ZASTOSOWANIA MATEMATYKI
APPLICATIONES MATHEMATICAE
XIII, 2 (1972)

BOGUSŁAWA BEDNAREK-KOZEK and A. KOZEK (Wrocław)

ON THE ROBUSTNESS OF PROPERTIES CHARACTERIZING THE NORMAL DISTRIBUTION

1. Introduction. In 1951 Sapogov [9] proved the robustness for Cramer's theorem. Certain types of extentions of this result were investigated in [4], [10]-[13].

In 1968 Hoang Huu Nhu [1] showed that, with some supplementary assumptions,

1. ε-independence of the random variables X_1+X_2 and X_1-X_2, where X_1 and X_2 are independent, implies $c/\sqrt{\ln 1/\varepsilon}$-normality of X_1 and X_2;

2. ε-independence of

$$\bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i \quad \text{and} \quad S^2 = \sum_{i=1}^{n}(X_i-\bar{X})^2,$$

where X_i are independent with the same distribution function F, implies $c/\sqrt{\ln 1/\varepsilon}$-normality of X_i $(i = 1, \ldots, n)$.

The purpose of this paper is to prove analogous theorems:

1. for some ε-independent linear and quadratic statistics without the assumption that the random variables have finite moments,

2. for some quadratic statistics which have ε-polynomial regression on the sum of the random variables with finite variance.

We also obtain similar results in the case of Lévy's metric.

2. Notation and definitions. Let us denote by $F_X(x)$ the distribution function and by $f_X(t)$ the characteristic function of the random variable X. Further, let $\mathrm{E}X$ and $\mathrm{Var}\, X$ stand for the *mean* and *variance* of X, respectively. Moreover, let

$$M_\Theta X = \mathrm{E}|X|^{2(1+\Theta)} \quad \text{for some } \Theta \in (0, \tfrac{1}{2}].$$

Let us write

$$\varphi_{XY}(x, y) = G_{XY}(x, y) - F_X(x)F_Y(y), \tag{1}$$

where $G_{XY}(x, y)$ is the joint distribution function of (X, Y), $q(\varepsilon, a_0, f) = \sup\{c: |f_X(t)| > \varepsilon^{a_0} \text{ for all } t \in (-c, c)\}$, with $1 > \varepsilon \geqslant 0$, $a_0 > 0$ and let

$$h_\gamma(x) = \begin{cases} -\gamma & \text{if } x < -\gamma, \\ x & \text{if } -\gamma \leqslant x \leqslant \gamma, \\ \gamma & \text{if } x > \gamma, \end{cases} \tag{2}$$

where γ is a positive constant.

The definitions 1 and 2 may be found, for example, in [1]. However, for the sake of completeness, they should be included here.

Definition 1. A distribution function F is *ε-normal* if there exist m and $\sigma > 0$ such that

$$\sup_x \left| F(x) - \Phi\left(\frac{x-m}{\sigma}\right) \right| \leqslant \varepsilon,$$

where

$$\Phi(x) = (2\pi)^{-1/2} \int_{-\infty}^{x} \exp\left(-\frac{x^2}{2}\right) dx.$$

Definition 2. The random variables X and Y are said to be *ε-independent* if for every x and y

$$|\varphi_{XY}(x, y)| \leqslant \varepsilon,$$

where $\varphi_{XY}(x, y)$ is defined by (1).

Furthermore, let us introduce the following definitions.

Definition 3. We say that the random variable Y has *ε-polynomial regression of order k on the random variable X* if there exists a function $\eta(x)$ such that

$$E(Y \mid X = x) = \sum_{j=1}^{k} \beta_j x^j + \eta(x),$$

where $\beta_k \neq 0$ and $\sup_x |\eta(x)| \leqslant \varepsilon$.

Definition 4. We say that *the random variable X belongs to the family $F(M, \Theta; \varepsilon)$* if there exists $\gamma > 0$ such that γ and $-\gamma$ are the points of continuity of $F_X(x)$ and the following conditions hold:

$$F_X(-\gamma) \leqslant \varepsilon, \quad F_X(\gamma) \geqslant 1 - \varepsilon,$$
$$M^{-1} \leqslant \operatorname{Var} h_\gamma(X) \leqslant M, \quad M_\Theta\big(h_\gamma(X)\big) \leqslant M,$$

where M and Θ are constants, $M > 1$, $0 < \Theta \leqslant \frac{1}{2}$.

The constants C_i appearing in the theorems below depend on M, Θ, n but not on ε. This is the reason for considering the families $F(M, \Theta; \varepsilon)$ with M and Θ fixed.

3. Theorems. Let X_1, X_2 be random variables and let

$$S_1 = a_{11}X_1 + a_{12}X_2, \qquad S_2 = a_{21}X_1 + a_{22}X_2 \tag{3}$$

where

$$-\infty < \frac{a_{11}a_{22}}{a_{12}a_{21}} < 0. \tag{4}$$

THEOREM 1. *Let $X_1, X_2 \in F(M, \Theta, \varepsilon)$ be independent random variables such that S_1 and S_2 given by* (3) *are ε-independent. If ε is sufficiently small, then the random variables X_1, X_2 are $\delta_i(\varepsilon)$-normally distributed, where*

$$\delta_i(\varepsilon) = \frac{C_1}{\sqrt{\ln 1/\varepsilon}}. \tag{5}$$

Remark. If X_1, X_2 are non-degenerate and independent random variables, then (4) is a necessery condition for the independence of S_1 and S_2.

