

AD-763 821

ON ESTIMATION OF FAILURE RATES FOR
CERTAIN CLASSES OF DISTRIBUTIONS

Sam C. Saunders

Washington State University

Prepared for:

Air Force Office of Scientific Research

May 1973

DISTRIBUTED BY:

NTIS

**National Technical Information Service
U. S. DEPARTMENT OF COMMERCE
5285 Port Royal Road, Springfield Va. 22151**

AD 763821

On Estimation of Failure Rates for Certain
Classes of Distributions

by

Sam C. Saunders

Department of Pure and Applied Mathematics
Washington State University
Pullman, Washington

Contract AFOSR-72-2281

for

Air Force Office of Scientific Research

May, 1973

DDC
RECEIVED
JUL 26 1973
RECEIVED
B

Approved for public release;
distribution unlimited.

§1. Introduction

In this article we shall consider random variables (r.v.'s) which take values on the closed interval $[0, \infty]$. We shall regard these as the service life of some mechanism according to a predetermined criterion or the time of death of a biological entity and we shall call these r.v.s *life lengths*.

Letting X denote such a random variable with distribution function F . We shall denote the survival probability (reliability) by \bar{F} where

$$P[X > t] = \bar{F}(t) = 1 - F(t) \quad \text{for } t \geq 0.$$

We shall classify such a r.v. by the behavior of its *failure rate* (or hazard rate) which is defined when a density $f = F'$ exists by

$$q = f/\bar{F}$$

or by letting $Q(t) = \int_0^t q(x) dx$ we have the well known

$$f = q e^{-Q} \quad \bar{F} = e^{-Q}$$

We shall always assume that $q \in \mathcal{L}$ where \mathcal{L} is some compact space of non-negative integrable functions and we address ourselves to the problem of estimating q from observations on the life length.

When one must make predictions about life or death as in reliability studies or replacement policies one is naturally interested in classifying the set of random variables from which the observations are obtained. This classification can often be done more naturally by the behavior of their failure rate rather than the density. Consequently there can be practical interest in estimating it.

In actuarial studies the estimation of failure rates, there called the force of mortality, has been done for years because of the interest in calculating the probability of death given survival to a fixed time. The famous

paper of Gompertz [6] in 1825 proposed an exponential form for the failure rate. This was modified in 1860 by Makeham through the addition of another parameter. Extensive applications of the Gompertz-Makeham hazard rate to human data have continued to be made, see Cramér, Wold [5]. More recently studies by Jones [8] and Taylor [10] on the mechanism of aging in humans was carried out by comparisons of the failure rate. Thus the use of failure rate as a means of classifying data has a long and continuing history. Moreover recent studies which classify distributions according to their failure rates have been shown to have interesting properties. In this regard we mention the work of Barlow, Marshall and Proschan who have shown e.g. that monotone failure rates generalize some of the properties of Polya frequency functions, see [2] and [5].

Here we consider the problems of estimating the failure rate for more general classes than parametric, examples of which are Gompertz-Makeham and Weibull, and which have not been previously studied.

Some of the useful classifications of failure rate, well known by their acronyms, are the IFR (increasing failure rate) and IFRA (increasing failure rate average). Following the notation previously introduced in [10] we say, for any positive functions q and r defined as $[0, \infty)$,

$$(1.1) \quad q \in S_r \text{ iff } q/r \text{ is increasing}$$

$$(1.2) \quad q \in C_r \text{ iff } q/r \text{ is convex, increasing and}$$

$$\lim_{x \rightarrow 0} \frac{q(x)}{r(x)} = 0.$$

In the special case where $r(x) = x^K$ for some $K \geq 0$ we shall merely write S_K and C_K . Thus we have

$$Q \in C_0 \text{ iff } q \text{ is IFR}$$

$$Q \in S_1 \text{ iff } q \text{ is IFRA.}$$

Other cases may arise when the failure rate decreases and this may occur in a variety of practical situations. We thus introduce the notation

Ω_0 = set of all non-negative integrable functions of such that

$$\int_0^{\infty} q(x)dx = \infty.$$

$$\Omega_1 = \{q \in \Omega_0 : q \in \uparrow\} = S_0$$

$$\Omega_2 = \{q \in \Omega_0 : q \text{ is convex}\}$$

Other subsets may be of interest in certain applications.

