A NOTE ON THE FAILURE RATES IN FINITE MIXED POPULATIONS

JI HWAN CHA,* *Ewha Womans University*

MASSIMILIANO GIORGIO,** *Second University of Naples*

Abstract

Almost all populations existing in the real world are finite populations. Specifically, in the areas relevant to lifetime modeling and analysis, finite populations are frequently encountered. However, descriptions of failure/survival patterns of elements in the finite population have not yet been properly established. In particular, it is questionable whether the ordinary failure rate can be defined for finite populations in the same way and whether the corresponding interpretations are still valid. In this paper we consider two kinds of finite mixed population and provide new definitions for their failure rates. Then we clarify the notion of failure rate in finite populations.

Keywords: Failure rate; proportion-based failure rate; heterogeneous population; finite mixed population

2010 Mathematics Subject Classification: Primary 60E05
Secondary 90B25

1. Introduction

The failure (hazard) rate function plays a crucial role in the areas of lifetime modeling and analysis such as reliability, survival analysis, actuarial sciences, and demographic and biological disciplines which deal with the 'lifetime' of items. Denote the lifetime of an item by X and its absolutely continuous cumulative distribution function (CDF) and probability distribution function (PDF) by $F(t)$ and $f(t)$, respectively. The failure rate function of X is defined by

$$\lambda(t) = \lim_{\Delta t \to 0} \frac{1}{\Delta t} \mathrm{P}(t < X \leq t + \Delta t \mid X > t) = \frac{f(t)}{\bar{F}(t)}, \tag{1}$$

where $\bar{F}(t) \equiv 1 - F(t)$. The failure rate in (1) is also called the 'force of mortality' or 'mortality rate' in actuarial and demographic disciplines. In what follows, we will use the term 'failure rate' and describe our problem in the context of reliability for a convenient formulation. However, applications of the results obtained in this paper should not be limited to only the reliability field but could generally be applied to fields relevant to lifetime modeling and analysis as well.

The failure rate function $\lambda(t)$ has the following very popular and interesting probabilistic interpretation (see Aven and Jensen (1999, p. 37)):

$$\lambda(t)\Delta t \approx \mathrm{P}(t < X \leq t + \Delta t \mid X > t).$$

Note that the failure rate function in (1) is generally defined under the basic assumption that the hypothetical population of interest consists of infinitely many items, i.e. infinite population.

However, almost all populations in the areas relevant to lifetime modeling and analysis, such as reliability, quality control, survival analysis, and biological sciences and demographic discipline, are finite. For example, consider items produced during specific years by a manufacturing company. In this case, the total produced items constitute a finite population. Also, in the areas of reliability and quality control, produced items may constitute lots of (finite) size N, with these lots sold to customers. Then, from the view point of a customer who purchased a lot, the reliability of this 'finite population' (i.e. lot) is of importance. Of course, when the population size is very large, a finite population can be well approximated by an infinite population, which is the main reason why infinite populations are frequently adopted in practice. However, when we have a relatively small population size, a more careful approach needs to be taken. In this case, it is questionable whether the failure rate function defined in (1) would still constitute a proper and comprehensive definition.

In this paper we consider finite mixed populations. A mixture is an important practical tool for modeling lifetimes under heterogeneity. For a detailed practical example, a population of items from several different production lines in a factory can constitute a mixed population due to different reliability characteristics among the production lines. Research on the failure rate function of the mixture distribution has been carried out by many authors; see, e.g. Finkelstein (2008), (2009), Finkelstein and Esaulova (2001), Navarro and Hernandez (2004), Jiang and Murthy (1995), Block *et al.* (2003), and Vaupel and Zhang (2010). However, most of the research is based on the assumption of infinite populations.

We consider two kinds of finite mixed population and provide new definitions of the failure rate in finite populations. Furthermore, the notion of the failure rate in a finite population is clarified.

The structure of this paper is as follows. In Section 2 we consider a fixed finite mixed population and introduce a new definition of its failure rate function. We show that, in the case of a finite population, the use of the ordinary failure rate function (1) may lead to misleading interpretations. In Section 3, a randomly selected finite mixed population is considered and the new definition of its failure rate function for the population is discussed. Furthermore, we show that the survival function of a finite mixed population does not uniquely determine the new type of failure rate. Finally, in Section 4 we give some concluding remarks.

2. Fixed finite mixed population

We start this section by considering the basic questions raised by the problem, which are the motivation and starting point of this research.

