The Annals of Probability
1973, Vol. 1, No. 3, 504-508

A NOTE ON THE RATE OF CONVERGENCE AND ITS APPLICATIONS[1]

BY RASUL. A. KHAN

University of Wisconsin

Let S_n denote the partial sums of i.i.d. random variables with mean zero and moment generating function existing in some neighborhood of the origin. We give explicit upper bounds for $P_m^+ = P(S_n \geqq a + bn$ for some $n \geqq m)$ and $P_m = P(|S_n| \geqq a + bn$ for some $n \geqq m)$, $a \geqq 0$, $b > 0$. These bounds immediately give the rate of convergence for the strong law of large numbers. An application is also made to a sequential selection procedure.

1. Introduction and summary. Let $X_1, X_2, \cdots$ be i.i.d. random variables with $EX_1 = 0$, and set $S_n = X_1 + \cdots + X_n$. Assuming the existence of moment generating function (mgf) in some neighborhood of the origin, we derive upper bounds for $P_m^+ = P(S_n \geqq a + bn$ for some $n \geqq m)$ and $P_m = P(|S_n| \geqq a + bn$ for some $n \geqq m)$, $a \geqq 0$, $b > 0$. This problem for more general boundaries $f(n)$ has been considered by Robbins (see [2] and the references therein). For $a = 0$, the bound for P_m gives the rate of convergence for the strong law of large numbers. These bounds are shown to be useful for certain sequential selection procedures.

2. Bounds for P_m^+ and P_m. We begin with

$$P_1^+ = P(S_n \geqq a + bn \text{ for some } n \geqq 1)\,, \qquad a > 0,\ b > 0\,.$$

Assume that $\phi(\theta) = Ee^{\theta X_1} < \infty$ for $0 < \theta \leqq C \leqq \infty$. Noting that $Z_n = e^{\theta S_n}[\phi(\theta)]^{-n}$ is a positive martingale with $EZ_1 = 1$, we obtain

$$P\left(S_n \geqq \frac{\log \varepsilon}{\theta} + n\frac{\log \phi(\theta)}{\theta} \text{ for some } n \geqq 1\right) \leqq 1/\varepsilon\,. \tag{1}$$

Setting $(\log \varepsilon)/\theta = a > 0$, $\psi(\theta) = \log \phi(\theta)$, (1) gives

$$P\left(S_n \geqq a + \frac{n\psi(\theta)}{\theta} \text{ for some } n \geqq 1\right) \leqq e^{-a\theta} \tag{2}$$

uniformly in $0 < \theta \leqq C$. From (2) it follows that

$$P(S_n \geqq a + bn \text{ for some } n \geqq 1) \leqq \inf_{\{0<\theta\leqq C:\psi(\theta)\leqq b\theta\}} e^{-a\theta} = e^{-a\alpha(b)} \tag{3}$$

where $\alpha(b) = \sup_{0<\theta\leqq C} \{\theta : \psi(\theta) \leqq b\theta\}$ $(b > 0)$. Note that $\alpha(b) > 0$.

LEMMA. *The function $\alpha(b)$ defined above is continuous and non-decreasing. Moreover, if there exists a $t_0(b)$ such that $\psi(t_0(b)) = bt_0(b)$ (i.e. if the equation $\psi(\theta) = b\theta$*

Received July 3, 1971; revised September 29, 1972.

[1] Sponsored by the United States Army under Contract No.: DA-31-124-ARO-D-462. A part of this paper was written while the author was at Columbia University.

AMS 1970 *subject classifications.* Primary 60G50; Secondary 60F99.

Key words and phrases. Upper bound, rate of convergence, strong law of large numbers, moment generating function, sequential selection procedure.

($b > 0$) has a solution, which is necessarily positive and unique), then $\alpha(b) = t_0(b)$. The equation $\psi(\theta) = b\theta$ has a solution if $\lim_{\theta\to C} \psi'(\theta) = \infty$.

PROOF. $\psi(\theta)$ is a continuous convex function with $\psi(0) = 0$ and $\psi'(0+) = 0$.

