The Annals of Mathematical Statistics
1970, Vol. 41, No. 3, 918–938

ASYMPTOTICALLY OPTIMAL TESTS IN MARKOV PROCESSES

BY RICHARD A. JOHNSON[1] AND GEORGE G. ROUSSAS[2]

University of Wisconsin

1. Introduction and summary. Let $(\mathscr{X}, \mathscr{A})$ be a measurable space and let Θ be an open subset of the k-dimensional Euclidean space $\mathscr{E}_k$. For each $\theta \in \Theta$, let P_θ be a probability measure on $\mathscr{A}$. Let $\{X_n, n \geqq 0\}$ be a discrete parameter Markov process defined on $(\mathscr{X}, \mathscr{A}, P_\theta)$, X_n taking values in the Borel real line $(R, \mathscr{B})$. Finally, let $\mathscr{A}_n$ be the σ-field induced by the first $n+1$ random variables $X_0, X_1, \cdots, X_n$ from the process and let $P_{n,\theta}$ be the restriction of P_θ to the σ-field $\mathscr{A}_n$.

Under suitable conditions on the process, the following results are derived. Let θ_0 be an arbitrary but fixed point in Θ and let $\Delta_n(\theta_0)$ be a k-dimensional vector defined in terms of the random variables $X_0, X_1, \cdots, X_n$; $\Delta_n{}^*(\theta_0)$ stands for a certain truncated version of $\Delta_n(\theta_0)$. By means of $\Delta_n{}^*(\theta_0)$ and $h \in \mathscr{E}_k$, one defines a probability measure $R_{n,h}$, $n \geqq 0$. The first main result is that the sequences $\{P_{n,\theta}\}$ and $\{R_{n,h}\}$ of probability measures with $h = n^{\frac{1}{2}}(\theta - \theta_0)$, $\theta \in \Theta$, are *differentially equivalent* at the point θ_0. (See Definition 5.1.) This is shown in Corollary 5.1. It is also proved in Corollary 5.2 that the sequence $\{\Delta_n{}^*(\theta_0)\}$ is *differentially sufficient* at θ_0 (see Definition 5.2) for the family $\{P_{n,\theta}; \theta \in \Theta\}$ of probability measures. Next, let $\{h_n\}$ be a bounded sequence of h's in $\mathscr{E}_k$ and set $\theta_n = \theta_0 + h_n n^{-\frac{1}{2}}$. Then for hypotheses testing problems, Theorem 6.1 allows one to restrict oneself to the class of tests depending on $\Delta_n(\theta_0)$ alone, at least as far as the asymptotic power of the test under alternatives of the form P_{n,θ_n} is concerned.

In Section 7, these results are applied to the case of testing hypotheses about a real-valued parameter. More specifically, asymptotically most powerful tests for testing the hypothesis $\theta = \theta_0$ against one-sided alternatives are constructed. This is covered in Theorem 7.1.1. Also an asymptotically most powerful unbiased test for testing the same hypothesis as above against two-sided alternatives is constructed in Theorem 7.1.2.

The first of these problems was also dealt with in Johnson and Roussas [2] but the approach is different here. The second problem is solved in Wald [8] for the independent identically distributed case. However, both the assumptions and approach are different here in addition to the Markovian character of the random variables involved.

Section 6 treats the general situation where Θ is an open subset of $\mathscr{E}_k$. Theorem 6.1 together with Theorem 6.3 provide a way for studying the corresponding

Received March 6, 1969; revised November 3, 1969.

[1] This author's research was partially supported by the Air Force Office of Scientific Research through Grant AFOSR-69-1789.

[2] The research of this author was supported by the National Science Foundation under Grant GP-11890.

hypothesis testing problem in the k-dimensional parameter case. Finally, at the end of the last section, an outline is presented of forthcoming results for that case.

These results extend, under substantially weaker conditions, the work of Wald [8], [9] to Markov processes. The method of proof relies heavily on the development in LeCam [3].

Unless otherwise stated, limits will be taken as the sequence $\{n\}$, or subsequences thereof, tends to ∞. Integrals without limits will extend over the entire appropriate space. For $h \in \mathscr{E}_k$, h' stands for its transpose. All bounding constants will be finite numbers.

2. Notation and assumptions. Let Θ be an open subset of the k-dimensional Euclidean space $\mathscr{E}_k$ and for each $\theta \in \Theta$, let $\{X_n, n \geqq 0\}$ be a discrete time parameter real-valued Markov process defined on the probability space $(\mathscr{X}, \mathscr{A}, P_\theta)$ with initial and transition distributions $p_\theta(\cdot)$ and $p_\theta(\cdot, \cdot)$, respectively. It will be assumed in the following that the probability measures $\{P_{n,\theta}, \theta \in \Theta\}$, $n \geqq 0$ are mutually absolutely continuous, where $P_{n,\theta}$ stands for the restriction of P_θ to the σ-field $\mathscr{A}_n$ induced by the random variables $X_0, X_1, \cdots, X_n$. Therefore for any $\theta, \theta' \in \Theta$, we will have $[dP_{0,\theta'}/dP_{0,\theta}] = q(X_0; \theta, \theta')$, $[dP_{1,\theta'}/dP_{1,\theta}] = q(X_0, X_1; \theta, \theta')$. Furthermore, let $q(X_1 \mid X_0; \theta, \theta') = q(X_0, X_1; \theta, \theta')/q(X_0; \theta, \theta')$, $\varphi_j(\theta, \theta') = [q(X_j \mid X_{j-1}; \theta, \theta')]^{\frac{1}{2}}$ and $f_j(\theta, \theta') = [q(X_{j-1}, X_j; \theta, \theta')]^{\frac{1}{2}}$, $j = 1, \cdots, n$, so that $\int \varphi_1{}^2(\theta, \theta')\, dP_{1,\theta} = 1$. There is no loss in generality by assuming that Θ contains the origin in the $\mathscr{E}_k$ space and we will do so.

The results in this paper are derived under the following set of assumptions and an additional one presented in Section 7.

ASSUMPTION 1. For each $\theta \in \Theta$, the Markov process $\{X_n, n \geqq 0\}$ is stationary and metrically transitive (ergodic). (See, e.g., Doob [1] page 457.)

ASSUMPTION 2. The probability measures $\{P_{n,\theta}, \theta \in \Theta\}$, $n \geqq 0$ are mutually absolutely continuous.

ASSUMPTION 3. (i) For each $\theta \in \Theta$, the random function $\varphi_1(\theta, \theta')$ is differentiable in quadratic mean (q.m.) with respect to θ' at (θ, θ) when P_θ is employed.

Let $\dot{\varphi}_1(\theta)$ be the derivative of $\varphi_1(\theta, \theta')$ with respect to θ' at (θ, θ). Then,

(ii) $\dot{\varphi}_1(\theta)$ is $\mathscr{A}_2 \times \mathscr{C}$-measurable, where $\mathscr{C}$ denotes the σ-field of Borel subsets of Θ.

Let $\Gamma(\theta)$ be the covariance function defined by $\Gamma(\theta) = 4\mathscr{E}_\theta[\dot{\varphi}_1(\theta)\, \dot{\varphi}_1{}'(\theta)]$. Then,

(iii) $\Gamma(\theta)$ is positive definite for every $\theta \in \Theta$.

ASSUMPTION 4. For every $\theta \in \Theta$, the random function $f_1(\theta, \theta')$ is continuous in $P_{1,\theta}$-probability at (θ, θ).

For an arbitrary but fixed $\theta \in \Theta$, we will be interested in sequences $\theta_n \to \theta$. Then, from Assumption 2 it follows that $[dP_{n,\theta_n}/dP_{n,\theta}] = q(X_0; \theta, \theta_n)\prod_{j=1}^n \varphi_j{}^2(\theta, \theta_n)$ is well defined outside a P_θ-null set for all $\theta \in \Theta$. Therefore outside this set we can define the random variable $\Lambda_n[P_{n,\theta_n}; P_{n,\theta}]$ as follows

$$\Lambda[P_{n,\theta_n}; P_{n,\theta}] = \log [dP_{n,\theta_n}/dP_{n,\theta}] = \log [q(X_0; \theta, \theta_n)\prod_{j=1}^n \varphi_j^2(\theta, \theta_n)].$$

In the sequel the probability measures $P_{n,\theta}$ and P_θ will be used interchangeably.

3. Preliminary results. For $h_n, h \in \mathscr{E}_k$ the following theorem holds true.

THEOREM 3.1. *Under Assumptions* A1, A2, A3(i), A3(ii) *and* A4, *we have*

(i) $\mathscr{L}\{\Lambda[P_{n,\theta_n}; P_{n,\theta}] \mid P_\theta\} \to N(-\frac{1}{2}h'\Gamma(\theta)h, h'\Gamma(\theta)h)$.

(ii) $\mathscr{L}\{\Lambda[P_{n,\theta_n}; P_{n,\theta}] \mid P_{\theta_n}\} \to N(\frac{1}{2}h'\Gamma(\theta)h, h'\Gamma(\theta)h)$,

where $\theta_n = \theta + h_n n^{-\frac{1}{2}}$ *and* $h_n \to h$.

The first conclusion of this theorem is Theorem 3.2.1 in Roussas [6] and the second is Theorem 2 in Roussas [7].

The k-dimensional random vector $\Delta_n(\theta)$ is defined by

$$\Delta_n(\theta) = 2n^{-\frac{1}{2}}\sum_{j=1}^n \dot{\varphi}_j(\theta). \tag{3.1}$$

The function $\Delta_n(\theta)$ plays a central role in the development.

The following theorems are also true.

THEOREM 3.2. *Under the same assumptions as those in Theorem* 3.1, *we have*

$$\mathscr{L}[\Delta_n(\theta) \mid P_\theta] \to N(0, \Gamma(\theta)).$$

This is Theorem 3.2.1 in Roussas [6].

THEOREM 3.3. *Under the same assumptions as those in Theorem* 3.1, *we have*

$$\Lambda[P_{n,\theta_n}; P_{n,\theta}] - h'\Delta_n(\theta) \to -A(h, \theta) \quad \text{in} \quad P_{n\theta}\text{-probability},$$

where $A(h, \theta) = \frac{1}{2}h'\Gamma(\theta)h$.

This is Theorem 3.1.1 in Roussas [6].

Heuristically, this last theorem states that $\exp[-A(h, \theta) + h'\Delta_n(\theta)]\, dP_{n,\theta}$ approximates dP_{n,θ_n}. However, the integral of the approximating measure may not even be finite. The next few sections are devoted to the construction of an exponential family which does approximate dP_{n,θ_n}.

Now we recall from LeCam [3] two of several equivalent definitions of the concept of contiguity which will be used in this paper.

DEFINITION 3.1. Two sequences $\{P_n\}$ and $\{Q_n\}$ of probability measures defined on $(\mathscr{X}, \mathscr{A})$ are said to be *contiguous* if for any sequence $\{T_n\}$ of random variables we have $T_n \to 0$ in P_n-probability if and only if $T_n \to 0$ in Q_n-probability.

DEFINITION 3.2. Let $\{P_n\}$ be as in Definition 3.1 and $\{T_n\}$ be a sequence of random variables. Then we say that the sequence $\{\mathscr{L}(T_n \mid P_n)\}$ is *relatively compact* if for every $\{n'\} \subset \{n\}$ there is a further $\{m\} \subset \{n'\}$ such that $\{\mathscr{L}(T_m \mid P_m)\}$ converges (in the weak sense) to a probability distribution.

