JOURNAL OF MATHEMATICAL ANALYSIS AND APPLICATIONS **123**, 281–291 (1987)

Limit Theorems for Logarithmic Means

A. O. PITTENGER*

Department of Mathematics, University of Maryland Baltimore County, Catonsville, Maryland 21228

Submitted by R. P. Boas

Received October 9, 1985

1. INTRODUCTION

Given two positive quantities x and y, there are a variety of ways of computing their "average," and the problem of comparing different averages has attracted mathematicians for years. The best known example is the inequality between the geometric mean $(xy)^{1/2}$ and the arithmetic mean $(x+y)/2$, while a slightly less familiar example is

$$(xy)^{1/2} \leqslant L(x, y) \leqslant (x+y)/2, \tag{1.1}$$

where $L(x, x) = x$ and $L(x, y) = (x-y)/\log(x/y)$ for $x \neq y$. An example of a more arcane relationship is

$$\frac{(xy)^2}{2} \cdot \left[\left(\frac{x}{y}\right)^{\sqrt{3}/6} + \left(\frac{y}{x}\right)^{\sqrt{3}/6}\right] \leqslant L(x, y) \leqslant \left(\frac{x^{1/3}+y^{1/3}}{2}\right)^3.$$

(See [8, 10].)

For a presentation of new results as well as a summary of work in this area, the reader can consult the recent papers on extended mean values by Leach and Sholander [5, 6]. One of the themes of those papers is the investigation of two-parameter families of means which give known one-parameter families as special cases. Examples of the latter are the usual power means

$$A_p(x, y) = \begin{cases} [(x^p + y^p)/2]^{1/p}, & p \neq 0 \\ (xy)^{1/2}, & p = 0, \end{cases} \tag{1.2}$$

* Research supported in part by NSF Grant DMS 840539.

0022-247X/87 $3.00

and the rth logarithmic means

$$L_r(x, y) = \phi_r^{-1}\left(\int_0^1 \phi_r(xv + y(1-v))\,dv\right), \tag{1.3}$$

where

$$\phi_r(t) = \begin{cases} t^r, & r \neq 0 \\ \log(t), & r = 0. \end{cases} \tag{1.4}$$

In analogy with (1.1) it is possible to determine conditions when $L_r(x, y)$ dominates or is dominated by $A_p(x, y)$, and the reader may find a discussion of these matters in [9]. Leach and Sholander [6] analyze this problem in their two-parameter context and also give references to related work.

Both indexed families A_p and L_r are non-decreasing in their index and both interpolate from $(xy)^{1/2}$ to $(x+y)/2$. This naturally suggests an effort to extend these investigations to multivariable means. In [11] the author described a family defined as "the" r-logarithmic mean for n positive variables, a family which is a special case of some general results of Carlson on hypergeometric means [1, 2]. At about the same time, Leach and Sholander published an interesting paper [7] discussing two-parameter, multivariate mean values. Using techniques of divided differences, they arrive at a formula introduced by Stolarsky [12] and obtain some general structural properties of this extended mean. The logarithmic means described in [11] constitute an important one-parameter subclass, as we shall see in Section 5.

In this paper we investigate the problem of limit values of extended multivariate means. Section 2 gives definitions and records relevant properties of multivariate logarithmic means. Our interest in that topic was stimulated by the question of a (probabilistic) strong law for logarithmic means, and the reader will find that perspective evident in Sections 3 and 4 in which we prove that the rth logarithmic mean of a sequence of positive, independent, identically distributed random variables converges almost surely to the mean of the distribution.

The only role played by the independence assumption in that work is that $\sum X_i(\omega)/n$ converges for almost all ω. If we assume instead that we have sequences whose averages converge, all of the results and techniques of Sections 3 and 4 carry over. With that in mind, we relate our work to that of Leach and Sholander in Section 5 and apply the asymptotic results to their two-parameter means. The main result of that section is that on suitable sequences, their means converge to the harmonic mean of the given sequence, assuming that exists.