For the random variables $X_1, \ldots, X_n$ let us write

$$S = \sum_{j=1}^{n} X_j, \tag{6}$$

$$S_3 = \sum_{\substack{j=1\\k=1}}^{n} a_{jk}X_jX_k + \sum_{j=1}^{n} b_jX_j, \tag{7}$$

where

$$\sum_{j=1}^{n} a_{jj} \neq 0 \tag{8}$$

and

$$\sum_{j,k=1}^{n} a_{jk} = 0, \qquad \sum_{j=1}^{n} b_j = 0. \tag{9}$$

THEOREM 2. *Let $X_1, \ldots, X_n \in F(M, \Theta, \varepsilon)$ be independent and identically distributed random variables such that S and S_3 given by* (6) *and* (7) *are ε-independent. If ε is sufficiently small, then the random variables $X_1, \ldots, X_n$ are $\delta(\varepsilon)$-normally distributed, where*

$$\delta(\varepsilon) = \frac{C_2}{\sqrt{\ln 1/\varepsilon}}. \tag{10}$$

Remark. If non-degenerate independent random variables $X_1, \ldots, X_n$ are identically normally distributed, then by virtue of the theorem of Laha [3] $S(X_1, \ldots, X_n)$ and $S_3(X_1, \ldots, X_n)$ are independent if and only if the statistics $S_3(X_1-\lambda, \ldots, X_n-\lambda)$ and $S_3(X_1, \ldots, X_n)$ are identically distributed for each $\lambda \in R$. The assumptions (9) are the necessery condition for this case.

Let us remark that if the random variables U and V are ε-independent, then the random variables U and V^2 are 2ε-independent. Hence and from Theorem 2 we obtain

COROLLARY TO THEOREM 2. *Let*

$$S_4 = \sum_{j=1}^{n} a_j X_j,$$

where

$$\sum_{j=1}^{n} a_j^2 \neq 0, \quad \sum_{j=1}^{n} a_j = 0.$$

Let $X_1, \ldots, X_n \in F(M, \Theta, 2\varepsilon)$ *be independent and identically distributed random variables such that* $S = \sum_{j=1}^{n} X_j$ *and* S_4 *are* ε-*independent. If* ε *is sufficiently small, then the random variables* $X_1, \ldots, X_n$ *are* $\delta(\varepsilon)$-*normally distributed, where* $\delta(\varepsilon)$ *is given by* (10) *with the constant* C_3 *instead of* C_2.

THEOREM 3. *Let* $X_1, \ldots, X_n$ *be independent and identically distributed random variables with finite variance and let*

$$S_5 = \sum_{j,k=1}^{n} a_{jk} X_j X_k + \sum_{j=1}^{n} b_j X_j.$$

If the statistic S_5 *has* ε-*polynomial regression of order* $\leqslant 2$ *on* S $\left(S = \sum_{j=1}^{n} X_j\right)$ *and if the regression coefficients satisfy conditions*

$$n\beta_2 \neq \sum_{j=1}^{n} a_{jj}, \quad n^2\beta_2 = \sum_{j,k=1}^{n} a_{jk}, \quad n\beta_1 = \sum_{j=1}^{n} b_j, \tag{11}$$
$$\frac{\beta_0}{\sum_{j=1}^{n} a_{jj} - n\beta_2} > 0,$$

then, for ε *sufficiently small, there exists a constant* C_4 *such that the random variables* $X_1, \ldots, X_n$ *are* $C_4/\sqrt{\ln 1/\varepsilon}$-*normally distributed.*

Remark. Lukacs and Laha [6] proved the following theorem: Under the assumption that $X_1, \ldots, X_n$ are independently and identically distributed random variables, they are non-degenerate and normally distributed if and only if the statistic S_5 has polynomial regression of order $\leqslant 2$ on S and the regression coefficients satisfy conditions (11).

4. Lemmas. In order to prove the theorems from Section 3 we use the well-known method [7] consisting in deduction of certain differential equations for characteristic functions and the method due to Hoang Huu Nhu [1] which allows us to use Esseen's theorem ([5], 20.3A).

To prove the theorems of Section 3 we need the following lemmas.

LEMMA 1. *Let $X_1, \ldots, X_n$ be independently distributed random variables and let γ_i and $-\gamma_i$ ($\gamma_i > 0, i = 1, \ldots, n$) be the points of continuity of $F_{X_i}(x)$ ($i = 1, \ldots, n$) such that*

$$F_{X_i}(-\gamma_i) \leqslant \varepsilon \quad \text{and} \quad F_{X_i}(\gamma_i) \geqslant 1-\varepsilon.$$

If the statistics $T_1(X_1, \ldots, X_n)$ and $T_2(X_1, \ldots, X_n)$ are ε-independent and if

$$Y_i = h_{\gamma_i}(X_i) \quad (i = 1, \ldots, n), \tag{12}$$

then

I. *Y_i ($i = 1, \ldots, n$) are independent random variables,*

II. *Y_i ($i = 1, \ldots, n$) have finite moments of all orders,*

III. $\sup_x |F_{X_i}(x) - F_{Y_i}(x)| \leqslant \varepsilon$ ($i = 1, \ldots, n$),

IV. *the statistics $T_1(Y_1, \ldots, Y_n)$ and $T_2(Y_1, \ldots, Y_n)$ are $(6n+1)\varepsilon$-independent.*

Proof. Properties I, II and III of the variables Y_i ($i = 1, \ldots, n$) follow immediately from (12) and from the definition (2) of the function h_{γ_i}. Thus, in order to prove Lemma 1, it is enough to prove property IV.

