A NOTE ON SUMS OF INDEPENDENT RANDOM VARIABLES

BY HIDENORI MORIMURA

§ 1. Introduction.

The problem discussed in this paper is the one being concerned indirectly with queuing theory, but it happens during the study of the method used in the other two articles on the theory [4, 5].

Let X_i $(i=1,2,\cdots)$ be a sequence of mutually independent real valued random variables with the common distribution. Put $S_n=\sum_{i=1}^{n}X_i$. Stein [6] shows that the probability that S_1, S_2, $\cdots$, S_n are lying in the interval (a, b) is of exponential order with n, where $a<0<b$ are arbitrary constants. In this case, we need not any conditions on X_i except the trivial condition as $P(X_i=0)<1$.

On the other hand, if X_i are bounded, from a result by Loève [3] it will be implied that $P(S_1>a, \cdots, S_n>a)$ is of exponential order with n provided $EX_i<0$.

When the existence of variance of X_i is assumed, the central limit theorem will be applicable. But, based on it, we can only see that the probability will converge to zero. Moreover, if a bounded condition in a sense is added, the order of the probability will be exponentially small with n by a result due to Feller [1].

By the way, if $EX_i<0$, using the results by Sparre Andersen [7] it is easily proved that

$$\sum_{n=1}^{\infty}P(S_1>0, \cdots, S_n>0)<\infty,$$

then we can conclude that

$$P(S_1>0, \cdots, S_n>0)=o\left(\frac{1}{n}\right).$$

This fact was noted by Prof. T. Kawata in his recent paper which discussed on queuing theory [2]. We shall investigate the problem when the boundedness of X_i is not assumed. In § 2 we shall give a counter example for the assersion that the order of the probability will be exponentially small with n, even if $EX_i<0$. From the example, we can see the following fact: If we assume only $-\infty<EX_i<0$, then we shall be able to assert only that the probability is $o(n^{-1})$ as $n\to\infty$. Furthermore, in § 3, we shall prove that the probability is smaller than $O(n^{-(1+\sqrt{5})/2+\varepsilon})$ under the condition that the mean and variance of X_i exist, and the former is less than zero, where ε is an arbitrary positive number.

Received September 9, 1961.

§ 2. An example.

Suppose that $\{X_i\}$ is a sequence of integral valued random variables with the common distribution defined as

$$\text{(2.1)} \qquad \begin{cases} P(X_i = k) = A \cdot k^{-(2+\varepsilon)} & (k \geqq 1), \\ P(X_i = -1) = q, \\ P(X_i = \text{other value}) = 0, \end{cases}$$

where ε is a positive number. Furthermore let A and q be some positive constants such that

$$\text{(2.2)} \qquad A \cdot \zeta(2+\varepsilon) + q = 1, \qquad A \cdot \zeta(1+\varepsilon) < q.$$

Thus, we have

$$\text{(2.3)} \qquad -\infty < E(X_i) < 0.$$

Now, since

$$(S_1 > 0, \cdots, S_n > 0) \subset (S_1 > 0) = (X_1 > 0),$$

we have

$$\text{(2.4)} \qquad \begin{aligned} P(S_1 > 0, \cdots, S_n > 0) &= \sum_{k=1}^{\infty} P(X_1 = k) P(S_2 > 0, \cdots, S_n > 0 \mid X_1 = k) \\ &\geqq A \cdot \sum_{k=1}^{\infty} \frac{1}{k^{2+\varepsilon}} \sum_{j=0}^{k-1} \binom{n-1}{j} q^j (1-q)^{n-1-j}. \end{aligned}$$

