The Annals of Mathematical Statistics
1969, Vol. 40, No. 6, 1980–1984

THE MARKOV INEQUALITY FOR SUMS OF INDEPENDENT RANDOM VARIABLES[1]

BY S. M. SAMUELS

Purdue University

The purpose of this paper is to prove the following theorem.

THEOREM. *Let $S_n = X_1 + \cdots + X_n$ be a sum of n independent, non-negative random variables with means*

$$\mathbf{v} = (\nu_1, \cdots, \nu_n) = (EX_1, \cdots, EX_n)$$
$$N = \nu_1 + \cdots + \nu_n;$$

and, for each $\lambda > N$, let

$$\psi_n(\lambda; \mathbf{v}) = \sup P(S_n \geqq \lambda),$$

where the supremum is taken over all such S_n. (We ignore $\lambda \leqq N$ since the supremum is trivially one.) *Then,*

$$\lambda \geqq [\max(4, n-1)]N \Rightarrow \psi_n(\lambda; \mathbf{v}) = 1 - \textstyle\prod_{1 \leqq i \leqq n} (1 - \nu_i/\lambda),$$

which is attained if and only if, for each i,

$$P(X_i = \lambda) = \nu_i/\lambda = 1 - P(X_i = 0).$$

Since these X_i's are identically distributed when the means are equal, we have an immediate

COROLLARY. *Let $\{X_i : 1 \leqq i \leqq n\}$ be* i.i.d., *non-negative, with common mean ν. If $\lambda > [\max(4n, (n-1)n)]\nu$, then*

$$P(X_1 + \cdots + X_n \geqq \lambda) \leqq 1 - (1 - \nu/\lambda)^n.$$

Equality holds if and only if $X_i \varepsilon \{0, \lambda\}$.

We shall present an outline of the proof as a series of lemmas. The first three lemmas show that, to prove the theorem, it suffices to prove the proposition following Lemma 3. The remaining three lemmas constitute a proof of that proposition. After stating the six lemmas, we sketch their proofs. Finally, there is a brief discussion of how the theorem may be improved.

LEMMA 1. *Without loss of generality we may assume that each X_i has at most two mass points—call them a_i and b_i—satisfying:*

$$(1) \qquad 0 \leqq a_i < \nu_i \leqq b_i \leqq \lambda,$$
$$P(X_i = b_i) = (\nu_i - a_i)/(b_i - a_i) = 1 - P(X_i = a_i).$$

Received 30 April 1968; revised 28 October 1968.

[1] This research was supported in part by the Office of Naval Research Contract NONR 1100(26) at Purdue University and in part by National Science Foundation Grant GP-9400, also at Purdue University.

Moreover, if we let

(2) $\psi_n(\lambda; \mathbf{v}, \mathbf{a}) = \sup_{\{S_n:(a_1,\cdots,a_n)=a\}} P(S_n \geqq \lambda), \qquad A = a_1 + \cdots + a_n,$

then $\psi_n(\lambda; \mathbf{v}, \mathbf{a})$ *is attained and*

(3) $\psi_n(\lambda; \mathbf{v}) = \max_{0 \leqq a < \nu} \psi_n(\lambda; \mathbf{v}, \mathbf{a}) = \max_{0 \leqq a < \nu} \psi_n(\lambda - A; \mathbf{v} - \mathbf{a}, \mathbf{0}).$

LEMMA 2. *For any* $H \subset \{1, 2, \cdots, n\}$, *let*

$$N_H = \sum_{i \varepsilon H} \nu_i, \qquad N_k = \nu_1 + \cdots + \nu_k \quad (1 \leqq k < n), \qquad N_0 = 0.$$

Then, for $n \geqq 2$,

$$\begin{aligned} &\max_{0 \leqq a < \nu} [1 - \textstyle\prod_{1 \leqq i \leqq n} (1 - (\nu_i - a_i)/(\lambda - A))] \\ (4) \qquad &= \max_{H \subset \{1,2,\cdots,n\}} [1 - \textstyle\prod_{i \notin H} (1 - \nu_i/(\lambda - N_H))] \\ &= \max_{0 \leqq k \leqq n-1} [1 - \textstyle\prod_{i > k} (1 - \nu_i/(\lambda - N_k))] \quad (\text{if } \nu_1 \leqq \nu_2 \leqq \cdots \leqq \nu_n) \\ &= 1 - \textstyle\prod_{i \leqq i \leqq n} (1 - \nu_i/\lambda) \qquad (\text{if } \lambda \geqq 2N). \end{aligned}$$

(It should be remarked that this lemma can be restated in terms of random variables as follows. Among all S_n with $b_i - a_i = \lambda - A$ for each i, $P(S_n \geqq \lambda)$ is maximized only if, for each i, either $a_i = 0$ or $b_i = \nu_i$. Moreover, the maximum is only attained if the i's for which $b_i = \nu_i$ correspond to the *smallest* means. Finally if λ is sufficiently large, all the a_i's must be zero.)

