The Annals of Statistics
1981, Vol. 9, No. 5, 1096–1106

ROBUST ASYMPTOTIC TESTS OF STATISTICAL HYPOTHESES INVOLVING NUISANCE PARAMETERS[1]

BY PAUL C. C. WANG

University of Calgary

A robust version of Neyman's optimal $C(\alpha)$ test is proposed for contamination neighborhoods. The proposed robust test is shown to be asymptotically locally maximin among all asymptotic level α tests. Asymptotic efficiency of the test procedure at the ideal model is investigated. An outlier resistant version of Student's t-test is proposed.

1. Introduction. Huber (1965) obtained a robust version of the probability ratio test by explicitly exhibiting a pair of least favorable distributions. Subsequently Huber-Carol (1970) and Rieder (1978) studied the corresponding asymptotic robust testing problem when the ideal model is completely specified by the parameter ξ tested. The purpose of this paper is to extend the asymptotic robust testing theory to the situation where the ideal model involves nuisance parameters. The problem considered is a robust analogue of the composite parametric testing problem considered by Neyman (1958), and the test procedure derived herein is a robust extension of the optimal $C(\alpha)$ test. For another robust extension, see Beran (1980).

Specifically let $X_1, X_2, \cdots, X_n$ be independent random variables each assuming values in $\mathscr{X} \subset \mathbb{R}^m$ for some m. Let $\mathscr{P} = \{P_\theta : \theta \in \Theta\}$ be a family of probability measures on the sample space, and suppose the parameter space Θ is an open product set $\Xi \times \mathscr{N}$ where $\Xi \subset \mathbb{R}$ and $\mathscr{N} \subset \mathbb{R}^k$. A generic element of Θ is denoted by $\theta = (\xi, \eta)$ where $\xi \in \Xi$ and $\eta \in \mathscr{N}$. For each n, $\epsilon_n \in (0, 1)$ and $\theta \in \Theta$ let $\mathscr{V}(\epsilon_n; \theta)$ denote the contamination neighborhood of P_θ with size ϵ_n, i.e., the set of probability measures of the form

$$Q = (1 - \epsilon_n)P_\theta + \epsilon_n R$$

for some probability measure R. Let $\mathscr{V}^{(n)}(\epsilon_n; \theta)$ denote the product of n-copies of $\mathscr{V}(\epsilon_n; \theta)$.

Based on the sample $X_1, X_2 \cdots, X_n$ we are interested in testing the hypothesis that $\xi = \xi_0$, while allowing ϵ_n contamination. To state the problem formally for a one-sided test, we wish to test

$$H_n : \mathscr{L}(X_1, X_2, \cdots, X_n) \in \mathscr{V}^{(n)}(\epsilon_n; \xi_0, \eta)$$

against the alternative

$$K_n : \mathscr{L}(X_1, X_2, \cdots, X_n) \in \mathscr{V}^{(n)}(\epsilon_n; \xi_n, \eta) \quad \text{for} \quad \xi_n \geq \xi_{n1}(\eta) > \xi_0,$$

where $\xi_{n1}(\eta)$ is a specified function of the unspecified vector of nuisance parameters η, and $\mathscr{L}(X_1, X_2, \cdots, X_n)$ denotes the law of $X_1, X_2, \cdots, X_n$. The indifference region $\{(\xi_0, \xi_{n1}(\eta)) : \eta \in \mathscr{N}\}$ is required because otherwise H_n and K_n are indistinguishable (see Remark 4 in Rieder, 1978). A precise criterion for choosing $\xi_{n1}(\eta)$ in a local framework will be given later (cf. (2.1) and (3.1)).

Received November 1979; revised December 10, 1980.

[1] Research supported in part by a Postgraduate Scholarship from NSERC of Canada and the office of Naval Research Contract N00014-75-C-0159. Revision supported under contract DE-AC02-76CH0016 to Brookhaven National Laboratory.

AMS 1970 *subject classification*: Primary 62G10, 62G35; secondary 62E20.

Key words and Phrases: least favorable distributions, optimal $C(\alpha)$ test, asymptotic local maximin test, ϵ-contamination.

EXAMPLE. A particular example of the above testing problem concerns the testing of the mean of a normal population in the presence of possible outliers. The problem is to test

$$H_{n0}: \mathscr{L}(X_1) = (1 - \epsilon_n)\Phi\{(x - \mu_0)/\sigma\} + \epsilon_n R_{n0}$$

with alternative

$$H_{n1}: \mathscr{L}(X_1) = (1 - \epsilon_n)\Phi\{(x - \mu_n)/\sigma\} + \epsilon_n R_{n1}, \quad \text{for} \quad \mu_n \geq \mu_{n1}(\sigma) > \mu_0,$$

with unknown σ. Here Φ is the cdf of the standard normal variable and R_{nj}'s are arbitrary distribution functions which in principle may depend upon the parameters μ and σ^2. We shall show that the robust procedure appropriate for this testing problem is to reject H_{n0} for large values of

$$n^{-1/2} \sum_{i=1}^{n} \min(V, \max((X_i - \mu_0)/\hat{\sigma}_n, -V)),$$

where $V > 0$ satisfies a certain equation and $\hat{\sigma}_n$ is any $n^{1/2}$-consistent and resistant estimate of σ such as the normalized interquantile range or the symmetrized interquartile range. For a discussion of the lack of robustness of Student's t-test against outliers, see Millar (1980).

OUTLINE. As in Huber-Carol (1970) and Rieder (1978) we shall employ a local approach and allow the neighborhoods in H_n and K_n to shrink at the same rate as ξ_n approaches ξ_0.

By fixing the nuisance parameter η, a necessary and sufficient condition is derived for a truncated likelihood ratio test to be an asymptotic level-α test free of the local nuisance parameters (Theorem 2.1). The condition requires the nuisance score functions to be uncorrelated with a certain truncated test function.

By making use of the orthogonality condition and Huber's first basic result on testing (see Theorem 1 of Huber, 1965), an asymptotic local maximin test at η is constructed (Theorem 2.2). This test, which is based on the likelihood ratio of a pair of asymptotically least favorable distributions, is not a genuine test since it depends upon η. By adapting the original results of LeCam (1969), the question of existence of $n^{1/2}$-consistent and resistant estimates of η is discussed in Section 2. Upon estimating the known η, a studentized version of the fixed-η sequence of tests is shown to satisfy an asymptotic optimality criterion defined in Section 2 (Theorem 2.3).

Sections 3 and 4 supply assumptions, technical details and proofs of results.

In Section 5 we investigate the performance of the proposed robust test with respect to the optimal $C(\alpha)$ test at the ideal model $\mathscr{P}$. We return to the example of the normal mean in Section 6.

