ON THE CONVEXITY OF MEAN VALUE FUNCTIONS

HAROLD SHNIAD

1. **Introduction.** Let (a) denote a set $a_1, a_2, \cdots, a_n$ of n distinct positive numbers, $n \geqq 2$, with the subscripts ν labeled so that $a_\nu < a_{\nu+1}$ for $\nu = 1, \cdots, n-1$. Let (ξ) denote a set of positive numbers $\xi_1, \xi_2, \cdots, \xi_n$ with $\sum_{\nu=1}^n \xi_\nu = 1$. The mean value function $M_t(a, \xi) = (\sum_{\nu=1}^n \xi_\nu a_\nu^t)^{1/t}$, $t \neq 0$, $\pm\infty$; $M_0(a, \xi) = \prod_{\nu=1}^n a_\nu^{\xi_\nu}$; $M_{-\infty}(a, \xi) = \min_{\nu=1,2,\cdots,n} a_\nu$ and $M_{+\infty}(a, \xi) = \max_{\nu=1,2,\cdots,n} a_\nu$; is a continuous and strictly increasing function of t for $-\infty \leqq t \leqq +\infty$.[1] For given fixed sets (a) and (ξ), let $M(t)$ denote $M_t(a, \xi)$ and $\Lambda(t)$ denote $\log M_t(a, \xi)$. Each of the functions $M(t)$ and $\Lambda(t)$ has horizontal asymptotes and consequently at least one point of inflection. We shall show that these functions may have more than one inflection point, but shall show that $\Lambda(t)$ is a convex function of t in a neighborhood of $-\infty$, and a concave function of t in a neighborhood of $+\infty$. A sufficient condition will be obtained for $\Lambda(t)$ to be convex for all negative t, and one for $\Lambda(t)$ to be concave for all positive t. Finally, the applicability of the methods used to more general weighted sums will be considered briefly.

2. **Notations and fundamental formulae.** Let

$$f(t) = \log\left(\sum_{\nu=1}^n \xi_\nu a_\nu^t\right), \tag{1}$$

$$\eta_\nu(t) = \frac{\xi_\nu a_\nu^t}{\sum_{\nu=1}^n \xi_\nu a_\nu^t}, \quad \lambda_\nu = \log a_\nu \qquad (\nu = 1, 2, \cdots, n), \tag{2}$$

$$S_k = \sum_{\nu=1}^n \eta_\nu(\lambda_\nu)^k. \tag{3}$$

Then

$$\frac{d\eta_\nu}{dt} = \eta_\nu\lambda_\nu - \eta_\nu S_1, \quad \frac{dS_k}{dt} = S_{k+1} - S_1 S_k,$$

Presented to the Society, December 31, 1947; received by the editors July 5, 1947, and, in revised form, August 25, 1947.

[1] See Hardy, Littlewood, and Pólya, *Inequalities*, Cambridge University Press, 1934, chap. 2, for the basic properties of the mean value function.

$$
\begin{aligned}
f' &= S_1, \\
f'' &= S_2 - S_1^2, \\
f''' &= S_3 - 3S_1S_2 + 2S_1^3, \qquad (4)\\
f^{\mathrm{IV}} &= S_4 - 4S_1S_3 - 3S_2^2 + 12S_1^2S_2 - 6S_1^4,
\end{aligned}
$$

$$f^{\mathrm{IV}} = (S_4 - 3S_2^2 + 2S_1^4) - 4S_1 f'''. \tag{5}$$

Since $\Lambda(t) = \log M(t) = f/t$, we have

$$\Lambda^{(k)} = (1/t)(f^{(k)} - k\Lambda^{(k-1)}); \tag{6}$$

letting $k=3$ in (6) and integrating, we obtain the result:

$$(t_2^3)\Lambda''(t_2) = (t_1^3)\Lambda''(t_1) + \int_{t_1}^{t_2} t^2 f''' dt, \tag{7}$$

or

$$t^3\Lambda''(t) = \int_0^t t^2 f''' dt \quad \text{and} \quad \Lambda''(0) = (1/3)f'''(0). \tag{8}$$

