Aplikace matematiky

van Huu Nguyen

Rank test of hypothesis of randomness against a group of regression alternatives

Aplikace matematiky, Vol. 17 (1972), No. 6, 422–447

Persistent URL: http://dml.cz/dmlcz/103435

Terms of use:

© Institute of Mathematics AS CR, 1972

Institute of Mathematics of the Czech Academy of Sciences provides access to digitized documents strictly for personal use. Each copy of any part of this document must contain these *Terms of use*.

This document has been digitized, optimized for electronic delivery and stamped with digital signature within the project *DML-CZ: The Czech Digital Mathematics Library* http://dml.cz

RANK TEST OF HYPOTHESIS OF RANDOMNESS AGAINST A GROUP OF REGRESSION ALTERNATIVES[1]

NGUYEN-VAN-HUU

(Received September 6, 1971)

1. SUMMARY

In this work the problem of testing the hypothesis of randomness against a group of alternatives of regression in a parameter involved in the distributions of random observations is investigated and a rank test for this problem is suggested. This problem is a generalization of the problem of detecting a shift in a location parameter of a distribution occurring at an unknown time point between consecutively taken observations. The latter problem was considered and a rank test for it was proposed by Bhattacharyya and Johnson (1968). The rank test in this work is shown to be locally average most powerful within the class of all possible rank tests in the sense of the definition in Section 3 below. The asymptotic normality of the rank test statistic and the asymptotic efficiency of the rank test are shown not only for the case of location and scale parameters but for the case of a general parameter.

The parametric test for a similar problem for the density of a one-parameter exponential family and a rank decision rule for a combined problem of testing and classification will appear in subsequent papers.

2. INTRODUCTION

Throughout this paper let $X_1, \ldots, X_N$ be independent observations which are supposed to have absolutely continuous distribution functions with densities $f_1(x), \ldots$ $\ldots, f_N(x)$ with respect to Lebesgue measure.

Let H_0 be the hypothesis under which

$$f_1(x) = \ldots = f_N(x) = f(x) \tag{1}$$

where $f(x)$ is an element of a certain family $\mathscr{F}$ of density functions.

---

[1]) This article is a part of author's thesis prepared during his stay in the Mathematical Institute of the Czechoslovak Academy of Sciences.

Let K_m, $m = 1, \ldots, s$, be the alternative under which

$$f_1(x) = f(x, \Delta C_{m1}), \ldots, f_N(x) = f(x, \Delta C_{mN}) \tag{2}$$

where $f(x, 0) = f(x)$; C_{mj} are the so-called regression constants, Δ is an unknown parameter. Then K_m is called the regression alternative.

It is required to test H_0 against $K_1, \ldots, K_s$. A special case of this problem where $f(x, \theta) = f(x - \theta)$ and

$$C_{mj} = 0, 1 \quad \text{if} \quad m \geqq j\,, \quad m < j\,, \quad \text{respectively,} \tag{3}$$

was investigated by Bhattacharyya and Johnson in [1].

The other special cases of this problem are as follows:

Putting

$$C_{ii} = 1\,, \quad C_{ij} = 0 \quad \text{for} \quad i \neq j\,, \quad i, j = 1, \ldots, N \tag{4}$$

we obtain the problem of slippage in a parameter.

Putting in (2)

$$C_{mj} = 0 \quad \text{or} \quad (j - N + m)/m$$
$$\text{if} \quad j \leqq N - m \quad \text{or} \quad j \geqq N - m + 1\,, \quad \text{respectively}\,, \tag{5}$$
$$\text{for} \quad m = 1, \ldots, s \quad (s \leqq N - 1)$$

we obtain the growth problem (I) where the alternatives K_m, $m = 1, \ldots, s$, express the fact that the parameter remains unchanged until the time point $k = N - m$ and then it grows linearly up to the value Δ so that the rates of growth are different for different alternatives (see Figure 1 with $s = 3$, $N = 6$).

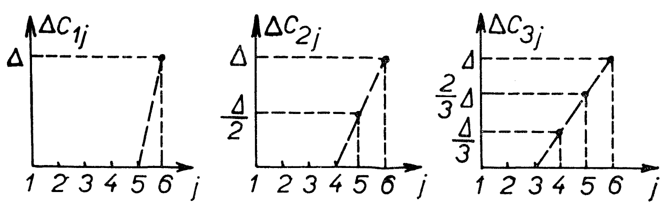

Fig. 1.

Similarly, putting in (2)

$$C_{mj} = 0 \quad \text{or} \quad j - N + m$$
$$\text{if} \quad j \leqq N - m \quad \text{or} \quad j \geqq N - m + 1\,, \quad \text{respectively,} \tag{6}$$
$$\text{for} \quad m = 1, \ldots, s \quad (s \leqq N - 1)\,,$$

then we obtain the growth problem (II) where the alternatives K_m corresponding to the regression constants (6) express the fact that the parameter remains unchanged until the time point $k = N - m$ and then it grows linearly up to the value s at the same rate of growth for all alternatives (see Figure 2 with $s = 3$, $N = 6$).

The problem of testing hypotheses of changes in parameters – a special case of the above problem with the regression constants given by (3) – was investigated

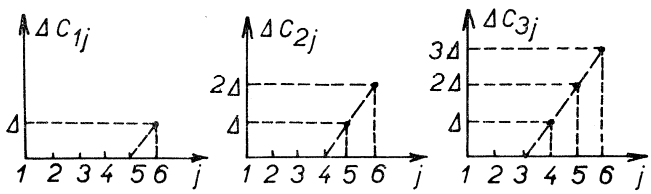

Fig. 2.

by Page [9], [10], Chernoff and Zacks [2] for the mean of normal distribution and by Kander and Zacks [6] for the parameter of the one-parameter exponential family. The tests suggested by these authors are based directly on observations and not on ranks. Bhattacharyya and Johnson are the first who proposed a test based on ranks.

3. RANK TESTS

1. Notations.

Let us denote the ordered sample from $X_1, \ldots, X_N$ by $X^{(1)} < X^{(2)} < \ldots < X^{(N)}$ and the ranks of $X_1, \ldots, X_N$ by $R_1, \ldots, R_N$.

Put $X^{(\cdot)} = (X^{(1)}, \ldots, X^{(N)})$, $R = (R_1, \ldots, R_N)$ and let $x^{(\cdot)} = (x^{(1)}, \ldots, x^{(N)})$, $r = (r_1, \ldots, r_N)$ be a realization of $X^{(\cdot)}$ and R, respectively.

Let U be the uniformly distributed random variable on $(0, 1)$ and $U^{(\cdot)} = (U^{(1)}, \ldots$ $\ldots, U^{(N)})$ the ordered sample from the observations $U_1, \ldots, U_N$ on U. E_0 denotes the expectation under H_0.

2. Locally average most powerful LAMP rank test of H_0 against $K_1, \ldots, K_s$.

Let T be any test of H_0 against $K_1, \ldots, K_s$ and β_T its power function under K_m. β_T depends on Δ and m, i.e. $\beta_T = \beta_T(\Delta, m)$.

Put

$$\bar{\beta}_T(\Delta, p) = \sum_{m=1}^{s} p_m \beta_T(\Delta, m) . \tag{7}$$

$\bar{\beta}_T(\Delta, p)$ is called the average power function of the test T with respect to the weights $p_1, \ldots, p_s$ where $p_m \geqq 0$, $\sum_{m=1}^{s} p_m = 1$.

It is required to find a test which maximizes $\bar{\beta}_T(\Delta, p)$ within the class of all possible tests for each fixed $p = (p_1, \ldots, p_s)$ and for all Δ. In general such a test does not exist. Let us confine ourselves to a narrower class – the class of all rank tests.

Definition. *The α-level test T^* possessing the property that there exists an $\varepsilon > 0$ such that T^* maximizes $\bar{\beta}_T(\Delta, p)$ within the class of all α-level tests for all $0 < \Delta \leqq \varepsilon$ is called the locally average most powerful test with respect to the weights $p_1, \ldots, p_s$.*

Theorem 1. *Assume that $f(x, \theta)$ involved in (2) has the following properties:*
(A_1) For each x, $f(x, \theta)$ is absolutely continuous in $\theta \in J$, where J is an open interval containing the point 0 and

$$\lim_{\theta \to 0} [f(x, \theta) - f(x)]/\theta = \dot{f}(x, 0)$$

$$(A_2) \quad \lim_{\theta \to 0} \int_{-\infty}^{\infty} |\dot{f}(x, \theta)| \, dx = \int_{-\infty}^{\infty} \dot{f}(x, 0)| \, dx$$

holds where $\dot{f}(x, \theta)$ denotes the partial derivative of $f(x, \theta)$ in θ.

Then the test with the rejection region

$$T_{N,p}(R) > C_\alpha \tag{8}$$

where

$$T_{N,p}(r) = \sum_{k=1}^{N} C_k(p) \, E_0[\dot{f}(X^{(r_k)}, 0)/f(X^{(r_k)})] \tag{9}$$

with

$$C_k(p) = \sum_{m=1}^{s} C_{mk} p_m \tag{10}$$

is the LAMP rank test at the level α within the class of all α-level tests depending only on R for testing H_0 against $K_1, \ldots, K_s$.

