Centrum voor Wiskunde en Informatica
Centre for Mathematics and Computer Science

N.G. Gamkrelidze

On a local limit theorem for lattice distribution

Department of Operations Research, Statistics, and System Theory Report BS-R9004 February

The Centre for Mathematics and Computer Science is a research institute of the Stichting Mathematisch Centrum, which was founded on February 11, 1946, as a nonprofit institution aiming at the promotion of mathematics, computer science, and their applications. It is sponsored by the Dutch Government through the Netherlands Organization for the Advancement of Research (N.W.O.).

On a Local Limit Theorem for Lattice Distribution

N.G. Gamkrelidze
Department of Probability Theory and Mathematical Statistics
Tbilisi Mathematical Institute
Academy of Sciences of the Georgian SSR
150A, D. Agmashenebeli Avenue, Tbilisi 380012, USSR

In this paper the straightforward extension of the classical local limit theorem (l.l.t.) is given. In particular a notion of the function of smoothness is introduced, its properties are investigated and the sufficient conditions for the l.l.t. are presented in terms of this function. Beforehand some counter examples are discussed i.e. examples where the plausible, at first site, conjecture that the l.l.t. holds, turns out to be false. Finally some numerical examples are given.

1980 Mathematics Subject Classification: 60F05
Keywords and Phrases: local limit theorem, asymptotically uniform distribution.
Note: This paper was completed while the author was visiting the Centre for Mathematics and Computer Science, Amsterdam. We wish to thank K. Dzhaparidze for discussions and editorial work.

1. Let

$$\xi_1, \cdots, \xi_n, \cdots \tag{1}$$

be a sequence of independent integer-valued random variables (i.v.r.v.).

Denote $S_n = \xi_1 + \cdots + \xi_n$, $A_n = ES_n$, $B_n^2 = DS_n$, $P_n(m) = P(S_n = m)$ and $P_{km} = P(\xi_k = m)$. Let P_η^{*m} be the m-fold convolution of the distribution of an independent integer-valued random variable η.

The sequence (1) is said to satisfy the local limit theorem (l.l.t.) if

$$P_n(m) = (2\pi B_n)^{-1/2} \exp\{-(m-A_n)^2/2B_n^2\} + o(B_n^{-1}) \tag{2}$$

as $n \to \infty$.

We will say that the strong local limit theorem (s.l.l.t.) holds if relation (2) is true for (1) as well as for every sequence, which differs from (1) only by a finite number of the first members [1]. As is known [1,2] it is necessary for the l.l.t. that (1) is asymptotically uniformly distributed (a.u.d.) in the sense that for any fixed integer valued $h > 0$

$$\lim_{n\to\infty} P(S_n \equiv j \bmod h) = h^{-1}, \quad j = 0, 1, 2, \ldots, h-1.$$

For uniformly distributed random variables in (1) (i.e. $P(\xi_n = k) = 1/m_n$, $k = 1, 2, \ldots, m_n$) the l.l.t. is equivalent to the central limit theorem (c.l.t.) [3]. It seems reasonable to ask whether this fact is true generally. Unfortunately the answer is negative. Furthermore in Example 1 we construct a sequence of a.u.d. independent i.v. random variables for which the c.l.t. holds and at the same time the l.l.t. fails to hold.

2. EXAMPLE 1 [4]. Let ξ_{2k-1} be a random variable with the symmetric Poisson distribution and the characteristic function $f(t, \xi_{2k-1}) = \exp\{\Lambda_k(\cos th_k - 1)\}$. The parameters Λ_k and h_k will be choosen below. Take an even uniformly distributed random variable ξ_{2k}, which takes each couple of values $-k, \ldots, k$ with probability $1/2k$. It should be noted that if we add an arbitrary random variable ξ to a random variable η distributed by modulo h, then the sum $\xi + \eta$ is distributed by modulo h. Thus $S_n = \xi_1 + \cdots + \xi_n$ is a.u.d.

The sum S_n / B_n satisfies the c.l.t., because even terms ξ_{2k} are asymptotically normal. Now, if we choose Λ_k, h_k as follows: $\Lambda_k = k^{1/2}$ and $h_k = 2[e^k e^{-1/k}]$, then the random variables increase very

Report BS-R9004
Centre for Mathematics and Computer Science
P.O. Box 4079, 1009 AB Amsterdam, The Netherlands

rapidly and $P(S_n=0)\not\sim P(S_n=h_m)$. Thus the l.l.t. fails to be valid.

Now one might suggest that the extra condition on ξ_k-infinite negligibility property of ξ_k-would help us. Unfortunately this is a false conjecture.

