KYBERNETIKA — VOLUME 22 (1986), NUMBER 6

INFLUENCE OF CONTAMINATION LEVEL DEVIATIONS ON THE TEST ERROR PROBABILITIES

JAN ÁMOS VÍŠEK

Limits of the ratio of the worst possible deviations of the most powerful test error probabilities with respect to the deviations of the contamination level parameters in Rieder's model of contaminacy are given for the finite sample size under weak regularity conditions on the deviations of contamination parameters. Asymptotic approximation of these limits in the local alternative set-up is presented, too.

1. INTRODUCTION

Recently a few attempts were performed to modify Hampel's influence curve idea for the hypotheses testing purposes (see [1], [2], [6] and [8]). Derived results may serve as characteristics of the sensitivity of the test risk with respect to the "throwing in" an extra datum into the sample, but they did not give, without a more detailed analysis, any hint at the solution of the problem of the test risk behaviour under small deviations of the contamination level parameters. This problem was for the first time attacked in [9] for Huber's model of contaminacy. In present paper Rieder's model of contaminacy is considered. Let us explain the problem in more details.

Having estimated contamination level parameters of given model of contaminacy one may find a least favourable pair of distributions, if any, and construct a test based on its likelihood ratio (i.e. one establishes the most powerful test). Now it is of interest to know what are the worst possible deviations in the error probabilities of this test in the case that the estimation of the contamination level parameters was wrong. This problem was solved by an infinitesimal approach which enabled to derive formulae for the size and power dependency, defined as the derivatives of the size and power of the (fixed, above mentioned) test with respect to the contamination parameters. Then asymptotic approximation of these formulae in the local alternative setting was found (for more extended discussion of the results of this type the reader is recommended to consult the last section of [8]).

2. NOTATIONS

Let $\mathbb{N}$ denote the set of all positive integers and $\mathbb{R}$ the real line being endowed by the Borel σ-algebra $\mathscr{B}$. Let $(\mathscr{X}, \mathscr{A})$ be a measurable space and $\mathscr{M}$ the class of all probability measures on it. Let for every $n \in \mathbb{N}$ and $i = 0, 1$ $P_{in} \in \mathscr{M}$, $P_{0n} \neq P_{1n}$ and let p_{in} denote a density of P_{in} with respect to a σ-finite measure μ. Let us recall Rieder's model of contaminacy. For $\varepsilon_{in} \geqq 0$, $\delta_{in} \geqq 0$, $0 < \varepsilon_{in} + \delta_{in} < 1$ put

$$\mathscr{P}_{in}(\varepsilon_{in}, \delta_{in}) = \{Q \in \mathscr{M}: Q(B) \geqq (1 - \varepsilon_{in}) P_{in}(B) - \delta_{in} \text{ for all } B \in \mathscr{A}\}.$$

Definition 1. Let $Q_{in} \in \mathscr{P}_{in}(\varepsilon_{in}, \delta_{in})$ be such that for $\pi_n \in \mathrm{d}Q_{0n}/\mathrm{d}Q_{1n}$ we have

$$Q_{0n}(\pi_n > t) = \sup\{Q'(\pi_n > t): Q' \in \mathscr{P}_{0n}(\varepsilon_{0n}, \delta_{0n})\}$$

and

$$Q_{1n}(\pi_n > t) = \inf\{Q''(\pi_n > t): Q'' \in \mathscr{P}_{1n}(\varepsilon_{1n}, \delta_{1n})\}$$

for all $t \in (0, \infty)$. Then (Q_{0n}, Q_{1n}), if any, is called the least favourable pair (LFP) for $(\mathscr{P}_{0n}(\varepsilon_{0n}, \delta_{0n}), \mathscr{P}_{1n}(\varepsilon_{1n}, \delta_{1n}))$.

One of possible solutions of the LFP problem for the above recalled model of contaminacy may be described by the densities q_{0n} and q_{1n} (with respect to μ) of Q_{0n} and Q_{1n}, respectively (see [4]):

Let $\Delta_n \in \mathrm{d}P_{1n}/\mathrm{d}P_{0n}$ and Δ_{0n} and Δ_{1n} be solutions of equations

$$(1) \qquad \Delta_{0n}P_{0n}(\Delta_n < \Delta_{0n}) - P_{1n}(\Delta_n < \Delta_{0n}) = \nu_{1n} + \omega_{0n}\Delta_{0n}$$

and

$$(2) \qquad P_{1n}(\Delta_{1n} < \Delta_n) - \Delta_{1n}P_{0n}(\Delta_{1n} < \Delta_n) = \nu_{0n}\Delta_{1n} + \omega_{1n},$$

where $\nu_{in} = (1 - \varepsilon_{in})^{-1}(\varepsilon_{in} + \delta_{in})$ and $\omega_{in} = \delta_{in}(1 - \varepsilon_{in})^{-1}$. Then

$$\frac{q_{1n}}{q_{0n}} = \frac{1 - \varepsilon_{1n}}{1 - \varepsilon_{0n}} \max\{\Delta_{0n}, \min\{\Delta_n, \Delta_{1n}\}\} \quad Q_{0n} + Q_{1n}\ a.e.$$

and

$$q_{0n} = \begin{cases} \dfrac{1 - \varepsilon_{0n}}{\nu_{1n} + \omega_{0n}\Delta_{0n}} (\nu_{1n}p_{0n} + \omega_{0n}p_{1n}) & \text{on } \{\Delta_n < \Delta_{0n}\}, \\ (1 - \varepsilon_{0n})\, p_{0n} & \text{on } \{\Delta_{0n} \leqq \Delta_n \leqq \Delta_{1n}\}, \\ \dfrac{1 - \varepsilon_{0n}}{\nu_{0n}\Delta_{1n} + \omega_{1n}} (\omega_{1n}p_{0n} + \nu_{0n}p_{1n}) & \text{on } \{\Delta_{1n} < \Delta_n\}. \end{cases}$$

