Distributions with Monotone Failure Rate

by

Richard E. Barlow\textsuperscript{1} and Frank Proschan\textsuperscript{2}

The Florida State University
Department of Statistics
Tallahassee, Florida 32306

November, 1981

FSU Statistics Report M-597
APOS R Technical Report No. 81-129

\textsuperscript{1}Department of Industrial Engineering and Operations Research, University of California, Berkeley. Research supported by Grant APOS R-81-0122.

\textsuperscript{2}Research supported by the Air Force Office of Scientific Research under Grant APOS R-81-0038. Reproduction in whole or in part is permitted for any purpose of the United States Government.

Key Words: Monotone failure rate, increasing failure rate, decreasing failure rate, IFR, DFR, coherent structures, bounds, reliability, failure rate, increasing failure rate average, IFRA, survival, life distribution, total positivity, mixtures of distributions, convolutions.
ABSTRACT

Distributions with Monotone Failure Rate

by

Richard E. Barlow and Frank Proschan
University of California at Berkeley and Florida State University

This survey of the class of life distributions possessing monotone failure rate is intended for publication in a forthcoming Handbook of Reliability, edited by Igor Ushakov, University of Moscow, and colleagues. Contributions are being made by mathematicians and statisticians from different countries on various aspects of reliability theory and practice.
DISTRIBUTIONS WITH MONOTONE FAILURE RATE

by

Richard E. Barlow and Frank Proschan

1. DEFINITIONS AND CHARACTERIZATIONS

Most statistical analyses are conducted in terms of the probability
density \( f(t) \) and/or the probability distribution function \( F(t) \). In
reliability analysis, however, the failure rate function plays the key role.
The failure rate function \( r(t) \) is defined by:

\[
r(t) = \frac{f(t)}{\bar{F}(t)},
\]

(1)

where \( \bar{F}(t) \equiv 1 - F(t) \) is the survival probability or reliability. Since
reliability analysis is concerned with lifelength, the argument \( t \) is taken
nonnegative throughout. Also, in (1) it is assumed that \( \bar{F}(t) > 0 \).

Equation (1) yields a fundamental identity. Integrating both sides
between \( 0 \) and \( x \) and then exponentiating, we obtain:

\[
\frac{x}{\int_0^x r(t)\,dt} \quad \bar{F}(x) \equiv e
\]

(2)

for \( 0 \leq x < \infty \). Thus from a knowledge of the failure rate function, we may
deduce the survival probability.

In the case that the density exists for all values of \( t \geq 0 \), we say
that the distribution \( F \) has increasing failure rate if \( r(t) \) is monotonic
increasing in \( t \geq 0 \); we write \( F \) is IFR. Note that as a consequence of
(2), \( -\log \bar{F}(t) = \int_0^x r(t)\,dt \) is convex. This property permits us to extend
the notion of \( F \) IFR to the case in which the density fails to exist for all
values of $t \geq 0$; in general, we say that the distribution $F$ is IFR if $-\log F$ is convex.

In dual fashion, we say that $F$ has decreasing failure rate if $r(t)$ is monotonic decreasing for $t \geq 0$; we write $F$ is DFR. More generally, we say that $F$ is DFR if $\log F$ is convex.

A useful characterization of IFR distributions may readily be derived. It follows from $-\log F$ convex that $\frac{F(t + x)}{F(t)}$ is decreasing in $t \geq 0$ for fixed $x > 0$; the interpretation is that the reliability of a unit of initial age $t$ for a mission of duration $x$ is decreasing in $t$. In a dual fashion, $F$ is DFR if $\frac{F(t + x)}{F(t)}$ is increasing in $t \geq 0$ for fixed $x > 0$.

A useful property holds for distributions with monotone failure rate.

1.1 Theorem

Let $F$ be IFR. Let $F(z) < 1$. Then $F$ is absolutely continuous on $(-\infty, z)$; that is $F$ has a probability density on $(-\infty, z)$.

Proof:

Let $R(z) = -\log F(z)$ denote the hazard function. Let $\varepsilon > 0$ and

$\alpha_1 < \beta_1 < \alpha_2 < \beta_2 < \ldots < \alpha_m < \beta_m \leq z$ satisfying $\sum_{i=1}^{m} (\beta_i - \alpha_i) < \varepsilon/r^+(z)$,

where $r^+(z) = \lim_{\delta \to 0} [R(z + \delta) - R(z)]/\delta$ exists finitely since $R \equiv -\log F$ is convex. Then

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

$$= \sum_{i=1}^{m} \left[ |R(\beta_i) - R(\alpha_i)|/(\beta_i - \alpha_i) \right] r^+(z) \sum_{i=1}^{m} (\beta_i - \alpha_i) \leq \varepsilon.$$
Thus \( R \) is absolutely continuous on \((-\infty, z)\), and the result follows.

Note, however, that an IFR distribution may have a jump at the right-hand endpoint of its interval of support, if finite.

A dual result exists for \( F \) DFR; in this case \( F \) may have a jump at the left-hand endpoint of its interval of support.
2. EXAMPLES OF MONOTONE FAILURE RATE

Mechanical parts, subject to wearout from continued use, generally have IFR life distributions.