Before our main discussions, it is useful to present some preliminaries on the interpretations of the failure rate and the corresponding applications. According to the aim of the analyses, the failure rate can usually be interpreted as

- a measure of the proneness to failure (death) of an item which was randomly selected at time 0 from the population and has survived up to time t, or

- a measure of the proneness to failure of an item which is randomly selected at time t from all the nonfailed items in the population.

Throughout this paper, the term 'randomly selected' is defined as 'selected with equal probabilities'. Obviously, the above two-fold interpretations are equivalent for infinite populations and can be expressed by the ordinary failure rate in (1). Hereafter, for the sake of clarity, we will

use the term 'failure rate' to refer to the former interpretation and the term 'proportion-based failure rate' to refer to the latter interpretation.

By virtue of this two-fold interpretation, engineers are attracted to the notion of the failure rate. In fact, the first interpretation captures their intuitive notions about wear, in the sense that items that wear out should have a failure rate function that increases with time. Based on the shape of the failure rate function, in the reliability field, reasonable decisions (e.g. maintenance actions) which optimize system performance measures can be reached.

On the other hand, the second interpretation is referred to by engineers when the whole population of nonfailed items is of concern. For example, burn-in procedures are developed based on this interpretation (see, e.g. Mi (1994) and Cha (2000), (2001), (2003)).

Similarly, in biometry, actuarial, and demographic disciplines the failure rate function is considered a useful tool for expressing opinions about ageing or attrition over time.

Now we start our main discussions. Consider identical items produced by k manufacturers (manufacturer 1, manufacturer 2, etc.) belonging to the same manufacturing company. In this situation, each of these k subpopulations can be considered to be an infinite population as infinite numbers of items can be produced (theoretically). Suppose that the failure rate function (defined in accordance with (1)) of an item from subpopulation i is given by

$$\lambda_i(t) = \frac{f_i(t)}{\bar{F}_i(t)}, \qquad i = 1, 2, \ldots, k,$$

where $F_i(t) = 1 - \bar{F}_i(t)$ and $f_i(t)$ are the corresponding absolutely continuous CDF and PDF of subpopulation i, respectively. During a 'fixed period', n_i items are produced from manufacturer i, i.e. n_i items are randomly selected from subpopulation i, $i = 1, 2, \ldots, k$. Then all of these produced items from k subpopulations constitute a mixed population of size N, where N is a fixed constant and $N = n_1 + n_2 + \cdots + n_k$. These items mixed from k manufacturers during the period then constitute a *fixed finite mixed population*.

As we have a 'mixed population' which is composed of k subpopulations (subpopulation 1, subpopulation 2, etc.), with corresponding proportions

$$\pi_1 = \frac{n_1}{N}, \qquad \pi_2 = \frac{n_2}{N}, \ldots, \qquad \pi_k = \frac{n_k}{N},$$

the mixture distribution of the population is given by

$$F(t) = \pi_1 F_1(t) + \pi_2 F_2(t) + \cdots + \pi_k F_k(t),$$

and the corresponding density function is given by

$$f(t) = \pi_1 f_1(t) + \pi_2 f_2(t) + \cdots + \pi_k f_k(t).$$

Then the 'ordinary failure rate' which corresponds to the definition in (1) is given by

$$\lambda(t) = \frac{\sum_{i=1}^{k} \pi_i f_i(t)}{\sum_{i=1}^{k} \pi_i \bar{F}_i(t)} = \sum_{i=1}^{k} \pi_i(t) \lambda_i(t), \qquad (2)$$

where the time-dependent probabilities $\pi_i(t)$, $i = 1, 2, \ldots, k$, are

$$\pi_i(t) = \frac{\pi_i \bar{F}_i(t)}{\sum_{j=1}^{k} \pi_j \bar{F}_j(t)}, \qquad i = 1, 2, \ldots, k.$$

Suppose now that this finite population 'lives' for time t (i.e. all the items in the finite population are operated during $(0, t]$) and we have an aged finite population at time t. This situation can be frequently realized in practice, for instance, when items are burned-in in the reliability field.

(Burn-in is a widely used engineering method applied to eliminate initial failures in field usage. See, e.g. Mi (1994), Cha (2000), (2001), (2003), and Ebrahimi (2004).) Suppose now that, at time t, an item is randomly selected among the 'survived items' in the finite population. Then the questions of interest are the following.

- Can the measure of the proneness to failure of the item randomly selected at time t be described by $\lambda(t)$ in (2)?

