This is a peer-reviewed, accepted author manuscript of the following research article: Cha, J. H., & Finkelstein, M. (2023). On age composition of dynamic heterogeneous populations. *Stochastic Models*. https://doi.org/10.1080/15326349.2023.2263538

On Age Composition of Dynamic Heterogeneous Populations

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

[a]Department of Statistics, Ewha Womans University, Seoul, 120-750, Rep. of Korea.
E-mail: jhcha@ewha.ac.kr

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

Abstract. Populations of items continuously manufactured and incepted into operation are characterized by the corresponding distribution of a random age at each instant of chronological time. In the current study, we consider heterogeneous populations with lifetime distributions indexed by a frailty parameter. Two approaches to defining the age composition for the described populations are described and the corresponding stochastic comparisons are considered.

Keywords: Frailty; heterogeneous populations; age composition; stochastic comparisons; production rate

2010 Mathematics Subject Classification: 60H3

1.Introduction

In recent years, quite a few studies have been reported in the literature that discuss the properties of items having random age at inception. Some relevant stochastic comparisons for random ages and remaining lifetimes can be found in, e.g., Li and Zuo (2004), Li and Xu (2006), Cai and Zheng (2012), Finkelstein and Vaupel (2015), Cha and Finkelstein (2018), Hazra et al (2018), Khaledi and Shaked (2010), Dewan and Khaledi (2014), Patra and Kundu (2021). Random age also naturally arises in population studies when items are continuously manufactured/born and we are interested in a distribution of age (to be called an *age composition*) in a population at a given chronological instant of time *t* or asymptotically when $t \to \infty$. Just for a feel, let us first recall the relevant description that is often used in demographic studies.

Denote by $M(x,t)$ the "approximated" total population size with age in $(0,x]$ at time t, which is approximated so that $M(x,t)$ is continuous and differentiable with respect to x. Thus, we can define

$$N(x,t) = \frac{d}{dx} M(x,t),$$

and $N(x,t)dx$ can be interpreted as the approximate population size with age in $(x, x+dx]$ at time t. Note that $N(x,t)$ is conventionally called the age-specific population size at time *t* in the demographic literature. See, e.g., Keiding (1990) and Arthur and Vaupel (1984) for discussion of

[1] Corresponding Author

this quantity. Let X_t denote a random age at time t of an item, which is picked out at random (with equal probabilities) from a population of size $M(t,t) = \int_0^t N(u,t)du$ (Cha and Finkelstein (2018)) and define the corresponding pdf

$$\pi_t(x) = \frac{N(x,t)}{\int_0^t N(u,t)du}, \quad 0 \le x \le t \tag{1}$$

to be called the "*age composition*". Thus, $\pi_t(x)dx$ can be interpreted as the approximate proportion of the population with age in $(x, x+dx]$ among total population at time t. Therefore,

$$\Pi_t(x) = \Pr(X_t \le x) = \frac{\int_0^x N(u,t)du}{\int_0^t N(u,t)du} \tag{2}$$

is the Cdf of X_t. Assuming that populations are sufficiently large, (1) and (2) can be used for description of X_t. In what follows, we will also use this natural intuitive reasoning that can be made easily mathematically strict.

The described populational definitions do not specify the underlying lifetimes of items that are usually assumed to be i.i.d. However, these distributions can be marginal masking the heterogeneous nature of items. Although many studies of heterogeneity in reliability and demography (mostly with respect to the mortality or failure rates) were triggered by the ground-braking paper by Vaupel et al. (1979), as far as we know, populational characteristics, i.e., the age composition of a population, *were discussed only for the homogeneous case*. The following definitions will be used throughout the paper.

Definition 1 (Shaked and Shanthikumar, 2007). Let X_1 and X_2 be two non-negative continuous (discrete) random variables with the corresponding cdfs $F_{X_1}(t)$ and $F_{X_2}(t)$, respectively. Denote their pdfs (pmfs) and failure rates as $f_{X_1}(t)$, $f_{X_2}(t)$, and $\lambda_{X_1}(t)$, $\lambda_{X_2}(t)$, respectively, if applicable.

(i) If $F_{X_1}(t) \ge F_{X_2}(t)$ for all $t \ge 0$, then X_1 is smaller than X_2 in the usual stochastic order, denoted by $X_1 \le_{st} X_2$.

(ii) If $\lambda_{X_1}(t) \ge \lambda_{X_2}(t)$ for all $t \ge 0$, then X_1 is smaller than X_2 in the hazard/ failure rate order, denoted by $X_1 \le_{hr} X_2$.

(iii) If $f_{X_1}(t)/f_{X_2}(t)$ decreases in $t \ge 0$, then X_1 is smaller than X_2 in the likelihood ratio order, denoted by $X_1 \le_{lr} X_2$.

2. Heterogeneous setting and assumptions

First, recall the setting for the homogeneous case. Assume that items are manufactured (incepted into operation) with rate $B(t)$ at time t meaning that $B(t)$ is the number of items incepted into operation in the small unit interval of time, whereas the process has started at $t=0$. In what follows, for convenience and having in mind the forthcoming integrations, we will (loosely) use notation $B(t)dt$ for the number of items manufactured in $[t,t+dt)$. For definiteness, let manufacturing of items and inception of them into operation be simultaneous (as for organisms). Assuming that the lifetimes of items manufactured at any $t\geq 0$ are i.i.d. with the Cdf $F(x)$ and survival function $\bar{F}(x)$. Then, $B(t-x)\bar{F}(x)dx = B(t-x)dx\bar{F}(x)$ can be interpreted as the expected number of alive items with age in $(x, x+dx]$ at time t, and , $\int_0^x B(t-u)\bar{F}(u)du$ can be interpreted as the total expected number of alive items with age $(0, x]$ at time t. Then, similar to (1), based on the expected numbers of alive items which "approximate" the exact numbers of alive items, the pdf of the age of the resulting population at time t is

$$\pi_t(x) = \frac{B(t-x)\bar{F}(x)}{\int_0^t B(t-u)\bar{F}(u)du} \cdot I(0 \leq x \leq t), \tag{3}$$

$$I(0\leq x\leq t) = \begin{cases} 1, & 0\leq x\leq t \\ 0, & x>t \end{cases}, \quad \bar{F}(x) \equiv 1-F(x), x\geq 0.$$

Thus, multiplied by dx, the numerator of $\pi_t(x)dx$ is the expected number of items alive and having the age in $(x, x+dx]$, whereas the denominator is the expected size of the population at time *t*. Therefore, $\pi_t(x)dx$ can be interpreted as the approximate proportion of items alive and having the age in $(x, x+dx]$ at time t among all the items alive at time t. Setting $B(t)=B, t\to\infty$, we arrive at the well-known density of the equilibrium distribution:

$$\pi_t(x) = \frac{\bar{F}(x)}{\int_0^t \bar{F}(u)du} \cdot I(0\leq x\leq t); \quad \pi_\infty(x) = \frac{\bar{F}(x)}{\int_0^\infty \bar{F}(u)du}. \tag{4}$$

The second relationship also characterizes the pdf of the remaining lifetime in this set up (see, e.g., Finkelstein and Vaupel (2015)), which is the same as asymptotic distributions of the forward and backward recurrence times for the corresponding renewal process (Ross (1996)).

