PROPERTIES OF PROBABILITY DISTRIBUTIONS WITH MONOTONE HAZARD RATE[1]

By Richard E. Barlow,[2] Albert W. Marshall and Frank Proschan

Boeing Scientific Research Laboratories

0. Summary. In this paper, we relate properties of a distribution function F (or its density f) to properties of the corresponding hazard rate q defined for $F(x) < 1$ by $q(x) = f(x)/[1 - F(x)]$. It is shown, e.g., that the class of distributions for which q is increasing is closed under convolution, and the class of distributions for which q is decreasing is closed under convex combinations. Using the fact that q is increasing if and only if $1 - F$ is a Pólya frequency function of order two, inequalities for the moments of F are obtained, and some consequences of monotone q for renewal processes are given. Finally, the finiteness of moments and moment generating function is related to limiting properties of q.

1. Introduction. The hazard rate is of interest because of its probabilistic interpretation: If, for example, F is a life distribution, $q(x)\,dx$ is the conditional probability of death in $[x, x + dx]$ given survival to age x. Because of this interpretation, unless otherwise indicated, F is assumed to be the distribution of a positive random variable, although for many of the results this is not necessary.

The hazard rate is important in a number of applications, and is known by a variety of names. It is used by actuaries under the name of "force of mortality" to compute mortality tables [22]. In statistics its reciprocal for the normal distribution is known as "Mill's ratio". It plays an important role in determining the form of extreme value distributions, and in extreme value theory is called the intensity function [9]. Tukey [23] obtains qualitative results concerning order statistics from distributions with monotonic hazard rates; he refers to such distributions as "subexponential". In reliability theory, an increasing hazard rate often corresponds to wearout; in models for replacement [4], checking [3], spare parts provisioning [5], etc., the assumption of an increasing hazard rate results in useful qualitative conclusions concerning the form of the solution. In the study of telephone traffic, a hazard rate $q(x)$ decreasing in $x \geqq 0$ is sometimes observed for the duration of a telephone call [15].

Although general results concerning hazard rate are obtained in this paper, particular attention is paid to distributions with monotone hazard rate. The definition of a distribution function with an increasing hazard rate is clearly equivalent to the statement that $1 - F$ is a Pólya frequency function of order 2

Received August 17, 1962.

[1] Part of the research of the first two authors was accomplished while they were at the Institute for Defense Analyses, Princeton, New Jersey.

[2] Now at San Jose State College.

(PF_2). First defined by Schoenberg [19], a PF_2 function is a non-negative measurable function $g(x)$ defined for all real x such that

$$\begin{vmatrix} g(x_1 - y_1) & g(x_1 - y_2) \\ g(x_2 - y_1) & g(x_2 - y_2) \end{vmatrix} \geqq 0 \tag{1.1}$$

whenever $x_1 < x_2$ and $y_1 < y_2$; in addition, $g(x) \not\equiv 0$ for at least two distinct values of x. This wide class of functions enjoys many useful properties such as variation diminishing, closure under convolution, unimodality [19], and certain moment properties [11]. (There is a substantial literature on variation diminishing transformations; see [19] for references.) When the underlying density is PF_2 (more generally a monotone likelihood ratio density) statistical decision procedures are particularly simple [12]. It is easy to verify that if the density f is PF_2 then the distribution has increasing hazard rate; the converse is not true, as we shall see in Section 3.

In Section 3 we confine attention to distributions with a monotone hazard rate. We prove that the increasing hazard rate property is closed under convolution, but not necessarily under convex combinations. On the other hand, the decreasing hazard rate property is not closed under convolution, while it is preserved under convex combinations. In Section 4, many of the moment properties of PF_2 densities obtained by Karlin, Proschan, and Barlow [11] are found to be possessed by the larger class of distributions with increasing hazard rate. In Section 5, applications are made to renewal processes yielding positivity properties for the distribution of the number of renewals, and to semi-Markov processes yielding an asymptotic bound on the mean occupation time. Finally in Section 6 results are obtained for general distributions of positive random variables, not assuming monotonicity of hazard rate.

2. Preliminaries. In this section we present definitions and results needed in later sections.

If $\{p_i\}$ is a discrete distribution, the hazard rate is defined on integers k satisfying $\sum_{i=k}^{\infty} p_i > 0$ by

$$q(k) = p_k / \sum_{i=k}^{\infty} p_i . \tag{2.1}$$

Note that in the discrete case $q(k) \leqq 1$.

The concept of Pólya frequency functions, where the argument is the difference of two variables in (1.1), was extended by Gantmacher and Krein [7] and by Karlin and Rubin [12] to functions of two separate variables. A function $f(x, y)$ of two real variables ranging over linearly ordered one-dimensional sets X and Y respectively is said to be totally positive of order k (TP_k) if for all $x_1 < x_2 < \cdots < x_m$, $y_1 < y_2 < \cdots < y_m$, ($x_i \varepsilon X$; $y_j \varepsilon Y$) and all $1 \leqq m \leqq k$, the determinant,

$$|f(x_i, y_j)| \geqq 0. \tag{2.2}$$

If $f(x, y)$ is TP_k for $k = 1, 2, \cdots$, then we say $f(x, y)$ is TP_∞. Typically X is an interval of the real line or a set of integers; similarly for Y. In the latter case the term "sequence" rather than "function" is used. In the special case in which the TP_k function $f(x, y)$ may be written as $g(x - y)$ with X and Y each consisting of the entire real line, g is said to be a *Pólya frequency function of order* k (PF_k).

A function $f(x, y)$ is *sign reverse regular of order* k (SRR_k) if $f(x, -y)$ is TP_k in $x \varepsilon X$, $y \varepsilon Y$. This means that the determinantal inequalities (2.2) hold, provided that the factor $(-1)^{m(m-1)/2}$ is inserted before the determinant. These definitions appear in [10] and [11].

In the continuous case, the "normalized" moments

$$\lambda_s = \frac{\mu_s}{\Gamma(s+1)} = \frac{1}{\Gamma(s+1)} \int_0^\infty x^s f(x)\, dx$$

are of considerable importance. In the discrete case, we often consider the binomial moments

$$B_0 = 1 \quad \text{and} \quad B_i = \sum_{k=0}^{\infty} \binom{k}{i} p_k, \qquad i = 1, 2, \cdots,$$

of distributions placing mass p_k at $k = 0, 1, 2, \cdots$, with $\sum p_k = 1$.

