

Minimax Error Selection of a Discrete Univariate Distribution With Prescribed Componentwise Bounds

A. J. Goldman and P. R. Meyers*

Institute for Basic Standards, National Bureau of Standards, Washington, D.C. 20234

(September 20, 1968)

The topic treated is that of finding a reproducible, plausible and computationally simple method of selecting a discrete frequency distribution with prescribed upper and lower bounds on its components. The problem is shown to be tractable when a minimax error selection criterion is employed, and "error" is measured by maximum absolute deviation between components. In this case one obtains a linear program of a special form admitting explicit solution. The vertices of the polyhedron of optimal solutions can also be found explicitly, and so their centroid can be calculated if unique specification is required.

Key Words: Linear programs; mathematical models; operations research.

1. Introduction

In the mathematical modelling efforts associated with an operations research study, one may well have only incomplete information on which to base a representation of the probabilities of the various outcomes of some pertinent chance event. Under these circumstances, one should of course examine the consequences of several alternative probability distributions, each consistent with the information at hand. It still seems desirable, however, to have a systematic and reproducible method for arriving at a *single* "nominal" distribution, to serve as a base-point for such sensitivity analyses.

This note works out the mathematics of one approach, based on a "minimax error" criterion, to the selection of a nominal distribution. The "incomplete information" is assumed to consist of upper and lower bounds on the individual terms of the probability distribution.

Let \mathbf{L} and \mathbf{U} be real n -vectors whose components satisfy

$$0 \leq L_i < U_i \leq 1. \quad (1.1)$$

A real n -vector \mathbf{x} will be called a *probability vector* if it has nonnegative entries which sum to 1. We will be concerned with the set $P(\mathbf{L}, \mathbf{U})$ of probability vectors \mathbf{x} whose components satisfy

$$L_i \leq x_i \leq U_i. \quad (1.2)$$

Our objective is to choose $\mathbf{x} \in P(\mathbf{L}, \mathbf{U})$ to minimize

$$F(\mathbf{x}) = \max \{d(\mathbf{x}, \mathbf{y}) : \mathbf{y} \in P(\mathbf{L}, \mathbf{U})\} \quad (1.3)$$

where d is the metric on n -space given by

$$d(\mathbf{x}, \mathbf{y}) = \max_i |x_i - y_i|. \quad (1.4)$$

*Present address: IBM World Trade Corporation, 821 United Nations Plaza, New York, New York 10017.

It would be interesting to consider the problem with the alternate metrics

$$d(\mathbf{x}, \mathbf{y}) = \sum_i (x_i - y_i)^2, \quad d(\mathbf{x}, \mathbf{y}) = \sum_i |x_i - y_i|, \quad (1.5)$$

but this will not be done here. The common features of (1.4) and (1.5) are invariance under translation—used in the next section—and convexity in \mathbf{y} for fixed \mathbf{x} . The latter property implies that the maximum in (1.3) occurs at an extreme point of $P(\mathbf{L}, \mathbf{U})$, and since the extreme points of this polyhedron can be found explicitly, $F(\mathbf{x})$ can be written as a discrete extremum. Thus the remainder of the problem—the minimization of F —has a strongly combinatorial flavor, but appears to be considerably harder for (1.5) than for (1.4).

The solution method is derived and justified in sections 2–4. In general there will be a convex polyhedron of optimal \mathbf{x} 's rather than a single one; section 5 proposes the centroid of this polyhedron's vertices as a plausible “representative” choice, and shows how these vertices can be calculated. If they are not too numerous, they might perhaps provide the basis for the sort of sensitivity analysis mentioned above. Section 6 shows how the solution method extends to the generalization

$$d(\mathbf{x}, \mathbf{y}) = \max_i w_i |x_i - y_i|$$

of (1.4), where the w_i are given positive numbers.