DFR life distributions occur in practice in several alternative fashions. Certain metals become tougher with continued usage. As a consequence, their life distributions are DFR. DFR life distributions describe also the life-lengths of businesses, television shows, telephone calls, and other assorted objects. A third class of DFR distributions arises in an interesting fashion, namely as a mixture of exponential distributions:

\[ F(t) = \int g_\lambda(t) d\mu(\lambda), \]

where \( g_\lambda(t) = 1 - e^{-\lambda t} \) and \( \mu \) is a probability measure over positive values of \( \lambda \). To verify this claim, apply Schwarz's inequality to obtain

\[ \int \tilde{g}_\lambda d\mu(\lambda) \int (-g'_\lambda) d\mu(\lambda) \geq \left\{ \int [\tilde{g}_\lambda(-g'_\lambda)]^2 d\mu(\lambda) \right\}^2. \]

Since \( g_\lambda / \tilde{g}_\lambda \) is constant, \( \tilde{g}_\lambda(-g'_\lambda) = g^2_\lambda \). Thus

\[ \left\{ \int [\tilde{g}_\lambda(-g'_\lambda)]^2 d\mu(\lambda) \right\}^2 = \left\{ \int g^2_\lambda d\mu(\lambda) \right\}^2, \]

implying \( \int \tilde{g}_\lambda d\mu(\lambda) \int (-g'_\lambda) d\mu(\lambda) \geq \left\{ \int g_\lambda d\mu(\lambda) \right\}^2 \), i.e., \( F(-f') \geq f^2 \). We conclude that \( F \) is DFR.
3. PRESERVATION OF MONOTONE FAILURE RATE UNDER RELIABILITY OPERATIONS

An interesting practical as well as theoretical question is the following: Under which reliability operations is the class of IFR (DFR) distributions closed? For example, is the convolution of IFR distributions an IFR distribution?

Convolution:

To answer this question we need to introduce some concepts from total positivity.

3.1 Definition

A function $h(x), -\infty < x < \infty$, is a Pólya frequency function of order 2 ($PF_2$) if

(a) $h(x) \geq 0$ for $-\infty < x < \infty$, and

(b) \[
\begin{vmatrix}
  h(x_1 - y_1) & h(x_1 - y_2) \\
  h(x_2 - y_1) & h(x_2 - y_2)
\end{vmatrix} \geq 0
\]  \hspace{1cm} (4)

for all $-\infty < x_1 < x_2 < \infty$ and $-\infty < y_1 < y_2 < \infty$, or equivalently,

(b') log $h(x)$ is concave on $(-\infty, \infty)$, or equivalently,

(b'') for fixed $\Delta > 0$, $h(x + \Delta)/h(x)$ is decreasing in $x$ for $a \leq x \leq b$,

where $a = \inf_{h(y) > 0} y$, $b = \sup_{h(y) > 0} y$.

Note that a Pólya frequency function $f$ of order 2 is not necessarily a probability frequency function, in that $\int_{-\infty}^{\infty} f(x)dx$ need not be 1, nor even finite.

Next we show that an IFR survival probability is $PF_2$. 
3.2 Theorem

Let \( F \) be a distribution function. Then \( F \) is IFR if and only if \( \bar{F} \) is PF.  

Proof:

By definition, \( F \) is IFR if and only if \( \frac{\bar{F}(t + \Delta)}{\bar{F}(t)} \) is decreasing in \( -\infty < t < \infty \) for each \( \Delta > 0 \). Thus by Definition 3.1 (b''), \( F \) is IFR if and only if \( \bar{F} \) is PF.  

Now we may prove that the sum of independent IFR random variables is IFR.

3.3 Theorem

Let \( F_1 \) and \( F_2 \) be IFR. Then their convolution \( F(t) = \int_0^t F_1(t - x)dF_2(x) \) is IFR.

Proof:

First assume \( F_1 \) has density \( f_1 \), \( F_2 \) has density \( f_2 \). For \( t_1 < t_2 \), \( u_1 < u_2 \), write  

\[
D = \begin{vmatrix} \bar{F}(t_1 - u_1) & \bar{F}(t_1 - u_2) \\ \bar{F}(t_2 - u_1) & \bar{F}(t_2 - u_2) \end{vmatrix}
= \begin{vmatrix} \int \bar{F}_1(t_1 - s)f_2(s - u_1)ds & \int \bar{F}_1(t_1 - s)f_2(s - u_2)ds \\ \int \bar{F}_1(t_2 - s)f_2(s - u_1)ds & \int \bar{F}_1(t_2 - s)f_2(s - u_2)ds \end{vmatrix}
= \int \int_{s_1 < s_2} \begin{vmatrix} \bar{F}_1(t_1 - s_1) & \bar{F}_1(t_1 - s_2) \\ \bar{F}_1(t_2 - s_1) & \bar{F}_1(t_2 - s_2) \end{vmatrix} f_2(s_1 - u_1) f_2(s_1 - u_2) ds_2 ds_1
\]
by the Basic Composition Formula\(^1\) (Karlin, 1968, page 17). Throughout the domain of integration is \((\sim, \omega)\). Integrating the inner integral by parts, we obtain

\[
D = \int \int_{s_1 < s_2} \begin{vmatrix} \overline{F}_1(t_1 - s_1) & f_1(t_1 - s_2) & f_2(s_1 - u_1) & f_2(s_1 - u_2) \\ F_1(t_2 - s_1) & f_1(t_2 - s_2) & \overline{F}_2(s_2 - u_1) & \overline{F}_2(s_2 - u_2) \end{vmatrix} ds_2 ds_1.
\]

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

\[
\frac{f_1(t_2 - s_2)}{\overline{F}_1(t_2 - s_2)} \overline{F}_1(t_2 - s_1) - \frac{f_1(t_1 - s_2)}{\overline{F}_1(t_1 - s_2)} \overline{F}_1(t_1 - s_1)
\]

assuming nonzero denominators. But

\[
\frac{f_1(t_2 - s_2)}{\overline{F}_1(t_2 - s_2)} > \frac{f_1(t_1 - s_2)}{\overline{F}_1(t_1 - s_2)}
\]

by hypothesis, while

\[
\frac{\overline{F}_1(t_2 - s_2)}{\overline{F}_1(t_2 - s_1)} > \frac{\overline{F}_1(t_1 - s_2)}{\overline{F}_1(t_1 - s_1)}
\]