Central to the theory that follows are the exponential distribution $F(x) = 1 - e^{-x/\mu}$, $x \geqq 0$, $\mu > 0$ and the geometric distribution $p_k = p(1-p)^k$, $k = 0, 1, 2, \cdots; 0 \leqq p \leqq 1$; both are characterized by having constant hazard rate. The exponential distribution is PF_∞ with normalized moments

$$\lambda_s = \mu^s, \qquad s > -1. \tag{2.3}$$

The geometric distribution is a PF_∞ sequence, with binomial moments

$$B_i = [(1-p)/p]^i. \tag{2.4}$$

In the continuous case we use repeatedly the easily verified relation

$$1 - F(t) = \exp\left[-\int_{-\infty}^{t} q(x)\, dx\right]. \tag{2.5}$$

3. Monotone hazard rate. As indicated in the introduction, distributions with monotone hazard rate are of considerable practical interest. Such distributions constitute a very large class. If $q(x) = f(x)/[1 - F(x)]$ is increasing in x, $-\infty < x < \infty$, then we say that F (or f) has an increasing hazard rate (IHR). (Throughout this paper we use "increasing" to signify "non-decreasing", and "decreasing" to signify "non-increasing".) It is not possible that $q(x)$ is decreasing for all x, since $q(x)$ decreasing at $x = t$ implies that the density $f(x)$ is also decreasing at t. However, if the support of F is bounded from below, say by a, then $q(x)$ may be decreasing in $x \geqq a$. If the support of F were a finite interval, say $[a, b]$, then (using (2.5)),

$$\lim_{x \to b}\left[1 - \exp\left(-\int_x^{x+\Delta} q(z)\, dz\right)\right]$$
$$= \lim_{x \to b} [F(x+\Delta) - F(x)]/[1 - F(x)] = 1,$$

and hence $\limsup_{x \to b} q(x) = \infty$ so that q is not decreasing. Hence we say that a distribution F has a decreasing hazard rate (DHR) if its support is of the form $[a, \infty)$, and $q(x)$ is decreasing in $x \geqq a$. For convenience, we shall take $a = 0$.

If the distribution F has a density, it can be verified by differentiating

$$\log [1 - F(x)]$$

that $q(x)$ is increasing if and only if the support of F is an interval, and if on that interval, $\log [1 - F(x)]$ is concave. Similarly, F is DHR if and only if the support of F is $[0, \infty)$ and $\log [1 - F(x)]$ is convex in $x \geqq 0$. For a mathematical convenience and added generality, we use this concavity (convexity) property as the definition of IHR (DHR) whether or not a density exists. (See Schoenberg [19].)

Well-known examples are the gamma distributions with density $f(t) = \lambda(\lambda t)^{\alpha-1} \exp(-\lambda t)/\Gamma(\alpha)$, $\lambda, \alpha, t \geqq 0$, and the Weibull distribution with density $f(t) = \lambda \alpha t^{\alpha-1} \exp(-\lambda t^{\alpha})$, $\lambda, \alpha, t \geqq 0$, which have increasing hazard rate (IHR) for $\alpha > 1$ and decreasing hazard rate (DHR) for $\alpha < 1$. For $\alpha = 1$, both coincide with the exponential distribution, which is characterized by a constant hazard rate. The IHR distributions cited above and most commonly used distributions have the additional property that their densities are PF_2.

In the discrete case, an example is provided by the negative binomial (Pascal) distribution

$$p_k = \binom{-r}{k} p^r (p-1)^k \quad r \geqq 0, 0 \leqq p \leqq 1, k = 0, 1, \cdots,$$

which is IHR for $r > 1$ and DHR for $r < 1$. With $r = 1$, this distribution coincides with the geometric distribution, characterized by a constant hazard rate.

An important subclass of distributions with IHR are those with densities that are PF_2. This fact follows by writing $q^{-1}(x) = \int_0^\infty f(x + \Delta)\, d\Delta / f(x)$ and noting that the integrand is decreasing in x when f is PF_2.

Several useful and essentially known results (see [19]) follow directly from this definition. If F is IHR (DHR) and $F(0-) = 0$, then $x^{-1} \log [1 - F(x)]$, and hence $[1 - F(x)]^{1/x}$, is decreasing (increasing) in x. Bounds on $1 - F(x)$ in terms of percentiles can be obtained using this observation.

If $G(y) = 1 - e^{-y}$ for $y \geqq 0$ and F is a distribution with $F(0-) = 0$, then F is IHR (DHR) if and only if there exists a non-negative convex (concave) increasing function h such that $F(x) = G(h(x))$. If F is IHR and h is a non-negative convex increasing function, not identically constant, then $F(h(x))$ is IHR.

From (2.5), it follows that F is IHR (DHR) if and only if for all $x \geqq 0$

$$[F(t + x) - F(t)]/[1 - F(t)]$$

is increasing in t (decreasing in $t \geqq 0$ and $F(0-) = 0$).

The following basic theorem follows directly from the definitions and a fundamental result of Schoenberg ([19] p. 337); we omit the proof.

THEOREM 3.1. *F is* IHR *if and only if* $1 - F$ *is* PF_2. *F is* DHR *if and only if the support of F is* $[0, \infty)$, *and* $1 - F(x + y)$ *is* TP_2 *for* $x + y \geqq 0$.

A dual to the hazard rate is the ratio

$$f(x)/F(x). \tag{3.1}$$

If X is a time variable and time is reversed then $f(-x)/F(-x)$ becomes the hazard rate. Thus a random variable X has increasing hazard rate if and only if $-X$ has decreasing (3.1) ratio. Replacing X by $-X$ we obtain from Theorem 3.1 that (3.1) is decreasing in x if and only if F is PF_2.

In the discrete case, it can be shown directly that $\{p_j\}_{j=0}^{\infty}$ is IHR if and only if $\sum_{k+\Delta}^{\infty} p_j / \sum_{k}^{\infty} p_j$ is decreasing in k for integer $\Delta \geqq 1$. Thus $\{p_j\}_{j=0}^{\infty}$ is IHR if and only if $\{\sum_{j=k}^{\infty} p_j\}_{k=0}^{\infty}$ is a PF_2 sequence. According to Schoenberg [20], this is equivalent to the following. There exist two integers α, β, $0 \leqq \alpha \leqq \beta \leqq \infty$ such that $p_j > 0$ if and only if $\alpha \leqq j \leqq \beta$, and the polygonal line of vertices $x = k$, $y = \log \sum_{j=k}^{\infty} p_j$, $\alpha \leqq k \leqq \beta$, is concave. Thus discrete analogs of the condition that $\log [1 - F(x)]$ is concave and of Theorem 3.1 are valid.

