

On the Kolmogorov-Rogozin Inequality for the Concentration Function

C. G. ESSEEN

Received May 10, 1965

1. Introduction

The concentration function $Q(X; \lambda)$ of a random variable X is defined by

$$Q(X; \lambda) = \sup_{-\infty < x < \infty} P(x \leq X \leq x + \lambda), \lambda > 0.$$

If X and Y are independent random variables it is easily seen that

$$(1.1) \quad Q(X + Y; \lambda) \leq \min(Q(X; \lambda); Q(Y; \lambda)).$$

Let X_1, X_2, \dots be a sequence of mutually independent random variables and set

$$S_n = X_1 + X_2 + \dots + X_n.$$

An inequality which relates the concentration function of S_n to the concentration functions of the summands X_k and which is much deeper than (1.1) has been given by KOLMOGOROV [1], [2]. ROGOZIN [3] has obtained the following refinement of the KOLMOGOROV inequality: For any positive $\lambda \leq L$, one has

$$(1.2) \quad Q(S_n; L) \leq \frac{CL}{\lambda} \left\{ \sum_{k=1}^n (1 - Q(X_k; \lambda)) \right\}^{-1/2}$$

where C is an absolute constant. (In the following we shall denote by C absolute, in general different positive constants.) ROGOZIN [4] has also generalized the inequality (1.2). His result is expressed by the following theorem.

Theorem 1. *For any positive $\lambda_1, \dots, \lambda_n \leq L$, one has*

$$(1.3) \quad Q(S_n; L) \leq CL \left\{ \sum_{k=1}^n \lambda_k^2 (1 - Q(X_k; \lambda_k)) \right\}^{-1/2}.$$

As is well known the characteristic function is an important tool in studying properties of S_n . The proofs of KOLMOGOROV and ROGOZIN, however, are based on quite different methods. A combinatorial lemma of SPERNER is for instance fundamental in ROGOZIN's refinement of the KOLMOGOROV inequality. The purpose of the present paper is to give a new proof of Theorem 1 using characteristic functions. It will also be shown that the multi-dimensional case can be treated in the same way, at least if a certain symmetry condition is satisfied. It should be remarked that ROSÉN [5], using characteristic functions, has proved the inequality

$$Q(S_n; L) \leq C(L; F) n^{-1/2}$$

in the case of identically distributed summands. Here $C(L; F)$ is a quantity independent of n but depending on L and on the distribution function F of the summands. No explicit determination, however, of this dependence is given in the paper of ROSÉN.

2. Proof of the Inequality

We introduce auxiliary continuous functions $H(x)$ and $h(t)$ defined for all real x and t and satisfying the conditions

$$\text{a) } H(x) \geq 0, \quad \text{b) } \int_{-\infty}^{\infty} H(x) dx = 2\pi, \quad \text{c) } h(t) = 0 \quad \text{for } |t| \geq 1,$$

$$\text{d) } h(t) = (2\pi)^{-1} \int_{-\infty}^{\infty} e^{-itx} H(x) dx.$$

Thus

$$(2.1) \quad |h(t)| \leq 1, \quad -\infty < t < \infty.$$

Functions satisfying the above conditions are for instance

$$(2.2) \quad H(x) = \left(\frac{\sin x/2}{x/2}\right)^2, \quad h(t) = \begin{cases} 1 - |t|, & |t| \leq 1 \\ 0 & |t| > 1 \end{cases}.$$

Let the characteristic function of X_k be $f_k(t)$. Then the characteristic function of S_n is $\prod_{k=1}^n f_k(t)$. The distribution function of S_n will be denoted by $F_n^*(x)$. Our starting point is the easily proved relation

$$(2.3) \quad \int_{-\infty}^{\infty} H(a(x - \xi)) dF_n^*(x) = a^{-1} \int_{-a}^a \left(\prod_{k=1}^n f_k(t)\right) h(t/a) e^{-it\xi} dt$$

where a and ξ are real parameters and a is positive.