Proof. Let

$$q_{\Delta m}(x) = \prod_{j=1}^{N} f(x_j, \Delta C_{mj}) \tag{11}$$

be the joint density of $X_1, \ldots, X_N$ under K_m and put

$$q_\Delta(x) = \sum_{m=1}^{s} p_m \, q_{\Delta m}(x) \, . \tag{12}$$

Let $Q_\Delta(.)$, $Q_{\Delta m}(.)$ be the probability measures with respect to the densities q_Δ and $q_{\Delta m}$. Consider the problem of testing H_0 against a simple alternative q_Δ with Δ fixed. According to Neyman-Pearson's Lemma (see [7]) the most powerful rank test at the level α for testing H_0 against q_Δ is given by the critical function

$$\Phi_\Delta(r) = 1, \gamma, 0 \quad \text{if} \quad Q_\Delta\{R = r\} >, = , < C'_\alpha \tag{13}$$

respectively, since the vector R is, under H_0, uniformly distributed. Hereafter $\Phi(r)$ denotes the probability of rejecting H_0 when r is a realization of R. The constants C'_α and γ are defined so that the test has the significance level α. Let $\Phi'(r)$ be the critical function of any rank test. Then the power function of $\Phi'(r)$ under q_Δ is given by:

$$\sum_r \Phi'(r)\, Q_\Delta\{R = r\} = \sum_{m=1}^{s} p_m \sum_r \Phi'(r)\, Q_{\Delta m}\{R = r\} = \tag{14}$$

$$= \sum_{m=1}^{s} p_m \beta_{\Phi'}(\Delta, m) = \bar{\beta}_{\Phi'}(\Delta, p)$$

where $\beta_{\Phi'}(\Delta, m)$ denotes the power of Φ' under $q_{\Delta m}$ and the summation in r is over all possible permutations of $\{1, 2, \ldots, N\}$. Consequently

$$\bar{\beta}_{\Phi_\Delta}(\Delta, p) \geqq \bar{\beta}_{\Phi'}(\Delta, p)\,. \tag{15}$$

Let us calculate $Q_\Delta\{R = r\}$. We have

$$Q_\Delta\{R = r\} = \int \ldots \int_{\{R=r\}} q_\Delta(x_1, \ldots, x_N)\, \mathrm{d}x_1 \ldots \mathrm{d}x_N = \tag{16}$$

$$= \int \ldots \int_{\{R=r\}} \prod_{i=1}^{N} f(x_i)\, \mathrm{d}x_i + \sum_{m=1}^{s} p_m \int \ldots \int_{\{R=r\}} \Big[\prod_{i=1}^{N} f(x_i, \Delta C_{mi}) - \prod_{i=1}^{N} f(x_i)\Big] \times$$

$$\times\, \mathrm{d}x_1 \ldots \mathrm{d}x_N = 1/N! + \sum_{m=1}^{s} p_m \sum_{k=1}^{N} \int \ldots \int_{\{R=r\}} [f(x_k, \Delta C_{mk}) - f(x_k)] \times$$

$$\prod_{j=k+1}^{N} f(x_j) \prod_{i=1}^{k-1} f(x_i, \Delta C_{mi})\, \mathrm{d}x_1 \ldots \mathrm{d}x_N = 1/N! + \sum_{m=1}^{s} \sum_{k=1}^{N} C_{mk} p_m g_{mk}(\Delta)$$

where

$$g_{mk}(\Delta) = \int \ldots \int_{\{R=r\}} [f(x_k, \Delta C_{mk}) - f(x_k)]\, (\Delta C_{mk})^{-1} \tag{17}$$

$$\prod_{j=k+1}^{N} f(x_j) \prod_{i=1}^{k-1} f(x_i, \Delta C_{mi})\, \mathrm{d}x_1 \ldots \mathrm{d}x_N\,.$$

In view of the conditions of Theorem 1 we obtain:

$$\lim_{\Delta\to 0} f(x_i, \Delta C_{mi}) = f(x_i)\,, \tag{18}$$

$$\lim_{\Delta\to 0} [f(x_k, \Delta C_{mk}) - f(x_k)]/\Delta C_{mk} = \dot{f}(x_k, 0)\,, \tag{19}$$

$$\limsup_{\Delta\to 0} \int_{-\infty}^{\infty}\dots\int_{-\infty}^{\infty} |f(x_k, \Delta C_{mk}) - f(x_k)|\,|\Delta C_{mk}|^{-1} \times \tag{20}$$
$$\times \prod_{j=k+1}^{N} f(x_j) \prod_{i=1}^{k-1} f(x_i, \Delta C_{mi})\,\mathrm{d}x_1 \dots \mathrm{d}x_N =$$
$$= \limsup_{\Delta\to 0} \int_{-\infty}^{\infty} |f(x_k, \Delta C_{mk}) - f(x_k)|\,|\Delta C_{mk}|^{-1}\,\mathrm{d}x_k =$$
$$= \limsup_{\Delta\to 0} \int_{-\infty}^{\infty} \left| \int_0^{\Delta C_{mk}} \dot{f}(x, \theta)\,\mathrm{d}\theta \right| |\Delta C_{mk}|^{-1}\,\mathrm{d}x \leqq$$
$$\limsup \frac{1}{|\Delta C_{mk}|} \int_0^{|\Delta C_{mk}|} \left(\int_{-\infty}^{\infty} |\dot{f}(x, \theta)|\,\mathrm{d}x \right) \mathrm{d}\theta = \int_{-\infty}^{\infty} |\dot{f}(x, 0)|\,\mathrm{d}x, \quad \text{by} \quad (\mathrm{A}_2)\,.$$

It follows from (17)–(20) and Theorem II. 4.2 in [3] that

$$\lim_{\Delta\to 0} g_{mk}(\Delta) = \int\dots\int_{\{R=r\}} [\dot{f}(x_k, 0)/f(x_k)] \prod_{i=1}^{N} f(x_i)\,\mathrm{d}x_i = \tag{21}$$
$$= E_0\{[\dot{f}(X_k, 0)/f(X_k)]|R = r\}\,P\{R = r\} =$$
$$= (1/N!)\,E_0[\dot{f}(X^{(r_k)}, 0)/f(X^{(r_k)})]$$

since $f(x, \theta) \geqq 0$ for all θ and in view of the condition (A_1), $f(x_k) = 0$ implies $\dot{f}(x_k, 0) = 0$ a.e., therefore the first equality in (20) holds; the last equality in (20) follows from the independence of $X^{(\cdot)}$ on R (see Theorem II. 1.2. a in [3]).

It follows from (21) and (16) that

$$\lim_{\Delta\to 0} [Q_\Delta\{R = r\} - 1/N!]/\Delta = (1/N!) \sum_{k=1}^{N} C_k(p)\,E_0\{\dot{f}(X^{(r_k)}, 0)/f(X^{(r_k)})\} =$$
$$= (1/N!)\,T_{N,p}(r)\,.$$

Consequently, there exists an $\varepsilon > 0$ such that $Q_\Delta\{R = r\}$ is a strictly increasing function of $T_{N,p}$ for all $0 < \Delta \leqq \varepsilon$ and hence there is a constant C_α such that (13) may be written in the form

$$\Phi(r) = 1, \gamma, 0 \quad \text{if} \quad T_{N,p}(r) >, =, < C_\alpha \quad \text{respectively.}$$

The function does not depend on $\Delta \in (0, \varepsilon]$. Q.E.D.

Corollary 1. *Suppose that $f(x, \theta) = f(x - \theta)$, i.e. θ is a location parameter and that*

(A_1') *$f(x)$ is absolutely continuous,*

$$(A_2') \quad \int_{-\infty}^{\infty} |f'(x)| \, dx < \infty ,$$

where $f'(x)$ denotes the almost everywhere derivative of $f(x)$. Then for testing H_0 against $K_1, \ldots, K_s$ there exists an α-level rank test defined by the rejection region

$$T_{N,p}^{(1)}(R) > C_\alpha \tag{22}$$

where

$$T_{N,p}^{(1)}(r) = \sum_{k=1}^{N} C_k(p) \, E_0[-f'(X^{(r_k)})/f(X^{(r_k)})] . \tag{23}$$

The test is LAMP within the class of all one-sided rank tests at the level α.

Proof. It is easy to see that the conditions (A_1), (A_2) of Theorem 1 are fulfilled provided (A_1'), (A_2') are satisfied, hence Corollary 1 follows from Theorem 1.

Remark 1. If the regression constants C_{mj} assume the form (3) with $s = N - 1$ then we obtain from Corollary 1 the test given by (22) with

$$T_{N,p}^{(1)}(r) = \sum_{k=2}^{N} P_{k-1} \, E_0[-f'(X^{(r_k)})/f(X^{(r_k)})] \tag{24}$$

where $P_k = \sum_{m=1}^{k} p_m$ are the cumulative weights. This test was suggested by Bhattacharyya and Johnson in [1].

Corollary 2. *Suppose that $f(x, \theta) = \exp(-\theta) f((x - \eta) \exp(-\theta))$ where η is a nuisance parameter, θ in an unknown scale parameter and that*

(A_1'') *$f(x)$ is absolutely continuous,*

$$(A_2'') \quad \int_{-\infty}^{\infty} |x f'(x)| \, dx < \infty .$$

Then the test given by the rejection region

$$T_{N,p}^{(2)}(R) > C_\alpha \tag{25}$$

where

$$T_{N,p}^{(2)}(r) = \sum_{k=1}^{N} C_k(p) \, E_0[-1 - X^{(r_k)} f'(X^{(r_k)})/f(X^{(r_k)})] \tag{26}$$

is the LAMP rank test for testing H_0 against $K_1, \ldots, K_s$.