Assume now that the lifetimes of items manufactured at time *t* are heterogeneous, as usually the case in practice due to non-stability of the production process, heterogeneous resources, changing environmental factors, etc. This means that the corresponding subpopulations' Cdfs can be modeled as

$$F_t(x,\alpha) = P(T_{t,\alpha}\leq x) = P(T_t \leq x \mid \mathrm{A}_t = \alpha),$$

where A_t is a random variable (frailty) with the pdf $\theta_t(\alpha)$ and support in $[0,\infty)$ defining variability in the manufactured/incepted items at time *t*; $T_{t,\alpha}$ is the lifetime of an item indexed by α, whereas T_t is a lifetime of an item with distribution defined as the following mixture:

$$F_t(x) = \int_0^\infty F_t(x,\alpha)\theta_t(\alpha)d\alpha. \tag{5}$$

In our model, heterogeneity in lifetimes is caused by heterogeneity in production processes. Thus, it is natural to assume that the production rate also depends on α, i.e., $B_{\alpha}(t)$. For example, in a binary case, the production rate of items of poor quality ('freak' items) is usually smaller than that of the normal quality. Similar to the homogeneous case, we can define $B_{\alpha}(t)dtd\alpha$ as the number of items with a frailty parameter in $[\alpha, \alpha + d\alpha)$ manufactured in $[t, t + dt)$.

Proposition 1. The pdf $\theta_t(\alpha)$ can be defined via the α-specific production rates as

$$\theta_t(\alpha) = \frac{B_{\alpha}(t)}{\int_0^{\infty} B_{\alpha}(t)d\alpha} . \tag{6}$$

Indeed, as stated, $B_{\alpha}(t)d\alpha dt$ is the number of items manufactured in $[t, t + dt) \times [\alpha, \alpha + d\alpha)$, whereas the denominator multiplied by dt is the overall number of items produced in $[t, t + dt)$. Thus, $\theta_t(\alpha)d\alpha$ defines the probability of $\mathrm{A}_t \in [\alpha, \alpha + d\alpha)$, i.e., $\theta_t(\alpha)$ is the pdf of A_t.

Thus, using the foregoing reasoning, the bivariate extension of the univariate pdf $\pi_t(x)$ in (3) to the heterogeneous case is

$$\pi_t(x, \alpha) = \frac{B_{\alpha}(t - x)\overline{F}_{t-x}(x, \alpha)}{\int_0^t B_{\alpha}(t - u)\overline{F}_{t-u}(u, \alpha)du} \cdot I(0 \le x \le t) , \tag{7}$$

which describes the age composition at time t of items belonging to the subpopulation characterized by the frailty variable α. Relationship (7) is rather general and the meaningful analysis of the corresponding stochastic properties can be performed after adopting important assumptions (some of them can hopefully be relaxed, which can constitute a topic for the future research).

Let $B_{\alpha}(t) \equiv B_{\alpha}$ do not depend on the chronological time, which describes the *stability* of the production process *in time*. Then the mixing distribution in (6), $\theta_t(\alpha) \equiv \theta(\alpha)(\mathrm{A}_t \equiv \mathrm{A})$ does not depend on *t*, while (7), if we further assume that $F_t(x, \alpha)$, similar to (3), also does not depend on *t,* reduces to.

$$\pi_t(x, \alpha) = \frac{\overline{F}(x, \alpha)}{\int_0^t \overline{F}(u, \alpha)du}; \; \pi_{\infty}(x, \alpha) = \frac{\overline{F}(x, \alpha)}{\int_0^{\infty} \overline{F}(u, \alpha)du} \tag{8}$$

for the finite and infinite time, respectively (the indicator is omitted for the notation sake).

For considering the heterogeneous case, similar to the lifetime modeling in heterogeneous populations (see, e.g., Finkelstein (2008)), we will need the following ordering in the family of lifetimes: usual stochastic order (Shaked and Shanthikumar (2007))

$$F(x, \alpha_1) \le F(x, \alpha_2), \alpha_1 \le \alpha_2, x \in [0, \infty) . \tag{9}$$

or the stronger hazard rate order

$$\lambda(x,\alpha_1) \le \lambda(x,\alpha_2),\ \alpha_1 \le \alpha_2,\ x \in [0,\infty)\,, \tag{10}$$

unless otherwise specified. Specifically, this natural ordering holds for the popular in applications proportional hazard (PH) model, i.e.,

$$\lambda(x,\alpha) = \alpha\lambda_b(x)\,, \tag{11}$$

where $\lambda_b(x)$ is a baseline failure rate that corresponds to some baseline Cdf $F_b(x)$.

3. Stochastic description of age composition

The age composition for the homogeneous populations was given in (3), whereas it is defined by the homogeneous subpopulation in (8). In order to 'integrate the frailty out', as it is done for the following 'natural' mixture (with respect to the frailty pdf $\theta_t(\alpha) \equiv \theta(\alpha)$) can be considered:

$$\tilde{\pi}_t(x) = \int_0^\infty \pi_t(x,\alpha)\theta(\alpha)d\alpha = \int_0^\infty \frac{\bar{F}(x,\alpha)}{\int_0^t \bar{F}(u,\alpha)du}\theta(\alpha)d\alpha. \tag{12}$$

Thus, $\tilde{\pi}_t(x)$ can be interpreted as the expectation of $\pi_t(x,\alpha)$ with respect to the pdf $\theta(\alpha)$. Is this a 'real' age composition of a population in accordance with the *general definition* (1) as it looks from the first sight? Note that, indeed $\pi_t(x,\alpha)$ is the age composition of the subpopulation indexed by α and $\theta(\alpha)$ represents by definition the proportion of the subpopulation indexed by α among items produced *at each time point*. However, the real proportion of the subpopulation indexed by α is not $\theta(\alpha)$ because some items already have failed before, whereas these probabilities of failures are different for subpopulations. It could be the same $\theta(\alpha)$ only if subpopulations are statistically identical (which means, in fact, homogeneity of a population) or if the items produced in accordance with $\theta(\alpha)$ do not fail at all. Denote the random age that corresponds to $\tilde{\pi}_t(x)$ by X_t.

Let us model now heterogeneous population not by mixing age compositions of subpopulations as in (12) but by 'mixing on the level of an item'. Thus, substituting $\bar{F}(x)$, by (5) for the case $\theta_t(\alpha) \equiv \theta(\alpha)$ and obtaining the age composition as for the corresponding homogeneous case,

$$\pi_t(x) = \frac{\int_0^\infty \bar{F}(x,\alpha)\theta(\alpha)d\alpha}{\int_0^\infty \int_0^t \bar{F}(u,\alpha)du\theta(\alpha)d\alpha} = \int_0^\infty \frac{\bar{F}(x,\alpha)}{\int_0^t \bar{F}(u,\alpha)du} \frac{\int_0^t \bar{F}(u,\alpha)du\theta(\alpha)}{\int_0^\infty \int_0^t \bar{F}(u,\alpha)du\theta(\alpha)d\alpha}d\alpha$$
$$= \int_0^\infty \pi_t(x,\alpha)\frac{\int_0^t \bar{F}(u,\alpha)du\theta(\alpha)}{\int_0^\infty \int_0^t \bar{F}(u,\alpha)du\theta(\alpha)d\alpha}d\alpha. \tag{13}$$