By (2.1) and (2.3) we get

$$\int_{-\infty}^{\infty} H(a(x - \xi)) dF_n^*(x) \leq a^{-1} \int_{-a}^a \prod_{k=1}^n |f_k(t)| dt.$$

Now let $H(x)$ be chosen such that

$$(2.4) \quad \min_{-L/2 \leq x \leq L/2} H(ax) \geq \mu > 0$$

and denote by I the closed interval $[\xi - L/2, \xi + L/2]$. Then

$$\int_{-\infty}^{\infty} H(a(x - \xi)) dF_n^*(x) \geq \mu P(S_n \in I)$$

or

$$P(S_n \in I) \leq (a\mu)^{-1} \int_{-a}^a \prod_{k=1}^n |f_k(t)| dt.$$

Since ξ is arbitrary we thus obtain

$$(2.5) \quad Q(S_n; L) \leq (a\mu)^{-1} \int_{-a}^a \prod_{k=1}^n |f_k(t)| dt.$$

From now on we choose $H(x)$ according to (2.2). Then (2.4) is satisfied by

$$(2.6) \quad aL = 2\pi, \quad \mu = 4/\pi^2.$$

From (2.5) and (2.6) we get

$$(2.7) \quad Q(S_n; L) \leq CL \int_{-2\pi/L}^{2\pi/L} \prod_{k=1}^n |f_k(t)| dt.$$

Thus far our arguments correspond to those of ROSÉN [5]. Our estimation of the right hand side of (2.7) will, however, be quite different. Applying the inequality

$$\prod_{k=1}^n |f_k(t)| = \left\{ \prod_{k=1}^n [1 - (1 - |f_k(t)|^2)] \right\}^{1/2} \leq \exp \left\{ -\frac{1}{2} \sum_{k=1}^n (1 - |f_k(t)|^2) \right\},$$

we get from (2.7)

$$(2.8) \quad Q(S_n; L) \leq C L \int_{-2\pi/L}^{2\pi/L} \exp \left\{ -\frac{1}{2} \sum_{k=1}^n (1 - |f_k(t)|^2) \right\} dt.$$

Let X'_k be a random variable independent of X_k and with the same distribution. Then $X_k - X'_k$ is symmetrically distributed with the characteristic function $|f_k(t)|^2$. Denoting the corresponding distribution function by $G_k(x)$ we have

$$(2.9) \quad 1 - |f_k(t)|^2 = \int_{-\infty}^{\infty} (1 - \cos tx) dG_k(x) \geq \int_{|x| > \lambda_k/2} (1 - \cos tx) dG_k(x).$$

From the definition of the concentration function and from (1.1) it follows that

$$(2.10) \quad \int_{|x| > \lambda_k/2} dG_k(x) \geq 1 - Q(X_k - X'_k; \lambda_k) \geq 1 - Q(X_k; \lambda_k).$$

(In the following we assume that $Q(X_k; \lambda_k) < 1$; it is easily seen that this does not imply any loss of generality.) Let $\varepsilon, \varepsilon_k$ be small but fixed positive quantities and $\varepsilon = \sum_{k=1}^n \varepsilon_k$. It is easily seen that the right member of (2.9) may be uniformly approximated for $|t| \leq 2\pi/L$ by a finite sum in the following way:

$$(2.11) \quad \int_{|x| > \lambda_k/2} (1 - \cos tx) dG_k(x) = \sum_{\nu=1}^{N_k} p_{\nu k} (1 - \cos t x_{\nu k}) + \theta_k$$

where

$$(2.12) \quad |x_{\nu k}| \geq \lambda_k/2, \quad |\theta_k| \leq \varepsilon_k, \quad p_{\nu k} > 0$$

and

$$(2.13) \quad \sum_{\nu=1}^{N_k} p_{\nu k} \geq 1 - Q(X_k; \lambda_k) - \varepsilon_k.$$

The last inequality follows from (2.10).

