The Annals of Statistics
1979, Vol. 7, No. 4, 729–741

ESTIMATION OF STARSHAPED SEQUENCES OF POISSON AND NORMAL MEANS[1]

BY MOSHE SHAKED

Indiana University

A vector $\boldsymbol{\mu} = (\mu_1, \cdots, \mu_n)$ is said to be upper [lower] starshaped if $\mu_{m+1} \geqslant \bar{\mu}_m \geqslant 0$ $[0 \leqslant \mu_{m+1} \leqslant \bar{\mu}_m]$ $m = 1, \cdots, n-1$, where $\bar{\mu}_m$ is a weighted average of $\mu_1, \cdots, \mu_m$. Obtained is the maximum likelihood estimate of $\boldsymbol{\mu}$ when the μ_i's are the means of n Poisson or normal populations and $\boldsymbol{\mu}$ is known to be starshaped. The method is applied to obtain estimators of IHRA (increasing hazard rate average) distribution functions.

1. Introduction. Let $\boldsymbol{\mu} = (\mu_1, \cdots, \mu_n)$ be a vector of n unknown parameters and let $\mathbf{w} = (w_1, \cdots, w_n)$ be n associated weights. Consider $\boldsymbol{\mu}$ and $\mathbf{w}$ as functions on the set $\{1, 2, \cdots, n\}$. The function (or the vector) $\boldsymbol{\mu}$ is called *isotonic* if $\mu_1 \leqslant \mu_2 \leqslant \cdots \leqslant \mu_n$ and *antitonic*, if $\mu_1 \geqslant \mu_2 \geqslant \cdots \geqslant \mu_n$ (Barlow et al. (1972)). Denote $\bar{\mu}_m = \Sigma_{i=1}^m w_i \mu_i)/\Sigma_{i=1}^m w_i$ and $\bar{\mu}_0 = 0$. We say that $\boldsymbol{\mu}$ is *upper-starshaped* if $\mu_{m+1} \geqslant \bar{\mu}_m$, $m = 0, 1, \cdots, n-1$, and *lower-starshaped* if $0 \leqslant \mu_{m+1} \leqslant \bar{\mu}_m$, $m = 1, 2, \cdots, n-1$. Note that if $\mu_{i+1} \geqslant \mu_i \geqslant 0$ (respectively, $0 \leqslant \mu_{i+1} \leqslant \mu_i$), $i = 1, 2, \cdots, n-1$ then $\boldsymbol{\mu}$ is upper-starshaped (respectively, lower-starshaped).

Starshaped vectors arise in a variety of applications. In reliability theory one can be induced by any IHRA distribution F (Birnbaum, et al. (1966)). Such distribution satisfies

$$t^{-1}R(t) \equiv t^{-1}[-\log(1 - F(t))]\uparrow t \geqslant 0. \tag{1.1}$$

A lifelength of a coherent system has an IHRA distribution if each of its independent components has an IHRA distribution—in particular, if each component has an exponential lifelength. Let $w_i > 0$, $i = 1, \cdots, n$ be n lengths of consecutive time intervals, $w_0 \equiv 0$, and define $\mu_i = w_i^{-1}[R(\Sigma_{j=0}^i w_j) - R(\Sigma_{j=0}^{i-1} w_j)]$, $i = 1, 2, \cdots, n$, then $\boldsymbol{\mu}$ is upper-starshaped. Similarly, if F is a DHRA distribution then $\boldsymbol{\mu}$ is lower-starshaped. The method that is developed in Section 3 is applied in Section 5 to estimate IHRA distributions.

The following paradigm includes a large class of potential applications for the results derived here. Consider a species consisting of k individuals each of which has a quantitative characteristic of interest X. Denote the expected value of X of the population in generation 1 by $\mu_1 \geqslant 0$. Assume that k new individuals are produced in each generation and are added to the existing population and that the characteristic of interest X is 'improving on the average'—that is, let $\mu_2 \geqslant \mu_1$ be the expected value of X for each offspring of the first generation, thus the expected value of X of

Received May 1977; revised May 1978.

[1]This work is prepared under contract 05-0695 from Sandia Laboratories.

AMS 1970 *subject classifications.* Primary 62F10; secondary 60K10.

Key words and phrases. Isotonic regression, IHRA distributions, maximum likelihood estimation, Poisson and normal extremum problems.

the population in generation 2 is $\bar{\mu}_2 = \frac{1}{2}(\mu_1 + \mu_2)$. The expected value of X for each offspring of the second generation is $\mu_3 \geqslant \bar{\mu}_2$ and the expected value of X of the population in generation 3 is $\bar{\mu}_3$. In general assume the expected value of X for each offspring of the ith generation is $\mu_{i+1} \geqslant \bar{\mu}_i$; the expected value of X in the $(i + 1)$st generation is $\bar{\mu}_{i+1}$. Clearly $\boldsymbol{\mu} = (\mu_1, \cdots, \mu_n)$ is upper-starshaped.

In Section 3 an explicit expression for the least square estimators of $\boldsymbol{\mu}$ as a starshaped parameter vector is derived; these are also the maximum likelihood estimators (m.l.e.'s) if the μ_i's are the means of n normal populations. In Section 2 the m.l.e's when the μ_i's are the means of n Poisson populations, are obtained explicitly. These two cases were the only ones we could work out explicitly. In Section 4 we discuss the reasons for our good luck with the normal and the Poisson extremum problem and conjecture a general method of obtaining starshaped m.l.e's. Some applications are discussed in Section 5.

2. The Poisson extremum problem. Let $\mu_1, \cdots, \mu_n$ be the means of n Poisson populations and assume $\boldsymbol{\mu}$ is upper-starshaped with respect to the positive integer weights $w_1, \cdots, w_n$; that is,

$$\boldsymbol{\mu} \in \mathcal{C}_u \equiv \{(\mu_1, \cdots, \mu_n) | \mu_{m+1} \geqslant \bar{\mu}_m, m = 0, 1, \cdots, n-1\}.$$

From a sample $X_{i1}, \cdots, X_{iw_i}$, $i = 1, 2, \cdots, n$ designating the ith population, the m.l.e. $\boldsymbol{\mu}^*$ of $\boldsymbol{\mu}$ is to be obtained. The likelihood function is

$$L(\boldsymbol{\mu}) = (\Pi_{i=1}^n \mu_i^{w_i g_i}) \exp\{-w_i\mu_i\} \Pi_{i=1}^n (\Pi_{j=1}^{w_i} X_{ij}!)^{-1}, \tag{2.1}$$

where $g_i = (\Sigma_{j=1}^{w_i} X_{ij})/w_i$, $i = 1, 2, \cdots, n$. It is easy to see that this problem is equivalent to

$$\text{maximize } \Sigma_{i=1}^n (g_i \log \mu_i - \mu_i) w_i \tag{2.2}$$

subject to $\boldsymbol{\mu} \in \mathcal{C}_u$.