Now consider $P(S_n \geqq bn$ for some $n \geqq 1)$, $b > 0$. The line $f_s(n) = sn + b - s$ $(0 < s \leqq b)$ passes through the point $(1, b)$, so that

$$P(S_n \geqq f_b(n) \text{ for some } n \geqq 1) \leqq P(S_n \geqq f_s(n) \text{ for some } n \geqq 1)$$

uniformly in $0 < s \leqq b$. Thus

$$P(S_n \geqq bn \text{ for some } n \geqq 1) \leqq \exp\{-\sup_{0<s\leqq b} (b - s)\alpha(s)\}\,. \tag{4}$$

Hence from (3) and (4) we have

$$\begin{aligned} P(S_n &\geqq a + bn \text{ for some } n \geqq 1) \\ &\leqq \min\,[\exp(-a\alpha(b)), \exp\{-\sup_{0<s\leqq b} (b - s)\alpha(s)\}] \end{aligned} \tag{5}$$

where $a \geqq 0$, $b > 0$.

Considering the line $f_s(n) = s(n - 1) + a + b$ $(0 < s \leqq b)$ through the point $(1, a + b)$, one also obtains

$$P(S_n \geqq a + bn \text{ for some } n \geqq 1) \leqq \exp[-\sup_{0<s\leqq b} (a + b - s)\alpha(s)]\,. \tag{6}$$

EXAMPLE. Let P denote the probability under which $X_1, X_2, \cdots$ are i.i.d. $N(0, 1)$ random variables. Then $\phi(\theta) = e^{\theta^2/2}$, and $\psi(\theta) = b\theta$ gives $\alpha(b) = 2b$. Thus (5) gives

$$P(S_n \geqq a + bn \text{ for some } n \geqq 1) \leqq \min\,[e^{-2ab}, e^{-b^2/2}]\,, \tag{7}$$

which is better than the usual bound e^{-2ab} $(a \geqq 0, b > 0)$. Further, (6) gives

$$P(S_n \geqq a + bn \text{ for some } n \geqq 1) \leqq \exp(-(a + b)^2/2) \qquad \text{if} \quad b \geqq a\,. \tag{8}$$

Note that the bound $\exp(-(a + b)^2/2)$ is better than e^{-2ab} or the improved bound (7) if $b \geqq a$. Also we have

$$\begin{aligned} P(S_n &\geqq a + bn \text{ for some } n \geqq 1) \\ &\leqq P(S_1 \geqq a + b) + P(S_n \geqq a + bn \text{ for some } n \geqq 2) \\ &\leqq 1 - \Phi(a + b) + P(S_n \geqq a + 2b + s(n - 2) \text{ for some } n \geqq 1) \end{aligned}$$

uniformly in $0 < s \leqq b$. Hence it follows that

$$P(S_n \geqq a + bn \text{ for some } n \geqq 1) \leqq 1 - \Phi(a + b) + e^{-2ab}f(a, b) \tag{9}$$

where $f(a, b) = \exp[-(a - 2b)^2/4]$, and

$$\Phi(u) = (2\pi)^{-\frac{1}{2}} \textstyle\int_{-\infty}^{u} \exp(-t^2/2)\, dt\,.$$

The bound given by (9) is crude when a and b are small. However, for $a = 0$, $b > 0$, we have

$$P(S_n \geqq bn \text{ for some } n \geqq 1) \leqq 1 - \Phi(b) + e^{-b^2}\,,$$

while the bound obtained by Robbins [2] is

$$P(S_n \geqq bn \text{ for some } n \geqq 1) \leqq 2(1 - \Phi(b)) .$$

Since $1 - \Phi(b) \sim (2b^2\pi)^{-\frac{1}{2}}e^{-b^2/2}$ (as $b \to \infty$), $\lim_{b\to\infty} e^{-b^2}/(1 - \Phi(b)) = 0$. Hence (9) is asymptotically better.