2. The Likelihood Function

Let an ordered sample $x_1 < x_2 < \dots < x_n$ of life lengths with failure rate q be given. The joint density of the sample is,

$$\prod_{i=1}^n q(x_i) e^{-Q(x_i)}.$$

Taking logarithms and setting $\xi = (x_1, \dots, x_n)$ we obtain

$$(2.1) \quad L^*(q|\xi) = \sum_{i=1}^n [\ln q(x_i) - Q(x_i)]$$

which is called the likelihood function. In the usual sense we want to maximize this with respect to $q \in \Omega_k$ (some subset of Ω_0).

It can be seen that the right hand side of (2.1) can be made arbitrarily large by taking $q(x_n)$ large but choosing q so that $Q(x_n)$ remains finite. Thus no maximum likelihood estimate of q exists, in the usual sense, over Ω_k .

Alternatively we consider the estimate \hat{q} , defined by

$$(2.2) \quad \hat{q}(x) = \infty \quad \text{for } x \geq x_n$$

and \hat{q} defined as $(0, x_n)$, from the almost likelihood equation

$$(2.3) \quad l(q|\xi) = \sum_{i=1}^{n-1} \ln q(x_i) - \sum_{i=1}^n Q(x_i),$$

by the inequality

$$L(\hat{q}|\xi) \geq L(q|\xi) \quad \text{for all } q \in \Omega_k.$$

We shall call \hat{q} so defined the *almost maximum likelihood estimate* (AMLE).

Remark 1: For a given sample ξ , $L(\cdot|\xi)$ is a concave function over the convex set Ω_0 and is almost surely continuous in the weak topology.

Proof: That Ω_0 is a convex set is obvious. Note that the second term in the definition of $L(\cdot|\xi)$ is linear. Thus to show for $t \in (0,1)$

$$L(tq_1 + (1-t)q_2|\xi) \geq tL(q_1|\xi) + (1-t)L(q_2|\xi)$$

it is sufficient to show that

$$\sum_1^{n-1} \ln[tq_1(x_i) + (1-t)q_2(x_i)] \geq t \sum_1^{n-1} \ln q_1(x_i) + (1-t) \sum_1^{n-1} \ln q_2(x_i).$$

But this follows since the logarithm is a concave function. Let $q_n \rightarrow q$ i.e., this implies $Q_n \rightarrow Q$ and since the logarithm is continuous, unless some component of ξ falls in the exceptional set (which will happen with probability zero), then $L(q_n|\xi) \rightarrow L(q|\xi)$.

Since for any given ξ the function $L(\cdot|\xi)$ is almost surely a continuous function over any one of the compact subsets Ω_k of Ω_0 , the corresponding AMLE, say \hat{q}_k , exists by the preceding remark and it is given by

$$(2.4) \quad L(\hat{q}_k|\xi) = \sup_{q \in \Omega_k} L(q|\xi).$$

Let $\eta = (y_1, \dots, y_{n-1})$ be an ordered $(n-1)$ -tuple of non-negative numbers and set

$$(2.5) \quad \Omega_k(\eta) = \{q \in \Omega_k : q(x_i) = y_i \quad i=1, \dots, n-1\},$$

which being a closed subset of a compact set is compact. Let

$E_k = \{\eta : \Omega_k(\eta) \neq \emptyset\}$, then

$$\begin{aligned}
 (2.6) \quad \sup_{q \in \Omega_k} L(q|\xi) &= \sup_{n \in E_k} \left\{ \sup_{q \in \Omega_k(n)} L(q|\xi) \right\} \\
 &= \sup_{n \in E_k} \left\{ \sum_{i=1}^{n-1} \ln y_i - \inf_{q \in \Omega_k(n)} \sum_{i=1}^n (n-i+1) A_i(q) \right\}
 \end{aligned}$$

where we set

$$A_i(q) = \int_{x_{i-1}}^{x_i} q(x) dx \quad \text{for } i = 1, \dots, n \text{ with } x_0 = 0.$$

