and only if $P = \tilde{P}$, and that $H(P, \tilde{P}) = 1$ if and only if $P \perp \tilde{P}$.

For the rest of this paper we use the notation f and h for densities and square roots of densities, respectively, along with subscripts or superscripts (f_θ corresponds to P_θ, etc.) to indicate their associated probability measures without further comment.

There are important relationships between the Hellinger metric and contiguity/asymptotic separation. The following equation is well known [cf. Strasser (1985), Lemma 2.15] and quite useful:

$$2H^2(P, \tilde{P}) \le \|P - \tilde{P}\| \le 2H(P, \tilde{P})\sqrt{2 - H^2(P, \tilde{P})}, \tag{9}$$

where $\|\cdot\|$ is the total variation norm. It is easy to see that $P^n \vartriangle \tilde{P}^n$ if and only if

$$\limsup_{n\to\infty} \|P^n - \tilde{P}^n\| = 2. \tag{10}$$

Thus

$$P^n \vartriangle \tilde{P}^n \quad \text{if and only if } \limsup_{n\to\infty} H(P^n, \tilde{P}^n) = 1 \tag{11}$$

by (9). For the infinite product situation, we can easily monitor the Hellinger distance by monitoring the Hellinger distances between the components. If $P^n = \prod_{i=1}^{\infty} P_{ni}$ and $\tilde{P}^n = \prod_{i=1}^{\infty} \tilde{P}_{ni}$, then

$$\begin{aligned} H^2(P^n, \tilde{P}^n) &= 1 - \prod_{i=1}^{\infty} \int h_{ni} \tilde{h}_{ni}\, d\nu_{ni} \\ &= 1 - \prod_{i=1}^{\infty} \left(1 - H^2\left(P_{ni}, \tilde{P}_{ni}\right)\right). \end{aligned} \tag{12}$$

By (11) and (12), $P^n \vartriangle \tilde{P}^n$ if and only if either

$$\limsup_{n\to\infty} \sum_{i=1}^{\infty} H^2\left(P_{ni}, \tilde{P}_{ni}\right) = \infty$$

(13) or

$$\limsup_{n\to\infty} \sup\left\{H\left(P_{ni}, \tilde{P}_{ni}\right) \colon i \in N\right\} = 1.$$

For further details regarding contiguity, differentiability and the Hellinger metric and their connection to asymptotic statistics see Strasser (1985).

Let $(\Omega, \mathscr{F}, \{P_\theta\colon \theta \in \Theta\})$ be an experiment. We now list some assumptions which will be invoked later on.

(A.1) $P_\theta \equiv P_{\theta'}$ for all $\theta, \theta' \in \Theta$ (homogeneity).

(A.2) $\lim_{\theta' \to \theta} H(P_\theta, P_{\theta'}) = 0$ for all $\theta \in \Theta$ (continuity). When there is no possible confusion we write $H(\theta, \theta')$ in place of $H(P_\theta, P_{\theta'})$.

(A.3) $\lim_{n\to\infty} H(\theta, \theta_n) = 1$ for each $\theta \in \Theta$ and sequence $\{\theta_n\} \in \Theta$ such that either $|\theta_n| \to \infty$ or $\theta_n \to t \in \overline{\Theta} \setminus \Theta$, where $\overline{\Theta}$ denotes the closure of Θ (asymptotic separation at the boundary).

(A.4) $P_\theta \neq P_{\theta'}$ for all $\theta \neq \theta'$ (identifiability).

Since $\Theta \subset \mathbb{R}^d$, (A.1), (A.2) and (A.3) imply $\overline{\Theta} \setminus \Theta$ is closed. This is stated and proved in the following proposition.

Proposition 2.1. *Let $E = (\Omega, \mathscr{F}, \{P_\theta\colon \theta \in \Theta\})$ be an experiment which satisfies* (A.1), (A.2) *and* (A.3). *Then $\overline{\Theta} \setminus \Theta$ is closed.*

Proof. If $\overline{\Theta} \setminus \Theta = \varnothing$, we are done. So suppose that it is not null and let $\{t_j\} \subset \overline{\Theta} \setminus \Theta$ be such that $t_j \to t \in \mathbb{R}^d$. Let $\theta_0 \in \Theta$. By (A.3) there exists $\{\theta_j\} \in \Theta$ such that $|t_j - \theta_j| \to 0$ and $H(\theta_j, \theta_0) \to 1$. Also if $t \in \Theta$, $H(\theta_j, t) \to 0$ by (A.2). This would then imply $H(\theta_0, t) = 1$, a contradiction to (A.1), since $H(\theta_0, t) \geq H(\theta_j, \theta_0) - H(\theta_j, t)$ for all j. Hence $t \notin \Theta$ and this implies that $t \in \overline{\Theta} \setminus \Theta$. □

Finally, whenever there is an infinite product measure it is implicitly assumed that the component measures are compact, that is, there exists a compact subclass [cf. Neveu (1965), page 26] in the component σ-field such that the measure of any set is the supremum of the sets in the subclass (this is needed to ensure the existence of the infinite product measures).

3. Sufficient conditions for contiguity / asymptotic separation dichotomy.

A corollary of a more general result in Liptser, Pukelsheim and Shiryaev (1982) which gives necessary and sufficient conditions for contiguity in the infinite product situation is stated below [this is actually a generalization of a result in Oosteroff and van Zwet (1979)].

PROPOSITION 3.1. *Let* $E_{ni} = (\Omega_{ni}, \mathscr{F}_{ni}, \{P_{ni}, \tilde{P}_{ni}\})$ *be an experiment for all* $n, i \in \mathbb{N}$, $P^n = \prod_1^\infty P_{ni}$ *and* $\tilde{P}^n = \prod_1^\infty \tilde{P}_{ni}$. *Then* $\tilde{P}^n \lhd \rhd P^n$ *if and only if*

$$\limsup_{n\to\infty} \sum_1^\infty H^2(P_{ni}, \tilde{P}_{ni}) < \infty, \tag{14}$$

$$\lim_{K\to\infty} \limsup_{n\to\infty} \sum_1^\infty \tilde{P}_{ni}(\tilde{f}_{ni} > K f_{ni}) = 0 \tag{15}$$

and

$$\lim_{K\to\infty} \limsup_{n\to\infty} \sum_1^\infty P_{ni}(f_{ni} > K\tilde{f}_{ni}) = 0. \tag{16}$$

Before stating and proving the main results in this section, we need to state a technical lemma which states that contiguity and asymptotic separation are not affected by measurable transformations which are one-to-one, onto and have measurable inverses. The proof is easy and is omitted.

LEMMA 3.1. *Let* $E_n = (\Omega_n, \mathscr{F}_n, \{P^n, \tilde{P}^n\})$ *be a sequence of experiments and let* I_n *be a measurable transformation from* $(\Omega_n, \mathscr{F}_n)$ *onto itself which is one-to-one, onto and has a measurable inverse. Let* $Q^n = P^n \circ I_n$ *and* $\tilde{Q}^n = \tilde{P}^n \circ I_n$. *Then* $\tilde{P}^n \lhd P^n$ *if and only if* $\tilde{Q}^n \lhd Q^n$ *and* $\tilde{P}^n \vartriangle P^n$ *if and only if* $\tilde{Q}^n \vartriangle Q^n$.

REMARK. Let $E_{ni} = (\Omega, \mathscr{F}, \{P_{ni}, \tilde{P}_{ni}\})$ be an experiment for all $n, i \in \mathbb{N}$, $P^n = \prod_1^\infty P_{ni}$ and $\tilde{P} = \prod_1^\infty \tilde{P}_{ni}$. Let Q^n and $\tilde{Q}^n$ be obtained by a common rearrangement of the component probability measures of P^n and $\tilde{P}^n$, respectively (rearrangement can be different for each n). By Lemma 3.1, mutual contiguity between P^n and $\tilde{P}^n$ is equivalent to mutual contiguity between Q^n and $\tilde{Q}^n$. A similar statement holds for asymptotic separation. This fact will prove useful for simplifying some arguments and calculations.

Based on the previous results we now state and prove a technical theorem from which the main results of this section follow.

Theorem 3.1. *Let* $E = (\Omega, \mathscr{F}, \{P_\theta\colon \theta \in \Theta\})$ *be an experiment satisfying* (A.1) *through* (A.4), $P^n = \prod_1^\infty P_{\theta_{ni}}$ *and* $\tilde{P}^n = \prod_1^\infty P_{\tilde{\theta}_{ni}}$, *where* $\{\theta_{ni}\colon n, i \in \mathbb{N}\} \subset \Theta_c$ *with* Θ_c *being a compact subset of* Θ. *Also assume that if* ν *is a dominating measure and* Θ_c' *is any compact subset of* Θ, *then*

$$\left\{\frac{(h_\theta - h_{\theta'})^2}{|\theta - \theta'|^2}\colon \theta \in \Theta_c,\, \theta' \in \Theta_c'\right\} \text{ is u.i.} \tag{17}$$

and

$$\lim_{\rho \to 0} \inf\left\{\frac{H(\theta, \theta')}{|\theta - \theta'|}\colon \theta \in \Theta_c,\, |\theta - \theta'| < \rho\right\} > 0. \tag{18}$$

Then $P^n \lhd \rhd \tilde{P}^n$ *or* $P^n \vartriangle \tilde{P}^n$ *with the former occurring if and only if*

$$\liminf_{n \to \infty} \inf\left\{\operatorname{dist}\left(\tilde{\theta}_{ni}, \overline{\Theta} \setminus \Theta\right)\colon i \in \mathbb{N}\right\} > 0 \tag{19}$$

and

$$\limsup_{n \to \infty} \sum_1^\infty |\theta_{ni} - \tilde{\theta}_{ni}|^2 = M < \infty, \tag{20}$$

where $\overline{\Theta}$ *is the closure of* Θ *and* $\operatorname{dist}(\tilde{\theta}_{ni}, \varnothing) = 1$ *by convention* ($\varnothing$ *denotes the null set*).