---

\(^1\)Basic Composition Formula. Let \(w(x,z) = \int u(x,y)v(y,z)d\sigma(y)\) converge absolutely, where \(d\sigma(y)\) is a sigma-finite measure. Then

\[
\begin{vmatrix} w(x_1, z_1) & \ldots & w(x_1, z_n) \\
\vdots & \ddots & \vdots \\
w(x_n, z_1) & \ldots & w(x_n, z_n) \end{vmatrix} = \int \int \begin{vmatrix} u(x_1, y_1) & \ldots & u(x_1, y_n) \\
\vdots & \ddots & \vdots \\
u(x_n, y_1) & \ldots & u(x_n, y_n) \end{vmatrix} \\
\begin{vmatrix} v(y_1, z_1) & \ldots & v(y_1, z_n) \\
\vdots & \ddots & \vdots \\
v(y_n, z_1) & \ldots & v(y_n, z_n) \end{vmatrix} d\sigma(y_1) \ldots d\sigma(y_n).
\]

by Theorem 3.2. Thus the first determinant is nonnegative. A similar argument
holds for the second determinant, so that $D \geq 0$. But by Theorem 3.2, this
implies that $F$ is IFR.

If $F_1$ or $F_2$ do not have a density at the right-hand endpoint(s), the
theorem may be proved in a similar fashion using limiting arguments.||

It is easy to verify that the convolution of DFR distributions is not
DFR. To see this, simply convolve a gamma distribution of order $\alpha \left( \frac{1}{2} < \alpha < 1 \right)$,
a DFR distribution, with itself. The convolution is a gamma distribution of
order $2\alpha (2\alpha > 1)$, a strictly IFR distribution.

Mixture:

Next we consider mixtures of distributions. We have already seen in
Section 2 that a mixture of exponential distributions is DFR. Actually, a
stronger result can be proved using a slight modification of the argument
proving the exponential mixture result:

3.4 Theorem

Let

$$F(t) = \int F_A(t)d\mu(\lambda) \quad (3)$$

be a mixture of DFR distributions $F_A$ with mixing measure $\mu(\lambda)$. Then $F$
is also DFR.

We omit the proof since it is essentially that given in Section 2 for
the exponential case (3).

What about mixtures of IFR distributions? It is immediately apparent
that such mixtures need not be IFR. To see this, take the simple case of
the mixture of exponential distributions, (3). We have already seen that this mixture is DFR (actually a slight refinement in the argument yields strictly DFR). Since the exponential distribution is the boundary case of an IFR distribution, we conclude that mixtures of IFR distributions are not necessarily IFR.

An interesting open question remains: Let \( F(t) = \int F_\lambda(t)d\mu(\lambda) \) be a mixture of IFR distributions \( F_\lambda \). Find a subclass \( I_0 \) of the class \( I \) of IFR distributions and a subset \( M_0 \) of the set \( M \) of probability measures such that when the \( F_\lambda \) belong to \( I_0 \) and the \( \mu \) belongs to \( M_0 \), then \( F \) is IFR. Simply put, under what conditions is a mixture of IFR distributions IFR?

**Formation of Coherent Systems**

The next reliability operation we consider is the formation of coherent systems using IFR components. We are tempted to reason that a coherent system "wears out" with increasing age since more and more components fail as time goes by. In addition, the individual IFR components are "wearing out" with increasing age since by assumption they each have increasing failure rate. Putting these two facts together, we might conclude (falsely!) that a coherent system of IFR components is wearing out in the sense that it has IFR life-length.

A simple counterexample shows that this reasoning is incorrect. Consider a system of two independent exponential components in parallel, with component \( i \) having failure rate \( \lambda_i, i = 1,2 \). The failure rate of the system is

\[
r(t) = \frac{\lambda_1e^{-\lambda_1t} + \lambda_2e^{-\lambda_2t} - (\lambda_1 + \lambda_2)e^{-(\lambda_1 + \lambda_2)t}}{e^{-\lambda_1t} + e^{-\lambda_2t} - e^{-(\lambda_1 + \lambda_2)t}} \tag{6}
\]
for $0 \leq t < \infty$. System failure rate $r(t)$ is plotted in Figure 1 for several pairs of $\lambda_1, \lambda_2$ values. Note that $r(t)$ is initially increasing and then decreasing with one exception: When $\lambda_1 = \lambda_2 = 0.5$, $r(t)$ is monotonic increasing for all nonnegative $t$. Of course, these graphical results can be verified analytically.

The question now arises: What is the smallest class of life distributions that contains the life distributions of coherent systems of independent IFR components? In a brilliant paper, Birnbaum, Esary, and Marshall (1966) show that the class of increasing failure rate on the average (IFRA) life distributions is closed under the formation of coherent systems.

A distribution $F(t)$ with failure rate $r(t)$ is said to have an increasing failure rate average if $\frac{1}{t} \int_0^t r(u) du$ is monotonic increasing on $(0, \infty)$. If the failure rate fails to exist, then the condition for IFRA is $-\frac{1}{t} \log \bar{F}(t)$ monotonic increasing on $(0, \infty)$. Since an IFR distribution is automatically IFRA, (an increasing function is clearly increasing on the average), the answer to our question is that the IFRA class of life distributions is the smallest class of life distributions that contains the life distributions of coherent systems of independent IFR components.

Birnbaum, Esary, and Marshall (1966) also proved an interesting converse theorem. They show that given an arbitrary IFRA distribution $F$ and an $\varepsilon > 0$, however small, there exist a coherent structure $\phi$ and a set of independent exponential life distributions $G_i(t) = 1 - e^{-\lambda_it}$, $i = 1, \ldots, n$, such that

$$|\bar{F}(t) - h_{\phi}(\bar{G}_{\lambda_1}(t), \ldots, \bar{G}_{\lambda_n}(t))| < \varepsilon$$

for $0 \leq t < \infty$, where $h_{\phi}(\cdot, \ldots, \cdot)$ is the structure $\phi$ reliability. In other words, every IFRA distribution can be approximated by an appropriate
FIGURE 1

REPRESENTATIVE SHAPES OF THE SYSTEM FAILURE RATE WHEN THE SYSTEM CONSISTS OF TWO EXPONENTIAL COMPONENTS IN PARALLEL WITH FAILURE RATES $\lambda_1, \lambda_2$ such that $\lambda_1 + \lambda_2 = 1$. 
coherent system of exponential components.

