



Max-Planck-Institut für demografische Forschung  
Max Planck Institute for Demographic Research  
Konrad-Zuse-Strasse 1 · D-18057 Rostock · GERMANY  
Tel +49 (0) 3 81 20 81 - 0; Fax +49 (0) 3 81 20 81 - 202;  
<http://www.demogr.mpg.de>

---

MPIDR WORKING PAPER WP 2011-022  
DECEMBER 2011

## **On Ordered Subpopulations and Population Mortality at Advanced Age**

Maxim Finkelstein (FinkelM@ufs.ac.za)

---

This working paper has been approved for release by: James W. Vaupel (jwv@demogr.mpg.de),  
Head of the Laboratory of Survival and Longevity.

© Copyright is held by the authors.

Working papers of the Max Planck Institute for Demographic Research receive only limited review.  
Views or opinions expressed in working papers are attributable to the authors and do not necessarily  
reflect those of the Institute.

# On Ordered Subpopulations and Population Mortality at Advanced Ages

Maxim Finkelstein

Department of Mathematical Statistics,  
University of the Free State, South Africa.

e-mail : FinkelM@ufs.ac.za

and

Max Planck Institute for Demographic Research, Rostock, Germany

**Abstract.** We consider hazard (mortality) rates in populations consisting of ordered (in the defined stochastic sense) subpopulations. This setting can be interpreted via the fixed frailty models with one or more frailty parameters. The shape of the hazard rate is of the main interest in this paper. Specifically, the deceleration and leveling off in the hazard rates (mortality plateaus) are discussed and some examples of lifetime distributions that can result in asymptotically flat hazard rates are considered. These examples are based on vitality models when an organism's initial vitality (resource) is 'consumed' in the course of life in accordance with a simple stochastic process (e.g., Wiener process with drift or the gamma process).

**Key words:** Hazard rate, frailty, mortality plateau, Wiener process, gamma process, inverse-Gaussian distribution, Birnbaum-Saunders distribution.

## 1. Introduction

The shape of the hazard 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 supercentenarians have been recorded. The International Database of Longevity offers the detailed information on thoroughly validated cases of supercentenarians. Gampe (2010) 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 most popular explanation of this fact is by considering the corresponding fixed frailty models that account for heterogeneity of populations. Vaupel et al (1979) (see also Beard (1959, 1971)) 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 (see later) frailty model, which can be considered as another justification of the uniqueness and importance of this distribution for human mortality modeling. Finkelstein and Esaulova (2006) have proved that the gamma distribution of frailty is not so critical in this respect: all densities that behave as  $z^\alpha, \alpha > 1$  when  $z \rightarrow 0$  are equivalent in this sense (see also Steinzalts and Wachter (2006)). Missov and Finkelstein (2011) have extended this result to the case of regularly varying at  $z \rightarrow 0$  densities.

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 populations with 'properly' (see Section 2) ordered subpopulations is defined by the small values of frailty, as the subpopulations with larger values of frailty (and,

therefore, larger values of the hazard rate) are dying out first. Note that, apart from statistical explanations of mortality plateaus based on frailty models (and *stochastic vitality* models to be discussed in Section 4), there also exist biologically motivated approaches (see, e.g., Mueller and Rose (1996)).

The first question to be answered is what *common* statistical distributions are characterized by the asymptotically flat hazard 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 non-monotone and even can be non-positive, 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 lifespan of organisms (Anderson, 2000; Li and Anderson, 2009), 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 (2001) and Steinsaltz and Evans (2004) as an example highlighting the meaning and properties of the corresponding quasistationary distributions for this particular case. Our goal in this paper is more modest: to exploit further some relevant distributional properties in the context of stochastic ordering of lifetimes of subpopulations in heterogeneous populations.

The other example of a distribution with asymptotically flat hazard rate is the Birnbaum-Saunders distribution (Birnbaum and Saunders, 1969) 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 hazard rate in this case is decreasing to 0 as  $t \rightarrow \infty$ . It should be noted, however, that the initial increase in hazard rates for all of 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’ (30-100 years) are obviously limited.

The paper is organized as follows. In Section 2 relevant stochastic orderings of subpopulations in terms of lifetimes are described. We also discuss the notions of *fixed* and *evolving* heterogeneity (frailty) in the context of heterogeneous populations. Section 3 deals with mixtures as a convenient tool for analysis of heterogeneity in the univariate and bivariate settings. In Section 4, we consider the first passage time problems for the vitality modeling and analyze it further from the ‘classical’ *distributional point of view*, which can complement to a certain extent the corresponding analysis from the *process point of view* that can be found in the literature (Aalen and Gjessing, 2001; Steinsaltz and Evans, 2004; Li and Anderson, 2009). Specifically, we use the fact that randomization of some parameters of lifetime distributions can be interpreted in terms of the corresponding heterogeneous populations with (fixed) frailties when the suitable stochastic ordering of lifetimes with respect to these parameters holds. We consider the inverse Gaussian and the Birnbaum-Saunders distributions with asymptotically flat hazard rates and show that when parameters of these distributions are randomized, the population hazard rate is vanishing as  $t \rightarrow \infty$ . In Section 5, some concluding remarks are given.