Denote the random age that corresponds to $\pi_t(x)$ by Y_t. Thus, $\pi_t(x)$, defined by the *intrinsic heterogeneity* in the population can be interpreted similar to the expectation of $\pi_t(x,\alpha)$ in (12), but with respect to the pdf

$$\theta(\alpha \mid t) \equiv \frac{\int_0^t \overline{F}(t-u,\alpha)du\theta(\alpha)}{\int_0^\infty \int_0^t \overline{F}(t-u,\alpha)du\theta(\alpha)d\alpha} = \frac{\int_0^t \overline{F}(u,\alpha)du\theta(\alpha)}{\int_0^\infty \int_0^t \overline{F}(u,\alpha)du\theta(\alpha)d\alpha}, \tag{14}$$

which is the conditional distribution of α given that the item produced in [0,t] is 'alive' at time *t*. To see this, it is sufficient to write the time-independent version of (6) as $\theta(\alpha) = B_\alpha / B$, where B is the production rate of a whole population and to substitute it in the numerator and denominator of (14) taking into account that $B_\alpha d\alpha$ is the number of items produced in the subpopulation with frailty in $[\alpha, \alpha + d\alpha)$ at any interval $[u, u+du), u \in [0,t]$. Then $\theta(a \mid t)$ will define the real proportion of items that are alive and belong to the subpopulation $[\alpha, \alpha + d\alpha)$ out of all items that are alive at time *t*.

Denote frailty with the pdf (14) by $\mathrm{A}(t)$. Then (12) and (13) can be written as the corresponding expectations

$$\tilde{\pi}_t(x) = E[\pi_t(x, \mathrm{A})];\ \pi_t(x) = E[\pi_t(x, \mathrm{A}(t))]. \tag{15}$$

The following simple proposition provides stochastic comparison of $\mathrm{A}(t)$ and A.

Proposition 2. Let the usual stochastic ordering (9) hold for lifetimes in the described heterogeneous populations of items. Then $\mathrm{A}(t)$ and A are ordered in the sense of the likelihood ratio ordering, i.e.,

$$\mathrm{A}(t) \le_{lr} \mathrm{A}, t > 0. \tag{16}$$

The *proof* follows immediately after considering the ratio as a function of α

$$\frac{\theta(\alpha \mid t)}{\theta(\alpha)} \propto \int_0^t \overline{F}(u,\alpha)du,$$

as due to assumption (9), the integral is decreasing in α for any fixed $t > 0$.

Thus, the age composition $\tilde{\pi}_t(x)$ given by (12) is obtained under the assumption that heterogeneity in the described population is fixed in time and is equal to heterogeneity in production rate defined by the pdf $\theta(\alpha)$. The *real*, heterogeneity, at time *t* is described by $\theta(a \mid t)$ and, due to the described above reasons, is stochastically smaller. This intuitively should lead to a stochastically larger random age, because the smaller values of α (at least for models (9) and (10)) result in stochastically larger lifetimes of the corresponding subpopulations. The latter means that the random age in this population is larger in some stochastic sense. Theorem 1 to follow proves this fact for our setting, whereas Remark 1, shows validity of $\tilde{\pi}_t(x)$ for some settings.

Remark 1. Age composition $\tilde{\pi}_t(x)$ can also have a practical meaning when mixing is executed on a different level and has different origin. To see this, as an example, consider a discrete setting with two subpopulations. Assume that we have two fleets of used items that are manufactured at different locations and operate in different environment (subpopulations). Therefore, the baseline Cdf for the manufactured items are different, i.e., $F_1(x)$ and $F_2(x)$. In accordance with (4), the age compositions (asymptotic) for each subpopulation is defined as

$$\pi_{1,\infty}(x) = \frac{\overline{F}_1(x)}{\int_0^\infty \overline{F}_1(u)du}, \pi_{2,\infty}(x) = \frac{\overline{F}_2(x)}{\int_0^\infty \overline{F}_2(u)du}$$

Let an item be picked up at random from a subpopulation, but the choice of a subpopulation is defined by probabilities φ_1 and $\varphi_2 = 1 - \varphi_1$ that can be defined by various factors (e.g., location, price, etc). Then the random age of the item picked up in this way is defined by the pdf

$$\tilde{\pi}_{1,2,\infty}(x) \equiv \pi_{1,\infty}(x)\varphi_1 + \pi_{2,\infty}(x)\varphi_2 ,$$

which is a discrete version of $\tilde{\pi}_\infty(x)$. On the other hand, the analogue of $\pi_\infty(x)$ is defined as

$$\tilde{\pi}_{1,2,\infty}(x) \equiv \frac{\varphi_1\overline{F}_1(x) + \varphi_2\overline{F}_2(x)}{\int_0^\infty (\varphi_1\overline{F}_1(u) + \varphi_2\overline{F}_2(u))du} ,$$

which describes the choice/preferences between items of two types.

We will now compare X_t (that corresponds to the age composition $\tilde{\pi}_t(x)$ given by (12)) and Y_t (that corresponds to the age composition $\pi_t(x)$ given by (13)). This stochastic comparison, for the sake of notation will be made for the asymptotic case, whereas the case of the finite t is similar. Note that, the proof of Theorem 1 'transforms' comparison of age compositions into the comparison of the corresponding mixture failure rates (Finkelstein (2008)).

Numerical illustration 1

In the following, the sensitivity analysis of the important measure of the distribution of the age composition $\pi_t(x)$ with respect to the modeling parameters is performed. In Figure 1.1, assuming that $\theta(\alpha) = 1, 1 \le \alpha \le 2$ (Uniform Distribution), the pdfs $\pi_t(x)$ have been obtained when $\overline{F}(x,\alpha) = exp\{-\alpha x\}$, $\overline{F}(x,\alpha) = exp\{-1.2\alpha x\}$, $\overline{F}(x,\alpha) = exp\{-1.4\alpha x\}$, $\overline{F}(x,\alpha) = exp\{-1.6\alpha x\}$ and $x \le t = 5$. The corresponding lifetimes are ordered in the sense of the usual stochastic ordering with $\overline{F}(x,\alpha) = exp\{-\alpha x\}$ defining the largest lifetime and $\overline{F}(x,\alpha) = exp\{-1.6\alpha x\}$-the smallest one. It can be seen that the corresponding densities, as expected, intersect and, e.g., the density that corresponds to the former distribution is first the smallest and then the largest (after intersection).