No corresponding result holds for a coherent system of DFR components.

The IFRA Closure Theorem

Let

\[
X_i(t) = \begin{cases} 
1 & \text{if component } i \text{ is functioning at time } t \\
0 & \text{otherwise.}
\end{cases}
\]

Let \( \phi(\cdot) \) be coordinatewise nondecreasing and

\[
\phi[X(t)] = \begin{cases} 
1 & \text{if system functions at } t \\
0 & \text{otherwise,}
\end{cases}
\]

where \( X(t) = (X_1(t), \ldots, X_n(t)) \). The class of such \( \phi \)'s are called coherent structures (with another condition not required for the closure theorem).

Assume for the moment \( F \) has density \( f \) and let

\[
r(t) = f(t)/(1 - F(t))
\]

be the failure rate. Then

\[
\frac{1}{t} \int_0^t r(u)du \leq \frac{1}{at} \int_0^at r(u)du \leq \frac{1}{t} \int_0^t r(u)du
\]

for all \( 0 \leq \alpha \leq 1 \) or

\[
\left(\frac{1}{[\bar{F}(at)]^{\alpha}}\right)^{1/\alpha} \geq \left(\frac{1}{[\bar{F}(t)]^{\alpha}}\right)^{1/\alpha}
\]

for \( 0 \leq \alpha \leq 1 \) or \([\bar{F}(at)]^{1/\alpha} \geq \bar{F}(t) \quad \forall \ t \geq 0 \).
Definitions

$F$ is IFRA iff $F(0) = 0$ and

$$\frac{1}{\bar{F}(at)} \geq \bar{F}(t) \quad \forall \ t \geq 0 \text{ and } 0 \leq \alpha \leq 1.$$ 

Savits' [Block and Savits (1976)] noticed that this is equivalent to

$$\left[ \int_0^\infty I_{(t,\infty)}(u) dF(u) \right]^{\frac{1}{\alpha}} \geq \int_0^\infty I_{(t,\infty)}(u) dF(u) \quad \forall 0 < \alpha < 1$$

for indicator functions $I_{(t,\infty)}$ with respect to $(t,\infty)$.

But the key observation was that also

$$\left[ \int_0^\infty I_{(t,\infty)}^\alpha(u) dF(u) \right]^{\frac{1}{\alpha}} \geq \int_0^\infty I_{(t,\infty)}(u) dF(u).$$

Rewriting, $T$ (with distribution $F$) is IFRA iff

$$\left[ EI\left(\frac{T}{\alpha}\right) \right]^\alpha \geq \left[ EI(T) \right]^\alpha \quad \forall 0 < \alpha < 1$$

and increasing, indicator functions, $I$.

3.5 Theorem

$T$ (with distribution $F$) is IFRA iff

$$Eh\left(\frac{T}{\alpha}\right) \geq \left[ Eh(T) \right]^\alpha \quad \forall h \geq 0 , \ h^+, \ 0 \leq \alpha \leq 1.$$
This implies
\[
\frac{1}{Eh(T)} \geq \frac{1}{Eh(T)} \geq (Eh(T))^a.
\] (7)

Proof:

The theorem follows from the Minkowski inequality; i.e., if

\[
\|f_i\|_a = \left(\frac{E_i(T)}{a}\right)^a, \quad f_i \geq 0,
\]

then

\[
\sum_{i=1}^{n} \|f_i\|_a = \sum_{i=1}^{n} f_i \leq 0 \leq 1.
\]

We wish to show \(\|h(\cdot)\|_a \geq \|h(\cdot)\|_{a=1}\).

Approximate \(h\) by \(\psi = \sum_{i=1}^{n} a_i I_i\) where \(a_i \geq 0\) and \(I_i\) are increasing indicators.

By the IFRA definition, \(\|I_i(\cdot)\|_a = \|I(\cdot)\|_{a=1}\) and

\[
\|\psi(\cdot)\|_{a=1} = \sum_{i=1}^{n} \|a_i I_i(\cdot)\|_{a=1} \leq \sum_{i=1}^{n} \|a_i I_i(\cdot)\|_{a=1} = \sum_{i=1}^{n} \|a_i I_i(\cdot)\|_{a=1}.
\] (8)

The first inequality in (8) follows from the IFRA definition and the second by the Minkowski inequality. Theorem 3.5 follows by the Lebesgue monotone convergence theorem.
Another key fact is that Theorem 3.5 has a vector generalization:

3.6 Theorem

If $T_1, T_2, \ldots, T_n$ are independent IFRA, then

$$\text{Eh}^\alpha\left(\frac{T_1}{\alpha}, \frac{T_2}{\alpha}, \ldots, \frac{T_n}{\alpha}\right) \geq \left[\text{Eh}(T_1, T_2, \ldots, T_n)\right]^\alpha \quad \forall \ h \geq 0, \ h^+, \text{ and } 0 \leq \alpha \leq 1.$$

The proof is by induction.

