$g(x)$ has a finite number of oscillations, this implies that there is an interval of positive length Δ in the interior of $(0, 1)$ on which $g(x) = 0$. But then the largest of the values $Z_1, \cdots, Z_{n+1}$ is certainly no smaller than Δ; therefore Y_n is certainly no smaller than $\Delta^u\Gamma(n + u + 1)/\Gamma(n + 2)$, and this last expression approaches infinity as n increases.

REFERENCES

[1] B. F. Kimball, "Some basic theorems for developing tests of fit for the case of nonparametric probability distribution functions," *Ann. Math. Stat.*, Vol. 18 (1947), pp. 540–548.
[2] B. Sherman, "A random variable related to the spacing of sample values," *Ann. Math. Stat.*, Vol. 21 (1950), pp. 339–361.
[3] M. Greenwood, "The statistical study of infectious diseases," *J. Roy. Stat. Soc.*, Vol. 109 (1946), pp. 85–110.
[4] L. Weiss, "The stochastic convergence of a function of sample successive differences," *Ann. Math. Stat.*, Vol. 26 (1955), pp. 532–536.
[5] D. A. Darling, "On a class of problems related to the random division of an interval," *Ann. Math. Stat.*, Vol. 24 (1953), pp. 239–253.
[6] J. Neyman and E. S. Pearson, "Contributions to the theory of testing statistical hypotheses," *Stat. Res. Memoirs*, Vol. 1 (1936), pp. 1–37.

ON THE PROBABILITY OF LARGE DEVIATIONS FOR SUMS OF BOUNDED CHANCE VARIABLES

By Harry Weingarten

Bureau of Ships, Navy Department

1. Summary. The following theorems are proved.

Theorem 1. *If $x_1, x_2, \cdots$ satisfy $-1 \leqq x_n \leqq a$, $a \leqq 1$ and*

$$E(x_n \mid x_1, \cdots, x_{n-1}) \leqq -u \max (|x_n| \mid x_1, \cdots, x_{n-1}),$$

$0 < u < 1$, *then for any positive t,*

$$\Pr \{x_1 + \cdots + x_n \geqq t \text{ for some } n\} \leqq \theta^t,$$

where θ is the positive root (other than $\theta = 1$) of

$$\frac{a + u}{a + 1}\theta^{a+1} - \theta^a + \frac{1 - u}{a + 1} = 0. \tag{1}$$

This choice of θ is the best possible.

Theorem 2. *If $x_1, x_2, \cdots$ satisfy $|x_n| \leqq 1$ and $E(x_n \mid x_1, \cdots, x_{n-1}) = 0$, then for all $N > 0$,*

$$\Pr\left\{\left|\frac{x_1 + \cdots + x_n}{n}\right| \geqq \epsilon \text{ for some } n \geqq N\right\} \leqq 2\varphi^N,$$

Received November 4, 1955. (Revised January 3, 1956).

where $\varphi = (1 + \epsilon)^{-(1+\epsilon)/2}(1 - \epsilon)^{-(1-\epsilon)/2}$. *This choice of* φ *is, for every* ϵ *between* 0 *and* 1, *the best possible.*

Both results are improvements of results of Blackwell [1], and the methods of proof are somewhat similar.

2. Proofs of the Theorems. Since the methods for Theorems 1 and 2 are similar to those in [1], we merely indicate the main steps.

For Theorem 1, let $\Phi(N, t)$ be the least upper bound, over all sequences $\{x_n\}$ satisfying the hypotheses of Theorem 1, of the probability

$$\Pr\{x_1 + \cdots + x_k \geqq t \text{ for some } k \leqq N\};$$

in particular $\Phi(0, t)$ is 1 for $t \leqq 0$ and 0 for $t > 0$. Then

$$\Phi(N + 1, t) = U\Phi(N, t),$$

where U is the transformation taking Borel-measurable functions of t into Borel-measurable functions of t, such that the value of Uf at t is

$$\sup_{x \in X} Ef(t - x),$$

where X consists of all chance variables satisfying $-1 \leqq x \leqq a$ and $EX \leqq -u \max |X|$. Now if θ satisfies (1), then $Ug = g$, where $g = \theta^t$. Also, $f_1 \geqq f_2$ for all t implies $Uf_1 \geqq Uf_2$ for all t. Repeated application of this to $g \geqq \Phi(0, t)$ yields $g \geqq \Phi(N, t)$ for all t, and letting $N \to \infty$ completes the proof of Theorem 1. To see that this choice is the best possible consider the sequence $x_1, x_2, \cdots$ independent, with the distribution of each $x_n = a$, and -1 with probabilities $(1 - u)/(a + 1)$, $(a + u)/(a + 1)$ respectively. This sequence satisfies the hypotheses of Theorem 1, and it will be shown that

$$\Pr\{x_1 + \cdots + x_n \geqq t \text{ for some } n\}^{1/t} \to \theta$$

as $t \to \infty$.

