The Annals of Mathematical Statistics
1969, Vol. 40, No. 4, 1207–1215

ASYMPTOTICALLY MOST POWERFUL TESTS IN MARKOV PROCESSES[1]

BY RICHARD A. JOHNSON AND GEORGE G. ROUSSAS

University of Wisconsin, Madison

1. Introduction and summary. Wald [8] treated the problem of testing $H_0: \theta = \theta_0$ against one-sided alternatives by giving conditions under which tests based on the maximum likelihood estimator are asymptotically most powerful. He defines a sequence of tests $\{\lambda_n\}$ to be an asymptotically most powerful test of H_0 against $H_1: \theta > \theta_0$, on level of significance α, if for any other sequence of tests $\{\omega_n\}$ of level α,

$$\limsup [\sup (\mathcal{E}_\theta \omega_n - \mathcal{E}_\theta \lambda_n ; \theta > \theta_0, \theta \in \Theta)] = 0, \tag{1.1}$$

where Θ is the parameter space. A similar expression holds for testing H_0 against $H_2: \theta < \theta_0$.

Wald's regularity conditions on the population density are quite strong. The maximum likelihood estimator, being the value which maximizes the likelihood function, is required to be a consistent estimator in the probability sense and the consistency must be uniform over certain intervals in Θ. Also, his conditions imply that the centered and scaled version converges uniformly in law, over certain intervals in Θ, to the standard normal. Wald formulates tests in terms of the maximum likelihood estimate.

In the present work, we extend Wald's results in two directions. First, the regularity conditions are substantially weakened through use of the techniques of LeCam and the tests need not be based on the maximum likelihood estimate. Secondly, tests concerning the parameter in the joint distribution of the random variables involved are shown to be asymptotically most powerful when the observations arise from a stationary Markov process. In order that an α-level test exist for any α, it was necessary to consider tests which are possibly randomized.

Section 2 contains the basic assumptions on the Markov process and the preliminary results appear in Section 3. The results through Section 3 hold for a k-dimensional parameter space and are presented in this general formulation. The main results are presented as Theorems 4.1 and 4.2 in Section 4. The following section treats the special case of independent identically distributed random variables. Four examples are presented in Section 6.

In order to avoid unnecessary repetition in this paper, all limits will be taken as the sequence $\{n\}$ of positive integers, or a subsequence, converges to infinity.

The present authors hope to be able to report soon on results concerning k-dimensional parameter versions of the main theorems in this paper.

Received 3 April 1968; revised 7 October 1968.
[1] This paper was prepared with partial support of the National Science Foundation Grant GP-8884.

2. Notation and assumptions. Set $(\mathfrak{X}, \mathfrak{A}) = \times_{i=0}^{\infty}(R_i, \mathfrak{B}_i)$, where $(R_i, \mathfrak{B}_i) = (R, \mathfrak{B})$ denotes the Borel real line. The parameter space Θ is assumed to be an open subset of the k-dimensional Euclidean space $\mathcal{E}_k$. According to Kolmogorov's consistency theorem, a probability measure P_θ will be induced on $\mathfrak{A}$ by a probability distribution $p_\theta(\cdot)$ defined on $\mathfrak{B}$ and a transition probability measure $p_\theta(\cdot, \cdot)$ defined on $R \times \mathfrak{B}$. For each $\theta \varepsilon \Theta$, the coordinate process $\{X_n, n \geqq 0\}$, n an integer, is a Markov process with initial measure $p_\theta(\cdot)$ and stationary transition measure $p_\theta(\cdot, \cdot)$.

Let $\mathfrak{A}_n$ denote the σ-field induced by the random variables $\{X_0, X_1, \cdots, X_n\}$ and let $P_{n,\theta}$ denote the restriction of P_θ to $\mathfrak{A}_n$. It will be assumed below that the probability measures $\{P_{n,\theta}, \theta \varepsilon \Theta\}$, $n \geqq 0$, are mutually absolutely continuous. It follows that the derivatives $[dP_{0,\theta'}/dP_{0,\theta}] = q(X_0; \theta, \theta')$ and $[dP_{1,\theta'}/dP_{1,\theta}] = q(X_0, X_1; \theta, \theta')$ are well defined for every $\theta, \theta' \varepsilon \Theta$. Also set $q(X_1 \mid X_0, \theta, \theta') = q(X_0, X_1; \theta, \theta')/q(X_0; \theta, \theta')$. For the joint measures $P_{n,\theta'}$ and $P_{n,\theta}$, we then have $[dP_{n,\theta'}/dP_{n,\theta}] = q(X_0; \theta, \theta') \prod_{j=1}^{n} q(X_j \mid X_{j-1}; \theta, \theta')$ on a set having P_θ measure one for all $\theta \varepsilon \Theta$. In what follows, discussion will be restricted to that set.

