This is a peer-reviewed, accepted author manuscript of the following research article: Finkelstein, M & Cha, JH 2025, 'Heterogeneity in populations and the paradoxes of survival: a tribute to Nozer Singpurwalla', Applied Stochastic Models in Business and Industry, vol. 41, no. 1. https://doi.org/10.1002/asmb.2919

Heterogeneity in populations and the paradoxes of survival: A tribute to Nozer Singpurwalla

Maxim Finkelstein[a,b] and Ji Hwan Cha[1,c]

[a] Department of Mathematical Statistics, University of the Free State, Bloemfontein, South Africa
E-mail: FinkelM@ufs.ac.za

[b]Department of Management Science, University of Strathclyde, Glasgow, UK
[c]Department of Statistics, Ewha Womans University, Seoul, 120-750, Rep. of Korea.
E-mail: jhcha@ewha.ac.kr

Abstract. We consider several survival models in heterogeneous settings. Heterogeneity in the failure rates of subpopulations results (as a specific case) in the famous failure rate paradox when the failure rate of a mixture of items with constant failure rates is decreasing. Random failure rate that is due to a point process that increases it at random times on fixed values also results in the 'bending down' of the population failure rate. Similar effect is observed while analysing the extreme shock models with shock processes that possess memory. Finally, another paradox when, due to heterogeneity in a vital parameter of a model, a terminating point process with decreasing rate after 'mixing' becomes a non-terminating one with increasing rate is described. Those are the impacts of heterogeneity that are discussed from the unified perspective that employs the 'principle': the weaker subpopulations are dying out first.

Keywords. Heterogeneous populations, frailty, extreme shock model, self-exciting point processes, self-regulating point processes

1. Introduction

In October 2001, Professor Nozer Singpurwalla visited the first author at the University of Free State, South Africa. In numerous conversations and discussions, Nozer was repeatedly coming back to the issue of the Bayesian interpretation of the failure rate that he called the subjective or the "predictive failure rate" along with the objective failure rate that was mostly referred by him as the "model failure rate" (Singpurwalla (2011), Singpurwalla and Wilson (1999)). He was always more interested in foundational, aspects of reliability theory rather than in straightforward applications. Needless to say, the work and personality of Prof. Singpurwalla had an impact on development of reliability theory in the last decades. These discussions eventually inspired the interest of the first author in the topic of mixture failure rates that can describe heterogeneous populations and are, in fact, 'predictive' in Nozer's terminology. However, obviously, they can be studied without a Bayesian flavour as well. As a result, a number of relevant publications in this area emerged that were co-authored with the second author as well.

The current paper, in a way, discusses and interprets some of our developments of the topic through the years from the unified viewpoint, providing some new results and insights as well.

---

[1] Corresponding Author

The general approaches considered (except the last recent part) were discussed in this or other way with Nozer at various conferences or personal communications.

This paper describes some effects of heterogeneity of populations on survival characteristics and, specifically and most importantly, on the corresponding failure rates. Due to instability of production processes, environmental and other factors, most populations of manufactured items in real life are heterogeneous (Finkelstein (2008)). Human populations are also heterogeneous within one country and among the countries or large areas as well.

As far as we know, the first model for the failure rate in heterogeneous populations in a reliability-based context was described in Barlow and Proschan (1975), whereas a Bayesian explanation of the corresponding paradox was given in Barlow (1985). A population of items with exponentially distributed lifetimes was considered. The uncertainty in the parameter of exponential distribution was modelled by a random prior. Although each item in a population was described by a constant failure rate, the predictive failure rate was decreasing in time. It can be considered as intuitively paradoxical, whereas the formal explanation of this fact was provided in these references: as the failure rate is a conditional characteristic, the prior distribution is being updated for the survived items. Later, Finkelstein (2008, 2009) was using for explanation the following principle: "*The weaker subpopulations are dying out first*". Intuitively, it is clear that the items from a subpopulation with a larger failure rate fail stochastically earlier, which results in the smaller failure rate for a population. However, for exhibiting this effect, the subpopulations (i.e., the items that are described by the same failure rate, or, equivalently, distribution function) in a population should be ordered appropriately, in order the terms "weaker" or "stronger" to have a proper stochastic sense.

Heterogeneity of populations is closely related to the term "frailty" intensively used in statistical literature for over 40 years (see, e.g., Cha and Finkelstein (2014) on mixture models with frailties in reliability context). Frailty is an unobserved random variable that can characterize uncertainty in a distribution or model parameters. This term was first introduced in a seminal paper by Vaupel at al. (1979) for a gamma-distributed frailty. See also Vaupel and Yashin (1985) for other counter-intuitive effects of heterogeneity. It is worth noting, however, that this specific case of the gamma-frailty model was, in fact, first introduced by the British actuary Robert Beard (1959) far before the corresponding reliability studies. The gamma mixing distribution was considered in this paper for the underlying Gompertz distribution with the exponentially increasing failure rate. The effect of heterogeneity was in 'slowing down' the increase of the mixture failure rate that eventually was tending to a constant (a plateau of human mortality).

The shape of the resulting failure is closely related to aging properties of items as, for instance, the increasing failure rate defines the simplest but most important in various applications class of IFR lifetime distributions. It can be versatile depending mostly on the shape of the underlying/baseline failure rate. The studies show that the mixing (prior) distribution has a less prominent effect in this respect. The decreasing pattern for the constant underlying failure rate was described above as well as the plateau of human mortality. Another popular example (see the next section) is the Weibull distribution with increasing failure rate. The mixture failure rate in this case is initially increasing and then decreasing exhibiting the corresponding bell shape. Throughout this paper, we call such initially increasing and then decreasing function a 'bell-shaped' function in a wider sense whether its limit is given by 0 or not. The bathtub shape for a specific additive mixing model with increasing underlying failure rate was described in Lynn and Singpurwalla (1997). Other patterns can also exist exhibiting the well-pronounced effect of populations heterogeneity on its statistical characteristics.