Let

$$A = \{(a_1, \ldots, a_n);\ -\gamma_i < a_i < \gamma_i,\ i = 1, \ldots, n\}, \qquad A' = R^n \setminus A,$$
$$B = B_x = T_1^{-1}[(-\infty, x)], \qquad D = D_y = T_2^{-1}[(-\infty, y)],$$

where $T_i^{-1}[W]$ ($i = 1, 2$) denote the inverse image of the set $W \subset R$.

Let P_X and P_Y be the joint distributions of the variables $X_1, \ldots, X_n$ and $Y_1, \ldots, Y_n$, respectively.

From the definition (2) of h_{γ_i}, we have

$$P_X(A') \leqslant 2n\varepsilon, \qquad P_Y(A') \leqslant 2n\varepsilon.$$

Moreover, by virtue of the definitions of P_X and P_Y, for each Borelian set $W \subset R^n$ we have

$$P_Y(W) = P_Y(W \cap A) + P_Y(W \cap A'),$$
$$P_Y(W \cap A) = P_X(W) - P_X(W \cap A').$$

Hence we obtain

$$|\varphi_{T_1T_2}(x, y)| = |P_Y(B \cap D) - P_Y(B)P_Y(D)| \leqslant (6n+1)\varepsilon.$$

Thus the statistics T_1 and T_2 are $(6n+1)\varepsilon$-independent, which completes the proof of Lemma 1.

LEMMA 2. *If the function $f(t)$ can be expressed by a formula*

$$f(t) = \exp\left\{-\frac{1}{2}\sigma^2 t^2 + imt + \int_0^t \int_0^\tau K(u)\,du\,d\tau\right\}$$

for $t \in [-q, q]$, where $q = q(\varepsilon, a_0, f)$, $\varepsilon > 0$, $a_0 > 0$ and $|K(t)|$ *is a function bounded by c for $t \in [-q, q]$, then*

$$q \geqslant \sqrt{\frac{2a_0}{\sigma^2 + c}} \sqrt{\ln\frac{1}{\varepsilon}}.$$

LEMMA 3. *Let U, V be random variables $a\varepsilon$-independent and let $\varphi(u, v) = \varphi_{UV}(u, v)$ be the function defined by* (1). *If*

$$J_1 = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} v e^{itu} d\varphi(u, v),$$

$$J_2 = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} iuve^{itu} d\varphi(u, v),$$

$$J_3 = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} v^2 e^{itu} d\varphi(u, v),$$

then for each $N \geqslant 1$ and for each t

(13) $$|J_1| \leqslant 8Na\varepsilon(1 + N|t|) + 2N^{-\Theta}(M_\Theta U + 2E|V|^{1+\Theta}),$$

(14) $$|J_1| \leqslant 8Na\varepsilon(1 + N|t|) + 2N^{-\Theta}(M_\Theta U + 2M_\Theta V),$$

(15) $$|J_2| \leqslant 8N^2 a\varepsilon(2 + N|t|) + 2N^{-\Theta}\left(M_\Theta U + M_\Theta V + \sqrt{(M_\Theta U)(M_\Theta V)}\right),$$

(16) $$|J_3| \leqslant 8N^2 a\varepsilon(1 + N|t|) + 2N^{-\Theta}(M_\Theta U + 2M_\Theta V).$$

Integrating by parts we obtain after some calculations inequalities (13)-(16) (see [1], [8]).

LEMMA 4. *If Y_1 and Y_2 are independently distributed random variables, then*

(17) $$M_\Theta(Y_1 + aY_2) \leqslant 2(1 + |a|)^3 \max_{j=1,2}(M_\Theta Y_j).$$

If $Y_1, \ldots, Y_n$ are independent and identically distributed random variables and if

$$\max\{|a_{11}|, |a_{12}|, \ldots, |a_{nn-1}|, |a_{nn}|, |b_1|, \ldots, |b_n|\} \leqslant 1,$$

then

$$\mathrm{E}\Big|\sum_{j,k=1}^{n} a_{jk}Y_jY_k+\sum_{j=1}^{n} b_jY_j\Big|^{1+\Theta}\leqslant 4n^4(M_\Theta Y_1+1) \tag{18}$$

and

$$M_\Theta\Big(\sum_{j=1}^{n} Y_j\Big)\leqslant n^4 M_\Theta Y_1. \tag{19}$$

Proof. We shall prove (18). One can prove the remaining inequalities in an analogous way.