To do this we consider a game between two players with fortunes, stakes, etc., as follows:

Players	P_1	P_2
Fortunes	t	b
Stakes	$a \leqq 1$	1
Probability of winning a game	$p = \dfrac{a + u}{a + 1}$	$q = \dfrac{1 - u}{a + 1}$

The probability of the ruin of P_1 which we are interested in is easily seen to be the same as $\Pr\{x_1 + \cdots + x_n \geqq t\}$ for the case of a sequence $x_1, x_2, \cdots$ independently and identically distributed with each $x_n = a, -1$ with probabilities $(1 - u)/(a + 1)$, $(a + u)/(a + 1)$ respectively.

Let us approximate a by some rational fraction r/s and then change the units in which the game is played. We will have

Players	P_1	P_2
Fortunes	st	sb
Stakes	r	s
Probability of winning a game	$p = \dfrac{a+u}{a+1}$	$q = \dfrac{1-u}{a+1}$

Using the results of [2], pages 144–146, we obtain

$$\theta_1^{st} \frac{\theta_1^{sb-s+1} - 1}{\theta_1^{st+sb-s+1} - 1} \leqq y_{st} \leq \theta_1^{st-r+1} \frac{\theta_1^{sb} - 1}{\theta_1^{st+sb-r+1} - 1},$$

where θ_1 is the root of $p\theta_1^{r+s} - \theta_1^r + q = 0$ and y_{st} is the probability of the ruin of P_1 when his fortune is st. If the fortune of P_2 becomes infinite, we have

$$\theta_1^{st} \leqq y_{st} \leqq \theta_1^{st-r+1}.$$

When we return to the original units of the game, we can state

$$[\theta_1^s]^t \leqq y_t \leqq [\theta_1^s]^{t-(r/s)+(1/s)}$$

or

$$\theta_2^t \leqq y_t \leqq \theta_2^{t-(r/s)+(1/s)},$$

where θ_2 is the root of $p\theta_2^{(r/s)+1} - \theta_2^{r/s} + q = 0$.

By choosing r and s large enough, we may come as close as we wish to a, and so we may finally write

$$\theta^t \leqq y_t \leqq \theta^{t-a}$$

where θ is the root of $p\theta^{a+1} - \theta^a + q = 0$. This is possible, since the probability of ruin in the game where the stakes are r and s is the general solution of the difference equation

$$y_x = py_{x+s} + qy_{x-r},$$

where y_x is the probability of ruin of P_1 when his fortune is x. Such solutions are known to be continuous functions of the stakes. That θ_2, the root of $p\theta^{(r/s)+1} - \theta^{r/s} + q = 0$, approaches θ, the root of $p\theta^{a+1} - \theta^a + q = 0$, follows from the fact that the solution of a polynomial is a continuous function of the coefficients. From this we may obtain

$$[\theta^t]^{1/t} \leqq \Pr\{x_1 + \cdots + x_n \geqq t\}^{1/t} \leqq [\theta^{t-a}]^{1/t}$$

and so

$$\Pr\{x_1 + \cdots + x_n \geqq t\}^{1/t} \to \theta \quad \text{as} \quad t \to \infty.$$

As a matter of fact we note that θ is really a lower bound, since

$$\theta^{1-(a/t)} > \theta.$$

Hence θ is best possible.