Proof. We observe that, under H_0, $f_1(x) = \ldots = f_N(x) = f(x - \eta)$ and from (A_1''), (A_2'') we obtain

$$\dot{f}(x, \theta) = -f[(x - \eta) \exp(-\theta)] \exp(-\theta) - (x - \eta) f'[(x - \eta) . \\ . \exp(-\theta)] \exp(-2\theta)$$

hence

$$\int_{-\infty}^{\infty} |\dot{f}(x, \theta)| \, dx = \int_{-\infty}^{\infty} |f(x - \eta) + (x - \eta) f'(x - \eta)| \, dx = \int_{-\infty}^{\infty} |\dot{f}(x, 0)| \, dx \leqq$$

$$\leqq \int_{-\infty}^{\infty} f(x) \, dx + \int_{-\infty}^{\infty} |x f'(x)| \, dx = 1 + \int_{-\infty}^{\infty} |x f'(x)| \, dx < \infty .$$

Thus the conditions (A_1), (A_2) of Theorem 1 are fulfilled. Consequently, we obtain from Theorem 1 a LAMP rank test defined by the rejection region $T_{Np}(R) > C_\alpha$ where

$$T_{Np}(r) = \sum_{k=1}^{N} C_k(p) \, E_0^*\{[-f(X^{(r_k)} - \eta) - (X^{(r_k)} - \eta) f'(X^{(r_k)} - \eta)]/f(X^{(r_k)} - \eta)\} =$$

$$= \sum_{k=1}^{N} C_k(p) \, E_0\{-1 - X^{(r_k)} f'(X^{(r_k)})/f(X^{(r_k)})\} = T_{Np}^{(2)}(r)$$

with E_0^*, E_0 denoting the expectation under the hypothesis that the common density is $f(x - \eta)$ and $f(x)$, respectively.

Remark 2. Let $F(x)$ be the distribution function with respect to the density $f(x)$ and let

$$F^{-1}(u) = \inf\{x : F(x) \geqq u\}, \quad 0 < u < 1,$$

(27) $$\varphi(u, f) = \dot{f}(F^{-1}(u), 0)/f(F^{-1}(u)),$$

(28) $$a_N(i, f) = E\varphi(U^{(i)}, f), \quad i = 1, 2, \ldots, N.$$

Then

(29) $$T_{Np}(R) = \sum_{k=1}^{N} a_n(R_k, f) \, C_k(p)$$

(see expression (3) of II. 4.3 in [3]). $\varphi(u, f)$ is called the score function, and $a_N(i, f)$ are called the scores.

3. Locally average most powerful rank tests of the hypothesis of randomness with a symmetric distribution

Consider the hypothesis H_0^* under which the densities of $X_1, \ldots, X_N$ satisfy

(30) $$f_1(x) = \ldots = f_N(x) = f(x) \quad \text{with} \quad f(x) = f(-x)$$

and consider the alternative K_m^*, $m = 1, 2, \ldots, s$, under which

$$f_1(x) = f(x, \Delta C_{m1}), \ldots, f_N(x) = f(x, \Delta C_{mN}) \tag{31}$$

with $f(x, 0) = f(x)$, C_{mj}'s known, Δ being an unknown parameter.

Let $|X|^{(1)} < \ldots < |X|^{(N)}$ be the ordered sample from $|X_1|, \ldots, |X_N|$ and $R_1^+, \ldots, R_N^+$ the ranks of $|X_1|, \ldots, |X_N|$; let $v = (v_1, \ldots, v_N)$ be a realization of the vector sign $X =$ $=$ (sign $X_1, \ldots,$ sign X_N) . v_i assume the values 1 or -1.

Theorem 2. *Suppose that $f(x, \theta)$ occurring in K_m^* satisfies the following conditions:* (A_1^*) *For each $x, f(x, \theta)$ is absolutely continuous in $\theta \in J$, where J is an open interval containing the point* 0, *and there exists*

$$\lim_{\theta \to 0} [f(x, \theta) - f(x)]/\theta = \dot{f}(x, 0)$$

where $\dot{f}(x, 0)$ may be expressed in the form

$\dot{f}(x, 0) = u(\operatorname{sign} x)\, t(|x|)$ *with u, t being some functions,*

$$(\mathrm{A}_2^*) \quad \lim_{\theta \to 0} \int_{-\infty}^{\infty} |\dot{f}(x, \theta)|\, \mathrm{d}x = \int_{-\infty}^{\infty} |\dot{f}(x, 0)|\, \mathrm{d}x\,.$$

Then the test with the rejection region

$$T_{Np}^*(R^+, \operatorname{sign} X) > C_\alpha \tag{32}$$

where

$$T_{Np}^*(r, v) = \sum_{k=1}^{N} C_k(p)\, u(v_k)\, E_0[t(|X|^{(r_k)})/f(|X|^{(r_k)})]$$

is LAMP within the class of all α-level tests depending only on R^+ and sign X *for testing H_0^* against $K_1^*, \ldots, K_s^*$.*

Proof. According to Theorem II. 1.3 in [3] the vector R^+ and sign X are, under H_0, mutually independent and

$$P\{\operatorname{sign} X = v\} = 2^{-N},\ P\{R^+ = r\} = 1/N!\,.$$

Let $q_{m\Delta}^* = \prod_{k=1}^{N} f(x_k, \Delta C_{mk})$ be the joint density of $X_1, \ldots, X_N$ under K_m^* and let $q_\Delta^* = \sum_{m=1}^{s} p_m\, q_{m\Delta}^*$.

Let $Q_{m\Delta}^*$, Q_Δ^* be the probability measures with respect to $q_{m\Delta}^*$, q_Δ^*, respectively.

By Neyman-Pearson's Lemma, the most powerful rank test within the class of all α-level tests depending only on R^+ and sign X for testing H_0^* against a simple alternative q_Δ^* with Δ fixed is defined by the following critical function:

$$\Phi_\Delta(r, v) = 1, \gamma, 0 \quad \text{if} \quad Q_\Delta^*\{R = r, \operatorname{sign} X = v\} >, =, < C'_\alpha \quad \text{respectively.} \tag{33}$$

It is easy to see that if $\Phi'(r, v)$ is any critical function depending only on r and v then $\bar{\beta}_{\Phi_\Delta}(\Delta, p) \geqq \bar{\beta}_{\Phi'}(\Delta, p)$ where $\bar{\beta}_{\Phi_\Delta}(\Delta, p)$ and $\beta_{\Phi'}(\Delta, p)$ denote the average powers of the tests defined by Φ_Δ and Φ'. Thus Φ_Δ defines the average most powerful rank test within the class of all α-level tests depending only on R^+ and sign X.

It is easy to prove that if $f(x, \theta)$ satisfies the conditions (A_1^*), (A_2^*) then

$$\lim_{\Delta \to 0} [2^N N! \, Q_\Delta\{R = r, \operatorname{sign} X = v\} - 1]/\Delta =$$

$$= \lim_{\Delta \to 0} 2^N N! \sum_{m=1}^{s} p_m \sum_{k=1}^{N} \int \dots \int_{\{R=r,\ \operatorname{sign} X = v\}} [f(x_k, \Delta C_{mk}) - f(x_k)] \, \Delta^{-1} \, .$$

$$\cdot \prod_{i=k+1}^{N} f(x_i) \prod_{j=1}^{k-1} f(x_j, \Delta C_{mj}) \, dx_1, \dots, dx_N =$$

$$= 2^N N! \sum_{m=1}^{s} p_m \sum_{k=1}^{N} C_{mk} u(v_k) \int \dots \int_{\{R=r,\ \operatorname{sign} X = v\}} [t(|x_k|)/f(|x_k|)] \prod_{i=1}^{N} f(x_i) \, dx_i =$$

$$= \sum_{k=1}^{N} C_k(p) \, u(v_k) \, E_0[t(|X|^{(r_k)})/f(|X^{(r_k)}|)] = T_{Np}^*(r, v) \, .$$

Consequently, there exists an $\varepsilon > 0$ such that (33) is equivalent to (32) for all $0 < \Delta \leqq \varepsilon$.

Corollary 3. *Suppose that $f(x, \theta)$ involved in K_m^*, $m = 1, \dots, s$, assumes the form $f(x, \theta) = f(x - \theta)$ with $f(x) = f(-x)$ and that $f(x)$ satisfies the conditions (A_1'), (A_2') of Corollary 1. Then the test with the rejection region*

$$T_{Np}^{(*)}(R^+, \operatorname{sign} X) > C_\alpha \tag{34}$$

where

$$T_{N,p}^{(*)}(r, \operatorname{sign} x) = \sum_{k=1}^{N} C_k(p) \operatorname{sign} x_k \, E_0[-f'(|X|^{(r_k)})/f(|X|^{(r_k)})] \tag{35}$$

is the LAMP rank test within the class of all α-level tests depending only on R^+ and sign X *for testing H_0^* against $K_1^*, \dots, K_s^*$.*

Proof. We observe that under the conditions of this corollary the conditions (A_1^*), (A_2^*) are fulfilled and $\dot{f}(x, 0) = -f'(x) = -\operatorname{sign} x \, f'(|x|)$ since $f(x)$ is symmetric, thus $u(\operatorname{sign} x) = -\operatorname{sign} X$, $t(|x|) = f'(|X|)$, hence (35) follows from (32).

Remark 1. Theorem 1 of Bhattacharyya and Johnson in [1] is a direct consequence of Corollary 3, by letting C_{mj} assume the form (3).

