This is to certify that the

dissertation entitled

Fisher Information and Dichotomies in
Contiguity/Asymptotic Separation

presented by

Brian J. Thelen

has been accepted towards fulfillment
of the requirements for

Ph.D. degree in Statistics

Major professor

Date November 13, 1986

FISHER INFORMATION AND DICHOTOMIES IN CONTIGUITY/ASYMPTOTIC SEPARATION

by

Brian J. Thelen

A DISSERTATION

Submitted to
Michigan State University
in partial fulfillment of the requirements
for the degree of

DOCTOR OF PHILOSOPHY

Department of Statistics and Probability
1986

ABSTRACT

FISHER INFORMATION AND DICHOTOMIES IN CONTIGUITY/ASYMPTOTIC SEPARATION

By

Brian J. Thelen

A contiguity/asymptotic separation dichotomy for sequences of product measures is proved under the assumption that the component measures belong to a differentiable experiment. 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 Fisher information and equivalence/singularity dichotomies between two product measures, one with a fixed component measure and the second with rigidly perturbed component measures. We then use the preceding two results for statistical applications. First we prove asymptotic normality of the likelihood ratio under a differentiability assumption and secondly under the differentiability assumption we give a necessary condition for consistency in nonlinear least squares estimation, thereby generalizing a result in Wu (1981).

To my wife Mary Ann and my children

Mark, Rachel, and Bethany

ACKNOWLEDGEMENTS

I would like to thank the Office of Naval Research for their support of me during the summer of 1985 and IBM for their support of me during the academic year of 1985-86.

I would like to express my appreciation to my thesis advisor Professor V. Mandrekar for his support and encouragement and to Professors R. Erickson and G. Lappan for serving on my guidance committee. Professionally I owe a great debt to Professor James Hannan who taught me much about statistical science and even more about being a statistical scientist.

Also thanks are due to Cathy Sparks for a superb job of typing this thesis under less than ideal conditions.

Finally I express my deep appreciation to my family; to my parents for teaching me the value of hard work and persistence in the face of obstacles, and to my wife whose total support of me in this pursuit has been constant even in the face of many trials.

TABLE OF CONTENTS

CHAPTER 0

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 $(\Omega, \mathcal{F}, \mathcal{P})$ is an experiment, (i.e. $(\Omega, \mathcal{F})$ is a measurable space and $\mathcal{P}$ is a set of probability measures on $(\Omega, \mathcal{F})$) we say that a equivalence/singularity dichotomy holds if $P, \tilde{P} \in \mathcal{P}$ implies $P \equiv \tilde{P}$ or $P \perp \tilde{P}$.

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

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

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

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

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 $\Omega = \mathbb{R}^\infty$, $\mathcal{F} = \mathfrak{B}(\mathbb{R}^\infty)$ (i.e. the Borel σ-field), and $\mathcal{P}$ is the set of all Gaussian probability measures. This was proved by Feldman (1958) and Hajek (1958). In the special case where $P = \prod_1^\infty \mathcal{N}(0,1)$ and $\tilde{P} = \prod_1^\infty \mathcal{N}(\mu_i,\sigma_i^2)$, they showed that $P \equiv \tilde{P}$ if and only if

(3) $$\sum_1^\infty (\mu_i^2+(1-\sigma_i)^2) < \infty.$$

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 $\{\mu_i\}_1^\infty \in \ell^2$. Note that $\mathcal{N}(\mu_i,1)$ is just a translate of $\mathcal{N}(0,1)$ by μ_i and an interesting question is; what other probability measures on $(\mathbb{R},\mathfrak{B}(\mathbb{R}))$ besides $\mathcal{N}(0,1)$ satisfy this property? This was answered by Shepp (1965) who showed that if P is a probability measure on $(\mathbb{R},\mathfrak{B}(\mathbb{R}))$ and P_t is the translate of P by t, then $\prod_1^\infty P \equiv \prod_1^\infty P_{t_i}$ for all $\{t_i\} \in \ell^2$ if and only if $P \equiv \lambda$ (λ is Lebesgue measure) and P has finite Fisher information i.e. there exists $f \in dP/d\lambda$ such that f is locally absolutely continuous and

$$(4) \qquad \int ((f')^2/f)\, d\lambda < \infty.$$

He also showed $\prod_1^\infty P \perp \prod_1^\infty P_{t_i}$ for all $\{t_i\} \notin \ell^2$. Thus a necessary and sufficient condition for an ℓ^2 type of equivalence/singularity 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 $(\mathbb{R}^k, \mathcal{B}(\mathbb{R}^k))$ in LeCam (1970), Proposition 2. Specifically, with P a probability measure on $(\mathbb{R}^k, \mathcal{B}(\mathbb{R}^k))$, it was shown that $\prod_1^\infty P \equiv \prod_1^\infty P_{t_i}$ for all $\{t_i\} \in \ell^2$ if and only if $P \equiv \lambda$ and the map

$$(5) \qquad t \in \mathbb{R}^k \to (f(\cdot+t))^{1/2} \in L^2(\lambda)$$

is differentiable 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 (4) and (5) are equivalent by comparing this with Shepp's result.

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

All of the above results are related to equivalence/singularity dichotomies between 2 measures. However in the asymptotic theory of statistics there are useful generalizations of equivalence and singularity which are applicable to 2 sequences of measures. These are contiguity and asymptotic separation. Eagleson (1981) proved that a contiguity/asymptotic separation dichotomy holds between any 2 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 thesis are generalizations and clarifications of the previous results of Eagleson (1981) and Steele (1986). The first result is to give sufficient conditions for a contiguity/asymptotic separation dichotomy in the case of 2 sequences of product measures where the component probability measures are from a dominated experiment $E = (\Omega, \mathcal{F}, \{P_\theta : \theta \in \Theta\})$ with Θ in some finite dimensional Euclidean space. The sufficient condition is 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. Also in the case of a differentiable experiment with some additional natural assumptions, we give necessary and sufficient conditions for the 2 sequences of product probability measures to be contiguous. These conditions are essentially that the ℓ^2 distance between the two sequences of component parameters is asymptotically bounded.

The second result 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 on $(\mathbb{R}^k, \mathcal{B}(\mathbb{R}^k))$ and all rigid motion perturbations of P. We parameterize this experiment by (t, R), where t is the translation vector, and R 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 the thesis 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 asymptotically bounded, then the experiment of P perturbed by rigid motions is L^2-differentiable. This clarifies Steele's result in that it shows that his unusual 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.

We also extend some of these results to the case of an experiment generated by P and all invertable affine perturbations of P. In this case we are able to prove a partial converse. Namely we show under some natural conditions, that contiguity between two sequences of product measures for all sequences of component parameters with asymptotically bounded ℓ^2 distance implies $P \equiv \lambda$ and an L^2-differentiability condition similar to that given in (5).

We then apply these results and some ideas in their proofs to a couple of statistical applications. The first is to give a simple proof of the asymptotic normality of the likelihood ratio of two triangular arrays under the assumption of differentiability. The second application is to verify necessary conditions for consistency in multivariate nonlinear least squares estimation under the differentiability assumption, thereby generalizing a result of Wu (1981).

CHAPTER I

NOTATION AND PRELIMINARIES

We first give the basic definitions of some concepts discussed in Chapter 0. An <u>experiment</u> E is a measurable space $(\Omega,\mathcal{F})$ along with a class of probability measures $\mathcal{P}$ and we write $E = (\Omega,\mathcal{F},\mathcal{P})$. Suppose $(\Omega,\mathcal{F},\{P,\tilde{P}\})$ is an experiment. Then P and $\tilde{P}$ are <u>equivalent</u> $(\tilde{P} \equiv P)$ if $\tilde{P}$ is absolutely continuous to P $(\tilde{P} \ll P)$ and P is absolutely continuous to $\tilde{P}$ $(P \ll \tilde{P})$. P and $\tilde{P}$ are <u>singular</u> $(P \perp \tilde{P})$ if there exists $F \in \mathcal{F}$ such that $P(F) = 1$ and $\tilde{P}(F) = 0$.

Let $(\Omega,\mathcal{F},\nu)$ be a σ-finite measure space and $\{f_\alpha: \alpha \in I\} \subset L^1(\Omega,\mathcal{F},\nu)$. Then $\{f_\alpha: \alpha \in I\}$ is <u>uniformly integrable</u> (u.i.) if for all $\epsilon > 0$ there exists $h \in L^1(\Omega,\mathcal{F},\nu)$ such that

$$\int (|f_\alpha|-h)_+ \, d\nu < \epsilon \qquad \text{for all } \alpha \in I,$$

where $(x)_+ = \max\{x,0\}$ for all $x \in \mathbb{R}$. For more details regarding uniform integrability in this general framework, see Fabian/Hannan (1985), Section 4.8, or Bauer (1981), Section 2.12. A simple but useful result, which is proved in the appendix as Proposition A.1, is that since ν is σ-finite, $\{f_\alpha: \alpha \in I\} \subset L^1(\nu)$ is u.i. if and only if for every sequence $\{\alpha_n\} \subset I$, there exists a subsequence $\{\alpha_{n'}\}$ such that $\{f_{\alpha_{n'}}\}$ is u.i..

Let $(\Omega_n, \mathcal{F}_n, \{P^n, \tilde{P}^n\})$ be a sequence of experiments. The sequence $\{\tilde{P}^n\}$ is <u>contiguous</u> to the sequence $\{P^n\}$ $(\tilde{P}^n \triangleleft P^n)$ if for each sequence $\{F_n\}$ where $F_n \in \mathcal{F}_n$ and $P^n(F_n) \to 0$ as $n \to \infty$, implies $\tilde{P}^n(F_n) \to 0$ as $n \to \infty$. The sequences $\{P^n\}$ and $\{\tilde{P}^n\}$ are <u>mutually contiguous</u> $(\tilde{P}^n \triangleleft\triangleright P^n)$ if $\{P^n\}$ is contiguous to $\{\tilde{P}^n\}$ and vice versa. The sequence $\{P^n\}$ is <u>asymptotically separated</u> from $\{\tilde{P}^n\}$ $(\tilde{P}^n \blacktriangle P^n)$ if there exists a subsequence $\{n'\}$ and a corresponding subsequence of sets $\{F_{n'}\}$ such that $F_{n'} \in \mathcal{F}_{n'}$, $P^{n'}(F_{n'}) \to 1$, and $\tilde{P}^{n'}(F_{n'}) \to 0$. It is easy to see in the special case of $(\Omega_n, \mathcal{F}_n) = (\Omega, \mathcal{F})$, $P^n = P$, and $\tilde{P}^n = \tilde{P}$ for all n, that $\tilde{P}^n \triangleleft\triangleright P^n$ $(\tilde{P}^n \blacktriangle P^n)$ if and only if $\tilde{P} \equiv P$ $(\tilde{P} \perp P)$, and in this sense contiguity (asymptotic separation) is a generalization of equivalence (singularity).

Let $(\Omega, \mathcal{F}, \{P_\theta : \theta \in \Theta\})$ be an experiment with $\Theta \subset \mathbb{R}^d$. The experiment is <u>dominated</u> if there exists a σ-finite measure ν on $(\Omega, \mathcal{F})$ 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 $h_\theta = f_\theta^{1/2} \in L^2(\Omega, \mathcal{F}, \nu)$ for all $\theta \in \Theta$. The experiment is <u>differentiable at $\theta = \theta_o$</u> if the mapping $\theta \in \Theta \to h_\theta$ is differentiable at $\theta = \theta_o$ as a mapping from Θ to the Hilbert space $L^2(\Omega, \mathcal{F}, \nu)$ i.e. there exists $\nabla h_{\theta_o} \in \prod_1^d L^2(\Omega, \mathcal{F}, \nu)$ such that

$$\lim_{\theta \to \theta_o} \| h_\theta - h_{\theta_o} - (\theta-\theta_o)^T \cdot \nabla h_{\theta_o} \| / |\theta-\theta_o| = 0 \tag{1.1}$$

where the limit is through $\theta \in \Theta$, $(\theta-\theta_o)^T$ is the transpose of $(\theta-\theta_o)$, 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 <u>differentiable</u> if it is differentiable at all points in Θ and E is <u>regular</u> if it is continuously differentiable. It is easy to see that differentiability is not dependent on the dominating measure. Throughout this thesis all parameter spaces are assumed to be in finite dimensional Euclidean space, and in any Euclidean space we use the usual norm which we denote by $|\cdot|$.

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

$$2H^2(P,\tilde{P}) = \int (h-\tilde{h})^2 \, d\nu \tag{1.2}$$

where $\nu = P+\tilde{P}$, $h \in (dP/d\nu)^{1/2}$, and $\tilde{h} \in (d\tilde{P}/d\nu)^{1/2}$. 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 $0 \le H(P,\tilde{P}) \le 1$ with $H(P,\tilde{P}) = 0$ ($H(P,\tilde{P}) = 1$) if and only if $P = \tilde{P}$ ($P \perp \tilde{P}$).

For the rest of this thesis 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.

$$(1.3) \qquad 2H^2(P,\tilde{P}) \leq \|P-\tilde{P}\| \leq 2H(P,\tilde{P})(2-H^2(P,\tilde{P}))^{1/2}$$

where $\|\cdot\|$ is the total variation norm. Since $P^n \blacktriangle \tilde{P}^n$ if and only if

$$(1.4) \qquad \limsup_{n\to\infty} \|P^n-\tilde{P}^n\| = 2,$$

(1.3) implies

$$(1.5) \qquad P^n \blacktriangle \tilde{P}^n \text{ if and only if } \limsup_{n\to\infty} H(P^n,\tilde{P}^n) = 1.$$

For the infinite product situation, we can easily monitor the Hellinger distance by monitoring the Hellinger distances between the components. If $P^n = \prod_1^\infty P_{ni}$ and $\tilde{P}^n = \prod_1^\infty \tilde{P}_{ni}$ then

$$H^2(P^n,\tilde{P}^n) = 1 - \prod_1^\infty \int h_{ni}\tilde{h}_{ni}\, d\upsilon_{ni}$$

(1.6)

$$= 1 - \prod_1^\infty (1-H^2(P_{ni},\tilde{P}_{ni})),$$

where $\upsilon_{ni} = P_{ni}+\tilde{P}_{ni}$. By (1.5) and (1.6), $P^n \blacktriangle \tilde{P}^n$ if and only if either

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

(1.7) or

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

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

Let $(\Omega,\mathcal{F},\{P_\theta : \theta \in \Theta\})$ be an experiment. We now list some assumptions which will be invoked later in this thesis.

