Kybernetika

Jan Ámos Víšek

On second order efficiency of a robust test and approximations of its error probabilities

Kybernetika, Vol. 19 (1983), No. 5, 387--407

Persistent URL: http://dml.cz/dmlcz/125787

KYBERNETIKA — VOLUME 19 (1983), NUMBER 5

ON SECOND ORDER EFFICIENCY OF A ROBUST TEST AND APPROXIMATIONS OF ITS ERROR PROBABILITIES

JAN ÁMOS VÍŠEK

Second order efficiency of asymptotically minimax robust test proposed by H. Rieder is proved. Approximations of error probabilities of the most powerful robust test obtained by means of second order Edgeworth expansion are compared with simulated values as well as with approximations derived from asymptotic distribution of this test found in a setting of local alternatives.

1. INTRODUCTION

In 1977 H. Rieder, following Huber and Strassen [4], has proposed a new model of contamination and found a least favourable pair of distribution (LFP), the likelihood ratio of which may be used as a test statistic for construction of the most powerful test. The form of contamination which is considered in Rieder's model implies avoidance or decrease of bad effects of the fact that a portion of population can be generated by a distribution different from the assumed one. Moreover, this model ensures us against errors caused by rounding numbers. The level of decrease (or in an ideal case, of avoidance) of the bad effects depends on the accuracy with which contamination parameters are estimated (or more precisely, guessed). But the applicability of Rieder's robust test was seriously influenced by the lack of knowledge of its distribution.

Let us assume that we study an i.i.d. model. Then the most powerful test is, in fact, based on the test statistic which is the sum of logaritms of likelihood ratio of LFP. (In what follows this test is denoted as LFP-test.) A distribution of LFP-test is therefore an nth convolution of a distribution which may be derived from LFP. The shape of this distribution is however generally a little more complicated than to be feasible, even in the simplest cases, to derive characteristic function or directly find convolutions even for not very large sizes of samples.

Let us assume to be faced with a real (and therefore finite) testing problem. In what follows the problem of testing a simple hypothesis against a simple alternative both being contaminated, so in fact, the problem of testing a composite hypothesis against a composive alternative, is considered. Our aim may be then to construct an appropriate asymptotic setting running "through" our problem (i.e. a sequence of testing problems incorporating the initial problem for some finite size of sample) and to use the asymptotic results obtained in such setting to approximate the characteristics (e.g. the error probabilities) which were required in the initial problem. To complete the work on such a task it is necessary to check the effectiveness of the approximations (see e.g. [2] and [7]).

In the present paper two well known asymptotic settings are recalled. In the first of them the hypothesis, the alternative and the level of contamination are fixed with only the number of observations running to infinity. The most powerful test is then the LFP-test, an analytic form of which was found in [5]. Approximations of the error probabilities are obtained by the Edgeworth or saddle-point expansions of distribution of the sum of logarithms of the likelihood ratio of LFP. The values gained by them were checked by simulations. The second order expansions proved to be very good for this purpose.

Sometimes it would be useful to know not only the error probabilities of the LFP-test when one of the distributions from LFP is true, but also the error probabilities of it when one of the initial distributions is true (i.e. when one of the distributions which were "centres" of composite hypothesis and alternative, is true). These values may help us to judge how great loss we shall suffer, when the LFP-test will be used instead of the most powerful test of the non-contaminated model, i.e. what is the price of superfluous ensurance against contamination.

Undoubtedly, it is also of interest to know how significant changes of error probabilities might be expected if our estimation of the level of contamination is unprecise. Some numerical examples offering a first rough image about it are presented at the end of the paper.

However an evaluation of the cumulants of the logarithm of the likelihood ratio of LFP up to the forth one (necessary for the second order Edgeworth expansion) is usually a tiresome task. So we would appreciate a possibility to approximate the error probabilities in a simpler way. Such possibility seemed to be offered by the second setting, the setting of local alternatives. As it is known, in this setting the sequence of hypotheses and the sequence of alternatives converge one to the other with increasing number of observations. In order to ensure the disjointness of the composite hypothesis and the composite alternative both having been produced by contamination, one must define the parameters of contamination as decreasing functions of the number of observations. For any size of sample we may then find an LFP-test and study their asymptotic distribution. This was done by Rieder in [6]. His results represent a simpler way of approximation of error probabilities in comparison with the saddle-point or Edgeworth expansion. Unfortunately, the approximations gained by the local alternative setting turned out to be rather bad. So we are still challenged to find another asymptotic setting giving good approximations of the error probabilities and on the other hand not requiring long and dreary computation.

In above mentioned paper [6] Rieder established an asymptotically most powerful test and showed its first order efficiency. The second order efficiency is proved in the present paper under a supplementary condition and numerically illustrated by means of Edgeworth approximations.

2. NOTATIONS

Let $\mathbb{R}$ denote the real line and $\mathcal{N}$ the set of all positive integers. Let $(\Omega, \mathcal{B})$ be a measurable space and let $\mathcal{M}$ be the set of all probability measures on it. For some $\tau > 0$ let $\{P_\theta : |\theta| \leqq \tau\}$ denote a real parameter family from $\mathcal{M}$. Let ε_j, δ_j be real numbers $0 \leqq \varepsilon_j, 0 \leqq \delta_j$, $0 < \varepsilon_j + \delta_j < 1$, $j = 0, 1$. Define for functions $f : \mathcal{N} \times \mathbb{R} \to [0, \tau]$ and $g, h : \mathcal{N} \times \mathbb{R} \to [0, 1)$

$$\tau_n = f(n, \tau), \quad \varepsilon_{jn} = g(n, \varepsilon_j), \quad \delta_{jn} = h(n, \delta_j),$$

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

$$(1) \qquad \mathcal{P}_{jn} = \{Q \in \mathcal{M} : Q(B) \geqq (1 - \varepsilon_{jn}) P_{jn}(B) - \delta_{jn} \quad \text{for all} \quad B \in \mathcal{B}\},$$

$$\mathcal{P}_{jn}^{\otimes n} = \{\bigotimes_{i=1}^{n} Q_i : Q_i \in \mathcal{P}_{jn} \quad \text{for} \quad i = 1, \ldots, n\},$$

$$H_j = \{w_n : w_n \in \mathcal{P}_{jn}^{\otimes n} \quad \text{for all} \quad n \in \mathcal{N}\},$$

where "$\bigotimes_{i=1}^{n}$" denotes the nth product and w_n may be considered to be the distributions of samples of sizes n. Let us recall the definition of LFP (Q_{0n}, Q_{1n}) for $(\mathcal{P}_{0n}, \mathcal{P}_{1n})$: Q_{0n} and Q_{1n} are the probability distributions satisfying

$$Q_{0n}(\{\pi < t\}) = \sup\{Q'(\{\pi < t) \ : Q' \in \mathcal{P}_{0n}\},$$

$$Q_{1n}(\{\pi < t\}) = \inf\ \{Q''(\{\pi < t\}) : Q'' \in \mathcal{P}_{1n}\}, \quad t \in (0, \infty),$$

where $\pi \in \mathrm{d}Q_{1n}/\mathrm{d}Q_{0n}$.