3. **The function** $f'''(t)$. The idea behind our analysis of the convexity or concavity of $\Lambda(t)$ in certain ranges of t is to deduce results from corresponding properties of the relatively simpler function $f'''(t)$. By comparing formula (4) with the following expression:

$$\Lambda''(t) = (1/t)(S_2 - S_1^2) - (2/t^2)(S_1) + (2/t^3)f,$$

it may be seen that a direct study of $\Lambda''(t)$ is considerably more complicated than that of $f'''(t)$. Formulae (7) and (8) will be used to pass from conclusions regarding the sign of $f'''(t)$ in an interval to similar conclusions regarding the sign of $\Lambda''(t)$.

Our analysis of $f'''(t)$ will be that of a function:

$$F(t) = \sum_{\nu=1}^{n} \eta_\nu(\lambda_\nu)^3 - 3\left(\sum_{\nu=1}^{n} \eta_\nu\lambda_\nu\right)\left[\sum_{\nu=1}^{n} \eta_\nu(\lambda_\nu)^2\right] + 2\left(\sum_{\nu=1}^{n} \eta_\nu\lambda_\nu\right)^3,$$

where the expressions η_ν, λ_ν ($\nu = 1, 2, \cdots, n$; $n \geqq 2$) are subject to the conditions:

(a) λ_ν is any fixed finite valued real number with $\lambda_\nu < \lambda_{\nu+1}$;

(b) $\eta_\nu = \eta_\nu(t)$ with $\eta_\nu > 0$, $\sum_{\nu=1}^{n} \eta_\nu(t) = 1$, and $\eta_\nu(t)$ is a continuous function of t defined for all real t;

(c) η_1 is a strictly decreasing function of t and $\lim_{t\to-\infty} \eta_1 = +1$, $\lim_{t\to+\infty} \eta_1 = 0$;

(d) η_n is a strictly increasing function of t and $\lim_{t\to-\infty} \eta_n = 0$, $\lim_{t\to+\infty} \eta_n = +1$.

That $f'''(t)$ is a function $F(t)$ follows easily from formula (4) and the definitions (2) and (3).

LEMMA. *The function $F(t)$ has the property that if each λ_ν is replaced by $\lambda_\nu+c$ (c independent of ν), then $F(t)$ remains invariant.*

PROOF. Let $S_k(t, c)=\sum_{\nu=1}^{n}\eta_\nu(\lambda_\nu+c)^k$. Then

$$\frac{\partial S_k(t, c)}{\partial c} = kS_{k-1}(t, c), \qquad \frac{\partial S_1(t, c)}{\partial c} = 1,$$

and

$$\frac{\partial}{\partial c}\left\{S_3(t, c) - 3S_1(t, c)S_2(t, c) + 2[S_1(t, c)]^3\right\} = 0.$$

THEOREM 1. (a) *If $\eta_1(t_1)=1/2$, then $t<t_1$ implies that $F(t)>0$.*

(b) *If $\eta_n(t_2)=1/2$, then $t>t_2$ implies that $F(t)<0$.*

(c) *We have $F(t_1)=0$ if and only if $n=2$; that is, if and only if $t_1=t_2$.*

PROOF. *Case* 1, $n=2$. We replace λ_ν by $\lambda_\nu-\lambda_2$ and obtain $F(t)=(\eta_1-3\eta_1^2+2\eta_1^3)(\lambda_1-\lambda_2)^3=(1-2\eta_1)N(t)$, where $N(t)<0$. Hence $F(t)>0$ for $t<t_1$, since $\eta_1(t)$ is a strictly decreasing function of t.