Throughout the rest of the paper the following notations and conventions will be used: $M(m) = \Sigma_{i=1}^m w_i\mu_i$, $W(m) = \Sigma_{i=1}^m w_i$, $\bar{\mu}_m = M(m)/W(m)$, similarly $G(m)$, $\bar{g}_m$, $M^*(m)$ and $\bar{\mu}_m^*$ are defined; $M(0) \equiv \bar{\mu}_0 \equiv G(0) \equiv \bar{g}_0 \equiv M^*(0) \equiv \bar{\mu}_0^* \equiv W(0) = 0$. By convention $0/0 = 1$ and $a/0 = \infty$, $a > 0$.

The next theorem gives an explicit expression for the m.l.e. $\boldsymbol{\mu}^*$. In practice, substitute the observed g's in (2.9) and (2.10) to obtain $\boldsymbol{\mu}^*$. To gain some insight into the nature of the solution of (2.2) note that $\boldsymbol{\mu}^*$ can also be obtained using the following algorithm: if

$$g_{m+1} \geqslant \bar{g}_m, \qquad m = 1, 2, \cdots, n-1 \tag{2.3}$$

then $\mu_m^* = g_m$, $m = 1, 2, \cdots, n$. If (2.3) does not hold then substitute in the object function of (2.2)

$$\mu_{m+1} = \bar{\mu}_m \tag{2.4}$$

for every 'violation' in (2.3)—that is, for every m such that $g_{m+1} < \bar{g}_m$, substitute (2.4) in (2.2) and thus obtain a function of $n - v$ variables (v is the number of violations) and maximize it by equating its partial derivatives to zero. The explicit formulas (2.9) and (2.10) have been thus obtained.

To illustrate, consider the problem (2.2) with $n = 3$ and assume $g_2 \geqslant g_1$ and $g_3 < \bar{g}_2$. According to the above algorithm substitute in the object function of (2.2) $\mu_3 = (w_1\mu_1 + w_2\mu_2)/(w_1 + w_2)$. Now maximize the bivariate function

$$(2.5)\quad (g_1 \log \mu_1 - \mu_1)w_1 + (g_2 \log \mu_2 - \mu_2)w_2 + \left[g_3 \log\left(\frac{w_1\mu_1 + w_2\mu_2}{w_1 + w_2} \right) - \left(\frac{w_1\mu_1 + w_2\mu_2}{w_1 + w_2} \right) \right] w_3.$$

By equating the partial derivatives to zero obtain

$$\mu_1^* = g_1\frac{\bar{g}_3}{\bar{g}_2}, \ \mu_2^* = g_2\frac{\bar{g}_3}{\bar{g}_2}$$

and

$$\mu_3^* = \bar{g}_3$$

in agreement with (2.9) and (2.10).

Before stating Theorem 2.1, a sufficient condition for a vector $\mathbf{u}$ to be a solution of (2.2) is introduced.

PROPOSITION 2.1. *If a vector* $\mathbf{u}$ *satisfies*

$$(2.6)\quad \mathbf{u} \in \mathcal{C}_u,$$

$$(2.7)\quad \Sigma_{m=1}^{n} w_m g_m = \Sigma_{m=1}^{n} w_m u_m$$

and

$$(2.8)\quad \Sigma_{m=1}^{n}(g_m\mu_m/u_m - \mu_m)w_m \leqslant 0 \forall \boldsymbol{\mu} \in \mathcal{C}_u,$$

then $\mathbf{u}$ *is a solution of* (2.2).

PROOF. Since for every $y > 0$, $\log y \leqslant y - 1$ it follows from (2.7) and (2.8) that

$$\Sigma_{i=1}^{n}(g_i \log \mu_i - \mu_i)w_i - \Sigma_{i=1}^{n}(g_i \log u_i - u_i)w_i \leqslant 0 \forall \boldsymbol{\mu} \in \mathcal{C}_u;$$

that is, $\mathbf{u}$ is a solution of (2.2). □

Note that as in Barlow, et al. (1972), page 44, it can be shown that a solution $\mathbf{u}$ for (2.2) must satisfy (2.8). To see this, note that if $\boldsymbol{\mu} \in \mathcal{C}_u$ then also $\rho\boldsymbol{\mu} \in \mathcal{C}_u$ whenever $\rho > 0$. Thus $\Sigma_{i=1}^{k}(g_i \log \rho u_i - \rho u_i)w_i$ achieves its maximum as a function of ρ at $\rho = 1$. On setting the derivative at $\rho = 1$ equal to zero (2.8) is obtained.

THEOREM 2.1. *If* $g_i \geqslant 0$, $i = 1, 2, \cdots, n$, *then a solution of* (2.2) *is*:

$$(2.9)\quad \mu_m^* = \max(g_m, \bar{g}_m)\Pi_{j=m+1}^{n} \min(1, \bar{g}_j/\bar{g}_{j-1}), \qquad m = 1, 2, \cdots, n-1,$$

$$(2.10)\quad \mu_n^* = \max(g_n, \bar{g}_n).$$

The solution is unique.

Proof. The proof is written under the assumption that $g_1 > 0$ (then $\bar{g}_i > 0$, $i = 1, \cdots, n$). Obvious, but tedious modifications are needed when $g_1 = 0$ (that is, $\bar{g}_i = 0$ for some i) and we omit it.