NOTATIONS. Throughout this paper the following notations are adopted. E denotes the expectation taken with respect to $P_{\xi_0,\eta}$ unless specified otherwise; $a^+ = \max(a, 0)$; $a^- = \max(-a, 0)$; $I(A)$ denotes the indicator function of the set A; $|\cdot|$ denotes the usual Euclidean norm; $\alpha \in (0, 1)$ denotes the level of significance and $z_{1-\alpha}$ denotes the $(1 - \alpha)$ quantile of the standard normal. For notational convenience quantities such as ξ_0 will be suppressed in formulae whenever this does not cause confusion. In particular the random element X in a random function $f(X; \theta)$ may also be suppressed and we shall write it as $f(\theta)$.

2. Framework, optimality and results. Following the usual asymptotic local approach, upon fixing $\theta \in \Theta$, we shall consider parameter points of the form

$$\theta_n = \theta + n^{-1/2} h_n \quad \text{where} \quad h_n, h \in \mathbb{R}^{k+1} \quad \text{and} \quad h_n \to h.$$

Then for each $n \geq 1$ and $\eta \in \mathscr{N}$ define

(2.1) $$\epsilon_n = n^{-1/2}\epsilon \quad \text{and} \quad \xi_{n1}(\eta) = \xi_0 + n^{-1/2}\tau_1(\eta),$$

where $\tau_1(\eta)$ is some positive function to be specified later. For now let us assume $\tau_1(\eta)$ has been chosen such that $\{H_n\}$ and $\{K_n\}$ are asymptotically disjoint.

For each sequence $\{\psi_n\}$ of tests and each $\theta \in \Theta$ denote

$$\alpha_n(\psi_n; \theta) = \sup\{E(\psi_n \mid Q_n) : Q_n \in \mathscr{V}^{(n)}(\epsilon_n; \theta)\},$$

and

$$\beta_n(\psi_n; \theta) = \inf\{E(\psi_n \mid Q_n) : Q_n \in \mathscr{V}^{(n)}(\epsilon_n; \theta)\}.$$

Also, for $\eta \in \mathscr{N}$, $t \in \mathbb{R}^k$ and $M < \infty$ denote

$$\boldsymbol{\beta}_n(\psi_n; t \mid \eta, M) = \inf\{\beta_n(\psi_n; \xi_0 + n^{-1/2}\tau, \eta + n^{-1/2}t) : \tau_1(\eta) \le \tau \le M\}.$$

A sequence $\{\psi_n\}$ of tests is called an asymptotic level α test for $\{H_n\}$ iff

(2.2) $$\sup_T \limsup_{n\to\infty} \sup_{|t|\le T} \alpha_n(\psi_n; \xi_0, \eta + n^{-1/2}t) \le \alpha \quad \text{for all} \quad \eta \in \mathscr{N}.$$

Let Ψ_α denote the set of asymptotic level α tests for $\{H_n\}$. A sequence of tests $\{\psi_n^*\}$ is called an asymptotic local maximin (for short, ALM) test for $\{H_n\}$ against $\{K_n\}$ iff $\{\psi_n^*\} \in \Psi_\alpha$ and

(2.3) $$\sup_{M,T} \limsup_{n\to\infty} \sup_{|t|\le T} (\boldsymbol{\beta}_n(\psi_n; t \mid \eta, M) - \boldsymbol{\beta}_n(\psi_n^*; t \mid \eta, M)) \le 0$$

for all $\eta \in \mathscr{N}$ and $\{\psi_n\} \in \Psi_\alpha$.

FURTHER NOTATIONS. Let $\dot\phi(\theta) = (\dot\phi_1(\theta), \cdots \dot\phi_{k+1}(\theta))'$ denote the derivative in quadratic mean of the parametric family $\mathscr{P}$; cf. LeCam (1960), Roussas (1972). For the moment just think of $\dot\phi$ as the vector of pointwise partial derivatives of the log of the ideal density multiplied by a half—i.e., the Fisher score functions. Denote $\Delta_j(\eta) = 2\dot\phi_j(\xi_0, \eta)$ for $j = 1, 2, \cdots k + 1$ and set $\Delta(\eta) = (\Delta_2(\eta), \cdots \Delta_{k+1}(\eta))'$. For each $a \in \mathbb{R}^k$ and $\eta \in \mathscr{N}$ define the random function $\Lambda(a, \eta)$ by

(2.4) $$\Lambda(a, \eta) = \Delta_1(\eta) + \textstyle\sum_{j=1}^k a_j \Delta_{j+1}(\eta).$$

On $\mathbb{R} \times \mathbb{R}^k \times \mathscr{N}$ define numerical functions h_0 and h_1 by

(2.5) $$h_0(w, a, \eta) = E(\Lambda(a, \eta) - w)^+ \quad \text{and} \quad h_1(w, a, \eta) = h_0(w, a, \eta) + w.$$

Some properties of h_0 are given in Lemma 4.1 and others can be found in Huber-Carol (1970); Quang (1974) and Rieder (1978). From these functions, define another pair of numerical functions $V(a, \eta)$ and $U(a, \eta)$ implicitly by

(2.6) $$h_0(V(a, \eta), a, \eta) = \frac{\epsilon}{\tau_1(\eta)} = h_1(U(a, \eta), a, \eta).$$

Denote the Huberized function of $\Lambda(a, \eta)$ by $H(a, \eta)$, i.e.,

(2.7) $$H(a, \eta) = \min(V(a, \eta), \max(\Lambda(a, \eta), U(a, \eta))).$$

Under the hypothesized ideal model the mean of $H(a, \eta)$ is zero and the variance is strictly positive.

Now let $\eta \in \mathscr{N}$ be fixed and adopt the following nomenclature. A "test" $\psi_n(\eta)$ is called an asymptotic level α test at η iff it satisfies (2.2) at the point η. Let $\Psi_\alpha(\eta)$ denote such a class of tests at η. A sequence $\{\psi_n^*(\eta)\}$ of tests is called an asymptotic local maximin test at η iff $\{\psi_n^*(\eta)\} \in \Psi_\alpha(\eta)$ and satisfies (2.3) at the point η.

For $t_{nj}, t_j \in \mathbb{R}^k$ such that $t_{nj} \to t_j (j = 0, 1)$, denote $\theta_{n0} = (\xi_0, \eta_{n0})$ and $\theta_{n1} = (\xi_{n1}(\eta), \eta_{n1})$ where $\eta_{nj} = \eta + n^{-1/2}t_{nj}$ for $j = 0, 1$. Let $\psi_{n,H}(t_0, t_1 \mid \eta)$ denote Huber's level α likelihood ratio test between $\mathscr{V}^{(n)}(\epsilon_n; \theta_{n0})$ and $\mathscr{V}^{(n)}(\epsilon_n; \theta_{n1})$.