Now assume that χ_n defined by $\chi_n = \Lambda[Q_n; P_n]$ exists. Then

DEFINITION 3.3. The sequences $\{P_n\}$ and $\{Q_n\}$ are said to be *contiguous* if $\{\mathscr{L}(\chi_n \mid P_n)\}$ is relatively compact. Furthermore, if $\{m\} \subset \{n\}$ is such that $\{\mathscr{L}(\chi_m \mid P_m)\}$ converges to a probability distribution $\mathscr{L}(\chi)$, say, then $\mathscr{E}(\exp \chi) = 1$, where the expectation is calculated under $\mathscr{L}(\chi)$.

The following result follows from Theorem 3.1(i) and Theorem 2.1(6) in LeCam [3] and is isolated here for easy reference.

PROPOSITION 3.1. *Under the same assumptions as those in Theorem* 3.1, *we have that the sequences* $\{P_{n,\theta_n}\}$ *and* $\{P_{n,\theta}\}$ *are contiguous.*

PROOF. By Theorem 3.1(i), $\{\mathscr{L}\{\Lambda[P_{n,\theta_n}; P_{n,\theta}] \mid P_\theta\}\}$ is, trivially, relatively compact and the limit law $\mathscr{L}(\chi)$ is $N(-\frac{1}{2}h'\Gamma(\theta)h, h'\Gamma(\theta)h)$. Furthermore it is easily seen that $\mathscr{E}(\exp \chi) = 1$ and Theorem 2.1(6) in LeCam [3] completes the proof.

4. Some lemmas. In this section, some lemmas are collected for use in later sections of the paper. Let W be a bounded convex symmetric neighborhood of θ_0 in $\mathscr{E}_k$. Without loss of generality—and with conceptual advantages—we may assume that W is the closed sphere of radius 1 centered at θ_0. On $\mathscr{E}_k$, define the real-valued function ρ by $\rho(t) = \sup\{h't; h \in W\}$. Then $\rho(t) = \theta_0't + \sup\{(h-\theta_0)'t; h \in W\} = \theta_0't + \|t\|$, where $\|\cdot\|$ is the usual norm in $\mathscr{E}_k$. Consequently, there is no restriction if it is assumed that $\theta_0 = 0$ when defining the function ρ and we will do so. Thus, it will be assumed that W is the unit sphere centered at the origin with the function ρ defined on $\mathscr{E}_k$ by

$$\rho(t) = \|t\|. \tag{4.1}$$

It follows that ρ is continuous and therefore Theorem 3.2 implies that

$$\mathscr{L}[\rho(\Delta_n) \mid P_0] \to \mathscr{L}[\rho(\Delta) \mid P_0], \tag{4.2}$$

where Δ is a k-dimensional random vector such that $\mathscr{L}(\Delta \mid P_0) = N(0, \Gamma)$, $\Delta_n = \Delta_n(0)$, and Γ denotes $\Gamma(0)$.

Now we are in a position to proceed with the lemmas.

LEMMA 4.1. *For every* $\lambda > 0$, *we have* $\mathscr{E}_0[\exp \lambda\, \rho(\Delta)] < \infty$.

PROOF. Let $\Delta = (\Delta_1, \cdots, \Delta_k)'$. Then $\|\Delta\| \leqq \sum_{j=1}^k |\Delta_j|$ and therefore

$$\begin{aligned} \mathscr{E}_0[\exp \lambda\rho(\Delta)] &= \mathscr{E}_0[\exp \lambda\|\Delta\|] \leqq \mathscr{E}_0[\exp(\lambda \textstyle\sum_{j=1}^k |\Delta_j|)] \\ &= \mathscr{E}_0(\textstyle\prod_{j=1}^k \exp \lambda|\Delta_j|) \leqq \textstyle\prod_{j=1}^k \mathscr{E}_0^{1/k}(\exp k\lambda|\Delta_j|) \end{aligned}$$

by a generalized version of the Hölder inequality. But $\mathscr{E}_0(\exp k\lambda|\Delta_j|) < \infty$ since Δ_j is normal, $j = 1, \cdots, k$. This completes the proof of the lemma.

Consider the random vectors Δ_n, Δ and for each $\alpha > 0$, define the following truncated random vectors Δ_n^α, Δ^α.

$$\begin{aligned} \Delta_n^\alpha &= \Delta_n, \quad \text{if } \rho(\Delta_n) < \alpha; \qquad & \Delta^\alpha &= \Delta, \quad \text{if } \rho(\Delta) < \alpha; \\ &= 0, \quad \text{otherwise} & & 0, \quad \text{otherwise.} \end{aligned} \tag{4.3}$$

Then the following lemma is true.

LEMMA 4.2. *For every* $\alpha > 0$, *we have* $\mathscr{L}(\Delta_n^\alpha \mid P_0) \to \mathscr{L}(\Delta^\alpha \mid P_0)$.

PROOF. For every bounded, continuous, real-valued function f on $\mathscr{E}_k$, it is easily seen that $\int f(t)\, d\mathscr{L}(\Delta_n^\alpha \mid P_0) \to \int f(t)\, d\mathscr{L}(\Delta^\alpha \mid P_0)$, since $\mathscr{L}(\Delta_n \mid P_0) \to \mathscr{L}(\Delta \mid P_0)$ and the set of discontinuities of $f\{tI_{[\rho(t)<\alpha]}(t)\}$ is a subset of the surface of the sphere of radius α, centered at the origin, which is assigned (k-dimensional) Lebesgue measure zero and hence $\mathscr{L}(\Delta \mid P_0)$ measure zero.

Lemma 4.2 and the continuity of ρ also imply that

$$\mathscr{L}[\rho(\Delta_n^\alpha) \mid P_0] \to \mathscr{L}[\rho(\Delta^\alpha) \mid P_0] \quad \text{for every} \quad \alpha > 0. \tag{4.4}$$

By the definition of ρ and Δ^α, Δ_n^α, it follows that $\rho(\Delta^\alpha) \leqq \alpha$, $\rho(\Delta_n^\alpha) \leqq \alpha$ for every $\alpha > 0$ and all n. Thus whereas $\mathscr{E}_0[\exp \lambda\rho(\Delta_n)]$ need not be finite, the expectations $\mathscr{E}_0[\exp \lambda\rho(\Delta^\alpha)]$ and $\mathscr{E}_0[\exp \lambda\rho(\Delta_n^\alpha)]$ are finite for every $\alpha, \lambda > 0$ and all n. What is more, the following lemma holds true.

LEMMA 4.3. *For every* $\alpha, \lambda > 0$, *we have*

$$\mathscr{E}_0[\exp \lambda\rho(\Delta_n^\alpha)] \to \mathscr{E}_0[\exp \lambda\rho(\Delta^\alpha)].$$

PROOF. Since the distributions $\mathscr{L}[\rho(\Delta^\alpha) \mid P_0]$ and $\mathscr{L}[\rho(\Delta_n^\alpha) \mid P_0]$ have support confined to the interval $[0, \alpha]$, the result follows from (4.4) and the Helly–Bray lemma.

LEMMA 4.4. *For every* $\lambda > 0$, *we have*

$$\mathscr{E}_0[\exp \lambda\rho(\Delta^\alpha)] \to \mathscr{E}_0[\exp \lambda\rho(\Delta)] \qquad \textit{as} \quad \alpha \to \infty.$$

PROOF. We have

$$\mathscr{E}_0[\exp \lambda\rho(\Delta^\alpha)] = \mathscr{E}_0\{\exp \lambda\rho(\Delta) I_{[\rho(\Delta)<\alpha]}(\Delta)\}.$$

Now $\exp\{\lambda\rho(\Delta) I_{[\rho(\Delta)<\alpha]}(\Delta)\}$ is bounded by $\exp \lambda\rho(\Delta)$ (independent of α) which is P_0-integrable. Also, since $\exp\{\lambda\rho(\Delta) I_{[\rho(\Delta)<\alpha]}(\Delta)\} \to \exp[\lambda\rho(\Delta)]$ as $\alpha \to \infty$, the Dominated convergence theorem applies and gives the result.

Now let $\{\alpha_\nu\}$, $\{\varepsilon_\nu\}$ be two sequences such that $0 < \alpha_\nu \uparrow \infty$ and $0 < \varepsilon_\nu \downarrow 0$ as $\nu \to \infty$. Let $\lambda > 0$ be specified later. For every α_ν, ε_ν and λ as above, there exists a positive integer $N_\nu = N(\alpha_\nu, \varepsilon_\nu, \lambda)$ such that

$$|\mathscr{E}_0[\exp \lambda\rho(\Delta_n^{\alpha_\nu})] - \mathscr{E}_0[\exp \lambda\rho(\Delta^{\alpha_\nu})]| \leqq \varepsilon_\nu, \qquad n \geqq N_\nu, \quad \nu = 1, 2, \cdots. \tag{4.5}$$

This is possible according to Lemma 4.3. Clearly $\{N_\nu\}$ can be chosen so that $N_\nu \uparrow \infty$. Then, we define a sequence $\{\Delta_n^*\}$ of k-dimensional random vectors by

$$\Delta_n^* = \Delta_n^{\alpha_\nu} \quad \text{for} \quad n \quad \text{such that} \quad N_\nu \leqq n < N_{\nu+1}, \qquad \nu = 1, 2, \cdots. \tag{4.6}$$

Clearly,

$$|\mathscr{E}_0[\exp \lambda\rho(\Delta_n^*)] - \mathscr{E}_0[\exp \lambda\rho(\Delta^{\alpha_\nu})]| \leqq \varepsilon_\nu, \qquad N_\nu \leqq n < N_{\nu+1}, \nu = 1, 2, \cdots.$$

by (4.5). Thus

$$\mathscr{E}_0[\exp \lambda\rho(\Delta_n^*)] - \mathscr{E}_0[\exp \lambda\rho(\Delta^{\alpha_\nu})] \to 0 \quad \text{as} \quad n \to \infty (\nu \to \infty).$$

But $\mathscr{E}_0[\exp \lambda\rho(\Delta^{\alpha_\nu})] \to \mathscr{E}_0[\exp \lambda\rho(\Delta)]$ as $\nu \to \infty$ by Lemma 4.4. Therefore $\mathscr{E}_0[\exp \lambda\rho(\Delta_n^*)] \to \mathscr{E}_0[\exp \lambda\rho(\Delta)]$.

The following result has been established.