Corollary 4. *Suppose that $f(x, \theta)$ occurring in K_m^*, $m = 1, \dots, s$, assumes the form*

$$f(x, \theta) = \exp(-\theta) f(x \exp(-\theta)) \quad \text{with} \quad f(x) = f(-x)$$

and that $f(x)$ satisfies the conditions (A_1''), (A_2'') *of Corollary* 2. *Then the test with the rejection region*

$$T^{(**)}_{N,p}(R^+) > C_\alpha \tag{36}$$

where

$$T^{(**)}_{N,p}(r) = \sum_{k=1}^{N} C_k(p)\, E_0[-1 - |X|^{(r_k)} f'(|X|^{(r_k)})/f(|X^{(r_k)}|)] \tag{37}$$

is the LAMP rank test within the class of all α-level tests depending only on R^+ and sign X *for testing H_0^* against $K_1^*, \dots, K_s^*$.*

Proof. It is easy to see that under the conditions of Corollary 4, the conditions (A_1^*), (A_2^*) are also satisfied and

$$\dot{f}(x, 0) = \dot{f}(|x|, 0) = -f(|x|) - |x| f'(|x|)\,.$$

Thus $u(\text{sign}\, x) \equiv 1$ and $t(|x|) = \dot{f}(|x|, 0)$ hence (37) follows from (32).

Remark 2. Let $F(x)$ be the distribution function with respect to $f(x)$, $F^{-1}(u) = \inf\{x : F(x) \geqq u\}$, $0 < u < 1$,

$$\varphi_1(u, f) = -f'(F^{-1}(u))/f(F^{-1}(u))\,, \tag{38}$$

$$\varphi_2(u, f) = -1 - F^{-1}(u) f'(F^{-1}(u))/f(F^{-1}(u)) \tag{39}$$

which are the special forms of $\varphi(u, f)$ given by (27). Putting

$$\varphi^+(u, f) = \varphi_1(\tfrac{1}{2} + \tfrac{1}{2}u, f)\,, \tag{40}$$

$$a^+_{1N}(i, f) = E\varphi^+_1(U^{(i)}, f)\,, \tag{41}$$

$$\varphi^+_2(u, f) = \varphi_2(\tfrac{1}{2} + \tfrac{1}{2}u, f)\,, \tag{42}$$

$$a^+_{2N}(i, f) = E\,\varphi^+_2(U^{(i)}, f) \tag{43}$$

then the test statistics given by (35), (37) may be written in the form:

$$T^{(*)}_{N,p}(R^+, \text{sign}\, X) = \sum_{k=1}^{N} C_k(p)\, \text{sign}\, X_k\, a^+_{1N}(R^+_k, f)\,, \tag{44}$$

$$T^{(**)}_{N,p}(R^+, \text{sign}\, X) = \sum_{k=1}^{N} C_k(p)\, a_{2N}(R^+_k, f)\,. \tag{45}$$

4. Unbiasedness of LAMP rank tests

In this section let us consider the rank test given by

$$T(R) = \sum_{k=1}^{N} C_k\, a_N(R_k) \tag{46}$$

where

$$a_N(i) = E\,\varphi(U^{(i)}) \tag{47}$$

or

$$a_N(i) = \varphi(i/(N+1)) \tag{48}$$

with φ being an arbitrary score function.

Consider the null hypothesis defined above and the alternative K defined by

$$f_1(x) = g(x, d_1), \ldots, f_N(x) = g(x, d_N) \tag{49}$$

where, as usual, $f_1, \ldots, f_N$ denote the densities of the observations $X_1, X_2, \ldots, X_N$.

Let $g(x, 0) = g(x)$ and let $G(x)$ be the distribution function with respect to the density $g(x)$.

Definition. *We say that the density $g(x, \theta)$ has the property T if for every θ there exists a transformation $T_\theta : X \to T_\theta X$ such that when X has the density $g(x, 0)$ then $T_\theta X$ has the density $g(x, \theta)$ and if $\theta_i \leqq \theta_j$, $X_i < X_j$ then*

$$T_{\theta_i} X_i < T_{\theta_j} X_j\,. \tag{50}$$

It is obvious that T_θ must satisfy $T_0 X \equiv X$.

Theorem 3. *Suppose that $g(x, \theta)$ has the property T, then any rank test rejecting H_0 as $T(R)$ is sufficiently large is unbiased for testing H_0 against K provided*

$$(C_i - C_j)(d_i - d_j) \geqq 0 \quad \textit{for all} \quad i, j = 1, 2, \ldots, N \tag{51}$$

and $\varphi(u)$ is non-decreasing.

Proof. With no loss of generality we can suppose that $d_1 \leqq \ldots \leqq d_N$.

Let $X_1, \ldots, X_N$ have the same density $g(x, 0)$ then $T_{d_1} X_1, \ldots, T_{d_N} X_N$ have the densities $g(x, d_1), \ldots, g(x, d_N)$, respectively. Let $R_1, \ldots, R_N$ be the ranks of $X_1, \ldots, X_N$ and $R'_1, \ldots, R'_N$ the ranks of $T_{d_1} X_1, \ldots, T_{d_N} X_N$. It is sufficient to show that

$$T(R') \geqq T(R)\,. \tag{52}$$

Assume that $R_i < R_j$, i.e. $X_i < X_j$ for $i < j$, then, by the property T, $T_{d_i} X_i < T_{d_j} X_j$ since $d_i \leqq d_j$, thus $R'_i < R'_j$. Consequently, $R'_1, \ldots, R'_N$ is better ordered than $R_1, \ldots, R_N$ (see Definition in [8]). Applying Corollary 2 of Theorem 5 in [8] with a slight generalization, we obtain (52) since (51) together with the assumption that $d_i \leqq d_j$ implies $C_i \leqq C_j$ for all $i < j$ and the assumption that $\varphi(u)$ is non-decreasing implies that $a_N(j) \geqq a_N(i)$ for all $i < j$. Q.E.D.

Example 1. Let $g(x, \theta) = g(x - \theta)$; then $g(x, \theta)$ has the property T with $T_\theta X = X + \theta$. Put in (49)

$$d_1 = \ldots = d_m = 0\,, \quad d_{m+1} = \ldots = d_N = 1 \tag{53}$$

with m arbitrary fixed ($m = 1, \ldots, N - 1$).

Then (51) is fulfilled provided $C_i \leqq C_j$ for all $i < j$ and Theorem 3.2 of Bhattacharyya and Johnson in [1] may be obtained from Theorem 3.

Example 2. Let $f(x, \theta) = \exp(-\theta) f(x \exp(-\theta))$ for $x > 0$
$= 0$ for $x \leqq 0$;

then $f(x, \theta)$ has the property T with $T_\theta X = (\exp(\theta)) X$. Consequently, Theorem 3 applies to this density.

Example 3. Let $g(x, \theta) = \exp(-x/(1 + \theta))/(1 + \theta)$ for $x > 0$,
$= 0$ otherwise, where $1 + \theta > 0$;

then $g(x, \theta)$ has the property T with $T_\theta X = (1 + \theta) X$ and Theorem 3 also applies to such a density.

5. Asymptotic normality of rank test statistics under H_0

In this section we shall show that under some conditions the test statistic $T_{Np}(R)$ given by (9) or (27)–(29) is asymptotically normal under H_0. However, the test statistic $T_{Np}(R)$ is only a special case of the following statistic:

$$T'_{Np}(R) = \sum_{k=1}^{N} C_k(p)\, a_N(R_k) \tag{54}$$

where the scores satisfy

$$\int_0^1 [a_N(1 + [uN]) - \varphi(u)]^2 \,\mathrm{d}u \to 0 \quad \text{as} \quad N \to \infty \tag{55}$$

($[uN]$ denotes the entier of uN) with $\varphi(u)$ square integrable and $C_k(p)$ are defined by (10).

Actually, if $\varphi(u, f)$ given by (27) is square integrable, then by Theorem V.1.4.b in [3] the scores of the test statistic $T_{Np}(R)$ given by (29) satisfy (55). Consequently, we shall consider the statistic $T'_{Np}(R)$ instead of $T_{Np}(R)$.

Theorem 4. *Assume that $\varphi(u)$ is square integrable and*

$$0 < \int_0^1 [\varphi(u) - \bar{\varphi}]^2 \,\mathrm{d}u < \infty \quad \textit{where} \quad \bar{\varphi} = \int_0^1 \varphi(u) \,\mathrm{d}u \tag{56}$$

and that

$$\sum_{k=1}^{N} [C_k(p) - \bar{C}(p)]^2 / \max_k [C_k(p) - \bar{C}(p)]^2 \to \infty \tag{57}$$

with $\bar{C}(p) = \sum_{k=1}^{N} C_k(p)/N$.

Then the test statistic $T'_{Np}(R)$ *given by* (54) *is under* H_0 *asymptotically normal* $N(\eta_{cp}, \sigma_{cp})$ *where*

$$\eta_{cp} = \sum_{k=1}^{N} C_k(p) \sum_{i=1}^{N} a_N(i)/N \doteq \sum_{k=1}^{N} C_k(p) \int_0^1 \varphi(u)\,\mathrm{d}u\,, \tag{58}$$

$$\sigma_{cp}^2 = \sum_{k=1}^{N} [C_k(p) - \bar{C}(p)]^2 \int_0^1 (\varphi(u) - \bar{\varphi})^2\,\mathrm{d}u\,. \tag{59}$$

This Theorem follows from Theorem V.1.5.a in [3].

Corollary 5. *Assume that* $\varphi(u)$ *is square integrable and* (56) *holds and that*

$$\sum_{k=1}^{N} (C_{mk} - \bar{C}_m)(C_{nk} - \bar{C}_n) \to b_{mn} \tag{60}$$

for all $m, n = 1, 2, \ldots, s$ *with* s *fixed, not depending on* N,

$$\max_{1 \leqq k \leqq N} (C_{mk} - \bar{C}_m) \to 0 \quad \textit{for all} \quad m = 1, \ldots, s \tag{61}$$

where $\bar{C}_m = \sum_{k=1}^{N} C_{mk}/N$.