(A1) $P_\theta \equiv P_{\theta'}$ for all θ, $\theta' \in \Theta$ (homogeneity).

(A2) $\lim_{\theta'\to\theta} H(P_\theta,P_{\theta'}) = 0$ for all $\theta \in \Theta$ (continuity). When there is no possible confusion we write $H(\theta,\theta')$ in place of $H(P_\theta,P_{\theta'})$.

(A3) $\lim_{n\to\infty} H(\theta,\theta_n) = 1$ for each $\theta \in \Theta$ and sequence $\{\theta_n\} \subset \Theta$ such that $|\theta_n| \to \infty$ or $\theta_n \to t \in \bar{\Theta}\backslash\Theta$. Here $\bar{\Theta}$ denotes the closure of Θ (asymptotic separation at the boundary).

(A4) $P_\theta \neq P_{\theta'}$ for all $\theta \neq \theta'$ (identifiability).

Since $\Theta \subset \mathbb{R}^d$, (A1), (A2), and (A3) imply $\bar{\Theta}\backslash\Theta$ is closed. This is stated and proved in the following proposition.

<u>Proposition 1.1</u>. Let $E = (\Omega,\mathcal{F},\{P_\theta: \theta \in \Theta\})$ be an experiment which satisfies (A1), (A2), and (A3). Then $\bar{\Theta}\backslash\Theta$ is closed.

Proof: If $\bar{\Theta}\backslash\Theta = \phi$, we are done. So suppose it is not null and let $\{t_j\} \subset \bar{\Theta}\backslash\Theta$ be such that $t_j \to t \in \mathbb{R}^d$. Let $\theta_o \in \Theta$. By (A3) there exists $\{\theta_j\} \subset \Theta$ such that $|t_j-\theta_j| \to 0$ and $H(\theta_j,\theta_o) \to 1$. Thus if $t \in \Theta$, $H(\theta_j,t) \to 0$ by (A2). This would then imply $H(\theta_o,t) = 1$, a contradiction to (A1), since $H(\theta_o,t) \geq H(\theta_j,\theta_o) - H(\theta_j,t)$ for all j. Hence $t \notin \Theta$ and this implies $t \in \bar{\Theta}\backslash\Theta$. □

Finally whenever there is an infinite product measure it is implicitly assumed the component probability measures are compact, i.e. there exists a compact subclass (cf. Neveu (1965), pg. 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 measure).

CHAPTER II

SUFFICIENT CONDITIONS FOR DICHOTOMY

2.1 Main results.

A corollary, of a more general result in Lipster, Pukel'sheim, and Shiryaev (1982), which gives necessary and sufficient conditions for contiguity in the infinite product situation is stated and proved below (this is actually a generalization of a result in Oosterhoff/Van Zwet (1979)).

Proposition 2.1. Let $E_{ni} = (\Omega_{ni}, \mathcal{F}_{ni}, \{P_{ni}, \tilde{P}_{ni}\})$ 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

$$(2.1) \qquad \limsup_{n\to\infty} \sum_1^\infty H^2(P_{ni}, \tilde{P}_{ni}) < \infty$$

and

$$(2.2) \qquad \lim_{K\to\infty} \limsup_{n\to\infty} \sum_1^\infty \tilde{P}_{ni}(\tilde{f}_{ni} > Kf_{ni}) = 0$$

and

$$(2.3) \qquad \lim_{K\to\infty} \limsup_{n\to\infty} \sum_{1}^{\infty} P_{ni}(f_{ni} > K\tilde{f}_{ni}) = 0.$$

Proof: By the remark following corollary 1 in Lipster, Pukel'sheim, and Shiryaev (1982), $\tilde{P}_n \triangleleft P_n$ if and only if (2.1) holds,

$$(2.4) \qquad \lim_{K\to\infty} \limsup_{n\to\infty} \tilde{P}_n(\sup_i(\tilde{f}_{ni}/f_{ni}) > K) = 0,$$

and

$$(2.5) \qquad \lim_{K\to\infty} \limsup_{n\to\infty} P_n(\sup_i(f_{ni}/\tilde{f}_{ni}) > K) = 0.$$

But (2.4) is equivalent to

$$(2.6) \qquad \lim_{K\to\infty} \limsup_{n\to\infty} [1-\prod_{1}^{\infty}(1-\tilde{P}_n(\tilde{f}_{ni} > Kf_{ni}))] = 0.$$

By taking the exponential of both sides, (2.6) is seen to be equivalent to (2.2). By symmetry, (2.3) is equivalent to (2.5). □.

Before stating and proving our main results in this chapter, we need a technical lemma which shows that contiguity and asymptotic separation are not affected by measurable transformations which are one-to-one, onto, and have a measurable inverse.

Lemma 2.1. Let $E_n = (\Omega_n, \mathcal{F}_n, \{P^n, \tilde{P}^n\})$ be a sequence of experiments and let I_n be a measurable transformation from $(\Omega_n, \mathcal{F}_n)$ 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 $\tilde{P}^n \triangleleft P^n$ ($\tilde{P}^n \blacktriangle P^n$) implies $\tilde{Q}^n \triangleleft Q^n$ ($\tilde{Q}^n \blacktriangle Q^n$).

Proof: Let $\{F_n\}$ be such that $Q^n(F_n) \to 0$. This is equivalent to $P^n(I_n^{-1}(F_n)) \to 0$ which implies $\tilde{P}^n(I_n^{-1}(F_n)) \to 0$ since $\tilde{P}^n \triangleleft P^n$. But this is equivalent to $\tilde{Q}_n(F_n) \to 0$ and hence $\tilde{Q}_n \triangleleft Q_n$. The proof of the asymptotic separation result is similar. □.

Remark. Let $E_{ni} = (\Omega, \mathcal{F}, \{P_{ni}, \tilde{P}_{ni}\})$ be an experiment for all $n, i \in \mathbb{N}$, $P^n = \prod_1^\infty P_{ni}$ and $\tilde{P}^n = \prod_1^\infty \tilde{P}_{ni}$. Let Q^n and $\tilde{Q}^n$ be obtained by a common juxtaposition of the component probability measures of P^n and $\tilde{P}^n$ respectively (juxtaposition can be different for each n). By Lemma 2.1 $\tilde{P}^n \triangleleft P^n$ ($\tilde{P}^n \blacktriangle P^n$) if and only if $\tilde{Q}^n \triangleleft Q^n$ ($\tilde{Q}^n \blacktriangle Q^n$). This fact will be useful for simplifying some arguments and calculations.

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

Theorem 2.1. Let $E = (\Omega, \mathcal{F}, \{P_\theta: \theta \in \Theta\})$ be an experiment satisfying (A1) through (A4), $P^n = \prod_1^\infty P_{\theta_{ni}}$, and $\tilde{P}^n = \prod_1^\infty P_{\tilde{\theta}_{ni}}$ where $\{\theta_{ni}: n, i \in \mathbb{N}\} \subset \Theta_c$, Θ_c a compact subset of Θ. Also assume that if ν is a dominating σ-finite measure and Θ_c' is any compact subset of Θ, then

$$(2.7) \qquad \{(h_\theta - h_{\theta'})^2/|\theta-\theta'|^2: \theta \in \Theta_c, \theta' \in \Theta_c'\} \text{ is u.i.}$$

and

$$(2.8) \qquad \lim_{\rho\to 0} \inf\{H(\theta,\theta')/|\theta-\theta'|: \theta \in \Theta_c, |\theta-\theta'| < \rho\} > 0.$$

Then $P^n \triangleleft\triangleright \tilde{P}^n$ or $P^n \blacktriangle \tilde{P}^n$ with the former occuring if and only if

$$(2.9) \qquad \liminf_{n\to\infty} \inf\{\text{dist}(\tilde{\theta}_{ni}, \bar{\Theta}\backslash\Theta): i \in \mathbb{N}\} > 0$$

and

$$(2.10) \qquad \limsup_{n\to\infty} \sum_1^\infty |\theta_{ni} - \tilde{\theta}_{ni}|^2 = M < \infty$$

where $\bar{\Theta}$ is the closure of Θ and $\text{dist}(\tilde{\theta}_{ni}, \phi) = 1$ by convention (ϕ is the empty set).

Proof: First suppose

$$(2.11)\quad \limsup_{n\to\infty} \sup\{|\tilde{\theta}_{ni}|: i \in \mathbb{N}\} = \infty.$$

By the remark following Lemma 2.1, without loss of generality there exist subsequences $\{\tilde{\theta}_{n'1}\}$ and $\{\theta_{n'1}\}$ such that $|\tilde{\theta}_{n'1}| \to \infty$ and $\theta_{n'1} \to \theta_o \in \Theta$. Then $H(\theta_o, \tilde{\theta}_{n'1}) \to 1$ by (A3) and $H(\theta_o, \theta_{n'1}) \to 0$ by (A2). But by the triangle inequality and algebra,

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

for all n' and on taking the limit infimum of both sides, $H(\theta_{n'1}, \tilde{\theta}_{n'1}) \to 1$. Thus by (1.7), $P^n \blacktriangle \tilde{P}^n$.

Now suppose (2.9) is false. By the above argument if (2.11) if true, then $P^n \blacktriangle \tilde{P}^n$, so assume that (2.11) is also false. By a previous remark, without loss of generality there exists subsequences $\{\tilde{\theta}_{n'1}\}$ and $\{\theta_{n'1}\}$ and a subset $\{t_{n'}\} \subset \bar{\Theta}\backslash\Theta$, such that $\tilde{\theta}_{n'1} \to t \in \mathbb{R}^d$, $\theta_{n'1} \to \theta_o \in \Theta$, and $|\tilde{\theta}_{n'1} - t_{n'}| \to 0$. By the triangle inequality $t_{n'} \to t$ and hence $t \notin \Theta$ by Proposition 1.1. Thus $P^n \blacktriangle \tilde{P}^n$ by (A2), (A3), (1.7), and an argument similar to the one used in the previous paragraph.

Now suppose (2.10) is false, and we want to show $P^n \blacktriangle \tilde{P}^n$. By previous arguments it suffices to prove $P^n \blacktriangle \tilde{P}^n$ under the additional assumptions that (2.11) is false and (2.9) is true. By these additional assumptions (note we are not using that (2.8) is false in this statement) there exists an $N \in \mathbb{N}$ and a compact set $\Theta_c' \subset \Theta$ such that $\{\tilde{\theta}_{ni}: n \geq N, i \in \mathbb{N}\} \subset \Theta_c'$. Let $\rho_o > 0$ and $\alpha_o > 0$ be such that

$$(2.12) \qquad H(\theta,\theta')/|\theta-\theta'| > \alpha_o$$

for $\theta \in \Theta_c$, $\theta' \in \Theta_c'$, and $|\theta-\theta'| < \rho_o$. Also $\{H(\theta,\theta')/|\theta-\theta'|: \theta \in \Theta_c, \theta' \in \Theta_c', |\theta-\theta'| \geq \rho_o\}$ is bounded away from 0. To show this, assume it is not bounded away from 0 and choose sequences $\{\theta_j\} \subset \Theta_c$ and $\{\theta_j'\} \subset \Theta_c$ such that $\theta_j \to \theta_o \in \Theta$, $\theta_j' \to \theta_o' \in \Theta$, $|\theta_j-\theta_j'| \geq \rho_o$ for all j, and $H(\theta_j,\theta_j') \to 0$. By (A2), $H(\theta_o,\theta_o') = 0$, which contradicts the identifiability assumption (A4). Hence combining this with (2.12), $\{H(\theta,\theta')/|\theta-\theta'|: \theta \in \Theta_c,\theta' \in \Theta_c'\}$ is bounded away from 0, and by (2.7), it is bounded above. Thus there exists $\alpha,\beta \in (0,\infty)$ such that

$$(2.13) \qquad \alpha|\theta-\theta'|^2 \leq H^2(\theta,\theta') \leq \beta|\theta-\theta'|^2$$

for $\theta \in \Theta_c$, $\theta' \in \Theta_c'$. Since (2.10) is false, (2.13) and (1.7) imply $P^n \vartriangle \tilde{P}^n$.

For the final case suppose (2.9) and (2.10) are true with M as the constant given in (2.10). By the remark following Lemma 2.1, without loss of generality

$$(2.14)\quad |\theta_{ni}-\tilde{\theta}_{ni}| \geq |\theta_{n,i+1}-\tilde{\theta}_{n,i+1}| \quad \text{for all } i,n \in \mathbb{N}.$$

As in the previous paragraph, since (2.10) implies (2.11) is false, there exists a compact set $\Theta_c' \subset \Theta$ and an $N_o \in \mathbb{N}$ such that $\{\tilde{\theta}_{ni}: n \geq N_o,\ i \in \mathbb{N}\} \subset \Theta_c'$. Since (2.13) is only dependent on the hypothesis given in the statement of theorem, it holds in this case as well. This combined with (2.10) implies (2.1) is true.

