ADAPTIVE TESTS

by

K. S. Mak
Purdue University
Technical Report #88-38

PURDUE UNIVERSITY

DEPARTMENT OF STATISTICS
ADAPTIVE TESTS
by
K. S. Mak
Purdue University
Technical Report #88-38

Department of Statistics
Purdue University
August 1988

This research was supported in part by the Office of Naval Research Contract N00014-88-K-0170 and NSF Grants DMS-8606964, DMS-8702620A1 at Purdue University. Reproduction in whole or in part is permitted for any purpose of the United States Government.
Adaptive Tests
by
K. S. Mak
Purdue University

Abstract

The problem of hypothesis testing when the distribution is specified only up to a nuisance parameter is considered. A test is said to be adaptive if it is asymptotically optimal regardless of the value of the nuisance parameter. The exponential rate of convergence to zero of the probability of type II error when the probability of type I error converges to zero exponentially fast at a fixed rate is used as the optimal criterion. A necessary and sufficient condition for the existence of adaptive test is obtained.

Research supported by a David Ross Fellowship of Purdue University. This research was also supported in part by the Office of Naval Research Contract N00014-88-K-0170 and NSF Grants DMS-8606964, DMS-8702620A1 at Purdue University. Reproduction in whole or in part is permitted for any purpose of the United States Government.
1 Introduction

Let \( x = (x_1, \ldots, x_n) \) be \( n \) independent, identically, distributed observations on a random variable \( X \) having distribution \( P \) or \( Q \). It is desired to test the null hypothesis that \( X \) has distribution \( P \) versus the alternative that \( X \) has distribution \( Q \).

Let \( \phi_n = \phi_n(x) \) be any test function. Let \( A \) be a non-negative number. As in Tusnády (1977), a sequence of test functions \( \{\phi_n\} \) is said to have exponential rate \( A \) if

\[
\limsup_{n \to \infty} E_P \phi_n < 1, \quad \text{when } A = 0,
\]

\[
\limsup_{n \to \infty} n^{-1} \log E_P \phi_n \leq -A, \quad \text{when } A > 0. \tag{1.1}
\]

Let \( \Phi_A(P) \) be the set of all sequences of tests that have exponential rate \( A \). Let

\[
B(A, P, Q) = -\inf \left\{ \liminf_{n \to \infty} n^{-1} \log E_Q (1 - \phi_n) : \{\phi_n\} \in \Phi_A(P) \right\}. \tag{1.2}
\]

In other words, \( B(A, P, Q) \) is the optimal exponential rate at which the probability of type II error can converge to zero. A sequence of tests \( \{\phi_n\} \) in \( \Phi_A(P) \) is said to be asymptotically optimal if

\[
\liminf_{n \to \infty} n^{-1} \log E_Q (1 - \phi_n) = -B(A, P, Q). \]

Let the densities of \( P, Q \) be denoted by \( p(x), q(x) \) respectively. It is assumed that \( E_P (q(x)/p(x))^t < \infty \) for all \( t \). It is known that (see Bahadur (1971)) if \( \lim_{n \to \infty} E_P \phi_n = \alpha, 0 < \alpha < 1 \) (so that the sequence of tests \( \{\phi_n\} \) is of exponential rate 0), then \( B(0, P, Q) = K(P, Q) \) where \( K(P, Q) = E_P \log(p(x)/q(x)) \) is the Kullback-Leibler information number.

When \( A > 0 \), it is shown in Blahut (1974), Tusnády (1977), and Birgé (1981) that

\[
-B(A, P, Q) = \inf_{t \geq 0} \left\{ tC + \log E_Q \left( \frac{p(x)}{q(x)} \right)^t \right\}. \tag{1.3}
\]

where \( C \) is determined by

\[
-A = \inf_{t \geq 0} \left\{ -tC + \log E_P \left( \frac{q(x)}{p(x)} \right)^t \right\}. \tag{1.4}
\]

Furthermore, if \( 0 < A < K(Q, P) \), then \( B = B(A, P, Q) > 0 \) and

\[
A = B + C. \tag{1.5}
\]
A similar notion of asymptotic optimality for composite hypotheses has been investigated in Bahadur (1960), Hoeffding (1965), Brown (1971), Tusnády (1977) and Birgé (1981).

Now assume that the distributions $P$ and $Q$ are not determined exactly, but only up to a finite-valued nuisance parameter $r, r = 1, \ldots, l$. For example, a message, in one of $l$ possible languages using the same alphabets, is to be transmitted through a noisy channel $n$ times and a choice has to be made between two possible messages or rather the probability distributions associated with the messages, without knowing which language is used. Another example is that there are $l$ measurement types and for each type $\alpha$, the measurement has two possible distributions $P_\alpha$ and $Q_\alpha$. Thus, one has to test $P_\alpha$ versus $Q_\alpha$ with $\alpha$ unknown.