The following examples show that $1 - F(x)$ may be PF_2 and the density $f(x)$ not PF_2, also $F(x)$ may be PF_2 while $f(x)$ is not PF_2. Let

$$f(t) = \begin{cases} \frac{2}{3}, & 0 \leqq t \leqq \frac{1}{2} \\ \frac{4}{3}, & \frac{1}{2} < t \leqq 1, \end{cases} \qquad f^*(t) = \begin{cases} \frac{4}{3}, & 0 \leqq t \leqq \frac{1}{2} \\ \frac{2}{3}, & \frac{1}{2} < t < 1. \end{cases}$$

It is easily verified that $f(t)/[1 - F(t)]$ and $F^*(t)/f^*(t)$ are increasing in t, but that $F(t)/f(t)$ and $f^*(t)/[1 - F^*(t)]$ are not increasing in t (at $t = \frac{1}{2}$) so that f and f^* are not PF_2.

The following example shows that densities with IHR need not be unimodal as are PF_2 densities. Let

$$f(x) = \begin{cases} 1 + a - 4ax, & 0 \leqq x \leqq \frac{1}{2} \\ 1 - 3a + 4ax, & \frac{1}{2} < x < 1 \end{cases} \qquad -1 \leqq a \leqq 1.$$

For $-1 \leqq a \leqq 0$, f is PF_2. For $-1 \leqq a \leqq 2 - 3^{\frac{1}{2}}$, f is IHR, and for $-1 \leqq a \leqq \frac{1}{3}$, $F(t)/f(t)$ is increasing. Of course, f is unimodal only for $-1 \leqq a \leqq 0$.

As another multimodal example, consider a distribution determined by (2.5) where q is an increasing step function. If q has a jump of magnitude Δ at x_0, then the density f has a jump of magnitude greater than Δ at x_0. Hence such an IHR distribution may have an infinite number of modes; in this case, it may also possess the additional property that $\limsup_{x\to\infty} f(x) > 0$.

The following examples show that the density $f(x)$ can fail to be TP_2 in sums of the argument and yet $1 - F(x)$ or $F(x)$ may still possess the TP_2 property. Let

$$1 - F(t) = \begin{cases} 1, & t < 0, \\ 2/[1 + (t + 1)^2], & t \geqq 0, \end{cases}$$

and

$$F^*(t) = \begin{cases} 2/[1 + (t - 1)^2], & t \leqq 0, \\ 1, & t \geqq 0. \end{cases}$$

It is easily verified that $\log [1 - F(t)]$ and $\log F^*(t)$ are convex for $t \geqq 0$ and $t \leqq 0$ respectively, but that $\log f(x)$ and $\log f^*(x)$ are not convex in the same range.

A key result is that the IHR property is preserved under convolution. The first proof of this theorem is due to Walter Weissblum; an alternate, somewhat more concise, proof is given here.

THEOREM 3.2. *If F and G are* IHR, *then their convolution H, given by $H(t) = \int_{-\infty}^{\infty} F(t - x)\, dG(x)$ is also* IHR.

PROOF. Assume F has density f, G has density g. For $t_1 < t_2$, $u_1 < u_2$, form

$$D = |1 - H(t_i - u_j)|_{i,j=1,2} = \left|\int [1 - F(t_i - s)]g(s - u_j)\, ds\right|$$

$$= \iint_{s_1 < s_2} |1 - F(t_i - s_k)|\; |g(s_k - u_j)|\, ds_2\, ds_1$$

by ([17] p. 48, prob. 68). (This representation has also been used by Karlin, Schoenberg, etc.) Integrating the inner integral by parts, we obtain

$$D = \iint_{s_1 < s_2} \begin{vmatrix} 1 - F(t_1 - s_1) & f(t_1 - s_2) \\ 1 - F(t_2 - s_1) & f(t_2 - s_2) \end{vmatrix} \begin{vmatrix} g(s_1 - u_1) & g(s_1 - u_2) \\ 1 - G(s_2 - u_1) & 1 - G(s_2 - u_1) \end{vmatrix} ds_2\, ds_1 .$$

The sign of the first determinant is the same as that of

$$\frac{f(t_2 - s_2)}{1 - F(t_2 - s_2)} \cdot \frac{1 - F(t_2 - s_2)}{1 - F(t_2 - s_1)} - \frac{f(t_1 - s_2)}{1 - F(t_1 - s_2)} \cdot \frac{1 - F(t_1 - s_2)}{1 - F(t_1 - s_1)},$$

assuming the denominators are non-zero. But $f(t_2 - s_2)/[1 - F(t_2 - s_2)] \geqq f(t_1 - s_2)/[1 - F(t_1 - s_2)]$ by hypothesis, while

$$[1 - F(t_2 - s_2)]/[1 - F(t_2 - s_1)] \geqq [1 - F(t_1 - s_2)]/[1 - F(t_1 - s_1)]$$

by Theorem 3.1. Thus the sign is non-negative. A similar argument holds for the second determinant, so that $D \geqq 0$. But by Theorem 3.1, this implies H is IHR. If F and/or G do not have densities, the theorem may be proved in a similar fashion using limiting arguments. ||

Theorem 3.2 holds also in the discrete case, and the proof is similar. Because no use was made of the assumption that X and Y are positive random variables, we obtain the following:

COROLLARY 3.3. *If $f(t)/F(t)$ and $g(t)/G(t)$ are decreasing in t, then $h(t)/H(t)$ is decreasing in t.*

PROOF. Replace X by $-X$ and Y by $-Y$ in Theorem 3.2 and use the remark at the end of Theorem 3.1.||

It is of interest to note that the DHR property is *not* preserved under convolution. A counter-example is obtained if f and g are gamma densities with $\frac{1}{2} \leqq \alpha < 1$. However, it is true that a mixture of DHR distributions is also DHR. The following theorem may be obtained as a consequence of the result in [1] which states that the sum of logarithmically convex functions is itself logarithmically convex. We present a somewhat different proof.

THEOREM 3.4.[3] *If* $F(t, \phi)$ *is a* DHR *distribution in* t *for each* ϕ *in* Φ, *then* $G(t) = \int_\Phi F(t, \phi)\, d\mu(\phi)$ *is* DHR *where* μ *is a probability measure in* Φ.