LEMMA 4.5. *Let the sequence* $\{\Delta_n^*\}$ *be defined by* (4.6). *For every* $\lambda > 0$, *we have*

$$\mathscr{E}_0[\exp \lambda\rho(\Delta_n^*)] \to \mathscr{E}_0[\exp \lambda\rho(\Delta)].$$

From the definition of Δ_n^*, we have $\Delta_n^* = \Delta_n^{\alpha_\nu} = \Delta_n I_{[\rho(\Delta_n)<\alpha_\nu]}(\Delta_n)$ for n such that $N_\nu \leqq n < N_{\nu+1}$. Thus $(\Delta_n^* \neq \Delta_n) = [\rho(\Delta_n) \geqq \alpha_\nu]$ and

$$P_0(\Delta_n^* \neq \Delta_n) = P_0[\rho(\Delta_n) \geqq \alpha_\nu]. \tag{4.7}$$

But $\mathscr{L}[\rho(\Delta_n) \mid P_0] \to \mathscr{L}[\rho(\Delta) \mid P_0]$ by (4.2) and $\alpha_\nu \to \infty$ as $\nu \to \infty$ (which implies $n \to \infty$). Therefore, taking the limits in (4.7) as $\nu \to \infty$, we get

$$P_0(\Delta_n^* \neq \Delta_n) \to 0. \tag{4.8}$$

Since $\mathscr{L}(\Delta_n \mid P_0) \to N(0, \Gamma)$ by Theorem 3.2, it follows from (4.8) that

$$\mathscr{L}(\Delta_n^* \mid P_0) \to N(0, \Gamma). \tag{4.9}$$

Now let $C_n = (\Delta_n^* \neq \Delta_n)$ and $Z_n = I_{C_n}$. Then, for $\varepsilon > 0$, we have $P_0(|Z_n| \geqq \varepsilon) = P_0(Z_n = 1) = P_0(C_n) \to 0$ by (4.8). That is, $Z_n \to 0$ in P_0-probability and this implies that $Z_n \to 0$ in P_{θ_n}-probability by Proposition 3.1. Here $\theta_n = h_n n^{-\frac{1}{2}}$, where $h_n, h \in \mathscr{E}_k$ and $h_n \to h$. Thus we have the following proposition.

PROPOSITION 4.1. *Let* Δ_n, Δ_n^* *be as above and set* $\theta_n = h_n n^{-\frac{1}{2}}$, *where* $h_n, h \in \mathscr{E}_k$ *and* $h_n \to h$. *Then*

$$P_0(\Delta_n^* \neq \Delta_n) \to 0 \quad \textit{and} \quad P_{\theta_n}(\Delta_n^* \neq \Delta_n) \to 0.$$

The following lemma will also be needed.

LEMMA 4.6. *Let* $\{Y_n\}$ *be a sequence of m-dimensional random vectors and let* $\{Z_n\}$ *be a sequence of random variables such that* $|Z_n| \leqq M$ *for all n. Let Q be a probability measure on* $\mathscr{A}$ *and let* $\mathscr{L}(Y_n \mid Q) \to \lambda^*$, *where* λ^* *is a probability measure on the m-dimensional Borel σ-field. Then if* $\{\mathscr{L}(Y_n, Z_n \mid Q)\}$ *converges (weakly), it converges to a probability measure on the* $(m+1)$*-dimensional Borel σ-field.*

PROOF. Denote by F_n the cdf of (Y_n, Z_n). Then the proof consists of showing that $\limsup \{1 - F_n(\alpha, \beta]\} \to 0$, where $F_n(\alpha, \beta]$ denotes the variation of F_n over the $(m+1)$-dimensional interval $(\alpha, \beta]$. However, this follows readily from our assumptions.

For simplicity, we set $\Lambda_n = \Lambda[P_{n,\theta_n}; P_{n,0}]$, and then recall that $\theta_n = h_n n^{-\frac{1}{2}}$, h_n, $h \in \mathscr{E}_k$, and $h_n \to h$. The following lemma will be needed in the sequel.

LEMMA 4.7. *Let* $\{Y_n\}$, $\{Z_n\}$ *be two sequences of random variables and let* Y, Z *be random variables such that* $\mathscr{L}(Y_n, Z_n \mid P_0) \to \mathscr{L}(Y, Z \mid P_0)$. *Assume that* $\Lambda_n - Y_n \to c$, *a constant, in* P_0*-probability. Then*

$$\mathscr{L}(Y_n, Z_n \mid P_{\theta_n}) \to \exp(y+c)\mathscr{L}(Y, Z \mid P_0),$$

where here and in the sequel the notation f$\mathscr{L}$ *stands for a probability measure which is absolutely continuous with respect to the probability measure* $\mathscr{L}$ *and has density* f.

PROOF. We have $\mathscr{L}(Y_n, Z_n \mid P_0) \to \mathscr{L}(Y, Z \mid P_0)$ if and only if $\mathscr{L}(c_1 Y_n + c_2 Z_n \mid P_0) \to \mathscr{L}(c_1 Y + c_2 Z \mid P_0)$ for every c_1, c_2 in R. Also $c_1 \Lambda_n - c_1 Y_n \to c_1 c$ in P_0-probability. Hence $\mathscr{L}(c_1 \Lambda_n + c_2 Z_n \mid P_0) \to \mathscr{L}[c_1 (Y+c) + c_2 Z \mid P_0]$ for every c_1, c_2 in R which is equivalent to $\mathscr{L}(\Lambda_n, Z_n \mid P_0) \to \mathscr{L}(Y+c, Z \mid P_0)$. By setting $Y+c = U$, we then have $\mathscr{L}(\Lambda_n, Z_n \mid P_0) \to \mathscr{L}(U, Z \mid P_0)$. Then Theorem 2.1 in LeCam [3] implies that

$$\mathscr{L}(\Lambda_n, Z_n \mid P_{\theta_n}) \to \mathscr{L}(U_0, Z_0 \mid P_0), \qquad \text{where}$$

$$\begin{aligned} dP(U_0 \leqq u, Z_0 \leqq z) &= \exp u \, dP(U \leqq u, Z \leqq z) \\ &= \exp u \, dP(Y \leqq u - c, Z \leqq z). \end{aligned}$$

The substitution $u - c = y$ completes the proof.

5. Differential equivalence of certain probability measures. In this section, $h \in \mathscr{E}_k$ and $\{h_n\}$ is taken to be a bounded sequence in $\mathscr{E}_k$. Then $\mathscr{E}_0(\exp h'\Delta_n^*)$ is finite as has been seen. Set

$$\exp B_n(h) = \mathscr{E}_0(\exp h'\Delta_n^*) \tag{5.1}$$

and define the probability measures $R_{n,h}$ on $\mathscr{A}_n$ as follows

$$R_{n,h}(A) = \exp[-B_n(h)] \int_A \exp(h'\Delta_n^*) \, dP_{n,0}, \qquad A \in \mathscr{A}_n. \tag{5.2}$$

Replacing h by h_n, we get R_{n,h_n} and then the probability measures P_{n,θ_n} and R_{n,h_n} are differentially equivalent. (See Definition 5.1.) More precisely, we have the following theorem.

THEOREM 5.1. *For any bounded sequence* $\{h_n\}$ *in* $\mathscr{E}_k$ *and* θ_n *defined by* $\theta_n = h_n n^{-\frac{1}{2}}$, *we have* $\|R_{n,h_n} - P_{n,\theta_n}\| \to 0$, *where* $\|\cdot\|$ *is the norm associated with convergence in variation*; *that is,*

$$\|R_{n,h_n} - P_{n,\theta_n}\| = 2 \sup[|R_{n,h_n}(A) - P_{n,\theta_n}(A)|; A \in \mathscr{A}_n].$$

PROOF. The proof is by contradiction. Assume that $\|R_{n,h_n} - P_{n,\theta_n}\| \nrightarrow 0$. Then there exists an $\varepsilon > 0$, $\{m\} \subseteq \{n\}$ and $\{h_m\} \subseteq \{h_n\}$ such that $\|R_{m,h_m} - P_{m,\theta_m}\| > 2\varepsilon$ for all m. Equivalently, $\sup\{|R_{m,h_m}(A) - P_{m,\theta_m}(A)|; A \in \mathscr{A}_m\} > \varepsilon$ for all m. Then there exists $A_m \in \mathscr{A}_m$ such that $|R_{m,h_m}(A_m) - P_{m,\theta_m}(A_m)| > \varepsilon$ for all m. Since $\{h_m\}$ is bounded, there exists $\{h_r\} \subseteq \{h_m\}$ such that $h_r \to h$, say. Thus we have $|R_{r,h_r}(A_r) - P_{r,\theta_r}(A_r)| > \varepsilon$ for all r and $h_r \to h$. By setting $Z_r = I_{A_r}$, the last relationship becomes

$$\left|\int Z_r \, dR_{r,h_r} - \int Z_r \, dP_{r,\theta_r}\right| > \varepsilon \quad \text{for all} \quad r. \tag{5.3}$$

The remaining part of the proof consists in arriving at a contradiction of (5.3). Since the process of doing so is rather long, it would be more suggestive to break the various auxiliary results into lemmas. Toward this end, we begin.

LEMMA 5.1. *There exists a subsequence* $\{s\}$ *of* $\{r\}$ *and a random variable* Z *with* $0 \leqq Z \leqq 1$ *such that*

$$\int Z_s\, dP_{s,\theta_s} \to \int Z \exp(h'\Delta + c)\, dP_0,$$

where $c = -A(h, 0)$, *to be shortened to* $A(h)$, *and* h *is the limit of* $\{h_s\}$.

PROOF. Let $\mu_r = \mathscr{L}(h_r'\Delta_r^*, Z_r \mid P_0)$. Then there exists $\{\mu_s\} \subseteq \{\mu_r\}$ such that $\mu_s \to \mu$. Lemma 4.6 applies with $m = 1$, $Y_s = h_s'\Delta_s^*$, $Z_s = I_{A_s}$, $Q = P_0$, and λ^* being the $N(0, h'\Gamma h)$ measure. Therefore μ is a probability measure and we let $\mu = \mathscr{L}(h'\Delta, Z \mid P_0)$, where $0 \leqq Z \leqq 1$. Then

$$\mathscr{L}(h_s'\Delta_s^*, Z_s \mid P_0) \to \mathscr{L}(h'\Delta, Z \mid P_0). \tag{5.4}$$

Now $\Lambda_s - h_s'\Delta_s \to c$ in P_0-probability by Theorem 3.3, while $h_s'\Delta_s - h_s'\Delta_s^* \to 0$ in P_0-probability by (4.8). Thus $\Lambda_s - h'\Delta_s^* \to c$ in P_0-probability and Lemma 4.7 applies and gives

$$\mathscr{L}(h_s'\Delta_s^*, Z_s \mid P_{\theta_s}) \to \exp(y + c)\mathscr{L}(h'\Delta, Z \mid P_0). \tag{5.5}$$

Next
$$\int Z_s\, dP_{\theta_s} = \int_{[0,1]} z\, d\mathscr{L}(Z_s \mid P_{\theta_s}) = \int_{[0,1]\times R} z\, d\mathscr{L}(h_s'\Delta_s^*, Z_s \mid P_{\theta_s}) \to \int_{[0,1]\times R} z \exp(y + c)\, d\mathscr{L}(h'\Delta, Z \mid P_0)$$

by (5.5) and this is equal to $\int Z \exp(h'\Delta + c)\, dP_0$. The proof is completed.

LEMMA 5.2. *Let* $\{s\}$ *be as in Lemma* 5.1. *Then*
(i) $\int (Z_s \exp h_s'\Delta_s^*)\, dP_0 \to \int (Z \exp h'\Delta)\, dP_0$, *and*
(ii) $\int \exp h_s'\Delta_s^*\, dP_0 \to \int \exp h'\Delta\, dP_0$.