Let us remark that

$$\begin{aligned}\mathrm{E}\Big|\sum_{j,k=1}^{n} a_{jk}Y_jY_k+\sum_{j=1}^{n} b_jY_j\Big|^{1+\Theta} &\leqslant n^3\mathrm{E}\Big|\max_{j=1,\ldots,n}(|Y_j|^2)+\max_{j=1,\ldots,n}(|Y_j|)\Big|^{1+\Theta}\\ &\leqslant 4n^3\mathrm{E}(\max\{|Y_1|^{2(1+\Theta)},\ldots,|Y_n|^{2(1+\Theta)},|Y_1|^{1+\Theta},\ldots,|Y_n|^{1+\Theta}\})\\ &\leqslant 4n^3\mathrm{E}(\max\{|Y_1|^{2(1+\Theta)},\ldots,|Y_n|^{2(1+\Theta)},1\})\leqslant 4n^4(M_\Theta Y_1+1),\end{aligned}$$

which completes the proof of Lemma 4.

LEMMA 5. *If $f(t)$ is the characteristic function of the distribution function $F(x)$ and if for each $t\in[-Q,Q]$*

$$f(t)=\exp\Big\{-\tfrac{1}{2}\sigma^2t^2+imt+\int_0^t\int_0^\tau K(u)\,du\,d\tau\Big\} \tag{20}$$

and the function $K(t)$ satisfies the condition

$$|K(t)|\leqslant c<\sigma^2 \quad \textit{for each } t\in[-Q,Q], \tag{21}$$

then

$$\sup_x\Big|F(x)-\Phi\Big(\frac{x-m}{\sigma}\Big)\Big|\leqslant\frac{1}{\pi}\frac{c}{\sigma^2-c}+\frac{24}{\pi}\frac{1}{\sqrt{2\pi}\,\sigma Q}.$$

Proof. Let

$$L=L(t)=|f(t)-\exp(-\tfrac{1}{2}t^2\sigma^2+imt)| \quad \text{for each } t\in[-Q,Q].$$

By the inequality $|e^z-1|\leqslant|z|\exp|z|$ we obtain

$$\begin{aligned}L&=\exp(-\tfrac{1}{2}t^2\sigma^2)\Big|\exp\Big(\int_0^t\int_0^\tau K(u)\,du\,d\tau\Big)-1\Big|\\ &\leqslant\Big|\int_0^t\int_0^\tau K(u)\,du\,d\tau\Big|\exp\Big(-\tfrac{1}{2}t^2\sigma^2+\Big|\int_0^t\int_0^\tau K(u)\,du\,d\tau\Big|\Big)\\ &\leqslant\frac{ct^2}{2}\exp\big(-\tfrac{1}{2}t^2(\sigma^2-c)\big).\end{aligned}$$

Hence it follows that

$$\int_0^Q \left| \frac{f(t)-\exp(-\frac{1}{2}t^2\sigma^2+imt)}{t} \right| dt \leqslant \int_0^Q \frac{1}{2} ct\exp\big(-\tfrac{1}{2}t^2(\sigma^2-c)\big)\, dt \leqslant \frac{1}{2}\,\frac{c}{\sigma^2-c}.$$

Thus, from Esseen's theorem ([5], 20.3.A) we have

$$\sup_x \left| F(x)-\Phi\left(\frac{x-m}{\sigma}\right)\right| \leqslant \frac{1}{\pi}\,\frac{c}{\sigma^2-c}+\frac{24}{\pi}\,\frac{1}{\sqrt{2\pi}\,\sigma Q},$$

which completes the proof of Lemma 5.

5. Proofs of the theorems. By ε-independence of the statistics $S_1(X_1, X_2)$ and $S_2(X_1, X_2)$ and by virtue of Lemma 1 it follows that there exist independent random variables Y_i $(i=1,2)$ such that Y_i have finite moments of all orders, $S_1(Y_1, Y_2)$ and $S_2(Y_1, Y_2)$ are 13ε-independent statistics and $\sup_x|F_{X_i}(x)-F_{Y_i}(x)| \leqslant \varepsilon$ $(i=1,2)$.

If Y_i $(i=1,2)$ are $c_5/\sqrt{\ln 1/\varepsilon}$-normal, then there exists a constant C_1 such that X_1, X_2 are $C_1/\sqrt{\ln 1/\varepsilon}$-normal.

Let us show that there exists a constant c_5 such that Y_1, Y_2 are $c_5/\sqrt{\ln 1/\varepsilon}$-normal. Let us remark that multiplication of random variables by numbers not equal to zero preserves their properties to be ε-independent and ε-normally distributed. Therefore it is enough to show that if

$$U = Y_1+Y_2, \qquad V = Y_1+aY_2, \tag{22}$$

where $a<0$ and U, V are 13ε-independently distributed statistics, then there exists a constant c_5 such that Y_1, Y_2 are $c_5/\sqrt{\ln 1/\varepsilon}$-normal. First we need the following