PROOF. First suppose that $F(t, \phi)$ has a differentiable density $f(t, \phi)$. Since the density of any DHR distribution must be a decreasing function, we have by Schwarz's inequality that

$$\int [1 - F(t, \phi)]\, d\mu(\phi) \int -f'(t, \phi)\, d\mu(\phi) \geqq \left\{\int \{1 - F(t, \phi)][-f'(t, \phi)]\}^{\frac{1}{2}}\, d\mu(\phi)\right\}^2.$$

Since $f(t, \phi)/[1 - F(t, \phi)]$ is decreasing in t, we must have

$$[1 - F(t, \phi)]f'(t, \phi) \leqq -[f(t, \phi)]^2.$$

Hence $\int [1 - F(t, \phi)]\, d\mu(\phi) \int - f'(t, \phi)\, d\mu(\phi) \geqq \{\int f(t, \phi)\, d\mu(\phi)\}^2$, that is, $[1 - G(t)]g'(t) \leqq -[g(t)]^2$, so that G is DHR.

If F does not have a differentiable density, the same result may be obtained by limiting arguments.||

Mixtures of IHR distributions are not necessarily IHR. For example, a mixture of two distinct exponentials is not IHR since it is not exponential, and by the above theorem it is DHR.

The variation diminishing properties of totally positive functions are well known [19]. Because of the relative weakness of the IHR property, we obtain a correspondingly weak variation diminishing property. We define the number of changes of trend of a function g by $T(g) = \lim_{\Delta \to 0} V\{g(x + \Delta) - g(x)\}$, when the limit exists, where V refers to number of changes of sign [16].

THEOREM 3.5. *Let* F *be a continuous* IHR *distribution*, g *a real absolutely continuous function on* $(-\infty, \infty)$ *with* $V(g) \leqq 1$, $T(g) \leqq 1$. *Suppose that* $h(x) = \int_{-\infty}^{\infty} g(x - y)\, dF(y)$ *exists. Then* $V(h) \leqq V(g)$. *Furthermore, if* $V(h) = 1$, h *changes sign in the same order as* g.

PROOF. If $V(g) = 0$ or $T(g) = 0$, the result is obvious.

Suppose $V(g) = 1$, $g(y) \geqq 0$ for $y < b$, $g(y) \leqq 0$ for $y > b$, and $g(y)$ is decreasing for $y < a$ and increasing for $y > a$ where $-\infty < b < a < \infty$. Suppose $h(x_1) \leqq 0$. Consider any $x_2 > x_1$. We may integrate by parts ([18] p. 102) to obtain:

$$h(x_2) = g(\infty) - \int_{-\infty}^{\infty} [1 - F(x_2 - y)]\, dg(y),$$

[3] Professor Karlin has pointed out that Theorem 3.4 can be extended to the TP_k case.

noting that $|g(\infty)| < \infty$ by assumption while $\lim_{u\to\infty} g(-u)\ [1 - F(u) = 0$ since $\int_{-\infty}^{\infty} g(x - y)\ dF(y)$ exists by hypothesis. Hence

$$h(x_2) = g(\infty) + \int_{-\infty}^{a} \frac{1 - F(x_2 - y)}{1 - F(x_1 - y)} [1 - F(x_1 - y)]\ d[-g(y)]$$

$$- \int_{a}^{\infty} \frac{1 - F(x_2 - y)}{1 - F(x_1 - y)} [1 - F(x_1 - y)]\ dg(y)$$

$$\leqq g(\infty) + \int_{-\infty}^{a} \frac{1 - F(x_2 - a)}{1 - F(x_1 - a)} [1 - F(x_1 - y)]\ d[-g(y)]$$

$$- \int_{a}^{\infty} \frac{1 - F(x_2 - a)}{1 - F(x_1 - a)} [1 - F(x_1 - y)]\ dg(y)$$

since by Theorem 3.1, $1 - F$ is PF_2, so that $[1 - F(x + \Delta)]/[1 - F(x)]$ is decreasing in x for $\Delta > 0$. Thus

$$h(x_2) \leqq g(\infty) - \frac{1 - F(x_2 - a)}{1 - F(x_1 - a)} \int_{-\infty}^{\infty} [1 - F(x_1 - y)]\ dg(y)$$

$$= g(\infty) + \frac{1 - F(x_2 - a)}{1 - F(x_1 - a)} [h(x_1) - g(\infty)]$$

$$= g(\infty)\left[1 - \frac{1 - F(x_2 - a)}{1 - F(x_1 - a)}\right] + \frac{1 - F(x_2 - a)}{1 - F(x_1 - a)} h(x_1) \leqq 0.$$

Thus $V(h) \leqq 1$. Moreover, if $V(h) = 1$, then h changes sign from $+$ to $-$, i.e., in the same order as g.

The remaining cases with $V(g) = 1$ can be reduced to the above case either by replacing g by $-g$, y by $-y$, or both.||

Note that the assumption of continuous F and g was used to justify integration by parts. Actually the result follows whenever integration by parts is valid, i.e., for g absolutely continuous even if F has the single discontinuity at the right hand of the interval of support, possible for IHR distributions.

4. Moment inequalities. Many of the moment inequalities obtained in [11] are true for the larger class of distributions studied in this paper. It will be convenient to define

$$(4.1)\qquad \gamma^{(s)}(t) = (-t)^{s-1}/\Gamma(s), \quad t \leqq 0$$
$$= 0, \quad t > 0,$$

$$(4.2)\qquad f_s(t) = \int_{-\infty}^{\infty} [\gamma^{(s)}(x)/\lambda_s] f(t - x)\ dx, \quad s > 0$$
$$= f(t), \quad s = 0,$$

where f is a density on the positive axis and $\lambda_s = \int_0^\infty [x^s f(x)/\Gamma(s + 1)]\ dx$.

We remark that if f is IHR, then $1 - F(x)$ tends to zero exponentially fast [19] and hence f has finite moments of all orders.

Using the identity $\gamma^{(r+s)}(t) = \int_{-\infty}^{\infty} \gamma^{(r)}(x)\gamma^{(s)}(t - x)\ dx$ it is easily verified that $f_s(t)$ is a density. Note that $f_1(x+y)$ SRR_2 in $x+y \geqq 0$ implies that f is IHR. Also $f_2(x+y)$ SRR_2 in $x+y \geqq 0$ implies f_1 is IHR, or equivalently, the "mean residual life" $\int_t^\infty [1-F(x)]\,dx/[1-F(t)]$ corresponding to the density f is decreasing. This class of distributions is of natural interest in reliability theory.

