JOURNAL OF APPROXIMATION THEORY 17, 57–65 (1976)

Error Measures and Their Associated Means

F. T. METCALF

Department of Mathematics, University of California, Riverside, Riverside, California 92502

Communicated by Oved Shisha

Received August 2, 1974

1. INTRODUCTION

In a characteristically interesting and well-motivated note [1], Pólya considered, under the assumptions

$$a \leqslant x \leqslant b \quad \text{and} \quad 0 < a < b, \tag{1}$$

the following two problems:

PROBLEM I. Find

$$\min_p (\max_x | p - x |)$$

and the value of p for which it is attained under conditions (1).

PROBLEM II. Find

$$\min_p (\max_x | p - x |/x)$$

and the value of p for which it is attained under conditions (1).

In a direct fashion, without explicit recourse to the Tchebyscheff alternation theorem, Pólya shows that the solutions to these problems are given by:

Solution I (well known).

$$p = (a + b)/2, \qquad \min_p (\max_x | p - x |) = (b - a)/2.$$

Solution II (not so well known).

$$p = 2ab/(a + b), \qquad \min_p (\max_x | p - x |/x) = (b - a)/(b + a).$$

The solution to Problem I states that if one wishes to minimize the extreme *absolute error* in estimating the unknown x, subject to $0 < a \leqslant x \leqslant b$, then one should take the *arithmetic mean* of a and b as the "best choice."

The solution to Problem II states that if one wishes to minimize the extreme of a certain *relative error* in estimating the unknown x, then one should take the *harmonic mean* of a and b as the "best choice."

Let $p^*(a, b)$ denote the best choice for a given minmax problem of the above type. The main points which may be extracted from Pólya's examples are the following.

(i) The minimum of the extreme error occurs at one of the end points of the interval $[a, b]$, and $p^*(a, b)$ is uniquely determined by equating the errors at the endpoints and specifying that $p^*(a, b) \in [a, b]$; for example, in Problem II

$$\min_p (\max_x | p - x |/x) = \frac{|(2ab/(a + b)) - a|}{a} = \frac{|(2ab/(a + b)) - b|}{b} .$$

(ii) The best choice $p^*(a, b)$, as a function of a and b, is a "mean."

For the familiar Problem I, observation (i) is simply the Tchebyscheff alternation procedure for the approximation of the function x by a consatnt. However, for Problem II, involving a relative error, the classical alternation theorem is not applicable. In papers subsequent to Pólya's, alternation theorems covering Problem II (as a very special case) have been established by Selfridge [2], Moursund [3], Dunham [4], and Griesel [5]. These authors have considered successively more general measures of the "error" involved in the approximation of one quantity by another. In Section 2 below, such a general "error measure" will be defined.

For both Problems I and II, observation (ii) conforms to the notion that a best approximation is in some sense a mean. However, in these cases, the means arising from the approximation problems are, in fact, two of the classical means of real numbers.

One purpose of the present paper is to demonstrate that, for general error measures, the simple approximation problem considered by Pólya always leads to a mean of a and b. A precise definition of mean will be given in Section 2.

A further example of the best approximation $p^*(a, b)$ is provided by the following problem involving a slightly different version of relative error. This example also illustrates the sensitivity of $p^*(a, b)$ with respect to changes in the error measure. By changing the denominator in the relative error Problem II, from x to $\max(p, x)$, the best approximation changes from the harmonic mean to the geometric mean.

PROBLEM III. Find

$$\min_p (\max_x | p - x |/\max(p, x))$$

and the value of p for which it is attained under conditions (1).

Solution III (easily verified).

$$p = (ab)^{1/2}, \qquad \min_p (\max_x | p - x |/\max(p, x)) = (b/a)^{1/2} - 1.$$

A further list of such problems could be given; however, such a list is available in Aissen [6]. The repeated appearance of the geometric mean in this list led Aissen to show that a certain large class of error measures give the geometric mean as the solution to the min max problem. These error measures are of the form $| p - x |/\varphi(p, x)$, where the conditions on φ which single out the geometric mean are: (i) $\varphi(p, x) = \varphi(x, p)$; (ii) φ homogeneous (order 1).

