PROBABILITY DISTRIBUTIONS WITH MONOTONE HAZARD RATE

by 45

SAMUEL EDWARD NEVIUS

B. A., University of Kansas, 1967

A MASTER'S REPORT

submitted in partial fulfillment of the

requirements for the degree

MASTER OF SCIENCE

Department of Statistics
and Computer Science

KANSAS STATE UNIVERSITY
Manhattan, Kansas

1969

Approved by:

A. M. Feyerherm

Major Professor

TABLE OF CONTENTS

1. INTRODUCTION

Suppose X is a random variable representing the service life to failure of a specified material, structure, or device. Such a "failure distribution" represents an attempt to describe mathematically the length of life of the material. On the basis of actual observations of times to failure, it is difficult to distinguish among various nonsymmetrical probability distributions. Differences among the gamma, Weibull, and lognormal distribution functions become significant only in the tails of the distribution, where actual observations are sparse because of limited sample sizes. In order to discriminate among probability functions that cannot be distinguished from each other within the range of actual observation, it is necessary to appeal to a concept that permits differentiation among distribution functions based on physical considerations, namely the hazard rate (Barlow and Proschan, 1965, pp. 9-10).

The hazard rate is defined for a random variable with distribution function F and density f as

$$q(x) = \frac{f(x)}{1 - F(x)}, \quad \text{for } x \text{ such that } F(x) < 1.$$

For a "failure distribution" as described above, $q(x)dx$ represents the conditional probability that an item of age x will fail in the interval $(x, x + dx)$. To see this, let $h = dx$ and note that

$$f(x)h = F(x+h) - F(x) + o(h), \text{ where } \lim_{h\to 0} \frac{o(h)}{h} = 0.$$

Because of this interpretation F will be assumed throughout this paper to be the distribution function of a positive random variable, although for many of the results this is not necessary.

The hazard rate is known by a variety of names in various applications. It is used by actuaries under the name of "force of mortality" to compute mortality tables. Its reciprocal for the normal distribution is known as "Mill's ratio." Papers in extreme value distribution theory call it the intensity function, while in reliability theory both hazard rate and failure rate are used.

Most of the properties discussed in this paper are due to the monotonicity of the hazard rate. The assumption that the life of a unit has increasing hazard rate is usually easy to accept; wear out intuitively would cause older units to have greater chance of failing in the next unit of time. Decreasing hazard rate could conceivably correspond to some physical characteristic of improvement with age, so that older units have less chance of failure in the next unit of time. Theorems given in this paper will explain other cases of decreasing hazard rate, where no process of improvement is involved. Proschan (1963) discusses such a case. Also bounds on survival probabilities based on the assumption of monotone hazard rate will be presented, as well as statistical tests for monotonicity of hazard rate.

2. BASIC PROPERTIES OF THE HAZARD FUNCTION

Let $F(x)$ be the distribution function corresponding to an absolutely continuous random variable X and let $f(x)$ be its probability density function. The hazard rate of X is defined for $F(x) < 1$ by

$$q(x) = \frac{f(x)}{1 - F(x)} \quad .$$

The following relationship holds between $F(x)$ and the hazard function q:

$$1 - F(x) = (1 - F(x_o)) \exp(-\int_{x_o}^{x} q(z)dz),$$

where x_o is an arbitrary value of x. To see this, note that

$$\frac{d}{dx} \log(1 - F(x)) = \frac{-f(x)}{1 - F(x)} .$$

Thus $\log(1 - F(x)) = \int_{x_o}^{x} -f(z)/(1 - F(z))dz + \text{"constant"}$,

where the constant is $\log(1 - F(x_o))$.

Thus $1 - F(x) = (1 - F(x_o)) \exp(-\int_{x_o}^{x} q(z)dz)$.

If there is a lower bound ε to the distribution (i.e., $F(\varepsilon) = 0$), then

$$1 - F(x) = \exp(-\int_{\varepsilon}^{x} q(z)dz) \text{ for } x > \varepsilon.$$

A lower bound $\varepsilon = 0$ will be assumed in this paper unless otherwise specified; this assumption gives

$$1 - F(x) = \exp(-\int_{0}^{x} q(z)dz) \text{ for } x > 0.$$

Also it will be assumed throughout this paper that

$F(0-) = 0$, $F(+\infty) = 1$, and that F is right continuous.

The above property can be used to verify that a constant hazard function

$$q(x) = \lambda \text{ for } x > 0$$

gives rise to an exponential distribution, since

$$F(x) = 1 - \exp(-\int_0^x \lambda dz) = 1 - \exp(-\lambda x),$$

which is the distribution function of an exponential distribution with mean $1/\lambda$. Distributions often assumed as failure laws are listed below, along with their hazard rates (when simple expressions are available). (Barlow and Proschan, 1965, p. 13).

(i) The exponential:

$$f(t) = \lambda e^{-\lambda t}, \quad q(t) = \lambda, \quad t \geq 0$$

(ii) The gamma:

$$f(t) = \frac{\lambda^{\alpha} t^{\alpha-1} e^{-\lambda t}}{\Gamma(\alpha)} \qquad \lambda, \alpha > 0, \ t \geq 0$$

(iii) The Weibull:

$$f(t) = \lambda\alpha t^{\alpha-1} e^{-\lambda t^{\alpha}}, \quad q(t) = \lambda\alpha t^{\alpha-1}, \quad \lambda, \alpha > 0, \ t \geq 0$$

(iv) The modified extreme value distribution:

$$f(t) = \frac{1}{\lambda} \exp\left(\frac{-e^{t}-1}{\lambda} + t\right), \quad q(t) = e^{t}/\lambda, \quad \lambda > 0, \ t \geq 0$$

(v) The truncated normal:

$$f(t) = \frac{1}{a\sigma\sqrt{2\pi}} \exp\left(-\frac{(t-\mu)^2}{2\sigma^2}\right), \qquad \begin{array}{l} \sigma > 0 \\ -\infty < \mu < \infty \\ 0 \leq t < \infty, \end{array}$$

where a is a normalizing constant.

(vi) The log normal:

$$f(t) = \frac{1}{t\sigma\sqrt{2\pi}} \exp\left(-\frac{1}{2\sigma^2}(\log t - \mu)^2\right), \quad \begin{array}{l} -\infty < \mu < \infty \\ \sigma > 0 \\ t \geq 0 \end{array}$$

The exponential has constant hazard rate, but the gamma and Weibull have increasing hazard rate for $\alpha > 1$ and decreasing hazard rate for $\alpha < 1$. For $\alpha > 1$ the hazard rate of the gamma is bounded above by λ, while the hazard rate of the Weibull is unbounded. For $\alpha = 1$, both coincide with the exponential. The modified extreme value distribution and truncated normal have increasing hazard rate. The log normal distribution has a decreasing hazard rate in the long life range; i.e., its hazard rate increases at first and then eventually decreases to zero.

The hazard rate of a discrete distribution $\{p_k\}_{k=0}^{\infty}$ is defined as follows:

$$q(k) = \frac{p_k}{\sum_{j=k}^{\infty} p_j} .$$

In the discrete case, $q(k) \leq 1$. Among typical discrete failure distributions, the binomial and Poisson distributions have increasing hazard rate, the geometric family ($p_k = p(1 - p)^k$, $0 < p < 1$, $k = 0,1,2,\cdots$) has constant hazard rate and the negative binomial distributions

$$\left(p_k = \binom{-\alpha}{k} p^{\alpha} (-q)^k,\ p = 1 - q > 0,\ \alpha \geq 0,\ k = 0,1,2,\cdots\right)$$

have increasing hazard rate for $\alpha > 1$ and decreasing hazard rate for $\alpha < 1$. For $\alpha = 1$, the negative binomial corresponds to the geometric and thus has constant hazard rate. (Barlow and Proschan, 1965, p. 18).