Finally, we want to mention that our intention is not to perform a comprehensive review of relevant results in the literature (see, e.g., Li and Anderson (2009) and

Yashin et al (2000) for that), but rather to refer to the sources that are necessary and important for our presentation.

## 2. Fixed and evolving (changing) heterogeneity

Let  $F(t)$ ,  $f(t)$  and  $\mu(t)$  be the Cdf, the pdf and the hazard rate (force of mortality) for some infinite *homogeneous* population that characterize the corresponding random lifetime  $T \geq 0$ . However, one can hardly find homogeneous populations in real life and therefore, neglecting existing heterogeneity can lead to substantial errors and misconceptions in stochastic analysis in demography, survival and risk analysis as well as other disciplines. 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. A popular way to model heterogeneity is based on a notion of a non-negative random unobserved parameter (frailty)  $Z$ . The term “frailty” was suggested in Vaupel *et al.* (1979) for the gamma-distributed  $Z$  and the multiplicative hazard rate model of the form  $\mu(t, Z) = Z\mu(t)$ , which should be understood in terms of realizations  $Z = z$ . We see that the difference between subpopulations is modeled *directly* by the differences in mortality rates, e.g., for two realizations  $z_2 > z_1$ , this difference is  $(z_2 - z_1)\mu(t)$ . Thus, the multiplicative frailty model describes the ordering of subpopulations in the sense of the *hazard rate ordering*:

$$\mu(t, z_1) \leq \mu(t, z_2); \quad z_1 \leq z_2, t \in [0, \infty), \quad (1)$$

which obviously leads to the (weaker) ordering in the sense of the *usual stochastic ordering* (stochastic dominance):

$$F(t, z_1) \leq F(t, z_2); \quad z_1 \leq z_2, t \in [0, \infty), \quad (2)$$

as

$$F(t, z_i) = 1 - \exp\left\{-\int_0^t \mu(u, z_i) du\right\}.$$

Thus, the smaller is the value of  $z$ , the larger is the lifetime of the subpopulation  $T_z$  in the appropriate stochastic sense (e.g., (1) or (2)):

$$T_{z_1} \geq T_{z_2}; \quad z_1 \leq z_2. \quad (3)$$

Thus, 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 (mostly, but not necessarily the hazard rate ordering (1)).*

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 hazard rate is:

$$\mu(t, a, b) = ae^{bt} \quad (4)$$

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 (1). We just want to emphasize the fact that in this specific model, frailty acts multiplicatively and ‘directly’ on the hazard rate, which is not the case in general even when ordering (1) holds. Some relevant aspects of frailty modeling for the bivariate case will be considered in the next section.

Finkelstein and Esaulova (2006) have extended multiplicative frailty model to a rather general class of survival models that includes also additive ( $\mu(t, Z) = \mu(t) + Z$ ) and accelerated life (AL) models. The shape of the hazard rate for the AL model that is defined by the scale transformation  $F(Zt)$  was not considered in the literature before, however the ‘physical’ motivation that the degradation processes in objects are heterogeneous due to ‘fluctuations’ in the time scale of the baseline distribution  $F(t)$ , is quite appealing. The striking result of this paper states that the hazard rate in this case behaves as  $1/t(t \rightarrow \infty)$  regardless of the distribution of frailty  $Z$ . This specifically means that the AL frailty model can never end up in the mortality plateau! Note that, as  $\mu(t, Z) = Z\mu(Zt)$  in this case, ordering (1) does not necessarily hold, however, ordering (2) always holds.

In accordance with our definition, the fixed heterogeneity (frailty) is described only by ordered subpopulation lifetimes. The hazard rate, which is of the main interest in this paper, is obviously uniquely defined by the population lifetime distribution. 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, 2009) 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 in Section 4. 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 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 (2009) 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*. Another model that usually is referred to in the literature as an example of changing frailty (Yashin et al, 1976) is the Le Bras (Le Bras, 1976) model, where the corresponding continuous Markov chain describes the ‘process of moving to the absorbing state’ (see also Horvitz and Tulpukar (2008) for somehow related approach). Note that, all three mentioned above models result in the asymptotically constant hazard rate.

In this study, 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 of Section 3, 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 the ordered subpopulations. As follows from the title of the paper, this is not our case.

As was stated in the introduction, we will mainly focus on the qualitative analysis of the hazard rates and specifically on the possibility of mortality plateaus. Before considering specific models, we must formulate some general results on mixture modeling.