We may write $[dP_{n,\theta'}/dP_{n,\theta}] = q(X_0; \theta, \theta') \prod_{j=1}^{n} \phi_j^2(\theta, \theta')$, where $\phi_j(\theta, \theta') = [q(X_j \mid X_{j-1}; \theta, \theta']^{\frac{1}{2}}$ each $j = 1, 2, \cdots, n$. Clearly, $\int \phi_1^2(\theta, \theta')\, dP_{1,\theta} = 1$. It will also prove convenient to introduce the notation $f_j(\theta, \theta') = [q(X_{j-1}, X_j; \theta, \theta')]^{\frac{1}{2}}$ each $j = 1, 2, \cdots, n$.

We now state the main assumptions.

Assumptions. (A1) For each $\theta \varepsilon \Theta$, the Markov process $\{X_n, n \geqq 0\}$ is stationary and metrically transitive (ergodic). (See, e.g., p. 460 [1]).

(A2) The probability measures $\{P_{n,\theta}, \theta \varepsilon \Theta\}$, $n \geqq 0$, are mutually absolutely continuous.

(A3) (i) For every $\theta \varepsilon \Theta$, the random function $\phi_1(\theta, \theta')$ is differentiable in quadratic mean (q.m.) with respect to θ' at (θ, θ) when P_θ is employed. (See, e.g., p. 470 [4], or p. 39 [2]). (ii) Let $\dot{\phi}_1(\theta)$ denote the derivative in q.m. of $\phi_1(\theta, \theta')$ with respect to θ' at (θ, θ). Then $\dot{\phi}_1(\theta)$ is $\mathfrak{A}_2 \times \mathfrak{C}$-measurable, where $\mathfrak{C}$ denotes the σ-field of Borel subsets of Θ.

Let $\Gamma(\theta)$ be the covariance function defined by $\Gamma(\theta) = 4\mathcal{E}_\theta\{\dot{\phi}_1(\theta) \cdot \dot{\phi}_1'(\theta)\}$, where $\dot{\phi}_1'$ is the transpose of $\dot{\phi}_1$.

(A4) For every $\theta \varepsilon \Theta$, the random function $f_1(\theta, \theta')$ is continuous in $P_{n,\theta}$-probability at (θ, θ) as $\theta' \to \theta$.

From (A2), it follows that for an arbitrary but fixed $\theta \varepsilon \Theta$ and $\theta_n \varepsilon \Theta$, $n = 1, 2, \cdots$, $[dP_{n,\theta_n}/dP_{n,\theta}] = q(X_0; \theta, \theta_n) \prod_{j=1}^{n} \phi_j^2(\theta, \theta_n)$ is well-defined except on a $P_{\theta'}$-null set for all $\theta' \varepsilon \Theta$. Disregarding this null set, we define the random variable $\Lambda[P_{n,\theta_n}; P_{n,\theta}]$ by

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

3. Preliminaries. We proceed by first recalling some known results.

Theorem 3.1. *Let $h_n, h \varepsilon \mathcal{E}_k$, where $h_n \to h$. Under assumptions* (A1) — (A3) (ii) *and* (A4),

$$\Lambda[P_{n,\theta+h_n n^{-1/2}}; P_{n,\theta}] - h'\Delta_n(\theta) \to -A(h, \theta)$$

in $P_{n,\theta}$*-probability, where* $A(h, \theta) = \frac{1}{2}h'\Gamma(\theta)h = 2\,\mathcal{E}_\theta|h'\dot\phi_1(\theta)|^2$ *and* $\Delta_n(\theta) = 2n^{-\frac{1}{2}}\sum_{j=1}^{n}\dot\phi_j(\theta)$.

This is a restatement of Theorem 4.1.1, page 980, of [5]. This result enables us to obtain the limiting distribution of $[P_{n,\theta+h_n n^{-\frac{1}{2}}}; P_{n,\theta}]$ under $P_{n,\theta}$.

THEOREM 3.2. *Under the conditions of Theorem* 3.1,

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

and consequently,

$$\mathcal{L}\{\Lambda[P_{n,\theta+h_n n^{-\frac{1}{2}}}; P_{n,\theta}] \mid P_{n,\theta}\} \to N(-\tfrac{1}{2}h'\Gamma(\theta)h,\ h'\Gamma(\theta)h).$$

This is Theorem 3.2. in [5], p. 984.

Next, we note that our main assumptions allow us to conclude that certain sequences of measures are contiguous. (For the definition of this concept, see p. 41 [2]).

LEMMA 3.1. *Under the conditions of Theorem* 3.1, *the sequences of measures* $\{P_{n,\theta}\}$ and $\{P_{n,\theta+h_n n^{-\frac{1}{2}}}\}$ *are contiguous.*

PROOF. According to Theorem 3.2, $\{\Lambda[P_{n,\theta+h_n n^{-\frac{1}{2}}}; P_{n,\theta}] \mid P_{n,\theta}\}$ converges to $N(-\frac{1}{2}h'\Gamma(\theta)h,\ h'\Gamma(\theta)h)$. Denote this limit by $\mathcal{L}(\chi)$. Then $\mathcal{E}[\exp(\chi)] = (2\pi)^{-\frac{1}{2}}\sigma^{-1}\int_{-\infty}^{\infty}\exp[-(x - \sigma^2/2)^2/2\sigma^2]\,dx = 1$, where σ^2 equals $h'\Gamma(\theta)h > 0$. By Theorem 2.1, page 40, of [2], this result is equivalent to the statement that the sequence of measures $\{P_{n,\theta}\}$ and $\{P_{n,\theta+h_n n^{-\frac{1}{2}}}\}$ are contiguous. The case where $h = 0$ is covered by Proposition 3.1 below.

