

Chapter 5

Heterogeneous Populations

Homogeneity of objects is a unique property that is very rare in nature and in industry. It can be created in the laboratory, but not outside it. Therefore, one can hardly find homogeneous populations in real life; however, most of reliability modeling deals with homogeneous cases. Due to instability of production processes, environmental and other factors, most populations of manufactured items in real life are heterogeneous. Similar considerations are obviously true for biological items (organisms). Neglecting heterogeneity can lead to serious errors in reliability assessment of items and, as a consequence, to crucial economic losses. Stochastic analysis of heterogeneous populations presents a significant challenge to developing mathematical descriptions of the corresponding reliability indices. On the other hand, everything depends on the definition, on what we understand by homogeneous and heterogeneous populations. From the statistical point of view, these terms mean the following.

In homogeneous populations, the lifetimes of items form a sequence of independent and identically distributed random variables (i.i.d.) with the common Cdf $F(t)$ pdf $f(t)$, and the failure rate, $\lambda(t)$. However, due to instability of production processes, environmental and other factors, most populations of manufactured items in real life (and biological organisms in nature as well) are heterogeneous. This means that these populations can be often considered as a *finite or non-finite collection of homogeneous subpopulations* [which are frequently ordered in some suitable stochastic sense, e.g., in the sense of the hazard rate ordering (2.70)].

As an illustrative discrete example, we can think about the collection of $n = 2$ subpopulations of statistically identical items produced at different facilities and mixed together in one population. Assume for simplicity, that each subpopulation consists of a sufficiently large (infinite) number of items. Let the first subpopulation be described by the failure rate $\lambda(t)$ (baseline failure rate), whereas the second subpopulation, due to the better production quality has a smaller failure rate $k\lambda(t)$, where k is a fixed constant such that $0 < k < 1$. Let the proportions of both subpopulations in the population be π_1 and π_2 , $\pi_1 + \pi_2 = 1$. An item is selected at random from the described heterogeneous population and therefore, we do not know to which subpopulation it belongs (although the proportions can be

known at some instances). This choice can be described by the discrete random variable Z (unobserved) with the possible values “1” and “ k ” and the corresponding probability masses $\pi(1) = \pi_1$, $\pi(k) = \pi_2$. Based on the description of Z , the failure rates of the subpopulation with $Z = z$ can be now specified as $\lambda(t, z)$: $\lambda(t, 1) = \lambda(t)$ and $\lambda(t, k) = k\lambda(t)$. In the literature, the random variable Z is often called “frailty”. Frailty describes the susceptibility to failures of items from different ordered subpopulations. Various frailty models have been studied in numerous statistical publications. However, as most of the settings that were considered in reliability theory and practice are homogeneous, the concept of frailty has not been sufficiently elaborated in the reliability literature so far.

Instability of production processes, environmental and other factors can obviously result in more than $n = 2$ ‘quality levels’ and in the continuous frailty model as well. Let, as previously, $\lambda(t)$ denote now the failure rate of some baseline subpopulation. For illustration of the continuous frailty concept, consider the multiplicative (proportional) frailty model. In this model, the failure rates of all other subpopulations are defined as $\lambda(t, z) \equiv z\lambda(t)$, where z is the realization of Z with support, e.g., in $[0, \infty)$. Thus, the failure rate is larger (smaller) for larger (smaller) values of z and we see here the explicit ordering of the corresponding subpopulations in the sense of the hazard rate ordering (2.70). The frailty Z is now the continuous random variable. The term “frailty” was introduced in Vaupel et al. [63] for the gamma-distributed frailty Z . It is worth noting, however, that this specific case of the gamma-frailty model was, in fact, first considered by the British actuary Robert Beard [7, 8].

Mixtures of distributions usually present an effective mathematical tool for modeling heterogeneity, especially when we are interested in the failure rate, which is the conditional characteristic. The introductory Sect. 2.3 was devoted to the shape of the failure rate in the homogeneous setting, which is really important in many applications (reliability, demography, risk analysis, etc.). In heterogeneous populations, the analysis of the shape of the mixture (population) failure rate starts to be even more meaningful. It is well known, e.g., that mixtures of decreasing failure rate (DFR) distributions are always DFR [6]. On the other hand, mixtures of increasing failure rate (IFR) distributions can decrease, at least in some intervals of time. Note that the IFR distributions are often used to model lifetimes governed by the aging processes. Therefore, the operation of mixing can dramatically change the pattern of population aging, e.g., from positive aging (IFR) to negative aging (DFR).

In Sects. 5.1–5.6, on the basis of Finkelstein [28, 29], we will present a brief survey of results relevant for our further discussion in this and in the subsequent chapters. In the rest of this chapter, some new applications of the mixture failure rate modeling will be considered.

5.1 Failure Rate of Mixture of Two Distributions

Suppose, for instance, that a population of some manufactured items consists of items with and without manufacturing defects. The time to failure of an item picked up at random from this population can be obviously described in terms of mixtures. We start with a mixture of two lifetime distributions $F_1(t)$ and $F_2(t)$ with the pdfs $f_1(t)$ and $f_2(t)$ and failure rates $\lambda_1(t)$ and $\lambda_2(t)$, respectively, whereas the Cdf, pdf, and the failure rate of the mixture itself are denoted by $F_m(t)$, $f_m(t)$ and $\lambda_m(t)$, accordingly.

Let the masses π and $1 - \pi$ define the discrete mixture distribution. The mixture survival function and the mixture pdf are

$$\begin{aligned}\bar{F}_m(t) &= \pi\bar{F}_1(t) + (1 - \pi)\bar{F}_2(t), \\ f_m(t) &= \pi f_1(t) + (1 - \pi)f_2(t),\end{aligned}\tag{5.1}$$

respectively. In accordance with the definition of the failure rate (2.4), the mixture failure rate in this case is

$$\lambda_m(t) = \frac{\pi f_1(t) + (1 - \pi)f_2(t)}{\pi\bar{F}_1(t) + (1 - \pi)\bar{F}_2(t)}.$$

As $\lambda_i(t) = f_i(t)/\bar{F}_i(t)$, $i = 1, 2$, this can be transformed into

$$\lambda_m(t) = \pi(t)\lambda_1(t) + (1 - \pi(t))\lambda_2(t),\tag{5.2}$$

where the time-dependent probabilities are

$$\pi(t) = \frac{\pi\bar{F}_1(t)}{\pi\bar{F}_1(t) + (1 - \pi)\bar{F}_2(t)}, \quad 1 - \pi(t) = \frac{(1 - \pi)\bar{F}_2(t)}{\pi\bar{F}_1(t) + (1 - \pi)\bar{F}_2(t)},\tag{5.3}$$

It follows from Eq. (5.2) that $\lambda_m(t)$ is contained between $\min\{\lambda_1(t), \lambda_2(t)\}$ and $\max\{\lambda_1(t), \lambda_2(t)\}$. Specifically, if the failure rates are ordered as $\lambda_1(t) \leq \lambda_2(t)$, then

$$\lambda_1(t) \leq \lambda_m(t) \leq \lambda_2(t).$$

Differentiating (5.1) results in [51]:

$$\lambda'_m(t) = \pi(t)\lambda'_1(t) + (1 - \pi(t))\lambda'_2(t) - \pi(t)(1 - \pi(t))(\lambda_1(t) - \lambda_2(t))^2.\tag{5.4}$$

Assume that $\lambda_i(t)$ $i = 1, 2$ are DFR. Then the mixture failure rate is also decreasing, which is the well-known fact for general mixtures [6].

As $\bar{F}_i(0) = 1$, $i = 1, 2$, the initial value of the mixture failure rate ($t = 0$) is just the ‘ordinary’ mixture of initial values of the two failure rates, i.e.,

$$\lambda_m(0) = \pi\lambda_1(0) + (1 - \pi)\lambda_2(0).$$

When $t > 0$, the conditional probabilities $\pi(t)$ and $1 - \pi(t)$ are obviously not equal to π and $1 - \pi$, respectively. Assume that $\lambda_1(t) \leq \lambda_2(t)$. Dividing the numerator and the denominator in the first equation in (5.3) by $\bar{F}_1(t)$ it is easy to see that the proportion of the survived up to t items in the mixed population, i.e., $\pi(t)$ is increasing ($(1 - \pi(t))$ is decreasing). This effect can be meaningfully interpreted in the following way: *the weakest items are dying out first*. Therefore,

$$\lambda_m(t) < \pi\lambda_1(t) + (1 - \pi)\lambda_2(t), \quad t > 0. \quad (5.5)$$

Thus, $\lambda_m(t)$ is always smaller than the expectation $\pi\lambda_1(t) + (1 - \pi)\lambda_2(t)$.

Assume now that both $\lambda_1(t)$ and $\lambda_2(t)$ are increasing for $t \geq 0$. Can the mixture failure rate initially (at, least, for small t) decrease? Equation (5.4) helps us to give the positive answer to this question. The corresponding sufficient condition is

$$\pi\lambda'_1(t) + (1 - \pi)\lambda'_2(t) - \pi(1 - \pi)(\lambda_1(0) - \lambda_2(0))^2 < 0, \quad (5.6)$$

where the derivatives are obtained at $t = 0$. Inequality (5.6), e.g., means that if $|\lambda_1(0) - \lambda_2(0)|$ is sufficiently large, then the mixture failure rate is initially decreasing no matter how fast the failure rates $\lambda_1(t)$ and $\lambda_2(t)$ are increasing in the neighborhood of 0, which is a remarkable fact, indeed. Let, for instance,

$$\lambda_1(t) = c_1t + a_1, \quad \lambda_2(t) = c_2t + a_2, \quad 0 < c_1 < c_2, \quad 0 < a_1 < a_2,$$

Then, if

$$a_2 - a_1 > \left(\frac{\pi c_1 + (1 - \pi)c_2}{\pi(1 - \pi)} \right)^{1/2},$$

$\lambda_m(t)$ is initially decreasing.

What about the asymptotic (for large t) behavior of $\lambda_m(t)$? Due to the *weakest populations are dying first* principle the intuitive guess would be: the mixture failure rate tends (in some suitable sense) to the failure rate of the strongest population as $t \rightarrow \infty$. Block and Joe [13] give some general conditions for this convergence. We will just consider here an important specific case of proportional failure rates that allows formulating these conditions explicitly:

$$\lambda_1(t) \equiv \lambda(t, z_1) = z_1\lambda(t), \quad \lambda_2(t) \equiv \lambda(t, z_2) = z_2\lambda(t), \quad z_2 > z_1,$$

where $\lambda(t)$ is some baseline failure rate. We will distinguish between the convergence

$$\lambda_m(t) - \lambda(t, z_1) \rightarrow 0 \text{ as } t \rightarrow \infty \quad (5.7)$$

and the asymptotic equivalence

$$\lambda_m(t) = \lambda(t, z_1)(1 + o(1)) \text{ as } t \rightarrow \infty, \quad (5.8)$$

which will mostly be used in the following alternative notation: $\lambda_m(t) \sim \lambda(t, z_1)$ as $t \rightarrow \infty$.

When $\lambda(t)$ has a finite limit as $t \rightarrow \infty$, these relationships coincide. The following theorem [32] specifies the corresponding conditions:

Theorem 5.1 Consider the mixture model (5.1)–(5.3), where

$$\lambda(t, z_1) = z_1\lambda(t), \quad \lambda(t, z_2) = z_2\lambda(t); \quad z_2 > z_1 > 0,$$

and $\lambda(t) \rightarrow \infty$ as $t \rightarrow \infty$. Then

- Relationship (5.8) holds;
- Relationship (5.7) holds if

$$\lambda(t) \exp\{-(z_2 - z_1) \int_0^t \lambda(u) du\} \rightarrow 0 \text{ as } t \rightarrow \infty. \quad (5.9)$$

The proof is straightforward and is based on considering the quotient $\lambda_m(t)/\lambda(t, z_1)$ as in Block and Joe [13]. □

Condition (5.9) is a rather weak one. In essence, it states that the pdf of a distribution with an ultimately increasing failure rate tends to 0 as $t \rightarrow \infty$. All distributions that are typically used in lifetime data analysis meet this requirement.

Similar reasoning can be used for describing the shape of the failure rate for the mixture of $n > 2$ distributions [13, 28].

We have described some approaches to analyze the general pattern of the shape of the mixture failure rate for two distributions focusing on initial and tail behavior. The concrete shapes can be versatile. We will just present here a few examples. More information on specific shapes of the mixture failure rate of two distributions can be found in Gurland and Sethuraman [40], Gupta and Waren [39], Block et al. [14, 18], Lai and Xie [43], Navarro and Hernandez [51], Finkelstein [28], and Block et al. [16]. Note that the different shapes of the mixture mortality rate were analyzed in various demographic applications.

- As follows from Gupta and Waren [39], the mixture of two gamma distributions with increasing failure rates (with the same scale parameter) can result either in the increasing mixture failure rate or in the modified bathtub (MBT) mixture failure rate (it first increases and then behaves like a bathtub (BT) failure rate). This shape agrees with our general reasoning of this section, as it can be easily verified that condition (5.6) does not hold in this case and therefore the initial decreasing is not possible.
- Similar shapes occur for the mixtures of two Weibull distributions with increasing failure rates. Note that in this case, MBT shape results when p in Eq. (5.1) is less than some ξ , $0 < \xi < 1$ and the mixture failure rate increases for $p \geq \xi$.

- Navarro and Hernandez [51] state that the mixture failure rate of two truncated normal distributions (we are dealing with lifetime random variables), depending on parameters involved, can also be increasing, BT-shaped or MBT-shaped. The BT shape obtained via the generalized mixtures (when p is a real number and not necessarily $p \in [0, 1]$) where studied in Navarro and Hernandez [52].
- Block et al. [18] give explicit conditions which describe the possible shapes of the mixture failure rate for two increasing linear failure rates. Again the possible shapes in this case are IFR, BT, and MBT (for the non-crossing linear failure rates).
- Block et al. [16] present an interesting generalization when one of the distributions is itself a continuous mixture of exponentials (and therefore, decreasing) and the other is a gamma distribution. It is shown that for the specific values of parameters involved the mixture failure rate has a BT shape. In essence, these authors are ‘constructing’ the BT shape using the specifically decreasing in $(0, \infty)$ to $\zeta > \lambda_0 > 0$ failure rate of the first distribution and the increasing to λ_0 failure rate of the second distribution. Note that, as follows from (5.3), $\lambda_m(t)$ is contained between these two failure rates. Block et al. [16] also prove that mixtures of DFR gamma distributions with an IFR gamma distribution are bathtub-shaped and mixtures of modified Weibull distributions (the failure rate is decreasing not to 0, as for ‘ordinary’ Weibull distribution, but to ζ) with an IFR gamma distribution have also the bathtub-shaped failure rate.

5.2 Continuous Mixtures

Let Z be now a continuous mixing random variable (frailty) with support in $[0, \infty)$ and the pdf $\pi(z)$. Other intervals of support can be also considered. Similar to the previous section, the mixture survival function and the mixture pdf are defined as the following expectations:

$$\begin{aligned}\bar{F}_m(t) &= \int_0^{\infty} \bar{F}(t, z)\pi(z)dz, \\ f_m(t) &= \int_0^{\infty} f(t, z)\pi(z)dz,\end{aligned}\tag{5.10}$$

respectively, where the notation for conditional functions $\bar{F}(t|Z = z) = \bar{F}(t, z)$ and $f(t|Z = z) = f(t, z)$ means that a lifetime distribution is indexed by parameter z . The corresponding conditional failure rate is denoted by $\lambda(t, z)$, whereas the mixture (observed) failure rate is

$$\lambda_m(t) = \frac{\int_0^{\infty} f(t, z)\pi(z)dz}{\int_0^{\infty} \bar{F}(t, z)\pi(z)dz}.\tag{5.11}$$

Equation (5.11) can be transformed to [47]:

$$\lambda_m(t) = \int_0^\infty \lambda(t, z)\pi(z|t)dz, \quad \pi(z|t) = \frac{\pi(z)\bar{F}(t, z)}{\int_0^\infty \bar{F}(t, z)\pi(z)dz}, \quad (5.12)$$

where $\pi(z|t)$ denotes the conditional pdf of Z on condition that $T > t$, i.e., an item described by a lifetime T with the Cdf $F_m(t)$ had survived in $[0, t]$. Denote this random variable by $Z|t$. Obviously the masses $\pi(t)$ and $1 - \pi(t)$ in (5.1) correspond to $\pi(z|t)$ in the continuous case.

Under the mild assumptions (see Theorem 5.2), a property that is similar to the discrete case (5.5) holds for the continuous case as well, i.e.,

$$\lambda_m(t) < \lambda_P(t) \equiv \int_0^\infty \lambda(t, z)\pi(z)dz, \quad t > 0; \quad \lambda_m(0) = \lambda_P(t) \quad (5.13)$$

meaning that the mixture failure rate is always smaller than the ‘ordinary’ expectation. Thus, owing to conditioning, the mixture failure rate is smaller than the unconditional one for each $t > 0$, which, as in the discrete case, can be interpreted via *the weakest populations are dying out first* principle. As time increases, those subpopulations that have larger failure rates have larger chances of dying and, therefore, the proportion of subpopulations with a smaller failure rate increases.

The following theorem [33] states also the condition for $\lambda_P(t) - \lambda_m(t)$ to increase:

Theorem 5.2 *Let the failure rate $\lambda(t, z)$ be differentiable with respect to both arguments and be ordered as*

$$\lambda(t, z_1) < \lambda(t, z_2), \quad z_1 < z_2, \forall z_1, z_2 \in [a, b], \quad t \geq 0. \quad (5.14)$$

Then

- Inequality (5.13) holds;
- If, additionally, $\partial\lambda(t, z)/\partial z$ is increasing in t , then $\lambda_P(t) - \lambda_m(t)$ is increasing.

We will consider now two important applications specific in cases of model (5.12). Let $\lambda(t, z)$ be indexed by parameter z in the following additive way:

$$\lambda(t, z) = \lambda(t) + z, \quad (5.15)$$

where $\lambda(t)$ is a deterministic, continuous, and positive function for $t > 0$. It can be viewed as some baseline failure rate. Equation (5.15) defines for $z \in [0, \infty)$ a family of ‘horizontally parallel’ functions. We will be interested in an increasing $\lambda(t)$. Applying (5.12) to this model results in

$$\lambda_m(t) = \lambda(t) + \frac{\int_0^\infty z \bar{F}(t, z) \pi(z) dz}{\int_0^\infty \bar{F}(t, z) \pi(z) d\theta} = \lambda(t) + E[Z|t], \quad (5.16)$$

where, in accordance with (5.12), $E[Z|t]$ denotes the expectation of the random variable $Z|t$. It can be easily shown by direct derivation that $E'[Z|t] = -\text{Var}(Z|t) < 0$. Differentiating (5.16) and using this property, we obtain the following result [32, 47].

Theorem 5.3 *Let $\lambda(t)$ be an increasing, convex function in $[0, \infty)$. Assume that $\text{Var}(Z|t)$ is decreasing in $t \in [0, \infty)$ and*

$$\text{Var}(Z|0) > \lambda'(0).$$

Then $\lambda_m(t)$ decreases in $[0, c)$ and increases in $[c, \infty)$, where c can be uniquely defined from the following equation:

$$\text{Var}(Z|t) = \lambda'(t).$$

It follows from this theorem that the corresponding model of mixing results in the bathtub shape of the mixture failure rate: it first decreases and then increases, converging to the failure rate of the strongest population, which is $\lambda(t)$ in our case. It seems that the conditional variance $\text{Var}(Z|t)$ should decrease, as the “weak populations are dying out first” when t increases. It turns out, however, that this intuitive reasoning is not true for the general case and some specific distributions can result in initially increasing $\text{Var}(Z|t)$. The corresponding counter-example can be found in Finkelstein and Esaulova [32]. It is also shown that $\text{Var}(Z|t)$ is always decreasing in $[0, \infty)$ when Z is gamma-distributed.