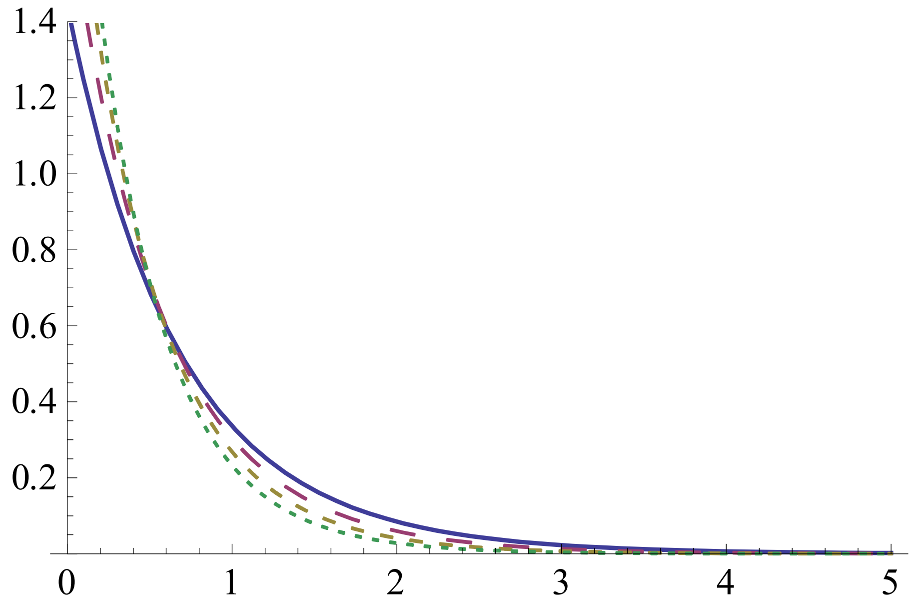

Figure 1.1 The pdfs $\pi_t(x)$ when $\overline{F}(x,\alpha) = exp\{-\alpha x\}$ (solid), $\overline{F}(x,\alpha) = exp\{-1.2\alpha x\}$ (large dotted), $\overline{F}(x,\alpha) = exp\{-1.4\alpha x\}$ (medium dotted), $\overline{F}(x,\alpha) = exp\{-1.6\alpha x\}$ (small dotted) and $t = 5$.

We have also performed numerical experiments for different uniform mixing distributions. Specifically, assuming that $\overline{F}(x,\alpha) = exp\{-\alpha x\}$, the pdfs $\pi_t(x)$ have been obtained when $\theta(\alpha) = 1, 1 \leq \alpha \leq 2$, $\theta(\alpha) = 1/1.2, 1 \leq \alpha \leq 2.2$, $\theta(\alpha) = 1/1.4, 1 \leq \alpha \leq 2.4$, $\theta(\alpha) = 1/1.6, 1 \leq \alpha \leq 2.6$ and $t = 5$ (see Figure 1.2). The results, as expected, have shown that the changes in the corresponding densities are not significant and can be neglected in practice. Thus, the pdf $\pi_t(x)$ is more robust with respect to the changes of the mixing distributions than to the changes of the subpopulation distributions.

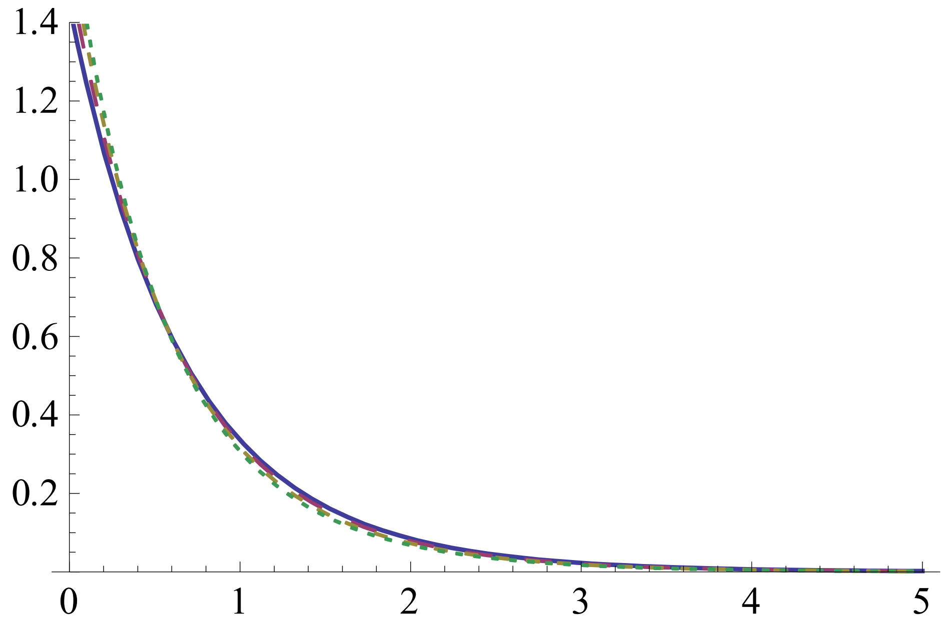

Figure 1.2 The pdfs $\pi_t(x)$ when $\theta(\alpha) = 1, 1 \le \alpha \le 2$ (solid), $\theta(\alpha) = 1/1.2, 1 \le \alpha \le 2.2$ (large dotted), $\theta(\alpha) = 1/1.4, 1 \le \alpha \le 2.4$ (medium dotted), $\theta(\alpha) = 1/1.6, 1 \le \alpha \le 2.6$ (small dotted) and $t = 5$.

4. Stochastic comparisons of random ages for two age compositions (one population)

Theorem 1. Let the usual stochastic ordering $F(x,\alpha_1) \le F(x,\alpha_2), \alpha_1 \le \alpha_2, x \in [0,\infty)$ hold for lifetimes in the described heterogeneous populations of items. Then

$$X_t \le_{lr} Y_t, t \ge 0 \text{ and } X_\infty \le_{lr} Y_\infty. \tag{19}$$

Proof. First, we show $X_\infty \le_{lr} Y_\infty$. Denote $\mu(\alpha) = \int_0^\infty \bar{F}(u,\alpha)du$. It is sufficient to show that the function $\Psi(x)$

$$\frac{\tilde{\pi}_\infty(x)}{\pi_\infty(x)} \propto \frac{\int_0^\infty \frac{\bar{F}(x,\alpha)}{\mu(\alpha)}\theta(\alpha)d\alpha}{\int_0^\infty \bar{F}(x,\alpha)\theta(\alpha)d\alpha} \equiv \Psi(x) \tag{20}$$

is decreasing in *x* as $\int_0^\infty \int_0^\infty \bar{F}(u,\alpha)\theta(\alpha)dud\alpha$ does not depend on *x*. Observe that

$$\Psi'(x) \equiv \frac{1}{\left(\int_0^\infty \bar{F}(x,\alpha)\theta(\alpha)d\alpha\right)^2} \times \left[\int_0^\infty (-f(x,\alpha))\frac{\theta(\alpha)}{\mu(\alpha)}d\alpha \int_0^\infty \bar{F}(x,\alpha)\theta(\alpha)d\alpha + \int_0^\infty f(x,\alpha)\theta(\alpha)d\alpha \int_0^\infty \bar{F}(x,\alpha)\frac{\theta(\alpha)}{\mu(\alpha)}d\alpha\right],$$

and thus, to prove the theorem, it is sufficient to show that

$$\frac{\int_0^\infty f(x,\alpha)\theta(\alpha)d\alpha}{\int_0^\infty \bar{F}(x,\alpha)\theta(\alpha)d\alpha} < \frac{\int_0^\infty f(x,\alpha)\frac{\theta(\alpha)}{\mu(\alpha)}d\alpha}{\int_0^\infty \bar{F}(x,\alpha)\frac{\theta(\alpha)}{\mu(\alpha)}d\alpha}$$

$$= \frac{\int_0^\infty f(x,\alpha)\frac{\theta(\alpha)}{\mu(\alpha)}\frac{1}{\int_0^\infty \frac{\theta(\alpha)}{\mu(\alpha)}d\alpha}d\alpha}{\int_0^\infty \bar{F}(x,\alpha)\frac{\theta(\alpha)}{\mu(\alpha)}\frac{1}{\int_0^\infty \frac{\theta(\alpha)}{\mu(\alpha)}d\alpha}d\alpha} = \frac{\int_0^\infty f(x,\alpha)\theta^*(\alpha)d\alpha}{\int_0^\infty \bar{F}(x,\alpha)\theta^*(\alpha)d\alpha},$$