We now want to verify that (2.2) and (2.3) hold. Let $\epsilon, \epsilon_o > 0$. Then let $\{\theta_j\} \subset \Theta_c$ and $\{\theta_j'\} \subset \Theta_c'$ be sequences such that $\theta_j \neq \theta_j'$ for all j and $|\theta_j-\theta_j'| \to 0$. Then there exists subsequences $\{\theta_{j'}\}$ and $\{\theta_{j'}'\}$ such that $\theta_{j'} \to \theta_o$, $\theta_{j'}' \to \theta_o$, $h_{\theta_{j'}} \to h_{\theta_o}$ a.e.-ν, and $h_{\theta_{j'}'} \to h_{\theta_o}$ a.e.-ν. This implies

$$\lim_{j'\to\infty} P_{\theta_{j'}}\{(h_{\theta_{j'}}-h_{\theta_{j'}'})^2 \geq \epsilon_o h_{\theta_{j'}}^2\}/H^2(\theta_{j'},\theta_{j'}')$$

$$= \lim_{j'\to\infty} \int 1_{\{(h_{\theta_{j'}}-h_{\theta_{j'}'})^2 \geq \epsilon_o h_{\theta_{j'}}^2\}} \{(h_{\theta_{j'}}-h_{\theta_{j'}'})^2/$$

$$|\theta_{j'}-\theta_{j'}'|^2\}\, d\nu \cdot (\epsilon_o H^2(\theta_{j'},\theta_{j'}')/|\theta_{j'}-\theta_{j'}'|^2)^{-1}$$

$$= 0$$

by (2.13), (2.7), and (A1), since (A1) implies that the integrand is converging to 0 a.e.-ν (cf. Fabian/Hannan

(1985), Lemma 4.8.3). By an exactly analogous argument

$$\lim_{j'\to\infty} P_{\theta'_{j'}}\{(h_{\theta_{j'}}-h_{\theta'_{j'}})^2 \geq \epsilon_o h^2_{\theta'_{j'}}\}/H^2(\theta_{j'},\theta'_{j'}) = 0.$$

Since the original sequences were arbitrary we have actually shown

$$\lim_{\rho\to 0} \sup\{P_\theta\{(h_\theta-h_{\theta'})^2 \geq \epsilon_o h^2_\theta\}/H^2(\theta,\theta'):\ |\theta-\theta'| \leq \rho\} = 0$$

and

$$\lim_{\rho\to 0} \sup\{P_{\theta'}\{(h_\theta-h_{\theta'})^2 \geq \epsilon_o h^2_{\theta'}\}/H^2(\theta,\theta'):\ |\theta-\theta'| \leq \rho\} = 0$$

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

(2.15) $$P_\theta\{(h_\theta-h_{\theta'})^2 \geq \epsilon_o h^2_\theta\}/H^2(\theta,\theta') < \epsilon$$

and

(2.16) $$P_{\theta'}\{(h_\theta-h_{\theta'})^2 \geq \epsilon_o h^2_{\theta'}\}/H^2(\theta,\theta') < \epsilon$$

for $\theta \in \Theta_c$ and $\theta' \in \Theta'_c$ such that $|\theta-\theta'| < \rho$. To show $P^n \lhd\rhd \tilde{P}^n$, it suffices to prove that any two subsequences $\{P^{n'}\}$ and $\{\tilde{P}^{n'}\}$ have mutually contiguous subsubsequences. Hence it suffices to prove $P^{n'} \lhd\rhd \tilde{P}^{n'}$ for two subsequences for which $\theta_{n'i} \to \theta_i \in \Theta$, $\tilde{\theta}_{n'i} \to \tilde{\theta}_i \in \Theta$, $f_{\theta_{n'i}} \to f_{\theta_i}$ a.e.-ν and $f_{\tilde{\theta}_{n'i}} \to f_{\tilde{\theta}_i}$ a.e.-ν for all $i \in \mathbb{N}$, since (A2) is true and since Θ_c and Θ_c' are both compact. Let $C = \max\{M/\rho^2, 1\}$. By (A1) there exists $K > (1+\epsilon_o)^2$ such that

$$(2.17) \qquad \lim_{n'\to\infty} P_{\theta_{n'i}}(f_{\theta_{n'i}} > Kf_{\tilde{\theta}_{n'i}}) = P_{\theta_i}(f_{\theta_i} > Kf_{\tilde{\theta}_i}) < \epsilon/C$$

and

$$(2.18) \qquad \lim_{n'\to\infty} P_{\tilde{\theta}_{n'i}}(f_{\tilde{\theta}_{n'i}} > Kf_{\theta_{n'i}}) = P_{\tilde{\theta}_i}(f_{\tilde{\theta}_i} > Kf_{\theta_i}) < \epsilon/C$$

for $i \le C$. There exists $N \in \mathbb{N}$ such that $n \ge N$ and $i > C$ implies $|\theta_{ni} - \tilde{\theta}_{ni}| < \rho$ by (2.10) and (2.14). By combining (2.17) and (2.18) with (2.13) and (2.16),

$$(2.19) \qquad \limsup_{n'\to\infty} \sum_1^\infty P_{\theta_{n'i}}(f_{\theta_{n'i}} > Kf_{\tilde{\theta}_{n'i}}) < \epsilon(1+M\beta)$$

and

$$(2.20) \qquad \limsup_{n'\to\infty} \sum_1^\infty P_{\tilde{\theta}_{n'i}}(f_{\tilde{\theta}_{n'i}} > Kf_{\theta_{n'i}}) < \epsilon(1+M\beta).$$

Since ϵ was arbitrary, this implies $P^{n'} \triangleleft\triangleright \tilde{P}^{n'}$ by Proposition 2.1. $\square$.

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

$$\int \nabla h^1_{\theta_o} \cdot (\nabla h^1_{\theta_o})^T \, d\nu_1 = \int \nabla h^2_{\theta_o} \cdot (\nabla h^2_{\theta_o})^T \, d\nu_2.$$

In particular $\int \nabla h^1_{\theta_o} \cdot (\nabla h^1_{\theta_o})^T \, d\nu_1$ is non-singular if and only if $\int \nabla h^2_{\theta_o} \cdot (\nabla h^2_{\theta_o})^T \, d\nu_2$ is non-singular i.e. non-singularity is independent of the dominating measure.

Theorem 2.2. Let $E = (\Omega, \mathcal{F}, \{P_\theta: \theta \in \Theta\})$ be an experiment differentiable at $\theta = \theta_o$ and assume E satisfies (A1) through (A4). Let ν be a dominating measure, ∇h_{θ_o} be the differential corresponding to ν, and assume that $\int \nabla h_{\theta_o} \cdot \nabla h^T_{\theta_o} \, d\nu$ is non-singular. If $P^n = \prod_1^\infty P_{\theta_o}$ and $\tilde{P}^n = \prod_1^\infty P_{\tilde{\theta}_{ni}}$, then $P^n \blacktriangleleft\!\blacktriangleright \tilde{P}^n$ or $P^n \blacktriangle \tilde{P}^n$ with the former being true if and only if

$$\text{(2.21)} \qquad \liminf_{n\to\infty} \inf\{\text{dist}(\tilde{\theta}_{ni}, \bar{\Theta}\backslash\Theta) : i \in \mathbb{N}\} > 0$$

and

(2.22) $$\limsup_{n\to\infty} \sum_{1}^{\infty} |\theta_o - \tilde{\theta}_{ni}|^2 < \infty.$$

Proof: Let $\Theta_c = \{\theta_o\}$ and note that

$$\lim_{\rho\to 0} \inf\{H(\theta_o,\theta')/|\theta_o-\theta'|: |\theta_o-\theta'| < \rho\}$$

$$= \lim_{\rho\to 0} \inf\{\|(\theta_o-\theta')^T\cdot\nabla h_{\theta_o}\|/|\theta_o-\theta'|: |\theta_o-\theta'| < \rho\}$$

$$\geq \inf\{t^T\cdot(\int \nabla h_{\theta_o}\cdot\nabla h_{\theta_o}^T\, d\nu)\cdot t: t \in \mathbb{R}^d, |t| = 1\}$$

$$> 0.$$

Thus (2.8) holds.

Let Θ_c' be a compact subset of Θ. In order to verify (2.9) it suffices, by Proposition A.1, to prove that if $\{\theta_j\}$ is a convergent sequence in $\Theta_c'\setminus\{\theta_o\}$ satisfying that $(\theta_j-\theta_o)/|\theta_j-\theta_o|$ converges to $t \in \mathbb{R}^d$, then $\{(h_{\theta_j}-h_{\theta_o})^2/|\theta_j-\theta_o|^2: j \in \mathbb{N}\}$ is u.i.. To show this we consider two cases. First suppose $\theta_j \to \theta_o$. Then $(h_{\theta_j}-h_{\theta_o})/|\theta_j-\theta_o| \to t^T\cdot\nabla h_{\theta_o}$ in $L^2(\nu)$ and hence $(h_{\theta_j}-h_{\theta_o})^2/|\theta_j-\theta_o|^2 \to (t^T\cdot\nabla h_{\theta_o})^2$ in $L^1(\nu)$ which implies u.i.. For the second case suppose $\theta_j \to \theta \neq \theta_o$. Then by (A2), $(h_{\theta_j}-h_{\theta_o})/|\theta_j-\theta_o| \to (h_\theta-h_{\theta_o})/|\theta-\theta_o|$ in $L^2(\nu)$ and hence $(h_{\theta_j}-h_{\theta_o})^2/|\theta_j-\theta_o|^2 \to (h_\theta-h_{\theta_o})^2/|\theta-\theta_o|^2$ in $L^1(\nu)$ which again implies u.i.. Hence we have verified (2.9), so by Theorem 2.1, the result follows.

□.

<u>Theorem 2.3.</u> Let $E = (\Omega,\mathcal{F},\{P_\theta\colon \theta \in \Theta\})$ be a regular experiment which satisfies (A1) through (A4). Let ν be a dominating σ-finite measure and assume that $\int \nabla h_\theta \cdot \nabla h_\theta^T \, d\nu$ is non-singular 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 $P^n \triangleleft\triangleright \tilde{P}^n$ or $P^n \perp \tilde{P}^n$ with the former being true if and only if (2.9) and (2.10) are true.

Proof: Let Θ_c' be a compact subset of Θ. To show (2.7) and (2.8) are true, and hence obtain the desired result, it suffices to prove that if $\{\theta_j\} \subset \Theta_c$ and $\{\theta_j'\} \subset \Theta_c'$, there exists 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 $\{\|h_{\theta_j}-h_{\theta_j'}\|/|\theta_j-\theta_j'|\}$ 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 $\theta_j \to \theta_o \in \Theta$, $\theta_j' \to \theta_o' \in \Theta$, and $(\theta_j-\theta_j')/|\theta_j-\theta_j'| \to t \in \mathbb{R}^2$.

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

$$(2.23) \qquad (h_{\theta_j} - h_{\theta_j'})/|\theta_j - \theta_j'| \rightarrow t^T \cdot \nabla h_{\theta_o} \quad \text{in} \quad L^2(\nu)$$

which implies $\{(h_{\theta_j} - h_{\theta_j'})^2/|\theta_j - \theta_j'|^2 : j \in \mathbb{N}\}$ is u.i.. Also (2.23) implies

$$(2.24) \qquad \liminf_{j\rightarrow\infty} \; H(\theta_j, \theta_j')/|\theta_j - \theta_j'| > 0$$

since $\int \nabla h_{\theta_o} \cdot \nabla h_{\theta_o}^T \, d\nu$ is non-singular.

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

$$(2.25) \qquad (h_{\theta_j} - h_{\theta_j'})/|\theta_j - \theta_j| \rightarrow (h_{\theta_o} - h_{\theta_o'})/|\theta_o - \theta_o'|$$

in $L^2(\nu)$ by (A2). Hence we again have $\{(h_{\theta_j} - h_{\theta_j'})^2/|\theta_j - \theta_j'|^2 : j \in \mathbb{N}\}$ is u.i.. Also in this case (2.24) again holds by (2.25) and (A4). The result now follows. □.

2.2 Application to Gaussian Sequences. In this section we use our results to prove the contiguity/asymptotic separation dichotomy for sequences of Gaussian processes with arbitrary index sets. This is a generalization of Corollary 4 in Eagleson (1981), which dealt only with triangular arrays. First we prove the dichotomy for the case of countable product measures in Corollary 2.1. Secondly in Corollary 2.2 we prove this generalizes the triangular array result. Finally in Corollary 2.3 we use these two results to obtain the general Gaussian contiguity/asymptotic separation dichotomy.

<u>Corollary 2.1.</u> Let $E = (\mathbb{R}^k, \mathcal{B}(\mathbb{R}^k), \{P_{\mu C}: \mu \in \mathbb{R}^k, C \in PD\})$ where PD is the set of all positive definite k by k matrices and $P_{\mu C}$ is multivariate normal with mean μ and covariance matrix C. Let $P^n = \prod_1^\infty P_{\mu_o C_o}$ and $\tilde{P}^n = \prod_1^\infty P_{\tilde{\mu}_{ni} \tilde{C}_{ni}}$. Then $P^n \blacktriangleleft\!\blacktriangleright \tilde{P}^n$ or $P^n \blacktriangle \tilde{P}^n$ with the former being true if and only if

$$(2.26) \qquad \liminf_{n\to\infty} \inf\{\det(\tilde{C}_{ni}): i \in \mathbb{N}\} > 0$$

and

$$(2.27) \qquad \limsup_{n\to\infty} \sum_1^\infty \{|\tilde{\mu}_{ni}-\mu_o|^2 + |\tilde{C}_{ni}-C_o|^2\} < \infty.$$

where $\det(\tilde{C}_{ni})$ is the determinant of $\tilde{C}_{ni}$ and 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 47-49 in Roussas (1972), the mapping $\mu \in \mathbb{R} \to h_{\mu C}$ is differentiable for all $C > 0$ and the mapping $C > 0 \to h_{\mu C}$ 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

$$\partial/\partial C \; h_{\mu C}(x) = ((-1/4C) + ((x-\mu)^2/2C^2)) \\ (2\pi C)^{-(1/4)} \exp(-(x-\mu)^2/4C)$$

and

$$\partial/\partial \mu \; h_{\mu C}(x) = ((x-\mu)/C)(2\pi C)^{-(1/4)} \exp(-(x-\mu)^2/4C).$$

It is easy to verify that the 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/Sternberg (1968), Theorem 3.9.3), the experiment is differentiable. The remaining assumptions in the hypothesis of Theorem 2.2 are easily verified and are left to the reader. □.