Case 2, $n>2$. Assume the theorem valid for all n satisfying $2\leqq n\leqq k$ and consider $n=k+1$. Suppose that $\eta_1(t)>1/2$, and let $d_\nu=\lambda_\nu-\lambda_2$. Then

$$\begin{aligned} F(t) &= \left(\eta_1 d_1^3 + \sum_{\nu=3}^{k+1}\eta_\nu d_\nu^3\right) - 3\left(\eta_1 d_1 + \sum_{\nu=3}^{k+1}\eta_\nu d_\nu\right)\left(\eta_1 d_1^2 + \sum_{\nu=3}^{k+1}\eta_\nu d_\nu^2\right) \\ &\quad + 2\left(\eta_1 d_1 + \sum_{\nu=3}^{k+1}\eta_\nu d_\nu\right)^3 \\ &= P_1 + P_2 + P_3 + P_4, \end{aligned}$$

where

$$\begin{aligned} P_1 &= (\eta_1 - 3\eta_1^2 + 2\eta_1^3)d_1^3, \\ P_2 &= \sum_{\nu=3}^{k+1}\eta_\nu d_\nu^3 - 3\left(\sum_{\nu=3}^{k+1}\eta_\nu d_\nu\right)\left(\sum_{\nu=3}^{k+1}\eta_\nu d_\nu^2\right) + 2\left(\sum_{\nu=3}^{k+1}\eta_\nu d_\nu\right)^3, \\ P_3 &= 3\eta_1 d_1^2(2\eta_1 - 1)\sum_{\nu=3}^{k+1}\eta_\nu d_\nu, \\ P_4 &= 3\eta_1 d_1\left[2\left(\sum_{\nu=3}^{k+1}\eta_\nu d_\nu\right)^2 - \sum_{\nu=3}^{k+1}\eta_\nu d_\nu^2\right]. \end{aligned}$$

We have $P_1>0$ for $\eta_1>1/2$, since $d_1<0$.

If we regard the numbers $0, d_3, \cdots, d_{k+1}$ as λ's and the numbers $(\eta_1+\eta_2), \eta_3, \cdots, \eta_{k+1}$ as weights, then the expression P_2 is a function $F(t)$ with $n=k$. Since $(\eta_1+\eta_2)>1/2$ whenever $\eta_1>1/2$, the induction hypothesis implies that $P_2>0$.

We have $P_3>0$ since $d_\nu>0$ for $\nu=3, \cdots, k+1$.

Cauchy's inequality gives:

$$\left(\sum_{\nu=3}^{k+1}\eta_\nu d_\nu\right)^2 \leqq \left(\sum_{\nu=3}^{k+1}\eta_\nu d_\nu^2\right)\left(\sum_{\nu=3}^{k+1}\eta_\nu\right) < \frac{1}{2}\left(\sum_{\nu=3}^{k+1}\eta_\nu d_\nu^2\right).$$

This inequality together with $d_1<0$ gives $P_4>0$. This concludes the proof of part (a).

If $\eta_1=1/2$ and $n>2$, then $P_1=P_3=0$, $P_2>0$ and $P_4>0$. Hence for $n>2$, we have $F(t_1)>0$. The rest of the theorem may be verified by using a similar procedure (for example, replace λ_ν by $\lambda_\nu-\lambda_{k-1}$, and so on).

4. **The behavior of $\Lambda''(t)$.** We shall establish the following result:

THEOREM 2. *There exists a t_1 such that $t<t_1$ implies that $\Lambda''(t)>0$.*

PROOF. By Theorem 1 and the fact that η_1 decreases continuously from plus 1 to 0, there is a value p such that $f'''(p)=0$ and such that $t<p$ implies $f'''(t)>0$.

(a) If $p>0$, then $\Lambda''(t)>0$ for all t on the range $-\infty<t\leqq p$. Since $\Lambda''(0)=(1/3)f'''(0)$, we have $\Lambda''(0)>0$. That $\Lambda''(t)>0$ for all nonzero t on the given range follows from formula (8).

(b) If $p=0$, then $\Lambda''(0)>0$ for all negative t, and $\Lambda''(0)=0$.