To avoid the degenerated case $Q_{0n} = Q_{1n}$ a.e. let us assume hereafter $\Delta_{0n} < \Delta_{1n}$. Hence the most powerful test for testing $H_n: P \in \mathscr{P}_{0n}(\varepsilon_{0n}, \delta_{0n})$ against $A_n: P \in \mathscr{P}_{1n}(\varepsilon_{1n}, \delta_{1n})$ for a sample of size m may be based on the statistics

$$T_{m,n}(x) = \sum_{i=1}^{m} IC_n(x_i),$$

where $x \in \mathscr{X}^\infty$ and $IC_n(x) = \log\{q_{1n}(x)/q_{0n}(x)\}$. Let us call a test Φ_n based on $T_{\cdot,n}$

LFP$(\varepsilon_{0n}, \varepsilon_{1n}, \delta_{0n}, \delta_{1n})$-test. Moreover it follows that

$$
\begin{aligned}
(3) \qquad & Q_{0n}(\Delta_n < \Delta_{0n}) = (1 - \varepsilon_{0n}) P_{0n}(\Delta_n < \Delta_{0n}) - \delta_{0n}, \\
& Q_{0n}(\Delta_{1n} < \Delta_n) = (1 - \varepsilon_{0n}) P_{0n}(\Delta_{1n} < \Delta_n) + \varepsilon_{0n} + \delta_{0n}, \\
& Q_{1n}(\Delta_n < \Delta_{0n}) = (1 - \varepsilon_{1n}) P_{1n}(\Delta_n < \Delta_{0n}) + \varepsilon_{1n} + \delta_{1n},
\end{aligned}
$$

and

$$Q_{1n}(\Delta_{1n} < \Delta_n) = (1 - \varepsilon_{1n}) P_{1n}(\Delta_{1n} < \Delta_n) - \delta_{1n}$$

(see (5.10), (5.14), (5.15) and (5.16) in [4]).

3. SENSITIVITY OF LFP-TEST WITH RESPECT TO PARAMETERS OF MODEL OF CONTAMINACY

Let us firstly explain the problem attacked in this section. Having prescribed contamination level to be given by the quadruple $(\varepsilon_{0n}, \varepsilon_{1n}, \delta_{0n}, \delta_{1n})$ we can find LFP $(\varepsilon_{0n}, \varepsilon_{1n}, \delta_{0n}, \delta_{1n})$-test then we would like to study the worst possible behaviour of the LFP $(\varepsilon_{0n}, \varepsilon_{1n}, \delta_{0n}, \delta_{1n})$-test under another quadruple $(\tilde{\varepsilon}_{0n}, \tilde{\varepsilon}_{1n}, \tilde{\delta}_{0n}, \tilde{\delta}_{1n})$. To cope with the problem we may try to find LFP of distributions for

$$\text{LFP } (\varepsilon_{0n}, \varepsilon_{1n}, \delta_{0n}, \delta_{1n})\text{-test with respect to } (\mathscr{P}_{0n}(\tilde{\varepsilon}_{0n}, \tilde{\delta}_{0n}), \mathscr{P}_{1n}(\tilde{\varepsilon}_{1n}, \tilde{\delta}_{1n}))$$

(definition see below) and then use an infinitesimal approach to establish a characteristics of sensitivity of the test risk with respect to changes of contamination level. Now, let us give the promised definition.

Definition 2. Let $Q_{in}^{IC} \in \mathscr{P}_{in}(\tilde{\varepsilon}_{in}, \tilde{\delta}_{in})$, $\tilde{\varepsilon}_{in} \geqq 0$, $\tilde{\delta}_{in} \geqq 0$, $0 < \tilde{\varepsilon}_{in} + \tilde{\delta}_{in} < 1$, be such that for every $t \in \mathbb{R}$ we have

$$Q_{0n}^{IC}(IC_n(x) > t) = \sup \{Q'(IC_n(x) > t) : Q' \in \mathscr{P}_{0n}(\tilde{\varepsilon}_{0n}, \tilde{\delta}_{0n})\}$$

and

$$Q_{1n}^{IC}(IC_n(x) > t) = \inf \{Q''(IC_n(x) > t) : Q'' \in \mathscr{P}_{1n}(\tilde{\varepsilon}_{1n}, \tilde{\delta}_{1n})\} .$$

Then $(Q_{0n}^{IC}, Q_{1n}^{IC})$, if any, is called the least favourable pair for $IC_n(x)$ (LFP$_{IC}$) with respect to $(\mathscr{P}_{0n}(\tilde{\varepsilon}_{0n}, \tilde{\delta}_{0n}), \mathscr{P}_{1n}(\tilde{\varepsilon}_{1n}, \tilde{\delta}_{1n}))$.

Remark 1. In the Definition 1 we have required $\varepsilon_{in} \geqq 0$, $\delta_{in} \geqq 0$, $0 < \varepsilon_{in} + \delta_{in} < 1$ which implies $\nu_{in} > 0$ and $\omega_{in} \geqq 0$. Since $\Delta_{0n} \geqq 0$ by (1) it implies

$$(4) \qquad P_{0n}(\Delta_n < \Delta_{0n}) > 0$$

and hence $\Delta_{0n} > 0$. Rewriting (2) as

$$(1 - \varepsilon_{1n}) P_{1n}(\Delta_{1n} < \Delta_n) - \delta_{1n} = \Delta_{1n}(1 - \varepsilon_{1n}) \{P_{0n}(\Delta_{1n} < \Delta_n) + \nu_{0n}\}$$

and taking into account $\Delta_{1n} > \Delta_{0n} > 0$ one obtains

$$(5) \qquad (1 - \varepsilon_{1n}) P_{1n}(\Delta_{1n} < \Delta_n) - \delta_{1n} > 0 .$$