The most popular and elaborated applications model of mixing is the multiplicative one:

$$\lambda(t, z) = z \lambda(t), \quad (5.17)$$

where, as previously, the baseline $\lambda(t)$ is a deterministic, continuous, and positive function for $t > 0$. In survival analysis, Eq. (5.17) is usually called a multiplicative frailty model (proportional hazards). The mixture failure rate in this case is

$$\lambda_m(t) = \int_0^\infty \lambda(t, z) \pi(z|t) dz = \lambda(t) E[Z|t]. \quad (5.18)$$

Differentiating both sides gives

$$\lambda'_m(t) = \lambda'(t) E[Z|t] + \lambda(t) E'[Z|t]. \quad (5.19)$$

Thus, when $\lambda(0) = 0$, the failure rate $\lambda_m(t)$ increases in the neighborhood of $t = 0$. Further behavior of this function depends on the other parameters involved. Similar to the additive case, $E'[Z|t] = -\lambda(t) \text{Var}(Z|t) < 0$, which means that $E[Z|t]$ is decreasing in t [38]. Therefore, it follows from Eq. (5.18) that the function

$\lambda_m(t)/\lambda(t)$ is a decreasing one, which imply that $\lambda(t)$ and $\lambda_m(t)$ cross at most at only one point. It immediately follows from Eq. (5.19) that when $\lambda(t)$ is decreasing, $\lambda_m(t)$ is also decreasing (another proof of this well-known property). When $\lambda(0) \neq 0$ and

$$\frac{\lambda'(0)}{\lambda^2(0)} \leq \frac{\text{Var}(Z)}{E[Z]},$$

the mixture failure rate is decreasing in $[0, \varepsilon)$, $\varepsilon > 0$ meaning, e.g., that for the fixed $E[Z]$ the variance of Z should be sufficiently large.

Asymptotic behavior of $\lambda_m(t)$ as $t \rightarrow \infty$ for this and other (more general models will be discussed in Sect. 5.4). Note that, the accelerated life model (ALM) to be studied in this section does not allow the foregoing reasoning based on considering expectation $E[Z|t]$.

5.3 Examples

5.3.1 Weibull and Gompertz Distributions

Consider multiplicative frailty model (5.17). Let Z be a gamma-distributed random variable with shape parameter α and scale parameter β and let $\lambda(t) = \gamma t^{\gamma-1}$, $\gamma > 1$ be the increasing failure rate of the Weibull distribution, $\lim_{t \rightarrow \infty} \lambda(t) = \infty$. The mixture failure rate $\lambda_m(t)$ in this case, can be obtained by the direct integration, as in Finkelstein [28] (see also [38]):

$$\lambda_m(t) = \frac{\alpha \beta \gamma t^{\gamma-1}}{1 + \beta t^\gamma}. \quad (5.20)$$

The shape of the mixture failure rate differs dramatically from the shape of the increasing baseline failure rate $\lambda(t)$. Thus $\lambda_m(t)$ is equal to 0 at $t = 0$, increases to a maximum at

$$t_{\max} = \left(\frac{\gamma - 1}{\beta} \right)^{\frac{1}{\gamma}}$$

and then decreases to 0 as $t \rightarrow \infty$ (Fig. 5.1).

Weibull distribution with $\gamma > 1$ is often used for modeling aging processes as its failure rate is increasing. Therefore the mixture model results in the dramatically different shape (the upside-down bathtub shape). This phenomenon should certainly be taken in account in reliability practice.

The described shape of the mixture failure rate was observed for a heterogeneous sample of miniature light bulbs [28]. The failure rate of the homogeneous population of these light bulbs, however, follows the Weibull law. Therefore the observed shape complies with the predicted one.

Fig. 5.1 The mixture failure rate for the Weibull baseline distribution, $\gamma = 2, \alpha = 1$

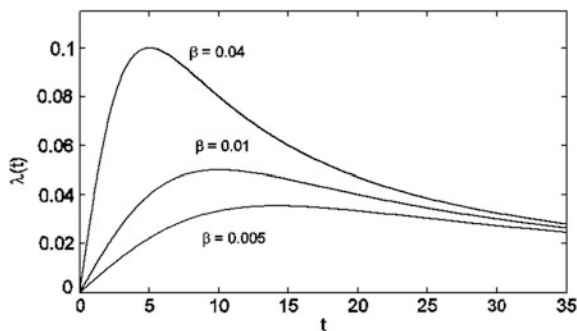
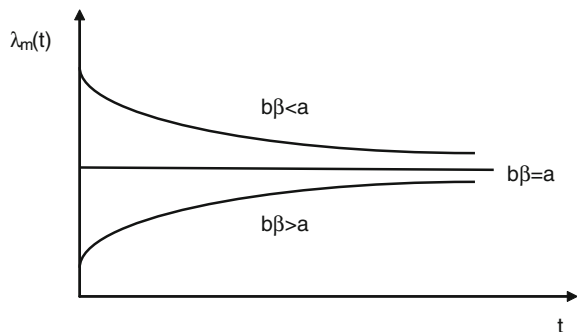


Fig. 5.2 Gamma-Gompertz mixture failure rate



Let again the mixing distribution be the gamma distribution with shape parameter c and scale parameter β , whereas the baseline distribution be the Gompertz distribution with the failure rate $\lambda(t) = a \exp\{bt\}$, $a, b > 0$. Owing to its computational simplicity, the gamma-frailty model is practically the only one widely used in applications so far. Direct computation in accordance with Eq. (5.12) for this baseline failure rate results in

$$\lambda_m(t) = \frac{bc \exp\{bt\}}{\exp\{bt\} + \left(\frac{b\beta}{a} - 1\right)}. \tag{5.21}$$

If $b\beta = a$, then $\lambda_m(t) \equiv bc$. However, if $b\beta > a$, then $\lambda_m(t)$ increases to bc and if $b\beta < a$, it decreases to bc (Fig. 5.2).

Thus, we are mixing exponentially increasing failure rates and as a result obtaining a slowly increasing (decreasing) mixture failure rate, which converges to a constant value.

5.3.2 Reliability Theory of Aging

Consider now a discrete frailty parameter, $Z = N$ with the Cdf $F_0(n) \equiv P(N \leq n)$. We will be interested in the following meaningful reliability interpretation.

Let N be a random number of initially (at $t = 0$) operating independent and identically distributed components with constant failure rates λ . Assume that these components form a parallel system, which, according to Gavrillov and Gavrillova [36], models the lifetime of an organism (generalization to the series-parallel structure is straightforward). These authors also provide a biological justification of the model. In each realization $N = n$, $n \geq 1$, the degradation process of pure death can be defined as just the number of failed components. When this number reaches n , the death of an organism occurs. Denote by $\lambda_n(t)$ the mortality (failure) rate, which describes T_n —the time to death for the fixed $N = n$, $n = 1, 2, \dots$ ($n = 0$ is excluded, as there should be at least one operating component at $t = 0$). It is shown in Gavrillov and Gavrillova [36] that as $t \rightarrow 0$, this mortality rate tends to an increasing power function (the Weibull law), which is a remarkable fact. On the other hand, for random N , similar to (5.2), (5.3) and (5.11, 5.12), the observed (mixture) mortality rate is given as the following conditional expectation with respect to N :

$$\lambda_m(t) = E[\lambda_N(t)|T > t], \quad (5.22)$$

where T , as usual, denotes the lifetime of interest. Therefore, as previously, $\lambda_m(t)$ is a conditional expectation (on condition that the system is operable at t) of a random mortality rate $\lambda_N(t)$. Note that, for small t , this operation can approximately result in the unconditional expectation

$$\lambda_m(t) \approx E[\lambda_N(t)] = \sum_{n=1}^{\infty} P_n \lambda_n(t), \quad (5.23)$$

where $P_n \equiv \Pr[N = n]$, but the limiting transition, as $t \rightarrow 0$, should be performed carefully in this case. As $t \rightarrow \infty$, we observe the following mortality plateau [34]:

$$\lambda_m(t) \rightarrow \lambda. \quad (5.24)$$

This is due to the fact that the conditional probability that only one component with the failure rate λ is operating tends to 1 as $t \rightarrow \infty$ (on condition that the system is operating).

Assume now that N is Poisson distributed with parameter η (on condition that the system is operable at $t = 0$). Therefore

$$P_n = \frac{\exp\{-\eta\} \eta^n}{n!(1 - \exp\{-\eta\})}, \quad n = 1, 2, \dots$$

It can be shown via direct integration that the time to death in our simplified model has the following Cdf [55]:

$$F(t) = \Pr[T \leq t] = \frac{1 - \exp\{-\eta \exp\{-\lambda t\}\}}{1 - \exp\{-\eta\}}. \quad (5.25)$$

The corresponding mixture mortality rate is

$$\lambda_m(t) = \frac{F'(t)}{1 - F(t)} = \frac{\eta\lambda \exp\{-\lambda t\}}{\exp\{\eta \exp\{-\lambda t\}\} - 1}. \quad (5.26)$$

Performing, as $t \rightarrow \infty$, the limiting transition in (5.26), we also arrive at the mortality plateau (5.5).

In fact, the mortality rate given by Eq. (5.26) is far from the exponentially increasing Gompertz law. The Gompertz law can erroneously follow (as in Gavrilov and Gavrilova [36]) from (5.23) if this approximation is used formally, without considering a proper conditioning in (5.23). However, for some specific values of parameters and sufficiently small t , exponential approximation can still hold. The relevant discussion can be found in Steinsaltz and Evans [55].

5.4 Mixture Failure Rate for Large t

The failure (mortality) rate behavior for large t , is important for objects at the last phase of their useful life (e.g., the above mentioned mortality plateaus). Among the first to consider the limiting behavior of mixture failure rates for the continuous mixtures were Clarotti and Spizzichino [23]. They showed that the mixture failure rate for a family of exponential distributions with parameter $\alpha \in [a, \infty)$ converges to the failure rate of the strongest population, which is a in this case. Block et al. [17], Block et al. [14], and Li [44] extended this to a general case (see also [15]). As the approach (and obtained important mathematical results) of these authors is very general and some assumptions are rather restrictive, it does not provide specific asymptotic relationship that can be used in practical analysis for mixed populations. In order to be able to perform this analysis, Finkelstein and Esaulova [33] developed an approach that was applied to reasonably general survival model that allows for explicit asymptotic relationships and covers (as specific cases) three most popular in survival analysis frailty models: additive, proportional, and accelerated life. The main results that were obtained using this approach are discussed below. The corresponding proofs that are quite technical can be found in this paper.

Let $T \geq 0$ be a lifetime with the cdf $F(t)$, pdf $f(t)$, and the failure rate $\lambda(t)$. Let, as previously, these functions be indexed by the realization of the frailty parameter $Z = z$, i.e., $F(t, z)$, $f(t, z)$, $\lambda(t, z)$, respectively. Consider the following general survival model:

$$\Lambda(t, z) = A(z\phi(t)) + \psi(t), \quad (5.27)$$

where $\Lambda(t, z) \equiv \int_0^t \lambda(t, z)$ denotes the corresponding cumulative failure rate and $A(\cdot)$, $\psi(\cdot)$ and $\phi(\cdot)$ are increasing differentiable functions of their arguments. The meaning of relationship (5.27): we perform a scale transformation $\phi(t)$ in the argument of the cumulated failure rate $\Lambda(t)$ and ‘insert’ a frailty parameter. An important feature of the model is that parameter z is a multiplier.

This model includes a number of well-known survival analysis and reliability specific cases, i.e.,

Additive Model: Let

$$A(u) \equiv u, \phi(t) = t, \psi(0) = 0.$$

Then

$$\lambda(t, z) = z + \psi'(t), \quad \Lambda(t, z) = zt + \psi(t). \tag{5.28}$$

PH (multiplicative) Model: Let

$$A(u) \equiv u, \phi(t) = \Lambda(t).$$

Then

$$\begin{aligned} \lambda(t, z) &= z\lambda(t), \\ \Lambda(t, z) &= z\Lambda(t) = z \int_0^t \lambda(u) du. \end{aligned} \tag{5.29}$$

Accelerated Life Model: Let

$$A(u) \equiv \Lambda(u), \phi(t) = t.$$

Then

$$\Lambda(t, z) = \int_0^{zt} \lambda(u) du = \Lambda(zt), \tag{5.30}$$

$$\lambda(t, z) = z\lambda(zt). \tag{5.31}$$

We are interested in asymptotic behavior (as $t \rightarrow \infty$) of $\lambda_m(t)$. For simplicity of notation (and, in fact, not loosing the generality), we will assume further that $\psi(t) = 0$.

Theorem 5.4 *Let the cumulative failure rate $\Lambda(t, z)$ be given by Eq. (5.27) ($\psi(t) = 0$) and let the mixing pdf $\pi(z)$, $z \in [0, \infty)$ be defined as*

$$\pi(z) = z^\alpha \pi_1(z), \tag{5.32}$$

where $\alpha > -1$ and $\pi_1(z)$, $\pi_1(0) \neq 0$ is a function bounded in $[0, \infty)$ and continuous at $z = 0$. Assume also that $\phi(t) \rightarrow \infty$ as $t \rightarrow \infty$ and that $A(s)$ satisfies

$$\int_0^\infty \exp\{-A(s)\} s^\alpha ds < \infty. \tag{5.33}$$

Then

$$\lambda_m(t) \sim (\alpha + 1) \frac{\phi'(t)}{\phi(t)}, \quad (5.34)$$

where, as usual, asymptotic notation $a(t) \sim b(t)$ as $t \rightarrow \infty$ means that $\lim_{t \rightarrow \infty} a(t)/b(t) = 1$. As we had mentioned, another possible notation for (5.34) is $\lambda_m(t) = (\alpha + 1)\phi'(t)/\phi(t)(1 + o(1))$.

The *proof* of this result is cumbersome and is based on Abelian-type theorems for the corresponding asymptotic integrals. That is why the multiplicative form in $A(z\phi(t))$ is so important. □

The specific case of this theorem for the multiplicative model (5.31) was independently considered by Steinsaltz and Wachter [56]. Assumption (5.32) just states the ‘form’ of the admissible mixing distribution and holds for the main lifetime distributions, such as Weibull, gamma, truncated normal, etc. However, it does not hold for a lognormal distribution, as the corresponding asymptote is proportional to $1/z$ when $z \rightarrow 0$. Assumption (5.33) is a very weak one (weaker than just having a finite expectation for a lifetime) and can be omitted in practical analysis.

A crucial feature of this result is that the asymptotic behavior of the mixture failure rate depends only on the behavior of the mixing distribution in the neighborhood of 0 and on the derivative of the logarithm of the scale function $\phi(t)$, i.e.,

$$(\log \phi(t))' = \phi'(t)/\phi(t).$$

When $\pi(0) \neq 0$ and $\pi(z)$ is bounded in $[0, \infty)$, the result does not depend on the mixing distribution at all, as $\alpha = 0$ in this case. Intuitively, the qualitative meaning is quite clear: as $t \rightarrow \infty$, only the most robust survivors are left and in, accordance with (5.27), this corresponds to the small values of z (weak populations are dying out first).

It is easy to see that for the multiplicative model (5.29), Eq. (5.34) reduces to

$$\lambda_m(t) \sim \frac{(\alpha + 1)\lambda(t)}{\int_0^t \lambda(u)du}. \quad (5.35)$$

and to

$$\lambda_m(t) \sim \frac{\alpha + 1}{t} \quad (5.36)$$

for the ALM (5.30), (5.31).

It should be noted that (5.36) is a really surprising result, as the shape of the mixture failure rate for large t does not depend on the baseline distribution $F(t)$. It is also dramatically different from the multiplicative case (5.35). This means that the ‘nature’ of the ALM is such that it ignores’ the baseline distribution for large t .

Comparing (5.35) and (5.36), we see that the latter never results in the asymptotically flat observed failure rate (the mortality plateau in human mortality studies), whereas the multiplicative model can have this possibility, as in the case of the gamma-frailty model for the Gompertz distribution (see Eq. 5.21).

Note that, by direct integration, Eq. (5.21) can be generalized to the case of an arbitrary (absolutely continuous) baseline distribution characterized by the failure rate $\lambda(t)$:

$$\lambda_m(t) = \frac{c\lambda(t)}{\beta + \Lambda(t)} = \frac{c\lambda(t)}{\beta + \int_0^t \lambda(u)du}. \quad (5.37)$$

It is clear that $c = \alpha + 1$ for the gamma pdf and this formula perfectly comply with the general asymptotic result (5.34) and a classical result by Vaupel et al. [63].

Let, for instance, $\pi(z)$ be the uniform density in $[0, 1]$ and let also $\lambda(t) = \exp\{t\}$ ($a, b = 1$ for simplicity of notation). Then $\lambda(t, z) = z \exp\{t\}$ and

$$\int_0^\infty \bar{F}(t, z)\pi(z)dz = \frac{1}{\omega}(1 - \exp\{-\omega\}),$$

$$\int_0^\infty f(t, z)\pi(z)dz = (\omega + 1) \left[-\frac{\exp\{-\omega\}}{\omega} + \frac{1}{\omega^2}(1 - \exp\{-\omega\}) \right],$$

where $\omega = \exp\{t\} - 1$ and $\omega \rightarrow \infty$ as $t \rightarrow \infty$. Therefore, in accordance with Eq. (5.11),

$$\lim_{t \rightarrow \infty} \lambda_m(t) = 1.$$

The same limit holds for $\lambda_m(t)$ in (5.37) for the considered specific values of parameters. This example illustrates the fact that the asymptotic value of the mixture failure rate does not depend on a mixing distribution if $\pi(0) \neq 0$.

Theorem 5.4 deals with the case when the support of a mixing distribution includes 0, i.e., $z \in [0, \infty)$. In this case, the strongest population cannot usually be properly defined. If, however, the support is separated from 0, the mixture failure rate can tend to the failure rate of the strongest population as $t \rightarrow \infty$. The following theorem [33] states reasonable conditions for this convergence (we assume, for simplicity, as previously, that $\psi(t) = 0$):

Theorem 5.5 *Let, as in Theorem 5.4, the class c by Eq. (5.27), where $\phi(t) \rightarrow \infty$, $\psi(t) = 0$ and let $A(s)$ be twice differentiable. Assume that, as $s \rightarrow \infty$*

$$\frac{A''(s)}{(A'(s))^2} \rightarrow 0 \quad (5.38)$$

and

$$sA'(s) \rightarrow \infty. \quad (5.39)$$

Also assume that for all $b, c > a$, $b < c$, the quotient $A'(bs)/A'(cs)$ is bounded as $s \rightarrow \infty$. Finally, let the mixing pdf $\pi(z)$ be defined in $[a, \infty)$, $a > 0$, bounded in this interval and continuous at $z = a$ and $\pi(a) \neq 0$. Then

$$\lambda_m(t) \sim a\phi'(t)A'(a\phi(t)). \quad (5.40)$$

The assumptions of this theorem are rather natural and hold at least for the specific models under consideration and for the main lifetime distributions. Assume additionally that the family of failure rates $\lambda(t, z)$ is ordered in z (as for additive or multiplicative models), i.e.,

$$\lambda(t, z_1) < \lambda(t, z_2), \quad z_1 < z_2, \forall z_1, z_2 \in [a, \infty), \quad a > 0. \quad (5.41)$$

The right-hand side of (5.40) can be interpreted in this case as the failure rate of the strongest population. Specifically, for the multiplicative model:

$$\lambda_m(t) \sim a\lambda(t). \quad (5.42)$$

Thus, as intuition suggests, the mixture failure rate asymptotically does not depend on a mixing distribution. A similar result holds also for the case when there is a singularity in the pdf of the mixing distribution of the form:

$$\pi(z) = (z - a)^\alpha \pi_1(z - a), \quad (5.43)$$

where $\alpha > -1$ and $\pi_1(z - a)$ is bounded, $\pi_1(0) \neq 0$.