where we introduce the pdf $\theta^*(\alpha) \equiv \frac{\theta(\alpha)}{\mu(\alpha)} / \int_0^\infty \frac{\theta(\alpha)}{\mu(\alpha)}d\alpha$. Note that,

$$\frac{\int_0^\infty f(x,\alpha)\theta(\alpha)d\alpha}{\int_0^\infty \bar{F}(x,\alpha)\theta(\alpha)d\alpha} = \int_0^\infty \frac{f(x,\alpha)}{\bar{F}(x,\alpha)}\frac{\bar{F}(x,\alpha)\theta(\alpha)}{\int_0^\infty \bar{F}(x,\alpha)\theta(\alpha)d\alpha}d\alpha = \int_0^\infty \lambda(x,\alpha)\frac{\bar{F}(x,\alpha)\theta(\alpha)}{\int_0^\infty \bar{F}(x,\alpha)\theta(\alpha)d\alpha}d\alpha$$

is the mixture failure rate function with respect to the mixing pdf $\theta(\alpha)$ and also is the expectation of $\lambda(x,\mathrm{A})$ with respect to the conditional pdf of $(\mathrm{A}\,|\,T>t)$, i.e., $\bar{F}(x,\alpha)\theta(\alpha)/\int_0^\infty \bar{F}(x,\alpha)\theta(\alpha)d\alpha$. On the other hand, similarly

$$\frac{\int_0^\infty f(x,\alpha)\theta^*(\alpha)d\alpha}{\int_0^\infty \bar{F}(x,\alpha)\theta^*(\alpha)d\alpha} = \int_0^\infty \lambda(x,\alpha)\frac{\bar{F}(x,\alpha)\theta^*(\alpha)}{\int_0^\infty \bar{F}(x,\alpha)\theta^*(\alpha)d\alpha}d\alpha$$

is the mixture failure rate function with respect to the mixing pdf $\theta^*(\alpha)$ and also is the expectation of $\lambda(x,\mathrm{A})$ with respect to the conditional pdf $\bar{F}(x,\alpha)\theta^*(\alpha)/\int_0^\infty \bar{F}(x,\alpha)\theta^*(\alpha)d\alpha$. Observe that the ratio of these two conditional pdfs

$$\frac{\dfrac{\bar{F}(x,\alpha)\theta(\alpha)}{\int_0^\infty \bar{F}(x,\alpha)\theta(\alpha)d\alpha}}{\dfrac{\bar{F}(x,\alpha)\theta^*(\alpha)}{\int_0^\infty \bar{F}(x,\alpha)\theta^*(\alpha)d\alpha}} = \frac{\int_0^\infty \bar{F}(x,\alpha)\theta^*(\alpha)d\alpha}{\int_0^\infty \bar{F}(x,\alpha)\theta(\alpha)d\alpha}\frac{\theta(\alpha)}{\theta^*(\alpha)} = \frac{\int_0^\infty \bar{F}(x,\alpha)\theta^*(\alpha)d\alpha}{\int_0^\infty \bar{F}(x,\alpha)\theta(\alpha)d\alpha}\int_0^\infty \frac{\theta(\alpha)}{\mu(\alpha)}d\alpha \cdot \mu(\alpha)$$

is decreasing in α as $\mu(\alpha)$ is decreasing due to (9). This means that the random variable that corresponds $\bar{F}(x,\alpha)\theta^*(\alpha)/\int_0^\infty \bar{F}(x,\alpha)\theta^*(\alpha)d\alpha$ is larger than that defined by the pdf $\bar{F}(x,\alpha)\theta(\alpha)/\int_0^\infty \bar{F}(x,\alpha)\theta(\alpha)d\alpha$ in the sense of the likelihood ratio, and, accordingly, in the sense of the usual stochastic order. Therefore, as $\lambda(x,\alpha)$ is increasing in α, we can conclude that

$$\int_0^\infty \lambda(x,\alpha)\frac{\bar{F}(x,\alpha)\theta(\alpha)}{\int_0^\infty \bar{F}(x,\alpha)\theta(\alpha)d\alpha}d\alpha < \int_0^\infty \lambda(x,\alpha)\frac{\bar{F}(x,\alpha)\theta^*(\alpha)}{\int_0^\infty \bar{F}(x,\alpha)\theta^*(\alpha)d\alpha}d\alpha,$$

which follows from the well-known results on ordering mixture failure rates for different mixing distributions (e.g., Finkelstein (2008). This inequality implies that $\Psi'(x)<0$, which completes the proof.

As the proof is the same for finite *t*, we also have under the same assumptions:

$$X_t \le_{lr} Y_t, t \ge 0.$$

■

Example 1. Let $\bar{F}(x,\alpha)=e^{-\alpha\lambda x}$, $\mu(\alpha)=1/\alpha\lambda$ (PH model) and $\theta(\alpha)=\beta e^{-\beta\alpha}$. Then, for illustration, we can obtain the integrals directly and show

$$\int_0^\infty \bar{F}(x,\alpha)\theta(\alpha)d\alpha=\frac{\beta}{\beta+\lambda x};\quad \int_0^\infty \frac{\bar{F}(x,\alpha)}{\mu(\alpha)}\theta(\alpha)d\alpha=\frac{\lambda\beta}{(\beta+\lambda x)^2}=\alpha(x)\frac{\beta}{\beta+\lambda x},$$

$$\frac{\int_0^\infty \frac{\bar{F}(x,\alpha)}{\mu(\alpha)}\theta(\alpha)d\alpha}{\int_0^\infty \bar{F}(x,\alpha)\theta(\alpha)d\alpha}=\alpha(x),$$

where $\alpha(x)=(\lambda\beta)/(\beta+\lambda x)$ is, obviously, decreasing with *x*. Therefore, (19) holds. The values of $\tilde{\pi}_t(x)$, $\pi_t(x)$ and $\tilde{\pi}_t(x)/\pi_t(x)$ for different values of *x* are given in Table 1 when $\lambda=1, \beta=1$, and t=5.

x	0	0.5	1.0	1.5	2.0	2.5	3.0	3.5	4.0	4.5	5.0
$\tilde{\pi}_t(x)$	1.050	0.490	0.291	0.197	0.145	0.113	0.092	0.077	0.066	0.057	0.051
$\pi_t(x)$	0.558	0.372	0.279	0.223	0.186	0.159	0.140	0.124	0.112	0.101	0.093
$\tilde{\pi}_t(x)/\pi_t(x)$	1.882	1.132	1.043	0.883	0.779	0.710	0.657	0.620	0.589	0.564	0.548

Table 1. The values of $\tilde{\pi}_t(x)$, $\pi_t(x)$ and $\tilde{\pi}_t(x)/\pi_t(x)$ for different of x when $\lambda=1$, $\beta=1$, and t=5

It can be seen that the ratio $\tilde{\pi}_t(x)/\pi_t(x)$ monotonically decreases and $X_t \le_{lr} Y_t$ holds.