To prove that $\boldsymbol{\mu}^*$ of (2.9) and (2.10) satisfies (2.6) use Identity (A.1) to obtain, for $2 \leq l \leq n-2$,

$$\begin{aligned}\bar{\mu}_l^* &= \Sigma_{m=1}^{l} w_m \mu_m^* / W(l) \\ &=_{(2.9),\,(2.10)} (W(l))^{-1} \\ &\quad \times \left[\Sigma_{m=1}^{l-1} w_m \max(g_m, \bar{g}_m) \Pi_{j=m+1}^{l} \min(1, \bar{g}_j/\bar{g}_{j-1}) + w_l \max(g_l, \bar{g}_l)\right] \\ &\quad \times \Pi_{j=l+1}^{n} \min(1, \bar{g}_j/\bar{g}_{j-1}) \\ &=_{(A.1)} \bar{g}_l \min(1, \bar{g}_{l+1}/\bar{g}_l) \Pi_{j=l+2}^{n} \min(1, \bar{g}_j/\bar{g}_{j-1}).\end{aligned}$$

Similarly, $\bar{\mu}_{l+1}^* = \bar{g}_{l+1} \Pi_{i=l+2}^{n} \min(1, \bar{g}_j/\bar{g}_{j-1})$. Using these expressions it is easily seen that if $\bar{g}_l \leq \bar{g}_{l+1}$ then $\bar{\mu}_l^* \leq \bar{\mu}_{l+1}^*$ and if $\bar{g}_l > \bar{g}_{l+1}$ then $\bar{\mu}_l^* = \bar{\mu}_{l+1}^*$. Thus, for $2 \leq l \leq n-2$

$$\bar{\mu}_{l+1}^* \geq \bar{\mu}_l^*. \tag{2.11}$$

The proofs of (2.11) for $l = 1$ and $l = n - 1$ are similar. Inequalities (2.11) show that $\boldsymbol{\mu}^*$ satisfies (2.6).

That $\boldsymbol{\mu}^*$ satisfies (2.7) follows from Identity (A.1) with $l = n$.

For real numbers a and b define $I(a < b) = 1$ if $a < b$ and $I(a < b) = 0$ if $a \geq b$. To prove that $\boldsymbol{\mu}^*$ satisfies (2.8) use Identity (A.3) with $l = n$ to obtain

$$\begin{aligned}&\Sigma_{m=1}^{n} (g_m \mu_m / \mu_m^* - \mu_m) w_m \\ &\quad =_{(2.9),\,(2.10)} \Sigma_{m=1}^{n-1} \left\{ w_m \mu_m \frac{g_m}{\max(g_m, \bar{g}_m)} \Pi_{j=m+1}^{n} \max\left(1, \frac{\bar{g}_{j-1}}{\bar{g}_j}\right)\right\} \\ &\qquad + w_n \mu_n \frac{g_n}{\max(g_n, \bar{g}_n)} - \Sigma_{m=1}^{n} w_m \mu_m \\ &\quad =_{(A.3)} \Sigma_{m=1}^{n-1} \left\{ I(g_m < \bar{g}_{m-1}) \frac{w_m W(m-1)}{G(m)} \right. \\ &\qquad\qquad \left. \times (\bar{\mu}_{m-1} - \mu_m)(\bar{g}_{m-1} - g_m) \Pi_{j=m+1}^{n} \max\left(1, \frac{\bar{g}_{j-1}}{\bar{g}_j}\right)\right\} \\ &\qquad + I(g_n < \bar{g}_{n-1}) \frac{w_n W(n-1)}{G(n)} (\bar{\mu}_{n-1} - \mu_n)(\bar{g}_{n-1} - g_n) \leq 0.\end{aligned}$$

The last inequality holds since $\bar{\mu}_{m-1} \leq \mu_m$, $m = 1, 2, \cdots, n$. This completes the proof of the first part of the theorem.

To prove uniqueness of the solution let $\boldsymbol{\mu}^*$ and $\boldsymbol{\mu}^{**}$ be two solutions of (2.2). Substitute $\boldsymbol{\mu} = \boldsymbol{\mu}^*$ and $\mathbf{u} = \boldsymbol{\mu}^{**}$ in (A.5) and then $\boldsymbol{\mu} = \boldsymbol{\mu}^{**}$ and $\mathbf{u} = \boldsymbol{\mu}^*$ and add the two resulting inequalities to obtain $\Sigma(\mu_i^* - \mu_i^*)^2(\mu_i^* \mu_i^{**})^{-1} w_i g_i \leq 0$. Thus $\mu_i^* = \mu_i^{**}$, $i = 1, 2, \cdots, n$. □

The next theorem gives the m.l.e. of a lower-starshaped $\boldsymbol{\mu}$. It is clear that the m.l.e. $\boldsymbol{\mu}^*$ is the solution of

$$\text{(2.2')} \qquad \text{maximize } \Sigma_{i=1}^{n}(g_i \log \mu_i - \mu_i)w_i$$

subject to $\boldsymbol{\mu} \in \mathcal{C}_l \equiv \{(\mu_1, \cdots, \mu_n) | \bar{\mu}_m \geqslant \mu_{m+1} \geqslant 0, m = 1, 2, \cdots, n\}$.

THEOREM 2.1'. *If $g_1 > 0$ and $g_i \geqslant 0$, $i = 2, \cdots, n$ then a solution of (2.2') is:*

$$\text{(2.9')} \qquad \mu_m^* = \min(g_m, \bar{g}_m)\Pi_{j=m+1}^{n} \max(1, \bar{g}_j/\bar{g}_{j-1}), \qquad m = 1, 2, \cdots, n-1,$$

$$\text{(2.10')} \qquad \mu_n^* = \min(g_n, \bar{g}_n).$$

The solution is unique.

The proof of this theorem parallels the proof of Theorem 2.1 and is omitted.

A slight modification of formulas (2.9') and (2.10') is needed if $g_1 = 0$. Let $i_0 = \min\{i : g_i > 0\}$, then for $m \geqslant i_0$, μ_m^* is given in (2.9') and (2.10') and $\mu_1^* = \cdots = \mu_{i_0-1}^* = \mu_{i_0}^*$.