Proof:

The implication holds for $n = 1$ by Theorem 3.5. Assume the implication holds for $n - 1$.

$$\text{Eh}^\alpha\left(\frac{T_1}{\alpha}, \frac{T_2}{\alpha}, \ldots, \frac{T_n}{\alpha}\right) = \text{E}_T\left(\text{E}_{T_1}, \ldots, \text{E}_{T_{n-1}} h^{\alpha}\left(\frac{T_1}{\alpha}, \frac{T_2}{\alpha}, \ldots, \frac{T_{n-1}}{\alpha}\right)\right),$$

and by induction,

$$\geq \text{E}_T\left(\text{E}_{T_1}, \ldots, \text{E}_{T_{n-1}} h\left(T_1, \ldots, T_{n-1}\right)\right)^\alpha.$$

Since $T_n$ is IFRA and $g(\cdot) = E_T h(T_1, \ldots, T_{n-1})$ is nonnegative and nondecreasing, then

$$\text{Eh}^\alpha\left(\frac{T_1}{\alpha}, \frac{T_2}{\alpha}, \ldots, \frac{T_n}{\alpha}\right) = \text{E}_T g^\alpha\left(\frac{T_n}{\alpha}\right) \geq \left(\text{E}_T g(T_n)\right)^\alpha$$

$$= \left(\text{Eh}(T_1, \ldots, T_n)\right)^\alpha.$$
\[ \frac{1}{\alpha} g(\alpha T_1, \alpha T_2, \ldots, \alpha T_n) \geq g(T_1, T_2, \ldots, T_n) \text{ for } 0 \leq \alpha \leq 1, \text{ then} \]
\[ g(T_1, T_2, \ldots, T_n) \text{ is IPRA.} \]

Proof:

Let \( I \) be an increasing indicator function. By definition of IPRA, we need only show

\[ \mathbb{E}[g(T_1, T_2, \ldots, T_n)]^{\alpha} \leq (\mathbb{E}[g(T_1, \ldots, T_n)])^{\alpha} \quad \forall 0 \leq \alpha \leq 1. \]

But

\[ \frac{1}{\alpha} g(\alpha T_1, \ldots, \alpha T_n) \geq g(T_1, \ldots, T_n) \]

implies

\[ \frac{1}{\alpha} g(T_1^*, \ldots, T_n^*) \geq g(T_1^*/\alpha, \ldots, T_n^*/\alpha) \quad \text{(change of variable)} \]

implies

\[ \mathbb{E}[g(T_1^*, \ldots, T_n^*)]^{\alpha} \geq \mathbb{E}[g(T_1^*/\alpha, \ldots, T_n^*/\alpha)]^{\alpha} \]
\[ \geq \mathbb{E}[I^*(\frac{T_1}{\alpha}, \ldots, \frac{T_n}{\alpha})]^{\alpha} \geq (\mathbb{E}[g(T_1, \ldots, T_n)])^{\alpha} \]

by Theorem 3.6 since \( I^*(\cdot) = I[g(\cdot)] \) is an increasing indicator function and \( T_1, T_2, \ldots, T_n \) are independent IPRA.||
Examples

Suppose $T_1, T_2, \ldots, T_n$ are independent and IFRA.

1. $g(T_1, T_2, \ldots, T_n) = \min_{1 \leq k \leq n} \max_{1 \leq i \leq k} T_i$ is IFRA (IFRA closure theorem).

2. $g(T_1, T_2, \ldots, T_n) = T_1 + T_2 + \ldots + T_n$ is IFRA (IFRA convolution theorem).

3. $g(T_1, T_2, \ldots, T_n) = \left( \frac{1}{n} \sum_{i=1}^{n} T_i^\beta \right)^{1/\beta}$ is IFRA for $0 < \beta$.

4. $g(T_1, T_2, \ldots, T_n) = \sqrt[n]{\prod_{i=1}^{n} T_i}$ is IFRA.

5. $g(T_1) = T_1^\beta$ for $0 < \beta \leq 1$ is IFRA.

6. Let $K_1, K_2, \ldots, K_k$ be minimum cut sets in a two terminal network. Then the maximum flow through the network is

$$\min_{K} \sum_{i \in K} Z_i$$

if $Z_i$ is the capacity of arc $i$. If $Z_1, \ldots, Z_n$ are independent IFRA, then the maximum flow through the network is IFRA.
\[ \lambda_s = \frac{\mu_s}{\Gamma(s+1)} = \frac{1}{\Gamma(s+1)} \int_0^\infty sx^{s-1}F(x)\,dx \]

\[ \geq (s) \int_0^\infty \frac{sx^{s-1}}{\Gamma(s+1)} \exp\left(-\frac{x}{\lambda_r^s}\right)\,dx = \frac{1}{\lambda_r^s}. \]

5.7 Theorem

(a) For fixed \( t \geq 0 \), the nontrivial lower bound \( \exp\left(-t/\lambda_r^s\right) \) of Corollary 5.4 is decreasing in \( r > 0 \).

(b) The domain \( \left[0, \frac{1}{\lambda_r^s}\right] \) of applicability of this bound is increasing in \( r > 0 \).

Proof:

(a) By Corollary 5.6, for \( r < s \), \( t \geq 0 \), we have \( \exp\left(-t/\lambda_s^s\right) \leq \exp\left(-t/\lambda_r^s\right) \), and so (a) holds.

(b) We leave the proof to the reader.

The special case \( r = 1 \) in Corollary 5.4 is important in reliability applications.

5.8 Bound Based on First Moment

Let \( F \) be IFR with mean \( \mu_1 \). Then

\[ F(t) \geq \begin{cases} \frac{t}{\mu_1} \exp(-t/\mu_1) & \text{for } t < \mu_1, \\ 0 & \text{for } t \geq \mu_1. \end{cases} \quad (16) \]

The bound is sharp.
5.9 Bounds on System Reliability

In the early stages of system design, it is often necessary to predict system reliability from quite minimal knowledge—say, of system structure, of component mean lives, and that each component is IFR, with all components independent. Using (16) we may make a conservative prediction of system reliability.

