Exchangeable models of heterogeneity, mixtures, and multivariate aging properties

Fabio Spizzichino
Dipartimento di Matematica
Università "La Sapienza"
Roma, Italy
fabio.spizzichino@uniroma1.it

Abstract

We compare multivariate aging notions of conditionally independent lifetimes with corresponding univariate aging notions of their one-dimensional marginal distributions.

Let $\overline{G}(t|\theta)$ $(t \geq 0)$ be a one-dimensional survival function indexed by θ; for simplicity's sake, we assume here that θ is a non-negative scalar quantity; let, furthermore, $\Pi^{(1)}$ be a probability distribution over $[0, \infty)$ and consider the (*mixture-type*) one-dimensional survival function defined by

$$\overline{F}^{(1)}(t) = \int_0^\infty \overline{G}(t|\theta) d\Pi^{(1)}(\theta). \tag{1}$$

In a Bayesian standpoint, $\overline{G}(\cdot|\theta)$ can be seen as the conditional survival function of the lifetime T of a component, conditional on the value θ, taken on by a quantity Θ, which affects the life distribution of that component. In this view, interpreting $\Pi^{(1)}$ as the *a priori* distribution of Θ, $\overline{F}^{(1)}(\cdot)$ in (1) becomes the "predictive" survival function of T. In the same view, we focus attention on classes of (exchangeable) joint distributions, whose one-dimensional marginal distribution has the survival function coinciding with $\overline{F}^{(1)}(\cdot)$. For the moment, we first consider lifetimes $T_1, ..., T_n$ of n different components, where the distribution of each lifetime T_i $(i = 1, ..., n)$ is affected by Θ according to the position

$$P\{T_i > t|\Theta = \theta\} = \overline{G}(t|\theta). \tag{2}$$

If Θ is distributed according to $\Pi^{(1)}$, the (predictive) one-dimensional survival function of T_i is $\overline{F}^{(1)}(\cdot)$. Assuming that $T_1, ..., T_n$ are conditionally i.i.d. given Θ, their (exchangeable) joint survival function is

$$\overline{F}^{(n)}(t_1, ..., t_n) = \int_0^\infty \prod_{i=1}^n \overline{G}(t_i|\theta) d\Pi^{(1)}(\theta). \tag{3}$$

As it is well known, a wide literature has been devoted to analysing aging properties of univariate mixture-type distributions of the form (1). Some basic, by now classic, references on this subject are, e.g. Barlow and Proschan (1981) and Vaupel and Yashin (1985); several other references are reviewed in Shaked and Spizzichino (2001).

Assuming that a property of aging holds for $\overline{G}(t|\theta)$ $(\forall \theta \geq 0)$, here we want to compare corresponding aging properties of $\overline{F}^{(1)}(\cdot)$ and *corresponding* multivariate aging properties of $\overline{F}^{(n)}(\cdot, ..., \cdot)$, in the sense that we are going to explain henceforth. Later on, we shall briefly consider also a wider class of exchangeable distributions for $T_1, ..., T_n$, which still admit $\overline{F}^{(1)}(\cdot)$ in (1), as the one-dimensional marginal survival function.

Let $\leq_*$ be a one-dimensional stochastic ordering, such as *usual stochastic (st), hazard rate (hr), likelihood-ratio (lr),* or *mean residual life (mrl) orderings* (see Shaked and Shantikumar (1994) as a basic reference for definitions, basic properties, relations with concepts of aging, etc.). We fix attention on concepts of univariate aging (such as, *IFR, DFR, NBU, DMRL, PF_2,...)* for a lifetime T, which can be defined by conditions of the type

$$\overline{F}_{T-t'}(\tau|T > t') \geq_* \overline{F}_{T-t''}(\tau|T > t''), \forall (t', t'') \in A \tag{4}$$

where $\geq_*$ is a suitable notion of univariate ordering, A is an appropriate subset of $[0, \infty) \times [0, \infty)$, and $\overline{F}_{T-t}(\tau|T > t) = \frac{\overline{F}(t+\tau)}{\overline{F}(t)}$ is the conditional survival function of the *residual lifetime* $T - t$ given $\{T > t\}$.

In correspondence of notions of the above type, concepts of multivariate aging for exchangeable lifetimes $T_1, ..., T_n$ can be defined in a natural way (Bassan and Spizzichino (1999)) as we briefly recall here (see Spizzichino (2001) for more details and for an extended discussion on these notions of aging); for brevity's sake, we consider the simple case $n = 2$. Let then T_1, T_2 be two exchangeable lifetime and set $D = \{T_1 - t', T_2 > t''\}$. For given $\geq_*$ and A as above, we can define a concept of bivariate aging by setting the condition

$$\overline{F}_{T_1-t'}(\tau|D) \geq_* \overline{F}_{T_2-t''}(\tau|D), \forall (t', t'') \in A. \tag{5}$$