Missov and Finkelstein [49] have generalized these results to the wider class of mixing distributions. It turned out that the mixing pdf (5.32) in Theorem 5.4 can be of a more general form

$$\pi(z) = z^\alpha G(z)\pi_1(z),$$

where $G(z)$ is a regularly varying function. Recall (Bingham et al. [11]) that a positive function $G(t)$ defined on $(0, \infty)$ is slowly varying at 0 if for every $k > 0$,

$$\lim_{t \rightarrow 0} \frac{G(kt)}{G(t)} = 1.$$

Moreover, a positive function $R(t)$ defined on $(0, \infty)$ is regularly varying at 0 with power $-\infty < p < \infty$, if

$$\lim_{t \rightarrow 0} \frac{R(t)}{t^p G(t)} = 1,$$

where the function $G(t)$ is slowly varying at 0.

5.5 Mortality Plateaus

As it was already mentioned, demographers had recently observed the deceleration in human mortality at advanced ages which eventually results in human mortality plateaus [58]. The most reasonable explanation of this fact is via the concept of heterogeneity of human population which obviously takes place. The following refers to the interpretation of our results for this application.

- As follows from Eq. (5.36), the ALM (5.31) never results in the asymptotically flat failure rate. Moreover, it asymptotically tends to 0 and does not depend on a baseline distribution, which is Gompertz for the case under consideration
- The only function $g(t)$, for which $g(t)/\int_0^t g(u)du$ tends to a constant as $t \rightarrow \infty$, is the exponential function. Therefore, as follows from Relationship (5.35), the asymptotically flat rate in the multiplicative model (5.29) can result via mixing of a random lifetime distributed *only* in accordance with the Gompertz distribution or in accordance with a distribution with the failure rate that asymptotically converges to an exponential function.
- In accordance with Theorem 5.4, the admissible mixing distributions (i.e., the distributions that can lead to the asymptotically flat mortality rate) are those with behavior as z^α , $\alpha > -1$ for $z \rightarrow 0$. The behavior outside the neighborhood of 0 does not contribute to asymptotic properties of the failure rate. Therefore, the power law (Weibull distribution), the gamma distribution, and some other distributions are admissible. Note that, when the mixing pdf is such that $\pi(0) \neq 0$ has a finite limit when $z \rightarrow 0$ (as, e.g., for the exponential distribution), relationship (5.35) reduces to

$$\lambda_m(t) \sim \frac{\lambda(t)}{\int_0^t \lambda(u)du}$$

- And, therefore, the mixture mortality rate does not depend on the mixing distribution at all! The same result holds for, e.g., the mixing density that is $1/a$, $a > 0$ in $[0, a]$ and is 0 in (a, ∞) (uniform distribution).

In view of the foregoing discussion, the asymptotically flat rate (as for human populations) can be viewed as an indication of:

- that the mixing model is multiplicative,
- that the underlying distribution is definitely Gompertz or asymptotically converges to the Gompertz distribution,
- that the mixing pdf is proportional to z^α , $\alpha > -1$, when $z \rightarrow 0$, e.g., the gamma distribution. The form of this distribution outside neighborhood of 0 has no influence on the asymptotic behavior of $\lambda(t)$.

5.6 Inverse Problem

There can be different approaches to considering the inverse problem in mixing. In view of the results of Sect. 5.4, one can be interested in defining the class of mixing distributions that ‘produce’ the mixture failure rate of the form given by (5.34). The following theorem [49] solves this problem.

Theorem 5.6 *Let conditions of Theorem 5.4 hold and, therefore, Relation (5.34) takes place. Then the pdf $\pi(z)$ of the mixing (frailty) distribution satisfies for $z \rightarrow 0$*

$$\frac{\int_0^\infty \exp\{-A(z\phi(t))\} z \pi'(z) dz}{\int_0^\infty \exp\{-A(z\phi(t))\} \pi(z) dz} \sim \alpha. \quad (5.44)$$

Condition (5.44) is not easy to check. However, the following theorem [49] gives a simple sufficient condition.

Theorem 5.7 *Let $\pi(z)$ be a regularly varying function defined by $\pi(z) = z^\alpha G(z)$, where $\alpha > -1$ and $\pi'(z)$ be asymptotically monotone as $z \rightarrow 0$. Then Relationship (5.44) holds.*

A well-known fact from survival analysis states that the failure data alone do not uniquely define a mixing distribution and additional information (e.g., on covariates) should be taken into account (a problem of nonidentifiability, as, e.g., in Tsiatis [59] and Yashin and Manton [66]). On the other hand, the following *specific* inverse problem can be solved analytically, at least for additive and multiplicative models of mixing [28]:

Given the mixture failure rate $\lambda_m(t)$ and the mixing pdf $\pi(z)$, obtain the failure rate $\lambda(t)$ of the baseline distribution.

This means that under certain assumptions any shape of the mixture failure rate can be constructed by the proper choice of the baseline failure rate. To illustrate this statement, consider the additive model (5.28):

$$\bar{F}(t, z) = \exp\{-\Lambda(t) - zt\}, \quad f(t, z) = (\lambda(t) + z) \exp\{-\Lambda(t) - zt\}. \quad (5.45)$$

Therefore, the mixture survival function in (5.10) can be written via the Laplace transform as

$$\bar{F}_m(t) = \exp\{-\Lambda(t)\} \int_0^\infty \exp\{-zt\} \pi(z) dz = \exp\{-\Lambda(t)\} \pi^*(t), \quad (5.46)$$

where, $\pi^*(t) = E[\exp\{-zt\}]$ is the Laplace transform of the mixing pdf $\pi(z)$. Therefore, Eq. (5.15) yields

$$\lambda_m(t) = \lambda(t) + \frac{\int_0^\infty z \exp\{-zt\} \pi(z) dz}{\int_0^\infty \exp\{-zt\} \pi(z) dz} = \lambda(t) - \frac{d}{dt} \log \pi^*(t) \quad (5.47)$$

and the solution of the inverse problem for this special case is given by the following relationship:

$$\lambda(t) = \lambda_m(t) + \frac{d}{dt} \log \pi^*(t) = \lambda_m(t) - E[Z|t]. \quad (5.48)$$

If the Laplace transform of the mixing distribution can be derived explicitly, then Eq. (5.48) gives a simple analytical solution for the inverse problem. Assume, e.g., that ‘we want’ the mixture failure rate to be constant, i.e., $\lambda_m(t) = c$. Then the baseline failure rate is obtained as

$$\lambda(t) = c - E[Z|t].$$

The corresponding survival function for the multiplicative model (5.17) is $\exp\{-z\Lambda(t)\}$ and the mixture survival function for this specific case is

$$\bar{F}_m(t) = \int_0^{\infty} \exp\{-z\Lambda(t)\} \pi(z) dz = \pi^*(\Lambda(t)). \quad (5.49)$$

It is obtained in terms of the Laplace transform of the mixing distribution as a function of the cumulative baseline failure rate $\Lambda(t)$. Therefore,

$$\lambda_m(t) = -\frac{d}{dt} \log \pi^*(\Lambda(t)). \quad (5.50)$$

The general solution to the inverse problem in terms of the Laplace transform is also simple in this case. Note that,

$$\pi^*(\Lambda(t)) = \exp\{-\Lambda_m(t)\}, \quad (5.51)$$

where $\Lambda_m(t)$ denotes the cumulative mixture failure rate. Applying the inverse Laplace transform $L^{-1}(\cdot)$ to both sides of this equation finally results in

$$\lambda(t) = \Lambda'(t) = \frac{d}{dt} L^{-1}(\exp\{-\Lambda_m(t)\}). \quad (5.52)$$

The Laplace transform methodology in multiplicative and additive models is usually very effective. It constitutes a convenient tool for dealing with mixture failure rates when the Laplace transform of the mixing distribution can be obtained explicitly. The exponential family [41] presents a wide class of such distributions. The corresponding pdf is defined in this case as

$$\pi(z) = \frac{\exp\{-\theta z\} g(z)}{\eta(\theta)}, \quad (5.53)$$

where $g(z)$ and $\eta(z)$ are some positive functions and θ is a parameter. The function $\eta(\theta)$ plays the role of a normalizing constant ensuring that the pdf integrates to 1. The gamma, the inverse Gaussian, and the stable distributions are relevant

examples. Note that, the Laplace transform of $\pi(z)$ depends only on the normalizing function $\eta(z)$ [41], i.e.,

$$\pi^*(s) \equiv \int_0^{\infty} \exp\{-sz\} \pi(z) dz = \frac{\eta(\theta + s)}{\eta(\theta)}.$$

This means that under certain assumptions any shape of the mixture failure rate can be constructed by the proper choice of the baseline failure rate. Specifically, for the exponential family of mixing densities and the multiplicative model under consideration, the mixture failure rate is obtained as

$$\begin{aligned} \lambda_m(t) &= -\frac{d}{dt} \log \frac{\eta(\theta + \Lambda(t))}{\eta(\theta)} \\ &= -\lambda(t) \frac{\frac{d}{d(\theta + \Lambda(t))} \eta(\theta + \Lambda(t))}{\eta(\theta + \Lambda(t))}. \end{aligned} \quad (5.54)$$

Therefore, the solution to the inverse problem can be obtained in this case as the derivative of the following function:

$$\Lambda(t) = \eta^{-1}(\exp\{-\lambda_m(t)\} \eta(\theta)) - \theta. \quad (5.55)$$

It can be easily calculated [28] that when the mixing pdf is gamma with parameters α and β , the solution of the inverse problem is obtained as

$$\lambda(t) = \frac{\beta}{\alpha} \lambda_m(t) \exp\left\{\frac{\Lambda_m(t)}{\alpha}\right\}. \quad (5.56)$$

Assume that the mixture failure rate is constant, i.e., $\lambda_m(t) = c$. It follows from (5.56) that for obtaining a constant $\lambda_m(t)$ the baseline $\lambda(t)$ should be exponentially increasing, i.e.,

$$\lambda(t) = \frac{\beta}{\alpha} c \exp\left\{\frac{ct}{\alpha}\right\}.$$

But this is what we would really expect. As we already mentioned, this result is really surprising: we are mixing the exponentially increasing family of failure rates and arriving at a constant mixture failure rate.

5.7 The Failure Rate Dynamics in Heterogeneous Populations

The mixture failure rate function and some other measures based on it (e.g., the reliability function, the mean residual life function, etc.) are conventionally considered as *measures of performance* (or quality) of items in heterogeneous

populations. However, if we pick an operable item at random from this population, its individual failure rate at each instant of time can be considered as a random variable, whereas the mixture failure rate is defined as its expectation. As in the case of ‘ordinary’ random variables, other than expectation characteristics are also important. The obvious first choice is the corresponding variance.

As an example, consider a system that should perform an important mission. The quality of its performance can be described by the probability of operation without failures during a mission time. If a mission is important and its failure results, e.g., in substantial economic loss, then not only the population (mixture) failure rate of a system that defines the average value of this probability, but the deviations from this value due to heterogeneity of a population are of considerable interest. As the weakest items are dying out first, the composition of the ordered heterogeneous population is improving in the sense that proportions of stronger items are increasing. However, does it mean that the ‘quality’ (from a broader perspective) of the entire population is improving? Not necessarily, as this quality can depend also on the variability characteristics to be discussed in this section. Furthermore, when we are dealing with failures that may result in serious consequences, more attention should be paid to the items with a high risk of failure, i.e., the items with large failure rates. Therefore, the measures for quality of these items should be also defined.

We consider a heterogeneous population of items (components) that consists of different homogeneous subpopulations, that are modeled via the frailty Z . The numbers of items in populations are supposed to be sufficiently large and thus our problems can be statistically described in terms of infinite populations. As time progresses, the failed items are discarded and therefore, the composition of the population of survived items (which is, in fact, the conditional frailty $Z|T > t$) changes. Alternatively, an item is chosen at random from our heterogeneous population and if it did not fail in $[0, t)$, then our initial knowledge about its ‘quality’ which is described by the frailty Z is changing in accordance with $Z|T > t$ (see Eq. (5.12) and the discussion after it).

For illustrating the dynamics in variability characteristics, consider the case of $n = 2$ subpopulations that can be generalized to the arbitrary finite n . Denote the lifetime of a component from the strong subpopulation by T_S and its absolutely continuous Cdf, pdf, and the failure rate function by $F_1(t)$, $f_1(t)$ and $\lambda_1(t)$, respectively. Similarly, the lifetime, the Cdf, the pdf, and the failure rate function of a weak component are T_W , $F_2(t)$, $f_2(t)$ and $\lambda_2(t)$, accordingly. Formal definitions of the strong and weak subpopulations will be given after presenting the necessary notation. The initial ($t = 0$) composition of our mixed population is as follows: the proportion of strong items is π , whereas the proportion of weak items is $1 - \pi$, which means that the distribution of the discrete frailty Z with realizations z_1 and z_2 in this case is

$$\pi(z) = \begin{cases} \pi, & z = z_1 \\ 1 - \pi, & z = z_2 \end{cases}$$

and z_1, z_2 ($z_1 < z_2$), correspond to the strong and the weak subpopulations, respectively. In accordance with Eqs. (5.1)–(5.3):

The mixture (population) survival function is

$$\bar{F}_m(t) = \pi \bar{F}_1(t) + (1 - \pi) \bar{F}_2(t).$$

The mixture (observed) failure rate is

$$\lambda_m(t) = \frac{\pi f_1(t) + (1 - \pi) f_2(t)}{\pi \bar{F}_1(t) + (1 - \pi) \bar{F}_2(t)} = \pi(t) \lambda_1(t) + (1 - \pi(t)) \lambda_2(t), \quad (5.57)$$

where the time-dependent probabilities are

$$\pi(t) = \frac{\pi \bar{F}_1(t)}{\pi \bar{F}_1(t) + (1 - \pi) \bar{F}_2(t)}, \quad 1 - \pi(t) = \frac{(1 - \pi) \bar{F}_2(t)}{\pi \bar{F}_1(t) + (1 - \pi) \bar{F}_2(t)}.$$

Thus, the composition of our population is changing in time in accordance with the following distribution of $Z|t \equiv Z|T > t$:

$$\pi(z|t) = \begin{cases} \pi(t), & z = z_1 \\ 1 - \pi(t), & z = z_2 \end{cases}.$$

Assume now that the populations are ordered (and therefore, the weak and the strong subpopulations are defined accordingly) in the sense of the failure rate ordering:

$$\lambda_2(t) \geq \lambda_1(t), \quad t \geq 0.$$

Then, it is easy to see that the proportion of strong items

$$\pi(t) = \frac{\pi}{\pi + (1 - \pi) \bar{F}_2(t) / \bar{F}_1(t)},$$

is increasing as t is increasing. In the context of burn-in, e.g., it means that the quality of a population in the defined sense is improving as the time of burn-in is increasing.

Equation (5.57) defines the observed (mixture) failure rate, which is obviously an averaged characteristic. However, the above mixture setting implies that an operable item at time t can be described by a *random failure rate* $\lambda_R(t)$ with realizations $\lambda_1(t)$ and $\lambda_2(t)$:

$$\lambda_R(t) = \begin{cases} \lambda_1(t), & \text{with probability } \pi(t), \\ \lambda_2(t), & \text{with probability } 1 - \pi(t). \end{cases} \quad (5.58)$$

Thus, we can also interpret (5.57) as the expectation of the random failure rate $\lambda_R(t)$

$$\lambda_m(t) = E[\lambda_R(t)].$$

Expectation is obviously an important characteristic, but, as in the case of ‘ordinary random variables’ we might be interested in moments and, first of all, in $\text{Var}[\lambda_R(t)]$ as the variability measure of the population structure. This measure is important as we want to know (or control) the ‘risks’ (i.e., large deviations from the mean) that can occur in field usage. Therefore, $\lambda_m(t)$ and $\text{Var}[\lambda_R(t)]$ can describe the quality of our heterogeneous population. It is reasonable to assume that the larger these characteristics are, the worse is the corresponding quality. Furthermore, at many instances, along with the absolute variability measure $\text{Var}[\lambda_R(t)]$, the relative variability is of interest. Thus, in addition to $\text{Var}[\lambda_R(t)]$, we will consider the measure for the ‘relative deviation’, i.e., the corresponding coefficient of variation:

$$CV[\lambda_R(t)] = \sqrt{\text{Var}[\lambda_R(t)]}/E[\lambda_R(t)] = \sqrt{\text{Var}[\lambda_R(t)]}/\lambda_m(t).$$

We will derive now general formulas for the measures of interest. In order to obtain $\text{Var}[\lambda_R(t)]$, in accordance with (5.58), it is easier to consider the supplementary random variable $\lambda_{RC}(t)$, which is equal to $\lambda_1(t) - \lambda_2(t)$ with probability $\pi(t)$ and to 0 with probability $1 - \pi(t)$. Then

$$\text{Var}[\lambda_R(t)] = \text{Var}[\lambda_{RC}(t)] = (\lambda_1(t) - \lambda_2(t))^2 \pi(t)(1 - \pi(t)), \quad (5.59)$$

and

$$CV[\lambda_R(t)] = \sqrt{\text{Var}[\lambda_R(t)]}/\lambda_m(t) = \frac{(\lambda_2(t) - \lambda_1(t))\sqrt{\pi(t)(1 - \pi(t))}}{\pi(t)\lambda_1(t) + (1 - \pi(t))\lambda_2(t)}. \quad (5.60)$$

As we know, the shape of the mixture failure rate is very important in describing heterogeneous populations. In accordance with the foregoing considerations, the shape of the functions $\text{Var}[\lambda_R(t)]$ and $CV[\lambda_R(t)]$ is also of interest. For simplicity, we consider first the mixture of two exponential distributions. Let $\lambda_2(t) = \lambda_2 > \lambda_1(t) = \lambda_1$. Then, as a special case of Eq. (5.59),

$$\text{Var}[\lambda_R(t)] = (\lambda_1 - \lambda_2)^2 \pi(t)(1 - \pi(t)),$$

and

$$\lambda'_m(t) = -(\lambda_1 - \lambda_2)^2 \pi(t)(1 - \pi(t)) = -\text{Var}[\lambda_R(t)]. \quad (5.61)$$

Thus, the slope of the mixture failure rate in this case is equal to the variance of the random failure rate (with the negative sign). We can consider the following two cases:

- (i) Let the initial proportion of strong components be larger than 0.5 ($\pi > 0.5$); then $\pi(t)(1 - \pi(t))$ strictly decreases in t from $\pi(0)(1 - \pi(0))$. Therefore, $\lambda_m(t)$ and $\text{Var}[\lambda_R(t)]$ strictly decrease and, therefore, the population becomes ‘better’ (the failure rate is smaller) and more ‘stable’ (the variance is smaller). Observe that

$$\begin{aligned}
 CV'[\lambda_R(t)] &= \frac{1}{2\sqrt{\pi(t)(1-\pi(t))(\lambda_1\pi(t) + \lambda_2(1-\pi(t)))^2}} \\
 &\quad \times [(\lambda_2 - \lambda_1)\pi'(t)\{1 - 2\pi(t)\}(\lambda_1\pi(t) + \lambda_2(1-\pi(t))) + 2(\lambda_2 - \lambda_1)^2\pi'(t)\pi(t)(1-\pi(t))] \\
 &= \frac{1}{2\sqrt{\pi(t)(1-\pi(t))(\lambda_1\pi(t) + \lambda_2(1-\pi(t)))^2}} (\lambda_2 - \lambda_1)\pi'(t)\{\lambda_2(1-\pi(t)) - \lambda_1\pi(t)\}.
 \end{aligned}$$

Therefore, as $\pi'(t)$ is positive ($\pi(t)$ is increasing):

$$CV'[\lambda_R(t)] > 0 \Rightarrow \frac{\lambda_2}{\lambda_1} > \frac{\pi(t)}{1-\pi(t)}.$$

Obviously, $\pi(t)/(1-\pi(t))$ strictly increases to ∞ as t increases. Thus, when

$$\frac{\lambda_2}{\lambda_1} > \frac{\pi(0)}{1-\pi(0)}, \quad (5.62)$$

$CV[\lambda_R(t)]$ increases and then decreases with one change point t^* such that $\lambda_2/\lambda_1 = \pi(t^*)/(1-\pi(t^*))$. When

$$\frac{\lambda_2}{\lambda_1} < \frac{\pi(0)}{1-\pi(0)},$$

then $CV[\lambda_R(t)]$ monotonically decreases.