Proof. First suppose

$$\limsup_{n \to \infty} \sup\left\{|\tilde{\theta}_{ni}|\colon i \in \mathbb{N}\right\} = \infty. \tag{21}$$

By a previous remark, without loss of generality there exist subsequence $\{n'\}$ such that $|\tilde{\theta}_{n'1}| \to \infty$ and $\theta_{n'1} \to \theta_0 \in \Theta$. Then $H(\theta_0, \tilde{\theta}_{n'1}) \to 1$ by (A.3) and $H(\theta_0, \theta_{n'1}) \to 0$ by (A.2). But by the triangle inequality and algebra,

$$H\left(\theta_{n'1}, \tilde{\theta}_{n'1}\right) \geq H\left(\theta_0, \tilde{\theta}_{n'1}\right) - H(\theta_0, \theta_{n'1})$$

for all n' and on taking the limit infimum of both sides, $H(\theta_{n'1}, \tilde{\theta}_{n'1}) \to 1$. Thus by (13), $P^n \vartriangle \tilde{P}^n$.

Now suppose that (19) is false. By the above argument if (21) is true, then $P^n \vartriangle \tilde{P}^n$, so assume that (21) is also false. By the previous remark, without loss of generality there exist a subsequence $\{n'\}$ and $\{t_{n'}\} \subset \overline{\Theta} \setminus \Theta$, such that $\tilde{\theta}_{n'1} \to t \in \mathbb{R}^d$, $\theta_{n'1} \to \theta_0 \in \Theta$ and $|\tilde{\theta}_{n'1} - t_{n'}| \to 0$. By the triangle inequality, $t_{n'} \to t$, and hence $t \notin \Theta$ by Proposition 2.1. Thus $P^n \vartriangle \tilde{P}^n$ by (A.2), (A.3), (13) and an argument similar to the one used in the previous paragraph.

Now suppose (20) is false, and we now want to show $P^n \vartriangle \tilde{P}^n$. By the previous arguments it suffices to prove $P^n \vartriangle \tilde{P}^n$ under the additional assumptions that (21) is false and (19) is true. By these additional assumptions, there exist an $N \in \mathbb{N}$ and a compact set $\Theta_c' \subset \Theta$ such that $\{\tilde{\theta}_{ni}\colon n \geq N,\ i \in \mathbb{N}\} \subset \Theta_c'$. Let

$\rho_0 > 0$ and $\alpha_0 > 0$ be such that

$$\frac{H(\theta, \theta')}{|\theta - \theta'|} > \alpha_0 \tag{22}$$

for $\theta \in \Theta_c$, $\theta' \in \Theta_c'$ and $|\theta - \theta'| < \rho_0$.

Also the LHS of (22) is bounded away from 0 when considering $\theta \in \Theta_c$, $\theta \in \Theta_c'$ and $|\theta - \theta'| \geq \rho_0$. To show this, assume that it is not bounded away from 0. Then we can choose sequences $\{\theta_j\} \subset \Theta_c$ and $\{\theta_j'\} \subset \Theta_c'$ such that $\theta_j \to \theta_0 \in \Theta$, $\theta_j' \to \theta_0' \in \Theta$, $|\theta_j - \theta_j'| \geq \rho_0$ and $H(\theta_j, \theta_j') \to 0$. By (A.2), $H(\theta_0, \theta_0') = 0$, which contradicts the identifiability assumption of (A.4). Thus the LHS of (22) is bounded away from 0 when $|\theta - \theta'| \geq \rho_0$.

By combining the above with (17) and (22), there exist $\alpha, \beta \in (0, \infty)$ such that

$$\alpha|\theta - \theta'|^2 \leq H^2(\theta, \theta') \leq \beta|\theta - \theta'|^2 \quad \text{for } \theta \in \Theta_c,\ \theta' \in \Theta_c'. \tag{23}$$

Since (20) is false, (23) and (13) imply $P^n \vartriangle \tilde{P}^n$.

For the final case suppose (19) and (20) are true with M as the constant given in (20). By the remark preceding this theorem, without loss of generality we can assume

$$|\theta_{ni} - \tilde{\theta}_{ni}| \geq |\theta_{n,i+1} - \tilde{\theta}_{n,i+1}| \quad \text{for all } i, n \in \mathbb{N}. \tag{24}$$

As in the previous paragraph, since (20) implies (21) is false, there exists a compact set $\Theta_c' \subset \Theta$ and $N_0 \in \mathbb{N}$ such that $\{\tilde{\theta}_{ni}: n \geq N_0, i \in \mathbb{N}\} \subset \Theta_c'$. Since (23) is only dependent on the hypothesis given in the statement of the theorem, it holds in this case as well. This combined with (20) implies (14) is true.

We now want to verify that (15) and (16) hold. Let $\varepsilon, \varepsilon_0 > 0$. Then let $\{\theta_j\} \in \Theta_c$ and $\{\theta_j'\} \in \Theta_c'$ be sequences such that $\theta_j \neq \theta_j'$ for all j and $|\theta_j - \theta_j'| \to 0$. Then there exist subsequences $\{\theta_{j'}\}, \{\theta_{j'}'\}$ such $\theta_{j'} \to \theta_0$ and $\theta_{j'}' \to \theta_0$ and by assumption (A.1) we can also assume that $h_{\theta_{j'}} \to h_{\theta_0}$ and $h_{\theta_{j'}'} \to h_{\theta_0}$ a.e.-ν. This implies

$$\begin{aligned}
&\lim_{j' \to \infty} \frac{P_{\theta_{j'}}\left\{\left(h_{\theta_{j'}} - h_{\theta_{j'}'}\right)^2 \geq \varepsilon_0 h_{\theta_{j'}}^2\right\}}{H^2\left(\theta_{j'}, \theta_{j'}'\right)} \\
&\quad = \lim_{j' \to \infty} \frac{\int 1_{\{(h_{\theta_{j'}} - h_{\theta'_{j'}})^2 \geq \varepsilon_0 h_{\theta_{j'}}^2\}} \dfrac{\left(h_{\theta_{j'}} - h_{\theta_{j'}'}\right)^2}{|\theta_{j'} - \theta_{j'}'|^2}\, d\nu}{\varepsilon_0 \dfrac{H^2\left(\theta_{j'}, \theta_{j'}'\right)}{|\theta_{j'} - \theta_{j'}'|^2}} \\
&\quad = 0
\end{aligned} \tag{25}$$

by (23) and (17) which implies u.i. of the integrands. By an exactly analogous argument

$$\lim_{j' \to \infty} \frac{P_{\theta_{j'}'}\left\{\left(h_{\theta_{j'}} - h_{\theta_{j'}'}\right)^2 \geq \varepsilon_0 h_{\theta_{j'}'}^2\right\}}{H^2\left(\theta_{j'}, \theta_{j'}'\right)} = 0. \tag{26}$$

Since the original sequences were arbitrary we have actually shown

$$\lim_{\rho\to 0} \sup\left\{\frac{P_\theta\{(h_\theta - h_{\theta'})^2 \geq \varepsilon_0 h_\theta^2\}}{H^2(\theta,\theta')} : |\theta - \theta'| \leq \rho\right\} = 0$$

and

$$\lim_{\rho\to 0} \sup\left\{\frac{P_{\theta'}\{(h_\theta - h_{\theta'})^2 \geq \varepsilon_0 h_{\theta'}^2\}}{H^2(\theta,\theta')} : |\theta - \theta'| \leq \rho\right\} = 0,$$

where both limits are over $\theta \in \Theta_c$ and $\theta' \in \Theta_c'$ with $\theta \neq \theta'$. Thus there exists a $\rho > 0$ such that

$$\frac{P_\theta\{(h_\theta - h_{\theta'})^2 \geq \varepsilon_0 h_\theta^2\}}{H^2(\theta,\theta')} < \varepsilon \tag{27}$$

and

$$\frac{P_{\theta'}\{(h_\theta - h_{\theta'})^2 \geq \varepsilon_0 h_{\theta'}^2\}}{H^2(\theta,\theta')} < \varepsilon \tag{28}$$

for $\theta \in \Theta_c$, $\theta' \in \Theta_c'$ such that $|\theta - \theta'| < \rho$. To show that $P^n \lhd \rhd \tilde{P}^n$, it suffices to prove that any subsequence $\{n'\}$ has a subsubsequence $\{n''\}$ such that $P^{n''} \lhd \rhd \tilde{P}^{n''}$. Hence it suffices to prove $P^{n'} \lhd \rhd \tilde{P}^{n'}$ for a subsequence $\{n'\}$ such that $\theta_{n'i} \to \theta_i \in \Theta$, $\tilde{\theta}_{n'i} \to \tilde{\theta}_i \in \Theta$, $f_{\theta_{n'i}} \to f_{\theta_i}$ a.e.-ν and $f_{\tilde{\theta}_{n'i}} \to f_{\tilde{\theta}_i}$ a.e.-ν for all $i \in \mathbb{N}$, since (A.2) is true and since Θ_c and Θ_c' are both compact. Let $C = \max\{M/\rho^2, 1\}$. By (A.1) there exist $K > (1 + \varepsilon_0)^2$ such that

$$\limsup_{n'\to\infty} P_{\theta_{n'i}}(f_{\theta_{n'i}} > K f_{\tilde{\theta}_{n'i}}) \leq P_{\theta_i}(f_{\theta_i} \geq K f_{\tilde{\theta}_i}) < \frac{\varepsilon}{C} \tag{29}$$

and

$$\limsup_{n'\to\infty} P_{\tilde{\theta}_{n'i}}(f_{\tilde{\theta}_{n'i}} > K f_{\theta_{n'i}}) \leq P_{\tilde{\theta}_i}(f_{\tilde{\theta}_i} \geq K f_{\theta_i}) < \frac{\varepsilon}{C} \tag{30}$$

for $i \leq C$. There exist $N \in \mathbb{N}$ such that $n \geq N$ and $i > C$ implies $|\theta_{ni} - \tilde{\theta}_{ni}| < \rho$ by (20) and (24). By combining (29) and (30) with (23), (27) and (28),

$$\limsup_{n'\to\infty} \sum_1^\infty P_{\theta_{n'i}}(f_{\theta_{n'i}} > K f_{\tilde{\theta}_{n'i}}) < \varepsilon(1 + M\beta) \tag{31}$$

and

$$\limsup_{n'\to\infty} \sum_1^\infty P_{\tilde{\theta}_{n'i}}(f_{\tilde{\theta}_{n'i}} > K f_{\theta_{n'i}}) < \varepsilon(1 + M\beta). \tag{32}$$

Since ε was arbitrary, this implies $P^{n'} \lhd \rhd \tilde{P}^{n'}$ by Proposition 3.1. □

REMARK. Let $E = (\Omega, \mathscr{F}, \{P_\theta: \theta \in \Theta\})$ be a dominated experiment and let ν_1, ν_2 be two dominating σ-finite measures. Suppose E is differentiable at $\theta = \theta_0$ with differential $\nabla h_{\theta_0}^j$ for measure ν_j, $j = 1, 2$. Then it is an easy exercise to show that

$$\int \nabla h^1_{\theta_0} \cdot \left(\nabla h^1_{\theta_0}\right)^t d\nu_1 = \int \nabla h^2_{\theta_0} \cdot \left(\nabla h^2_{\theta_0}\right)^t d\nu_2. \tag{33}$$

In particular, $\int \nabla h^1_{\theta_0} \cdot (\nabla h^1_{\theta_0})^t d\nu_1$ is nonsingular if and only if $\int \nabla h^2_{\theta_0} \cdot (\nabla h^2_{\theta_0})^t d\nu_2$ is nonsingular, that is, nonsingularity is independent of the dominating measure. We say that the differential is nonsingular at $\theta = \theta_0$ if the matrix as given in (33) is nonsingular.