3. The AMLE in the case of increasing hazard rate

In this case, clearly, $A_i(q)$ is minimized over $\Omega_1(n)$ by setting, for each $i = 1, \dots, n-1$

$$q(x) = y_{i-1} \quad \text{for } x_{i-1} \leq x < x_i \quad \text{with } y_0 = 0.$$

Therefore

$$A_i(q) = y_{i-1}(x_i - x_{i-1}) \quad i = 1, \dots, n.$$

Thus we want to evaluate, from (2.6),

$$(3.1) \quad \sup_{n \in E_1} \left\{ \sum_{i=1}^{n-1} \ln y_i - (n-i)y_i(x_{i+1} - x_i) \right\}$$

where

$$E_1 = \{(y_1, \dots, y_{n-1}) : y_1 \leq y_2 \leq \dots \leq y_{n-1}\}.$$

To do this we quote a result due to Brunk et al [1].

If $h(a_1, \dots, a_n)$ is a bounded strictly concave function for $-\infty < a_i < \infty$ $i=1, \dots, n$ which is to be maximized over the convex set

$$(3.2) \quad S_n = \{(a_1, \dots, a_n) : a_1 \leq a_2 \leq \dots \leq a_n\}$$

then it has a unique maximum which we designate by $\bar{a} = (\bar{a}_1, \dots, \bar{a}_n)$. Moreover, if $\alpha^0 = (\alpha_1^0, \dots, \alpha_n^0)$ is the unconstrained maximum then we have the following relationship holding: If $\bar{a}_k < \bar{a}_{k+1}$ for some $k = 1, \dots, n$ then

$$a_k^0 \leq \bar{a}_k < \bar{a}_{k+1} \leq \alpha_{k+1}^0 \quad \text{and} \quad \bar{a}_n \geq \alpha_n^0, \quad \bar{a}_1 \leq \alpha_1^0.$$

The usefulness of this theorem lies in the following procedure which is called "Brunkizing" see [3] and [4]: First obtain α^0 the unconstrained maximum of h over R_n . If $\alpha^0 \in S_n$, it is the solution \bar{a} sought.

If it does not fall in S_n then for some k we must have $a_k^0 > a_{k+1}^0$ and hence by the theorem above follows $\bar{a}_k = \bar{a}_{k+1}$. Thus we identify the variables a_k and a_{k+1} to obtain a new function of $(n-1)$ variables. We again maximize this function over R_{n-1} , checking to see if this unconstrained maximum falls in S_{n-1} . If it does we are through. If not, we repeat this procedure until at the final stage (which must occur in at most n steps) we have maximized in unconstrained fashion a function of a reduced number of variables. We then set \bar{a}_1 equal to the final value of the pooled variate to which that coordinate has contributed.

In many instances the unconstrained maximum may be a stationary point of the function h which can be found by setting the partial derivatives equal to zero, i.e.,

$$D_i h(\alpha) = 0 \quad \text{for } i = 1, \dots, n$$

and solving for the solution $\alpha^0 = (a_1^0, \dots, a_n^0)$. The procedure then is clear. This is especially useful when h is separable.

Lemma: If a convex function f is separable, i.e., of the form

$$f(x_1, \dots, x_n) = \sum_1^n f_i(x_i)$$

where each f_i is convex over $[0, \infty]$ and we let $u(i-k, i+j)$ be the value of x which minimizes

$$\sum_{\alpha=i-k+1}^{i+j-1} f_\alpha(x)$$

then the minimizing point of f over the region $0 \leq x_1 \leq \dots \leq x_n \leq \infty$, call it (a_1, \dots, a_n) is given by

$$a_i = \max_{k \geq 1} \min_{j \geq 1} u(i-k, i+j).$$

This is essentially the method used by Brunk loc. cit. to find some maximum likelihood estimates of parameters constrained in the same manner as (3.2).