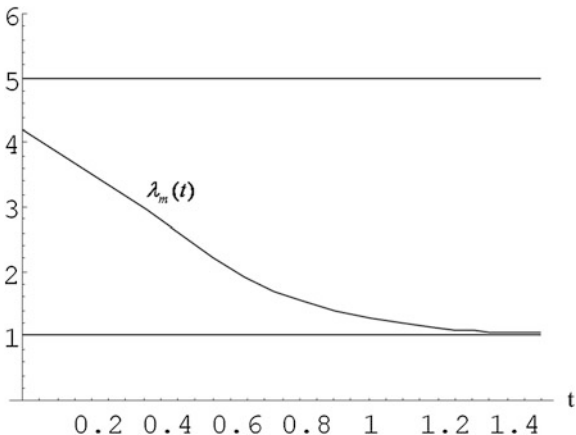
- (ii) Let the initial proportion of strong components be smaller or equal to 0.5 ($\pi \leq 0.5$). As it was stated, the proportion of remaining weak components $1 - \pi(t)$ is always decreasing in time. Therefore, the first guess based on intuition would be that $\text{Var}[\lambda_R(t)]$ (similar to (i)) is also decreasing. However, it is easy to see that at time t such that $\pi(t) = 0.5$, the function, $\text{Var}[\lambda_R(t)]$ (and as follows from (5.61), $|\lambda'_m(t)|$ as well) has its maximum and only after this point it strictly decreases. In this case, Inequality (5.62) always holds and thus $CV[\lambda_R(t)]$ increases and then decreases with one change point t^* such that $\lambda_2/\lambda_1 = \pi(t^*)/(1-\pi(t^*))$.

Equation (5.59) can be used for analyzing the shape of $\text{Var}[\lambda_R(t)]$ for time-dependent failure rates. Specifically, when $\lambda_2(t) - \lambda_1(t)$ is increasing and $\pi \leq 0.5$, then $\pi(t)(1-\pi(t))$ first strictly increases and then decreases. Therefore, $\text{Var}[\lambda_R(t)]$ initially strictly increases.

When $\lambda_1(t) - \lambda_2(t)$ is decreasing:

- (i) If $\pi > 0.5$, then $\pi(t)(1-\pi(t))$ strictly decreases and $\text{Var}[\lambda_R(t)]$ strictly decreases.
- (ii) If $\pi \leq 0.5$, then, $\pi(t)(1-\pi(t))$ strictly increases in $[0, t^*)$ and decreases in $[t^*, \infty)$, where t^* is the solution of the following equation: $\pi(t) = 0.5$. Thus $\text{Var}[\lambda_R(t)]$ strictly decreases in $[t^*, \infty)$.

Fig. 5.3 Mixture Failure Rate $\lambda_m(t)$



Equation (5.60) can be used for analyzing the shape of $CV[\lambda_R(t)]$. For instance, if $\lambda_2(t) - \lambda_1(t)$ is decreasing and $\lambda_m(t)$ is increasing, then $CV[\lambda_R(t)]$ is strictly decreasing or it initially increases and then monotonically decreases.

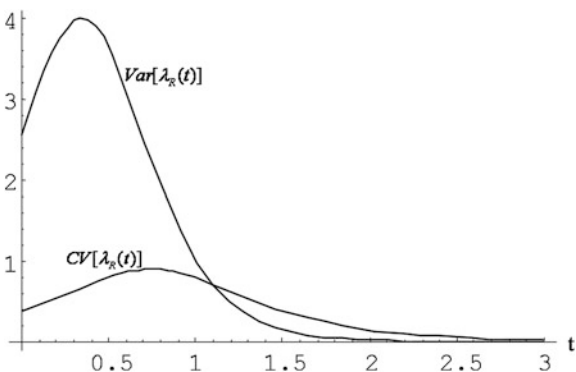
Example 5.1 Let $\lambda_1(t) = 1$, $\lambda_2(t) = 5$ and $\pi = 0.2$. Then the mixture failure rate $\lambda_m(t)$ is given by Fig. 5.3.

Assume that an item has survived to age 0.4. As follows from the graph: $\lambda_m(0.4) \approx 3.0$. How much can we rely on this value? To answer this question, it is reasonable to consider $Var[\lambda_R(t)]$ given by Fig. 5.4.

We can see that $Var[\lambda_R(t)]$ has a maximum at $t \approx 0.4$ ($\pi(0.4) \approx 0.5$). This means that at $t = 0.4$, approximately 50 % of survived items have the failure rate with realization 5.0, and the other 50 % will have it 1.0, whereas the observed (mixture) failure rate $\lambda_m(t)$ is 3.0. However, as t increases from 0.4, we may more and more ‘rely’ on $\lambda_m(t)$ as variability decreases.

The above example is rather interesting: We may think that the population would become more and more ‘stable’ (monotonically) as $\lambda_m(t)$ (monotonically)

Fig. 5.4 $Var[\lambda_R(t)]$ and $CV[\lambda_R(t)]$



approaches the failure rate of the strongest subpopulation. However, it is not true, as the variance is not monotonic. The similar conclusion follows when considering $CV[\lambda_R(t)]$ (Fig. 5.4).

Similar considerations can be applied to continuous mixtures defined by Eqs. (5.10)–(5.12). Let our subpopulations be ordered in the sense of the failure rate ordering:

$$\lambda(t, z_1) \leq \lambda(t, z_2), \quad z_1 < z_2, \forall z_1, z_2 \in [0, \infty), t \geq 0. \tag{5.63}$$

Denote the Cdfs of $\pi(z)$ and $\pi(z|t)$ by $\Pi(z)$ and $\Pi(z|t)$, respectively, and by $Z|t$ the conditional frailty (on condition that the item did not fail in $[0, t)$). The following simple result describes the important property of the family $\{Z|t\}_{t \geq 0}$.

Theorem 5.8 *Let our subpopulations be ordered in the sense of the failure rate ordering (5.64). Then the family of random variables $Z|t \equiv Z|T > t$ is DLR (decreasing in the sense of the likelihood ratio) in $t \in [0, \infty)$.*

Proof Recall that a random variable X (with the pdf $f(t)$) is smaller than a random variable Y (with the pdf $g(t)$) in the sense of the likelihood ratio ordering (LRO) if $f(t)/g(t)$ is decreasing in t (see also (2.71)). Therefore, the DLR property of the family $\{Z|t\}_{t \geq 0}$ means that for all $t_2 > t_1$, $Z|t_2$ is smaller than $Z|t_1$ in the sense of the LRO.

In accordance with the definition of the conditional mixing distribution (5.12) in the mixing model (5.11), the ratio of the corresponding densities for different instants of time is

$$L(z, t_1, t_2) = \frac{\pi(z|t_2)}{\pi(z|t_1)} = \frac{\bar{F}(t_2, z) \int_0^\infty \bar{F}(t_1, z) \pi(z) dz}{\bar{F}(t_1, z) \int_0^\infty \bar{F}(t_2, z) \pi(z) dz}.$$

Therefore, monotonicity in z of $L(z, t_1, t_2)$ is defined by the function

$$\frac{\bar{F}(t_2, z)}{\bar{F}(t_1, z)} = \exp \left\{ - \int_{t_1}^{t_2} \lambda(u, z) du \right\},$$

which, owing to ordering (5.63), is decreasing in z for all $t_2 > t_1$. □

As the LRO ordering is stronger than the usual stochastic ordering, it means that $\Pi(z|t)$ is increasing in t for each $z > 0$. Therefore, in accordance with (5.63), the proportion of ‘better’ (with smaller failure rates) items is increasing.

For tractability, consider now the important specific case of the multiplicative model: $\lambda(t, z) = z\lambda(t)$. Therefore,

$$\lambda_R(t) = Z_t \lambda(t),$$

where $Z_t = Z|t$ and

$$\lambda_m(t) = E[\lambda_R(t)] = \lambda(t) \int_0^{\infty} z\pi(z|t)dz = \lambda(t)E[Z|t].$$

Observe that

$$\text{Var}[\lambda_R(t)] = (\lambda(t))^2 \text{Var}[Z_t] = (\lambda(t))^2 \text{Var}[Z|t],$$

and thus,

$$CV[\lambda_R(t)] = \frac{\sqrt{\text{Var}[Z|t]}}{E[Z|t]} = CV[Z|t].$$

Furthermore, as $E'[Z_t] = E'[Z|t] = -\lambda(t)\text{Var}[Z|t] < 0$,

$$\lambda'_m(t) = \lambda'(t)E[Z|t] - (\lambda(t))^2 \text{Var}[Z|t].$$

Specifically, when the population is a mixture of exponential distributions, we have

$$\lambda'_m(t) = -(\lambda(t))^2 \text{Var}[Z|t].$$

Example 5.2 Consider continuous mixture of exponentials. Let the conditional failure rate and the mixing distribution be $\lambda(t, z) = z$ and $\pi(z) = \theta \exp\{-\theta z\}$, respectively. Then

$$\lambda_m(t) = E[\lambda_R(t)] = E[Z|t] = 1/(\theta + t),$$

and

$$\text{Var}[\lambda_R(t)] = \text{Var}[Z|t] = 1/(\theta + t)^2.$$

Thus

$$CV[\lambda_R(t)] = 1.$$

Obviously, the quality of the population is defined only by $E[Z|t]$, which is decreasing in t . Therefore, the failure rates are ‘improving’ and the variance as well. However, the CV is constant, and this characteristic often more adequately describes variability especially when both the failure rate and its variance are decreasing in time.

5.8 Stochastic Intensity for Minimal Repairs in Heterogeneous Populations

In Sect. 2.5, we have defined and described the crucial for the reliability of repairable systems notion of minimal repair. This was done for items from homogeneous populations. It is really a challenge to define and study minimal repair in heterogeneous populations.

Consider a system with an absolutely continuous time to failure Cdf $F(t)$ and the failure rate $\lambda(t)$, which starts operating at $t = 0$. Assume that the repair action is performed instantaneously upon failure. Recall that the repair is usually qualified as *perfect* if the Cdf of the repaired object is $F(t)$ (as good as new) and as *minimal* at time x , if its Cdf is:

$$F(t|x) \equiv 1 - \frac{1 - F(t+x)}{1 - F(x)} \quad (5.64)$$

(as bad as old), which is equivalent to Eq. (2.26). Thus the minimal repair restores our system (in terms of the corresponding distribution) to the state it had prior to the failure.

Sometimes, upon failure, we can observe additional information about the state of an object (e.g., the structure of a system). This can allow us to define a more general type of repair, which is usually called the *information-based (or physical) minimal* repair. The information-based minimal repair brings our object back to the state (to be defined by the relevant information) it had just prior to the failure [4, 5, 10, 19, 26, 27, 50].

It is really challenging to generalize the notion of minimal repair to items from *heterogeneous populations*. The corresponding attempt was performed in Finkelstein [27] and further elaborated in Cha and Finkelstein [20]. Our presentation in this section will mostly follow the latter paper.

Let failures of repairable items be repaired instantaneously. Then the process of repairs can be described by a stochastic point process. A convenient way of mathematical description of these processes is using the concept of the stochastic intensity (the intensity process) $\lambda_t, t \geq 0$ defined by Relationship (2.12). A classical example of λ_t is the intensity process generated by the renewal process (perfect, instantaneous repairs):

$$\lambda_t = \sum_{n=0}^{\infty} \lambda(t - T_n) I(T_n \leq t < T_{n+1}), \quad T_0 = 0,$$

where $T_1 < T_2 < T_3 < \dots$, are the random failure times. Another standard example is the ‘deterministic stochastic intensity’ $\lambda_t = \lambda(t)$ which defines the nonhomogeneous Poisson process (NHPP) of repairs with rate (intensity) $\lambda(t)$. It is well known that this example can also be interpreted as the process of *minimal* repairs.

As in the previous sections, we formally describe heterogeneous populations in the following way. Let $T \geq 0$ be a lifetime r.v. with the Cdf $F(t)$ ($\bar{F}(t) \equiv 1 - F(t)$). Assume that $F(t)$ is indexed by a r.v. Z , i.e.,

$$P(T \leq t|Z = z) \equiv P(T \leq t|z) \equiv F(t, z)$$

and that the pdf $f(t, z)$ exists. Then the corresponding failure rate $\lambda(t, z)$ is $f(t, z)/\bar{F}(t, z)$. Let Z be a frailty with support in $[a, b]$, $0 \leq a < b \leq \infty$, and the pdf $\pi(z)$. The above setting leads naturally to considering mixtures of distributions, which are useful for describing heterogeneity [see Eqs. (5.10–5.12)].

We can now define two types (scenarios) of minimal repair for heterogeneous populations, but in a more general context than in Finkelstein [27]. The first type of minimal repair does not employ any additional information and, therefore, the failed item is replaced by the statistically identical item. As the failure time distribution in this case is just the mixture (5.10), the stochastic intensity for the corresponding *process of minimal repairs* of this type is obviously equal to the mixture failure rate, i.e.,

$$\lambda_t = \lambda_m(t), \quad t \geq 0.$$

The second type of minimal repair (already information-based) restores an item to a statistically identical item with the same value of frailty Z . It can be realized in practice by performing the second ‘operation’ resulting in the ‘classical’ minimal repair when during the repair only a small part of a large system is replaced. It is natural to suggest that the state of an item is also defined by the corresponding realization of the frailty parameter (i.e., if $Z = z$ before the failure, it should be z after the failure). Thus (5.64) is modified to:

$$F(t, z|x) \equiv 1 - \frac{1 - F(t + x, z)}{1 - F(x, z)}.$$

Our main attention here focuses on this type of minimal repair, as it is the most ‘interesting’ from both a practical and a theoretical points of view.

Let us come back to the definition of the intensity process (2.12) and modify it with respect to the ‘heterogeneous’ case when the orderly point process is indexed by the frailty parameter Z . Observe that the stochastic intensity λ_t (unconditional with respect to frailty Z) can be specified now as:

$$\begin{aligned} \lambda_t &= \lim_{\Delta t \rightarrow 0} \frac{E[\Pr[N(t, t + \Delta t) = 1|H_t, Z]]}{\Delta t} \\ &= E \left[\lim_{\Delta t \rightarrow 0} \frac{\Pr[N(t, t + \Delta t) = 1|H_{t-}, Z]}{\Delta t} \right] \\ &= E[\lambda_{t,Z}], \end{aligned} \tag{5.65}$$

where the expectation is with respect to the conditional distribution $Z|H_t$ and

$$\lambda_{t,Z} \equiv \lim_{\Delta t \rightarrow 0} \frac{\Pr[N(t, t + \Delta t) = 1 | H_t, Z]}{\Delta t}. \quad (5.66)$$

Then $\lambda_{t,z}(Z = z)$ in (5.66) can be interpreted as the conditional (with respect to Z) stochastic intensity of the orderly point process, indexed by frailty Z .

We will specify now our point process. As before, let Z be the frailty of an item randomly selected at time $t = 0$ from our heterogeneous population. Upon each failure we perform the minimal repair of the second type. Note that, in this case, if $Z = z$ at time $t = 0$, then the corresponding realization is $\lambda_{t,z} = \lambda(t, z)$ for all $t \geq 0$. Therefore, for the second type of minimal repair, $\lambda_{t,Z}$ in (5.66) is now given by

$$\lambda_{t,Z} = \lambda(t, Z), \quad t \geq 0,$$

and, in accordance with (5.65), the corresponding stochastic intensity λ_t is the expectation of $\lambda(t, Z)$ with respect to the distribution of $Z | H_t$. This operation means that, although the value of Z is chosen at $t = 0$ and is fixed, its distribution is updated with time as information about failures and survival times emerges (see the detailed procedure in what follows).

We see that stochastic modeling for the second type of minimal repair is dramatically different from that for the first type, as information about the operational history (failure times and survival times) updates the conditional frailty distribution $Z | H_t$.

In accordance with our considerations, it is clear that the stochastic intensity $\lambda_t = E[\lambda_{t,Z}]$ defined in (5.65) for $t \in [0, t_1)$, where t_1 is the realization of the failure time T_1 , is just the mixture failure rate (5.12), i.e., $\lambda_m^1(t) = \lambda_m(t)$, as the information at hand is just the initial distribution $\pi(z)$ (and the fact that the item has survived in $[0, t)$).

Consider now the next interval $[t_1, t_2)$. Given the additional information (in addition to the initial distribution $\pi(z)$) that an item has failed at $t = t_1$, the pdf of frailty $Z = z$ (we repair an item to the state, defined by the same value of frailty) is

$$\pi_{02}(z) \equiv \frac{\lambda(t_1, z) \exp\left\{-\int_0^{t_1} \lambda(s, z) ds\right\} \cdot \pi(z)}{\int_a^b \lambda(t_1, z) \exp\left\{-\int_0^{t_1} \lambda(s, z) ds\right\} \cdot \pi(z) dz}. \quad (5.67)$$

Thus the ‘initial frailty distribution’ (at the start of the second cycle) just after the minimal repair is given by (5.67). Furthermore, the ‘remaining survival function’ at time $t = t_1$ is given by $[\bar{F}(t_1 + u, z) / \bar{F}(t_1, z)]$. Then, the conditional frailty distribution $Z | H_t$ in $[t_1, t_2)$ is

$$\frac{[\bar{F}(t, z) / \bar{F}(t_1, z)] \cdot \pi_{02}(z)}{\int_a^b [\bar{F}(t, z) / \bar{F}(t_1, z)] \cdot \pi_{02}(z) dz} = \frac{\lambda(t_1, z) \exp\left\{-\int_0^t \lambda(s, z) ds\right\} \cdot \pi(z)}{\int_a^b \lambda(t_1, z) \exp\left\{-\int_0^t \lambda(s, z) ds\right\} \cdot \pi(z) dz},$$

and the corresponding stochastic intensity is, in accordance with (5.65),

$$\lambda_m^2(t) = \int_a^b \lambda(t, z) \cdot \frac{\lambda(t_1, z) \exp\{-\int_0^{t_1} \lambda(s, z) ds\} \cdot \pi(z)}{\int_a^b \lambda(t_1, z) \exp\{-\int_0^{t_1} \lambda(s, z) ds\} \cdot \pi(z) dz} dz, \text{ in } [t_1, t_2]. \quad (5.68)$$

Using another useful (Bayesian) interpretation, we can say that the item fails at time t_1 and, after repair, survives in $[t_1, t]$. Thus, the corresponding probability (conditional probability given $Z = z$ at $t = 0$) is

$$\begin{aligned} & \lambda(t_1, z) \exp\left\{-\int_0^{t_1} \lambda(s, z) ds\right\} \cdot \exp\left\{-\int_{t_1}^t \lambda(s, z) ds\right\} dt_1 \\ &= \lambda(t_1, z) \exp\left\{-\int_0^t \lambda(s, z) ds\right\} dt_1. \end{aligned}$$

Given this information, the conditional frailty distribution $Z|H_t$ should be updated as

$$\frac{\lambda(t_1, z) \exp\{-\int_0^{t_1} \lambda(s, z) ds\} \cdot \pi(z)}{\int_a^b \lambda(t_1, z) \exp\{-\int_0^{t_1} \lambda(s, z) ds\} \cdot \pi(z) dz},$$

which yields (5.68).