Theorem 3.2. *Let $E = (\Omega, \mathscr{F}, \{P_\theta: \theta \in \Theta\})$ be an experiment differentiable at $\theta = \theta_0$ and assume E satisfies* (A.1) *through* (A.4). *Let ν be a dominating measure and assume that the differential ∇h_{θ_0} is nonsingular. If $P^n = \prod_1^\infty P_{\theta_0}$ and $\tilde{P}^n = \prod_1^\infty P_{\tilde{\theta}_{ni}}$, the $P^n \lhd \rhd \tilde{P}^n$ or $P^n \vartriangle \tilde{P}^n$ with the former being true if and only if*

$$\liminf_{n\to\infty} \inf\left\{\operatorname{dist}\left(\tilde{\theta}_{ni}, \overline{\Theta} \setminus \Theta\right): i \in \mathbb{N}\right\} > 0 \tag{34}$$

and

$$\limsup_{n\to\infty} \sum_1^\infty |\theta_0 - \tilde{\theta}_{ni}|^2 < \infty. \tag{35}$$

Proof. Let $\Theta_c = \{\theta_0\}$ and note that

$$\begin{aligned} &\lim_{\rho\to 0} \inf\left\{\frac{H^2(\theta_0, \theta')}{|\theta_0 - \theta'|^2}: |\theta_0 - \theta'| < \rho\right\} \\ &= \lim_{\rho\to 0} \inf\left\{\frac{\left\|(\theta_0 - \theta')^t \cdot \nabla h_{\theta_0}\right\|^2}{|\theta_0 - \theta'|^2}: |\theta_0 - \theta'| < \rho\right\} \\ &\geq \inf\left\{t^t \cdot \left(\int \nabla h_{\theta_0} \cdot \left(\nabla h_{\theta_0}\right)^t d\nu\right) \cdot t: t \in \mathbb{R}^d, |t| = 1\right\} \\ &> 0. \end{aligned} \tag{36}$$

Thus (18) holds.

Let Θ_c' be a compact subset of Θ. In order to verify (17) it suffices to prove that if $\{\theta_j\}$ is a convergent sequence in $\Theta_c' \setminus \{\theta_0\}$ with

$$\frac{\theta_j - \theta_0}{|\theta_j - \theta_0|} \to t \in \mathbb{R}^d,$$

then

$$\left\{\left(h_{\theta_j} - h_{\theta_0}\right)^2 \middle/ |\theta_j - \theta_0|^2: j \in \mathbb{N}\right\} \text{ is u.i.}$$

To show this, we consider two cases. First suppose $\theta_j \to \theta_0$. Then $(h_{\theta_j} - h_{\theta_0})/|\theta_j - \theta_0| \to t^t \cdot \nabla h_{\theta_0}$ in $L^2(\nu)$ and hence $(h_{\theta_j} - h_{\theta_0})^2/|\theta_j - \theta_0|^2 \to (t^t \cdot \nabla h_{\theta_0})^2$ in

$L^1(\nu)$ which implies u.i. For the second case suppose $\theta_j \to \theta \neq \theta_0$. Then by (A.2),

$$\frac{h_{\theta_j} - h_{\theta_0}}{|\theta_j - \theta_0|} \to \frac{h_\theta - h_{\theta_0}}{|\theta - \theta_0|} \quad \text{in } L^2(\nu)$$

and hence

$$\frac{\left(h_{\theta_j} - h_{\theta_0}\right)^2}{|\theta_j - \theta_0|^2} \to \frac{\left(h_\theta - h_{\theta_0}\right)^2}{|\theta - \theta_0|^2} \quad \text{in } L^1(\nu),$$

which again implies u.i. Hence we have verified (17) and so by Theorem 3.1, the result follows. □

THEOREM 3.3. *Let $E = (\Omega, \mathscr{F}, \{P_\theta\colon \theta \in \Theta\})$ be a regular experiment which satisfies* (A.1) *through* (A.4). *Let ν be a dominating σ-finite measure and assume that the differential ∇h_θ is nonsingular for all $\theta \in \Theta$. Furthermore, assume that Θ is locally convex and $\{\theta_{ni}\colon i, n \in \mathbb{N}\} \subset \Theta_c \subset \Theta$, where Θ_c is compact. If $P^n = \prod_1^\infty P_{\theta_{ni}}$ and $\tilde{P}^n = \prod_1^\infty P_{\tilde{\theta}_{ni}}$, then $P^n \lhd\rhd \tilde{P}^n$ or $P^n \vartriangle \tilde{P}^n$ with the former being true if and only if* (19) *and* (20) *are true.*

PROOF. Let Θ_c' be a compact subset of Θ. To show (17) and (18) are true, and hence by Theorem 3.1 obtain the desired result, it suffices to prove that if $\{\theta_j\} \subset \Theta_c$ and $\{\theta_j'\} \subset \Theta_c'$, there exist subsequences $\{\theta_{j'}\}$ and $\{\theta_{j'}'\}$ such that $\{\|h_{\theta_{j'}} - h_{\theta_{j'}'}\|/|\theta_{j'} - \theta_{j'}'|\}$ is bounded away from 0 and $\{(h_{\theta_{j'}} - h_{\theta_{j'}'})^2/|\theta_{j'} - \theta_{j'}'|^2\}$ is u.i. Thus dropping the subsequence notation for convenience, it suffices to prove that $\{\|h_{\theta_j} - h_{\theta_j'}\|/|\theta_j - \theta_j'|\}$ is bounded away from 0 and $\{(h_{\theta_j} - h_{\theta_j'})^2/|\theta_j - \theta_j'|^2\}$ is u.i. under the additional assumptions that $\theta_j \to \theta_0 \in \Theta$, $\theta_j' \to \theta_0' \in \Theta$ and $(\theta_j - \theta_j')/|\theta_j - \theta_j'| \to t \in \mathbb{R}^d$.

If $\theta_0 = \theta_0'$, then by the local convexity of Θ, the continuity of the differential and a standard differential calculus result for normed linear spaces [cf. Loomis and Sternberg (1968), page 149],

$$\frac{h_{\theta_j} - h_{\theta_j'}}{|\theta_j - \theta_j'|} \to t^t \cdot \nabla h_{\theta_0} \quad \text{in } L^2(\nu), \tag{37}$$

which implies $\{(h_{\theta_j} - h_{\theta_j'})^2/|\theta_j - \theta_j'|^2\colon j \in \mathbb{N}\}$ is u.i. Also (37) implies

$$\liminf_{j\to\infty} \frac{H(\theta_j, \theta_j')}{|\theta_j - \theta_j'|} > 0, \tag{38}$$

since ∇h_{θ_0} is nonsingular at $\theta = \theta_0$.

In the case $\theta_0 \neq \theta_0'$,

$$\frac{h_{\theta_j} - h_{\theta_j'}}{|\theta_j - \theta_j'|} \to \frac{h_{\theta_0} - h_{\theta_0'}}{|\theta - \theta_0|} \quad \text{in } L^2(\nu) \tag{39}$$

by (A.2). Hence we again have $\{(h_{\theta_j} - h_{\theta_j'})^2/|\theta_j - \theta_j'|^2\colon j \in \mathbb{N}\}$ is u.i. Also in this case (38) holds by (39) and (A.4). The result follows. □

4. Applications of sufficiency results. In this section we use the previous results to prove a contiguity/asymptotic separation dichotomy for sequences of Gaussian processes with arbitrary index sets, for sequences from infinite triangular arrays of a multinomial experiment, and for sequences from infinite triangular arrays of an exponential family. The first result is a generalization of Corollary 4 in Eagleson (1981), which dealt only with finite triangular arrays of a Gaussian experiment. To prove the general Gaussian dichotomy, we first prove the dichotomy for the case of countable product measures in Corollary 4.1. We then use this result to obtain the general Gaussian contiguity/asymptotic separation dichotomy.

COROLLARY 4.1. *Let* $E = (\mathbb{R}^k, \mathscr{B}(\mathbb{R}^k), \{P_{\mu\Sigma}: \mu \in \mathbb{R}^k, \Sigma \in PD\})$, *where* PD *is the set of all positive definite k by k matrices and* $P_{\mu\Sigma}$ *is multivariate normal with mean* μ *and covariance matrix* Σ. *Let* $P^n = \prod_1^\infty P_{\mu_0\Sigma_0}$ *and* $\tilde{P}^n = \prod_1^\infty P_{\tilde{\mu}_{ni}, \tilde{\Sigma}_{ni}}$. *Then* $P^n \lhd \rhd \tilde{P}^n$ *or* $P^n \vartriangle \tilde{P}^n$ *with the former occurring if and only if*

$$\liminf_{n\to\infty} \inf\{\det(\tilde{\Sigma}_{ni}) : i \in \mathbb{N}\} > 0 \tag{40}$$

and

$$\limsup_{n\to\infty} \sum_{i=1}^{\infty} \{|\tilde{\mu}_{ni} - \mu_0|^2 + |\tilde{\Sigma}_{ni} - \Sigma_0|^2\} < \infty, \tag{41}$$

where the norm on matrices is as elements of $\mathbb{R}^{k^2}$.