Now consider P_m^+ and P_m. The system of lines $f_s(n) = s(n - m) + a + bm$ $(0 < s \leqq b)$ pass through the point $(m, a + bm)$. It follows that

$$\begin{aligned} P_m^+ &= P(S_n \geqq a + bn \text{ for some } n \geqq m) \\ &\leqq \exp[-\sup_{0<s\leqq b} \{a + bm - sm\}\alpha(s)] . \end{aligned} \tag{10}$$

Note that

$$\begin{aligned} P_m^+ &\leqq P(S_n \geqq bn \text{ for some } n \geqq m) \\ &\leqq \exp[-m \sup_{0<s\leqq b} (b - s)\alpha(s)] \to 0 \quad \text{as } m \to \infty . \end{aligned}$$

As a bound for the right hand side of the inequality (10) (which is sharp if $\phi(\theta) \leqq e^{\theta^2/2}$), we have

$$\begin{aligned} P_m^+ &\leqq \exp[-\tfrac{1}{2}(a + bm)\alpha((a + bm)/2m)] \quad &\text{if } bm \geqq a \\ &\leqq \exp[-\tfrac{1}{2}(a + bm)\alpha(b)] \quad &\text{if } bm < a . \end{aligned}$$

Assuming that $\phi(-\theta) < \infty$ for $0 < \theta \leqq D \leqq \infty$, it follows that

$$P(S_n \leqq -a - bn \text{ for some } n \geqq 1) \leqq e^{-a\beta(b)} \tag{11}$$

where $\beta(b) = \sup_{0<\theta\leqq D} \{\theta : \psi_-(\theta) \leqq b\theta\}$ $(b > 0)$, and where $\psi_-(\theta) = \log \phi(-\theta)$. Note that $\beta(b) > 0$. Thus

$$\begin{aligned} &P(S_n \leqq -a - bn \text{ for some } n \geqq 1) \\ &\quad\leqq \min [\exp(-a\beta(b)), \exp\{-\sup_{0<s\leqq b} (b - s)\beta(s)\}] . \end{aligned} \tag{12}$$

It follows from (5) and (12) that

$$\begin{aligned} &P(|S_n| \geqq a + bn \text{ for some } n \geqq 1) \\ &\quad\leqq \min [\exp(-a\alpha(b)), \exp\{-\sup_{0<s\leqq b} (b - s)\alpha(s)\}] \\ &\qquad + \min [\exp(-a\beta(b)), \exp\{-\sup_{0<s\leqq b} (b - s)\beta(s)\}] . \end{aligned} \tag{13}$$

It further follows that

$$\begin{aligned} P_m &\leqq \exp[-\sup_{0<s\leqq b} (a + m(b - s))\alpha(s)] \\ &\quad + \exp[-\sup_{0<s\leqq b} (a + m(b - s))\beta(s)] . \end{aligned} \tag{14}$$

The inequality (14) leads to the following.

THEOREM. *Let* $X_1, X_2, \cdots$ *be* i.i.d. *random variables with* $EX_1 = 0$, *and* mgf $\phi(\theta)$ *existing in some neighborhood of* $\theta = 0$. *Then for every* $\varepsilon > 0$, *there exists a number* ρ $(0 \leqq \rho < 1)$ *depending only on* ε *and* ϕ *such that*

$$P(|S_n| \geqq \varepsilon n \textit{ for some } n \geqq m) \leqq 2\rho^m$$

where $\rho = \exp[-\min \{\sup_{0<s\leqq\varepsilon} (\varepsilon - s)\alpha(s), \sup_{0<s\leqq\varepsilon} (\varepsilon - s)\beta(s)\}]$.