LEMMA 6. *Let U, V given by* (22) *be 13ε-independently distributed statistics and let $f_i(t)=f_{Y_i}(t)$ $(i=1,2)$. If*

$$q_1 = q\left(\varepsilon, \frac{\Theta}{20}, f_1\cdot f_2\right)$$

and

$$Q_1 = \min(q_1, \varepsilon^{-1/5}), \tag{23}$$

then there exists a function $K_2(t)$ such that for each $t\in[-Q_1, Q_1]$

$$|K_2(t)| \leqslant c_6\cdot\varepsilon^{\Theta/10} \tag{24}$$

and

$$f_2(t) = \exp\Big(-\tfrac{1}{2}\sigma_2^2t^2+im_2t+\int_0^t\int_0^\tau K_2(u)\,du\,d\tau\Big),$$

where

$$\sigma_2^2 = \frac{\mathrm{Var}(Y_1 + aY_2)}{a(a-1)}, \qquad m_2 = \mathrm{E}\, Y_2.$$

Proof. Let $\varphi(u, v) = \varphi_{U,V}(u, v)$ and

$$(25)\qquad \begin{aligned} A_1 &= A_1(t) = \mathrm{E}[(Y_1 + aY_2)^2 \exp\{it(Y_1 + Y_2)\}], \\ A_2 &= A_2(t) = \mathrm{E}[(Y_1 + aY_2)\exp\{it(Y_1 + Y_2)\}]. \end{aligned}$$

By independence of the random variables Y_1, Y_2 we have

$$A_1 = -f_1''(t)f_2(t) - 2af_1'(t)f_2'(t) - a^2 f_1(t)f_2''(t).$$

On the other hand, we have

$$A_1 = \mathrm{E}(V^2)f_1(t)f_2(t) + \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} v^2 e^{itu}\, d\varphi(u, v).$$

Hence we obtain

$$(26)\qquad -R_2(t) = f_1''(t)f_2(t) + a^2 f_1(t)f_2''(t) + 2af_1'(t)f_2'(t) + \mathrm{E}(V^2)f_1(t)f_2(t),$$

where

$$R_2(t) = \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} v^2 e^{itu}\, d\varphi(u, v).$$

Similarly, it follows from (25) that

$$(27)\qquad f_1'(t)f_2(t) + af_1(t)f_2'(t) = i\mathrm{E}(V)f_1(t)f_2(t) + \bar{R}_2(t),$$

where

$$\bar{R}_2(t) = \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} ive^{itu}\, d\varphi(u, v).$$

After the differentiation of (27) we obtain from (26) that for every $t \in [-q_1, q_1]$

$$[\ln f_2(t)]'' = \frac{-\mathrm{Var}\, V}{a(a-1)} + \frac{H_2(t)}{f_1(t)f_2^2(t)},$$

where

$$(28)\qquad H_2(t) = \frac{f_2(t)}{a(a-1)}\left[\mathrm{E}(V)\bar{R}_2(t) - i\bar{R}_2'(t) - i(a-1)\frac{\bar{R}_2(t)f_2'(t)}{f_2(t)} - R_2(t)\right].$$

Hence, by using the facts that

$$a < 0, \qquad f_2'(0) = im_2, \qquad \frac{\mathrm{Var}(V)}{a(a-1)} = \sigma_2^2,$$

we obtain by integration that

$$f_2(t) = \exp\left(-\tfrac{1}{2}\sigma_2^2t^2+im_2t+\int_0^t\int_0^\tau \frac{H_2(u)}{f_1(u)f_2^2(u)}\,dud\tau\right)$$

for every $t\in[-q_1, q_1]$.

Let

$$K_2(t) = \frac{H_2(t)}{f_1(t)f_2^2(t)}.$$

We shall show that there exists a constant c_6' such that

$$|K_2(t)| \leqslant c_6'\max(E|Y_1|, E|Y_2|, 1)\left[1+\max_{i=1,2}\left(M_\Theta(Y_i)\right)\right]\varepsilon^{\Theta/10}$$

for every $t\in[-Q_1, Q_1]$.

Since

$$\text{(30)}\qquad \mathrm{E}|V| \leqslant \mathrm{E}|Y_1|-a\mathrm{E}|Y_2| \quad \text{and} \quad |f_2'(t)| \leqslant E|Y_2|$$

we see from (28) that

$$|H_2(t)| \leqslant \frac{1}{a(a-1)}\left[\left(\mathrm{E}|Y_1|+(1-2a)\mathrm{E}|Y_2|\right)|\bar{R}_2(t)|+|\bar{R}_2'(t)|+|R_2(t)|\right].$$

By 13ε-independence of U, V and by virtue of (30), Lemma 3 ((14), (16)) and Lemma 4 ((17)) we obtain

$$|H_2(t)| \leqslant \max(E|Y_1|, E|Y_2|, 1)\{104N\varepsilon(1-a)[2+2N|t|+3N+2N^2|t|]+ + 4N^{-\Theta}[\max_{j=1,2}(M_\Theta Y_j)][32(1-a)+11(1-a)^4]\}\frac{1}{a(a-1)}$$

for every $t\in[-q_1, q_1]$.