The second purpose of this paper is to provide a representation for all of the error measures which yield a given mean. One of the special cases will show how conditions (i) and (ii), just above, can be replaced by the condition

$$\varphi(p, x) = (x/p)\, \varphi(p, p^2/x).$$

Also, this condition on φ will be shown to be necessary and sufficient for the error measure $| p - x |/\varphi(p, x)$ to yield the geometric mean.

Remark 1. As pointed out in Aissen [6], error measures of the form $| p - x |/\varphi(p, x)$ are suggested in Huntington [7], a paper dealing with the apportionment of representatives in Congress. The utilization of a relative error in attacking this problem served to eliminate some inequities arising from procedures based on the absolute error.

Remark 2. It should also be mentioned here that Beckenbach [8] has treated "least squares" versions of Problems I and II, and especially the mean arising from that version of Problem II. However, this type of problem will not be dealt with here.

2. Error Measures and Means

As mentioned earlier, generalizations of the usual error have recently been provided by a number of authors. Moursund used the term "weight function," and Dunham the term "order function." The terminology "error measure" will be adopted here, as it seems to be more descriptive of the role this type of function plays.

Definition. A real-valued function $\mathscr{E}$, defined on $[A, B] \times [A, B]$, where $-\infty < A < B < +\infty$, will be said to be an *error measure* if:

(E1) $\mathscr{E}$ is continuous;

(E2) $0 \leqslant \mathscr{E}(p, x)$ for $p, x \in [A, B]$;

(E3) $\mathscr{E}(p, x) = 0$ if and only if $p = x$;

(E4) for $p < x$, $\mathscr{E}(p, x)$ (strictly) decreases in p and increases in x; for $p > x$, $\mathscr{E}(p, x)$ increases in p and decreases in x.

Condition (E4) simply gives to $\mathscr{E}(p, x)$ the same monotonicity properties as possessed by $| p - x |$, while the first three conditions are motivated by obvious considerations. The min max problem for this general error measure is to find

$$\min_{A \leqslant p \leqslant B} (\max_{a \leqslant x \leqslant b} \mathscr{E}(p, x)),$$

and the value of p for which it is attained, where $[a, b] \subseteq [A, B]$.

THEOREM 1. *Let $\mathscr{E}$ be an error measure. Then*

$$\min_{A \leqslant p \leqslant B} (\max_{a \leqslant x \leqslant b} \mathscr{E}(p, x)) = \mathscr{E}(p^*, a) = \mathscr{E}(p^*, b),$$

where $p^* = p^*(a, b)$ is unique, $a < p^* < b$ if $a < b$, and $p^* = a$ if $a = b$.

Proof. From the monotonicity properties of $\mathscr{E}$ it follows that

$$\max_{a \leqslant x \leqslant b} \mathscr{E}(p, x) = \max[\mathscr{E}(p, a), \mathscr{E}(p, b)].$$

Since the right-hand side is continuous in p, one has the existence of $p^* \in [A, B]$ such that the min max problem is solved. Further, one must have $\mathscr{E}(p, a) = \mathscr{E}(p, b)$ at $p = p^*$.

The monotonicity properties of $\mathscr{E}$ also show that $\max[\mathscr{E}(p, a), \mathscr{E}(p, b)]$ is decreasing for $A \leqslant p < a$, and increasing for $b < p \leqslant B$ (if $A < a$ and $b < B$). Thus, one has $p^* \in [a, b]$, and $a < p^* < b$ when $a < b$.

Finally, since $\mathscr{E}(p, a)$ increases and $\mathscr{E}(p, b)$ decreases, for $a \leqslant p \leqslant b$, there can exist at most one p^* such that $\mathscr{E}(p^*, a) = \mathscr{E}(p^*, b)$. ∎

The following lemmas complete the requirements for $p^*(a, b)$ to be a mean. They verify several reasonable properties of the optimal proximate p^*. However, it should first be noted that $p^*(a, b)$ has been defined only for $A \leqslant a \leqslant b \leqslant B$. It will now be assumed that p^* has been extended *symmetrically* to all of $[A, B] \times [A, B]$.