In Example 2 below we construct a sequence of independent a.u.d. integer valued random variables which satisfies the c.l.t., has the infinite negligibility property, but the l.l.t. fails to be valid.

EXAMPLE 2 [5]. Let $[1,1,\ldots,1]$ be a continued fraction of the number $\alpha=(1+\sqrt{5})/2$. Represent a suitable fraction by means of the following table

j	1	2	3	
p_j	3	5	8	$\cdots$
q_j	2	3	5	$\cdots$

It then follows that p_j $(j=1,2,\cdots)$ is the Fibonachi sequence i.e. $p_j=p_{j-1}+p_{j-2}$ and $q_j=p_{j-1}$ $(j\geqslant 2)$. Let us consider a sequence of independent random variables which we represent by the following table

$$\xi_1,\cdots,\xi_{n_1},$$
$$\xi_{n_1+1},\cdots,\xi_{n_1+n_2},$$
$$\cdots\ \cdots\ \cdots$$
$$\xi_{n_1+\cdots+n_j+1},\cdots,\xi_{n_1+n_2+\cdots+n_{j+1}}$$
$$\cdots\ \cdots\ \cdots\ \cdots\ \cdots$$
$$\xi_{n_1+\cdots+n_{k-1}+1},\cdots\ \xi_{n_1+\cdots+n_k}$$

Every line with the number j consists of i.i.d. random variables which take the values $0,q_j,p_j$ with probabilities $\frac{p_{j-2}}{p_j}$, $\frac{1}{p_j}$, $\frac{1}{p_j}$; the number n_j of random variables in the line equal to $[p_j^{3/2}]+1$. (We denote by [a] the integer part of a.)

We shall verify the property of infinite negligibility as follows. For arbitrary n we can choose a number k, so that $N_{k-1}\leqslant n\leqslant N_k$ where $N_k=n_1+\cdots n_k$, hence

$$\max_{1\leqslant j\leqslant n}|\xi_j-E\xi_j|\leqslant 2p_n,\qquad B_n^2>\frac{1}{3}\sum_{j=1}^{k-1}\frac{p_j^2+q_j^2}{p_j}$$

and thus

$$\max_{1\leqslant j\leqslant n}|\xi_j-E\xi_j|/B_n=\frac{\text{const}}{p_k^{1/2}}\to 0\ \text{ as }\ n\to\infty.$$

Obviously Liapunov's condition and the property of infinite negligibility hold.

Now we shall investigate the a.u.d. property by the Dvoretzky-Wolfowitz test [2], which states that for an arbitrary fixed $h>0$ and $r=0,1,\ldots,h-1$ the characteristic function of sums of i.d. random variables tend to zero.

We obtain

$$|f(2\pi\frac{r}{h},\xi_{N_j})|\leqslant 1-\frac{\eta}{p_j}\quad\text{where}\quad \eta=\eta(h)>0.$$

Finally we shall show that $f(t,S_n)$ does not tend to zero near the point $2\pi\alpha$.

Let us consider Taylor's expansion when $|t-2\pi\alpha|<\frac{1}{B_{n_k}}$. We can write

$$|f(t,\xi_j)|^2 \geqslant 1-\frac{c}{p_j^3}-\frac{c}{p_jB_{N_k}}-c\frac{p_j}{B_{N_k}^2}$$

and we obtain

$$\prod_{j=1}^{k}|f(t,\xi_j)|^{2n_j} \geqslant \exp\left\{-c\sum_{j=1}^{k}\frac{n_j}{p_j^3}-\frac{c}{B_{N_k}}\sum_{j=1}^{k}\frac{n_j}{p_j}-\frac{c}{B_{N_k}^2}\sum_{j=1}^{k}n_jp_j\right\} \geqslant e^{-c}.$$

Hence we may conclude that if K is sufficiently large then

$$J_{N_k}=B_{N_k}\int_{\epsilon\leqslant|t|\leqslant\pi}\prod_{j=1}^{k}|f(t,\xi_j)|^{2n_j}dt>B_{N_k}\int_{|t-2\pi\alpha|<B_{N_k}^{-1}}e^{-c}dt>e^{-c}.$$

It should be noted that a necessary condition for the l.l.t. is the following fact [6]:

$$J_n=B_n\int_{\epsilon_n\leqslant|t|\leqslant\pi}\prod_{j=1}^{n}|f(t,\xi_j)|^2dt\to 0$$

for any positive ϵ_n which tends to zero as $n\to\infty$.