2. Definitions

We assume throughout that x_i denotes a strictly positive real number and that x denotes the $n+1$-vector $(x_0, \ldots, x_n)$. We let A_n stand for the simplex

$$A_n = \left\{ (v_0, \ldots, v_{n-1}), \quad 0 \leqslant v_i, \quad \sum_0^{n-1} v_i \leqslant 1 \right\}, \tag{2.1}$$

and routinely use $v_n = 1 - \sum_0^{n-1} v_i$ and the inner product

$$x \cdot v = \sum_{i=0}^{n} x_i \cdot v_i. \tag{2.2}$$

The volume of A_n is $(n!)^{-1}$, and its calculation is a special case of a Dirichlet integral. (See, for example, [13, p. 258].) Letting dv stand for the volume measure in A_n, we use $d\rho_n$ to denote the probability $(n)!\, dv$. With ϕ_r as given in (1.4), we can now define the logarithmic means.

(2.3) Definition. $L_r(x) = \phi_r^{-1}(\int_{A_n} \phi_r(x \cdot v)\, d\rho_n)$. As an example, if $r = -1$, $\phi_r^{-1} = t^{-1} = \phi_r$, and

$$L(x_0, \ldots, x_n) \equiv L_{-1}(x) = \left(\int_{A_n} (x \cdot v)^{-1}\, d\rho_n \right)^{-1}. \tag{2.4}$$

(2.5) Proposition [11]. *Let x be fixed. Then $L_r(x)$ is increasing in r, strictly so if the x_i are not all equal, and*

$$\left(\prod_0^n x_i \right)^{1/n+1} = L_{-n-1}(x) \leqslant L_{-1}(x) \leqslant L_1(x) = \sum_0^n x_i/(n+1).$$

Explicit representations of L_r are given in [11] for integer values of r in $[-n-1, -1]$, but these will not be necessary here.

3. A Limit Theorem in a Restricted Case

Suppose that the given sequence x_i is restricted by $0 < a < x_i < b$ and that the limit

$$\mu = \lim_n \sum_{i=0}^{n} x_i/(n+1) \tag{3.1}$$

exists. Letting $x(n)$ denote the vector defined by the first $n+1$ values in the sequence, we show that for every fixed r, $L_r(x(n))$ also converges to μ. A different problem, and one which is related to the two-parameter means of Leach and Sholander, would be to allow r to vary with n. Thus, for example, instead of getting μ we could obtain the smaller value

$$\lim_n L_{-n-1}(x(n)) = \exp\left(\lim_n \sum \log(x_i)/(n+1)\right), \tag{3.2}$$

assuming $\sum \log(x_i)/n$ converged to a limit.

The original proof of (3.3) below used a weak convergence argument which was straightforward but rather technical. After reading that proof, Bill Pruitt suggested a different technique which eliminated many of the complications, and we follow his suggestion below.

(3.3) THEOREM. *Suppose* $0 < a < x_i < b$ *and* $\mu = \lim \sum x_i/(n+1)$ *exists. Then for fixed* r, $\lim L_r(x(n)) = \mu$.

Proof. In defining L_r we used the dot product which can be interpreted as a convex combination of the x_i. However, another way to view $x \cdot v$ is as an integral. Using 1_A to denote the characteristic function of a set, define

$$F_n(t; v, x) = \sum_{i=0}^{n} v_i 1_{[0,t]}(x_i), \tag{3.4}$$

so that F_n is a v-parametrized distribution function which is piecewise constant with a jump of size v_i when $t = x_i$. It is easy to check that

$$x \cdot v = \int tF_n(dt; v, x),$$

so that

$$L_r(x(n)) = \phi_r^{-1}\left(\int_{A_n} \phi_r\left(\int tF_n(dt; v, x)\right) d\rho_n\right).$$