PROOF. Since $\{h_n\}$ is bounded, one can choose $\lambda > 0$ sufficiently large so that $\lambda^{-1}\|h_n\| \leqq 1$ for all n. Then $\lambda^{-1} h_n \in W$ for all n, which implies that $\lambda^{-1} h_n' t \leqq \rho(t)$ for all n, and this is equivalent to

$$h_n' t \leqq \lambda\rho(t) \quad \text{for all} \quad n. \tag{5.6}$$

We may assume that this is the λ used in defining Δ_n^* in (4.6). Now let $B = \{t \in \mathscr{E}_k;\ \|t\| \leqq r\}$. Then since $\mathscr{E}_0[\exp \lambda\rho(\Delta)] < \infty$ by Lemma 4.1, one can choose r sufficiently large so that

$$\int_{B^c} \exp[\lambda\rho(t)]\, d\mathscr{L}(\Delta \mid P_0) < \varepsilon. \tag{5.7}$$

Let $\tau_n = \mathscr{L}(\Delta_n^*, Z_n \mid P_0)$. Then, by Lemma 4.6, there exists $\{\tau_{m'}\} \subseteq \{\tau_n\}$ such that $\tau_{m'} \to \tau$ and τ is a probability measure. Let $\tau = \mathscr{L}(\Delta, Z \mid P_0)$. Since $\tau_{m'} \to \tau$ implies that $\mathscr{L}(h'\Delta_{m'}^*, Z_{m'} \mid P_0) \to \mathscr{L}(h'\Delta, Z \mid P_0)$, we may assume that the sequence $\{s\}$ of Lemma 5.1 is a subsequence of $\{m'\}$. Thus we have $\tau_s \to \tau$. Also, since $\mathscr{L}(\Delta_s^* \mid P_0) \to \mathscr{L}(\Delta \mid P_0)$ and $I_B(t) \exp \lambda\rho(t)$ is bounded and its discontinuity set is a subset of the surface of B which is assigned Lebesgue measure zero and hence $\mathscr{L}(\Delta \mid P_0)$-measure zero, we have

$$\int_B \exp \lambda\rho(t)\, d\mathscr{L}(\Delta_s^* \mid P_0) \to \int_B \exp \lambda\rho(t)\, d\mathscr{L}(\Delta \mid P_0). \tag{5.8}$$

Now

$$\mathscr{E}_0[\exp \lambda\rho(\Delta_s^*)] = \int \exp \lambda\rho(t)\, d\mathscr{L}(\Delta_s^* \mid P_0) \to \int \exp \lambda\rho(t)\, d\mathscr{L}(\Delta \mid P_0)$$
$$= \mathscr{E}_0[\exp \lambda\rho(\Delta)]$$

by Lemma 4.5. This result together with relation (5.8) then imply that

$$\int_{B^c} \exp \lambda\rho(t)\, d\mathscr{L}(\Delta_s^* \mid P_0) \to \int_{B^c} \exp \lambda\rho(t)\, d\mathscr{L}(\Delta \mid P_0). \tag{5.9}$$

Relations (5.7) and (5.9) imply that

$$\int_{B^c} \exp \lambda\rho(t)\, d\mathscr{L}(\Delta_s^* \mid P_0) \leqq 2\varepsilon \quad \text{for} \quad s \quad \text{sufficiently large.} \tag{5.10}$$

Next $|\int_{[0,1]\times B^c} (z \exp h_s't)\, d\tau_s - \int_{[0,1]\times B^c} (z \exp h't)\, d\tau| \leqq \int_{[0,1]\times B^c} \exp \lambda\rho(t)\, d\tau_s + 2\int_{[0,1]\times B^c} \exp \lambda\rho(t)\, d\tau$ by (5.6) and the fact that $h_s \to h$. The bound is equal to $\int_{B^c} \exp \lambda\rho(t)\, d\mathscr{L}(\Delta_s^* \mid P_0) + 2\int_{B^c} \exp \lambda\rho(t)\, d\mathscr{L}(\Delta \mid P_0)$. Relations (5.7) and (5.10) imply that

$$\left|\int_{[0,1]\times B^c} (z \exp h_s't)\, d\tau_s - \int_{[0,1]\times B^c} (z \exp h't)\, d\tau\right| \leqq 4\varepsilon \tag{5.11}$$

for s sufficiently large.

Now we concentrate on the set $[0, 1] \times B$. On this set, we have

$$|z \exp h_s' t - z \exp h' t| \leqq |z|\, |\exp h_s' t - \exp h't| \leqq |\exp h_s't - \exp h't| \leqq M_1 \|h_s - h\|,$$

where M_1 is a constant. Therefore

$$\begin{aligned}
&\left|\int_{[0,1]\times B} (z \exp h_s't)\, d\tau_s - \int_{[0,1]\times B} (z \exp h't)\, d\tau\right| \\
&\quad \leqq \left|\int_{[0,1]\times B} (z \exp h_s't)\, d\tau_s - \int_{[0,1]\times B} (z \exp h't)\, d\tau_s\right| \\
&\qquad + \left|\int_{[0,1]\times B} (z \exp h't)\, d\tau_s - \int_{[0,1]\times B} (z \exp h't)\, d\tau\right| \\
&\quad \leqq M_1 \|h_s - h\| + \left|\int_{[0,1]\times B} (z \exp h't)\, d\tau_s - \int_{[0,1]\times B} (z \exp h't)\, d\tau\right| \to 0,
\end{aligned}$$

since $h_s \to h$, $\tau_s \to \tau$, and the integrand is bounded and continuous except on a subset of the surface of $[0, 1] \times B$, where B can be chosen so that the surface of $[0,1] \times B$ has τ-measure zero. Thus we have

$$\left|\int_{[0,1]\times B} (z \exp h_s't)\, d\tau_s - \int_{[0,1]\times B} (z \exp h't)\, d\tau\right| \leqq \varepsilon \tag{5.12}$$

for s sufficiently large. Relations (5.11) and (5.12) then imply that

$$\int_{[0,1]\times \mathscr{E}_k} (z \exp h_s't)\, d\tau_s \to \int_{[0,1]\times \mathscr{E}_k} (z \exp h't)\, d\tau. \tag{5.13}$$

Finally, relation (5.13) is rewritten as

$$\int (Z_s \exp h_s'\Delta_s^*)\, dP_0 \to \int (Z \exp h'\Delta)\, dP_0 \tag{5.14}$$

which is the first part of the lemma. From an argument entirely similar to that employed to establish relation (5.13) (Z is merely removed), it is seen that

$$\int_{[0,1]\times B} \exp h_s't\, d\tau_s \to \int_{[0,1]\times B} \exp h't\, d\tau$$

which is equivalent to $\int \exp h_s'\Delta_s^*\, dP_0 \to \int \exp h'\Delta\, dP_0$ and this is the second part of the lemma.

LEMMA. 5.3. *Let $\{s\}$ and $\{h_s\}$ be as in Lemma* 5.1. *Then*

$$\int \{Z_s \exp[h_s'\Delta_s^* - A(h_s)]\}\, dP_0 \to \int [Z \exp(h'\Delta + c)]\, dP_0.$$

PROOF. We know that $\mathscr{L}(h'\Delta \mid P_0)$ is $N(0, h'\Gamma h)$ and $h'\Gamma h = -2c$. Thus the right-hand side of (ii) in Lemma 5.2, being the moment generating function of $h'\Delta$ evaluated at $t = 1$, is equal to $\exp(\frac{1}{2}h'\Gamma h) = \exp(-c)$. That is, $\int \exp h'\Delta\, dP_0 = \exp(-c)$.

It follows that $A(h_s) \to -c$, and $\exp[-A(h_s)] \to \exp c$. This result together with relation (5.14) gives

$$\int \{Z_s \exp[h_s'\Delta_s^* - A(h_s)]\}\, dP_0 \to \int [Z \exp(h'\Delta + c)]\, dP_0$$

as was to be shown.

LEMMA 5.4. *Let $\{s\}$ and $\{h_s\}$ be as in Lemma* 5.1. *Then*

$$\int Z_s\, dR_{s,h_s} - \int Z_s\, dP_{s,\theta_s} \to 0.$$

PROOF. From Lemma 5.1 and Lemma 5.3, one obtains

$$\int Z_s\, dP_{\theta_s} - \int \{Z_s \exp[h_s'\Delta_s^* - A(h_s)]\}\, dP_0 \to 0. \tag{5.15}$$

Now, from the definition of $R_{s,\,h_s}$

$$\begin{aligned} \int Z_s\, dR_{s,h_s} &= \int \{Z_s \exp[-B_s(h_s)] \exp h_s'\Delta_s^*\}\, dP_0 \\ &= \exp[-B_s(h_s)] \int [Z_s \exp h_s'\Delta_s^*]\, dP_0. \end{aligned} \tag{5.16}$$

But $\exp B_s(h_s) = \int \exp h_s'\Delta_s^*\, dP_0 \to \int \exp h'\Delta\, dP_0$ by Lemma 5.2(ii), and $\int \exp h'\Delta\, dP_0 = \exp(-c)$ as was seen in the proof of Lemma 5.3. Furthermore $A(h_s) \to -c$ and hence

$$\exp B_s(h_s) - \exp A(h_s) \to 0. \tag{5.17}$$

Employing relation (5.16) and Lemma 5.2(i), together with relation (5.17)

$$\int Z_s\, dR_{s,h_s} - \int \{Z_s \exp[h_s'\Delta_s^* - A(h_s)]\}\, dP_0 \to 0. \tag{5.18}$$

Finally, relations (5.15) and (5.18) imply that

$$\int Z_s\, dR_{s,h_s} - \int Z_s\, dP_{s,\theta_s} \to 0$$

as asserted.

Now we complete the proof of Theorem 5.1 by observing that Lemma 5.4 contradicts relation (5.3) (with r replaced by s).

This theorem has two important corollaries whose formulation requires the following definitions.

DEFINITION 5.1. Let $\{Q_{n,\theta}^{(1)}\}$ and $\{Q_{n,\theta}^{(2)}\}$ be two sequences of probability measures on $\mathscr{A}_n$. Then we say that these two sequences are *differentially* (*asymptotically*) *equivalent* at the point θ_0 if, for each bounded set C in $\mathscr{E}_k$,

$$\sup\{\|Q_{n,\theta}^{(1)} - Q_{n,\theta}^{(2)}\|;\, \theta \in \Theta,\, n^{\frac{1}{2}}(\theta - \theta_0) \in C\} \to 0.$$

COROLLARY 5.1. *The sequences* $\{P_{n,\theta}\}$, $\{R_{n,h_n}\}$ *with* $h_n = \theta n^{\frac{1}{2}}$ *are differentially (asymptotically) equivalent at the point* 0.

PROOF. Let $R_{n,h_n} = R'_{n,\theta}$, where $h_n = \theta n^{\frac{1}{2}}$, and set $Q_{n,\theta}^{(1)} = P_{n,\theta}$, $Q_{n,\theta}^{(2)} = R'_{n,\theta}$. Assume that for a bounded set C of $\mathscr{E}_k$, $\sup\{\|Q_{n,\theta}^{(1)} - Q_{n,\theta}^{(2)}\|;\ \theta\in\Theta,\ \theta n^{\frac{1}{2}}\in C\}$ does not converge to zero. Then there exists $\{m\}\subseteq\{n\}$ and a sequence $\{\theta_m\}$ with $\theta_m m^{\frac{1}{2}}\in C$ such that $\|Q_{m,\theta_m}^{(1)} - Q_{m,\theta_m}^{(2)}\| > \varepsilon$, some $\varepsilon > 0$ and for all m. Reverting to the original measures, this becomes $\|P_{m,\theta_m} - R_{m,h_m}\| > \varepsilon$ for all m, where $\theta_m = h_m m^{-\frac{1}{2}}$ and $h_m\in C$ for all m; but this contradicts the theorem and completes the proof of the corollary.