- Can the proneness of future failure of the item randomly selected at time t be measured by $\lambda(t + u)$, $u \geq 0$?

Note that, according to the two-fold interpretation described above, $\pi_i(t)$ in (2) can be interpreted in two ways: (i) the probability that the item randomly selected at $t = 0$ and operational at time t will be from subpopulation i; (ii) the probability that the item randomly selected at time t among survived items in an infinite population will be from subpopulation i. According to this consideration, if an item in the above mixed finite population is randomly selected at time $t = 0$, then we can answer the above two questions in the affirmative. However, in our setting, the item is selected at time t among survived items in a finite population and a more careful consideration is needed. In this case, the answer to the first question is given in Corollary 1, whereas the following theorem gives a general result for obtaining it.

Theorem 1. *Consider a finite population with N different items whose lifetimes are denoted by $X_1, X_2, \ldots, X_N$. Let $r_i(t)$ be the failure rate function (defined in accordance with (1)) of X_i, $i = 1, 2, \ldots, N$. Then the proportion-based failure rate of this finite population is given by*

$$\sum_{i=1}^{N} \mathrm{E}\left[\frac{M_i'}{M_1' + M_2' + \cdots + M_N'} \,\middle|\, N' > 0 \right] r_i(t),$$

where N' is the total number of items that survived up to time t and M_i', $i = 1, 2, \ldots, N$, is 1 if the ith item survived up to time t and 0 otherwise.

Proof. As we are dealing with a finite population, it is clear that the proportion-based failure rate should be obtained by applying the limiting measure

$$\lim_{\Delta t \to 0} \frac{1}{\Delta t} \mathrm{E}\left[\frac{\sum_{i=1}^{N} \mathbf{1}(t < X_i \leq t + \Delta t)}{\sum_{i=1}^{N} \mathbf{1}(X_i > t)} \,\middle|\, \text{at least one unit is alive at time } t \right],$$

which is called the 'engineering notion of hazard rate' in Ebrahimi (1996).

Define the sets A_1, A_2, and A_3 by $A_1 \equiv [0, t]$, $A_2 \equiv (t, t + \Delta t]$, $A_3 \equiv (t + \Delta t, \infty)$, and the set C by

$$C \equiv \{(c_1, c_2, \ldots, c_N): c_i = 1, 2, 3, \ i = 1, 2, \ldots, N, \ (c_1, c_2, \ldots, c_N) \neq (1, 1, \ldots, 1)\}.$$

Then observe that

$$\mathrm{E}\left[\frac{\sum_{i=1}^{N} \mathbf{1}(t < X_i \leq t + \Delta t)}{\sum_{i=1}^{N} \mathbf{1}(X_i > t)} \,\middle|\, \text{at least one unit is alive at time } t \right]$$

$$= \sum_{(j_1, j_2, \ldots, j_N) \in C} \frac{M_2(j_1, j_2, \ldots, j_N)}{M_1(j_1, j_2, \ldots, j_N)}$$

$$\times \mathrm{P}(X_i \in A_{j_i}, \ i = 1, 2, \ldots, N \mid X_i \notin A_1 \text{for at least one } i), \quad (3)$$

where $M_1(j_1, j_2, \ldots, j_N)$ is the number of '2's or '3's in $(j_1, j_2, \ldots, j_N)$ and $M_2(j_1, j_2, \ldots, j_N)$ is the number of '2's in $(j_1, j_2, \ldots, j_N)$. Then, in deriving (3) in a more detailed form, it is sufficient to compute the conditional probabilities only for the sets $j_1, j_2, \ldots, j_N$ for which there is exactly one i such that $X_i \in A_2$, as the probability of observing more than one failure in the interval A_2 will vanish when we take the limit '$\Delta t \to 0$'. Excluding such 'eventually vanishing' terms, if we focus on the conditional probabilities with '$X_1 \in A_2$' (unconditional part) then the sum of these terms in (3) is given by

$$\frac{1}{N} P(t < X_1 \leq t + \Delta t, \, X_j > t + \Delta t, \, j = 2, 3, \ldots, N \mid \text{at least one } X_i > t)$$

$$+ \frac{1}{N-1} [P(t < X_1 \leq t + \Delta t, \, X_2 < t, \, X_j > t + \Delta t, \, j = 3, \ldots, N \mid \text{at least one } X_i > t)$$

$$+ P(t < X_1 \leq t + \Delta t, \, X_2 > t + \Delta t, \, X_3 < t, \, X_j > t + \Delta t,$$
$$j = 4, \ldots, N \mid \text{at least one } X_i > t)$$