Remark 1. Taking $f(n, z) = g(n, z) = h(n, z) = z$ we shall obtain a classical setting with the fixed hypothesis, alternative, level of contamination and with the size of sample running to infinity. On the other hand putting $f(n, z) = g(n, z) = h(n, z) = z/\sqrt{n}$ we shall arrive at the local alternative setting.

3. ASSUMPTIONS

AS 1. There is $\tau > 0$, such that $P_\theta \ll P_0$ for all $|\theta| < \tau$. Let p_θ denote a suitable version of $\mathrm{d}P_\theta / \mathrm{d}P_0$.

AS 2. $\theta \to p_\theta(x)$ is twice differentiable for all $x \in \Omega$. Put

$$\Lambda(x) = \left[\frac{\partial}{\partial\theta} \log p_\theta(x)\right]_{\theta=0}.$$

AS 3.
$$0 < \int \Lambda^2(x)\,\mathrm{d}P_0 < \infty\,.$$

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

AS 5. There is a function $\bar{h}$ in $L^1(P_0)$ such that

$$\sup_{|\theta|\leqq\tau} \left|\frac{\partial^2}{\partial\theta^2}\, p_\theta(x)\right| \leqq \bar{h}(x) \quad \text{for all} \quad x \in \Omega\,.$$

AS 6.
$$(\varepsilon_0 + \delta_0 + \delta_1)/2\tau < \int \left(\Lambda(x) - \frac{\varepsilon_1 - \varepsilon_0}{2\tau}\right)^+ \mathrm{d}P_0\,.$$

Let d_0 and d_1 be real numbers given by

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

and

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

Further let $\Delta_n \in \mathrm{d}P_{1n}/\mathrm{d}P_{0n}$ and Δ_{0n} and Δ_{1n} be real numbers defined by

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

(2)

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

Let for $A, B \in \mathscr{B}$

$$A \div B := (A \cap B^c) \cup (A^c \cap B)$$

and

$$D_0 = \{x \in \Omega : d_0 \geqq \Lambda(x)\}\,, \quad D_1 = \{x \in \Omega : d_1 \leqq \Lambda(x)\}$$

and for each $n \in \mathscr{N}$

$$C_{0n} = \{x \in \Omega : \Delta_n(x) < \Delta_{0n}\}\,, \quad C_{1n} = \{x \in \Omega : \Delta_{1n} < \Delta_n(x)\}\,.$$

Finally put

$$G_n = \{(D_0 \cap C_{0n}) \cup (D_1 \cap C_{1n})\}^c$$

and

$$\varphi_n(x) = \sup_{|\theta| \leq \tau} \frac{\partial^2}{\partial \theta^2} \log p_o(x) .$$

Now we may give a supplementary assumption.

AS 7. $$\limsup_{n \to \infty} \sup \{\varphi_n(x) : x \in G_n\} < \infty .$$

Remark 2. As 7 may be weakened in the following way:

AS 7'. $$\limsup_{n \to \infty} \sup \{\varphi_n(x) : x \in (D_0 \div C_{0n}) \cup (D_1 \div C_{1n})\} < \infty$$

with the random variable

$$\varphi_n(x) \cdot 1_{\{D_0{}^c \cap C_{0n}{}^c \cap C_{1n}{}^c \cap D_1{}^c\}}(x)$$

having finite fourth moment (with respect to Q_{0n} – see the proof of Theorem 2).

Remark 3. H. Rieder has shown that an exponential family satisfies the above regularity assumptions AS 1 – AS 5 (AS 6 is fulfilled for small ε_j and δ_j). Let $\{f_o(x) = C(\theta) \cdot h(x) \cdot \exp\{Q(\theta) \cdot x\}$. It is a reasonable requirement that from

$$f_{o_1}(x) \equiv f_{o_2}(x)$$

follows $\theta_1 = \theta_2$. Now let $\theta_1 \neq \theta_2$. Let us assume that $Q(\theta_1) = Q(\theta_2)$. From it we have $C(\theta_1) = C(\theta_2)$ and finally $f_{o_1}(x) \equiv f_{o_2}(x)$. But it contradicts the above requirement. So we have a one-to-one mapping $Q(\theta) : \Theta \to R$; further, let us assume $\tilde{\Theta} \subset R$ and use a reparametrization

$$\tilde{\theta} = Q(\theta)$$

and

$$g_{\tilde{\theta}}(x) = C(Q^{-1}(\tilde{\theta}))\, h(x) \exp\{\tilde{\theta}x\} = D(\tilde{\theta})\, h(x) \exp\{\tilde{\theta}x\} .$$

Then

$$\Delta_{\tilde{\theta}}(x) := g_{\tilde{\theta}}(x)/g_0(x) = \mathscr{E}(\tilde{\theta}) \exp\{\tilde{\theta}x\}$$

and

$$\frac{\partial^2}{\partial \tilde{\theta}^2} \log \Delta_{\tilde{\theta}}(x) = \frac{\mathscr{E}''(\tilde{\theta})\, \mathscr{E}(\tilde{\theta}) - [\mathscr{E}'(\tilde{\theta})]^2}{\mathscr{E}^2(\tilde{\theta})} .$$

So the assumptions AS 7 may be written for the exponential family into the form

$$\sup_{|\tilde{\theta}| \leq \tilde{\tau}} \left| \frac{\mathscr{E}''(\tilde{\theta})\, \mathscr{E}(\tilde{\theta}) - [\mathscr{E}'(\tilde{\theta})]^2}{\mathscr{E}^2(\tilde{\theta})} \right| < \infty .$$

AS 7 is fulfilled for rather large range of probability families having

$$\frac{\partial^2}{\partial \theta^2} \log p_o(x)$$

continous while G_n being usually contained in a bounded closed set.

Remark 4. Also H. Rieder proved in [5] that

$$\pi_n = \frac{1 - \varepsilon_{1n}}{1 - \varepsilon_{0n}} \max \{\Delta_{0n}, \min \{\Delta_n, \Delta_{1n}\}\}$$

(Δ_{0n} and Δ_{1n} given in (2)) is the likelihood ratio of LFP (Q_{0n}, Q_{1n}) for $(\mathscr{P}_{0n}, \mathscr{P}_{1n})$ and, thus

$$W_n(x) = \sum_{i=1}^{n} \log \pi_n(x_i) \tag{3}$$

generates the minimax test of $H_{0n} : w \in \mathscr{P}_{0n}$ against $H_{1n} : w \in \mathscr{P}_{1n}$. In [6] he established that for

$$IC^*(x) := \max \{d_0, \min \{\Lambda(x), d_1\}\} ,$$

the statistic

$$Z_n(x) = \frac{1}{\sqrt{n}} \sum_{i=1}^{n} IC^*(x_i) \tag{4}$$

yields an asymptotically minimax test ψ_n of H_0 against H_1 (see (1)).

4. SECOND ORDER EFFICIENCY OF ψ_n

Throughout this section the local alternative setting is considered.

We are now going to use the result of Bickel, Chibisov and van Zwet presented in [1]. Let us quote at first their definition of the v_n-efficiency and their result.

Let for every $n \in \mathscr{N}$, P_{0n} and P_{1n} be two possible distributions of X_n (in an arbitrary sample space) and let $W_n(X_n)$ and $Z_n(X_n)$ be the logarithm of the likelihood ratio $\mathrm{d}P_{1n}(X_n)/\mathrm{d}P_{0n}(X_n)$ and a statistic, respectively.