For the proof of Theorem 2 we have:

$$\Pr\left\{\frac{x_1 + \cdots + x_n}{n} \geqq \epsilon \quad \text{for some} \quad n \geqq N\right\}$$

$$\leqq \Pr\{x_1 - k\epsilon + \cdots + x_n - k\epsilon \geqq N\epsilon(1-k) \text{ for some } n\}$$

$$\leqq [(1+\epsilon)^{-(1+\epsilon)/2} \cdot (1-\epsilon)^{-(1-\epsilon)/2}]^N,$$

where the last inequality is obtained by applying Theorem 1 to the sequence

$$y_n = \frac{x_n - k\epsilon}{1 + k\epsilon}.$$

Here we have taken

$$a = \frac{1 - k\epsilon}{1 + k\epsilon} \qquad \text{and} \qquad u = \frac{k\epsilon}{1 + k\epsilon}$$

and found

$$\Pr\left\{y_1 + \cdots + y_n \geqq \frac{N\epsilon(1-k)}{1+k\epsilon}\right\} \leqq \theta^{N\epsilon(1-k)/(1+k\epsilon)},$$

where θ is the root of

$$\tfrac{1}{2}\theta^{2/(1+k\epsilon)} - \theta^{(1-k\epsilon)/(1+k\epsilon)} + \tfrac{1}{2} = 0. \tag{2}$$

To find the smallest value of $\theta^{N\epsilon(1-k)/(1+k\epsilon)}$ we may procede in the following manner: Beginning with (2), we write

$$\theta^{(1-k\epsilon)/(1+k\epsilon)}(\theta - 2) + 1 = 0, \tag{3}$$

and solving for $k\epsilon$, we find

$$k\epsilon = \frac{\log\theta + \log(2-\theta)}{\log\theta - \log(2-\theta)}. \tag{4}$$

Giving $k\epsilon$ the value from (4), we find that $R = \theta^{N\epsilon(1-k)/(1+k\epsilon)}$ gives

$$R^{1/N} = \theta^{-[(1-\epsilon)\log\theta + (1+\epsilon)\log(2-\theta)]/2\log\theta} \tag{5}$$

If we take logarithms in both sides of (5) and simplify, we may rewrite (5) to obtain

$$R^{1/N} = \theta^{-(1-\epsilon)/2}(2-\theta)^{-(1+\epsilon)/2}, \tag{6}$$

and it is very easy to find the value of θ which makes R a minimum to be

$$\theta = 1 - \epsilon. \tag{7}$$

The same inequality holds for

$$\Pr\left\{\frac{x_1 + \cdots + x_n}{n} \leqq -\epsilon \text{ for } n \geqq N\right\},$$

and hence Theorem 2 is proved.

To see that φ is the best possible, consider the case of the sequence $\{x_n\}$ independently distributed, each taking the values ± 1 with probabilities $\frac{1}{2}$, $\frac{1}{2}$. It follows from a result of Chernoff [3] that

$$\Pr\{x_1 + \cdots + x_n \geqq n\epsilon\}^{1/n} \to \varphi \quad \text{as} \quad n \to \infty,$$

so that our φ is exact.

REFERENCES

[1] D. BLACKWELL, "On optimal systems," *Ann. Math. Stat.*, Vol. 25 (1954), pp. 394–397.
[2] J. V. USPENSKY, *Introduction to Mathematical Probability*, McGraw-Hill, 1937.
[3] H. CHERNOFF, "A measure of asymptotic efficiency for test of a hypothesis based on a sum of observations," *Ann. Math. Stat.*, Vol. 23 (1952), pp. 493–507.

A REMARK ON THE ROOTS OF THE MAXIMUM LIKELIHOOD EQUATION

BY C. KRAFT AND L. LECAM[1]

University of California, Berkeley

1. Introduction and summary. The statistical literature combines two types of investigations concerning the consistency of maximum likelihood (M.L.) estimates. A few of these, such as the most excellent paper of A. Wald [1], do prove directly the consistency of M.L. estimates. However, most investigators seem to have concentrated their efforts on proving the existence and consistency of suitably selected roots of the successive likelihood equations. Some authors, see [2], for example, add the supplementary remark that such consistent roots will eventually be unique in suitably small neighborhoods of the true value and will achieve a local maximum.

It is the purpose of the present note to point out by means of examples that this second mode of attack is not adequate. In the examples given below, the "usual regularity conditions" of Cramér [3] or Wald [4] are satisfied, but the M.L. estimates are not consistent. It should also be pointed out that the direct proofs of existence of roots, simple in the case of a unidimensional parameter, become unwieldy in more than one dimension. On the other hand, if one has proved the consistency of the M.L. estimates, the existence of roots follows trivially from the fact that when a differentiable function reaches its maximum in an open set, the derivatives vanish at that point.

2. Examples with independent identically distributed variables. The first example given below has the following characteristics:

Received September 21, 1955.
[1] This paper was prepared with the support of the Office of Ordnance Research, U. S. Army, under Contract DA-04-200-ORD-171.