THEOREM 2.1. *The sequence* $\{\psi_{n,H}(t_0, t_1 \mid \eta)\}$ *of tests is an asymptotic level* α *test at* η *iff the vector* $a = (t_1 - t_0)/\tau_1(\eta)$ *satisfies the condition*

$$EH(a, \eta)\Delta_j(\eta) = 0 \quad \text{for} \quad j = 2, \cdots k + 1. \tag{2.8}$$

Let $a^*(\eta)$ denote a solution of (2.8) (cf. Lemma 3.4) and denote $\sigma_*^2(\eta) = EH^2(a^*(\eta), \eta)$. For $\eta \in \mathcal{N}$ let $C_\alpha(\eta) = z_{1-\alpha} + \sigma_*(\eta)^{-1}\epsilon V(a^*(\eta), \eta)$ and define

$$T_n(\eta) = n^{-1/2}\sigma_*(\eta)^{-1} \textstyle\sum_{i=1}^n H(X_i; a^*(\eta), \eta). \tag{2.9}$$

THEOREM 2.2. *The sequence* $\{\psi_n^*(\eta)\}$, *where* $\psi_n^*(\eta) = I(T_n(\eta) \geq C_\alpha(\eta))$, *is a* ALM *sequence of tests at* η. *The asymptotic minimum power of* $\psi_n^*(\eta)$ *at the point* η *is given by*

$$\beta(\tau \mid \eta) = \lim_{n\to\infty} \beta_n(\psi_n^*(\eta); \xi_n, \eta_n) = 1 - \Phi\{z_{1-\alpha} - [\tau_1(\eta)\sigma_*(\eta) + (\tau - \tau_1(\eta))S(\eta)]\}, \tag{2.10}$$

where

$$S(\eta) = \sigma_*(\eta) + \frac{\epsilon(V(a^*(\eta), \eta) - U(a^*(\eta), \eta))}{\tau_1(\eta)\sigma_*(\eta)}, \tag{2.11}$$

whenever $n^{1/2}(\xi_n - \xi_0) \to \tau$ *and* $\{n^{1/2}(\eta_n - \eta)\}$ *is bounded.*

So far the ALM procedure obtained is at the point η. Since the value of η is not specified by the hypothesis tested, we need to substitute an estimate $\hat{\eta}_n$ for η. From the local set-up it is desirable to have estimates $\hat{\eta}_n$ satisfying the following condition.

Condition C. For every bounded subset B of $\mathbb{R}^{k+1}$ and every $\theta = (\xi, \eta)$,

$$\lim_{b\to\infty} \limsup_{n\to\infty} \sup\{Q_n(n^{1/2} \mid \hat{\eta}_n - \eta \mid > b) : Q_n \in \mathcal{V}^{(n)}(\epsilon_n; \theta + n^{-1/2}h), h \in B\} = 0.$$

Note that, because the alternative test sequence $\{K_n\}$ is not contiguous to $\{H_n\}$ (cf. Quang, 1974; Rieder, 1978), we shall require the sequence $\{n^{1/2}(\hat{\eta}_n - \eta)\}$ to be uniformly bounded in probability over both the hypothesis and alternative neighborhoods.

THEOREM 2.3. *Let* $\{\hat{\eta}_n\}$ *be a sequence of estimates satisfying condition C. Then the sequence* $\{\psi_n^*(\hat{\eta}_n)\}$ *of tests is an* ALM *sequence for testing* $\{H_n\}$ *against* $\{K_n\}$. *Its asymptotic minimum power function is given by* (2.10).

REMARKS.

1. For the case when the data is i.i.d., the existence of $n^{1/2}$-consistent and resistant estimates can be immediately obtained from LeCam's (1969) construction, pages 105–107. By $n^{1/2}$-consistent and resistant estimates we meant estimates $\hat{\theta}_n$ having the property that for every $\delta > 0$ and every compact $K \subset \Theta$ there is a number b such that for all n

$$\sup\{Q_n(n^{1/2} \mid \hat{\theta}_n - \theta \mid > b) : Q_n \in \mathcal{V}^{(n)}(\epsilon_n; \theta), \theta \in K\} < \delta. \tag{2.12}$$

When the data is non-i.i.d., LeCam's construction will still work as long as $n^{1/2}$ times the Kolmogorov distance between the empiric and the cdf of the ideal model is bounded in probability uniformly over shrinking neighborhoods. In particular if we have for every $\delta > 0$ there exists $b(\delta)$ such that for every array of df's $F_1, F_2, \cdots F_n$ and every n

$$P_F(n^{1/2}\sup_x \mid \hat{F}_n(x) - \bar{F}_n(x)| \geq b(\delta)) \leq \delta \quad \text{where} \quad \bar{F}_n = n^{-1} \textstyle\sum_{j=1}^n F_j \tag{2.13}$$

and $\hat{F}_n$ is the empirical df, then the required result will follow. In a discusion with Professor LeCam, he showed that (2.13) can be obtained by using his Poissonization technique, cf. Theorem 3 of LeCam (1970). For the case when observations take their values in $\mathbb{R}$ or when the observations are vector valued with continuous df's, (2.13) can also be obtained from the results of Van Zuijlen (1978) and Neuhaus (1975) respectively. In any event, if the observations are independent, ϵ_n is of the order $O(n^{-1/2})$ and if some conditions on the ideal models are met (such as the first five assumptions listed in Section 3), there always exist estimates satisfying (2.12).

2. It can be shown that as $\epsilon \to 0$ the test statistic $T_n(\eta)$ at η tends to the optimal $C(\alpha)$ test statistic at η and $C_\alpha(\eta)$ tends to $z_{1-\alpha}$. Therefore the procedure $\psi_n^*(\hat{\eta}_n)$ can be regarded as a robust extension of the optimal $C(\alpha)$ test.

3. The test procedure at the point η can also be obtained by restricting attention to the class of tests based on sums of the form $n^{-1/2}\sum_{i=1}^n f(X_i, \eta)$ for some test function f. In this case a necessary and sufficient condition for a member of this class to be an asymptotic level α test at η is given by

$$Ef(\eta)\Delta_j(\eta) = 0 \text{ for } j = 2, \cdots k + 1.$$

This condition is also required in the corresponding parametric setup, cf. Theorem 1 of Neyman (1958). Note that the proposed procedure is a member of the class $C(\alpha)$ of tests. Hence the relationship between our robust procedure and the optimal $C(\alpha)$ test is further strengthened.

3. Asymptotic local maximin tests at η. The following conditions are assumed to hold throughout the paper unless specified otherwise.

(A.1) $P_{\theta_1} \neq P_{\theta_2}$ whenever $\theta_1 \neq \theta_2$.