The random variables $\Lambda[P_{n,\theta+h_n n^{-\frac{1}{2}}}; P_{n,\theta}]$, $n \geqq 0$, also have a limiting distribution under the moving measures.

THEOREM 3.3. *Under the conditions of Theorem* 3.1,

$$\mathcal{L}\{\Lambda[P_{n,\theta+h_n n^{-\frac{1}{2}}}; P_{n,\theta}] \mid P_{n,\theta+h_n n^{-\frac{1}{2}}}\} \to N(\tfrac{1}{2}h'\Gamma(\theta)h,\ h'\Gamma(\theta)h).$$

This is Theorem 2 of [7].

Although the results contained in the following two lemmas are known, there is no convenient reference.

LEMMA 3.2. *Let* $\{Y_n\}$ *and* $\{Z_n\}$ *be two sequences of random variables satisfying*

$$\lim \mathcal{L}[Y_n \mid Q] = \lim \mathcal{L}[Z_n \mid Q] = N(\mu, \sigma^2), \qquad \sigma^2 > 0.$$

For a fixed α *with* $0 < \alpha < 1$, *let the sequences of real numbers* $\{c_n\}$ *and* $\{d_n\}$ *be defined as the* sup *over numbers* c *and* d *satisfying* $1 - \alpha \geqq P[Y_n \leqq c \mid Q]$ *and* $1 - \alpha \geqq P[Z_n \leqq d \mid Q]$ *respectively. Then* $\lim c_n$ *is finite and* $\lim (c_n - d_n) = 0$.

PROOF. Since $\lim P[Y_n \leqq y \mid Q] = \Phi[(y - \mu)/\sigma]$, it follows that the convergence is uniform in $y \in R$. Consequently, c_n converges to $\mu + \sigma\xi_\alpha$, where ξ_α is the upper α-th quantile of Φ. The same argument applied to $\{Z_n\}$ yields $\lim (c_n - d_n) = 0$.

LEMMA 3.3. *Let* $\{Y_n\}$, $\{Z_n\}$, $\{c_n\}$ *and* $\{d_n\}$ *be defined as in Lemma* 3.2. *Assume that* $\lim \mathcal{L}[Y_n \mid Q_n] = \lim \mathcal{L}[Z_n \mid Q_n] = N(\mu, \sigma^2)$, $\sigma^2 > 0$. *Then*

$$P[Y_n \leqq c_n \mid Q_n] - P[Z_n \leqq d_n \mid Q_n] \to 0.$$

PROOF. Clearly, $P[Y_n \leqq y \mid Q_n]$ and $P[Z_n \leqq y \mid Q_n]$ converge to $\Phi[(y - \mu)/\sigma]$ uniformly in $y \varepsilon R$. Thus, $|P[Y_n \leqq c_n \mid Q_n] - P[Z_n \leqq d_n \mid Q_n]| \leqq |P[Y_n \leqq c_n \mid Q_n] - \Phi[(c_n - \mu)/\sigma]| + |P[Z_n \leqq d_n \mid Q_n] - \Phi[(d_n - \mu)/\sigma]| + |\Phi[(c_n - \mu)/\sigma] - \Phi[(d_n - \mu)/\sigma]|$. Repeating the argument of Lemma 3.2 for the sequences of cdf's $P[Y_n \leqq c_n \mid Q_n]$ and $P[Z_n \leqq d_n \mid Q_n]$ determined by the moving measures, we obtain the desired result.

The following result was presented without proof by LeCam [3].

PROPOSITION 3.1. *Let P and Q be two probability measures on a σ-field $\mathfrak{A}$. Let Z be the logarithm of the likelihood ratio of Q relative to P. For every $\epsilon > 0$,*

$$\|P - Q\| \leqq 2(1 - e^{-\epsilon}) + 2P[|Z| > \epsilon],$$

where $\|\cdot\|$ is the norm associated with convergence in variation.

PROOF. Set $B = [f - g > 0]$ where f and g are the densities of P and Q, respectively, with respect to $P + Q$. Then

$$\|P - Q\| = 2 \sup (|P(A) - Q(A)|; A \varepsilon \mathfrak{A}) = 2[P(B) - Q(B)].$$

Set $D = [|Z| > \epsilon]$ and note that $P(B) - Q(B) = P(B \cap D) - Q(B \cap D) + P(B \cap D^c) - Q(B \cap D^c) \leqq P(D) + P(B \cap D^c) - \int_{B \cap D^c} e^z \, dP$. On D^c, we have $-\epsilon \leqq Z \leqq \epsilon$ and consequently $P(B \cap D^c) - \int_{B \cap D^c} e^z \, dP \leqq 1 - e^{-\epsilon}$. The proof is complete.