$$+ \cdots$$

$$+ P(t < X_1 \leq t + \Delta t, \, X_j > t + \Delta t, \, j = 2, \ldots, N-1,$$
$$X_N < t \mid \text{at least one } X_i > t)]$$

$$+ \frac{1}{N-2} [P(t < X_1 \leq t + \Delta t, \, X_2 < t, \, X_3 < t, \, X_j > t + \Delta t,$$
$$j = 4, \ldots, N \mid \text{at least one } X_i > t)$$

$$+ P(t < X_1 \leq t + \Delta t, \, X_2 < t, \, X_3 > t + \Delta t, \, X_4 < t, \, X_j > t + \Delta t,$$
$$j = 5, \ldots, N \mid \text{at least one } X_i > t)$$

$$+ \cdots$$

$$+ P(t < X_1 \leq t + \Delta t, \, X_j > t + \Delta t, \, j = 2, \ldots, N-2, \, X_{N-1} < t,$$
$$X_N < t \mid \text{at least one } X_i > t)]$$

$$+ \cdots + \frac{1}{1} P(t < X_1 \leq t + \Delta t, \, X_j < t, \, j = 2, \ldots, N \mid \text{at least one } X_i > t).$$

If we take the limit of the above equation after dividing by Δt, we obtain

$$r_1(t) \frac{1}{N} P(X_1 > t, \, X_j > t, \, j = 2, 3, \ldots, N \mid \text{at least one } X_i > t)$$

$$+ r_1(t) \frac{1}{N-1} [P(X_1 > t, \, X_2 < t, \, X_j > t, \, j = 3, \ldots, N \mid \text{at least one } X_i > t)$$

$$+ P(X_1 > t, \, X_2 > t, \, X_3 < t, \, X_j > t, \, j = 4, \ldots, N \mid \text{at least one } X_i > t)$$
$$+ \cdots$$
$$+ P(X_1 > t, \, X_j > t, \, j = 2, \ldots, N-1, \, X_N < t \mid \text{at least one } X_i > t)]$$

$$+ r_1(t) \frac{1}{N-2} [P(X_1 > t, \, X_2 < t, \, X_3 < t, \, X_j > t, \, j = 4, \ldots, N \mid \text{at least one } X_i > t)$$

$$+ P(X_1 > t, \, X_2 < t, \, X_3 > t, \, X_4 < t, \, X_j > t,$$
$$j = 5, \ldots, N \mid \text{at least one } X_i > t)$$
$$+ \cdots$$
$$+ P(X_1 > t, \, X_j > t, \, j = 2, \ldots, N-2, \, X_{N-1} < t,$$
$$X_N < t \mid \text{at least one } X_i > t)]$$

$$+ \cdots + r_1(t)\frac{1}{1}\,\mathrm{P}(X_1 > t,\, X_j < t,\, j = 2,\ldots,N \mid \text{at least one } X_i > t)$$

$$= \mathrm{E}\!\left[\frac{M_1'}{M_1' + M_2' + \cdots + M_N'} \,\middle|\, N' > 0\right] r_1(t),$$

where N' is the total number of items which survived up to time t and M_i', $i = 1, 2, \ldots, N$, is 1 if the ith item survived up to time t and 0 otherwise. Now, as the structure of (3) is symmetric,

$$\lim_{\Delta t \to 0} \frac{1}{\Delta t}\,\mathrm{E}\!\left[\frac{\sum_{i=1}^{N} \mathbf{1}(t < X_i \le t + \Delta t)}{\sum_{i=1}^{N} \mathbf{1}(X_i > t)} \,\middle|\, \text{at least one unit is alive at time } t\right]$$

$$= \sum_{i=1}^{N} \mathrm{E}\!\left[\frac{M_i'}{M_1' + M_2' + \cdots + M_N'} \,\middle|\, N' > 0\right] r_i(t).$$

Corollary 1. *For the fixed finite mixed population described above, the proportion-based failure rate at time t is given by*

$$\sum_{i=1}^{k}\Bigg[\frac{1}{1 - \prod_{j=1}^{k}(F_j(t))^{n_j}}$$
$$\times \sum_{m_1=0}^{n_1}\sum_{m_2=0}^{n_2}\cdots\sum_{m_i=1}^{n_i}\cdots\sum_{m_k=0}^{n_k}\left(\frac{m_i}{\sum_{j=1}^{k} m_j}\right)\prod_{j=1}^{k}\binom{n_j}{m_j}(\bar{F}_j(t))^{m_j}(F_j(t))^{n_j-m_j}\Bigg]\lambda_i(t). \tag{4}$$