Heterogeneity described by different failure rates of subpopulations shows, in a way, an aggregated pattern, as the failure rate itself is an unobserved 'aggregated' characteristic. It is usually described by a single frailty (can be multivariate as well). However, in the described sense, it is a fixed frailty assigned at t=0. On the other hand, the effect of e.g., environment on the baseline failure rate of an item can be in the form of a stochastic process and not a random variable, thus defining time-dependent or evolving heterogeneity (see Li and Andersen (2009), Finkelstein (2012) and Aalen et al. (2008)). Thus, the failure rate can become the full-fledged stochastic process and not just a path process as in the case of the described fixed frailty. The effects of heterogeneity of this kind for lifetime models can be also explained intuitively in a similar manner to that used for the fixed frailty modelling.

The paper is organized as follows. In Section 2 some well-known mixture failure rate models are briefly introduced and some basic notions are discussed. Section 3 describes the hazard rate process induced by external shocks and the corresponding effect of this heterogeneity. Section 4 deals with the survival model based on the generalized Polya process. Section 5 shows how heterogeneity in parameter of the extended generalized Polya process results in the generalized Polya process. Concluding remarks are given in Section 6.

2. 'Traditional' mixing with fixed frailty

Consider a homogeneous population of i.i.d. items with lifetimes T described by the absolutely continuous cdf $F(t)$, pdf $f(t)$ and failure rate, $\lambda(t)$. Denote also the corresponding survival function by $\bar{F}(t) = 1 - F(t)$. However, due to instability of production processes, environmental and other factors, most populations of manufactured items in real life are heterogeneous. Let the lifetime of an item chosen at random from this heterogenous population be T_h .

The simplest and the most popular way to model this heterogeneity in applications is via the unobserved random variable that is often called "frailty". Consider for definiteness, a continuous frailty $Z \geq 0$ with support in $[0, \infty)$ and the pdf $\pi(z)$. The baseline cdf, pdf and the failure rate in this case are indexed by parameter Z, i.e., $F(t,z) = P(T_h \leq t | Z = z)$, $\bar{F}(t,z) = 1 - F(t,z), f(t,z) = \frac{\partial}{\partial t} F(t,z)$ and $\lambda(t,z) = \frac{f(t,z)}{\bar{F}(t,z)}$ accordingly, where $Z = z$ is the corresponding realization. (Note that the equivalent notations $F(t|z), f(t|z)$ and $\lambda(t|z)$ can be also used). On the other hand, the population (mixture) characteristics are defined, obviously, as the following expectations

$$F_m(t) = \int_0^\infty F(t,z)\pi(z)dz,$$

$$f_m(t) = \int_0^\infty f(t,z)\pi(z)dz.$$

As the failure rate is a conditional characteristic, the population (observed, or mixture) failure rate should be defined also as the conditional expectation (Lynn and Singpurwalla (1997)):

$$\lambda_m(t) = \int_0^\infty \lambda(t,z)\pi(z|t)dz, \tag{1}$$

$$\pi(z|t) = \frac{\pi(z)\bar{F}(t,z)}{\int_0^\infty \bar{F}(t,z)\pi(z)dz}, \tag{2}$$

where $\pi(z|t)$ denotes the conditional pdf of Z on condition that $T_h > t$. On the other hand, an unconditional expectation or the 'model/baseline mixture failure rate' is defined as

$$\lambda_b(t) = \int_0^\infty \lambda(t,z)\pi(z)dz. \tag{3}$$

It is also important to note that the above setting describes the random failure rate, which is the simplest stochastic process, namely, the path process λ_{bt}

$$\lambda_{bt} = \lambda(t,Z), t \geq 0. \tag{4}$$

This relation can define, e.g., an impact of a random environment on some baseline failure rate The more general form of this failure rate process to be called *the model failure rate process* (MFRP) will be considered in the next sections. It describes the failure rates of the corresponding subpopulations in a heterogeneous population. However, if we want to describe the failure rate of an item chosen at random from this heterogeneous population, conditioning on $T_h > t$ defines the corresponding *hazard rate process* (HRP)

$$\lambda_t \equiv \lambda_{bt}|T_h > t, t \geq 0$$

and, in what follows, we will use this important terminology for our paper.

To arrive at certain important properties for the shape of the population (mixture) failure rate, additional assumptions should be imposed. The main one is the corresponding ordering assumption

$$\lambda(t,z_1) < \lambda(t,z_2), \quad z_1 < z_2, \forall z_1, z_2 \in (0,\infty),\ t \geq 0. \tag{5}$$

Thus, we have an ordered family of failure rates of 'subpopulations'. The following result is obtained (Finkelstein (2008)):

Let ordering (5) take place for the described model of mixing. Then the following inequality holds

$$\lambda_m(t) < \lambda_b(t), \qquad t > 0. \tag{6}$$

Thus, owing to conditioning, the mixture failure rate is smaller than the unconditional one for each $t > 0$. This can be intuitively explained via the principle: "the weaker subpopulations are dying out first". Indeed, the subpopulations with larger failure rates terminate stochastically earlier 'pushing' the resulting unconditional failure rate down. Moreover, if $\lambda(t,z)$ is differentiable in both arguments and $\partial\lambda(t,z)/\partial z$ is increasing in t, then $\lambda_b(t) - \lambda_m(t)$ is increasing (Finkelstein and Cha (2013)).