With the notation $\mu_r^{(s)} = \int_0^\infty x^r f_s(x)\,dx$, $\lambda_r^{(s)} = \mu_r^{(s)}/\Gamma(r+1)$, it also follows from the definition of f_s that

$$\lambda_r^{(s)} = \lambda_{r+s}/\lambda_s\,. \tag{4.3}$$

Using (4.3) we obtain $\prod_{i=0}^{j-1} \mu_1^{(i)} = \lambda_j/\lambda_0$, so that for integral i and n,

$$\lambda_n^{(i)} = \prod_{j=1}^{i+n-1} \mu_1^{(j)}. \tag{4.4}$$

Proceeding in the manner of Theorem 1 of [11], p. 1025, we obtain

THEOREM 4.1. *If* $f_s(x+y)$ *is* SRR_k *in* $x+y \geqq 0$, *then*

(i) $f_{n+s}(x+y)$ *is* SRR_k *in* $x+y \geqq 0$ *for all positive integers* n, *and*

(ii) $f_{r+s}(x)$ *is* SRR_k *in* $r \geqq 0$ *and* $x \geqq 0$.

(i) implies that if f is IHR then f_n is IHR, $n = 1, 2, \cdots$. A converse can be given in the case $k = 2$.

THEOREM 4.2. f *is* IHR *if and only if* $f_i(x)$ *is* SRR_2 *in* $i = 0, 1, 2, \cdots$ *and* $x \geqq 0$.

PROOF. Suppose that f is IHR. The support of f_n coincides with that of f, and it is sufficient to show

$$\begin{vmatrix} f_i(t) & f_i(t+\Delta) \\ f_{i+j}(t) & f_{i+j}(t+\Delta) \end{vmatrix} \leqq 0 \tag{4.5}$$

for t and $t+\Delta$ $(\Delta > 0)$ in the support of f. By (i) of Theorem 4.1, f_n is IHR, $n = 1, 2, \cdots$; i.e., $f_n(t)/f_{n+1}(t) \leqq f_n(t+\Delta)/f_{n+1}(t+\Delta)$. This means that $f_n(t)/f_n(t+\Delta)$ is increasing in n, and yields (4.5). Of course it is trivial that f is IHR if $f_i(x)$ is SRR_2 in i and x.

Note that by integrating on Δ from 0 to infinity in (4.5) we obtain that $q_i(t)$ is increasing in $i = 0, 1, \cdots$ for all t.

The following theorem extends Theorem 1 of [11] and the proof is similar.

THEOREM 4.3. *If* $f_s(x+y)$ *is* SRR_k *in* $x+y \geqq 0$ *then* λ_{r+s+t} *is* SRR_k *in* $r+t \geqq -1$.

In particular if f is IHR, then λ_{r+s} is SRR_2 in $r+s \geqq 0$. Further extensions are possible in terms of integer moments. Define

$$\begin{aligned} a_{i+j} &= \lambda_{i+j-1}\,, && i+j > 0 \\ &= f(0), && i+j = 0. \end{aligned}$$

Again following the method of proof of Theorem 1 of [11], p. 1025, it is not hard to show that

THEOREM 4.4. *If* $f_k(x+y)$ *is* SRR_k *in* $x+y \geqq 0$, *then* a_{i+j} *is* SRR_k *in* $i+j = 0, 1, \cdots$.

In particular, if f has decreasing mean residual life, then a_{i+j} is SRR_2 in $i + j = 0, 1, 2, \cdots$.

If f is IHR, another moment inequality can be obtained from (16) of [11] and the fact that each f_i is PF_2. Specifically,

$$[\lambda_{i+t}/\lambda_i]^s = [\lambda_t^{(i)}]^s \leqq [\lambda_s^{(i)}]^t = [\lambda_{i+s}/\lambda_i]^t,$$

$t > s > 0$, or $\lambda_a^{b-c}\lambda_c^{a-b} \leqq \lambda_b^{a-c}$, $a \geqq b \geqq c \geqq 1$, c an integer. This inequality is to be compared with Lyapounov's inequality, $\mu_a^{b-c}\mu_c^{a-b} \geqq \mu_b^{a-c}$, $a \geqq b \geqq c \geqq 0$.

5. Applications to renewal processes. Let $X_1, X_2, \cdots$ be a renewal process, that is, a sequence of independent, non-negative, and identically distributed random variables which are not zero with probability one. Write F for the distribution of X_1, $S_k = X_1 + X_2 + \cdots + X_k$, and F^k for the distribution of S_k.

Let $N(t)$ be the maximum index k for which $S_k \leqq t$, subject to the convention $N(t) = 0$ if $X_1 > t$. The first moment $M(t) = E[N(t)]$ satisfies

$$M(t) = F(t) + \int_0^t M(t - x)\, dF(x). \tag{5.1}$$

If F has density f, the derivative of $M(t)$ will be denoted by $m(t)$ and satisfies $m(t) = f(t) + \int_0^t m(t - x)f(x)\, dx$ (see, e.g., [21]). To obtain bounds on $m(t)$, define the "shortage" random variable, $\delta_t = t - S_{N(t)}$, and denote its distribution by G_t. Then

$$m(t) = \int_0^t \{f(x)/[1 - F(x)]\}\, d_x G_t(x)$$

and hence $\inf_{0 \leqq x \leqq t} q(x) \leqq m(t) \leqq \sup_{0 \leqq x \leqq t} q(x)$. If $q(t)$ is increasing,

$$tf(0) \leqq M(t) \leqq \int_0^t q(x)\, dx.$$

Equality holds in the exponential case.

If F is IHR, the distribution of $N(t)$ has important positivity properties.