Specifically, let \( h \) be the reliability function of a coherent system of \( n \) independent components. Let component \( i \) have (unknown) life distribution \( F_i \), IFR, with mean \( \mu_i \) (known), \( i = 1, \ldots, n \). Then a lower bound on system reliability \( h(F_1(t), \ldots, F_n(t)) \) for a mission of duration \( t \) is given by

\[
h(F_1(t), \ldots, F_n(t)) \geq h\left(e^{-\frac{t}{\mu_1}}, \ldots, e^{-\frac{t}{\mu_n}}\right) \text{ for } t < \min (\mu_1, \ldots, \mu_n). \tag{17}\]

Actually, we may make a conservative reliability prediction even when the exact system structure and the corresponding reliability function are not yet known. Furthermore, the components need not be independent, but merely associated.

Specifically, assume (a) that we have a coherent system with (unknown) structure function \( \Phi \), and (b) that the components are associated, with IFR marginal distributions with (known) means \( \mu_1, \ldots, \mu_n \), respectively. Then we may make a conservative reliability prediction for system reliability \( \bar{F}(t) \):

\[
\bar{F}(t) \geq \exp \left[-t \sum_{i=1}^{n} \left(\frac{1}{\mu_i}\right)\right] \text{ for } t < \min (\mu_1, \ldots, \mu_n). \tag{18}\]
Proof:

\[ P(t) = P[\psi(X_1(t), ..., X_n(t)) = 1] \geq \prod_{i=1}^{n} P[X_i(t) = 1] \geq \prod_{i=1}^{n} e^{-\frac{t}{\mu_i}} = \exp \left[ -t \prod_{i=1}^{n} \left( \frac{1}{\mu_i} \right) \right]. \]

Bound (18), resulting from the application of two successive inequalities, may turn out to be quite conservative in many applications.

5.10 Remark

In practice in the early stages of system design, it is often assumed that all components

(a) Are connected in series.
(b) Are independent.
(c) Have exponential life distributions with known mean lives \( \mu_1, \mu_2, ..., \mu_n \).

Then (18) yields a reliability estimate which is conservative if in fact all components

(a) Are part of a coherent system.
(b) Are associated.
(c) Have IFR marginal life distributions with known mean lives \( \mu_1, \mu_2, ..., \mu_n \).

5.11 Example

Suppose an electronic circuit consists of 10 diodes, 4 transistors, 20 resistors, and 10 capacitors. We assume that the system is coherent.
We do not assume that components are independent, but only that they are associated. Finally, component lives are IFR with mean lives given by:

Silicon diode: $\mu_d = 500,000$.
Silicon transistor: $\mu_t = 100,000$.
Composition resistor: $\mu_n = 1,000,000$.
Ceramic capacitor: $\mu_c = 500,000$.

Then (18) yields the conservative predicted system reliability:

$$\bar{F}(t) \geq e^{-0.0001t} \quad \text{for} \quad t < 100,000,$$

where $0.001$ represents the failure rate computed on a parts-count basis, and $100,000$ represents the minimum mean life among components.

For DFR distribution $F$, upper bounds on $\bar{F}(t)$ can be given in terms of a single moment, as shown in

5.12 Theorem

Let $F$ be DFR with mean $\mu_1$. Then

$$\bar{F}(t) \leq \begin{cases} 
\frac{t}{\mu_1} & t < \mu_1 \\
\frac{\mu_1 \cdot e^{-1}}{t} & t \geq \mu_1.
\end{cases} \quad (19)$$

The inequality is sharp.

Proof:

Let $L = \log \bar{F}(t)$. Because $\log \bar{F}(x)$ is convex, there exists $a$, 


\( L \leq a \leq 0 \), such that

\[
\log \bar{F}(x) \geq \frac{L - \alpha}{t} x + x \quad x \geq 0,
\]

or

\[
\bar{F}(x) \geq \exp \left( \frac{L - \alpha}{t} x + a \right).
\]

Thus for some \( a \), \( L \leq a \leq 0 \),

\[
\mu_1 = \int_0^\infty \bar{F}(x)dx \geq \int_0^\infty \exp \left( \frac{L - \alpha}{t} x + a \right)dx
\]

\[
= \tau \frac{x}{a - L}
\]

or

\[
L \leq a - \frac{\tau a}{\mu_1} = \phi(a)
\]

Subject to \( a \leq 0 \), \( \phi(a) \) is maximized by

\[
a_0 = \begin{cases} -\log \left( \frac{t}{\mu_1} \right), & t > \mu_1 \\ 0, & t \leq \mu_1 \end{cases}
\]

---

**FIGURE 2**
and \( L \leq \phi(\alpha_0) \) yields (19). Equality is attained uniquely by the distribution \( G \), where

\[
\ln \bar{G}(x) = \begin{cases} 
0 & , \quad x < 0 \\
-\frac{x}{t} + \log \frac{\mu_1}{t} & , \quad x \geq 0 \quad \text{for} \quad t \geq \mu_1.
\end{cases}
\]

Bounds in terms of the \( r \)-th moment can be easily obtained by the same method.

The sharp lower bound on \( \bar{F}(t) \) for DFR distributions is zero. To see this, let \( \varepsilon > 0 \) be arbitrarily small and

\[
\bar{g}(x) = \begin{cases} 
1 & , \quad x < 0 \\
\varepsilon e^{-\alpha x} & , \quad x \geq 0,
\end{cases}
\]

where \( \alpha = \varepsilon/\mu_1 \). Then \( G \) is DFR with a jump at the origin and \( \bar{G}(x) \leq \varepsilon \) for all \( x \geq 0 \).
6. ESTIMATION OF A MONOTONE FAILURE RATE

Given a set of ordered observations \( X_1 \leq X_2 \leq \ldots \leq X_n \) from an IFR distribution \( F \), Grenander (1956) derives by maximum likelihood methods, an estimator for \( F \) which itself is IFR. Marshall and Proschan (1965) derive by isotonic regression the same estimator and show that it is consistent.