Moreover applying this procedure to (3.1) leads to estimates \hat{q} which have been obtained by Grenander [7] and studied in detail by Proschan and Marshall [9]. We state these results as

Theorem 1: If $Q \in C_0$, i.e., q is IFR, then based on the sample ξ the AMLE of Q is $\hat{Q}(\cdot|\xi)$, which is the greatest convex minorant of $\{t, -\ln[1-\tilde{F}_n(t)]: t > 0\}$ where \tilde{F}_n is the empirical cumulative distribution based on ξ and \hat{q} is given explicitly by

$$\hat{q}(x_n|\xi) = \infty$$

$$\hat{q}(x_k|\xi) = \min_{j \geq k+1} \max_{i \leq k} \left[\sum_{\alpha=i+1}^j \frac{(n-\alpha+1)\Delta x_\alpha}{j-i} \right]^{-1} \quad \text{for } k = 1, \dots, n-1.$$

It is quite apparent that the estimation problems which arise for certain of the separate subclasses, as defined in Section 1, are related. We now state

Theorem 2: If $q \in S_r$, with r known, then the AMLE of Q , based on ξ is

$$\tilde{Q}(t|\xi) = \hat{Q}(R(t)|R(\xi))$$

where $R(\xi) = (R(x_1), \dots, R(x_n))$.

Proof: By definition, and a slight abuse of notation,

$$L(Q|\xi) = \sum_{i=1}^{n-1} \ln q(x_i) - \sum_1^n Q(x_i)$$

$$= \sum_{i=1}^{n-1} \ln \left(\frac{q(x_i)}{r(x_i)} \right) + \sum_1^{n-1} \ln r(x_i) - \sum_1^n (n-j+1) \int_{x_{j-1}}^{x_j} \frac{q(t)}{r(t)} dR(t).$$

Note that

$$L(Q|\xi) = L(QR^{-1}:R(\xi)) - \sum_1^{n-1} \ln rR^{-1}(x_1)$$

where juxtaposition of function denotes composition.

Since $q \in S_r$ iff $QR^{-1} \in C_0$ it follows that

$$\sup_{\{Q:q \in S_r\}} L(Q|\xi) = \sup_{\psi \in C_0} L(\psi|R(\xi)) - \sum_1^{n-1} \ln rR^{-1}(x_1).$$

Hence $\tilde{Q}(\cdot|\xi)$ is found at the AMLE $\hat{\psi}(\cdot|R(\xi))$ in C_0 but since $\psi R = Q$ we have $\hat{\psi}(R(t)|R(\xi)) = \tilde{Q}(t|\xi)$. ||

Thus having an explicit expression for the AMLE of q when it is known to be increasing enables one, using this result, to immediately calculate the AMLE for $q \in S_r$ for any known function r .

As an illustration we suppose $q \in S_1$ i.e. we know the failure rate divided by time is an increasing function. Thus

$$r(x) = x, \quad R(x) = \frac{x^2}{2} \quad \text{for } x > 0.$$

By the theorem above we have our estimate \tilde{q} defined for $t > 0$ by

$$\tilde{q}(t|\xi) = \sum_{k=1}^n \hat{q}[R(x_k)|R(\xi)] r(t) \{x_k \leq t < x_{k+1}\}$$

where we set $x_{n+1} = \infty$ and $\{P\}$ is the indicator of the relation P being one if true and zero otherwise. More explicitly

$$\tilde{q}(t|\xi) = \begin{cases} t \sum_{k=1}^{n-1} \min_{j \geq k+1} \max_{i \leq k} \sum_{\alpha=i+1}^j \left[\frac{(n-\alpha+1)(x_\alpha^2 - x_{\alpha-1}^2)}{2(j-i)} \right]^{-1} \{x_k \leq t < x_{k+1}\} \\ \infty & t \geq x_n \end{cases}$$

Because $Q \in S_R$ iff $QR^{-1} \in S_1$ and $Q \in S_1$ iff q is IFRA we see that if we could find a maximum likelihood estimate for q when it is IFRA we could, in a manner similar to the theorem above, find a corresponding estimate for $Q \in S_R$ for any known R . Unfortunately we cannot do so, as we now show, because the likelihood equation contains, for each observation

$$\ln q(x_k) - Q(x_k) = \ln[x_k G'(x_k) + G(x_k)] - x_k G(x_k)$$

where we define the increasing function $G(x) = Q(x)/x$. Clearly we can make the slope $G'(x_k)$ arbitrarily large while holding $G(x_k)$ finite. Thus no likelihood estimate exists since this holds for every observation.