PROOF. For the sake of brevity and clarity we only prove this for the case $k = 1$, since conceptually the proof in the multivariate case is the same. Let λ, Lebesgue measure, be the dominating measure. By Examples 3.1 and 3.2 on pages 47 to 49 in Roussas (1972), the mapping $\mu \in \mathbb{R} \to h_{\mu\Sigma}$ is differentiable for all $\Sigma > 0$ and the mapping $\Sigma \in \mathbb{R}_+ \to h_{\mu\Sigma}$ is differentiable for all $\mu \in \mathbb{R}$ with the L^2-derivatives coinciding with the L^2-equivalence classes containing the pointwise partial derivatives

$$\frac{\partial}{\partial\Sigma} h_{\mu\Sigma}(x) = \left(\frac{-1}{4\Sigma} + \frac{(x-\mu)^2}{4\Sigma^2}\right) \frac{\exp\left(\dfrac{-(x-\mu)^2}{4\Sigma}\right)}{(2\pi\Sigma)^{1/4}}$$

and

$$\frac{\partial}{\partial\mu} h_{\mu\Sigma}(x) = \left(\frac{x-\mu}{2\Sigma}\right) \frac{\exp\left(\dfrac{-(x-\mu)^2}{4\Sigma}\right)}{(2\pi\Sigma)^{1/4}}.$$

It is easy to verify that above are continuous as mappings from $\Theta = \mathbb{R} \times \mathbb{R}_+$ to $L^2(\lambda)$. Thus by a standard result in differential calculus for normed linear spaces [e.g., Loomis and Sternberg (1968), Theorem 3.9.3], the experiment is differentiable. The remaining assumptions in the hypothesis of Theorem 3.2 are easily verified and are left to the reader. □

REMARK. Corollary 4.1 is subsumed by an upcoming example in this section. Specifically, we show in the example that under fairly general conditions, exponential families generate regular experiments and hence Theorem 3.3 is applicable. However in the Gaussian example, since one can always translate and rescale, we only need Theorem 3.2 to prove the dichotomy.

We now state Eagleson's result on Gaussian triangular arrays as a corollary to the previous corollary. The proof is easy and is omitted.

COROLLARY 4.2. *Let E be as in Corollary 4.1 and let $P^n = \prod_1^n P_{\mu_0 \Sigma_0}$ and $\tilde{P}^n = \prod_1^n P_{\tilde{\mu}_{ni} \tilde{\Sigma}_{ni}}$. Then $P^n \lhd \rhd \tilde{P}^n$ or $P^n \vartriangle \tilde{P}^n$ with the former occurring if and only if*

$$\liminf_{n\to\infty} \inf\left\{\det\left(\tilde{\Sigma}_{ni}\right): 1 \le i \le n\right\} > 0 \tag{42}$$

and

$$\limsup_{n\to\infty} \sum_1^n \left\{|\tilde{\mu}_{ni} - \mu_0|^2 + |\tilde{\Sigma}_{ni} - \Sigma_0|^2\right\} < \infty. \tag{43}$$

We now want to state and prove the general Gaussian dichotomy. First however we must state a lemma which is needed in the proof.

LEMMA 4.1. *Let $(\Omega_n, \mathscr{F}_n, \{P^n, \tilde{P}^n\})$ be a sequence of experiments and for each n let $\mathscr{F}_n^0$ be a field generating $\mathscr{F}_n$. Then $\tilde{P}^n \lhd P^n$ ($\tilde{P}^n \vartriangle P^n$) if and only if $\tilde{P}^n$ is contiguous (asymptotically separated) to P^n on $\mathscr{F}_n^0$ (take the natural extensions of the definitions of contiguity and asymptotic separation for fields).*

PROOF. By a straightforward corollary to the proof of the approximation property of $\mathscr{F}_n^0$ to $\mathscr{F}_n$ it is possible to approximate simultaneously any set in $\mathscr{F}_n$ by a set in $\mathscr{F}_n^0$ with respect to P^n and $\tilde{P}^n$ [cf. Billingsley (1979), Theorem 11.4]. The result follows easily. □

COROLLARY 4.3. *Let $E = (\mathbb{R}^S, \mathscr{B}(\mathbb{R}^S), \{P^n, \tilde{P}^n\})$ be an experiment, where S is an arbitrary index set and P^n and $\tilde{P}^n$ are Gaussian probability measures. Then $P^n \lhd \rhd \tilde{P}^n$ or $P^n \vartriangle \tilde{P}^n$.*

PROOF. By Lemma 4.1 it suffices to prove the dichotomy for the sequence of experiments $E_n = (\mathbb{R}^n, \mathscr{B}(\mathbb{R}^n), \{P^n, \tilde{P}^n\})$, where P^n and $\tilde{P}^n$ are Gaussian probability measures. If the cardinality of the set $\{n \in \mathbb{N}: P^n \perp \tilde{P}^n\}$ is infinite, then clearly $P^n \vartriangle \tilde{P}^n$. If it is finite we can assume, without loss of generality, that P^n and $\tilde{P}^n$ are nondegenerate. By translating the components and invoking Lemma 3.1 we can assume without loss of generality that P^n has mean 0. Next by diagonalizing the covariance matrix of P^n, and then rescaling the components of P^n, and again invoking Lemma 3.1, we can assume that P^n has an identity covariance matrix. Finally, we diagonalize the covariance matrix of $\tilde{P}^n$ (this does not affect the covariance matrix of P^n since it is the identity) and again invoking

Lemma 3.1, we can assume without loss of generality that $\tilde{P}^n$ is a product ofone-dimensional Gaussian measures. Thus without loss of generality we can assume that P^n is an n-fold product of $\mathscr{N}(0,1)$ and $\tilde{P}^n$ is an n-fold product of $\{\mathscr{N}(\mu_{ni}, \sigma_{ni}^2): 1 \le i \le n\}$. By Corollary 4.2 the result now follows. □

Examples. We now give two more examples where the dichotomy results in Theorems 3.2 and 3.3 apply. The second subsumes the first, but the first example is presented because of its simplicity and transparency relative to the theory in Section 3.

EXAMPLE 1 (Multinomial). Let $\Omega = \{1, \dots, d\}$ and $E = (\Omega, 2^\Omega, \{P_\theta: \theta \in \Theta\})$, where $\Theta = \{\theta \in \mathbb{R}_+^\Omega: \theta(j) > 0 \ \forall\ j, \Sigma_1^d \theta(j) = 1\}$, $f_\theta = \Sigma_1^d \theta(j) 1_{\{j\}}$, and where ν, the dominating measure, is assumed to be the counting measure on $(\Omega, 2^\Omega)$. For $j \in \Omega$, let $e_j \in \mathbb{R}^\Omega$ be defined by

$$e_j(j') = \begin{Bmatrix} 1 & \text{if } j' = j \\ 0 & \text{if } j' \neq j \end{Bmatrix}.$$

Then for $\theta \in \mathbb{R}_+^\Omega$,

$$\begin{aligned} \frac{\sqrt{\theta + \varepsilon e_j} - \sqrt{\theta}}{\varepsilon} &= \left(\frac{\sqrt{\theta(j) + \varepsilon} - \sqrt{\theta(j)}}{\varepsilon} \right) 1_{\{j\}} \\ &\to \frac{1}{2\sqrt{\theta(j)}} 1_{\{j\}} \quad \text{as } \varepsilon \to \infty, \end{aligned} \tag{44}$$

where the last convergence is in $L^2(\nu)$. Thus the mapping $\theta \in \mathbb{R}_+^\Omega \to \sqrt{\theta}$ is partially differentiable and it is easy to see that the partial derivatives are continuous as functions from $\mathbb{R}_+^\Omega$ to $L^2(\nu)$. Thus by a standard differential calculus result for normed linear spaces [e.g., Loomis and Sternberg (1968), Theorem 3.9.3] and since $\theta \in \Theta \to h_\theta$ is just a restriction of the above mapping, the experiment is regular with

$$2\nabla h_\theta = \left[\frac{1}{\sqrt{\theta(1)}} 1_{\{1\}}, \dots, \frac{1}{\sqrt{\theta(d)}} 1_{\{d\}} \right]^t.$$

Let $P^n = \prod_1^\infty P_{\theta_{ni}}$ and $\tilde{P}^n = \prod_1^\infty P_{\tilde{\theta}_{ni}}$, and assume there exist a $\rho > 0$ such that

$$\rho < \theta_{ni}(j) < 1 - \rho \quad \text{for all } n, i \in \mathbb{N} \text{ and } j \in \Omega.$$

Then by Theorem 3.3, $P^n \lhd \rhd \tilde{P}^n$ or $P^n \vartriangle \tilde{P}^n$ with the former being true if and only if

$$\liminf_{n \to \infty} \inf\left\{ |\tilde{\theta}_{ni}(j)|, |1 - \tilde{\theta}_{ni}(j)| : i \in \mathbb{N},\ j \in \Omega \right\} > 0$$

and

$$\limsup_{n \to \infty} \sum_1^\infty \sum_1^d |\theta_{ni}(j) - \tilde{\theta}_{ni}(j)|^2 < \infty.$$

EXAMPLE 2 (Exponential family). Let $E = (\Omega, \mathscr{F}, \{P_\theta: \theta \in \Theta\})$ be an experiment where Θ is an open subset of $\mathbb{R}^d$. Assume there exist a σ-finite measure ν which dominates E and there exists random variables $\{T_j: 1 \le j \le m\}$ such that

$$c(\theta)\exp\left(\sum_1^m \theta_j T_j\right) \in \frac{dP_\theta}{d\nu} \quad \text{for all } \theta \in \Theta.$$

This is the exponential family with the natural parameterization, and we further assume that the random variables $\{T_j\}$ are linearly independent, that is, $\sum_1^m t_j T_j = 0$ a.e.-ν implies $t_j = 0$ for all j.

Then (A.1) holds since $P_\theta \equiv \nu$ for all $\theta \in \Theta$. Also if $P_\theta = P_{\theta'}$, then $\theta = \theta'$ by the linear independence of $\{T_j\}$ and hence (A.4) holds. By Theorem 78.2 in Strasser (1985), E is differentiable with

$$\nabla h_\theta = \tfrac{1}{2}\left[(T_1 - P_\theta T_1)h_\theta, \dots, (T_d - P_\theta T_d)h_\theta\right]^t,$$

where $P_\theta T_j = \int T_j \, dP_\theta$ for all j. By well-known properties of exponential families [cf. Lemma 2.1 and Theorem 2.2 in Brown (1986)] the mapping $\theta \to \nabla h_\theta$ is continuous and hence E is regular.