As we have seen above $J_{N_k}>$const, so that the necessary condition $J_n\to 0$ is violated and sequence (1) of independent a.u.d. random variables satisfying the c.l.t. has infinite negligibility property, but the l.l.t. fails to be valid.

3. *Main statement.* For any, integer-valued random variable η define the function of smoothness by $\delta(P_\eta)=\sum_Z|P(\eta=m)-P(\eta=m-1)|$, where Z is the set of all integers.

Let us list some properties of $\delta(P_\eta)$ [7,8].

1) Obviously $\delta(P_\eta)\leqslant 2$.

2) THEOREM. *If* $\delta(P_\eta)<2$ *then the maximal step of the distribution of* η *is equal to one.*

The converse assertion may be false: it may happen that $h_{\max}=1$ while $\delta(P_\eta)=2$. For example, for the random variable η which takes the values 0,3 and 5 with probabilities 1/3 we have $\delta(P_\eta)=2$.

3) ASSERTION. $\delta(P_\eta)\geqslant 2\max_{m\in Z}P(\eta=m)$.

If the distribution is unimodal (i.e. if there is an m_0 such that $P(m+1)\leqslant P(m)$ for $m\geqslant m_0$ and $P(m-1)\leqslant P(m)$ for $m\leqslant m_0$, then $\delta(P_\eta)=2P(m_0)=2\max_{m\in Z}P(m)$).

If in addition the distribution is symmetric or has a non-negative characteristic function, then $m_0=0$ and

$$\delta(P_\xi)=2P(0)=\frac{1}{\pi}\int_{-\pi}^{\pi}f(t)dt.$$

4) ASSERTION. *If* η_1 *and* η_2 *are independent integer valued random variables, then* $\delta(P_{\eta_1+\eta_2})\leqslant \min_j\delta(P_{\eta_j})$.

From property 4) it follows that $\delta(P_{s_n})$ is a monotone non-increasing function of n.

It should be noted that if ξ_k are identically distributed with $h_{\max}=1$ then, starting from some n, the inequality $\delta(P_{S_n})<2$ holds, although, as is already noted, $\delta(P_{\xi_1})$ may equal 2. But if ξ_n are distributed differently and $h_{\max}=1$, then the equality $\delta(P_{S_n})=2$ may hold for each n. Indeed, let ξ_1 take the values 0 and 8^{k-1} with probabilities 1/2. Then the difference of any pair of possible values is $\geqslant 2$, and so $\delta(P_{s_n})=2$.

The properties of $\delta(P_{\eta_n})$ listed below describe the behaviour of this function as $n\to\infty$.

5) THEOREM. *If* $\lim_{n\to\infty}\delta(P_{\eta_n})=0$ *then the random variable* η_n *is asymptotically uniformly distributed, i.e.*

$$\lim_{n\to\infty}(P_{\eta_n} \equiv j \bmod l) = 1/l, \quad j=0,1,\ldots,l-1$$

for an arbitrary integer $l>1$.

The converse assertion is false.

6) THEOREM. *Let* η_n *be an arbitrary i.v.r.v. Let* $g(x)$ *be a nonnegative bounded piecewise monotone function (the number of monotonicity intervals is finite), and for some* A_n *and* B_n *let*

$$\sum_{m\in Z}|(P_{\eta_n}=m)-\frac{1}{B_n}g\left[\frac{m-A_n}{B_n}\right]| \to 0, \quad \text{as} \quad n\to\infty$$

where $B_n\to\infty$. *Then* $\delta(P_{\eta_n})\to 0$ *as* $n\to\infty$.

7) Let us show that the sum S_n of independent identically distributed i.v. random variables ξ_k with $h_{\max}=1$ always satisfies $\delta(P_{S_n})\to 0$ as $n\to\infty$ (no matter whether their distribution is attracted to a stable law or not). Moreover, for each sequence ξ_k satisfying the indicated conditions, there is a constant $A>0$ such that the estimate $\delta(P_{S_n})=A_n^{-1/2}+o(n^{-1/2})$ is true ($n\to\infty$). For simplicity we suppose that the considered i.v. random variables are bounded by one and the same constant.

THEOREM. *If* $\{\xi_k\}_{k=1}^{n}$ *is a sequence of independent bounded* ($\xi_n\leqslant L$) *identically distributed i.v. random variables with* $h_{\max}=1$, *then*

$$\delta(P_{S_n}) = \frac{2}{\sigma\sqrt{2\pi n}} + o(\frac{1}{n})$$

where σ^2 *is the variance of* ξ_n.