§ 4. The AMLE in the Case of Convex Hazard Rate

Let Ω'_2 be the subset of Ω_2 each element of which has a derivative existing a.e. Since this subset is dense in Ω_2 , from (2.4), we have that

$$(4.1) \quad L(\hat{q}_2|\xi) = \sup_{q \in \Omega'_2} L(q|\xi).$$

Let Ω_2^* be the subset of Ω'_2 each element of which is piecewise linear, then

$$(4.2) \quad \sup_{q \in \Omega'_2} L(q|\xi) \geq \sup_{q \in \Omega_2^*} L(q|\xi).$$

Now

$$(4.3) \quad L(\hat{q}_2|\xi) = \sup_{\eta \in E'_2} \left\{ \sum_{i=1}^{n-1} \xi y_i - \inf_{q \in \Omega'_2(\eta)} \sum_{i=1}^n (n-i+1) A_i(q) \right\}$$

where

$$E'_2 = \{ \eta : \Omega'_2(\eta) \neq \emptyset \}.$$

Remark: If $q \in \Omega'_k(\eta)$ then there exists a piecewise linear convex function, say p_q , defined by

$$p_q(x) = \max\{y_i + q'(x_i)(x-x_i) : i=1, \dots, n-1\}$$

and such that for each $i = 1, \dots, n$

$$(4.4) \quad A_i(q) \geq A_i(p_q) \quad \text{with equality iff } q'(x_i) = q'(x_{i-1}).$$

Proof: The inequality (4.4) is geometrically obvious. If $q'(x_{i-1}) < q'(x_i)$ then the lines $d_i(x) = y_i + q'(x_i)(x-x_i)$ and $d_{i-1}(x)$ intersect at some point s_i and $x_{i-1} < s_i < x_i$. Now

$$q'(t) \geq q'(x_{i-1}) = p'_q(x_{i-1}) \quad \text{for } x_{i-1} < t < s_i$$

and therefore

$$q(x) - q(x_{i-1}) = \int_{x_{i-1}}^x q'(t) dt \geq q'_p(x_{i-1})(x-x_{i-1}) = q_p(x) - q_p(x_{i-1}).$$

By similar reasoning we obtain the same inequality on (s_i, x_i) . Thus

$$q(x) \geq q_p(x) \quad \text{for } x \in (x_{i-1}, x_i)$$

and integration yields the result. The claim for equality is easily seen. ||

However from this remark we notice that

$$\sup_{q \in \Omega'_2} L(q|\xi) \leq \sup_{p \in \Omega^*_2} L(p|\xi) .$$

and thus by (4.2) we have

$$\begin{aligned} L(\hat{q}_2|\xi) &= \sup_{q \in \Omega^*_2} L(q|\xi) \\ (4.5) \quad &= \sup_{\eta \in E^*_2} \sup_{p \in \Omega^*(\eta)} \left| \sum_{i=1}^{n-1} \ln y_i - \sum_{i=1}^n (n-i+1) A_i(p) \right| \end{aligned}$$

where

$$E^*_2 = \{ \eta : \Omega_k^*(\eta) \neq \emptyset \} .$$

For given $\xi = (x_1, \dots, x_n)$, let $\eta = (y_1, \dots, y_{n-1}) \in E^*_2$ be such that there exists $p \in \Omega^*_2(\eta)$, which is a piecewise, linear, convex non-negative function with slope m_i when passing through the points (x_i, y_i) for $i = 1, \dots, n-1$. Consequently we must have, as a necessary and sufficient condition,

$$(4.6) \quad m_1 \leq \frac{\Delta y_2}{\Delta x_2} \leq m_2 \leq \dots \leq \frac{\Delta y_{n-1}}{\Delta x_{n-1}} \leq m_{n-1} .$$