2. Simplification and Feasibility Analysis

From (1.3) and the form of d , it is apparent that the constraint set $P(\mathbf{L}, \mathbf{U})$ can be replaced by any of its translates without changing the problem. Our *first* simplification is a translation through $(-\mathbf{L})$, which replaces $P(\mathbf{L}, \mathbf{U})$ by the set

$$Q(S, \mathbf{V}) = \{\mathbf{x} : 0 \leq x_i \leq V_i \text{ (all } i), \sum_i x_i = S\} \quad (2.1)$$

where $\mathbf{V} = \mathbf{U} - \mathbf{L} > 0$ and $S = 1 - \sum_i L_i$. Now (1.3) is replaced by

$$F(\mathbf{x}) = \max \{d(\mathbf{x}, \mathbf{y}) : \mathbf{y} \in Q(S, \mathbf{V})\}. \quad (2.2)$$

We turn next to a feasibility analysis, i.e., to determining when $Q(S, \mathbf{V})$ will be nonempty. The condition

$$0 \leq S \leq \sum_i V_i \quad (2.3)$$

is obviously necessary; we will show it is also sufficient. If equality holds throughout (2.3), then clearly $Q(S, \mathbf{V}) = \{\mathbf{0}\}$. If equality does not hold, then $(S/\sum_i V_i)\mathbf{V}$ is in $Q(S, \mathbf{V})$. From now on we assume (2.3) holds, in fact that

$$0 < S < \sum_i V_i, \quad (2.4)$$

since otherwise $Q(S, \mathbf{V})$ is a single point and the whole problem is trivial.

3. Evaluation of F

$F(\mathbf{x})$, as given by (1.4) and (2.2), is readily evaluated in terms of the quantities

$$M_i = \max \{y_i : \mathbf{y} \in Q(S, \mathbf{V})\}, \quad (3.1)$$

$$m_i = \min \{y_i : \mathbf{y} \in Q(S, \mathbf{V})\}. \quad (3.2)$$

Indeed, we have

$$F(\mathbf{x}) = \max_y \max_i \max \{y_i - x_i, x_i - y_i\} = \max_i \max \{\max_y (y_i - x_i), \max_y (x_i - y_i)\}$$

or finally

$$F(\mathbf{x}) = \max_i \max \{M_i - x_i, x_i - m_i\}. \quad (3.3)$$

We next find explicit expressions for M_i and m_i , so that (3.3) is also explicit. It will be shown that

$$M_i = \min \{V_i, S\}, \quad (3.4)$$

$$m_i = \max \{0, S - \sum_{j \neq i} V_j\}. \quad (3.5)$$

If M_i^* and m_i^* denote the right-hand sides of (3.4) and (3.5), then clearly any $\mathbf{y} \in Q(S, \mathbf{V})$ satisfies $m_i^* \leq y_i \leq M_i^*$. It suffices to show that $y_i = M_i^*$ for some $\mathbf{y} \in Q(S, \mathbf{V})$, and that $y_i = m_i^*$ for some $\mathbf{y} \in Q(S, \mathbf{V})$.

For the first purpose, if $V_i \leq S$ we note from (2.4) that

$$0 \leq \theta = (S - V_i) / \sum_{j \neq i} V_j < 1,$$

and set

$$y_i = V_i, y_j = \theta V_j \text{ for } j \neq i.$$

If $S < V_i$, we set

$$y_i = S, y_j = 0 \quad \text{for } j \neq i.$$

For the second purpose, if $S \geq \sum_{j \neq i} V_j$ we set

$$y_i = S - \sum_{j \neq i} V_j, y_j = V_j \text{ for } j \neq i.$$

If $S < \sum_{j \neq i} V_j$ we set

$$y_i = 0, y_j = V_j S / \sum_{k \neq i} V_k \text{ for } j \neq i.$$

This completes the proof of (3.4) and (3.5).

We now make a second simplification, namely translation of the constraint set $Q(S, \mathbf{V})$, through $(-\mathbf{m})$ where \mathbf{m} is the vector with components m_i . The result is a new constraint set $Q(S', \mathbf{V}')$, where

$$S' = S - \sum_i m_i, \mathbf{V}' = \mathbf{V} - \mathbf{m}.$$

However, the primes will be dropped. We now have all $m_i = 0$, so that

$$F(\mathbf{x}) = \max_i \{M_i - x_i, x_i\}, \quad (3.6)$$

$$S \geq \sum_{j \neq i} V_j \quad (\text{all } i). \quad (3.7)$$

Note that (3.4) still holds.