Then the statistic $T'_{Np}(R)$ *is under* H_0 *asymptotically normal* $N(\eta_{cp}, \sigma_{cp})$ *for any* $p = (p_1, \ldots, p_s)$ *which are arbitrary real numbers.*

Proof. First we suppose that $\sum_m \sum_n p_m p_n b_{mn} > 0$, then (57) is fulfilled since

$$\sum_{k=1}^{N} [C_k(p) - \bar{C}(p)]^2 / \max_k [C_k(p) - \bar{C}(p)]^2 \geqq$$

$$\geqq \sum_m \sum_n (C_{mk} - \bar{C}_m)(C_{nk} - \bar{C}_n)\, p_m p_n / \max_{k,m} (C_{mk} - \bar{C}_m)^2 \sim$$

$$\sim \sum_m \sum_n p_m p_n b_{mn} / \max_{k,m} (C_{mk} - \bar{C}_m)^2 \to \infty\,.$$

Consequently, the asymptotic normality of $T'_{Np}(R)$ in this case follows from Theorem 4.

Suppose now that $\sum_m \sum_n p_m p_n b_{mn} = 0$, then according to Theorems II.3.1.c and

II.4.3 in [3] we obtain:

$$\operatorname{var}(T'_{Np}) = (N-1)^{-1}\sum_{i=1}^{N}(a_N(i)-\bar{a}_N)^2\sum_{j=1}^{N}[C_j(p)-\bar{C}(p)]^2 \leqq$$

$$\leqq (N/(N-1))\,N^{-1}\sum_{i=1}^{N}a_N^2(i)\sum_{j=1}^{N}[C_j(p)-\bar{C}(p)]^2 \to$$

$$\to \sum_m\sum_n p_m p_n b_{mn}\int_0^1 \varphi(u)\,\mathrm{d}u = 0$$

with $\bar{a}_N = \sum_i a_N(i)/N$. Consequently, $T'_{Np}(R)$ has the asymptotically degenerate normal distribution.

Remark 1. Theorem 4 and Corollary 5 remain true for the test statistic $T_{Np}^{(**)}(R^+)$ given by (45) provided H_0 is replaced by H_0^* since R^+ is under H_0^* uniformly distributed.

6. Asymptotic distribution of the test statistic under contiguous alternatives

Consider a sequence $\{p_\nu, q_\nu\}$ of simple alternatives q_ν's and simple hypotheses p_ν's defined on measure spaces $\{\mathscr{X}_\nu, \mathscr{A}_\nu\}$ respectively.

Definition 1. *We say that the sequence of densities $\{q_\nu\}$ is contiguous to $\{p_\nu\}$ if for any sequence of events $\{A_\nu\}$ $(A_\nu \in \mathscr{A}_\nu)$, $P_\nu\{A\} \to 0$ implies $Q_\nu\{A\} \to 0$ where P_ν and Q_ν are the probability measures corresponding to p_ν, q_ν, respectively. If H_ν and K_ν are simple or composite hypotheses and alternatives, we say that the sequence $\{K_\nu\}$ is contiguous to $\{H_\nu\}$ if for each ν there exists a $p_\nu \in H_\nu$ and $q_\nu \in K_\nu$ such that q_ν is contiguous to p_ν.*

Definition 2. *We say that the density $g(x, \theta)$ has the property U if for every θ there exists a transformation $X \to U_\theta(X)$ such that if X has the density $g(x, \theta)$ then $U_\theta(X)$ has the density $g(x, 0)$ and vice versa*; we denote this briefly by

$$[X \to g(x, \theta)] \Leftrightarrow [U_\theta(X) \to g(x, 0)] \, .$$

Moreover, suppose that U_θ has the following properties:

1) $U_0(X) = X$;
2) *$U_\theta(x)$ is a strictly increasing function of x for each θ;*
3) *For every θ and h there exists a function $V_\theta(h)$ such that*

$$U_{\theta+h}(x) = U_{V_\theta(h)}[U_\theta(x)] \quad \textit{with} \quad V_\theta(0) = 0 \quad \textit{for all} \quad \theta \, .$$

Theorem 5. *Consider the alternative K_ν defined by (49) with $N = N_\nu$, $d_k = d_{k\nu}$ and consider the test statistic $T'_{Np}(R)$ with the scores satisfying (55).*

Suppose that the conditions of Theorem 4 are fulfilled and that the density $g(x, \theta)$ occurring in K_ν satisfies the conditions (A_1), (A_2) of Theorem 1 and

$$0 < I(g) = \int_0^1 \varphi^2(u, g)\,du = \int_{-\infty}^{\infty} [\dot{g}(x, 0)/g(x, 0)]^2\, g(x, 0)\,dx < \infty \tag{62}$$

with $\varphi(u, g)$ given by (27) where g, G play the role of f, F.

Further, assume that one of the following conditions is satisfied:

(i) $$I(g) \sum_{k=1}^{N} d_k \to b^2 > 0,\ \max_k d_k^2 \to 0\,;$$

(ii) $$g(x, \theta) \quad \textit{has the property } U \textit{ and}$$

$$I(g) \sum_{k=1}^{N} V_{\bar{d}}(d_k - \bar{d}) \to b^{*2} > 0,\ \max_k V_{\bar{d}}(d_k - \bar{d}) \to 0\,. \tag{63}$$

Then the statistic $T'_{Np}(R)$ given by (54) is, under K_ν, asymptotically normal $N(\mu^{(i)}_{dcp}, \sigma_{cp})$ under the condition (i) and $N(\mu^{(ii)}_{dcp}, \sigma_{cp})$ under the condition (ii), where

$$\mu^{(i)}_{dcp} = \mu_{cp} + \sum_{k=1}^{N} [C_k(p) - \bar{C}(p)]\, d_k \int_0^1 \varphi(u)\, \varphi(u, g)\,du\,, \tag{64}$$

$$\mu^{(ii)}_{dcp} = \mu_{cp} + \sum_{k=1}^{N} [C_k(p) - C(p)]\, V_{\bar{d}}(d_k - \bar{d}) \int_0^1 \varphi(u)\, \varphi(u, g)\,du \tag{65}$$

with μ_{cp}, σ_{cp} given by (58), (59), $\bar{d} = N^{-1} \sum d_k^0$.

Proof. First we shall show that the assertion about the asymptotic normality of $T'_{Np}(R)$ under the condition (ii) holds provided it holds under the condition (i).

As a matter of fact, the distribution of $T'_{Np}(R)$ does not change if we carry out the transformations $X_k \to U(X_k)$ for all $k = 1, \ldots, N$, where $U(x)$ is a strictly increasing function, namely $U(x) = U_{\bar{d}}(x)$.

We have, by the property U,

$$[X_k \to g(x, d_k)] \Leftrightarrow [U_{\bar{d}}(X_k) \to g(x, V_{\bar{d}}(d_k - \bar{d}))]$$

since

$$U_{d_k}(X_k) = U_{\bar{d} + (d_k - \bar{d})}(X_k) = U_{V_{\bar{d}}(d_k - \bar{d})}[U_{\bar{d}}(X_k)]$$

and

$$[X_k \to g(x, d_k)] \Leftrightarrow [U_{d_k}(X_k) \to g(x, 0)]$$

imply

$$[U_{V_{\bar{d}}(d_k - \bar{d})}[U_{\bar{d}}(X_k)] \to g(x, 0)] \Leftrightarrow [U_{\bar{d}}(X_k) \to g(x, V_{\bar{d}}(d_k - \bar{d}))]\,.$$

Consequently, we can suppose without any loss of generality that X_k has the density $g(x, d_k')$ with $d_k' = V_{\bar{d}}(d_k - \bar{d})$ for all $k = 1, \ldots, N$. It follows from (ii) that

$$I(g) \sum_{k=1}^{N} d_k' \to b^{*2} > 0 \,, \max_k \, (d_k') \to 0 \,.$$

Thus the condition (ii) reduces to the condition (i).

Let us now prove the assertion of this theorem under (i). We need the following propositions:

Proposition 1. *Denote the expectation with respect to the density* $p(x) = \prod_{i=1}^{N} g(x_i, 0)$ *by* E_0 *and put*

$$W_d = 2 \sum_{k=1}^{N} \{[g(X_k, d_k)/g(X_k, 0)]^{1/2} - 1\} \,,$$

$$T_d = - \sum_{k=1}^{N} d_k \dot{g}(X_k, 0)/g(X_k, 0) \,.$$

Then under the condition (i) *we have*

$$E_0 W_d \to - b^2/4 \,, \tag{66}$$

$$\operatorname{var}(W_d - T_d) \to 0 \,. \tag{67}$$

Proof. We omit it since it is carried out quite similarly as the proof of Lemma VI.2.1.a, b in [3].

Proposition 2. *Assume that the condition* (i) *is satisfied, then* $\log L_d - T_d + b^2/2$ *with* $L_d = \sum_{k=1}^{N} \log [g(X_k, d_k)/g(X_k, 0)]$ *converges, under* $p(x)$, *in probability to zero. Furthermore,* $\log L_d$ *is, under* $p(x)$, *asymptotically normal* $N(-b^2/2, b)$ *and* $q_\nu(x) = \sum_{i=1}^{N_\gamma} g(x_i, d_{i\nu})$ *is contiguous to* $p(x) = p_\nu(x) = \prod_{i=1}^{N_\nu} g(x_i, 0)$.

