Probability in the Engineering and Informational Sciences, **15**, 2001, 383–400. Printed in the U.S.A.

MODELING A FAILURE RATE FOR A MIXTURE OF DISTRIBUTION FUNCTIONS

MAX S. FINKELSTEIN
University of the Orange Free State
9300 Bloemfontein, Republic of South Africa
E-mail: MSF@wwg3.uovs.ac.za

VERONICA ESAULOVA
Saint Petersburg State University
St. Petersburg, Russia

It is well known that mixtures of decreasing failure rate (DFR) distributions are always DFR. It turns out that, very often, mixtures of increasing failure rate (IFR) distributions can decrease at least in some intervals of time. Usually, this property can be observed asymptotically as $t \to \infty$. In this article, several types of underlying continuous IFR distribution are considered. Two models of mixing are studied: additive and multiplicative. The limiting behavior of a mixture failure rate function is analyzed. It is shown that the conditional characteristics (expectation and variance) of the mixing parameter are crucial for the limiting behavior. Several examples are presented and possible generalizations are discussed.

1. INTRODUCTION

Mixtures of decreasing failure rate (DFR) distributions are always DFR. It turns out that, very often, mixtures of increasing failure rate (IFR) distributions can decrease at least in some intervals of time. It means that the operation of mixing can change the corresponding pattern of aging. This fact is rather surprising and should be taken into account in various applications.

Consider a lifetime random variable (r.v.) $T \geq 0$ with distribution function (DF) $F(t)$ and survival function $\bar{F}(t) \equiv 1 - F(t)$. Assume that $F(t)$ is indexed by a pa-

© 2001 Cambridge University Press 0269-9648/01 $12.50

https://doi.org/10.1017/S0269964801153076 Published online by Cambridge University Press

rameter θ so that $P(T \leq t|\theta) = F(t,\theta)$ and the probability density function $f(t,\theta)$ exists. Then, the corresponding failure rate $\lambda(t,\theta)$ can be defined in a usual way as $f(t,\theta)/\bar{F}(t,\theta)$. Let θ be interpreted as a nonnegative r.v. with support in $[0,\infty)$ and probability density function $\pi(\theta)$. (For the sake of convenience and hoping that it will not lead to confusion, we use the same notation for the random variable and its realization.) A finite support $[a,b]$, $a \geq 0$, $b < \infty$, can be considered as well. A mixture DF is defined by

$$F_m(t) = \int_0^\infty F(t,\theta)\pi(\theta)\,d\theta, \tag{1}$$

whereas the mixture failure rate in accordance with the definition is

$$\lambda_m(t) = \frac{\int_0^\infty f(t,\theta)\pi(\theta)\,d\theta}{\int_0^\infty \bar{F}(t,\theta)\pi(\theta)\,d\theta}. \tag{2}$$

One can also consider the discrete mixture

$$F_m(t) = \sum_k F(t,\theta_k)\Pi(\theta_k), \tag{3}$$

where $\Pi(\theta_k)$ is the probability mass of θ_k. The failure rate in this case is defined similar to (2).

The origin of mixing in practice can be "physical" when, for instance, a number of devices of different (heterogeneous) types, performing the same function and not distinguishable in operation, are mixed together. They are chosen for operation in accordance with probabilities $\Pi(\theta_k), k = 1,2,\ldots,n$. A similar situation arises when data from different distributions are pooled to enlarge the sample size, for instance. Mixing different distribution functions in a subjective approach is suggested by an unknown θ and is performed via a prior $\pi(\theta)$ on θ (Lynn and Singpurwalla [8]). Our main interest in this article is in describing the aging properties of $F_m(t)$ based on the aging properties of the family $F(t,\theta)$.

A well-known closure (under mixtures) result in Barlow and Proschan [2] states that if $\lambda(t,\theta), \forall\theta \geq 0$, is decreasing in t (DFR), then $\lambda_m(t)$ is also decreasing in t. The "borderline" example in Proschan [10] shows the DFR property of the mixture of exponentials. Barlow [1] gives a Bayes explanation of this fact and Mie [9] presents another elegant proof. A mixture of distributions with IFR is not necessarily IFR. Furthermore, some mixtures of IFR distributions can exhibit the DFR property and this is a surprising result. At least, there can be certain intervals in $[0,\infty)$ where $\lambda_m(t)$ decreases. A number of examples and thoughtful considerations on these topics were given by Gurland and Sethuraman [7] and Lynn and Singpurwalla [8]. Specifically, the main interest of the first article is in mixtures of IFR distributions that are ultimately DFR. This means that $\lambda_m(t)$ is decreasing in $[c,\infty)$ for c sufficiently large.

https://doi.org/10.1017/S0269964801153076 Published online by Cambridge University Press

We use some helpful examples of this article in our study. The latter article deals with the bathtub DFs when $\lambda_m(t)$ decreases in $[0, c)$ and increases in $[c, \infty)$, where c is not necessarily large. The bathtub distributions were also studied in Block and Savits [4] and the limiting behavior of $\lambda_m(t)$ as $t \to \infty$ in Block et al. [5].

Similar problems also arise in survival analyses. Specifically, in a number of publications on analyses of mortality, it was found that after a certain age, further aging apparently does not lead to an increase in the risk of death, and the failure (hazard) rates were observed to decline for the "oldest-old" group [12]. This experimental fact was considered surprising by many specialists. The most probable explanation for the deceleration of mortality is heterogeneity of the population. Models that take heterogeneity into account are often referred to as frailty models (see, e.g., [11]). What is more surprising is that this effect can be observed even for a very sharply increasing failure rate $\lambda(t, \theta) = \theta \exp\{t\}$ of the truncated extreme value distribution

$$F(t, \theta) = 1 - \exp\{-\theta(\exp\{t\} - 1)\}, \qquad t \geq 0, \tag{4}$$

which in a slightly more generalized form is often used for modeling of mortality.

The goal of our article is to analyze mixtures of IFR distributions. We shall discuss the above-mentioned and other relevant facts by analyzing piecewise monotonicity properties of the resulting failure rate $\lambda_m(t)$. We are also interested in situations in which the IFR closure under mixture property holds. In practice it can be quite undesirable for the modeling process when the aging properties of the mixture DF $F_m(t)$ differ from those of the initial family $F(t, \theta)$.