Let $A_1, \ldots, A_l$ be fixed positive numbers. A test $\phi_n^\alpha$ is called an adaptive test if it is an asymptotically optimal test of rate $A_\alpha$ for each value of the nuisance parameter $\alpha$. That is, $\phi_n^\alpha$ is adaptive if for each $\alpha$,

$$\limsup_{n \to \infty} n^{-1} \log E P_n^\alpha \phi_n^\alpha \leq -A_\alpha \quad \text{and} \quad \liminf_{n \to \infty} n^{-1} \log E Q_n^\alpha (1 - \phi_n^\alpha) = -B(A_\alpha, P_\alpha, Q_\alpha). \quad (1.6)$$

The existence of adaptive tests of rate 0 has been investigated in Rukhin (1982, 1986). A necessary and sufficient condition for the existence of such a test is that $K(P_\alpha, Q_\beta) \geq K(P_\beta, Q_\beta)$ for all $\alpha, \beta$. It is shown in section 2 that for adaptive tests to exist, any two distributions $P_\alpha, Q_\alpha, \alpha \neq \beta$ cannot be more 'difficult' to distinguish than $P_\alpha, Q_\alpha$. Here, 'difficulty' in distinguishing two distributions is measured by the rate of convergence of the type II error. When an adaptive test exists, it is shown that a weighted likelihood ratio test with weights depending on the rates of convergence of the type I and II errors is always adaptive. An overall maximum likelihood ratio test may not be adaptive.

2 Condition for the Existence of Adaptive Tests

Let $p_\alpha(x), q_\alpha(x)$ be the densities of $P_\alpha, Q_\alpha$ respectively. Let $A_1, \ldots, A_l$ be positive constants and denote $B(A_\alpha, P_\alpha, Q_\beta), C(A_\alpha, P_\alpha, Q_\beta)$ by $B_{\alpha\beta}, C_{\alpha\beta}$ respectively. Also, let $\psi(t|P_\alpha, Q_\beta) = E^P_\beta (q_\beta(x)/p_\alpha(x))'$. Assume that $\psi(t|P_\alpha, Q_\beta) < \infty$ for all $t$, for all $\alpha, \beta$. Note that $\psi(t|P_\alpha, Q_\beta)$ is the moment generating function of $\log(q_\beta(x)/p_\alpha(x))$, thus by the finiteness assumption, $\psi(t|P_\alpha, Q_\beta)$ is strictly convex.
and differentiable, indeed differentiation can be carried out under the expectation sign. Let

\[ T_n = T_n(x) = n^{-1} \log \frac{\max_a e^{n \beta \alpha} \prod_{j=1}^n q_{\beta}(x_j)}{\max_a e^{n \alpha \alpha} \prod_{j=1}^n p_{\alpha}(x_j)} \tag{2.1} \]

Let \( \phi_n = \phi_n(x) \) be a test with critical region given by \( T_n \geq 0 \), i.e.

\[ \phi_n = \begin{cases} 1 & \text{if } T_n \geq 0 \\ 0 & \text{otherwise}. \end{cases} \tag{2.2} \]

**Theorem 2.1** Assume that \( 0 < A_\alpha < \min_\beta K(Q_\beta, P_\alpha) \), for all \( \alpha \). If \( B_\beta \geq B_\alpha \) for all \( \beta, \alpha \), then \( \phi_n \) is an adaptive test.

**Proof.** We first show that \( \phi_n \) has exponential rate \( A_\alpha \) for every \( \alpha \). Since,

\[
E_p^\alpha \phi_n \leq P_\alpha \left( n^{-1} \log \frac{\max_\beta e^{n \beta \alpha} \prod_{j=1}^n q_{\beta}(x_j)}{e^{n \alpha \alpha} \prod_{j=1}^n p_{\alpha}(x_j)} \geq 0 \right)
\]

\[
\leq \sum_{\beta=1}^l P_\alpha \left( n^{-1} \log \frac{e^{n \beta \alpha} \prod_{j=1}^n q_{\beta}(x_j)}{e^{n \alpha \alpha} \prod_{j=1}^n p_{\alpha}(x_j)} \geq 0 \right)
\]

\[
\leq l \cdot \max_\beta P_\alpha \left( n^{-1} \sum_{j=1}^n \log \frac{q_{\beta}(x_j)}{p_{\alpha}(x_j)} \geq A_\alpha - B_\beta \right),
\]

and

\[
P_\alpha \left( n^{-1} \sum_{j=1}^n \log \frac{q_{\beta}(x_j)}{p_{\alpha}(x_j)} \geq A_\alpha - B_\beta \right) \leq e^{-nt(A_\alpha-B_\beta)} \left[ E_p^\beta \left( \frac{q_{\beta}(x)}{p_{\alpha}(x)} \right)^{t^n} \right]
\]

for any \( t > 0 \), it follows that,

\[
E_p^\alpha \phi_n \leq l \cdot \max_\beta \inf_{t>0} e^{-nt(A_\alpha-B_\beta)}(\psi(t|P_\alpha, Q_\beta))^n.
\]