### 3. Mixtures for modeling heterogeneity

First, we describe our setting in a formal way. Consider a population of identically distributed items with lifetimes  $T_i, i=1,2,\dots$ . Each item is ‘affected’ by a non-observable univariate frailty parameter  $Z_i$  and the lifetimes  $T_i$  are conditionally independent given the values of parameters  $Z_i = z_i$ . Assume that these parameters are i.i.d with a common pdf  $\pi(z)$  and with support in  $[0, \infty)$ . (The general support  $[a, b), 0 \leq a < b \leq \infty$  can be considered as well). Then, obviously,  $T_i, i=1,2,\dots$  are also i.i.d.. For convenience, the sub index “ $i$ ” will be omitted and therefore, the lifetimes and frailties for all items will be denoted by  $T$  and  $Z$ , respectively. Thus, the lifetime  $T$  is characterized by the following *mixture* (population) Cdf and pdf

$$\begin{aligned} F(t) &= \int_0^{\infty} F(t, z)\pi(z)dz, \\ f(t) &= \int_0^{\infty} f(t, z)\pi(z)dz, \end{aligned} \quad (4)$$

respectively, where  $F(t, z) \equiv F(t|z) = \Pr[T \leq t | Z = z]$ ,  $f(t, z) = F'(t, z)$  are the corresponding conditional characteristics for realization  $Z = z$ , which can be interpreted as the Cdf and the pdf for the corresponding homogeneous subpopulation, respectively. Then the mixture (population) hazard rate  $\mu(t)$  is defined as (see, e.g., Finkelstein (2008))

$$\begin{aligned} \mu(t) &= \frac{f(t)}{F(t)} \\ &= \frac{\int_0^{\infty} f(t, z)\pi(z)dz}{\int_0^{\infty} \bar{F}(t, z)\pi(z)dz} = \int_0^{\infty} \mu(t, z)\pi(z|t)dz, \end{aligned} \quad (5)$$

where  $\pi(z|t)$  is the conditional pdf of  $Z$  (on condition that  $T > t$ )

$$\pi(z|t) \equiv \pi(z) \frac{\bar{F}(t, z)}{\int_0^{\infty} \bar{F}(t, z) \pi(z) dz}, \quad (6)$$

and  $\bar{F} \equiv 1 - F$ . Accordingly, denote by  $Z|t$  the conditional frailty at time  $t$  (on condition that  $T > t$ ). Another (Bayesian) interpretation of  $\pi(z|t)$  (for a single item) is the posterior pdf of frailty  $Z$ , which corresponds to the prior pdf  $\pi(z)$ .

The foregoing definitions describe the ‘standard’ statistical mixture (or fixed frailty) model for an item and for the collection of items (population) as well. As was discussed in the previous section, we understand heterogeneity as the property of a population that consists of ordered homogeneous subpopulations (ordered lifetimes  $T_z$ , as defined in (3)). But what type of ordering is sufficient for our reasoning? As we are looking at hazards rates, the first guess would be ordering (1). How in this context can we mathematically interpret the well-known and intuitively clear property: “*the weakest populations are dying out first*” and the corresponding mortality deceleration with time? To answer this question, denote 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  $\Pi(z|t)$  (Finkelstein, 2008). 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 1.** *Let ordering (1) hold. Then  $\Pi(z|t)$  is an increasing function of  $t$  for each fixed  $z$ .*

*Proof.* It follows from (6) that

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

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

$$\frac{\int_a^z \bar{F}'_t(t, u) \pi(u) du}{\int_a^z \bar{F}(t, u) \pi(u) du} > \frac{\int_a^b \bar{F}'_t(t, u) \pi(u) du}{\int_a^b \bar{F}(t, u) \pi(u) du}.$$

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

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

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

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

which obviously follows from ordering (1). ■

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

$$\Pr[T \leq t \mid Z_1 = z, Z_2 = z_2] \equiv \Pr[T \leq t \mid z_1, z_2] = F(t, z_1, z_2)$$

and

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

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 (Finkelstein, 2008):

$$\begin{aligned} \mu(t) &= \frac{f(t)}{F(t)} = \frac{\int_0^\infty \int_0^\infty f(t, z_1, z_2) \pi(z_1, z_2) dz_1 dz_2}{\int_0^\infty \int_0^\infty \bar{F}(t, z_1, z_2) \pi(z_1, z_2) dz_1 dz_2} \\ &= \int_0^\infty \int_0^\infty \mu(t, z_1, z_2) \pi(z_1, z_2 \mid t) dz_1 dz_2, \end{aligned} \quad (7)$$

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

$$\pi(z_1, z_2 \mid t) = \pi(z_1, z_2) \frac{\bar{F}(t, z_1, z_2)}{\int_0^\infty \int_0^\infty \bar{F}(t, z_1, z_2) \pi(z_1, z_2) dz_1 dz_2}. \quad (8)$$

Equation (7) 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 hazard rate is just the sum  $\mu(t) = \mu_1(t) + \mu_2(t)$  of the corresponding ‘univariate hazard rates’.