Proof. According to Theorem V.1.2 in [3] and (i), T_d is asymptotically normal $N(0, b)$, since by the assumption that $g(x, \theta)$ satisfies the conditions (A_1), (A_2), $\int_{-\infty}^{\infty} \dot{g}(x, 0)\, dx = 0$. This together with (66), (67) implies that W_d is, under $p(x)$, asymptotically normal $N(-b^2/4, b)$. By LeCam's second Lemma (see VI.1.3 in [3]) Proposition 2 is proved since the condition (i) entails (4) of Section VI.1.3. in [3].

Proposition 3. *Assume that the condition* (i) *and the conditions of Theorem* 4 *are fulfilled. Then* $(T'_{Np}(R), \log L_d)$ is, *under* $p(x)$, *asymptotically jointly normal*

$$N(\mu_1, \mu_2, \sigma_1^2, \sigma_2^2, \sigma_{12}) = N(\mu_{cp}, - b^2/2, \sigma_{cp}^2, b^2, \mu_{dcp}^{(i)} - \mu_{cp}) \,.$$

Proof. We shall denote $X_\nu \sim Y_\nu$ if $(X_\nu - Y_\nu)(\operatorname{var}(Y_\nu))^{-1/2} \to 0$ in probability under $p(x)$ as $\nu \to \infty$. Let

$$S^\varphi = \sum_{k=1}^N C_k(p)\, a_N^\varphi(R_k) \quad \text{where} \quad a_N^\varphi(i) = E\varphi(U^{(i)})$$

for $i = 1, 2, \ldots, N$. Note that

$$E_0 S^\varphi = \bar{C}(p) \sum_{i=1}^N a_N^\varphi(i) \doteq N\bar{C}(p) \int_0^1 \varphi(u)\, du = \mu_{cp} \doteq E_0\, T'_{Np}(R)\,.$$

It is easy to see that

$$T'_{Np}(R) - \mu_{cp} \sim S^\varphi - E_0 S^\varphi = \sum_{k=1}^N [C_k(p) - \bar{C}(p)]\, a_N^\varphi(R_k) \tag{68}$$

(see the proof of Theorem V.1.6.a in [3]) and that

$$S^\varphi - E_0 S^\varphi \sim \sum_{k=1}^N [C_k(p) - \bar{C}(p)]\, \varphi(U_k) = T_c^* \quad \text{(say)} \tag{69}$$

(see the proof of Theorem V.1.5.a in [3]).

On the other hand, it follows from Proposition 1, 2 that

$$\log L_d \sim T_d - b^2/2 \tag{70}$$

where T_d was defined in Proposition 1 or equivalently by

$$T_d = -\sum_{k=1}^N d_k\, \varphi(U_k, g)\,.$$

Consequently, $(T'_{Np}(R) - \mu_{cp}, \log L_d) \sim (T_c^*, T_d - b^2/2)$. Moreover, we can show that $(T_c^*, T_d - b^2/2)$ is asymptotically two-variate normal $N(0, -b^2/2, \sigma_{cp}^2, b^2, \mu_{dcp}^{(i)} - \mu_{cp})$. Q.E.D.

Finally, we observe that the assertion of Theorem 5 under the condition (i) follows from Proposition 3 and LeCam's third Lemma (see Section VI.1.4. in [3]).

Remark 1. Assume that $g(x, \theta) = g(x - \theta)$, i.e. θ is a location parameter. Then $g(x, \theta)$ has the property U with $U_\theta(x) = x - \theta$ since the function is strictly increasing and $U_{\theta+h}(x) = x - (\theta + h) = U_h(U_\theta(x))$, thus $V_\theta(h) = h$ and (63), (65) reduce respectively to

$$I(g) \sum_{k=1}^N (d_k - \bar{d})^2 \to b^{*2} > 0\,, \quad \max_k\, (d_k - \bar{d})^2 \to 0\,, \tag{71}$$

$$\mu_{dcp}^{(ii)} = \sum_{k=1}^N [C_k(p) - \bar{C}(p)]\,(d_k - \bar{d}) \int_0^1 \varphi(u)\, \varphi(u, g)\, du\,. \tag{72}$$

Remark 2. Assume that $g(x, \theta) = \exp(-\theta)\, g(x \exp(-\theta))$, i.e. θ is the scale parameter. Then $g(x, \theta)$ has the property U with $U_\theta(x) = x \exp(-\theta)$ since $U_{\theta+h}(x) =$

$= x \exp(-(\theta + h)) = U_h(U_\theta(x))$, thus $V_\theta(h) = \mathrm{h}$ and (63), (65) reduce to (71), (72), respectively.

Remark 3. Let $g(x, \theta) = [2\pi(1 + \theta)]^{-1/2} \exp(-x^2/2(1 + \theta))$ with $1 + \theta > 0$. Then $g(x, \theta)$ has the property U with $U_\theta x = x/(1 + \theta)$ since $U_{\theta+h}(x) = x/\sqrt{(1 + \theta + h)} = U_{h/(1+\theta)}[U_\theta(x)]$, thus $V_\theta(h) = h/(1 + \theta)$ and (63), (65) reduce respectively to

$$I(g) \sum_{k=1}^{N} (d_k - \bar{d})^2/(1 + \bar{d})^2 \to b^{*2} > 0 \quad (\text{for this density} \quad I(g) = 2), \tag{73}$$

$$\max_k (d_k - \bar{d})^2/(1 + \bar{d})^2 \to 0,$$

$$\mu_{dcp}^{(ii)} = (1 + \bar{d})^{-1} \sum_{k=1}^{N} [C_k(p) - \bar{C}(p)] (d_k - \bar{d}) \int_0^1 \varphi(u) \varphi(u, g) \, du. \tag{74}$$

An analogous remark applies to the exponential density

$$g(x, \theta) = (1 + \theta) \exp(-(1 + \theta) x) \quad \text{for} \quad x > 0,$$
$$= 0 \quad \text{for} \quad x \leqq 0, \quad \text{where} \quad 1 + \theta > 0.$$

Remark 4. Theorem VI.2.4 in [3] may be obtained from Theorem 5 and Remarks 1,2.

Remark 5. Theorem 4.1 of Bhattacharyya and Johnson in [1] is only a special case of Theorem 5.

As a matter of fact, the test statistic given by (24) and the alternative considered by these authors is only a special case of the statistic $T'_{Np}(R)$ and of the alternative K defined by (49) with $g(x, \theta) = g(x - \theta)$ and $d_i = 0$ for $i \leqq m$; $d_i = \theta/N^{1/2}$ for $i \geqq m + 1, 1 \leqq m \leqq N - 1$.

Assume that the conditions (A_1), (A_2) in [1] and $\lim_{N\to\infty} m/N = \lambda$ are fulfilled, then

$$\sum_{i=1}^{N} (d_i - \bar{d})^2 = \theta^2[1 - m/N] \, m/N \to \theta^2 \lambda(1 - \lambda) > 0,$$

$$\max_i (d_i - \bar{d})^2 = (\theta^2/N) \max \{(1 - m/N)^2, \ (m/N)^2\} \to 0.$$

This together with (A_1) in [1] entails (57) and (71). Thus the conditions of Theorem 5 are fulfilled, hence the asymptotic normality of the test statistic (24) under the alternative considered by Bhattacharyya and Johnson follows from Theorem 5.

7. Asymptotic distribution of the signed rank test statistic

In this section we shall show the asymptotic normality of the following statistic:

$$T^+_{Np}(R, \operatorname{sign} X) = \sum_{k=1}^{N} C_k(p) \operatorname{sign} X_k \, a^+_N(R^+_k) \tag{75}$$

with $C_k(p)$ given by (10) and the scores satisfying

$$\int_0^1 [a_N^+(1 + [uN]) - \varphi^+(u)]^2 \, du \to 0 \text{ as } N \to \infty \tag{76}$$

where $\varphi(u)$ is square integrable on $(0, 1)$. The signed rank test statistic given by (35) or (44) is only a special case of the statistic (75) with the scores satisfying (76) (see Theorem V.1.4.b in [3]).

Consider the hypothesis H_0^* defined by (30) and the alternative K_v^* defined by

$$f_1(x) = g(x - d_1), \ldots, f_N(x) = g(x - d_N) \tag{77}$$

with $d_k = d_{kv}$, $N_v = N$ and $g(x) = g(-x)$.

Theorem 6. *Consider the statistic T_{Np}^+ given by (75) with the scores satisfying (76). Assume that*

$$\sum_{k=1}^N C_k^2(p) / \max_k C_k^2(p) \to \infty \, ; \tag{78}$$

then $T_{Np}^+(R, \operatorname{sign} X)$ is, under H_0^*, asymptotically normal $N(0, \sigma_{cp}^+)$ where

$$(\sigma_{cp}^+)^2 = \int_0^1 [\varphi^+(u)]^2 \, du \sum_{k=1}^N C_k^2(p) \, . \tag{79}$$

Proof. Theorem 6 follows from Theorem 1.1 of Hušková [4].

Corollary 6. *Assume that s does not depend on N and that the regression constants C_{mk}'s satisfy*

$$\sum_{k=1}^N C_{mk} C_{nk} \to b'_{mn} \, , \quad \max_k C_{mk}^2 \to 0 \quad \text{for all} \quad m, n = 1, 2, \ldots, s \, . \tag{80}$$

Then the statistic $T_{Np}^+(R, \operatorname{sign} X)$ with scores satisfying (76) is, under H_0^*, asymptotically normal $N(0, \sigma_{cp}^+)$ for any real numbers $p_1, p_2, \ldots, p_s$.