Remark. Corollary 2.1 is subsumed by an upcoming example in Section 2.3. Specifically we show in the example that under fairly general conditions, exponential families generate regular experiments and hence Theorem 2.3 is applicable. However in the Gaussian example, since one can always translate and rescale, we only needed Theorem 2.2 to prove the dichotomy.

Corollary 2.2. Let E be as in Corollary 2.1 and let $P^n = \prod_1^n P_{\mu_o C_o}$ and $\tilde{P}^n = \prod_1^n P_{\tilde{\mu}_{ni}\tilde{C}_{ni}}$. Then $P^n \triangleleft\triangleright \tilde{P}^n$ or $P^n \perp \tilde{P}^n$ with the former occuring if and only if

$$(2.28) \qquad \liminf_{n\to\infty} \inf\{\det(\tilde{C}_{ni}): 1 \le i \le n\} > 0$$

and

$$(2.29) \qquad \limsup_{n\to\infty} \sum_1^n \{|\tilde{\mu}_{ni}-\mu_o|^2+|\tilde{C}_{ni}-C_o|^2\} < \infty.$$

Proof: Let $Q^n = P^n \times \prod_{n+1}^{\infty} P$ and $\tilde{Q}^n = \tilde{P}^n \times \prod_{n+1}^{\infty} P$ where P is multivariate normal with mean 0 and identity covariance matrix. By Proposition A.2 in the appendix, $Q^n \triangleleft\triangleright \tilde{Q}^n$ ($Q^n \perp \tilde{Q}^n$) if and only if $P^n \triangleleft\triangleright \tilde{P}^n$ ($P^n \perp \tilde{P}^n$). Thus by Corollary 2.1, the result follows. □.

In the appendix we prove a proposition (Proposition A.3) which essentially says contiguity and asymptotic separability can be monitored on fields which generate the σ-fields. Using this result and Corollary 2.2 we have the following corollary which gives the dichotomy for Gaussian processes with arbitrary index sets.

Corollary 2.3. Let $E = (\mathbb{R}^S, \mathfrak{B}(\mathbb{R}^S), \{P^n, \tilde{P}^n\})$ be an experiment where S is an arbitrary index set and P^n and $\tilde{P}^n$ are Gaussian probability measures. Then $P^n \blacktriangleleft\!\blacktriangleright \tilde{P}^n$ or $P^n \blacktriangle \tilde{P}^n$.

Proof: By Propositions A.2 and A.3, it suffices to prove the dichotomy for the sequence of experiments $E_n = (\mathbb{R}^n, \mathfrak{B}(\mathbb{R}^n), \{P^n, \tilde{P}^n\})$ 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 $P^n \blacktriangle \tilde{P}^n$. If it is finite we can assume without loss of generality that P^n and $\tilde{P}^n$ are non-degenerate. By translating the components and invoking Lemma 2.1 we can assume without loss of generality that P^n has mean 0. Next by diagonalizing the covariance matrix of P^n, rescaling the components of P^n, and invoking Lemma 2.1, we can assume without loss of generality that P^n has an identity covariance matrix. Finally we diagonalize the covariance matrix of $\tilde{P}^n$ (this doesn't affect the covariance matrix of P^n since it is the identity) and again invoking Lemma 2.1, we can assume without loss of generality that $\tilde{P}^n$ is a product of one-dimensional normal distributions. Thus we can assume that P^n is an n-fold product of $\mathcal{N}(0,1)$ and $\tilde{P}^n$ is an n-fold product of $\{\mathcal{N}(\mu_{ni},\sigma_{ni}^2): 1 \le i \le n\}$. By Corollary 2.2, the result now follows. □.

<u>2.3 Examples</u>. We now give two examples where the dichotomy results in Theorems 2.2 and 2.3 apply. The second subsumes the first, but the first example is presented because of its simplicity and transparency relative to the previous theory.

(1) <u>Multinomial.</u> Let $\Omega = \{1,\ldots,d\}$ and, $E = (\Omega, 2^\Omega, \{P_\theta: \theta \in \Theta\})$ where $\Theta = \{\theta \in \mathbb{R}_+^\Omega: \theta(j) > 0$ for all $j, \sum_1^d \theta(j) = 1\}$, and $dP_\theta/d\nu = \sum_1^d \theta(j) 1_{\{j\}}$ where ν is the counting measure on $(\Omega, 2^\Omega)$. For $j \in \Omega$, let $e_j \in \mathbb{R}^\Omega$ be defined by

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

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

$$((\theta+\epsilon e_j)^{1/2}-\theta^{1/2}) = \{((\theta(j)+\epsilon)^{1/2}-(\theta(j))^{1/2})/\epsilon\}1_{\{j\}}$$
$$\rightarrow (1/2)(\theta(j))^{-1/2}1_{\{j\}} \qquad \text{as} \quad \epsilon \rightarrow 0,$$

where the last convergence is in $L^2(\nu)$. Thus the mapping $\theta \in \mathbb{R}_+^\Omega \to \theta^{1/2}$ 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 for normed linear spaces (e.g. Loomis/Sternbert (1968), Theorem 3.9.3), and since $\theta \in \Theta \to h_\theta$ is just a restriction of the above mapping, the mapping $\theta \in \Theta \to h_\theta$ is continuously differentiable with

$$2\nabla h_\theta = [(\theta(1))^{-1/2} 1_{\{1\}}, \ldots, (\theta(d))^{-1/2} 1_{\{d\}}]^T .$$

Hence E is regular. Let $P^n = \prod_1^\infty P_{\theta_{ni}}$ and $\tilde{P}^n = \prod_1^\infty P_{\tilde{\theta}_{ni}}$ and assume there exists a $\rho > 0$ such that

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

Then by Theorem 2.3, $P^n \triangleleft\triangleright \tilde{P}^n$ or $P^n \vartriangle \tilde{P}^n$, with the former being true if and only if

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

and

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

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

$$c(\theta)\exp(\sum_1^n \theta_j T_j) \in dP_\theta/d\nu \qquad \text{for all} \quad \theta \in \Theta.$$

This is the exponential family with the natural parameterization, and we further assume that the random variables $\{T_j\}$ are affinely linearly independent, i.e. $\Sigma t_j T_j = 0$ a.e.-ν implies $t_j = 0$ for all j, and there does not exist $\{t_j\}$ and $c \in \mathbb{R}$ such that $\Sigma t_j T_j = c$ a.e.-ν.

Then (A1) holds since $P_\theta \equiv \nu$ for all $\theta \in \Theta$. Also $P_\theta = P_{\theta'}$ implies $\theta = \theta'$ by the affine linear independence of $\{T_j\}$ and hence (A4) holds. By Theorem 78.2 in Strasser (1985), E is differentiable with

$$\nabla h_\theta = [(T_1 - P_\theta T_1)h_\theta, \ldots, (T_d - P_\theta T_d)h_\theta]^T$$

where $P_\theta T_j = \int T_j \, dP_\theta$ for all j. Also the mapping $\theta \in \Theta \to P_\theta T_j$ is continuous by Lemma 3.5.5 of Fabian/Hannan (1985). By using this continuity and applying the lemma again, the mapping $\theta \in \Theta \to \nabla h_\theta$ is continuous i.e. 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 (A3) is true then we have a contiguity/asymptotic separation dichotomy by Theorem 2.3. If we do not assume (A3) is true but instead only assume $\{\tilde{\theta}_{ni}: n, i \in \mathbb{N}\}$ is also contained in some compact subset of Θ we again get a contiguity/asymptotic separation dichotomy by Theorem 2.3. 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 $\{\theta_{ni}\} \cup \{\tilde{\theta}_{ni}\}$, and hence (A3) is satisfied trivially.

CHAPTER III

NECESSARY AND SUFFICIENT CONDITIONS FOR DICHOTOMY

3.1 Preliminaries and Auxiliary Results. In this chapter, we prove a converse of the sufficiency result in Chapter 2 for a specific experiment $E = (\mathbb{R}^k, \mathscr{B}(\mathbb{R}^k), \{P_{tR}: t \in \mathbb{R}^k, R \in \mathscr{R}\})$. 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 $R \in \mathscr{R}$ and $t \in \mathbb{R}^k$. For this experiment E, the parameter space $\Theta = \mathbb{R}^k \times \mathscr{R} \subset \mathbb{R}^{k(k+1)}$ when $k \geq 2$ and $\Theta = \mathbb{R}$ when $k = 1$.

We also prove a partial converse for an experiment E based on an underlying probability measure P on $(\mathbb{R}^k, \mathscr{B}(\mathbb{R}^k))$ and invertable affine perturbations of P. Thus $P_{(t,A)} = P \circ AT_t$ where $A \in \mathscr{A}$, the set of all invertable linear transformations from $\mathbb{R}^k$ to $\mathbb{R}^k$. For this experiment E, the parameter space $\Theta = \mathbb{R}^k \times \mathscr{A} \subset \mathbb{R}^{k(k+1)}$.

In this context for $k = 1$, Shepp (1965) proved a very interesting result. Specifically he showed that $\prod_1^\infty P \perp \prod_1^\infty P_{t_i}$ for all $\{t_i\} \notin \ell^2$. Even more interesting is that he showed

$\prod_1^\infty P \equiv \prod_1^\infty P_{t_i}$ for all $\{t_i\} \in \ell^2$ if and only if $P \equiv \lambda$ (where λ is Lebesque measure) and there exists $f \in dP/d\lambda$ which is locally absolutely continuous and such that

$$\int ((f')^2/f)\, d\lambda < \infty \tag{3.1}$$

i.e. f has finite Fisher information. This result was generalized to a more abstract setting by LeCam (1970), Proposition 2. This result included as a corollary a generalization of Shepp's translation result to the multivariate setting.

Steele (1986) generalized both of the above multivariate results by including all rigid motions. First Steele proved that $\prod_1^\infty P \perp \prod_1^\infty P_{\theta_i}$ for all $\{\theta_i\} \notin \ell^2$ such that $\theta_i \to 0$ as $i \to \infty$. Secondly and more importantly he showed that $\prod_1^\infty P \equiv \prod_1^\infty P_{\theta_i}$ for all $\{\theta_i\} \in \ell^2$ such that $\theta_i \to 0$ if and only if $P \equiv \lambda$ and for all one-parameter groups $\{p(s): s \in \mathbb{R}\}$ in the space of rigid motions, there exists a number K such that

$$\left|\int h(x)\, [d/ds(\varphi(p(s)x))|_{s=0}]\, d\lambda(x)\right| \leq K\|\varphi\|_2 \tag{3.2}$$

for all $\varphi \in C_c^\infty(\mathbb{R}^k)$. Here $h \in (dP/d\lambda)^{1/2}$ 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. We now show that if E is dominated by λ and is differentiable at $(0,I)$ (I is identity operator on $\mathbb{R}^k$), then (3.2) is satisfied.

<u>Proposition 3.1</u>. Let E be an experiment as given above and suppose E is dominated by λ and is differentiable at $\theta = (0,I)$. Then (3.2) is satisfied.

Proof: Let $h \in dP/d\lambda$, and $\varphi \in C_c^\infty(\mathbb{R}^k)$. Then

$$|\int h(x)[d/ds(\varphi(p(s)x))|_{s=0}]\, d\lambda(x)|$$

$$= \left|\lim_{\epsilon\to 0} \int h(x)[\varphi(p(\epsilon)x)-\varphi(x)](\epsilon^{-1})\, d\lambda(x)\right|$$

$$= \left|\lim_{\epsilon\to 0} \int [h(p(-\epsilon)x)-h(x)](\epsilon^{-1})\varphi(x)\, d\lambda(x)\right|$$

$$= \left|\int (\nabla p_{(0)}\cdot\nabla h_{(0,I)})\, \varphi\, d\lambda\right|$$

$$\leq \|\nabla p_{(0)}\cdot\nabla h_{(0,I)}\|_2\|\varphi\|_2.$$

In the above, we have used the Lebesque dominated convergence theorem, a change of variables, and the chain rule in differential calculus for normed linear spaces. Also we have used that one-parameter groups in a Lie group are differentiable (cf. Warner (1983), pages 102-103). □.

By Proposition 3.1, (3.2) 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. Thus we will actually show that if $P \equiv \lambda$ and satisfies (3.2), then E is differentiable. In proving the main theorem we need two results due to LeCam (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 chapter 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 3.2.__ Let $\psi: U \subset \mathbb{R}^d \to \mathbb{H}$ where $\mathbb{H}$ is a Hilbert space and U is a Borel subset. Suppose

$$(3.3) \qquad \limsup_{u' \to u} \|\psi(u') - \psi(u)\| / |u' - u| < \infty$$

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

__Proposition 3.3.__ Let $E = (\Omega, \mathcal{F}, \{P_\theta: \theta \in \Theta\})$ be an experiment satisfying (A1) and $\theta_o \in \Theta$. If $\prod_1^\infty P_{\theta_o} \equiv \prod_1^\infty P_{\theta_i}$ for all $\{\theta_i\}$ such that $\{(\theta_o - \theta_i)\} \in \ell^2$, then

$$(3.4) \qquad \limsup_{\theta\to\theta_o} H(\theta_o,\theta)/|\theta_o-\theta| < \infty .$$

Lemma 3.1. Let $h \in L^2(\mathbb{R}^k)$. If $\mathcal{A}$ is the set of all invertable linear transformations from $\mathbb{R}^k$ to $\mathbb{R}^k$, then

$$(3.5) \qquad \lim_{(t,A)\to(0,I)} \|h\circ AT_t-h\|_2 = 0$$

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

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

$$(3.6) \qquad \|h\circ AT_t-h\|_2 \le \|h\circ AT_t-g\circ AT_t\|_2 + \|g\circ AT_t-g\|_2 + \|g-h\|_2.$$

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 and the invariance of λ under translation. 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 (3.6),

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

Since ϵ was arbitrary, the result follows. □.