4. Minimization of F

The minimization of F over $Q(S, \mathbf{V})$ can be viewed as a linear program, namely to select z and \mathbf{x} so as to minimize z , subject to $\mathbf{x} \in Q(S, \mathbf{V})$ and to

$$z \geq M_i - x_i; z \geq x_i \quad (\text{all } i).$$

These last constraints are equivalent to

$$M_i - z \leq x_i \leq z,$$

so that the linear program's full set of constraints reads

$$\max \{0, M_i - z\} \leq x_i \leq \min \{z, V_i\} \quad (\text{all } i), \quad (4.1)$$

$$\sum_i x_i = S. \quad (4.2)$$

For fixed z , the existence of an \mathbf{x} satisfying (4.1) and (4.2) clearly implies that

$$\sum_i \max \{0, M_i - z\} \leq S \leq \sum_i \min \{z, V_i\}, \quad (4.3)$$

$$z \geq \frac{1}{2} \max_i M_i > 0. \quad (4.4)$$

We now show that the converse also holds. Note first that (3.4) and (4.4) yield

$$V_i \geq M_i > M_i - z,$$

which together with (4.4) yields

$$\max \{0, M_i - z\} \leq \min \{z, V_i\} \quad (\text{all } i). \quad (4.5)$$

If equality holds in the second part of (4.3), then (4.1) and (4.2) are satisfied by taking

$$x_i = \min \{z, V_i\} \quad (\text{all } i).$$

If strict inequality holds, then $k = n$ has the property

$$\sum_{j \leq k} \min \{z, V_j\} + \sum_{j > k} \max \{0, M_j - z\} > S, \quad (4.6)$$

but $k = 0$ does not, and so there is a smallest $k \in \{1, 2, \dots, n\}$ with this property; for that k ,

$$\sum_{j < k} \min \{z, V_j\} + \sum_{j \geq k} \max \{0, M_j - z\} \leq S. \quad (4.7)$$

We now set

$$x_j = \min \{z, V_j\} \quad (\text{all } j < k),$$

$$x_j = \max \{0, M_j - z\} \quad (\text{all } j > k),$$

$$x_k = S - \sum_{j \neq k} x_j;$$

that (4.1) is satisfied for $i = k$ follows from (4.6) and (4.7).

The problem has now been reduced to minimizing z subject to (4.3) and (4.4). For this purpose, renumber so that

$$V_1 \geq V_2 \geq \dots \geq V_n > V_{n+1} = 0, \quad (4.8)$$

which by (3.4) implies

$$M_1 \geq M_2 \geq \dots \geq M_n > M_{n+1} = 0. \quad (4.9)$$

If z^* is the minimum value of $z > 0$ satisfying the first half of (4.3),

$$\sum_i \max \{0, M_i - z\} \leq S, \quad (4.10)$$

and z^{**} is the minimum value of $z > 0$ satisfying the second half,

$$\sum_i \min \{z, V_i\} \geq S, \quad (4.11)$$

then the desired minimum value of z is

$$z_{\min} = \max \{z^*, z^{**}, \frac{1}{2} \min \{S, V_1\}\}. \quad (4.12)$$

Thus it suffices to determine z^* and z^{**} .