Therefore,

\[
\limsup_{n \to \infty} n^{-1} \log E_p^\alpha \phi_n \leq \max_\beta \inf_{t>0} \left\{ -t(A_\alpha - B_\beta) + \log \psi(t|P_\alpha, Q_\beta) \right\}
\]

\[
\leq \max_\beta \inf_{t>0} \left\{ -t(A_\alpha - B_\alpha) + \log \psi(t|P_\alpha, Q_\beta) \right\}
\]

\[
= -A_\alpha
\]

as \( A_\alpha - B_\alpha = C_\alpha \) by (1.5).
Since $\phi_n$ is a test of exponential rate $A_\alpha$ for each $\alpha$, to show that $\phi_n$ is an adaptive test, it is enough to show that $\liminf_{n \to \infty} \log E_Q^\alpha (1 - \phi_n) \leq -B_{a\alpha}$. Now,
\[ E_Q^\alpha (1 - \phi_n) \leq \max_\beta Q_\alpha \left( n^{-1} \sum_{j=1}^n \log \frac{p_\beta(x_j)}{q_\alpha(x_j)} > B_{a\alpha} - A_\beta \right), \]
and so, by a similar argument, one obtains
\[
\liminf_{n \to \infty} n^{-1} \log E_Q^\alpha (1 - \phi_n) \leq \max_\beta \{ -t(B_{a\alpha} - A_\beta) + \log \psi(t|Q_\alpha, P_\beta) \}. \tag{2.3}
\]
Since the function $\log \psi(t|Q_\alpha, P_\beta)$ is strictly convex, there exist $0 < t_0 < 1$ such that $-B_{a\alpha} = t_0 CG_{a\alpha} + \log \psi(t_0|Q_\alpha, P_\beta)$. Therefore,
\[
\inf_{t>0} \{ -t(B_{a\alpha} - A_\beta) + \log \psi(t|Q_\alpha, P_\beta) \} \\
\leq -t_0(B_{a\alpha} - A_\beta) + \log \psi(t_0|Q_\alpha, P_\beta) \\
= -B_{a\alpha} + (1 - t_0)B_{a\alpha} + t_0A_\beta - B_{a\alpha} - t_0(A_\beta - B_{a\alpha}) \\
= -B_{a\alpha} + (1 - t_0)(B_{a\alpha} - B_{a\beta}) \\
\leq -B_{a\alpha}.
\]
Thus, from (2.3), $\liminf_{n \to \infty} n^{-1} \log E_Q^\alpha (1 - \phi_n) \leq -B_{a\alpha}$, for all $\alpha$ and $\phi_n$ is an adaptive test.

A necessary condition for the existence of adaptive tests is obtained by considering the asymptotic behaviour of the most powerful test of the simple hypothesis $\Pi_{j=1}^n p_\alpha(x_j)$ versus the simple alternative $w_\alpha \Pi_{j=1}^n q_\alpha(x_j) + w_\beta \Pi_{j=1}^n q_\beta(x_j), \alpha \neq \beta$, where $w_k = e^{nk}/(e^{nk} + e^{n\beta k}), k = \alpha, \beta$, and $b_\alpha, b_\beta$ are real numbers.

For the remainder of this section, we assume that $P_\alpha, Q_\alpha$ are absolutely continuous with respect to the Lebesgue measure for all $\alpha$. Let $\hat{Q}_n$ denote the distribution with density $w_\alpha \Pi_{j=1}^n q_\alpha(x_j) + w_\beta \Pi_{j=1}^n q_\beta(x_j)$. Let $\phi_n^\alpha$ be the following likelihood ratio test:
\[
\phi_n^\alpha = \begin{cases} 1 & \text{if } \frac{w_\alpha \Pi_{j=1}^n q_\alpha(x_j) + w_\beta \Pi_{j=1}^n q_\beta(x_j)}{\Pi_{j=1}^n p_\alpha(x_j)} \geq c_n \\ 0 & \text{otherwise} \end{cases}
\]
where $c_n$ is a constant.

Lemma 2.1 For a fixed $\alpha$, if $\lim_{n \to \infty} n^{-1} \log E_P^\alpha \phi_n^\alpha = -A_\alpha$, then
\[
\lim_{n \to \infty} n^{-1} \log E_{\hat{Q}_n} (1 - \phi_n^\alpha) = \max_{k=\alpha, \beta} (b_k + \rho_{a_k}(b_\alpha, b_\beta, C)) - \delta,
\]
where \( \rho_{ak}(b_\alpha, b_\beta, C) = \inf_{t > 0} \{ s(C - b_\alpha) + t(C - b_\beta) + \log E_Q^k(p_{ak}(x)) \} \) and \( C = \max_{k = \alpha, \beta} (C_{ak} + b_k). \)

