Z. Wahrscheinlichkeitstheorie verw. Geb. 8, 332–342 (1967)

A Sharpening of the Inequality of Berry-Esseen

V. M. ZOLOTAREV

Received August 27, 1966

1. Introduction

K. ESSEEN's inequality [4] which is well known in probability theory is an extremely important instrument for the investigation of various asymptotic properties in the uniform metric. The same is true for the inequality of A. BERRY [2] which applies to the same situation and was presented some years before. ESSEEN's inequality is easier to apply (and therefore used more frequently) and is, in addition, applicable to a broader class of cases. At the same time experimental work with ESSEEN's inequality (especially in the problem of absolute estimation of the remainder in A. M. LJAPUNOV's theorem) showed that BERRY's inequality is sharper than ESSEEN's.

This leads to the attempt to construct an analogue of the inequalities of BERRY and ESSEEN which on one hand is more economical than BERRY's inequality, and on the other hand does not yield to ESSEEN's inequality as to scope and ease of applicability. The present paper deals with the solution of this problem.

In how far the new inequality is really sharper than BERRY's inequality can be judged from the results of its application (in connection with a new method) to the above mentioned classical problem of estimation of the remainder in LJAPUNOV's theorem.

Let $\xi_1, \dots, \xi_n$ be independent random variables with mean zero and finite absolute third moments $\beta_1, \dots, \beta_n$. By $\sigma_1^2, \dots, \sigma_n^2$ we denote the variances and by $\sigma^2 = \sum_{k=1}^{n} \sigma_k^2$ and $\varepsilon = \left(\sum_{k=1}^{n} \beta_k\right) / \sigma^3$ Ljapunov's quantities.

We form the normed sum $\zeta = (\xi_1 + \cdots + \xi_n)/\sigma$ and let $F(x)$ denote the distribution function of ζ and $\Phi(x)$ the distribution function of the standard normal law. According to LJAPUNOV's theorem there exists a minimal numerical constant C such that

$$\sup_x |F(x) - \Phi(x)| \leqq C\varepsilon .$$

The true value of C has not been found so far, but there is reasonto expect that

$$C = C^* = \frac{\sqrt{10} + 3}{6\sqrt{2\pi}} = 0.40974 .$$

The best two-sided estimate of C so far has been the following: $C \geqq C^*$ (K. ESSEEN [4]), $C < 4.8$ (G. BERGSTRÖM [1]). In the particular case of identically distributed terms ξ_k the upper estimate of C has been improved: $C < 2.031$ (K. TAKANO [6]) ·

An improved version of the inequality of BERRY-ESSEEN permitted to lower the upper estimate of C (see [7]). In fact we could show in the general case that $C < 1.322$, and in the above mentioned particular case that $C < 1.301$.

We can now lower these estimates even more on account of sharper estimates of the characteristic functions. We formulate here the latest result; the proofs are based on methods from [7] and will be published elsewhere.

Theorem 1. *In the general case we can state that*

$$C < 0.9051\,,$$

and in the case of identically distributed terms

$$C < 0.82\,.$$

2. The Basic Theorem

Let $L(x)$ and $H(x)$ be functions of bounded variation defined on the entire real axis and such that they simultaneously are either right continuous or left continuous, and $l(t)$, $h(t)$ their Fourier-Stieltjes transforms. We put

$$\Delta = \sup_x |L(x) - H(x)|\,.$$

Our aim is to construct upper estimates of Δ in terms of l and h.

We will use the following functions and qualities: $p(x)$ is the density of an absolutely continuous symmetric distribution (which we can choose arbitrarily) with a characteristic function $\omega(t)$ which is absolutely integrable over the entire real axis;

$$m(t) = l(t) - h(t)\,.$$

For positive values of x and y we put

$$V(x) = x \int\limits_{|u| < x} p(u)\,du\,, \quad Q(y) = \frac{y}{2\pi} \int\limits_{-\infty}^{\infty} \left| \omega\left(\frac{t}{y}\right) \frac{m(t)}{t} \right| dt;$$

α is the unique positive zero of the function $2V(x) - x$.