Now let $0 < p < \frac{1}{4}$ and define a subset of A_n by

$$S_n = \{v: \sup_t(|F_n(t, v, x) - F_n(t; x)|) > n^{-p}\},$$

where

$$F_n(t; x) = \sum_{i=0}^{n} (n+1)^{-1} 1_{[0,t]}(x_i).$$

Suppose we could show that $\lim_n \rho_n(S_n) = 0$. By the bounds on the x_i, the contribution of the set S_n to the integral could be ignored in the limit, while for v in S_n^c it is easy to check (using integration by parts) that

$$\left|\int t\, F_n(dt; v, x) - \int t\, F_n(dt; x)\right| \leqslant n^{-p}(b-a).$$

Hence $L_r(x(n))$ will have the same limit as

$$\phi_r^{-1}\left(\int_{A_n} \phi_r\left(\int t\, F_n(dt; x)\right) d\rho_n\right) = \sum x_i/(n+1),$$

and that is the assertion of the theorem.

It thus remains to show that S_n is asymptotically negligible, and for that purpose the symmetry of A_n and ρ_n allows us to assume for the calculation below that the x_i are non-decreasing. If $x_k < x_{k+1}$, then

$$F_n(x_k; v, x) - F_n(x_k; x) = V_k - (k+1)/(n+1),$$

where $V_k = \sum_0^k v_i$. Hence by Chebyshev's inequality [3, p. 46]

$$\begin{aligned} \rho_n(S_n) &\leqslant \sum_{k=0}^{n-1} \rho_n(\{v: |V_k - (k+1)/(n+1| > n^{-p}\}) \\ &\leqslant n^{4p} \sum_{k=0}^{n-1} \int_{A_n} (V_k - (k+1)/(n+1))^4\, d\rho_n. \end{aligned} \tag{3.5}$$

One can interpret V_k as the $(k+1)$th order statistic of n independent, uniformly distributed random variables on [0, 1], and that enables us to compute the distribution and required means of V_k for $0 \leqslant k \leqslant n-1$. From Feller [4, p. 23], we have

$$\rho_n(\{v: V_k \leqslant t\}) = \sum_{j=k+1}^{n} \binom{n}{j} t^j (1-t)^{n-j}. \tag{3.6}$$

One verifies that the right-hand side of (3.6) equals

$$n\binom{n-1}{k} \int_0^t x^k (1-x)^{n-k}\, dx,$$

and from that it is easy to identify the rth mean of V_k as a beta integral, leading to

$$\int_{A_n} V_k^r\, d\rho_n = \frac{(k+1)(k+2)\cdots(k+r)}{(n+1)(n+2)\cdots(n+r)}. \tag{3.7}$$

The remainder of the calculation consists of evaluating the right-hand side of (3.5) using (3.7). After combining terms we obtain the expression

$$\sum_{k=1}^{n} \frac{3k^4(n-5)-6k^3(n+1)(n-5)+3k^2(n+1)^2(n-7)+6k(n+1)^3}{(n+1)^4(n+2)(n+3)(n+4)},$$

and it is then immediate that

$$\rho_n(S_n) \leqslant n^{4p} \cdot A/n.$$

Since $p < \frac{1}{4}$, S_n is asymptotically negligible, and the proof is complete.

We should note that the fourth moment is indeed necessary, since the second moment yields $\frac{1}{6}$ as the upper bound, and that is not particularly helpful.

As we remarked in the Introduction, an open problem is the comparison of different families of parametrized means. Theorem 3.3 shows that asymptotically the $p=1$ arithmetic mean will be the smallest power mean dominating the multivariate logarithmic means L_r, $r \leqslant 1$, and the largest power mean dominated by L_r, $1 \leqslant r$.