Proof. We first show that the assumption implies that \( \lim_{n \to \infty} c_n = C - \bar{b}. \) From the definition of \( \phi_n^\alpha \), one obtains

\[
\lim_{n \to \infty} n^{-1} \log E_P^\alpha \phi_n^\alpha \\
\leq \lim_{n \to \infty} \inf_{n > 0} n^{-1} \log P_a \left( n^{-1} \log 2 \cdot \max_{k = \alpha, \beta} w_k \prod_{j=1}^{n} \frac{q_k(x_j)}{p_a(x_j)} \geq c_n \right) \\
\leq \max_{k = \alpha, \beta} \lim_{n \to \infty} \inf_{n > 0} n^{-1} \log P_a \left( n^{-1} \sum_{j=1}^{n} \log \frac{q_k(x_j)}{p_a(x_j)} \geq c_n - b_k + n^{-1} \log \frac{w}{2} \right) \\
\leq \max_{k = \alpha, \beta} \inf_{t > 0} \left\{ -t(C_k - b_k + \bar{b}) + \log \psi(t|P_a, Q_k) \right\}
\]

where \( w = e^{n b_\alpha} + e^{n b_\beta}, C = \sup_{n \to \infty} c_n. \) Therefore, \( \inf_{t > 0} \left\{ -t(C - b_k + \bar{b}) + \log \psi(t|P_a, Q_k) \right\} \geq \inf_{t > 0} \left\{ -tC_{ak} + \log \psi(t|P_a, Q_k) \right\} \) for \( k = \alpha \) or \( \beta. \) This implies that \( C - b_k + \bar{b} \leq C_{ak}, \) for \( k = \alpha \) or \( \beta, \) i.e.

\[
C \leq \max_{k = \alpha, \beta} (C_{ak} + b_k) - \bar{b}. \tag{2.4}
\]

Also, \( E_P^\alpha \phi_n^\alpha \geq P_a(n^{-1} \log w_k \prod_{j=1}^{n} \frac{q_k(x_j)}{p_a(x_j)} \geq c_n) \) for \( k = \alpha \) and \( \beta. \) Thus

\[
E_P^\alpha \phi_n^\alpha \geq \max_{k = \alpha, \beta} P_a \left( n^{-1} \sum_{j=1}^{n} \log \frac{q_k(x_j)}{p_a(x_j)} \geq c_n - b_k + n^{-1} \log \frac{w}{2} \right).
\]

Using a similar argument as above, one obtains, for \( k = \alpha \) and \( \beta, \)

\[
\inf_{t > 0} \left\{ -tC_{ak} + \log \psi(t|P_a, Q_k) \right\} \geq \inf_{t > 0} \left\{ -t(C - b_k + \bar{b}) + \log \psi(t|P_a, Q_k) \right\},
\]

where \( C = \liminf_{n \to \infty} c_n. \) Hence, \( C_{ak} \leq C - b_k + \bar{b} \) for each \( k, \) i.e.

\[
C \geq \max_{k = \alpha, \beta} (C_{ak} + b_k) - \bar{b}. \tag{2.5}
\]

Equations (2.4) and (2.5) imply that \( \lim_{n \to \infty} c_n = C - \bar{b}. \) Now,

\[
E_Q^k(1 - \phi_n^\alpha) \leq Q_k \left( n^{-1} \log \max_{k = \alpha, \beta} w_k \prod_{j=1}^{n} \frac{q_k(x_j)}{p_a(x_j)} < c_n \right) \\
= \left( n^{-1} \sum_{j=1}^{n} \log \frac{q_k(x_j)}{p_a(x_j)} < c_n - b_k + n^{-1} \log \frac{w}{2} \right) \\
\leq Q_k \left( n^{-1} \sum_{j=1}^{n} \log \frac{q_k(x_j)}{p_a(x_j)} < c_n - b_k + n^{-1} \log \frac{w}{2} \right) \tag{2.6}
\]
Similarly,

$$E^k_Q(1 - \phi_n^\alpha) \geq Q_k \left( n^{-1} \sum_{j=1}^{n} \log \frac{q_k(x_j)}{p_{a}(x_j)} < c_n - b_k + n^{-1} \log \frac{w}{2}, k = \alpha, \beta \right). \tag{2.7}$$

Applying Theorem 5.1 of Groeneboom et al. (1979), (2.6) and (2.7) imply that

$$\lim_{n \to \infty} n^{-1} \log E^k_Q(1 - \phi_n^\alpha) = \rho_{ak}(b_\alpha, b_\beta, C)$$

for $k = \alpha, \beta$. Since $E_{\alpha}^n(1 - \phi_n^\alpha) = w_\alpha E_{\alpha}^n(1 - \phi_n^\alpha) + w_\beta E_{\beta}^n(1 - \phi_n^\alpha)$, we have

$$\lim_{n \to \infty} n^{-1} \log E_{\alpha}^n(1 - \phi_n^\alpha) = \max_{k=\alpha, \beta} \lim_{n \to \infty} n^{-1} \log E_k^Q(1 - \phi_n^\alpha)$$

and the result follows.