As examples, we will consider two models of mixing which are popular in applications.

a. The additive model

The model failure rate process (MFRP) defined in (4) in this specific case is

$$\lambda_{bt} = \lambda(t) + Z, t \geq 0. \tag{7}$$

whereas the corresponding hazard rate process (HRP), in accordance with our definition, is

$$\lambda_{th} \equiv \lambda_{bt}|T_h > t = \lambda(t) + Z|T_h > t, \tag{8}$$

where the random variable $Z|T_h > t$ has the conditional pdf $\pi(z|t)$ (at each insant of time t) defined in (2). Then, taking expectations of both sides of (8) results in the population failure rate

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

Equation (7) defines for $z \in [0, \infty)$ a family of 'horizontally parallel' functions. Lynn and Singpurawalla (1997) had shown, under some mild additional assumptions, that when $\lambda(t)$ is increasing, the mixture failure rate can have a *bathtub failure rate*.

b. The multiplicative model

This model is, in a way, the proportional hazard (PH) model used in statistics to model the effect of covariates on the baseline failure rate. In our notation, it is defined for $Z = z$ as

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

whereas the corresponding mixture failure rate and the hazard rate process are

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

$$\lambda_{th} \equiv \lambda_{bt}|T_h > t = (Z|T_h > t)\lambda(t),$$

accordingly.

The effect of multiplication can result in more dramatic changes as compared with an additive model. For instance (Finkelstein (2008)), let Z be the 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. Then the mixture failure rate has a bell-shape, which is illustrated by Figure 1.

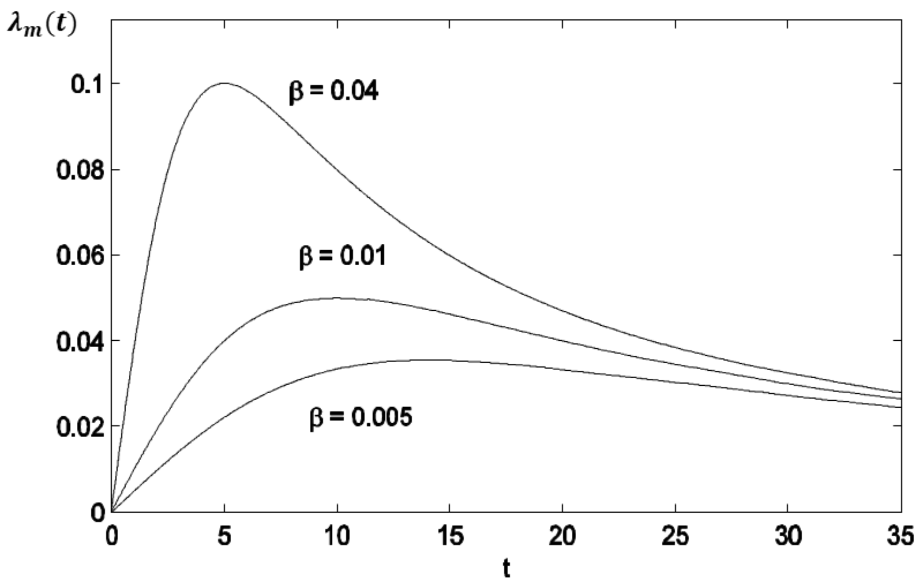

Figure 1. The mixture failure rate for the Weibull baseline distribution, $\gamma = 2$, $\alpha = 1$

This follows as a specific case of the mixture failure rate for the multiplicative model with the general baseline distribution (Beard (1971), Vaupel et al (1979)), i.e., $\lambda(t,z) = z\lambda(t)$, with a gamma-distributed frailty $\pi(z) = \frac{\beta^\alpha z^{\alpha-1}}{\Gamma(\alpha)}\exp\{-\beta z\}, z > 0$. In this case, the explicit form for the mixture survival function is given by

$$\bar{F}_m(t) = \frac{\beta^\alpha}{(\beta+\Lambda(t))^\alpha}$$

and its mixture failure rate is obtained by

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

where $\Lambda(t) = \int_0^t \lambda(x)dx$ is the corresponding cumulative baseline failure rate. The corresponding cumulative mixture failure rate is given by $\Lambda_m(t) = \int_0^t \lambda_m(x)dx = \ln(\beta + \Lambda(t))^\alpha - \ln\beta^\alpha$.

Thus, the shape of the mixture failure rate can dramatically differ from that of the model (baseline) failure rate. This presents a *paradox*, which is well-explained nowadays due to conditioning when obtaining the mixture failure rate.

As mentioned, the path process (4) presents the simplest model for the random failure rate with fixed, 'assigned' frailty. However, this is only a small class of models, where heterogeneity plays important role in describing the corresponding survival model. The next section will consider heterogeneity induced by a more general stochastic process and the heterogeneity in its sample paths can be considered in terms of the evolving frailty (Li and Anderson (2009)).