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

$$E_1^P = (\mathbb{R}^k, \mathcal{B}(\mathbb{R}^k), \{P_\theta: \theta \in \mathbb{R}^k\}),$$

$$E_2^P = (\mathbb{R}^k, \mathcal{B}(\mathbb{R}^k), \{P_\theta: \theta \in \Theta = \mathbb{R}^k \times \mathfrak{A}\}) \quad (k \geq 2), \quad \text{and}$$

$$E_3^P = (\mathbb{R}^k, \mathcal{B}(\mathbb{R}^k), \{P_\theta: \theta \in \Theta = \mathbb{R}^k \times \mathcal{A}\}),$$

where E_1^P corresponds to the translation experiment, E_2^P is the rigid motion experiment, and E_3^P is the invertable affine transformation experiment. We sometimes suppress the superscript P for notational convenience when the underlying probability measure is clear.

If $E \in \{E_1^P, E_2^P, E_3^P\}$, let $\ell(E) = \{\theta \in \Theta$: (3.4) holds at $\theta\}$. Thus if E is dominated, $\ell(E)$ represents the points $\theta \in \Theta$ at which the mapping $\theta \in \Theta \to h_\theta$ is Lipschitz.

Let $\theta \in \Theta$. As given in the above, θ represents an element in some Euclidean space. However we will sometimes find it convenient to let θ also represent the transformation i.e. 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 overtly distinguish between the two interpretations of θ in the rest of this thesis.

3.2 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 prove a partial result in this direction for the invertable affine transformation experiment. These are given in Theorems 3.2 and 3.3. Before proving them we need some further results, related to E_1, E_2, and E_3, which are interesting in their own right.

First in Lemma 3.2, we show that in all 3 cases if $P \ll \lambda$ and if (3.4) is satisfied at one point then (3.4) is satisfied at all points. This is equivalent to saying that if $P \ll \lambda$ and the mapping $\theta \in \Theta \to h_\theta$ is Lipschitz at one point, then it is Lipschitz on all of Θ. The usefulness of this result is derived from Proposition 3.2, and is given in Theorem 3.1. Specifically we show that if $P \ll \lambda$ and (3.4) holds at one point, then E_i is differentiable for $i = 1,2,3$.

<u>Lemma 3.2.</u> Let $P \ll \lambda$, $E \in \{E_1, E_2, E_3\}$, and suppose $\ell(E) \neq \phi$. Then $\ell(E) = \Theta$.

Proof: Let $E = E_1$, $\theta_o \in \ell(E)$, and $\theta_1 \in \Theta = \mathbb{R}^k$. Then

$$\limsup_{\theta\to\theta_1} \|h_\theta - h_{\theta_1}\| / |\theta-\theta_1|$$

$$= \limsup_{\theta\to\theta_1} \|h\circ\theta - h\circ\theta_1\| / |\theta-\theta_1|$$

$$= \limsup_{\theta\to\theta_1} \|h\circ\theta\theta_1^{-1}\theta_o - h\circ\theta_o\| / |\theta\theta_1^{-1}\theta_o-\theta_o| < \infty,$$

since $\theta_o \in \ell(E_1)$. Note that the second equality follows by the invariance of Lebesgue measure under translation, and the invariance of the absolute value norm on $\mathbb{R}^k$ under translation. Thus $\theta_1 \in \ell(E_1)$.

Let $E = E_3$, $\theta_o \in \ell(E_3)$, and $\theta_1 \in \Theta = \mathbb{R}^k \times \mathcal{A}$. Now temporarily substitute (t_o, A_o) for θ_o, (t_1, A_1) for θ_1, and (t, A) for θ. Then

$$\begin{aligned}\theta\theta_1^{-1}\theta_o &= AT_tT_{-t_1}A_1^{-1}A_oT_{t_o} \\ &= AA_1^{-1}A_oT_{(A_o^{-1}A_1(t-t_1)+t_o)}.\end{aligned}$$

Thus

$$\begin{aligned}(3.7)\qquad |\theta\theta_1^{-1}\theta_o-\theta_o|^2 &= |(A_o^{-1}A_1(t-t_1);(AA_1^{-1}-I)A_o)|^2 \\ &\le |A_o^{-1}A_1|^2|t-t_1|^2 + |A_1^{-1}A_o|^2|A-A_1|^2 \\ &\le K^2(A_o,A_1)|\theta-\theta_1|^2\end{aligned}$$

where $K(A_o,A_1) = \max\{|A_o^{-1}A_1|, |A_1^{-1}A_o|\}$. Note that the first inequality follows since the norm of matrix as elements of $\mathbb{R}^{k^2}$ is greater than or equal to the norm of a matrix as a linear operator from $\mathbb{R}^k$ to $\mathbb{R}^k$. By (3.7), there exists $K < \infty$ such that

$$(3.8) \qquad \begin{aligned} &\|h_\theta - h_{\theta_1}\| / |\theta - \theta_1| \\ &\quad \le [\|h_{\theta\theta_1^{-1}\theta_o} - h_{\theta_o}\| \cdot |A_1^{-1}A_o|^{1/2} / |\theta\theta_1^{-1}\theta_o - \theta_o|] \\ &\quad\quad \cdot [|\theta\theta_1^{-1}\theta_o - \theta_o| / |\theta - \theta_1|] \\ &\quad \le K\|h_{\theta\theta_1^{-1}\theta_o} - h_{\theta_o}\| / |\theta\theta_1^{-1}\theta_o - \theta_o| \end{aligned}$$

for all $\theta \in \Theta$. This implies $\theta_1 \in \ell(E_3)$.

The case of $E = E_2$ follows by a similar and slightly easier argument than that just given for $E = E_3$. □.

<u>Remark.</u> Notice in (3.8), if we restrict ourselves to $\mathbb{R}\times\mathcal{A}_c$ where $\mathcal{A}_c$ is a compact subset of $\mathcal{A}$, then we can find a K for which (3.8) holds for all $\theta_1 \in \mathbb{R}\times\mathcal{A}_c$. In particular this is true when $\mathcal{A}_c = \mathfrak{K}$. This will be used later on.

In Theorem 3.1, by using Lemma 3.2 and Proposition 3.1, we prove that a sufficient condition for differentiability of E_i is that $\ell(E_i) \ne \phi$, for $i = 1,2,3$. The result for $i = 1$ was previously known (c.f. the remark following Proposition 2 in LeCam (1970)). The purpose of proving it again in the case $i = 1$, is that it is useful in understanding the proofs for the cases $i = 2,3$.

Before stating and proving the first theorem, we outline some of the key ideas of the proof. The essential idea behind the proofs of all 3 cases is to use Proposition 3.2 and Lemma 3.2 to show that E is differentiable except on a null set, and then to use the differentiability of E on a dense set in Θ to get differentiability on all of Θ. The case $E = E_1$ is proven easily this way since the differentials are all translates of each other. In the case $E = E_2$, there are two main difficulties. First Θ is already a null set so Proposition 3.2 doesn't even give 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 3.2 for this new experiment. We then transform back to the original experiment E_2 to get differentiability on a dense set in Θ. The second difficulty is that differentiability at a point θ_o is not directly transferable to other points in Θ. However locally and asymptotically it is transferable. This last difficulty is also inherent in the case $E = E_3$.

<u>Theorem 3.1.</u> Let $(\Omega,\mathcal{F},P)$ be an experiment dominated by λ, $E \in \{E_1^P, E_2^P, E_3^P\}$ and suppose $\ell(E) \neq \phi$. Then

(a) E is differentiable in the case $i = 2$ and in this case there exists a family of differentials $\{\nabla h_\theta : \theta \in \Theta\}$ which is uniformly bounded over compact subsets of Θ.

(b) E is regular in the case $i = 1, 3$.

Proof: Let $E = E_1$. Then $\ell(E_1) = \Theta = \mathbb{R}^k$ by Lemma 3.2. By Proposition 3.2, there exists $\Theta_o \subset \Theta$ such that E_1 is differentiable at $\theta = \theta_o$ for all $\theta_o \in \Theta_o$, and $\lambda(\Theta \backslash \Theta_o) = 0$. Let $\theta_o \in \Theta_o$, $\theta_1 \in \Theta$. Then for $\theta \in \Theta$

$$\|h_\theta - h_{\theta_1} - (\theta-\theta_1)^T \cdot \nabla h_{\theta_o} \circ \theta^{-1}\theta_1\|$$
$$= \|h_{\theta\theta_1^{-1}\theta_o} - h_{\theta_o} - (\theta-\theta_1)^T \cdot \nabla h_{\theta_o}\|$$

and

$$|\theta-\theta_1| = |\theta\theta_1^{-1}\theta_o - \theta_o|.$$

Hence on letting $\theta \to \theta_1$, we see E_1 is differentiable at θ_1 with $\nabla h_{\theta_1} = \nabla h_{\theta_o} \circ \theta_o^{-1}\theta_1$. By Lemma 3.1, the map $\theta \in \Theta \to \nabla h_\theta$ is continuous and hence E is regular.

Let $E = E_2$. Then exactly as in the previous case, $\ell(E_2) = \Theta = \mathbb{R}^k \times \mathfrak{R}$. Now let $L = k(k-1)/2$. By the general implicit function theorem (e.g. Auslander/MacKenzie (1963)), for each $R \in \mathfrak{R}$, there exists neighborhoods V of $R \in \mathfrak{R}$ and U of $0 \in \mathbb{R}^L$, and $\psi \in C^1(U,V)$, such that $\psi(0) = R$ and ψ is a homeomorphism with the differential of ψ at $u \in U$, $d\psi_u$, having rank L for all $u \in U$. Also there exist a neighborhood W of $R \in \mathbb{R}^{k^2}$, where $V = W \cap \mathfrak{R}$, 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, $\bar{U}$ is compact, and the above is true on a neighborhood of $\bar{U}$ with the same η and ψ.

Now fix an $R_o \in \mathfrak{R}$ and let η, ψ, W, U, and V be as above. By the compactness of $\bar{U}$,

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

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

$$(3.9) \qquad |\psi(u')-\psi(u)| \le \beta|u'-u| \qquad \text{for all } u,u' \in U.$$

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. Let $\{u_j\}$ and $\{u_j'\}$ be arbitrary sequences in U such that $u_j \neq u_j'$ for all j. It suffices to prove that $\{|\psi(u_j')-\psi(u_j)|/|u_j'-u_j|\}$ is bounded away from 0. Since this is true if and only if every subsequence has a subsubsequence which is bounded away from 0, it suffices to prove that the above sequence is bounded away from 0 under the additional assumptions that $u_j \to u \in \bar{U}$ and $u_j' \to u' \in \bar{U}$. If $u \neq u'$, then $\{|\psi(u_j')-\psi(u_j)|/|u_j'-u_j|\}$ is bounded away from 0 by the continuity and injectivity of ψ. If $u = u'$, then

$$(3.10) \qquad \lim_{j\to\infty} |\psi(u_j')-\psi(u_j)-d\psi_u(u_j'-u_j)|/|u_j'-u_j| = 0$$

by the continuous differentiability of ψ, the convexity of $\bar{U}$, and standard differential calculus. But $d\psi_u$ has rank L so (3.10) implies

$$(3.11) \qquad \liminf_{j\to\infty} |\psi(u_j')-\psi(u_j)|/|u_j'-u_j| > 0.$$

Thus combining this with (3.9), there exists positive constants α,β such that

$$(3.12) \qquad \alpha|u'-u| \le |\psi(u')-\psi(u)| \le \beta|u'-u|$$

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

Now consider a new experiment $E_2^* = (\mathbb{R}^k, \beta(\mathbb{R}^k), \{P_{\theta^*}^*: \theta^* \in \Theta^*\}$ where $\Theta^* = \mathbb{R}^k \times U$, $P_{(t,u)}^* = P_{(t,\psi(u))}$, and $h_{\theta^*}^* \in (dP_{\theta^*}^*/d\lambda)^{1/2}$. We define a new map $\psi_1: \mathbb{R}^k \times U \to \Theta$ by $\psi_1(t,u) = (t,\psi(u))$. By (3.12),

$$(3.13) \qquad |\psi_1(\theta^*)-\psi_1(\theta_o^*)| \le (1+\beta)|\theta^*-\theta_o^*|$$

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

$$(3.14)\qquad \limsup_{\theta^*\to\theta^*_o} \|h_{\theta^*}-h_{\theta^*_o}\|/|\theta^*-\theta^*_o| < K(1+\beta)$$

by dividing and multiplying by $|\psi_1(\theta^*)-\psi_1(\theta^*_o)|$, invoking (3.13), and recalling the remark following Lemma 3.2. Thus $\ell(E^*_2) = \Theta^*$. By Proposition 3.2, there exists $\Theta^*_o \subset \Theta^*$ such that E^*_2 is differentiable on Θ^*_o and $\lambda(\Theta^*\setminus\Theta^*_o) = 0$.