Definition 1. Let $\Phi_n(W_n, \alpha_n)$ and $\psi_n(Z_n, \alpha_n)$ be two sequences of the test functions such that

$$\mathsf{E}_{0n}\, \Phi_n(W_n, \alpha_n) = \alpha_n \quad \text{and} \quad \mathsf{E}_{0n}\, \psi_n(Z_n, \alpha_n) = \alpha_n \, .$$

For a sequence $v_n \in (0, 1]$, we shall say that the sequence $\psi_n(Z_n, \alpha_n)$ is v_n-efficient if, for $n \to \infty$,

$$\mathsf{E}_{1n}[\Phi_n(W_n, \alpha_n) - \psi_n(Z_n, \alpha_n)] = o(v_n) \, .$$

Let us denote

$$\begin{aligned} \lambda_n &= 0 \quad \text{if} \quad W_n = Z_n \, , \\ &= W_n - Z_n \quad \text{otherwise} \, . \end{aligned}$$

Theorem 1. (Bickel, Chibisov, van Zwet, [1] p. 171). Suppose that

$$\liminf_{n \to \infty} \alpha_n > 0 \tag{B 1}$$

and that there exists $A > 0$ such that for every $x_0 \in \mathbb{R}$, every $\gamma > 0$ and $n \to \infty$

$$\sup_{x \leqq x_0} P_{0n}(x - v_n^{1/2} \leqq W_n \leqq x) = O(v_n^{1/2}) \, , \tag{B2}$$

(B3) $$\mathsf{E}_{0n}\{|\lambda_n|\, 1_{(\gamma v_n^{1/2}, A)}(|\lambda_n|)\} = o(v_n)\,,$$

(B4) $$P_{0n}(\lambda_n \geqq A) = o(v_n)\,,$$

(B5) $$P_{1n}(\lambda_n \leqq -A) = o(v_n)\,.$$

Then the sequence of test $\psi_n(Z_n, \alpha_n)$ is v_n-efficient.

Throughout the rest of the paper AS 1 – AS 7 are assumed to be fulfilled.

Theorem 2. Let $\psi_n(Z_n, \alpha_n)$ be given as follows:

$$\begin{aligned}\psi_n(Z_n, \alpha_n) &= 0 \quad Z_n < C_n(\alpha_n)\,.\\ &= 1 \quad Z_n > C_n(\alpha_n)\,,\end{aligned}$$

where Z_n is given in (4) and $C_n(\alpha_n)$ is chosen so that

$$\mathsf{E}\psi_n(Z_n, \alpha_n) = \alpha_n\,,$$

where the mean is taken with respect to $Q_{0n}^{\otimes n}$ (cf. (1)). Then the tests $\psi_n(Z_n, \alpha_n)$ are second order efficient.

To prove the theorem we shall give two lemmas. The first of them is due to Rieder (in [6], but it was not isolated there). Let us put

$$d_{jn} = \frac{1}{2\tau_n} \log \Delta_{jn}\,.$$

Lemma 1. For any $w_n \in H_0 \cup H_1$ (see (1)) we have

$$\lim_{n\to\infty} \mathrm{var}_{w_n} \frac{1}{2\tau_n} \log \pi_n(x) = \mathsf{E}_{P_0} IC^{*2}(x)$$

and

$$\lim_{n\to\infty} d_{jn} = d_j\,, \quad |d_j| < \infty\,, \quad j = 0, 1\,.$$

For the proof see [6], p. 1088, relation (*).

Lemma 2.

$$d_{jn} = d_j + O(n^{-1/2})\,.$$

Proof. Let us recall that $p_\theta(x)$ is the density of the probability measure P_θ, $|\theta| < \tau$, with respect to P_0. Let for real u, d and x

$$f(u, d, x) = \frac{\exp\{2ud\}\, p_{-u}(x) - p_u(x)}{u(1 + \exp\{2ud\})}\,,$$

$$F(d, u) = \int f^+(d, u, x)\, \mathrm{d}P_0(x)$$

and

$$G(d) = \int (d - \Lambda(x))^+ \, \mathrm{d}P_0(x) .$$

Then d_0 and d_{0n} may be redefined by

$$F(d_{0n}, \tau_n) = \frac{\varepsilon_1 + \delta_1}{\tau(1 - \varepsilon_{1n})(1 + \exp\{2\tau_n d_{0n}\})} + \frac{\delta_0 \exp\{2\tau_n d_{0n}\}}{\tau(1 - \varepsilon_{0n})(1 + \exp\{2\tau_n d_{0n}\})}$$

and

$$G(d_0) = \frac{\varepsilon_1 + \delta_0 + \delta_1}{2\tau} .$$

Let us recall that

(5) $$D_0 = \{x \in \Omega : d_0 - \Lambda(x) \geqq 0\}$$

and

$$C_{0n} = \{x \in \Omega : \Lambda_n(x) < \Lambda_{0n}\} .$$

From the following chain of inequalities:

$$\Lambda_n(x) < \Lambda_{0n} , \quad \log \frac{p_{\tau_n}}{p_{-\tau_n}} < \log \Lambda_{0n} ,$$

$$\log p_{\tau_n} < 2\tau_n d_{0n} + \log p_{-\tau_n} ,$$

$$p_{\tau_n} < \exp\{2\tau_n d_{0n}\} \, p_{-\tau_n} ,$$

$$0 < \frac{\exp\{2\tau_n d_{0n}\} \, p_{-\tau_n} - p_{\tau_n}}{\tau_n(1 + \exp\{\tau_n d_{0n}\})}$$

(remember that $\tau > 0$ and $\tau_n = \tau/\sqrt{n}$), we can see that

(6) $$C_{0n} = \{x \in \Omega : f(\tau_n, d_{0n}, x) < 0\} .$$

So the relations defining d_0 and d_{0n} may be rewritten

(7) $$\int_{D_0} (d_0 - \Lambda(x)) \, \mathrm{d}P_0 = \frac{\varepsilon_1 + \delta_0 + \delta_1}{2\tau}$$

(8) $$\frac{1}{\tau_n(1 + \exp\{2\tau_n d_{0n}\})} \int_{C_{0n}} [\exp\{2\tau_n d_{0n}\} \, p_{-\tau_n} - p_{\tau_n}] \, \mathrm{d}P_0 =$$

$$= \frac{\varepsilon_1 + \delta_1}{\tau(1 - \varepsilon_{1n})(1 + \exp\{2\tau_n d_{0n}\})} + \frac{\delta_0 \exp\{2\tau_n d_{0n}\}}{\tau(1 - \varepsilon_{0n})(1 + \exp\{2\tau_n d_{0n}\})} .$$