The hazard rate as defined for continuous distributions assumes the existence of the probability density function of the distribution. A more general definition of increasing (decreasing) hazard rate distributions will now be given, one that does not assume the existence of the density function, but that will reduce to increasing (decreasing) $q(t)$ in case the density exists. In this paper the words "increasing" and "decreasing" will not be used in the strict sense; i.e., increasing is used as non-decreasing, decreasing as non-increasing. Thus the exponential distribution is both an increasing hazard rate distribution and a decreasing hazard rate distribution. Also, the symbol IHR will denote an increasing hazard rate distribution, similarly for DHR.

Definition: A nondiscrete distribution F is IHR (DHR) iff

$$\frac{F(t + x) - F(t)}{1 - F(t)}$$

is increasing (decreasing) in t for $x > 0$, $t \geq 0$ such that $F(t) < 1$. (Barlow and Proschan, 1965, p. 23).

Of course, no generality is lost by considering discrete distributions to be IHR (DHR) when $q(k)$ is increasing (decreasing).

The next result will show the equivalence of IHR (DHR) distributions and distributions for which $q(t)$ is increasing (decreasing).

<u>Theorem</u>: Assume F has a density f with $F(0-) = 0$ as usual. Then F is IHR (DHR) iff $q(t)$ is increasing (decreasing).

<u>Proof</u>: Since $q(t) = \lim\limits_{x \to 0} \dfrac{F(t + x) - F(t)}{x(1 - F(t))}$,

it suffices to show q(t) increasing (decreasing) in t implies

$$\frac{F(t + x) - F(t)}{1 - F(t)} \quad \text{increasing (decreasing) in } t.$$

Without loss of generality, consider the case of q(t) increasing. Then for $t_1 \leq t_2$

$$q(t_1) \leq q(t_2)$$

Thus $$\int_0^x q(t_1 + u)du \leq \int_0^x q(t_2 + u)du.$$

Let $v = t_1 + u$, $w = t_2 + u$.

Then

$$\int_{t_1}^{t_1+x} q(v)dv \leq \int_{t_2}^{t_2+x} q(w)dw .$$

This implies

$$\exp\left(-\int_{t_2}^{t_2+x} q(w)dw\right) \leq \exp\left(-\int_{t_1}^{t_1+x} q(v)dv\right) .$$

But $$1 - F(t) = \exp\left(-\int_0^t q(x)dx\right) .$$

So

$$\frac{1 - F(t_2 + x)}{1 - F(t_2)} \leq \frac{1 - F(t_1 + x)}{1 - F(t_1)}$$

Adding one to both sides,

$$\frac{1 - F(t_1)}{1 - F(t_1)} + \frac{F(t_1 + x) - 1}{1 - F(t_1)} \leq \frac{1 - F(t_2)}{1 - F(t_2)} + \frac{F(t_2 + x) - 1}{1 - F(t_2)} ,$$

and rewriting, gives

$$\frac{F(t_1 + x) - F(t_1)}{1 - F(t_1)} \leq \frac{F(t_2 + x) - F(t_2)}{1 - F(t_2)}$$

Q.E.D

(Barlow and Proschan, 1965, p. 23).

Further properties of the hazard function depend on mathematical properties of convex functions, Polya frequency functions, and totally positive functions. The following definitions will be useful:

Definition: Let u be a function defined on an open interval I, and $P = (\xi, u(\xi))$ a point on its graph. A line L passing through P is said to support u at ξ if the graph of u lies entirely above or on L. (This excludes vertical lines.) (Feller, 1966, p. 151).

Definition: The function u is called convex in I if a supporting line exists at each point of I. (The function u is concave if -u is convex). (Feller, 1966, p. 151).

Definition: A function p(x) defined for x in $(-\infty,\infty)$ is a Polya frequency function of order 2 (PF_2) iff $p(x) \geq 0$ for all x and

$$\begin{vmatrix} p(x_1 - y_1) & p(x_1 - y_2) \\ p(x_2 - y_1) & p(x_2 - y_2) \end{vmatrix} \geq 0$$

whenever $-\infty < x_1 \leq x_2 < \infty$ and $-\infty < y_1 \leq y_2 < \infty$. (Barlow and Proschan, 1965, p. 24).

<u>Theorem</u>: A necessary and sufficient condition that f be a PF_2 density is that the ratio

$$\frac{f(t)}{F(t + \Delta) - F(t)}$$

be increasing in t for all Δ. For proof, see Barlow and Proschan (1965, p. 229).

Definition: A function $p(x,y)$ defined for $x \in X$ and $y \in Y$ (X and Y linearly ordered sets) is <u>totally positive of order 2</u> (TP_2) iff $p(x,y) \geq 0$ for all $x \in X$, $y \in Y$ and

$$\begin{vmatrix} p(x_1, y_1) & p(x_1, y_2) \\ p(x_2, y_1) & p(x_2, y_2) \end{vmatrix} \geq 0$$

whenever $x_1 \leq x_2$ and $y_1 \leq y_2$ $(x_1, x_2 \in X, y_1, y_2 \in Y)$ (Barlow and Proschan, 1965, p. 25).

The next theorem gives two necessary and sufficient conditions for a distribution to be IHR (DHR) which will prove useful later.

<u>Theorem</u>: The following statements are equivalent:

(a) F is an IHR distribution. (F is a DHR distribution.)

(b) Log $(1 - F(t))$ is concave for t in $(t|F(t) < 1, t \geq 0)$ (Log $(1 - F(t))$ is convex for t in $(t|F(t) < 1, t \geq 0)$)

(c) $1 - F(t)$ is PF_2. ($1 - F(x + y)$ is TP_2 in x and y for $x + y \geq 0$)

<u>Proof</u>: (a)$\Leftrightarrow$(b). Let $1 - F(t) = e^{-R(t)}$ for some function R (i.e., $R(t) = -\log(1 - F(t))$). Then

$$\frac{F(t+x)-F(t)}{1-F(t)} = \frac{1-F(t)-1+F(t+x)}{1-F(t)} = \frac{(1-F(t)-(1-F(t+x))}{1-F(t)} =$$

$$1 - \frac{1-F(t+x)}{1-F(t)} = 1 - e^{-(R(t+x)-R(t))}.$$

Thus F is IHR iff $1 - e^{-(R(t+x)-R(t))}$ is increasing in t for all $x > 0$, i.e. iff $R(t+x) - R(t)$ is increasing in t for all $x > 0$.
Thus F is IHR iff $R(t)$ is convex iff $-\log(1 - F(t))$ is convex iff $\log(1-F(t))$ is concave. Similarly for DHR.

In the case F has a density, this result is immediate, since

$$q(t) = -\frac{d}{dt}\log(1 - F(t)).$$

(a)$\Leftrightarrow$(c): F is IHR iff for $t_1 \le t_2$ and $x \ge 0$.

$$[1 - F(t_1) - (1 - F(t_1 + x))](1 - F(t_2))$$

$$- [(1 - F(t_2) - (1 - F(t_2 + x))](1 - F(t_1)) \le 0$$

i.e. iff

$$\begin{vmatrix} 1 - F(t_1) - (1 - F(t_1 + x)) & 1 - F(t_2) - (1 - F(t_2 + x)) \\ 1 - F(t_1) & 1 - F(t_2) \end{vmatrix} \le 0$$

Subtracting the second row from the first, we see that this is the condition that $1 - F(t)$ be PF_2. (Barlow and Proschan, 1965, p. 25).