Let us set

$$(4.7) \quad x_0 = 0, \quad y_0 = y_1 - m_1 x_1, \quad y_n = y_{n-1} + m_{n-1} \Delta x_n, \quad m_n = m_{n-1}.$$

Define the salient points of p as

$$(4.8) \quad s_i = \frac{-\Delta y_i + \Delta(m_i x_i)}{\Delta m_i} \quad \text{for all } i \text{ such that } m_i > m_{i-1}.$$

Now

$$A_i(p) = \int_{x_{i-1}}^{x_i} p(t) dt = t p(t) \Big|_{x_{i-1}}^{x_i} - \int_{x_{i-1}}^{x_i} t dp(t),$$

and using the fact that $p'(t) = m_i$ for $t \in (s_{i-1}, s_i)$, where we set

$$s_1 = x_0 = 0, \quad s_n = x_n,$$

we have for $i = 1, \dots, n$,

$$\int_{x_{i-1}}^{x_i} t dp(t) = \frac{m_{i-1}}{2} (s_i^2 - x_{i-1}^2) + \frac{m_i}{2} (x_i^2 - s_i^2).$$

Substituting into equation (4.5) we find the expression in braces, which is to be maximized, becomes

$$(4.9) \quad \sum_1^{n-1} \ln y_i - \sum_1^n \left(x_i y_i - m_i \frac{x_i^2}{2} \right) - \sum_2^{n-1} \frac{n-i+1}{2} s_i^2 \Delta m_i.$$

Remark: Equation (4.9) defines a concave function of the vector

$$\alpha = (y_1, \dots, y_{n-1}, m_1, \dots, m_{n-1}).$$

Proof: That the first summation is concave and the second linear is obvious.

We show that the third, it being subtracted, is convex, by examining each term, namely

$$s_i^2 \Delta m_i = \frac{(y_{i-1} - y_i + m_i x_i - m_{i-1} x_{i-1})^2}{m_i - m_{i-1}} \quad \text{for } i = 2, \dots, n-1.$$

To see this is convex we need only see that for $0 < t < 1$ and $b_1 > b_2 > 0$ we have

$$\frac{(ta_1 + (1-t)a_2)^2}{tb_1 + (1-t)b_2} \leq t \frac{a_1^2}{b_1} + (1-t) \frac{a_2^2}{b_2}$$

which can be easily checked by elementary methods. ||

We have just argued that

$$L(\hat{q}_2 | \xi) = \sup_{\alpha \in S_{2n-2}^*} h(\alpha)$$

where S_{2n-2}^* is the convex set of vectors α satisfying (4.6) and

$$\begin{aligned} h(\alpha) = & \sum_1^{n-1} \ln y_i - \sum_1^{n-1} x_i y_i + \sum_1^{n-1} \frac{m_i x_i^2}{2} - \sum_2^{n-1} \frac{n-i+1}{2} s_i^2 \Delta m_i \\ & - x_n y_{n-1} - \frac{m_{n-1} x_n^2}{2} + x_n x_{n-1} m_{n-1}. \end{aligned}$$

Lemma: From (4.8) there follows

$$\frac{\partial s_i}{\partial y_k} = \begin{cases} 0 & i \neq k, k+1 \\ -1/\Delta m_k & i = k \\ -1/\Delta m_{k+1} & i = k+1 \end{cases}$$

and

$$\frac{\partial s_j}{\partial m_j} = \frac{x_j - s_j}{\Delta m_j}, \quad \frac{\partial s_{j+1}}{\partial m_j} = \frac{s_{j+1} - x_j}{\Delta m_{j+1}}$$

Remark: We now find, recalling $s_1 = 0$, $s_n = x_n$, that

$$\frac{\partial h}{\partial m_j} = \frac{n-j+1}{2} (x_j - s_j)^2 - \frac{n-j}{2} (s_{j+1} - x_j)^2 \quad j = 1, \dots, n-1$$

$$\frac{\partial h}{\partial y_j} = \frac{1}{y_j} - (n-j)\Delta s_{j+1} - (x_j - s_j) \quad j = 1, \dots, n-1.$$