3. A sequential selection procedure. Let $\pi_1, \pi_2, \cdots, \pi_k$ be k populations with distribution functions $F(x - \theta_1), \cdots, F(x - \theta_k)$ where θ_i's are unknown location parameters. Assuming the usual Δ-configuration: $\theta_1 \leqq \theta_2 \leqq \cdots \leqq \theta_{k-1} \leqq \theta_k - \Delta$ $(\Delta > 0)$, we want to select π_k with a given probability of error. Let X_i be the random variable associated with the population π_i $(1 \leqq i \leqq k)$. Since the distribution of $X_i - \theta_i$ is independent of θ_i, denote it by $F(y)$. We assume that $g(t) = \int \exp(ty)\, dF(y) < \infty$ for $|t| \leqq t_1 \leqq \infty$, and hence

$$-\infty < E_{\theta_i} X_i = \theta_i + M < \infty, \qquad 0 < \sigma_{X_i}^2 = \sigma^2 = \int (y - M)^2\, dF(y) < \infty .$$

Let $\{X_{ij} : j = 1, 2, \cdots, n\}$ $(1 \leqq i \leqq k)$ be k independent sequences of independent random variables from respective populations π_i. We want to select the population with the largest location parameter such that for a given ε, $0 < \varepsilon < 1$,

$$P_\Delta(CS) \geqq 1 - \varepsilon \tag{15}$$

where $CS \equiv$ correct selection. Using the results of Section 2, we will show that Paulson's elimination procedure [1] can be used to accomplish (15).

The elimination procedure. Let $f(n) \geqq 0$ be such that $f(n)/n \to 0$ as $n \to \infty$. At the nth stage eliminate π_i if

$$\max_\nu \textstyle\sum_{j=1}^n X_{\nu j} - \sum_{j=1}^n X_{ij} > f(n) \tag{16}$$

where ν runs over all populations not previously eliminated. After $(k - 1)$ populations have been eliminated (if ever), stop and assert the remaining population as the one with the largest location parameter.

Clearly, the procedure terminates with probability one. To compute the probability of error, it suffices to consider $\theta_1 = \cdots = \theta_{k-1} = \theta_k - \Delta$. From (16) we have

$$P_\Delta(\text{error}) \leqq (k - 1) P_\Delta\left(\textstyle\sum_{j=1}^n X_{1j} - \sum_{j=1}^n X_{kj} > f(n) \text{ for some } n \geqq 1\right) . \tag{17}$$

Setting $Y_j = (X_{1j} - X_{kj} + \Delta)/\sigma 2^{\frac{1}{2}}$, $\theta_1 = \cdots = \theta_{k-1} = \theta_k - \Delta$, $Y_1, Y_2, \cdots$ are i.i.d. with $EY_1 = 0$, $EY_1^2 = 1$. Let $\phi_X(t)$ denote the mgf of X. Then

$$\begin{aligned} \phi_{Y_1}(t) &= \exp[t\Delta/\sigma 2^{\frac{1}{2}}]\phi_{X_1}(t/\sigma 2^{\frac{1}{2}})\phi_{X_k}(-t/\sigma 2^{\frac{1}{2}}) \\ &= g(t/\sigma 2^{\frac{1}{2}}) g(-t/\sigma 2^{\frac{1}{2}}) , \quad 0 < t \leqq \sigma 2^{\frac{1}{2}} t_1 = t^* \leqq \infty . \end{aligned}$$

Let P denote the probability under which the distribution of Y_1 corresponds to $\phi_{Y_1}(t)$. Setting $S_n = Y_1 + \cdots + Y_n$, (17) gives

$$P_\Delta(\text{error}) \leqq (k - 1) P\left(S_n \geqq \frac{n\Delta}{\sigma 2^{\frac{1}{2}}} + \frac{f(n)}{\sigma 2^{\frac{1}{2}}} \text{ for some } n \geqq 1\right) . \tag{18}$$