Let $P^n = \prod_1^\infty P_{\theta_{ni}}$ and $\tilde{P}^n = \prod_1^\infty P_{\tilde{\theta}_{ni}}$ with $\{\theta_{ni}: n, i \in \mathbb{N}\}$ restricted to some compact subset of Θ. If we assume (A.3) is true, then we have a contiguity/asymptotic separation dichotomy by Theorem 3.3. If we do not assume (A.3) is true but instead only assume that $\{\tilde{\theta}_{ni}: n, i \in \mathbb{N}\}$ is also contained in some compact subset of Θ, we again get a contiguity/asymptotic separation dichotomy by Theorem 3.3. This last statement follows by the fact that we could without loss of generality assume Θ is compact by taking it to be the closure of $\{\theta_{ni}, \tilde{\theta}_{ni}: n, i \in \mathbb{N}\}$, and hence (A.3) is satisfied trivially.

5. Necessary and sufficient conditions for an l^2 dichotomy. In this section we prove a converse of the sufficiency result in Section 3 for a specific experiment $E = (\mathbb{R}^k, \mathscr{B}(\mathbb{R}^k), \{P_{tR}: t \in \mathbb{R}^k, R \in \mathscr{R}\})$, $k \ge 2$. This experiment E is based on an underlying probability measure P on $(\mathbb{R}^k, \mathscr{B}(\mathbb{R}^k))$ and rigid motion perturbations of P. Thus $P_{tR} = P \circ RT_t$ where T_t is the translation operator by the vector t and R is in $\mathscr{R}$, the set of all orthogonal transformations on $\mathbb{R}^k$. Note that all rigid motions can uniquely be expressed as RT_t for some $t \in \mathbb{R}^k$ and $R \in \mathscr{R}$. Considering $\mathscr{R}$ as a subset of all $k \times k$ matrices, the parameter space for this experiment E is $\Theta = \mathbb{R}^k \times \mathscr{R} \subset \mathbb{R}^{k(k+1)}$.

For this experiment, Steele (1986) proved that $\prod_1^\infty P \perp \prod_1^\infty P_{\theta_i}$ for all sequences $\{\theta_i\} \notin l^2$ and which converge to the identity. Secondly and more importantly he showed that $\prod_1^\infty P \equiv \prod_1^\infty P_{\theta_i}$ for all $\{\theta_i\} \in l^2$ if and only if $P \equiv \lambda$ and for all one-parameter subgroups $\{p(s): s \in \mathbb{R}\}$ in the space of rigid motions, there exist a number K such that

$$\left|\int h(x)\left[d/ds(\phi(p(s)x))\right]_{s=0} d\lambda(x)\right| \le K\|\phi\|_2 \tag{45}$$

for all $\phi \in C_c^\infty(\mathbb{R}^k)$. Here $h \in \sqrt{dP/d\lambda}$ and $C_c^\infty(\mathbb{R}^k)$ is the set of all infinitely differentiable functions with compact support. Steele defines finite Fisher information by this last condition. It is easy to show that if E is dominated by λ and is differentiable at $(0, I)$ (I is the $k \times k$ identity matrix), then (45) is satisfied.

Hence (45) appears to be a weaker condition than differentiability. However we will prove that the hypothesis of $P \equiv \lambda$ and E being differentiable are necessary for an l^2 dichotomy. In proving the main theorem we will need two results due to Le Cam (1970) (Theorem 1 and Proposition 2) which we state as propositions for ease of reference and a technical lemma, which is stated and proved. In the rest of this section λ will always denote Lebesgue measure on $\mathbb{R}^k$ (or sometimes $\mathbb{R}^d$), where we have suppressed the index k (or sometimes d) for notational convenience.

PROPOSITION 5.1. *Let ψ: $U \subset \mathbb{R}^d \to \mathscr{H}$, where $\mathscr{H}$ is a Hilbert space and U is a Borel subset. Suppose*

$$\limsup_{u' \to u} \frac{\|\psi(u') - \psi(u)\|}{|u' - u|} < \infty \tag{46}$$

for λ-a.e. $u \in U$. Then ψ is strongly differentiable at λ-a.e. $u \in U$.

PROPOSITION 5.2. *Let $E = (\Omega, \mathscr{F}, \{P_\theta: \theta \in \Theta\})$ be an experiment satisfying* (A.1) *and let $\theta_0 \in \Theta$. If $\prod_1^\infty P_{\theta_0} \equiv \prod_1^\infty P_{\theta_i}$ for all $\{\theta_i\}$ such that $\{(\theta_0 - \theta_i)\} \in l^2$, then*

$$\limsup_{\theta \to \theta_0} \frac{H(\theta_0, \theta)}{|\theta_0 - \theta|} < \infty. \tag{47}$$

LEMMA 5.1. *Let $h \in L^2(\mathbb{R}^k)$. If $\mathscr{A}$ is the set of all invertible linear transformations from $\mathbb{R}^k$ to $\mathbb{R}^k$, then*

$$\lim_{(t, A) \to (0, I)} \|h \circ AT_t - h\|_2 = 0, \tag{48}$$

where the limit is over $t \in \mathbb{R}^k$ and $A \in \mathscr{A}$.

PROOF. Let $\varepsilon > 0$ and g be a continuous function from $\mathbb{R}^k$ to $\mathbb{R}$ with compact support such that $\|g - h\|_2 < \varepsilon$. Then

$$\|h \circ AT_t - h\|_2 \le \|h \circ AT_t - g \circ AT_t\|_2 + \|g \circ AT_t - g\|_2 + \|g - h\|_2. \tag{49}$$

But $\|h \circ AT_t - g \circ AT_t\|_2 \to \|g - h\|_2$ as $(t, A) \to (0, I)$, by a change of variables. Also since g has compact support, $\|g \circ AT_t - g\|_2 \to 0$ as $(t, A) \to (0, I)$ by the Lebesgue dominated convergence theorem. Thus by taking the limit supremum in (49),

$$\limsup_{(t, A) \to (0, I)} \|h \circ AT_t - h\|_2 \le 2\varepsilon. \tag{50}$$

Since ε was arbitrary, the result follows. □

We now set out some notation before going to the main results of this section. For P, a probability measure on $(\mathbb{R}^k, \mathscr{B}(\mathbb{R}^k))$, let

$$E_1^P = \left(\mathbb{R}^k, \mathscr{B}(\mathbb{R}^k), \{P_\theta\colon \theta \in \Theta = \mathbb{R}^k\}\right)$$

and

$$E_2^P = \left(\mathbb{R}^k, \mathscr{B}(\mathbb{R}^k), \{P_\theta \in \Theta = \mathbb{R}^k \times \mathscr{R}\}\right),$$

where E_1^P corresponds to the translation experiment and E_2^P corresponds to the rigid motion experiment. We sometimes suppress the superscript P for notational convenience when the underlying probability measure is clear.

Let $\theta \in \Theta$. As given in the previous paragraph, θ represents an element in some Euclidean space. However we will sometimes find it convenient to let θ also represent the transformation, that is, in E_1, $\theta = T_\theta$, etc. It will always be clear from the context whether θ represents a transformation or an element of the parameter space, and hence we will not overly distinguish between the two interpretations of θ.

Main results. The main result of this section is necessary and sufficient conditions for the contiguity/asymptotic separation dichotomy in the rigid motion experiment. We also give an analogous result for necessary and sufficient conditions for a contiguity/asymptotic separation dichotomy in the translation experiment since this result (to the best of our knowledge) has never appeared in the literature and because the proofs for the translation experiment dichotomy help to illuminate the much more difficult proofs used in the rigid motion dichotomy result. First we need some preliminary results related to E_1^P and E_2^P which are interesting in their own right. In Lemma 5.2, we show in both experiments that if (47) is satisfied at one point, then (47) is satisfied at all points. This is equivalent to saying that if $P \equiv \lambda$ and the mapping $\theta \in \Theta \to h_\theta$ is Lipschitz at one point, then it is Lipschitz at all points in Θ. The usefulness of this result is derived from Proposition 5.2, and is given in Theorem 5.1. Specifically, we show that if $P \equiv \lambda$ and the mapping, as given above, is Lipschitz at one point, then the experiment generated by P is differentiable. For the translation experiment this was stated in Le Cam (1970).

DEFINITION. Let $E = (\Omega, \mathscr{F}, \{P_\theta\colon \theta \in \Theta\})$ be an experiment. Then $l(E)$ denotes all the points $\theta \in \Theta$ for which (47) is true, that is, all the points where the Hellinger distance is Lipschitz.

LEMMA 5.2. *Let $P \equiv \lambda$, $E \in \{E_1^P, E_2^P\}$ and suppose that $l(E) \neq \varnothing$. Then $l(E) = \Theta$.*

PROOF. Let $E = E_2$, $\theta_0 \in l(E_2)$ and $\theta_1 \in \Theta = \mathbb{R}^k \times \mathscr{R}$. Now temporarily substitute (t_0, R_0), (t_1, R_1) and (t, R) for θ_0, θ_1 and θ, respectively. Then

$$\begin{aligned}\theta\theta_1^{-1}\theta_0 &= RT_tT_{-t_1}R_1^{-1}R_0T_{t_0}\\ &= RR_1^{-1}R_0T_{(R_0^{-1}R_1(t-t_1)+t_0)}.\end{aligned}$$

Thus

$$\begin{aligned} |\theta\theta_1^{-1}\theta_0 - \theta_0|^2 &= \left|\left(R_0^{-1}R_1(t - t_1), \left(RR_1^{-1} - I\right)R_0\right)\right|^2 \\ &= |t - t_1|^2 + \left|(R - R_1)R_1^{-1}R_0\right|^2 \\ &= |\theta - \theta_1|^2. \end{aligned} \tag{51}$$

By (51),

$$\frac{\|h_\theta - h_{\theta_1}\|}{|\theta - \theta_1|} = \frac{\|h_{\theta\theta_1^{-1}\theta_0} - h_{\theta_0}\|}{|\theta\theta_1^{-1}\theta_0 - \theta_0|} \tag{52}$$

for all $\theta \in \Theta$. Here we have also used the invariance of Lebesgue measure under rigid motions. This implies that $\theta_1 \in l(E)$.

The proof in the case of $E = E_1$ is exactly analogous to the above and is even easier. □

In Theorem 5.1, by using Lemma 5.2 and Proposition 5.1, we prove that a sufficient condition for differentiability of E in either the translation or the rigid motion case is that $l(E) \neq \varnothing$. Again the result for the translation experiment was previously known [cf. the remark following Proposition 2 in Le Cam (1970)]. For the sake of completeness and for purposes of helping to clarify the proof in the rigid motion case we give a proof for the translation case.