Substituing N instead of $|t|$ and then $\varepsilon^{-1/5}$ instead of N we obtain

$$|H_2(t)| \leqslant \max(E|Y_1|, E|Y_2|, 1)[936(1-a)+ + \max_{i=1,2}(M_\Theta Y_i)172(1-a)^4]\frac{1}{a(a-1)}\varepsilon^{\Theta/5}.$$

It follows from (23) that there exists a constant c_6' such that for each $t\in[-Q_1, Q_1]$

$$|K_2(t)| \leqslant c_6'\max(\mathrm{E}|Y_1|, \mathrm{E}|Y_2|, 1)[1+\max_{i=1,2}(M_\Theta Y_i)]\cdot\varepsilon^{\Theta/10},$$

hence there exists a constant c_6 such that for $t\in[-Q_1, Q_1]$ formula (24) holds. This completes the proof of Lemma 6.

We turn now to the proof of Theorem 1. Let us exchange Y_1 for Y_2 and let us put a^{-1} instead of a. Then, by Lemma 6, there exists a function

$K_1(t)$ and a constant c_7 such that for every $t \in [-Q_1, Q_1]$

(31) $$|K_1(t)| \leqslant c_7 \varepsilon^{\Theta/10}$$

and

(32) $$f_1(t) = \exp\left(-\tfrac{1}{2}\sigma_1^2 t^2 + i m_1 t + \int_0^t \int_0^\tau K_1(u)\,du\,d\tau\right),$$

where

$$\sigma_1^2 = \frac{\operatorname{Var} V}{1-a}, \quad m_1 = \mathrm{E}\, Y_1.$$

It follows from Lemma 6 and from formulas (31) and (32) that the assumptions of Lemma 2 are satisfied if $f(t) = f_1(t) f_2(t)$, $a_0 = \Theta/20$ and $q = Q_1$. Hence, we have

$$q_1 \geqslant \left[\frac{10}{\Theta}\left(\sigma_1^2 + \sigma_2^2 + (c_6 + c_7)\cdot \varepsilon^{\Theta/10}\right)\right]^{-1/2} \cdot \sqrt{\ln \frac{1}{\varepsilon}}\,.$$

By virtue of (20) and (24) we obtain that for ε sufficiently small condition (21) is satisfied. Thus, Lemmas 5 and 6, and formulas (31) and (32) imply that there exists a constant c_5 such that in both cases $Q_1 = q_1$ and $Q_1 = \varepsilon^{-1/5}$ the inequality

$$\left|F_{Y_i}(x) - \Phi\left(\frac{x - m_i}{\sigma_i}\right)\right| \leqslant \frac{c_5}{\sqrt{\ln 1/\varepsilon}} \qquad (i = 1,2)$$

holds for each $x \in R$.

This completes the proof of Theorem 1.

Proof of Theorem 2. It follows from Lemma 1 that in order to prove Theorem 2 it is enough to show that there exists a constant c_8 such that if $T_1 = S$, $T_2 = S_3$ and Y_i $(i = 1, 2, \ldots, n)$ are defined by (12), then $Y_1, \ldots, Y_n$ are $c_8/\sqrt{\ln 1/\varepsilon}$-normally distributed random variables.

Let us assume $\varphi(u, v) = \varphi_{SS_3}(u, v)$, $f(t) = f_{Y_1}(t)$ and $A_3 = A_3(t) = \mathrm{E}\left(S_3 \exp\,(itS)\right)$.

The independence of the random variables Y_i $(i = 1, \ldots, n)$ and properties (8) and (9) of S_3 imply that

$$A_3 = -\sum_{j=1}^{n} a_{jj}\left(\ln f(t)\right)'' f^n(t) \quad \text{and} \quad A_3 = \operatorname{Var}(Y_1) \sum_{j=1}^{n} a_{jj} f^n(t) + R(t),$$

where

$$R(t) = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} v e^{itu}\, d\varphi(u, v).$$

Hence if $\mathrm{Var}(Y_1) = \sigma^2$, $E(Y_1) = m$ and $q_2 = (\varepsilon, \Theta/8n, f)$, then

$$[\ln f(t)]'' = -\sigma^2 - \frac{1}{\sum_{j=1}^{n} a_{jj}} \frac{R(t)}{f^n(t)} \quad \text{for each } t \in [-q_2, q_2].$$

Thus for each $t \in [-q_2, q_2]$

$$f(t) = \exp\left\{ -\tfrac{1}{2}t^2\sigma^2 + imt - \int_0^t \int_0^\tau \frac{R(u)}{\sum_{j=1}^{n} a_{jj} f^n(u)}\, du\, d\tau \right\}.$$

Without loss of generality we may assume that

$$\max\{|a_{11}|, \ldots, |a_{nn}|, |b_1|, \ldots, |b_n|\} \leqslant 1.$$

Since $S(Y_1, \ldots, Y_n)$, $S_3(Y_1, \ldots, Y_n)$ are $(6n+1)$ ε-independently distributed statistics, we see from Lemma 3 (13) and Lemma 4 ((18), (19)) that for every $t \in [-q_2, q_2]$

(35) $$|R(t)| \leqslant 8N(6n+1)\varepsilon(1+N|t|) + 18n^4 N^{-\Theta} M_\Theta(Y_1) + 16n^4 N^{-\Theta}.$$