Combining the inequalities (2.8), (2.9) and (2.11) we get

$$(2.14) \quad Q(S_n; L) \leq C L e^\varepsilon \int_{-2\pi/L}^{2\pi/L} \prod_{k=1}^n \prod_{\nu=1}^{N_k} \exp \left\{ -\frac{1}{2} p_{\nu k} (1 - \cos t x_{\nu k}) \right\} dt.$$

Set

$$\alpha_{\nu k} = \lambda_k^2 p_{\nu k} \left(\sum_{k=1}^n \sum_{\nu=1}^{N_k} \lambda_k^2 p_{\nu k} \right)^{-1} = \lambda_k^2 p_{\nu k} A^{-1}$$

where from (2.13), and since $L \geq \max_k \lambda_k$,

$$(2.15) \quad A = \sum_{k=1}^n \sum_{\nu=1}^{N_k} \lambda_k^2 p_{\nu k} \geq \sum_{k=1}^n \lambda_k^2 (1 - Q(X_k; \lambda_k)) - \varepsilon L^2.$$

We now apply the Hölder inequality to the right member of (2.14) and obtain

$$(2.16) \quad Q(S_n; L) \leq C e^\varepsilon \prod_{k=1}^n \prod_{\nu=1}^{N_k} \left\{ L \int_{-2\pi/L}^{2\pi/L} \exp \left\{ -\frac{1}{2} A \lambda_k^{-2} (1 - \cos t x_{\nu k}) \right\} dt \right\}^{\alpha_{\nu k}}.$$

Let us estimate the integrals

$$(2.17) \quad J_{\nu k} = L \int_{-2\pi/L}^{2\pi/L} \exp \left\{ -\frac{1}{2} A \lambda_k^{-2} (1 - \cos t x_{\nu k}) \right\} dt.$$

We may of course suppose that $x_{vk} > 0$. Then

$$J_{vk} = L x_{vk}^{-1} \int_{-2\pi x_{vk}/L}^{2\pi x_{vk}/L} \exp\left\{-\frac{1}{2} A \lambda_k^{-2} (1 - \cos t)\right\} dt.$$

Two cases may occur:

1. $x_{vk} \leq L$. Then

$$\begin{aligned} J_{vk} &\leq L x_{vk}^{-1} \int_{-2\pi}^{2\pi} \exp\left\{-\frac{1}{2} A \lambda_k^{-2} (1 - \cos t)\right\} dt \\ &\leq C L x_{vk}^{-1} \int_{-\pi/4}^{\pi/4} \exp\left\{-\frac{1}{2} A \lambda_k^{-2} (1 - \cos t)\right\} dt \leq C L x_{vk}^{-1} \int_{-\pi/4}^{\pi/4} \exp\left\{-\frac{1}{5} A \lambda_k^{-2} t^2\right\} dt \\ &\leq C L x_{vk}^{-1} \int_{-\infty}^{\infty} \exp\left\{-\frac{1}{5} A \lambda_k^{-2} t^2\right\} dt \leq C L A^{-1/2} (x_{vk}^{-1} \lambda_k) \leq C L A^{-1/2}. \end{aligned}$$

The last inequality follows from (2.12).

2. $x_{vk} > L$. Denoting by $[x]$ the integer part of x we obtain

$$\begin{aligned} J_{vk} &\leq [x_{vk} L^{-1}]^{-1} \int_{-2\pi([x_{vk} L^{-1}] + 1)}^{2\pi([x_{vk} L^{-1}] + 1)} \exp\left\{-\frac{1}{2} A \lambda_k^{-2} (1 - \cos t)\right\} dt \\ &\leq 2 \int_{-2\pi}^{2\pi} \exp\left\{-\frac{1}{2} A \lambda_k^{-2} (1 - \cos t)\right\} dt \leq C \lambda_k A^{-1/2} \leq C L A^{-1/2}. \end{aligned}$$

In both cases

$$(2.18) \quad J_{vk} \leq C L A^{-1/2}.$$

From (2.16), (2.17), (2.18) and (2.15) we finally obtain

$$(2.19) \quad Q(S_n; L) \leq C e^\varepsilon L \left\{ \sum_{k=1}^n \lambda_k^2 (1 - Q(X_k; \lambda_k)) - \varepsilon L^2 \right\}^{-1/2}.$$

Since, however, ε is arbitrarily small and C is an absolute constant we may set $\varepsilon = 0$ in (2.19) and the KOLMOGOROV-ROGOZIN inequality (1.3) is proved.