5. Stochastic comparisons of random ages (two populations)

As discussed in the previous sections, the age composition that adequately describes the 'structure' of the defined heterogeneous population is given by the pdf of Y_t as

$$\pi_t(x) = \frac{\int_0^\infty \bar{F}(x,\alpha)\theta(\alpha)d\alpha}{\int_0^\infty \int_0^t \bar{F}(u,\alpha)du\theta(\alpha)d\alpha}$$

whereas, the corresponding limiting distribution is

$$\pi_\infty(x) = \frac{\int_0^\infty \bar{F}(x,\alpha)\theta(\alpha)d\alpha}{\int_0^\infty \int_0^\infty \bar{F}(u,\alpha)du\theta(\alpha)d\alpha}.$$

We will compare now age compositions for two heterogeneous populations, Population 1 and Population 2. Assume first that they share the same mixing distribution $\theta(\alpha)$. Denote by $f_i(x,\alpha)$, $F_i(x,\alpha)$, $\bar{F}_i(x,\alpha)$ and $\lambda_i(x,\alpha)$, $i=1,2$, the pdf, the Cdf, the survival function and the failure rate for the i-th population, respectively. Let Y_{1t}, $Y_{1\infty}$ and Y_{2t}, $Y_{2\infty}$ be the corresponding random ages.

Theorem 2. Suppose that
(i) $\lambda_i(x,\alpha)$, $i=1,2$, is increasing (decreasing) in α,
(ii) $\lambda_1(x,\alpha) > \lambda_2(x,\alpha)$, $\forall x>0$, for each $\alpha>0$,
(iii) $\int_0^x \lambda_1(u,\alpha)-\lambda_2(u,\alpha)du$ is decreasing (increasing) in α, for each $x>0$.
Then

$$Y_{1t} \le_{lr} Y_{2t}, t \ge 0 \text{ and } Y_{1\infty} \le_{lr} Y_{2\infty}.$$

Proof.
Denote by $\pi_{i\infty}(x)$, the pdf of $Y_{i\infty}$, $i=1,2$, respectively. Then

$$\frac{\pi_{1\infty}(x)}{\pi_{2\infty}(x)} = \frac{\int_0^\infty \int_0^\infty \bar{F}_2(u,\alpha)du\theta(\alpha)d\alpha}{\int_0^\infty \int_0^\infty \bar{F}_1(u,\alpha)du\theta(\alpha)d\alpha} \frac{\int_0^\infty \bar{F}_1(x,\alpha)\theta(\alpha)d\alpha}{\int_0^\infty \bar{F}_2(x,\alpha)\theta(\alpha)d\alpha}.$$

Observe that

$$\frac{d}{dx}\left(\frac{\int_0^\infty \bar{F}_1(x,\alpha)\theta(\alpha)d\alpha}{\int_0^\infty \bar{F}_2(x,\alpha)\theta(\alpha)d\alpha}\right) = \frac{1}{\left(\int_0^\infty \bar{F}_2(x,\alpha)\theta(\alpha)d\alpha\right)^2}$$

$$\times\left[\int_0^\infty (-f_1(x,\alpha))\theta(\alpha)d\alpha\int_0^\infty \bar{F}_2(x,\alpha)\theta(\alpha)d\alpha+\int_0^\infty f_2(x,\alpha)\theta(\alpha)d\alpha\int_0^\infty \bar{F}_1(x,\alpha)\theta(\alpha)d\alpha\right].$$

Thus, to show $Y_{1\infty}\le_{lr}Y_{2\infty}$, it is sufficient to show that

$$\frac{\int_0^\infty f_2(x,\alpha)\theta(\alpha)d\alpha}{\int_0^\infty \bar{F}_2(x,\alpha)\theta(\alpha)d\alpha}<\frac{\int_0^\infty f_1(x,\alpha)\theta(\alpha)d\alpha}{\int_0^\infty \bar{F}_1(x,\alpha)\theta(\alpha)d\alpha}.$$

Note that

$$\frac{\int_0^\infty f_i(x,\alpha)\theta(\alpha)d\alpha}{\int_0^\infty \bar{F}_i(x,\alpha)\theta(\alpha)d\alpha}=\int_0^\infty \lambda_i(x,\alpha)\frac{\bar{F}_i(x,\alpha)\theta(\alpha)}{\int_0^\infty \bar{F}_i(x,\alpha)\theta(\alpha)d\alpha}d\alpha\ ,\ i=1,2, \tag{21}$$

and the expression in (21) is the expectation of $\lambda_i(x,\alpha)$ wit respect to the distribution $\hat{\theta}_i(\alpha)\equiv\frac{\bar{F}_i(x,\alpha)\theta(\alpha)}{\int_0^\infty \bar{F}_i(x,\alpha)\theta(\alpha)d\alpha}$, $i=1,2$. Observe that the ratio of these pdfs

$$\frac{\hat{\theta}_1(\alpha)}{\hat{\theta}_2(\alpha)}\propto\frac{\bar{F}_1(x,\alpha)}{\bar{F}_2(x,\alpha)}=\exp\left\{-\int_0^x \lambda_1(u,\alpha)-\lambda_2(u,\alpha)du\right\}$$

is increasing in α due to assumption (iii), which means that the random variable corresponding to $\hat{\theta}_1(\alpha)$ is larger than that corresponds to $\hat{\theta}_2(\alpha)$ in the sense of likelihood ratio order and, accordingly, in the sense of the usual stochastic order. Thus,

$$\int_0^\infty \lambda_1(x,\alpha)\frac{\bar{F}_1(x,\alpha)\theta(\alpha)}{\int_0^\infty \bar{F}_1(x,\alpha)\theta(\alpha)d\alpha}d\alpha>\int_0^\infty \lambda_2(x,\alpha)\frac{\bar{F}_1(x,\alpha)\theta(\alpha)}{\int_0^\infty \bar{F}_1(x,\alpha)\theta(\alpha)d\alpha}d\alpha$$

$$\ge\int_0^\infty \lambda_2(x,\alpha)\frac{\bar{F}_2(x,\alpha)\theta(\alpha)}{\int_0^\infty \bar{F}_2(x,\alpha)\theta(\alpha)d\alpha}d\alpha, \tag{22}$$

where the first inequality holds due to assumption (ii) and the last inequality holds due to assumption (i). Inequality (22) means that $\frac{\pi_{1\infty}(x)}{\pi_{2\infty}(x)}$ is decreasing in x, which eventually implies that $Y_{1\infty}\le_{lr}Y_{2\infty}$. The result $Y_{1t}\le_{lr}Y_{2t}, t\ge 0$ can be proved in the same way.

■

As mentioned earlier, the proportional hazard (PH) model is practically important and is widely used in reliability and lifetime analysis. In this case, Theorem 2 can be simplified as follows.

Corollary 1. Let $\lambda_i(x,\alpha) = c_i(\alpha)\lambda(x)$, $i = 1,2$. If
(i) $c_i(\alpha)$, $i = 1,2$, is increasing (decreasing) in α
(ii) $c_1(\alpha) > c_2(\alpha)$, $\forall \alpha > 0$
(iii) $c_1(\alpha) - c_2(\alpha)$ is decreasing (increasing) in α.
Then

$$Y_{1t} \le_{lr} Y_{2t}, t \ge 0 \text{ and } Y_{1\infty} \le_{lr} Y_{2\infty}.$$

Proof.
The proof is straightforward, as assumptions (i)-(iii) of Theorem 2 are, obviously, met in this case.
■

Example 2. Suppose that $\lambda_i(x,\alpha) = c_i(\alpha)\lambda(x)$, $i = 1,2$.
(I) Let $c_1(\alpha) = 2 - e^{-\alpha}$ and $c_2(\alpha) = 2 - 2e^{-\alpha}$. Then $c_i(\alpha)$, $i = 1,2$, is increasing in α, $c_1(\alpha) - c_2(\alpha) = e^{-\alpha} > 0$, $\forall \alpha > 0$, and $c_1(\alpha) - c_2(\alpha)$ is decreasing in α.
(II) Let $c_1(\alpha) = e^{-\alpha}$ and $c_2(\alpha) = 0.5e^{-\alpha} + 0.5$. Then $c_i(\alpha)$, $i = 1,2$, is decreasing in α, $c_1(\alpha) - c_2(\alpha) = 0.5(1 - e^{-\alpha}) > 0$, $\forall \alpha > 0$, and $c_1(\alpha) - c_2(\alpha)$ is increasing in α.