On the other hand

$$1 \geqq (1 - \varepsilon_{1n}) \{P_{1n}(\Delta_n \leqq \Delta_{0n}) + P_{1n}(\Delta_{1n} \leqq \Delta_n)\} + \varepsilon_{1n}$$

and hence (see (5))

$$1 > 1 - \{(1 - \varepsilon_{1n}) P_{1n}(\Delta_{1n} \leqq \Delta_n) - \delta_{1n}\} \geqq (1 - \varepsilon_{1n}) P_{1n}(\Delta_n \leqq \Delta_{0n}) + \varepsilon_{1n} + \delta_{1n} .$$

But then there is a $\tilde{\delta}_{1n} > \delta_{1n}$ such that

$$(1 - \varepsilon_{1n}) P_{1n}(\Delta_n \leqq \Delta_{0n}) + \varepsilon_{1n} + \tilde{\delta}_{1n} < 1 ,$$

i.e.

$$\varepsilon_{1n}(1 - P_{1n}(\Delta_n \leqq \Delta_{0n})) < 1 - P_{1n}(\Delta_n \leqq \Delta_{0n}) - \tilde{\delta}_{1n}$$

and therefore there is a $\tilde{\varepsilon}_{1n} > \varepsilon_{1n}$ such that

$$\tilde{\varepsilon}_{1n}(1 - P_{1n}(\Delta_n \leqq \Delta_{0n})) < 1 - P_{1n}(\Delta_n \leqq \Delta_{0n}) - \tilde{\delta}_{1n} ,$$

i.e.

(6)
$$\tilde{\varepsilon}_{1n} + \tilde{\delta}_{1n} < 1 - (1 - \tilde{\varepsilon}_{1n}) P_{1n}(\Delta_n \leqq \Delta_{0n}) .$$

Likewise it is possible to arrive at

(7)
$$P_{1n}(\Delta_{1n} < \Delta_n) > 0$$

and find $\tilde{\varepsilon}_{0n} > \varepsilon_{0n}$ and $\tilde{\delta}_{0n} > \delta_{0n}$ such that

(8)
$$(1 - \tilde{\varepsilon}_{0n}) P_{0n}(\Delta_{1n} \leqq \Delta_n) + \tilde{\varepsilon}_{0n} + \tilde{\delta}_{0n} < 1 .$$

The following lemma gives the form of the least favourable pair under the conditions entitled by (5)–(8).

Lemma 1. Let $\varepsilon_{in} \leqq \tilde{\varepsilon}_{in}, \delta_{in} \leqq \tilde{\delta}_{in}, \tilde{\varepsilon}_{0n} + \tilde{\delta}_{0n} < 1 - (1 - \tilde{\varepsilon}_{0n}) P_{0n}(\Delta_{1n} < \Delta_n)$, $\tilde{\varepsilon}_{1n} + \tilde{\delta}_{1n} < (1 - \tilde{\varepsilon}_{1n}) P_{1n}(\Delta_n < \Delta_{0n}), \tilde{\delta}_{0n} < (1 - \tilde{\varepsilon}_{0n}) P_{0n}(\Delta_n < \Delta_{0n})$ and $\tilde{\delta}_{1n} < (1 - \tilde{\varepsilon}_{1n}) P_{1n}(\Delta_{1n} < \Delta_n)$. For any $A, B \in \mathscr{A}$ put

$$w_A(B) = 1 \quad \text{for} \quad A \cap B \neq \emptyset , \qquad = 0 \quad \text{otherwise} .$$

Define for $B \in \mathscr{A}$

$$Q_{0n}^{IC}(B) = \{(1 - \tilde{\varepsilon}_{0n}) P_{0n}(\Delta_n \leqq \Delta_{0n}) - \tilde{\delta}_{0n}\} w_{\{\Delta_n \leqq \Delta_{0n}\}}(B) + (1 - \tilde{\varepsilon}_{0n}) P_{0n}(B \cap \{\Delta_{0n} < \Delta_n < \Delta_{1n}\}) + \{(1 - \tilde{\varepsilon}_{0n}) P_{0n}(\Delta_{1n} \leqq \Delta_n) + \tilde{\varepsilon}_{0n} + \tilde{\delta}_{0n}\} w_{\{\Delta_{1n} \leqq \Delta_n\}}(B)$$

and

$$Q_{1n}^{IC}(B) = \{(1 - \tilde{\varepsilon}_{1n}) P_{1n}(\Delta_n \leqq \Delta_{0n}) + \tilde{\varepsilon}_{1n} + \tilde{\delta}_{1n}\} w_{\{\Delta_n \leqq \Delta_{0n}\}}(B) + (1 - \tilde{\varepsilon}_{1n}) P_{1n}(B \cap \{\Delta_{0n} < \Delta_n < \Delta_{1n}\}) + \{(1 - \tilde{\varepsilon}_{1n}) P_{1n}(\Delta_{1n} \leqq \Delta_n) - \tilde{\delta}_{1n}\} w_{\{\Delta_{1n} \leqq \Delta_n\}}(B) .$$

Then $(Q_{0n}^{IC}, Q_{1n}^{IC})$ is the LFP$_{IC}$ with respect to $(\mathscr{P}_{0n}(\tilde{\varepsilon}_{0n}, \tilde{\delta}_{0n}), \mathscr{P}_{1n}(\tilde{\varepsilon}_{1n}, \tilde{\delta}_{1n}))$.