Although it is difficult to analyze  $\mu(t)$  in (7) in full generality, certain qualitative considerations that will be very helpful in the next section can be stated. Indeed, let us first fix the second frailty  $Z_2 = z_2$ . Then the corresponding hazard rate is defined by the univariate frailty model (5):

$$\mu(t, z_2) = \int_0^{\infty} \mu(t, z_1, z_2) \pi(z_1, z_2 | t) dz_1. \quad (9)$$

Thus, at the *first stage*, we’ve 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 hazard rate. As our goal is to analyze the hazard rate, at the *second stage*, we consider our overall population as a ‘continuous collection’ of *homogeneous subpopulations* with hazards rates given by (9). Then we can analyze  $\mu(t)$  again in the univariate manner. For instance, assume that the family  $\mu(t, z_2)$  is ordered in  $z_2$  (the smaller values of  $z_2$  correspond to the smaller values of  $\mu(t, z_2)$ ). Therefore, the deceleration in mortality due ‘to the weakest populations are dying out first’ takes place. Specifically, let  $\mu(t, z_2)$  for each  $z_2$  decreases (non-increases) at least asymptotically, as  $t \rightarrow \infty$ . It is well-known that the corresponding population (mixture) hazard rate is strictly decreasing in this case (see, e.g., Ross (1996)). Thus, we have described the following result:

**Theorem 2.** *Let frailty  $Z_1 = z_1$  ( $Z_2 = z_2$ ) in the bivariate frailty mode be fixed. Assume that the corresponding univariate frailty model (with respect to  $Z_2$  ( $Z_1$ )) results in decreasing (non-increasing) ordered hazard rates for all values of  $z_1$  ( $z_2$ ). Then ‘releasing (randomizing)’  $Z_1$  ( $Z_2$ ), results in the strictly decreasing population hazard 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 (keeping in mind our notation (e.g.,  $\pi(z | T > t) \equiv \pi(z | t)$ )) from equations (6)-(9). Theorem 2 then follows, as the (univariate) mixture of distributions with decreasing (non-increasing) hazard rates is characterized by the strictly decreasing hazard rate.

**Example 1.** An important conclusion that illustrates Theorem 2 deals with the Gompertz law of mortality (4). It is well known that randomization of  $a$  (e.g., via the gamma distribution of the frailty) results in the mortality plateau. 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’ obviously can be applied to the case when there are more than 2 frailties or parameters of distributions that can be randomized in a similar way. Note that, 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 hazard rate will also tend to this curve (vanishing fixed frailty), which will also be illustrated in the next section.

**Remark 1.** As we emphasized, heterogeneity of a population in *this paper* is understood as the *property that characterizes the differences in lifetimes of stochastically ordered homogeneous subpopulations*. A natural question arises: what is the measure for these differences? Obviously, the variance of frailty  $Z$  can be considered as such. More generally, consider stochastic comparison of two frailty random variables  $Z_1$  and  $Z_2$  via the *variability ordering*: we say that  $Z_1$  is more variable than  $Z_2$  and write  $Z_1 \geq_v Z_2$  if (Ross, 1996):

$$E[h(Z_2)] \geq E[h(Z_1)] \quad \text{for all increasing convex } h(\cdot). \quad (10)$$

Specifically, when  $E[Z_1] = E[Z_2]$  it can be shown that (10) holds for all convex functions  $h(\cdot)$ . For instance, as  $h(x) = x^2$ , we would have that  $\text{Var}(Z_2) \geq \text{Var}(Z_1)$ .

**Remark 2.** Variability dynamics of the conditional frailty  $Z|t$  (i.e., frailty among survivors at time  $t$ ) can be rather complex and is not fully investigated for the general case so far. However, it follows from Theorem 1 that  $E[Z|t]$  is decreasing with time, which is obviously also an indication of the “weakest populations are dying out first” process. It seems from the first sight that  $\text{Var}(Z|t)$  should also decrease, but surprisingly (at least, for us) this is not true for a general case. The counter-example can be found in Esaulova (2006), which shows that the conditional variance for some specific distribution ( $\pi(z) = 2z; z \in [0,1]$ ) of  $Z|t$  and for the additive frailty model is increasing in the neighbourhood of 0 regardless of the baseline distribution. It is also shown that  $\text{Var}(Z|t)$  is decreasing in  $[0, \infty)$  when  $Z$  is gamma-distributed.

The discussion of this section will help us to analyze the shape of the hazard rate for some examples of vitality models. We will focus mostly on the vitality model described by the Wiener process with drift (Anderson, 2000; Weitz and Frazer, 2001; Li and Anderson, 2009). The parameters of the corresponding lifetime distribution after randomization will act as fixed frailties that define the corresponding ordered subpopulations. We hope that this interpretation will add some simple useful additional reasoning from the *distributional point of view* to the *process point of view* approach developed by by Aalen and Gjessing (2001) and Steinsaltz and Evans (2004).

## 4. Vitality models and lifetime distributions

### a. 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:

$$V_t = v_0 - Rt, \quad (11)$$

where  $R$  is a positive random variable with the Cdf  $S(t)$ . For each realization  $R = r$ , (11) can model the linear decline in physiological functions of organisms noted by Strehler and Mildvan (1960) 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  $F_R(t) = \Pr[T_R \leq t] = \Pr[R \geq v_0/t] = 1 - S(v_0/t)$ . 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 an inverse gamma distribution form:

$$f_R(t) = \frac{(v_0 a)^\eta}{\Gamma(\eta)} t^{-\eta-1} e^{-v_0 a/t}. \quad (12)$$

We will analyze the shape of the corresponding hazard rate using the ‘classic’ Glazer’s theorem (Glazer, 1980), formulated in a slightly more general form by Marshall and Olkin (2006). 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 hazard rate  $\mu(t)$  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 rationale behind this statement becomes apparent when  $\lim_{t \rightarrow \infty} f(t) = 0$ . Indeed, by using L’Hopital’s rule:  $\lim_{t \rightarrow \infty} \mu(t) = \lim_{t \rightarrow \infty} f(t) / \bar{F}(t) = \lim_{t \rightarrow \infty} -f'(t) / f(t)$ .