4. Logarithmic Means and the Strong Law of Large Numbers

Part of the incentive for studying a logarithmic mean in n variables arises from possible applications in statistics. Assume that we have an underlying probability space and the x_i are $X_i(\omega)$, where X_i, $0 \leqslant i$, is a sequence of positive, independent, identically distributed random variables with mean μ. To see that $L \equiv L_{-1}(x(n))$ might be of statistical interest, we note that if $\bar{X}$ is the sample mean, then

$$0 \leqslant \bar{X} - L = \sum_{i=0}^{n} (\bar{X} - X_i)\, C_i(\omega, n), \tag{4.1}$$

where

$$C_i(\omega, n) = L \cdot \int_{A_n} v_i \left(\sum v_j X_j \right)^{-1} d\rho_n.$$

Thus $\bar{X} - L$ is a convex combination of the differences $\bar{X} - X_i$, and the weights are such that L is a biased estimator of the population mean μ:

$$E[L] \leqslant E(\bar{X}) = \mu.$$

(As usual, $E(X)$ denotes the expectation of X.)

Under these hypotheses and with a little extra work we can extend the results of (3.3) to this context.

(4.2) THEOREM. *Let X_i, $0 \leqslant i < \infty$, be positive, independent, identically distributed random variables with expectation μ. If $1 \leqslant r$, then almost surely*

$$\lim_n L_r(X_0, ..., X_n) = \mu. \tag{4.3}$$

If $E[-\log(X_0 \wedge 1)] < \infty$, then (4.3) also holds for $r < 1$. ($X_0 \wedge 1$ denotes $\min(X_0, 1)$.)

Proof. Suppose the X_i take values in $[a, b]$ with $0 < a < b < \infty$. Define $F_n(t; v, X)$ and $F_n(t; X)$ as in Section 3. In this context, the latter distribution is the empirical distribution function, and a standard result from probability theory is the Glivenko–Cantelli theorem [3, p. 124]

$$\lim_n \sup_t |F_n(t, X) - F(t)| = 0, \qquad \text{a.s.,}$$

where F is the distribution function of the given random variables. It thus follows as in Section 3 that L_r converges to μ.

Now we begin to eliminate some of the restrictions on the random variables. First assume $0 < X_i < b$. If $1 \leqslant r$, the proof in Section 3 applies without change. To obtain (4.3) for negative r, it suffices by (2.5) to prove the assertion for $r = -k$, k a positive integer. Using the mean value theorem and monotonicity we obtain the following estimate: if $0 < x_i < y_i$

$$\begin{aligned} 0 &\leqslant L_{-k}(y) - L_{-k}(x) \\ &= L_{-k}(y)\, L_{-k}(x) \left[\left(\int (x \cdot v)^{-k}\, d\rho \right)^{-1/k} - \left(\int (y \cdot v)^{-k}\, d\rho \right)^{-1/k} \right] \\ &\leqslant (L_{-k}(y))^{k+1} \cdot \max(y_i - x_i, 1 \leqslant i \leqslant n) \cdot \exp\left(\frac{-(k+1)}{n} \sum \log(x_i) \right). \end{aligned}$$

Choosing $x_i = X_i(\omega)$ and $y_i = X_i(\omega) + \varepsilon$, the foregoing inequality and the constraint on $E[-\log(X_i \wedge 1)]$ give a lower bound for $\underline{\lim} L_{-k}(X_0, ..., X_n)$ of the form $\mu - \varepsilon C$, where C is a constant. Since $\overline{\lim}\, L_{-k}$ is bounded above by μ, almost surely, we have extended (4.3) to this class of bounded, positive random variables.