Proof. It is evident that $Q_{in}^{IC} \in \mathscr{P}_{in}(\tilde{\varepsilon}_{in}, \tilde{\delta}_{in})$. Let us denote

$$r_{in} = \log \left\{ \frac{1 - \varepsilon_{1n}}{1 - \varepsilon_{0n}} \Delta_{in} \right\}. \tag{9}$$

Then for any $Q_n' \in \mathscr{P}_{0n}(\tilde{\varepsilon}_{0n}, \tilde{\delta}_{0n})$ we have

$$Q_n'(IC_n(x) \in [r_{0n}, r_{1n}]) = 1$$

and hence we may restrict ourselves on $t \in (r_{0n}, r_{1n})$. But for every $t \in (r_{0n}, r_{1n})$ we have $\{IC_n(x) > t\} \cap \{\Delta_{1n} \leqq \Delta_n\} \neq \emptyset$ which yields

$$Q_{0n}^{IC}(IC_n(x) > t) \geqq Q_n'(IC_n(x) > t).$$

The proof for $i = 1$ is similar (compare Lemma 1 of [9]). □

Remark 2. It was shown in Lemma 2 of [9] that $\mathrm{LFP}_{T_{m,n}}$ with respect to $(\mathscr{P}_{0n}^{\otimes m}(\tilde{\varepsilon}_{0n}, \tilde{\delta}_{0n}), \mathscr{P}_{1n}^{\otimes m}(\tilde{\varepsilon}_{1n}, \tilde{\delta}_{1n}))$ ($\mathscr{P}^{\otimes m}$ denotes the set of product probability measures on $(\mathscr{X}^m, \mathscr{A}^m)$ generated by $\mathscr{P}$) consists of $((Q_{0n}^{IC})^{\otimes m}, (Q_{1n}^{IC})^{\otimes m})$. Therefore the following formula for $(Q_{in}^{IC})^{\otimes m}$ will be useful.

Lemma 2. For any $B \in \mathscr{A}$, put

$$\tilde{Q}_{in}(B) = (1 - \tilde{\varepsilon}_{in}) P_{in}(B \cap \{\Delta_{0n} < \Delta_n < \Delta_{1n}\})$$

and denote

$$L_{in} = Q_{in}^{IC}(IC_n(x) = r_{0n})$$

and

$$U_{im} = Q_{in}^{IC}(IC_n(x) = r_{1n}).$$

Then for any $t \in (-\infty, \infty)$ and $x \in \mathscr{X}^\infty$

$$(Q_{in}^{IC})^{\otimes m}(T_{m,n}(x) > t) = \sum_{j=0}^{m} \binom{m}{j} \sum_{k=0}^{m-j} \binom{m-j}{k} L_{in}^j U_{in}^k \tilde{Q}_{in}^{\otimes(m-j-k)}(T_{m-j-k,n}(x) > t - jr_{0m} - kr_{1n}),$$

where $\tilde{Q}_{in}^{\otimes 0}(A) = 1$ for all $A = \mathscr{A}$.

The proof is similar to the proof of Assertion 1 of [8] and will be omitted.

According to the outline given above we shall try to establish

$$\lim_{\tilde{\varepsilon}_{in} \to \varepsilon_{in}+, \tilde{\delta}_{in} \to \delta_{in}+} (\tilde{\varepsilon}_{in} + \tilde{\delta}_{in} - \varepsilon_{in} - \delta_{in})^{-1} \{(Q_{in}^{IC})^{\otimes n}(T_{n,n}(x) > t) - Q_{in}^{\otimes n}(T_{n,n}(x) > t)\},$$

once we have found $(Q_{in}^{IC})^{\otimes n}(T_{n,n}(x) > t)$. We shall work with the following model of contaminacy; a convenience of this model will be apparent from the context below.

Assumptions AS 1. Let $\tilde{\varepsilon}_{in} = \varepsilon_{in} + \xi_i(\tau, n)$ and $\tilde{\delta}_{in} = \delta_{in} + \zeta_i(\tau, n)$ where $\xi_i: [0, \infty) \times \mathbb{N} \to [0, 1]$, $\zeta_i: [0, \infty) \times \mathbb{N} \to [0, 1]$ are mappings which have right derivatives at $\tau = 0$ (denoted by $\partial/\partial\tau\, \xi_i(0, n)$, and $\partial/\partial\tau\, \xi_i(0, n)$, respectively), and are non-decreasing in τ and satisfy $\xi_i(0, n) = 0$ and $\zeta_i(0, n) = 0$, respectively, for all $n \in \mathbb{N}$.

Definition 3. Denote

$$\mathsf{SD}(t, n) = \lim_{\tau \to 0+} \tau^{-1}\{(Q_{0n}^{IC})^{\otimes n} (T_{n,n}(x) > t) - Q_{0n}^{\otimes n}(T_{n,n}(x) > t)\}$$

and

$$\mathsf{PD}(t, n) = \lim_{\tau \to 0+} \tau^{-1}\{(Q_{1n}^{IC})^{\otimes n} (T_{n,n}(x) > t) - Q_{1n}^{\otimes n}(T_{n,n}(x) > t)\}$$

for $i \in \mathbb{R}$ and $n = \mathbb{N}$.

These limits, if exist, will be called the dependency of the size and of the power of test on the changes of the contamination level, respectively, or briefly size and power dependency.

Theorem 1. Let Assumptions AS 1 be fulfilled. Then

$$\begin{aligned}(10)\qquad \mathsf{SD}(t, n) = & -n\{(1 - \varepsilon_{0n})^{-1} (\partial/\partial\tau)\, \xi_0(0, n)\, Q_{0n}^{\otimes n}(T_{n,n}(x) > t) + \\ & + [\omega_{0n}(\partial/\partial\tau)\, \xi_0(0, n) + (\partial/\partial\tau)\, \zeta_0(0, n)]\, Q_{0n}^{\otimes n-1}(T_{n-1,n}(x) > t - r_{0n}) - \\ & - [(1 + \nu_{0n})(\partial/\partial\tau)\, \xi_0(0, n) + (\partial/\partial\tau)\, \zeta_0(0, n)]\, Q_{0n}^{\otimes n-1}(T_{n-1,n}(x) > t - r_{1n})\}\end{aligned}$$

and

$$\begin{aligned}(11)\qquad \mathsf{PD}(t, n) = & -n\{(1 - \varepsilon_{1n})^{-1} (\partial/\partial\tau)\, \xi_1(0, n)\, Q_{1n}^{\otimes n}(T_{n,n}(x) > t) + \\ & + [\omega_{1n}(\partial/\partial\tau\ \xi_1(0, n) + (\partial/\partial\tau)\, \zeta_1(0, n)]\, Q_{1n}^{\otimes n-1}(T_{n-1,n}(x) > t - r_{1n}) - \\ & - [(1 + \nu_{1n})(\partial/\partial\tau)\, \xi_1(0, n) + (\partial/\partial\tau)\, \zeta_1(0, n)]\, Q_{1n}^{\otimes n-1}(T_{n-1,n}(x) > t - r_{0n})\}\end{aligned}$$

(for r_{in} see (9)).