Since $\phi_n^\alpha$ is the most powerful test of $\Pi_{j=1}^{\nu} p_\alpha(x_j)$ versus $w_\alpha \Pi_{j=1}^{\nu} q_\alpha(x_j) + w_\beta \Pi_{j=1}^{\nu} q_\beta(x_j)$, this yields the following

**Corollary 2.1** For any test $\phi_n^\alpha$ such that $\limsup_{n \to \infty} n^{-1} \log E_k^Q \phi_n^\alpha \leq -A_\alpha$, for a fixed $\alpha$

$$\max_{k=\alpha, \beta} (b_k + \liminf_{n \to \infty} n^{-1} \log E_k^Q(1 - \phi_n^\alpha)) \geq \max_{k=\alpha, \beta} (b_k + \rho_{ak}(b_\alpha, b_\beta, C))$$

for any $\beta = 1, \ldots, \nu$ and any $b_1, \ldots, b_{\nu}$.

**Corollary 2.2** $\rho_{ak}(b_\alpha, b_\beta, C) \geq -B(A_\alpha, P_\alpha, Q_k), k = \alpha, \beta$.

Proof. As shown in the proof of lemma 2.1, $\lim_{n \to \infty} n^{-1} \log E_k^Q \phi_n^\alpha = -A_\alpha$ implies that $\lim_{n \to \infty} n^{-1} \log E_k^Q(1 - \phi_n^\alpha) = \rho_{ak}(b_\alpha, b_\beta, C)$ for $k = \alpha, \beta$. Since, as a test of $P_\alpha$ versus $Q_k, \phi_n^\alpha$ has exponential rate $A_\alpha$, the result follows from (1.2).

**Theorem 2.2** If an adaptive test exists, then for all $\alpha, \beta$

$$B_{\alpha \alpha} \leq B_{\beta \alpha}.$$

Proof. Let $\phi_n^\alpha$ be an adaptive test. Since $\phi_n^\alpha$ satisfies the condition in corollary 2.1, by letting $b_\alpha = B_{\alpha \alpha}$ and $b_\beta = B_{\beta \beta}$, one obtains

$$0 = \max_{k=\alpha, \beta} (B_{kk} - B_{kk}) \geq \max_{k=\alpha, \beta} (B_{kk} + \rho_{ak}(B_{\alpha \alpha}, B_{\beta \beta}, C)),$$

i.e. $\rho_{ak}(B_{\alpha \alpha}, B_{\beta \beta}, C) \leq -B_{kk}$ for $k = \alpha, \beta$. But, from corollary 2.2, with $k = \beta, \rho_{\alpha \beta}(B_{\alpha \alpha}, B_{\beta \beta}, C) \leq -B_{\alpha \beta}$, so that $B_{\beta \beta} \leq B_{\alpha \beta}$. Since $\alpha, \beta$ are arbitrary, the result follows.

Hence from Theorems 2.1 and 2.2 we have established
Corollary 2.3 If $P_{\alpha}, Q_{\alpha}$ are absolutely continuous with respect to the Lebesgue measure for all $\alpha$, then an adaptive test exists iff for all $\alpha, \beta$,

$$B_{aa} \leq B_{\beta \alpha}. \quad (2.8)$$

From the proof of theorem 2, we see that if an adaptive test exists, then 

$$\rho_{\alpha \beta}(B_{aa}, B_{\beta \beta}, C) \leq -B_{\beta \beta} \forall \alpha, \beta.$$ 

Suppose that $P_{\alpha} = Q_{\beta}$ for some $\alpha \neq \beta$. Assume that $A_{\beta}$ satisfies the condition in theorem 2.1, in particular $A_{\beta} < K(Q_{\beta}, P_{\beta})$, i.e. $B_{\beta \beta} > 0$. Consider

$$\rho_{\alpha \beta}(B_{aa}, B_{\beta \beta}, C) = \inf_{s, t > 0} \left\{ s(C - B_{aa}) + t(C - B_{\beta \beta}) + \log E_{Q}^{\beta} \left( \frac{p_{a}(x)}{q_{a}(x)} \right)^{s} \left( \frac{p_{a}(x)}{q_{\beta}(x)} \right)^{s} \right\}$$

$$\geq \inf_{s > 0} \left\{ s(C - B_{aa}) + \log E_{Q}^{\beta} \left( \frac{p_{a}(x)}{q_{a}(x)} \right)^{s} \right\} + \inf_{t > 0} \{ t(C - B_{\beta \beta}) \}. \quad (2.9)$$

The second term of (2.9) is zero because $C - B_{\beta \beta} = \max_{k=\alpha, \beta}(C_{ak} + B_{kk}) - B_{\beta \beta} \geq C_{\alpha \beta} > 0$ when $P_{\alpha} = Q_{\beta}$. Let $f(s)$ be the function in the paranthesis of the first term in (2.9). Then $f'(0) = C - B_{aa} + E_{Q}^{\beta} \log p_{a}(x)/q_{a}(x) = C - B_{aa} + K(P_{\alpha}, Q_{a}) \geq C_{aa} + K(P_{\alpha}, Q_{a}) \geq 0$. Since $f(s)$ is a convex function, this implies that $\inf_{s > 0} f(s) = 0$. That is $\rho_{\alpha \beta}(B_{aa}, B_{\beta \beta}, C) = 0$. But, by assumption $B_{\beta \beta} > 0$, and inequality $\rho_{\alpha \beta}(B_{aa}, B_{\beta \beta}, C) \leq -B_{\beta \beta}$ is impossible, and hence adaptive test cannot exist. This yields