Q.E.D.

In the case of DHR distributions, it is not possible that $q(x)$ is decreasing for all x, since $q(x)$ decreasing at $x = t$ implies the density

$f(x)$ is also decreasing at $x = t$. If the support of F is bounded below, say by zero, then $q(x)$ may be decreasing for $x \geq 0$. Barlow, Marshall, and Proschan (1963, p. 377) show that the support of a DHR distribution F cannot be a finite interval. Thus the support of a DHR distribution must be $[0,\infty)$, where $q(x)$ is decreasing for $x \geq 0$. Another interesting result obtained in this article is that F is IHR (DHR) iff there exists a non-negative convex (concave) increasing function h such that $F(x) = G(h(x))$, where

$$G(x) = 1 - e^{-y},\ y \geq 0.$$

Also if F is IHR and h is a non-negative convex increasing function, not identically constant, then $F(h(x))$ is IHR.

Barlow and Proschan (1965, p. 25) remark that, due to a measurable convex function being continuous in the interior of the region of its definition, F cannot have a jump in the interior of its support if F is IHR or DHR. If F is IHR, F may possess a jump only at the right-hand end of its interval of support (which implies $F(0-) = 0 = F(0)$). If F is DHR, F may possess a jump only at the origin. We may, in fact, show that the continuous part of F in either the IHR or DHR case is absolutely continuous. Recall that a function $f(x)$ is absolutely continuous on a closed interval $[a,b]$ if for all $\varepsilon > 0$ there exists $\eta > 0$ such that if (a_1,b_1), (a_2,b_2), ..., (a_n,b_n) is any finite set of nonoverlapping intervals such that the sum of the lengths of the intervals is less than η, then

$$\sum_{i=1}^{n} |f(a_i) - f(b_i)| < \varepsilon .$$

To show that the continuous part of an IHR distribution function F is absolutely continuous, let $\varepsilon > 0$. Then choose z such that

$$u(z) = -\log(1 - F(z)) < \infty$$

Let $0 \le \alpha_1 < \beta_1 < \alpha_2 < \beta_2 < \cdots < \alpha_m < \beta_m \le z$

be points satisfying

$$\sum_{i=1}^{m} (\beta_i - \alpha_i) < \frac{\varepsilon}{u^+(z)} \quad , \quad \text{where}$$

$$u^+(z) = \lim_{\delta \searrow 0} \frac{u(z + \delta) - u(z)}{\delta} < \infty$$

since U is convex. Then

$$\sum_{i=1}^{m} |U(\beta_i) - U(\alpha_i)| = \sum_{i=1}^{m} \frac{u(\beta_i) - u(\alpha_i)}{\beta_i - \alpha_i} (\beta_i - \alpha_i)$$

$$\le U^+(z) \sum_{i=1}^{m} (\beta_i - \alpha_i) < \varepsilon.$$

Thus U is absolutely convergent on [0,z], as was to be shown. Similarly if F is DHR. (Barlow and Proschan, 1965, p. 26)

Similar results hold in the discrete case. Barlow, Marshall, and Proschan (1963, p. 379) state that

$$\{p_j\}_{j=0}^{\infty} \text{ is IHR iff } \frac{\sum_{j=k+\Delta}^{\infty} p_j}{\sum_{j=k}^{\infty} p_j}$$

is decreasing in k for integer $\Delta \ge 1$. Thus $\{p_j\}_{j=0}^{\infty}$ is IHR iff $\left\{\sum_{j=k}^{\infty} p_j\right\}_{k=0}^{\infty}$

is a PF_2 sequence (where, of course, a PF_2 sequence is simply a PF_2 function defined on a set of integers). This is equivalent to an analog of the condition that $\log(1 - F(x))$ be concave, namely, that there exist two integers α, β, $0 \leq \alpha \leq \beta \leq \infty$ such that $p_j > 0$ iff $\alpha \leq j \leq \beta$ and the polygonal line of vertices $x = k$,

$$y = \log \sum_{j=k}^{\infty} p_j, \quad \alpha \leq k \leq \beta, \text{ is concave.}$$

3. PRESERVATION OF MONOTONE HAZARD RATE

The next property of monotone hazard rate distributions to be considered is their behavior under convolution and convex combination.

Barlow, Marshall, and Proschan (1963, p. 380) prove that if F and G are IHR, then their convolution H, given in the continuous case by

$$H(t) = \int_{-\infty}^{\infty} F(t - x)dG(x), \quad \text{is also IHR.}$$

This result also holds in the discrete case. But the DHR property is not preserved under convolution. For a counter example consider gamma densities f and g with $1/2 \leq \alpha < 1$ (Barlow, Marshall, and Proschan, 1963, p. 380).

Mixtures of DHR distributions are however, DHR. The following result is given in Barlow, Marshall, and Proschan (1963).

<u>Theorem</u>: 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.

A similar theorem given in Proschan (1963), states that convex combinations of DHR distributions are DHR:

<u>Theorem</u>: If $F_i(t)$ is DHR for $i = 1, 2, \ldots, n$, then

$$G(t) = \sum_{i=1}^{n} p_i F_i(t) \text{ is DHR where each } p_i > 0, \quad \sum_{i=1}^{n} p_i = 1.$$

Proof: Suppose $F_i(t)$ has differentiable density $f_i(t)$, $i=1,2,\ldots,n$. Since the density of any DHR distribution must be a decreasing function, we have $-f_i'(t) > 0$ so that by the Schwarz inequality

$$\sum_{i=1}^{n} p_i(1 - F_i(t)) \sum_{i=1}^{n} (-p_i f_i'(t)) \geq \left\{\sum_{i=1}^{n} p_i[(1 - F_i(t))(-f_i'(t))]^{1/2}\right\}^2 .$$

Since $\dfrac{f_i(t)}{1 - F_i(t)}$ is decreasing in t,

$$\frac{d}{dt}\left\{\frac{f_i(t)}{1 - F_i(t)}\right\} = \frac{(1 - F_i(t))f_i'(t) + f_i^2(t)}{[1 - F_i(t)]^2} \leq 0 .$$

Thus $(1 - F_i(t))f_i'(t) \leq -(f_i(t))^2$.

$$\text{Thus} \quad \sum_{i=1}^{n} p_i(1 - F_i(t)) \sum_{i=1}^{n} (-p_i f_i'(t)) \geq \left\{\sum_{i=1}^{n} p_i f_i(t)\right\}^2 .$$

$$\text{i.e.,} \quad \sum_{i=1}^{n} p_i(1 - F_i(t)) \sum_{i=1}^{n} p_i f_i'(t) \leq -\left\{\sum_{i=1}^{n} p_i f_i(t)\right\}^2 .$$

$$\text{This implies} \quad \left(\sum_{i=1}^{n} p_i - \sum_{i=1}^{n} p_i F_i(t)\right) \sum_{i=1}^{n} p_i f_i'(t) \leq -\left\{\sum_{i=1}^{n} p_i f_i(t)\right\}^2,$$

and so $(1 - G(t))g'(t) \leq -g^2(t)$.