In Theorem 2, the mixing distribution for the two heterogeneous populations was the same, whereas the subpopulation distributions of lifetimes were different. Assume now that the two heterogeneous populations share the same subpopulation distributions, but they have different mixing distributions. Denote by $\theta_1(\alpha)$ and $\theta_2(\alpha)$ the mixing distributions of the first and the second populations that describe the mixing random variables A_1 and A_2 respectively. We also denote by Y_{1t}, $Y_{1\infty}$ and Y_{2t}, $Y_{2\infty}$ the corresponding random ages for the two populations.

Theorem 3. Let $\lambda(x,\alpha_1) \le \lambda(x,\alpha_2), \alpha_1 \le \alpha_2, x \in [0,\infty)$ and the likelihood ratio ordering hold for mixing random variables: $\mathrm{A}_1 \ge_{lr} \mathrm{A}_2$. Then

$$Y_{1t} \le_{lr} Y_{2t}, t \ge 0 \text{ and } Y_{1\infty} \le_{lr} Y_{2\infty}..$$

Proof.
Similar to the proof of Theorem 2, it is sufficient to show that

$$\frac{\int_0^\infty f(x,\alpha)\theta_2(\alpha)d\alpha}{\int_0^\infty \bar{F}(x,\alpha)\theta_2(\alpha)d\alpha} < \frac{\int_0^\infty f(x,\alpha)\theta_1(\alpha)d\alpha}{\int_0^\infty \bar{F}(x,\alpha)\theta_1(\alpha)d\alpha}.$$

Due to the assumption that $A_1 \geq_{lr} A_2$, the ratio of the corresponding pdfs $\frac{\bar{F}(x,\alpha)\theta_1(\alpha)}{\int_0^\infty \bar{F}(x,\alpha)\theta_1(\alpha)d\alpha}$ and $\frac{\bar{F}(x,\alpha)\theta_2(\alpha)}{\int_0^\infty \bar{F}(x,\alpha)\theta_2(\alpha)d\alpha}$,

$$\frac{\int_0^\infty \bar{F}(x,\alpha)\theta_2(\alpha)d\alpha}{\int_0^\infty \bar{F}(x,\alpha)\theta_1(\alpha)d\alpha} \frac{\bar{F}(x,\alpha)\theta_1(\alpha)}{\bar{F}(x,\alpha)\theta_2(\alpha)}$$

is increasing in α, which means the likelihood ratio order. Thus,

$$\frac{\int_0^\infty f(x,\alpha)\theta_1(\alpha)d\alpha}{\int_0^\infty \bar{F}(x,\alpha)\theta_1(\alpha)d\alpha} = \int_0^\infty \lambda(x,\alpha)\frac{\bar{F}(x,\alpha)\theta_1(\alpha)}{\int_0^\infty \bar{F}(x,\alpha)\theta_1(\alpha)d\alpha}d\alpha \geq \int_0^\infty \lambda(x,\alpha)\frac{\bar{F}(x,\alpha)\theta_2(\alpha)}{\int_0^\infty \bar{F}(x,\alpha)\theta_2(\alpha)d\alpha}d\alpha$$

$$= \frac{\int_0^\infty f(x,\alpha)\theta_2(\alpha)d\alpha}{\int_0^\infty \bar{F}(x,\alpha)\theta_2(\alpha)d\alpha},$$

which completes the proof.

■

Numerical illustration 2

Suppose that the subpopulation distributions are exponential distributions with $\lambda(x,\alpha) = \alpha$, which fulfills the condition in Theorem 3. Further assume that the distribution of A_1 is given by $\theta_1(\alpha) = 1, 1.5 \leq \alpha \leq 2.5$, and that of A_2 is given by $\theta_2(\alpha) = 1, 1 \leq \alpha \leq 2$. Then the condition $A_1 \geq_{lr} A_2$ obviously holds. Fixing t=5, the graphs for the pdfs and Cdf of Y_{1t} and Y_{2t} are given in Figure 2.1 and 2.2, respectively. Furthermore, the graph for the ratio of pdfs $\tilde{\pi}_{1t}(x)/\pi_{2t}(x)$ is given in Figure 2.3.

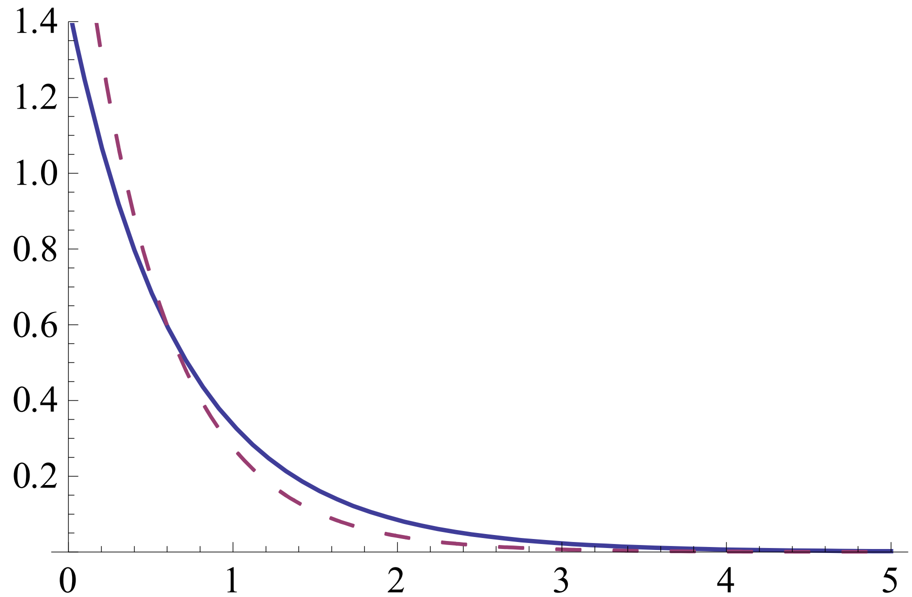

Figure 2.1. The pdfs of Y_{1t} (dotted) and Y_{2t} (solid)

From Figure 2.1, we can see that the distribution of Y_{2t} places more distribution weights to larger values and less distribution weights to smaller values.