DEFINITION 5.2. A sequence $\{V_n\}$ of $\{\mathscr{A}_n\}$-measurable functions, or the sequence $\{\mathscr{B}_n\}$ of the σ-fields induced by them, is said to be *differentially (asymptotically) sufficient* at θ_0 for the family $\{P_{n,\theta};\ \theta\in\Theta\}$ if there is a family $\{Q_{n,\theta};\ \theta\in\Theta\}$ such that the sequences $\{P_{n,\theta}\}$ and $\{Q_{n,\theta}\}$ are differentially equivalent at θ_0 and for each n, V_n, or the σ-field $\mathscr{B}_n$, is sufficient for the family $\{Q_{n,\theta};\ \theta\in\Theta\}$.

COROLLARY 5.2. *The sequence* $\{\Delta_n{}^*\}$ *is differentially (asymptotically) sufficient at* 0 *for the family* $\{P_{n,\theta};\ \theta\in\Theta\}$.

PROOF. From the definition of $R_{n,h}$ and for each n, one has $[dR_{n,h}/dP_{n,0}] = \exp[-B_n(h)]\exp(h'\Delta_n{}^*)$, $h\in C_n = \{h\in\mathscr{E}_k;\ h = \theta n^{\frac{1}{2}},\ \theta\in\Theta\}$; equivalently

$$[dR'_{n,\theta}/dP_{n,0}] = \exp[-B_n(\theta n^{\frac{1}{2}})]\exp(n^{\frac{1}{2}}\theta'\Delta_n{}^*),\ \theta\in\Theta,$$

where $R'_{n,\theta} = R_{n,h}$ with $h = \theta n^{\frac{1}{2}}$. Thus, for each n, $\Delta_n{}^*$ is sufficient for the family $\{R'_{n,\theta};\ \theta\in\Theta\}$ or, equivalently, for the family $\{R_{n,h};\ h\in C_n\}$. Since $\{P_{n,\theta}\}$, $\{R_{n,h}\}$ are differentially equivalent at the point 0 by Corollary 5.1, the proof is completed.

6. Asymptotic properties of tests based on Δ_n and asymptotic distribution of $\{\Delta_n\}$. In this section, two main results are presented. The first asserts that from the asymptotic power viewpoint, any test may be based on Δ_n alone. The second provides the asymptotic distribution of $\{\Delta_n\}$ under the moving measure P_{n,θ_n}. More precisely, we have

THEOREM 6.1. *Let* $\{Z_n\}$ *be a sequence of random variables such that* $|Z_n| \leqq 1$, $n \geqq 1$ *and set* $\bar{Z}_n = \mathscr{E}_{P_{n,0}}(Z_n \mid \Delta_n)$. *Then*

$$\sup|\mathscr{E}(Z_n| P_{n,\theta_n}) - \mathscr{E}(\bar{Z}_n \mid P_{n,\theta_n})| \to 0,$$

where $\theta_n = hn^{-\frac{1}{2}}$ *and the* sup *is taken over all sequences* $\{Z_n\}$ *of random variables bounded by* 1 *in absolute value and over all* h's *in a bounded set* C.

PROOF. We have, $\mathscr{E}(Z_n \mid P_{n,\theta_n}) - \mathscr{E}(\bar{Z}_n \mid P_{n,\theta_n}) = I_1(n, h) + I_2(n, h) + I_3(n, h)$, where

$$I_1(n, h) = \mathscr{E}(Z_n \mid P_{n,\theta_n}) - \mathscr{E}(Z_n \mid R_{n,h})$$
$$I_2(n, h) = \mathscr{E}(Z_n \mid R_{n,h}) - \mathscr{E}(\bar{Z}_n \mid R_{n,h})$$
$$I_3(n, h) = \mathscr{E}(\bar{Z}_n \mid R_{n,h}) - \mathscr{E}(\bar{Z}_n \mid P_{n,\theta_n}).$$

Now, with the sup as above, $\sup|I_1(n, h)| \leqq \sup\{\|R_{n,h} - P_{n,\theta_n}\|;\ h\in C\}$. If the right-hand side of this last relation does not converge to zero, there exists

$\{m\} \subseteq \{n\}$ and a sequence $\{h_m\}$ of h's in C such that $\|R_{m,h_m} - P_{m,\theta_m}\| > \varepsilon$ for all m; but this contradicts Theorem 5.1 and thus $\sup|I_1(n, h)| \to 0$. In the same way it is seen that $\sup|I_3(n, h)| \to 0$. Thus we need only show that $\sup|I_2(n, h)| \to 0$. To see this, we observe that

$$\begin{aligned}\mathscr{E}(\bar{Z}_n \mid R_{n,h}) &= \mathscr{E}[\mathscr{E}_{P_{n,0}}(Z_n \mid \Delta_n) \mid R_{n,h}] = \int [\mathscr{E}_{P_{n,0}}(Z_n \mid \Delta_n)]\, dR_{n,h} \\ &= \int [\mathscr{E}_{P_{n,0}}(Z_n \mid \Delta_n)] \exp[-B_n(h)] \exp(h'\Delta_n^*)\, dP_{n,0} \\ &= \mathscr{E}_{P_{n,0}}\{\mathscr{E}_{P_{n,0}}(Z_n \mid \Delta_n) \exp[-B_n(h)] \exp(h'\Delta_n^*)\} \\ &= \mathscr{E}_{P_{n,0}}[\mathscr{E}_{P_{n,0}}\{Z_n \exp[-B_n(h)] \exp(h'\Delta_n^*) \mid \Delta_n\}]\end{aligned}$$

because Δ_n^* is a function of Δ_n. Thus the last expression reduces to

$$\mathscr{E}_{P_{n,0}}\{Z_n \exp[-B_n(h)] \exp(h'\Delta_n^*)\}$$

and this is equal to $\mathscr{E}(Z_n \mid R_{n,h})$. Therefore $I_2(n, h) = 0$ and the proof of the theorem is completed.

According to this last result, if $\{Z_n\}$ is any sequence of tests, where Z_n has level α_n, then one can always replace Z_n by $\bar{Z}_n = \mathscr{E}_{P_{n,0}}(Z_n \mid \Delta_n)$ which has the same level. Furthermore, the asymptotic power will remain unchanged. Thus in searching for an optimal test, for example in the sense of maximizing the power under the alternatives considered above, one may restrict attention to tests based only on Δ_n.

Next let $h, h_n \in \mathscr{E}_k$ with $h_n \to h$ and, for $\theta \in \Theta$, set $\theta_n = \theta + h_n n^{-\frac{1}{2}}$. We also set $\Lambda_n(\theta) = \Lambda[P_{n,\theta_n}; P_{n,\theta}]$. Then Theorem 3.2 provides the asymptotic distribution of $\Delta_n(\theta)$ under $P_{n,\theta}$. In statistical applications, the asymptotic distribution of $\Delta_n(\theta)$ under P_{n,θ_n} is also needed. With this in mind, we establish the following theorem.

THEOREM 6.2. *Let $h, h_n \in \mathscr{E}_k$ with $h_n \to h$ and let $\theta_n = \theta + h_n n^{-\frac{1}{2}}$, $\theta \in \Theta$. Then*

$$\mathscr{L}[\Delta_n(\theta) \mid P_{n,\theta_n}] \to N(\Gamma(\theta)h, \Gamma(\theta)).$$

PROOF. Let $c' = (c_1, \cdots, c_k)$ and $c_0, c_j \in R, j = 1, \cdots, k$. Then

$$c_0 \Lambda_n(\theta) + c'\Delta_n(\theta) = c_0[\Lambda_n(\theta) - h_n'\Delta_n(\theta)] + [c_0 h_n' + c']\Delta_n(\theta). \tag{6.1}$$

By Theorem 3.2, we then have

$$\mathscr{L}\{[c_0 h_n' + c']\Delta_n(\theta) \mid P_{n,\theta}\} \to N(0, \Gamma^*(\theta, h)), \tag{6.2}$$

where $\Gamma^*(\theta, h) = (c_0 h' + c')\Gamma(\theta)(c_0 h + c)$. Therefore, relation (6.1), together with Theorem 3.2 and Theorem 3.3, gives

$$\mathscr{L}[c_0 \Lambda_n(\theta) + c'\Delta_n(\theta) \mid P_{n,\theta}] \to N(-\tfrac{1}{2}c_0 h'\Gamma(\theta)h, \Gamma^*(\theta, h)). \tag{6.3}$$

Let $d' = (c_0, c_1, \cdots, c_k)$. Then

$$c_0 \Lambda_n(\theta) + c'\Delta_n(\theta) = d' \begin{pmatrix} \Lambda_n(\theta) \\ \Delta_n(\theta) \end{pmatrix}$$

while $-\frac{1}{2}c_0\, h'\Gamma(\theta)h = d'\begin{pmatrix} -\frac{1}{2}h'\Gamma(\theta)h \\ 0 \\ \vdots \\ 0 \end{pmatrix}$ and $\Gamma^*(\theta, h) = d'\begin{pmatrix} h'\Gamma(\theta)h & h'\Gamma(\theta) \\ \Gamma(\theta)h & \Gamma(\theta) \end{pmatrix} d.$

Thus relation (6.3) states that

$$\mathscr{L}\left[d'\begin{pmatrix} \Lambda_n(\theta) \\ \Delta_n(\theta) \end{pmatrix} \Big| P_{n,\theta} \right] \to N\left(d'\begin{pmatrix} -\frac{1}{2}h'\Gamma(\theta)h \\ 0 \\ \vdots \\ 0 \end{pmatrix}, d'\begin{pmatrix} h'\Gamma(\theta)h & h'\Gamma(\theta) \\ \Gamma(\theta)h & \Gamma(\theta) \end{pmatrix} d \right)$$

and this is equivalent to

$$\mathscr{L}\{[\Lambda_n(\theta), \Delta_n(\theta)] \mid P_{n,\theta}\} \to N\left(\begin{pmatrix} -\frac{1}{2}h'\Gamma(\theta)h \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \begin{pmatrix} h'\Gamma(\theta)h & h'\Gamma(\theta) \\ \Gamma(\theta)h & \Gamma(\theta) \end{pmatrix} \right). \tag{6.4}$$

Let the distribution on the right-hand side of relation (6.4) be the joint distribution, under P_θ, of the random variables Λ^* and Δ, that is, $\mathscr{L}(\Lambda^*, \Delta \mid P_\theta)$ where $\mathscr{L}(\Lambda^* \mid P_\theta) = N(-\frac{1}{2}h'\Gamma(\theta)h,\ h'\Gamma(\theta)h)$ and $\mathscr{L}(\Delta \mid P_\theta) = N(0, \Gamma(\theta))$. Then relation (6.4) implies, by Theorem 2.1(6) in LeCam [3],

$$\mathscr{L}\{[\Lambda_n(\theta), \Delta_n(\theta)] \mid P_{n,\theta_n}\} \to \exp(\lambda)\mathscr{L}(\Lambda^*, \Delta \mid P_\theta). \tag{6.5}$$