Consider now the intensity process in $[t_2, t_3)$. As we know that the item has failed at times t_1 and t_2 and after minimal repairs has survived to t_1-t_2 , the corresponding probability (conditional probability given $Z = z$ at $t = 0$, divided by $dt_1 dt_2$) is

$$\begin{aligned} & \lambda(t_1, z) \exp\left\{-\int_0^{t_1} \lambda(s, z) ds\right\} \cdot \lambda(t_2, z) \exp\left\{-\int_{t_1}^{t_2} \lambda(s, z) ds\right\} \cdot \exp\left\{-\int_{t_2}^t \lambda(s, z) ds\right\} \\ &= \lambda(t_1, z) \lambda(t_2, z) \exp\left\{-\int_0^t \lambda(s, z) ds\right\}. \end{aligned}$$

Given this information, the conditional frailty distribution $Z|H_t$ should be updated as

$$\frac{\lambda(t_1, z) \lambda(t_2, z) \exp\{-\int_0^t \lambda(s, z) ds\} \cdot \pi(z)}{\int_a^b \lambda(t_1, z) \lambda(t_2, z) \exp\{-\int_0^t \lambda(s, z) ds\} \cdot \pi(z) dz}.$$

Thus, in $[t_2, t_3)$, as before,

$$\lambda_m^3(t) = \int_a^b \lambda(t, z) \cdot \frac{\lambda(t_1, z) \lambda(t_2, z) \exp\{-\int_0^t \lambda(s, z) ds\} \cdot \pi(z)}{\int_a^b \lambda(t_1, z) \lambda(t_2, z) \exp\{-\int_0^t \lambda(s, z) ds\} \cdot \pi(z) dz} dz, \text{ in } [t_2, t_3).$$

More generally, for $t \in [t_{n-1}, t_n)$, the conditional frailty distribution $Z|H_t$ is defined by

$$\pi^n(z|t_1, \dots, t_{n-1}) \equiv \frac{\lambda(t_1, z) \cdots \lambda(t_{n-1}, z) \exp\{-\int_0^t \lambda(s, z) ds\} \cdot \pi(z)}{\int_a^b \lambda(t_1, z) \cdots \lambda(t_{n-1}, z) \exp\{-\int_0^t \lambda(s, z) ds\} \cdot \pi(z) dz} \quad (5.69)$$

and, therefore,

$$\lambda_m^n(t) = \int_a^b \lambda(t, z) \pi^n(z|t_1, \dots, t_{n-1}) dz \text{ in } [t_{n-1}, t_n). \quad (5.70)$$

Based on (5.69) and (5.70), the corresponding stochastic intensity can now be defined as

$$\lambda_t = \sum_{n=1}^{\infty} \lambda_m^n(t) I(T_{n-1} \leq t < T_n), \quad T_0 \equiv 0. \quad (5.71)$$

The following result presents a useful ordering of stochastic intensities for minimal repairs of the first and the second types (Cha and Finkelstein [20]).

Theorem 5.9 *Let the values of $\lambda(t, z)$ be ordered with respect to z : for all $z_1, z_2 \in [a, b]$, $t \geq 0$*

$$\lambda(t, z_1) < \lambda(t, z_2), \text{ if } z_1 < z_2.$$

Then

$$\lambda_m(t) \leq \lambda_t, \quad t \geq 0,$$

where λ_t is the stochastic intensity for the second type of minimal repair in (5.71).

Proof Note that if $X \leq_{st} Y$ and $g(\cdot)$ is any increasing function, then $g(X) \leq_{st} g(Y)$ and, accordingly, $E[g(X)] \leq E[g(Y)]$. Observe that both $\lambda_m(t)$ and λ_t are expectations of $\lambda(t, Z)$ with respect to the mixing distributions

$$\pi(z|t) = \pi(z) \frac{\bar{F}(t, z)}{\int_a^b \bar{F}(t, z) \pi(z) dz}$$

and

$$\pi^n(z|t_1, \dots, t_{n-1}) \equiv \frac{\lambda(t_1, z) \cdots \lambda(t_{n-1}, z) \exp\{-\int_0^t \lambda(s, z) ds\} \cdot \pi(z)}{\int_a^b \lambda(t_1, z) \cdots \lambda(t_{n-1}, z) \exp\{-\int_0^t \lambda(s, z) ds\} \cdot \pi(z) dz},$$

respectively. Then it is sufficient to show that

$$\Pi(v|t) \geq \Pi^n(v|t_1, \dots, t_{n-1}), \quad (5.72)$$

for all $n \geq 1$, $0 < t_1 < \dots < t_{n-1} < t$, where $\Pi(z|t)$ and $\Pi^n(v|t_1, \dots, t_{n-1})$ are the corresponding Cdfs. Observe that

$$\begin{aligned} \pi^n(z|t_1, \dots, t_{n-1}) &\equiv \frac{\lambda(t_1, z) \cdots \lambda(t_{n-1}, z) \bar{F}(t, z) \pi(z)}{\int_a^b \lambda(t_1, z) \cdots \lambda(t_{n-1}, z) \bar{F}(t, z) \pi(z) dz} \\ &= \frac{\lambda(t_1, z) \cdots \lambda(t_{n-1}, z) \cdot \pi(z|t)}{\int_a^b \lambda(t_1, z) \cdots \lambda(t_{n-1}, z) \pi(z|t) dz}. \end{aligned}$$

It is clear that there exist $a \leq z^*(a, v) \leq v$ and $v \leq z^*(v, b) \leq b$ such that

$$\int_a^v \lambda(t_1, z) \cdots \lambda(t_{n-1}, z) \cdot \pi(z|t) dz = \lambda(t_1, z^*(a, v)) \cdots \lambda(t_{n-1}, z^*(a, v)) \int_a^v \pi(z|t) dz$$

and

$$\int_v^b \lambda(t_1, z) \cdots \lambda(t_{n-1}, z) \cdot \pi(z|t) dz = \lambda(t_1, z^*(v, b)) \cdots \lambda(t_{n-1}, z^*(v, b)) \int_v^b \pi(z|t) dz.$$

Thus,

$$\begin{aligned} \Pi^n(v|t_1, \dots, t_{n-1}) &= \frac{\lambda(t_1, z^*(a, v)) \cdots \lambda(t_{n-1}, z^*(a, v)) \cdot \int_a^v \pi(z|t) dz}{\lambda(t_1, z^*(a, v)) \cdots \lambda(t_{n-1}, z^*(a, v)) \cdot \int_a^v \pi(z|t) dz + \lambda(t_1, z^*(v, b)) \cdots \lambda(t_{n-1}, z^*(v, b)) \cdot \int_v^b \pi(z|t) dz} \\ &\leq \int_a^v \pi(z|t) dz = \Pi(v|t). \end{aligned}$$

Since $\lambda(t_1, z) \cdots \lambda(t_{n-1}, z)$ is an increasing function of z ,

$$\lambda(t_1, z^*(a, v)) \cdots \lambda(t_{n-1}, z^*(a, v)) \leq \lambda(t_1, z^*(v, b)) \cdots \lambda(t_{n-1}, z^*(v, b)),$$

and, therefore, Inequality (5.72) is justified. \square

Example 5.3 Suppose that $F(t, z)$ is an exponential distribution with parameter $\lambda(t, z) = z\lambda$ and let $\pi(z)$ be an exponential pdf in $[0, \infty)$ with parameter θ . Then direct integration in (5.11) gives: $\lambda_m(t) = \lambda/(\lambda t + \theta)$. Observe that

$$\begin{aligned} \pi^n(z|t_1, \dots, t_{n-1}) &\equiv \frac{\lambda(t_1, z) \cdots \lambda(t_{n-1}, z) \exp\{-\int_0^t \lambda(s, z) ds\} \cdot \pi(z)}{\int_a^b \lambda(t_1, z) \cdots \lambda(t_{n-1}, z) \exp\{-\int_0^t \lambda(s, z) ds\} \cdot \pi(z) dz} \\ &= \frac{(z\lambda)^{n-1} \exp\{-z\lambda t\} \cdot \theta \exp\{-\theta z\}}{\int_0^\infty (z\lambda)^{n-1} \exp\{-z\lambda t\} \cdot \theta \exp\{-\theta z\} dz}, \end{aligned}$$

and, from (5.69) and (5.70),

$$\lambda_m^n(t) = \frac{\int_0^\infty (z\lambda)^n \exp\{-(\lambda t + \theta)z\} dz}{\int_0^\infty (z\lambda)^{n-1} \exp\{-(\lambda t + \theta)z\} dz} = n \frac{\lambda}{\lambda t + \theta}.$$

Finally,

$$\lambda_1 = \sum_{n=1}^{\infty} n \frac{\lambda}{\lambda t + \theta} I(T_{n-1} \leq t < T_n), \quad T_0 \equiv 0.$$

Thus, $\lambda_m(t) \leq \lambda_t$, $t \geq 0$, holds.

Denote by $H_m(t)$ and $H_\lambda(t)$ the mean numbers of repairs (failures) in $[0, t)$ that correspond to the minimal repair processes of type 1 and type 2, respectively. The following result obviously follows from Theorem 5.9: $H_m(t) \leq H_\lambda(t)$.

5.9 Preventive Maintenance in Heterogeneous Populations

The previous section dealt with the minimal repair as a specific type of corrective maintenance (CM). Now we will consider the preventive maintenance in heterogeneous populations. Our presentation mostly follows Cha and Finkelstein [21], whereas the developed approach is related to that of Sect. 5.8.

Preventive maintenance (PM) for non-repairable systems is a schedule of planned maintenance actions aimed at the prevention of breakdowns and failures of deteriorating systems. By “non-repairable” in this context we mean that the failure of a system is considered as an ‘end event’ and, therefore, the CM is not performed. We shall use this term in the defined sense throughout this section. Detailed surveys on the PM models for deteriorating systems can be found in, e.g., Valdez-Flores and Feldman [60] and Wang [65]. However, almost all models, procedures, and approaches described in the literature and those applied in reliability practice deal *only* with the case when the items come from homogeneous populations. Therefore, as in the case of the minimal repair in the previous section, it is quite a challenge to generalize PM to the case of heterogeneous populations of items.

As previously, we deal with the population described by the continuous mixtures setting (5.10)–(5.12). If the items are not maintained during operation, then their susceptibility to failures can be described by the ‘ordinary’ failure rate (2.4) (homogeneous case) or (5.12) (heterogeneous case). However, when maintenance actions that can affect reliability of items are performed, the corresponding effects should be taken into account. In the following, we will assume that the times of maintenance are negligible.

Consider first, reliability of a non-repairable item from a homogeneous population under PM (without CM). As PM affects its lifetime, we need to define new reliability measures in this case. Let T_p be the time to failure of item ‘under preventive maintenance’ and H_t be the maintenance history in $[0, t)$, i.e., the times of maintenance actions and the stochastic effects of the corresponding maintenances. Then, in order to describe the susceptibility to failure at time t , it is natural to define the following conditional failure rate:

$$\lambda_c(t) \equiv \lim_{\Delta t \rightarrow 0} \frac{\Pr[t < T_P \leq t + \Delta t | H_t, T_P > t]}{\Delta t}, \quad t \geq 0 \quad (5.73)$$

Note that when maintenance is deterministic (times and effect), $\lambda_c(t)$ is also deterministic. However, if, e.g., times of maintenances are random, then $\lambda_c(t)$ is the stochastic process. The following example for the ‘homogeneous items’ is crucial for our further discussion:

Example 5.4 A non-repairable item with a lifetime described by the *increasing* failure rate $\lambda(t)$ starts its operation at $t = 0$. If it operable, it is preventively maintained at times kt_{PM} , $k = 1, 2, \dots$. Assume that each preventive maintenance does not change the ‘shape’ of the function $\lambda(t)$, but the age of the item is reduced in accordance with the factor $0 < \alpha < 1$ (the reduced age is called the ‘virtual age’). Therefore, PM has the effect of decreasing the failure rate as compared to an item that is not preventively maintained [28, 42]). Under these assumptions, the ‘virtual age’ of the item just after the first PM is αt_{PM} , just after the second PM is $\alpha(\alpha t_{PM} + t_{PM}) = \alpha^2 t_{PM} + \alpha t_{PM}, \dots$, and the virtual age just after the $(n - 1)$ th PM, is

$$\begin{aligned} t_{n-1} &= \alpha t_{PM} + \alpha^2 t_{PM} + \dots + \alpha^{n-1} t_{PM} \\ &= [\alpha(1 - \alpha^{n-1}) / (1 - \alpha)] t_{PM}, \quad n = 2, 3, \dots \end{aligned} \quad (5.74)$$

Suppose that the item under this PM schedule has not failed until time t , $t \in [(n - 1)t_{PM}, nt_{PM})$ meaning that it has been preventively maintained for $(n - 1)$ times at kt_{PM} , $k = 1, 2, \dots, (n - 1)$, whereas the last PM was performed at $(n - 1)t_{PM}$. Thus, the virtual age of this item at time t is given by $t_{n-1} + (t - (n - 1)t_{PM})$. Due to the PM assumptions, the statistical state of the maintained item at time t is the same as that of an identical (without maintenance) item with age $t_{n-1} + (t - (n - 1)t_{PM})$. Accordingly, the conditional failure rate (5.73) that takes into account the described specific history H_t is given by

$$\lambda_c(t) = \lambda(t_{n-1} + (t - (n - 1)t_{PM})), \quad t \in [(n - 1)t_{PM}, nt_{PM}).$$

or, equivalently, letting $t_0 \equiv 0$:

$$\lambda_c(t) = \sum_{n=1}^{[t/t_{PM}]+1} \lambda(t_{n-1} + (t - (n - 1)t_{PM})) I((n - 1)t_{PM} \leq t < nt_{PM}). \quad (5.75)$$

where $I(\cdot)$ is the corresponding indicator and $[t/t_{PM}]$ denotes the integer part of t/t_{PM} . Therefore, if the original failure rate $\lambda(t)$ is increasing, then $\lambda_c(t) \leq \lambda(t)$, for all t and accordingly, PMs increase reliability of our item, i.e.,

$$\exp \left\{ - \int_0^t \lambda_c(u) du \right\} \leq \exp \left\{ - \int_0^t \lambda(u) du \right\}.$$

We will now study the PM considered in Example 5.4, but for items from a heterogeneous population described by (5.10)–(5.12). Suppose that an item is randomly selected from this population and is preventively maintained at times kt_{PM} , $k = 1, 2, \dots$. Preventive maintenance does not change the shape of the failure rate of an item but reduces its age in the same way as described by (5.74). Then, following the similar reasoning as in Example 5.4, one may construct the conditional failure rate by simply replacing $\lambda(t)$ in (5.75) with $\lambda_m(t)$:

$$\lambda_c(t) = \sum_{n=1}^{\lceil t/t_{\text{PM}} \rceil + 1} \lambda_m(t_{n-1} + (t - (n-1)t_{\text{PM}}))I((n-1)t_{\text{PM}} \leq t < nt_{\text{PM}}). \quad (5.76)$$

However, distinct from the homogeneous case, it is now not clear at all how this age reducing operation can be performed. In what follows, we will investigate the appropriateness of $\lambda_c(t)$ in (5.76) in defining the actual susceptibility of the survived item to failure at time t . For this purpose, we will suggest the operational profile that results in (5.76) and explain why it is unrealistic in practice. Then, we will suggest alternative profile with a different form of the conditional failure rate, which can be already justified in practice. Finally, the corresponding comparison of two profiles will be performed.

Operation profile 1 *An item is chosen at random from our population and starts operation at $t = 0$. Furthermore, a statistically identical “NEW” population is ‘switched on’ at time $t_{\text{PM}} - \alpha t_{\text{PM}}$ (the delayed start). At time $t = t_{\text{PM}}$, if the selected item has not failed yet, it is replaced by an item randomly selected from the “delayed” population with age αt_{PM} . Then the replaced one starts its operation. At time $t = 2t_{\text{PM}}$, if the replaced item has not failed yet, it is replaced by an item randomly selected from another ‘delayed’ population that started its operation at $2t_{\text{PM}} - (\alpha^2 t_{\text{PM}} + \alpha t_{\text{PM}})$ and, therefore, its age is now $\alpha^2 t_{\text{PM}} + \alpha t_{\text{PM}}$. Then the replaced item starts its operation, and so on.*

We will construct the corresponding conditional failure rate for the described Operation profile 1 and will show that it is eventually given by Eq. (5.76). First, it is necessary to have in mind that the conditional failure rate defined in (5.73) can be expressed for the heterogeneous case as

$$\begin{aligned} \lambda_c(t) &= \lim_{\Delta t \rightarrow 0} \frac{E[\Pr[t < T_P \leq t + \Delta t | H_t, T_P > t, Z]]}{\Delta t} \\ &= E \left[\lim_{\Delta t \rightarrow 0} \frac{\Pr[t < T_P \leq t + \Delta t | H_t, T_P > t, Z]}{\Delta t} \right] \\ &= E[\lambda_{t,Z}], \end{aligned} \quad (5.77)$$

where the expectation is with respect to the conditional distribution $Z|(H_t, T_P > t)$ and

$$\lambda_{t,Z} \equiv \lim_{\Delta t \rightarrow 0} \frac{\Pr[t < T_P \leq t + \Delta t | H_t, T_P > t, Z]}{\Delta t}. \quad (5.78)$$

Then $\lambda_{t,z}(Z = z)$ in (5.78) can be interpreted as the conditional (with respect to Z in addition to H_t) failure rate of the item, indexed by the frailty Z .

Denote by $\lambda_m^1(t)$ the failure rate $\lambda_c(t)$ in the interval $[0, t_{PM})$ (defined by (5.77) for the Operation profile 1). It obviously equals the mixture failure rate in this interval, i.e.,

$$\lambda_c(t) \equiv \lambda_m^1(t) = \lambda_m(t), \text{ in } [0, t_{PM}),$$

as information at hand is just the initial distribution $\pi(z)$ (and the fact that the item has survived in $[0, t)$).

As the survived item is replaced by an item *randomly selected* from the statistically identical population (but with the initial age αt_{PM}) at $t = t_{PM}$, the conditional failure rate $\lambda_{t,Z}$ in $[t_{PM}, 2t_{PM})$ is

$$\lambda_{t,Z} = \lambda(\alpha t_{PM} + (t - t_{PM}), Z), \tag{5.79}$$

where Z is the frailty randomly selected at the previous PM. Consider now the conditional distribution $Z|(H_t, T_P > t)$. Note that at $t = t_{PM}$, the initial distribution of Z is

$$\pi_{0Z}(z) = \frac{\bar{F}(\alpha t_{PM}, z)\pi(z)}{\int_0^\infty \bar{F}(\alpha t_{PM}, z)\pi(z)dz} \tag{5.80}$$

and we know that the item has additionally survived in $(t_{PM}, t]$. Therefore, the corresponding survival function (for $Z = z$) is

$$\frac{\bar{F}(\alpha t_{PM} + (t - t_{PM}), z)}{\bar{F}(\alpha t_{PM}, z)}.$$

After updating, the conditional distribution $Z|(H_t, T_P > t)$ becomes

$$\frac{\bar{F}(\alpha t_{PM} + (t - t_{PM}), z)\pi(z)}{\int_0^\infty \bar{F}(\alpha t_{PM} + (t - t_{PM}), z)\pi(z)dz}. \tag{5.81}$$

Therefore, in accordance with (5.78), the failure rate $\lambda_m^2(t)$ in $[t_{PM}, 2t_{PM})$ for the described operation is

$$\begin{aligned} \lambda_c(t) &\equiv \lambda_m^2(t) \\ &= \int_0^\infty \lambda(\alpha t_{PM} + (t - t_{PM}), z) \frac{\bar{F}(\alpha t_{PM} + (t - t_{PM}), z)\pi(z)}{\int_0^\infty \bar{F}(\alpha t_{PM} + (t - t_{PM}), z)\pi(z)dz} dz, \text{ in } [t_{PM}, 2t_{PM}). \end{aligned}$$

Similar to (5.81), the conditional distribution $Z|(H_t, T_P > t)$ for the interval $t \in [(n - 1)t_{PM}, nt_{PM})$ is

$$\frac{\bar{F}(t_{n-1} + (t - (n - 1)t_{PM}), z)\pi(z)}{\int_0^\infty \bar{F}(t_{n-1} + (t - (n - 1)t_{PM}), z)\pi(z)dz}$$

and we eventually arrive at

$$\begin{aligned}\lambda_c(t) &\equiv \lambda_m^n(t) \\ &= \int_0^\infty \lambda(t_{n-1} + (t - (n-1)t_{PM}), z) \frac{\bar{F}(t_{n-1} + (t - (n-1)t_{PM}), z)\pi(z)}{\int_0^\infty \bar{F}(t_{n-1} + (t - (n-1)t_{PM}), z)\pi(z)dz} dz, \text{ in } [(n-1)t_{PM}, nt_{PM}),\end{aligned}$$

$n = 1, 2, 3, \dots$, where $t_0 \equiv 0$ and t_{n-1} are defined in (5.74).