**Theorem 3** (Marshall and Olkin, 2007). *Let the density  $f(t)$  of a lifetime random variable be strictly positive and differentiable on  $(0, \infty)$ , such that  $\lim_{t \rightarrow \infty} f(t) = 0$ . Then*

- (i) *If  $g(t)$  is increasing, then the hazard rate  $\mu(t)$  is also increasing.*
- (ii) *If  $g(t)$  is decreasing, then  $\mu(t)$  is also decreasing.*
- (iii) *If there exists  $t_1$  for which  $g(t)$  is decreasing in  $t \leq t_1$  and increasing in  $t \geq t_1$ , then there exists  $t_2: 0 \leq t_2 \leq t_1$ , such that  $\mu(t)$  is decreasing in  $t \leq t_2$  and increasing in  $t \geq t_2$ .*
- (iii) *If there exists  $t_1$  for which  $g(t)$  is increasing in  $t \leq t_1$  and decreasing in  $t \geq t_1$ , then there exists  $t_2: 0 \leq t_2 \leq t_1$ , such that  $\mu(t)$  is increasing in  $t \leq t_2$  and decreasing in  $t \geq t_2$ .*

The hazard rate  $\mu_R(t)$  that corresponds to (12) can be easily analyzed with the help of this theorem. Indeed, as  $\lim_{t \rightarrow 0} f_R(t) = 0$ , it follows that  $\lim_{t \rightarrow 0} \mu_R(t) = 0$ , whereas

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

and  $\mu_R(t)$  is “bell-shaped” with a maximum at  $t < v_0 a / (\eta + 1)$ .

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 (11), we can identify the corresponding subpopulations for each value of  $R = r$  that will be definitely ordered in the sense (3) (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 hazard rate. If we are interested in explaining mortality plateaus that has been observed in human and other populations, then we must look at more realistic vitality models. The first candidate for that is when the simplest stochastic process in (11)  $Rt$  is substituted by the more advanced stochastic model given by the Wiener process with drift.

#### 4b. Wiener process with drift.

We modify the degradation model (11) with the fixed initial vitality  $v_0$  to

$$\begin{aligned} V_t &= v_0 - R_t, \\ R_t &= rt + W_t, \end{aligned} \quad (13)$$

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., Cox and Miller, 1965) 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:

$$f_R(t) \equiv f_R(t; v_0, r, \sigma) = \frac{v_0}{\sigma \sqrt{2\pi}} t^{-3/2} \exp\left\{-\frac{(v_0 - rt)^2}{2\sigma^2 t}\right\}. \quad (14)$$

The exact expression for the corresponding hazard rate,  $\mu_R(t) \equiv \mu_R(t; v_0, r, \sigma)$ , is complicated and therefore, as our goal is just to analyze its *shape*, we will use Theorem 3. It is easy to derive from (14) that

$$g_R(t) = -\frac{d \log f_R(t)}{dt} = \frac{3}{2t} + \frac{r^2}{2\sigma^2} - \frac{v_0^2}{2\sigma^2 t^2}. \quad (15)$$

Note that, (14) is written in parameterization  $v_0, r, \sigma$ . 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):

$$f_R(t; \lambda, \omega) = \frac{\lambda \omega}{\sigma \sqrt{2\pi}} (\lambda t)^{-3/2} \exp\left\{-\frac{(\omega - \lambda t)^2}{2\lambda t}\right\}. \quad (16)$$

It immediately follows from (15) that the hazard rate tends to a constant when  $t \rightarrow \infty$  (mortality plateau):

$$\lim_{t \rightarrow \infty} \mu_R(t) = \lim_{t \rightarrow \infty} -\frac{d \log f_R(t)}{dt} = \frac{\lambda}{2} = \frac{r^2}{2\sigma^2} \quad (17)$$

It is also obvious that  $\lim_{t \rightarrow 0} \mu_R(t) = 0$ . The ‘rest of the shape’ of  $\mu_R(t)$  is defined by Theorem 3:  $\mu_R(t)$  is increasing for  $t \in [0, t_2]$ , 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 (1977) using straightforward calculus and asymptotic bounds. We, however, rely on a general Theorem 3 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, unlike Subsection 4a, we cannot identify 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. Moreover, as the sample paths of the Wiener process are non-monotonic, this process is often criticized in the engineering literature as a stochastic model for monotonic deterioration (see the next subsection for the Gamma process that is monotonic). Note that, the process point of view with evolving heterogeneity was briefly discussed in Section 2.