The determination of z^* goes as follows. Let

$$M_j^* = \sum_i^j M_i - jM_j. \quad (4.13)$$

Then $M_1^* = 0 < S$, and by (2.4) we have $M_{n+1}^* \geq S$. Moreover, the calculation

$$M_{j+1}^* - M_j^* = j(M_j - M_{j+1}) \geq 0$$

shows that the sequence $\{M_j^*\}_1^{n+1}$ is nondecreasing. Thus there is a unique $J \in \{1, 2, \dots, n\}$ such that

$$0 = M_1^* \leq M_2^* \leq \dots \leq M_J^* < S \leq M_{J+1}^* \leq \dots \leq M_{n+1}^*. \quad (4.14)$$

Now if $0 < z < M_{J+1}$, then

$$\sum_i \max \{0, M_i - z\} \geq \sum_i^{J+1} M_i - (J+1)z > M_{J+1}^* \geq S,$$

so that z does not satisfy (4.10). And if $M_J \geq z \geq M_{J+1}$, then (4.10) becomes

$$\sum_i \{M_i - Jz\} \leq S, \quad (4.15)$$

which is equivalent to

$$z \geq z^* = (\sum_i \{M_i - S\})/J. \quad (4.16)$$

By use of (4.14), the value of z^* proposed in (4.16) is easily verified to satisfy $M_J \geq z^* \geq M_{J+1}$, and so is indeed the smallest $z > 0$ satisfying (4.10).

The determination of z^{**} goes as follows. Let

$$V_k^{**} = kV_k + \sum_{i=k+1}^n V_i. \quad (4.17)$$

Then by (2.4), we have $V_1^{**} > S$, and $0 = V_{n+1}^{**} < S$. Also, the calculation

$$V_{k+1}^{**} - V_k^{**} = k(V_{k+1} - V_k) \leq 0$$

shows that the sequence $\{V_k^{**}\}_1^{n+1}$ is nonincreasing. Thus there is a unique $K \in \{1, 2, \dots, n\}$ such that

$$V_1^{**} \geq V_2^{**} \geq \dots \geq V_K^{**} \geq S > V_{K+1}^{**} \geq \dots \geq V_{n+1}^{**}. \quad (4.18)$$

Now if $z \leq V_{K+1}$, then

$$\sum_i \min \{z, V_i\} \leq (K+1)z + \sum_{i=K+2}^n V_i \leq V_{K+1}^{**} < S,$$

so that z does not satisfy (4.11). And if $V_K \geq z > V_{K+1}$, then (4.11) becomes

$$Kz + \sum_{K+1}^n V_i \geq S,$$

which is equivalent to

$$z \geq z^{**} = (S - \sum_{K+1}^n V_i)/K. \quad (4.19)$$

By use of (4.18), the value of z^{**} proposed in (4.19) is easily verified to satisfy $V_K \geq z^{**} > V_{K+1}$, and so is indeed the smallest $z > 0$ satisfying (4.11).

The process for finding K , and thus z^{**} , can be shortened as follows. First observe that the constraint set $Q(S, \mathbf{V})$ can clearly be replaced by $Q(S, \mathbf{M})$, where the components M_i of \mathbf{M} are given by (3.4). Thus V_k 's are replaced by M_k 's, and V_k^{**} 's by M_k^{**} 's. From (4.13) and (4.17) we have

$$M_j^* + M_j^{**} = \sum_1^n M_i,$$

so that (4.18) is equivalent to

$$M_K^* \leq \sum_1^n M_i - S < M_{K+1}^*,$$

and K can be evaluated using the same sequence $\{M_j^*\}_1^n$ used to locate J . This sequence can be generated using the initial condition $M_1^* = 0$ and the recursion

$$M_{j+1}^* = M_j^* + j(M_j - M_{j+1}).$$

To conclude this section, we recapitulate the solution process developed above. The input data are the n -vectors \mathbf{L} and \mathbf{U} , satisfying $0 \leq \mathbf{L} < \mathbf{U}$. (If some $L_i = U_i$, we set $x_i = L_i$, drop the i th component from consideration, and rescale \mathbf{x} by $(1 - L_i)^{-1}$.)

STEP 1: Set $V_i = U_i - L_i$ for $1 \leq i \leq n$. Set $S = 1 - \sum_i L_i$.

STEP 2: If $S < 0$ or $S > \sum_i V_i$, STOP; the problem is infeasible. If $S = 0$ set $\mathbf{x} = \mathbf{0}$ and go to Step 10; if $S = \sum_i V_i$ set $\mathbf{x} = \mathbf{V}$ and go to Step 10.

