The Annals of Mathematical Statistics
1968, Vol. 39, No. 5, 1747–1750

ON THE ROBUSTNESS OF SOME CHARACTERIZATIONS OF THE NORMAL DISTRIBUTION

BY L. D. MESHALKIN

University of Moscow and University of California, Berkeley

1. Introduction. Let us introduce some definitions.

DEFINITION 1. Two distribution functions F and G are ϵ-coincident if

$$\sup_x |F(x) - G(x)| \leqq \epsilon.$$

DEFINITION 2. A distribution function F is ϵ-normal if there exist $a > 0$ and b such that

$$\sup_x |F(x) - \Phi(ax + b)| \leqq \epsilon,$$

where $\Phi(x)$ is $(2\pi)^{-\frac{1}{2}} \int_{-\infty}^{x} \exp\{-x^2/2\}\, dx$.

DEFINITION 3. Two random variables η and ζ are ϵ-independent if for every a, b, c, d, e, f

$$|\int_{ay+bz<c, dy+ez<f} dQ(y, z)| \leqq \epsilon, \tag{1}$$

where

$$Q(y, z) = P\{\eta < y, \zeta < z\} - P\{\eta < y\}P\{\zeta < z\}. \tag{2}$$

In 1956 N. A. Sapogov (Leningrad) [3] showed, that if $F_3 = F_1 * F_2$ is ϵ-normal, and if $F_1(0) = \frac{1}{2}$,

$$\int_{-N}^{N} x\, dF_1 = a_1, \qquad \int_{-N}^{N} x^2\, dF_1(x) - a_1^2 = \sigma_1^2 > 0, \qquad N = (2 \log (1/\epsilon))^{\frac{1}{2}} + 1,$$

then

$$\sup_x |F_1(x) - \Phi((x - a_1)/\sigma_1)| < C\sigma_1^{-3}(\log (1/\epsilon))^{-\frac{1}{2}}.$$

This study was continued by Hoang Hiu Nye (Moscow) [2] who showed in 1966 that, with some supplementary assumptions,

(a) ϵ-independence of the random variables of $\xi + \eta$ and $\xi - \eta$, where ξ and η are independent, implies $\beta_1(\epsilon)$-normality of the ξ and η;

(b) ϵ-independence of

$$\bar{\xi} = \sum \xi_i/n \quad \text{and} \quad S^2 = \sum (\xi_i - \bar{\xi})^2,$$

where the ξ_i are independent and have the same distribution function F, implies $\beta_2(\epsilon)$-normality of F. In his theorems the $\beta(\epsilon)$ are of the order of

$$(\log (1/\epsilon))^{-\frac{1}{2}}.$$

The purpose of this paper is to show that in some cases we can obtain a much better order of magnitude of the $\beta(\epsilon)$.

Received 19 April 1967.

THEOREM 1. *If* ξ_1 *and* ξ_2 *are independent identically distributed and* $E\xi_i = 0$, $E\xi_i^2 = 1$, $E|\xi_i|^3 < M < \infty$, *then* ϵ*-independence of* $\xi_1 + \xi_2$ *and* $\xi_1 - \xi_2$ *implies* $C_1(M)\epsilon^{\frac{1}{3}}$*-normality of* ξ_i.

THEOREM 2. *If* ξ_1 *and* ξ_2 *are independent identically distributed and* $E\xi_i = 0$, $E\xi_i^2 = 1$, $E|\xi_i|^3 < M < \infty$, *then* ϵ*-coincidence of the distribution functions of* $(\xi_1 + \xi_2)/2^{\frac{1}{2}}$ *and* ξ_i *implies* $C_2(M)\epsilon^{\frac{1}{3}}$*-normality of* ξ_i.

2. Proof. As the proofs of both theorems are very similar we shall prove only the first of them. All constant C_i $(i = 1, \cdots, 8)$ will be functions of M. Assume also that ϵ is so small that $\epsilon < C_2^3 \cdot C_8^{-4}$, where C_2 and C_8 will be defined by (7) and (20), and that inequality (18) holds.