From (16) it follows that  $\lambda$  is the scale parameter. Therefore, obviously, the corresponding lifetimes are decreasing in  $\lambda$  in the sense of the usual stochastic ordering (2), i.e., for the fixed  $\omega$ :

$$F_R(\lambda_1 t; \omega) \leq F_R(\lambda_2 t; \omega); \quad \lambda_1 \leq \lambda_2, t \in [0, \infty). \quad (18)$$

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 (1) also takes place (Marshall and Olkin, 2007), which means:

$$\mu_R(t; \lambda_1, \omega) = \lambda_1 \mu_R(\lambda_1 t; \omega_1) \leq \lambda_2 \mu_R(\lambda_2 t; \omega_2) = \mu_R(t; \lambda_2, \omega); \quad \lambda_1 \leq \lambda_2, t \in [0, \infty).$$

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  $\sigma$  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  $\sigma$ . This is a rather unexpected observation, however as stated, it is a consequence of the considered specific setting.

After discussing the issue of stochastic ordering, we can now *qualitatively* analyze the shape of  $\mu_R(t; \lambda, \omega)$  for large  $t$  with respect to the randomized parameters  $r$  and  $\sigma$  ( $\nu_0$  is fixed so far) to be denoted by  $R$  and  $\Sigma$ , respectively. Note that, Aalen and Gjessing (2001), have performed the necessary derivations assuming that  $R$  is normally distributed and  $\sigma$  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  $\Sigma$  are non-negative random variables with supports in  $[0, \infty)$ . Thus, the bivariate frailty model discussed in Section 3 can be applied. We proceed as described there: fixing  $\Sigma = \sigma$  and considering subpopulations with one frailty parameter  $R$  (see equation (9) and the description after it). 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 hazard 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 hazard rates described above (increasing and then decreasing to a plateau), this heterogeneous subpopulation has asymptotically decreasing to 0 hazard rate (Ross, 1996). Now, at the second stage, as these hazard rates are ordered with respect to the values of the second frailty  $\Sigma = \sigma$ , we can use Theorem 2, which means that the population hazard 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  $\Sigma$  are modified to  $[a, \infty]$  and  $[0, b]$ , respectively. Then the population hazard rate tends to the failure rate of the strongest subpopulation which is, in accordance with (17) (Finkelstein, 2009),

$$\lim_{t \rightarrow \infty} \mu_R(t) = \frac{a^2}{2b^2}. \quad (19)$$

Alternatively, we can define the ordered subpopulations for the specific case under consideration in the ‘univariate way’ via parameter  $r/\sigma$ . Note that, for our qualitative considerations, we do not need the distributions of neither  $R, \Sigma$ , nor  $R/\Sigma$ . Thus, the points in a positive  $r, \sigma$  quadrant are ordered in accordance with the coefficient  $k$ , i.e.,  $r = k\sigma, k > 0$ . The shape of the hazard rate for each fixed  $k$  has the shape of the hazard rate of the inverse Gaussian distribution. Therefore, the population hazard rate is decreasing as  $t \rightarrow \infty$ , as it should be, because the result should be the same. Similar, when the supports of frailties  $R$  and  $\Sigma$  are modified to  $[a, \infty]$  and  $[0, b]$ , respectively, then (19) holds. Aalen and Geising (2001) (see also Whitmore (1986)) have shown explicitly that for the specific case of normally distributed  $R$  and fixed  $\sigma$ , the population hazard rate decreases to 0 as  $1/t^2$ . This perfectly conforms with our qualitative results for a general setting.

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 (17) that, distinct from 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 hazard rate. This was analytically shown and discussed using the concept of quasisationary distributions in (Aalen and G, 2000, Steinsaltz and Evans (2004), Li and Anderson, 2009). There is not much more to say about that in our analysis except

the fact that the subpopulations that correspond to different values of vitality (when other parameters are fixed) are also ordered in the sense of the hazard rate ordering: the larger values of vitality correspond to the smaller values of the hazard rate. As  $\omega = rv_0/\sigma^2$ , we can extend the bivariate frailty procedure of the previous section to the case of three frailties, but as stated, adding frailty  $V_0$  will not change the *asymptotic behavior* of the population hazard rate.

### c. Gamma process and the Birnbaum-Saunders distribution.

The Wiener process is often criticized as a model for degradation and ageing 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  $t$ :

$$f_{R_t}(x) = Ga(x | r^2t/\sigma^2, r/\sigma^2), \quad \mu, \sigma > 0, \quad (20)$$

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

where  $Ga(x | \alpha, \beta)$  denotes the gamma distribution with shape parameter  $\alpha$  and scale parameter  $\beta$ . 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

$$\begin{aligned} F_{R_t}(t) &= \Pr[T_R \leq t] = \Pr[R_t \geq v_0] \\ &= \int_{v_0}^{\infty} f_{R_t}(x) dx = \frac{\Gamma(r^2t/\sigma^2, v_0r/\sigma^2)}{\Gamma(r^2t/\sigma^2)}, \end{aligned} \quad (21)$$

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 (Noortwijk, 2009). It is shown by Liao et al (2006) that the corresponding hazard rate is increasing, whereas Abdel-Hameed (1975) 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 (2005) have derived a very complex exact expression for the pdf  $f_R(t)$ . Therefore, a simpler meaningful approximation for (21) 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

$$F_{BS}(t; \lambda, \alpha) = \Phi(\alpha^{-1}h(\lambda t)), \quad t > 0, \quad (22)$$