Let $\mathfrak{S}_1$, $\mathfrak{S}_2$ be two sets of the real line. We complete the set $\mathfrak{S} = \mathfrak{S}_1 \cup \mathfrak{S}_2$ by the points $-\infty$ and ∞ and form the quantity

$$\beta(\mathfrak{S}_1, \mathfrak{S}_2) = \inf(a'' - a') \tag{1}$$

where the greatest lower bound is taken with respect to all possible pairs of points $a' < a''$ from the completed set $\mathfrak{S}$.

When constructing the inequality we are aiming at we will use additional conditions on the functions L and H. These conditions will be selected from the following three.

Condition $\mathbf{A_1}$. We denote by $\mathfrak{A}_1$ the set of all points of discontinuity of the function $H(x)$ and by $\mathfrak{B}_1$ its complement. Then $H(x)$ has a derivative in all points of $\mathfrak{B}_1$, and

$$q_H = \sup_{\mathfrak{B}_1} |H'(x)| < \infty\,.$$

Condition $\mathbf{A_2}$. We denote by $\mathfrak{A}_2$ the set of all points of discontinuity of the function $L(x)$ and by $\mathfrak{B}_2$ its complement. Then $L(x)$ has a derivative in all points of $\mathfrak{B}_2$, and

$$q_L = \sup_{\mathfrak{B}_2} |L'(x)| < \infty .$$

Condition A_3. The function $L(x)$ is monotone. In this case we formally put $q_L = 0$ and take for $\mathfrak{A}_3$ the empty set.

In the following the set $\mathfrak{A}_3$ will play in condition A_3 the same role as the sets of points of discontinuity did in the conditions A_1, A_2.

Theorem 2. *Suppose that A_1 and one condition A_j, $j = 2, 3$, are fulfilled. Accordingly we form from the sets $\mathfrak{A}_1, \mathfrak{A}_j$ by (1) the quantity $\beta = \beta(\mathfrak{A}_1, \mathfrak{A}_j)$. We assume that $\beta > 0$. Then for all positive x, y, satisfying the requirements*

$$x > \alpha, \quad y \geqq \frac{4}{\beta} x,$$

the following inequality

$$\Delta \leqq I(x, y) = \frac{x[qV(x) + Q(y)]}{y[2V(x) - x]} \tag{2}$$

holds, where $q = q_H + q_L$.

Proof. In general, there need not exist an x for which the absolute value of the function $M(x) = L(x) - H(x)$ takes its extremum Δ. However for arbitrary $\delta > 0$ there is a point x_δ such that $|M(x_\delta)| \geqq \Delta - \delta$. Without restricting the generality we can assume that

$$M(x_\delta) \leqq -\Delta + \delta .$$

Let $\mathfrak{A}_1$ and $\mathfrak{A}_j$ be the sets corresponding to the pair of conditions satisfied. We denote by $\mathfrak{A}$ the union of $\mathfrak{A}_1$ and $\mathfrak{A}_j$, completed by $-\infty$ and ∞. We set

$$a = \sup\{a' : a' \in \mathfrak{A},\ a' \leqq x_\delta\},$$
$$b = \inf\{a'' : a'' \in \mathfrak{A},\ a'' > x_\delta\}.$$

Suppose that the function $L(x)$ does not decrease. Then it is not hard to verify that, whichever condition A_2 or A_3 we choose together with A_1, the following inequality

$$M(x_\delta - \gamma + \nu) \leqq M(x_\delta) + (\gamma - \nu) q \leqq -\Delta + \delta + (\gamma - \nu) q \tag{3}$$

will be satisfied for any non-negative finite number $\gamma \leqq b - x_\delta$ and all ν with $|\nu| \leqq \gamma$.