4. Main results.

Although the preceding developments are for a k-dimensional parameter space, to obtain asymptotically most powerful one-sided tests, we restrict ourselves to the case $k = 1$ or equivalently it is assumed that Θ is an open subset of R. The quantity $\Delta_n(\theta)$ is then a real-valued random variable with asymptotic variance $\sigma^2(\theta)$ where $\sigma^2(\theta) = 4\,\mathcal{E}_\theta \mid \dot{\phi}(\theta)|^2$.

Let $\theta_0 \varepsilon \Theta$ and define the sequence of critical functions $\{\varphi_n\}$ by

$$(4.1) \qquad \begin{aligned} \varphi_n(X_0, X_1, \cdots, X_n) &= 1, && \text{if } \Delta_n(\theta_0) > c_n^*, \\ &= \gamma_n, && \text{if } \Delta_n(\theta_0) = c_n^*, \\ &= 0, && \text{otherwise,} \end{aligned}$$

where the sequences $\{c_n^*\}$ and $\{\gamma_n\}$ are determined by the requirement $\mathcal{E}_{\theta_0}\varphi_n = \alpha$ $(0 < \alpha < 1)$.

We require one further assumption.

(A5) Let $\{\theta_n\}$ be a sequence of elements of Θ with $\theta_n > \theta_0$ for each n. The condition $\lim n^{\frac{1}{2}}(\theta_n - \theta_0) = \infty$ implies that $\Delta_n(\theta_0) \rightarrow \infty$ in P_{θ_n}-probability.

THEOREM 4.1. *Under assumptions* (A1)–(A5), *the test φ_n defined by* (4.1) *is asymptotically most powerful for testing $H_0:\theta = \theta_0$ against the alternative $H_1:\theta > \theta_0, \theta \varepsilon \Theta$.*

PROOF. Let $\{\varphi_n\}$ be defined by (4.1) and let $\{\omega_n\}$ be any other sequence of α-level tests. Assume, for contradiction, that the right hand side of (1.1) takes the value $\delta > 0$. Then there exist sequences $\{n'\} \subset \{n\}$ and $\{\theta_{n'}\}$ with $\theta_{n'} \varepsilon \Theta$

and $\theta_{n'} > \theta$ for all n', satisfying

$$\lim \left(\mathcal{E}_{\theta_{n'}}\omega_{n'} - \mathcal{E}_{\theta_{n'}}\varphi_{n'}\right) = \delta. \tag{4.2}$$

Consider the sequence $\{n'^{\frac{1}{2}}(\theta_{n'} - \theta_0)\}$. Suppose first that it is unbounded. Then there exists a subsequence $\{n''\} \subset \{n'\}$ with $[n'']^{\frac{1}{2}}(\theta_{n''} - \theta_0) \to \infty$. According to (A5), $\Delta_{n''}(\theta_0) \to \infty$ in $P_{\theta_{n''}}$-probability and consequently $\lim \mathcal{E}_{\theta_{n''}}\varphi_{n''} = \lim P_{\theta_{n''}}[\Delta_{n''}(\theta_0) > c^*_{n''}] = 1$ by Theorem 3.2 and Lemma 3.2 which show that the $c^*_{n''}$ are bounded. Thus $\{(n')^{\frac{1}{2}}(\theta_{n'} - \theta_0)\}$ must be bounded.

If this last sequence is bounded, there exists a subsequence $\{m\} \subset \{n'\}$ such that $\lim m^{\frac{1}{2}}(\theta_m - \theta_0) = t \geqq 0$. First consider the case $t > 0$. Setting $m^{\frac{1}{2}}(\theta_m - \theta_0) = t_m$, we write $\theta_m = \theta_0 + t_m m^{-\frac{1}{2}}$ where $t_m \to t$. Make the identification $Y_m = t\Delta_m(\theta_0)$, $Z_m = \Lambda[P_{m,\theta_0+t_m m^{-\frac{1}{2}}}; P_{m;\theta_0}] + \frac{1}{2}t^2\sigma^2(\theta_0)$, $Q = P_{\theta_0}$ and $c_m = tc_m^*$ in the statement of Lemma 3.2. We have set $\sigma^2(\theta_0) = 4\mathcal{E}_{\theta_0}|\dot{\phi}_1(\theta_0)|^2$. Also define the sequence of critical functions $\{\psi_n\}$ by

$$\begin{aligned} \psi_n(X_0, X_1, \cdots, X_n) &= 1, && \text{if } \Lambda[P_{n,\theta_0+t_n n^{-\frac{1}{2}}}; P_{n,\theta_0}] + \tfrac{1}{2}t^2\sigma^2(\theta_0) > d_n \\ &= \gamma_n^*, && \text{if } \Lambda[P_{n,\theta_0+t_n n^{-\frac{1}{2}}}; P_{n,\theta_0}] + \tfrac{1}{2}t^2\sigma^2(\theta_0) = d_n \\ &= 0, && \text{otherwise,} \end{aligned} \tag{4.3}$$

where the sequences $\{d_n\}$ and $\{\gamma_n^*\}$ are determined by the requirement $\mathcal{E}_{\theta_0}\psi_n = \alpha$.