2. CONDITIONAL CHARACTERISTICS AND MODELS

A useful and elegant tool for analyzing $\lambda_m(t)$ was considered by Lynn and Singpurwalla [8] in the form of the conditional probability density function (p.d.f.) $\pi(\theta|t)$ of θ given $T \geq t$:

$$\pi(\theta|t) = \frac{\pi(\theta)\bar{F}(t, \theta)}{\displaystyle\int_0^\infty \bar{F}(t, \theta)\pi(\theta)\, d\theta}. \tag{5}$$

Then, $\lambda_m(t)$ can be written as

$$\lambda_m(t) = \int_0^\infty \lambda(t, \theta)\pi(\theta|t)\, d\theta. \tag{6}$$

Denote by $E[\theta|t]$ the conditional expectation of θ given $T \geq t$:

$$E[\theta|t] = \int_0^\infty \theta\pi(\theta|t)\, d\theta.$$

https://doi.org/10.1017/S0269964801153076 Published online by Cambridge University Press

An important characteristic for further consideration is $E'[\theta|t]$, the derivative with respect to t:

$$E'[\theta|t] = \int_0^\infty \theta \pi'(\theta|t)\, d\theta,$$

where

$$\pi'(\theta|t) = -\frac{f(t,\theta)\pi(\theta)}{\int_0^\infty \bar{F}(t,\theta)\pi(\theta)\, d\theta} + \frac{\bar{F}(t,\theta)\pi(\theta)\lambda_m(t)}{\int_0^\infty \bar{F}(t,\theta)\pi(\theta)\, d\theta}$$

$$= \lambda_m(t)\pi(\theta|t) - \frac{f(t,\theta)\pi(\theta)}{\int_0^\infty \bar{F}(t,\theta)\pi(\theta)\, d\theta}. \quad \textbf{(7)}$$

While obtaining Eq. (7), Eq. (2) was used. Eventually, we obtain the following lemma.

LEMMA 1: *The following equation for $E'[\theta|t]$ holds:*

$$E'[\theta|t] = \lambda_m(t)E[\theta|t] - \frac{\int_0^\infty \theta f(t,\theta)\pi(\theta)\, d\theta}{\int_0^\infty \bar{F}(t,\theta)\pi(\theta)\, d\theta}. \quad \textbf{(8)}$$

Now we shall consider two specific cases, where the mixing r.v. θ can be entered directly into the failure rate model.

2.1. Additive Model

Suppose that

$$\lambda(t,\theta) = \alpha(t) + \theta, \quad \textbf{(9)}$$

where $\alpha(t)$ is some deterministic continuous increasing function ($\alpha(t) \geq 0, t \geq 0$) to be specified later. Then, noting that $f(t,\theta) = \lambda(t,\theta)\bar{F}(t,\theta)$, and applying definition (2) for this concrete model,

$$\lambda_m(t) = \alpha(t) + \frac{\int_0^\infty \theta \bar{F}(t,\theta)\pi(\theta)\, d\theta}{\int_0^\infty \bar{F}(t,\theta)\pi(\theta)\, d\theta} = \alpha(t) + E[\theta|t]. \quad \textbf{(10)}$$

https://doi.org/10.1017/S0269964801153076 Published online by Cambridge University Press

Using this relation and Lemma 1, the specific form of $E'[\theta|t]$ can be obtained:

$$E'[\theta|t] = (\alpha(t) + E[\theta|t])E[\theta|t] - \frac{\int_0^\infty (\theta\alpha(t)\bar{F}(t,\theta) + \theta^2\bar{F}(t,\theta))\pi(\theta)\,d\theta}{\int_0^\infty \bar{F}(t,\theta)\pi(\theta)\,d\theta}$$

$$= [E[\theta|t]]^2 - \int_0^\infty \theta^2\pi(\theta|t)\,d\theta = -\mathrm{Var}(\theta|t), \tag{11}$$

where $\mathrm{Var}(\theta|t)$ denotes the variance of θ given $T \geq t$. This result can be formulated in the form of the following lemma.

LEMMA 2: *The conditional expectation of θ for the additive model is a decreasing function of $t \in [0,\infty)$, which follows from*

$$E'[\theta|t] = -\mathrm{Var}(\theta|t) < 0.$$

Upon differentiating (10) and using relation (11), we can obtain now the result that was stated (without proof) in Lynn and Singpurwalla [8].

THEOREM 1: *Let $\alpha(t)$ be an increasing (nondecreasing) convex function in $[0,\infty)$. Assume that $\mathrm{Var}(\theta|t)$ is decreasing in $t \in [0,\infty)$ and*

$$\mathrm{Var}(\theta|0) > \alpha'(0).$$

Then, $\lambda_m(t)$ decreases in $[0,c)$ and increases in $[c,\infty)$, where c can be uniquely defined by $\mathrm{Var}(\theta|t) = \alpha'(t)$ (bathtub shape).

In addition to Lynn and Singpurwalla [8], we have included an assumption that $\mathrm{Var}(\theta|t)$ should decrease in t for all values of t. Intuitively, it seems that similar to the fact that $E[\theta|t]$ is decreasing in $[0,\infty)$ (which follows from (11)), $\mathrm{Var}(\theta|t)$ should also decrease, as the "weak populations are dying out" while t increases. However, this is not true for the general case. The counterexample is presented in the Appendix, showing that the conditional variance is increasing in the neighborhood of 0. It is shown also that $\mathrm{Var}(\theta|t)$ is decreasing in $[0,\infty)$ when θ is exponentially distributed.

Lynn and Singpurwalla [8] were primarily interested in the bathtub shape of $\lambda_m(t)$. We, along with other features, are interested in preservation of IFR properties.

COROLLARY: *Let all other assumptions of Theorem 1 hold while*

$$\mathrm{Var}(\theta|0) \leq \alpha'(0).$$

Then, $\lambda_m(t)$ is increasing in $[0,\infty)$.