3. Least squares estimation (the normal extremum problem). Let μ_i be the mean of a normal population whose known variance is σ_i^2, $i = 1, 2, \cdots, n$. Assume that $\boldsymbol{\mu}$ is upper-starshaped with respect to the weights $\mathbf{w}$ where w_i is a positive integer multiple of σ_i^{-2}, $i = 1, \cdots, n$. From a sample $X_{i1}, \cdots, X_{i\tilde{w}}$ (where $\tilde{w} = w_i\sigma_i^2$), $i = 1, \cdots, n$, designating the ith population, the m.l.e. $\boldsymbol{\mu}^*$ is obtained. Let $g_i = \tilde{w}_i^{-1}\Sigma_{j=1}^{\tilde{w}_i}X_{ij}$, then $\mathbf{g}$ is a sufficient statistic for $\boldsymbol{\mu}$. The likelihood function associated with $\mathbf{g}$ is

$$\text{(3.1)} \qquad L(\boldsymbol{\mu}) = \left\{\Pi_{i=1}^{n}\left[\left(2\pi\sigma_i^2\right)^{-1/2}\tilde{w}_i^{1/2}\right]\right\} \exp\left[-\frac{1}{2}\Sigma_{i=1}^{n}(g_i - \mu_i)^2 w_i\right].$$

Clearly $\boldsymbol{\mu}^*$ is a solution of

$$\text{(3.2)} \qquad \text{minimize } \Sigma_{i=1}^{n}(g_i - \mu_i)^2 w_i$$

subject to $\boldsymbol{\mu} \in \mathcal{C}_u$ (the notations of the preceding section are used here, too). Theorem 3.1 gives the unique solution of (3.2) when

$$\text{(3.3)} \qquad \bar{g}_i \geqslant 0, \qquad i = 1, 2, \cdots, n;$$

thus, if, in a particular experiment, some of the g_i's are negative such that (3.3) does not hold then (3.9) and (3.10) do not give $\boldsymbol{\mu}^*$. However, in many applications, distributions of nonnegative random variables are approximated by normal distributions—then $g_i \geqslant 0$, $i = 1, 2, \cdots, n$, hence (3.3) holds and (3.9) and (3.10) apply. Further applications of Theorem 3.1 are given in Section 5.

In practice $\boldsymbol{\mu}^*$ is found by substituting the observed g's in (3.7) and (3.8). To throw some light on the nature of $\boldsymbol{\mu}^*$ note that $\boldsymbol{\mu}^*$ can be obtained using an algorithm similar to the one described in Section 2: if

$$\text{(3.4)} \qquad g_{m+1} \geqslant \bar{g}_m, \qquad m = 1, 2, \cdots, n-1$$

then $\mu_m^* = g_m$, $m = 1, 2, \cdots, n$. If (3.4) fails, substitute in the object function of

(3.2) $\mu_{m+1} = \bar{\mu}_m$ for every m that violates (3.4) and get a function of $n - v$ variables (v is the number of violations). This function is maximized by equating its partial derivatives to zero. Here the computations are neater than those in Section 2 since here one gets a set of $n - v$ linear equations.

To illustrate, consider the problem (3.2) with $n = 4$ and assume $g_2 < g_1$, $g_3 \geqslant \bar{g}_2$ and $g_4 < \bar{g}_3$. According to the above algorithm substitute in the object function of (3.2) $\mu_1 = \mu_2$ and $\mu_4 = (w_1\mu_1 + w_2\mu_2 + w_3\mu_3)/(w_1 + w_2 + w_3) = [(w_1 + w_2)\mu_1 + w_3\mu_3]/(w_1 + w_2 + w_3)$. Now maximize the bivariate function

$$(3.5)\quad w_1(g_1 - \mu_1)^2 + w_2(g_2 - \mu_1)^2 + w_3(g_3 - \mu_3)^2 + w_4\{g_4 - [(w_1 + w_2)\mu_1 + w_3\mu_3]/(w_1 + w_2 + w_3)\}^2.$$

By equating the partial derivatives to zero obtain

$$\mu_3^* = g_3 - w_4(W(4))^{-1}(\bar{g}_3 - g_4),$$
$$\mu_1^* = g_1 - w_2(W(2))^{-1}(g_1 - g_2) - w_4(W(4))^{-1}(\bar{g}_3 - g_4),$$

and

$$\mu_2^* = \mu_1^*,$$
$$\mu_4^* = \bar{g}_4,$$

in agreement with (3.9) and (3.10).

Before stating Theorem 3.1 a necessary and sufficient condition for a vector $\mathbf{u}$ to be a solution of (3.2) is given.

PROPOSITION 3.1. *A vector* $\mathbf{u}$ *is the unique solution of* (3.2) *if and only if*

$$(3.6)\qquad \mathbf{u} \in \mathcal{C}_u,$$

$$(3.7)\qquad \Sigma_{i=1}^n(g_i - u_i)u_iw_i = 0$$

and

$$(3.8)\qquad \Sigma_{i=1}^n(g_i - u_i)\mu_iw_i \leqslant 0 \forall \boldsymbol{\mu} \in \mathcal{C}_u.$$

PROOF. Since $\mathcal{C}_u$ is a convex cone the result follows from Theorems 1.3–1.5 of Barlow, et al. (1972). □

For every real a denote $a^+ \equiv \max(0, a)$ and $a^- \equiv \min(0, a) = (-a)^+$.

THEOREM 3.1. *Assume that* (3.3) *holds, then a solution of* (3.2) *is*:

$$(3.9)\qquad \mu_m^* = \max(g_m, \bar{g}_m) - \Sigma_{j=m+1}^n w_j(W(j))^{-1}(\bar{g}_{j-1} - g_j)^+, \quad m = 1, 2, \cdots, n - 1,$$

$$(3.10)\qquad \mu_n^* = \max(g_n, \bar{g}_n).$$

The solution is unique.