Remark. From the proof, especially from (2.10), it follows that $1 - Q(X_k; \lambda_k)$ in (1.3) may be replaced by $1 - Q(X_k - X'_k; \lambda_k)$ which generally gives a better estimate.

3. The Multi-dimensional Case

In this section we consider random vectors $X = (X_1, \dots, X_r)$. The following definitions and notations will be used: Set

$$Q(X; \lambda_1, \dots, \lambda_r) = \sup_x P(x_1 \leq X_1 \leq x_1 + \lambda_1, \dots, x_r \leq X_r \leq x_r + \lambda_r)$$

where $x = (x_1, \dots, x_r)$. The domain $\bar{D}(\xi; \lambda_1, \dots, \lambda_r)$ is defined by $|x_k - \xi_k| > \lambda_k/2$ ($k = 1, \dots, r$) where $\xi = (\xi_1, \dots, \xi_r)$ is fixed. The complement of \bar{D} is denoted by $D(\xi; \lambda_1, \dots, \lambda_r)$. Set

$$Q(X; D(\lambda_1, \dots, \lambda_r)) = \sup_{\xi} P(X \in D(\xi; \lambda_1, \dots, \lambda_r)).$$

By $C(r)$ we denote different positive constants depending only on r . A random vector independent of X and with the same distribution will be denoted by X' .

Consider a sequence of mutually independent random vectors

$$X^{(k)} = (X_1^{(k)}, \dots, X_r^{(k)}) \quad \text{with sum } S_n = X^{(1)} + \dots + X^{(n)}.$$

We want to estimate

$$Q(S_n; L_1, \dots, L_r) \quad \text{where } L_i \geq \max_k \lambda_i^{(k)} \quad (k = 1, \dots, n; i = 1, \dots, r)$$

and where $\lambda_1^{(k)}, \dots, \lambda_r^{(k)}$ are given sequences of positive quantities. Applying Theorem 1 we at once obtain

$$(3.1) \quad \begin{aligned} Q(S_n; L_1, \dots, L_r) &\leq Q(X_i^{(1)} + \dots + X_i^{(n)}; L_i) \\ &\leq C L_i \left\{ \sum_{k=1}^n (\lambda_i^{(k)})^2 (1 - Q(X_i^{(k)}; \lambda_i^{(k)})) \right\}^{-1/2}. \end{aligned}$$

An inequality containing the concentration functions of the random vectors themselves can be obtained in the following way. Our starting point is the inequality

$$(3.2) \quad \begin{aligned} &Q(S_n; L_1, \dots, L_r) \\ &\leq C(r) L_1 \dots L_r \int \dots \int \exp \left\{ -\frac{1}{2} \sum_{k=1}^n \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} (1 - \cos(t_1 x_1 + \dots + t_r x_r)) \times \right. \\ &\quad \left. \times dG_k(x_1, \dots, x_r) \right\} dt_1 \dots dt_r \end{aligned}$$

where G_k is the distribution function of $X^{(k)} - X^{(k)'}$. This inequality is analogous to (2.8) and is proved in a similar way by means of the auxiliary function $h(t_1) \dots h(t_r)$, $h(t)$ being defined by (2.2). A straight forward generalization of the method of proof used in the one-dimensional case yields

Theorem 2. For any positive $\lambda_i^{(k)} \leq L_i$, one has

$$(3.3) \quad \begin{aligned} &Q(S_n; L_1, \dots, L_r) \\ &\leq C(r) L_1 \dots L_r \left\{ \sum_{k=1}^n (\lambda_1^{(k)} \dots \lambda_r^{(k)})^2 (1 - Q(X^{(k)}; D(\lambda_1^{(k)}, \dots, \lambda_r^{(k)}))) \right\}^{-1/2}. \end{aligned}$$

Remark. As in the one-dimensional case, $Q(X^{(k)}; D)$ in (3.3) may be replaced by $Q(X^{(k)} - X^{(k)'}; D)$.