It means that the IFR-closure property exists under this assumption. We shall call it the IFR-stability property. In certain situations, it can be very important to know that nonparametric properties of distributions (e.g., characteristics of aging) are not changed after the operation of mixing. It is worthwhile noting that the governing

https://doi.org/10.1017/S0269964801153076 Published online by Cambridge University Press

factor in this analyses is the conditional variance $\mathrm{Var}(\theta|t)$: If it is sufficiently small, then the mixture is IFR stable. This conclusion also follows from some general considerations: The variance is responsible for possible decreasing of $\lambda_m(t)$. When $\mathrm{Var}(\theta|t) = 0$ (θ is deterministic), $\lambda_m(t)$ is increasing due to the assumption that $\alpha(t)$ is increasing.

2.2. Multiplicative Model

Suppose that

$$\lambda(t,\theta) = \theta\alpha(t), \tag{12}$$

where $\alpha(t)$ is some deterministic increasing (nondecreasing) at least for sufficiently large t continuous function ($\alpha(t) \geq 0, t \geq 0$). Thus, in this article, we are mostly interested in the case of ultimately increasing (nondecreasing) baseline functions $\alpha(t)$, especially when $\alpha(t) \to \infty$ as $t \to \infty$. The model (12) is usually called a proportional hazards (PH) model and the previous one is called an additive hazards model. Applying relation (6),

$$\lambda_m(t) = \int_0^\infty \lambda(t,\theta)\pi(\theta|t)\,d\theta = \alpha(t)E[\theta|t]. \tag{13}$$

After differentiating

$$\lambda_m'(t) = \alpha'(t)E[\theta|t] + \alpha(t)E'[\theta|t]. \tag{14}$$

It follows immediately from this equation that when $\alpha(0) = 0$, the failure rate $\lambda_m(t)$ increases in the neighborhood of $t = 0$. Further behavior of this function depends on the other parameters involved.

Example 1: Consider the specific type of the Weibull distribution with linear failure rate $\lambda(t,\theta) = 2\theta t$ and assume that $\pi(\theta)$ is a gamma probability density function:

$$\pi(\theta) = \frac{\vartheta^\beta \theta^{\beta-1}\exp\{-\theta\vartheta\}}{\Gamma(\beta)}, \qquad \beta, \vartheta > 0, \theta \geq 0.$$

The mixture failure rate can be easily obtained via direct integration:

$$\lambda_m(t) = \frac{2\beta t}{\vartheta + t^2}.$$

It is equal to zero at $t = 0$ and tends to zero as $t \to \infty$, with a single maximum at $t = \sqrt{\vartheta}$. Hence, the mixture of IFR distributions has a decreasing (tending to zero!) failure rate for sufficiently large t and this is rather surprising. Furthermore, the same result asymptotically holds for arbitrary Weibull distributions with increasing failure rate. We shall come back to discussing IFR-reversibility or IFR-stability properties in the forthcoming sections.

https://doi.org/10.1017/S0269964801153076 Published online by Cambridge University Press

It follows from (14) that the mixture will be IFR stable if and only if for $\forall t \in [0,\infty)$,

$$\frac{\alpha'(t)}{\alpha(t)} \geq -\frac{E'[\theta|t]}{E[\theta|t]}. \tag{15}$$

Substituting $\lambda_m(t)$ and $f(t,\theta) = \lambda(t,\theta)\bar{F}(t,\theta) = \theta\alpha(t)\bar{F}(t)$ in (8), similar to (11), we obtain the following lemma.

LEMMA 2a. *The conditional expectation of θ for the multiplicative model is a decreasing function of $t \in [0,\infty)$ as follows from*

$$E'[\theta|t] = -\alpha(t)\,\mathrm{Var}(\theta|t) < 0. \tag{16}$$

Combining this result with (12) and (13) for the specific case $\alpha(t) = \text{const}$, we arrive at the well-known result on the DFR property of the mixture of exponentials. Thus, the foregoing can be considered as a *new proof* of this fact.

With the help of (16), inequality (15) can be written as

$$\frac{\alpha'(t)}{\alpha^2(t)} \geq \frac{\mathrm{Var}(\theta|t)}{E[\theta|t]}. \tag{17}$$

Thus, the first two conditional moments and the function $\alpha(t)$ are responsible for the IFR (DFR) properties of the mixture distribution. A similar result was obtained in a different way by Gurland and Sethuraman [7]. In this article, the right-hand side of inequality (17) was written in terms of the moment generating function of the mixing DF $\Pi(\theta)$. Considering these problems with the help of conditional characteristics of θ can be more convenient for the analysis of the corresponding monotonicity properties.

3. LIMITING BEHAVIOR

Relations (5) and (6) can be also used for analyzing the limiting behavior of $\lambda_m(t)$ as $t \to \infty$. We shall start with the discrete mixture case (3), and for simplicity, let $k = 2$; $\theta_1 = a_1$, $\theta_2 = a_2$, $a_2 \geq a_1$; $\Pi(a_1) = p_1$, $\Pi(a_2) = p_2$; $p_1 + p_2 = 1$. Then,

$$\lambda_m(t) = \lambda(t,a_1)\Pi(a_1|t) + \lambda(t,a_2)\Pi(a_2|t), \tag{18}$$

where

$$\Pi(a_i|t) = \frac{p_i\bar{F}(t,a_i)}{p_1\bar{F}(t,a_1) + p_2\bar{F}(t,a_2)}, \qquad i = 1,2. \tag{19}$$

Example 2: Consider the Weibull distribution of the following form:

$$\bar{F}(t,a_i) = \exp\{-a_i t^b\}; \qquad \lambda(t,a_i) = a_i b t^{b-1};\ b > 1;\ i = 1,2.$$

https://doi.org/10.1017/S0269964801153076 Published online by Cambridge University Press

Thus, in accordance with (18) and (19), for the multiplicative model,

$$\lambda_m(t) = a_1 b t^{b-1} \frac{p_1 \exp\{-a_1 t^b\}}{p_1 \exp\{-a_1 t^b\} + p_2 \exp\{-a_2 t^b\}} + a_2 b t^{b-1} \frac{p_2 \exp\{-a_2 t^b\}}{p_1 \exp\{-a_1 t^b\} + p_2 \exp\{-a_2 t^b\}}. \quad \textbf{(20)}$$