Specializing to linear functions (i) $f(n) = d > 0$, and (ii) $f(n) = (c - \lambda n)^+$, $0 < c < \infty$, $0 < \lambda < \Delta < \infty$, (18) reduces to

$$P_\Delta(\text{error}) \leqq (k - 1) P(S_n \geqq a + bn \text{ for some } n \geqq 1) \tag{19}$$

where $a = d/\sigma 2^{\frac{1}{2}}$, $b = \Delta/\sigma 2^{\frac{1}{2}}$ for (i), and $a = c/\sigma 2^{\frac{1}{2}}$, $b = (\Delta - \lambda)/\sigma 2^{\frac{1}{2}}$ for (ii). The inequalities (3) and (19) lead to

$$(20)\qquad P_\Delta(\text{error}) \leqq (k-1)\exp\left[-\frac{d}{\sigma 2^{\frac{1}{2}}}\,\alpha\left(\frac{\Delta}{\sigma 2^{\frac{1}{2}}}\right)\right] \quad \text{for (i)}, \quad \text{and}$$

$$P_\Delta(\text{error}) \leqq (k-1)\exp\left[-\frac{c}{\sigma 2^{\frac{1}{2}}}\,\alpha\left(\frac{\Delta-\lambda}{\sigma 2^{\frac{1}{2}}}\right)\right] \quad \text{for (ii)}.$$

These inequalities lead to (15) by proper choice of d, c and λ. The inequalities (20) can be improved by using the improved inequality (6). We now give some examples.

EXAMPLE 1. $dF(x-\theta) = (2\pi)^{-\frac{1}{2}}\exp(-(x-\theta)^2/2)\,dx$, $-\infty < \theta < \infty$. Here $\sigma = 1$, $g(t) = e^{t^2/2}$, so that $\phi(t) = g(t/2^{\frac{1}{2}})g(-t/2^{\frac{1}{2}}) = e^{t^2/2}$. Hence the inequality (8) gives Paulson's result [1]. The error bounds are given by

$$P_\Delta(\text{error}) \leqq (k-1)\min\{e^{-d\Delta}, e^{-\Delta^2/4}\}, \quad \text{and}$$

$$P_\Delta(\text{error}) \leqq (k-1)\min\{e^{-c(\Delta-\lambda)}, e^{-(\Delta-\lambda)^2/4}\}.$$

Note that the bound $\exp(-(a+b)^2/2)$ in (8) can be used by properly choosing d, c and λ.

EXAMPLE 2. $dF(x-\theta) = \frac{1}{2}e^{-|x-\theta|}\,dx$, $-\infty < \theta < \infty$. Here $E_\theta X = \theta$, $\sigma^2 = 2$, $g(t) = (1-t^2)^{-1}$, $|t| < 1$. Thus $\phi(t) = g(t/2)g(-t/2) = (1-t^2/4)^{-1}$, $0 < t < 2$, and $\psi(t) = \log\phi(t) = -2\log\{1-t^2/4\}$. Since $\psi'(t) \to \infty$ as $t \to 2$, $\alpha(b)$ is the solution of $\psi(t) = bt$ $(b > 0)$. The probability of error can be made arbitrarily small by choosing the constants involved.

EXAMPLE 3. $dF(x-\theta) = e^{-(x-\theta)}\,dx$, $x \geqq \theta$, $-\infty < \theta < \infty$. Here $E_\theta X = \theta + 1$, $\sigma^2 = 1$, $g(t) = (1-t)^{-1}$, $0 < t < 1$; $\phi(t) = g(t/2^{\frac{1}{2}})g(-t/2^{\frac{1}{2}}) = (1-t^2/2)^{-1}$, $0 < t < 2^{\frac{1}{2}}$. Hence $\alpha(b)$ is the solution of $\psi(t) = \log\phi(t) = bt$ $(b > 0)$.

Acknowledgment. I am thankful to Professor Herbert Robbins for his inspiring ideas.

REFERENCES

[1] PAULSON, EDWARD (1964). A sequential procedure for selecting the population with the largest mean from k normal populations. *Ann. Math. Statist.* **35** 174–180.

[2] ROBBINS, HERBERT (1970). Statistical methods related to the law of the iterated logarithm. *Ann. Math. Statist.* **41** 1397–1409.

[3] ROBBINS, H. and SIEGMUND D. (1968). Iterated logarithm inequalities and related statistical procedure. *Mathematics of the Decision Sciences* **2** American Mathematical Society, Providence, 267–279.

215 DALE DRIVE
APT. 202
KENT, OHIO 44240