THEOREM 5.1. *If F is* IHR, *then*

(a) *$P[N(t) \leqq n]$ is TP_2 in $t \geqq 0$ and $n \geqq 0$;*

if, in addition, $F(0) = 0$, then

(b) *$P[N(t) \leqq n]$ is PF_2 in integer n.*

PROOF. By Theorems 3.1 and 3.2

$$\begin{vmatrix} 1 - F^n(t) & 1 - F^{n+1}(t) \\ 1 - F^n(t-x) & 1 - F^{n+1}(t-x) \end{vmatrix} = \int \begin{vmatrix} 1 - F^n(t) & 1 - F^n(t-u) \\ 1 - F^n(t-x) & 1 - F^n(t-x-u) \end{vmatrix} dF(u) \leqq 0$$

which proves (a) since $P[N(t) \leqq n] = 1 - F^n(t)$. Convolving terms in the bottom row of the second determinant with F we obtain (b).||

COROLLARY 5.2. *If F is PF_2 and $F(0) = 0$, then*

(a) *$P[N(t) \geqq n]$ is SRR_2 in $t \geqq 0$ and $n \geqq 0$.*

(b) $P[N(t) \geqq n]$ *is* PF_2 *in* n.

PROOF. (a) Under the hypotheses of Theorem 5.1, $1 - F^n(t)$ is TP_2 for $-\infty < t < \infty$ and $n \geqq 0$. Hence if F is PF_2 and X has distribution F, $-X$ is IHR and $P[-X_1 - X_2 - \cdots - X_n \geqq t]$ is TP_2 in $-\infty \leqq t < \infty$ and $n \geqq 0$. Hence $F^n(t)$ is SRR_2 in $-\infty \leqq t < \infty$ and $n \geqq 0$ and proof proceeds as in Theorem 5.1.||

REMARKS. (a) of Theorem 5.1 and (a) of Corollary 5.2 are still true even if the X_i are not identically distributed. (b) of Theorem 5.1 may be used to weaken the assumptions made in [10] for solving an inventory problem.

It is well known (see, e.g., [21]) that

$$M(t) - t/\mu_1 = (\lambda_2 - \lambda_1^2)/\lambda_1^2 + o(1).$$

Hence if F is IHR (DHR), t/μ_1 overestimates (underestimates) $M(t)$ for large t.

The semi-Markov process ([21] p. 260) is a generalization of the renewal process. These processes have countable state spaces, and the sequence of states forms a stationary Markov chain with transition matrix (p_{ij}). The waiting times between transitions are independent, and the wait in state i given that the next state is j has distribution F_{ij}. The density g_{ij} of the first passage time from state i to state j satisfies

$$g_{ij}(t) = p_{ij}f(t) + \sum_{k \neq j} p_{ik} \int_0^\infty f(t - \theta)g_{kj}(\theta)\, d\theta \tag{5.2}$$

where f is the density of the unconditional waiting time in state i. We have the following generalization of Theorem 2.5 of [2].

THEOREM 5.3. *If* f *is* IHR

$$\begin{vmatrix} \mu_r & l_{ij}^{(r)} \\ \mu_s & l_{ij}^{(s)} \end{vmatrix} \geqq 0$$

where $1 \leqq r \leqq s$ *and* $l_{ij}^{(r)}$ *is the* rth *moment of* $g_{ij}(t)$.

PROOF. We need only show

$$\begin{vmatrix} \gamma^{(r)} * f(t) & \gamma^{(r)} * g_{ij}(t) \\ \gamma^{(s)} * f(t) & \gamma^{(s)} * g_{ij}(t) \end{vmatrix} \geqq 0,$$

where $\gamma^{(r)}$ is defined in (4.1). From (5.1) it is clearly sufficient to show

$$\int_0^\infty \begin{vmatrix} \gamma^{(r)} * f(t) & \gamma^{(r)} * f(t - \theta) \\ \gamma^{(s)} * f(t) & \gamma^{(s)} * f(t - \theta) \end{vmatrix} g_{kj}(\theta)\, d\theta \geqq 0$$

which is true by assumption.||

The limiting error made in estimating the mean occupation time before t in state i starting in state i by $\mu_1 t/l_{ii}$ is [2]

$$\lim_{t\to\infty} \int_0^t [P_{ii}(x) - \mu_1/l_{ii}]dx = (\mu_1 l_{ii}^2 - l_{ii}\mu_2)/2l_{ii}^2$$

where $P_{ii}(x)$ is the probability that the process is in state i after x time units if it starts in state i at $t = 0$. If f is IHR this error term is positive by Theorem 5.3.

6. General results using the hazard rate. In this section we relate properties of the distribution function to properties of the hazard rate at infinity, and prove a limit theorem for $F_i(x)$.

THEOREM 6.1. *Let* $F(0) = 0$, *and let* $a \leqq 1$. *If* $\lim_{t\to\infty} t^a q(t) = L(a)$ *exists (finite or infinite), then for all* $b \geqq 0$,

$$\begin{aligned} \lim_{t\to\infty} [1 - F(t)]/[1 - F(t + bt^a)] &= e^{bL(a)}, & a < 1, \\ &= (1 + b)^{L(a)}, & a = 1. \end{aligned} \tag{6.1}$$

PROOF. We prove the theorem only for $L < \infty$; modifications required for $L = \infty$ should be clear. Choose $\epsilon > 0$. Then there exists $T < \infty$ such that $x > T$ implies $(L - \epsilon)x^a < q(x) < (L + \epsilon)x^a$. Since

$$\int_t^{t+bt^a} q(x)\, dx = \log [1 - F(t)] - \log [1 - F(t + bt^a)],$$

it follows that for $t > T$,

$$\int_t^{t+bt^a} \frac{L - \epsilon}{x^a}\, dx < \log [1 - F(t)] - \log [1 - F(t + bt^a)] < \int_t^{t+bt^a} \frac{L + \epsilon}{x^a}\, dx.$$

The result follows by performing the indicated integration and letting $t \to \infty$. ||

REMARK. Let $X_1, X_2, \cdots$ be a sequence of independent random variables with distribution function F. Then $\xi_i = \max (X_1, X_2, \cdots, X_i)$ are said to satisfy the law of large numbers if there exist real numbers $\{A_i\}$ such that $P\{|\xi_i - A_i| < \epsilon\} \to 1$ as $i \to \infty$ for every $\epsilon > 0$. It is shown in [8] that this condition is equivalent to $\lim_{t\to\infty} [1 - F(t + \epsilon)]/[1 - F(t)] = 0$. With $a = 0$, we obtain from (6.1) that is is equivalent to $\lim_{t\to\infty} q(t) = \infty$. Similarly, in [8] the condition that ξ_i be relatively stable is defined and shown to be equivalent to $\lim_{t\to\infty} [1 - F(kt)]/[1 - F(t)] = 0$ for all $k > 1$ ($a = 1$ in (6.1)).