Now define a map $\eta_1: \mathbb{R}^k\times W \to \mathbb{R}^k\times U$ 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 $\mathbb{R}^k\times U$. This last fact also implies $\psi_1\circ\eta_1$ is the identity map on $\mathbb{R}^k\times V$. Thus by the chain rule E is differentiable on $\Theta_o = \psi_1(\Theta^*_o)$ with

$$(3.15)\quad \nabla h_\theta = \nabla\eta_{1\theta}\nabla h^*_{\eta_1(\theta)} \qquad \text{for } \theta \in \Theta_o.$$

Let $t_o \in \mathbb{R}^k$ be fixed. Then there exists a sequence $\{(t_j,u_j)\} \in \Theta^*_o$ such that $(t_j,u_j) \to (t_o,0)$. Letting $\theta_j = \psi_1(t_j,u_j)$, we see that $\theta_j \to \theta_o = (t_o,R_o)$. Now for a fixed j and letting $\theta \in \mathbb{R}^k\times V$, we have

$$(3.16)\qquad \|h_\theta-h_{\theta_o}-(\theta-\theta_o)^T\cdot\nabla h_{\theta_j}\circ\theta_j^{-1}\theta_o\|$$

$$= \|h\circ\theta\theta_o^{-1}\theta_j-h\circ\theta_j-(\theta-\theta_o)^T\cdot\nabla h_{\theta_j}\|$$

$$\le \|h\circ\theta\theta_o^{-1}\theta_j-h\circ\theta_j-(\theta\theta_o^{-1}\theta_j-\theta_j)^T\cdot\nabla h_{\theta_j}\|$$
$$+ \|(\theta\theta_o^{-1}\theta_j-\theta_j-\theta+\theta_o)^T\cdot\nabla h_{\theta_j}\|.$$

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

$$(3.17) \quad \theta\theta_o^{-1}\theta_j = RT_tT_{-t_o}R_o^{-1}R_jT_{t_j}$$

$$= RR_o^{-1}R_jT_{(R_j^{-1}R_o(t-t_o)+t_j)}.$$

Thus corresponding to the first term in (3.16),

$$(3.18) \quad |\theta\theta_o^{-1}\theta_j-\theta_j| = |(R_j^{-1}R_o(t-t_o), RR_o^{-1}R_j-R_j)|$$

$$= |(t-t_o, RR_o^{-1}-I)|$$

$$= |(t-t_o, R-R_o)|$$

$$= |\theta-\theta_o|.$$

For the second term in (3.16), by (3.17) and the Cauchy-Schwarz inequality,

$$(3.19) \quad \|(\theta\theta_o^{-1}\theta_j-\theta_j-\theta+\theta_o)^T \cdot \nabla h_{\theta_j}^T\|^2$$

$$\leq |\theta\theta_o^{-1}\theta_j-\theta_j-\theta+\theta_o|^2 \|(\nabla h_{\theta_j}^T \cdot \nabla h_{\theta_j})^{1/2}\|^2$$

$$= |((R_j^{-1}R_o-I)(t-t_o), RR_o^{-1}R_j-R_j-R+R_o)|^2 \|(\nabla h_{\theta_j}^T \cdot \nabla h_{\theta_j})^{1/2}\|^2$$

$$= |((R_o-R_j)(t-t_o), (I-RR_o^{-1})(R_o-R_j))|^2 \|(\nabla h_{\theta_j}^T \cdot \nabla h_{\theta_j})^{1/2}\|^2$$

$$\le |R_o-R_j|^2 |(t-t_o, R-R_o)|^2 \|(\nabla h_\theta^T \cdot \nabla h_{\theta_j})^{1/2}\|^2$$

where in the last inequality, we have used that the norm of R_o-R_j as a linear operator is less than or equal to $|R_o-R_j|$, and we have also used that the inverse of an element in $\mathfrak{R}$ is equal to its adjoint. On dividing the quantity in (3.16) by $|\theta-\theta_o|$, and letting $\theta \to \theta_o$, we see that the first term goes to 0 by the differentiability of E_2 at θ_j and by (3.18). Thus by (3.17) and (3.19), if $g^j = \nabla h_{\theta_j} \circ \theta_j^{-1}\theta_o$,

$$\limsup_{\theta\to\theta_o} \|h_\theta - h_{\theta_o} - (\theta-\theta_o)^T \cdot g^j\| / |\theta-\theta_o| \tag{3.20}$$

$$\le |R_j-R_o| \|(\nabla h_{\theta_j}^T \cdot \nabla h_{\theta_j})^{1/2}\|.$$

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

$$\|(\nabla h_{\theta_j}^T \cdot \nabla h_{\theta_j})^{1/2}\| \tag{3.21}$$

$$= \|(\nabla h^{*T}_{\eta_1(\theta_j)} \nabla \eta_{1\theta_j}^T \nabla \eta_{1\theta_j} \nabla h^*_{\eta_1(\theta_j)})^{1/2}\|.$$

But $\nabla\eta_{1\theta_j} \to \nabla\eta_{1\theta_o}$ by 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 (3.14). Thus the quantity in (3.21) is uniformly bounded in j, and hence by (3.20),

$$(3.22) \quad \lim_{j\to\infty} \limsup_{\theta\to\theta_o} \|h_\theta - h_{\theta_o} - (\theta-\theta_o)^T \cdot g^j\| / |\theta-\theta_o| = 0.$$

Let $M = \mathbb{R}^k \times (d\psi_0(\mathbb{R}^L)) \subset \mathbb{R}^{k(k+1)}$. Then M is an $(L+k)$ dimensional subspace since $d\psi_0$ has rank L. Now let $\theta^*_o = (t_o, 0)$. Then by the differentiability of ψ_1,

$$(3.23) \quad \lim_{\theta^*\to\theta^*_o} |\psi_1(\theta^*) - \psi_1(\theta^*_o) - d\psi_{1\theta^*_o}(\theta^*-\theta^*_o)| / |\theta^*-\theta^*_o| = 0.$$

By dividing and multiplying the expression in (3.23) by $|\psi_1(\theta^*)-\psi_1(\theta^*_o)|$, noting that ψ_1 is a homeomorphism, and invoking (3.11),

$$(3.24) \quad \lim_{\theta\to\theta_o} |\theta-\theta_o - Pr_M(\theta-\theta_o)| / |\theta-\theta_o| = 0,$$

where Pr_M is the projection operator onto M.

Also if $y \in M$ and $|y| = 1$, there exists $\theta^*_1 \in \Theta^*$ such that $d\psi_{1\theta^*_o}(\theta^*_1-\theta^*_o) = s_o y$, where $s_o > 0$. Then

$$(3.25)\quad \lim_{s\to 0} |\psi_1(\theta_o^*+s(\theta_1^*-\theta_o^*))-\psi_1(\theta_o^*)-sd\psi_{1\theta_o^*}(\theta_1^*-\theta_o^*)|/s|\theta_1^*-\theta_o^*| = 0,$$

where the limit is over $s \in \mathbb{R}\backslash\{0\}$. Hence if we again divide and multiply by $|\psi_1(\theta_o^*+s(\theta_1^*-\theta_o^*))-\psi_1(\theta_o^*)|$ and invoke (3.11), we obtain

$$(3.26)\quad \liminf_{\theta\to\theta_o} |((\theta-\theta_o)/|\theta-\theta_o|)-y| = 0$$

for all $y \in M$ such that $|y| = 1$.

Let $\epsilon > 0$. Then by (3.22) there exists $J \in \mathbb{N}$ such that $j \geq J$ implies

$$(3.27)\quad \limsup_{\theta\to\theta_o} \|h_\theta-h_{\theta_o}-(\theta-\theta_o)^T\cdot g^j\|/|\theta-\theta_o| < \epsilon.$$

Let $y \in M$ be such that $|y| = 1$, and let $\{\theta_n\} \subset \Theta$ be a sequence such that $\theta_n \to \theta_o$ and

$$(3.28)\quad \lim_{n\to\infty} |((\theta_n-\theta_o)/|\theta_n-\theta_o|)-y| = 0.$$

Then for $j, j' \geq J$ and $n \in \mathbb{N}$,

$$(3.29)\quad \|y\cdot(g^j-g^{j'})\| \leq \|y\cdot g^j-((h_{\theta_n}-h_{\theta_o})/|\theta_n-\theta_o|)\| + \|((h_{\theta_n}-h_{\theta_o})/|\theta_n-\theta_o|)-y\cdot g^{j'}\|$$

Now triangulate the first term on the right hand side of (3.29) using the term $((\theta_n-\theta_o)^T\cdot g^j)/|\theta_n-\theta_o|$, and triangulate the second using the term $((\theta_n-\theta_o)^T\cdot g^{j'})/|\theta_n-\theta_o|$. Then invoke (3.28) and (3.27), and apply the Lebesgue dominated convergence theorem as $n \to \infty$ to get

$$\|y\cdot(g^j-g^{j'})\| < 2\epsilon. \tag{3.30}$$

Since this is true for any $j, j' \geq J$, we have just shown $\{y\cdot g^j\}$ is a Cauchy sequence in $L^2(\lambda)$. Since y was arbitrary, this shows $\{Pr_M g^j\}$ is a Cauchy sequence in $\prod_1^{k(k+1)} L^2(\lambda)$. Let g be the limit of $\{Pr_M g^j\}$ in $\prod_1^{k(k+1)} L^2(\lambda)$. By (3.24) and (3.30),

$$\limsup_{\theta\to\theta_o} \|h_\theta-h_{\theta_o}-(\theta-\theta_o)^T\cdot g\|/|\theta-\theta_o| \leq 2\epsilon$$

after triangulating the numerater on $(\theta-\theta_o)^T\cdot g^j$ and $Pr_M(\theta-\theta_o)\cdot g^j$ for large j. Thus E_2 is differentiable at θ_o. Since θ_o 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 remark following Lemma 3.2, and invoking (3.26), we also obtain a family of differentials

$\{\nabla h_\theta : \theta \in \Theta\}$ which are uniformly bounded over compact subsets of Θ.

Let $E = E_3$. By Lemma 3.2, $\Theta = \ell(E_3)$, and by Proposition 3.2, there exists a measurable subset $\Theta_o \subset \Theta$ such that E_3 is differentiable on Θ_o and $\lambda(\Theta \setminus \Theta_o) = 0$. Again let $\theta_o \in \Theta$ and $\{\theta_j\} \subset \Theta_o$ be such that $\theta_j \to \theta_o$. A similar argument as done in the case $E = E_2$, shows that if $g^j = \nabla h_{\theta_j} \circ \theta_j^{-1} \theta_o$, then

$$(3.31) \qquad \lim_{j\to\infty} \limsup_{\theta\to\theta_o} \| h_\theta - h_{\theta_o} - (\theta-\theta_o)^T \cdot g^j \| = 0.$$

The only difference is that it is easier since one can carry out the argument directly instead of transforming the experiment. Since Θ is open it follows that $\{g^j\}$ is Cauchy in $\prod_1^{k+k^2} L^2(\nu)$. By (3.31) the limit g of $\{g^j\}$ is seen to be the differential of h_θ at $\theta = \theta_o$. Thus E is differentiable at $\theta = \theta_o$. Also by Holder's inequality,

$$\| \nabla h_{\theta_j} - \nabla h_{\theta_o} \| \leq \| \nabla h_{\theta_j} - \nabla h_{\theta_o} \circ \theta_o^{-1} \theta_j \| + \| \nabla h_{\theta_o} \circ \theta_o^{-1} \theta_j - \nabla h_{\theta_o} \|.$$

By a change of variables and Lemma 3.1, $\nabla h_{\theta_j} \to \nabla h_{\theta_o}$. Hence by an easy argument, this implies the map $\theta \in \Theta \to \nabla h_\theta$ is continuous and E is regular.

Remark. Let $E \in \{E_1^P, E_2^P, E_3^P\}$. We will sometimes write E^P and drop the subscript i in order to conveniently keep track of several experiments generated by different perturbations of the same probability measure.

By Theorem 3.1, we now have a useful sufficient condition for checking when E^P is differentiable. Namely we need only check whether $\ell(E^P) \neq \phi$. Most often this will be done by checking if $\ell(E^P)$ contains the identity. We apply this in the next result, Proposition 3.3, to show that if P_o, P_1 are two probability measures with E^{P_o} being differentiable, then $E^{P_o * P_1}$ is differentiable.

Proposition 3.3. Let $i \in \{1,2,3\}$ and $(\mathbb{R}^k, \mathcal{B}(\mathbb{R}^k), \{P_o, P_1\})$ be an experiment such that $P_o \ll \lambda$ and $E_i^{P_o}$ is differentiable. Then $E_i^{P_o * P_1}$ is differentiable.

Proof. Let $P = P_o * P_1$ and note that by Fubini, $P \ll \lambda$. Let Θ be the appropriate parameter space depending on whether $i = 1,2,$ or 3, and let $h_\theta \in (dP_\theta/d\lambda)^{1/2}$ for $\theta \in \Theta$, and $f_o \in dP_o/d\lambda$. Then

$$h_\theta(y) = \left(\int f_o(\theta y - x)\, dP_1(x)\right)^{1/2}(K(\theta)) \tag{3.32}$$

for all $y \in \mathbb{R}^k \setminus B$ where $\lambda(B) = 0$, $K(\theta) = 1$ for $i = 1$ and 2, and $K(\theta) = |A|^{1/2}$ for $i = 3$ where $\theta = (t, A)$.

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

(3.33) $$(h_{\theta'}-h_\theta)^2 = h^2_{\theta'}-2h_{\theta'}h_\theta+h^2_\theta$$

by algebra. But by (3.32), if $h_o \in (dP_o/d\lambda)^{1/2}$,

$$h_{\theta'}(y)h_\theta(y) = (\int f_o(\theta'y-x)\, dP_1(x) \int f_o(\theta y-x)\, dP_1(x))^{1/2}(K(\theta)K(\theta'))$$

$$\geq (\int h_o(\theta'y-x)h_o(\theta y-x)\, dP_1(x))(K(\theta)K(\theta'))$$

for a.e.-λ y by Holder's inequality. Combining this with (3.33),

(3.34) $$(h_{\theta'}(y)-h_\theta(y))^2$$

$$\leq \int (K(\theta')h_o(\theta'y-x)-K(\theta)h_o(\theta y-x))^2\, dP_1(x)$$

for a.e.-λ y. Thus letting $\theta' \to \theta$, we obtain that $\theta \in \ell(E^P_i)$ by (3.34) and Fubini. By Theorem 3.1, E^P_i is differentiable. □.