(A.2) $\mathscr{P}$ is a collection of mutually absolutely continuous probabilities.

(A.3) For each $\theta \in \Theta$ the random function $\phi(\theta, \theta') = [dP_{\theta'}/dP_\theta]^{1/2}$ is differentiable in quadratic mean $[P_\theta]$ with respect to θ' at point θ. Let $\dot{\phi}(\theta)$ denote the derivative in quadratic mean.

(A.4) Let E_θ denote the expectation taken w.r.t. P_θ. The covariance function $\Gamma(\theta) = 4E_\theta\dot{\phi}(\theta)\dot{\phi}'(\theta)$ is finite, positive definite and continuous.

(A.5) If $\theta' \to \theta$, $\dot{\phi}(\theta') \to \dot{\phi}(\theta)$ in P_θ-probability.

(A.6) The positive function $\tau_1(\eta)$ satisfies the condition

$$\frac{\epsilon}{\tau_1(\eta)} < \min\{E\Lambda^+(a, \eta) : a \in \mathbb{R}^k\} \text{ for every } \eta \in \mathscr{N}. \tag{3.1}$$

(A.7) For each η, the matrix $\bar{\Gamma}_{22}(a, \eta) = E\Delta(\eta)\Delta'(\eta)I(U(a, \eta) < \Lambda(a, \eta) < V(a, \eta))$ is positive definite at a solution of the orthogonality condition (2.8).

(A.8) Let $H^*(\eta) = H(a^*(\eta), \eta)$. For every $\eta \in \mathscr{N}$ and every compact set K of $\mathbb{R}^k$ there exist a number M_0 such that

$$\sup_x |H^*(x, \eta + s) - H^*(x, \eta + t)| \le M_0 |s - t| \quad \text{for all} \quad s, t \in K.$$

REMARKS.

1. The condition (3.1) is the asymptotic disjointness condition of Rieder (1978) modified for the nuisance parameter case. This condition will guarantee the sequence $\{H_n\}$ and the alternative sequence $\{K_n\}$ to be asymptotically disjoint in the local framework. By Fatou's lemma $E\Lambda^+(a, \eta) \to \infty$, for each η, as $|a| \to \infty$. Hence the right side of (3.1) is positive. This quantity is also continuous in η. Therefore w.l.o.g. we shall henceforth assume $\tau_1(\eta)$ is a continuous function of η.

2. It will be seen later that assumption (A.7) guarantees the condition (2.8) has only one solution. This assumption is readily satisfied if ϵ is sufficiently small. Since as ϵ tends to 0 the matrix $\bar{\Gamma}_{22}$, evaluated at $a^*(\eta)$, tends to the corresponding $k \times k$ subpartition matrix of Γ.

3. Assumption (A.8) is used for the purpose of studentization. If we are to use discretized estimates $\tilde{\eta}_n$ obtained from $\hat{\eta}_n$ (see LeCam, 1960), (A.8) can be replaced by the following weaker assumption.

(A.8)′ For each $\eta \in \mathscr{N}$ and every sequence $\{s_n\}$ converging to $s \in \mathbb{R}^k$, the uniform difference $\sup_x |H^*(x, \eta + n^{-1/2}s_n) - H^*(x, \eta)|$ converges to zero.

This remark will become clear at the end of Section 4.

To prove the results stated at the point (ξ_0, η) let us first collect some useful facts. Proofs of these facts are omitted. For related results, interested readers may wish to consult Huber-Carol (1970), Quang (1974) and Rieder (1978).

Let $\Lambda_n(t_0, t_1 | \eta)$ denote the likelihood ratio statistic associated with $\psi_{n,H}(t_0, t_1 | \eta)$. It is given by the sum $\sum_{i=1}^n \log \pi_n(X_i; t_0, t_1 | \eta)$ where $\pi_n(t_0, t_1 | \eta) = \min\{C_{n0}^{-1}, \max\{\phi_{n1}^2/\phi_{n0}^2, C_{n1}\}\}$, C_{n0}^{-1} and C_{n1} are solutions of certain equations (cf. (3) of [4]) and $\phi_{nj} = [dP_{\theta_{nj}}/dP_{\xi_0,\eta}]^{1/2}$.

LEMMA 3.1. *Let* $a = (t_1 - t_0)/\tau_1(\eta)$ *then*

$$n^{1/2}\log \pi_n(t_0, t_1 | \eta)/\tau_1(\eta) \to H(a, \eta) \quad in \quad L_2[P_{\xi_0,\eta}] \tag{3.2}$$

and

$$E\Lambda_n(t_0, t_1 | \eta) = -\epsilon\tau_1(\eta)V(a, \eta) - 2\tau_1(\eta)t_0'EH(a, \eta)\Delta(\eta) - \tfrac{1}{2}\tau_1^2(\eta)EH^2(a, \eta) + o(1). \tag{3.3}$$

LEMMA 3.2. *Let probability measures* P_n *and* P_n' *be given. Let* $\epsilon_n \in [0, 1)$ *and let* $\mathscr{V}(\epsilon_n; P_n')$ *denote the usual contamination neighborhood of* P_n' *with size* ϵ_n. *Let* f_n *be bounded real valued measurable functions on* $\mathscr{X}$. *Define* $T_n(f_n) = n^{-1/2}\sum_{i=1}^n f_n(X_i)$ *and assume*

i) $\{P_n'^{(n)}\}$ *is contiguous to* $\{P_n^{(n)}\}$.
ii) $E(f_n(X)| P_n) = 0$, $E(f_n^2(X)| P_n) = 1$.
iii) $n^{1/2}\epsilon_n \to \epsilon$; $\sup_x f_n(x) \to d_1$; $\inf_x f_n(x) \to d_2$, *where* d_1 *and* d_2 *are finite.*
iv) $n^{1/2}E(f_n(x)| P_n') \to c$.

Then it holds, uniformly in $z \in \mathbb{R}$, *that*

$$\limsup_{n\to\infty} \{Q_n(T_n(f_n) > z): Q_n \in \mathscr{V}^{(n)}(\epsilon_n; P_n')\} = 1 - \Phi(z - c - \epsilon\, d_1)$$

and

$$\liminf_{n\to\infty} \{Q_n(T_n(f_n) > z): Q_n \epsilon \mathscr{V}^{(n)}(\epsilon_n; P_n')\} = 1 - \Phi(z - c - \epsilon\, d_2).$$

PROOF OF THEOREM 2.1. From the definition of the asymptotic level α test at η it is clear that we need to examine the maximum tail behavior of $\Lambda_n(t_0, t_1 | \eta)$ over another neighborhood $\mathscr{V}^{(n)}(\epsilon_n; \tilde\theta_{n0})$, where $\tilde\theta_{n0} = (\xi_0, \eta + n^{-1/2}\tilde t_{n0})$ and $\tilde t_{n0}$ converges to some $\tilde t_0 \in \mathbb{R}^k$.