3. Hazard rate processes induced by shocks

Let an item operate in a random environment modelled by a point process of external shocks $\{N(t), t \geq 0\}$, where $N(t)$ is the number of shocks in $[0, t]$. A simple failure rate path model in (4) is now replaced by the following MFRP

$$\lambda_{bt} = \lambda_0(t) + \eta N(t), \tag{9}$$

whereas 'conditioning on survivors' defines the corresponding HRP $\{\lambda_{bt}|T_h > 0, t \geq 0\}$,

$$\lambda_{th} \equiv \lambda_{bt}|T_h > t = \lambda_0(t) + \eta(N(t)|T_h > t), \tag{10}$$

where $\lambda_0(t)$ is some deterministic function that can be considered as a background failure rate;

-$\{N(t), t \geq 0\}$ is a nonhomogeneous Poisson process (NHPP) of shocks with rate $r(t)$.

-η is a deterministic jump size on each event from the point process.

Thus, heterogeneity is now induced by the random number of shocks that occurred in each interval of time. It follows from (10) that for stochastic description, the properties of $N(t)|T_h > t, t \geq 0$ should be analysed. The following result provides the distribution of $N(t)|T_h > t$.

The conditional distribution of $(N(t)|T_h > t)$ *for each* $t > 0$ *is given by the Poisson distribution with mean* $\int_0^t exp\{-\eta(t-x)\}r(x)dx$*, that is,*

$$P(N(t) = n|T_h > t) = \frac{\left(\int_0^t exp\{-\eta(t-x)\}r(x)dx\right)^n}{n!} exp\left\{-\int_0^t exp\{-\eta(t-x)\}r(x)dx\right\},$$

$n = 0,1,2, \ldots$.

The *proof* of this supplementary result can be found in Cha and Finkelstein (2016a).

Taking expectations of both sides of (10) results in the following expression for the population failure rate, $\lambda_m(t)$

$$\lambda_m(t) = \lambda_0(t) + \eta E[N(t)|T_h > t]. \quad (11)$$

Using relationship for $P(N(t) = n|T_h > t)$,

$$\lambda_m(t) = \lambda_0(t) + \eta \int_0^t exp\{-\eta(t-x)\}r(x)dx. \quad (12)$$

On the other hand, the unconditional (model or baseline) population failure rate, similar to (3) and for the NHPP with rate $r(t)$ can be defined as

$$\lambda_b(t) = \lambda_0(t) + \eta \int_0^t r(x)dx \quad (13)$$

It follows from (12) that

$$\int_0^t exp\{-\eta(t-x)\}r(x)dx < \int_0^t r(x)dx$$

and, therefore, inequality (6) holds in this case as well.

Intuitively, this effect can be also explained because the weaker subpopulations (with the larger realization of $N(t)$) are dying out first, as opposed to the stronger subpopulations (with the smaller realization of $N(t)$). This is 'bending' the unconditional (model) failure rate $\lambda_b(t)$ down to result in the 'mixture' failure rate (12). However, the conditioning now is applied not to the assigned (fixed) random variable (frailty), but to the corresponding stochastic process (evolving heterogeneity). Thus, the considered heterogeneity is due to the variability of sample paths of the process of shocks in (9) (see, Anderson (2000); Li and Anderson (2009) and Finkelstein (2012) for the similar type of heterogeneity due to variability of the sample paths of the Wiener process and also for the discussion and examples of fixed and evolving heterogeneity).

Example 1. (Cha and Finkelstein (2016a)). Let $r(t) = r$. Then $E[N(t)] = rt$. Assume also that $\lambda_0(t) = 0$. Then

$$\eta E[N(t)|T_h > t] = \eta \int_0^t exp\{-\eta(t-x)\}r(x)dx = r(1 - exp\{-\eta t\}).$$

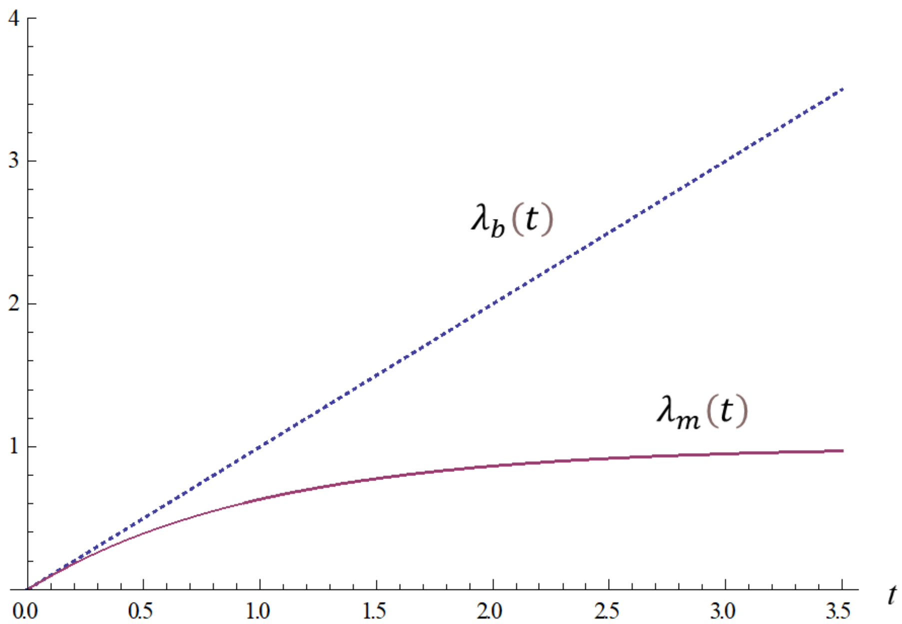

Figure 2. Population failure rate with $\eta = r = 1$.

Fig. 2 illustrates the resulting curves for the specific case when $\eta = r = 1$.