LEMMA 3. *To prove the theorem it suffices to prove the following.*

PROPOSITION.

$$\lambda \geqq [\max (4, n - 1)]N \Rightarrow \psi_n(\lambda; \mathbf{v}, \mathbf{0}) = 1 - \textstyle\prod_{1 \leqq i \leqq n} (1 - \nu_i/\lambda).$$

Since the proposition is well-known to be true for $n = 1$, we shall proceed by induction.

LEMMA 4. *Assume that the proposition is true for* $n - 1$ *and that* $\lambda \geqq [\max (4, n - 1)]N$. *Suppose* S_n *attains* $\psi_n(\lambda; \mathbf{v}, \mathbf{0})$—*with each* $a_i = 0$, *of course. Then, for each* i, $b_i < N - \nu_i$ *or* $b_i > \frac{1}{2}\lambda$; *hence*

$$\sum_{\{i: b_i < N - \nu_i\}} b_i < \sum_{1 \leqq i \leqq n} (N - \nu_i) = (n - 1)N < \lambda.$$

If, for some i, $b_i = \lambda$, *then*

$$(5) \qquad P(S_n \geqq \lambda) = 1 - \textstyle\prod_{1 \leqq i \leqq n} (1 - \nu_i/\lambda).$$

LEMMA 5. *Suppose* S_n *is of the following form, for some* $H \subset \{1, 2, \cdots, n\}$:

$$a_i = 0 \quad \text{for each} \quad i, \qquad \sum_{i \varepsilon H} b_i < \lambda, \qquad i \notin H \Rightarrow \tfrac{1}{2}\lambda < b_i < \lambda.$$

If $\lambda > 4N$, *then*

$$P(S_n \geqq \lambda) \leqq 1 - \textstyle\prod_{i \notin H} (1 - \nu_i/(\lambda - N_H)).$$

LEMMA 6. *The proposition is true.*

PROOF OF LEMMA 1. The first statement is standard (see, e.g., [2]), as is the fact that (2) is attained. To prove the second equality of (3), we suppose $S_n = \sum X_i$ and $T_n = \sum Y_i$ are appropriate random variables each attaining one of the upper bounds. Then, by definition (2),

$$\begin{aligned}\psi_n(\lambda; \mathbf{v}, \mathbf{a}) &= P(\sum X_i \geqq \lambda) = P(\sum (X_i - a_i) \geqq \lambda - A) \\ &\leqq \psi_n(\lambda - A; \mathbf{v} - \mathbf{a}, \mathbf{0}) = P(\sum Y_i \geqq \lambda - A) = P(\sum (Y_i + a_i) \geqq \lambda) \\ &\leqq \psi_n(\lambda; \mathbf{v}, \mathbf{a}).\end{aligned}$$

Proof of Lemma 2. The first equality of (4) is given by Lemma 2.3 of [3], while the second follows from formula (4.3) of [3]. To prove the last equality we first note that, since $\lambda > N$ and $\nu_1 \leqq \nu_2 \leqq \cdots \leqq \nu_n$, we have

$$\begin{aligned}&[\textstyle\prod_{i>1} (1 - \nu_i/(\lambda - \nu_i))]/[\prod_{i>0} (1 - \nu_i/\lambda)] \\ (6) \quad &\geqq \lambda^n(\lambda - 2\nu_1)^{n-2}(\lambda - N + (n-2)\nu_1)/(\lambda - \nu_1)^{2n-2}(\lambda - N + (n-1)\nu_1) \\ &= [C_0 + (C_1 + \lambda - 2N)\nu_1 + O(\nu_1^2)]/[C_0 + C_1\nu_1 + O(\nu_1^2)] \\ &> \lambda^2(\lambda - N)/(\lambda - \nu_1)^2(\lambda - N + \nu_1).\end{aligned}$$

The first and second expressions are equal when $\nu_2 = \cdots = \nu_{n-1} = \nu_1$, $\nu_n = N - (n-1)\nu_1$. From the third expression we see that, if $\lambda < 2N$ and $\nu_1 = \cdots = \nu_{n-1}$ is positive but sufficiently close to zero, then (6) is less than one. If, on the other hand, $\lambda \geqq 2N$, then the fourth expression (which is obtained by setting $\nu_2 = \cdots = \nu_{n-1} = 0$, $\nu_n = N - \nu_1$) is an increasing function of ν_1; hence (6) is greater than one.