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 (21), the corresponding approximation reads (Noortwijk, 2009):

$$F_{R_t}(t) \approx \Phi\left(\sqrt{\frac{v_0 r}{\sigma^2}} \left[ \sqrt{\frac{rt}{v_0}} - \sqrt{\frac{v_0}{rt}} \right]\right). \quad (23)$$

It was obtained by Park and Padget (2005) 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 hazard rate functions are also approximated. Therefore, given our interest in asymptotic behavior of hazard rates, why not to start directly from distribution (23) 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 (13) 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 (Marshall and Olkin, 2006) one can obtain the lifetime distribution (22), where

$$\alpha = \sigma / \sqrt{\mu v_0}, \quad \lambda = \mu / v_0. \quad (24)$$

Differentiation of (22) results in the following density

$$f_{BS}(t; \lambda, \alpha) = \frac{\lambda}{2\alpha\sqrt{2\pi}} \left[ \frac{1}{\sqrt{\lambda t}} \left( 1 + \frac{1}{\lambda t} \right) \right] \exp \left\{ -\frac{1}{2\alpha^2} \left( \lambda t - 2 + \frac{1}{\lambda t} \right) \right\}. \quad (25)$$

Obviously,  $\lim_{t \rightarrow 0} \mu_{BS}(t; \lambda, \alpha) = 0$ . Using Theorem 3, it can be shown that the hazard rate is bell-shaped and is decreasing to a constant as  $t \rightarrow \infty$  (mortality plateau):

$$\begin{aligned} \lim_{t \rightarrow \infty} \mu_{BS}(t; \lambda, \alpha) &= \lim_{t \rightarrow \infty} -\frac{d \log f_{BS}(t; \lambda, \alpha)}{dt} \\ &= \frac{\lambda}{2\alpha^2} = \frac{\mu^2}{2\sigma^2} \end{aligned} \quad (26)$$

It follows from (26) 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 (25) 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 larger (smaller) values of the hazard rate in  $[0, \infty)$ . Alternatively, as previously, we can order with respect to the quotient  $\mu/v_0 = k$ . However, when  $t \rightarrow \infty$ , the ordering is only in accordance with the values of  $\mu$ .

The possibility of ordering with respect to the values of  $\sigma$  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 (26), this ordering exists *asymptotically*. Assume now that  $\mu$  is a realization of a random variable  $M$ , whereas  $\sigma$  is a realization of a random variable  $\Sigma$  with supports in  $[0, \infty]$ . Then, similar to the case of the inverse Gaussian distribution, the randomization results in the asymptotically decreasing to 0

population hazard rate. Mortality plateaus are theoretically possible in this model only when the supports of the frailties  $M$  and  $\Sigma$  are  $[a, \infty]$  and  $[0, b]$ , respectively.

## 5. Concluding Remarks

We define the fixed heterogeneity as a property of populations that consist of subpopulations with stochastically ordered subpopulations, as given by equation (3). As far as we know, this formal definition of heterogeneity was not considered in the literature so far. As we are discussing the shapes of hazard rates (with an emphasis on asymptotic behavior and mortality plateaus), it is natural to apply here the specific hazard rate ordering (1). When it holds, we prove that the conditional frailty distribution  $\Pi(z|t)$  tends to be more concentrated around small values of the frailty  $Z \geq 0$  as time increases, which corresponds to stronger populations and can be interpreted as “the weakest subpopulations are dying out first”.

Our Theorem 2 justifies the two-stage procedure in considering heterogeneity that corresponds to the bivariate frailty models. Under certain assumptions, it means that at the first stage, one can ‘create’ a heterogeneous subpopulation of the entire population that correspond to the univariate frailty (with the fixed second frailty) and then at the second stage, consider the collection of these subpopulations. Specifically, when the hazard rate at the first stage is asymptotically flat (mortality plateau), the population hazard rate is decreasing as  $t \rightarrow \infty$  (no mortality plateau). 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  $b$ !

As examples for ordered subpopulations, we consider several vitality models with the main focus on the Wiener process with drift that describes the decrease of vitality of organisms. This model was intensively studied in the literature, but we interpret the results on the shape of the hazard rate via the randomization of the parameters of the inverse Gaussian distribution and the corresponding ordering of subpopulations. We also show that the initial vitality has no effect on asymptotic behavior of the hazard rate and that only a non-randomized version of this distribution can end up in the asymptotically flat rate. We also briefly discuss the notion of evolving heterogeneity as *variability in sample paths of the underlying process of deterioration*, which obviously differs from our understanding of the fixed heterogeneity as *variability of lifetimes in homogeneous subpopulations*.

It should be noted that our discussion on asymptotically flat (or decreasing) hazard rates applies usually to really advanced ages of organisms, as, e.g., in the case of human populations, when the deceleration in mortality rates start to be obvious only after the age of 100. However, as follows, e.g., from Carey et al (1992), the decrease in mortality rates can happen relatively earlier for fruitflies. In Finkelstein (2006), it has been reported on the experiment with a population (heterogeneous) of only 700 miniature light bulbs with recorded times of failures. It is known that the lifetime of these objects for a homogeneous case is described by the Weibull distribution with the increasing hazard rate. However, the experiment showed the bell-shaped hazard rate with ‘observed’ maximum, which totally conforms with the fixed frailty model for this case.