Theorem 3.1 together with Lemma 3.1 and Theorem 3.3 give

$$\mathcal{L}[t\Delta_m(\theta_0) \mid P_{m,\theta_0+t_m m^{-\frac{1}{2}}}] \to N(t^2\sigma^2(\theta_0), t^2\sigma^2(\theta_0)).$$

This result together with Theorem 3.2 verifies that Lemma 3.3 is applicable with $\{Y_m\}$, $\{Z_m\}$, $\{c_m\}$ and $\{d_m\}$ defined above with Q replaced by P_{θ_m}. Using $\lim P_{\theta_m}[Z_m = d_m] = \lim P_{\theta_m}[Y_m = c_m] = 0$,

$$\lim \left(\mathcal{E}_{\theta_m}\varphi_m - \mathcal{E}_{\theta_m}\psi_m\right) = 0. \tag{4.4}$$

Also, since $\mathcal{E}_{\theta_{n'}}\omega_{n'} - \mathcal{E}_{\theta_{n'}}\varphi_{n'}$ tends to δ, it follows that

$$\lim \left(\mathcal{E}_{\theta_m}\omega_m - \mathcal{E}_{\theta_m}\varphi_m\right) = \delta. \tag{4.5}$$

From (3.4) and (3.5), we have

$$\lim \left(\mathcal{E}_{\theta_m}\omega_m - \mathcal{E}_{\theta_m}\psi_m\right) = \delta. \tag{4.6}$$

Therefore, for all sufficiently large m

$$\mathcal{E}_{\theta_m}\psi_m \leqq \mathcal{E}_{\theta_m}\omega_m - \delta/2. \tag{4.7}$$

However, for testing $H_0:\theta = \theta_0$ against a specific $\theta_m > \theta$, ψ_m is most powerful according to the Neyman-Pearson fundamental lemma. This contradicts $m^{\frac{1}{2}}(\theta_m - \theta_0) \to t > 0$.

Finally, suppose the only convergent subsequence satisfies $m^{\frac{1}{2}}(\theta_m - \theta_0) \to t = 0$. Set $Z_m = \Lambda[P_{m,\theta_0+t_m m^{-\frac{1}{2}}}; P_{m,\theta_0}]$. According to Theorem 3.1, $P_{\theta_0}[|Z_m| > \epsilon] \to 0$ for every $\epsilon > 0$. Applying Proposition 3.1 with $Q = P_{m,\theta_m}$, $P = P_{m,\theta_0}$ and $Z = Z_m$, it follows that $\|P_{m,\theta_0} - P_{m,\theta_m}\| \to 0$. Also, the power of the test (4.1) is asymptotically α since

$$(4.8)\quad |P_{m,\theta_0}[\varphi_m > c_m^*] + \gamma_m P_{m,\theta_0}[\varphi_m = c_m^*] - P_{m,\theta_m}[\varphi_m > c_m^*] - \gamma_m P_{m,\theta_m}[\varphi_m = c_m^*]| \leqq 2\,\|P_{m,\theta_0} - P_{m,\theta_m}\|.$$

A similar calculation shows that the test (4.3) also has asymptotic power α and consequently $\mathcal{E}_{\theta_m}\varphi_m - \mathcal{E}_{\theta_m}\psi_m \to 0$. As above, this is a contradiction to (4.2) since ψ_m is most powerful. This concludes the proof.

It is clear from the previous proof that even without assumption (A5), the test based on Δ_n is locally best.

COROLLARY 4.1. *Under the assumptions* (A1)–(A4), *the test* φ_n *defined by* (4.1) *is asymptotically locally most powerful.*

To test the hypothesis $H_0:\theta = \theta_0$ against the hypothesis $H_2:\theta < \theta_0$, we need to modify the last assumption.

(A′5) Let $\{\theta_n\}$ be a sequence of elements of Θ with $\theta_n < \theta_0$ for each n. The condition $\lim n^{\frac{1}{2}}(\theta_n - \theta_0) = -\infty$ implies that $\Delta_n(\theta_0) \to -\infty$ in P_{θ_n}-probability.

Define a sequence of tests $\{\varphi_n'\}$ by

$$(4.9)\quad \begin{aligned} \varphi_n'(X_0, X_1, \cdots, X_n) &= 1, && \text{if } \Delta_n(\theta_0) < c_n', \\ &= \gamma_n', && \text{if } \Delta_n(\theta_0) = c_n', \\ &= 0, && \text{otherwise,} \end{aligned}$$

where the sequences $\{c_n'\}$ and $\{\gamma_n'\}$ are determined by $\mathcal{E}_{\theta_0}\varphi_n' = \alpha$ for a fixed $\alpha(0 < \alpha < 1)$.

The following two results are verified in the same manner as Theorem 4.1 and Corollary 4.1.