Let us recall that

$$\lim_{n \to \infty} d_{0n} = d_0 , \quad |d_0| < \infty ,$$

(see Lemma 1) and so

$$\exp\{2\tau_n d_{0n}\} = 1 + 2\tau_n d_{0n} + O(n^{-1}).$$

From it follows

$$\int_{C_{0n}} [\exp\{2\tau_n d_{0n}\}\, p_{-\tau_n} - p_{\tau_n}]\, \mathrm{d}P_0 = 2\tau_n \int_{C_{0n}} \frac{p_{-\tau_n} - p_{\tau_n}}{2\tau_n}\, \mathrm{d}P_0 +$$

$$+ 2\tau_n d_{0n} \int_{C_{0n}} \mathrm{d}P_0 + 2\tau_n d_{0n} \int_{C_{0n}} (\mathrm{d}P_{-\tau_n} - \mathrm{d}P_0) + O(n^{-1}).$$

Applying AS 3 and AS 5 and using the second order Taylor expansion one can find that

$$2\tau_n d_{0n} \int_{C_{0n}} (\mathrm{d}P_{-\tau_n} - \mathrm{d}P_0) = O(n^{-1}).$$

So we have

$$\int_{C_{0n}} [\exp\{2\tau_n d_{0n}\}\, p_{-\tau_n} - p_{\tau_n}]\, \mathrm{d}P_0 =$$

$$= -2\tau_n \int_{C_{0n}} \left[\frac{\partial p_\theta}{\partial \theta}\right]_{\theta=0} \mathrm{d}P_0 + (2\tau_n)^2 \int_{C_{0n}} \left[\frac{\partial^2 p_\theta}{\partial \theta^2}\right]_{\theta=\xi} \mathrm{d}P +$$

$$+ 2\tau_n d_{0n} \int_{C_{0n}} \mathrm{d}P_0 + O(n^{-1}), \quad \xi \in (-\tau_n, \tau_n).$$

Because of AS 5 we have

$$(2\tau_n)^2 \int \left[\frac{\partial^2 p_\theta}{\partial \theta^2}\right]_{\theta=\xi} \mathrm{d}P_0 = O(n^{-1}).$$

Let us recall that $p_0(x) \equiv 1$ so

$$\left[\frac{\partial p_\theta}{\partial \theta}\right]_{\theta=0} = \left[\frac{\partial p_\theta}{\partial \theta}\frac{1}{p_\theta}\right]_{\theta=0} = \left[\frac{\partial}{\partial \theta} \log p_\theta\right]_{\theta=0} = \Lambda(x).$$

Using it we may rewrite our result as follows

$$\int_{C_{0n}} [\exp\{2\tau_n d_{0n}\}\, p_{-\tau_n} - p_{\tau_n}]\, \mathrm{d}P_0 = 2\tau_n \int_{C_{0n}} (d_{0n} - \Lambda)\, \mathrm{d}P_0 + O(n^{-1}).$$

Moreover, from Lemma 1 and AS 3, it follows that

$$\int_{C_{0n}} (d_{0n} - \Lambda)\, \mathrm{d}P_0$$

is finite (at least for large n). Making use of Lemma 1 and the definitions of τ_n, ε_{jn} and δ_{jn} we may rewrite (8) into the form

$$\int_{C_{0n}} (d_{0n} - \Lambda)\, \mathrm{d}P_0 = \frac{\varepsilon_1 + \delta_1 + \delta_0}{2\tau} + O(n^{-1/2}). \tag{9}$$

Subtracting (9) from (7) we obtain

(10)

$$\int_{D_0 - C_{0n}} (d_0 - \Lambda)\, \mathrm{d}P_0 - \int_{C_{0n} - D_0} (d_{0n} - \Lambda)\, \mathrm{d}P_0 + \int_{D_0 \cap C_{0n}} (d_0 - d_{0n}) \mathrm{d}P_0 = O(n^{-1/2}).$$

Now let us split $\mathcal{N}$ into three disjoint sets

$$\mathcal{N}_1 = \{n \in \mathcal{N} : d_0 \geqq d_{0n} \;\&\; C_{0n} - D_0 \neq \emptyset\},$$

$$\mathcal{N}_2 = \{n \in \mathcal{N} : d_0 < d_{0n} \;\&\; D_0 - C_{0n} \neq \emptyset\}$$

and

$$\mathcal{N}_3 = \mathcal{N} - (\mathcal{N}_1 \cup \mathcal{N}_2).$$

Let $n \in \mathcal{N}_1$. Then $d_0 - d_{0n} \geqq 0$ and there is $x_n \in C_{0n} - D_0$, i.e. (cf. (5) and (6))

$$\mathrm{d}_{0n} \geqq (\log p_{\tau_n} - \log p_{-\tau_n})/2\tau_n$$

and

$$d_0 < \Lambda .$$

From the last three inequalities we derive

$$\Lambda(x_n) - \frac{1}{2\tau_n}\left[\log p_{\tau_n}(x_n) - \log p_{-\tau_n}(x_n)\right] > d_0 - d_{0n} \geqq 0,$$

$$\Lambda(x_n) - \left[\frac{\partial}{\partial\theta} \log p_\theta(x_n)\right]_{\theta=\xi} > d_0 - d_{0n} \geqq 0,$$

where $\xi \in (-\tau_n, \tau_n)$. Finally, because of

$$\Lambda = \left[\frac{\partial}{\partial\theta} \log p_\theta\right]_{\theta=0}$$

we get

$$\left[\frac{\partial^2}{\partial\theta^2} \log p_\theta(x_n)\right]_{\theta=\eta} \xi > d_0 - d_{0n} \geqq 0.$$

Applying AS 7 we obtain

$$\mathcal{K}\tau_n \geqq \mathcal{K}\xi > d_0 - d_{0n} \geqq 0.$$

Analogously for any $n \in \mathcal{N}_2$ we could find that

$$0 \geqq d_0 - d_{0n} > -\mathcal{K}\tau_n .$$

Now let $n \in \mathcal{N}_3$. Then either $d_0 \geqq d_{0n} \;\&\; C_{0n} - D_0 = \emptyset$ or $d_0 < d_{0n} \;\&\; D_0 - C_{0n} = \emptyset$. Let the first possibility be true. Then we have

$$\int_{D_0 - C_{0n}} (d_0 - \Lambda)\, \mathrm{d}P_0 + \int_{D_0 \cap C_{0n}} (d_0 - d_{0n})\, \mathrm{d}P_0 = O(n^{-1/2})$$

(cf. (10)). Because of nonnegativity of both integrals we have

$$\int_{D_0 \cap C_{0n}} (d_0 - d_{0n})\, \mathrm{d}P_0 = O(n^{-1/2}) .$$

The assumption that the second possibility is true leads to the same result. To conclude the proof it suffices to show that there is $a > 0$ and $n_0 \in \mathcal{N}$ so that for any $n \in \mathcal{N}$, $n \geqq n_0$

$$P_0(D_0 \cap C_{0n}) > a .$$

To do it let us define $f(x) = d_0 - \Lambda(x)$ and

$$f_n(x) = d_{0n} - \frac{1}{2\tau_n} (\log p_{\tau_n} - \log p_{-\tau_n}) .$$

Then we have

$$\lim_{n\to\infty} f_n(x) = \lim_{n\to\infty} \left\{ d_{0n} - \tfrac{1}{2} \left(\frac{\log p_{\tau_n} - \log p_0}{\tau_n} + \frac{\log p_{-\tau_n} - \log p_0}{-\tau_n} \right) \right\} =$$

$$= d_0 - \left[\frac{\partial}{\partial \theta} \log p_\theta(x) \right]_{\theta=0} = d_0 - \Lambda(x) = f(x) .$$