STEP 3: Form the vector \mathbf{m} with components $m_i = \max \{0, S - \sum_{j \neq i} V_j\}$. Replace \mathbf{V} by $\mathbf{V} - \mathbf{m}$, and S by $S - \sum_i m_i$.

STEP 4: Renumber the components of \mathbf{V} to be in nonincreasing order. Record the corresponding permutation π of $\{1, 2, \dots, n\}$. Set $V_{n+1} = 0$.

STEP 5: Calculate $M_i^* = \min \{V_i, S\}$ for $1 \leq i \leq n+1$.

STEP 6: Calculate M_j^* , from the initial condition $M_1^* = 0$ and the recursion $M_{j+1}^* = M_j^* + j(M_j - M_{j+1})$, until a first one M_{j+1}^* is reached which is $\geq S$ and a first one M_{K+1}^* is reached which is $> \sum_1^n M_i - S$. Set $z^* = (\sum_1^n M_i - S)/J$ and $z^{**} = (S - \sum_{K+1}^n M_i)/K$.

STEP 7: Form $z = \max \{z^*, z^{**}, \frac{1}{2}M_1\}$. (This corresponds to the z_{\min} of the previous text.)

STEP 8: (Assumes any optimal \mathbf{x} will do.) Calculate S_j^* , from the initial condition $S_n^* = \sum_1^n \min \{z, V_i\}$ and the backwards recursion

$$S_{j-1}^* = S_j^* + \max \{0, M_j - z\} - \min \{z, V_j\},$$

until a first one S_{k-1}^* is reached which is $\leq S$. Define \mathbf{x} by

$$x_j = \min \{z, V_j\} \quad (\text{all } j < k),$$

$$x_j = \max \{0, M_j - z\} \quad (\text{all } j > k),$$

$$x_k = S - \sum_{j \neq k} x_j.$$

STEP 9: Permute the components of \mathbf{x} by π^{-1} . Then replace \mathbf{x} by $\mathbf{x} + \mathbf{m}$.

STEP 10: Replace \mathbf{x} by $\mathbf{x} + \mathbf{L}$.

5. The Centroid Solution

Once z_{\min} has been found, the optimal \mathbf{x} 's are defined by the constraints (4.1) and (4.2), which we repeat here,

$$\sum_i x_i = S, \quad (5.1)$$

$$a_i \leq x_i \leq b_i \quad (\text{all } i) \quad (5.2)$$

where $a_i = \max \{0, M_i - z_{\min}\}$ and $b_i = \min \{z_{\min}, V_i\}$. One member of this convex polyhedron of solutions can be found by the construction used to show that (4.3) and (4.4) implied the existence of a solution to (4.1) and (4.2). But this is somewhat arbitrary. In view of the original intention to fasten on a single \mathbf{x} , it seems less arbitrary (at least to the writers) to select the *centroid* of the vertices of the polyhedron. Fortunately, these extreme points can be found explicitly. In what follows we assume

$$\sum_i a_i < S < \sum_i b_i,$$

thus excluding the trivial cases in which the polyhedron degenerates to a single point.

The polyhedron lies in the $(n-1)$ -dimensional hyperplane (5.1); its extreme points are characterized as intersections of (5.1) with some $n-1$ of the $2n$ "bounding hyperplanes" $x_i = a_i$ and $x_i = b_i$. Such a point of intersection, $\mathbf{x}(j, A)$, is associated to each $j \in \{1, \dots, n\}$ and each of the 2^{n-1} subsets A of $\{1, 2, \dots, n\} - \{j\}$. The coordinates of $\mathbf{x}(j, A)$ are given by

$$x_i = a_i \quad i \in A,$$

$$x_i = b_i \quad i \in \{1, 2, \dots, n\} - \{j\} - A,$$

$$x_j = S - \sum_{i \neq j} x_i.$$

The extreme points of the polyhedron are precisely those $\mathbf{x}(j, A)$ for which

$$a_j \leq x_j \leq b_j.$$

6. The Weighted Version

We now generalize the objective function (1.3) to

$$F(\mathbf{x}) = \max \{ \max_i w_i |x_i - y_i| : \mathbf{y} \in P(\mathbf{L}, \mathbf{U}) \} \quad (6.1)$$

where $\{w_i\}_1^n$ is a set of positive "weights" with $\sum_i w_i = 1$. This corresponds to the case in which the accuracy, with which \mathbf{x} approximates the "true" distribution, is more important for some components than for others.