Denote $g_n = n^{1/2} \log \pi_n(t_0, t_1 | \eta)/\tau_1(\eta)$, $\sigma_n^2 = Eg_n^2 - (Eg_n)^2$ and $\sigma^2 = EH^2(a, \eta)$. For each n define the random function $f_n = (g_n - Eg_n)/\sigma_n$ and form the sum $T_n(f_n) = n^{-1/2} \cdot\sum_{i=1}^n f_n(X_i)$. Hence Λ_n can be written as $\Lambda_n = \sigma_n\tau_1(\eta)T_n(f_n) + \sigma_n\tau_1(\eta)n^{1/2}Eg_n$. By Lemma 3.2

$$\lim_{n\to\infty} \sup\{Q_n(T_n(f_n) > y_n): Q_n \in \mathscr{V}^{(n)}(\epsilon_n; \tilde\theta_{n0})\} = 1 - \Phi(y - c - \epsilon V(a, \eta)/\sigma)$$

for every $y_n \to y$, where $c = \lim_{n\to\infty} n^{1/2}E_{\tilde\theta_{n0}} f_n$. By (3.2) and (A.3) it can be shown that $c = 2\tilde t_0'EH(a, \eta)\Delta(\eta)/\sigma$. Put these together with (3.3) and the fact that the critical value of the $\psi_{n,H}$-test necessarily tends to $z_{1-\alpha}\tau_1(\eta)(EH^2(a, \eta))^{1/2} - \tfrac{1}{2}\tau_1^2(\eta)EH^2(a, \eta)$ with $a = (t_1 - t_0)/\tau_1(\eta)$, the required results follows. □

As proof shows, Theorem 2.1 implies in particular that Huber's pair (Q_{n0}^*, Q_{n1}^*) of least favorable distributions between $\mathscr{V}(\epsilon_n; \xi_0, \eta - n^{-1/2}\tau_1(\eta)a^*(\eta))$ and $\mathscr{V}(\epsilon_n; \xi_{n1}(\eta), \eta)$ is asymptotically least favorable for testing

$$U\{\mathscr{V}^{(n)}(\epsilon_n; \xi_0, \eta + n^{-1/2}t): | t | \le T\} \text{ versus } \mathscr{V}^{(n)}(\epsilon_n; \xi_{n1}(\eta), \eta).$$

The next proposition shows that the pair is also asymptotically least favorable for testing against the following larger nested neighborhood:

$$U\{\mathscr{V}^{(n)}(\epsilon_n; \xi_0 + n^{-1/2}t, \eta + n^{-1/2}t): \tau_1(\eta) \le \tau \le M, | t | \le T\}.$$

PROPOSITION 3.3. *Let* $\mathscr{U}_{n0}$ *and* $\mathscr{U}_{n1}$ *denote the hypothesis and the alternative nested neighborhoods respectively. Let* $\{\psi_n^*\}$ *be any probability ratio test between* Q_{n0}^* *and* Q_{n1}^*. *Then*

$$\limsup_{n\to\infty} \sup\{E(\psi_n^* \mid Q_n) - E(\psi_n^* \mid Q_{n0}^*): Q_n \in \mathscr{U}_{n0}\} \le 0 \tag{3.4}$$

and

$$\liminf_{n\to\infty} \inf\{E(\psi_n^* \mid Q_n) - E(\psi_n^* \mid Q_{n1}^*) : Q_n \in \mathscr{U}_{n1}\} \geq 0. \tag{3.5}$$

Denote

$$\psi_{n,H}^* = \psi_{n,H}(-\tau_1(\eta)a^*(\eta), 0 \mid \eta). \tag{3.6}$$

The asymptotic minimum power of $\psi_{n,H}^*$ at the point η is given by the expression (2.10).

PROOF. Let $\tau_n \to \tau \geq \tau_1(\eta)$ and let $\{t_n\}$ be a bounded sequence in $\mathbb{R}^k$. Denote $\theta_n = (\xi_0 + n^{-1/2}\tau_n, \eta + n^{-1/2}t_n)$. By Lemma 3.2 and the orthogonality condition (2.8) we have

$$\lim_{n\to\infty} \inf\{Q_n(\Lambda_n(-\tau_1(\eta)a^*(\eta), 0 \mid \eta) > \lambda_n) : Q_n \in \mathscr{V}^{(n)}(\epsilon_n; \theta_n)\} = 1 - \Phi(y - c - \epsilon\, d_2) \tag{3.7}$$

for every $\lambda_n \to \lambda$, where

$$\sigma_*(\eta)\tau_1(\eta)(y - c - \epsilon\, d_2) = \lambda + \tfrac{1}{2}\tau_1^2(\eta)\sigma_*^2(\eta) - \tau\tau_1(\eta)\sigma_*^2(\eta) - \epsilon(\tau - \tau_1(\eta))[V(a^*(\eta), \eta) - U(a^*(\eta), \eta)]. \tag{3.8}$$

Now since the last term on the right side of (3.8) is positive, it follows that

$$\sigma_*(\eta)\tau_1(\eta)(y - c - \epsilon\, d_2) \leq \lambda - \tfrac{1}{2}\tau_1^2(\eta)\sigma_*^2(\eta).$$

By the same argument, with θ_n replaced by $(\xi_{n1}(\eta), \eta)$, we have from (3.7) that

$$\lim_{n\to\infty} E(\psi_n^* \mid Q_{n1}^*) = 1 - \Phi\{(\tau_1(\eta)\sigma_*(\eta))^{-1}(\lambda - \tfrac{1}{2}\tau_1^2(\eta)\sigma_*^2(\eta))\}.$$

This proves (3.5). The last assertion follows from (3.7) and (3.8). □

This proposition only proves that the sequence $\{\psi_{n,H}^*\}$ of (3.6) maximizes the minimum power over $\mathscr{U}_{n1}$ among all asymptotic level α tests at the point η. We will now show the $\psi_{n,H}^*$-test has the ALM property.