Now suppose the X_i are not bounded above. From the preceding results we have, with the obvious notation,

$$E[X_0 \wedge b] \leqslant \underline{\lim} L_r(X) \leqslant \overline{\lim}\, L_r(X), \tag{4.4}$$

so that if $\mu = \infty$, (4.3) is immediate. If $r \leqslant 1$, we know $\overline{\lim}\, L_r(X)$ is bounded above by μ almost surely, and that fact combines with (4.4) to prove (4.2). It thus remains to consider the case when $1 \leqslant r$ and $\mu < \infty$, and this situation seems to require a different approach. We assume r is a positive integer, expand $(\sum v_i X_i)^r$, and collect terms according to the exponents of X_i. Evaluating the integrals over A_n, it is possible to show there is a constant $A(r)$, independent of n, such that

$$L_r^r(X_0,\dots, X_n) \leqslant \left(\frac{\sum X_i}{n+1}\right)^r + A(r)\left(\frac{\sum X_i}{n+1}\right)^{r-2} \frac{\sum X_i^2}{(n+1)^2}.$$

The required bound, $\overline{\lim}\, L_r(X) \leqslant \mu$, then follows from the usual strong law and the following, elementary result.

(4.5) LEMMA. *Suppose* $0 \leqslant x_i$ *and* $\lim \sum x_i/n = \mu < \infty$. *Then* $\lim_n (x_n/n) = 0$ *and* $\lim_n \sum x_i^2/n^2 = 0$.

5. ASYMPTOTICS IN THE TWO-PARAMETER CASE

In their paper on multivariate extended mean values [7], Leach and Sholander give an interesting derivation and analysis of a class of general multivariate means. If we translate their notation to that used in this paper, we can write their means as

$$E(r, s, x) = \left(\int_{A_n} (x \cdot v)^{s-n}\, d\rho_n \Big/ \int_{A_n} (x \cdot v)^{r-n}\, d\rho_n\right)^{1/(s-r)}$$
$$= [(L_{s-n}(x))^{s-n}/(L_{r-n}(x))^{r-n}]^{1/(s-r)}, \tag{5.1}$$

in the case when $r \neq s$. When $r = s$

$$E(r, r, x) = \exp\left(\frac{\int_{A_n}(v \cdot x)^{r-n} \log(v \cdot x)\, d\rho_n}{\int_{A_n}(v \cdot x)^{r-n}\, d\rho_n}\right), \tag{5.2}$$

so that

$$E(n, n, x) = L_0(x). \tag{5.3}$$

Equation (5.1) enables us to immediately compute that $E(n, n+1, x)$ is the arithmetic mean and $E(-1, n, x)$ is the geometric mean. Leach and Sholander obtain those results as a consequence of an analysis which also gives

$$E(-2, -1, x) = \left(\sum x_i^{-1}/(n+1)\right)^{-1}, \tag{5.4}$$

the harmonic mean. Combining (5.4) with (5.1) enables us to express

$L_{-2-n}(x)$ in terms of the harmonic and geometrix means, while the results of [11] yield evaluations for $E(k, j, x)$ for an integer $j \neq k$ in $\{-1, 0, \ldots, n-1\}$ when the $(n+1)$ variables $x_0, \ldots, x_n$ are distinct.

Equation (5.1) shows that the logarithmic means form an important subclass of the two-parameter means,

$$L_r(x) = E(n, r+n, x),$$

and we should expect their asymptotic properties to be at least as important.

(5.5) PROPOSITION. *Let x_k, $0 \leqslant k < \infty$, be a sequence of positive real numbers such that $\mu = \lim \sum x_k/(n+1)$ exists and is finite and such that*

$$\overline{\lim}\left(-\sum \log(x_k \wedge 1)\Big/(n+1)\right) < \infty.$$

Then for all r and s

$$\lim_n E(r+n, s+n, x(n)) = \mu.$$

Proof. Assume $r \neq s$ and use $r+n$ and $s+n$ for r and s in (51). The hypotheses on the x_k are precisely what was required in Sections 3 and 4 to prove the strong law, and thus $E(r+n, s+n, x(n))$ converges to μ. Since E increases with increasing r and s [7, Theorem 4], the case of $r = s$ follows as well.