Before stating and proving the theorem, we outline some of the key ideas of the proof. The proof of differentiability for the translation experiment (where it is assumed that λ is the dominating measure) basically consists of noting that being Lipschitz at one point implies it is true for all points in the parameter space. By Proposition 5.1, the experiment is differentiable everywhere except for a λ-null subset of Θ. By the translation invariance of λ, this easily implies that the experiment is differentiable everywhere. For the rigid motion experiment there are two main difficulties to this approach. First Θ is already a null set, so Proposition 5.2 does not even imply differentiability at any points. To overcome this difficulty, we locally transform E_2 into a new experiment E_2^* with a new parameter space Θ^* which is not null, and then we invoke Proposition 5.2 for this new experiment. We then transform back to the original experiment E_2 and thus show that E_2 is differentiable on a dense set in Θ. The second difficulty is that, unlike the translation case, differentiability at a point θ_0 is not directly transferable to other points in Θ. However locally and asymptotically it is transferable.

THEOREM 5.1. *Let $(\Omega, \mathscr{F}, P)$ be an experiment dominated by λ, $E \in \{E_i^P\}_{i=1}^2$, and suppose $l(E) \neq \varnothing$. Then:*

(a) *E is regular in the case $i = 1$.*

(b) *E is differentiable in the case $i = 2$ and in this case the family of differentials $\{\nabla h_\theta : \theta \in \Theta\}$ is uniformly bounded in the L^2-norm.*

PROOF. The proof of differentiability for the translation experiment was essentially given in the remarks prior to this theorem. Regularity follows by Lemma 5.1.

Let $E = E_2$. Then by Lemma 5.2, $l(E) = \Theta = \mathbb{R}^k \times \mathscr{R}$. Now let $L = k(k-1)/2$. By the general implicit function theorem [e.g., Auslander and MacKenzie (1963)] for each $R \in \mathscr{R}$ there exist neighborhoods V of R and U of $0 \in \mathbb{R}^L$, and $\psi \in C^1(U, V)$ such that $\psi(0) = R$ and ψ is homeomorphism with the differential of ψ at $u \in U$, $d\psi_u$, having rank L for all $u \in U$. Also there exists a neighborhood W of $R \in \mathbb{R}^{k^2}$ and $\eta \in C^1(W, U)$ such that η is onto, and $\eta \circ \psi$ is the identity mapping on U. Without loss of generality we can assume U is convex, $\overline{U}$ is compact and the above is true on a neighborhood of $\overline{U}$ with the same η and ψ.

Now fix an $R_0 \in \mathscr{R}$ and let η, ψ, W, U and V be as above for this particular R_0. By the compactness of $\overline{U}$ and the continuity of the differential,

$$\sup\{\|d\psi_u - d\psi_{u'}\|: u, u' \in U\} = \beta < \infty.$$

Thus using a standard differential calculus result [e.g., Loomis and Sternberg (1968), Theorem 3.7.4],

$$|\psi(u') - \psi(u)| \le \beta|u' - u| \quad \text{for all } u, u' \in U. \tag{53}$$

Thus $\{|\psi(u') - \psi(u)|/|u' - u|: u, u' \in U\}$ is bounded above. We now want to show it is bounded away from 0. It suffices to prove that

$$\liminf_{j\to\infty} \frac{|\psi(u_j') - \psi(u_j)|}{|u_j' - u_j|} > 0 \tag{54}$$

for arbitrary sequences $\{u_j\}, \{u_j'\}$ in U such that $u_j \ne u_j'$ for all j. Since this is true if and only if every subsequence has a subsubsequence which is bounded away from 0, it suffices to prove (54) under the additional assumptions that $u_j \to u \in \overline{U}$ and $u_j' \to u' \in \overline{U}$. If $u \ne u'$, then (54) is true by the continuity and the injectivity of ψ. If $u = u'$, then

$$\lim_{j\to\infty} \frac{|\psi(u_j') - \psi(u_j) - d\psi_u(u_j' - u_j)|}{|u_j' - u_j|} = 0 \tag{55}$$

by the continuous differentiability of ψ, the convexity of $\overline{U}$ and standard differential calculus. But $d\psi_u$ has rank L so (55) implies that (54) holds in this case also.

Thus there exist positive constants α, β such that

$$\alpha|u' - u| \le |\psi(u') - \psi(u)| \le \beta|u' - u| \tag{56}$$

for all $u, u' \in U$.

Now consider a new experiment $E_2^* = (\mathbb{R}^k, \mathscr{B}(\mathbb{R}^k), \{P_{\theta^*}^*: \theta^* \in \Theta^*\})$, where $\Theta^* = \mathbb{R}^k \times U$, $P_{(t,u)}^* = P_{(t,\psi(u))}$ and $h_{\theta^*}^* \in \sqrt{dP_{\theta^*}^*/d\lambda}$. We define a new map $\psi_1: \mathbb{R}^k \times U \to \Theta$ by $\psi_1(t, u) = (t, \psi(u))$. By (56)

$$|\psi_1(\theta^*) - \psi_1(\theta_0^*)| \le (1 + \beta)|\theta^* - \theta_0^*| \tag{57}$$

for all $\theta_0^*, \theta^* \in \Theta^*$. If Θ_c^* is a compact subset of Θ^*, then there exist $K < \infty$ such that for all $\theta_0^* \in \Theta_c^*$,

$$\limsup_{\theta^* \to \theta_0^*} \frac{\|h_{\theta^*} - h_{\theta_0^*}\|}{|\theta^* - \theta_0^*|} < K(1 + \beta) \tag{58}$$

by dividing and multiplying by $|\psi_1(\theta^*) - \psi_1(\theta_0^*)|$, and by invoking (57) and invoking the proof of Lemma 5.2. Thus $l(E_2^*) = \Theta^*$. By Proposition 5.1, there exist $\Theta_0^* \subset \Theta^*$ such that E_2^* is differentiable on Θ_0^* and $\lambda(\Theta^* \setminus \Theta_0^*) = 0$.

Now define a map η_1: $\mathbb{R}^k \times W \to \mathbb{R}^k \times U$ by $\eta_1(t, w) = (t, \eta(w))$. Note that $\eta_1 \in C^1(\mathbb{R}^k \times W)$ and $\eta_1 \circ \psi_1$ is the identity map on $\mathbb{R}^k \times V$. Thus by the chain rule, E is differentiable on $\Theta_0 = \psi_1(\Theta_0^*)$ with

$$\nabla h_\theta = \nabla \eta_{1\theta} \nabla h^*_{\eta_1(\theta)} \quad \text{for } \theta \in \Theta_0. \tag{59}$$

Let $t_0 \in \mathbb{R}^k$ be fixed. Then there exist a sequence $\{(t_j, u_j)\} \subset \Theta_0^*$ such that $(t_j, u_j) \to (t_0, 0)$. Letting $\theta_j = \psi_1(t_j, u_j)$, we see that $\theta_j \to \theta_0 = (t_0, R_0)$. Now for a fixed j and letting $\theta \in \mathbb{R}^k \times V$, we have

$$\begin{aligned} &\left\| h_\theta - h_{\theta_0} - (\theta - \theta_0)^t \cdot \left(\nabla h_{\theta_j} \circ \theta_j^{-1}\theta_0\right)\right\| \\ &\quad = \left\| h \circ \theta\theta_0^{-1}\theta_j - h \circ \theta_j - (\theta - \theta_0)^t \cdot \nabla h_{\theta_j}\right\| \\ &\quad \le \left\| h \circ \theta\theta_0^{-1}\theta_j - h \circ \theta_j - \left(\theta\theta_0^{-1}\theta_j - \theta_j\right)^t \cdot \nabla h_{\theta_j}\right\| \\ &\qquad + \left\| \left(\theta\theta_0^{-1}\theta_j - \theta_j - \theta + \theta_0\right)^t \cdot \nabla h_{\theta_j}\right\|. \end{aligned} \tag{60}$$

If we let $\theta_j = (t_j, R_j)$ and $\theta = (t, R)$, and if we view $\theta\theta_0^{-1}\theta_j$ as an operator, we have

$$\begin{aligned} \theta\theta_0^{-1}\theta_j &= RT_t T_{-t_0} R_0^{-1} R_j T_{t_j} \\ &= RR_0^{-1} R_j T_{R_j^{-1}R_0(t-t_0)+t_j}. \end{aligned} \tag{61}$$

Corresponding to the first term in (60), note that

$$|\theta\theta_0^{-1}\theta_j - \theta_j| = |\theta - \theta_0| \tag{62}$$

by (51). For the second term in (60), by (61) and the Cauchy–Schwarz inequality,

$$\begin{aligned} &\left\| \left(\theta\theta_0^{-1}\theta_j - \theta_j - \theta + \theta_0\right)^t \cdot \nabla h_{\theta_j}\right\|^2 \\ &\quad \le |\theta\theta_0^{-1}\theta_j - \theta_j - \theta + \theta_0|^2 \left\| \left(\nabla h_{\theta_j}^t \cdot \nabla h_{\theta_j}\right)^{1/2}\right\|^2 \\ &\quad = \left|\left(\left(R_j^{-1}R_0 - I\right)(t - t_0), RR_0^{-1}R_j - R_j - R + R_0\right)\right|^2 \left\| \left(\nabla h_{\theta_j}^t \cdot \nabla h_{\theta_j}\right)^{1/2}\right\|^2 \\ &\quad = \left|\left((R_0 - R_j)(t - t_0), \left(I - RR_0^{-1}\right)(R_0 - R_j)\right)\right|^2 \left\| \left(\nabla h_{\theta_j}^t \cdot \nabla h_{\theta_j}\right)^{1/2}\right\|^2 \\ &\quad \le |R_j - R_0|^2 |(t - t_0, R - R_0)|^2 \left\| \left(\nabla h_{\theta_j}^t \cdot \nabla h_{\theta_j}\right)^{1/2}\right\|^2, \end{aligned} \tag{63}$$

where in the last inequality, we have used that the norm of $R - R_0$ as a linear operator is less than or equal to $|R - R_0|$, and we also used that the inverse of an element of $\mathscr{R}$ is equal to its transpose. On dividing the quantity in (60) by $|\theta - \theta_0|$, and letting $\theta \to \theta_0$, we see that the first term goes to 0 by the differentiability of E_2 at θ_j and by (62). Thus by (61) and (63), if $g^j = \nabla h_{\theta_j} \circ \theta_j^{-1}\theta_0$,

$$(64)\quad \limsup_{\theta \to \theta_0} \frac{\left\| h_\theta - h_{\theta_0} - (\theta - \theta_0)^t \cdot g^j \right\|}{|\theta - \theta_0|} \le |R_j - R_0| \left\| \left(\nabla h_{\theta_j}^t \cdot \nabla h_{\theta_j} \right)^{1/2} \right\|.$$