It can be easily seen that

$$\lambda_m(\infty) = lim_{t\to\infty} \eta\, E[N(t)|T_h > t] = r.$$

This is a remarkable fact showing that this specific model describes not only the *deceleration in population failure* rate (in the specified sense) but the *monotonic approaching to the plateau* as well. In accordance with our discussion in Section 2, this is happening due to evolving heterogeneity in the sample paths of the process: the most vulnerable items are dying out first and only survivors 'contribute' to the population failure rate, which is a conditional characteristic.

Example 2. Consider the decreasing in time rate of the NHPP of shocks $r(t) = r/(t + c)$, $c > 0$. This happens, when the environment is gradually becoming less severe. Then for $\lambda_0(t) = 0$,

$$\lambda_m(t) = \eta r\, exp\{-\eta t\} \int_0^t \frac{exp\{\eta x\}}{x + c} dx,$$

where the exponential integral can be calculated numerically. The corresponding bell-shaped is shown in Figure 3.

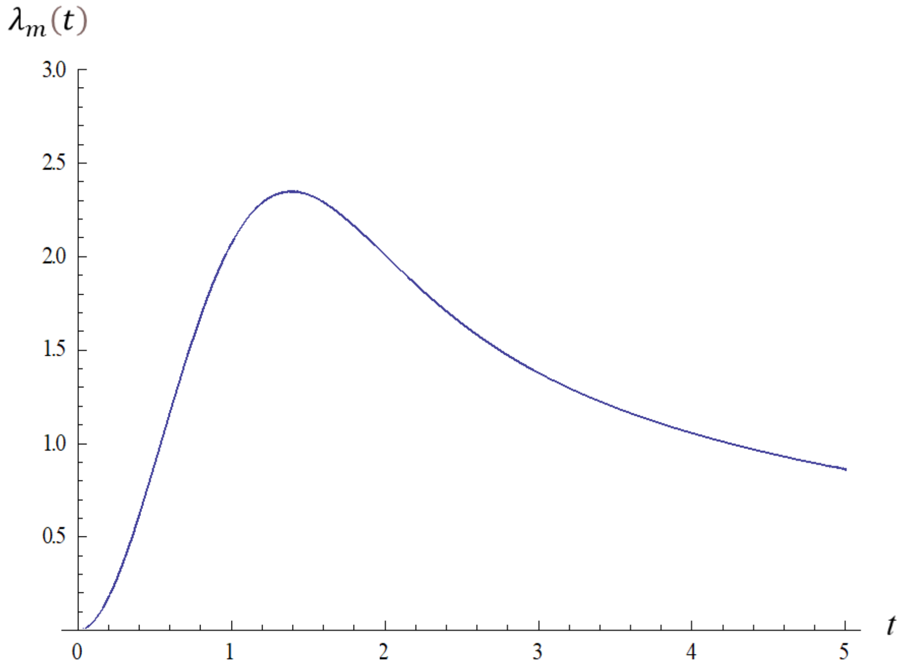

Figure 3. Population failure rate with $\eta = r = 5$, $c = 1$.

4. Survival model induced by the shock process with memory.

In the previous sections, heterogeneity was imposed directly on the failure rates via the fixed frailty (Section 2) or the evolving frailty (Section 3). For the latter case, it was due to heterogeneous sample paths of the NHPP $\{N(t), t \geq 0\}$. We will now consider the extreme shock survival model (Cha and Finkelstein (2018)) for an item subject to the process of external shocks. It can be seen that in this case, when the process is NHPP, there is no induced heterogeneity in the considered model (distinct from the previous section, where the NHPP process of shocks acts directly on the failure rate). However, if the process of shocks with memory (filtration) is considered, heterogeneity can arise in the framework of the extreme shock survival model, as it will be shown in what follows in this section.

Let $H_{t-} \equiv \{N(u), 0 \leq u < t\}$ be the history (internal filtration) of the orderly point process in $[0, t)$, i.e., the number of events in $[0, t)$ denoted by $N(t-)$ and the corresponding arrival times $T_0 \equiv 0 \leq T_1 \leq T_2 \leq \ldots \leq T_{N(t-)} < t$. The useful characterization of the point processes is via the stochastic intensity (Aven and Jensen (1999), Finkelstein and Cha (2013))

$$\lambda_t = lim_{\Delta t \to 0} \frac{Pr[N(t, t + \Delta t) = 1 | H_{t-}]}{\Delta t} = lim_{\Delta t \to 0} \frac{E[N(t, t + \Delta t) | H_{t-}]}{\Delta t} \quad (14)$$

where $N(t_1, t_2)$, $t_1 < t_2$, is the number of events in $[t_1, t_2)$. It should be noted that although some similarities can be seen, the MFRP in (9) is not formally a stochastic intensity in the defined sense as (14) defines the point process, whereas (9) just includes the point process into a model.

For our modelling, the generalized Polya process with explicit filtration will be used.

Definition 1. *(Cha (2014)).*

A counting process $\{N(t), t \geq 0\}$ *is called the Generalized Polya Process (GPP) with the set of parameters* $(\lambda(t), \alpha, \beta)$, $\alpha \geq 0, \beta > 0$, *if*

(i) $N(0) = 0$;

(ii) $\lambda_t = (\alpha N(t-) + \beta)\lambda(t)$.