Proof. In the fixed finite mixed population considered in Theorem 1, we suppose that $r_i(t) = \lambda_1(t)$, $i = 1, 2, \ldots, n_1$; $r_i(t) = \lambda_2(t)$, $i = n_1+1, n_1+2, \ldots, n_1+n_2$; ...; $r_i(t) = \lambda_k(t)$, $i = \sum_{j=1}^{k-1} n_j + 1, \sum_{j=1}^{k-1} n_j + 2, \ldots, \sum_{j=1}^{k-1} n_j + n_k$, where $n_1 + n_2 + \cdots + n_k = N$. Then this setting corresponds to the fixed finite mixed population we are concerned with in this section and the proportion-based failure rate at time t is given by

$$\sum_{i=1}^{k} \mathrm{E}\!\left[\frac{N_i'}{N_1' + N_2' + \cdots + N_k'} \,\middle|\, N' > 0\right]\lambda_i(t),$$

where N_i' is the number of items from subpopulation i in the finite population which has survived up to time t and $N' = \sum_{i=1}^{k} N_i'$. It can be seen that the conditional expectation can be expressed as

$$\mathrm{E}\!\left[\frac{N_i'}{N_1' + N_2' + \cdots + N_k'} \,\middle|\, N' > 0\right]$$
$$= \frac{1}{1 - \prod_{j=1}^{k}(F_j(t))^{n_j}}$$
$$\times \sum_{m_1=0}^{n_1}\sum_{m_2=0}^{n_2}\cdots\sum_{m_i=1}^{n_i}\cdots\sum_{m_k=0}^{n_k}\left(\frac{m_i}{\sum_{j=1}^{k} m_j}\right)\prod_{j=1}^{k}\binom{n_j}{m_j}(\bar{F}_j(t))^{m_j}(F_j(t))^{n_j-m_j},$$

which completes the proof.

In the following discussion, in order to distinguish the proportion-based failure rate in a finite population from the ordinary failure rate $\lambda(t)$, we will use the notation $\lambda_{F_N}(t)$.

The following example illustrates the difference between $\lambda(t)$ and $\lambda_{F_N}(t)$.

Example 1. Consider a finite mixed population with two subpopulations, and let $n_1 = n_2 = 1$. Suppose that the corresponding failure rate functions are given by $\lambda_1(t) = \lambda_1$, $t > 0$, and $\lambda_2(t) = \lambda_2$, $t > 0$. Then the failure rate function in (2) is defined by

$$\lambda(t) = \frac{e^{-\lambda_1 t}}{e^{-\lambda_1 t} + e^{-\lambda_2 t}}\lambda_1 + \frac{e^{-\lambda_2 t}}{e^{-\lambda_1 t} + e^{-\lambda_2 t}}\lambda_2.$$

Now we consider the proportion-based failure rate in Corollary 1. In this case, it is not difficult to verify that the conditional expectations are given by

$$\mathrm{E}\left[\frac{N_1'}{N_1' + N_2'} \;\middle|\; N_1' + N_2' > 0\right] = \frac{(1/2)e^{-\lambda_1 t}e^{-\lambda_2 t} + e^{-\lambda_1 t}(1 - e^{-\lambda_2 t})}{e^{-\lambda_1 t} + e^{-\lambda_2 t} - e^{-\lambda_1 t}e^{-\lambda_2 t}}$$

and

$$\mathrm{E}\left[\frac{N_2'}{N_1' + N_2'} \;\middle|\; N_1' + N_2' > 0\right] = \frac{(1/2)e^{-\lambda_1 t}e^{-\lambda_2 t} + e^{-\lambda_2 t}(1 - e^{-\lambda_1 t})}{e^{-\lambda_1 t} + e^{-\lambda_2 t} - e^{-\lambda_1 t}e^{-\lambda_2 t}}.$$

Thus, we have

$$\lambda_{F_N}(t) = \frac{(1/2)e^{-\lambda_1 t}e^{-\lambda_2 t} + e^{-\lambda_1 t}(1 - e^{-\lambda_2 t})}{e^{-\lambda_1 t} + e^{-\lambda_2 t} - e^{-\lambda_1 t}e^{-\lambda_2 t}}\lambda_1 + \frac{(1/2)e^{-\lambda_1 t}e^{-\lambda_2 t} + e^{-\lambda_2 t}(1 - e^{-\lambda_1 t})}{e^{-\lambda_1 t} + e^{-\lambda_2 t} - e^{-\lambda_1 t}e^{-\lambda_2 t}}\lambda_2,$$

and, therefore, $\lambda(t) \neq \lambda_{F_N}(t)$.