Replacing λ by $\lambda - N_k$ and repeating the argument we find that, if $\lambda - N_k \geqq 2(N - N_k)$—which is implied by $\lambda \geqq 2N$—then,

$$[\textstyle\prod_{i>k+1} (1 - \nu_i/(\lambda - N_{k+1}))]/[\prod_{i>k} (1 - \nu_i/(\lambda - N_k))] > 1.$$

This not only completes the proof of Lemma 2, but also shows that in the theorem itself, the quantity "max $(4, n-1)$" cannot be replaced by anything smaller than 2. This point will be elaborated upon below.

Proof of Lemma 3. If the proposition is true as stated, then since

$$\lambda \geqq [\max(4, n-1)]N \Rightarrow \lambda - A \geqq [\max(4, n-1)](N - A),$$

the same hypothesis implies

$$\psi_n(\lambda - A; \mathbf{v} - \mathbf{a}, \mathbf{0}) = 1 - \textstyle\prod_{1 \leqq i \leqq n} (1 - (\nu_i - a_i)/(\lambda - A))$$

for all $\mathbf{a} < \mathbf{v}$. The theorem then follows from (3) and (4).

Proof of Lemma 4. By hypothesis and by definition (2), we have, for each i,

$$\begin{aligned}&1 - (1 - \nu_i/\lambda) \textstyle\prod_{j \neq i} (1 - \nu_j/\lambda) \\ (7) \quad &\leqq \psi_n(\lambda; \mathbf{v}, \mathbf{0}) = P(S_n \geqq \lambda) \\ &= P(X_i = 0)P(S_n - X_i \geqq \lambda) + P(X_i = b_i)P(S_n - X_i \geqq \lambda - b_i) \\ &\leqq P(X_i = 0)\psi_{n-1}(\lambda - \nu_i; \mathbf{v}^*, \mathbf{0}) + P(X_i = b_i)P(S_n - X_i \geqq \lambda - b_i),\end{aligned}$$

where $\mathbf{v}^* = (\nu_1, \cdots, \nu_{i-1}, \nu_{i+1}, \cdots, \nu_n)$. But

$$[\max(4, n-1)]N \geqq [\max(4, n-2)](N - \nu_i),$$

so, by the induction hypothesis,

$$\psi_{n-1}(\lambda - \nu_i; \mathbf{v}^*, \mathbf{0}) = 1 - \prod_{j\neq i}(1 - \nu_j/\lambda). \tag{8}$$

If $b_i = \lambda$, then, using (1), we obtain (5) immediately. If $b_i < \lambda$, we use the ordinary Markov inequality, which gives

$$P(S_n - X_i \geqq \lambda - b_i) \leqq (N - \nu_i)/(\lambda - b_i). \tag{9}$$

Rewriting (7), we have, using (1) and (9),

$$\prod_{j\neq i}(1 - \nu_j/\lambda) \geqq \lambda(\lambda - b_i - N + \nu_i)/(\lambda - b_i)^2. \tag{10}$$

We now use the elementary bound

$$\prod_{j\neq i}(1 - \nu_j/\lambda) \leqq 1 - (N - \nu_i)/\lambda + (N - \nu_i)^2/2\lambda^2$$

and the notation $\Gamma_i = (N - \nu_i)/\lambda$, $x_i = b_i/\lambda$. Substituting into (10) and rewriting it we obtain

$$(1 - \Gamma_i + \tfrac{1}{2}\Gamma_i^2)x_i^2 - (1 - \Gamma_i)^2 x_i + \tfrac{1}{2}\Gamma_i^2 \geqq 0. \tag{11}$$

By hypothesis, $\Gamma_i \leqq \frac{1}{4}$. It is easy to check that the left side of (11) is negative for $\Gamma_i \leqq x_i \leqq \frac{1}{2}$, which proves the lemma.

Proof of Lemma 5. By hypothesis, for each $i \not\in H$,

$$\begin{aligned} P(S_n < \lambda) = {} & P(X_i = b_i)P(\textstyle\sum_{k\varepsilon H} X_k < \lambda - b_i)\prod_{j\not\in H, j\neq i} P(X_j = 0) \\ & + P(X_i = 0)[\textstyle\prod_{j\not\in H, j\neq i} P(X_j = 0)] \\ & \cdot\{1 + \textstyle\sum_{j\not\in H, j\neq i} P(X_j = b_j)P(\sum_{k\varepsilon H} X_k < \lambda - b_j)/P(X_j = 0)\}. \end{aligned}$$