If the random vectors are identically distributed and $L_i = \lambda_i^{(k)} = \lambda$, the right member of (3.3) is of order $n^{-1/2}$. From the example $r = 2$, $X^{(k)} = (1, 1)$ with prob. 1/2, $X^{(k)} = (-1, -1)$ with prob. 1/2 it is seen that

$$P(S_{2n} = 0) = \binom{2n}{n} 2^{-2n} \sim (\pi n)^{-1/2}$$

and hence that the inequality (3.3) cannot generally be improved. In the above example, however, the distribution of $X^{(k)}$ is actually one-dimensional. If the distributions of the random vectors $X^{(k)}$ are non-degenerate, the inequality (3.3) is not satisfactory. This may be seen from the case where the vectors $X^{(k)}$ have the same non-degenerate multinomial distribution. Then $\sup_x P(S_n = x) = O(n^{-r/2})$.

Thus, if the random vectors $X^{(k)}$ have non-degenerate r -dimensional distributions we want to replace (3.3) by an inequality the right member of which is of order $n^{-r/2}$ in the case of identically distributed vectors.

In a special case an inequality of the desired type is easily obtained by our method. The random vector $Y = (Y_1, \dots, Y_r)$ is called sign-invariant if all the random vectors $(\pm Y_1, \dots, \pm Y_r)$ have the same distribution. We now introduce the following symmetry condition (S): *The random vector X satisfies the condition (S) if $X - X'$ is sign-invariant.*

Theorem 3. *Let $X^{(1)}, \dots, X^{(n)}$ be mutually independent random vectors with r components. If all the $X^{(k)}$ satisfy the condition (S) then for any positive $\lambda_i^{(k)} \leq L_i$ ($i = 1, \dots, r; k = 1, \dots, n$) one has*

$$(3.4) \quad \begin{aligned} & Q(S_n; L_1, \dots, L_r) \\ & \leq C(r) L_1 \dots L_r \left\{ \sum_{k=1}^n (\lambda_1^{(k)} \dots \lambda_r^{(k)})^{2/r} (1 - Q(X^{(k)}; D(\lambda_1^{(k)}, \dots, \lambda_r^{(k)}))) \right\}^{-r/2}. \end{aligned}$$

Proof. For the sake of simplicity we confine ourselves to the case where the vectors are identically distributed, $r = 2$, $L_1 = L_2 = L$, $\lambda_1^{(k)} = \lambda_2^{(k)} = \lambda$ ($k = 1, \dots, n$), $L \geq \lambda$. The inequality (3.2) then becomes

$$(3.5) \quad \begin{aligned} Q(S_n; L, L) & \leq C L^2 \int_{-2\pi/L}^{2\pi/L} \int_{-2\pi/L}^{2\pi/L} \times \\ & \times \exp \left\{ -\frac{n}{2} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (1 - \cos(t_1 x_1 + t_2 x_2)) dG(x_1, x_2) \right\} dt_1 dt_2. \end{aligned}$$

In analogy to (2.9), using the condition (S), we get uniformly for $|t_1| \leq 2\pi/L$, $|t_2| \leq 2\pi/L$

$$(3.6) \quad \begin{aligned} & \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (1 - \cos(t_1 x_1 + t_2 x_2)) dG(x_1, x_2) \\ & \geq \int_{|x_1| > \lambda/2} \int_{|x_2| > \lambda/2} (1 - \cos(t_1 x_1 + t_2 x_2)) dG(x_1, x_2) \\ & = 2 \sum_{v_1=1}^M \sum_{v_2=1}^N p_{v_1 v_2} (2 - \cos(t_1 x_{1v_1} + t_2 x_{2v_2}) - \cos(t_1 x_{1v_1} - t_2 x_{2v_2})) + \theta \end{aligned}$$

where

$$(3.7) \quad p_{v_1 v_2} > 0, \quad x_{1v_1} \geq \lambda/2, \quad x_{2v_2} \geq \lambda/2, \quad |\theta| \leq 2\varepsilon/n.$$

(We assume of course that $Q(X; D(\lambda, \lambda)) < 1$, X having the same distribution as the random vectors $X^{(k)}$.) Further

$$(3.8) \quad 4 \sum_{v_1=1}^M \sum_{v_2=1}^N p_{v_1 v_2} \geq 1 - Q(X - X'; D(\lambda, \lambda)) - \varepsilon \geq 1 - Q(X; D(\lambda, \lambda)) - \varepsilon.$$