Taking into account Eq. (5.12),

$$\lambda_m^n(t) = \lambda_m(t_{n-1} + (t - (n-1)t_{PM})), \quad n = 1, 2, 3, \dots \quad (5.82)$$

and thus, $\lambda_c(t)$ for the Operation profile 1 is given by (5.76). However, this strategy can hardly be realized in the PM practice for many reasons. For instance, even if the item selected at time $t = 0$ has been described by the frailty $Z = z_1$, its value can be changed to $Z = z_2$, $z_1 \neq z_2$ just after the first PM at t_{PM} , which is unrealistic.

Then, what is the proper conditional failure rate for our PM policy? It is more realistic to assume that the original frailty variable $Z = z$ selected at time $t = 0$ is *preserved* throughout the whole operation of an item:

Operation profile 2 *An item is chosen at random from our population and starts operation at $t = 0$. The original frailty that is ‘acquired’ at $t = 0$ is preserved during the PM actions that follow the pattern of the ‘virtual age structure’ defined in (5.74).*

As the PMs are applied to the same item, this operation profile is definitely more adequate than the first one. However, the construction of the corresponding failure rate is completely different in this case.

In $[0, t_{PM})$, the failure rate is still the same:

$$\lambda_c(t) \equiv \lambda_m^1(t) = \lambda_m(t), \text{ in } [0, t_{PM}),$$

as the information at hand is the same as before.

Consider now the second cycle $[t_{PM}, 2t_{PM})$. As the survived item was randomly selected at time $t = 0$ from the heterogeneous population, the conditional failure rate $\lambda_{t,Z}$ in $[t_{PM}, 2t_{PM})$ is given by (5.79), where Z is the frailty ‘randomly selected’ at $t = 0$. At $t = t_{PM}$, the survived item has the frailty $Z = z$ with the pdf that in accordance with (5.12) is

$$\pi_{02}(z) \equiv \frac{\bar{F}(t_{PM}, z)\pi(z)}{\int_0^\infty \bar{F}(t_{PM}, z)\pi(z)dz}.$$

We also have the information that the item with the decreased age αt_{PM} after the PM has additionally survived in $(t_{PM}, t]$. Therefore, the corresponding survival function is

$$\frac{\bar{F}(t_1 + (t - t_{PM}), z)}{\bar{F}(t_1, z)}.$$

In accordance with (5.12), the conditional distribution $Z|(H_r, T_P > t)$ is given now by

$$\frac{\overline{F}(t_1 + (t - t_{PM}), z) / \overline{F}(t_1, z) \cdot \pi_{02}(z)}{\int_0^\infty [\overline{F}(t_1 + (t - t_{PM}), z) / \overline{F}(t_1, z)] \cdot \pi_{02}(z) dz}$$

and the failure rate $\lambda_m^2(t)$ in $[t_{PM}, 2t_{PM})$, in accordance with (5.77), is

$$\begin{aligned} \lambda_m^2(t) &= \int_0^\infty \lambda(t_1 + (t - t_{PM}), z) \cdot \frac{\overline{F}(t_1 + (t - t_{PM}), z) / \overline{F}(t_1, z) \cdot \pi_{02}(z)}{\int_0^\infty [\overline{F}(t_1 + (t - t_{PM}), z) / \overline{F}(t_1, z)] \cdot \pi_{02}(z) dz} dz \\ &= \int_0^\infty \lambda(t_1 + (t - t_{PM}), z) \cdot \frac{\overline{F}(t_1 + (t - t_{PM}), z) / \overline{F}(t_1, z) \cdot \overline{F}(t_{PM}, z) \pi(z)}{\int_0^\infty [\overline{F}(t_1 + (t - t_{PM}), z) / \overline{F}(t_1, z)] \cdot \overline{F}(t_{PM}, z) \pi(z) dz} dz. \end{aligned}$$

In a similar way, for $t \in [(n - 1)t_{PM}, nt_{PM})$,

$$\begin{aligned} \lambda_m^n(t) &= \int_0^\infty \lambda(t_{n-1} + (t - (n - 1)t_{PM}), z) \cdot \frac{\overline{F}(t_{n-1} + (t - (n - 1)t_{PM}), z) / \overline{F}(t_{n-1}, z)}{\int_0^\infty [\overline{F}(t_{n-1} + (t - (n - 1)t_{PM}), z) / \overline{F}(t_{n-1}, z)]} \\ &\quad \times \frac{\overline{F}(t_{PM}, z) \cdot \frac{\overline{F}(t_1 + t_{PM}, z)}{\overline{F}(t_1, z)} \dots \frac{\overline{F}(t_{n-2} + t_{PM}, z)}{\overline{F}(t_{n-2}, z)} \pi(z)}{\overline{F}(t_{PM}, z) \cdot \frac{\overline{F}(t_1 + t_{PM}, z)}{\overline{F}(t_1, z)} \dots \frac{\overline{F}(t_{n-2} + t_{PM}, z)}{\overline{F}(t_{n-2}, z)} \pi(z) dz}, \end{aligned} \tag{5.83}$$

where $t_{n-1}, n = 1, 2, 3, \dots, (t_0 \equiv 0)$ are defined in (5.74).

Observe that conditional failure rates for both operation profiles can now be uniformly written as

$$\lambda_c^J(t) = \sum_{n=1}^{\lceil t/t_{PM} \rceil + 1} \lambda_{m,J}^n(t) I((n - 1)t_{PM} \leq t < nt_{PM}), \quad J = I, II,$$

where $I(\cdot)$ is the corresponding indicator and $J = I, II$ refers to the number of the profile. Thus, $\lambda_{mI}^n(t)$ corresponds to $\lambda_m^n(t)$ in (5.82) and $\lambda_{mII}^n(t)$ to $\lambda_m^n(t)$ in (5.83).

Therefore, in practice, $\lambda_c^I(t)$ (not $\lambda_c^I(t)$) should be applied for the described type of PM. However, assume that the user, who is performing the PM (via reducing the age of items by the method described previously), does not know (or does not take into account) the heterogeneity structure of the population and considers it as homogeneous with the corresponding time to failure distribution $F_m(t)$ and the failure rate $\lambda_m(t)$. Then he is using the failure rate $\lambda_c^I(t)$ to assess reliability of items in operation. What is the consequence of this error? The following theorem answers to this question.

Theorem 5.10 *Let the values of $\lambda(t, z)$. be ordered with respect to z : for all $z_1, z_2 \in [0, \infty]$, $t \geq 0$.*

$$\lambda(t, z_1) < \lambda(t, z_2) \text{ if } z_1 < z_2.$$

Then

$$\lambda_c^{II}(t) \leq \lambda_c^I(t), \text{ for all } t \geq 0.$$

The *proof* of this theorem is rather straightforward (although technical) and can be found in Cha and Finkelstein [21]. □

It follows from this theorem that using $\lambda_c^I(t)$ instead of the ‘proper’ $\lambda_c^{II}(t)$ eventually results in the overestimation of the failure rate of items under operation. Practically, this may cause unnecessary frequent PMs and therefore, additional redundant costs.

Example 5.5 Suppose that $\lambda(t, z)$ is strictly increasing in t for each z (e.g., $\lambda(t, z) = z\lambda t, \lambda > 0$). An item is randomly selected from the heterogeneous population and it is preventively maintained at times $kt_{PM}, k = 1, 2, \dots$. Let τ be the mission time of the item in field operation. If the mission is successful, a gain $K > 0$ is obtained, whereas if the mission is not completed (a failure in $[0, \tau)$), a cost $c_f > 0$ is incurred ($K > c_f$). Furthermore, the cost for each PM is $c_p > 0$. Then, the following cost function, which is the function of t_{PM} , can be constructed.

$$\begin{aligned} c(t_{PM}) &= \left\langle \frac{\tau}{t_{PM}} \right\rangle c_p + c_f \cdot P(T_p \leq \tau) - K \cdot P(T_p > \tau) \\ &= \left\langle \frac{\tau}{t_{PM}} \right\rangle c_p - (K + c_f) \cdot \exp \left\{ - \int_0^\tau \lambda_c^{II}(u) du \right\} + c_f, \end{aligned} \tag{5.84}$$

where $\langle \tau/t_{PM} \rangle$ is the largest integer which is *strictly* less than τ/t_{PM} . The problem is to find the optimal t_{PM}^* which satisfies

$$c(t_{PM}^*) = \min_{t_{PM} \in (0, \infty)} c(t_{PM})$$

It is reasonable to consider only $t_{PM} \in (0, \tau]$ as $c(t_{PM}) = c(\tau)$, for all $t_{PM} \in (\tau, \infty)$. When $t_{PM} \rightarrow 0$ ($[\tau/t_{PM}] \rightarrow \infty$), obviously, $\exp\{-\int_0^\tau \lambda_c^{II}(u) du\} \rightarrow \exp\{-\lambda_m(0)\tau\}$, which implies that $\lim_{t_{PM} \rightarrow 0} c(t_{PM}) = \infty$. On the other hand, $c(\tau) = c_f - (K + c_f) \exp\{-\int_0^\tau \lambda_m(u) du\}$. Therefore, there should be an optimal $t_{PM}^* \in (0, \tau)$ depending on the parameters involved, e.g., when τ is large enough and K is relatively large compared with c_f and c_p .

5.10 Population Mortality at Advanced Ages (Demographic Application)

In Sects. 5.4 and 5.5, we have briefly discussed asymptotic behavior of mixture failure rates as $t \rightarrow \infty$. In the current section, we will deal with this problem from a different view point and in more detail [31].

The shape of the failure rate (force of mortality) at advanced ages especially for human populations has attracted a considerable interest in the last decades when more and more centenarians and super centenarians have been recorded. The International Database on Longevity (<http://www.supercentenarians.org/>) offers the detailed information on thoroughly validated cases of super centenarians. Gampe [35] has used these data to estimate the human force of mortality after the age of 110. Her analysis revealed that human mortality between ages 110 and 114 levels off regardless of gender. The widely used explanation of this fact is by employing the corresponding fixed frailty models that account for heterogeneity of populations. Beard [7, 8] (see also Vaupel et al. [63]) has considered the Gompertz (baseline)-gamma-frailty model, which results in the asymptotically flat hazard rate. Note that, the exponentially increasing hazard rate of the Gompertz distribution is the only baseline function that can ‘produce’ this shape in the framework of the multiplicative frailty model (see Sect. 5.3.1), which can be considered as another justification of the uniqueness and importance of this distribution for human mortality modeling. As follows from the results of Sect. 5.4, the gamma distribution of frailty is not so unique in this respect and all probability density functions $f(z)$ that behave as z^α , $\alpha > 1$ when $z \rightarrow 0$ are equivalent in this sense.

The intuitive meaning of the deceleration of mortality at advanced ages in this context is simple and meaningful at the same time: the oldest-old mortality in *heterogeneous* populations with properly ordered subpopulations is defined by the small values of frailty, as the subpopulations with larger values of frailty (and, therefore, larger values of the failure rate) are dying out first.

The first question to be answered is what *common* statistical distributions are characterized by the asymptotically flat failure rate? The exponential distribution that is often used for statistical analysis of non-degrading objects is obviously not relevant for our topic. The most popular distribution of the desired type is the inverse Gaussian distribution. It is well known that it describes the distribution of the first passage time for the Wiener process with drift. Although its sample paths are nonmonotone and even can be nonpositive, the inverse Gaussian distribution was widely used, e.g., in reliability analysis of stochastic deterioration (aging) in engineering objects. It was also applied in vitality models for modeling the life-span of organisms [3, 45], where the initial vitality (resource) of organisms is ‘consumed’ in the course of life in accordance with the Wiener process with drift. This model was also studied in the path-breaking papers by Aalen and Gjessing [1] and Steinsaltz and Evans [55] as an example highlighting the meaning and properties of the corresponding quasistationary distributions for this particular case. Our goal in this section is more modest: to exploit further some relevant

distributional properties in the context of stochastic ordering of lifetimes of subpopulations in heterogeneous populations. However, the combination of these two approaches can hopefully be considered as the basis for the future research on statistical inference in heterogeneous populations with underlying stochastic processes (e.g., the Wiener process).

The other example of a distribution with asymptotically flat hazard rate is the Birnbaum-Saunders distribution [12] that was also derived as a distribution of the first passage time for the corresponding deterioration process and, therefore, is a good candidate for vitality models as well. We also consider the gamma process as a possible model of deterioration (with monotone sample paths), although the failure rate in this case is decreasing to 0 as $t \rightarrow \infty$. It should be noted, however, that the initial increase in the failure rates for all these models is not exponential, as in the case of the Gompertz distribution and, therefore, the possibilities of the corresponding mortality modeling for *human populations* for intermediate ages (30–90 years) are obviously limited.

5.10.1 Fixed and Evolving (Changing) Heterogeneity

Let $F(t)$, $f(t)$, and $\lambda(t)$ be the Cdf, the pdf, and the failure rate (force of mortality) for some infinite *homogeneous* population that characterize the corresponding random lifetime $T \geq 0$. As previously, by heterogeneity of a population we mean that it consists of a finite or non-finite number of homogeneous subpopulations that differ in some respect to be discussed. For instance, in the multiplicative frailty model of the form $\lambda(t, Z) = Z\lambda(t)$, the difference between subpopulations is modeled directly by the differences in failure rates: for two realizations $z_2 > z_1$, this difference is $(z_2 - z_1)\lambda(t)$. Thus, the multiplicative frailty model describes the ordering of subpopulations in the sense of the *hazard rate ordering* (2.70). More generally, the smaller is the value of z , the larger is the lifetime of the subpopulation T_z in the appropriate stochastic sense (e.g., (2.69), (2.70) or (2.71)):

$$T_{z_1} \geq T_{z_2}; \quad z_1 \leq z_2. \quad (5.85)$$

As previously in this chapter, we will understand the fixed heterogeneity (frailty) of a population as:

Heterogeneity in lifetimes of the corresponding homogeneous subpopulations that is defined by the appropriate stochastic ordering.

This also means that, if randomization of a parameter (parameters) of a lifetime distribution leads to the corresponding stochastic ordering, which formally is not always the case, then this operation can be also interpreted in terms of the fixed frailty modeling. For example, the Gompertz Cdf $F(t, a, b)$ is a function of two parameters, and the corresponding failure rate is:

$$\lambda(t, a, b) = ae^{bt}. \quad (5.86)$$

If we randomize a , whereas b is fixed, then (taking care, of course, of the corresponding baseline constant), we obviously arrive at the multiplicative frailty model (and to the asymptotically flat rate when the distribution of frailty is, e.g., gamma), which illustrates ordering (2.70). We just want to emphasize the fact that in this specific model, frailty acts multiplicatively and directly on the failure rate, which is not the case in general even when the hazard rate ordering (2.70) holds. Some relevant aspects of frailty modeling for the bivariate case will be considered later.

In accordance with our definition, the fixed heterogeneity (frailty) is described only by ordered subpopulation lifetimes. What can happen, if apart from the information on failure times (the black box point of view), we possess some information or adopt a model on a failure process or mechanism (the process point of view)? In this case, another type of heterogeneity, which is usually referred to as evolving (or changing) (see, e.g., Li and Anderson [45]) comes into play. This type of heterogeneity usually does not lead to ordering of lifetimes in the described here sense. However, it characterizes an important feature of a model, which can be useful for further analysis.

In order to illustrate our point, consider the model for vitality loss (fixed initial value) that will be treated in detail further in this section. The loss of vitality of an organism (deterioration) is modeled by the Wiener process with negative drift, in which the time to death is determined by the first passage time to the zero boundary. It is well known that the variance of the Wiener process is increasing linearly in time and if the drift is positive, the mean is also linearly increasing. However, due to the boundary, the most vulnerable organisms (or items in reliability engineering applications) are dying out first and linear functions that correspond to the non-boundary case ‘decelerate’. Actual shapes depend on parameters of the model (see the graphs in Li and Anderson [45] for the corresponding shapes for the specific values of parameters). Thus we do not see here any frailty parameters or ordered (in the defined in this section sense) lifetimes, but we observe the changing in time mean and variability in the survived population. And this is how the evolving heterogeneity should be understood:

Variability in sample paths of the underlying process of deterioration.

In this section, however, we are mostly interested in the fixed heterogeneity of lifetimes and the evolving heterogeneity of processes will be ‘hidden’ in lifetime distributions. We feel that this ‘distributional approach’ in the context of randomization of parameters and of the corresponding ordering of lifetimes was not sufficiently elaborated in the literature so far. For instance, for the first passage time models to be considered further, randomization of the initial vitality of an organism and of the corresponding drift parameter of the Brownian motion definitely illustrates this ordering, as the larger is the vitality and (or) the smaller is the drift parameter, the larger is the lifetime in some suitable stochastic sense to be discussed. Note that, there can be other situations when randomization is relevant but does not lead to the ordered subpopulations.

5.10.2 Fixed Heterogeneity

Equations (5.10)–(5.12) describe the standard statistical mixture (or the fixed frailty) model for an item and for the collection of items (population) as well. As was discussed in the previous subsection, we understand heterogeneity as the property of a population that consists of ordered homogeneous subpopulations (ordered lifetimes T_z , defined by Inequality (5.85)). But what type of ordering is sufficient for our reasoning? As we are looking at the failure rates, the first guess would be that this should be (2.70). How can we interpret in mathematical terms the well-known and intuitively clear property: “the weakest populations are dying out first” and the resulting mortality deceleration with time? To answer these questions, denote, as previously, by $\Pi(z)$ the Cdf of Z and by $\Pi(z|t)$ the Cdf that corresponds to the density $\pi(z|t)$. Therefore, the deceleration can be a consequence of the increasing in t distribution function $\Pi(z|t)$ [28]. This would mean that $\Pi(z|t)$ tends to be more concentrated around small values of $Z \geq 0$ as time increases, which corresponds to stronger populations. The following theorem proves this result.