Remark 1. In Example 1, $\lambda(t)$ corresponds to the proportion-based failure rate function of an infinite mixed population composed of two subpopulations ($\lambda_1(t) = \lambda_1$ and $\lambda_2(t) = \lambda_2$) with equal proportions $\pi_1 = \pi_2 = 0.5$. On the other hand, $\lambda_{F_N}(t)$ is the proportion-based failure rate function of a finite mixed population with the same structure (i.e. with the same subpopulations and the same proportions). Thus, intuitively, these two populations might be expected to share the same probabilistic properties of the lifetime of the components. For example, at time t, we randomly select one item among the survivors in the infinite population and one among the survivors in the finite population. Then, intuitively, reliability characteristics of these two items could be expected to be the same. However, contrary to our intuition, this is not true.

Remark 2. For the setting considered in this section,

$$\pi_i(t) = \frac{\pi_i \bar{F}_i(t)}{\sum_{j=1}^{k} \pi_j \bar{F}_j(t)} = \frac{n_i \bar{F}_i(t)}{\sum_{j=1}^{k} n_j \bar{F}_j(t)} = \frac{\mathrm{E}[N_i']}{\mathrm{E}[N_1' + N_2' + \cdots + N_k']}.$$

Therefore, for the type of finite mixed population considered above, $\lambda_{F_N}(t)$ is, as illustrated in Example 1, different from the failure rate $\lambda(t)$ defined in accordance with the ordinary definition in (1). In order to explain the origin of the difference between $\lambda(t)$ and $\lambda_{F_N}(t)$, it is useful to understand that, 'probabilistically', the two-fold interpretations lead to different types of finite population, and, thus, they define two different random variables and distribution functions. Indeed, in the case of the first interpretation, the items randomly selected at time 0 from the finite mixed population are probabilistically the same as the items randomly selected from an infinite population with the failure rate given in (2). On the other hand, in the case of the

proportion-based failure rate $\lambda_{F_N}(t)$ in (4), this is not true. In fact, $\lambda_{F_N}(t)$ in (4) is defined based on the knowledge of the finite (combinatorial) structure of the survived population (see also Theorem 2), whereas this knowledge of the structure is not taken into consideration by the ordinary failure rate $\lambda(t)$ in (2).

We are now ready to answer the second question raised at the start of this section. It is now obvious that the proneness of future failure of the item selected at time t can be described by (2) with $\lambda_i(t)$, $\bar{F}_i(t)$, and π_i replaced by $\lambda_i(t+u)$, $\bar{F}_i(t+u)/\bar{F}_i(t)$, and

$$
\left[\frac{1}{1 - \prod_{j=1}^{k}(F_j(t))^{n_j}} \right.
$$

$$
\left. \times \sum_{m_1=0}^{n_1} \sum_{m_2=0}^{n_2} \cdots \sum_{m_i=1}^{n_i} \cdots \sum_{m_k=0}^{n_k} \left(\frac{m_i}{\sum_{j=1}^{k} m_j} \right) \prod_{j=1}^{k} \binom{n_j}{m_j} (\bar{F}_j(t))^{m_j} (F_j(t))^{n_j - m_j} \right],
$$

respectively, $i = 1, 2, \ldots, k$. Therefore, the answers to the two questions for the fixed finite mixed population are 'no'.

3. Randomly selected finite mixed population

In this section we consider a finite population constructed from an infinite mixed population which is composed of k subpopulations (subpopulation 1, subpopulation 2, etc.), with the corresponding proportions π_1, π_2, ..., π_k, respectively. The failure rate function, CDF, and PDF of the corresponding subpopulations are given by $\lambda_i(t)$, $F_i(t) = 1 - \bar{F}_i(t)$, and $f_i(t)$, $i = 1, 2, \ldots, k$, respectively. Now a finite population of size N (fixed constant) is constructed by randomly selecting items from the infinite mixed population. Thus, we now have a *randomly selected finite mixed population*. Note that the failure rate function for the original 'infinite mixed population' described above is given by (2). Let N_i be the number of items from subpopulation i in this randomly selected finite mixed population, $i = 1, 2, \ldots, k$. Then, obviously, $N_i \sim \text{Bin}(N, \pi_i)$ and the conditional survival function for this finite population is given by

$$
\bar{F}(t \mid N_i, \, i = 1, 2 \ldots, k) = \sum_{i=1}^{k} \frac{N_i}{N} \bar{F}_i(t).
$$