Remark 3. Notice that for $\delta_{in} = 0$ and $\zeta_i(\tau, n) = 0$ $(i = 0, 1)$ the relation (10) and (11) turn into (4) and (5) of [9]. (10) and (11), though seemingly complicated allow a simple asymptotic approximation in the local alternative set-up (see Corrollary 1).

Proof of Theorem 1. Let us denote

$$M_{0n} = Q_{0k}(IC_n(x) = r_{0n}),$$
$$V_{0n} = Q_{0k}(IC_n(x) = r_{1n})$$

and for any $B \in \mathscr{A}$

$$\hat{Q}_{0n}(B) = (1 - \varepsilon_{0n})\, P_{0n}(B \cap \{\Delta_{0n} < \Delta_n < \Delta_{1n}\}).$$

Then we have (compare Lemma 2) for $m, n \in \mathbb{N}$ and $x \in \mathscr{X}^\infty$, $i \in \mathbb{R}$

$$Q_{0n}^{\otimes m}((T_{m,n}(x) > t) =$$
$$= \sum_{j=0}^{m} \binom{m}{j} \sum_{k=0}^{m-j} \binom{m-j}{k} M_{0n}^j V_{0n}^k \hat{Q}_{0n}^{\otimes(m-j-k)} (T_{m-j-k,n}(x) > i - jr_{0n} - kr_{1m}))$$

and

$$L_{0n} = M_{0n} - \lambda(\tau, n),$$

and

$$U_{0n} = V_{0n} + \varkappa(\tau, n)$$

where

$$(12)\qquad \lambda(\tau, n) = \xi_0(\tau, n)\, P_{0n}(\Delta_n \leqq \Delta_{0n}) + \zeta_0(\tau, n)$$

and

(13) $$\varkappa(\tau, n) = \xi_0(\tau, n)\left[1 - P_{0n}(\Delta_{1n} \leqq \Delta_n)\right] + \zeta_0(\tau, n)\,.$$

Moreover

$$\tilde{Q}_{0n} = (1 - \psi(\tau, n))\,\hat{Q}_{0n}\,,$$

where

(14) $$\psi(\tau, n) = (1 - \varepsilon_{0n})^{-1}\,\xi_i(\tau, n)\,.$$

Now for fixed j and k write

$$\binom{m}{j} L_{0n}^j \binom{m-j}{k} U_{0n}^k\, \tilde{Q}_{0n}^{\otimes m-j-k}(T_{m-j-k,n}(x) > t - jr_{0n} - kr_{1n}) -$$

$$- \binom{m}{j} M_{0n}^j \binom{m-j}{k} V_{0n}^k \hat{Q}_{0n}^{\otimes m-j-k}(T_{m-j-k,n}(x) > t - jr_{0n} - kr_{1n}) =$$

$$= \binom{m}{j}\binom{m-j}{k}\{[(M_{0n} - \lambda(\tau, n))^j - M_{0n}^j]\,U_{0n}^k \tilde{Q}_{0n}^{\otimes m-j-k}(T_{m-j-k,n}(x) > t - jr_{0n} - kr_{1n})$$

$$- M_{0n}^j[(V_{0n} + \varkappa(\tau, n))^k - V_{0n}^k]\,\tilde{Q}_{0n}^{\otimes m-j-k}(T_{m-j-k,n}(x) > t - jr_{0n} - kr_{1n}) -$$

$$- M_{0n}^j V_{0n}^k[\tilde{Q}_{0n}^{\otimes m-j-k}(T_{m-j-k,n}(x) > t - jr_{0n} - kr_{1n}) -$$

$$- \hat{Q}_{0n}^{\otimes m-j-k}(T_{m-j-k,n}(x) > t - jr_{0n} - kr_{1n})]\}\,.$$

A straightforward computation gives

$$\lim_{\tau \to 0+} \frac{1}{\tau}\,[(M_{0n} - \lambda(\tau, n))^j - M_{0n}^j]\,U_{0n}^k \tilde{Q}_{0r}^{\otimes m-j-k}(T_{m-j-k,n}(x) > t - jr_{0n} - kr_{1n}) =$$

$$= -j\,\frac{\partial}{\partial \tau}\,\lambda(0, n)\,M_{0n}^{j-1} V_{0n}^k \hat{Q}_{0n}^{\otimes m-j-k}(T_{m-j-k,n}(x) > t - jr_{0n} - kr_{1n})\,,$$

$$\lim_{\tau \to 0+} \frac{1}{\tau}\,M_{0n}^j[(V_{0k} + \varkappa(\tau, n))^k - V_{0n}^k)\,\tilde{Q}_{0n}^{\otimes m-j-k}(T_{m-j-k,n}(x) > t - jr_{0n} - kr_{1n}) =$$

$$= k\,\frac{\partial}{\partial \tau}\,\varkappa(0, n)\,M_{0n}^j V_{0n}^{k-1} \hat{Q}_{0n}^{\otimes m-j-k}(T_{k-j-k,n}(x) > t - jr_{0n} - kr_{1n})$$

and

$$\lim_{\tau \to 0+} \frac{1}{\tau}\,M_{0n}^j V_{0n}^k\,[\tilde{Q}_{0n}^{\otimes m-j-k}(T_{m-j-k,n}(x) > t - jr_{0n} - kr_{1n}) -$$

$$- \hat{Q}_{0n}^{\otimes m-j-k}(T_{m-j-k,n}(x) > t - jr_{0n} - kr_{1n})] =$$

$$= -(m - j - k)\,\frac{\partial}{\partial \tau}\,\psi(0, n)\,M_{0n}^j V_{0n}^k \hat{Q}_{0n}^{\otimes m-j-k}(T_{m-j-k,n}(x) > t - jr_{0n} - kr_{1n})\,.$$