The reasoning of sections 2 and 3 go through unchanged, with occasional insertions of " w_i " in appropriate formulas. We find that (4.1) is replaced by

$$\max \{0, M_i - z/w_i\} \leq x_i \leq \min \{z/w_i, V_i\} \quad (\text{all } i) \quad (6.2)$$

so that (4.3) and (4.4) are replaced by

$$\sum_i \max \{0, M_i - z/w_i\} \leq S \leq \sum_i \min \{z/w_i, V_i\}, \quad (6.3)$$

$$z \geq \frac{1}{2} \max_i M_i w_i. \quad (6.4)$$

The problem is now that of minimizing z subject to (6.3) and (6.4). As before, we write

$$z_{\min} = \max \{z^*, z^{**}, \frac{1}{2} \max_i M_i w_i\}, \quad (6.5)$$

where now z^* is the minimum value of $z \geq 0$ satisfying

$$\sum_i \max \{0, M_i - z/w_i\} \leq S \quad (6.6)$$

and z^{**} is the minimum value of $z \geq 0$ satisfying

$$\sum_i \min \{z/w_i, V_i\} \geq S. \quad (6.7)$$

The renumbering now is to be such that

$$V_1 w_1 \geq V_2 w_2 \geq \dots \geq V_n w_n > 0 \quad (6.8)$$

which by (3.1) implies a similar condition for $\{M_i w_i\}_1^n$. As before we set $V_{n+1} = 0$ (implying $M_{n+1} = 0$), and now also choose any $w_{n+1} > 0$.

The determination of z^* goes as follows. Let

$$M_j^* = \sum_1^j M_i - w_j M_j \sum_1^j (1/w_i).$$

Then $M_1^* = 0 < S$, and by (2.4) we have $M_{n+1}^* \geq S$. The calculation

$$M_{j+1}^* - M_j^* = (w_j M_j - w_{j+1} M_{j+1}) \sum_1^j (1/w_i) \geq 0$$

shows that sequence $\{M_j^*\}_1^{n+1}$ is nondecreasing. Thus there is a unique $J \in \{1, 2, \dots, n\}$ such that

$$0 = M_1^* \leq M_2^* \leq \dots \leq M_J^* < S \leq M_{J+1}^* \leq \dots \leq M_{n+1}^*. \quad (6.9)$$

Now if $0 < z < M_{J+1} w_{J+1}$, then

$$\sum_i \max \{0, M_i - z/w_i\} \geq \sum_1^{J+1} (M_i - z/w_i) > M_{J+1}^* \geq S,$$

so that z does not satisfy (6.6). And if $M_J w_J \geq z \geq M_{J+1} w_{J+1}$, then (6.6) becomes

$$\sum_1^J (M_i - z/w_i) \leq S,$$

which is equivalent to

$$z \geq z^* = (\sum_1^J M_i - S) / \sum_1^J (1/w_i). \quad (6.10)$$

By use of (6.9), the value of z^* proposed in (6.10) is easily verified to satisfy $M_J w_J \geq z^* \geq M_{J+1} w_{J+1}$, and so is indeed the smallest $z > 0$ satisfying (6.6).

The determination of z^{**} is similar, involving

$$V_k^{**} = w_k V_k \sum_1^k (1/w_i) + \sum_{k+1}^n V_i,$$

$$V_1^{**} \geq V_2^{**} \geq \dots \geq V_K^{**} \geq S > V_{K+1}^{**} \geq \dots \geq V_{n+1}^{**},$$

$$z^{**} = (S - \sum_{k+1}^n V_i) / \sum_1^k (1/w_i). \quad (6.11)$$

(Paper 72A4-276)