Let $\eta = \xi_1 + \xi_2$, $\zeta = \xi_1 - \xi_2$ and $Q(y, z)$ is defined by (2); then we have that $\varphi(t)$, the characteristic function of ξ_i, satisfies the equation

$$(3) \quad \varphi(2t) = E \exp\{it(\eta + \zeta)\} = E \exp\{it\eta\} \cdot \mathrm{E} \exp\{it\zeta\} + f(t) = \varphi^3(t)\varphi(-t) + f(t),$$

where
$$f(t) = \int \exp\{it(y + z)\}\, dQ(y, z).$$

Our purpose now is to estimate $|f(t)|$. Let $x_1 = (y + z)/2$, $x_2 = (y - z)/2$ and

$$(4) \qquad R(x_1, x_2) = \int_{y+z<2x_1,\, y-z<2x_2} dQ(y, z);$$

then

$$f(t) = \int \exp\{2itx_1\}\, dR(x_1, x_2) = \int \exp\{2itx\}\, dS(x) = \int_{|x| \leqq a} + \int_{|x|>a} = \mathfrak{J}_1 + \mathfrak{J}_2,$$

where $S(x) = R(x, \infty) - R(x, -\infty)$. From (1) and (4) we obtain that $|S(x)| \leqq 2\epsilon$. Now we can estimate $|\mathfrak{J}_1|$ by the integration by parts

$$(5) \quad |\mathfrak{J}_1| = |S(a)\exp\{2ita\} - S(-a)\exp\{-2ita\} - 2it\int_{-a}^{a} S(x)\exp\{2itx\}\, dx| \leqq 4\epsilon(1 + 2ta)$$

It follows from the Tchebysheff's inequality for 3rd moments that

$$(6) \qquad |\mathfrak{J}_2| \leqq C_1/a^3.$$

Let now $a = \epsilon^{-\frac{1}{3}}$, then from (5) and (6) we have

$$(7) \qquad |f(t)| \leqq C_2\epsilon(1 + \epsilon^{-\frac{1}{3}}t).$$

Let

$$(8) \qquad \varphi(t) = \exp\{-t^2/2\} + h(t),$$

then according to Essen's well-known theorem (see, p. 196–197 of [1]), in order to prove the desired result it is enough to show that for one $T \geqq \epsilon^{-\frac{1}{3}}$

$$(9) \qquad U \equiv \int_{-T}^{T} |h(t)/t|\, dt \leqq C_3\epsilon^{\frac{1}{3}}.$$

Let $t_i = \epsilon^{\frac{1}{3}} 2^i$, $\epsilon^{-\frac{1}{3}} \leqq t_k < 2\epsilon^{-\frac{1}{3}}$,

$$\gamma_i = \max_{t_{i-1} \leqq t \leqq t_i} |h(t)| \qquad (i = 1, 2, \cdots, k),$$

then

$$\int_{t_{i-1}}^{t_i} |h(t)/t| \, dt \leqq \int_{t_{i-1}}^{t_i} (\gamma_i/t_{i-1}) \, dt = \gamma_i . \tag{10}$$

The conditions on the moments of the ξ give us an estimate of $|h(t)|$ for small t

$$|h(t)| \leqq C_4 |t|^3,$$

then

$$\int_0^{t_0} |h(t)/t| \, dt \leqq C_4 t_0^3 = C_4 \epsilon. \tag{11}$$

Since $|h(t)| = |h(-t)|$ from (9)–(11) we obtain

$$U \leqq 2C_4\epsilon + 2 \sum_{i=1}^{k} \gamma_i . \tag{12}$$

From (3) and (8) we have inequality

$$|h(2t)| \leqq \sum_{l=0}^{3} \binom{4}{l} \exp\{-lt^2/2\} |h(t)|^{4-l} + |f(t)|. \tag{13}$$

Then from (13) and (7)

$$\gamma_{i+1} \leqq 4a_i\gamma_i(1 + 1.5\gamma_i + \gamma_i^2) + \gamma_i^4 + C_2\epsilon(1 + 2^i), \tag{14}$$

where $$a_i = \exp\{-t_{i-1}^2/2\}.$$

Let us show that if for all $i \leqq l$ $(l \leqq k - 1)$ we have

$$\gamma_i^4 \leqq C_2\epsilon(1 + 2^i), \tag{15}$$

then (15) holds for $i = l + 1$ too.