For convenience, if \( F \) is IFR we define its failure rate \( r(x) = \infty \) for all \( x \) such that \( F(x) = 1 \). It is not possible to obtain a maximum likelihood estimator for \( F \in F \), the class of IFR distributions. We cannot maximize \( \prod_{i=1}^{n} f(X_i) \), since for \( F \in F \), \( f(X_i) \) can be arbitrarily large. Consequently, we first consider the subclass \( F^M \) of distributions \( F \) in \( F \) with corresponding failure rates bounded by \( M \), obtaining \( \sup_{F \in F^M} \prod_{i=1}^{n} f(X_i) \leq M^n \).

We shall see that there is a unique distribution \( F_n^M \) in \( F^M \) at which the supremum is attained. These conventional maximum likelihood estimators \( F_n^M \) for \( F^M \) converge in distribution as \( M \to \infty \) (i.e., as \( F_n^M \to F \)) to an estimator \( F_n \in F \) which we call maximum likelihood for \( F \). Furthermore, the density \( f_n^M \) and failure rate \( r_n^M \) of \( F_n^M \) converge in a natural way to the density \( f_n \) and failure rate \( r_n \) (of the continuous part) of \( F_n \) as is shown below.

**Derivation of the Estimators:**

Using (2) we obtain that the log likelihood \( L = L(F) \) is given, for \( F \in F^M \), by

\[
L = \frac{n}{2} \log r(X_1) - \frac{n}{2} \int_{-\infty}^{X_1} r(z) \, dz.
\]

(20)

\( L \) is maximized over \( F_n^M \) by a distribution with failure rate constant between
observations, as shown by Grenander (1956) as follows: Let \( F \in F^M \) have failure rate \( r^* \) and let \( F^* \) be the distribution with failure rate

\[
\begin{align*}
    r^*(x) &= 0, & x < X_i \\
    &= r(X_i), & X_i \leq x \leq X_{i+1}, \quad i = 1, 2, \ldots, n - 1 \\
    &= r(X_n), & x \geq X_n.
\end{align*}
\]

Then \( F^* \in F^M \), and \( r(x) \geq r^*(x) \) so that \( -\int_0^{X_i} r(z)dz \leq -\int_0^{X_i} r^*(z)dz \) for all \( i \); we conclude that \( L(F) \leq L(F^*) \). Thus, we may replace \( L \) by the function

\[
\sum_{i=1}^{n} \log r(X_i) - \sum_{i=1}^{n-1} (n - i)(X_{i+1} - X_i)r(X_i).
\]

The maximization of (22) subject to \( r(X_1) \leq \ldots \leq r(X_n) = M \) is performed in Grenander (1956); it can also be performed as a direct application of Brunk (1958), Corollary 2.1 and the discussion following (see also van Eeden (1956, 1957)). This yields for \( r \) (corresponding to \( F \in F^M \)) the estimator

\[
\hat{r}_n^M(X_i) = \min \left( \min_{v \geq i+1} \max_{u \leq i} \left\{ \frac{1}{v-u} \left[ r^{-1} + \ldots + r_{v-1}^{-1} \right] \right\}^{-1}, M \right),
\]

where \( r_n = M \) and

\[
r_j = [(n - j)(X_{j+1} - X_j)]^{-1} \quad \text{for} \quad j = 1, 2, \ldots, n - 1.
\]

The estimator \( \hat{r}_n^M \) given in (23) differs in form from the one given in Grenander (1956) but is equivalent to it. We may use the form given in (23) to establish consistency.
The maximization procedure which yields (23) may be described as follows. First, find the maximum of (22) obtaining (24). If there is a reversal, say \( r_i > r_{i+1} \), then set \( r(X_i) = r(X_{i+1}) \) in (22) and repeat the procedure. After, at most, \( n \) steps of this kind, a monotone estimator is obtained. The maximum derived with \( r(X_i) = r(X_{i+1}) \) can be directly obtained by replacing \( r_i \) and \( r_{i+1} \) by their harmonic mean \( 2\left(\frac{1}{r_i} + \frac{1}{r_{i+1}}\right)^{-1} \). Succeeding steps amount to further such averaging which is extended just to the point necessary to eliminate all reversals. It can be seen that this is exactly what is called for in (23) taking into account \( r(X) \leq M \). (In this connection see also Ayer, Brunk, Ewing, Reid, and Silverman (1955) and Brunk (1955).) The resulting estimator \( \hat{r}_n^M \) is of the form

\[
\hat{r}_n^M(x) = \begin{cases} 
0, & x \leq X_1 \\
\min \left( n_1 + 1, n_{i+1}, M \right), & X_{n_1+1} \leq x \leq X_{n_{i+1}+1} \\
M, & x > X_n,
\end{cases}
\]

where \( r_1, n_1 \leq r_{n_1+1}, n_2 \leq \cdots \leq r_{n_k+1}, n_{k-1} \), \( 0 = n_0 < n_1 < \cdots < n_k < n-1 \), and \( r_{n_1+1}, n_{i+1} \) is the harmonic mean of \( r_{n_1+1}, r_{n_1+2}, \ldots, r_{n_{i+1}} \). Then \( n_i \) are determined by the rule which determines the extent of the averaging.

The estimator for \( r \) corresponding to \( F \in \mathcal{F} \) is obtained by letting \( N \to \infty \) in (23), and is given by

\[
\hat{r}_n(X_i) = \min_{v \geq i+1} \max_{u \leq i} \left( v - u \right) \left( (n-u)(X_{i+1} - X_u) + \cdots + (n-v+1)(X_i - X_{i-1}) \right)^{-1},
\]

\[i = 1, 2, \ldots, n-1, \text{ and } \hat{r}_n(X_n) = \infty.\]
For the remaining values of \( x \), \( \hat{r}_n(x) \) is determined by (21) with \( \hat{r}_n \) replacing \( r \) and \( r^* \). The corresponding estimators \( \hat{F}_n \) and \( \hat{f}_n \) for \( F \) and \( f \) are obtained from \( \hat{r}_n \) using (2) and the relation \( \hat{f}_n(x) = \hat{r}_n(x)[1 - \hat{F}_n(x)] \).