But

$$\frac{d}{dt}\left\{\frac{g(t)}{1 - G(t)}\right\} = [1 - G(t)]g'(t) + g^2(t) \leq 0.$$

Thus the ratio $\frac{g(t)}{1 - G(t)}$ is decreasing, which means G is DHR.

If the F_i do not have differentiable densities, Proschan remarks that the same results may be obtained by limiting arguments. (Proschan, 1963, p. 382)

Q.E.D.

However mixtures of IHR distributions are not necessarily IHR. As a counter example, consider two distinct exponentials. A mixture of two distinct exponentials must be DHR by the above theorem. But such a mixture is obviously not exponential. Thus it cannot be IHR, since a constant hazard rate (which must be the case if a distribution is both IHR and DHR) gives rise to an exponential distribution.

Proschan (1963) uses the above results to give a theoretical explanation of decreasing hazard rate in situations where no physical characteristic of improvement with age is present. His specific example concerns the distribution of failure intervals for the air conditioning systems of a fleet of jet airplanes. Using the test for monotone hazard rate developed by Proschan and Pyke (to be presented later in this paper), Proschan concludes that it seems safe to accept the exponential distribution for each plane in his experiment, although to each plane may correspond a different hazard rate. Applying this test to the pooled data, he concludes that the failure distribution for the entire fleet is DHR. Since pooling of failure intervals in this situation may be shown to correspond approximately to forming convex combinations of the individual exponential distributions, the above theorem gives a theoretical explanation of the observed DHR. A number of commonly occurring situations of life distributions with DHR could arise in the same way. For example, a reasonable explanation for the DHR property of semiconductors would be that units within a given production lot exhibit a constant hazard rate since semiconductors do not seem to wear out, but that the hazard rate varies from lot to lot as a result of inescapable manufacturing variability. Thus the life distribution of units for the various lots combined would be DHR.

4. BOUNDS FOR DISTRIBUTIONS WITH MONOTONE HAZARD RATE

This section will be primarily concerned with giving bounds on the quantity $1 - F(t)$ for distributions F such that $F(0^-) = 0$. If F is a failure distribution, then $1 - F(t)$ is the survival probability, i.e. $1 - F(t)$ is the probability of survival until time t.

Many of the results will be improvements on Markov's Inequality, which states that for a probability distribution F such that $F(0^-) = 0$, and for

$$\mu_r = \int_0^\infty x^r dF(x) < \infty \qquad , \quad r,\ t > 0, \text{ that}$$

$$0 \leq 1 - F(t) \leq \mu_r/t^r$$

(This result is incorrectly stated in Barlow and Marshall (1964, p. 1234).) To prove the above form, first note that in the case $r = 1$,

$$1 - F(t) = P(X > t) = \int_t^\infty dF(x) = \int_t^\infty \frac{t}{t}\, dF(x) \leq \frac{1}{t}\int_t^\infty x dF(x) \leq \frac{1}{t}\int_0^\infty x dF(x) = \frac{\mu_1}{t},$$

For $r > 1$, let $Y = X^r$.

Then $P(Y > t^r) \leq \dfrac{\mu_r}{t^r}$, which implies

$$1 - F(t) = P(X > t) = P(Y > t^r) \leq \frac{\mu_r}{t^r}$$

since X is a positive valued random variable. Barlow and Marshall (1964, p. 1234) remark that Markov's inequality is sharp in the sense that without strengthened hypotheses, no tighter bounds can be given. In fact, for each positive r and t there exist distributions which attain equality. Of course, the added hypothesis which will be added here will be that of monotone hazard rate.

The exponential distribution with constant hazard rate, being the boundary distribution between IHR and DHR distributions, provides natural bounds on the survival probability of IHR and DHR distributions. It can be shown that if F is IHR with mean μ_1, then $1 - F(t)$ must cross e^{-t/μ_1} exactly once, and the crossing is necessarily from above (Barlow and Proschan, 1965, p. 26). This implies that $1 - F(t)$ tails off exponentially fast. A more precise statement of this exponential rate of decrease is contained in the following lemma.

<u>Lemma</u>: If F is IHR (DHR), then

$$[1 - F(t)]^{1/t}$$

is decreasing (increasing) in t.

<u>Proof</u>: Suppose F is IHR. Then $\log(1 - F(t))$ is concave in t, so

$$\frac{\log\,[1 - F(t)] - \log\,[1 - F(0^-)]}{t - 0}$$

is decreasing in t. Thus

$$e^{\frac{\log [1 - F(t)]}{t}} = e^{\log [1 - F(t)]^{1/t}}$$

is decreasing in t, which implies $[1 - F(t)]^{1/t}$ is decreasing in t. Similarly for DHR. Barlow and Proschan (1965, p. 27)

Q.E.D.

This lemma also provides a proof of the fact that IHR distributions have finite moments of all orders. To see this, we need the following proposition.

<u>Proposition</u>: If X is a non-negative random variable and if μ_r exists, then

$$\mu_r = r \int_0^\infty x^{r-1} [1 - F(x)]dx.$$

<u>Proof</u>: Result is immediate if X has a density f.

Then

$$r \int_0^\infty x^{r-1} [1 - F(x)]\, dx = r \int_0^\infty x^{r-1} \int_x^\infty f(u)du\, dx = r \int_0^\infty \int_0^u f(u)x^{r-1}\, dx\, du =$$

$$\int_0^\infty u^r f(u)du = \mu_r.$$

If X has no density, recall that if α and f are complex functions of bounded variations on [a,b], and f is also continuous, then

$$\int_a^b f d\alpha = f(b)\alpha(b) - f(a)\alpha(a) - \int_a^b \alpha df$$

(Rudin, 1964, p. 122)

Thus

$$\int_0^T x^r dF(x) = T^r F(T) - \int_0^T F(x) d(x^r) = T^r F(t) + r \int_0^T (1 - F(x))\, x^{r-1} dx - r \int_0^T x^{r-1} dx$$

$$= -T^r\ (1 - F(t)) + r \int_0^T (1 - F(x))\, x^{r-1} dx$$

But

$$T^r\ [1 - F(t)] = T^r \int_T^\infty dF(x) < \int_T^\infty x^r dF(x) \quad \text{for } T > 1 \to 0 \text{ as } T \to \infty$$

Thus

$$\int_0^\infty x^r dF(x) = r \int_0^\infty [1 - F(x)]\ x^{r-1} dx. \quad \text{(Feller, 1966, pp. 148-149)}$$

Q.E.D.

Since if F is IHR, $[1 - F(t)]^{1/t}$ is decreasing in t, $1 - F(x) \leq [1 - F(t)]^{x/t}$ for $x > t$.

Then

$$r \int_t^\infty x^{r-1}\ [1 - F(x)]\ dx \leq r \int_t^\infty x^{r-1}\ [1 - F(t)]^{x/t}\ dx < \infty$$

when $1 - F(t) < 1$ and $r \geq 0$. Hence, IHR distributions have finite moments of all orders. (Barlow and Proschan, 1965, p. 27)

Bounds on survival probabilities depending on percentiles and means will be presented next.