Since $\nabla h_{\theta_j} = \nabla \eta_{1\theta_j} \nabla h^*_{\eta_1(\theta_j)}$,

$$(65)\quad \left\| \left(\nabla h_{\theta_j}^t \cdot \nabla h_{\theta_j} \right)^{1/2} \right\| = \left\| \left(\nabla h^{*t}_{\eta_1(\theta_j)} \nabla \eta_{1\theta_j}^t \nabla \eta_{1\theta_j} \nabla h_{\eta_1(\theta_j)} \right)^{1/2} \right\|.$$

But $\nabla \eta_{1\theta_j} \to \nabla \eta_{1\theta_0}$ by the continuity of the differential and $\{\|(\nabla h^{*t}_{\eta_1(\theta_j)} \cdot \nabla h^*_{\eta_1(\theta_j)})^{1/2}\|\}$ is uniformly bounded by (58). Thus the quantity in (65) is uniformly bounded in j, and hence by (64),

$$(66)\quad \lim_{j \to \infty} \limsup_{\theta \to \theta_0} \frac{\left\| h_\theta - h_{\theta_0} - (\theta - \theta_0)^t \cdot g^j \right\|}{|\theta - \theta_0|} = 0.$$

Let $M = \mathbb{R}^k \times (d\psi_0(\mathbb{R}^L)) \subset \mathbb{R}^{k+k^2}$. Then M is an $(L + k)$-dimensional subspace since $d\psi_0$ has rank L. Now let $\theta_0^* = (t_0, 0)$. Then by the differentiability of ψ_1,

$$(67)\quad \lim_{\theta^* \to \theta_0^*} \left| \psi_1(\theta^*) - \psi_1(\theta_0^*) - d\psi_{1\theta_0^*}(\theta^* - \theta_0^*) \right| = 0.$$

By dividing and multiplying the expression in (67) by $|\psi_1(\theta^*) - \psi_1(\theta_0^*)|$, noting that ψ_1 is a homeomorphism and invoking (56),

$$(68)\quad \lim_{\theta \to \theta_0} \frac{|\theta - \theta_0 - \mathrm{Pr}_M(\theta - \theta_0)|}{|\theta - \theta_0|} = 0,$$

where Pr_M is the projection operator onto the subspace M. By (66) and (68), $\{\mathrm{Pr}_M g^j\}$ is a Cauchy sequence in $\prod_1^{k(k+1)} L^2(\lambda)$. Let g be the limit of $\{\mathrm{Pr}_M g^j\}$ in $\prod_1^{k(k+1)} L^2(\lambda)$. By (68)

$$(69)\quad \limsup_{\theta \to \theta_0} \frac{\left\| h_\theta - h_{\theta_0} - (\theta - \theta_0)^t \cdot g \right\|}{|\theta - \theta_0|} \le 2\varepsilon$$

after triangulating the numerator on $(\theta - \theta_0)^t \cdot g^j$ and $\mathrm{Pr}_M(\theta - \theta_0) \cdot g^j$. Thus E_2 is differentiable at θ_0. Since θ_0 was arbitrary, E_2 is differentiable. Also by projecting the differentials onto the appropriate subspaces (as was done previously with the subspace M), invoking the proof of Lemma 5.2 and (68), we also obtain a family of differentials $\{\nabla h_\theta \colon \theta \in \Theta\}$ which are uniformly bounded. □

REMARK. By Theorem 5.1 we now have a useful sufficient condition for checking when E_i^P is differentiable for $i = 1, 2$. Namely, we only need to check whether $l(E_i^P) \neq \varnothing$. Most often this will be done by checking whether the identity is contained in $l(E_i^P)$. We apply this in the next result, Proposition 5.3, a result we will need in proving our main theorem.

PROPOSITION 5.3. *Let* $i \in \{1,2\}$ *and* $(\mathbb{R}^k, \mathscr{B}(\mathbb{R}^k), \{P_0, P_1\})$ *be an experiment such that* $P_0 \equiv \lambda$ *and* $E_i^{P_0}$ *is differentiable. Then* $E_i^{P_0 * P_1}$ *is differentiable.*

PROOF. Let $P = P_0 * P_1$ and note that by Fubini, $P \equiv \lambda$. Let Θ be the appropriate parameter space depending on whether $i = 1, 2$ and let $h_\theta \in (dP_\theta/d\lambda)^{1/2}$ for $\theta \in \Theta$, and $f_0 \in dP_0/d\lambda$. Then

$$h_\theta(y) = \left(\int f_0(\theta y - x)\, dP_1(x)\right)^{1/2} \tag{70}$$

for all $y \in \mathbb{R}^k \setminus B$ where $\lambda(B) = 0$.

Let $\theta \in \Theta$ be fixed. For $\theta' \in \Theta \setminus \{\theta\}$,

$$(h_{\theta'} - h_\theta)^2 = h_{\theta'}^2 - 2h_{\theta'}h_\theta + h_\theta^2 \tag{71}$$

by algebra. But by (70)

$$\begin{aligned} h_{\theta'}(y)h_\theta(y) &= \left[\left(\int f_0(\theta' y - x)\, dP_1(x)\right)\left(\int f_0(\theta y - x)\, dP_1(x)\right)\right]^{1/2} \\ &\geq \int h_0(\theta' y - x) h_0(\theta y - x)\, dP_1(x) \end{aligned}$$

for λ-a.e. y, by Holder's inequality where $h_0 = (f_0)^{1/2}$. Combining this with (71),

$$(h_{\theta'}(y) - h_\theta(y))^2 \leq \int (h_0(\theta' y - x) - h_0(\theta y - x))^2\, dP_1(x) \tag{72}$$

for λ-a.e. y. Thus letting $\theta' \to \theta$ we obtain that $\theta \in l(E_i^P)$ by (72) and Fubini. By Theorem 5.1, E_i^P is differentiable. □

We are now ready to state and prove the main theorem of this section, which gives a partial converse to Theorem 3.2 in the case of translation and rigid motion experiments. Some of the ideas in the proof are in Steele (1986). Before we do this we state a simple lemma, without proof, which we will need.

LEMMA 5.3. *Let* $E = (\mathbb{R}^k, \mathscr{B}(\mathbb{R}^k), \{P_{ni}, \tilde{P}_{ni}, Q\colon n, i \in \mathbb{N}\})$ *be an experiment. If* $\prod_1^\infty P_{ni} * Q \vartriangle \prod_1^\infty \tilde{P}_{ni} * Q$, *then* $\prod_1^\infty P_{ni} \vartriangle \prod_1^\infty \tilde{P}_{ni}$.

THEOREM 5.2. *Let* $(\mathbb{R}^k, \mathscr{B}(\mathbb{R}^k), P)$ *be an experiment and let* $E \in \{E_1^P, E_2^P\}$ *and in the case* $E = E_2^P$, E *satisfies* (A.4). *Then the following are true*:

(a) $P^n = \prod_1^\infty P_{\theta_{ni}} \lhd \rhd \tilde{P}^n = \prod_1^\infty P_{\tilde{\theta}_{ni}}$ *for all* $\{\theta_{ni}\}, \{\tilde{\theta}_{ni}\}$ *such that*

$$\limsup_{n\to\infty} \sum_1^\infty |\theta_{ni} - \tilde{\theta}_{ni}|^2 < \infty \tag{73}$$

if and only if $P \equiv \lambda$ *and* E *is differentiable.*

(b) $P^n = \prod_1^\infty P_{\theta_{ni}} \vartriangle \tilde{P}^n = \prod_1^\infty P_{\tilde{\theta}_{ni}}$ *for all* $\{\theta_{ni}\}, \{\tilde{\theta}_{ni}\}$ *such that*

$$\limsup_{n\to\infty} \sum_1^\infty |\theta_{ni} - \tilde{\theta}_{ni}|^2 = \infty. \tag{74}$$

PROOF. By Lemma 3.1, we can without loss of generality assume $\{\theta_{ni}\} \subset \{0\} \times \mathscr{R}$ in the case $E = E_2$.

(a) $\prod_1^\infty P \equiv P_{\tilde{\theta}} \times \prod_2^\infty P$ for all $\tilde{\theta} \in \Theta$ and hence (A.1) is true. This implies that $P \equiv \lambda$ [cf. Steele (1986), Lemma 4.1] and henceforth let λ be the dominating measure. By Proposition 5.2, $l(E) \neq \varnothing$ and thus by Theorem 5.1, E is differentiable.

For the converse clearly (A.1) and (A.3) are true, and by Lemma 5.1, (A.2) holds. Also (A.4) holds in the case of $E = E_1$ by a straightforward argument and for $E = E_2$ the assumption holds by the hypothesis. Thus by Theorem 3.1, it suffices to prove

$$\lim_{\rho\to 0} \inf\left\{\frac{\|h_{\theta'} - h_\theta\|}{|\theta' - \theta|} : \theta \in \Theta_c, |\theta' - \theta| < \rho\right\} > 0, \tag{75}$$

where $\Theta_c = \{0\}$ for $E = E_1$ and $\Theta_c = \{0\} \times \mathscr{R}$ in the case $E = E_2$.

We first prove (75) for the case of $E = E_1$. Note that $\|h_{\theta'} - h_\theta\| = \|h_{\theta'-\theta} - h_0\|$ by the translation invariance of Lebesgue measure. Thus it suffices to show

$$\liminf_{\theta\to 0} \frac{\|h_\theta - h_0\|}{|\theta|} > 0. \tag{76}$$

Suppose (76) is false. Then by the regularity of E and standard differential calculus, there exist $\theta_0 \neq 0$ such that

$$\lim_{s\to 0} \frac{\|h_{s\theta_0} - h_0\|}{|s|} = 0, \tag{77}$$

where the limit is through $s \in \mathbb{R}$. Thus

$$\begin{aligned} \|h_{\theta_0} - h_0\| &\leq \sum_1^m \left\| h_{(i\theta_0)/m} - h_{((i-1)\theta_0)/m} \right\| \\ &= \sum_1^m \left\| h_{(\theta_0/m)} - h_0 \right\| \\ &= \frac{\left\| h_{(\theta_0/m)} - h_0 \right\|}{1/m}, \end{aligned} \tag{78}$$

where the first equality is by the translation invariance of Lebesgue measure. On letting $m \to \infty$ in (78) and invoking (77), $\|h_{\theta_0} - h_0\| = 0$. This implies that $P_{\theta_0} = P_0$, a contradiction to (A.4).