Substituting from (1) and using the Markov inequality, which gives

$$P(\textstyle\sum_{k\varepsilon H} X_k < \lambda - b_j) \geqq \max[0, 1 - N_H/(\lambda - b_j)],$$

we obtain—assuming $b_i < \lambda - N_H$:

$$P(S_n < \lambda) \geqq C_1 - C_2[N_H + (\lambda - b_i)D_i]/b_i(\lambda - b_i) \tag{12}$$

where $C_2 > 0$ and

$$\begin{aligned} D_i &= \textstyle\sum_{j\not\in H, j\neq i}[\nu_j/(b_j - \nu_j)]\max[0, 1 - N_H/(\lambda - b_j)] \\ &\leqq [(N - N_H)/(\lambda/2 - N_H)][1 - 2N_H/\lambda] \\ &= 2(N - N_H)/\lambda. \end{aligned}$$

It is easy to verify that the minimum of the right side of (12) in the interval $\lambda/2 \leqq b_i \leqq \lambda - N_H$ is attained for $b_i = \lambda - N_H$, provided $\lambda > 4N$.

Thus, under the hypothesis,

$$\begin{aligned} P(S_n < \lambda) &\geqq \textstyle\prod_{i\not\in H}(1 - \nu_i/b_i) \qquad (b_i \geqq \lambda - N_H) \\ &\geqq \textstyle\prod_{i\not\in H}(1 - \nu_i/(\lambda - N_H)) \end{aligned}$$

as was to be proved.

PROOF OF LEMMA 6. Lemmas 4 and 5 insure that $\lambda > [\max(4, n-1)]N$ implies

$$\psi_n(\lambda; \mathbf{v}, \mathbf{0}) \leqq \max_{H\subset\{1,2,\cdots,n\}} [1 - \prod_{i\notin H} (1 - \nu_i/(\lambda - N_H))].$$

The proposition then follows from Lemma 2.

This completes the proof of the theorem.

IMPROVING THE THEOREM. Let us define

$$C_n(\mathbf{v}) = \inf\{C: \lambda \geqq C \Rightarrow \psi_n(\lambda; \mathbf{v}) = 1 - \prod_{1\leqq i\leqq n} (1 - \nu_i/\lambda)\}.$$

The theorem states that

$$C_n(\mathbf{v}) \leqq [\max(4, n-1)]N. \tag{13}$$

In the proof of Lemma 2, we showed that

$$\sup_\nu C_n(\mathbf{v})/N \geqq 2 \quad \text{for} \quad n \geqq 2. \tag{14}$$

(Of course $C_1(\nu) = \nu$ by the Markov inequality.) If, as we have conjectured in [3], $\psi_n(\lambda; \mathbf{v})$ is given by (4) for all $\lambda > N$, then not only does equality hold in (14), but also it can be shown that

$$[C_n(\mathbf{v}), \infty) = \{\lambda: \psi_n(\lambda; \mathbf{v}) = 1 - \prod_{1\leqq i\leqq n} (1 - \nu_i/\lambda)\}.$$

The conjecture is known to be true for $n \leqq 4$ (see [4]); hence in (13) and in the theorem, the "4" can be replaced by "2".

What about $\inf_\nu C_n(\mathbf{v})/N$? For $n = 2$, it is $\frac{1}{4}(3 + 5^{\frac{1}{2}})$. We suspect that, by taking $\nu_1 = \cdots = \nu_2 = N/n$ and $\lambda = (1 + \alpha)N$ in (6), we can show that, if the conjecture is true, then

$$\inf_\nu C_n(\mathbf{v})/N \to 1 \quad \text{as} \quad n \to \infty.$$

But we have not done so.

Acknowledgment. The proof owes a great deal to J. R. B. Kemperman. In [4], I showed that $C_n(\mathbf{v}) < \infty$, but the form of my proof did not allow me to estimate it. Kemperman showed that more mileage could be gotten from the Markov inequality than I had dared to hope. He used it in [1] to obtain the bound $C_n(\mathbf{v}) < 9N^3/(\min \nu_i)^2$. Lemma 4, which is the crux of my proof, is merely a refinement of his basic idea.

REFERENCES

[1] KEMPERMAN, J. R. B. (Written Communication.)

[2] MULHOLLAND, H. P. and ROGERS, C. A. (1958). Representation theorems for distribution functions. *Proc. London Math. Soc.* **8** 177–223.

[3] SAMUELS, S. M. (1966). On a Chebyshev-type inequality for sums of independent random variables. *Ann. Math. Statist.* **37** 248–259.

[4] SAMUELS, S. M. (1968). More on a Chebyshev-type inequality for sums of independent random variables. *Purdue Stat. Dept. Mimeo. Series* No. 155.