But

$$\int \exp(\lambda)\, d\mathscr{L}(\Lambda^*, \Delta \mid P_\theta) = \int\int \exp(\lambda)\, d\mathscr{L}[\Lambda^* \mid \Delta \,|\, P_\theta]\, d\mathscr{L}(\Delta \mid P_\theta)$$

and $\mathscr{L}[\Lambda^* \mid \Delta \,|\, P_\theta]$ is normal with mean

$$\begin{aligned} \mathscr{E}(\Lambda^* \mid P_\theta) + \Sigma_{12}\Sigma_{22}^{-1}[\Delta - \mathscr{E}(\Delta \mid P_\theta)] &= -\tfrac{1}{2}h'\Gamma(\theta)h + h'\Gamma(\theta)\Gamma^{-1}(\theta)\,\Delta \\ &= -\tfrac{1}{2}h'\Gamma(\theta)h + h'\Delta \end{aligned}$$

and variance

$$\Sigma_{11} - \Sigma_{12}\Sigma_{22}^{-1}\Sigma_{21} = h'\Gamma(\theta)h - h'\Gamma(\theta)\Gamma^{-1}(\theta)\Gamma(\theta)h = h'\Gamma(\theta)h - h'\Gamma(\theta)h = 0.$$

(See, e.g., Rao [5] page 441(v)). Hence

$$\int_{(-\infty, \lambda']} \exp(\lambda)\, d\mathscr{L}[\Lambda^* \mid \Delta \,|\, P_\theta] = \exp[-\tfrac{1}{2}h'\Gamma(\theta)h + h'\Delta] \quad \text{if} \quad \lambda' \geqq -\tfrac{1}{2}h'\Gamma(\theta)h + h'\Delta$$

and zero otherwise. Letting $(-\infty, t]$ stand for a k-dimensional interval, relation (6.5) states that

$$P[\Lambda_n(\theta) \leqq \lambda',\ \Delta_n(\theta) \leqq t \mid P_{n,\theta_n}] \to \int_{(-\infty, t]}\int_{(-\infty, \lambda']} \exp(\lambda)\, d\mathscr{L}(\Lambda^*, \Delta \mid P_\theta),$$

which implies that

$$P[\Delta_n(\theta) \leqq t \mid P_{n,\theta_n}] \to \int_{(-\infty, \infty)} [\textstyle\int_{(-\infty, t]} \exp(\lambda)\, d\mathscr{L}(\Lambda^*, \Delta \mid P_\theta)]\, d\lambda.$$

But the right-hand side of the last relation is equal to

$$\int_{(-\infty, t]} \exp\left[-\tfrac{1}{2}h'\Gamma(\theta)h + h'\delta\right] d\mathscr{L}(\Delta \mid P_\theta)$$

and by taking into consideration that $\mathscr{L}(\Delta \mid P_\theta) = N(0, \Gamma(\theta))$, this integral becomes

$$\int_{(-\infty, t]} \exp\left[-\tfrac{1}{2}h'\Gamma(\theta)h + h'\delta\right](2\pi)^{-k/2} \left|\Gamma(\theta)\right|^{-\frac{1}{2}} \exp\left[-\tfrac{1}{2}\delta'\Gamma^{-1}(\theta)\delta\right] d\delta$$
$$= \int_{(-\infty, t]} (2\pi)^{-k/2} \left|\Gamma(\theta)\right|^{-\frac{1}{2}} \exp\left\{-\tfrac{1}{2}[\delta - \Gamma(\theta)h]'\Gamma^{-1}(\theta)[\delta - \Gamma(\theta)h]\right\} d\delta.$$

Therefore $\mathscr{L}[\Delta_n(\theta) \mid P_{n, \theta_n}] \to N(\Gamma(\theta)h, \Gamma(\theta))$ as was to be seen.

Theorem 6.1 shows that, from an asymptotic power viewpoint, any test can be based on Δ_n. Proceeding further along the same path, the next theorem asserts that one can actually base all tests on Δ_n^* rather than Δ_n. The significance of this fact is apparent: the original family of distributions can be treated asymptotically as if it were an exponential family, where Δ_n^* plays the all important role of the statistic appearing in the exponent of an exponential family.

THEOREM 6.3. *Let $\{Z_n\}$, $n \geqq 1$ be a sequence of test functions and set $\theta_n = hn^{-\frac{1}{2}}$, $h \in \mathscr{E}_k$. Then, we have*

$$\sup\{|\mathscr{E}_{\theta_n} Z_n(\Delta_n) - \mathscr{E}_{\theta_n} Z_n(\Delta_n^*)| ; h \in K\} \to 0,$$

where K is any compact subset of $\mathscr{E}_k$.

PROOF. We recall that Δ_n^* is defined as follows

$$\Delta_n^* = \Delta_n, \quad \text{if} \quad \rho(\Delta_n) < \alpha_\nu$$
$$= 0, \quad \text{otherwise,} \qquad \text{where} \quad \alpha_\nu \uparrow \infty.$$

Let $Q^h_{1,n}(B) = P_{\theta_n}(\Delta_n \in B)$, $Q^h_{2,n}(B) = P_{\theta_n}(\Delta_n^* \in B)$, where B varies over the k-dimensional Borel σ-field $\mathscr{B}_k$. Set $B_n = \{\rho(z) < \alpha_\nu\} - \{0\}$. Then $\Delta_n^* = \Delta_n$ pointwise on $\Delta_n^{*-1}(B_n) = \Delta_n^{-1}(B_n)$ so that $P_{\theta_n}(\Delta_n \in B_n) = P_{\theta_n}(\Delta_n^* \in B_n)$. Next for any $B \in \mathscr{B}_k$, we have

$$Q^h_{1,n}(B \cap B_n) = Q^h_{2,n}(B \cap B_n),$$
$$Q^h_{1,n}(B \cap B_n^c) \leqq Q^h_{1,n}(B_n^c) = P_{\theta_n}(\Delta_n \neq \Delta_n^*) + P_{\theta_n}(\Delta_n = 0), \qquad \text{and}$$
$$Q^h_{2,n}(B \cap B_n^c) \leqq Q^h_{2,n}(B_n^c) = P_{\theta_n}[\rho(\Delta_n^*) \geqq \alpha_\nu] + P_{\theta_n}(\Delta_n^* = 0) = P_{\theta_n}(\Delta_n^* = 0),$$

since $[\rho(\Delta_n^*) \geqq \alpha_\nu] = \emptyset$.

In terms of these relations,

$$\begin{aligned} |Q^h_{1,n}(B) - Q^h_{2,n}(B)| &= |Q^h_{1,n}(B \cap B_n^c) - Q^h_{2,n}(B \cap B_n^c)| \\ &\leqq Q^h_{1,n}(B \cap B_n^c) + Q^h_{2,n}(B \cap B_n^c) \\ &= P_{\theta_n}(\Delta_n \neq \Delta_n^*) + P_{\theta_n}(\Delta_n = 0) + P_{\theta_n}(\Delta_n^* = 0). \end{aligned} \tag{6.6}$$

Clearly, $(\Delta_n^* = 0) = (\Delta_n \neq \Delta_n^*) + (\Delta_n = 0)$. Therefore (6.6) becomes

$$|Q^h_{1,n}(B) - Q^h_{2,n}(B)| \leqq 2[P_{\theta_n}(\Delta_n \neq \Delta_n^*) + P_{\theta_n}(\Delta_n = 0)]$$

and hence

$$\sup\{|Q_{1,n}^h(B)-Q_{2,n}^h(B)|;\,B\in\mathscr{B}_k\}\leqq 2[P_{\theta_n}(\Delta_n\neq\Delta_n^*)+P_{\theta_n}(\Delta_n=0)].$$

Also

$$\sup\{|Q_{1,n}^h(B)-Q_{2,n}^h(B)|;\,B\in\mathscr{B}_k,\ h\in K\}$$
$$\leqq 2\sup\{[P_{\theta_n}(\Delta_n\neq\Delta_n^*)+P_{\theta_n}(\Delta_n=0)];\,h\in K\}.$$

Now the right-hand side of this last relation converges to zero for, if not, there is $\{m\}\subseteq\{n\}$ and $h_m\in K$ with $h_m\to h$ such that at least one of the sequences $\{P_{\theta_m^*}(\Delta_m\neq\Delta_m^*)\}$, $\{P_{\theta_m^*}(\Delta_m=0)\}$ does not converge to zero. Here $\theta_m^*=h_m m^{-\frac{1}{2}}$. This is a contradiction since $P_0(\Delta_m\neq\Delta_m^*)\to 0$, $P_0(\Delta_m=0)\to 0$ and the same is true for the above sequences by contiguity. Now

$$|\mathscr{E}_{\theta_n}Z_n(\Delta_n)-\mathscr{E}_{\theta_n}Z_n(\Delta_n^*)|=\left|\int z(t)\,d\mathscr{L}(\Delta_n\,|\,P_{\theta_n})-\int z(t)\,d\mathscr{L}(\Delta_n^*\,|\,P_{\theta_n})\right|,$$

where $0\leqq z(t)\leqq 1$, $t\in\mathscr{E}_k$. The right-hand side of this last relation is bounded by $\|Q_{1,n}^h-Q_{2,n}^h\|$ which converges to zero by the previous result. This completes the proof of the theorem.

7. Some applications to hypotheses testing problems. In this section, we further extend the results of Wald [8] along the lines suggested by LeCam [3]. More general hypotheses are considered than in Johnson and Roussas [2] although the null hypothesis is required to be simple. As expected from the terminology of the previous sections, the asymptotically optimal tests may be based on the sequence $\{\Delta_n\}$. In the present paper, we treat in detail the one-dimensional parameter situation considered in Johnson and Roussas [2] from the viewpoint of differential (asymptotic) sufficiency. At the end of this section, some preliminary remarks are also made in anticipation of k-dimensional parameter results.

7.1. *One-dimensional* Θ. Modifying slightly the definitions of Wald [8] with respect to the significance level requirement, we have

DEFINITION 7.1.1. A sequence of tests $\{\lambda_n\}$ with $\mathscr{E}_{\theta_0}\lambda_n\to\alpha$ is said to be an *asymptotically most powerful* test of $H_0:\theta=\theta_0$ against $H_1:\theta>\theta_0$, of asymptotic size α, if for any other sequence of tests $\{\omega_n\}$ with $\mathscr{E}_{\theta_0}\omega_n\to\alpha$, we have

$$\limsup\,[\sup(\mathscr{E}_\theta\omega_n-\mathscr{E}_\theta\lambda_n;\,\theta>\theta_0,\,\theta\in\Theta)]\leqq 0. \tag{7.1.1}$$

A similar expression is required to hold with $\theta>\theta_0$ replaced by $\theta<\theta_0$ if the alternative is $H_2:\theta<\theta_0$.

DEFINITION 7.1.2. A sequence of tests $\{\lambda_n\}$ is defined to be an *asymptotically most powerful unbiased* test of the hypothesis $H_0:\theta=\theta_0$, of asymptotic size α, if $\mathscr{E}_{\theta_0}\lambda_n\to\alpha$ with $\liminf\,[\inf(\mathscr{E}_\theta\lambda_n;\,\theta\neq\theta_0;\,\theta\in\Theta)]\geqq\alpha$, and if for any other sequence of tests $\{\omega_n\}$ satisfying $\mathscr{E}_{\theta_0}\omega_n\to\alpha$ and $\liminf\,[\inf(\mathscr{E}_\theta\omega_n;\,\theta\neq\theta_0;\,\theta\in\Theta)]\geqq\alpha$, we have

$$\limsup\,[\sup(\mathscr{E}_\theta\omega_n-\mathscr{E}_\theta\lambda_n;\,\theta\neq\theta_0,\,\theta\in\Theta)]\leqq 0. \tag{7.1.2}$$

Once more, without loss of generality, we take $\theta_0=0$.