Proof. First suppose that $\sum_{m,n} p_m p_n b'_{mn} > 0$. It is easy to see that (80) entails (78) and the conclusion of this corollary follows from Theorem 6.

The case where $\sum_m \sum_n p_m p_n b'_{mn} = 0$ may be treated similarly as in the proof of Corollary 5.

Theorem 7. *Consider the alternative K_v^* defined by (77). Assume that (78) and*

$$\sum_{i=1}^N d_i^2 \to b_1^2 > 0 \, , \quad \max_i d_i^2 \to 0 \tag{81}$$

hold and that the density $g(x)$ satisfies the conditions (A_1'), (A_2') *of Corollary* 1. *Moreover, if*

$$(82) \qquad 0 < \int_0^1 \varphi^2(u, g)\, du < \infty \quad \textit{with} \quad \varphi(u, g) = -\, g'(G^{-1}(u))/g'(G^{-1}(u))$$

where G denotes the distribution function with respect to g, then the statistic $T^+_{Np}(R^+, \operatorname{sign} X)$ with the scores satisfying (76) *is, under K^*_ν, asymptotically normal $N(\mu^+_{dcp}, \sigma^+_{cp})$ where*

$$(83) \qquad \mu^+_{dcp} = \sum_{k=1}^{N} C_k(p)\, d_k \int_0^1 \varphi^+(u)\, \varphi^+(u, g)\, du$$

and σ^+_{cp} is given by (79), $\varphi^+(u, g) = \varphi(\frac{1}{2} + \frac{1}{2}u, g)$.

Proof. It follows from Proposition 2 that the condition (81) is sufficient for the contiguity of K^*_ν to H_0. On the other hand, under the conditions of Theorem 7 the conditions of Theorem 17 in [5] or Theorem 2.2 in [4] are fulfilled and the assertion of Theorem 7 follows from the cited theorems.

Remark. Theorem 4.3 of Bhattacharyya and Johnson in [1] may be obtained from Theorem 7.

8. Asymptotic efficiency of rank test

We say that an α-level test T^* is based on a statistic T if the critical region of the test assumes the form $\{T > C_\alpha\}$.

Suppose that T^*_1, T^*_2 are based on T_1, T_2, respectively, and that T_1, T_2 are asymptotically normal $N(0, \sigma_1)$, $N(0, \sigma_2)$ under H_0 and $N(\mu_1, \sigma_1)$, $N(\mu_2, \sigma_2)$ under the alternative K^*_ν. Then the asymptotic powers of T^*_1, T^*_2 under K_ν are given by

$$(84) \qquad 1 - \phi(k_{1-\alpha} - \mu_1/\sigma_1)\,, \quad 1 - \phi(k_{1-\alpha} - \mu_2/\sigma_2)\,,$$

respectively, where $k_{1-\alpha}$ is the $100(1-\alpha)$ percentage point of the standardized normal dictribution function $\phi(x)$.

The quantity

$$(85) \qquad e[T_2 : T_1] = [(\mu_2/\sigma_2)/(\mu_1/\sigma_1)]^2 = (\mu_2\sigma_1/\mu_1\sigma_2)^2$$

is called the asymptotic relative efficiency of the test T^*_2 compared to T^*_1. If T^*_1 is asymptotically most powerful with respect to the definition below, then $e[T_2 : T_1] = e[T_2]$ is called the asymptotic efficiency of the test T^*_2. Note that the definition of the relative efficiency is meaningful only as μ_1, μ_2 are positive since if, for example, $\mu_1 < 0$ the test T^*_1 is worse than the test defined by the critical function $\Phi(x) \equiv \alpha$.

Definition. *A test with the probability $\Phi_\nu(x)$ of rejecting the hypothesis is called the asymptotically maximin most powerful for testing H_ν against K_ν at the level α if*

$$\text{(A)} \qquad \limsup_{\nu\to\infty} \left\{ \sup_{P_\nu \in H} \int \Phi_\nu(x)\, \mathrm{d}P_\nu(x) \right\} \leqq \alpha \,,$$

$$\text{(B)} \qquad \lim_{\nu\to\infty} \left[\beta(\alpha, H_\nu, K_\nu) - \inf_{Q_\nu \in K_\nu} \int \Phi_\nu(x)\, \mathrm{d}Q_\nu(x) \right] = 0$$

where

$$\beta(\alpha, H_\nu, K_\nu) = \sup_{\Phi_\nu' \in \Psi_\nu(\alpha)} \inf_{Q_\nu \in K_\nu} \int \Phi_\nu'(x)\, \mathrm{d}Q_\nu(x)$$

with $\Psi_\nu(\alpha)$ being the class of all tests satisfying $\sup_{P_\nu \in H_\nu} \int \Phi_\nu' \, \mathrm{d}P \leqq \alpha$.

For the sake of simplicity, let us delete the subscript ν in what follows, writing for example

$$\sum_{i=1}^{N} d_i^2 \to b^2 \quad \text{for} \quad \lim_{\nu\to\infty} \sum_{i=1}^{N_\nu} d_{i\nu}^2 = b^2 \,.$$

Let $H_0 (H_0 = H_{0\nu})$ be the hypothesis defined by (1) with respect to the sample size N_ν.

Theorem 8. *Consider the problem of testing H_0 against the $K (K = K_\nu)$ defined by* (49). *Assume that the conditions* (i), (ii) *of Theorem 5 are fulfilled. Then the following relations hold*:

$$\text{(86)} \qquad \beta(\alpha, H_0, q) \to 1 - \phi(k_{1-\alpha} - b) \quad \textit{under } \text{(i)} \,,$$

$$\text{(87)} \qquad \beta(\alpha, H_0, q) \to 1 - \phi(k_{1-\alpha} - b^*) \quad \textit{under } \text{(ii)}$$

where $q(x) = \prod_{i=1}^{N} g(x, d_i)$ and b and b^* are defined by the conditions (i), (ii).

The maximum powers (86), (87) *are asymptotically attained by the rank tests based on the following statistics*:

$$\text{(88)} \qquad S = \sum_{i=1}^{N} d_i \, a_N(R_i, g) \,,$$

$$\text{(89)} \qquad S' = \sum_{i=1}^{N} V_{\bar{d}}(d_i - \bar{d}) \, a_N(R_i, g) \,,$$

respectively, where $a_N(i, g)$ are defined by (27), (28) *with f, F replaced by g, G and $V_\theta(h)$ are defined by Definition 2 of Section 6.*

Proof. First suppose that the condition (i) of Theorem 5 is fulfilled. It is clear that

$$\text{(90)} \qquad \beta(\alpha, H_0, q) \leqq \beta(\alpha, p_0, q)$$

where $p_0(x) = \prod_{i=1}^{N} g(x_i, 0)$.

On the other hand, from LeCam's third Lemma (see VI.1.4 in [3]) and from Proposition 2 of Section 6 it follows that $\log q/p_0$ is asymptotically normal $N(-b/2, b)$ under p_0 and $N(b/2, b)$ under q. Consequently, a test based on q/p_0 has the following power:

$$\beta(\alpha, p_0, q) \to 1 - \phi(k_{1-\alpha} - b)\,. \tag{91}$$

On the other hand, this asymptotic power belongs to the test based on S, according to Theorem 5. Consequently

$$\liminf \beta(\alpha, H_0, q) \geqq 1 - \phi(k_{1-\alpha} - b) \tag{92}$$

and (86) follows from (90)–(92).

Suppose now that the condition (ii) of Theorem 5 is fulfilled. Note that $U_\theta(x)$ has an almost everywhere derivative $U'_\theta(x)$ in x and its inverse function $U_\theta^{-1}(x)$ exists for each θ since $U_\theta(x)$ is strictly increasing in x. First, it is clear that

$$\beta(\alpha, H_0, q) \leqq \beta(\alpha, \bar{p}_0, q) \tag{93}$$

where $\bar{p} = \prod_{i=1}^{N} g(x_i), \bar{d})$.

On the other hand, we have, under the condition (ii) of Theorem 5,

$$[X_k \to g(x, d_k)] \Leftrightarrow [U_{\bar{d}}(X_k) \to g(x, V_{\bar{d}}(d_k, \bar{d}))]$$

for $k = 1, \ldots, N$, hence

$$\begin{aligned} P\{U_{\bar{d}}(X_k) < y\} &= P\{X_k < U_{\bar{d}}^{-1}(y)\} \quad \text{entails} \\ g(y, V_{\bar{d}}(d_k - \bar{d})) &= g(U_{\bar{d}}^{-1}(y), d_k)\, [U_{\bar{d}}^{-1}(y)]\,. \end{aligned}$$

Putting $y = U_{\bar{d}}(x)$ we obtain

$$U'_{\bar{d}}(x)\, g(U_{\bar{d}}(x), V_{\bar{d}}(d_k - \bar{d})) = g(x, d_k)\,.$$

Consequently

$$\begin{aligned} \log q(X)/\bar{p}(X) &= \sum_{k=1}^{N} \log\,(g(X_k, d_k)/g(X_k, \bar{d})) = \\ &= \sum_{k=1}^{N} \log g(U_{\bar{d}}(X_k), V_{\bar{d}}(d_k - \bar{d}))/g(U_{\bar{d}}(X_k), 0) \end{aligned}$$

since $V_{\bar{d}}(0) = 0$, $U'_{\bar{d}}(x) > 0$.