PROOF. First note that when either $g_m \geqslant \bar{g}_m$ or $g_m \leqslant \bar{g}_m$,

(3.11) $$\max(g_m, \bar{g}_m) - W(m-1)(W(m))^{-1}(\bar{g}_{m-1} - g_m)^+ = g_m, \quad m = 1, 2, \cdots, n.$$

Thus for $m \leqslant n-2$

$$\begin{aligned}\Sigma_{i=1}^m w_i \mu_i^* &= \Sigma_{i=1}^m w_i \max(g_i, \bar{g}_i) - \Sigma_{i=1}^m w_i \Sigma_{j=i+1}^{m+1}(W(j))^{-1} w_j(\bar{g}_{j-1} - g_j)^+ \\ &\quad - \Sigma_{i=1}^m w_i \Sigma_{j=m+2}^n (W(j))^{-1} w_j(\bar{g}_{j-1} - g_j)^+ \\ &= \Sigma_{i=1}^m w_i \max(g_i, \bar{g}_i) - \Sigma_{j=1}^{m+1} W(j-1) w_j (W(j))^{-1}(\bar{g}_{j-1} - g_j)^+ \\ &\quad - \Sigma_{j=m+2}^n W(m) W(j)^{-1} w_j(\bar{g}_{j-1} - g_j)^+ \\ &= \Sigma_{i=1}^m w_i\left[\max(g_i, \bar{g}_i) - W(i-1)(W(i))^{-1}(\bar{g}_{i-1} - g_i)^+\right] \\ &\quad - W(m)\Sigma_{j=m+1}^n w_j (W(j))^{-1}(\bar{g}_{j-1} - g_j)^+ .\end{aligned}$$

Using (3.11),

$$\bar{\mu}_m^* = \bar{g}_m - w_{m+1}(W(m+1))^{-1}(\bar{g}_m - g_{m+1})^+ - \Sigma_{j=m+2}^n w_j (W(j))^{-1}(\bar{g}_{j-1} - g_j)^+ .$$

Using (3.9),

(3.12)

$$\mu_{m+1}^* - \bar{\mu}_m^* = \max(g_{m+1}, \bar{g}_{m+1}) - \bar{g}_m + w_{m+1}(W(m+1))^{-1}(\bar{g}_m - g_{m+1})^+ .$$

If $g_{m+1} \geqslant \bar{g}_{m+1}$ it is clear that $\mu_{m+1}^* - \bar{\mu}_m^* \geqslant 0$. If $g_{m+1} < \bar{g}_{m+1}$ then $\mu_{m+1}^* = \bar{\mu}_m^*$. That $\mu_n^* \geqslant \bar{\mu}_{n-1}^*$ is shown similarly, and the proof that $\boldsymbol{\mu}^*$ satisfies (3.6) is complete.

Using Identity (A.7) note that

$$\begin{aligned}&\Sigma_{m=1}^n (g_m - \mu_m^*)\mu_m w_m \\ &\quad = \Sigma_{m=1}^{n-1} w_m \mu_m\left[(g_m - \bar{g}_m)^- + \Sigma_{j=m+1}^n w_j (W(j))^{-1}(\bar{g}_{j-1} - g_j)^+\right] \\ (3.13) \qquad &\qquad + w_n \mu_n (g_n - \bar{g}_n)^- \\ &\quad =_{(A.7)} \Sigma_{m=1}^n I(\bar{g}_{m-1} > g_m) w_m W(m-1) \\ &\qquad (W(m))^{-1}(\bar{g}_{m-1} - g_m)(\bar{\mu}_{m-1} - \mu_m) \leqslant 0\end{aligned}$$

where the last inequality follows from the restriction on $\boldsymbol{\mu}$. Thus $\boldsymbol{\mu}^*$ satisfies (3.8).

To show that $\boldsymbol{\mu}^*$ satisfies (3.7) note that, if $\bar{g}_{m-1} > g_m$ then from (3.12)

(3.14) $$\bar{\mu}_{m-1}^* - \mu_m^* = \bar{g}_{m-1} - w_m (W(m))^{-1}(\bar{g}_{m-1} - g_m) - \bar{g}_m = 0.$$

Substitute $\boldsymbol{\mu}^*$ for $\boldsymbol{\mu}$ in (3.13) and use (3.14) to obtain that $\boldsymbol{\mu}^*$ satisfies (3.7). □

If it is known that $\boldsymbol{\mu}$ is lower-starshaped then the m.l.e. is the $\boldsymbol{\mu}^*$ that

(3.2′) $$\text{minimize } \Sigma_{i=1}^n (g_i - \mu_i)^2 w_i$$

subject to $\boldsymbol{\mu} \in \mathcal{C}_l$. The next theorem, the proof of which is similar to the proof of Theorem 3.1 gives $\boldsymbol{\mu}^*$.

THEOREM 3.1′. *Assume $g_i \geqslant 0$, $i = 1, 2, \cdots, n$, then a solution of* (3.2′) *is*:

$$\mu_m^* = \min(g_m, \bar{g}_m) - \Sigma_{j=m+1}^n w_j (W(j))^{-1} (\bar{g}_{j-1} - g_j)^-, \quad m = 1, 2, \cdots, n-1, \tag{3.7′}$$

$$\mu_n^* = \min(g_n, \bar{g}_n). \tag{3.8′}$$

The solution is unique.

4. Discussion. Both solutions given in Theorems 2.1 and 3.1 can be obtained by similar algorithms that were described above. According to these algorithms, one substitutes $\mu_{m+1} = \bar{\mu}_m$ in the objective function whenever $g_{m+1} < \bar{g}_m$. One then has a function of $m - v$ variables and one gets the desired solution $\boldsymbol{\mu}^*$ by equating partial derivatives to zero.

Maximization of expressions such as (2.5) and (3.5) usually reduces to solving $n - v$ equations with $n - v$ unknown. It is fortunate that in the two cases which are discussed in this paper the solution $\boldsymbol{\mu}^*$ can be obtained explicitly.

Once the explicit expressions for the μ_i^*'s are given one can proceed as in the proofs of Theorems 2.1 and 3.1 to show that the μ_i^*'s are indeed the maximizing constants. However, in problems such as (4.1) and (4.2) below we have not been able to obtain explicit expressions for the constants that maximize functions such as (2.5) and (3.5). Thus, we have not been able to show that an algorithm such as described above applies to problems such as the "geometric extremum problem":

$$\text{maximize } \Sigma_{i=1}^n [g_i \log \mu_i - (1 + g_i) \log(1 + \mu_i)] w_i, \tag{4.1}$$

subject to $\boldsymbol{\mu} \in \mathcal{C}_u$. And the "gamma extremum problem":

$$\text{minimize } \Sigma_{i=1}^n [\log \mu_i + g_i/\mu_i] w_i, \tag{4.2}$$

subject to $\boldsymbol{\mu} \in \mathcal{C}_u$.

We conjecture, however, that such an algorithm applies to (4.1) and (4.2), too.