<u>Theorem</u>: If F is IHR and $F(\xi_p) = p$ (i.e., ξ_p is a p^{th} percentile), then

$$1 - F(t) \geq e^{-\alpha t} \qquad t \leq \xi_p$$

$$\leq e^{-\alpha t} \qquad t \geq \xi_p$$

$$\text{where } \alpha = -\frac{\log(1 - p)}{\xi_p}$$

Similarly, if F is DHR and $F(\xi_p) = p$

then

$$1 - F(t) \leq e^{-\alpha t} \qquad t \leq \xi_p$$

$$\geq e^{-\alpha t} \qquad t \geq \xi_p$$

<u>Proof</u>: If $t_1 < t_2$ and F is IHR, then

$$[1 - F(t_1)]^{1/t_1} \geq [1 - F(t_2)]^{1/t_2}$$

So if $t \leq \xi_p$, then

$$[1 - F(t)]^{1/t} \geq (1 - p)^{1/\xi_p}$$

or

$$1 - F(t) \geq (1 - p)^{t/\xi_p} = e^{(t/\xi_p)(\log(1-p))} = e^{-\alpha t}$$

Similarly for the other cases . (Barlow and Proschan, 1965, p. 27)

Q.E.D.

To get a lower bound on $1 - F(t)$ when F is IHR, the following lemma is needed.

<u>Lemma</u>: (Jensen's Inequality) Let F be a probability distribution concentrated on an open interval I with finite mean E(X). If u is a convex function defined on I, then

$$E(u(X)) \geq u(E(X)).$$

<u>Proof</u>: In analytical terms the definition of a convex function u requires that $u(x) \geq u(\xi) + \lambda(x - \xi)$ for all x in I, where $(\xi, u(\xi))$ is a point on the graph of u and λ is the slope of the support line at $(\xi, u(\xi))$. Letting $\xi = E(X)$ and taking expectations, we get

$$E(u(X)) \geq E\,[u(E(X))] + \lambda E[x - E(X)] = u(E(X))..$$

(Feller, 1966, p. 152) Q.E.D.

<u>Theorem</u>: If F is IHR with mean μ_1, then

$$1 - F(t) \geq \begin{cases} e^{-t/\mu_1} & t < \mu_1 \\ 0 & t \geq \mu_1 \end{cases}$$

(The inequality is sharp.)

<u>Proof</u>: Barlow and Proschan (1965, p. 27) state that since an IHR distribution can always be approximated arbitrarily closely by a continuous IHR distribution, it can be assumed WLOG that F is continuous. Lex X be a random variable with distribution F. Then log [1 - F(t)] is concave in t and Jensen's inequality shows that

$$E(-\log(1 - F(X)) \geq -\log(1 - F(\mu_1)) \quad \text{or that}$$

$$E(\log(1 - F(X)) \leq \log[1 - F(\mu_1)].$$

Since F is continuous, F(X) is uniformly distributed on the unit interval (Hogg and Craig, 1965, p. 178). Then

$$P(1 - F(X) \leq x) = P(F(X) \geq 1 - x) = 1 - P(F(X) \leq 1 - x) = 1 - (1-x) = x.$$

Thus 1 - F(X) is uniformly distributed on the unit interval. So

$$E[\log (1 - F(X))] = \int_0^1 \log x \, dx = x \log x \Big|_0^1 - \int_0^1 dx, \text{ integrating by parts.}$$

Thus $E[\log (1 - F(X))] = -1$, which implies

$$\log [1 - F(\mu_1)] \geq -1, \text{ or } 1 - F(\mu_1) \geq e^{-1}.$$

For $t < \mu_1$, $[1 - F(t)]^{1/t} \geq [1 - F(\mu_1)]^{1/\mu_1}$.

Thus

$$1 - F(t) \geq e^{-t/\mu_1} \quad \text{for } t < \mu_1.$$

(Barlow and Proschan, 1965, p. 27)

Q.E.D.

Barlow and Proschan (1965, p. 28) point out that the exponential distribution with mean μ_1 attains the lower bound for $t < \mu_1$, while the degenerate distribution concentrating at μ_1 attains the lower bound for $t \geq \mu_1$. Also, the inequality is actually strict for $0 < t < \mu_1$ unless F is the exponential e^{-t/μ_1}.

The next theorem gives upper bounds for IHR distributions, given the mean μ_1.

<u>Theorem</u>: If F is IHR with mean μ_1, then

$$1 - F(t) \leq \begin{cases} 1, & t \leq \mu_1 \\ e^{-wt}, & t > \mu_1 \end{cases}$$

where w depends on t and satisfies $1 - w\mu_1 = e^{-wt}$.

For proof of this theorem, see Barlow and Proschan (1965, p. 29).

Similar theorems can be given for DHR distributions. The following sharp upper bound on $1 - F(t)$ is given in terms of a single moment, whereas the sharp lower bound on $1 - F(t)$ is zero for DHR distributions.

Theorem: If F is DHR with mean μ_1, then

$$1 - F(t) \leq \begin{cases} e^{-t/\mu_1}, & t \leq \mu_1 \\ \dfrac{\mu_1 e^{-1}}{t}, & t \geq \mu_1 . \end{cases}$$

For proof, see Barlow and Proschan (1965, p. 32).

Various generalizations of the above theorems concerning bounds on survival probabilities for monotone hazard rate distributions have been derived by Barlow and Marshall (1964). However, many of them can be characterized only through solutions of transcendental equations and are therefore inaccessible directly. Barlow and Marshall (1965) give a listing of these results along with tables for those without explicit forms. It should be noted that the conditions under which non-vacuous hypotheses can be made are limited by the inequalities

$$\mu_1^2 \leq \mu_2 \leq 2\mu_1^2 \quad \text{for IHR distributions}$$

and $\mu_2 \leq 2\mu_1^2$ for DHR distributions, where $\mu_i = \int_0^\infty x^i dF(x)$ for $i = 1,2$.

(These inequalities will be proved later in this section.)

Upper and lower bounds for survival probabilities of monotone hazard rate distributions are given for the case μ_r known for some r and for the case μ_1 and μ_2 known. For those cases in which no explicit expression is available, tables are given for some values of the parameters. An example is the upper bound for $1 - F(t)$ in the IHR case with μ_1 and μ_2 known. Table IV of Barlow and Marshall (1965) gives an upper bound for $1 - F(t)$ for the case $\mu_1 = 1$ for values of t from .30 to 1.25 and values of μ_2 from 1.05 to 1.50.

The following four theorems give bounds on the survival probability assuming a variety of conditions on the hazard rate function q(t).

<u>Theorem</u>: If $q(t) \geq \alpha$ for all $t \geq 0$ and $\int_0^\infty xf(x)dx = \mu_1$, then

$$1 - F(t) \leq \begin{cases} e^{-\alpha t} & , t \leq -(1/\alpha)\log(1 - \alpha\mu_1) = t_o \\[2ex] \dfrac{\alpha\mu_1 e^{-\alpha t}}{1 - e^{-\alpha t}} & , t \geq t_o \end{cases}$$

$$1 - F(t) \geq \begin{cases} \alpha\mu_1 - 1 + e^{-\alpha t} & , t \leq t_o \\[2ex] 0 & , t \geq t_o \end{cases}$$

These inequalities are sharp.

For proof, see Barlow and Marshall (1964, p. 1247). They remark that $q(t) \geq \alpha$ implies $\alpha\mu_1 \leq 1$ so t_o is well defined.