PROOF OF THEOREM 2.2. Let $t, t_n \in \mathbb{R}^k$ such that $t_n \to t$. Let $M < \infty$ be given and let $\{\psi_n\} \in \Psi_\alpha(\eta)$. By (3.5), it is enough to show

$$\lim_{n\to\infty} \sup(\beta_n(\psi_n; \xi_{n1}(\eta), \eta + n^{-1/2}t_n) - \beta_n(\psi_{n,H}^*; \xi_{n1}(\eta), \eta)) \leq 0. \tag{3.9}$$

For this, let $t_{n0} = t_n - \tau_1(\eta)a^*(\eta)$ and consider the testing problem: $\mathscr{V}^{(n)}(\epsilon_n; \xi_0, \eta + n^{-1/2}t_{n0})$ versus $\mathscr{V}^{(n)}(\epsilon_n; \xi_{n1}(\eta), \eta + n^{-1/2}t_n)$. By Theorem 2.1, an argument similar to the one used in the proof of Proposition 3.3 and Theorem 1 of Huber (1965) it is easily seen that (3.9) is true. This proves $\{\psi_{n,H}^*\}$ is an ALM sequence at the point η. To complete the proof we simply verify that the sequence $\{\psi_n^*(\eta)\}$ is an asymptotic level α test and has the same asymptotic minimum power as the $\psi_{n,H}^*$-test at η. □

To conclude this section we will show that the system (2.8) has a unique solution. For the remaining part of this section, dependencies on η shall be suppressed.

For $j = 2, \cdots, k + 1$ let $g_j(a) = EH(a)\Delta_j$. By (iii) of Lemma 4.1, g_j's are continuous.

LEMMA 3.4. (i) *Let $\{r_m\}$ be a sequence of numbers in $[1, \infty)$ such that $r_m \to \infty$. Also let $e, e_m \in E = \{e \in \mathbb{R}^k : |e| = 1\}$ such that $e_m \to e$. Let $\gamma_j = (E\Delta_2\Delta_j, \cdots E\Delta_{k+1}\Delta_j)'$ for $j = 2, \cdots k + 1$. Then*

$$\lim_{m\to\infty} r_m^{-1} g_j(r_m e_m) = \langle e, \gamma_j \rangle \quad \text{for} \quad j = 2, \cdots k + 1.$$

where $\langle \cdot, \cdot \rangle$ is the usual inner product on $\mathbb{R}^k$.

(ii) *The system* (2.8) *has a unique solution.*

PROOF.

(i) Denote $a_m = r_m e_m$ then

$$r_m^{-1} g_j(a_m) = E\Delta_j[r_m^{-1}H(a_m)]^+ - E\Delta_j[r_m^{-1}H(a_m)]^-. \tag{3.10}$$

By assumption (A.4) and since E is compact, $\Delta_j[r_m^{-1}H(a_m)]^+$ is uniformly integrable. By definitions of $V(a)$ and $U(a)$ we also have $[r_m^{-1}H(a_m)]^+ \to \langle e, \Delta\rangle^+$ in $P_{\xi_0,\eta}$-probability. Hence it follows that the first term on the right side of (3.10) tends to $E\Delta_j\langle e, \Delta\rangle^+$. Similarly the second term tends to $E\Delta_j\langle e, \Delta\rangle^-$. This proves (i).

(ii) For each constant c, define a continuous function $h_c: \mathbb{R}^k \to \mathbb{R}^k$ by $h_c(a) = a - cg(a)$, where $g(a) = (g_2(a) \cdots g_{k+1}(a))$. We will show the function h_c has fixed points for some $c \neq 0$. By Brouwer's Fixed Point Theorem it is enough to show that there exist $\delta > 0$ and $c \neq 0$ such that

$$\lim_{m\to\infty} \sup r_m^{-1} \sup\{| h_c(r_m e)| : e \in E\} \leq 1 - \delta$$

for any sequence $\{r_m\}$ of numbers such that $r_m \to \infty$.

Let e, $e_m \in E$ such that $e_m \to e$. Then it follows from (i) that

$$\lim_{m\to\infty} r_m^{-2} \mid h_c(r_m e_m)|^2 \leq 1 + c^2 \left(\textstyle\sum_{j=2}^{k+1} E\Delta_j^2\right)^2 - 2c \inf_{e\in E} e'\Gamma_{22}e,$$

where $\Gamma_{22} = E\Delta\Delta'$. This proves the system (2.8) has solutions.

Let $\tilde{a}$ be another solution. By Theorem 2.2 one can show that

$$H(a^*) = H(\tilde{a}) \text{ a.s. } [P_{\xi_0,\eta}],\ U(a^*) = U(\tilde{a}) \quad \text{and} \quad V(a^*) = V(\tilde{a}).$$

Hence $0 = E(H(a^*) - H(\tilde{a}))^2 = (a^* - \tilde{a})'\bar{\Gamma}_{22}(a^*)(a^* - \tilde{a})$. By assumption (A.7) the uniqueness part follows. □.

4. Studentized robust tests. In this section we shall show that the studentized sequence $\{\psi_n^*(\hat{\eta}_n)\}$ of tests is an asymptotic local maximin sequence. Relevant results on the continuities of various quantities as functions of η are summarized below.

LEMMA 4.1.

(i) $h_0(w, a, \eta)$ *of* (2.5) *is continuous on* $\mathbb{R} \times \mathbb{R}^k \times \mathcal{N}$.

(ii) *The implicit functions* $U(a, \eta)$ *and* $V(a, \eta)$ *of* (2.6) *are continuous in* (a, η).

(iii) *The mapping* $g: \mathbb{R}^k \times \mathcal{N} \to \mathbb{R}^k$ *defined by* $g(a, \eta) = EH(a, \eta)\Delta(\eta)$ *is continuous in* (a, η).

(iv) $a^*(\eta)$, $V(a^*(\eta), \eta)$, $U(a^*(\eta), \eta)$ *and* $\sigma_*^2(\eta)$ *are all continuous in* η.

The proof of this lemma is omitted since it follows readily from the assumptions and such basic techniques as the Dominated Convergence Theorem, Scheffe's Theorem, and Vitali's Theorem.

To perform the studentization it now suffices to show that for every $b < \infty$ and $\eta \in \mathcal{N}$ the quantity defined by

$$Z_n^* = \sup_{|s|\leq b} \mid n^{-1/2} \textstyle\sum H^*(\eta + n^{-1/2}s) - n^{-1/2} \sum H^*(\eta)|,$$

tends to zero in probability uniformly over the hypothesis and the alternative nested neighborhoods. In what follows we shall only check the convergence under the hypothesis. The other case can be checked in a similar way. Denote $\mathscr{U}_{n0}(\eta, T) = U\{\mathscr{V}^{(n)}(\epsilon_n; \xi_0, \eta + n^{-1/2}t) : |t| \leq T\}$.