These relations suggest that for $t \to \infty$,

$$\lambda_m(t) - \lambda(t, a_1) = (a_2 - a_1) b t^{b-1} \frac{p_2}{p_1} \exp\{-(a_2 - a_1) t^b\}(1 + o(1)) \to 0 \quad \textbf{(21)}$$

and the mixture failure rate is asymptotically "converging" to the failure rate of the strongest population from above. When $b = 1$, the setting reduces to a well-known exponential case. In addition to this convergence result (which is of main interest to us), the corresponding piecewise monotonicity properties can be analyzed in $[0,\infty)$. Thus, $\lambda_m'(t) > 0$ implies that (see also Theorem 1 in Gurland and Sethuraman [7])

$$(p_1 \exp\{-a_1 t^b\} + p_2 \exp\{-a_2 t^b\}) \times (p_1 \exp\{-a_1 t^b\} a_1 b(b-1) t^{b-2} + p_2 \exp\{-a_2 t^b\} a_2 b(b-1) t^{b-2}) > p_1 p_2 \exp\{-(a_1 + a_2) t^b\} (b t^{b-1}(a_1 - a_2))^2.$$

It is clear that for $t \in [c,\infty)$, where c is sufficiently large, this inequality holds. Hence, the mixture failure rate is increasing in this interval. On the other hand, it also holds for $t \in [0, d)$, where d is sufficiently small. The behavior in the intermediate interval depends on the parameters involved. For instance, one can always find a sufficiently small $\varepsilon > 0$ such that if $a_2 - a_1 < \varepsilon$, then the inequality holds in $[0,\infty)$. In this case, the mixture is IFR-stable.

It should be noted that the condition

$$\lim_{t\to\infty} \Pi(a_2|t) = 0 \left(\lim_{t\to\infty} \Pi(a_1|t) = 1 \right)$$

is not sufficient for the convergence result (21).

When dealing with the limiting behavior of the mixture failure rate $\lambda_m(t)$ for a general case of the multiplicative model with increasing $\alpha(t)$ ($\alpha(t) \to \infty$ as $t \to \infty$), we are first interested in the asymptotic convergence of the mixture failure rate to the failure rate of the strongest population:

$$\lambda_m(t) - \lambda(t, a_1) \to 0 \quad \text{as } t \to \infty. \quad \textbf{(22)}$$

On the other hand, a weaker asymptotic equivalence

$$\lambda_m(t) = \lambda(t, a_1)(1 + o(1)) \quad \text{as } t \to \infty \quad \textbf{(23)}$$

can be also important.

https://doi.org/10.1017/S0269964801153076 Published online by Cambridge University Press

The following theorem is a simple consequence of the above considerations.

THEOREM 2: *Let*

$$\lambda(t,a_1) = a_1\alpha(t), \qquad \lambda(t,a_2) = a_2\alpha(t), \qquad a_2 > a_1 > 0,$$

where $\alpha(t) \to \infty$ *as* $t \to \infty$ *in the multiplicative model for the mixture of two distributions. Then,*

(a) *Relation (23) takes place.*
(b) *Relation (22) takes place if and only if*

$$\alpha(t)\exp\left\{-(a_2 - a_1)\int_0^t \alpha(u)\,du\right\} \to 0 \quad \text{as } t \to \infty. \tag{24}$$

The *proof* of this result is absolutely straightforward. Indeed, denote for convenience, $c \equiv a_2/a_1 > 1$. Using simple transformations, similar to Block and Joe [3],

$$\frac{\lambda_m(t)}{\lambda(t,a_1)} = 1 + \frac{p_1(c-1)}{p_1 + p_2(\bar{F}_1(t,a_1))^{1-c}}.$$

As $\bar{F}(t,a_1) \to 0$ for $t \to \infty$, we immediately arrive at (23), whereas the condition

$$\lambda(t,a_1)(\bar{F}_1(t,a_1))^{c-1} \to 0 \qquad \text{as } t \to \infty,$$

which is equivalent to (24), leads to the convergence result (22).

Remark 1: Condition (24) is a rather weak one. In essence, it states that the p.d.f. of a distribution with (ultimately increasing) failure rate $(a_2 - a_1)\alpha(t)$ tends to 0 as $t \to 0$. All distributions which are typically used in a lifetime data analysis meet this condition. However, one can consider some "bizarre" distributions for which relation (24) does not hold. Let, for instance,

$$\alpha(t) = \beta_{n+1} \quad \text{for } t \in [n, n+1),\ n = 0,1,2,\ldots,$$

and

$$\beta_1 = 1, \qquad \beta_{n+1} = \exp\left\{\sum_{i=1}^{n} \beta_i\right\}, \qquad n = 1,2,3,\ldots.$$

The defined $\alpha(t)$ is piecewise continuous, but it can be made a continuous increasing (nondecreasing) function in an obvious way. It is easy to verify in this case that (24) does not hold. Therefore, there is no convergence defined by relation (22). (We want to thank the anonymous referee for this example.)

Remark 2: A general convergence result for the general mixture of distributions was obtained in Block et al. [5, Thm. 4.1]. Because we are dealing with the specific model of mixing in a direct way, only one assumption on the growth of $\alpha(t)$ in the form of (24) is considered. This assumption is weaker than the corresponding condition (C) in Block et al. [5], which states for our case that there exists nonnegative

https://doi.org/10.1017/S0269964801153076 Published online by Cambridge University Press

constants L and T such that $\alpha(t) \le \exp\{Lt\}$ for all $t \ge T$. Indeed, let, for instance, $\alpha(t) = \exp\{t^2\}$ and assumption (24) holds while condition (C) is not satisfied.