Putting $n' = [(n-1)q] + 1$, we further get

$$\text{(2.5)} \qquad \geqq A \cdot \sum_{k=n'}^{\infty} \frac{1}{k^{2+\varepsilon}} \sum_{j=0}^{k-1} \binom{n-1}{j} q^j (1-q)^{n-1-j},$$

where the second summation for $k > n$ will be regarded as unity. Hence, by the Laplace's theorem

$$P(S_1 > 0, \cdots, S_n > 0) \geqq A \cdot C \sum_{k=n'}^{\infty} \frac{1}{k^{2+\varepsilon}}$$

where C is a constant such as $0 < C < 1$. Then we get

$$\text{(2.6)} \qquad P(S_1 > 0, \cdots, S_n > 0) \geqq O\left(\frac{1}{n'^{1+\varepsilon}}\right) = O\left(\frac{1}{n^{1+\varepsilon}}\right).$$

Thus, the assertion that $P(S_1 > 0, \cdots, S_n > 0)$ will be a infinitesimal quantity of exponential order of n is neglected, even if $E(X_i) < 0$ is assumed. Furthermore, combining with the fact that $P(S_1 > 0, \cdots, S_n > 0) = o(1/n)$ which was stated in § 1, we can say that this result is rather severe, the assertion of smaller order than this is impossible without some additional restriction.

In above axample, we may substitute $3 + \varepsilon$ instead of $2 + \varepsilon$, without changing almost calculation. Then the example shows us that $P(S_1 > 0, \cdots, S_n > 0)$ cannot be of smaller order than $O(n^{-2-\varepsilon})$ under an additional restriction:

$$E(X_i^2) < \infty. \tag{2.7}$$

In the following, we shall prove that, for every $\delta < (1+\sqrt{5})/2 = 1.618\cdots$,

$$P(S_1 > 0, \cdots, S_n > 0) \leqq O(n^{-\delta})$$

in this case. We cannot verify yet that

$$P(S_1 > 0, \cdots, S_n > 0) \leqq O(n^{-\delta}), \quad \frac{1+\sqrt{5}}{2} \leqq \delta \leqq 2$$

is true or not.

§ 3. An evaluation.

In this section, we shall prove the following

THEOREM. *Let $\{X_i\}$ $(i=1,2,\cdots)$ be a sequence of independent random variables with common distribution and denote as $S_n = \sum_{i=1}^n X_i$.*

If $-\infty < E(X_i) \equiv \mu < 0$ and $E(X_i^2) < \infty$, then

$$P(S_1 > a,\ S_2 > a,\ \cdots,\ S_n > a) \leqq M_n,$$

where

$$M_n = O(n^{-\delta}) \qquad (n \to \infty)$$

for every $\delta < (1+\sqrt{5})/2$ and a is an arbitrary number.

Proof. Since, for $a' > a$, it is evident that

$$P(S_1 > a',\ S_2 > a',\ \cdots,\ S_n > a') \leqq P(S_1 > a,\ S_2 > a,\ \cdots,\ S_n > a),$$

it is sufficient to prove the theorem for $a < 0$. Thus hereafter we shall assume that a is a negative number. Choosing a constant c such as $c > a$, we have

$$\begin{aligned} &P(S_1 > a,\ S_2 > a,\ \cdots,\ S_n > a) \\ &= P(S_1 > a,\ S_2 > a,\ \cdots,\ c \geqq S_\nu > a,\ \cdots,\ S_n > a) \\ &\quad + P(S_1 > a,\ S_2 > a,\ \cdots,\ S_\nu > c,\ \cdots,\ S_n > a) \\ &= I_1 + I_2, \qquad \text{say}. \end{aligned} \tag{3.1}$$

Now, if we put as $Z_{\nu,n} = \sum_{i=\nu+1}^n X_i$, we can evaluate as follows:

$$\begin{aligned} I_1 &\leqq P(S_1 > a,\ S_2 > a,\ \cdots,\ c \geqq S_\nu > a,\ S_\nu + Z_{\nu,n} > a) \\ &\leqq P(Z_{\nu,n} > a - c) P(S_1 > a,\ S_2 > a,\ \cdots,\ c \geqq S_\nu > a) \\ &\leqq P(Z_{\nu,n} > a - c) \cdot \varphi_\nu, \end{aligned} \tag{3.2}$$

where

$$\varphi_\nu = P(S_1 > a,\ S_2 > a,\ \cdots,\ S_\nu > a). \tag{3.3}$$

Moreover, we get

$$I_2 = P(S_1 > a,\ S_2 > a,\ \cdots,\ S_{\nu-1} > a,\ S_\nu > c,\ \cdots,\ S_n > a) \tag{3.4}$$
$$\leqq P(S_\nu > c) = P(S_\nu - \nu\mu > c - \nu\mu).$$