PROPOSITION 4.2. *Let* b, $T < \infty$ *and* $\eta \in \mathcal{N}$ *be given. Then for every* $\delta > 0$ *the quantity* $\sup\{Q_n(|Z_n^*| > \delta) : Q_n \in \mathscr{U}_{n0}(\eta, T)\}$ *tends to zero as n tends to infinity.*

PROOF. For simplicity, let us assume the data is i.i.d. and let Q_n be the product of n-copies of $(1 - \epsilon_n)P_{\theta_n} + \epsilon_n R_n$, where $\theta_n = (\xi_0, \eta + n^{-1/2}t_n)$ with $t_n \to t \in \mathbb{R}^k$ and $\{R_n\}$ is any sequence of probability measures. For this Q_n and for $s \in \mathbb{R}^k$ define

$$\mu_n(s) = E(H^*(\eta + n^{-1/2}s)| Q_n) \quad \text{and} \quad Y_n(s) = n^{-1/2} \textstyle\sum [H^*(\eta + n^{-1/2}s) - \mu_n(s)].$$

It is enough to show

$$\sup_{|s|\leq b} n^{1/2} \mid \mu_n(s) - \mu_n(0)| \to 0 \tag{4.1}$$

and

(4.2) $$\sup_{|s|\le b} \mid Y_n(s) - Y_n(0) \mid \to 0 \text{ in } Q_n\text{-probability.}$$

Let us first prove (4.1). Let s, $s_n \in \mathbb{R}^k$ such that $s_n \to s$ and denote $\eta_n = \eta + n^{-1/2}s_n$. Because of (A.8) (or (A.8)′) and the fact that $EH^*(\eta) = 0$, we need only to show $n^{1/2}EH^*(\eta_n) \to 0$. Denote $\phi_n = [dP_{\xi_0,\eta_n}/dP_{\xi_0,\eta}]^{1/2}$ then by the orthogonality condition (2.8) we have

$$n^{1/2}EH^*(\eta_n) \le E \mid H^*(\eta_n)[n^{1/2}(\phi_n^2 - 1) - s'\Delta(\eta)] \mid + E \mid s'\Delta(\eta)[H^*(\eta_n) - H^*(\eta)] \mid$$

The required result now follows.

Next, we sketch the proof for (4.2). By the well known technique of decomposing a bigger cube as the union of bunch smaller cubes with vertices on a certain grid of points (cf. Bickel, 1975), one can easily demonstrate that the maximum difference between $Y_n(s)$ and $Y_n(0)$ on vertices tends to zero in Q_n-probability for each fixed grid width δ. Because of the assumption (A.8) the largest absolute difference in between, over any cube of the partition, is bounded by some constant times δ. Hence (4.2) follows. □

The next proposition supplements the last remark made at the beginning of Section 3.

PROPOSITION 4.3. *Let (A.1) to (A.7) be satisfied. Let $T < \infty$ and $\eta \in \mathcal{N}$ be given. For each $s \in \mathbb{R}^k$ let $Z_n(s)$ be the random function defined by: $Z_n(s) = n^{-1/2} \sum [H^*(\eta + n^{-1/2}s) - H^*(\eta)]$. Then*

i) *For every $\delta > 0$ and every sequence $\{s_n\}$ converging to $s \in \mathbb{R}$ the quantity $\sup\{Q_n(|Z_n(s_n)| > \delta) : Q_n \in \mathscr{U}_{n0}(\eta, T)\}$ tends to zero iff (A.8)′ is satisfied.*

ii) *(A.8)′ is necessary for the sequence $\{Z_n[n^{1/2}(\hat{\eta}_n - \eta)]\}$ to converge in probability to zero for all* p.m.s. *in $\mathscr{U}_{n0}(\eta, T)$ and for all $\eta \in \mathcal{N}$.*

iii) *Let $\tilde{\eta}_n$ be a version of the discretized estimates obtained from $\hat{\eta}_n$ (cf. LeCam* 1960, *page* 92). *Then $\{\psi_n^*(\tilde{\eta}_n)\}$ is another* ALM *sequence for testing $\{H_n\}$ against $\{K_n\}$.*

PROOF. i) Let Q_n be the same as in the proof of Proposition 4.2. By similar argument as in the proof of (4.1) and Chebyschev's inequality

$$Z_n(s_n) = n^{1/2}\epsilon_n E[H^*(\eta + n^{-1/2}s_n) - H^*(\eta) \mid R_n] + \Delta_n,$$

where $\Delta_n \to 0$ in Q_n-probability. Therefore the necessity of (A.8)′ follows. Evidently (A.8)′ is also sufficient.

The second assertion follows from the first assertion and Proposition A1 and A2 of LeCam (1960).

The last assertion follows from i), Lemma 4.1 and Proposition A3 of LeCam (1960). □

5. Efficiency at the ideal model. In this section, we investigate the efficiency of the ALM test procedure $\psi_n^*(\hat{\eta}_n)$ with respect to the optimal $C(\alpha)$ test at the ideal model $\mathscr{P}$. The purpose of this investigation is to have some definite ideas as to how much loss is incurred if we insist on using the robust test $\psi_n^*(\hat{\eta}_n)$ when the ideal model is correct.

Let $T_n(\eta)$ be as defined by (2.9). Let $\theta_n = (\xi_n, \eta_n)$ where $n^{1/2}(\xi_n - \xi_0) \to \tau > 0$ and $\{n^{1/2}(\eta_n - \eta)\}$ is bounded. By LeCam's third lemma (cf. Hajek and Šidák, 1967) and the orthogonality condition (2.8), it is true that the law of $T_n(\eta)$ under P_{θ_n} tends to the normal variable with mean $\tau S(\eta)$, and variance 1, where $S(\eta)$ is given by (2.11).

Denote $\sigma_J^2(\eta) = E\Delta_1^2(\eta) - \{E\Delta(\eta)\Delta_1(\eta)\}'\{E\Delta(\eta)\Delta'(\eta)\}^{-1}\{E\Delta(\eta)\Delta_1(\eta)\}$. Then the quantity defined by

(5.1) $$e = \sigma_J(\eta)^{-2}[\sigma_*(\eta) - \epsilon\ \{\tau_1(\eta)\sigma_*(\eta)\}^{-1}\ U(a^*(\eta), \eta)]^2,$$

can be used as a measure of the relative performance of the procedures at the ideal model. This quantity is easily seen to be the asymptotic ratio of the respective sample sizes needed for the corresponding test procedures so that they will achieve the same asymptotic minimum power at the ideal model for test H: $\xi = \xi_0$ versus K: $\xi \ge \xi_{n1}(\eta)$, $\eta \in \mathcal{N}$. This measure of efficiency is conservative since the asymptotic size of $\psi_n^*(\hat{\eta}_n)$ at the ideal model is strictly less than α.

The Pitman efficiency of the procedure: $\psi_{n,A}^*(\hat{\eta}_n) = I(T_n(\hat{\eta}_n) \geq z_{1-\alpha})$, is given by

$$e_A = \frac{S^2(\eta)}{\sigma_J^2(\eta)}, \tag{5.2}$$

where $S(\eta)$ is given by (2.11). It will be demonstrated by the example in Section 6 that $\psi_{n,A}^*(\hat{\eta}_n)$ performs well at the ideal model. However this procedure will not maintain the nominal rejection rates when the data is contaminated with outliers.