Putting $Y_k = U_d(X_k)$, we obtain

$$\log\,(q/\bar{p}) = \sum_{k=1}^{N} \log g(Y_k, d'_k)/g(Y_k, 0)$$

with $d'_k = V_{\bar{d}}(d_k - \bar{d})$. Note that Y_k has the density $g(x, 0)$ under $\bar{p}$ and $g(x, d'_k)$ under q and d'_k satisfy the conditions

$$I(g) \sum_{k=1}^{N} d'^2_k \to b^{*2} > 0\,, \quad \max_k d'^2_k \to 0\,, \quad \text{by (63).}$$

It follows from LeCam's third Lemma in [3] and from Proposition 2 of Section 6 that $\log(q/\bar{p})$ is asymptotically normal $N(-b^{*2}/2, b^*)$ under p and $N(b^{*2}/2, b^*)$ under q. Consequently, the test based on $q/\bar{p}$ has the following power:

$$\beta(\alpha, \bar{p}, q) \to 1 - \phi(k_{1-\alpha} - b^*)\,. \tag{94}$$

On the other hand, this asymptotic power belongs to the test based on S', according to Theorem 5 under the condition (ii), hence

$$\liminf \beta(\alpha, H_0, q) \geqq 1 - \phi(k_{1-\alpha} - b^*)\,. \tag{95}$$

Finally, (87) follows from (93)–(95).

Remark. By the argument of this proof and LeCam's second Lemma in [3], $q(x) = \prod_{i=1}^{N} g(x, d_i)$ is contiguous to H_0 under the condition (ii) of Theorem 5.

Let us now show the asymptotic efficiency of the test based on $T_{Np}(R)$ given by (9) or (29).

According to Theorems 5,8 and the definition of the asymptotic efficiency we have

$$e[T_{Np}(R)] = e[T_{Np}(R) : S] = \varrho^2 \varrho_0^2 \tag{96}$$

under K defined by (49) and under the condition (i),

$$e[T_{Np}(R)] = e[T_{Np}R : S'] = \varrho^2 \bar{\varrho}^2 \tag{97}$$

under K and under the condition (ii), provided

$$\sum_{k=1}^{N} (C_k(p) - \bar{C}(p))\, d_k / \{\sum_{k=1}^{N} [C_k(p) - \bar{C}(p)]^2 \sum_{k=1}^{N} d_k^2\}^{1/2} \to \varrho_0\,, \tag{98}$$

$$\sum_{k=1}^{N} (C_k(p) - \bar{C}(p))\, V_{\bar{d}}(d_k - \bar{d}) / \{\sum_{k=1}^{N} (C_k(p) - \bar{C}(p))^2 \sum_{k=1}^{N} (V_{\bar{d}}(d_k - \bar{d}))^2\}^{1/2} \to \bar{\varrho} \tag{99}$$

and with

$$\varrho = \int_0^1 \varphi(u, f)\, \varphi(u, g)\, \mathrm{d}u (I(f)\, I(g))^{-1/2}\,. \tag{100}$$

It is of interest to study the sensitivity of the asymptotic relative efficiency of the tests on $T_{Np}(R)$ corresponding to different choices of the weights.

Let us restrict ourselves to the cases where the alternative K defined by (49) satisfies the condition (i) of Theorem 5 or the parameter θ involved in the density $g(x, \theta)$ under K is a location or a scale parameter and (74) holds.

Denote by $T_u(R)$ the special form of $T_{Np}(R)$ with respect to the uniform weights, by $T_{dm}(R)$ the special form of $T_{Np}(R)$ with respect to the weights degenerate at m, i.e. $p_m = 1$, $p_i = 0$ for $i \neq m$, then we have

$$(101) \qquad T_u(R) = \sum_{k=1}^{N} (\bar{C}_{\cdot k} - \bar{\bar{C}})\, a_N(R_k, f)\,,$$

$$(102) \qquad T_{dm}(R) = \sum_{k=1}^{N} (C_{mk} - \bar{C}_{m\cdot})\, a_N(R_k, f)$$

where

$$\bar{C}_{\cdot k} = \sum_{i=1}^{s} C_{ik}/s, \quad \bar{C}_{m\cdot} = \sum_{k=1}^{N} C_{mk}/N\,,$$

$$\bar{\bar{C}} = \sum_{k=1}^{N} \bar{C}_{\cdot k}/N = \sum_{m=1}^{s} \bar{C}_{m\cdot}/s\,.$$

Assume that

$$(103) \qquad \{\max_k (\bar{C}_{\cdot k} - \bar{\bar{C}})^2\}^{-1} \sum_{k=1}^{N} (\bar{C}_{\cdot k} - \bar{\bar{C}})^2 \to \infty\,,$$

$$(104) \qquad \{\max_k (C_{mk} - \bar{C}_{m\cdot})^2\}^{-1} \sum_{k=1}^{N} (C_{mk} - \bar{C}_{m\cdot})^2 \to \infty\,.$$

Then, according to Theorem 4,5, the statistics $T_u(R)$ and $T_{dm}(R)$ are asymptotically normal both under H_0 and under K and we obtain the asymptotic relative efficiency under K

$$(105) \qquad e[T_u : T_{dm}] \doteq [\sum_{k=1}^{N} (\bar{C}_{\cdot k} - \bar{\bar{C}})(d_k - \bar{d})]^2 \sum_{k=1}^{N} (C_{mk} - \bar{C}_{m\cdot})^2\,.$$
$$\cdot \{[\sum_{k=1}^{N} (\bar{C}_{\cdot k} - \bar{C}_{m\cdot})(d_k - \bar{d})]^2 \sum_{k=1}^{N} (\bar{C}_{\cdot k} - \bar{\bar{C}})^2\}^{-1}$$

with the convention that $\bar{d} = 0$ if the condition (i) is fulfilled.

If $d_k = \Delta C_{mk}$ for all k, i.e. K coincides with alternative K_m defined by (2) and $p_m = 1$, $p_i = 0$ for $i \neq m$ are the correct degenerate weights, then we obtain

$$(106) \qquad e(T_u : T_{dm}) = \{\sum_{k=1}^{N} (\bar{C}_{\cdot k} - \bar{\bar{C}})(C_{mk} - \bar{C}_{m\cdot})\}^2 \{\sum_{k=1}^{N} (C_{mk} - \bar{C}_{m\cdot})^2\,.$$
$$\cdot \sum_{k=1}^{N} (\bar{C}_{\cdot k} - \bar{\bar{C}})^2\}^{-1} \leqq 1\,.$$

Especially, putting $C_{mj} = 0, 1$ if $j \leqq m, j \geqq m + 1$, respectively, i.e. K is the alternative of shift occuring at m, and $m/N \to \lambda$ $(0 < \lambda < 1)$, then (106) becomes

$$(107) \qquad e[T_u : T_{dm}] = 3\lambda(1 - \lambda) \leqq \tfrac{3}{4}\,.$$

This was shown in [1].

Acknowledgement. The author wishes to express his indebtedness to Dr. Z. Šidák under whose guidance the thesis was written.

References

[1] *G. K. Bhattacharyya, R. A. Johnson:* Nonparametric tests for shift at unknown time point. Annals of Math. Stat. 39 (1968) No. 5, 1731—1743.

[2] *H. Chernoff, S. Zacks:* Estimating the current mean of a normal distribution which is subjected to changes in time. Annals of Math. Stat. 35 (1964), 999—1018.

[3] *J. Hájek, Z. Šidák:* Theory of rank tests. Academia, Publishing house of the Czechoslovak Academy of Sciences, Praha 1967.

[4] *M. Hušková:* Asymptotic distribution of simple linear rank statistics used for testing symmetry hypotheses. (Czech.) Thesis, Prague 1968.

[5] *M. Hušková:* Asymptotic distribution of simple linear rank statistic for testing of symmetry. Z. Wahrscheinlichkeitstheorie. Geb. 14, (1970), 308—322.

[6] *Z. Kander, S. Zacks:* Test procedure for possible changes in parameters of statistical distribution occurring at unknown time point. Annals of Math. Stat. 37(1966), 1196—1210.

[7] *E. L. Lehmann:* Testing statistical hypotheses. J. Wiley, New York, 1959.

[8] *E. L. Lehmann:* Some concepts of independence. Annals of Math. Stat. 37 (1966) No. 2, 1137—1153.

[9] *E. S. Page:* Continuous inspection schemes. Biometrika 41 (1954), 100—116.

[10] *E. S. Page:* A test for a change in parameter occurring at an unknown point. Biometrika 42 (1955), 523—526.

Souhrn

POŘADOVÉ TESTY HYPOTÉZY NÁHODNOSTI PROTI SKUPINĚ REGRESNÍCH ALTERNATIV

NGUYEN-VAN-HUU

V článku se studuje problém testování hypotézy náhodnosti proti skupině regresních alternativ v neznámém parametru. Pro tento problém je navržen pořadový test. Jde o zobecnění problému testování posunutí v parametru lokace, které se objevuje v neznámém časovém bodě v řadě postupně pozorovaných veličin. Pro tento poslední problém pořadový test byl nalezen Bhattacharyyou a Johnsonem (1968). Pořadový test navržený v naší práci je lokálně průměrově nejmohutnější ve třídě všech možných pořadových testů ve smyslu definice v § 3. Dále je studována asymptotická normalita statistiky našeho pořadového testu a jeho asymptotická vydatnost nejen pro případ parametru lokace a škály, ale i pro případ obecného parametru.

Author's address: Dr. *Nguyen-van-Huu,* Can bo giang day Khoa Toan, truong Dai hoc Tong Hop, Ha-Noi, Viet-Nam.