Theorem 5.11 *Let stochastic ordering (5.85) in the sense of the failure rates hold. Then $\Pi(z|t)$ is a non-decreasing function of t for each fixed z .*

Proof. It follows from (5.12) that

$$\Pi(z|t) = \frac{\int_0^z \bar{F}(t, u)\pi(u)du}{\int_0^\infty \bar{F}(t, u)\pi(u)du}.$$

It is easy to see that the derivative of this function is nonpositive if

$$\frac{\int_0^z \bar{F}'_t(t, u)\pi(u)du}{\int_0^z \bar{F}(t, u)\pi(u)du} \geq \frac{\int_0^\infty \bar{F}'_t(t, u)\pi(u)du}{\int_0^\infty \bar{F}(t, u)\pi(u)du}.$$

Therefore, it is *sufficient* to show that the function:

$$A(t, z) = \frac{\int_0^z \bar{F}'_t(t, u)\pi(u)du}{\int_0^z \bar{F}(t, u)\pi(u)du}$$

is nonincreasing in z . As $\bar{F}'_t(t, z) = -\mu(t, z)\bar{F}(t, z)$, inequality $A'_z(t, z) \leq 0$ is equivalent to the following one:

$$\mu(t, z) \int_0^z \bar{F}(t, u)\pi(u)du \geq \int_0^z \mu(t, u)\bar{F}(t, u)\pi(u)du,$$

which obviously follows from Ordering (5.85) which should be understood in the sense of the hazard rate ordering. □

Consider now the bivariate frailty model. We will need the following considerations for analyzing asymptotic failure rates for vitality models of the next subsection. Let Z_1 and Z_2 be interpreted as non-negative random variables with supports in $[0, \infty)$. Similar to the univariate case,

$$P(T \leq t | Z_1 = z, Z_2 = z_2) \equiv P(T \leq t | z_1, z_2) = F(t, z_1, z_2)$$

and

$$\lambda(t, z_1, z_2) = \frac{f(t, z_1, z_2)}{\bar{F}(t, z_1, z_2)}$$

Assume that Z_1 and Z_2 have the joint pdf $\pi(z_1, z_2)$. The mixture failure rate is defined in this case as [28]:

$$\begin{aligned} \lambda(t) &= \frac{f(t)}{F(t)} = \frac{\int_0^\infty \int_0^\infty f(t, z_1, z_2) \pi(z_1, z_2) dz_1 dz_2}{\int_0^\infty \int_0^\infty \bar{F}(t, z_1, z_2) \pi(z_1, z_2) dz_1 dz_2} \\ &= \int_0^\infty \int_0^\infty \lambda(t, z_1, z_2) \pi(z_1, z_2 | t) dz_1 dz_2, \end{aligned} \tag{5.87}$$

where the corresponding conditional pdf (on condition $T > t$) is

$$\pi(z_1, z_2 | t) = \pi(z_1, z_2) \frac{\bar{F}(t, z_1, z_2)}{\int_0^\infty \int_0^\infty \bar{F}(t, z_1, z_2) \pi(z_1, z_2) dz_1 dz_2}. \tag{5.88}$$

Equation (5.87) is a general result and can be analyzed for some specific cases. For instance, it can be easily shown that when we assume the independence of frailties:

$$\pi(z_1, z_2) = \pi_1(z_1) \pi_2(z_2)$$

and the competing risks for the failure model:

$$F(t, z_1, z_2) = 1 - \bar{F}_1(t, z_1) \bar{F}_2(t, z_2),$$

the population failure rate is just the sum $\lambda(t) = \lambda_1(t) + \lambda_2(t)$ of the corresponding ‘univariate failure rates’.

Although it is difficult to analyze $\lambda(t)$ in (5.87) in full generality, certain qualitative considerations that will be very helpful in the next subsection can be stated. Indeed, let us first fix the second frailty $Z_2 = z_2$. Then the corresponding failure rate is defined by the univariate frailty model

$$\lambda(t, z_2) = \int_0^\infty \mu(t, z_1, z_2) \pi(z_1, z_2 | t) dz_1. \tag{5.89}$$

Thus, at the *first stage*, we have selected from our overall heterogeneous population the *heterogeneous subpopulation* that corresponds to $Z_2 = z_2$

$(z_2 < Z_2 \leq z_2 + dz_2)$ and have defined its failure rate. As our goal is to analyze the failure rate, at the *second stage*, we consider our overall population as a ‘continuous collection’ of *homogeneous subpopulations* with failure rates given by (5.89). Then we can analyze $\lambda(t)$ again in the univariate manner. For instance, assume that the family $\lambda(t, z_2)$ is ordered in z_2 (the smaller values of z_2 correspond to the smaller values of $\lambda(t, z_2)$). Therefore, the deceleration in mortality due to ‘the weakest populations are dying out first’ takes place. Specifically, let $\lambda(t, z_2)$ for each z_2 decreases (nonincreases) at least, asymptotically when $t \rightarrow \infty$. It is well known that the corresponding population (mixture) failure rate is strictly decreasing in this case (see, e.g., Ross [54]). Thus, we have described the following result [31]:

Theorem 5.12 *Let frailty $Z_1 = z_1$ ($Z_2 = z_2$) in the bivariate frailty model be first fixed. Assume that the corresponding univariate frailty model (with respect to Z_2 (Z_1)) results in the decreasing ordered failure rates for all subpopulations.*

Then ‘allowing’ random Z_1 (Z_2), results in the strictly decreasing population failure rate.

The formal *proof* of the validity of the two-stage procedure is straightforward and is based on the representation of the bivariate density $\pi(z_1, z_2)$ as a product $\pi_1(z_1|Z_2 = z_2)\pi_2(z_2)$ and on the similar representation for the conditional density:

$$\pi(z_1, z_2|t) = \pi_1(z_1|Z_2 = z_2, T > t)\pi_2(z_2|t).$$

The latter seems intuitively evident, and can be immediately obtained formally from Eqs. (5.87), (5.88). Theorem 2 then follows, as the (univariate) mixture of distributions with decreasing (nonincreasing) failure rates is characterized by the strictly decreasing failure rate. □

Example 5.6 An important application that illustrates Theorem 2 deals with the Gompertz law of mortality (5.86). It is well known that randomization of a (e.g., via the gamma distribution of the frailty) results in the mortality plateau as $t \rightarrow \infty$. Thus, randomization of b (second stage) results in the decreasing force of mortality as $t \rightarrow \infty$. Therefore, if we observe the mortality plateau for some population that follows the Gompertz-gamma model, then there should not be noticeable heterogeneity in this population due to parameter b .

The described multistage approach can be applied in a similar way to the case when there are more than 2 frailties or parameters of distributions that can be randomized. It is possible that all failure rates from the ordered family converge asymptotically (as $t \rightarrow \infty$) to one curve (specifically, to a constant). Therefore, the population failure rate also tends to this curve which will also be illustrated in the next subsection.

The foregoing discussion will help us to analyze the shape of the failure rate for some examples of vitality models. We will focus mostly on the vitality model described by the Wiener process with drift [3, 45, 64]. Parameters of lifetime distributions after randomization will act as fixed frailties that define the

corresponding ordered subpopulations. This interpretation adds some simple and useful additional reasoning from the *distributional point of view* to the *process point of view* approach developed by Aalen and Gjessing [1] and Steinsaltz and Evans [55].

5.10.3 Vitality Models and Lifetime Distributions

Linear process of degradation. We start with the simplest vitality model that will be used as an explanatory example for highlighting certain properties and approaches.

Let $v_0 > 0$ be the deterministic initial (at $t = 0$) vitality of an organism, which is monotonically decreasing with t in accordance with the simplest stochastic process:

$$V_t = v_0 - Rt, \quad (5.90)$$

where R is a positive random variable with the Cdf $S(t)$. For each realization $R = r$, (5.90) can model the linear decline in physiological functions of organisms noted by Strehler and Mildvan [57] and in numerous subsequent publications. However, exponential and logarithmic models for this decline can be also considered.

Death occurs when V_t reaches 0. Denote the corresponding lifetime by T_R . Therefore, the Cdf that describes this lifetime is

$$F_R(t) = \Pr[T_R \leq t] = \Pr[R \geq v_0/t] = 1 - S(v_0/t).$$

Assume that R is gamma-distributed with the pdf $a^\eta x^{\eta-1} e^{-ax} / \Gamma(\eta)$ with the scale parameter $a > 0$ and the shape parameter $\eta > 0$. Then the pdf $f_R(t) = F'_R(t)$ has the form of the inverse gamma distribution:

$$f_R(t) = \frac{(v_0 a)^\eta}{\Gamma(\eta)} t^{-\eta-1} e^{-v_0 a/t}. \quad (5.91)$$

We will analyze the shape of the corresponding hazard rate using the ‘classic’ Glazer’s theorem [37], formulated in a slightly more general form by Marshall and Olkin [48] as can be seen from Theorem 2.1 in Chap. 2. We will intensively use this result and other relevant considerations in what follows.

The essential fact to be exploited is that the behavior of the failure rate $\lambda(t)$ is related to the behavior of the derivative of the logarithm of the density of a lifetime distribution $F(t)$, namely,

$$g(t) = -\frac{d \log f(t)}{dt} = -\frac{f'(t)}{f(t)}.$$

The failure rate $\lambda_R(t)$ that corresponds to (5.91) can be easily analyzed with the help of Theorem 2.1. Indeed, as $\lim_{t \rightarrow \infty} f_R(t) = 0$, it follows that $\lim_{t \rightarrow \infty} \lambda_R(t) = 0$, whereas

$$\lim_{t \rightarrow \infty} \lambda_R(t) = \lim_{t \rightarrow \infty} f_R(t)/\bar{F}_R(t) = \lim_{t \rightarrow \infty} -\frac{d \log f_R(t)}{dt} = 0$$

and $\lambda_R(t)$ is bell-shaped with a maximum at $t = 2v_0a/(\eta + 1)$.

This simple example, however, can be helpful for discussing the notion of heterogeneity that we adopt. If we consider the model as a black box with the lifetime described by the Cdf $F_R(t)$, then by definition, the corresponding population is homogeneous. However, in view of the model (5.90), we can identify the corresponding subpopulations for each value of $R = r$ that will be definitely ordered (in this case the lifetimes that correspond to each realization $R = r$ are deterministic, and therefore, can be ordered accordingly). Thus, our infinite population can be considered as heterogeneous in the described sense.

The considered vitality model results in the vanishing at the infinity failure rate. If we are interested in explaining mortality plateaus that has been observed in human and other populations, then we must look at other, more realistic vitality models. The first candidate for that is when the simplest stochastic process Rt is substituted by the more advanced stochastic model given by the Wiener process with drift.

Wiener process with drift. We modify the degradation model (5.90) with the fixed initial vitality v_0 to

$$\begin{aligned} V_t &= v_0 - R_t, \\ R_t &= rt + W_t, \end{aligned} \tag{5.92}$$

where $R_t, t \geq 0$ is the Wiener process with drift, r is a drift parameter and $W_t, t \geq 0$ is the standard Wiener process with normally distributed values (for each fixed t) with mean 0 and variance $\sigma^2 t$.

It is well known (see, e.g., [24]) that the probability distribution for the first passage time (when R_t reaches the boundary v_0 for the first time) is defined by the inverse Gaussian distribution with the pdf:

$$f_R(t) \equiv f_R(t; v_0, r, \sigma) = \frac{v_0}{\sigma\sqrt{2\pi}} t^{-3/2} \exp\left\{-\frac{(v_0 - rt)^2}{2\sigma^2 t}\right\}. \tag{5.93}$$

The exact expression for the corresponding failure rate, $\lambda_R(t) \equiv \mu_R(t; v_0, r, \sigma)$, is complicated and, therefore, as our goal is just to analyze its *shape*, we will use Theorem 2.1. It is easy to derive from (5.93) that

$$g_R(t) = -\frac{d \log f_R(t)}{dt} = \frac{3}{2t} + \frac{r^2}{2\sigma^2} - \frac{v_0^2}{2\sigma^2 t^2}. \tag{5.94}$$

Note that, (5.93) is written in parameterization v_0, r, σ . However, reparameterization: $\lambda = r^2/\sigma^2, \omega = rv_0/\sigma^2$ leads to the standard two-parameter form of the inverse Gaussian distribution (which we need for stating some useful properties):

$$f_R(t; \lambda, \omega) = \frac{\lambda\omega}{\sigma\sqrt{2\pi}} (\lambda t)^{-3/2} \exp\left\{-\frac{(\omega - \lambda t)^2}{2\lambda t}\right\}. \quad (5.95)$$

It immediately follows from (5.94) that the failure rate tends to a constant when $t \rightarrow \infty$ (mortality plateau):

$$\lim_{t \rightarrow \infty} \lambda_R(t) = \lim_{t \rightarrow \infty} -\frac{d \log f_R(t)}{dt} = \frac{\lambda}{2} = \frac{r^2}{2\sigma^2}. \quad (5.96)$$

It is also obvious that $\lim_{t \rightarrow 0} \lambda_R(t) = 0$. The ‘rest of the shape’ of $\lambda_R(t)$ is defined by Theorem 2.1: $\lambda_R(t)$ is increasing for $t \in [0 \leq t_2]$, where $t_2 \leq t_1 = 2v_0^2/3\sigma^2$ and is asymptotically decreasing to the plateau for $t \geq t_2$. This form of the hazard rate for the inverse Gaussian distribution was first described by Chhikara and Folks [22] using straightforward calculus and asymptotic bounds. We, however, rely on a general Theorem 2.1 that can be used for analysis of other distributions as well.

Although the ‘underlying physics’ of the inverse Gaussian distribution is given by the Wiener process with drift, we cannot identify now the corresponding subpopulations in the sense that we have defined earlier. Therefore, the corresponding population in this ‘black-box’ analysis should be considered as *homogeneous* and there is no (fixed) heterogeneity in the defined sense so far.

From (5.95) it follows that λ is the scale parameter. Therefore, obviously, the corresponding lifetimes are decreasing in λ in the sense of the usual stochastic ordering (2.69), i.e., for the fixed ω :

$$F_R(\lambda_1 t; \omega) \leq F(\lambda_2 t; \omega); \quad \lambda_1 \leq \lambda_2, t \in [0, \infty). \quad (5.97)$$

This is a simple general fact. However, for the specific case of inverse Gaussian distribution, it can be shown that the stronger hazard rate ordering (2.70) also takes place [48], which means:

$$\begin{aligned} \mu_R(t; \lambda_1, \omega) &= \lambda_1 \mu_R(\lambda_1 t; \omega_1) \leq \lambda_2 \mu_R(\lambda_2 t; \omega_2) = \lambda_R(t; \lambda_2, \omega); \\ \lambda_1 &\leq \lambda_2, t \in [0, \infty). \end{aligned}$$

As $\lambda = r^2/\sigma^2$, the distribution of the first passage time $f_R(t; \lambda, \omega)$ does not change when we change r and σ proportionally. Thus the mechanism of the failure process driven by the Wiener process with drift is such that, e.g., the increase in the drift parameter is compensated by the proportional increase in the standard deviation σ . This is a rather unexpected observation, however, as stated, it is a consequence of the considered specific setting. Strictly speaking, as parameters λ and ω are ‘dependent’ the foregoing orders hold only asymptotically for large t and this is how we will understand it in what follows.

After discussing the issue of stochastic ordering, we can now *qualitatively* analyze the shape of $\lambda_R(t; \lambda, \omega)$ for large t with respect to the randomized parameters r and σ (v_0 is fixed so far) to be denoted by R and Σ , respectively. Note that, Aalen and Gjessing [1], have performed the necessary derivations assuming that R is normally distributed and σ is fixed. However, as the drift ($-r$) can be positive in this case, the resulting survival distribution is defective. These distributions are often used for describing the corresponding ‘cure models’.

Assume that R and Σ are non-negative random variables with supports in $[0, \infty)$. Thus, the bivariate frailty model discussed in Sect. 3 can be applied. We proceed as described there: fixing $\Sigma = \sigma$ and considering subpopulations with one frailty parameter R . At the *first stage*, we select from the overall heterogeneous population the heterogeneous (with respect to different values of r) *subpopulation* that corresponds to $\Sigma = \sigma$ and define its failure rate. As the corresponding homogeneous ‘sub-subpopulations’ (for different fixed values of r) are ordered in the sense of the hazard rate ordering and ‘have’ the shapes of the failure rates described above (increasing and then decreasing to a plateau), this heterogeneous subpopulation has asymptotically decreasing to 0 failure rate [54]. Now, at the second stage, as these failure rates are ordered with respect to the values of the second frailty $\Sigma = \sigma$, we can use Theorem 5.12, which means that the population failure rate is also decreasing as $t \rightarrow \infty$ (and in our specific case, it is decreasing to 0).

Thus, *mortality plateaus cannot occur in the described frailty model*. However, this can still happen, if the supports of frailties R and Σ are modified to $[a, \infty]$ and $[0, b]$, respectively. Then the population failure rate tends to the failure rate of the strongest subpopulation which is, in accordance with (5.96) [31],

$$\lim_{t \rightarrow \infty} \lambda_R(t) = \frac{a^2}{2b^2}. \quad (5.98)$$

We are ready now to add variability to the initial vitality. Denote the corresponding random variable by $V_0 \geq 0$ (fixed frailty). It immediately follows from (5.96) that, in contrast to the other considered fixed frailties, the effect of the initial vitality vanishes as $t \rightarrow \infty$. Therefore, it has no effect asymptotically on the shape of the failure rate. This was analytically shown and discussed using the concept of quasistationary distributions in Aalen and Gjessing [1], Steinsaltz and Evans [55], and Li and Anderson [45].

Gamma process and the Birnbaum-Saunders distribution. The Wiener process is often criticized as a model for degradation and aging as its sample paths are not necessarily positive and strictly increasing. On the other hand, the gamma process always possesses these properties. Therefore, let $R_t, t \geq 0$ be now the stationary gamma process with the following density for each t :

$$f_{R_t}(x) = Ga(x|r^2t/\sigma^2, r/\sigma^2), \quad \mu, \sigma > 0, \quad (5.99)$$

$$E[R_t] = rt, \quad \text{Var}(R_t) = \sigma^2t,$$

where $Ga(x|\alpha, \beta)$ denotes the gamma distribution with shape parameter α and scale parameter β . We see that the mean and the variance of this process have the same functional form as for the corresponding Brownian motion with drift. The first passage time distribution function for the vitality model with initial value v_0 is

$$\begin{aligned}
 F_{R_t}(t) &= \Pr[T_R \leq t] = \Pr[R_t \geq v_0] \\
 &= \int_{v_0}^{\infty} f_{R_t}(x) dx = \frac{\Gamma(r^2 t / \sigma^2, v_0 r / \sigma^2)}{\Gamma(r^2 t / \sigma^2)}, \tag{5.100}
 \end{aligned}$$

where $\Gamma(a, x) = \int_x^{\infty} z^{a-1} e^{-z} dz$ is the incomplete gamma function for $x \geq 0$ and $a > 0$. This function can be calculated numerically [61]. It is shown by Liao et al. [46] that the corresponding failure rate is increasing, whereas Abdel-Hameed [2] proves that it tends to infinity as $t \rightarrow \infty$, which means that the *mortality plateau cannot occur* in accordance with this model.

Park and Padgett [53] have derived a very complex exact expression for the pdf $f_R(t)$. Therefore, a simpler meaningful approximation for (5.100) was suggested by these authors in the form of the Birnbaum-Saunders distribution that can be already effectively analyzed. In a general form, this distribution is given by

$$F_{BS}(t; \lambda, \alpha) = \Phi(\alpha^{-1} h(\lambda t)), \quad t > 0, \tag{5.101}$$

where $\lambda, \alpha > 0$; $\Phi(\cdot)$ is a standard normal distribution function and $h(t) = t^{1/2} - t^{-1/2}$. For our specific case, the corresponding approximation reads [61]:

$$F_{R_t}(t) \approx \Phi\left(\sqrt{\frac{v_0 r}{\sigma^2}} \left[\sqrt{\frac{rt}{v_0}} - \sqrt{\frac{v_0}{rt}} \right]\right). \tag{5.102}$$

It was obtained by Park and Padgett [53] via discretization of the first passage time and then using the central limit theorem. The error of the approximation was not assessed, however, it was stated that it can be used at least for the case when $r \gg \sigma$. On the other hand, it should be noted that approximation of distribution functions does not necessarily mean that the tails of the failure rate functions are also approximated. Therefore, given our interest in asymptotic behavior of failure rates, why not to start directly from distribution (5.102) that, similar to the inverse Gaussian distribution, also has a meaningful *process point of view* interpretation. To see this, consider the following damage accumulation model. Let R_t in (5.92) be modeled by the following shock process: suppose that shocks occur at regular intervals at times $\Delta, 2\Delta, 3\Delta, \dots$. Let each shock causes a random damage $Y_i > 0$: i.i.d with $E[Y_i] = \Delta\mu, \text{Var}(Y_i) = \Delta\sigma^2$. Damages accumulate additively and the k -th shock is survived if the accumulated damage is less than the initial vitality v_0 , i.e., $\sum_1^k Y_i \leq v_0$. Then, letting $\Delta \rightarrow 0$ and using the central limit theorem, after straightforward derivations [48] one can obtain the lifetime distribution (5.100), where

$$\alpha = \sigma/\sqrt{\mu v_0}, \quad \lambda = \mu/v_0. \quad (5.103)$$

Differentiation of (5.101) results in the following density

$$f_{BS}(t; \lambda, \alpha) = \frac{\lambda}{2\alpha\sqrt{2\pi}} \left[\frac{1}{\sqrt{\lambda t}} \left(1 + \frac{1}{\lambda t} \right) \right] \exp \left\{ -\frac{1}{2\alpha^2} \left(\lambda t - 2 + \frac{1}{\lambda t} \right) \right\}. \quad (5.104)$$

Obviously, $\lim_{t \rightarrow 0} \lambda_{BS}(t; \lambda, \alpha) = 0$. Using Theorem 2.1, it can be shown now that the failure rate is bell-shaped [9] and is decreasing to a constant as $t \rightarrow \infty$ (mortality plateau):

$$\begin{aligned} \lim_{t \rightarrow \infty} \lambda_{BS}(t; \lambda, \alpha) &= \lim_{t \rightarrow \infty} -\frac{d \log f_{BS}(t; \lambda, \alpha)}{dt} \\ &= \frac{\lambda}{2\alpha^2} = \frac{\mu^2}{2\sigma^2}. \end{aligned} \quad (5.105)$$

It follows from (5.105) that, as previously, the effect of initial vitality v_0 is vanishing as $t \rightarrow \infty$. Similar to the case of the inverse Gaussian distribution, it can be seen from (5.104) that $\lambda = \mu/v_0$ is a *scale parameter* and, therefore, the usual stochastic ordering (and the hazard rate ordering) holds, i.e., if $v_0(\mu)$ is fixed, then the larger values of $\mu(v_0)$ will result in the larger (smaller) values of the failure rate in $[0, \infty)$.