So one obtains

$$\lim_{\tau \to 0+} \frac{1}{\tau} \sum_{j=0}^{m} \binom{m}{j} \sum_{k=0}^{m-j} \binom{m-j}{k} [(M_{0n} - \lambda(\tau, n))^j - M_{0n}^j]\,.$$

$$.\,U_{0n}^k \tilde{Q}_{0n}^{\otimes m-j-k}(T_{m-j-k,n}(x) > t - jr_{0n} - kr_{1n}) =$$

$$= -m\frac{\partial}{\partial\tau}\lambda(0,n)\sum_{j=0}^{m-1}\binom{m-1}{j}\sum_{k=0}^{m-1-j}\binom{m-1-j}{k}M_{0n}^{j}V_{0n}^{k}\hat{Q}_{0n}^{\otimes m-j-k}(T_{m-1-k,n}(x)>t-$$

$$-r_{0n}-jr_{0n}-kr_{1n}) = -m\frac{\partial}{\partial\tau}\lambda(0,n)\,Q_{0n}^{\otimes m-1}(T_{m-1,n}(x)>t-r_{0n})$$

and likewise

$$\lim_{\tau\to 0+}\frac{1}{\tau}\sum_{j=0}^{m}\binom{m}{j}\sum_{k=0}^{m-j}\binom{m-j}{k}M_{0n}^{j}[(V_{0n}+\varkappa(\tau,n))^{k}-V_{0n}^{k}]\,\tilde{Q}_{0n}^{\otimes m-j-k}(T_{m-j-k,n}(x)>t-$$

$$-jr_{0n}-kr_{1n}) =$$

$$= m\frac{\partial}{\partial\tau}\varkappa(0,n)\,Q_{0n}^{\otimes m-1}(T_{m-1,n}(x)>t-r_{1n}).$$

Finally

$$\lim_{\tau\to 0+}\frac{1}{\tau}\sum_{j=0}^{m}\binom{m}{j}\sum_{k=0}^{m-j}\binom{m-j}{k}M_{0n}^{j}V_{0n}^{k}[\tilde{Q}_{0n}^{\otimes m-j-k}(T_{m-j-k,n}(x)>t-jr_{0n}-kr_{1n})-$$

$$-\hat{Q}_{0n}^{\otimes m-j-k}(T_{m-j-k,n}(x)>t-jr_{0n}-kr_{1n})] =$$

$$-m\frac{\partial}{\partial\tau}\psi(0,n)\{Q_{0n}^{\otimes m}(T_{m,n}(x)>t)-M_{0n}Q_{0n}^{\otimes m-1}(T_{m-1,n}(x)>t-r_{0n})-$$

$$-V_{0n}Q_{0n}^{\otimes m-1}(T_{m-1,n}(x)>t-r_{1n})\}.$$

Taking into account that (see (3), (12), (13) and (14))

$$\frac{\partial}{\partial\tau}\lambda(0,n) = \frac{\partial}{\partial\tau}\xi_0(0,n)\,P_{0n}(\Delta_n\leqq\Delta_{0n})+\frac{\partial}{\partial\tau}\zeta_0(0,n),$$

$$\frac{\partial}{\partial\tau}\varkappa(0,n) = \frac{\partial}{\partial\tau}\xi_0(0,n)\,[1-P_{0n}(\Delta_{1n}\leqq\Delta_n)]+\frac{\partial}{\partial\tau}\zeta_i(0,n),$$

$$\frac{\partial}{\partial\tau}\psi(0,n) = (1-\varepsilon_{0n})^{-1}\frac{\partial}{\partial\tau}\xi_0(0,n),$$

$$M_{0n}\frac{\partial}{\partial\tau}\psi(0,n) = \frac{\partial}{\partial\tau}\xi_0(0,n)\,\{P_{0n}(\Delta_n\leqq\Delta_{0n})-\omega_{0n}\}$$

and

$$V_{0n}\frac{\partial}{\partial\tau}\psi(0,n) = \frac{\partial}{\partial\tau}\xi_0(0,n)\,\{P_{0n}(\Delta_{1n}\leqq\Delta_n)+v_{0n}\},$$

one concludes the proof of the first assertion of Theorem 1.

The proof of the second one may be performed in a similar way. □

Now we would like to derive asymptotic approximations for (10) and (11).

For some $h>0$ let $\{P_\theta\colon |\theta|\leqq h\}$ denote a one-real-parameter family in $\mathscr{M}$ and parameters $\varepsilon_i, \delta_i\in[0,1)$ be given so that $0<\varepsilon_i+\delta_i<1$, $i=0,1$. For each $n\in\mathbb{N}$

and $i = 0, 1$ put

$$h_n = h \,.\, n^{-1/2}\,, \quad \varepsilon_{in} = \varepsilon_i n^{-1/2}\,, \quad \delta_{in} = \delta_i n^{-1/2}\,,$$

$$P_{0n} = P_{-h_n}\,, \quad P_{1n} = P_{h_n}\,.$$

Further we shall assume that the family $\{P_\theta : |\theta| \leqq h\}$ and parameters ε_i and δ_i fulfil the following regularity conditions (see [5]).