Likewise, when using together with A_1 one of the conditions A_2 and A_3, where the function $L(x)$ does not increase, we obtain

$$M(x_\delta + \gamma - \nu) \leqq M(x_\delta) + (\gamma - \nu) q \leqq -\Delta + \delta + (\gamma - \nu) q \tag{4}$$

for any non-negative finite $\gamma \leqq x_\delta - a$ and all ν with $|\nu| \leqq \gamma$.

Therefore, under the conditions of the theorem at least one of the inequalities (3) and (4) will be satisfied for all finite γ from the interval

$$0 < \gamma \leqq \tfrac{1}{4}\beta . \tag{5}$$

Suppose that the inequality (4) is fulfilled. We choose any γ from the intervall (5) and any $y > \alpha/\gamma$. We multiply both members of (4) by $y p(y\nu)$ and integrate the inequality thus obtained with respect to ν over the intervall $|\nu| \leqq \gamma$. Taking into account that p is an even function we obtain

$$\int_{|\nu|\leqq\gamma} M(x_\delta+\gamma-\nu)\,y\,p(y\nu)\,d\nu \leqq (-\Delta+\delta+\gamma q)\int_{|\nu|\leqq\gamma} y\,p(y\nu)\,d\nu\,.$$

This inequality, in turn, permits to obtain the following one:

$$\lambda = \Big|\int_{-\infty}^{\infty} M(x_\delta+\gamma-\nu)\,y\,p(y\nu)\,d\nu\Big| \geqq \Big|\int_{|\nu|\leqq\gamma}\Big| - \Big|\int_{|\nu|>\gamma}\Big|\Big| \geqq$$
$$\geqq |\Delta-\delta-\gamma q|\int_{|\nu|\leqq\gamma y} p(\nu)\,d\nu - \Delta\int_{|\nu|>\gamma y} p(\nu)\,d\nu\,.$$

Next two cases are possible: $\Delta \leqq \delta+\gamma q$, or $\Delta > \delta+\gamma q$. We consider the second case only. We have

$$\lambda \geqq (2\Delta-\gamma q-\delta)\frac{V(\gamma y)}{\gamma y} - \Delta\,.$$

Hence we find that

$$\Delta \leqq \frac{(q\gamma+\delta)\,V(\gamma y)+\lambda\gamma y}{2\,V(\gamma y)-\gamma y}\,. \tag{6}$$

Since $V(x) < x$ for every $x > 0$, the right-hand member of the inequality (6) obviously is not smaller than $\delta+\gamma q$. Therefore the estimate (6) remains true in both cases mentioned before. We introduce a new variable $x = \gamma y$. Then the inequality (6) takes the form (2). Also the conditions $0 < \gamma \leqq \frac{1}{4}\beta$, $y > \alpha/\gamma$ turn out to be equivalent to $x > \alpha$, $y \geqq \frac{4}{\beta}x$.

3. Choice of the Density Function $p(x)$

In concrete problems the function $|m(t)|$ can usually be estimated in a non-trivial way only in some interval $|t| < y$. It is then advantageous to choose for the density function $p(x)$ a function whose characteristic function $\omega(t)$ vanishes outside of the interval $|t| \leqq 1$.

We indicate here two classes of such distributions. One of them is a subclass of the class of Pólya distributions and can be described as follows.

Lemma 1. *Each function $\omega(t)$, which is defined on the entire real axis and has the properties*

1. $\omega(t) = \omega(-t) \geqq 0\,,$
2. $\omega(0) = 1;\ \omega(t) = 0\,,\quad t \geqq 1\,,$
3. $\omega\left(\frac{t'+t''}{2}\right) \geqq \frac{1}{2}\{\omega(t')+\omega(t'')\}$ *for all non-negative t', t'',*

is a characteristic function of an absolutely continuous distribution.

A representative of this class is the distribution which was used by Berry, when he proved his inequality [2]:

$$p(x) = \frac{1-\cos x}{\pi x^2}\,. \tag{7}$$

For this distribution $\omega(t) = 1-|t|$ if $|t| \leqq 1$, and $\omega(t) = 0$ if $|t| > 1$.