Let $\gamma > 0$. Having used Egorov's theorem we can choose $E \subset \Omega$ such that

$$P_0(E^c) < \gamma$$

and $f_n(x)$ converges uniformly on E to $f(x)$ for $n \to \infty$. Let

$$F_k = \{x \in \Omega : d_0 - \Lambda > 1/k\} .$$

Then

$$D_0 = \bigcup_{k=1}^{\infty} F_k , \quad F_k \subset F_{k+1}$$

and so there is $k_0 \in \mathcal{N}$ such that

$$0 \leqq P_0(D_0) - P_0(F_{k_0}) \leqq \gamma .$$

Now let us find $n_1 \in \mathcal{N}$ so that for all $n \in \mathcal{N}$, $n \geqq n_1$ and for all $x \in E$

$$|f(x) - f_n(x)| < 1/k_0 .$$

For any $n \in \mathcal{N}$, $n \geqq n_1$ and $x \in F_{k_0} \cap E$ we have

$$1/k_0 < f(x) < f_n(x) + 1/k_0 .$$

It gives

$$f_n(x) > 0$$

and so $x \in C_{0n}$, i.e. $F_{k_0} \cap E \subset C_{0n}$. Simultaneously we have $F_{k_0} \subset D_0$ and finally

$$F_{k_0} \cap E \subset D_0 \cap C_{0n} .$$

Straightforward computation yields

$$P_0(D_0 \cap C_{0n}) \geqq P_0(F_{k_0} \cap E) = P_0(F_{k_0}) - P_0(F_{k_0} \cap E^c) \geqq$$
$$\geqq P_0(D_0) - \gamma - P(E^c) \geqq P_0(D_0) - 2\gamma .$$

Because of $G(d)$ having been increasing in d, $P_0(D_0)$ is positive (see (7)). Then it is sufficient to put $\gamma < \frac{1}{2} P_0(D_0)$.

Analogously it is possible to show that $d_{1n} = d_1 + O(n^{-1/2})$. □

Proof of Theorem 2. Let us assume for simplicity that $\varepsilon_0 = \varepsilon_1$. The logarithm of the likelihood ratio $(1/2\tau_n) \log \pi_n(x)$ has all moments finite (even uniformly with respect to n) because of $(1/2\tau_n) \log \pi_n(x) \in [d_{0n}, d_{2n}]$ and Lemma 1. Taking an Edgeworth expansion of a distribution of W_n one may easy find that (B1) is fulfilled. To verify (B2)–(B4) it is necessary to study λ_n more in details. Let us write at first

$$\frac{1}{2\tau_n} \log \frac{dP_{1n}}{dP_{0n}} = \frac{1}{2\tau_n} \left[\log \frac{dP_{1n}}{dP_0} - \log \frac{dP_{0n}}{dP_0} \right] = \tag{11}$$
$$= \frac{1}{2\tau_n} [\log p_{\tau_n}(x) - \log p_{-\tau_n}(x)] =$$
$$= \frac{1}{2\tau_n} \left\{ \Lambda(x) . 2\tau_n + \left[\frac{\partial^2}{\partial \theta^2} \log p_\theta(x) \right]_{\theta = \theta'} \tau_n^2 + \left[\frac{\partial^2}{\partial \theta^2} \log p_\theta(x) \right]_{\theta = \theta''} \tau_n^2 \right\},$$

where $\theta', \theta'' \in (-\tau_n, \tau_n)$. Further let us use a partition $\mathbb{R} = (D_0 \cap C_{0n}) \cup (D_0 \div C_{0n}) \cup (D_0^c \cap C_{0n}^c \cap D_1^c \cap C_{1n}^c) \cup (D_1 \div C_{1n}) \cup (D_1 \cap C_{1n})$ and because of $D_0 \cap D_1 = \emptyset$ and $C_{0n} \cap C_{1n} = \emptyset$, this decomposition consists of the disjoint sets. So we have

$$IC^*(x) - \frac{1}{2\tau_n} \log \pi_n(x) = \begin{cases} d_0 - d_{0n} & x \in D_0 \cap C_{0n} , \\ d_0 - \dfrac{1}{2\tau_n} \log \pi_n(x) & x \in D_0 \cap C_{0n}^c , \\ \Lambda(x) - d_{0n} & x \in D_0^c \cap C_{0n} , \\ \left[\dfrac{\partial^2}{\partial \theta^2} \log p_\theta(x) \right]_{\theta = \theta'} 2\tau_n & x \in D_0^c \cap C_{0n}^c \cap C_{1n}^c \cap D_1^c , \\ \Lambda(x) - d_{1n} & x \in D_1^c \cap C_{1n} , \\ d_1 - \dfrac{1}{2\tau_n} \log \pi_n(x) & x \in D_1 \cap C_{1n}^c , \\ d_1 - d_{1n} & x \in D_1 \cap C_{1n} . \end{cases} \tag{12}$$

Let $x \in D_0 \cap C_{0n}^c$. Then we may distinguish the following cases:

(i) $d_0 \leqq d_{0n}$, then

$$\Lambda(x) < d_0 \leqq d_{0n} \leqq \frac{1}{2\tau_n} \log \pi_n(x) = \frac{1}{2\tau_n} \log \frac{dP_{1n}}{dP_{0n}}$$

and so

$$\left|d_0 - \frac{1}{2\tau_n} \log \pi_n\right| \leqq \left|\Lambda(x) - \frac{1}{2\tau_n} \log \frac{\mathrm{d}P_{1n}(x)}{\mathrm{d}P_{0n}(x)}\right|,$$

(ii) $d_0 > d_{0n}$ then

$$\left|d_0 - \frac{1}{2\tau_n} \log \pi_n\right| \leqq \max\left\{\left|\Lambda(x) - \frac{1}{2\tau_n} \log \frac{\mathrm{d}P_{1n}(x)}{\mathrm{d}P_{0n}(x)}\right|,\ d_0 - d_{0n}\right\}.$$

Using (11) we derive

$$\left|d_0 - \frac{1}{2\tau_n} \log \pi_n\right| \leqq \max\left\{\sup_{|\theta|<\tau}\left|\frac{\partial^2}{\partial\theta^2} \log p_\theta(x)\right| \tau_n,\ d_0 - d_{0n}\right\}.$$

From AS 7 and the previous lemma we have

$$\sup_{x\in D_0\cap C_{0n}{}^c}\left|d_0 - \frac{1}{2\tau_n} \log \pi_n(x)\right| = O(n^{-1/2}).$$