Let

(36) $$N = \varepsilon^{-1/4}, \quad Q = \min(q_2, \varepsilon^{-1/4}), \quad K_3(t) = \frac{R(t)}{\sum_{j=1}^{n} a_{jj} f^n(t)},$$

$$c_9 = \frac{16(6n+1) + 18n^4}{\left|\sum_{j=1}^{n} a_{jj}\right|}.$$

It follows from (33)-(36) that for every $t \in [-Q_2, Q_2]$

(37) $$|K_3(t)| \leqslant c_9\big(1 + M_\Theta(Y_1)\big)\varepsilon^{\Theta/8}.$$

Moreover, (33), (34), (37) and Lemma 2 imply that

$$q_2 \geqslant \left[\frac{4n}{\Theta}\Big(\sigma^2 + c_9\big(1 + M_\Theta(Y_1)\big)\varepsilon^{\Theta/8}\Big)\right]^{-1/2} \sqrt{\ln\frac{1}{\varepsilon}}.$$

Since condition (21) is satisfied for sufficiently small values of ε, we deduce from Lemma 5 that there exists a constant c_8 such that in both cases, $Q_2 = q_2$ and $Q_2 = \varepsilon^{-1/4}$, the inequality

$$\sup_x \left| F_{Y_1}(x) - \Phi\left(\frac{x-m}{\sigma}\right)\right| \leqslant \frac{c_8}{\sqrt{\ln 1/\varepsilon}}$$

holds.

This completes the proof of Theorem 2.

Proof of Theorem 3. To prove Theorem 3 it is enough to show that the assumptions of Lemma 5 are satisfied.

Since the statistics S_5 has ε-polynomial regression of order $\leqslant 2$ on S, we have

$$\mathrm{E}\Big(\sum_{j,k=1}^{n} a_{jk}X_jX_k+\sum_{j=1}^{n} b_jX_j \,|\, \sum_{j=1}^{n} X_j = x\Big) = \beta_0+\beta_1 x+\beta_2 x^2+\eta(x)$$

and

$$\sup_x|\eta(x)| \leqslant \varepsilon.$$

If one multiplies the above expression by $\exp(itx)$ and takes the expectation, then

$$\mathrm{E}\Big[\Big(\sum_{j,k=1}^{n} a_{jk}X_jX_k+\sum_{j=1}^{n} b_jX_j\Big)\exp\Big(it\sum_{j=1}^{n}X_j\Big)\Big]$$
$$= \beta_0\mathrm{E}\Big[\exp\Big(it\sum_{j=1}^{n}X_j\Big)\Big]+\beta_1\mathrm{E}\Big[\sum_{j=1}^{n}X_j\exp\Big(it\sum_{j=1}^{n}X_j\Big)\Big]+$$
$$+\beta_2\mathrm{E}\Big[\Big(\sum_{j=1}^{n}X_j\Big)^2\exp\Big(it\sum_{j=1}^{n}X_j\Big)\Big]+\mathrm{E}\Big[\eta\Big(\sum_{j=1}^{n}X_j\Big)\exp\Big(it\sum_{j=1}^{n}X_j\Big)\Big].$$

We put $f(t) = f_{X_1}(t)$, $q_3 = q(\varepsilon, 1/2n, f)$,

$$\sigma^2 = \frac{\beta_0}{\sum_{j=1}^{n} a_{jj}-n\beta_2} \quad \text{and} \quad \eta_1(t) = \mathrm{E}\Big[\eta\Big(\sum_{j=1}^{n}X_j\Big)\exp\Big(it\sum_{j=1}^{n}X_j\Big)\Big].$$

It follows from (11) that for every $t\in[-q_3, q_3]$

$$[\ln f(t)]'' = -\sigma^2-\frac{\eta_1(t)}{\Big(\sum_{j=1}^{n} a_{jj}-n\beta_2\Big)f^n(t)}.$$

By integration we have for each $t\in[-q_3, q_3]$

$$f(t) = \exp\Big(-\tfrac{1}{2}\sigma^2t^2+imt-\int_0^t\int_0^\tau \frac{\sigma^2}{\beta_0}\frac{\eta_1(u)}{f^n(u)}\,du\,d\tau\Big).$$

Since for each $t\in[-q_3, q_3]$

$$\left|\frac{\sigma^2}{\beta_0}\,\frac{\eta_1(t)}{f^n(t)}\right| \leqslant c_{10}\varepsilon^{1/2},$$

where $c_{10} = |\sigma^2/\beta_0|$, the assumptions of Lemma 5 are satisfied. Thus, there exists a constant C_4 such that

$$\sup_x\left|F_{X_1}(x)-\Phi\Big(\frac{x-m}{\sigma}\Big)\right| \leqslant \frac{C_4}{\sqrt{\ln 1/\varepsilon}}.$$

The proof of Theorem 3 is therefore completed.

6. The case of Lévy's metric. $L(F_1, F_2)$ is said to be the *distance of the distribution functions* F_1, F_2 *in Lévy's metric*, i.e. the metric defined by the formula

$$L(F_1, F_2) = \inf\{h\colon F_1(x-h)-h \leqslant F_2(x) \leqslant F_1(x+h)+h \text{ for all } x\}.$$

It is well-known that the convergence of distribution functions in Lévy's metric is equivalent to the weak convergence.