We have not attempted to achieve maximum generality with our hypotheses, since our motivation is primarily one of discovering the general sort of asymptotic behavior of these means as n diverges. Indeed, the point of (5.5) is that along any half-line $(r+n, s+n)$ in the (r, s) plane, the multivariate means resemble the usual arithmetic mean, at least on sequences whose averages converge to a finite limit.

It may appear that the asymptotic behavior of $E(r, s, x(n))$ is too delicate a question for our techniques, but the surprising fact is that with a key result from [7] we can identify the limit of $E(r, s, x(n))$ as the harmonic mean of the sequence, assuming that mean exists.

(5.6) THEOREM. *Suppose $0 < x_k$, $0 \leqslant k < \infty$, and that both*

$$\lambda = \lim \sum \log(x_k)/(n+1)$$

and

$$0 < H(x) = \lim\left(\sum x_k^{-1}/(n+1)\right)^{-1}$$

exist and are finite. Then for all fixed r and s

$$\lim_n L_{s-n}(x(n)) = \lim_n \left(\prod_0^n x_k\right)^{1/(n+1)} = e^{\lambda}, \tag{5.7}$$

$$\lim_n E(r+n, s, x(n)) = e^{\lambda} \tag{5.8}$$

and

$$\lim_n E(r, s, x(n)) = H(x). \tag{5.9}$$

Remark. Note that the theorem says that asymptotically all $E(r, s)$ means look like the $E(-2, -1)$ mean on sequences with the requisite limits.

Proof. Let G_n denote the geometric mean

$$G_n = \left(\prod_0^n x_k\right)^{1/(n+1)}.$$

Then the analysis in [7, (33) and following] translates to

$$\int_{A_n} (x \cdot v)^{s-n} \, d\rho_n = (G_n)^{-(n+1)} \int_{A_n} \left(\frac{1}{x} \cdot v\right)^{-1-s} d\rho_n$$

or

$$L_{s-n}^{s-n}(x(n)) = (G_n)^{-(n+1)} L_{-1-s}^{-1-s}(x(n)). \tag{5.10}$$

Under the hypotheses (5.7) is immediate, and the use of (5.10) in equation (5.1) gives the remaining two assertions.

References

1. B. C. Carlson, A hypergeometric mean value, *Proc. Amer. Math. Soc.* **16** (1965), 759–766.
2. B. C. Carlson, Some inequalities for hypergeometric functions, *Proc. Amer. Math. Soc.* **17** (1966), 32–39.
3. K. L. Chung, "A Course in Probability Theory," Harcourt, Brace & World, New York, 1968.
4. W. Feller, "An Introduction to Probability Theory and its Applications, Vol. II," Wiley, New York, 1971.
5. E. Leach and M. Sholander, Extended mean values, *Amer. Math. Monthly* **85** (1978), 84–90.
6. E. Leach and M. Sholander, Extended mean values II, *J. Math. Anal. Appl.* **92** (1983), 207–223.

7. E. LEACH AND M. SHOLANDER, Multi-variable extended mean values, *J. Math. Anal. Appl.* **104** (1984), 390–407.
8. T. P. LIN, The power and the logarithmic mean, *Amer. Math. Monthly* **81** (1974), 879–883.
9. A. O. PITTENGER, Inequalities between arithmetic and logarithmic means, *Univ. Beograd Publ. Elektrotehn Fak, Ser. Mat. Fiz.* **678–715** (1980), 15–18.
10. A. O. PITTENGER, The symmetric, logarithmic and power means, *Univ. Beograd Publ. Elektrotehn. Fak. Ser. Mat. Fiz.* **678–715** (1980), 19–23.
11. A. O. PITTENGER, The logarithmic mean in *n* variables, *Amer. Math Monthly* **92** (1985), 99–104.
12. K. B. STOLARSKY, Generalizations of the logarithmic mean, *Math. Mag.* **48** (1975), 87–92.
13. E. T. WHITTAKER AND G. N. WATSON, "Modern Analysis," 4th ed., Cambridge Univ. Press, Cambridge, 1946.