Analogous relation may be proved for $\Lambda(x) - d_{0n}$ on $D_0^c \cap C_{0n}$, for $\Lambda(x) - d_{1n}$ on $D_1^c \cap C_{1n}$ and for $d_1 - (1/2\tau_n) \log \pi_n$ on $D_1 \cap C_{1n}^c$. From it and from (12) we may find that

$$IC^*(x) = \frac{1}{2\tau_n} \log \pi_n(x) + \frac{\xi_n(x)}{\sqrt{n}},$$

where $\xi_n(x)$ is a bounded random variable. So we have the difference λ_n of W_n and Z_n, where W_n is given in (3) and Z_n in (14), in the form

$$\lambda_n = \frac{1}{n} \sum_{j=1}^{n} \xi_{nj}(x_j),$$

where ξ_{nj} are n i.i.d. bounded (uniformly in n) random variables. Hence we may write

$$\lambda_n = \frac{1}{\sqrt{n}} S_n,$$

where

$$S_n = \frac{1}{\sqrt{n}} \sum_{j=1}^{n} \xi_{nj}(x_j).$$

Now let $A > 0$. Then

$$\mathsf{E}_{0n}|\lambda_n|\, 1_{(\gamma n^{-1/4}, A)}(|\lambda_n|) \leqq A \,.\, Q_{0n}^{\otimes n}\left(\frac{1}{\sqrt{n}} S_n \geqq \gamma n^{-1/4}\right) = A \,.\, Q_{0n}^{\otimes n}(S_n \geqq \gamma n^{1/4}).$$

Now one may take the Edgeworth expansion of an appropriate order to estimate $Q_{0n}^{\otimes n}(S_n \geqq \gamma n^{1/4})$. (The order should be such that a rest will be $o(n^{-1/2})$.) Since $\mathsf{E}\lambda_n$ is of order $O(n^{-1/2})$, $\limsup_{n\to\infty} \mathsf{E}S_n < \infty$. The leading term of the Edgeworth expansion will be a normal distribution and e.g. by means of an upper bound for this probability (see [1], Theorem 1.1) one may find that this probability is of order $o(n^{-1/2})$. Similarly

$$Q_{0n}^{\otimes n}(\lambda_n \geqq A) = o(n^{-1/2})$$

and

$$Q_{1n}^{\otimes n}(\lambda_n \leqq -A) = o(n^{-1/2})\,.$$

So (B2)–(B4) are fulfilled and we may use the Bickel-Chibisov-van Zwet theorem for $v_n = n^{-1/2}$ and it means in a usual terminology second order efficiency of ψ_n. $\square$

5. APPROXIMATION OF ERROR PROBABILITIES OF ROBUST TESTS

As it was said in the introduction, to work out the possibility of the robust testing means not only to establish the most powerful test in its analytic form, but also to find a way how to approximate its distribution, to be able to assess the critical values. One of a classical possibility to approximate distribution function of a sum of i.i.d. random variables is to use an Edgeworth (or a saddle-point) expansion Naturally it is necessary to check the reliability of such expansion by simulation.

In what follows the numerical results of such approximation are presented and checked by simulation for one symmetric distribution, represented by the normal law, and for one asymmetric – the one-gamma distribution.

In many papers in which a simulation is used to illustrate the properties of statistics or tests, the characterizations of the source of (pseudo) random numbers are omitted or, at the best, only a brief remark that such and such source of random numbers is included into the software of a given computer is made. Such remark is considered to imply that the source was tested by the producers. But their tests might be irrelevant to the purpose for which the author of paper has used the given source. So to offer the reader the possibility to make himself an idea about the trustworthiness of the illustrative example it is useful to display numerical results produced by the used source of the random numbers in a situation analogous to that just checked, but for which the simulated values may be simultaneously analytically (i.e. precisely) evaluated. (It is done in the first two rows in Tables 1 and 2.)

Now it seems to be useful to say a few words about the values of parameters of contamination which were chosen in the next examples. Let us take $\varepsilon_0 = \varepsilon_1 = \cdot 05$. This value represents a rather great homogenity of data (see [3]). To assign a value to δ, one may proceed as follows. Wishing to fulfill a recommendation for χ^2-test to have $np_i \geqq 5$ and assuming to posses 30 observations, one is lead to the conclusion to take the partition with $p_i \geqq \frac{1}{6}$. For the normal law the shortest interval having this probability is approximately $[-\cdot 2, \cdot 2]$. So it turns out, in some sense, as useless to measure with a greater precision than $\pm\cdot 2$. On the other hand, it seems unreasonable not to use the information carried by the observations measured by a more precise scale. So we may accept a compromise and decide to measure with precision $\cdot 1$.

(At a first glance it may look a very low precision, but realizing that the bulk of probability of the normal law lies in the interval $[-3, 3]$, we shall find that we have 60 points to assign an observed value to; so at least, after 30 observations, half of them will be left "empty".) Performing the observations with precision ·1, we shall round-up at the worst case about ·05. An interval with the length ·05 has probability (for $N(0, 1)$) not greater than ·02. But the observations may be contaminated (in the Huber sense) and so, let us put $\delta_0 = \delta_1 = \cdot 025$.

In the sequel we shall assume to be faced (for some finite sizes of samples) with two testing problems. In the first one the hypothesis is equal to $N(0, 1)$ and the alternative to $N(1, 1)$ and in the second one the hypothesis is represented by $G(1, 3)$ and the alternative by $G(2, 3)$, where $G(a, p)$ is a gamma distribution with the density

$$g(a, p, x) = \frac{a^p}{\Gamma(p)} x^{p-1} \exp\{-ax\}.$$

Rieder's model of contamination is applied.

Let the tests used in the following text be constructed to minimize the sum of the error probabilities, i.e. the tests of the form

$$\sum_{i=1}^{n} \log \frac{dQ_1(x_i)}{dQ_0(x_i)} \gtreqless 0$$

Table 1. Normal model.

n	5	10	15	20	25	30	35	40	45	50
PNT	·13178	·05692	·02640	·01267	·00621	·00309	·00155	·00078	·00040	·00020
PNS	·12256	·04564	0·2769	·01077	·00461	·00308	·00154	·00051	·00030	·00010
PCA	·25442	·17635	·12794	·09498	·07153	·05440	·04168	·03211	·02485	·01930
PCS	·25692	·17025	·11846	·09128	·07026	·05539	·04103	·03231	·02051	·01641

(Letters N and C on the second position of *PNT*, *PCA*, etc. in the tables indicate that the values are given for the normal and contaminated model, respectively. Similarly, letters T, S and A on the last position of *PNT*, *PCA*, etc. are related to the theoretical, simulated and approximated values, respectively.)

Table 2. Gamma model.

n	5	10	15	20	25	30	35	40	45	50
PGT	·18337	·05919	·02072	·00753	·00279	·00100	·00049	·00014	·00006	·00002
PGS	·18700	·05850	·02400	·00800	·00150	·00050	·00050	·00000	—	—
PCA	·40220	·24814	·16095	·10711	·07244	·04953	·03415	·02369	·01652	·01156
PCS	·41650	·22900	·15200	·10100	·06950	·04150	·02850	·01850	·01500	·00550

(Naturally letter G stands now instead of N to remember that the values were computed for the gamma model.)

will be considered. (The error probabilities of such tests are easy to simulate and it was one of the reasons why these tests have been chosen.) The next two tables show the possibility of approximation (of the sum) of the error probabilities by the second order Edgeworth expansion. The simulation results were obtained performing 2000 samples of the given size n.

Having confirmed trustworthiness of the second order Edgeworth expansion for approximation of the error probabilities of the robust tests we may use it for numerical studies of problems of robust testing which are interesting and important from practical point of view. We shall do it for the following three situation:

(i) behaviour of the robust test under the assumption that one of the initial (non-contaminated) probability measures is true,

(ii) study of reliability of the local alternative setting approximations of error probabilities,

(iii) influence of a biased estimation of the level of contamination on the error probabilities of the robust test.