LEMMA 1. *The function p^* is continuous on $[A, B] \times [A, B]$.*

Proof. Suppose $(a, b) \in [A, B] \times [A, B]$. Then there exists a sequence $\{(a_i, b_i)\}_{i=1}^{\infty}$ such that

$$(a_i, b_i) \in [A, B] \times [A, B], \qquad i = 1, 2, \ldots,$$
$$(a_i, b_i) \to (a, b) \qquad \text{as} \quad i \to \infty,$$

and

$$\lim_{i\to\infty} p^*(a_i, b_i) \qquad \text{exists} \qquad (= \xi).$$

Then

$$\mathscr{E}(p^*(a_i, b_i), a_i) \to \mathscr{E}(\xi, a) \qquad \text{(by (E1))},$$
$$\mathscr{E}(p^*(a_i, b_i), b_i) \to \mathscr{E}(\xi, b) \qquad \text{(by (E1))},$$

and

$$\mathscr{E}(p^*(a_i, b_i), a_i) = \mathscr{E}(p^*(a_i, b_i), b_i) \qquad \text{(definition of } p^*\text{)},$$

giving $\mathscr{E}(\xi, a) = \mathscr{E}(\xi, b)$. Thus $\xi = p^*(a, b)$ by the uniqueness of $p^*(a, b)$, and p^* is continuous at (a, b). ∎

LEMMA 2. *The function p^* is increasing in each of its arguments (the other argument being held fixed).*

Proof. Fix $b \in [A, B]$, and suppose that $A \leqslant a_1 < a_2 \leqslant B$ (the case of fixed a is treated similarly, or by using the symmetry of p^*). Suppose, to the contrary, that $p^*(a_2, b) \leqslant p^*(a_1, b)$. Then three cases arise:

(i) $a_1 < a_2 \leqslant b$. In this case one has

$$\begin{aligned} \mathscr{E}(p^*(a_1, b), a_1) &= \mathscr{E}(p^*(a_1, b), b) && \text{(definition of } p^*\text{)} \\ &\leqslant \mathscr{E}(p^*(a_2, b), b) && \text{(property (E4))} \\ &= \mathscr{E}(p^*(a_2, b), a_2) && \text{(definition of } p^*\text{)} \\ &\leqslant \mathscr{E}(p^*(a_1, b), a_2) && \text{(property (E4))}. \end{aligned}$$

However, $a_1 < a_2 \leqslant p^*(a_2, b) \leqslant p^*(a_1, b)$ and the extremes of the above chain of inequalities contradict the monotonicity of $\mathscr{E}$ in this situation.

(ii) $b \leqslant a_1 < a_2$. Using arguments similar to those in Case (i), one has

$$\begin{aligned} \mathscr{E}(p^*(a_1, b), a_1) &= \mathscr{E}(p^*(a_1, b), b) \\ &\geqslant \mathscr{E}(p^*(a_2, b), b) \\ &= \mathscr{E}(p^*(a_2, b), a_2) \\ &\geqslant \mathscr{E}(p^*(a_1, b), a_2). \end{aligned}$$

However, $p^*(a_1, b) \leqslant a_1 < a_2$ and the extreme inequality above again contradict the monotonicity of $\mathscr{E}$ in this situation.

(iii) $a_1 \leqslant b \leqslant a_2$. From Cases (i) and (ii) it follows that

$$\begin{aligned} p^*(a_1, b) &\leqslant p^*(b, b) && \text{(Case (i) with } a_2 = b\text{)} \\ &\leqslant p^*(a_2, b) && \text{(Case (ii) with } a_1 = b\text{)}. \end{aligned}$$

At least one of these last inequalities must be strict, since $a_1 < a_2$; and hence, $p^*(a_1, b) < p^*(a_2, b)$. ∎

Summarizing the above, the optimal proximate p^* has the following properties for $(a, b) \in [A, B] \times [A, B]$:

(P1) p^* is continuous;

(P2) $p^*(a, b) = p^*(b, a)$;

(P3) $\min(a, b) < p^*(a, b) < \max(a, b)$ for $a \neq b$, and $p^*(a, a) = a$;

(P4) p^* is increasing in each variable.

Thus, p^* possesses those properties which might reasonably be expected of a mean; and, for the purposes of this paper, the above properties will constitute the definition of a *mean*.

