



Soviet-era science, translated into English

MATHEMATICS

V. V. PETROV

1957

SovietRxiv

View the original and related papers at <https://sovietrxiv.org/items/ru-195701.42564>

Source: Math-Net.Ru and CyberLeninka. Machine translation. Verify with the original.

Abstract

Full Text

MATHEMATICS

V. V. PETROV

A LOCAL THEOREM FOR LATTICE DISTRIBUTIONS

(Presented by Academician A. N. Kolmogorov, 28 I 1957)

1. Consider a sequence of independent, not identically distributed random variables X_1, X_2, \dots , each of which can take only integer values. We shall use the notation:

$$a_m = EX_m, \quad \sigma_m^2 = DX_m, \quad A_n = \sum_{m=1}^n a_m, \quad s_n^2 = \sum_{m=1}^n \sigma_m^2;$$

$$P_n(N) = P\{X_1 + \dots + X_n = N\}.$$

For the special case of identical distributions, B. V. Gnedenko ^(1,2) obtained conditions necessary and sufficient in order that, as $n \rightarrow \infty$, uniformly with respect to N , $-\infty < N < \infty$, the relation

$$s_n P_n(N) - \frac{1}{\sqrt{2\pi}} e^{-(N-A_n)^2/2s_n^2} \rightarrow 0 \quad (1)$$

hold.

The local limit theorem for non-identical lattice distributions was studied by P. Mises ^(3,4), G. M. Bavli ^(5,6), Yu. V. Prokhorov ⁽⁷⁾, and G. A. Freiman ⁽⁸⁾. In works ⁽³⁻⁶⁾ the existence of finite absolute moments of the third order was assumed, which, together with certain other restrictions, makes it possible to investigate the problem of estimating the remainder term in relation (1). In ^(7,8) the case of uniformly bounded variables ($|X_m| \leq \text{const}$) was studied.

We shall assume that there exist finite absolute moments $E|X_m - a_m|^{2+\delta}$ of order $2 + \delta$, where $0 < \delta \leq 1$ ($m = 1, 2, \dots$). Introduce for consideration the Lyapunov fraction

$$L_n = \frac{1}{s_n^{2+\delta}} \sum_{m=1}^n E|X_m - a_m|^{2+\delta}.$$

In the present note we indicate conditions under which the following refinement of the local theorem for non-identical lattice distributions holds:

$$\left| s_n P_n(N) - \frac{1}{\sqrt{2\pi}} e^{-(N-A_n)^2/2s_n^2} \right| \leq CL_n, \quad (2)$$

where C is a constant. An analogous estimate for the deviation of the distribution function of the normalized sum of independent, non-identically distributed variables from the normal distribution function was obtained by Lyapunov for the case $\delta < 1$, by Cramér and Berry for $\delta = 1$ under special assumptions, and by Esseen ⁽⁹⁾ under the sole assumption of the existence of finite moments of order $2 + \delta$, $0 < \delta \leq 1$. The estimate

of order L_n for the difference of the corresponding distribution densities is obtained in (10).

2. Put $p_{mj} = P\{X_m = j\}$. Without loss of generality one may assume that $p_{m0} \geq p_{mj}$ for all j and $m = 1, 2, \dots$. We shall suppose that the sequence of random variables X_1, X_2, \dots satisfies the following conditions:

α) The greatest common divisor of all integers j for which

$$\frac{1}{\log n} \sum_{m=1}^n p_{m0} p_{mj} \rightarrow \infty$$

as $n \rightarrow \infty$ is equal to one;

$$\beta) \quad s_n \rightarrow \infty \quad \text{as } n \rightarrow \infty; \quad \sum_{m=1}^n E|X_m - a_m|^{2+\delta} \leq Bs_n^2$$

for all n and some constant B .

Theorem 1. *If conditions α) and β) are satisfied, then (2) holds, where C is a positive constant independent of n and N .*

In the case of identical distributions the conditions of the theorem reduce to the existence of a nonzero variance and of a finite absolute moment of order $2 + \delta$ ($0 < \delta \leq 1$), and to the requirement that the greatest common divisor of the values assumed with positive probability be equal to one.

For another special case—the case of uniformly bounded variables—condition β) reduces to the requirement $s_n \rightarrow \infty$ as $n \rightarrow \infty$, while condition α) is stronger than the corresponding condition of Yu. V. Prokhorov's theorem (7), but it makes it possible to obtain the estimate (2) for the remainder term.