From (3.5), (3.6) and (3.7) it follows that

$$(3.9) \quad \begin{aligned} Q(S_n; L, L) & \leq C L^2 e^\varepsilon \int_{-2\pi/L}^{2\pi/L} \int_{-2\pi/L}^{2\pi/L} \prod_{v_1, v_2} \times \\ & \times \exp \{ -n p_{v_1 v_2} (2 - \cos(t_1 x_{1v_1} + t_2 x_{2v_2}) - \cos(t_1 x_{1v_1} - t_2 x_{2v_2})) \} dt_1 dt_2. \end{aligned}$$

Setting

$$\alpha_{v_1 v_2} = p_{v_1 v_2} / \sum_{v_1=1}^M \sum_{v_2=1}^N p_{v_1 v_2} = p_{v_1 v_2} / A,$$

where on account of (3.8)

$$(3.10) \quad 4A \geq 1 - Q(X; D(\lambda, \lambda)) - \varepsilon,$$

we get from (3.9) and the Hölder inequality

$$(3.11) \quad Q(S_n; L, L) \leq C e^\varepsilon \prod_{\nu_1, \nu_2} (J_{\nu_1 \nu_2})^{\alpha_{\nu_1 \nu_2}}$$

where

$$J_{\nu_1 \nu_2} = \frac{L^2}{x_{1\nu_1} x_{2\nu_2}} \int_{-2\pi x_{1\nu_1}/L}^{2\pi x_{1\nu_1}/L} \int_{-2\pi x_{2\nu_2}/L}^{2\pi x_{2\nu_2}/L} \times \\ \times \exp\{-nA(2 - \cos(t_1 + t_2) - \cos(t_1 - t_2))\} dt_1 dt_2.$$

Making the transformation $t_1 + t_2 = u_1$, $t_1 - t_2 = u_2$ we easily find as in the one-dimensional case that

$$(3.12) \quad J_{\nu_1 \nu_2} \leq CL^2 \lambda^{-2} A^{-1} n^{-1}.$$

Since ε is arbitrarily small it follows from (3.11), (3.12) and (3.10) that

$$Q(S_n; L, L) \leq CL^2 \{n \lambda^2 (1 - Q(X; D(\lambda, \lambda)))\}^{-1}$$

which is the desired inequality.

Remark. In Theorem 3 we may replace $Q(X^{(k)}; D)$ by $Q(X^{(k)} - X^{(k)'}; D)$.

If the random vectors have non-degenerate distributions but the condition (S) is not satisfied it seems difficult to obtain a general and at the same time simple inequality analogous to (3.4) using the above method of proof. As in (3.6) we have to add r appropriately chosen terms of the form $p \cdot (1 - \cos(t_1 x_1 + \dots + t_r x_r))$ in order to obtain the desired order of magnitude. Under a more general condition than (S) one can in this way obtain an inequality where $1 - Q(X; D)$ is replaced by a certain Hellinger-integral. It may also be convenient to consider rectangular regions the sides of which are not parallel with the coordinate axes. It is also easy to generalize the method of ROSÉN to the multi-dimensional case. This leads, however, to types of inequalities not containing the concentration functions and thus extraneous to the subject of this paper.

References

- [1] KOLMOGOROV, A.: Two uniform limit theorems for sums of independent random variables. *Theor. Probab. Appl.* **1**, 384–394 (1956).
- [2] — Sur les propriétés des fonctions de concentration de M. P. Lévy. *Ann. Inst. Henri Poincaré XVI*, **1**, 27–34 (1958).
- [3] ROGOZIN, B. A.: An estimate for concentration functions. *Theor. Probab. Appl.* **6**, 94–97 (1961).
- [4] — On the increase of dispersion of sums of independent random variables. *Theor. Probab. Appl.* **6**, 97–99 (1961).
- [5] ROSÉN, B.: On the asymptotic distribution of sums of independent random variables. *Ark. Mat.* **4**, **24**, 323–332 (1961).

Department of Mathematics
The Royal Institute of Technology
Stockholm 70