The GPP with $(\lambda(t), \alpha = 0, \beta = 1)$ reduces to the NHPP with the rate $r(t) = \lambda(t)$. The following result was obtained in Cha (2014):

The probability of occurrence of n events for the GPP is given by

$$P(N(t) = n) = \frac{\Gamma(\beta/\alpha + n)}{\Gamma(\beta/\alpha) n!} (1 - exp\{-\alpha\, \Lambda(t)\})^n (exp\{-\alpha\, \Lambda(t)\})^{\frac{\beta}{\alpha}}, n = 0,1,2,\ldots, \quad (15)$$

where $\Lambda(t) \equiv \int_0^t \lambda(u)du$.

It follows from (15) that

$$E[N(t)] = \sum_1^n nP(N(t) = n) = \frac{\beta}{\alpha}(exp\{\, \alpha\, \Lambda(t)\} - 1). \quad (16)$$

Therefore, the rate of the GPP is

$$r_{GPP}(t) = \frac{d}{dt} E[N(t)] = \beta\lambda(t)\, exp\{\, \alpha\, \Lambda(t)\}. \quad (17)$$

Now we will define the corresponding survival model (Cha and Finkelstein (2016b)). Consider an item subject to the GPP process of shocks. Assume, for simplicity, that shocks constitute the only cause of failure. Let each shock result in item's failure with probability $p(t)$ and is survived with probability $q(t) = 1 - p(t)$ independently of 'everything else'. This defines the corresponding extreme shock model (see, e.g., Gut and Husler (2005)) that is well-known for the case of the NHPP of shocks with rate $r(t)$. In this specific case, the survival probability of an item $P(T > t)$ is defined as (Finkelstein (2008))

$$P(t) = P(T > t) = exp\left\{- \int_0^t p(u)r(u)\, du\right\} \quad (18)$$

with the corresponding failure rate of an item

$$\lambda_{it}(t) = p(t)r(t). \quad (19)$$

On the other hand, for the described GPP extreme shock model (Cha and Finkelstein (2018)):

$$P(T > t) = \frac{1}{\left(1 + \int_0^t \alpha p(v)\lambda(v)\, exp\{\, \alpha\, \Lambda(v)\}\, dv\right)^{\frac{\beta}{\alpha}}} \quad (20)$$

whereas the corresponding failure rate is

$$\lambda_{it}(t) = -\frac{dP(T > t)/dt}{P(T > t)} = \frac{\beta p(t)\lambda(t)\, exp\{\, \alpha\, \Lambda(t)\}}{\left(1 + \int_0^t \alpha p(v)\lambda(v)\, exp\{\, \alpha\, \Lambda(v)\}\, dv\right)}. \quad (21)$$

If in Definition 1, we fix the history as $N(t-)=n(t-)$, then the corresponding rate defines the NHPP and, in accordance with (18)-(19), the failure rate in this conditional failure model is $p(t)(\alpha n(t-)+\beta)\lambda(t)$, whereas $p(t)(\alpha N(t-)+\beta)\lambda(t)$ is, in fact, the corresponding MFRP $\lambda_{bt}, t \geq 0$, similar to (9) in the previous section (for a different model). As in (3) and (13), obtaining expectation of stochastic intensity $(\alpha N(t-)+\beta)\lambda(t)$, with respect to $N(t-)$ and using (16), we arrive at the model/baseline population failure rate

$$\lambda_b(t) = p(t)\beta\lambda(t)\, exp\{\,\alpha\Lambda(t)\}, \tag{22}$$

It immediately follows from (21) and (22) that

$$\lambda_{it}(t) < \lambda_b(t), \qquad t > 0.$$

We are now ready to interpret $\lambda_{it}(t)$ as the mixture failure rate $\lambda_m(t)$ similar to that in (1) and (12), whereas mixing in $\lambda_b(t)$ is obtained by taking expectation with respect to $N(t-)$. Thus, a similar effect of bending down of the failure rate due to induced heterogeneity takes place. It is important that we were able to interpret this effect of heterogeneity from the same viewpoint for three different models.

Example 3. Recall that the Gompertz lifetime distribution is described by the exponentially increasing failure/mortality rate.

$$\mu(t) = ae^{bt}, a, b > 0.$$

We will now derive this curve (even in a more 'advanced' form) using the extreme shock model developed in this section. Let $p(t) \equiv p;\ \lambda(t) \equiv \lambda; \beta = 1$. Then from (21)

$$\lambda_{it}(t) \equiv \lambda_m(t) = p(t) \cdot \frac{\lambda(t)\, exp\{\,\alpha\Lambda(t)\}}{\left(1 + \int_0^t \alpha p(v)\lambda(v)\, exp\{\,\alpha\Lambda(v)\}\, dv\right)}$$

$$= \frac{p\lambda e^{\alpha\lambda t}}{1 + p(1 - e^{\alpha\lambda t})},$$

which has the shape of the 'logistic curve' that is used in demography for modelling heterogeneous populations with exponentially increasing subpopulations failure rates (Vaupel et al., 1979). Note that the 'traditional mixing' of Section 2 will result also in the logistic curve with the corresponding parameters.

5. Stochastic intensity paradox

Recently Cha (2022) has proposed a new point process that is somewhat dual to the GPP described in the previous section. It is also defined via the stochastic intensity.

Definition 2.

A counting process $\{N(t), t \geq 0\}$ *is called the Extended Generalized Polya Process (EGPP) with the set of parameters* $(\lambda(t), \alpha, l)$, $\alpha \geq 0$, *l-integer, if*

(i) $N(0) = 0$*;*

(ii) $\lambda_t = (-\alpha N(t-) + \alpha l)\lambda(t)$.