THEOREM 4.2. *Under assumptions* (A1)–(A4) *and* (A′5), *the test defined by* (4.9) *is asymptotically most powerful for testing* $H_0:\theta = \theta_0$ *against the alternative* $H_2:\theta < \theta_0$.

COROLLARY 4.2. *Under assumptions* (A1)–(A4), *the test defined by* (4.9) *is asymptotically locally most powerful for testing* $H_0:\theta = \theta_0$ *against the alternative* $H_2:\theta < \theta_0$.

5. Independent identically distributed case. Because of its special interest, we present the simplified assumptions for the case when the random variables $\{X_n, n \geqq 0\}$ are independently distributed as X_0 where X_0 has probability density $[dP_{0,\theta'}/dP_{0,\theta}] = q(X_0; \theta, \theta')$ with respect to $P_{0,\theta}$.

Condition (A1) is always satisfied and (A4) will be implied by (A3)(i) in the obvious manner. The simplified assumptions become

(B2) The probability measures $\{P_{n,\theta}, \theta \,\varepsilon\, \Theta\}, n \geqq 0$, are mutually absolutely continuous.

(B3) (i) For every $\theta \,\varepsilon\, \Theta$, the random function $[q(X_0; \theta, \theta')]^{\frac{1}{2}}$ is differentiable in quadratic mean with respect to θ' at (θ, θ) when P_θ is employed.

(ii) Let $\phi_1(\theta)$ denote the derivative of $[q(X_0; \theta, \theta')]^{\frac{1}{2}}$ with respect to θ' at

(θ, θ). Then ϕ_1 is $\mathcal{A}_2 \times \mathcal{C}$-measurable, where $\mathcal{C}$ denotes the σ-field of Borel subsets of Θ.

It is then the quadratic mean derivative of $[q(X_0\,;\,\theta,\theta')]^{\frac{1}{2}}$ which enters Δ_n in the statement of assumptions (A5) and (A′5).

6. Examples. Four examples are presented. The conditions (A1)–(A4) have been verified for the first three examples in [6] and it remains to verify assumption (A5) or (A′5). In establishing the differentiability assumption, Vitali's theorem has been used. This theorem states that, for $r > 0$, $Y_n \to Y$ in the rth mean if and only if $Y_n \to Y$ in probability and $\mathcal{E}\,|Y_n|^r \to \mathcal{E}\,|Y|^r$ finite.

In the first two examples, UMP tests exist for each n.

EXAMXLE 1. $\{X_n\}$, $n > 0$, are independent with common distribution $N(\theta, 1)$, $\theta \in R$. Here

$$\Delta_n(\theta_n) = n^{-\frac{1}{2}} \textstyle\sum_{j=1}^n (X_j - \theta_0) = n^{-\frac{1}{2}} \sum_{j=1}^n (X_j - \theta_n) + n^{\frac{1}{2}}(\theta_n - \theta_0).$$

Under P_{θ_n}, the first term is distributed as $N(0, 1)$ and consequently $\Delta_n(\theta_0) \to \infty\,(-\infty)$ in P_{θ_n}-probability whenever $n^{\frac{1}{2}}(\theta_n - \theta_0) \to \infty\,(-\infty)$.

EXAMPLE 2. $\{X_n\}$, $n \geqq 0$, are independent with common distribution $N(0, \theta^2)$, $\theta > 0$. In this example,

$$\begin{aligned}\Delta_n(\theta_0) &= n^{-\frac{1}{2}} \textstyle\sum_{j=1}^n (-\theta_0^{-1} + \theta_0^{-3}X_j^2)\\ &= \theta_0^{-3}n^{-\frac{1}{2}} \textstyle\sum_{j=1}^n (X_j^2 - \theta_n^2) + \theta_0^{-3}n^{\frac{1}{2}}(\theta_n^2 - \theta_0^2).\end{aligned}$$

The first term $\theta_0^{-3}\theta_n^2 n^{-\frac{1}{2}} \sum_{j=1}^n [(X_j/\theta_n)^2 - 1]$ is bounded in P_{θ_n}-probability according to the Berry-Esseen theorem whenever $\{\theta_n\}$ is bounded. For $\{\theta_n\}$ unbounded and $\theta_n > \theta_0$, $\Delta_n(\theta_0)$ is stochastically larger than under the alternatives $\{\theta_0 + n^{-\frac{1}{2}}\}$. Thus $\Delta_n(\theta_0) \to \infty\,(-\infty)$ as $n^{\frac{1}{2}}(\theta_n - \theta_0) \to \infty\,(-\infty)$.

EXAMPLE 3. $\{X_n\}$, $n \geqq 0$, are normally distributed with $\mathcal{E}_\theta X_n = 0$ and covariance function $R(m, n) = \mathcal{E}_\theta(X_m X_n) = \exp[-\theta\,|m - n|]$, $\theta > 0$. Set $\rho = \rho(\theta) = \exp(-\theta)$.