5. A numerical example and an application. Clevenson and Zidek (1975) give the 36 numbers of oilwell discoveries in Alberta for the 3rd month of each half year from 1953 to 1970 (Tables 5.1 and 5.2). They assume that each of these random variables has a Poisson distribution. The accumulated experience in oilwell discoveries throughout the years may be expressed by the assumption that the parameters vector $\boldsymbol{\lambda} = (\lambda_1, \cdots, \lambda_{36})$ is upper-starshaped (this assumption is weaker than the possible alternative assumption that $\lambda_1 \leqslant \lambda_2 \leqslant \cdots \leqslant \lambda_{36}$). The m.l.e. $\boldsymbol{\lambda}^*$, computed from (2.9) and (2.10), is compared in Table 5.1 with the unrestricted m.l.e. $\mathbf{X}$. To be able to compare the performance of the two estimates using the loss function $\Sigma_{i=1}^{36}(\hat{\lambda}_i - \lambda_i)^2$, where $\hat{\boldsymbol{\lambda}}$ is some estimate of $\boldsymbol{\lambda}$, we need to know the actual values λ_i, $i = 1, \cdots, 36$. Clevenson and Zidek (1975) give a vector of "actual values" which is not upper-starshaped. Transforming it to an upper-starshaped vector using (2.9) and (2.10) with $g_i = \lambda_i$ (the λ_i of Clevenson and Zidek (1975)) and $w_i = 1$, $i = 1, \cdots, 36$ obtain the "actual values" λ_i's in Table 5.1. It is seen that the loss incurred using $\boldsymbol{\lambda}^*$ is much smaller than the loss using $\mathbf{X}$.

TABLE 5.1
Comparison of X and an upper-starshaped $\boldsymbol{\lambda}^*$ *as estimators of population monthly means for* 36 *selected months.*

i	X_i	λ_i	λ_i^*	$(X_i - \lambda_i)^2$	$(\lambda_i^* - \lambda_i)^2$
1	0	.49	0	.2401	.2401
2	0	.49	0	.2401	.2401
3	0	.49	0	.2401	.2401
4	1	.58	.36	.1764	.0484
5	2	.51	.72	2.2201	.0441
6	1	.51	.36	.2401	.0225
7	0	.69	.24	.4761	.2025
8	2	.54	.84	2.1316	.0900
9	0	.54	.32	.2916	.0484
10	0	.54	.32	.2916	.0484
11	0	.54	.32	.2916	.0484
12	1	.54	.58	.2116	.0016
13	3	1.16	1.74	3.3856	.3364
14	0	.58	.45	.3364	.0169
15	0	.93	.45	.8649	.2304
16	3	1.05	2.00	3.8025	.9025
17	0	.66	.54	.4356	.0144
18	2	.75	1.42	1.1025	.2209
19	1	.65	.71	.1225	.0036
20	2	.65	1.42	1.8225	.5929
21	0	1.12	.64	1.2544	.2304
22	0	.70	.64	.4900	.0036
23	0	.68	.64	.4624	.0016
24	1	1.30	.82	.0900	.2304
25	5	1.15	4.08	14.8225	8.5849
26	0	.72	.78	.5184	.0036
27	1	.72	.85	.0784	.0169
28	0	.72	.79	.5184	.0049
29	0	.72	.79	.5184	.0049
30	1	.72	.91	.0784	.0361
31	0	.72	.79	.5184	.0049
32	1	.78	.94	.0484	.0256
33	0	.72	.79	.5184	.0049
34	1	.72	.97	.0784	.0625
35	0	.72	.80	.5184	.0064
36	1	.72	1.00	.0784	.0784
Total	29	26.02	29.02	39.5152	12.8426

The columns $\boldsymbol{\lambda}$ and $\boldsymbol{\lambda}^*$ of Table 5.1 are typical numerical examples of Theorem 2.1. To obtain numerical examples of Theorem 2.1′, Table 5.2 is given. This table is constructed as Table 5.1 under the assumption that $\boldsymbol{\lambda}$ is lower-starshaped. As noted after Theorem 2.1′ here $\lambda_4^*, \cdots, \lambda_{36}^*$ are computed from (2.9′) and (2.10′) and $\lambda_1^* = \lambda_2^* = \lambda_3^* = \lambda_4^*$. Also in this example, the loss using $\boldsymbol{\lambda}^*$ is smaller than the loss using $\mathbf{X}$.

TABLE 5.2
Comparison of X and a lower-starshaped λ^* as estimators of population monthly means for 36 selected months.

i	X_i	λ_i	λ_i^*	$(X_i - \lambda_i)^2$	$(\lambda_i^* - \lambda_i)^2$
1	0	1.74	2.23	3.0276	.2041
2	0	1.23	2.23	1.5129	1.0000
3	0	.74	2.23	.5476	2.2201
4	1	1.24	2.23	.0576	.9801
5	2	1.18	2.23	.6724	1.1025
6	1	1.18	2.23	.0324	1.1025
7	0	1.22	0	1.4884	1.4884
8	2	1.12	1.91	.7744	.6241
9	0	.90	0	.8100	.8100
10	0	.23	0	.0529	.0529
11	0	0	0	0	0
12	1	.44	1.39	.3136	.9025
13	3	.93	1.39	4.2849	.2116
14	0	.62	0	.3844	.3844
15	0	.91	0	.8281	.8281
16	3	.91	1.21	4.3681	.0900
17	0	.57	0	.3249	.3249
18	2	.89	1.14	1.2321	.0625
19	1	.55	1.14	.2025	.3481
20	2	.55	1.14	2.1025	.3481
21	0	.86	0	.7396	.7396
22	0	.86	0	.7396	.7396
23	0	.35	0	.1225	.1225
24	1	.84	.99	.0256	.0225
25	5	.84	.99	17.3056	.0225
26	0	.67	0	.4489	.4489
27	1	.67	.95	.1089	.0324
28	0	.33	0	.1089	.1089
29	0	.33	0	.1089	.1089
30	1	.33	.88	.4489	.3025
31	0	.50	0	.2500	.2500
32	1	.76	.86	.0576	.0100
33	0	.67	0	.4489	.4489
34	1	.33	.83	.4489	.2500
35	0	0	0	0	0
36	1	.50	.81	.2500	.0961
Total	29	25.99	29.01	44.6301	16.8242

5.2. Estimation of IHRA distribution functions. Assume that the lifelength of a device has an IHRA distribution F (see Section 1). Suppose that, due to logistical reasons, samples of n_i devices, $i = 1, 2, \cdots, n$ are put on a life test simultaneously at n different locations. Suppose further that there is only one instrument (or person) that can detect failures and this instrument is used at time T_i to determine the number, r_i, of devices that are still functioning at time T_i in the ith location, $i = 1, 2, \cdots, n$, (this is the situation, e.g., if the information concerning failures of the devices is 'top secret' and only one person has the necessary clearance). It is desired to estimate $F(T_1), \cdots, F(T_n)$ subject to restriction (1.1). It is complicated in this case to obtain the m.l.e.'s. A naive method is to estimate $F(T_i)$ by $(n_i - r_i)/n_i$, but these estimates may give rise to an F which is not IHRA (in fact, if, for some i, $r_i/n_i > r_{i+1}/n_{i+1}$ then the estimates do not give even a distribution function).