We are now ready to state and prove the two main theorems of this chapter, which give a partial converse to Theorem 2.2 in the case $E \in \{E_1, E_2\}$ and a partial converse in the case $E = E_3$.

Theorem 3.2. Let $(\mathbb{R}^k, \mathfrak{B}(\mathbb{R}^k), P)$ be an experiment with $E \in \{E_1^P, E_2^P\}$, and in the case of $E = E_2$, assume E satisfies (A4).

Then the following are true:

$$\text{(a)} \quad P^n = \prod_1^\infty P_{\theta_{ni}} \triangleleft\triangleright \tilde{P}^n = \prod_1^\infty P_{\tilde{\theta}_{ni}} \quad \text{for all} \quad \{\theta_{ni}\}, \{\tilde{\theta}_{ni}\}$$

such that

$$(3.35) \qquad \limsup_{n\to\infty} \sum_1^\infty |\theta_{ni} - \tilde{\theta}_{ni}|^2 < \infty$$

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

$$\text{(b)} \quad P^n = \prod_1^\infty P_{\theta_{ni}} \blacktriangle \tilde{P}^n = \prod_1^\infty P_{\tilde{\theta}_{ni}} \quad \text{for all} \quad \{\theta_{ni}\}, \{\tilde{\theta}_{ni}\}$$

such that

$$(3.36) \qquad \limsup_{n\to\infty} \sum_1^\infty |\theta_{ni} - \tilde{\theta}_{ni}|^2 = \infty.$$

Proof: By Lemma 2.1, we can without loss of generality assume $\theta_{ni} = 0$ for $E = E_1$ and $\{\theta_{ni}\} \subset \{0\}\times\mathfrak{R}$ for $E = E_2$. We will prove both cases simultaneously since many of the arguments for both cases are the same. The major differences will be pointed out when they occur.

(a) $\prod_1^\infty P \equiv P_{\tilde{\theta}} \times \prod_2^\infty P$ for all $\tilde{\theta} \in \Theta$ and hence (A1) is true. This implies that $P \equiv \lambda$ (cf. Steele (1986), Lemma 4.1) and henceforth let λ be the dominating measure. By Proposition 3.3, $\ell(E) \neq \phi$ and by Theorem 3.1, E is differentiable.

For the converse clearly (A1) and (A3) are true, and by Lemma 3.1, (A2) holds. Also (A4) holds for $E = E_1$ by a straightforward argument and for $E = E_2$ by hypothesis. Thus by Theorem 2.1, it suffices to prove

$$(3.37) \quad \lim_{\rho \to 0} \inf \{\|h_{\theta'} - h_\theta\| / |\theta' - \theta| : \theta \in \Theta_c,\ |\theta' - \theta| < \rho\} > 0$$

where $\Theta_c = \{0\}$ for $E = E_1$ and $\Theta_c = \{0\} \times \mathfrak{R}$ for $E = E_2$.

We first prove (3.37) for the case $E = E_1$. Note that $\|h_{\theta'} - h_\theta\| = \|h_{\theta' - \theta} - h_0\|$ by the translation invariance of Lebesgue measure. Thus it suffices to show

$$(3.38) \quad \lim_{\theta \to 0} \|h_\theta - h_0\| / |\theta| > 0.$$

Note that in (3.38), we can use lim instead of limit infimum because E is differentiable. Suppose (3.38) is false. Then

$$(3.39)\qquad \|h_1-h_0\| \le \sum_1^n \|h_{(k/n)}-h_{((k-1)/n)}\|$$

$$= \sum_1^n \|h_{(1/n)}-h_0\|$$

$$= \|h_{(1/n)}-h_0\|/(1/n)$$

where the first equality is by the translation invariance of Lebesgue measure. On letting $n \to \infty$ in (3.39) and invoking the falsity of (3.38), $\|h_1-h_0\| = 0$. This implies $P_0 = P_1$ a contradiction to (A4).

We now prove (3.37) for 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 I.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

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

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

$$(3.40)\qquad |G(R,t) - I_o| = |(t,R) - (0,I)|.$$

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

By the theory of matrix Lie groups there exists a ball of radius r about 0, $B_r(0)$, in the tangent space and 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, $\exp(S)$, is defined as

$$\exp(S) = \sum_{0}^{\infty}(S^n/n!)$$

where we remind the reader that S is in the tangent space $T(0,I)$. In order to prove (3.37), it suffices to prove

$$\liminf_{\theta\to(0,I)} \|h_\theta - h_{(0,I)}\|/|\theta-(0,I)| > 0 \tag{3.41}$$

by the remark following Lemma 3.2. So suppose (3.41) is false and there exists $\theta_j \to (0,I)$ such that

$$\|h_{\theta_j} - h_{(0,I)}\|/|\theta_j-(0,I)| \to 0. \tag{3.42}$$

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

Then

$$(3.43) \qquad |(I_o)-\exp(\alpha_j S_j)|/|\alpha_j| = |\sum_1^\infty(\alpha_j S_j)^n|/|\alpha_j|$$

$$\leq \sum_1^\infty |\alpha_j|^{n-1}|S_j|^n$$

$$= e^{|\alpha_j|}$$

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

$$(3.44) \qquad \limsup_{j\to\infty} |(0,I)-\theta_j|/|\alpha_j| \leq 1$$

by (3.43), since $|\alpha_j| \to 0$. Also

$$(3.45) \qquad \lim_{j\to\infty} \|h\circ\exp(\alpha_j S_j)-h\circ\exp(\alpha_j S_o)\|/|\alpha_j| = 0$$

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

$$\|h\circ\exp(\alpha_j S_o)-h\|/|\alpha_j|$$

$$\leq \{\|h\circ\exp(\alpha_j S_o)-h\circ\exp(\alpha_j S_j)\| + \|h\circ\exp(\alpha_j S_j)-h\|\}/|\alpha_j|,$$

we have

$$(3.46)\qquad \lim_{j\to\infty} \|h\circ\exp(\alpha_j S_o)-h\|/|\alpha_j| = 0$$

by (3.42), (3.44), and (3.45). But the map $\alpha \in (-r,r) \to h\circ\exp(\alpha S_o)$ is differentiable by the chain rule. Hence by (3.46)

$$(3.47)\qquad \lim_{\alpha\to 0} \|h\circ\exp(\alpha S_o)-h\|/|\alpha| = 0.$$

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

$$(3.48)\qquad \|h\circ\exp(\alpha_o S_o)-h\| \le \sum_1^N \|h\circ\exp(n\alpha_o S_o/N) - h\circ\exp((n-1)\alpha_o S_o/N)\|$$

$$= \|h\circ\exp(\alpha_o S_o/N)-h\|/(1/N)$$

for all $N \in \mathbb{N}$ where the inequality is by Minkowski's inequality, and the equality is from the invariance of λ under rigid motions. Thus

$$\|h\circ\exp(\alpha_o S_o)-h\| = 0$$

by (3.47), and letting $N \to \infty$ in (3.48). Hence if $\theta_o = \exp(\alpha_o S_o)$, $P = P_{\theta_o}$, a contradiction to the identifiability assumption in (A4). Thus (3.37) is true in the case $E = E_2$ and we have proven the converse portion of part (a).

(b) First we convolute P with $\mathcal{N}(0,I)$. Then note that by characteristic functions, Proposition A.4, and the invariance of $\mathcal{N}(0,I)$ under rotations, (A4) is satisfied for the experiment $E^{P*\mathcal{N}(0,I)}$. Also by Proposition 3.3 $E^{P*\mathcal{N}(0,I)}$ is differentiable. Hence by Proposition A.4 in the appendix, it suffices to prove the result under the additional assumptions that $P \equiv \lambda$ and E is differentiable.

By the proof of (a), (3.37) holds. Hence by (1.7) the result follows. □.

<u>Theorem 3.4.</u> Let $(\mathbb{R}^k, \mathcal{B}(\mathbb{R}^k), P)$ be an experiment and $E = E_3^P$. If E satisfies (A4) then the following are true:

(a) $P^n = \prod_1^\infty P_{\theta_{ni}} \Leftrightarrow \tilde{P}^n = \prod_1^\infty P_{\tilde{\theta}_{ni}}$ for all $\{\theta_{ni}\}$, $\{\tilde{\theta}_{ni}\}$ such that $\{\theta_{ni}\} \subset \Theta_c \subset \Theta$, where Θ_c is compact,

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

and

$$(3.50) \qquad \liminf_{n\to\infty} \inf\{\det(\tilde{A}_{ni}) : i \in \mathbb{N}\} > 0,$$

where $(\tilde{t}_{ni}, \tilde{A}_{ni}) = \tilde{\theta}_{ni}$, if and only if $P \equiv \lambda$ and E is differentiable.

(b) If $P \equiv \lambda$ and E is differentiable, then $P^n = \prod_1^\infty P_{\theta_{ni}} \blacktriangle \tilde{P}^n = \prod_1^\infty P_{\tilde{\theta}_{ni}}$ for all $\{\theta_{ni}\}$, $\{\tilde{\theta}_{ni}\}$ such that $\{\theta_{ni}\} \subset \Theta_c \subset \Theta$, where Θ_c is compact, and

$$(3.51) \qquad \limsup_{n\to\infty} \sum_1^\infty |\theta_{ni} - \tilde{\theta}_{ni}|^2 = \infty$$

or

$$(3.52) \qquad \liminf_{n\to\infty} \inf\{\det(\tilde{A}_{ni}) : i \in \mathbb{N}\} = 0.$$

Proof: (a) Suppose (3.49) and (3.50) both hold. Then without loss of generality, if $\theta_{ni} = (t_{ni}, A_{ni})$, we can assume $t_{ni} = 0$ for all $n, i \in \mathbb{N}$ and $\{\tilde{\theta}_{ni}\} \subset \Theta_c' \subset \Theta$, where Θ_c' is compact. Exactly analogous to the proof of (a) in Theorem 3.3, $P \equiv \lambda$ and E is differentiable.

For the converse, (A1) and (A3) are true trivially while (A2) follows by Lemma 3.1. Since invertable affine transformations again form a Lie group, an exactly analogous argument as in Theorem 3.2 shows that (3.37) holds in this case also. Thus by Theorem 2.1, the converse follows.

(b) By the proof of (a), (3.37) holds. The result now follows by (1.7) □.

<u>Remark</u>. We conjecture that a more complete converse result also holds in the case $i = 3$. Namely we conjecture that if P is any measure on $(\mathbb{R}^k, \mathcal{B}(\mathbb{R}^k))$, and $E = E_3^P$, and if $\{\theta_{ni}\} \subset \mathbb{R}^k \times \mathcal{A}_c$, where $\mathcal{A}_c$ is a compact subset of $\mathcal{A}$, then $\prod_1^\infty P_{\theta_{ni}} \blacktriangle \prod_1^\infty P_{\tilde{\theta}_{ni}}$ for all $\{\tilde{\theta}_{ni}\}$ satisfying (3.51) or (3.52). The main difficulty is that the technique used in the case $i = 2$ does not work here. Namely we are unable to prove (or disprove) that if E_3^P satisfies (A4), then $E_3^{P*\mathcal{N}(0,I)}$ also satisfies (A4).

CHAPTER IV

STATISTICAL APPLICATIONS

4.1 Asymptotic Normality of Likelihood Ratio for Triangular Arrays

Part of the proofs in both Theorems 2.2 and 2.3 can be combined with a result of Oosterhoff/Van Zwet (1979) (Theorem 2, pg 162) to prove asymptotic normality of the likelihood ratio of a triangular array with asymptotically negligible components from a differentiable experiment. More specifically under the above assumptions, the likelihood ratio converges weakly to $\mathcal{N}(-\sigma^2/2, \sigma^2)$. The statistical importance of this is well known (cf. Hajek/Sidak (1967), pages 208-210). For completeness and ease of reference, the necessary result of Oosterhoff/Van Zwet is stated below as a proposition.

Proposition 4.1. Let $E_{ni} = (\Omega_{ni}, \mathcal{F}_{ni}, \{P_{ni}, \tilde{P}_{ni}\})$ be an experiment for $n \in \mathbb{N}$ and $1 \leq i \leq n$, $P^n = \prod_1^n P_{ni}$, $\tilde{P}^n = \prod_1^n \tilde{P}_{ni}$, and $L_n = \sum_1^n \log(\tilde{f}_{ni}/f_{ni})$. Then

$$(4.1) \qquad \mathcal{L}(L_n | P^n) \xrightarrow{\mathcal{D}} \mathcal{N}(-\sigma^2/2, \sigma^2)$$

and

$$(4.2) \qquad \lim_{n\to\infty} \inf\{P_{ni}(|\log(\tilde{f}_{ni}/f_{ni})| \geq \epsilon): 1 \geq i \geq n\} = 0$$

for all $\epsilon > 0$ if and only if

$$(4.3) \qquad \lim_{n\to\infty} \sum_1^n H^2(P_{ni}, \tilde{P}_{ni}) = \sigma^2/4$$

and

$$(4.4) \qquad \lim_{n\to\infty} \sum_1^n \int 1_{\{|f_{ni}-\tilde{f}_{ni}|>\epsilon f_{ni}\}}(h_{ni}-\tilde{h}_{ni})^2 \, d\nu_{ni} = 0$$

for all $\epsilon > 0$ where $\nu_{ni} = P_{ni}+\tilde{P}_{ni}$.