Let us note that conditions α) and β) are not violated if the distributions of a finite number of initial terms of the sequence of variables X_1, X_2, \dots are changed

(provided that the integrality condition and the condition of existence of finite moments of order $2 + \delta$ are observed).

3. Proof. Denote the characteristic function (c.f.) of the random variable X_m by $f_m(t)$, and the c.f. of the variable $s_n^{-1}(X_1 + \dots + X_n)$ by $\bar{f}_n(t)$. Put $z_N = s_n^{-1}(N - A_n)$ and

$$2\pi \left[s_n P n(N) - \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2} z_N^2} \right] = I_1 + I_2 + I_3,$$

where

$$I_1 = \int_{|t| \leq (24L_n)^{-1/\delta}} e^{-iz_N t} [\bar{f}_n(t) - e^{-\frac{1}{2} t^2}] dt,$$

$$I_2 = \int_{(24L_n)^{-1/\delta} < |t| \leq \pi s_n} e^{-iz_N t} \bar{f}_n(t) dt,$$

$$I_3 = - \int_{|t| \leq (24L_n)^{-1/\delta}} e^{-iz_N t - \frac{1}{2} t^2} dt.$$

We shall assume that $(24L_n)^{-1/\delta} < \pi s_n$; otherwise the estimates are simplified.

As is known (see, for example, (9)), for $|t| \leq (24L_n)^{-1/\delta}$

$$|\bar{f}_n(t) - e^{-\frac{1}{2} t^2}| \leq c_1 L_n |t|^{2+\delta} e^{-\frac{1}{4} t^2}.$$

Hence $|I_1| \leq c_2 L_n$. The positive constants c_1, c_2, \dots depend on δ , but not on n and N . It is easy to show that $|I_3| \leq c_3 L_n$.

It remains to estimate the integral I_2 . Obviously,

$$|I_2| \leq s_n \int_{s_n^{-1}(24L_n)^{-1/\delta} < |t| \leq \pi} \prod_{m=1}^n |f_m(t)| dt.$$

By condition β we have $L_n \leq B s_n^{-\delta}$, and therefore $s_n^{-1}(24L_n)^{-1/\delta} \geq c_4$. Let K_0 be a positive number such that the greatest common divisor of all integers j , $|j| < K_0$, for which

$$\frac{1}{\log n} \sum_{m=1}^n p_{m0} p_{mj} \rightarrow \infty$$

is equal to one. Such a K_0 exists by condition α . Put $K = \max(K_0, (2c_4)^{-1})$. We have

$$|I_2| \leq 2s_n \int_{(2K)^{-1}}^{\pi} \prod_{m=1}^n |f_m(t)| dt = R_2.$$

We estimate the integral R_2 by means of the method used in (7). In view of the fact that $|f_m(t)|^2 \leq \exp\{|f_m(t)|^2 - 1\}$, we have

$$\prod_{m=1}^n |f_m(t)| \leq \exp \left\{ \frac{1}{2} \sum_{m=1}^n (|f_m(t)|^2 - 1) \right\}.$$

Further,

$$|f_m(t)|^2 - 1 = \sum_j \tilde{p}_{mj} (\cos jt - 1), \quad \text{where} \quad \tilde{p}_{mj} = \sum_s p_{m,j+s} p_{ms}.$$

Denote by t_1, \dots, t_ν the points of the form $2\pi r/j$ lying on the segment $[(2K)^{-1}, \pi]$ (r and j relatively prime, $1 \leq r \leq [j/2]$, $2 \leq j < K$), taken in increasing order of their abscissae. Obviously, $t_\nu = \pi$, $t_1 > (2K)^{-1}$. Putting $\Delta_1 = [(2K)^{-1}, (t_1 + t_2)/2]$, $\Delta_k = [(t_{k-1} + t_k)/2, (t_k + t_{k+1})/2]$ for $k = 2, \dots, \nu - 1$, $\Delta_\nu = [(t_{\nu-1} + t_\nu)/2, \pi]$, we represent R_2 in the form of a sum of integrals over the Δ_k . Consider a fixed segment Δ_k containing the point t_k . Let $t_k = 2\pi r_0/j_0$. Obviously,

$$|f_m(t)|^2 - 1 \leq -2 \sum' \tilde{p}_{mj} \sin^2 \frac{jt}{2} - 2 \sum'' \tilde{p}_{mj} \sin^2 \frac{jt}{2},$$

where \sum' is the sum over all integers j , $|j| < K$, $j \not\equiv 0 \pmod{j_0}$, and \sum'' is the sum over all integers j , $|j| < K$, $j \equiv 0 \pmod{j_0}$, $j \neq 0$.