If $c - \nu\mu > 0$, the Chebyshev's inequality implies that

$$I_2 \leqq P(S_\nu - \nu\mu > c - \nu\mu) \leqq \frac{\nu\sigma^2}{(c-\nu\mu)^2}, \tag{3.5}$$

where $\sigma^2 = E(X_i^2) - \mu^2$.

Similarly, we can evaluate $P(Z_{\nu,n} > a - c)$ using the inequality when $a - c - (n-\nu)\mu > 0$. We have

$$P(Z_{\nu,n} > a - c) = P(Z_{\nu,n} - (n-\nu)\mu > a - c - (n-\nu)\mu) \tag{3.6}$$
$$\leqq \frac{(n-\nu)\sigma^2}{\{a - c - (n-\nu)\mu\}^2}.$$

Combining (3.1), (3.2), (3.3), (3.5) and (3.6) we can get

$$\varphi_n \leqq \varphi_\nu \frac{(n-\nu)\sigma^2}{\{a-c-(n-\nu)\mu\}^2} + \frac{\nu\sigma^2}{(c-\mu\nu)^2}. \tag{3.7}$$

Now, for any $\varepsilon > 0$ we can choose a positive integer N such as $\beta^N/(1-\beta) < \varepsilon$, where $\beta < 1$ is a positive number which will be determined suitably later. And, for this integer N, there exist an integer $n_0(N)$ and a sequence $\{\nu_k\}$ $(k = 0, 1, 2, \cdots, N)$ such as for $n \geqq n_0(N)$

$$\begin{cases} \nu_0 = n, \\ \nu_{k+1} = [\nu_k{}^\beta], \\ \nu_k > \dfrac{a}{\mu(1-\gamma)}, \qquad k = 0, 1, \cdots, N, \end{cases} \tag{3.8}$$

where $\gamma < 1$ is a positive number which will be chosen suitably such as

$$(\nu_{k+1} - \gamma\nu_k)\mu > 0 \qquad \text{for } k = 0, 1, \cdots, N. \tag{3.9}$$

Thus, we can find easily that

$$\nu_0 > \nu_1 > \cdots > \nu_N > 0$$

and that

$$c > a, \qquad a - c - (\nu_k - \nu_{k+1})\mu > 0,$$

setting the left hand of (3.9) as c.

Then, we can evaluate φ_n using the relation (3.7), (3.8) and (3.9) as follows:

$$\varphi_n \leqq \varphi_{\nu_1} \frac{(\nu_0 - \nu_1)\sigma^2}{\{a - (-\gamma\nu_0 + \nu_1 + \nu_0 - \nu_1)\mu\}^2} + \frac{\nu_1\sigma^2}{\{(\nu_1 - \gamma\nu_0 - \nu_1)\mu\}^2}$$
$$= \varphi_{\nu_1} \frac{(\nu_0 - \nu_1)\sigma^2}{\{a - (1-\gamma)\nu_0\mu\}^2} + \frac{\nu_1\sigma^2}{\gamma^2\nu_0^2\mu^2}$$
$$\leqq \frac{\sigma^2}{\gamma^2\mu^2}\frac{\nu_1}{\nu_0^2} + \frac{\sigma^2}{\gamma^2\mu^2}\frac{\nu_2}{\nu_1^2}\frac{(\nu_0 - \nu_1)\sigma^2}{\{a - (1-\gamma)\nu_0\mu\}^2} \tag{3.10}$$