4. Using now the function $\delta(P_\xi)$ we will present a wide class of "smooth" distributions (in sense of function $\delta(P_\xi)$) for which the strong version of the l.l.t. holds [9]:

THEOREM (M). *Let* $\xi_1,\ldots,\xi_n$ *be a sequence of independent (not necessarily identically distributed) integer valued random variables. Assume that*

1) *there exists a natural number* n_0 *and a positive number* $\lambda<\sqrt{2}$ *such that* $\delta(P_{\xi_k}^{*n_0})\leqslant\lambda$ *uniformly in k;*
2) *the c.l.t. holds true for the sequence (1);*
3) $B_n^2=O(n)$ *as* $n\to\infty$.

Then the sequence (1) satisfy the s.l.l.t.

The last theorem is essentially based on the following lemmas. The first one gives the connection between the characteristic function and the function of smoothness $\delta(P_\eta)$.

LEMMA 1. *Let* η *be an integer valued random variables. Then we have*

$$|f(t,\eta)| \leqslant \frac{\delta(P_\eta)}{2|\sin\frac{t}{2}|}$$

for every $t\neq 2\pi h$.

On the other hand, by Lemma 1 and Cramer's inequality we have

LEMMA 2. *Let* n_0 *be a natural number which satisfy condition* $\delta(P_\eta^{*n_0})<\sqrt{2}$. *(This is possible iff the distribution has a maximal step equal to 1). Then for* t: $|t|\leqslant\pi$ *we have*

$$|f(t,\eta)| \leqslant \exp\{-ct^2\}$$

where

$$c = \left[1 - \frac{\delta^2(P_\eta^{*n_0})}{2}\right] \frac{1}{2\pi^2 n_0}$$

To prove the l.l.t., one usually starts with the following representation

$$\Delta_n = 2\pi[B_n P_n(k) - \frac{1}{\sqrt{2\pi}} e^{-z_{n,m}^2/2}] = J_1 + J_2 + J_3$$

where $z = z_{n,m} = (m - ES_n)/2B_n^2$ and

$$J_1 = \int_{|t| \leqslant A} e^{-izt}[f(\frac{t}{B_n}, S_n) - e^{-t^2/2}]dt,$$

$$J_2 = \int_{|t| \geqslant A} e^{-izt - t^2/2} dt, \quad J_3 = \int_{A \leqslant |t|} e^{-izt} f(\frac{t}{B_n}, S_n) dt.$$

For brevity we omit here the details of estimating J_1 and J_2. We observe only that J_1 tends to zero because the c.l.t. is true. Concerning J_2 we note that it may be made arbitrary small by choosing A. The main problem is to show that $J_3 \to 0$ as $n \to \infty$.

Suppose that there exists a natural number n_0 and positive number $\lambda < \sqrt{2}$ such that $\delta(P_{\xi_n^{*n_0}}) \leqslant \lambda$ holds uniformly in k. Thus according to Lemma 2 we obtain

$$J_3 \leqslant \int_{A \leqslant |t| \leqslant \pi B_n} \prod_{j=1}^{n} |f(\frac{t}{B_n}, \xi_j)| dt \leqslant 2B_n \int_{A/B_n} e^{-ct^2 n} dt \leqslant \frac{2B_n^2 \pi^2}{A[\frac{n}{n_2}](1 - \frac{\lambda^2}{2})} \exp\{-[\frac{n}{n_0}]A^2(1 - \frac{\lambda^2}{2})/2B_n^2 \pi^2\}.$$

It should be noted that here the parameters n_0 and λ ensure the existence of sequence of i.i.d. random variables $\zeta_1^{(k)}, \ldots, \zeta_{n_0}^{(k)}$ which are distributed like ξ_k for every k $(k = \overline{1,n})$ and they are "sufficiently smooth" in the sense of the function of smoothness $\delta(P_{\xi_{1}(k)}^{*n_0}) \leqslant \lambda < \sqrt{2}$.

5. Now we consider basic properties of the function $\delta(P_\xi)$ in the multidimensional case [8].

Let $\xi = (\xi^{(1)}, \ldots, \xi^{(s)})$ be a random variable in the s-dimensional space. Let $\xi^{(k)}$ take only integer values $m^{(k)}, k = 1, \ldots, s$. Denote $m = (m^{(1)}, \ldots, m^{(s)})$, $e_1 = (1, 0, \ldots, 0), \ldots,$ $e_s = (0, \ldots, 1)$. The multidimensional function of smoothness may be defined as follows

$$\delta(P_\xi) = \sum_{k=1}^{s} \sum_{m} |P(\xi = m) - P(\xi = m - e_k)|$$

The second sum is taken over all integers.