Remark 3: It should be noted that relation (22) states only that the distance $\lambda_m(t) - \lambda(t, a_1)$ is positive and tends to zero as $t \to \infty$. The monotonicity properties should be studied additionally as in Example 2. It can be shown that for all IFR distributions, which are typically used in a lifetime data analyses, $(\lambda_m(t) - \lambda(t, a_1))'$ is ultimately negative as $t \to \infty$ and $\lambda_m(t)$ converges monotonically to $\lambda(t, a_1)$ as $t \to \infty$. On the other hand, some nontypical distributions can be constructed exhibiting a nonmonotone convergence (see Example 11 in Gurland and Sethuraman [7]). In this example, $\alpha(t)$ is increasing and bounded and $\alpha'(t)$ unboundedly oscillates as $t \to \infty$.

A continuous case will be studied on the basis of Eq. (14), where the necessary notation is given by Eqs. (5) and (6). From Lemma 2a, we know already that $E[\theta|t]$ is decreasing as $t \to \infty$. Thus, intuitively, it is reasonable to suggest that $E[\theta|t]$ is converging to a. This statement is proved in the following theorem.

THEOREM 3: *Consider the multiplicative failure rate model (12), where $\alpha(t)$ is an ultimately increasing function. Let $\Pi(\theta|t)$ denote the* DF *corresponding to the conditional* p.d.f. *$\pi(\theta|t)$; $\theta \in [a, b)$, $a \ge 0$, $b \le \infty$.*

Then, as $t \to \infty$, $\Pi(\theta|t)$ converges in distribution to the degenerate DF *with a mass in a.*

PROOF: It is clear that it is sufficient to prove the convergence of the conditional density to the Dirac δ-function

$$\pi(\theta|t) \to \delta(a) \quad \text{as } t \to \infty. \tag{25}$$

At first, we shall show that

$$\lim_{t\to\infty} \pi(a|t) = \frac{\pi(a)\bar{F}(t,a)}{\int_a^b \bar{F}(t,\theta)\pi(\theta)\,d(\theta)} = \frac{\pi(a)\exp\left\{-a\int_0^t \alpha(u)\,du\right\}}{\int_a^b \exp\left\{-\theta\int_0^t \alpha(u)\,du\right\}\pi(\theta)\,d\theta} = \infty. \tag{26}$$

Indeed, the integral

$$\int_a^b \exp\left\{(a-\theta)\int_0^t \alpha(u)\,du\right\}\pi(\theta)\,d\theta$$

monotonically tends to zero as $t \to \infty$ for each realization of θ (except for the point $\theta = a$ with probability measure equal to 0 due to our assumptions). Of course, this pointwise convergence to zero is not uniform, but it does not matter due to integration. For instance, if $\alpha(t) = 1$ (and this is"the worst case" for us), the integral tends to zero $\propto (t)^{-1}$ both for specific cases of uniform in $[a, b]$ or exponential in $[0,\infty)$

https://doi.org/10.1017/S0269964801153076 Published online by Cambridge University Press

distribution functions. The other explanation is that the mixture is a distribution function itself. Thus, the corresponding survival function should tend to zero as $t \to \infty$.

The second step is to prove that, for $\forall c \in (0, b-a)$,

$$\lim_{t\to\infty} \pi(a+c|t) = \lim_{t\to\infty} \frac{\pi(a+c)\exp\left\{-(a+c)\int_0^t \alpha(u)\,du\right\}}{\int_a^b \exp\left\{-\theta\int_0^t \alpha(u)\,du\right\}\pi(\theta)\,d\theta} = 0. \tag{27}$$

It can be clearly seen that, for $\forall h \in (0, b-a)$ and sufficiently large t,

$$\int_a^b \exp\left\{-\theta\int_0^t \alpha(u)\,du\right\}\pi(\theta)\,d\theta = \int_a^{a+h} \exp\left\{-\theta\int_0^t \alpha(u)\,du\right\}\pi(\theta)\,d\theta(1+o(1)). \tag{28}$$

Indeed, for $d < h$,

$$\int_a^{a+h} \exp\left\{-\theta\int_0^t \alpha(u)\,du\right\}\pi(\theta)\,d\theta > \int_a^{a+d} \exp\left\{-\theta\int_0^t \alpha(u)\,du\right\}\pi(\theta)\,d\theta$$
$$> \Pi(a, a+d)\exp\left\{-(a+d)\int_0^t \lambda(u)\,du\right\},$$

where

$$\Pi(a, a+d) = \int_a^{a+d} \pi(\theta)\,d\theta.$$

On the other hand,

$$\int_{a+h}^b \exp\left\{-\theta\int_0^t \alpha(u)\,du\right\}\pi(\theta)\,d\theta < \Pi(a+h, b)\exp\left\{-(a+h)\int_0^t \lambda(u)\,du\right\}$$

and (28) follows, as $d < h$.

Thus, due to the multiplicative nature of the model with increasing $\alpha(t)$, the interval of external integration $(a, a+h)$ asymptotically defines integration over the whole interval $[a, b)$. Substituting (28) into the ratio in (27) and choosing $h = c/2$, for instance, we can obtain, for $t \to \infty$,

$$\frac{\pi(a+c)\exp\left\{-(a+c)\int_0^t \alpha(u)\,du\right\}}{\int_a^b \exp\left\{-\theta\int_0^t \alpha(u)\,du\right\}\pi(\theta)\,d\theta} = \frac{\pi(a+c)\exp\left\{-(a+c)\int_0^t \alpha(u)\,du\right\}}{\int_a^{a+c/2} \exp\left\{-\theta\int_0^t \alpha(u)\,du\right\}\pi(\theta)\,d\theta}$$
$$\times (1+o(1)). \tag{29}$$

https://doi.org/10.1017/S0269964801153076 Published online by Cambridge University Press

Taking into account that the inequality

$$\exp\left\{-\theta\int_0^t \alpha(u)\,du\right\} > \exp\left\{-(a+c/2)\int_0^t \alpha(u)\,du\right\},$$

is valid for the internal integral in the denominator of the right-hand side in (29), relation (29) can be continued as

$$\frac{\pi(a+c)\exp\left\{-(a+c)\int_0^t \alpha(u)\,du\right\}}{\int_a^{a+c/2}\exp\left\{-\theta\int_0^t \alpha(u)\,du\right\}\pi(\theta)\,d\theta}$$

$$< \frac{\pi(a+c)\exp\left\{-(a+c)\int_0^t \alpha(u)\,du\right\}}{\Pi(a,a+c/2)\exp\left\{-(a+c/2)\int_0^t \alpha(u)\,du\right\}}$$

$$= \frac{\pi(a+c)}{\Pi(a,a+c/2)}\exp\left\{-c/2\int_0^t \alpha(u)\,du\right\} \to 0 \quad \text{as } t\to\infty.$$