However the class of distributions described in lemma 1 has the unpleasant property that sometimes $\omega(t)$ is not differentiable at zero. The following lemma describes a class of distributions of the type we are interested in which contains representatives for which the function ω is infinitely differentiable at zero.

Lemma 2. *Let the function $\nu(x)$ be defined on the entire real line with the property that*

1. $\nu(x) = \nu(-x)$,
2. $\nu(x) = 0$ for $x \geqq \frac{1}{2}$,
3. $\int\limits_{-\infty}^{\infty} \nu^2(x)\,dx = 1$.

Then the convolution $\omega(t) = (\nu * \nu)(t)$ is the characteristic function of an absolutely continuous distribution with density

$$p(x) = \frac{2}{\pi}\left\{\int\limits_0^{\frac{1}{2}} \cos(tx)\,\nu(t)\,dt\right\}^2.$$

Proof. We have

$$\omega(t) = (\nu * \nu)(t) = \int\limits_m^M \nu(t-y)\,\nu(y)\,dy$$

where $m = \max(-\frac{1}{2}, t-\frac{1}{2})$ and $M = \min(\frac{1}{2}, t+\frac{1}{2})$. Since the function $\omega(t)$ is even (being a convolution of even functions), we can confine ourselves to the case $t > 0$. Then obviously

$$\omega(t) = \int\limits_{t-\frac{1}{2}}^{\frac{1}{2}} \nu(t-y)\,\nu(y)\,dy. \tag{8}$$

In virtue of condition 1 and 3 of the lemma we have $\omega(0) = 1$, i.e. ω is normalized. It remains to show that $\omega(t)$ is a Fourier transform of a non-negative function $p(x)$.

Taking into account that ω is real we have

$$p(x) = \frac{1}{2\pi}\int\limits_{-\infty}^{\infty} \cos(tx)\,\omega(t)\,dt = \frac{1}{2\pi}\int\limits_{-\infty}^{\infty} \cos(tx)\,(\nu * \nu)(t)\,dt$$
$$= \frac{1}{2\pi}\left\{\int\limits_{-\infty}^{\infty} \cos(tx)\,\nu(t)\,dt\right\}^2 = \frac{2}{\pi}\int\limits_0^{\frac{1}{2}} \cos(tx)\,\nu(t)\,dt^2.$$

Condition 2 of the lemma guarantees that $\omega(t) = 0$ for $t \geqq 1$, as required.

Remark. If the conditions 1—3 of the lemma are completed by the condition

4. $\nu(x_1) \geqq \nu(x_2)$ for $x_2 \geqq x_1 \geqq 0$,

then it can be seen from the representation (8) that $\omega(t)$ will also be a non-increasing function on the half-line $t \geqq 0$.

We supplement lemma 2 by two examples of distributions from the class described in it. The first was given by ESSEEN in [3]:

a) $$p(x) = \frac{3}{8\pi}\left(\frac{\sin\frac{x}{4}}{\frac{x}{4}}\right)^4.$$

This density corresponds to the function

$$\omega(t)=\begin{cases}1-6t^2(1-|t|) & \text{if} \quad |t|\leqq\frac{1}{2},\\ 2(1-|t|)^3 & \text{if} \quad \frac{1}{2}<|t|\leqq 1,\\ 0 & \text{if} \quad |t|>1.\end{cases}$$

b) $$p(x)=\frac{16\pi^3}{3}\left\{\frac{\sin\frac{x}{2}}{\frac{x}{2}(x^2-4\pi^2)}\right\}^2.$$

This density corresponds to the function

$$\omega(t)=\begin{cases}\frac{1}{3}(1-|t|)(2+\cos(2\pi t))+\frac{1}{2\pi}\sin(2\pi|t|) & \text{if} \quad |t|\leqq 1,\\ 0 & \text{if} \quad |t|>1.\end{cases}$$