<u>Theorem 4.1.</u> Let $E = (\Omega, \mathcal{F}, \{P_\theta: \theta \in \Theta\})$ be an experiment satisfying (A1) through (A4) and which is differentiable at $\theta = \theta_o \in \Theta$ with the matrix $\int \nabla h_\theta \cdot \nabla h_\theta^T \, d\nu$ being non-singular. Let $P^n = \prod_1^n P_{\theta_o}$, and L_n be as in the previous proposition, and assume

$$(4.5) \qquad \lim_{n\to\infty} \sup\{|\theta_o-\tilde{\theta}_{ni}|: 1 \leq i \leq n\} = 0$$

and

$$(4.6) \qquad \lim_{n\to\infty} \sum_1^n \|(\tilde{\theta}_{ni}-\theta_o)^T \cdot \nabla h_{\theta_o}\|^2 = \sigma^2/4 > 0.$$

Then

$$(4.7) \qquad \mathcal{L}(L_n | P^n) \xrightarrow{\mathcal{D}} \mathcal{N}(-\sigma^2/2, \sigma^2).$$

Proof: By (4.5), (4.6), and the non-singularity of $\int (\nabla h_\theta \cdot \nabla h_\theta^T)\, d\nu$,

$$\limsup_{n\to\infty} \sum_1^n \left| \| (\tilde{\theta}_{ni}-\theta_o)^T \cdot \nabla h_{\theta_o} \|^2 - H^2(\theta_o, \tilde{\theta}_{ni}) \right|$$

$$= \limsup_{n\to\infty} \sum_1^n |\tilde{\theta}_{ni}-\theta_o|^2 \left| \| (\tilde{\theta}_{ni}-\theta_o)^T \cdot \nabla h_{\theta_o} \|^2 - H^2(\theta_o, \tilde{\theta}_{ni}) \right| / |\tilde{\theta}_{ni}-\theta_o|^2$$

$$= 0$$

and thus (4.3) is true. Also by (4.5), the proof of Theorem 2.2, and (2.15), (4.4) is true. Thus Proposition 4.1 implies the desired result. □.

<u>Remark.</u> If the hypothesis in Theorem 4.1 is strengthened to include regularity of E and local convexity of Θ, then θ_o can be replaced by $\{\theta_{ni}\}$ where $\{\theta_{ni}\} \subset \Theta_c \subset \Theta$, with Θ_c being compact, to get a stronger result implying a kind of uniform asymptotic normality. The proof is exactly analogous to the one just given except for the use of the proof of Theorem 2.3 in place of the proof of Theorem 2.2.

4.2 Necessary Conditions for Consistency

By using Theorem 3.2 (or Theorem 3.3) we can prove a necessary condition for consistency of estimation in a model which generalizes that of the nonlinear least squares given in Wu (1981) which was concerned only with translation. The result is stated precisely below but the proof is omitted since it is similar to that given in Wu.

Theorem 4.2 Let $(\Omega,\mathcal{F},P)$ be a probability measure space, Θ be an arbitrary parameter space, and $\epsilon_1,\epsilon_2,\ldots$ be i.i.d. k-dimensional random vectors. Assume that $Y_1,Y_2,\ldots$ are k-dimensional random vectors given by

$$Y_i = f_i(\theta)\epsilon_i \qquad \text{for} \qquad i \in \mathbb{N}$$

where $f_i:\Theta \to \{RT_t: t \in \mathbb{R}^k, R \in \mathfrak{R}\}$. Suppose the experiment $E = (\mathbb{R}^k,\mathcal{B}(\mathbb{R}^k),\{P_{(t,R)}: (t,R) \in \mathbb{R}^k\times\mathfrak{R}\})$ satisfies (A4), $P \equiv \lambda$, and E is differentiable at $(0,I)$. If there exists an estimate $\hat{\theta}_n(Y_1,\ldots,Y_n)$ such that $\hat{\theta}_n(Y_1,\ldots,Y_n) \to \theta$ in P_θ-probability for $\theta \in \Theta$, then

$$(4.8) \qquad \sum_1^\infty |t_i(\theta)-t_i(\theta')|^2+|R_i(\theta)-R_i(\theta')|^2 = \infty$$

for all $\theta' \neq \theta$ where $f_i(\theta) = R_i(\theta)T_{t_i(\theta)}$ and $f_i(\theta') = R_i(\theta')T_{t_i(\theta')}$.

Remark. Theorem 4.2 could be extended to allowing $f_i: \Theta \to \{AT_t: A \in \mathcal{A}, t \in \mathbb{R}^k\}$ again assuming the corresponding experiment satisfies the conditions stated above and $\{\det(A_i(\theta)): i \in \mathbb{N}\}$ is bounded away from 0 and is bounded above. In this case the necessary condition for consistency is that for all $\theta' \neq \theta$, an analogous condition to (4.8) holds.

APPENDIX

Proposition A.1. Let $(\Omega, \mathcal{F}, \nu)$ be a σ-finite measure space and let $\{f_\alpha: \alpha \in I\} \subset L^1(\nu)$. Then $\{f_\alpha: \alpha \in I\}$ is u.i. if and only if for all sequences $\{\alpha_n\} \subset I$ there exists a subsequence $\{\alpha_{n'}\}$ such that $\{f_{\alpha_{n'}}\}$ is u.i..

Proof: The "if" part is immediate by the definition of uniform integrability. For the "only if" part suppose $\{f_\alpha: \alpha \in I\}$ is not u.i.. Let $g \in L^1(\nu)$ be such that $g > 0$ everywhere. The existence of such a g is guaranteed by the σ-finiteness of ν. Now there exists $\epsilon > 0$ such that for each $n \in \mathbb{N}$ there is an $\alpha_n \in I$ such that

$$\text{(A.1)} \quad \int (|f_n|-ng)_+ \, d\nu > \epsilon.$$

Let $h \in L^1(\nu)$. Then by the Lebesgue dominated convergence theorem, $\int (h-ng)_+ \, d\nu \to 0$. On combining this with (A1),

$$\text{(A.2)} \quad \liminf_{n\to\infty} \int (|f_n|-h)_+ \, d\nu > \epsilon.$$

Since h was arbitrary, this implies there does not exists a subsequence $\alpha_{n'}$ such that $\{f_{\alpha_{n'}}\}$ is u.i.. □.

Proposition A.2. Let $E_n = (\Omega_n, \mathcal{F}_n, \{P^n, \tilde{P}^n\})$ and $E'_n = (\Omega'_n, \mathcal{F}'_n, Q^n)$ be two sequences of experiments. Then $\tilde{P}^n \triangleleft P^n$ ($\tilde{P}^n \blacktriangle P^n$) if and only if $\tilde{P}^n \times Q^n \triangleleft P^n \times Q^n$ ($\tilde{P}^n \times Q^n \blacktriangle P^n \times Q^n$).

Proof: Let $P^n(F_n) \to 0$. Then $P^n \times Q^n(F_n \times \Omega'_n) \to 0$ which in turn implies $\tilde{P}^n \times Q^n(F_n \times \Omega'_n) = \tilde{P}^n(F_n) \to 0$.

For the converse let $P^n \times Q^n(A_n) \to 0$ where $A_n \in \mathcal{F}_n \times \mathcal{F}'_n$. Then let $A_n|_{\omega'} = \{\omega \in \Omega_n : (\omega, \omega') \in A_n\}$ and by Fubini

$$P^n \times Q^n(A_n) = \int P^n(A_n|_{\omega'})\, dQ^n(\omega').$$

Thus $P^n(A_n|.) \to 0$ in Q^n-probability. Let $\epsilon > 0$. Then there exists $\rho > 0$ and $N \in \mathbb{N}$ such that $P^n(F_n) < \rho$ and $n \geq N$ implies $\tilde{P}^n(F_n) < \epsilon$. Thus

$$\begin{aligned}\limsup_{n\to\infty} \tilde{P}^n \times Q^n(A_n) &\leq \limsup_{n\to\infty} \int 1_{\{P^n(A_n|.)\geq\rho\}} \tilde{P}^n(A_n|.)\, dQ^n \\ &\quad + \limsup_{n\to\infty} \int 1_{\{P^n(A_n|.)<\rho\}} \tilde{P}^n(A_n|.)\, dQ^n \\ &< \epsilon.\end{aligned}$$

The proof of the asymptotic separability result is similar. □.

Proposition A.3. Let $(\Omega_n, \mathcal{F}_n, \{P^n, \tilde{P}^n\})$ be a sequence of experiments and for each n, let $\mathcal{F}_n^o$ be a field generating $\mathcal{F}_n$. Then $\tilde{P}^n \triangleleft P^n$ ($\tilde{P}^n \blacktriangle P^n$) if and only if $\tilde{P}^n$ is contiguous (asymptotically separated) to P^n on $\mathcal{F}_n^o$ (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 $\mathcal{F}_n^o$ to $\mathcal{F}_n$ it is possible to simultaneously approximate with respect to P^n and $\tilde{P}^n$ (cf. Billingsley (1979), Theorem 11.4). Hence for any sequence $\{A_n\}$ there exists a sequence $\{A_n^o\}$ such that $A_n^o \in \mathcal{F}_n^o$ and

$$\max\{P^n(A_n \Delta A_n^o), \tilde{P}^n(A_n \Delta A_n^o)\} \leq 1/n \quad \text{for all} \quad n,$$

where Δ denotes the symmetric difference. The result follows easily. □.

<u>Proposition A.4.</u> Let $E = (\mathbb{R}^k, \mathcal{B}(\mathbb{R}^k), \{P_{ni}, \tilde{P}_{ni}, Q\colon n, i \in \mathbb{N}\})$ be an experiment. If $\prod_1^\infty P_{ni} * Q \blacktriangle \prod_1^\infty \tilde{P}_{ni} * Q$, then $\prod_1^\infty P_{ni} \blacktriangle \prod_1^\infty \tilde{P}_{ni}$.

Proof: By Proposition A.3, it suffices to show that $\prod_1^n P_{ni} * Q \blacktriangle \prod_1^n \tilde{P}_{ni} * Q$ implies $\prod_1^n P_{ni} \blacktriangle \prod_1^n \tilde{P}_{ni}$. Let $\{n'\}$ be a subsequence and $\{F_{n'}\}$ be such that $F_{n'} \in \mathcal{B}(\mathbb{R}^{n'k})$, $\prod_1^{n'} P_{n'i} * Q(F_{n'}) \to 1$, and $\prod_1^{n'} \tilde{P}_{n'i} * Q(F_{n'}) \to 0$. Then

$$(A.1) \qquad \iint 1_{F_{n'}}(x+y)\, d\prod_1^{n'} Q(y)\, d\prod_1^{n'} P_{n'i}(x) \to 1$$

and

$$(A.2) \qquad \iint 1_{F_{n'}}(x+y)\, d\prod_1^{n'} Q(y)\, d\prod_1^{\tilde{n}'} P_{n'i}(x) \to 0.$$

Let $f_{n'}: \mathbb{R}^{n'k} \to \mathbb{R}$ be defined by

$$f_{n'}(x) = \int 1_{F_{n'}}(x+y)\, d\prod_1^{n'} Q(y).$$

By (A.1) and (A.2), $\prod_1^{n'} P_{n'i}(f_{n'} \geq .5) \to 1$ and $\prod_1^{\tilde{n}'} P_{n'i}(f_{n'} \geq .5) \to 0$. The result now follows. □.

REFERENCES

1. Auslander, L. (1967). Differential Geometry, Harper & Row, New York.

2. Auslander, L. and MacKenzie, R. (1963). Introduction to Differentiable Manifolds, McGraw-Hill, New York.

3. Bauer, H. (1981). Probability Theory and Elements of Measure Theory, Academic Press, New York.

4. Billingsley, P. (1979). Probability and Measure, Wiley, New York.

5. Eagleson, G.K. (1981). An extended dichotomy theorem for sequence of pairs of Gaussian measures, The Annals of Probability, 9, 453-459.

6. Fabian, V. and Hannan, J. (1985). Introduction to Probability and Mathematical Statistics, Wiley, New York.

7. Feldman, J. (1958). Equivalence and perpendicularity of Gaussian processes, Pacific Journal of Mathematics, 9, 699-708.

8. Hajek, J. (1958). On a property of normal distribution of any stochastic processes, Mathematical Statistics and Probability, 1, 245-252.

9. Hajek, J. and Sidak, Z. (1967). Theory of Rank Tests, Academic Press, New York.

10. Kakutani, S. (1948). On equivalence of infinite product measures, Annals of Mathematics, 49, 214-224.

11. LeCam, L. (1970). On the assumptions used to prove asymptotic normality of maximum likelihood estimates, The Annals of Statistics, 41, 802-828.

12. Lipster, R. Sh., Pukel'sheim, F., and Shiryaev, A.N. (1982). Necessary and sufficient conditions for contiguity and entire asymptotic separation of probability measures, Russian Mathematical Surveys, 37:6, 107-136.

13. Loomis, L. and Sternberg, S. (1968). Advanced Calculus, Addison-Wesley, Reading, Massachusetts.

14. Neveu, J. (1965). Mathematical Foundations of the Calculus of Probability, Holden-Day, San Francisco.

15. Oosterhoff, J. and Van Zwet, W.R. (1979). A note on contiguity and Hellinger distance, Contribution to Statistics, Reidel, Dorecht, 157-166.

16. Roussas, G. (1972). Contiguity of Probability Measures: Some Applications in Statistics, Cambridge University Press, Cambridge.

17. Shepp, L. A. (1965). Distinguishing a sequence of random variables from a translate of itself, The Annals of Mathematical Statistics, 36, 1107-1112.

18. Steele, J. M. (1986). Fisher information and detection of a Euclidean perturbation of an independent stationary process, The Annals of Probability, 14, 326-335.

19. Strasser, H. (1985). Mathematical Theory of Statistics, Walter de Gruyter, New York.

20. Warner, F. W. (1983). Foundations of Differentiable Manifolds and Lie Groups, Springer-Verlag, New York.

21. Wu, C. F. (1981). Asymptotic theory of nonlinear least squares estimation, The Annals of Mathematical Statistics, 9, 501-513.