Eventually, relations (26) and (27),

$$\lim_{t\to\infty}\pi(a|t) = \infty, \qquad \lim_{t\to\infty}\pi(a+c|t) = 0, \qquad \forall c\in(a,b-a), \tag{30}$$

combined with

$$\int_a^b \pi(\theta|t)\,d\theta = 1, \qquad \forall t\geq 0,$$

define the Dirac δ-function. Thus, (25) takes place. ■

Remark: As it was stated, the result of Theorem 3 could be expected due to the fact that $E[\theta|t]$ is decreasing as $t\to\infty$. The specific multiplicative model for the failure rate (12) leads to the *multiplicative representation* (13) for the mixture failure rate. This multiplicative representation creates the possibility to analyze the asymptotic behavior of $\lambda_m(t)$ *directly* via $E[\theta|t]$ and $\alpha(t)$. It is important that no additional conditions are imposed on $\alpha(t)$ in this approach. As can be easily seen, Theorem 3 automatically leads to a relation similar to (23), whereas the convergence result similar to (22) needs some assumptions to be considered later in Theorem 4.

The general convergence result of Theorem 4.1 in Block et al. [5] deals mostly with the case of a finite limit for the failure rate; our main interest is in $\alpha(t)\to\infty$ as $t\to\infty$. It is worth noting that the conditional expectation $E[t|\theta]$ does not necessarily decrease for the general case as it does for the specific models under consideration. It seems, however, that Theorem 3 can be proved on the basis of this general

https://doi.org/10.1017/S0269964801153076 Published online by Cambridge University Press

convergence result (with the corresponding assumptions, whereas our proof is free of them), but it is more reasonable to consider the specific case directly and to obtain simple conditions for convergence via the functions $E[\theta|t]$ and $\alpha(t)$.

First, we want, using (25), to perform a limit transition in Eq. (6) for the specific case of a constant failure rate: $\lambda_m(t,\theta) = \lambda(\theta) = \theta\alpha$. The well-known properties of the δ-function imply that

$$\lim_{t\to\infty} \lambda_m(t) = \lim_{t\to\infty} \int_a^b \lambda(t,\theta)\pi(\theta|t)\,d\theta$$

$$= \lim_{t\to\infty} \alpha \int_a^b \theta\pi(\theta|t)\,d\theta = \alpha a = \lambda(a). \tag{31}$$

Thus, if $a = 0$, the failure rate of a mixture of exponentials monotonically converges to 0, as follows from (31), (13), and Lemma 2a. For a general case,

$$\lim_{t\to\infty} \lambda_m(t) = \lim_{t\to\infty} \alpha(t) \int_a^b \theta\pi(\theta|t)\,d\theta = \lim_{t\to\infty} \alpha(t)E[\theta|t]. \tag{32}$$

It was proved in (16) for the multiplicative model that $E[\theta|t]$ is a decreasing function. Taking into account properties of the δ-function,

$$\lim_{t\to\infty} E[\theta|t] = a. \tag{33}$$

Example 1 and some general considerations suggest the possibility of the following convergence result, which is the continuous analog of (22) for our specific model:

$$\lambda_m(t) - a\alpha(t) \to 0, \qquad t \to \infty. \tag{34}$$

When $a = 0$,

$$\lim_{t\to\infty} \lambda_m(t) = 0 \tag{35}$$

and this equation was obtained directly for the specific case of the Weibull distribution in Example 1. We shall also formally call Eq. (35) the convergence to the failure rate of the strongest population. What additional assumptions should be imposed on $\alpha(t)$ for relation (34) ((35)) to be valid? Thus, the following difference should be analyzed as $t \to \infty$:

$$\lambda_m(t) - a\alpha(t) = \alpha(t)(E[\theta|t] - a). \tag{36}$$

Due to our assumptions, $\alpha(t)$ is increasing, whereas according to (33), the second multiplier tends to zero as $t \to \infty$. Hence, it is necessary and sufficient for the convergence (34) that $E[\theta|t]$ should decrease to 0 more sharply than $\alpha(t)$ increases. By using L'Hôspital's rule ($\alpha(t) \to \infty$) and Lemma 2, the following theorem is obtained:

https://doi.org/10.1017/S0269964801153076 Published online by Cambridge University Press

THEOREM 4: *Let the assumptions of Theorem 3 hold. Then:*

(a) *Relation (34), showing the property of asymptotic convergence of the mixture failure rate to the failure rate of the strongest population, takes place if and only if*

$$\lim_{t\to\infty} \alpha(t)(E[\theta|t] - a) = 0, \qquad a \geq 0,$$

or, equivalently,

$$\lim_{t\to\infty}\left(-\frac{\alpha^2(t)E'[\theta|t]}{\alpha'(t)}\right) = \lim_{t\to\infty}\left(\frac{\alpha^3(t)\,\mathrm{Var}(\theta|t)}{\alpha'(t)}\right) = 0. \tag{37}$$

(b) *If $a = 0$, the mixture failure rate converges to a constant $c \neq 0$ if and only if*

$$\lim_{t\to\infty} \alpha(t)E[\theta|t] = c.$$

The following corollary results from differentiating relation (36).

COROLLARY 1: *The mixture failure rate asymptotically converges to the failure rate of the strongest population monotonically if, in addition to (38),*

$$\frac{\alpha'(t)}{\alpha^2(t)} < \frac{\mathrm{Var}(\theta|t)}{E[\theta|t] - a}. \tag{38}$$

Directly from (36), we have the following corollary.