In [13] Zolotarev proved that for any distribution functions F_1, F_2 and for every number $Q > e$

$$L(F_1, F_2) \leqslant \frac{1}{\pi}\int_0^Q \frac{|f_1(t)-f_2(t)|}{t}\,dt + 2e\frac{\log Q}{Q},$$

where $f_i(t)$ are the characteristic functions of F_i $(i = 1,2)$.

Using this theorem instead of Esseen's one we obtain Theorems 1′-3′.

Let K_N be the family of distribution functions of the normal non-degenerate law and let

$$\varrho_N(X) = \inf_{F \in K_N}\{L(F_X, F)\}.$$

THEOREM 1′. *Under the assumptions of Theorem* 1

$$\varrho_N(X_i) \leqslant C_{11}\frac{\log\log\frac{1}{\varepsilon}}{\sqrt{\log\frac{1}{\varepsilon}}} \quad (i = 1,2).$$

THEOREM 2′. *If the assumptions of Theorem* 2 *are satisfied, then*

$$\varrho_N(X_i) \leqslant C_{12}\frac{\log\log\frac{1}{\varepsilon}}{\sqrt{\log\frac{1}{\varepsilon}}} \quad (i = 1, \ldots, n).$$

COROLLARY. *If the assumptions of the Corollary to Theorem* 2 *are satisfied, then*

$$\varrho_N(X_i) \leqslant C_{13}\frac{\log\log\frac{1}{\varepsilon}}{\sqrt{\log\frac{1}{\varepsilon}}} \quad (i = 1, \ldots, n).$$

THEOREM 3′. *If the assumptions of Theorem 3 are satisfied, then*

$$\varrho_N(X_i) \leqslant c_{14} \frac{\log\log \frac{1}{\varepsilon}}{\sqrt{\log \frac{1}{\varepsilon}}} \qquad (i = 1, \ldots, n).$$

As the proofs of the above theorems are very similar to the proofs given in Section 5, they are omitted.

Acknowledgment. We wish to thank Dr. W. Klonecki for his helpful discussions.

References

[1] Хоанг Хыу Ньы, *Об устойчивости некоторых характеристических свойств нормальной совокупности*, Теория вероятн. прим. 13 (1968), p. 308-314.

[2] — *Оценка устойчивости одной характеризации экспоненциального закона*, Лит. мат. сборник 8 (1968), p. 175-177.

[3] R. G. Laha, *On some properties of the normal and gamma distributions*, Proc. Amer. Math. Soc. 7 (1956), p. 172-174.

[4] Ю. В. Линник, *Общие теоремы о разложении безгранично делимых законов*, Теория вероятн. прим. 3 (1958), p. 3-40.

[5] M. Loève *Probability theory*, Van Nostrand, New York 1955.

[6] E. Lukacs, R. G. Laha, *On some characterization problems connected with quadratic regression*, Biometrica 47 (1960), p. 335-343.

[7] — *Applications of characteristic functions*, Griffin, London 1964.

[8] L. D. Meshalkin, *On the robustness of some characterizations of the normal distribution*, Ann. Math. Statist. 39 (1968), p. 1747-1750.

[9] Н. А. Сапогов, *Проблема устойчивости для теоремы Крамера*, Изв. АН СССР, сер. матем. 15 (1951), p. 205-218.

[10] — *О независимых слагаемых суммы случайных величин распределенной приближенно нормально*, Вестник Ленингр. Чнив. 19 (1959), p. 78-105.

[11] О. В. Шалаевский, *О устойчивости теоремы Д. А. Райкова*, Вестник Ленингр. Унив. 14 (1959), p. 41-49.

[12] В. М. Золотарев, *К вопросу об устойчивости разложения нормального закона распределения на компоненты*, Теория вероятн. прим. 13 (1968), p. 738-742.

[13] — *Несколько новых неравенств связанных с метрикой Леви*, Докл. АН СССР 190 (1970), p. 1019-1021.

MATHEMATICAL INSTITUTE
POLISH ACADEMY OF SCIENCES

Received on 15. 6. 1971;
revised version on 10. 2. 1972

BOGUSŁAWA BEDNAREK-KOZEK i A. KOZEK (Wrocław)

O STABILNOŚCI WŁASNOŚCI CHARAKTERYZUJĄCYCH ROZKŁAD NORMALNY

STRESZCZENIE

Jeśli $X_1, \ldots, X_n$ są niezależnymi zmiennymi losowymi, to ze stochastycznej niezależności lub z wielomianowej regresji jednej statystyki względem drugiej w wielu wypadkach wynika, że rozkłady zmiennych losowych $X_1, \ldots, X_n$ są już określone.

W pracy tej badamy stabilność pewnych własności charakteryzujących rozkład normalny. Przez zastąpienie niezależności statystyk słabszym założeniem ich ε-niezależności otrzymujemy charakteryzację rozkładów wyjściowych zmiennych losowych, jako „bliskich" odpowiednim rozkładom normalnym. Jako miary odległości między dwoma dystrybuantami F_1, F_2 używamy metryki $\varrho(F_1, F_2) = \sup_x |F_1(x) - F_2(x)|$ oraz metryki Lévy'ego.