From (15) it follows that for $i \leqq l < k$

$$1.5\gamma_i + \gamma_i^2 < C_5 \epsilon^{1/12} = \delta.$$

Repeating the inequality (14) l times we obtain

$$\gamma_{l+1} < 2C_2\epsilon(1 + 2^l) + 4a_l\gamma_l(1 + \delta) < 2C_2\epsilon(1 + 2^l) + 2C_2\epsilon(1 + 2^{l-1}) \cdot 4a_l(1 + \delta)$$
$$+ 4^2 a_l a_{l-1}(1 + \delta)^2 < \cdots < 2C_2\epsilon \sum_{j=1}^{l-1} (1 + 2^{l-j}) 4^j (1 + \delta)^j \Pi_j$$
$$+ 4^l(1 + \delta)^l \Pi_l \gamma_1 ,$$

where $\Pi_0 = 1$ and

$$\Pi_j = a_l a_{l-1} \cdots a_{l-j+1} = \exp\{-\sum_{m=1}^{j} t_{l-m}^2/2\} = \exp\{-t_0^2(4^l - 4^{l-j})/6\}.$$

$$\gamma_{l+1} < 4C_2\epsilon(1 + \delta)^l \sum_{j=0}^{l-1} 2^{l+j}\Pi_j + 4^l(1 + \delta)^l \Pi_l \gamma_1 \tag{16}$$
$$< (4C_2\epsilon + \gamma_1)(1 + \delta)^l \sum_{j=0}^{l} 2^{l+j}\Pi_j .$$

It is not difficult to calculate that

$$\sum_{j=0}^{l} 2^{l+j}\Pi_j < 2^{2l+1} \exp\{-t_0^2 \, 2^{2l-3}\} < C_6 \epsilon^{-\frac{2}{3}}. \tag{17}$$

According to our assumption about ϵ

$$(1 + \delta)^k \leqq 2. \tag{18}$$

Then from (14) it follows that

$$\gamma_1 \leqq C_7\epsilon. \tag{19}$$

Therefore from (16)–(19) we have

$$\gamma_{l+1} < 2(4C_2 + C_7)C_6\epsilon^{\frac{1}{3}} = C_8\epsilon^{\frac{1}{3}}. \tag{20}$$

From (20) and the assumption about ϵ it follows (15) for $i = l + 1$. Since we have shown that (15) is valid for all $l \leqq k$, we can use (16) to estimate $\sum \gamma_i$. According to (16), (17), (18), (19)

$$\sum_{i=1}^{k} \gamma_i < 2(4C_2 + C_7)\epsilon \sum_{l=1}^{k} 2^{2l+1} \exp\{-t_0^2 2^{2l-3}\} < C_3 t_0^{-2}\epsilon = C_3\epsilon^{\frac{1}{3}}.$$

By (12) the inequality (9) is proved.

REFERENCES

[1] Gnedenko, B. V. and Kolmogorov, A. N. (1954). *Limit Distributions for Sums of Independent Random Variables*. Addison-Wesley, Cambridge.

[2] Nye, Hoang Hiu (1966). On the stability of certain theorems characterizing normal distribution. International Congress of Mathematicians, Moscow 16–26 August 1966, Information Bulletin No. 6, 5.

[3] Sapogov, N. A. (1956). Stability problem for Cramér's theorem (Russian). *Leningrad Univ. Vestnik* No. II, 60–64.