COROLLARY 2: *If an increasing function $\alpha(t)$ is bounded (e.g., as in the case of the gamma distribution), then the mixture failure rate automatically converges to the failure rate of the strongest population.*

Remark: The direct asymptotic analysis of the mixture failure rate based on relation (32) and Theorem 4 enables one to obtain simple results in a straightforward way. It is usually easy to derive $E[\theta|t]$ in the specific cases. In Example 1, for instance, $E[\theta|t] = 2\beta/(\vartheta + t^2)$. For the mixture of exponentials (see Eqs. (31)) with exponential mixing distribution with parameter ϑ: $E[\theta|t] = (\alpha t + \vartheta)$. We are mostly interested in the case of convergence to the failure rate of the strongest population when $a = 0$ and $\lambda_m(t) \to 0$ as $t \to \infty$. It is worth noting that this case cannot be considered by the general convergence Theorem 4.1 in Block et al. [5], because assumption A (uniform convergence of the failure rates in $\theta \in [0,\infty)$) does not hold.

On the other hand, the "direct approach" also succeeds when a different convergence (not to the failure rate of the strongest population) takes place, as in Theorem 4(b) and in the following important example.

https://doi.org/10.1017/S0269964801153076 Published online by Cambridge University Press

Example 3: Consider the truncated extreme value distribution defined in a manner similar to (4):

$$F(t,k,\theta) = 1 - \exp\{-\theta k(\exp\{t\} - 1)\}, \qquad t \geq 0, k > 0,$$
$$\lambda(t,\theta) = \theta\alpha(t) = \theta k \exp\{t\}.$$

Assume that $\pi(\theta)$ is the gamma probability density function with scale parameter ϑ and shape parameter β. Then, it follows from Gurland and Sethuraman [7] that for $\vartheta < k$, the mixture failure rate is decreasing as $t \to \infty$, whereas for $\vartheta > k$, it is increasing. This is a formally obtained fact, but the two possibilities intuitively look rather strange. We can clarify this example and, in addition, consider the convergence properties in a direct way. For simplicity, let $\pi(\theta)$ be exponential with parameter ϑ. Then, direct integration in (2) for this specific case leads to

$$\lambda_m(t) = \frac{k\exp\{t\}}{\omega} = \frac{k\exp\{t\}}{k\exp\{t\} - k + \vartheta} = 1 + \frac{k - \vartheta}{k\exp\{t\} - k + \vartheta}.$$

It follows from this equation that $\lambda_m(0) = k/\vartheta$. When $k > \vartheta$, the mixture failure rate is monotonically decreasing asymptotically converging to 1. For $k < \vartheta$, it is monotonically increasing asymptotically converging to 1. When $k = \vartheta$, the mixture failure rate is equal to 1: $\lambda_m(t) \equiv 1$, $\forall t \in [0,\infty)$. The most amazing fact is that this kind of mixing can result in a constant failure rate for $k = \vartheta$!

The different behavior for different values of k in this specific nontypical example is the result of the "interaction" between $\alpha(t)$ and $E[\theta|t](a = 0)$. We can define it directly:

$$\alpha(t) = k\exp\{t\}; \qquad E[\theta|t] = \frac{1}{k\exp\{t\} - k + \vartheta}.$$

For a more general setting, the limit in (37) and the inequality (38) should be analyzed.

The mixture failure rate of the commonly used lifetime distributions usually exhibits the property of convergence to the failure rate of the strongest population (e.g., Example 1). As it follows from (36) and (37) and illustrated by Example 3, this is not always the case. The failure rate of the governing distribution should increase very sharply for exhibiting a different pattern.

It can be also obtained that, for instance, for the Weibull DF $F(t,\theta)$ with increasing $\lambda(t,\theta)$ and bounded support for $\theta : [a,b]$, the upper bound b can be chosen sufficiently close to a in such a way that the resulting $\lambda_m(t)$ will be increasing for $\forall t \in [0,\infty)$ (IFR stability). The similar situation in the discrete case was considered in Example 1.

Theorem 4 in a simplified way is valid for the additive model. Indeed,

$$\lambda_m(t) = \int_a^b (\alpha(t) + \theta)\pi(\theta|t)\,d\theta = \alpha(t) + \int_a^b \theta\pi(\theta|t)\,d\theta = \alpha(t) + E[\theta|t].$$

https://doi.org/10.1017/S0269964801153076 Published online by Cambridge University Press

As follows from Lemma 2, $E[\theta|t]$ is decreasing in time. Taking into consideration (33), we can conclude that *monotonically* as $t \to \infty$,

$$\lambda_m(t) - (\alpha(t) + a) \to 0.$$

Hence, $\lambda_m(t)$, as $t \to \infty$, is asymptotically converging to the increasing function $\alpha(t) + a$.

4. CONCLUDING REMARKS

The change in the pattern of the mixture failure rate behavior compared with the pattern of failure rates, which generate this mixture, is usually undesirable. Pooling from heterogeneous populations can often lead to this outcome.

In many cases, the mixture failure rate $\lambda_m(t)$ of the multiplicative model exhibits a tendency to decrease, at least for substantially large t, and eventually converges to 0. Some surprising examples show that this takes place even when $\lambda(t,\theta)$ is sharply increasing for large t as t^k for any $k > 1$. What is the reason for this kind of tendency? The origin is in combination of a decreasing conditional expectation $E[\theta|t]$ with increasing $\alpha(t)$ in such a way that $\alpha(t)E[\theta|t]$ tends to zero as $t \to \infty$.

Simplifying the situation, one can say that the mixture failure rate tends to increase, but at the same time, another process of "sliding down" due to the effect of dying out of the weakest populations takes place. Interaction between these two tendencies can result in different shapes of $\lambda_m(t)$. This interaction can lead even to the more surprising results, as in Example 3. On the other hand, preservation of the IFR property under the operation of mixing is possible if inequality (38) (with an opposite sign) holds.

We have considered only the two simplest important models for $\lambda(t,\theta)$: additive and multiplicative. The latter coincides with the proportional hazards model with a random multiplicative factor and is very important in different applications. The results can be easily generalized for the accelerated life model (see, e.g., Finkelstein [6]). As was mentioned earlier, a more general case can be approached as well. Some ordering in the family of $\lambda(t,\theta)$ should be imposed and additional problems dealing with convergence should be addressed. This is a topic for future work.