Because of the final remark in Section 2 and the fact that $V(a^*(\eta), \eta) > 0$, it is clear that $e < e_A \leq 1$. It can be shown that these quantities tend to 1 as ϵ decreases to 0. Therefore if one believes that the total amount of contamination is small but not zero, it is safe to use the robust test $\psi_n^*(\hat{\eta}_n)$ even if the data is actually generated from the ideal model $\mathscr{P}$. However the penalty is severe if one mistakenly assumes the existence of a large amount of contamination when there is none.

6. A robust *t*-test.

Consider the example introduced in Section 1 and let the minimum alternative $\mu_{n1}(\sigma)$ be of the form $\mu_{n1}(\sigma) = \mu_0 + \mu_{n1}^*\sigma$. The asymptotic disjointness condition (3.1) in this example now reduces to

$$\lim_{n\to\infty} \epsilon_n/\mu_{n1}^* < 1/\sqrt{2\pi}. \tag{6.1}$$

Let δ denote the limit in (6.1) and let $[z]_b^a$ denote $\min(a, \max(z, b))$ for real numbers a, b and z. For each δ, define $V(\delta)$ by $E(Z - V(\delta))^+ = \delta$ and let $\Delta^2(\delta) = E([Z]_{-V(\delta)}^{V(\delta)})^2$, where Z is distributed according to the standard normal variable. Then if $\hat{\sigma}_n$ is any $n^{1/2}$-consistent and resistant estimate of σ, our ALM procedure would reject H_{n0} whenever

$$\Delta(\delta)^{-1}n^{-1/2}\sum_{i=1}^n\left[\frac{X_i-\mu_0}{\hat{\sigma}_n}\right]_{-V(\delta)}^{V(\delta)} \geq z_{1-\alpha} + n^{1/2}\epsilon_n\frac{V(\delta)}{\Delta(\delta)}. \tag{6.2}$$

Note that the form of the test statistic is the same as the corresponding one obtained for the case when σ^2 is known; cf. Huber-Carol (1970), Rieder (1978). This is the case because the symmetry of the normal variable implies $a^*(\sigma) = 0$.

In the present example the quantities defined by (5.1) and (5.2) become

$$e(\delta) = \left[\Delta(\delta) + \delta\frac{V(\delta)}{\Delta(\delta)}\right]^2 \quad\text{and}\quad e_A(\delta) = \left[\Delta(\delta) + 2\delta\frac{V(\delta)}{\Delta(\delta)}\right]^2,$$

respectively. Values of these quantities, $V(\delta)$ and $\Delta(\delta)$ are given in the following table.

TABLE 1
$V(\delta)$, $\Delta(\delta)$ and Efficiences

δ	0	.01	.02	.03	.04	.05	.08	.1	.2	.3	.36	$1/\sqrt{2\pi}$
$V(\delta)$	∞	1.94	1.66	1.49	1.36	1.26	1.02	.9	.49	.22	.081	0
$\Delta(\delta)$	1.	.95	.92	.88	.85	.82	.73	.67	.43	.20	.079	0
$e(\delta)$	1.	.95	.91	.87	.83	.80	.71	.65	.43	.27	.20	$1/2\pi$
$e_A(\delta)$	1.	.99	.98	.96	.95	.94	.91	.89	.79	.71	.66	$2/\pi$

From the table we see that the procedure (6.2) performs reasonably well at the normal model for any ϵ_n and μ_{n1}^* such that their ratio ϵ_n/μ_{n1}^* is of an order no larger than 0.5. The severe loss of efficiency at the normal model as $\delta \uparrow 1/\sqrt{2\pi}$ illustrates a remark made in Section 5.

Also note that the procedure which rejects H_{n0} whenever the right side of (6.2) is larger than $z_{1-\alpha}$ performs well at the normal model. In the limit, as δ tends to $1/\sqrt{2\pi}$ it can be readily shown that this procedure tends to the Sign test. This explains why we get the efficiency of $2/\pi$ in the limit for this procedure. Of course, as already remarked in Section 5, this modified procedure will not maintain the nominal level of significance when the data might in fact contain outliers or when the data is non-homogeneous.

Acknowledgment. I would like to thank my advisor, P. J. Bickel, for his guidance, L. LeCam for many helpful discussions, and J. Neyman for giving me the opportunity to work on some of the weather modification problems for which the problem considered in this article originated. Suggestions for improvements by the referee, the Associate Editor and the Editor are also gratefully acknowledged.

REFERENCES

BERAN, R. J. (1980). Efficient robust tests in parametric models. Preprint.

BICKEL, P. J. (1975). One-step Huber estimates in the linear model. *J. Amer. Statist. Assoc.* **70** 428–434.

HAJEK, J., ŠIDÁK, Z. (1967). *Theory of Rank Tests*, Academic Press, New York.

HUBER, P. J. (1965). A robust version of the probability ratio test. *Ann. Math. Statist.* **36** 1753–1758.

HUBER-CAROL, C. (1970). Étude asymptotique de tests robustes. Thèse, E.T.H. Zürich.

LE CAM, L. (1960). Locally asymptotically normal families distributions. *Univ. Calif. Publ. Statist.* **3** 37–98.

LE CAM, L. (1969). *Théorie Asymptotique de la Decision Statistique.* Les Presses de l'Université de Montréal.

LE CAM, L. (1970). Remarques sur le théorème limite central dans les espaces localement convexes. *Probabilités sur les Structures Algébriques.* CNRS Paris, 233–248.

MILLAR, P. W. (1980). Robust tests of statistical hypotheses. Preprint.

NEUHAUS, G. (1975). Convergence of the reduced empirical process for non-i.i.d. random vectors. *Ann. Statist.* **3** 528–531.

NEYMAN, J. (1958). Optimal asymptotic tests of composite statistical hypotheses. *Probability and Statistics: The Cramér Volume.* Almquist and Wiksells, Uppsala, Sweden, 213–234.

QUANG, P. X. (1974). Robust sequential testing. Thesis, University of California, Berkeley.

RIEDER, H. (1978). A robust asymptotic testing model. *Ann. Statist.* **6** 1080–1094.

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

VAN ZUIJLEN, M. C. A. (1978). Properties of the empirical distribution function for independent nonidentically distributed random variables. *Ann. Probability.* **6** 250–266.

DEPARTMENT OF MATHEMATICS AND STATISTICS
THE UNIVERSITY OF CALGARY
CALGARY, ALBERTA, T2N IN4
CANADA