Proof: By straightforward calculation we find

$$\begin{aligned} \frac{\partial h}{\partial m_j} &= \frac{x_j^2}{2} - \frac{n-j+1}{2} \left(2\Delta m_j s_j \frac{\partial s_j}{\partial m_j} + s_j^2 \right) \\ &\quad - \frac{n-j}{2} \left(2\Delta m_{j+1} s_j \frac{\partial s_{j+1}}{\partial m_j} - s_{j+1}^2 \right) \end{aligned}$$

which simplification shows to be equal to the result given, and

$$\frac{\partial h}{\partial y_k} = \frac{1}{y_k} - i_k - \sum_{i=2}^{n-1} \frac{n-i+1}{2} \Delta m_i s_i \frac{\partial s_i}{\partial y_k} \quad k = 2, \dots, n-2$$

from which the presented result follows. The special cases $k = 1$ and $k = n-1$ can also be shown to yield the result given by using the boundary conditions. ||

We want to find the solution to the simultaneous equations

$$\frac{\partial h}{\partial m_j} = 0 \quad j = 1, \dots, n-1; \quad \frac{\partial h}{\partial y_j} = 0 \quad j = 1, \dots, n-1.$$

This is equivalent with finding the solution, in the variables (y_1, \dots, y_{n-1}) (s_2, \dots, s_{n-1}) , to the equations

$$(4.10) \quad x_j - s_j = \sqrt{\frac{n-j}{n-j+1}} (s_{j+1} - x_j) \quad j = 1, \dots, n-1$$

$$(4.11) \quad \frac{1}{y_j} = (n-j)(s_{j+1} - x_j) + (n-j+1)(x_j - s_j) \quad j = 1, \dots, n-1$$

Is it possible that such a solution exists? The answer is no if one imposes the conditions $s_1 = 0$, $s_n = x_n$. This means that h does not attain a maximum over the unconstrained variables and so Brunkizing the function is not possible in the usual sense.

Let n be fixed and set $\lambda_j = \sqrt{\frac{n-j}{n-j+1}}$ for $j = 1, \dots, n-1$, $\lambda_n = 0$.

Then the set of equations (4.10) is equivalent with the system

$$(4.12) \quad s_j = x_j(1+\lambda_j) - \lambda_j s_{j+1} \quad j = 1, \dots, n-1.$$

Introduce the new variables $t_i = a_i s_i$ for $i = 1, \dots, n$, where $a_{j+1} = -a_j \lambda_j$ for $j = 1, \dots, n-1$. By taking the product of both sides for $j = 1, \dots, k-1$ we obtain

$$a_k = a_1 (-1)^{k-1} \frac{\sqrt{n-k+1}}{n} \quad k = 2, \dots, n$$

where $a_1 > 0$ is unspecified. Multiplying (4.12) by a_j yields $\Delta t_{j+1} = a_j (1 + \lambda_j) x_j$ for $j = 1, \dots, n-1$. Summing over $j = 1, \dots, k-1$, we obtain, recalling $s_1 = 0$

$$s_k = \sum_{j=1}^{k-1} \frac{a_j (1 + \lambda_j)}{a_k} x_j \quad k = 2, \dots, n-1.$$

Substituting the definitions of the constants, simplifying and rearranging the summation yields

$$(4.13) \quad s_k = x_{k-1} + \sum_{j=0}^{k-2} (-1)^{j-k} \frac{\sqrt{n-j}}{n-k+1} \Delta x_{j+1} \quad k = 2, \dots, n-1.$$

On the other hand summing over $j = k, \dots, n-1$ in the same way and using the fact that $s_n = x_n$ we find

$$(4.14) \quad s_k = x_k - \sum_{j=k}^{n-1} (-1)^{j-k} \frac{\sqrt{n-j}}{n-k+1} \Delta x_{j+1} \quad k = 2, \dots, n-1.$$

For consistency we must have these two expressions yield the same solution. This is equivalent with

$$\sum_{j=0}^{n-1} (-1)^{j-k} \frac{\sqrt{n-j}}{n-k+1} \Delta x_{j+1} = 0 \quad \text{for all } 1 \leq k \leq n < \infty,$$

which will be false except for a set of probability zero.