Acknowledgments

The authors thank Professor N. Singpurwalla for drawing attention to the problem of mixing in the considered aspect, Professor H. Block and Professor T. Savits for helpful discussion, and an anonymous referee whose comments greatly improved the presentation of this article.

References

1. Barlow, R.E. (1985). A Bayes explanation of an apparent failure rate paradox. *IEEE Transactions on Reliability* R-34: 107–108.
2. Barlow, R.E. & Proschan, F. (1975). *Statistical theory of reliability: Probability models*. New York: Holt, Rinehart & Winston.
3. Block, W.H. & Joe, H. (1997). Tail behavior of the failure rate functions of mixtures. *Lifetime Data Analyses* 3: 269–288.

https://doi.org/10.1017/S0269964801153076 Published online by Cambridge University Press

4. Block, H.W. & Savits, T.H. (1997). Burn-in. *Statistical Science* 12: 1–19.
5. Block, H.W., Mie, J., & Savits, T.H. (1993). Burn-in and mixed populations. *Journal of Applied Probability* 30: 692–702.
6. Finkelstein, M.S. (1999). Wearing-out of components in a variable environment. *Reliability Engineering and System Safety* 66: 235–242.
7. Gurland, J. & Sethuraman, J. (1995). How pooling failure data may reverse increasing failure rate. *Journal of the American Statistical Association* 90: 1416–1423.
8. Lynn, N.J. & Singpurwalla, N.D. (1997). Comment: "Burn-in" makes us feel good. *Statistical Science* 12: 13–19.
9. Mie, J. (1998). A new explanation of decreasing failure rate of a mixture of distributions. *IEEE Transactions on Reliability* R-47: 460–462.
10. Proschan, F. (1963). Theoretical explanation of observed decreasing failure rate. *Technometrics* 5: 373–383.
11. Vaupel, J.W. & Carey, J.R. (1993). Compositional interpretation of medfly mortality. *Science* 260: 1666–1667.
12. Wang, J., Muller, H., & Capra, W.B. (1998). Analyses of oldest-old mortality: Lifetimes revisited. *Annals of Statistics* 26: 126–133.

APPENDIX
Monotonicity Properties of the Conditional Variance for Some Specific Cases

Substituting Eq. (9) into Eq. (5),

$$\pi(\theta|t) = \frac{\pi(\theta)\bar{F}(t,\theta)}{\int_0^\infty \bar{F}(t,\theta)\pi(\theta)\,d(\theta)} = \frac{\exp\{-\theta t\}\pi(\theta)}{\int_0^\infty \exp\{-\theta t\}\pi(\theta)\,d(\theta)}.$$

The sign of $E''[\theta|t]$ is of interest. Taking into account Eq. (11) and the definition of $E[\theta|t]$,

$$E''[\theta|t] = \int_0^\infty \theta\pi''(\theta|t)\,d\theta,$$

$$\pi'(\theta|t) = \pi(\theta|t)(-\theta + E[\theta|t]),$$

then

$$\pi''(\theta|t) = \pi(\theta|t)(\theta^2 - 2\theta E[\theta|t] + E[\theta|t]^2 + E'[\theta|t])$$

and

$$E''[\theta|t] = 2E[\theta|t]^3 + \int_0^\infty \theta^3\pi(\theta|t)\,d\theta - 3E[\theta|t]\int_0^\infty \theta^2\pi(\theta|t)\,d\theta. \quad \textbf{(A.1)}$$

The sign of $E''[\theta|t]$ in (A.1) is not necessarily positive. Let $t = 0$ and this value is important for the result stated in Section 2:

$$E''[\theta|0] = 2\left(\int_0^\infty \theta\pi(\theta)\,d\theta\right)^3 + \int_0^\infty \theta^3\pi(\theta)\,d\theta - 3\int_0^\infty \theta\pi(\theta)\,d\theta\int_0^\infty \theta^2\pi(\theta)\,d\theta.$$

https://doi.org/10.1017/S0269964801153076 Published online by Cambridge University Press

Let, for instance, $\pi(\theta) = 2\theta$ with support in $[0, 1]$. Then

$$2\left(2\int_0^1 \theta^2\, d\theta\right)^3 + 2\int_0^1 \theta^4\, d\theta - 3\left[4\int_0^1 \theta^2\, d\theta \int_0^1 \theta^3 \pi(\theta)\, d\theta\right] = \frac{106}{135} - 1 < 0.$$

This means that conditional variance is increasing in the neighborhood of 0.

Consider, now, the exponential distribution of θ: $\pi(\theta) = \exp\{-\theta\}$, $\theta \in [0,\infty)$:

$$\pi(\theta|t) = \frac{\exp\{-\theta(t+1)\}}{\displaystyle\int_0^\infty \exp\{-\theta(t+1)\}\, d\theta} = (t+1)\exp\{-\theta(t+1)\}.$$

Taking into account that for $k = 1,2,\ldots,$

$$\int_0^\infty \theta^k \pi(\theta|t)\, d\theta = \int_0^\infty \theta^k (t+1)\exp\{-\theta(t+1)\}\, d\theta = \frac{1}{(t+1)^k}\int_0^\infty \theta^k \exp\{-\theta\}\, d\theta$$

$$= \frac{k!}{(t+1)^k},$$

we arrive at

$$2\left(\int_0^\infty \theta\pi(\theta|t)\, d\theta\right)^3 + \int_0^\infty \theta^3 \pi(\theta|t)\, d\theta - 3\int_0^\infty \theta\pi(\theta|t)\, d\theta \int_0^\infty \theta^2 \pi(\theta|t)\, d\theta$$

$$= \frac{2}{(t+1)^3} > 0.$$

Hence, in this case, the conditional variance is decreasing for all values of $t \in [0,\infty)$.

Finally, the example of the uniform in $[0, 1]$ distribution of θ can be studied. Similar to the previous examples, it can be shown that the conditional variance in this case is decreasing, at least for small values of t.

https://doi.org/10.1017/S0269964801153076 Published online by Cambridge University Press