Thus, the marginal survival function of this finite population is given by

$$
\bar{F}(t) = \sum_{i=1}^{k} \text{E} \left[\frac{N_i}{N} \right] \bar{F}_i(t) = \sum_{i=1}^{k} \pi_i \bar{F}_i(t).
$$

The following theorem implies that the proportion-based failure rate function for the above randomly selected finite population $\lambda_{F_N}(t)$ is also given by $\lambda(t)$ in (2).

Theorem 2. *For the randomly selected finite mixed population described above, the proportion-based failure rate function $\lambda_{F_N}(t)$ is given by*

$$
\lambda_{F_N}(t) = \sum_{i=1}^{k} \pi_i(t) \lambda_i(t), \tag{5}
$$

where

$$
\pi_i(t) = \frac{\pi_i \bar{F}_i(t)}{\sum_{j=1}^{k} \pi_j \bar{F}_j(t)}, \qquad i = 1, 2, \ldots, k.
$$

Proof. Let N_i' be the number of units from subpopulation i, $i = 1, 2, \ldots, k$, in the finite population which has survived up to time t, and let $N' \equiv \sum_{i=1}^{k} N_i'$. As the survival function of an item randomly selected from the infinite mixed population is given by

$$\bar{F}(t) = \pi_1 \bar{F}_1(t) + \pi_2 \bar{F}_2(t) + \cdots + \pi_k \bar{F}_k(t),$$

it follows that $N' \sim \text{Bin}(N, \bar{F}(t))$. Thus, the probability mass function of N' given $N' > 0$ is

$$\text{P}(N' = n \mid N' > 0) = \frac{\binom{N}{n}(\bar{F}(t))^n (F(t))^{N-n}}{1 - (F(t))^N}, \qquad n = 1, 2, \ldots, N.$$

On the other hand, given $N' = n$, $n = 1, 2, \ldots, N$, the conditional distribution of N_i' is given by

$$\text{P}(N_i' = m \mid N' = n)$$
$$= \binom{n}{m} \left(\frac{\pi_i \bar{F}_i(t)}{\sum_{j=1}^{k} \pi_j \bar{F}_j(t)} \right)^m \left(1 - \frac{\pi_i \bar{F}_i(t)}{\sum_{j=1}^{k} \pi_j \bar{F}_j(t)} \right)^{n-m}, \qquad m = 1, 2, \ldots, n,$$

for $i = 1, 2, \ldots, k$. Thus, we have

$$\text{E}\left[\frac{N_i'}{N'} \,\Big|\, N' = n \right] = \frac{\pi_i \bar{F}_i(t)}{\sum_{j=1}^{k} \pi_j \bar{F}_j(t)}, \qquad i = 1, 2, \ldots, k,$$

and, therefore, the conditional expectation is given by

$$\text{E}\left[\frac{N_i'}{N'} \,\Big|\, N' > 0 \right] = \sum_{n=1}^{N} \text{E}\left[\frac{N_i'}{N'} \,\Big|\, N' = n \right] \text{P}(N' = n \mid N' > 0)$$
$$= \frac{\pi_i \bar{F}_i(t)}{\sum_{j=1}^{k} \pi_j \bar{F}_j(t)}, \qquad i = 1, 2, \ldots, k,$$

which completes the proof.

The intuitive meaning of Theorem 2 is as follows. We select N items at random from the infinite mixed population. As the N_is are not observed (without any further information), each has the failure rate given in (2). In fact, this infinite population can be considered homogeneous with items having the failure rate given in (2). As mentioned before, for an infinite population, the surviving item's failure rate at time t is then given by (2) whether it is selected at time 0 (the first interpretation) or at time t among survivors (the second interpretation). Therefore, $\lambda_{F_N}(t)$ in (5) is the same as the failure rate given in (2).

Remark 3. It is now clear that, for the randomly selected finite mixed population considered in this section, the two questions raised in Section 2 can be answered in the affirmative.

Remark 4. Suppose that we have an infinite mixed population composed of k infinite subpopulations (subpopulation 1, subpopulation 2, etc.) with corresponding proportions

$$\pi_1 = \frac{n_1}{N}, \qquad \pi_2 = \frac{n_2}{N}, \ldots, \qquad \pi_k = \frac{n_k}{N}. \tag{6}$$

The reliability measures for each subpopulation are assumed to be the same as before. From this infinite population, finite mixed populations are constructed in the following two models.