(c) If $p<0$ and $\Lambda''(\bar{t})\leqq 0$ for some $\bar{t}<p$, then from formula (7) we have $\Lambda''(t)<0$ for all t satisfying $\bar{t}<t\leqq p$. But $\Lambda(t)$ is an increasing and bounded function of t. Hence for some $t_1\leqq p$, $\Lambda''(t_1)>0$ and for such a t_1, $t<t_1$ will imply that $\Lambda''(t)>0$.

COROLLARY 1. *There exists a t_1 such that $M(t)$ is convex for $t\leqq t_1$.*

COROLLARY 2. *If $\xi_1\geqq 1/2$, then both $\Lambda(t)$ and $M(t)$ are convex for all negative t.*

PROOF. Since $\eta_1(0)=\xi_1$, the inequality $\xi_1\geqq 1/2$ implies that $f'''(t)>0$ for all negative t. The conclusion for $\Lambda(t)$ follows from parts (a) and (b) of the proof of Theorem 2. Since the convexity of the logarithm of a function always implies the convexity of the function, the result for $M(t)$ follows.

The same methods may be used to prove:

THEOREM 2′. *There exists a* t_2 *such that* $t>t_2$ *implies that* $\Lambda''(t)<0$.

COROLLARY 2′. *If* $\xi_n \geqq 1/2$, *then* $\Lambda(t)$ *is concave for all positive t.*

THEOREM 3. *If* $f'''(t)$ *has exactly one zero, then so does* $\Lambda''(t)$.

PROOF. Let $f'''(\bar{t})=0$. Then we have $f'''(t)>0$ for $t<\bar{t}$ and $f'''(t)<0$ for $t>\bar{t}$. If $\bar{t}>0$, the proofs of Theorems 2 and 2′ show that $\Lambda''(t)>0$ on the range $-\infty<t\leqq\bar{t}$ and that $\Lambda''(t)$ has one zero on the range $\bar{t}<t<+\infty$. The other cases $\bar{t}=0$ and $\bar{t}<0$ are similar.

5. **Convex-concave functions.** A function $g(t)$ defined on the range $-\infty<t<+\infty$ is convex-concave providing there exists a point $\bar{t}$ such that $g(t)$ is convex for $t\leqq\bar{t}$ and concave for $\bar{t}\leqq t$.

The question may be raised whether either of the functions $\Lambda(t)$ or $M(t)$ is necessarily convex-concave. By Theorems 1 and 3 the answer is affirmative for $\Lambda(t)$ in the case $n=2$. We now develop an example to show that this conclusion regarding either $\Lambda(t)$ or $M(t)$ is not true for $n>2$.

Consider the case $a_\nu=e^\nu$ $(\nu=1, 2, 3)$, $\xi_1=\xi_3<1/6$, $\lambda_\nu=\nu$. If λ_ν is replaced by $\lambda_\nu+c$, $f'(t)$ is increased by c and every higher order derivative of $f(t)$ remains invariant. Therefore, to calculate both $f'''(0)$ and $f^{\text{IV}}(0)$, the set (1, 2, 3) may be replaced by the set $(-1, 0, +1)$. We have that $f'''(0)$ is equal to $-\xi_1+\xi_3=0$ and, by formula (5), that $f^{\text{IV}}(0)$ is equal to $(\xi_1+\xi_3)-3(\xi_1+\xi_3)^2>0$. Therefore in a neighborhood of zero, $f'''(t)$ has the same sign as t. From formula (8), it follows that $\Lambda''(t)$ agrees with $f'''(t)$ in sign throughout this neighborhood and consequently that $\Lambda''(t)$ has at least three zeros. In particular for the mean value function $M(t)=(.1e^t+.8e^{2t}+.1e^{3t})^{1/t}$, the following table shows that there are values t_1 and t_2 such that $t_1<t_2$ and $M''(t_1)<0<M''(t_2)$ or that $M(t)$ is not convex-concave.