Assumptions AS 2. Let $P_\theta \ll P_0$ for all $|\theta| \leqq h$ and denote a suitable version of $\mathrm{d}P_\theta/\mathrm{d}P_0$ by p_θ. Let $p_\theta(x)$ be twice differentiable in θ for all $x \in \mathscr{X}$ and put $\Lambda(x) = \partial/\partial\theta\,[\log p_\theta(x)]_{\theta=0}$. Further let

$$0 < \lim_{\theta \to 0} \int \left(\frac{p_\theta^{1/2} - 1}{\theta}\right)^2 \mathrm{d}P_0 = \frac{1}{4}\int \Lambda^2(x)\,\mathrm{d}P_0 < \infty$$

$$\text{and} \quad \varepsilon_0 + \delta_0 + \delta_1 < \int (2h\,\Lambda(x) - \varepsilon_1 + \varepsilon_0)^+\,\mathrm{d}P_0\,,$$

where $(l(x))^+ = \max\{l(x), 0\}$. Finally, let there be $f \in \mathscr{L}^1(P_0)$ such that

$$\sup_{|\theta| \leqq h} \left|\frac{\partial^2}{\partial\theta^2}\,p_\theta(x)\right| \leqq f(x)$$

for all $x \in \mathscr{X}$.

Lemma 4. (Rieder). Put $d_{in} = (2h)^{-1}\,n^{1/2} r_{in}$ and $L_n(x) = (2h)^{-1}\,n^{1/2} I C_n(x)$. Then under the Assumptions AS 2

$$\lim_{n\to\infty} d_{in} = d_i \quad \text{and} \quad \lim_{n\to\infty} L_n(x) = L(x) \quad \text{for all} \quad x \in \mathscr{X}$$

where d_i are the unique solution of the equations

$$\int (d_0 - \Lambda(x))^+\,\mathrm{d}P_0 = (2h)^{-1}(\varepsilon_1 + \delta_0 + \delta_1)$$

and

$$\int (\Lambda(x) - d_1)^+\,\mathrm{d}P_0 = (2h)^{-1}(\varepsilon_0 + \delta_0 + \delta_1)$$

and

$$L(x) = \max\{d_0, \min\{\Lambda(x), d_1\}\} - (2h)^{-1}(\varepsilon_1 - \varepsilon_0)\,.$$

Denote $\mathsf{E}_{P_0}\,L^2(x) = \sigma^2$, then

$$\lim_{n\to\infty} \mathsf{E}_{Q_{0n}}\,T_{n,n}(x) = -2h^2\sigma^2\,,$$

$$\lim_{n\to\infty} \mathsf{E}_{Q_{1n}}\,T_{n,n}(x) = 2h^2\sigma^2$$

and

$$\mathrm{var}_{Q_{in}}\,T_{n,n}(x) = \mathrm{var}_{Q_{in}}\{n^{1/2} I C_n(x)\} = 4h^2\sigma^2 + O(n^{-1/2})\,.$$

For the proof see [5], Theorem 4.1 and its proof.

Let hereafter Φ and φ denote the standard normal distribution and its density with respect to Lebesque measure, respectively. Further let $v_n(t)$ denote the characteristic function of $IC_n(x)$ with respect to P_0 and $\sigma_{in} = \mathrm{var}_{Q_{in}} \{n^{1/2} IC_n(x)\}$.

Lemma 5 (Petrov, Edgeworth). Let us denote $\mu_{in} = \mathsf{E}_{Q_{in}} T_{m,n}(x)$ and let $\limsup_{n\to\infty} \limsup_{|t|\to\infty} |v_n(t)| < 1$. Then under Assumptions AS 2 we have for any $t \in \mathbb{R}$ and $z_i = (t - \mu_{in}) \sigma_{in}^{-1}$

$$Q_{in}^{\otimes n}(T_{n,n}(x) \leqq t) - \Phi(z_i) + n^{-1/2} R(z_i) = o(n^{-1/2}),$$

where

$$R(u) = (2\pi)^{-1/2} \exp\left\{-\frac{u^2}{2}\right\} (48h^3\sigma^3)^{-1} \gamma(u^2 - 1)$$

and γ is the third cumulant of $2h\, L(x)$ with respect to P_0.

Proof. Due to uniform (with respect to $n \in \mathbb{N}$) boundedness of $n^{1/2} IC_n(x)$ the proof follows directly from Theorem 1, Chapter VI, § 3 of [3] (see Lemma 4 of this paper and also Lemma 2 and Corollary 1 of [8] and compare also [7]). In what follows let us put for $\alpha \in (0, 1)$

$$t_\alpha(n) = \inf\{t \in \mathbb{R}: Q_{0n}^{\otimes n}(T_{n,n}(x) > t) \leqq \alpha\}. \qquad \square$$

Corollary 1. Let Assumptions AS 1 and AS 2 be fulfilled. Then

$$\mathsf{SD}(t_\alpha(n), n) = n^{1/2} . \{(\partial/\partial\tau)\, \xi_0(0, n) (1 - \varepsilon_{0n})^{-1} [d_1 + h\sigma^2] +$$
$$+ (\omega_{0n}(\partial/\partial\tau)\, \xi_0(0, n) + (\partial/\partial\tau)\, \zeta_0(0, n)) (d_1 - d_0)\} . \{\varphi(u_\alpha) . \sigma^{-1} + o(1)\}$$

and

$$\mathsf{PD}(t_\alpha(n), n) = n^{1/2} . \{\partial/\partial\tau)\, \xi_1(0, n) (1 - \varepsilon_{1n})^{-1} [d_0 - h\sigma^2] +$$
$$+ (\omega_{1n}(\partial/\partial\tau)\, \xi_1(0, n) + (\partial/\partial\tau)\, \zeta_1(0, n)) (d_0 - d_1)\} . \{\varphi(u_\alpha - 2h\sigma)\, \sigma^{-1} + o(1)\}.$$