Obviously we have $\delta(P_\xi) \leqslant 2s$.

It is not difficult to generalize Theorem 4 in R^s. For simplicity we consider the 2 dimensional case.

Let $S_1, S_2, \ldots,$ be a sequence of integer valued random vectors $S_n = (S_n^{(1)}, S_n^{(2)})$.

THEOREM. *If* $\lim_{n \to \infty} \delta(P_{S_n}) = 0$, *then the random vector* S_n *is asymptotically uniformly distributed.*

The next step is to establish the connection between the multidimensional function of smoothness and the maximal step.

THEOREM. *Let* ξ *be an integer valued vector* $\xi \in R^s$. *If* $\delta(P_\xi) < 2$, *then the maximal step of the distribution of* ξ *is equal to one.*

The following properties are concerned with the asymptotic behaviour of $\delta(P_{S_n})$, where S_n is the sum of random vectors $\xi_1, \dots, \xi_n$.

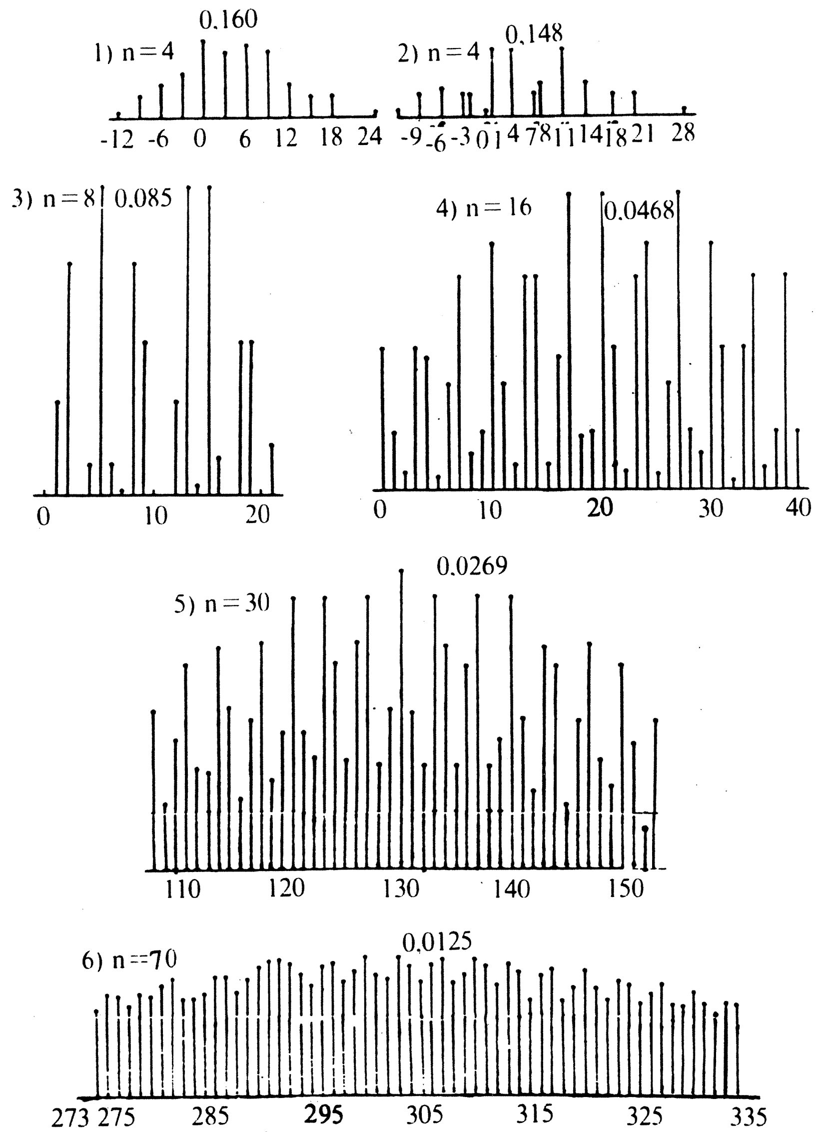
Denote $a^{(k)}=E\xi_j^{(k)}$, $\sigma_k^2=D\xi_1^{(k)}$, $\lambda_{jk}=E(\xi_1^{(j)}-a^{(j)})(\xi_1^{(k)}-a^{(k)})$ and $\Lambda=\det\{\lambda_{jk}\}_{j,k=1,\dots,s}$. Let Λ_{jk} be the co-factor of the element λ_{jk}.