For reliability interpretation, let $\alpha \equiv 1$. We can think in this case, e.g., about some debugging process in an item (program) when with each failure, the elimination of the corresponding defect occurs and the stochastic intensity decreases. However, our goal in this section is not in application, but rather in describing a fundamental property in context of mixing and heterogeneity. Note that, in this interpretation, l has a meaning of a number of defects (bugs) that are initially present in an item.

When $N(t-) = l$, the stochastic intensity becomes zero, which implies that additional events cannot occur anymore and, thus, the process terminates in this case. Equivalently, stochastic intensity can be defined as

$$\lambda_t = \sum_{i=0}^{l-1} (l-i)\lambda(t)\, I(T_i < t \le T_{i+1});\ \lambda_t = 0,\ t > T_{l,} \qquad (23)$$

where $T_i, i = 1,2,\dots,l$ is the arrival time of the i-th event and $T_0 = 0$.

It follows from Cha (2022) and also can be shown employing relevant reliability interpretation of the parallel system of l i.i.d. components that (compare with (15)):

$$P(N(t) = n) = \binom{l}{n}(1 - exp\{-\Lambda(t))^n (exp\{-\Lambda(t)\})^{l-n}, n = 0,1,2,\dots,l \qquad (24)$$

We see that this is the binomial distribution with the 'probability of success' $(1 - exp\{-\Lambda(t)\})$. Therefore,

$$E[N(t)] = l(1 - exp\{-\Lambda(t)\}). \qquad (25)$$

and the rate of this process is given by

$$r_{EGPP(t)} = d[E[N(t)]]/dt = l\lambda(t)\, exp\{-\Lambda(t)\} \qquad (26)$$

(compare with (16) and (17), respectively).

Example 4. Let $\lambda(t) = \lambda$, then (26) reduces to $l\lambda\, exp\{-\lambda t\}$ showing the exponential decline in the rate of occurrence of defects/bugs as in the NHPP-based models of software reliability. See, e.g., Musa and Okumoto (1984) and Goel (1985) for some initial *empirical* models of this kind. On the other hand, $r_{GPP}(t) = \beta\lambda\, exp\{\alpha\lambda t\}$ is exhibiting exponential growth.

We will now address some general properties of the GPP and the EGPP. A wide class of point processes is the class of self-exciting point processes. They describe random recurrent events when the occurrence of an event increases the likelihood of the occurrence of the subsequent events (Hawkes (1971, 2018)). Self-exciting point processes have, obviously, positively dependent increments and they are very useful in a wide range of applications. It follows from Definition 1 that GPP is the self-exciting process. On the other hand, there are point processes with a negative dependence between increments when the occurrence of an event decreases the likelihood of the occurrences of the subsequent events. Sometimes these processes are called the "self-regulating" (or self-correcting) point processes (Ertekin et al. (2015)). Obviously, as follows from Definition 2, the EGPP is the self-regulating process.

Thus, the GPP and EGPP are in a way dual and possess some opposite properties. However, an operation of appropriate mixing can 'turn' the terminating EGPP into non-terminating GPP! We will show this *paradoxica*l result using an example, whereas the general framework is reported in Cha and Finkelstein (2024). As the number of defects in an item is usually unknown, it is reasonable to assume that it is a random variable L. Let it be geometrically distributed with the corresponding pmf (other discrete distributions can be considered as well)

$$f_L(l) = \theta(1-\theta)^l,\ l = 0,1,2,\dots,\quad 0 < \theta < 1.$$

It can be shown using general results obtained in Cha and Finkelstein (2024), that the stochastic intensity for this mixed EGPP (in fact, it is a new type of a point process) can be obtained in the following form

$$\lambda_t = (N(t-)+1) + \frac{(1-\theta)\exp\{-\alpha\Lambda(t)\}}{1-(1-\theta)\exp\{-\alpha\Lambda(t)\}}\alpha\Lambda(t). \tag{27}$$

It follows from Definition 1 that this is the GPP just with a different parameter set, namely,

$$\left(\frac{(1-\theta)\exp\{-\alpha\Lambda(t)\}}{1-(1-\theta)\exp\{-\alpha\Lambda(t)\}}\alpha\Lambda(t), 1,1\right)$$

Thus, the terminating EGPP with decreasing rate after mixing with respect to L becomes the GPP with increasing rate, changing the 'pattern' to an opposite one. This counter-intuitive observation is due to heterogeneity induced via randomization and, in a way, has a similar origin as described in the previous sections via the principle "the weaker subpopulations are dying out first". Indeed, each realization of L can define the corresponding subpopulation with a 'lifetime' defined as the time when the process terminates. The larger realization corresponds to a stronger subpopulation with stochastically longer lifetime. As the weaker subpopulations are dying out first, the process with time the lifetimes are 'pushed' *upwards* eventually resulting in infinite lifetimes with probability 1. The latter corresponds to the non-terminating GPP. We intend to further justify this loose intuitive explanation and to report it elsewhere. Another topic to be explored in this respect is defining properties of mixing distributions that are 'responsible' for the described effect.

6. Concluding remarks

One can hardly find homogeneous populations in real life, however, most of reliability modelling, usually for simplicity, deals with a homogeneous case. Due to instability of production processes, environmental and other factors, most populations of manufactured items in real life are heterogeneous.

Mixtures of distributions usually present an effective mathematical tool for modeling heterogeneity, as they implement in the model an unobservable random variable, which is usually called "frailty". Obtaining expectations (or conditional expectations, as in the case of the failure rate) results in the population characteristics that can have properties different from those of the subpopulations. The convincing example of the latter is the possible change in aging properties of the population as compared with aging properties of subpopulations (e.g., the increasing failure rate turns to the bell-shaped). This presents the conventional approach in modeling heterogeneity considered in Section 2 via the fixed frailty assigned at $t = 0$.