3. Representation for Error Measures Yielding a Given Mean

The preceding section established that every error measure has an associated mean p^* (see Properties (P1)–(P4)). An interesting problem is to classify, in some way, those error measures associated with a given mean. This problem was partially attacked by Aissen in [6], where it was shown that error measures of the form

$$\mathscr{E}(p, x) = | p - x | / \varphi(p, x) \tag{2}$$

(φ satisfying certain natural conditions) are always associated with the *geometric mean* if

$$\varphi(p, x) = \varphi(x, p) \qquad \text{and} \qquad \varphi(\lambda p, \lambda x) = \lambda \varphi(p, x) \tag{3}$$

(i.e., φ symmetric and homogeneous of order 1). A representation for the error measures corresponding to a given mean is given below, with some special cases considered as examples.

In the following, M will denote a mean defined on $[A, B] \times [A, B]$, i.e., a function satisfying (P1)–(P4). Since M is increasing in both variables, there is a unique function

$$\alpha_M : \left\{ \begin{matrix} M(A, b) \leqslant p \leqslant M(B, b) \\ A \leqslant b \leqslant B \end{matrix} \right\} \to [A, B]$$

such that

$$M(\alpha_M(p, b), b) = p \qquad \text{for} \quad \begin{matrix} M(A, b) \leqslant p \leqslant M(B, b) \\ A \leqslant b \leqslant B. \end{matrix}$$

That is to say, the equation $M(a, b) = p$ has a unique solution in the variable a for the indicated ranges on p and b.

THEOREM 2. *Suppose $\mathscr{E}$ is an error measure and M a mean, both defined on $[A, B] \times [A, B]$. Then M is the mean associated with $\mathscr{E}$ if and only if*

$$\mathscr{E}(p, x) = \mathscr{S}(\alpha_M(p, x), x), \qquad M(A, x) \leqslant p \leqslant M(B, x), \quad A \leqslant x \leqslant B, \tag{4}$$

where $\mathscr{S}$ is a symmetric *error measure.*

Proof. ($\Rightarrow$) Suppose that M is the mean associated with $\mathscr{E}$; so that

$$\mathscr{E}(M(u, x), u) = \mathscr{E}(M(u, x), x) \tag{5}$$

for all $u, x \in [A, B]$. For $u, x \in [A, B]$ define $\mathscr{S}$ by

$$\mathscr{S}(u, x) = \mathscr{E}(M(u, x), x).$$

From (5) and the symmetry of M it follows that $\mathscr{S}$ is symmetric. Also, $\mathscr{S}$ is continuous, and $\mathscr{S}(u, x) \geqslant 0$ with equality if and only if $u = x$. Thus, only the appropriate monotonicity conditions need be verified in order to show that $\mathscr{S}$ is an error measure.

If $u < x$, then $M(u, x) < x$ and the monotonicity properties of $\mathscr{E}$ give that $\mathscr{S}$ is decreasing in u. From (5) and $u < M(u, x)$, it also follows that $\mathscr{S}$ is increasing in x. The cases for $x < u$ follow in a similar fashion, or by using the symmetry of $\mathscr{S}$. Hence, $\mathscr{S}$ is a symmetric error measure.

For $M(A, x) \leqslant p \leqslant M(B, x)$ and $A \leqslant x \leqslant B$, set $u = \alpha_M(p, x)$. Then

$$\begin{aligned} \mathscr{S}(\alpha_M(p, x), x) &= \mathscr{E}(M(\alpha_M(p, x), x), x) \\ &= \mathscr{E}(p, x) \end{aligned}$$

from the definition of α_M.

($\Leftarrow$) Given that (4) holds, it will now be shown that the unique solution of $\mathscr{E}(p, a) = \mathscr{E}(p, b)$ is $p = M(a, b)$. Since uniqueness is given by Theorem 1, it remains to show $\mathscr{E}(M(a, b), a) = \mathscr{E}(M(a, b), b)$. But

$$\begin{aligned} \mathscr{E}(M(a, b), a) &= \mathscr{S}(\alpha_M(M(a, b), a), a) \\ &= \mathscr{S}(b, a) = \mathscr{S}(a, b) \\ &= \mathscr{S}(\alpha_M(M(a, b), b), b) \\ &= \mathscr{E}(M(a, b), b), \end{aligned}$$

where (4), the definition of α_M, and the symmetry of $\mathscr{S}$ have all been used.