<u>Theorem</u>: If F is IHR, $q(t) \leq \beta$ for all $t \geq 0$, and

$\int_0^\infty xf(x)dx = \mu_1$, then

$$1 - F(t) \leq \begin{cases} 1 , & t \leq \mu_1 - 1/\beta \\ w_o , & t \geq \mu_1 - 1/\beta \end{cases}$$

where w_o is the unique solution of

$$\mu_1 = - \frac{t(1 - w)}{\log w} + \frac{w}{\beta} ,$$

$$1 - F(t) \geq \begin{cases} e^{-t/\mu_1} , & t \leq \mu_1 \\ e^{-\beta t + \beta\mu_1 - 1} , & t \geq \mu_1 \end{cases}$$

These inequalities are sharp.

For proof, see Barlow and Marshall (1964, p. 1251).

<u>Theorem</u>: If $q(t) \leq \beta < \infty$ for all $t \geq 0$ and

$$\int_0^\infty xf(x)dx = \mu_1,$$

then

$$1 - F(t) \leq \begin{cases} e^{-\beta z_o} , & t > \mu_1 - 1/\beta \\ 1 , & t \leq \mu_1 - 1/\beta \end{cases}$$

where z_o is the unique solution of $(t - z)e^{-\beta z} = \mu_1 - 1/\beta$

satisfying $0 \leq z_o \leq t$;

$$1 - F(t) \geq e^{-\beta t}$$

These inequalities are sharp.

For proof, see Barlow and Marshall (1964, p. 1250).

<u>Theorem</u>: If F is IHR, $q(t) \geq \alpha$ for all $x \geq 0$ and $\int_0^\infty xf(x)dx = \mu_1$, then

$$1 - F(t) \leq \begin{cases} e^{-\alpha t}, & t \leq -(1/\alpha)\log(1 - \alpha\mu_1) = t_o \\ e^{-yt}, & t \geq t_o \end{cases}$$

where y is determined by

$$\frac{1 - e^{-yt}}{y} = \mu_1;$$

$$1 - F(t) \geq \begin{cases} e^{-t/\mu_1}, & t \leq \mu_1 \\ e^{-(\alpha z + 1)}, & \mu_1 < t < t_o \\ 0, & t \geq t_o \end{cases}$$

where z is determined by $1 - \alpha\mu_1 = [1 - \alpha(t-z)]\, e^{-\alpha z}$.

These inequalities are sharp.

For proof, see Barlow and Marshall (1964, p. 1248).

The next theorem will present bounds on percentiles in terms of the mean.

Theorem: Assume F is IHR. If $p \leq 1 - e^{-1}$, then

$$[-\log(1 - p)]\,\mu_1 \leq \xi_p \leq \left(-\frac{\log(1 - p)}{p}\right)\mu_1;$$

if $p \geq 1 - e^{-1}$, then

$$\mu_1 \leq \xi_p \leq \left(-\frac{\log(1 - p)}{p}\right)\mu_1,$$

where $\xi_p = \sup\{t \mid F(t) \leq p\}$.

The inequalities are sharp.

Proof: To obtain the upper bound, use the fact that $[1 - F(t)]^{1/t}$ is decreasing in t since F is IHR.

Then for $x < \xi_p$,

$$1 - F(x) \geq [1 - F(\xi_p)]^{x/\xi_p} .$$

Then

$$\mu_1 = \int_0^\infty [1 - F(x)]dx \geq \int_0^{\xi_p} [1 - F(\xi_p)]^{x/\xi_p} dx = \int_0^{\xi_p} (1 - p)^{x/\xi_p} dx =$$

$$\left.\frac{(1 - p)^{x/\xi_p}}{\frac{1}{\xi_p}\log(1 - p)}\right|_0^{\xi_p} = \frac{\xi_p(1 - p)}{\log(1-p)} - \frac{\xi_p}{\log(1-p)} .$$

Thus

$$\xi_p \leq \mu_1 \frac{-\log(1 - p)}{p} \quad \text{for all } 1 > p > 0.$$

To obtain the lower bound, for $p \leq 1 - e^{-1}$, note that if

$\xi_p < \mu_1$, then $1 - p = 1 - F(\xi_p) \geq e^{-\xi_p/\mu_1}$

(This is the lower bound for an IHR distribution with mean μ_1.) This implies

$$\frac{-\xi_p}{\mu_1} \leq \log(1 - p), \text{ or } \quad \xi_p \geq \mu_1 \, [-\log(1 - p)].$$

And if

$$\xi_p \geq \mu_1, \quad 1 - F(\xi_p) = 1 - p \geq e^{-1} \geq e^{-\xi_p/\mu_1} \quad .$$

So

$$\log(1 - p) \geq \frac{-\xi_p}{\mu_1} \quad , \text{ or } \xi_p \geq \mu_1 \, [-\log(1 - p)] \text{ where } p \leq 1 - e^{-1}.$$

In the case $p \geq 1 - e^{-1}$, assume $\xi_p < \mu_1$.

Then

$$1 - F(\mu_1) \geq e^{-1} \geq 1 - p = 1 - F(\xi_p),$$

so

$$\xi_p \geq \mu_1, \text{ which is a contradiction.}$$

Thus

$$\mu_1 \leq \xi_p \ .$$

Barlow and Proschan remark that IHR distributions with the prescribed percentiles attaining the bounds can be easily constructed. (Barlow and

Proschan, 1965, p. 30)

Q.E.D.

This theorem also gives bounds on the mean of IHR distributions in terms of percentiles. For example, if M is the median, then

$$\frac{M}{2 \log 2} \leq \mu_1 \leq \frac{M}{\log 2} \ .$$

The final result of this section gives some useful moment inequalities such as those stated earlier to limit the conditions under which non-vacuous hypotheses can be made (i.e. $\mu_1^2 \leq \mu_2 \leq 2\mu_1^2$ for IHR distributions and $\mu_2 \geq 2\mu_1^2$ for DHR distributions).

<u>Lemma</u>: If

(a) F is IHR with mean μ_1 and $1 - G(x) = e^{-x/\mu_1}$,

(b) $\phi(x)$ is increasing (decreasing),

then

$$\int_0^\infty \phi(x)\ [1 - F(x)]\ dx \underset{(\geq)}{\leq} \int_0^\infty \phi(x)\ [1 - G(x)]\ dx \ .$$

A similar theorem is true for DHR distributions with all inequalities reversed.

<u>Proof</u>: Suppose ϕ is increasing and F is not identically equal to G. Since F is IHR, $1 - F(x)$ crosses $1 - G(x)$ exactly once from above at, say, t_o (i.e., $1 - F(t_o) = 1 - G(t_o)$].

Then

$$\int_0^\infty \phi(x)\ [1 - F(x)]\ dx - \int_0^\infty \phi(x)\ [1 - G(x)]\ dx$$

$$= \int_0^\infty [\phi(x) - \phi(t_o)]\ [(1 - F(x)) - (1 - G(x))]\ dx \leq 0$$

since

$$\mu_1 = \int_0^\infty (1 - F(x))dx = \int_0^\infty (1 - G(x))dx.$$

To prove for ϕ decreasing, replace ϕ by $-\phi$.

(Barlow and Proschan, 1965, p. 32)

Q.E.D.

<u>Theorem</u>: If F is IHR (DHR) with r^{th} moment μ_r, then

$$\mu_r \underset{(\geq)}{\leq} \Gamma(r + 1)\mu_1^r \quad , \quad r \geq 1$$

$$\underset{(\leq)}{\geq} \Gamma(r + 1)\mu_1^r \quad , \quad 0 \leq r \leq 1$$

<u>Proof</u>: Let $\phi(x) = x^{r-1}$ in lemma. Then ϕ is increasing for $r \geq 1$,

so

$$\mu_r = r \int_0^\infty x^{r-1} [1 - F(x)]\, dx \leq r \int_0^\infty x^{r-1} [1 - G(x)]\, dx.$$

For $0 \leq r \leq 1$, $\phi(x)$ is decreasing so inequality is reversed.