In view of (6.1) and known results, the following theorem should not be unexpected.

THEOREM 6.2. *Let* $r > 0$, *and let* $F(0) = 0$. *Then* $\mu_r < \infty$ *if*

$$r < \liminf_{t\to\infty} tq(t),$$

and $\mu_r = \infty$ *if* $r > \limsup_{t\to\infty} tq(t)$.

To prove this result, obtain bounds for $\log[1 - F(x)]$ as in the proof of (6.1) and use the representation $\mu_r = \int_0^\infty rx^{r-1}[1 - F(x)]\, dx$.

We conclude from Theorem 6.2 that if F is IHR, $\mu_r < \infty$ for all $r > 0$, but that there exist DHR distributions for which $\mu_r = \infty$ for all $r > 0$, namely those for which $\lim_{t\to\infty} tq(t) = 0$.

THEOREM 6.3. *If* $s < \liminf_{t\to\infty} q(t)$, *then* $\varphi(s) = \int_0^\infty e^{st}\, dF(x) < \infty$. *If* $s > \limsup_{t\to\infty} q(t)$, *then* $\varphi(s) = \infty$.

The proof of this criterion for finiteness of the moment generating function is similar to the proof of Theorem 6.2.

It is clear from the proofs of the above theorems that bounds on the hazard rate immediately yield various other bounds. In particular, if $0 < \alpha \leqq q(t) \leqq \beta \leqq \infty$ for all t, then

$$\exp(-\beta t) \leqq 1 - F(t) \leqq \exp(-\alpha t) \tag{6.2}$$

$$\alpha \exp(-\beta t) \leqq f(t) \leqq \beta \exp(-\alpha t), \tag{6.3}$$

$$\beta^{-s} \leqq \lambda_s \leqq \alpha^{-s}, \qquad s > -1. \tag{6.4}$$

Of course analogous statements are true in the discrete case.

THEOREM 6.4. *If* $\lim_{i\to\infty} \mu_1^{(i)} = \mu < \infty$ *exists, then the limit in distribution of* $F_i(x) = \int_0^x f_i(t)\,dt$ *is*

$$\begin{aligned} F^*(x) &= 1 - e^{-x/\mu}, && \mu \neq 0, \quad x \geqq 0, \\ &= 1, && \mu = 0, \quad x \geqq 0, \\ &= 0, && x < 0. \end{aligned}$$

PROOF. Using (4.4), we conclude that $\lim_i \mu_n^{(i)} = n!\mu^n$, the nth moment of F^*. F^* is the only distribution with moments $n!\mu^n$, and this uniqueness implies that $F_i \to F^*$ in distribution ([14] p. 185).

Writing $f_1(t) = [1 - F(t)]/\mu_1$, it is seen that Theorem 3.2 is true whether or not F has a density. In spite of this there is a discrete analog to Theorem 6.4. The analogous transformation of $\{p_j\}_{j=0}^{\infty}$ yields the probability distribution $\{p_j^{(1)}\}$ whre $p_j^{(1)} = \sum_{k=j}^{\infty} p_k/(B_0 + B_1)$ and in general,

$$p_j^{(i)} = \sum_{k=j}^{\infty} p_k^{(i-1)}/D_{i-1}\,.$$

The normalizing factor

$$D_{i-1} = B_0^{(i-1)} + B_1^{(i-1)} = \left[\sum_{k=0}^{i+1} \binom{i+1}{k} B_k\right] \Big/ \left[\sum_{k=0}^{i} \binom{i}{k} B_k\right]$$

($B_k^{(i)}$ is the kth binomial moment of $\{p_j^{(i)}\}$) can be obtained after first establishing by induction that

$$B_k^{(i)} = \left[\sum_{j=0}^{i+1} \binom{i+1}{j} B_{j+k}\right] \Big/ \left[\sum_{j=0}^{i+1} \binom{i+1}{j} B_j\right].$$

If $\{p_j\}$ is PF_2 then D_i is non-increasing in i and $\lim_{t\to\infty} D_i$ exists, for since $\{B_k\}$ and $\binom{i}{k}$ are PF_2 sequences in k, the convolution $\sum_{k=0}^{i} \binom{i}{k} B_{i-k}$ is PF_2.

Using the recursion $P_n^{(i+1)} - P_n^{(i)}/D_i = P_{n+1}^{(i+1)}$, we obtain by induction

THEOREM 6.4$'$. *If* $\lim_{i\to\infty} D_i = 1/B < \infty$ *exists, then* $\lim_{i\to\infty} p_j^{(i)} = B(1 - B)^j$.

The following lemma enables us to easily relate $\lim_{t\to\infty} q(t)$ and $\lim_{i\to\infty} \mu_1^{(i)}$.

LEMMA 6.5. (i) $q_1(t) \geqq q(t)$ *if and only if* q_1 *is non-decreasing at* t; *moreover* $q_1'(t) = 0$ *if and only if* $q_1(t) = q(t)$. (ii) *If* $\lim_{t\to\infty} q_i(t) = a$ *exists for some* i, *then* $\lim_{t\to\infty} q_i(t) = a$ *uniformly in* i.

PROOF. (i) follows by differentiating $q_1(t)$. To prove (ii), note that

$$\lim_{t\to\infty} q_i(t) = a$$

implies $\lim_{t\to\infty} q(t) = a$ by l'Hospital's rule. If we prove

$$\sup_{s\geqq t} q(s) \geqq \sup_{s\geqq t} q_i(s) \geqq \inf_{s\geqq t} q_i(s) \geqq \inf_{s\geqq t} q(s) \quad \text{for all } i, \tag{6.5}$$

then the uniformity follows. Suppose that $q_1(r) < \inf_{s>t} q(s) - \epsilon$ for some $r > t$ and $\epsilon > 0$. Then by (i), $q(u) \leqq \inf_{s\geqq t} q(s) - \epsilon$ for all $u \geqq r$ so that $\lim_{s\to\infty} q_1(s) < \lim_{s\to\infty} q(s)$, a contradiction. The proof for the supremum is similar. ||

THEOREM 6.6. *If* $\lim_i \mu_1^{(i)} = \mu < \infty$ *and* $\lim_{t\to\infty} q(t) = a$ *both exist, then* (i) $a = \mu^{-1}$, (ii) $r > \mu$ *implies* $\lambda_k = o(r^k)$, (iii) $\mu > 1(<1)$ *implies* $\lim_k \lambda_k = \infty\,(0)$.