However, at many instances, heterogeneity is imposed in a more complex way, which reflects, e.g., the impact of a random environment on a performance of operating items. In sections 3 and 4 this is modeled by external shocks that occur in accordance with the Poisson point process (direct impact on the baseline failure rate) and the generalized Polya process (extreme shock model), respectively. In both cases, the corresponding heterogeneity is not fixed and evolving with time (evolving heterogeneity). The end effects of this type of heterogeneity are similar to those for a fixed frailty, e.g., the population failure rate bends down as compared to the baseline: the weaker subpopulations are dying out first.

Finally, another paradox is outlined when, due to heterogeneity in a vital parameter of a model, a terminating point process (EGPP) with the decreasing rate after mixing becomes a non-terminating one with the increasing rate (GPP).

Acknowledgements

The authors would like to thank the reviewers for helpful comments and advice. The work of the second author was supported by the National Research Foundation of Korea (NRF) grant funded by the Korea government (MSIP) (No. 2023R1A2C1003238). The work of the second author was also supported by Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Education (Grant Number: 2019R1A6A1A11051177).

References

Aalen OO, Borgan O, Gjessing HK. *Survival and Event History Analysis: A Process Point of View*. Springer; 2008.

Anderson, J.J. A vitality-based model relating stressors and environmental properties to organism survival. *Ecological Monographs*,7093): 445-470, 2000.

Aven T, Jensen U. *Stochastic Models in Reliability*. Springer-Verlag, New York, 1999.

Barlow RE. A Bayesian explanation of the apparent failure rate paradox. *IEEE Transaction on Reliability*, 34: 107-108, 1985.

Barlow RE. and Proschan F. *Statistical Theory of Reliability and Life Testing. Probability Models*. New-York: Holt, Rinehart and Winston; 1975.

Beard, RE. Note on some mathematical mortality models. In C.E.W. Wolstenholme and M. O. Connor (eds). *The Lifespan of Animals*. Boston: Little, Brown, 302-311.

Cha JH. Characterization of the generalized Polya process and its applications. *Advances in Applied probability*, 46(4): 1148-1171, 2014.

Cha JH. An extended class of univariate and multivariate generalized Polya processes. *Advances in Applied Probability*, 54(3): 974-997, 2022.

Cha JH, Finkelstein M. Some notes on unobserved parameters (frailties) in reliability modeling. *Reliability Engineering and System Safety*,123: 99-103, 2014.

Cha JH, Finkelstein M. Point Processes for Reliability Analysis: Shocks and Repairable Systems. Springer; 2018.

Cha JH. and Finkelstein M. On some mortality rate processes and mortality deceleration with age. *Journal of Mathematical Biology*, 72:331-342, 2016a

Cha JH. and Finkelstein M. Justifying the Gompertz law of mortality via the generalized process of shocks. Theoretical Population Biology, 109: 54-62, 2016b.

Cha JH. and Finkelstein M. On the mixed extended generalized Polya process and its stochastic intensity paradox. Submitted.

Ertekin, S, Rudin, C, & McCormick, T H. Reactive point processes: A new approach to predicting power failures in underground electrical systems. *The Annals of Applied Statistics 9*: 122-144, 2015.

Finkelstein M, Cha JH. Stochastic Modelling for Reliability (shocks, burn-in and heterogeneous populations). Springer; 2013.

Finkelstein M. Failure Rate Modelling for Reliability and Risk. Springer; 2008.

Finkelstein M. Undersatnding the shape of mixture failure rate (With engineering and demographic applications). *Applied Stochastic Models in Business and Industry*, 25: 643-663, 2009.

Finkelstein M. On ordered subpopulations and population mortality at advanced ages. *Theoretical Population Biology*, 81: 292-299, 2012.

Goel AL. Software reliability models: assumptions, limitations and applicability. *IEEE Transactions on Software Engineering*, SE-11: 1411-1423, 1985.

Gut A, Husler J. Realistic variation of shock models. *Statistics and Probability Letters*, 74, 187-204, 2005.

Hawkes, AG. . Spectra of some self-exciting and mutually exciting point processes. Biometrika 58: 83-90, 1971.

Hawkes, AG. Hawkes processes and their applications to finance: a review. Quantitative Finance 18: 193-198, 2018.

Li T. and Anderson, JJ. The vitality model: a way to understand population survival and demographic heterogeneity. *Theoretical Population Biology*, 76: 118-131, 2009.

Lynn NJ and Singpurwalla ND. Comment: “Burn-in” makes us feel good. *Statistical Science*, 12: 13-19, 1997.

Musa, JD, Okumoto K. A logarithmic Poisson execution time model for software reliability measurement. In: Proceedings of the Seventh International Conference on Software Engineering. Orlando, FL, 230-237, 1984.

Singpurwalla ND. Anatomy of failure Rate. A mathematical dissection. *Applied Stochastic Models in Business and Industry*, 27:164-171, 2011.

Singpurwalla ND, Wilson SP. *Statistical Methods in Software Engineering*. Springer; 1999.

Vaupel, J.W., Manton, K.G., and Stallard, E. The impact of heterogeneity in individual frailty on the dynamics of mortality. *Demography*,16: 439-454, 1979.

Vaupel, J. W. and Yashin, A.I. Heterogeneity's ruses: some surprising effects of selection on population dynamics. *The American Statistician* 39:176-185, 1985.