THEOREM. *Let $\xi_1, \dots, \xi_n$ be a sequence of integer valued i.i.d. random vectors with maximal step 1 and $|\xi_k|<L$. Then*

$$\delta(P_{S_n}) = \sqrt{\frac{2}{\pi n\Lambda}}\,(\sqrt{\Lambda_{11}}+\cdots+\sqrt{\Lambda_{ss}})+o(\frac{1}{\sqrt{n}}).$$

6. SOME NUMERICAL EXAMPLES

Let η be a random variable which takes values 0,3,10 with probabilities 1/3. Tables 3,4,5 and 6 illustrate the behaviour of the probability $P_n(m)$ for $n=8,16,30,70$ and m: $ES_n-B_n\leqslant m\leqslant ES_n+B_n$. Table 2 corresponds to values -3,0,7 and $n=4$. For comparison Table 1 gives the distribution $P_n(m)$ of η taking values -3,0,6 with probabilities 1/3.



Concerning these examples we are interested in the following problems.

1. The estimation problem of the absolute value of approximation by the local limit theorem. This problem is solved by applying theorem (M).
2. The problem of approximating $P_n(m)$ as n is not large (i.e. $n=10,20$).

The situation in 1 and 2 is different. If n is 'large' ($n\geqslant 100$) the limit behaviour of $P_n(m)$ has the smoothness property. But if $n\simeq 10$ the probability $P_n(m)$ behaves irregularly and the addition to 'normal' approximation consequent terms does not improve essentially the approximation:

$$P_n(m) = \frac{1}{\sqrt{2\pi n}\,\sigma} e^{-\frac{z^2}{2}} \{1+\frac{\mu_3}{6\sigma^3\sqrt{n}}(z^3-3z)+\cdots\} \qquad (n_1)$$

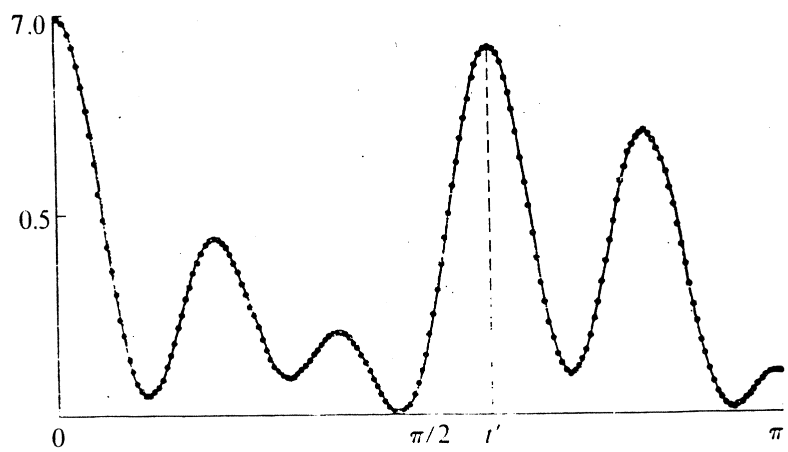
To improve the approximation we introduce correcting terms [10], and the l.l.t. with these corrections approximates $P_n(m)$ better than usually. For simplicity we assume that $f(t)$ is real and nonnegative. Let $t_1,t_2,\ldots,t_s$ be points of local maximums of characteristic function $f(t)$, $t_j\in(0,\pi)$. If $E\xi_j^2<\infty$, then

$$P_n(m) = \frac{1}{\sqrt{2\pi n}\,\sigma} e^{-\frac{m^2}{2n\sigma^2}} + 2\sum_{j=1}^{s} \frac{f^n(t_j)}{\sigma_j\sqrt{2\pi n}} \cos(t_j m) e^{-\frac{m^2}{2n\sigma_j^2}} + o(\frac{1}{n}). \qquad (n_2)$$

It should be noted that formula (n_2) has no advantage over (n_1) because the remainder terms in (n_2) and (n_1) are of one and the same order, however the approximation in (n_2) is better than (n_1). Indeed let $\tilde{\eta}$ be a random variable such that