The form of the asymptotic distribution of $\{\Delta_n\}$, under $P_{n,0}$ and P_{n,θ_n}, where $\theta_n = hn^{-\frac{1}{2}}$, suggests a form for tests based on Δ_n. Namely, for a fixed $\alpha (0 < \alpha < 1)$, let

(7.1.3)
$$\varphi_{1,n}(\Delta_n) = 1, \quad \text{if} \quad \Delta_n > c_n$$
$$= 0, \quad \text{if} \quad \Delta_n < c_n,$$

where $c_n \to \xi_\alpha$, the upper αth quantile of the $N(0, \Gamma)$, and $\mathscr{E}_0\, \varphi_{1,n}(\Delta_n) \to \alpha$. Further, set

(7.1.4)
$$\varphi_{2,n}(\Delta_n) = 1, \quad \text{if} \quad \Delta_n < a_n \quad \text{or} \quad \Delta_n > b_n$$
$$= 0, \quad \text{if} \quad a_n < \Delta_n < b_n,$$

where $\lim(-a_n) = \lim b_n = \xi_{\alpha/2}$, so that $\mathscr{E}_0\, \varphi_{2n}(\Delta_n) \to \alpha$.

For both sequences of tests, a critical function may assume any value on the boundary of its critical region.

In order to show that the first and the second tests are optimal in the sense of Definition 7.1.1 and Definition 7.1.2, respectively, we must employ one further assumption which was verified for a number of interesting cases in Johnson and Roussas [2].

ASSUMPTION 5. Let $\{\theta_n\}$ be a sequence of elements of Θ with $\theta_n > 0$ for each n. The condition $\lim \theta_n n^{\frac{1}{2}} = \infty$ implies that $\Delta_n \to \infty$ in P_{θ_n}-probability.

ASSUMPTION 5′. Let $\{\theta_n\}$ be a sequence of elements of Θ with $\theta_n < 0$ for each n. The condition $\lim \theta_n n^{\frac{1}{2}} = -\infty$ implies that $\Delta_n \to -\infty$ in P_{θ_n}-probability.

We first restate the result for one-sided tests from Johnson and Roussas [2] which was established without the concept of differential (asymptotic) sufficiency.

THEOREM 7.1.1. *Under Assumptions* 1–5, *the test* $\varphi_{1,n}$ *defined by* (7.1.3) *is asymptotically most powerful for testing* $H_0: \theta = 0$ *against the alternative* $H_1: \theta > 0$, $\theta \in \Theta$.

We have an analogous theorem for the alternative $H_2: \theta < 0, \theta \in \Theta$, under Assumptions 1–5′.

The proof of these results parallels the proof of Theorem 7.1.2 below and therefore will be omitted.

To Theorem 7.1.1, we also have the following Corollary.

COROLLARY 7.1.1. *Under Assumptions* 1–4, *the test* $\varphi_{1,n}$ *defined by* (7.1.3) *is asymptotically locally most powerful. That is, the test satisfies* (7.1.1) *when the alternatives are further restricted to* θ*'s for which* $n^{\frac{1}{2}}(\theta - \theta_0) \leqq C$ *for an arbitrary fixed* C.

Of course, we again have an analogous result for the alternative $H_2: \theta < 0, \theta \in \Theta$.

The proof of the next theorem requires the results on the differential (asymptotic) sufficiency of the truncated version Δ_n^*.

THEOREM 7.1.2. *Under Assumptions* 1–5, 5′, *the test* $\varphi_{2,n}$ *defined by* (7.1.4) *is asymptotically most powerful unbiased for testing* $H_0: \theta = 0$ *against the alternative* $H_3: \theta \neq 0, \theta \in \Theta$.

For the proof of this theorem some auxiliary results will be needed which are formulated below as lemmas. The first is one version of the extended Helly–Bray lemma and the proof is included because it is brief.

LEMMA 7.1.1. *Let* $G_n \to G$, *a continuous* cdf *of a random variable, and let* $\{g_n\}$, g *and* h *be continuous functions satisfying*

(i) $|g_n(x)| \leqq h(x)$ *for all* x,

(ii) $g_n(x) \to g(x)$ *uniformly on finite intervals, and*

(iii) $\int h\,dG_n \to \int h\,dG$.

Then

$$\int g_n\,dG_n \to \int g\,dG. \tag{7.1.5}$$

Relation (7.1.5) also holds for the modified functions $g_n^* = I_{(a_n, b_n)}\,g_n$ and $g^* = I_{(a,b)}\,g$, where $a_n \to a$, $b_n \to b$ with a, b, finite or infinite.

PROOF. Given an $\varepsilon > 0$, select c and d such that

$$\int_{(-\infty, c]} h\,dG + \int_{(d,\infty)} h\,dG < \varepsilon. \tag{7.1.6}$$

Then

$$\int_{(-\infty, c]} |g_n|\,dG_n + \int_{(d,\infty)} |g_n|\,dG_n \leqq \int_{(-\infty, c]} h\,dG_n + \int_{(d,\infty)} h\,dG_n \tag{7.1.7}$$

and the r.h.s. converges to the l.h.s. of (7.1.6) according to condition (iii) and the Helly–Bray lemma. Then on the interval (c, d),

$$\left|\int_{(c,d]} g_n\,dG_n - \int_{(c,d]} g\,dG\right| \leqq \int_{(c,d]} |g_n - g|\,dG_n + \left|\int_{(c,d]} g\,dG_n - \int_{(c,d]} g\,dG\right| \tag{7.1.8}$$

and the first term on the r.h.s. converges to zero by condition (ii) and the second by the Helly–Bray lemma.

Note that the bound (7.1.7) also holds for the modified functions g_n^* and g^* so that it is clearly sufficient to treat the case where a and b are finite. Since G is continuous, there exist an $a_0 > 0$ such that

$$\int_{(a-a_0, a+a_0]} h\,dG + \int_{(b-a_0, b+a_0]} h\,dG < \varepsilon. \tag{7.1.9}$$

For sufficiently large n, $0 = g_n^* = g^*$ outside the interval $(a-a_0, b+a_0)$ and $g_n^* = g_n$, $g^* = g$ on $(a+a_0, b-a_0)$. For the latter interval, use the method leading to (7.1.8). Also

$$\int_{(a-a_0, a+a_0]} |g_n^*|\,dG_n + \int_{(b-a_0, b+a_0]} |g_n^*|\,dG_n \leqq \int_{(a-a_0, a+a_0]} h\,dG_n + \int_{(b-a_0, b+a_0]} h\,dG_n$$

and this converges to the l.h.s. of (7.1.9) which completes the proof.

LEMMA 7.1.2. *Consider the sequence of tests* $\{\varphi_n(\Delta_n^*)\}$, $n \geqq 1$, *defined as follows*

$$\varphi_n(\Delta_n^*) = \begin{cases} 1, & \text{if } \Delta_n^* < a_n^* \text{ or } \Delta_n^* > b_n^* \\ \gamma_1^* \text{ or } \gamma_2^*, & \text{if } \Delta_n^* = a_n^* \text{ or } \Delta_n^* = b_n^*, \text{ respectively}, \\ 0, & \text{if } a_n^* < \Delta_n^* < b_n^*, \end{cases} \tag{7.1.10}$$

where the constants a_n^*, b_n^*, γ_1^*, *and* γ_2^* *are defined through the relations*

$$(7.1.11) \qquad \mathscr{E}_0\,\varphi_n(\Delta_n^*) = \alpha, \qquad \mathscr{E}_0[\Delta_n^*\varphi_n(\Delta_n^*)] = \alpha\mathscr{E}_0\,\Delta_n^*.$$

Then $a_n^* \to -\xi_{\alpha/2}$ *and* $b_n^* \to \xi_{\alpha/2}$, *where* ξ_p *is the upper pth quantile of the* $N(0, \Gamma)$.

PROOF. The proof is by contradiction. In the first place one may neglect asymptotically the contribution due to randomization. In fact, if $x_n \to x$, where x may be finite or $\pm\infty$, one has $P_0(\Delta_n^* \leqq x_n) - \Phi^*(x_n) \to 0$; here Φ^* stands for the distribution function of the $N(0, \Gamma)$ distribution. Thus by continuity of Φ^*, $P_0(\Delta_n^* \leqq x_n) \to \Phi^*(x)$. Therefore

$$P_0(\Delta_n^* = x_n) \leqq P_0(\Delta_n^* \leqq x_n + \varepsilon) - P_0(\Delta_n^* \leqq x_n - \varepsilon) \to \Phi^*(x+\varepsilon) - \Phi^*(x-\varepsilon)$$

and letting $\varepsilon \to 0$, one obtains $P_0(\Delta_n^* = x_n) \to 0$. Now assume that both $\{a_n^*\}$ and $\{b_n^*\}$ are unbounded. Then there are three cases to consider: there exist subsequences $\{a_m^*\}$ and $\{b_m^*\}$ such that $a_m^* \to -\infty$ and $b_m^* \to \infty$, or $a_m^* \to \infty$ and $b_m^* \to \infty$, or $a_m^* \to -\infty$ and $b_m^* \to -\infty$. In the first case one has, for m sufficiently large,

$$\begin{aligned}\alpha &\leqq \tfrac{1}{4}\alpha + \left[\int_{(-\infty,\, a_m^*+1)} d\mathscr{L}(\Delta_m^* \mid P_0) + \int_{(b_m^*-1,\, \infty)} d\mathscr{L}(\Delta_m^* \mid P_0)\right] \\ &\leqq \tfrac{1}{4}\alpha + \int_{(-\xi_{\alpha/4},\, \infty)} d\mathscr{L}(\Delta_m^* \mid P_0) + \int_{(\xi_{\alpha/4},\, \infty)} d\mathscr{L}(\Delta_m^* \mid P_0)\end{aligned}$$

and this converges to $\frac{3}{4}\alpha$. Thus we arrive at a contradiction. For the second case, one has $\alpha \geqq \int_{-\infty}^{a_m^*-1} d\mathscr{L}(\Delta_m^* \mid P_0)$ and this converges to 1. The third case is treated in a similar fashion. Now let one of the sequences $\{a_n^*\}$, $\{b_n^*\}$ be unbounded, for example, the sequence $\{a_n^*\}$ and the other be bounded. Then there exist subsequences $\{a_m^*\}$ and $\{b_m^*\}$ such that $a_m^* \to -\infty$ and $b_m^* \to b$ finite. Thus

$$\begin{aligned}\alpha &= \lim\left[\int_{(-\infty,\, a_m^*)} d\mathscr{L}(\Delta_m^* \mid P_0) + \int_{(b_m^*,\, \infty)} d\mathscr{L}(\Delta_m^* \mid P_0)\right] \\ &= 1 - \lim \int_{[a_m^*,\, b_m^*]} d\mathscr{L}(\Delta_m^*, P_0) = 1 - \Phi^*(b)\end{aligned}$$