PROOF. The uniformity of $\lim_{t\to\infty} q_i(\epsilon) = a$ justifies interchange of limits in $\lim_{t\to\infty} q_i(t) = \lim_{i\to\infty} \lim_{t\to\infty} q_i(t) = \lim_{t\to\infty} \lim_{i\to\infty} q_i(t) = \mu^{-1}$, the last equality following from Theorem 6.4. To prove (ii), note that the radius of convergence of $\sum_{k=0}^{\infty} \lambda_k z^k$ is $\lim_{i\to\infty} \lambda_{i-1}/\lambda_i = \mu^{-1}$ so that $\lim_{k\to\infty} \lambda_k r^{-k} = 0$ when $r > \mu$. To prove (iii), suppose $\mu > 1$, choose $\epsilon > 0$ and t_0 such that $q(t) \leqq \mu^{-1} + \epsilon < 1$ for $t > t_0$. The result then follows by truncation of the density at t_0. A similar proof can be given for $\mu < 1$, but the result follows from (ii). ||

Proofs of the following discrete analogs are omitted.

LEMMA 6.5′. (i) $q_1(k) \geqq q(k)$ *if and only if* $q_1(k) \leqq q_1(k + 1)$. (ii) *If*

$$\lim_{k\to\infty} q_i(k) = a$$

exists for some i, *then* $\lim_{k\to\infty} q_i(k) = a$ *uniformly in* i.

THEOREM 6.6′. *If* $\lim_{i\to\infty} D_i = B^{-1}$ *and* $\lim_{k\to\infty} q(k) = a$ *both exist, then* $a = B$.

We conclude this section with a remark concerning the hazard rates of convolutions. In the following q_f refers to the failure rate corresponding to density f, and similarly for q_g and q_h, where $h(t) = \int_0^t f(t - x)g(x)\,dx$.

REMARK. If $u(t) = \sup_{0\leqq x\leqq t} q_f(x)$ and $v(t) = \sup_{0\leqq x\leqq t} q_g(x)$, then

$$q_h(t) \leqq \min\,(u(t), v(t)), \tag{6.6}$$

$$\lim_{t\to\infty} q_h(t) = \min\,[\lim_{t\to\infty} q_f(t), \lim_{t\to\infty} q_g(t)], \tag{6.7}$$

providing the limits exist.

Acknowledgments. We would like to acknowledge the helpful suggestions of Samuel Karlin, George Weissblum, and Lloyd Welch.

REFERENCES

[1] ARTIN, E. (1931). *Einfuhrung in die Theorie der Gamma Funktion.* Hamburger Mathematische Einzelschriften, 11. Heft.

[2] BARLOW, RICHARD E. (1962). Applications of semi-Markov processes to counter and reliability problems. *Studies in Applied Probability and Management Science*, ed. by K. J. Arrow, S. Karlin and H. Scarf, Stanford Univ. Press.

[3] BARLOW, RICHARD E., HUNTER, LARRY C. and PROSCHAN, FRANK (1961). Optimum monitoring procedures. Boeing Scientific Research Laboratories Document D1-82-0104.

[4] BARLOW, RICHARD E. and PROSCHAN, FRANK (1962). Planned replacement. *Studies in Applied Probability and Management Science*, ed. by K. J. Arrow, S. Karlin and H. Scarf, Stanford Univ. Press.

[5] BARLOW, RICHARD E., HUNTER, LARRY C. and PROSCHAN, FRANK (in preparation). *Probabilistic Models in Reliability Theory*. Wiley, New York.

[6] ESARY, JAMES D. and PROSCHAN, FRANK (1962). Relationship between system failure rate and component failure rates. Boeing Scientific Research Laboratories Document D1-82-0161, to appear in *Technometrics*.

[7] GANTMACHER, F. and KREIN, M. (1950). *Oscillation Matrices and Small Vibrations of Mechanical Systems*. (2nd ed. in Russian). English translation by U. S. Atomic Energy Comm., Office of Technical Information (AEC-tr-4481).

[8] GNEDENKO, B. (1943). Sur la distribution limite du terme maximum d'une série aleatoire. *Ann. of Math.* **44** 423–453.

[9] GUMBEL, E. J. (1958). *Statistics of Extremes*. Columbia Univ. Press.

[10] KARLIN, S. and PROSCHAN, F. (1960). Pólya type distributions of convolutions. *Ann. Math. Statist.* **31** 721–736.

[11] KARLIN, S., PROSCHAN, F. and BARLOW, R. E. (1961). Moment inequalities of Pólya frequency functions. *Pacific J. Math.* **11** 1023–1033.

[12] KARLIN, S. and RUBIN, H. (1956). The theory of decision procedures for distributions with monotone likelihood ratio. *Ann. Math. Statist.* **27** 272–299.

[13] KARLIN, S. (1957). Pólya type distributions, II. *Ann. Math. Statist.* **28** 281–308.

[14] LOÈVE, MICHEL (1960). *Probability Theory*. (2nd ed.) Van Nostrand, Princeton.

[15] PALM, C. (1957). Research on telephone traffic carried by full availability groups. *Tele.* **1**. The Royal Board of Swedish Telecommunications, Stockholm.

[16] PÓLYA, G. and SCHOENBERG, I. J. (1958). Remarks on de la Vallee Poussin means and convex maps. *Pacific J. Math.* **8** 295–334.

[17] PÓLYA, G. and SZEGÖ, G. (1925). *Aufgaben und Lehrsatze aus der Analysis*, **1**. Springer, Berlin.

[18] SAKS, STANISLAW (1937). *Theory of the Integral*. Hafner, New York.

[19] SCHOENBERG, I. J. (1951). On Pólya frequency functions. *J. Analyse Math.* **1** 331–374.

[20] SCHOENBERG, I. J. (1955). On the zeros of the generating functions of multiply positive sequences and functions. *Ann. of Math.* **62** 447–471.

[21] SMITH, WALTER L. (1958). Renewal theory and its ramifications. *J. Roy. Statist. Soc. Ser. B* **20** 243–284.

[22] STEFFENSEN, J. F. (1930). *Some Recent Researches in the Theory of Statistics and Actuarial Science*. Cambridge Univ. Press.

[23] TUKEY, J. W. (1958). A problem of Berkson, and minimum variance orderly estimators. *Ann. Math. Statist.* **29** 588–592.