$\tilde{\eta}$	0	± 3	± 7	± 10
	1/3	1/9	1/9	1/9

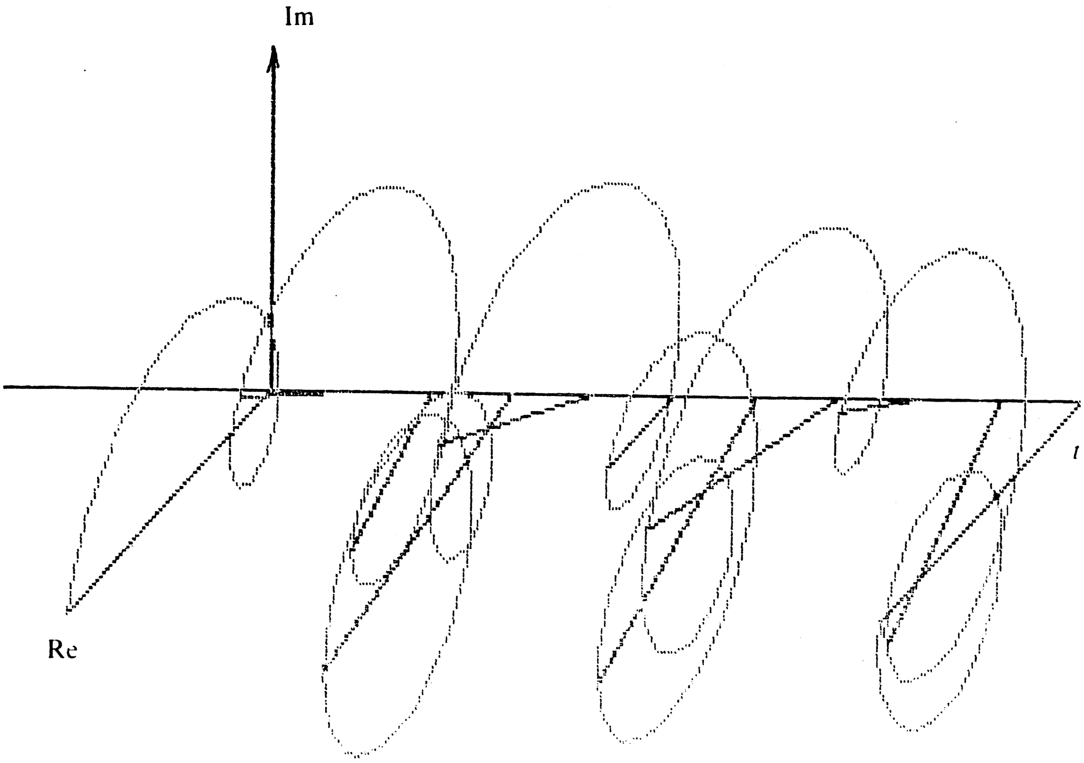
We give below the graph of its characteristic function.



For comparison we give also the graph of the characteristic function of the random variable