Table of M and its derivatives for $M(t)=(.1e^t+.8e^{2t}+.1e^{3t})^{1/t}$

t	-1.2	-0.8	-0.4	0.0
$M(t)$	6.519	6.810	7.098	7.389
$M'(t)$	0.7311	0.7215	0.7209	0.7389
$M''(t)$	-0.026	-0.017	$+0.018$	$+0.074$

6. **Generalized weighted sums.** Let (a) again denote a set of distinct positive numbers $a_1, a_2, \cdots, a_n$ and (ξ) a set of arbitrary positive numbers $\xi_1, \xi_2, \cdots, \xi_n$, $n\geqq 2$. Consider the function $S_t(a, \xi)$

$=[\sum_{\nu=1}^{n}\xi_\nu a_\nu^t]^{1/t}$ for $t\neq 0, \pm\infty$;

$$S_0(a, \xi) = \prod_{\nu=1}^{n} a_\nu^{\xi_\nu} \quad \text{if} \quad \sum_{\nu=1}^{n} \xi_\nu = 1;$$

$$S_0(a, \xi) = 0 \quad \text{if} \sum_{\nu=1}^{n} \xi_\nu \neq 1;$$

$$S_{-\infty}(a, \xi) = \min_{\nu=1,2,\cdots,n} a_\nu;$$

and

$$S_{+\infty}(a, \xi) = \max_{\nu=1,2,\cdots,n} a_\nu.$$

For fixed sets (a) and (ξ), the function $S_t(a, \xi)$ is continuous for $-\infty \leqq t < 0$ and for $0 < t \leqq +\infty$. If $\sum_{\nu=1}^{n}\xi_\nu \leqq 1$, then $S_t(a, \xi)$ is continuous from the right at $t=0$; and if $\sum_{\nu=1}^{n}\xi_\nu \geqq 1$, then $S_t(a, \xi)$ is continuous from the left at $t=0$. However, if we define f, η_ν, and S_k as in formulae (1), (2), and (3), then all of the results of §3 regarding $f'''(t)$ are valid without modification. Formula (7) will hold as long as t_1 and t_2 are both positive or both negative.

THEOREM 4. (a) *There is a point* p_1 *such that* $\log S_t(a, \xi)$ *is either a concave function of* t *for all* $t < p_1$ *or a convex function for all* $t < p_1$.

(b) *There is a point* p_2 *such that* $\log S_t(a, \xi)$ *is convex for all* $t > p_2$, *or concave for all* $t > p_2$.

PROOF OF THEOREM 4a. Assume that for all p, $\log S_t(a, \xi)$ is not convex in the range $-\infty < t < p$. Let $p_1 < 0$ be such that $t < p_1$ implies that $f'''(t) > 0$. Consider any number $t_0 < p_1$. Then, by assumption, $d^2 \log S_t(a, \xi)/dt^2 < 0$ has a solution for some $t' < t_0$. Formula (7) implies that this second derivative is negative for all t satisfying $t' < t \leqq p_1$, or that $\log S_t(a, \xi)$ is concave for all $t \leqq p_1$.

The proof of Theorem 4b is similar.

To show the applicability of Theorem 4 consider a function $S_t(a, \xi)$ where $\sum_{\nu=1}^{n}\xi_\nu < 1$. For all such sets (ξ) and for every set (a), the function $S_t(a, \xi)$ has the following properties:

(a) $t_1 < t_2$ and $t_1 t_2 > 0$ imply that $S_{t_1}(a, \xi) < S_{t_2}(a, \xi)$;

(b) $\min_{\nu=1,2,\cdots,n} a_\nu < S_t(a, \xi)$ for $t < 0$;

(c) $S_t(a, \xi) < \max_{\nu=1,2,\cdots,n} a_\nu$ for $t > 0$.

Hence there is no point p such that $t < p$ implies that $\log S_t(a, \xi)$ is concave nor a point p such that $t > p$ implies that $\log S_t(a, \xi)$ is convex. We conclude from Theorem 4 that if $\sum_{\nu=1}^{n}\xi_\nu < 1$, $\log S_t(a, \xi)$ is convex in a neighborhood of $-\infty$ and concave in a neighborhood of $+\infty$.