Proof. Put $z_0(\alpha, n) = (t_\alpha(n) - \mu_{0n-1})\, \sigma_{0n-1}$. From Lemma 4 and Lemma 5 it follows that

$$z_0(\alpha, n) = u_\alpha + O(n^{-1/2}). \tag{15}$$

Due to the equality

$$1 + v_{0n} = (1 - \varepsilon_{0n})^{-1} (1 + \delta_{0n})$$

we may write (10) in the form

$$\mathsf{SD}(t, n) = -n(1 - \varepsilon_{0n})^{-1} \{(\partial/\partial\tau)\, \xi_0(0, n)\, [Q_{0n}^{\otimes n}(T_{n,n}(x) > t) - Q_{0n}^{\otimes n-1}(T_{n-1,n}(x) >$$
$$> t - r_{1n})] + (\delta_{0n}(\partial/\partial\tau)\, \xi_0(0, n) + (1 - \varepsilon_{0n}) (\partial/\partial\tau)\, \zeta_0(0, n)) .$$
$$. [Q_{0n}^{\otimes n-1}(T_{n-1,n}(x) > t - r_{0n}) - Q_{0n}^{\otimes n-1}(T_{n-1,n}(x) > t - r_{1n})]\} .$$

From the preceding lemma we have

$$Q_{0n}^{\otimes n-1}(T_{n-1,n}(x) > t_\alpha(n) - r_{0n}) - Q_{0n}^{\otimes n-1}(T_{n-1,n}(x) > t_\alpha(n) - r_{1n}) =$$
$$= \Phi(z_0(\alpha, n) - \sigma_{0n}^{-1} r_{0n}) - \Phi(z_0(\alpha, n) - \sigma_{0n}^{-1} r_{1n}) -$$
$$- n^{-1/2}[R(z_0(\alpha, n) - \sigma_{0n}^{-1} r_{0n}) - R(z_0(\alpha, n) - \sigma_{0n}^{-1} r_{1n})] + o(n^{-1/2}) =$$
$$= \varphi(v_0(\alpha, n))\, [r_{0n} - r_{1n}]\, (2h\sigma)^{-1} + o(n^{-1/2}),$$

where $v_0(\alpha, n) \in (z_0(\alpha, n) - \sigma_{0n}^{-1} r_{1n}, \; z_0(\alpha, n) - \sigma_{0n}^{-1} r_{0n})$ and the facts that

$$\lim_{n \to \infty} n^{1/2}(r_{0n} - r_{1n}) = 2h(d_0 - d_1) \tag{16}$$

and

$$\lim_{n \to \infty} \sigma_{0n} = 2h\sigma$$

(which follow from Lemma 4) was taken into account. Now from (15) and (16) one derives that

$$v_0(\alpha, n) = u_\alpha + O(n^{-1/2})$$

and hence

$$n^{1/2}[Q_{0n}^{\otimes n-1}(T_{n-1,n}(x) > t_\alpha(n) - r_{0n}) - Q_{0n}^{\otimes n-1}(T_{n-1,n}(x) > t_\alpha(n) - r_{1n})] =$$
$$= \varphi(u_\alpha)\, \sigma^{-1}(d_0 - d_1) + o(1)\,.$$

The difference

$$Q_{0n}^{\otimes n}(T_{n,n}(x) > t_\alpha(n)) - Q_{0n}^{\otimes n-1}(T_{n-1,n}(x) > t_\alpha(n) - r_{1n})$$

may be attacked in the similar way (it needs more technicalities but the idea is the same – see the proof of Corollary 1 in [8] and note that $\mathsf{E}_{Q_{0n}}\, T_{n-1,n}(x) = (n-1)\,.\, n^{-1}\mu_{in}$). We obtain

$$n^{1/2}[Q_{0n}^{\otimes n}(T_{n,n}(x) > t_\alpha(n)) - Q_{0n}^{\otimes n-1}(T_{n-1,n}(x) > t_\alpha(n) - r_{1n})] =$$
$$= \varphi(u_\alpha)\, \sigma^{-1} d_1 + o(1)\,,$$

which concludes the proof of the first assertion of corollary. The proof of the second one may be performed likewise. □

Remark 4. The results of the corollary imply that to keep the size and power dependency bounded uniformly with respect to the sample size it requires

$$\frac{\partial}{\partial \tau}\, \xi_i(0, n) = O(n^{-1/2}) \quad \text{and} \quad \frac{\partial}{\partial \tau}\, \zeta_i(0, n) = O(n^{-1/2})\,.$$

(For the more detailed discussion of these equalities see the last section of [8].) On the other hand it is evident that the results presented in Theorem 1, although they are a little complicated, are (together with a numerical approximation method) much more useful for creating an idea about the sensitivity of the LFP-test with respect to failing estimation of the contamination level than those of Corollary 1 because of the numerical unreliability of the approximation formulae derived in the local alternative set-up (see also [7]).

(Received April 19, 1985.)

REFERENCES

[1] W. J. R. Eplett: An influence curve for two-sample rank tests. J. Roy. Statist. Soc. Ser. *B 42* (1980), 67–70.

[2] D. Lambert: Influence function for testing. J. Amer. Statist. Assoc. *76* (1981), 649–657.

[3] В. В. Петров: Суммы независимых случайных величин. Наука, Москва 1972.

[4] H. Rieder: Least favourable pairs for special capacities. Ann. Statist. *5* (1977), 909—921.
[5] H. Rieder: A robust asymptotic testing model. Ann. Statist. *6* (1978), 1080—1096.
[6] E. Ronchetti: Robust Testing in Linear Models: The Infinitesimal approach. Ph. D. Dissertation, SFIT Zurich 1982.
[7] J. Á. Víšek: On second order efficiency of a robust test and approximation of its error probabilities. Kybernetika *19* (1983), 5, 387—407.
[8] J. Á. Víšek: Sensitivity of the test risk with respect to contamination. Commun. Statist. — Sequential Analysis *2* (3) (1983), 243—258.
[9] J. Á. Víšek: On the dependence of the test error probabilities on the contamination level. Statistics *17* (1986), 1, 25—32.

RNDr. Jan Ámos Víšek CSc., Ústav teorie informace a automatizace ČSAV (Institute of Information Theory and Automation — Czechoslovak Academy of Sciences), Pod vodárenskou věží 4, 182 08 Praha 8, Czechoslovakia.