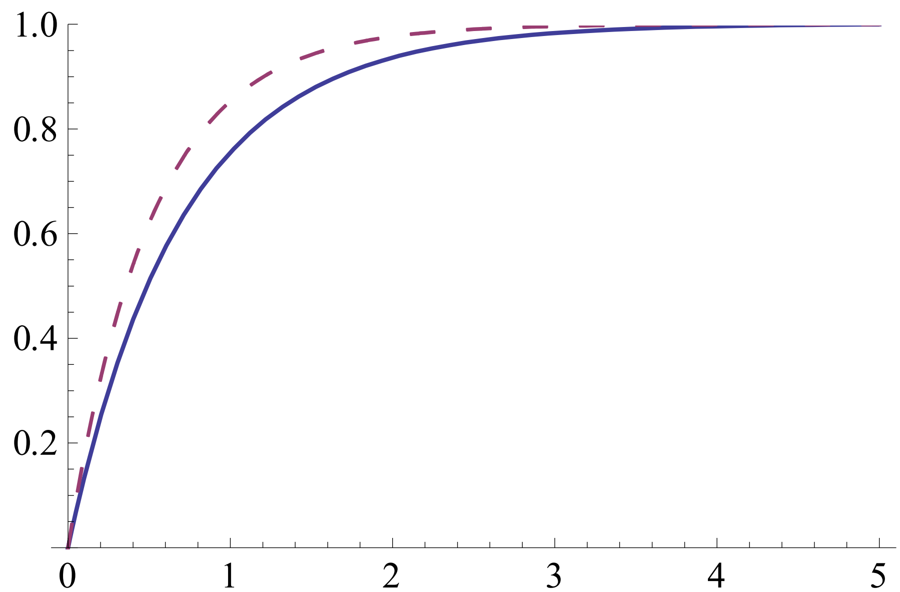

Figure 2.2. The Cdfs of Y_{1t} (dotted) and Y_{2t} (solid)

From Figure 2.2, we can see that the corresponding Cdfs are ordered, which is the result of $Y_{1t} \leq_{lr} Y_{2t}$, as $Y_{1t} \leq_{lr} Y_{2t}$ implies that $Y_{1t} \leq_{st} Y_{2t}$.

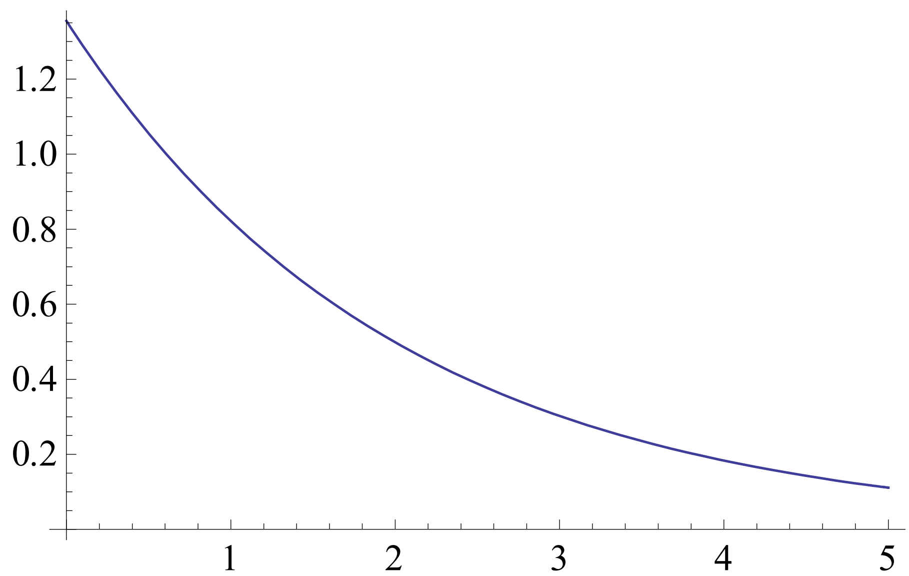

Figure 2.3. The ratio of pdfs $\tilde{\pi}_{1t}(x)/\pi_{2t}(x)$

From Figure 2.3, it can be seen that the ratio of pdfs $\tilde{\pi}_{1t}(x)/\pi_{2t}(x)$ is monotonically descreasing.

6. Concluding remarks

Population's age composition plays important role in describing stochastic properties of populations of items that are manufactured and continuously incepted into operation, as in the cases of cars, road machines, etc. For instance, having many 'aged' items in a population can result in the increased costs for maintenance and poor quality of performance.

Previous studies on age composition dealt with homogeneous populations. However, most of the real-life populations of manufactured items are heterogeneous due to non-stability of the production process, heterogeneous resources, changing environmental factors, etc. Therefore, in this paper, we study the effect of heterogeneity that is modeled by the relevant frailty parameters on the age composition of populations.

We discuss two approaches to defining the age composition for heterogeneous populations via mixing on different levels and obtain stochastic comparisons for the corresponding random variables. Furthermore, the relevant stochastic comparisons are obtained for two practically sound settings: when the mixing (frailty) distribution for the two populations with the same subpopulation distributions are different and when the mixing (frailty) distribution is the same but subpopulation distributions are different.

One of the possible directions for the future research in this area could focus on more practical applications in reliability and population studies.

Acknowledgements

The authors greatly thank the reviewers for their helpful comments and advice, which have improved the presentation of this paper. The work of the first author was supported by Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Education (Grant Number: 2019R1A6A1A11051177). The work of the first author was also supported by the National Research Foundation of Korea (NRF) grant funded by the Korea government (MSIP) (No. 2023R1A2C1003238).

References

Arthur, W.B.; Vaupel, J.W. Some general relationships in population dynamics. *Population Index* 1984, 50, 214-226.

Cai, N.; Zheng, Y. Preservation of generalized ageing class on the residual life at random time. *Journal of Statistical Planning and Inference* 2012, 142, 148–154.

Cha, J.H.; Finkelstein, M. On stochastic comparisons for population age structure and remaining lifetime. *Statistical Papres* 2018, 59, 199-213.

Dewan, I.; Khaledi, B.E. On stochastic comparisons of residual lifetime at a random time. *Statistics and Probability Letters* 2014, 88, 73-79.

Finkelstein, M. Modeling lifetimes with unknown initial age. *Reliability Engineering and System Safety* 2002, 76, 75-80.

Finkelstein, M.; Vaupel, J. On random age and remaining lifetime for population of items. *Applied Stochastic Models in Business and Industry* 2015, 31, 681-687.

Finkelstein, M. *Failure Rate Modeling for Reliability and Risk;* Springer: London, 2008.

Hazra, N.; Finkelstein, M.; Cha, J.H. Stochastic ordering for populations of manufactured items. *Test* 2018, 27, 173-196.

Khaledi, B.E.; Shaked, M. Stochastic comparisons of multivariate mixture. *Journal of Multivariate Analysis* 2010, 101, 2486–2498.

Keiding, N. Statistical inference in the Lexis diagram. *Philosophical Transactions of the Royal Society of London A* 1990, 332, 487-509.

Li, X.; Zuo, M.J. Stochastic comparison of residual life and inactivity time at a random time. *Stochastic Models* 2004, 20, 229–235.

Li, X.; Xu, M. Some results about MIT order and IMIT class of life distributions. *Probability in the Engineering and Information Sciencees* 2006, 20, 481–496.

Misra, N.; Naqvi, S. Stochastic comparison of residual lifetime mixture models. *Operations Research Letters* 2018, 46,122–127.

Patra, A.; Kundu, C. Stochastic comparisons and aging properties of residual lifetimes mixture models. *Mathematical Methods of Operations Research* 2021, 94, 123-143.

Ross, S. *Stochastic Processes*; New York: Wiley, 1996

Shaked, M.; Shanthikumar, J. *Stochastic Orders;* Springer: New York, 2007.

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