Thus it suffices to note that

$$r \int_0^\infty x^{r-1} [1 - G(x)]\, dx = r \int_0^\infty x^{r-1} e^{-x/\mu_1}\, dx = r\Gamma(r)\mu_1^r$$

$$= \Gamma(r + 1)\mu_1^r \text{ since } 1 = \int_0^\infty \frac{x^{\alpha-1}}{\Gamma(\alpha)\beta^\alpha} e^{-x/\beta} dx \quad \text{for} \quad \alpha,\beta > 0 .$$

(Barlow and Proschan, 1965, p. 33)

Q.E.D.

Thus for an IHR distribution,

$$\mu_2 \leq \Gamma(3)\mu_1^2 = 2\mu_1^2,$$

and for a DHR distribution

$$\mu_2 \geq 2\mu_1^2,$$

as stated earlier.

Also

$$\sigma^2 \leq \mu_1^2 \text{ in a IHR distribution since}$$

$$\sigma^2 = \mu_2 - \mu_1^2 \leq 2\mu_1^2 - \mu_1^2 = \mu_1^2 .$$

Thus the coefficient of variation σ/μ_1 is less than or equal to one. The inequality is reversed for DHR distributions.

A related topic which will not be discussed in detail here is that of tolerance limits for the class of distributions with monotone hazard rate. A lower tolerance limit is a function L of sample data

$$\underline{X} = (X_1, X_2, \ldots, X_n) \text{ such that } P_F\ [1 - F(L(\underline{X})) \geq 1 - q] \geq 1 - \alpha,$$

where $X_1, X_2, \ldots, X_n$ comprise a random sample from the distribution F. Then $1 - q$ is the population coverage for the interval $[L(\underline{X}), \infty)$ and $1 - \alpha$ is the confidence coefficient. An upper tolerance limit is a function $U(\underline{X})$ such that $P_F[F(U(\underline{X})) \geq q] \geq 1 - \alpha$.

Distribution-free tolerance limits exist based on order statistics, but these have the unfortunate disadvantage that for given α, q, k, there is a minimum sample size $N(\alpha, q, k)$ such that $P_F[1 - F(Y_k) \geq 1 - q] \geq 1 - \alpha$

is true only if $N \geq N(\alpha, q, k)$, where Y_k is the k^{th} order statistic. Hanson and Koopmans (1964) obtain upper tolerance limits of the form $Y_{N-k-j} + b(Y_{N-k} - Y_{N-k-j})$ for IHR distributions which are valid for all $N \geq 2$ and all $0 < 1 - q$, $1 - \alpha < 1$. Barlow and Proschan (1966) obtain upper and lower confidence limits for IHR and DHR distributions, along with confidence intervals for means and percentiles. The tolerance limits obtained by Barlow and Proschan are conservative and are of greatest value when the sample size is small enough so that the distribution-free tolerance limits do not exist. For very large sample size the distribution-free tolerance limit is close to the percentile providing the covering desired, whereas the conservative limit they obtain is not.

5. TESTS FOR MONOTONE HAZARD RATE

This paper has thus far been concerned exclusively with probabilistic arguments concerning monotone hazard rate probability distributions. A related statistical problem is that of testing if a sample comes from a population with monotone hazard rate. Perhaps the best known such test is a rank test based on normalized spacings derived by Proschan and Pyke (1967). Bickel and Doksum (1967) and Nadler and Eilbott (1968) show, however, that a uniform conditional test is asymptotically superior to the Proschan-Pyke test, while Barlow (1968) uses results obtained by Marshall and Proschan (1965) to derive a likelihood ratio test, concentrating primarily on small sample results. Only the tests derived by Proschan-Pyke and Nadler-Eilbott will be discussed here.

Suppose that $F(x)$ is a continuous distribution function satisfying $F(0) = 0$. A random sample of size n from $F(x)$ is denoted by $X_1, X_2, \ldots, X_n$, and the order statistics formed from such a sample are written as $X_{1,n} \leq X_{2,n} \leq \ldots \leq X_{n,n}$. The normalized sample spacings $D_{1,n}, D_{2,n}, \ldots, D_{n,n}$ are defined by $D_{i,n} = (n - i + 1)(X_{i,n} - X_{i-1,n})$, using the convention $X_{o,n} = 0$. Only the problem of testing for increasing hazard rate will be discussed here; the case of decreasing hazard rate can be treated by analogy. The hypothesis to be tested is

$$H_o\text{:}\quad F(x) = 1 - e^{-\lambda x} \text{ for some (unknown) } \lambda > 0$$

against the alternative

$$H_1\text{:}\quad F(x) \text{ is strictly IHR.}$$

Proschan and Pyke (1967) use a rank test based on the normalized spacings $D_{1,n}, D_{2,n}, \ldots, D_{n,n}$. The test statistic is as follows:

$$V_n = \sum_{1 \leq i < j \leq n} V_{ij}$$

where

$$V_{ij} = \begin{cases} 1, & \text{if } D_{i,n} \geq D_{j,n} \\ 0 & \text{otherwise} \end{cases}$$

The null hypothesis is rejected at the α level of significance if $V_n > v_{n,\alpha}$, where $v_{n,\alpha}$ is determined so that $P[V_n > v_{n,\alpha} | H_o] = \alpha$. To give a heuristic justification of this procedure, observe that under the null hypothesis that the normalized spacings $D_{1,n}, D_{2,n}, \ldots, D_{n,n}$ are independent and identically distributed with density $\lambda e^{-\lambda t}$ for some $\lambda > 0$, so that $P[V_{ij} = 1 | H_o] = \frac{1}{2}$ for $1 \leq i < j \leq n$ (Proschan and Pyke, 1967, p. 295). Also $D_{i,n}$ is stochastically larger than $D_{j,n}$ under H_1 when $i<j$; thus

$P[V_{ij} = 1|H_1] > \frac{1}{2}$ for these values of i and j. Each V_{ij} and, therefore, V_n tend to be larger under H_1, so that rejection of the null hypothesis occurs for large values of V_n.

The authors show that this test is unbiased

(i.e., $P(V_n \geq v_{n,\alpha} \mid H_1) \geq \alpha$ for $0 < \alpha \leq 1$, $n = 2, 3, \ldots$),

give references to tables of $P[V_n \leq k]$ for $n \leq 10$, and prove that V_n, suitably normalized, is asymptotically normal under both the null hypothesis and a large class of alternatives. They show that the exact distribution of V_n is independent of λ and has mean $\frac{n(n-1)}{4}$ and variance $\frac{(2n+5)(n-1)n}{72}$.

Examples of the use of this test in an actual industrial situation can be found in Proschan (1963). Here use is made of the asymptotic normality of the test statistic to compute the rejection region. Also it can be seen from this article that the same procedure as outlined above can be used to test for decreasing hazard rate, except that the null hypothesis is rejected for <u>small</u> values of the test statistic.