The possibility of ordering with respect to the values of σ for a general case is not clear (it is an open question in the theory of this distribution). On the other hand, as follows from (5.105), this ordering exists *asymptotically*. Assume now that μ is a realization of a random variable M , whereas σ is a realization of a random variable Σ with support to $[0, \infty]$. Then, similar to the case of the inverse Gaussian distribution, the randomization results in the asymptotically decreasing to 0 population failure rate. Mortality plateaus are theoretically possible in this model only when the supports of the frailties M and Σ are $[a, \infty]$ and $[0, b]$, respectively.

5.11 On the Rate of Aging in Heterogeneous Populations

In this section, we will consider another application of heterogeneity modeling to demography [30]. It should be noted that because of the existing heterogeneity, e.g., in populations for different countries, statistical models describing this property are crucial for this discipline.

Non-parametric classes of lifetime distributions were extensively studied in numerous publications of the last decades (see e.g., the excellent encyclopedic monograph by Lai and Xie [43] and the references therein). One of the main properties of a lifetime random variable that defines the corresponding non-parametric class is a property of *stochastic aging*. This notion can be understood in many ways. The most intuitively evident and the first to be considered in the

literature was the class of aging distributions with increasing (nondecreasing) failure rate (IFR) (see, e.g., Barlow and Proschan [6] for this and other basic classes).

Let $T \geq 0$ be a lifetime with an absolutely continuous Cdf $F(t)$, pdf $f(t)$ and the failure rate $\lambda(t) = f(t)/(1 - F(t))$. As in the previous section, we will use the terms *failure rate* and *mortality rate* interchangeably employing the first one mostly for a more general reasoning and the second one in a demographic context. Assume that the derivative $\lambda'(t)$ exists. Then, obviously, $F(t) \in IFR$, if $\lambda'(t) \geq 0, t \geq 0$. We can compare the ‘extent of aging’ described by different IFR distributions by the value of this derivative at each instant of time. However, this is not always the right thing to do, as intuitively, it is clear that at many instances in order to compare aging for different lifetimes some ‘relative reasoning’ should be also employed.

In life sciences (e.g., in demography), the *rate of aging* $R(t)$ is usually defined as

$$R(t) \equiv \frac{d \ln \lambda(t)}{dt} = \frac{\lambda'(t)}{\lambda(t)}. \quad (5.106)$$

This characteristic already describes the *relative* change in the failure (mortality) rate in an infinitesimally small unit interval of time. It takes into account the value of $\lambda(t)$, as intuition prompts that this measure should often depend not only on the derivative but on the value of the failure rate itself. Indeed, consider, for instance, two failure rates $\lambda(t)$ and $\lambda(t) + c$, where c is a constant. It is clear that the relative change for the second failure rate decreases as c increases and when c is large, the change in the failure rate can be negligible compared with the failure rate itself.

Thus, not only the change in the derivative is important, but also the level of the failure rate as well. Formal definition (5.106) is the simplest way to implement this relative concept. As most of simple definitions that are trying to describe complex properties, it has its *pros* and *contras* (e.g., De Gray [25] mostly focuses on the *contras*). However, this approach to defining the rate of aging is well justified in demography, as for the Gompertz law of mortality (5.86) that describes mortality rate at adult ages, it is a constant, i.e., $R(t) = b$. Thus, in practical demography, b is usually estimated as the slope of the Gompertz regression, i.e., the slope of $\ln \lambda(t)$. It should be understood, however, that $R(t)$ is just a useful (at least, for the Gompertz law) statistical measure, which describes in some ‘integrated way’ the real aging processes that are manifested by the changes in probabilities of failure (death) over time.

The foregoing considerations refer to the homogeneous populations, where obviously, b can be also regarded as the *individual rate of aging*. However, human populations are heterogeneous, and it is interesting to consider the rate of aging for this case. The general mixture model is described in Sect. 5.1 given by Eqs. (5.10)–(5.12). In what follows, we will focus on the specific multiplicative model (5.17). We will also need the following example:

Example 5.6 Let the frailty Z be a gamma-distributed random variable with shape parameter α and scale parameter β , whereas the baseline distribution be an arbitrary distribution with the failure rate $\lambda(t)$. It is well known [28] that (5.21) is generalized in this case to

$$\lambda_m(t) = \frac{\alpha\lambda(t)}{\beta + \Lambda(t)}, \quad (5.107)$$

where $\Lambda(t)$ is the cumulative failure rate $\Lambda(t) = \int_0^t \lambda(u)du$. Therefore,

$$E[Z|t] = \frac{\alpha}{\beta + \Lambda(t)}.$$

As $E[Z] = \alpha/\beta$ and $\text{Var}(Z) = \alpha/\beta^2$, Eq. (5.107) can now be written in terms of $E[Z]$ and $\text{Var}(Z) \equiv \sigma^2$ in the following way:

$$\lambda_m(t) = \lambda(t) \frac{E^2[Z]}{E[Z] + \sigma^2\Lambda(t)}, \quad (5.108)$$

which, for the specific case $E[Z] = 1$, gives the result of Vaupel et al. [63] that is widely used in demography:

$$\lambda_m(t) = \frac{\lambda(t)}{1 + \sigma^2\Lambda(t)}. \quad (5.109)$$

We will use Eq. (5.109) for analyzing the rate of aging as a function of parameters of the baseline and frailty distributions.

We start analyzing the rate of aging in heterogeneous populations with the specific gamma-Gompertz multiplicative model with the failure rate given by Eq. (5.21). Therefore,

$$\ln \lambda_m(t) = \ln a + bt - \ln [1 + (a\sigma^2/b)(\exp\{bt\} - 1)] \quad (5.110)$$

and the corresponding rate of aging is

$$R_m(t) = (\ln \lambda_m(t))' = b - \frac{a\sigma^2 \exp\{bt\}}{1 + (a\sigma^2/b)(\exp\{bt\} - 1)}. \quad (5.111)$$

Equation (5.111) states a simple and expected fact that the observed (population) rate of aging $R_m(t)$ is smaller than the individual rate of aging b . The latter, as was stated, corresponds to the homogeneous case. It can be also clearly seen that when σ^2 increases, $R_m(t)$ decreases. Therefore, the following hypothesis makes sense: *the increase in the rate of aging observed in the previous century in the developed countries could be attributed to the decreasing heterogeneity in mortality of populations in these countries.*

Another important feature that follows from (5.111) is that the increase in parameter a also results in the decrease in $R_m(t)$, which can be interpreted as some

kind of negative correlation between a of the Gompertz mortality law and the rate of aging.

In the case of arbitrary lifetimes, (5.109) results in

$$\begin{aligned}
 R_m(t) &= (\ln \lambda_m(t))' \\
 &= \frac{\lambda'(t)}{\lambda(t)} - \sigma^2 \frac{\lambda(t)}{1 + \sigma^2 \Lambda(t)} = R(t) - \sigma^2 \lambda_m(t)
 \end{aligned}
 \tag{5.112}$$

and, obviously, the rate of aging is also decreasing as a function of variance of the gamma-distributed frailty (for the fixed expectation $E[Z] = 1$). The similar conclusion was made in Yashin et al. [67].

Consider now a general case of the multiplicative model (5.17) not restricting ourselves to the gamma-distributed frailty. It can be shown [30] that

$$\begin{aligned}
 R_m(t) &= (\ln \lambda_m(t))' = \frac{\lambda'(t)E[Z|T > t] + \lambda(t)E'[Z|T > t]}{\lambda(t)E[Z|T > t]} \\
 &= \frac{\lambda'(t)}{\lambda(t)} + \frac{E'[Z|T > t]}{E[Z|T > t]} \\
 &= R(t) - \lambda(t) \frac{\text{Var}(Z|T > t)}{E[Z|T > t]}.
 \end{aligned}
 \tag{5.113}$$

Thus, as previously, the observed (mixture) rate of aging $R_m(t)$ is smaller than the individual rate of aging $R(t)$ defined for the baseline distribution with the failure rate $\lambda(t)$. A similar result using a different approach for derivations was independently recently obtained by Vaupel and Zhang [62]. As we are focusing on the specific multiplicative model (5.17), Eq. (5.113) is very helpful in analyzing a ‘proportional effect of environment’ on mortality rates.

Suppose now we have two heterogeneous populations with the *same baseline* $\lambda(t)$ and different frailties Z_1, Z_2 . In other words, compositions of populations are different. Let

$$\frac{\text{Var}(Z_2|T > t)}{E[Z_2|T > t]} \leq \frac{\text{Var}(Z_1|T > t)}{E[Z_1|T > t]}, \quad t > 0.
 \tag{5.114}$$

Then it is easy to see that the corresponding rates of aging are ordered as $R_{2m}(t) \geq R_{1m}(t)$. Thus, the rate of aging decreases as the *relative variance* increases, i.e.,

$$R_{2m}(t) - R_{1m}(t) = \lambda(t) \left[\frac{\text{Var}(Z_1|T > t)}{E[Z_1|T > t]} - \frac{\text{Var}(Z_2|T > t)}{E[Z_2|T > t]} \right] \geq 0, \quad \forall t \geq 0.$$

Inequality (5.114) defines a new class of stochastic ordering of random variables that can be called ordering in the sense of the relative variance [30]. The corresponding measure depends not only on the variance (variability), but on the mean as well.

References

1. Aalen OO, Gjessing HK (2001) Understanding the shape of the hazard rate: a process point of view. *Stat Sci* 16:11–22
2. Abdel-Hameed M (1975) A gamma wear process. *IEEE Trans Reliab* 24:152–153
3. Anderson JJ (2000) A vitality-based model relating stressors and environmental properties to Arjas E, Norros I (1989). Change of life distribution via a hazard transformation: an inequality with application to minimal repair. *Math Oper Res* 14:355–361
4. Aven T, Jensen U (1999) *Stochastic models in reliability*. Springer, New York
5. Aven T, Jensen U (2000) A general minimal repair model. *J Appl Probab* 37:187–197
6. Barlow R, Proschan F (1975). *Statistical theory of reliability and life testing*. Holt, Rennerhart & Winston, New York
7. Beard RE (1959) Note on some mathematical mortality models. In: Woolstenholme GEW, O'Connor M (eds) *The lifespan of animals*. Little, Brown and Company, Boston, pp 302–311
8. Beard RE (1971) Some aspects of theories of mortality, cause of death analysis, forecasting and stochastic processes. In: Brass W (ed) *Biological aspects of demography*. Taylor & Francis, London, pp 57–68
9. Bebbington M, Lai CD, Zitikis R (2008) A proof of the shape of the Birnbaum-Saunders hazard rate function. *Math Sci* 33:49–56
10. Bergman B (1985) *Reliability theory and its applications*. *Scand J Stat* 12:1–41
11. Bingham NH, Goldie CM, Teugels JL (1987) *Regular Variation*. Cambridge University Press, Cambridge
12. Birnbaum ZW, Saunders SC (1969) A new family of life distributions. *J Appl Probab* 6:319–327
13. Block HW, Joe H (1997) Tail behaviour of the failure rate functions of mixtures. *Lifetime Data Anal* 3:269–288
14. Block HW, Li Y, Savits TH (2003) Initial and final behavior of failure rate functions for mixtures and systems. *J Appl Probab* 40:721–740
15. Block HW, Li Y, Savits TH (2003) Preservation of properties under mixture. *Probab Eng Inf Sci* 17:205–212
16. Block HW, Li Y, Savits TH, Wang J (2008) Continuous mixtures with bathtub-shaped failure rates. *J Appl Probab* 45:260–270
17. Block HW, Mi J, Savits TH (1993) Burn-in and mixed populations. *J Appl Probab* 30:692–702
18. Block HW, Savits TH, Wondmagegnehu ET (2003) Mixtures of distributions with increasing linear failure rates. *J Appl Probab* 40:485–504
19. Boland PJ, El-Newehi E (1998) Statistical and information based minimal repair for k out of n systems. *J Appl Probab* 35:731–740
20. Cha JH, Finkelstein M (2011) Stochastic intensity for minimal repairs in heterogeneous populations. *J Appl Probab* 48:868–876
21. Cha JH, Finkelstein M (2012) Stochastic analysis of preventive maintenance in heterogeneous populations. *Oper Res Lett* 40:416–421
22. Chhikara RS, Folks JL (1977) The inverse Gaussian distribution as a lifetime model. *Technometrics* 19:461–468
23. Clarotti CA, Spizzichino F (1990) Bayes burn-in and decision procedures. *Probab Eng Inf Sci* 4:437–445
24. Cox DR, Miller HD (1965) *Theory of stochastic processes*. Methuen & Co., London
25. De Gray DNJ (2005) “The rate of aging”: a counterproductively undefinable term. *Rejuvenation Res* 8(2):77–78
26. Finkelstein M (1992) Some notes on two types of minimal repair. *Adv Appl Probab* 24:226–228
27. Finkelstein M (2004) Minimal repair in heterogeneous populations. *J Appl Probab* 41:281–286

28. Finkelstein M (2008) Failure rate modelling for reliability and risk. Springer, London
29. Finkelstein M (2009) Understanding the shape of the mixture failure rate (with engineering and demographic applications). *Appl Stoch Models Bus Ind* 25:643–663
30. Finkelstein M (2011) On the ‘rate of aging’ in heterogeneous populations. *Math Biosci* 22:20–23
31. Finkelstein M (2012) On ordered subpopulations and population mortality at advanced ages. *Theor Popul Biol* 81:292–299
32. Finkelstein M, Esaulova V (2001) Modelling a failure rate for the mixture of distribution functions. *Probab Eng Inf Sci* 15:383–400
33. Finkelstein M, Esaulova V (2006) Asymptotic behavior of a general class of mixture failure rates. *Adv Appl Probab* 38:244–262
34. Finkelstein M, Vaupel JW (2006) The relative tail of longevity and the mean remaining lifetime. *Demogr Res* 14(6):111–138
35. Gampe J (2010) Human mortality beyond age 110. In: Maier H, Gampe J, Jeune B, Robine J-M, Vaupel JW (eds) *Supercentenarians, demographic research monographs*. Springer, Berlin, pp 219–230
36. Gavrilo NA, Gavrilo NS (2001) The reliability theory of ageing and longevity. *J Theor Biol* 213:527–545
37. Glaser RE (1980) Bathtub and related failure rate characterizations. *J Am Stat Assoc* 75:667–672
38. Gupta RC, Gupta PL (1996) Ageing characteristics of the Weibull mixtures. *Probab Eng Inf Sci* 10:591–600
39. Gupta RC, Warren R (2001) Determination of change points of nonmonotonic failure rates. *Commun Stat Theory Methods* 30:1903–1920
40. Gurland J, Sethuraman J (1995) How pooling failure data may reverse increasing failure rate. *J Am Stat Assoc* 90:1416–1423
41. Hougaard P (2000) *Analysis of multivariate survival data*. Springer-Verlag, New York
42. Kijima M (1989) Some results for repairable systems with general repair. *J Appl Probab* 26:89–102
43. Lai CD, Xie M (2006) *Stochastic ageing and dependence for reliability*. Springer, New York
44. Li Y (2005) Asymptotic baseline of the hazard rate function of mixtures. *J Appl Probab* 42:892–901
45. Li T, Anderson JJ (2009) The vitality model: a way to understand population survival and demographic heterogeneity. *Theor Popul Biol* 76:118–131
46. Liao H, Elsayed EA, Chan LY (2006) Maintenance of continuously monitored degrading systems. *Eur J Oper Res* 175:821–835
47. Lynn NJ, Singpurwalla ND (1997) Comment: “burn-in” makes us feel good. *Stat Sci* 12:13–19
48. Marshall AW, Olkin I (2007) *Life distributions*. Springer, New York
49. Missov TI, Finkelstein M (2011) Admissible mixing distributions for a general class of mixture survival models with known asymptotics. *Theor Popul Biol* 80:64–70
50. Natvig B (1990) On information-based minimal repair and reduction in remaining system lifetime due to a failure of a specific module. *J Appl Probab* 27:365–375
51. Navarro J, Hernandez PJ (2004) How to obtain bathtub-shaped failure rate models from normal mixtures. *Probab Eng Inf Sci* 18:511–531
52. Navarro J, Hernandez PJ (2008) Mean residual life functions of finite mixtures, order statistics and coherent systems. *Metrica* 67:277–298
53. Park C, Padgett WJ (2005) Accelerated degradation models for failure based on geometric Brownian motion and gamma process. *Lifetime Data Anal* 11:511–527
54. Ross SM (1996) *Stochastic processes*, 2nd edn. Wiley, New York
55. Steinsaltz D, Evans S (2004) Markov mortality models: implications of quasistationarity and varying initial distributions. *Theor Popul Biol* 65:319–337
56. Steinsaltz D, Wachter KW (2006) Understanding mortality rate deceleration and heterogeneity. *Math Popul Stud* 13:19–37

57. Strehler L, Mildvan AS (1960) General theory of mortality and aging. *Science* 132:14–21
58. Thatcher RE (1999) The long-term pattern of adult mortality and the highest attained age. *J Roy Stat Soc* 162:5–43
59. Tsiatis A (1975) A nonidentifiability aspect of the problem of competing risks. *Proc Natl Acad Sci U S A* 72:20–22
60. Valdez-Flores C, Feldman RM (1989) A survey of preventive maintenance models for stochastically deteriorating single-unit systems. *Naval Res Logist* 36:419–446
61. Van Nortwijk JM (2009) A survey of the application of gamma processes in maintenance. *Reliab Eng Syst Saf* 94:2–21
62. Vaupel JW, Zhang Z (2010) Attrition in heterogeneous cohorts. *Demogr Res* 23:737–749
63. Vaupel JW, Manton KG, Stallard E (1979) The impact of heterogeneity in individual frailty on the dynamics of mortality. *Demography* 16:439–454
64. Weitz JS, Fraser HB (2001) Explaining mortality rates plateaus. *PNAS* 98:15383–15386
65. Wang H (2002) A survey of maintenance policies of deteriorating systems. *Eur J Oper Res* 139:469–489
66. Yashin AI, Manton KG (1997) Effects of unobserved and partially observed covariate processes on system failure: a review of models and estimation strategies. *Stat Sci* 12:20–34
67. Yashin AI, Ukraintseva SV, Boiko SI, Arbeev KG (2002) Individual aging and mortality rate: how are they related. *Soc Biol* 49:206–217