This will be presented in the next section.

6. NUMERICAL STUDY OF BEHAVIOUR OF ROBUST TESTS

At first, behaviour of the robust test, when the assumption about contamination is false, is studied. An application of the robust tests is a superfluous ensurance against a possible contamination in this situation and we suffer a loss not using, the most powerful (likelihood ratio) test.

The next tables numerically describe the just mentioned situation. We assume that the contamination model is true. Then we use the LFP-test with critical value ensuring that this test will have the size given on the upper margin of the tables.
The II type error probabilities then will be such as presented in the second rows. If our assumption is false and the noncontaminated model will be true the error probabilities of LFP-test (with just prescribed critical value) will change to the corresponding ones exposed in the third and fourth rows. All values were obtained by the second order Edgeworth expansion. The Greek letters α and β denote the probability of the error of the I type and of the II type, respectively. Indices P and Q added to α and β indicate that the values are taken with respect to the one of non-contaminated probabilities ("centre" of hypothesis or alternative generated by contamination) or with respect to the one of distributions from LFP.

To get an idea about the loss caused by an application of the robust test instead of the most powerful (likelihood ratio) test an additional line is given. It collects second type error probabilities β^* of the most powerful test, level of which, with respect to noncontaminated hypothesis, is equal to α_Q (see the upper margin of tables).

Further application of Edgeworth approximation of the error probabilities of the

Table 3. Normal model.

					$\alpha_Q = \cdot05$					
n	5	10	15	20	25	30	35	40	45	50
β_Q	·6563	·4277	·2714	·1683	·1025	·0614	·0362	·0212	·0122	·0070
α_P	·0240	·0163	·0118	·0088	·0067	·0052	·0041	·0033	·0026	·0021
β_P	·4901	·2277	·1002	·0422	·0172	·0068	·0026	·0010	·0004	·0001
β^*	·2773	·0646	·0129	·0023	·0004	$6 . 10^{-5}$	$9 . 10^{-6}$	10^{-6}	$2 . 10^{-7}$	$2 . 10^{-8}$

					$\alpha_Q = \cdot01$					
n	20	30	40	45	50	55	60	65	70	75
β_Q	·4052	·1998	·0897	·0586	·0377	·0239	·0150	·0093	·0057	·0034
α_P	·0012	·0007	·0004	·0003	·0002	·0002	·0001	·0001	·0001	·00007
β_P	·1481	·0356	·0075	·0033	·0014	·0006	·0003	·0001	·00004	·00001
β^*	·0159	·00081	·00003	$6 . 10^{-6}$	10^{-6}	$2 . 10^{-7}$	$3 . 10^{-8}$	$5 . 10^{-9}$	10^{-9}	—

(Hypothesis $N(0, 1)$, alternative $N(1, 1)$, level of contamination as above)

Table 4.

					$\alpha_Q = \cdot05$					
n	5	10	15	20	25	30	35	40	45	50
β_Q	·5191	·2623	·1239	·0552	·0238	·0099	·0041	·0017	·0007	·0003
α_P	·0272	·0192	·0148	·0118	·0096	·0080	·0067	·0057	·0049	·0042
β_P	·3560	·0960	·0129	·0047	·0009	·0002	·00002	$4 . 10^{-6}$	10^{-6}	—
β^*	·1776	·0145	·0008	$3 . 10^{-5}$	10^{-6}		less than 10^{-6}			

					$\alpha_Q = \cdot01$					
n	5	10	15	20	25	30	35	40	45	50
β_Q	·8701	·5818	·3528	·2018	·1085	·0556	·0274	·0131	·0061	·0028
α_P	·0044	·0029	·0021	·0159	·0012	·0010	·0008	·0006	·0005	·0004
β_P	·7926	·3492	·1248	·0378	·0104	·0027	·0006	·0001	·00002	·000005
β^*	·4704	·0923	·0110	·0009	·0007	·000004		less than 10^{-6}		

Table 4 sums up analogous values as Table 3, but for the gamma model (for $a = 1$), as a hypothesis, and $a = 2$, as an alternative and $p = 3$, with level of contamination described above).

robust test is a study of reliability of other possibility of error probability approximation.

So the next collection of tables (5, 6) is presented to offer a possibility to compare the approximations of error probabilities yielded by the above mentioned setting namely by a classical and a local alternative ones. They also illustrate how the second order efficiency works. Let us describe the meaning of the values in tables. The size α

of tests is again given on the upper margin of tables and inside of them the II type error probabilities β's are introduced only. The first row contains probabilities β_{LFP}'s for the LFP-test, the second one the analogous values β_{IC*}'s for Rieder's asymptotically minimax test (see Section 3). To get an idea how the second order efficiency works, the percentage differences $D_1\%$ of the first and second row are given in the third one. The just mentioned two tests have, in the setting of the local alternatives, the same asymptotic distribution. The approximation of error probabilities β_{LA}'s obtained by this asymptotic distribution are gathered in the fourth row and again the percentage differences $D_2\%$ of these values and those ones from the first row are available in the fifth row. The last row sums up the approximation β_{CLT}'s for the LFP-test produced by the central limit theorem. A comparison of the first and the last row leads to the conclusion that only in a case, when rather precise values of the error probabilities are required, we have to use the Edgeworth or the saddle-point expansion and in the others we may put up with the central limit theorem approximation. (But sometimes, in a symmetric situation, it is easier to find out a saddle-point expansion than central limit theorem approximation.) The parameters of models and contamination are the same as above.

Remark 4. It may seem queer that the local alternative setting approximations of error probabilitirs worsen with increasing size n of samples. But it is necessary to realize that for every n we must find a special local alternative setting so that P_{0n} and P_{1n} as well as ε_{0n}, ε_{1n}, δ_{0n} and δ_{1n} are equal to those in our problem, in which naturally the hypothesis, the alternative and the level of contamination are fixed. E.g. if we have estimated the level of contamination $\varepsilon_0' = \varepsilon_1' = {\cdot}05$, we must put for $n = 20$ in (1) $\varepsilon_0 = \varepsilon_1 = {\cdot}05 \,.\, \sqrt{20} = {\cdot}2236$ etc. and then compute the asymptotic approximations.

The last example of utilization of the Edgeworth approximation of error probabilities is the study of influence of a wrong estimation of the level of contamination on the error probabilities of the robust test.

The next tables (7, 8) collect the following values:

We assume the level of contamination to be as is given in the first rows ($\varepsilon_0 = \varepsilon_1 = 2\delta_0 = 2\delta_1$), we construct the most powerful test for this situation and we expect that α-level test (α given on the upper margin of tables) will have the second type error probability β_E as given in the second rows. But the real level of contamination is $\varepsilon_{0R} = \varepsilon_{1R} = 2\delta_{0R} = 2\delta_{1R} = {\cdot}05$ and therefore both error probabilities of our test will be different from expected ones. They are given in the third and fourth rows (and denoted by α_R and β_R). The study was performed for three sizes of sample, namely $n = 20, 30, 40$ as is pointed also on the upper margin of the tables. Normal and Gamma model were considered with the parameters as above (see page 401).