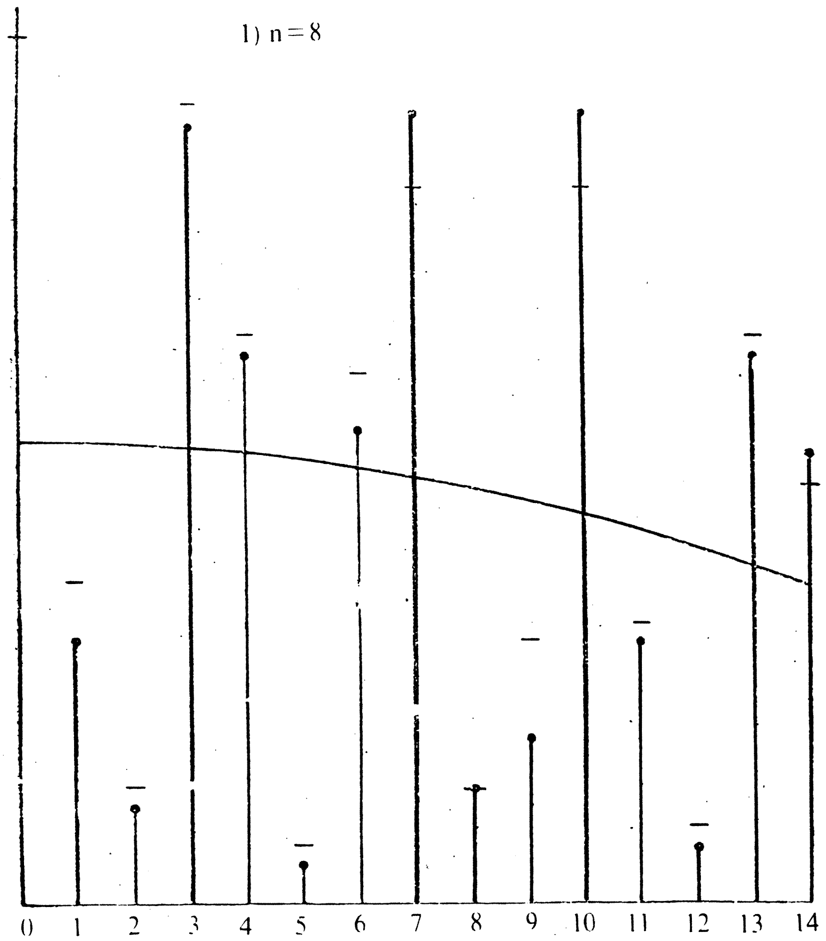
η	0	3	0
	1/3	1/3	1/3

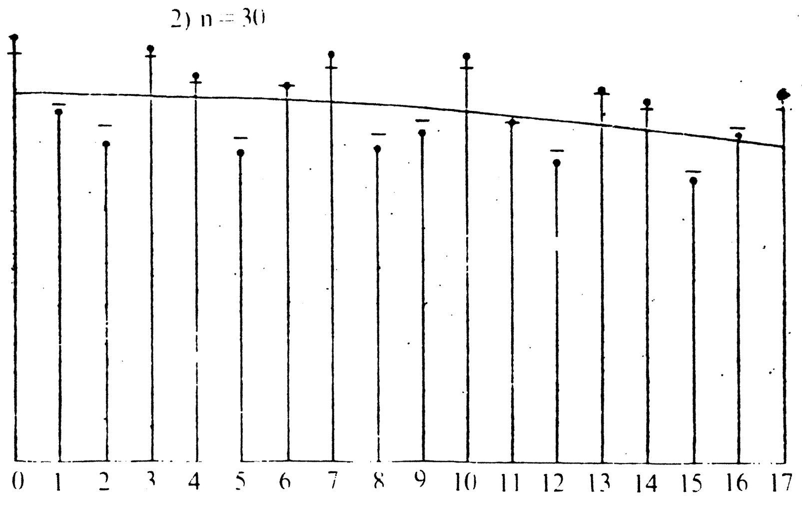
the symmetrization of which leads to $\tilde{\eta}$.



The next table presents the $n=8,30$ fold convolution of the distribution of the random variable η with itself (bold dots), the approximation by the classical local limit theorem (continuous curve) and the l.l.t. with correction (dashes).

This numerical material shows that formula (n_2) which takes in account only one local maximum of characteristic function describes the behaviour of probability $P_n(m)$ better then the l.l.t. To get the idea of the effect of correcting terms in (m_2) observe that the graph of the characteristic function is very close to 1. It is this local maximum that improves the accuracy of the approximation.


1) n=8
0
1
2
3
4
5
6
7
8
9
10
11
12
13
14


2) n=30
0
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17

REFERENCES

1. PROKHOROV U.V., On a local limit theorem for lattice distributions. *Doklady Academy of Sciences USSR,* vol. XCVIII (1954), N4, 535-538.
2. DVORETZKY, J. WOLFOWITZ, Sums of random integers reduced modulo *m*. *Duke Math. J.* vol. 18 (1951), N2, 501-507.
3. PROKHOROV U.V., On a local theorem. *In 'Limit Theorems in Probability Theory',* Academy of Sciences UzSSR Publ., Tashkent, 1963, 75-80.
4. GAMKRELIDZE N.G., On a local limit theorem for lattice random variables. *Theory Probab. Appl.,* vol. IX (1964), N4, 662-665.
5. GAMKRELIDZE N.G., On the connection between the local and the integral theorem for lattice distributions. *Theory Probab. Appl.,* vol. XIII (1968), N1, 174-179.
6. GAMKRELIDZE N.G., On a lower estimate of the speed of convergence in the local theorem. *Lit. Math. Rink.,* vol. VII (1968), N3, 405-408.
7. GAMKRELIDZE N.G., On 'smoothing' the probabilities for sums of independent integer-valued variables. *Theory Probab. Appl.,* vol. 26 (1981), N4, 823-828.
8. GAMKRELIDZE N.G., A measure of 'smoothness' of multidimensional distributions of integer-valued random vectors. *Theory Probab. Appl.,* vol. 30 (1986), N1, 427-431.
9. GAMKRELIDZE N.G., On the application of a smoothness function in proving a local limit theorem. *Theory Probab. Appl.,* vol. 33 (1988), N2, 352-355.
10. GAMKRELIDZE N.G., U.V. PROKHOROV. On the approximation of the distributions of sums of lattice random variables when the number of summands is small. *Theory Probab. Appl.,* vol. 16 (1971), N1, 144-146.