**Corollary 2.4** If $P_{\alpha} = Q_{\beta}$ for some $\alpha \neq \beta$, then an adaptive test does not exist.

**Remarks.**

1. If $P_{\alpha} = P$ for some $P$ for all $\alpha$, then (2.8) always holds, i.e. an adaptive test always exists.

2. By interchanging the roles of $P_{\alpha}, Q_{\alpha}$, we can define a similar notion of adaptation, i.e. a test is adaptive if the type II error converges to zero at a guaranteed rate while the type I error converges to zero at the optimal rate. Thus a test with the following properties:

$$\limsup_{n \to \infty} n^{-1} \log E_{Q}^{\alpha}(1 - \phi_{n}) \leq -A_{\alpha} \text{ and}$$

$$\liminf_{n \to \infty} n^{-1} \log E_{P}^{\beta} \phi_{n} = -B(A_{\alpha}, Q_{\alpha}, P_{\alpha})$$

exists iff

$$B(A_{\alpha}, Q_{\alpha}, P_{\alpha}) \leq B(A_{\beta}, Q_{\beta}, P_{\alpha}), \text{ for } \alpha, \beta = 1, \ldots, l. \quad (2.10)$$
It is easy to see that \( \lim_{A_\beta \to 0^+} B(A_\beta, Q_\beta, P_\alpha) = K(Q_\beta, P_\alpha) \). Thus, by setting \( A_1 = \cdots = A_i = 0 \), i.e. by letting the type II error converges to a positive constant, condition (2.10) becomes \( K(Q_\alpha, P_\alpha) \leq K(Q_\beta, P_\alpha) \). This result is obtained in Rukhin (1986).

3. Corollary 2.3 can be extended to the case when the nuisance parameter has countably many values. Assume that \( \inf_\beta K(Q_\beta, P_\alpha) > 0 \) and \( 0 < A_\alpha < \inf_\beta K(Q_\beta, P_\alpha) \) for every \( \alpha \). Let \( \{k_n\} \) be a non-decreasing sequence such that \( n^{-1} \log k_n \to 0 \) as \( n \to \infty \). Consider the following test:

\[
\phi_n^* = \begin{cases} 
1 & \text{if } n^{-1} \log \frac{\max_{\beta \leq k_n} e^{\nu_\beta} \prod_{j=1}^n q_\beta(x_j)}{\max_{\beta \leq k_n} e^{\nu_\beta} \prod_{j=1}^n p_\beta(x_j)} \geq 0 \\
0 & \text{otherwise.}
\end{cases}
\]

Using corollaries 2.1 and 2.2, it is clear that if an adaptive test exists, then (2.8) holds for all \( \alpha, \beta \). We claim that \( \phi_n^* \) is an adaptive test if (2.8) holds. We first show that \( \phi_n^* \) is a test of rate \( A_\alpha \) for every \( \alpha \). Pick \( n \) large enough so that \( \alpha \leq k_n \), then

\[
E_P^2 \phi_n^* \leq k_n \cdot \max_{\beta \leq k_n} \inf_{t>0} e^{-t(A_\alpha-B_\beta)} \psi(t|P_\alpha, Q_\beta)
\]

From the proof of theorem 2.1. \( \inf_{t>0} \{ e^{-t(A_\alpha-B_\beta)} \psi(t|P_\alpha, Q_\beta) \} \leq e^{-A_\alpha} \), for any \( \beta \). Thus,

\[
\limsup_{n \to \infty} n^{-1} \log E_P^2 \phi_n^* \leq \lim_{n \to \infty} n^{-1} \log k_n - A_\alpha
\]

Now, we show that type II error of the test \( \phi_n^* \) converges to zero at the optimal rate. Using the same argument as above

\[
\liminf_{n \to \infty} n^{-1} \log E_P^2 (1 - \phi_n^*) \\
\leq \sup_{\beta} \inf_{t>0} \{ -t(B_{aa} - A_\beta) + \log \psi(t|Q_\alpha, P_\beta) \} + \lim_{n \to \infty} n^{-1} \log k_n \\
\leq -B_{aa}.
\]

Hence, \( \phi_n^* \) is an adaptive test.