so that $b = \xi_\alpha$. Next, $\mathscr{E}_0\Delta_n^* = \int x\, d\mathscr{L}(\Delta_n^* \mid P_0) \to \int x\, d\Phi^*$ by Lemmas 7.1.1 and 4.5 and this is equal to zero. On the other hand, $\lim \mathscr{E}_0[\Delta_m^*\varphi_m(\Delta_m^*)] = \lim\left[\int_{(-\infty,\, a_m^*)} x\, d\mathscr{L}(\Delta_m^* \mid P_0) + \int_{(b_m^*,\, \infty)} x\, d\mathscr{L}(\Delta_m^* \mid P_0)\right] = \lim \int_{(b_m^*,\, \infty)} x\, d\mathscr{L}(\Delta_m^* \mid P_0)$ by Lemma 7.1.1 and this last limit equals $\int_{(b,\, \infty)} x\, d\Phi^*$ by the same lemma. Therefore, taking the limits of both sides of the second relation in (7.1.11) through the subsequence $\{m\}$, one has $0 = \int_{(b,\, \infty)} x\, d\Phi^*$ which is a contradiction. The case that $\{a_n^*\}$ is bounded and $\{b_n^*\}$ unbounded is treated similarly. Finally, consider the case that both $\{a_n^*\}$ and $\{b_n^*\}$ are bounded and let $\{a_m^*\}$, $\{b_m^*\}$ be subsequences such that $a_m^* \to a$, $b_m^* \to b$, where both a and b are finite. In the first place, $\lim \mathscr{E}_0\,\Delta_m^* = 0$ as was seen above. Next

$$\begin{aligned}\lim \mathscr{E}_0[\Delta_m^*\varphi_m(\Delta_m^*)] &= \lim\left[\int_{(-\infty,\, a_m^*)} x\, d\mathscr{L}(\Delta_m^* \mid P_0) + \int_{[b_m^*,\, \infty)} x\, d\mathscr{L}(\Delta_m^* \mid P_0)\right] \\ &= -\lim \int_{[a_m^*,\, b_m^*]} x\, d\mathscr{L}(\Delta_m^* \mid P_0) = -\int_{[a,\, b]} x\, d\Phi^*\end{aligned}$$

by Lemma 7.1.1. Therefore, by taking the limits of both sides of the second relation in (7.1.11) through the subsequence $\{m\}$, one has $\int_{[a,\, b]} x\, d\Phi^* = 0$ which can happen only if $a = -b$. By taking the limit of the first relation in (7.1.11) through the subsequence $\{m\}$, one also has $\Phi^*(b)-\Phi^*(a) = 1-\alpha$ or $\Phi^*(b)-\Phi^*(-b) = 1-\alpha$ which implies that $b = \xi_{\alpha/2}$. The proof of the lemma is completed.

It is worthwhile to remark that the power satisfies

$$\mathscr{E}_{\theta_m}\varphi_m(\Delta_m^*) \to \int_{|x|\geqq \xi_{\alpha/2}} d\Phi_h, \tag{7.1.12}$$

under alternatives $\theta_m = h_m m^{-\frac{1}{2}}$ with $h_m \to h$, where Φ_h denotes the distribution function of $N(\Gamma h, \Gamma)$. This follows directly from the last two lemmas.

PROOF OF THEOREM 7.1.2. Let $\{\varphi_{2,n}\}$ be defined by (7.1.4) and let $\{\omega_n\} = \{\omega_n(X_0, X_1, \cdots, X_n)\}$ be any other sequence of tests with $\mathscr{E}_0\,\omega_n \to \alpha$ and

$$\liminf \inf(\mathscr{E}_\theta\,\omega_n;\, \theta \neq 0, \theta\in\Theta) \to \alpha.$$

Assume, for contradiction, that the left-hand side of (7.1.2) takes the value $10\delta > 0$. That is, there exist sequences $\{m\}\subseteq\{n\}$ and $\{\theta_m\}$ with $\theta_m\in\Theta$ and $\theta_m \neq 0$ for all m, satisfying

$$\lim(\mathscr{E}_{\theta_m}\omega_m - \mathscr{E}_{\theta_m}\varphi_{2,m}) = 10\delta. \tag{7.1.13}$$

Passing to a further subsequence if necessary, it is clear that at least one of the following cases must occur: (a) $\theta_m m^{\frac{1}{2}} \to \infty$ (or $-\infty$), (b) $\theta_m m^{\frac{1}{2}} \to 0$, or (c) $\theta_m m^{\frac{1}{2}} \to h \neq 0$.

Consider case (c) first. According to Theorem 6.1, we can restrict ourselves to the conditional expectation $\bar{\omega}_n$ of ω_n, given Δ_n. Thus, writing θ_m for $h_m m^{-\frac{1}{2}}$ where $h_m \to h$, we have

$$\mathscr{E}_{\theta_m}\bar{\omega}_m(\Delta_n) - \mathscr{E}_{\theta_m}\varphi_{2,m}(\Delta_m) > 9\delta \quad \text{for} \quad m > N_1. \tag{7.1.14}$$

Further, for every test function $\psi_n(\Delta_n)$, we have

$$|\mathscr{E}_{\theta_n}\psi_n(\Delta_n) - \mathscr{E}_{\theta_n}\psi_n(\Delta_n^*)| \leqq 2P_{n,\theta_n}(\Delta_n \neq \Delta_n^*). \tag{7.1.15}$$

By Proposition 4.1, the right-hand side of (7.1.15) with n replaced by m converges to zero.

Specializing this to $\bar{\omega}_n$, $\varphi_{2,n}$ and utilizing (7.1.14), we have

$$\mathscr{E}_{\theta_m}\bar{\omega}_m(\Delta_m^*) - \mathscr{E}_{\theta_m}\varphi_{2,m}(\Delta_m^*) > 8\delta \quad \text{for} \quad m > N_2. \tag{7.1.16}$$

By Theorem 6.2 and Proposition 4.1, we have $\mathscr{L}(\Delta_m^* \mid P_{m,\theta_m}) \to N(\Gamma h, \Gamma)$. Thus

$$\left|\mathscr{E}_{\theta_m}\varphi_{2,m}(\Delta_m^*) - \int_{|x|\geqq\xi_{\alpha/2}} d\Phi_h\right| < \delta \quad \text{for} \quad m > N_3, \tag{7.1.17}$$

where Φ_h denotes the distribution function of $N(\Gamma h, \Gamma)$. Select $\alpha_1 > \alpha$ such that

$$\left|\int_{|x|\geqq\xi_{\alpha/2}} d\Phi_h - \int_{|x|\geqq\xi_{\alpha_1/2}} d\Phi_h\right| < \delta. \tag{7.1.18}$$

For $m > N_4$, $\mathscr{E}_0[\bar{\omega}_m(\Delta_m^*)] < \alpha_1$ since $\mathscr{E}_0\bar{\omega}_m(\Delta_m^*) \to \alpha$.

The test $\varphi_n(\Delta_n^*)$ defined in (7.1.10) with α replaced by α_1 is most powerful unbiased of level α_1. (See, e.g., Lehmann [4] page 126.) Therefore it is most powerful among all unbiased tests of level $\leqq \alpha_1$. For each m, any test that does not reject outside of an interval can be improved, in terms of power, by a test of that form (see Ferguson, T. S. 1967, *Mathematical Statistics*, Academic Press for a specialized version of a more general result due to Karlin). The asymptotic unbiasedness then leads to the inequality $\mathscr{E}_{\theta_m}\varphi_m(\Delta_m^*) \geqq \mathscr{E}_{\theta_m}\bar{\omega}_m(\Delta_m^*)$ for $m > N_5$. Also $\mathscr{E}_{\theta_m}\varphi_m(\Delta_m^*) \to \int_{|x| \geqq \xi_{\alpha 1/2}} d\Phi_h$ according to the remark following Lemma 7.1.2, so that

$$\left|\mathscr{E}_{\theta_m}\varphi_m(\Delta_m^*) - \int_{|x| \geqq \xi_{\alpha 1/2}} d\Phi_h\right| < \delta \quad \text{for} \quad m > N_6.$$

This, together with (7.1.17) and (7.1.18), gives

$$\mathscr{E}_{\theta_m}\varphi_m(\Delta_m^*) - \mathscr{E}_{\theta_m}\varphi_{2,m}(\Delta_m^*) < 3\delta \quad \text{for} \quad m > N_7.$$

Finally combining this result with the fact that $\mathscr{E}_{\theta_m}\bar{\omega}_m(\Delta_m^*) - \mathscr{E}_{\theta_m}\varphi_m(\Delta_m^*) \leqq 0$, we get

$$\mathscr{E}_{\theta_m}\bar{\omega}_m(\Delta_m^*) - \mathscr{E}_{\theta_m}\varphi_{2,m}(\Delta_m^*) < 3\delta \quad \text{for} \quad m > N_8$$

which contradicts (7.1.16).

If case (a) holds, the contradiction is obtained from Assumptions 5 and 5′ since $\mathscr{E}_{\theta_m}\varphi_{2,m}(\Delta_m) \to 1$.

If case (b) holds, the argument given in Theorem 4.1 of Johnson and Roussas [2] leads to a contradiction of (7.1.13). Namely, an application of Proposition 3.1 from that paper shows that $\|P_{m,\theta_m} - P_{m,\theta_0}\| \to 0$ so that all tests have asymptotic power equal to their level.

The above proof shows that the test (7.1.4) is asymptotically locally most powerful unbiased without Assumptions 5 and 5′. That is,

COROLLARY 7.1.2. *Under Assumptions* 1–4, *the test* $\varphi_{2,n}$ *defined by* (7.1.4) *is asymptotically locally most powerful unbiased.* (See also Corollary 7.1.1.)

7.2. *k-dimensional* Θ. It is also possible to obtain results similar to Wald [9] for testing a simple hypothesis in the multi-parameter situation. On the basis of Theorem 6.1 and Theorem 6.3, one may construct a sequence of tests which is asymptotically optimal. The criterion could be that of best average power over a family of surfaces, best constant power over a family of surfaces, or that of a most stringent test. These problems will be treated in detail in a forthcoming paper.

REFERENCES

[1] DOOB, J. L. (1953). *Stochastic Processes.* Wiley, New York.

[2] JOHNSON, R. A. and ROUSSAS, G. G. (1969). Asymptotically most powerful tests in Markov processes. *Ann. Math. Statist.* **40** 1207–1215.

[3] LECAM, L. (1960). Locally asymptotically normal families of distributions. *Univ. California Publ. Statist.* **3** 37–98.

[4] LEHMANN, E. L. (1959). *Testing Statistical Hypotheses.* Wiley, New York.

[5] RAO, C. R. (1965). *Linear Statistical Inference and its Applications.* Wiley, New York.

[6] ROUSSAS, G. G. (1965). Asymptotic inference in Markov processes. *Ann. Math. Statist.* **36** 978–992.

[7] ROUSSAS, G. G. (1968). Some applications of likelihood functions to the asymptotic efficiency of estimates. *Z. Wahrscheinlichkeitstheorie und Verw. Gebiete* **10** 252–260.

[8] WALD, A. (1941). Asymptotically most powerful tests of statistical hypotheses. *Ann. Math. Statist.* **12** 1–19.

[9] WALD, A. (1943). Tests of statistical hypotheses concerning several parameters when the number of observations is large. *Trans. Amer. Math. Soc.* **54** 426–482.