Estimates of $F(T_i)$, $i = 1, 2, \cdots, n$ can be obtained by minimizing some sum of squares, subject to restriction (1.1). Denote $h_i = -\log(r_i/n_i)$, $w_i = T_i - T_{i-1}$, and let $g_i = w_i^{-1}(h_i - h_{i-1})$, $i = 1, 2, \cdots, n$ ($h_0 = T_0 \equiv 0$), then g_i is the naive estimate of $\mu_i \equiv w_i^{-1}(R(T_i) - R(T_{i-1}))$. From (1.1) it follows that ($\mu_0 \equiv 0$)

$$\mu_i \geqslant \bar{\mu}_{i-1}, \qquad i = 1, 2, \cdots, n. \tag{5.1}$$

The estimate $\boldsymbol{\mu}^*$ of $\boldsymbol{\mu}$ can be obtained now from (3.9) and (3.10) and the estimate of $F(T_i)$ is

$$F^*(T_i) = 1 - \exp\{-\Sigma_{m=1}^{i} w_m \mu_m^*\}, \qquad i = 1, 2, \cdots, n.$$

Note that it is possible that some of the g_i's are negative. However, for all i, $\bar{g}_i = h_i/W(i) \geqslant 0$, and thus Theorem 3.1 applies.

APPENDIX

The Appendix consists of statements and proofs of three identities and an inequality which are used in the proofs of Theorem 2.1 and 3.1.

Identity A.1. If $g_1 > 0$ then

$$\Sigma_{m=1}^{l-1} w_m \max(g_m, \bar{g}_m) \Pi_{j=m+1}^{l} \min(1, \bar{g}_j/\bar{g}_{j-1}) + w_l \max(g_l, \bar{g}_l) = G(l), \qquad l = 2, 3, \cdots, n. \tag{A.1}$$

PROOF. The proof is by induction. If $l = 2$ then the left-hand side of (A.1) $= w_1 \bar{g}_2 + w_2 \bar{g}_2 = G(2)$ if $g_2 < g_1$ and the left-hand side of (A.1) $= w_1 g_1 + w_2 g_2 = G(2)$ if $g_2 \geqslant g_1$. Assume now that for some l ($2 < l \leqslant n - 1$)

$$\Sigma_{m=1}^{l-2} w_m \max(g_m, \bar{g}_m) \Pi_{j=m+1}^{l-1} \min(1, \bar{g}_j/\bar{g}_{j-1}) + w_{l-1} \max(g_{l-1}, \bar{g}_{l-1}) = G(l-1); \tag{A.2}$$

then

$$\begin{aligned} \text{LHS(A.1)} &= (\text{LHS(A.2)}) \min(1, \bar{g}_l, \bar{g}_{l-1}) + w_l \max(g_l, \bar{g}_l) \\ &=_{(\text{A.2})} G(l-1) \min(1, \bar{g}_l, \bar{g}_{l-1}) + w_l \max(g_l, \bar{g}_l) \\ &= G(l-1) + w_l g_l = G(l) \qquad \text{if} \quad \bar{g}_l \geqslant \bar{g}_{l-1} \\ &= G(l-1)\bar{g}_l/\bar{g}_{l-1} + w_l \bar{g}_l = G(l) \qquad \text{if} \quad \bar{g}_l < \bar{g}_{l-1} \end{aligned}$$

and the proof of (A.1) is complete. □

Identity A.2. If $g_1 > 0$ and $\boldsymbol{\mu} \in \mathcal{C}_u$ then

$$\Sigma_{m=1}^{l-1}\left\{w_m\mu_m\left(\frac{g_m}{\max(g_m, \bar{g}_m)}\right)\Pi_{j=m+1}^{l}\max(1, \bar{g}_{j-1}/\bar{g}_j)\right\} + w_l\mu_l\left(\frac{g_l}{\max(g_l, \bar{g}_l)}\right) - \Sigma_{m=1}^{l}w_m\mu_m$$

$$\text{(A.3)} \quad = \Sigma_{m=1}^{l-1}I(g_m < \bar{g}_{m-1})\frac{w_m W(m-1)}{G(m)} \times(\bar{\mu}_{m-1} - \mu_m)(\bar{g}_{m-1} - g_m)\Pi_{j=m+1}^{l}\max(1, \bar{g}_{j-1}/\bar{g}_j) + I(g_l < \bar{g}_{l-1})\frac{w_l W(l-1)}{G(l)}(\bar{\mu}_{l-1} - \mu_l)(\bar{g}_{l-1} - g_l), \qquad l = 2, \cdots, n.$$

PROOF. The proof is by induction. When $l = 2$ then the left-hand side of (A.3) = 0 if $g_2 \geqslant g_1$ and the left-hand side of (A.3) = $w_1 w_2(G(2))^{-1}(g_1 - g_2)(\mu_1 - \mu_2)$ if $g_2 < g_1$. Assume now that for some l $(2 < l \leqslant n - 1)$

$$\text{(A.4)} \quad \Sigma_{m=1}^{l-2}w_m\mu_m\left(\frac{g_m}{\max(g_m, \bar{g}_m)}\right)\Pi_{j=m+1}^{l-1}\max(1, \bar{g}_{j-1}/\bar{g}_j) + w_{l-1}\mu_{l-1}\left(\frac{g_{l-1}}{\max(g_{l-1}, \bar{g}_{l-1})}\right) - \Sigma_{m=1}^{l-1}w_n\mu_n$$

$$= \Sigma_{m=1}^{l-2}\left\{I(g_m < \bar{g}_{m-1})\frac{w_m W(m-1)}{G(m)}(\bar{\mu}_{m-1} - \mu_m) \times(\bar{g}_{m-1} - g_m)\Pi_{j=m+1}^{l-1}\max(1, \bar{g}_{j-1}/\bar{g}_j)\right\} + I(g_{l-1} < \bar{g}_{l-2})\frac{w_{l-1}W(l-2)}{G(l-1)}(\bar{\mu}_{l-2} - \mu_{l-1})(\bar{g}_{l-2} - g_{l-1}).$$