$$+\frac{\sigma^2}{\gamma^2\mu^2}\frac{\nu_3}{\nu_2^2}\frac{(\nu_1-\nu_2)\sigma^2}{\{a-(1-\gamma)\nu_1\mu\}^2}\frac{(\nu_0-\nu_1)\sigma^2}{\{a-(1-\gamma)\nu_0\mu\}^2}+\cdots$$
$$+\varphi_{\nu_N}\frac{(\nu_{N-1}-\nu_N)\sigma^2}{\{a-(1-\gamma)\nu_{N-1}\mu\}^2}\cdots\frac{(\nu_0-\nu_1)\sigma^2}{\{a-(1-\gamma)\nu_0\mu\}^2}.$$

Here, we wish to choose β suitably in order to get the best evaluation of φ_n in this case. Since, such will give the same order of n for all terms in the last expression of (3.10), first of all, we shall examine to make equal the order of the first two terms.

From (3.8) we have

$$\text{1st term} = O(n^{-(2-\beta)}),$$
$$\text{2nd term} = O(n^{-\{1+\beta(2-\beta)\}}),$$

and

$$2-\beta=1+\beta(2-\beta), \tag{3.11}$$

or

$$\beta^2-3\beta+1=0. \tag{3.12}$$

Thus, if we choose β as this, since

$$\text{3rd term} = O(n^{-\{1+\beta+\beta^2(2-\beta)\}})$$

we have

$$\text{3rd term} = O(n^{-[1+\beta\{1+\beta(2-\beta)\}]})$$
$$= O(n^{-[1+\beta(2-\beta)]}) = O(n^{-(2-\beta)}).$$

Quite similarly, we can see that all terms in (3.10) except the last term are $O(n^{-(2-\beta)})$. The last term in (3.10) will be evaluate as

$$O(n^{-(1+\beta+\beta^2+\cdots+\beta^{N-1})}) = O(n^{-(1-\beta^N)/(1-\beta)}) < O(n^{-\{1/(1-\beta)-\varepsilon\}}).$$

Thus, we can evaluate it for sufficiently large n by an order which is arbitrary near to the order

$$O(n^{-1/(1-\beta)}) = O(n^{-(2-\beta)}), \tag{3.13}$$

because (3.12) will be rewritten as

$$(1-\beta)(2-\beta)=1.$$

Thus, we can see that the positive root of the quadratic equation (3.12) less than 1 will give the best evaluation of φ_n in our case. It is

$$\beta=\frac{3-\sqrt{5}}{2}.$$

Thus, we have the following evaluation

$$\varphi_n \leqq O(n^{-\delta}),$$

where

$$\delta < 2 - \frac{3-\sqrt{5}}{2} = \frac{1+\sqrt{5}}{2}.$$

This is the assertion of our theorem.

In conclusion, the author expresses his sincerest thanks to Prof. T. Kawata and Dr. H. Hatori who have given valuable advices to him.

REFERENCES

[1] FELLER, W., Generalization of a probability limit theorem of Cramer. Trans. Amer. Math. Soc. 54 (1943), 361–372.
[2] KAWATA, T., On the imbedded queuing process of general type. Bull. Inst. Intern. Statist. 38 (Actes de la 32 session) 3e Liv. (1961), 445–455.
[3] LOÈVE, M., Probability theory. Van Nostrand, 1955.
[4] MORIMURA, H., On the number of served customers in a busy period. J. Oper. Res. Soc. Japan 4 (1961), 68–76.
[5] MORIMURA, H., On the relation between the distributions of the queue size and the waiting time. To appear in Kōdai Math. Sem. Rep.
[6] STEIN, C., A note on cumulative sums. Ann. Math. Statist. 17 (1946), 498–499.
[7] SPARRE ANDERSEN, E., On the fluctuation of sums of random variables, II. Math. Scand. 2 (1954), 195–223.

DEPARTMENT OF MATHEMATICS,
TOKYO INSTITUTE OF TECHNOLOGY.