The applicability of (4) is a consequence of the fact that

$$M(A, a) \leqslant M(a, b) \leqslant M(B, a)$$

and

$$M(A, b) \leqslant M(a, b) \leqslant M(B, b). \quad \blacksquare$$

To illustrate Theorem 2, the particular cases of the arithmetic mean and the geometric mean will be considered in the following corollaries. Other examples can readily be worked out, since the only "effort" required is to find the function α_M.

COROLLARY 2.1. *The arithmetic mean* $M_{\mathscr{A}}(a, b) = (a + b)/2$, $-\infty < a$, $b < +\infty$, *is associated with an error measure* $\mathscr{E}$ *if and only if*

$$\mathscr{E}(p, x) = \mathscr{S}(2p - x, x), \qquad -\infty < p, \quad x < +\infty,$$

where $\mathscr{S}$ *is a* symmetric *error measure.*

Proof. Solving $p = (a + b)/2$ for a gives

$$\alpha_{M_{\mathscr{A}}}(p, b) = 2p - b.$$

The restriction on the range of p is not of consequence here since one could take $A = -\infty$ and $B = +\infty$, with no essential change in the above proofs. ∎

COROLLARY 2.2. *The geometric mean* $M_{\mathscr{G}}(a, b) = (ab)^{1/2}$, $0 < a$, $b < +\infty$, *is associated with an error measure* $\mathscr{E}$ *if and only if*

$$\mathscr{E}(p, x) = \mathscr{S}(p^2/x, x), \qquad 0 < p, \quad x < +\infty,$$

where $\mathscr{S}$ *is a* symmetric *error measure.*

Proof. In this case $\alpha_{M_{\mathscr{G}}}(p, b) = p^2/b$, and the range of p may again be modified, this time to $0 < p < +\infty$. ∎

It is interesting to consider this last corollary in the light of Aissen's result (see (2) and (3)). Restricting Corollary 2.2 to error measures of the form (2), one has that the geometric mean is associated with $\mathscr{E}$ if and only if

$$|p - x|/\varphi(p, x) = \mathscr{S}(p^2/x, x), \qquad 0 < p, \quad x < +\infty, \tag{6}$$

where $\mathscr{S}$ is a symmetric error measure. Rewriting the arguments on the left of (6) in terms of p^2/x and x, and using the symmetry of $\mathscr{S}$, gives

$$\frac{|((p^2/x) \cdot x)^{1/2} - x|}{\varphi(((p^2/x) \cdot x)^{1/2}, x)} = \frac{|(x \cdot (p^2/x))^{1/2} - (p^2/x)|}{\varphi((x \cdot (p^2/x))^{1/2}, p^2/x)}.$$

This relation reduces to

$$\varphi(p, x) = (x/p)\,\varphi(p, p^2/x), \qquad 0 < p, \quad x < +\infty, \tag{7}$$

as the necessary and sufficient condition for the geometric mean to be associated with error measures of the form (2). It is readily seen that φ symmetric and homogeneous (order 1) is sufficient for (7) to hold.

References

1. G. Pólya, On the harmonic mean of two numbers, *Amer. Math. Monthly* **57** (1950), 26–28.
2. R. G. Selfridge, Approximations with least maximum error, *Pacific J. Math.* 3 (1953), 247–255.
3. David G. Moursund, Chebyshev approximation using a generalized weight function, *SIAM J. Numer. Anal.* **3** (1966), 435–450.
4. Charles B. Dunham, Chebyshev approximation with respect to a weight function, *J. Approximation Theory* **2** (1969), 223–232.
5. Martha Ann Griesel, Uniform approximation with respect to a general error function, Ph.D. dissertation, Univ. California, Riverside, 1971.
6. Michael Aissen, Means and the minimization of errors, *J. Res. Nat. Bur. Stand.* **72B** (1968), 1–4.
7. E. V. Huntington, The apportionment of representatives in Congress, *Trans. Amer. Math. Soc.* **30** (1928), 85–110.
8. E. F. Beckenbach, A class of mean value functions, *Amer. Math. Monthly* **57** (1950), 1–6.