Of course this means that no maximum exists for the function h over the positive orthant. It must be that the maximum will occur on the boundary of S_{2n-2}^* and thus at some point we must consider the equations with certain of the variables identified.

Let us suppose that

$$m_k = m_{k+1} = \dots = m_{k+r}$$

and

$$y_{k+i} = y_k + m_k (x_{k+i} - x_k) \quad i = 0, \dots, r.$$

Then writing the function h in terms of these reduced variables we find, after lengthy but straightforward calculations similar to those done previously,

$$(4.15) \quad \frac{\partial h}{\partial m_k} = \sum_{i=0}^r \frac{x_{k+i} - x_k}{y_k + m_k (x_{k+i} - x_k)} - \frac{n-k+1}{2} (x_k - s_k)^2 - \frac{n-k-r}{2} (s_{k+r+1} - x_k)^2 - \frac{1}{2} \sum_{i=0}^r (x_{k+i} - x_k)^2.$$

$$(4.16) \quad \frac{\partial h}{\partial y_k} = \sum_{i=0}^r \frac{1}{y_k + m_k (x_{k+i} - x_k)} - \sum_{i=0}^r x_{k+i} + (n-k+1)s_k - (n-k-r)s_{k+r+1}.$$

§5. Some Critical Remarks

It is now clear that those concave likelihood functions which are to be maximized over the region $0 \leq x_1 \leq \dots \leq x_n < \infty$ and which are not separable so that Lemma 1 can not be utilized, and further do not possess a maximum over the positive orthant, so that Brunkizing can not be utilized, pose special difficulties in obtaining explicit formulae for the maximizing functions. Unfortunately this is the case for convex failure rates in general. It is clear from the preceding equations that in each instance the AMLE estimates could be found using a machine program. This would begin by assuming the slope was constant and then successively relax equality until the maximum was reached. Clearly the unavailability of an explicit formula carries with it other shortcomings.

References

- [1] Ayer, M., Brunk, H. D., Ewing, G. M., Reid, W. T., Silverman, E., (1955). An Empirical Distribution Function for Sampling with Incomplete Information. *Ann. of Math. Stat.*, 26, 641-647.
- [2] Barlow, R. E. and Proschan, F., (1965). *Mathematical Theory of Reliability*. John Wiley and Sons, Inc. New York.
- [3] Brunk, H. D., (1955). Maximum Likelihood Estimates of Monotone Parameters. *Ann. of Math. Stat.*, 26, 607-616.
- [4] Brunk, H. D., (1958). On the Estimation of Parameters Restricted by Inequalities. *Ann. of Math. Stat.*, 29, 437-454.
- [5] Cramér, H. and Wold, F., (1935). Mortality Variations in Sweden. *Skand. Aktuarietidskr.*, 18, 161-241.
- [6] Gompertz, B., (1825). On the Nature of the Function Expressive of the Law of Human Mortality and on a New Mode of Determining Value of Life Congingencies. *Philos. Trans. Roy. Soc. London*, 115, 513-585.
- [7] Grenander, Ulf, (1956). On the Theory of Mortality Measurement, Part II. *Skand. Aktuarietidskr.*, 39, 125-153.
- [8] Jones, Hardin B., (1961). Mechanism of Aging Suggested from a Study of Altered Death Risks in *Proceedings of the Fourth Berkeley Symposium on Mathematical Statistics and Probability*, University of California Press, Berkeley, 267-292.
- [9] Marshall, A. W., Proschan, F., (1965). Maximum Likelihood Estimation for Distributions with Monotone Failure Rate. *Ann. of Math. Stat.*, 36, 69-77.
- [10] Saunders, Sam C., (1968). On the Determination of a Safe Life for Classes of Distributions Classified by Failure Rate. *Tech.*, 10, 361-377.
- [11] Taylor, William F., (1961). On the Methodology of Studying Aging in Humans in *Proceedings of the Fourth Berkeley Symposium on Mathematical Statistics and Probability*, University of California, Berkeley, 347-368.