4. When an adaptive test exists, then the test defined in (2.2) is always adaptive. However, an overall maximum likelihood ratio test, i.e. a test with critical region \( \{ x : \max_{\alpha} \prod_{j=1}^n q_\alpha(x_j)/\max_{\alpha} \prod_{j=1}^n p_\alpha(x_j) \geq e^{c_n} \} \), for some constant \( c_n \), is not necessarily adaptive even when an adaptive test exists. Let \( \hat{\phi}_n \) denote such a
test. If $c_n = \max_{k=\alpha, \beta} C_{\alpha, \beta}$, then it can be shown that $\hat{c}_n$ is a test of rate $A_{\alpha}$ for each $\alpha$. If $A_{\alpha}, \alpha = 1, \ldots, l$ are picked such that $C_{11} = \ldots = C_{ll}$, then a sufficient condition for the test $\hat{c}_n$ to be adaptive is that $B_{\alpha\beta} \geq B_{\beta\alpha} + t_{\alpha}(C - C_{\alpha\beta})$ where $C = \max_{\alpha, \beta} C_{\alpha, \beta}$, $t_{\alpha} = \max_{\beta} t_{\beta\alpha}$, and $t_{\beta\alpha}$ is the point where $-t_{\alpha\beta} + \log \psi(t|P_{\alpha}, Q_{\beta})$ attains its minimum.

3 Example

Let $P_{\alpha}, Q_{\alpha}$ belong to an exponential family with densities $p_{\alpha}, q_{\alpha}$ given by

$$p_{\alpha}(x) = \exp\{\xi_{\alpha}x - \chi(\xi_{\alpha})\},$$

$$q_{\alpha}(x) = \exp\{\eta_{\alpha}x - \chi(\eta_{\alpha})\}.$$ 

Let $g(t, u) = -tu + \log \psi(t|P_{\alpha}, Q_{\beta})$. By a straightforward calculation,

$$\psi(t|P_{\alpha}, Q_{\beta}) = \int \exp t\{\eta_{\beta}x - \xi_{\alpha}x - \chi(\eta_{\beta}) + \chi(\xi_{\alpha})\} \exp\{\xi_{\alpha}x - \chi(\xi_{\alpha})\}dx = \exp\{-t\chi(\xi_{\alpha}) - t\chi(\eta_{\beta}) + t(\eta_{\beta} + (1-t)\xi_{\alpha})\}.$$ 

Thus, $g(t, u) = -tu - (1-t)\chi(\xi_{\alpha}) - t\chi(\eta_{\beta}) + \chi(\eta_{\beta} + (1-t)\xi_{\alpha})$. Therefore we have

$$g'(t) = -u + \chi(\xi_{\alpha}) - \chi(\eta_{\beta}) + \chi'(\eta_{\beta} + (1-t)\xi_{\alpha}),$$

where $g', \chi'$ are derivatives of $g, \chi$ with respect to $t$.

If $s_{\alpha\beta}$ satisfies

$$s_{\alpha\beta}\chi'(s_{\alpha\beta}\eta_{\beta} + (1-s_{\alpha\beta})\xi_{\alpha}) - \chi(s_{\alpha\beta}\eta_{\beta} + (1-s_{\alpha\beta})\xi_{\alpha}) + \chi(\xi_{\alpha}) = A_{\alpha}, \quad (3.1)$$

and let

$$C_{\alpha\beta}(A_{\alpha}) = \chi(\eta_{\beta}) - \chi(\xi_{\alpha}) - \chi'(s_{\alpha\beta}\eta_{\beta} + (1-s_{\alpha\beta})\xi_{\alpha}), \quad (3.2)$$

then $g(s_{\alpha\beta}, C_{\alpha\beta}) = -A_{\alpha}$ and $g'(s_{\alpha\beta}, C_{\alpha\beta}) = 0$.

Suppose $0 < A_{\alpha} < \min_{\beta} K(Q_{\beta}, P_{\alpha})$, for each $\alpha$, then an adaptive test exists iff

$$A_{\alpha} - \chi(\eta_{\beta}) + \chi(\xi_{\alpha}) + \chi'(s_{\alpha\beta}\eta_{\beta} + (1-s_{\alpha\beta})\xi_{\alpha}) \geq A_{\beta} - \chi(\eta_{\beta}) + \chi(\xi_{\beta}) + \chi'(s_{\beta\alpha}\eta_{\beta} + (1-s_{\beta})\xi_{\beta})$$

or,

$$\chi(\xi_{\alpha}) - \chi(\xi_{\beta}) + \chi'(s_{\alpha\beta}\eta_{\beta} + (1-s_{\alpha\beta})\xi_{\alpha}) - \chi'(s_{\beta\alpha}\eta_{\beta} + (1-s_{\beta})\xi_{\beta}) \geq A_{\beta} - A_{\alpha}. \quad (3.3)$$
Suppose \( P_\alpha \sim N(\theta_\alpha, 1) \), \( Q_\alpha \sim N(\mu_\alpha, 1) \), then \( K(Q_\beta, P_\alpha) = 2^{-1}(\mu_\beta - \theta_\alpha)^2 \) and

\[
\chi(t\eta_\beta + (1-t)\xi_\alpha) = \frac{1}{2}(t\mu_\beta + (1-t)\theta_\alpha)^2,
\]

\[
\chi'(t\eta_\beta + (1-t)\xi_\alpha) = (t\mu_\beta + (1-t)\theta_\alpha)(\mu_\beta - \theta_\alpha).
\]