Thus, if $\bar{g}_l \geqslant \bar{g}_{l-1}$ then the left-hand side of (A.3) = the left-hand side of (A.4) and the right-hand side of (A.3) = the right-hand side of (A.4) and if $\bar{g}_l < \bar{g}_{l-1}$ then

$$\text{LHS(A.3)} = [\text{LHS(A.4)}]\frac{\bar{g}_{l-1}}{\bar{g}_l} + \frac{\bar{g}_{l-1}}{\bar{g}_l}\Sigma_{m=1}^{l-1}w_m\mu_m - \Sigma_{m=1}^{l-1}w_m\mu_m + \frac{w_l g_l \mu_l}{\bar{g}_l} - w_l\mu_l$$

$$= [\text{RHS(A.4)}]\frac{\bar{g}_{l-1}}{\bar{g}_l} + \frac{w_l W(l-1)}{G(l)}(\bar{\mu}_{l-1} - \mu_l)(\bar{g}_{l-1} - g_l) = \text{RHS(A.3)},$$

and the proof of (A.3) is complete. □

Inequality A.1. If $\mathbf{u}$ is a solution of (2.2) then

$$\text{(A.5)} \qquad \Sigma_{i=1}^{n}w_i(\mu_i - u_i)(g_i u_i^{-1} - 1) \leqslant 0 \forall \boldsymbol{\mu} \in \mathcal{C}_u.$$

PROOF. An argument of Barlow, et al. (1972), page 25 is used. Let $\mathbf{u}$ be a solution of (2.2) and $\boldsymbol{\mu} \in \mathcal{C}_u$. If $0 \leqslant \alpha \leqslant 1$ then $(1-\alpha)\mathbf{u} + \alpha\boldsymbol{\mu} \in \mathcal{C}_u$. Therefore

$$\text{(A.6)} \qquad \Sigma_{i=1}^{n}\{g_i \log[(1-\alpha)u_i + \alpha\mu_i] - [(1-\alpha)u_i + \alpha\mu_i]\}w_i$$

assumes its maximum at $\alpha = 0$. Thus, the derivative of (A.6) at $\alpha = 0$ is nonpositive, that is (A.5) holds. □

Identity A.3. If (3.3) holds and $\boldsymbol{\mu} \in \mathcal{C}_u$ then

$$\Sigma_{m=1}^{l-1} w_m \mu_m\Big[(g_m - \bar{g}_m)^- + \Sigma_{j=m+1}^{l} w_j(W(j))^{-1}(\bar{g}_{j-1} - g_j)^+\Big] + w_l\mu_l(g_l - \bar{g}_l)^-$$

$$= \Sigma_{m=1}^{l} I(\bar{g}_{m-1} > g_m) w_m W(m-1)(W(m))^{-1}(\bar{g}_{m-1} - g_m)(\bar{\mu}_{m-1} - \mu_m), \quad l = 2, 3, \cdots. \tag{A.7}$$

PROOF. The proof is by induction. When $l = 2$, then the left-hand side of (A.7) $= 0$ if $g_2 \geqslant g_1$, and the left-hand side of (A.7) $= w_1 w_2(W(2))^{-1}(g_1 - g_2)(\mu_1 - \mu_2)$ if $g_2 < g_1$. Assume now that for $l > 2$ ($l \leqslant n$)

$$\begin{aligned} &\Sigma_{m=1}^{l-2} w_m \mu_m\Big[(g_m - \bar{g}_m)^- + \Sigma_{j=m+1}^{l-1} w_j(W(j))^{-1}(\bar{g}_{j-1} - g_j)^+\Big] \\ &\qquad + w_{l-1}\mu_{l-1}(g_{l-1} - \bar{g}_{l-1})^- \\ &= \Sigma_{m=1}^{l-1} I(\bar{g}_{m-1} > g_m) w_m W(m-1)(W(m))^{-1} \\ &\qquad \times (\bar{g}_{m-1} - g_m)(\bar{\mu}_{m-1} - \mu_m). \end{aligned} \tag{A.8}$$

Then

$$\begin{aligned} \text{LHS(A.7)} &= \text{LHS(A.8)} + w_l\mu_l(g_l - \bar{g}_l)^- + (\Sigma_{m=1}^{l-1} w_m\mu_m) w_l(W(l))^{-1}(\bar{g}_{l-1} - g_l)^+ \\ &= \text{RHS(A.8)} + 0 \text{ if } g_l \geqslant \bar{g}_{l-1} \\ &= \text{RHS(A.8)} + w_l W(l-1)(W(l))^{-1}(\bar{g}_{l-1} - g_l)(\bar{\mu}_{l-1} - \mu_l) \text{ if } g_l < \bar{g}_{l-1} \\ &= \text{RHS(A.7)}. \end{aligned}$$

□

Acknowledgment. I would like to thank William Zimmer, who has pointed out to me the application which is discussed in Section 5.2, and a referee for very useful comments.

REFERENCES

[1] BARLOW, R. E., BARTHOLOMEW, D. J., BREMNER, J. M. and BRUNK, H. D. (1972). *Statistical Inference Under Order Restrictions*. Wiley, New York.

[2] BIRNBAUM, Z. W., ESARY, J. D. and MARSHALL, A. W. (1966). Stochastic characterization of wear-out for components and systems. *Ann. Math. Statist.* **37** 816–825.

[3] CLEVENSON, M. L. and ZIDEK, J. V. (1975). Simultaneous estimation of the means of independent Poisson laws. *J. Amer. Statist. Assoc.* **70** 698–705.

DEPARTMENT OF MATHEMATICS
INDIANA UNIVERSITY
BLOOMINGTON, INDIANA 47401