Remark 5. It follows from the Tables 7 and 8 that small inaccuracies in an estima-

Table 5. Normal model.

$\alpha = \cdot 05$

n	5	10	15	20	25	30	35	40	45	50
β_{LFP}	·6563	·4277	·2714	·1683	·1025	·0614	·0362	·0212	·0122	·0070
β_{IC^*}	·6618	·4350	·2781	·1733	·1057	·0633	·0373	·0218	·0125	·0072
D_1 %	·84	1·71	2·48	2·94	3·15	3·17	3·04	2·82	2·54	2·23
β_{LA}	·5560	·3148	·1684	·0863	·0428	·0207	·0098	·0045	·0021	·0009
D_2 %	15·28	26·40	37·95	48·71	58·20	66·26	73·05	78·60	83·11	86·73
β_{CLT}	·6522	·4289	·2731	·1690	·1022	·0606	·0354	·0204	·0116	·0065

$\alpha = \cdot 01$

n	20	30	40	45	50	55	60	65	70	75
β_{LFP}	·4052	·1998	·0897	·0586	·0377	·0239	·0150	·0093	·0057	·0034
β_{IC^*}	·4067	·2012	·0911	·0596	0·384	·0244	·0153	·0095	·0059	·0036
D_1 %	·35	·72	1·53	1·81	2·01	2·15	2·23	2·27	2·28	2·28
β_{LA}	·2476	·0872	·0269	·0144	·0075	·0038	·0020	·00097	·00047	·00023
D_2 %	39·11	56·37	70·01	75·43	79·99	83·80	86·95	89·53	91·64	93·34
β_{CLT}	·4074	·2005	·0894	·0580	·0371	·0233	·0145	·0089	·0054	·0033

Table 6. Gamma model.

$\alpha = \cdot 05$

n	5	10	15	20	25	30	35	40	45	50
β_{LFP}	·5191	·2623	·1239	·0552	0·238	·0099	·0041	·0017	·0007	·0003
β_{IC^*}	·5339	·2704	·1285	·0578	·0249	·0105	·0043	·0018	·0007	·0003
D_1 %	2·78	3·00	3·54	4·46	5·15	5·49	5·54	5·44	5·40	5·40
β_{LA}	·0187	·0001	$7 . 10^{-7}$	$5 . 10^{-8}$	—	—	—	—	—	—
D_2 %	96·48	99·94	99·99	99·999	—	—	—	—	—	—
β_{CLT}	·5609	·2782	·1241	·0514	·0201	·0076	·0027	·0010	·0003	·0001

$\alpha = \cdot 03$

n	5	10	15	20	25	30	35	40	45	50
β_{LFP}	·8701	·5818	·3528	·2018	·1085	·0556	·0274	·0131	·0061	·0028
β_{IC^*}	·8715	·5950	·3661	·2111	·1146	·0593	·0295	·0142	·0067	·0031
D_1 %	·16	2·21	3·61	4·38	5·28	6·17	7·11	7·61	7·97	8·48
β_{LA}	·0810	·0016	·00001	$2 . 10^{-7}$	—	—	—	—	—	—
D_2 %	90·70	99·72	99·99	99·999	—	—	—	—	—	—
β_{CLT}	·8632	·6117	·3797	·2132	·1107	·0540	·0250	·0111	·0047	·0019

tion of the level of contamination lead to not significant changes of the error probabilities. On the other hand a heavy underestimation of this level may cause very unpleasant deviations of the size and the power of the test.

(Received December 2, 1982.)

Table 7. Normal model.

$\alpha = \cdot 095$ $\qquad n = 20$

ε_0	·050	·045	·040	·035	·030	·025	·020	·015	·010	·005
β_E	·097	·075	·057	·042	·030	·021	·014	·009	·005	·003
α_R	·095	·108	·122	·138	·156	·177	·202	·231	·266	·311
β_R	·097	·086	·076	·066	·058	·049	·042	·035	·029	·024

$\alpha = \cdot 050$ $\qquad n = 30$

ε_0	·050	·045	·040	·035	·030	·025	·020	·015	·010	·005
β_E	·061	·043	·030	·019	·012	·007	·004	·002	·001	·0003
α_R	·050	·059	·070	·083	·100	·119	·143	·172	·209	·259
β_R	·061	·051	·043	·036	·029	·024	·019	·015	·011	·008

$\alpha = \cdot 025$ $\qquad n = 40$

ε_0	·050	·045	·040	·035	·030	·025	·020	·015	·010	·005
β_E	·042	·027	·017	·010	·006	·003	·001	·0006	·0002	·00005
α_R	·025	·030	·038	·049	·061	·077	·097	·124	·159	·209
β_R	·042	·034	·027	·022	·017	·013	·010	·007	·005	·003

Table 8. Gamma model.

$\alpha = \cdot 050$ $\qquad n = 20$

ε_0	·050	·045	·040	·035	·030	·025	·020	·015	·010	·005
β_E	.055	·042	·029	·020	·013	·008	·004	·002	·001	·0003
α_R	·050	·055	·063	·070	·079	·090	·102	·117	·135	·151
β_R	·055	·050	·042	·037	·032	·028	·025	·022	·019	·018

$\alpha = \cdot 050$ $\qquad n = 30$

ε_0	·050	·045	·040	·035	·030	·025	·020	·015	·010	·005
β_E	·010	·006	·003	·001	·0009	·0004	·0001	·00005	10^{-5}	10^{-5}
α_R	·050	·058	·067	·077	·089	·104	·121	·138	·162	·184
β_R	·010	·008	·006	·005	·004	·003	·003	·002	·002	·002

$\alpha = \cdot 050$ $\qquad n = 40$

ε_0	·050	·045	·040	·035	·030	·025	·020	·015	·010	·005
β_E	·002	·0008	·0004	·0001	$5 \cdot 10^{-5}$	10^{-5}	$4 \cdot 10^{-6}$	10^{-6}	less than 10^{-6}	
α_R	·050	·059	·070	·083	·097	·114	·134	·157	·185	·216
β_R	·002	·001	·0009	·0007	·0005	·0003	·0002	·0002	·0002	·0002

REFERENCES

[1] R. R. Bahadur: Some Limit Theorems in Statistics. Society for Industrial and Applied Mathematics, Philadelphia, Pennsylvania 19103.

[2] P. J. Bickel, D. M. Chibisov and W. R. van Zwet: On efficiency of first and second order. Internat. Statist. Rev. *49* (1981), 169—175.

[3] J. Hájek: Sampling from Finite Population. M. Dekker, Inc., New York 1981.

[4] F. R. Hampel: The influence curve and its role in robust estimation. J. Amer. Statist. Assoc. *69* (1974), 383—393.

[5] P. J. Huber and V. Strassen: Minimax tests and Neyman-Pearson lemma for capacities. Ann. Statist. *1* (1973), 251—263.

[6] H. Rieder: Least favourable pairs for special capacities. Ann. Statist. *5* (1977), 909—912.

[7] H. Rieder: A robust asymptotic testing model. Ann. Statist. *6* (1978), 1080—1094.

[8] J. Stone: Adaptive maximum likelihood estimators of a location parameter. Ann. Statist. *3* (1975), 267—284.

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.