It is easy to check that \( s_{\alpha\beta} = \sqrt{\frac{2A_\alpha}{(\mu_\beta - \theta_\alpha)^2}} \) satisfies (3.1), and it follows from (3.2) that

\[
C_{\alpha\beta}(A_\alpha) = \frac{(\mu_\beta - \theta_\alpha)^2}{2} \left( \frac{2\sqrt{2A_\alpha}}{|\mu_\beta - \theta_\alpha|} - 1 \right).
\]

Thus, if \( 0 < A_\alpha < \min_\beta 2^{-1}(\mu_\beta - \theta_\alpha)^2 \), then an adaptive test exists iff for all \( \alpha \neq \beta \)

\[
(\mu_\alpha - \theta_\alpha)^2 \left( \frac{\sqrt{2A_\alpha}}{|\mu_\alpha - \theta_\alpha|} - \frac{1}{2} \right) - (\mu_\alpha - \theta_\beta)^2 \left( \frac{\sqrt{2A_\beta}}{|\mu_\alpha - \theta_\beta|} - \frac{1}{2} \right) \geq A_\beta - A_\alpha.
\]

For example, when \( \theta_1 = -1, \mu_1 = 1; \theta_2 = -1.5, \mu_2 = 0.5, A_1 = 0.5, A_2 = 1.125 \), then \( C_{11} = 0, C_{12} = 0.375, C_{21} = 0.625, C_{22} = 1 \) and the above condition is satisfied. (Here \( B_{11} = B_{21} = 0.5, B_{22} = B_{12} = 0.125 \)).

Consider the following overall maximum likelihood ratio test with the above parametric values:

\[
\hat{\phi}_n(x) = \begin{cases} 
1 & \text{if } n^{-1} \log \prod_{a=1,2} \prod_{j=1}^n q_a(x_j) \geq C \\
0 & \text{otherwise}.
\end{cases}
\]

If the critical constant \( C \) is chosen such that \( C = \max_{\alpha\beta} C_{\alpha\beta}(A_\alpha) = 1 \), then \( \hat{\phi}_n \) is a test of exponential rate \( A_\alpha \) for \( \alpha = 1, 2 \). Now, consider the rate of convergence of the probability of type two error:

\[
\lim_{n \to \infty} n^{-1} \log E_Q^\alpha(1 - \hat{\phi}_n) = \max_{k=1,2} \inf_{s,t>0} \left\{ (s + t) + \log E_Q^\alpha \left( \frac{p_k(x)}{q_1(x)} \right)^s \left( \frac{p_k(x)}{q_2(x)} \right)^t \right\}.
\]

When \( \alpha = 2 \),

\[
E_Q^2 \left( \frac{p_k(x)}{q_1(x)} \right)^s \left( \frac{p_k(x)}{q_2(x)} \right)^t = \exp \left\{ -\frac{1}{2} \{ s(1-s)(\theta_k - \mu_1)^2 + t(1-t)(\theta_k - \mu_2)^2 + 2t(\theta_k - \mu_1)(\mu_1 - \mu_2) - 2st(\theta_k - \mu_2)(\theta_k - \mu_1) \} \right\}.
\]

Consider \( k = 1 \) and let the expression in the parenthesis of (3.5) be \( g(s,t) \), i.e.

\[
g(s,t) = s + t - \frac{9}{8} t(1-t) - 2s(1-s) + s + 3st.
\]
By differentiating \( g \) with respect to \( s \) and \( t \), one can show that \( (\partial/\partial s)g \) and \( (\partial/\partial t)g \) do not vanish simultaneously and therefore

\[
\inf_{s,t>0} g(s,t) = \min\{\inf_{s>0} g(s,0), \inf_{t>0} g(0,t)\}.
\]

By a straightforward calculation, \( \inf_{s,t>0} g(s,t) \approx -.0035 \). Hence,

\[
\lim_{n \to \infty} n^{-1} \log E_{\tilde{\phi}_n}^2 (1 - \phi_n) > -.125 = -B_2,
\]

and it follows that \( \tilde{\phi}_n \) is not an adaptive test.

Note that in this example, the critical constant \( C = 1 \) is the 'best' in the sense that \( \tilde{\phi}_n \) is a test of exponential rate \( A_n \) for reach \( \alpha \). If \( C \) is replaced by \( c_n \) in the definition of \( \phi_n \) such that \( \lim \inf_{n \to \infty} c_n = C' < 1 \), then by a similar calculation as above, one can show that \( \lim \sup_{n \to \infty} n^{-1} \log E_{\hat{\phi}_n}^2 > -A_2 \).

References


This paper considers the problem of hypothesis testing when the distribution is specified only up to a nuisance parameter. A test is said to be adaptive if it is asymptotically optimal regardless of the value of the nuisance parameter. The exponential rate of convergence to zero of the type II error when the type I error converges to zero exponentially fast at a fixed rate is used as the optimal criterion. A necessary and sufficient condition for the existence of adaptive test is obtained.