Nadler and Eilbott remark that despite the several nice properties of the Proschan-Pyke test, it suffers from the shortcoming that it does not fully utilize the distributional properties of the $D_{i,n}$ under the null hypothesis (Nadler and Eilbott, 1968, p. 7). Since the independence of $X_{1,n}$ and $X_{2,n} - X_{1,n}$ characterizes the exponential distribution, V_n uses the properties of the null distribution of F only insofar as it requires the normalized sample spacings to be independent and identically distributed. They improve upon the power of V_n by a procedure which uses the fact that under the null hypothesis the $D_{i,n}$ are exponentially distributed as well as the property that they comprise a random sample. The test statistic to be used is the following:

$$E_n = \frac{\sum_{i=1}^{n} (i - 1)D_{n-i+1,n}}{\sum_{i=1}^{n} D_{i,n}} = 2\,\frac{\sum_{i=1}^{n} (n - i)X_{i,n}}{\sum_{i=1}^{n} X_i} ,$$

where H_o is rejected for large values of E_n (as before the alternative hypothesis is IHR). The heuristic justification given for considering this statistic is that the fact that under H_1 $D_{i,n}$ tends to be larger than $D_{j,n}$ for $i < j$ should be reflected by a positive slope in the linear regression of the $D_{n-i+1,n}$ on the values $i - 1$. Under H_o, this regression has zero slope. To make the statistic scale invariant, the slope is divided by the average of the $D_{i,n}$, giving $\frac{12E_n - 6(n - 1)}{n^2 - 1}$. The statistic E_n is discussed for simplicity (Nadler and Eilbott, 1968, p. 8).

Theorem 1 and its corollaries in Nadler and Eilbott (1968) provide the critical values of E_n. They are as follows:

Theorem 1: Let $X_1, \ldots, X_n$ be a random sample from some exponential population. Then

$$Y_n = \frac{\sum_{i=1}^{n} (i - 1)X_i}{n\,\bar{X}}$$

is distributed as the sum of $n - 1$ independent uniform random variables.

For proof, see Nadler and Eilbott (1968, p. 10).

Corollary 1: Under H_o the statistic E_n has the same distribution as Y_n.

For proof, see Nadler and Eilbott (1968, p. 10).

Corollary 2: Under H_o the limiting distribution of $\sqrt{n-1}\,(E_n/(n-1) - 1/2)$ is a normal with mean zero and variance $\frac{1}{12}$. (Nadler and Eilbott, 1968, p. 10)

In practice the limiting normal distribution is approached so rapidly that the distribution of E_n is well approximated by a normal distribution with mean $\frac{n-1}{2}$ and variance $\frac{n-1}{12}$ using sample sizes as small as n = 10. The critical value of a size-α test of H_o vs. H_1 is thus $\frac{(n-1)}{2} + z_\alpha\sqrt{\frac{n-1}{12}}$, where z_α is the $1 - \alpha$ percentile of the standard normal.

Nadler and Eilbott present Monto Carlo studies which suggest that their procedure is more powerful than the Proschan-Pyke procedure. They estimated the power of the two tests for sample sizes of n = 15 and n = 50 under gamma and Weibull (IHR) alternatives. In every case in which a difference between the power functions was found significant at the .05 level, the estimated power of the E_n criterion was greater than that of the V_n criterion. It was also calculated that with large samples a test using the V_n criterion requires 33% more observations to achieve the power attained by the Nadler-Eilbott procedure (in the event that F(x) is actually a gamma or Weibull IHR distribution). The E_n test also has the properties of unbiasedness and consistency possessed by the Proschan-Pyke test and is asymptotically normal for a wide range of IHR alternatives.

ACKNOWLEDGMENTS

I am indebted to Dr. A. M. Feyerherm for his help in the preparation of this report. I am also indebted to Dr. W. J. Conover for originally suggesting this research area as a possible source for a master's report topic and to Miss Anna Bruce for her help in typing the rough draft.

BIBLIOGRAPHY

Barlow, Richard E. (1968). Likelihood ratio tests for restricted families of probability distributions. Ann. of Math. Statist. 39 547.

Barlow, Richard E., and Marshall, Albert W. (1964). Bounds for distributions with monotone hazard rate, I and II. Ann. of Math. Statist. 35 1234 and 1258.

Barlow, Richard E., and Marshall, Albert W. (1965). Tables of bounds for distributions with monotone hazard rate. Journal of ASA 60 872.

Barlow, Richard E., Marshall, Albert W., and Proschan, Frank. (1963). Properties of Probability distributions with monotone hazard rate. Ann. of Math. Statist. 34 375.

Barlow, Richard E., and Proschan, Frank (1965). Mathematical Theory of Reliability. John Wiley, New York.

Barlow, Richard E., and Proschan, Frank. (1966). Tolerance and confidence limits for classes of distributions based on failure rate. Ann. of Math. Statist. 37 1593.

Bickel, Peter J., and Doksum, Kjell. (1967). Tests for monotone failure rate based on normalized spacings. Operations Research Center Report ORC 67-37, University of California, Berkeley.

Feller, William. (1966) An Introduction to Probability Theory and Its Applications, vol 2. John Wiley, New York.

Hanson, D. L., and Koopmans, L. H. (1964). Tolerance limits for the class of distributions with increasing hazard rates. Ann. of Math. Statist. 37 816.

Hogg, Robert V., and Craig, Allen T. (1965). Introduction to Mathematical Statistics, Second Edition. The Macmillan Co., New York.

Marshall, Albert W., and Proschan, Frank. (1965). Maximum likelihood estimation for distributions with monotone failure rate. Ann. of Math. Statist. 36 69.

Nadler, Jack, and Eilbott, Joan. (1968). Testing for monotone failure rates. Unpublished technical report at Bell Telephone Laboratories, Inc., Whippany, New Jersey. (Submitted to the Journal of ASA, 1968).

Parzen, Emanuel. (1962). Stochastic Processes. Holden-Day, San Francisco.

Proschan, Frank. (1963). Theoretical explanation of observed decreasing failure rate. Technometrics 5 375.

Proschan, Frank, and Pyke, Ronald. (1967). Tests for monotone failure rate. Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability. University of California Press, Berkeley.

Rudin, Walter, (1964). Principles of Mathematical Analysis, Second Edition. McGraw-Hill, New York.

PROBABILITY DISTRIBUTIONS WITH MONOTONE HAZARD RATE

by

SAMUEL EDWARD NEVIUS

B. A., University of Kansas, 1967

AN ABSTRACT OF A MASTER'S REPORT

submitted in partial fulfillment of the

requirements for the degree

MASTER OF SCIENCE

Department of Statistics
and Computer Science

KANSAS STATE UNIVERSITY
Manhattan, Kansas

1969

Suppose X is a random variable representing the service life to failure of a specified material, structure, or device. Such a "failure distribution" represents an attempt to describe mathematically the length of life of the material. If X is a continuous random variable with distribution function F and density f, then the hazard rate q(x) of X is defined as

$$q(x) - \frac{f(x)}{1 - F(x)}, \quad \text{for } x \text{ such that } F(x) < 1.$$

Thus q(x)dx represents the conditional probability that an item of age x will fail in the interval (x, x + dx).

It has been found that the hazard rate of many failure distributions is either monotone increasing or monotone decreasing. This paper presents basic properties of the hazard function, and these properties are used to obtain results applicable to probability distributions with monotone hazard rate. The preservation of monotone hazard rate under convolution and convex combination is examined and bounds for the survival probability 1 - F(t) are obtained under the assumption of monotone hazard rate. Finally, two statistical hypothesis tests are presented to test the null hypothesis of constant hazard rate versus the alternative hypothesis of increasing (or decreasing) hazard rate.