We make use of the following result:

$$\mathcal{E}_\theta(X_{i_1}X_{i_2}X_{i_3}X_{i_4}) = 2\rho(\theta)^{i_4+i_3-i_2-i_1} + \rho(\theta)^{i_4-i_3+i_2-i_1}$$

for $0 \leqq i_1 \leqq i_2 \leqq i_3 \leqq i_4$. Setting $\rho(\theta_0) = \rho_0$ and $\rho(\theta_n) = \rho_n$, we use the expression for Δ_n on page 43 of [6] to write

$$\begin{aligned}\Delta_n(\theta_0) = n^{-\frac{1}{2}}\rho_0(1 - \rho_0^2)^{-2}\{&-(1 + \rho_0^2)[\textstyle\sum_{j=1}^n (X_{j-1}X_j - \rho_n)]\\ &+ \rho_0[\textstyle\sum_{j=1}^n (X_{j-1}^2 - 1)] + \rho_0[\sum_{j=1}^n (X_j^2 - 1]\\ &+ (1 + \rho_0^2)n(\rho_0 - \rho_n) + \rho_0^3 - \rho_0\}.\end{aligned}$$

Using the inequality $P[|Z| > M] \leqq M^{-2}\operatorname{Var}(Z)$, provided $\mathcal{E}Z = 0$, and bounding the variance of each of the first three terms, we show that each is bounded in P_{θ_n}-probability. More precisely,

$$\operatorname{Var}(\textstyle\sum_{j=1}^n X_j^2) = \operatorname{Var}(\sum_{j=1}^n X_{j-1}^2) \leqq n[2 + 4\rho_n^2(1 - \rho_n^2)^{-1}]$$

and $$\operatorname{Var}(\textstyle\sum_{j=1}^n X_jX_{j-1}) \leqq n[1 + \rho_n^2 + 4\rho_n^2(1 - \rho_n^2)^{-1}].$$

For testing $\theta = \theta_0$ against $\theta > \theta_0$ or $\rho = \rho_0$ against $\rho < \rho_0$, the variances of $-\rho_0(1 - \rho_0^2)^{-1}n^{-\frac{1}{2}} \sum_{j=1}^n (X_jX_{j-1} - \rho)$, $\rho_0^2(1 - \rho_0^2)^{-2}n^{-\frac{1}{2}} \sum_{j=1}^n (X_j^2 - 1)$ and $\rho_0^2(1 - \rho_0^2)^{-2}n^{-\frac{1}{2}} \sum_{j=1}^n (X_{j-1}^2 - 1)$ are each bounded by $[2 + 4\rho_0^2(1 - \rho_0^2)^{-1}]\cdot \rho_0^2(1 - \rho_0^2)^{-4}$. Since $n^{\frac{1}{2}}(\rho_0 - \rho_n) \geqq e^{-\theta_0}n^{\frac{1}{2}}(\theta_n - \theta_0)$, the condition $n^{\frac{1}{2}}(\theta_n - \theta_0) \to \infty$ implies that $\Delta_n \to \infty$ in P_{θ_n}-probability. A similar argument establishes (A′5) when ρ is bounded away from one under the alternative.

EXAMPLE 4. $\{X_n, n \geqq 0\}$ are independent with common probability density, the double exponential $p(x; \theta) = \frac{1}{2} \exp[-|x - \theta|]$, $x \in R$, $\theta \in R$. Here

$$\phi_1(\theta, \theta') = \exp[\tfrac{1}{2}|X_1 - \theta| - \tfrac{1}{2}|X_1 - \theta'|].$$

Now for each $\theta \in \Theta$, define the random variables

$$\begin{aligned} Z_j(\theta) &= -\tfrac{1}{2} \quad \text{if } X_j < \theta, \\ &= 0 \quad \text{if } X_j = \theta, \qquad j = 0, 1, \cdots, n. \\ &= \tfrac{1}{2} \quad \text{if } X_j > \theta, \end{aligned}$$

Then $\mathcal{E}_\theta Z_j(\theta) = \mathcal{E}_\theta Z_1(\theta) = 0$ and $\mathcal{E}_\theta Z_j^2(\theta) = \mathcal{E}_\theta Z_1^2(\theta) = \frac{1}{4}$, since $P_\theta(X_1 < \theta) = P_\theta(X_1 > \theta) = \frac{1}{2}$. Furthermore, it is easily seen that

$$(6.1) \quad h^{-1}[\phi_1(\theta, \theta + h) - 1] \to Z_1(\theta) \quad \text{in } P_\theta\text{-probability as } (0 \neq)h \to 0.$$

Next

$$(6.2) \qquad \mathcal{E}_\theta\{h^{-1}[\phi_1(\theta, \theta + h) - 1]\}^2 = 2h^{-2}[1 - \mathcal{E}_\theta\phi_1(\theta, \theta + h)],$$

where $\mathcal{E}_\theta\phi_1(\theta, \theta + h) = \frac{1}{2}(2 - h)\cdot\exp(h/2)$ if $h < 0$

and $\mathcal{E}_\theta\phi_1(\theta, \theta + h) = \frac{1}{2}(2 + h)\cdot\exp(-h/2)$ if $h > 0$.