4. The Problem of Minimization of $I(x, y)$

Having constructed the inequality (2) our next problem naturally is to select x and y from the domain of its admissible values so as to give (perhaps in an asymptotic sense) the minimum of the function I which is defined by (2). In the general situation, when only little is known about the properties of the functions $p(x)$, $\omega(t)$ and $m(t)$, it is of course impossible to obtain a complete solution of this analytical problem. However, we can essentially restrict the class of possible choices by setting up a representation of the surface $z = I(x, y)$ in various, a priori admissible cases. We need the following notations:

$$R(y)=y\,Q'(y),\quad P(y)=R(y)-Q(y),\quad \mu=\frac{1}{2\pi}\int\limits_{-\infty}^{\infty}\left(\frac{m(t)}{t}\right)dt;$$

$$W(x)=[2\,V(x)-x]\,[V(x)+2x^2p(x)]/(4x^2p(x)),$$

$$U(x)=x\,W(x)/[2\,V(x)-x];$$

$\mathfrak{M}=\{(x, y): x>\alpha,\ y>0\}$ is the domain of the function $I(x, y)$; $\mathfrak{N}=\{(x, y): x>\alpha,\ y>0,\ y\geqq\frac{4}{\beta}x\}$ is the set where the inequality (2) holds; $\varrho=(x_0, y_0)$ is a point, where $I(x, y)$ attains its absolute minimum in $\mathfrak{N}$ (if such a point exists).

In the following we will exclude the case $m(t)\equiv 0$, $Q(1)=\infty$ and the case $q=0$, $\beta=\infty$ which do not interest us much.

We write $I(x, y)$ in the form

$$I(x, y)=q\frac{x}{y}\frac{V(x)}{[2\,V(x)-x]}+\frac{Q(y)}{y}\frac{x}{2\,V(x)-x}=\left(q\frac{x}{y}+\frac{Q(y)}{y}\right)\left(1+\vartheta\frac{x-V(x)}{2\,V(x)-x}\right),\quad (9)$$

where $1<\vartheta<2$. Therefore, we have in the domain $\mathfrak{M}$:

$$I(x,y)>\frac{x}{2\,V(x)-x}\left(\frac{q\alpha}{2y}+\frac{Q(y)}{y}\right).\quad (10)$$

The following fact can easily be deduced from these relations: as $y\to\infty$ and $x=0(1)$,

$$\liminf I(x, y)\geqq\mu,$$

where in the case of a finite μ the inequality is strict; as $x \to \infty$ and $y \to \infty$,

$$\liminf \left\{ I(x, y) - q \frac{x}{y} \right\} \geqq \mu ,$$

and moreover, if x and y vary in such a way that $x = o(y)$, then

$$\lim I(x, y) = \mu .$$

Let $q = 0$. Then we note that in this case the exclusions which we stipulated imply $\beta < \infty$. Therefore if $q = 0$, it is easy to see that the function $I(x, y)$ for fixed y does not increase in x. It follows that

$$\inf_{\mathfrak{N}} I(x, y) = \inf_{x > \alpha} I\left(x, \frac{4}{\beta} x\right). \tag{11}$$

If however, $q > 0$, then it may be seen from the relations (9) and (10) that the function $I(x, y)$ will increase unboundedly regardless of the way we approach the boundaries $x = \alpha$, $y = 0$ of $\mathfrak{M}$ or move along a path for which $y = O(1)$, $x \to \infty$.

Exploiting the properties mentioned before we can give the following general arguments. We remain, of course, within the framework of the conditions of theorem 2.

1. If $\beta = \infty$ (this is only possible for $q > 0$) and $\mu = \infty$, then $\mathfrak{N} = \mathfrak{M}$; the point ϱ exists and satisfies the system of equations

$$\frac{\partial I}{\partial x} = 0 , \qquad \frac{\partial I}{\partial y} = 0 . \tag{12}$$

2. If $\beta = \infty$ (and consequently $q > 0$) and $\mu < \infty$, then $\mathfrak{N} = \mathfrak{M}$. The point ϱ may or may not exist. In the former case it is a solution of the system (12). In the latter case $\inf I(x, y)$ is found as the limit at $x \to \infty$, $y \to \infty$, $x = o(y)$ and turns out to be equal to μ.