It is of interest to note that the estimator \( \hat{r}_n \) can also be written in the form

\[
\hat{r}_n(x) = \inf_{F_n(v) \geq F_n(x)} \sup_{F_n(u) < F_n(x)} \frac{\int F_n(y)dy}{u} \tag{26}
\]

where \( F_n \) is the empirical distribution. Similarly, since \( r \) is increasing, \( r(x) \) is given by (26) with \( F \) replacing \( F_n \). Note that infimum of (26) may be taken over the set \( v > x \), but the supremum may be taken over \( u < x \) only if \( x = X_i \) for some \( i \) (in the corresponding formula for \( r \) with \( F \) replacing \( F_n \), \( F(u) < F(x) \) can without restriction be replaced by \( u < x \)).

It is easily seen from (25) or (26) that

\[
[r_n(x_i)]^{-1} = \sup_{v > i+1} \inf_{u < i} [\phi(v) - \phi(u)]/(v - u), \tag{27}
\]

where \( \phi(j) = n \int_0^{X_j} F_n(x)dx \). Let \( \phi^* \) be the convex minorant of \( \phi \) (i.e., \( \phi^* \) is the supremum of convex functions which at each \( j \) do not exceed \( \phi(j) \)).

Then as shown in Grenander (1956), \( [r(x)]^{-1} \) is the right-hand derivative of \( \phi^* \) at \( x \). This representation may be of some importance for computation, since \( \phi^* \) is easily obtained graphically from \( \phi \).

**Decreasing Failure Rate:**

Let \( F \) have decreasing failure rate on the interval of support, \([a, \infty)\).
\( \alpha \geq 0 \). If \( F \) is DFR, then by an argument similar to that used in the IFR case, it is absolutely continuous except possibly for a discontinuity at the point \( \alpha \). Thus, the measure determined by \( F \) is absolutely continuous with respect to \( \mu_\alpha = \delta_\alpha + \lambda \), where \( \delta_\alpha \) places unit mass on \( \{\alpha\} \) and \( \lambda \) is Lebesgue measure; we denote the density of \( F \) with respect to \( \mu_\alpha \) by \( f \) and again define the failure rate of \( F \) by \( r(x) = f(x)/\bar{F}(x^-) \). If \( F \) is DFR, we always take a version of \( f \) for which \( r \) is decreasing in \( (\alpha, \infty) \).

Allowing for the fact that \( f \) is a density with respect to \( \mu_\alpha \), we see that (2) is replaced by

\[
\bar{F}(x) = \left[1 - r(\alpha)\right] \exp \left[-\int_\alpha^x r(z) \, dz\right]. \tag{28}
\]

Estimation in the DFR case parallels that in the IFR case, but with some interesting differences. The first of these is that there are really two problems in the DFR case, depending on whether the point \( \alpha \) is known.

First consider the case that \( \alpha \) is known and suppose \( \alpha = X_1 = \ldots = X_k < X_{k+1} < \ldots < X_n \) (in case \( k = 0 \), we define \( X_0 = \alpha \)). Using (28), \( f(x) = r(x)\bar{F}(x^-) \) and the relations \( r(x) = f(x) = F(\alpha^+) \), we write the log likelihood in the form

\[
k \log r(\alpha) + (n - k) \log (1 - r(\alpha)) + \sum_{i=k+1}^{n} \log r(X_i) - \sum_{i=k+1}^{n} \int r(z) \, dz.
\]

Maximization of the first two terms yields \( \hat{r}_n(\alpha) = k/n = F(\alpha^+) \). Maximization of the last two terms is quite analogous to that in the IFR case, and yields for \( r \) the estimator \( \hat{r}_n(x) = \hat{r}_n(X_i) \), \( X_{i-1} < x < X_i \), \( i = k+1, \ldots, n \), where
\[ \hat{r}_n(X_i) = \max_{v \geq 1} \min_{u \leq i-1} \left\{ (v-u)^{-1}(n-u)(X_{u+1} - X_u) \right. \\
\left. + \ldots + (n-v+1)(X_v - X_{v-1}) \right\}^{-1}, \]
and \( X_0 = \alpha \) in case \( k = 0 \).

Contrary to the IFR case, this DFR estimator is not unique; it is determined by the likelihood equation only for \( x \leq X_n \), and may be extended beyond \( X_n \) in any manner that preserves the DFR property.

Consider now the case that \( \alpha \) is unknown, and assume for the moment that \( F \) is absolutely continuous with respect to Lebesgue measure. If \( F \) is DFR on \( (\alpha, \infty) \) for \( \alpha > X_1 \), then the likelihood \( \Lambda(F) = \Pi f(X_i) = 0 \). If \( F \) is DFR on \( (\alpha, X) \) for \( \alpha < X_1 \), then \( \Lambda(F) < \Lambda(F) \), where \( F \) is defined by

\[
F(x) = \begin{cases} 
\frac{(F(x) - F(X_1))}{\overline{F}(X_1)}, & x \geq X_1 \\
0, & x < X_1.
\end{cases}
\]

Thus the maximum likelihood estimator for \( \alpha \) unknown is found among those DFR distributions with support \( [X_1, \infty) \), and the problem reduces to the case of known \( \alpha \).
REFERENCES


Distributions with Monotone Failure Rate

Richard E. Barlow
Frank Proschan

The Florida State University
Department of Statistics
Tallahassee, Florida 32306

Air Force Office of Scientific Research
Bolling Air Force Base, D.C. 20332

November, 1981

Approved for public release; distribution unlimited.

This survey of the class of life distributions possessing monotone failure rate is intended for publication in a forthcoming Handbook of Reliability, edited by Igor Ushakov, University of Moscow, and colleagues. Contributions are being made by mathematicians and statisticians from different countries on various aspects of reliability theory and practice.