To prove (75) in the case $E = E_2$, we first embed $\mathbb{R}^k$ into $\mathbb{R}^{k+1}$ by the mapping $(x_1, \ldots, x_k) \in \mathbb{R}^k \to (x_1, \ldots, x_k, 1) \in \mathbb{R}^{k+1}$. On this embedded space in $\mathbb{R}^{k+1}$, all rigid motions can be represented by a set of linear transformations forming a matrix Lie group [cf. Auslander (1967), Theorem 1.6.6]. In matrix form the element in the matrix Lie group representing RT_t is denoted by $G(R, t)$ and is given by

(79) $$G(R, t) = \begin{bmatrix} R & Rt \\ 0 & 1 \end{bmatrix}.$$

Also letting I_0 be the identity on $\mathbb{R}^{k+1}$, note that

(80) $$|G(R, t) - I_0| = |(t, R) - (0, I)|.$$

This matrix Lie group has for each $G(R, t)$, a tangent space $T(t, R) \subset \mathbb{R}^{(k+1)^2}$ which is a $(k + L)$-dimensional subspace with $L = k(k-1)/2$.

By the theory of matrix Lie groups there exist a ball of positive radius r about 0, $B_r(0)$, in the tangent space and a neighborhood of $(0, I)$ in Θ such that the map $S \in B_r(0) \to \exp(S)$ is a continuously differentiable homeomorphism [cf. Warner (1983), Theorem 3.3.1 and Definition 3.8]. The exponential function, $\exp(S)$, is defined as

$$\exp(S) = \sum_0^\infty \frac{S^k}{k!},$$

where S is in the tangent space $T(0, I)$. By (52), in order to prove (75), it suffices to prove

(81) $$\liminf_{\theta \to (0, I)} \frac{\|h_\theta - h_{(0, I)}\|}{|\theta - (0, I)|} > 0.$$

So suppose that (81) is false and there exist a sequence $\{\theta_j\}$ such that $\theta_j \to (0, I)$ and

(82) $$\lim_{j \to \infty} \frac{\|h_{\theta_j} - h_{(0, I)}\|}{|\theta_j - (0, I)|} = 0.$$

For large j, $\theta_j = \exp(\alpha_j S_j)$, where $S_j \in T(0, I)$, $|S_j| = 1$ and $\alpha_j \in (0, r)$. By choosing a convergent subsequence, we can assume without loss of generality that $S_j \to S_0 \in T(0, I)$. Then

(83) $$\frac{|(0, I) - \exp(\alpha_j S_j)|}{|\alpha_j|} \le \exp(|\alpha_j|),$$

since $|S_j| = 1$ for all j. Thus

(84) $$\limsup_{j \to \infty} \frac{|(0, I) - \theta_j|}{|\alpha_j|} \le 1$$

by (83), since $|\alpha_j| \to 0$. Also

$$\lim_{j\to\infty} \frac{\|h\circ\exp(\alpha_j S_j) - h\circ\exp(\alpha_j S_0)\|}{|\alpha_j|} = 0 \tag{85}$$

by continuous differentiability of the map $S \in B_r(0) \to h\circ\exp(S)$ (which we get by the chain rule) and standard differential calculus. Thus since

$$\frac{\|h\circ\exp(\alpha_j S_0) - h\|}{|\alpha_j|} \le \frac{\|h\circ\exp(\alpha_j S_0) - h\circ\exp(\alpha_j S_j)\| + \|h\circ\exp(\alpha_j S_j) - h\|}{|\alpha_j|},$$

we have

$$\lim_{j\to\infty} \frac{\|h\circ\exp(\alpha_j S_o) - h\|}{|\alpha_j|} = 0 \tag{86}$$

by (82), (83) and (84). But the map $\alpha \in (-r, r) \to h\circ\exp(\alpha S_0)$ is differentiable by the chain rule. Hence by (86)

$$\lim_{\alpha\to 0} \frac{\|h\circ\exp(\alpha S_0) - h\|}{|\alpha|} = 0. \tag{87}$$

Let $\alpha_0 \in (0, r)$. Then

$$\begin{aligned} \|h\circ\exp(\alpha_0 S_0) - h\| &\le \sum_1^m \|h\circ\exp(i\alpha_0 S_0/m) - h\circ\exp((i-1)\alpha_0 S_0/m)\| \\ &= \frac{\|h\circ\exp(\alpha_0 S_0/m) - h\|}{1/m} \end{aligned} \tag{88}$$

for all n, where the inequality is by Minkowski's inequality, and the equality is from the invariance of λ under rigid motions. Thus

$$\|h\circ\exp(\alpha_0 S_0) - h\| = 0$$

by (87), and letting $m \to \infty$ in (88). Hence if $\theta_0 = \exp(\alpha_0 S_0)$, $P = P_{\theta_0}$, a contradiction to the identifiability assumption in (A.4). Thus (75) is true in the case of $E = E_2$ and we have proven the converse portion of part (a).

(b) First we convolute P with $\mathscr{N}(0, I)$. Then note by characteristic functions and invariance of $\mathscr{N}(0, I)$ under rotations, (A.4) is satisfied for the experiment $E^{P*\mathscr{N}(0, I)}$. Also by Proposition 5.3, this experiment is differentiable. By Lemma 5.3, it suffices to prove the result under the additional assumptions that $P \equiv \lambda$ and E is differentiable. By the proof of (a), (75) holds. Hence by (13) the result follows. □

REMARK. It is natural to consider generalizations of the above to the case of invertible affine transformations. In this framework we again start with a probability measure P on $\mathbb{R}^k$, $k \ge 1$. Let $\mathscr{A}$ be the set of all invertible linear transformations from $\mathbb{R}^k$ to $\mathbb{R}^k$. The experiment of all invertible affine perturbations of P is $E = (\Omega, \mathscr{F}, \{P_{tA}: t \in \mathbb{R}^k, A \in \mathscr{A}\})$, where

$$P_{tA} = P \circ AT_t.$$

The result in Lemma 5.2 holds exactly as stated for this experiment also. A result similar to Theorem 5.1(b) holds for E if the statement is modified to state that for all $\theta_0 \in \Theta$ the family of differentials $\{\nabla h_\theta: \theta \in \Theta\}$ is uniformly bounded in some open neighborhood of θ_0. This result can be used to prove a result exactly analogous to Theorem 5.2(a) for this invertible affine transformation experiment. In fact the proof is easier in that we do not need to transform the experiment since $\mathbb{R}^k \times \mathscr{A}$ is not a null set in $\mathbb{R}^{k+k^2}$. We have been unable to prove an analogous result to Theorem 5.2(b), mainly because it is not clear whether the trick of convolving P with $\mathscr{N}(0, I)$ preserves identifiability as was the case in the rigid motion experiment.

Acknowledgments. The paper is essentially based on my Ph.D. thesis. I wish to thank Professor V. Mandrekar for posing the problem and his many valuable suggestions. I also would like to thank Professor James Hannan for his careful reading of my thesis and his helpful suggestions. Finally, I would like to thank an Associate Editor for constructive suggestions which substantially improved the clarity of presentation in this article.

REFERENCES

AUSLANDER, L. (1967). *Differential Geometry*. Harper and Row, New York.

AUSLANDER, L. and MACKENZIE, R. (1963). *Introduction to Differentiable Manifolds*. McGraw-Hill, New York.

BAUER, H. (1981). *Probability Theory and Elements of Measure Theory*. Academic, New York.

BILLINGSLEY, P. (1979). *Probability and Measure*. Wiley, New York.

BROWN, L. (1986). *Fundamentals of Statistical Exponential Families*. IMS, Hayward, Calif.

EAGLESON, G. K. (1981). An extended dichotomy theorem for sequences of pairs of Gaussian measures. *Ann. Probab.* **9** 453–459.

FABIAN, V. and HANNAN, J. (1985). *Introduction to Probability and Mathematical Statistics*. Wiley, New York.

FELDMAN, J. (1958). Equivalence and perpendicularity of Gaussian processes. *Pacific J. Math.* **9** 699–708.

HÁJEK, J. (1958). On a property of normal distributions of any stochastic processes. *Czechoslovak Math. J.* **8** 610–618. (In Russian; translation *Select. Transl. Math. Statist. Probab.* **1** (1961) 245–252.)

KAKUTANI, S. (1948). On equivalence of infinite product measures. *Ann. of Math.* (2) **49** 214–224.

LE CAM, L. (1970). On the assumptions used to prove asymptotic normality of maximum likelihood estimates. *Ann. Math. Statist.* **41** 802–828.

LIPTSER, R. SH., PUKELSHEIM, F. and SHIRYAEV, A. N. (1982). Necessary and sufficient conditions for contiguity and entire asymptotic separation of probability measures. *Russian Math. Surveys* **37** 107–136.

LOOMIS, L. and STERNBERG, S. (1986). *Advanced Calculus*. Addison-Wesley, Reading, Mass.

NEVEU, J. (1985). *Mathematical Foundations of the Calculus of Probability*. Holden-Day, San Francisco.

OOSTERHOFF, J. and VAN ZWET, W. R. (1979). A note on contiguity and Hellinger distance. In *Contributions to Statistics. Jaroslav Hájek Memorial Volume* (J. Jurečková, ed.) 157–166. Reidel, Dordrecht.

ROUSSAS, G. (1972). *Contiguity of Probability Measures: Some Applications in Statistics.* Cambridge Univ. Press, Cambridge.

SHEPP, L. (1965). Distinguishing a sequence of random variables from a translate of itself. *Ann. Math. Statist.* **36** 1107–1112.

STEELE, J. (1986). Fisher information and detection of a Euclidean perturbation of an independent stationary process. *Ann. Probab.* **14** 326–335.

STRASSER, H. (1985). *Mathematical Theory of Statistics.* de Gruyter, New York.

THELEN, B. (1986). Fisher information and dichotomies in contiguity/asymptotic separation. Ph.D. dissertation, Michigan State Univ.

WARNER, F. W. (1983). *Foundations of Differentiable Manifolds and Lie Groups.* Springer, New York.

DEPARTMENT OF STATISTICS
1447A MASON HALL
UNIVERSITY OF MICHIGAN
ANN ARBOR, MICHIGAN 48109