For $h < 0$, the right hand side of (6.2) becomes

$$2h^{-2}[1 - \exp(h/2)] + h^{-1}\exp(h/2)$$
$$= \tfrac{1}{2}h^2(1 - e^h) + \tfrac{1}{2}h^{-1}e^h \to -\tfrac{1}{4} + \tfrac{1}{2} = \tfrac{1}{4}, \qquad \text{as } h \to 0.$$

That is, for $h < 0$, $h \to 0$, we have

$$(6.3) \qquad \mathcal{E}_\theta\{h^{-1}[\phi_1(\theta, \theta + h) - 1]\}^2 \to \tfrac{1}{4}$$

and a similar expression holds for $h > 0$, $h \to 0$. Relations (6.1) and (6.3) imply that $\dot{\phi}_1(\theta)$ exists and is equal to $Z_1(\theta)$. Since the remaining portions of the assumptions (A1)–(A4) are clearly true, it only remains for us to verify (A5).

With P_θ-probability equal to one, we have

$$\dot{\phi}_j(\theta) = \tfrac{1}{2}I_{(\theta,\infty)}(X_j) - \tfrac{1}{2}I_{(-\infty,\theta)}(X_j) \quad \text{and} \quad \Delta_n(\theta) = 2n^{-\frac{1}{2}} \textstyle\sum_{j=1}^n \dot{\phi}_j(\theta).$$

Let $\theta_n \in \Theta$, $\theta_n > \theta$, and $n^{\frac{1}{2}}(\theta_n - \theta) \to \infty$. Since a location parameter is involved, it is sufficient to consider the case $\theta = 0$. Thus,

$$(6.4) \qquad \Delta_n(0) = n^{-\frac{1}{2}} \textstyle\sum_{j=1}^n \operatorname{sgn}(X_j).$$

Now $P_{\theta_n}[\text{sgn}\ (X_1) = 1] = P_{\theta_n}[X_1 > 0] = P_0[X_1 > -\theta_n] = P_0[\text{sgn}\ (X_1 + \theta_n) = 1]$ so that the P_{θ_n}-distribution of $\Delta_n(0)$ is identical to the P_0-distribution of $n^{-\frac{1}{2}} \sum_{j=1}^n \text{sgn}\ (X_j + \theta_n)$. Consequently,

$$P_{\theta_n}[\Delta_n(0) > M] = P_0[n^{-\frac{1}{2}} \sum_{j=1}^n \text{sgn}\ (X_j + \theta_n) > M] \tag{6.5}$$

and we need only estimate the right hand side.

Under P_0, the variables sgn $(X_j + \theta_n)$ have mean $1 - e^{-\theta_n}$ and variance 1. For each n, set

$$Y_{nj} = \text{sgn}\ (X_j + \theta_n) - 1 + e^{-\theta_n} \qquad \text{for} \quad j = 1, 2, \cdots, n, \tag{6.6}$$

where the Y_{nj} are, for fixed n, independent with $\mathcal{E}_0 Y_{nj} = 0$ and $\mathcal{E}_0 Y_{nj}^2 = 1$. Therefore

$$\mathcal{L}[n^{-\frac{1}{2}} \sum_{j=1}^n Y_{nj} \mid P_0] \to N(0, 1)$$

by the normal convergence criterion (p. 295 [4]) and hence $n^{-\frac{1}{2}} \sum_{j=1}^n Y_{nj}$ is bounded in probability. Clearly,

$$n^{-\frac{1}{2}} \sum_{j=1}^n \text{sgn}\ (X_j + \theta_n) = n^{-\frac{1}{2}} \sum_{j=1}^n Y_{nj} + n^{\frac{1}{2}}(1 - e^{-\theta_n}) \tag{6.7}$$

so that the right hand side of (6.5) converges to one if $n^{\frac{1}{2}}(1 - e^{-\theta_n}) \to \infty$. This latter condition holds since $n^{\frac{1}{2}}(1 - e^{-\theta_n}) \geqq n^{\frac{1}{2}}\theta_n/2$ when $\theta_n \leqq \ln 2$ and $n^{\frac{1}{2}}(1 - e^{-\theta_n}) \geqq n^{\frac{1}{2}}/2$ otherwise.

Assumption (A′5) may be verified in an entirely analogous manner.

7. Acknowledgment We would like to express our thanks to the referee and an Associate Editor for their comments which helped improve the presentation.

REFERENCES

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

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

[3] LECAM, L. (1966). Likelihood functions for large numbers of independent observations. *Research papers in statistics*. (F. N. David, editor). Wiley, New York.

[4] LOÈVE, M. (1963). *Probability Theory* (3rd Ed.). Van Nostrand, New York.

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

[6] ROUSSAS, G. G. (1965). Asymptotic inference in Markov Processes, II. *Bull. de la Soc. Mathématique de Grèce Nouvelle Série* **6** 37–50.

[7] ROUSSAS, G. G. (1967). Some applications of likelihood functions to the asymptotic efficiency of estimates. (to appear in Wahrscheinlichkeitstheorie). Technical Report No. 131, Univ. Wisconsin.

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