3. If $\beta < \infty$, $\mu = \infty$, then $\mathfrak{N} \subset \mathfrak{M}$ and the point ϱ exists. In the case $q = 0$ this point ϱ lies on the straight line $y = \frac{4}{\beta} x$ and is determined by the value x_0 which satisfies

$$\frac{d}{dx} I\left(x, \frac{4}{\beta} x\right) = 0. \tag{13}$$

In the case $q > 0$ the point ϱ is either an interior point of $\mathfrak{N}$, and is then a solution of the system (12), or lies on the straight line $y = \frac{4}{\beta} x$ and, therefore, satisfies (13).

4. If $\beta < \infty$, $\mu < \infty$, then $\mathfrak{N} \subset \mathfrak{M}$. Here the point ϱ may or may not exist. If it does we have the same situation as in the case $\beta < \infty$, $\mu = \infty$. If it does not we have

$$\inf_{\mathfrak{N}} I(x, y) = \lim_{x \to \infty} I\left(x, \frac{4}{\beta} x\right) = \frac{1}{4} \beta q + \mu .$$

Lemma 3. *a) The system* (12) *is equivalent to the system*

$$P(y) = q\, V(x) , \quad R(y) = q\, W(x) . \tag{14}$$

If (x_*, y_*) *is a solution of* (14), *then*

$$I(x_*, y_*) = \frac{R(y_*)}{y_*[2\, V(x_*) - x_*]} = \frac{q}{y_*} U(x_*) . \tag{15}$$

b) The equation (13) *is equivalent to the equation*

$$4x^2 p(x)\, Q\left(\frac{4}{\beta}x\right) + 2x^3 p(x)\, q = (2\,V(x) - x)\, P\left(\frac{4}{\beta}x\right). \tag{16}$$

If x_* *is a solution of* (16), *then*

$$I\left(x_*, \frac{4}{\beta}x_*\right) = \frac{\beta}{8}\left(q + \frac{P\left(\frac{4}{\beta}x_*\right)}{2x_*^2 p(x_*)}\right). \tag{17}$$

The proof of the lemma amounts to simple transformations of the equation (12) and (13) and of $I(x_*, y_*)$ which we will not indicate here.

Remark 1. When estimating Δ we usually do not deal with the function $|m(t)|$ itself, but rather with its upper estimate which is non-trivial only in some finite interval. Thus the typical cases are 1 and 3, i.e. $\mu = \infty$.

Remark 2. The quantity $I(x, y)$ which appears in the inequality (2) can be made smaller not only by the choice of x and y for an already given density $p(x)$, but also by the choice of the density itself. Which density is optimal in this sense, and whether such a universal density for all functions $|m(t)|$ exists ist not yet known. Thus all we can do is to try out particular densities and compare, e.g., the classes of distributions described above (lemmas 1, 2).

The distribution (7) of A. Berry deserves our special attention, since it is the simpliest one and most suitable for application. For this particular density we have $\alpha = 1.69957\ldots$,

$$P(y) = \frac{1}{\pi}\int_0^y |m(t)|\, dt, \quad R(y) = \frac{y}{\pi}\int_0^y \frac{|m(t)|}{t}\, dt;$$

$$V(x) = \frac{2}{\pi}\{x\,\mathrm{Si}(x) - (1 - \cos x)\};$$

$$W(x) = \frac{x\,\mathrm{Si}(x)(2\,V(x) - x)}{2(1 - \cos x)}; \quad U(x) = \frac{x^2\,\mathrm{Si}(x)}{2(1 - \cos x)},$$

where $\mathrm{Si}(x) = \int_0^x \frac{\sin u}{u}\, du$ is the integral sinus function and $U(x)$ appears in the construction of the extreme value of I in formula (15).