The minimal distance between the points t_k is no less than $2\pi K^{-2}$; therefore, for any $t \in \Delta_k$, $j \not\equiv 0 \pmod{j_0}$, $|j| < K$, and any integer s , we have $|jt - 2\pi s| > \varepsilon_1$; hence $\sin^2 \frac{1}{2} jt > \varepsilon_2$ ($\varepsilon_1, \varepsilon_2, \dots$ are positive constants depending only on K). For $t \in \Delta_k$, $j \equiv 0 \pmod{j_0}$, $|j| < K$, $j \neq 0$, we obtain $\sin^2 \frac{1}{2} jt \geq \varepsilon_3 (t - t_k)^2$. Therefore

$$\frac{1}{2} \sum_{m=1}^n (|f_m(t)|^2 - 1) \leq -g_n - h_n (t - t_k)^2,$$

where

$$g_n = \varepsilon_2 \sum_{m=1}^n \sum' \tilde{p}_{mj}, \quad h_n = \varepsilon_3 \sum_{m=1}^n \sum'' \tilde{p}_{mj}.$$

By condition α , the choice of the number K , and the inequality $\tilde{p}_{mj} \geq p_{m0}p_{mj}$, we have

$$g_n \geq \varepsilon_2 \sum' \sum_{m=1}^n p_{m0}p_{mj} \geq M \log n$$

for any $M > 0$ and all sufficiently large n .

As is not difficult to see, $L_n \gg n^{-\delta/2}$ for all n (see, for example, the lemma in (10)). On the other hand, $L_n \ll Bs_n^{-\delta}$ by condition β . Consequently, $s_n \ll B^{1/\delta} \sqrt{n}$.

For sufficiently large n ,

$$s_n \int_{\Delta_k} \sum_{m=1}^n |f_m(t)| dt \ll s_n e^{-g_n} \int_{\Delta_k} e^{-h_n(t-t_k)^2} dt \ll c_5 e^{-\frac{M}{2} \log n} \ll c_5 L_n,$$

where c_5 is a constant not depending on n and N .

From the estimates obtained, inequality (2) follows.

4. A sufficient condition for β is the condition β' : there exist positive constants g and G such that, for all n ,

$$s_n^2 \gg ng, \quad \sum_{m=1}^n E |X_m - a_m|^3 \ll nG.$$

From Theorem 1 it follows that

Theorem 2. *If conditions α and β' are satisfied, then*

$$\left| s_n P_n(N) - \frac{1}{\sqrt{2\pi}} e^{-(N-A_n)^2/2s_n^2} \right| \ll \frac{C_0}{\sqrt{n}},$$

where C_0 is a positive constant not depending on n and N .

An example of a sequence of random variables each of which takes only two values, 0 and 1, with probabilities $p \neq 0$ and $q = 1 - p$, shows (2) that, in the general case, the order of the estimate given by Theorems 1 and 2 cannot be improved.

Leningrad State University
named after A. A. Zhdanov

Received
9 I 1957

REFERENCES

1. B. V. Gnedenko, *Uspekhi Mat. Nauk*, **3**, no. 3, 187 (1948).
2. B. V. Gnedenko, A. N. Kolmogorov, *Limit Distributions for Sums of Independent Random Variables*, 1949.
3. R. Mises, *Wahrscheinlichkeitsrechnung*, 1931.
4. R. Mises, *Giorn. Ist. Ital. degli Attuari*, **5**, 483 (1934).
5. G. M. Bavli, *Uch. Zap. Sverdlovsk. Univ.*, no. 2, 7 (1937).
6. G. Baluy, *Rev. Fac. Sci. Univ. Istanbul*, **2**, fasc. 2, 79 (1937).
7. Yu. V. Prokhorov, *DAN*, **98**, 535 (1954).
8. G. A. Freiman, *Vestn. LGU*, no. 1, 57 (1956).
9. C.-G. Esseen, *Acta Math.*, **77**, 1 (1945).
10. V. V. Petrov, *Theory of Probability and Its Applications*, **1**, 349 (1956).

Note: Figure translations are in progress. See original paper for figures.

Source: Math-Net.Ru and CyberLeninka. Machine translation. Verify with the original.