## References

Aalen, O.O., Gjessing H.K. (2001). Understanding the shape of the hazard rate: A process point of view. *Statistical Science*, 16, 11-22.

- Abdel-Hameed, M. (1975). A gamma wear process. *IEEE Transactions on Reliability*, 24, 152-153.
- Anderson, J.J. (2000). A vitality-based model relating stressors and environmental properties to organism survival. *Ecol. Monograph* 70, 445-470.
- Beard, R.E. (1959). Note on some mathematical mortality models. In: Woolstenholme GEW, O'Connor M (eds.) *The Lifespan of Animals*. Little, Brown and Company, Boston, 302-311.
- Beard, R.E. (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, 57-68.
- Birnbaum, Z.W., Saunders, S.C. (1969). A new family of life distributions. *Journal of Applied Probability*, 6, 319-327.
- Chhikara, R.S., Folks, J.L. (1977). The inverse Gaussian distribution as a lifetime model. *Technometrics*, 19, 461-468.
- Cox, D.R., Miller, H.D. (1965). *Theory of Stochastic processes*. Methuen & Co., London.
- Gampe, J. (2010). Human mortality beyond age 110. In: H. Maier, J. Gampe, B. Jeune, J.-M. Robine, and J.W. Vaupel, editors. *Supercentenarians, Demographic Research Monographs*, Springer, 219-230.
- Esaulova, V. (2006). Failure rates modeling for heterogeneous populations. PhD Dissertation, University of Magdeburg.
- Finkelstein, M. (2005). On some reliability approaches to human aging. *The International Journal of Reliability, Quality and Safety Engineering*, 12, 337-346.
- Finkelstein, M, Esaulova, V. (2006). Asymptotic behaviour of a general class of mixture failure rates. *Advances in Applied Probability*, 38, 244-262.
- Finkelstein, M. (2008). *Failure Rate Modeling for Reliability and Risk*. Springer, London.
- Finkelstein, M. (2009). Understanding the Shape of the Mixture Failure Rate (with Engineering and Demographic Applications). Discussion Paper. *Applied Stochastic Models in Business and Industry*, 25, 643-663.
- Glazer, R.E. (1980). Bathtub and related failure rate characterizations. *Journal of the American Statistical Association*, 76, 667-672.
- Horvitz, C.C., Tuljapulkar, S. (2008). Stage dynamics, period survival and mortality plateaus. *Amer. Nat.* 172, 203-215.
- Le Bras, H. (1976). Lois de mortalite et age limite. *Population*, 31, 655-692.
- Liao, H., Elsayed, E.A., Chan L.Y. (2006). Maintenance of continuously monitored degrading systems. *European Journal of Operations Research*, 175, 821-835.
- Li, T., Anderson, J.J. (2009). The vitality model: a way to understand population survival and demographic heterogeneity.
- Marshall, A.W., Olkin, I. (2007). *Life Distributions*. Springer.
- Missov, T., Finkelstein, M. (2011). Admissible frailty distributions for a general class of mixture mortality models with known asymptotics. *Theoretical Population Biology* (in press).
- Muller, F.L., Rose, M.R. (1996). Evolutionary theory predicts late-life mortality plateaus. *PNAS*, 93, 15249-15253.
- Park, C., Padgett, W.J. (2005). Accelerated degradation models for failure based on geometric Brownian motion and gamma process. *Lifetime Data Analysis*, 11, 511-527.
- Ross, S. (1996). *Stochastic Processes*. John Wiley, New York.

- Steinsaltz, D, Evans, S. (2004). Markov mortality models: implications of quasistationarity and varying initial distributions. *Theoretical Population Biology*, 65, 319-337.
- Steinsaltz, D., Wachter, K.W. (2006). Understanding mortality rate deceleration and heterogeneity. *Mathematical Population Studies*, 13, 19-37.
- Strehler, L., Mildvan, A.S. (1960). General theory of mortality and aging. *Science*, 132, 14-21.
- Thatcher, R.E. (1999). The long-term pattern of adult mortality and the highest attained age. *Journal of the Royal Statistical Society*, 162, 5-43.
- Van Nortwijk, J.M. (2009). A survey of application of gamma processes in maintenance. *Reliability Engineering and System safety*, 94, 2-21.
- Vaupel, J.W., Manton, K.G, Stallard, E. (1979). The impact of heterogeneity in individual frailty on the dynamics of mortality. *Demography*, 16, 439-454.
- Weitz, J.S., Fraiser, H.B. (2001). Explaining mortality rates plateaus. *PNAS*, 98, 15383-15386.
- Whitmore, C.A. (1986). First-passage-time models for duration data: regression structures and competing risks. *The Statistician.*, 35, 207-219.
- Yashin, A.I., Iachin, I.A., Begun, A.S. (2000). Mortality modeling: a review. *Mathematical Population Studies*, 8, 305-332.