When solving the extremal problem completely the following two facts may be useful which hold when $p(x)$ is the Berry function:

a) the system (14) has a solution in the stripe $\alpha < x < 2\pi$. In fact since

$$W(\alpha) = 0, \quad W(2\pi) = \infty; \qquad V(\alpha) > 0, \quad V(2\pi) < \infty,$$

the zero level lines of the surfaces $z_1 = P(y) - q\,V(x)$ and $z_2 = R(y) - q\,W(x)$ necessarily intersect in this stripe. Here it may be worth while to remark that in the concrete examples considered by the author which refer to the case $\beta = \infty$, $\mu = \infty$, the stripe $\alpha < x < 2\pi$ contained only one solution of the system (14) which also provided the absolute minimum for the function $I(x, y)$.

b) Solutions of the system (12) cannot give local maxima of $I(x, y)$, for if (x_*, y_*) is any such solution, then

$$\frac{\partial^2 I}{\partial y^2}\bigg|_{\substack{x=x_*\\ y=y_*}} = \frac{x_*}{[2\,V(x_*) - x_*]}\,\frac{P'(y_*)}{y_*^2} > 0.$$

5. Asymptotic Estimates of Δ

We now consider a sequence of pairs of functions (L^*, H^*) for which $\Delta^* \to 0$; this situation is, indeed, typical for probability theory. Everything referring to the sequence of functions and their corresponding characteristics will be marked with an asteric. The question naturally arises how fast Δ^* tends to zero. Although the estimate of Δ^* in terms of the minimum of the function I formally gives us some answer to this question, the complexity of the solution of the extremal problem makes us look for simpler methods.

Here we merely analyze the case where β^* turns out to be infinite for all pairs of sequences of functions L^*, H^* in terms of which it is given. Apart from the less interesting cases mentioned in the preceding paragraph this means that $q^* > 0$. For this reason we write the inequality (2) in the following form:

$$\Delta^* \leqq I^*(x, y q^*) = \frac{x[V(x) + Q^*(y q^*)/q^*]}{y[2V(x) - x]}. \tag{18}$$

This representation of I^* makes it appear natural to impose the additional assumption

$$\frac{1}{q^*} Q^*(y q^*) \to 0 \tag{19}$$

for every fixed $y > 0$, since we want the right side member of (18) tend to zero; in fact otherwise we could in general not exploit the inequality for an estimate of the speed of convergence of Δ^* to 0.

If the condition (19) is satisfied, we can find a constant $A > 0$ and a sequence $T^* \to \infty$ such that

$$\frac{1}{q^*} Q^*(T^* q^*) \leqq A + \delta^*, \tag{20}$$

where $\delta^* \to 0$. The inequality sign here allows us to include the cases where we can choose A so as to have $\delta^* \equiv 0$.

From (18) and (19) we obtain the inequality

$$T^* \Delta^* \leqq \frac{x[V(x) + A]}{2V(x) - x} + \frac{x}{2V(x) - x} \delta^* = B_1(x) + B_2(x)\delta^*.$$

We minimize the first term by choosing x. It is not difficult to verify that the value x_0 which yields an absolute minimum of the function B_1 is a solution of the equation

$$W(x) - V(x) = A. \tag{21}$$

We now formulate the following statement.

Theorem 3. *If the sequence of pairs of functions (L^*, H^*) is such that $\beta^* = \infty$ and the condition (19) is satisfied, then by choosing $T^* \to \infty$ and $A > 0$ in accordance with (20) we have*

$$T^* \Delta^* \leqq U(x_0) + B_2(x_0)\delta^*,$$

where x_0 is a solution of the equation (21).

We illustrate the proposed method by sharpening the central limit theorem. We use the same notations as in the introduction.

Theorem 4. *In the general case of independent terms* ξ_k *we have*

$$\sup_x |F(x) - \Phi(x)| \leqq 0.81967\,\varepsilon + 0.05894\,\varepsilon^{4/3} + O(\varepsilon^{5/3}),$$

and in the case of identically distributed terms ξ_k

$$\sup_x |F(x) - \Phi(x)| \leqq 0.81967\,\varepsilon - 0.99951\,\varepsilon^2 + O(\varepsilon^3);$$

the constants are given here up to one unit of the last digit.