Model A. From subpopulation i, n_i, $i = 1, 2, \ldots, k$, items are randomly selected and all of these selected items from the k subpopulations constitute a finite mixed population of size N, where N is a fixed constant and $N = n_1 + n_2 + \cdots + n_k$.

Model B. A finite mixed population of size N (fixed constant) is constructed by randomly selecting items from the infinite mixed population.

Then, the proportion-based failure rate functions for Models A and B are given by (4) and (5), respectively, with the π_is given in (6). Note that, for both of the above two types of finite mixed population (i.e. Models A and B), the corresponding survival functions are given by $\bar{F}(t) = \sum_{i=1}^{k} \pi_i \bar{F}_i(t)$. However, the proportion-based failure rate functions are different, depending on the way in which the finite populations are constructed. In the case of infinite populations there is one-to-one correspondence between the survival function and the proportion-based failure rate, that is, for the same survival function, the same proportion-based failure rate is obtained. This is not true in the case of finite populations. In finite mixed populations, the survival function does not uniquely define the proportion-based failure rate. How the finite population is constructed is also a crucial factor.

4. Concluding remarks

In this paper we considered two types of finite mixed population: fixed and randomly selected finite mixed populations. For these two types of finite mixed population, we considered proper definitions of the proportion-based failure rate function, and investigated the relationships between these definitions and those for infinite populations. Defining the proportion-based failure rate function of a fixed finite mixed population in terms of the failure rate functions of the subpopulations, we showed that the use of failure rate function defined for the infinite population instead of the appropriately proposed proportion-based failure rate function may give misleading information. The proportion-based failure rate function for a randomly selected finite mixed population was also discussed. Lastly, we revealed the important fact that the survival function of a finite population does not uniquely define the corresponding proportion-based failure rate function.

In this paper, for convenience, we described and discussed our problems from the viewpoint of reliability theory. However, applications of the results obtained in this paper should not be limited to reliability only but could also be generally applied to areas relevant to lifetime modeling and analysis, such as survival analysis, actuarial sciences, and demographic and biological disciplines.

Acknowledgements

References

AVEN, T. AND JENSEN, U. (1999). *Stochastic Models in Reliability*. Springer, New York.

BLOCK, H. W., SAVITS, T. H. AND WONDMAGEGNEHU, E. T. (2003). Mixtures of distributions with increasing linear failure rates. *J. Appl. Prob.* **40,** 485–504.

CHA, J. H. (2000). On a better burn-in procedure. *J. Appl. Prob.* **37,** 1099–1103.

CHA, J. H. (2001). Burn-in procedures for a generalized model. *J. Appl. Prob.* **38,** 542–553.

CHA, J. H. (2003). A further extension of the generalized burn-in model. *J. Appl. Prob.* **40,** 264–270.

EBRAHIMI, N. (1996). Engineering notion of mean-residual-life & hazard-rate for finite populations with known distributions. *IEEE Trans. Reliab.* **45,** 362–368.

EBRAHIMI, N. (2004). Burn-in and covariates. *J. Appl. Prob.* **41,** 735–745.

FINKELSTEIN, M. (2008). *Failure Rate Modelling for Reliability and Risk*. Springer, London.

FINKELSTEIN, M. (2009). Understanding the shape of the mixture failure rate (with engineering and demographic applications). *Appl. Stoch. Models Business Industry* **25,** 643–663.

FINKELSTEIN, M. S. AND ESAULOVA, V. (2001). Modelling a failure rate for the mixture of distribution functions. *Prob. Eng. Inf. Sci.* **15,** 383–400.

JIANG, R. AND MURTHY, D. N. P. (1995). Modelling failure data by mixture of two Weibull distributions: a graphical approach. *IEEE Trans. Reliab.* **44,** 477–488.

MI, J. (1994). Burn-in and maintenance policies. *Adv. Appl. Prob.* **26,** 207–221.

NAVARRO, J. AND HERNANDEZ, P. J. (2004). How to obtain bathtub-shaped failure rate models from normal mixtures. *Prob. Eng. Inf. Sci.* **18,** 511–531.

VAUPEL, J. W. AND ZHANG, Z. (2010). Attrition in heterogeneous cohorts. *Demographic Res.* **23,** 737–748.