The proof of this theorem as well as the proof of theorem 1 will be published in a separate paper.

To conclude we give another example which in our opinion has some independent interest. In his paper [6] P. B. PATNAIK pointed out that the distribution function $F_n(x, a)$ of the non-central $\chi_n^2(a)$ can be approximated for very large or very small values of the parameter of non-centrality by the distribution function $F_s(Cx, 0)$ of the normalized central $C\chi_s^2(0)$ if

$$C = \frac{n+a}{n+2a} \quad \text{and} \quad s = \frac{(n+a)^2}{n+2a}.$$

The general method which we have proposed permits to construct a relatively simple estimate of the deviation Δ of F_n from F_s, i.e.

$$\Delta = \sup_x |F_n(x, a) - F_s(Cx, 0)|.$$

In fact, since the characteristic functions of the absolutely continuous distributions F_n, F_s are equal to

$$f_n(t, a) = (1 - 2it)^{-n/2} \exp\left\{\frac{ita}{1-2it}\right\},$$
$$f_s(t/C, 0) = (1 - 2it/C)^{-s/2},$$

respectively, it is obvious that we are dealing with the case 2 ($\beta = \infty$, $\mu < \infty$) of our classification, and therefore can take the value

$$\mu = \frac{1}{\pi} \int_0^\infty |f_n(t, a) - f_s(t/C, 0)| \frac{dt}{t}$$

as an upper estimate of Δ. Although an estimate of this integral may be cumbersome it does not present any analytical difficulty.

As a result we obtain the following absolute estimates which are true for any $a > 0$ and $n > 0$:

$$\Delta < \frac{3}{8\pi} B\left(\frac{1}{2}, \frac{n}{4}\right) \frac{a^2}{(n+a)},$$

and denoting by $D^2 = 2(n + 2a)$ the dispersion of the distribution F_n, F_s,

$$\Delta < \frac{2}{3}\sqrt{\frac{2}{\pi}}\left(1 + \frac{9}{4}k\right)^{\frac{1}{2}} (1 + 4k)^3 D^{-1} + \frac{16}{\pi}\left(1 + \frac{1}{k}\right) D^{-2} \exp\left\{-\frac{D^2}{32}\frac{k}{(1+k)}\right\} +$$
$$+ \frac{1}{\pi}\left(1 + \frac{1}{k}\right) \exp\left\{-\frac{D^2}{8}\frac{k}{(1+k)}\right\},$$

where k is an arbitrary positive number. Therefore for large D we have

$$\Delta < \frac{2}{3}\sqrt{\frac{2}{\pi}} D^{-1} + O(D^{-3}\log D).$$

In particular it follows that uniformly for all a

$$\Delta < \frac{2}{3\sqrt{\pi}} n^{-1/2} + O(n^{-3/2} \log n).$$

Bibliography

1. BERGSTRÖM, H.: On the central limit theorem in the case of not equally distributed random variables. Skand. Aktuarietidskr. **33**, 37–62 (1949).
2. BERRY, A. C.: The accuracy of the Gaussian approximation to the sum of independent variates. Trans. Amer. math. Soc. **49**, 122–136 (1941).
3. ESSEEN, C. G.: Fourier analysis of distribution functions. Acta math. **77**, 1–125 (1945).
4. — A moment inequality with an application to the central limit theorem. Skand. Aktuarietidskr. **3–4**, 160–170 (1956).
5. PATNAIK, P. B.: The non-central χ^2 and F distributions and their applications. Biometrika **36**, 202–232 (1949).
6. TAKANO, K.: A remark to a result of A. C. BERRY. Res. Mem. Inst. statist. Math. **9**, 6 (1951).
7. ZOLOTAREV, V. M.: An absolute estimate of the remainder in the central limit theorem. Teor. Verojatn. Primen **11**, 108–119 (1966).

Mathematical Steklov Institute
Vavilov 42 B-333
Moscow (USSR)