It is immediate that, when T_1, T_2 are in particular i.i.d., (5) holds if and only if their common one-dimensional survival function satisfies (4), which it is convenient here to rewrite in the equivalent form

$$\overline{F}_{T_1-t'}(\tau|T_1 > t') \geq_* \overline{F}_{T_2-t''}(\tau|T_2 > t''), \forall (t', t'') \in A. \tag{6}$$

The conditions (5) and (6) are however not equivalent in the more general case of exchangeability; actually, in the cases of stochastic dependence among T_1 and T_2, (5) is often a more interesting inequality than (6). We now concentrate attention on the case $A = \{(t', t'') \,|0 \leq t' < t''\}$ and list some existing results for different possible choices for $\geq_*$. In any case we consider the hypothesis

$$\text{The condition (4) is satisfied by } \overline{G}(\cdot|\theta), \forall \theta \in [0, \infty) \tag{7}$$

a) $\geq_* = \geq_{st}$ (i.e. (6) $\equiv$ IFR). In this case (7) implies the validity of (5) (this is well-known, see e.g. the discussion in Spizzichino(2001) and references therein). On the contrary, (6) is not implied by (7) (this is very well-known, also); further conditions on $\{\overline{G}(t|\theta)\}_{\theta \geq 0}$ and on $\Pi^{(1)}$, under which (7) implies (6) are given e.g. in Lynch (1999) and Finkelstein and Esaulova (2001).

b) $\geq_* = \geq_{hr}$ (i.e (6) $\equiv$ IFR, again). In this case (5) is not implied by (7) (see Bassan and Spizzichino (1999)). Additional conditions which guarantee (5) are given in Bassan, Kochar and Spizzichino (2002); such conditions are similar to those in Lynch (1999). Notice that the condition (6) under the ordering $\geq_{hr}$ just coincides with the same condition under the ordering $\geq_{st}$.

c) $\geq_* = \geq_E$ (letting $U \geq_E V$ if $\mathbb{E}(U) \geq \mathbb{E}(V)$; under this position, (6) coincides with the notion of marginal Decreasing Mean Residual Life). Now (7) implies the validity of (5) (see Bassan, Kochar and Spizzichino (2002)), while (6) is not implied by (7) (easy examples can be found); further conditions on $\{\overline{G}(t|\theta)\}_{\theta \geq 0}$ and on $\Pi^{(1)}$ can be found, under which (7) implies (6).

d) $\geq_* = \geq_{mrl}$ (letting $U \geq_{mrl} V$ if $\mathbb{E}(U - t|U > t) \geq \mathbb{E}(V - t|V > t), \forall t \geq 0$; also under this position, (6) coincides with the notion of Decreasing Mean Residual Life). In this case (5) is not implied by (7). Additional conditions which guarantee (5) are given in Bassan, Kochar and Spizzichino (2002); such conditions also imply (6).

More generally, in order to better understand both the differences and the analogies existing between the two conditions (5) and (6) in the case described by (3), notice that, for one-dimensional survival functions of the type (1), the function $\overline{F}_{T-t}(\tau|T > t)$ and the failure rate function can respectively be rewritten in the form

$$\overline{F}_{T-t}(\tau|T > t) = \int_0^\infty \frac{\overline{G}(t + \tau|\theta)}{\overline{G}(t|\theta)} d\Pi(\theta|T > t), \tag{8}$$

$$r(t) = \int_0^\infty \rho(t|\theta) d\Pi(\theta|T > t) \tag{9}$$

where, $\forall\theta \in [0, +\infty)$, $\rho(t|\theta)$ denotes the failure rate function of $\overline{G}(t_i|\theta)$ and $\Pi(\theta|T > t)$ is the conditional distribution of Θ, given the observation $\{T > t\}$. These formulae are particularly useful to analyze some aspects of the behavior of the function $r(t)$ in terms of monotonicity properties of the functions $\rho(\cdot|\theta)$ (this is specially true when it is assumed that $\rho(\cdot|\theta)$ is an increasing function of θ, which is appropriate when Θ is interpreted as a *frailty* for the components). Equations (8) and (9) were considered by the author in a paper (Spizzichino (1992)) mainly concerned with the concepts of IFR and DFR (i.e. with the case $\geq_* = \geq_{st}$) and with their multivariate extensions; for the same case the same equations also motivated interesting developments presented in Lynn and Singpurwalla (1997) and Finkelstein and Esaulova (2001). More in general we notice here that, using the representation (8), the condition (6) can be rewritten as

$$\int_0^\infty \frac{\overline{G}(t' + \tau|\theta)}{\overline{G}(t'|\theta)} d\Pi(\theta|T > t') \geq_* \int_0^\infty \frac{\overline{G}(t'' + \tau|\theta)}{\overline{G}(t''|\theta)} d\Pi(\theta|T > t''), \forall(t', t'') \in A. \tag{10}$$

Similarly (5) can be rewritten in the form

$$\int_0^\infty \frac{\overline{G}(t' + \tau|\theta)}{\overline{G}(t'|\theta)} d\Pi(\theta|D) \geq_* \int_0^\infty \frac{\overline{G}(t'' + \tau|\theta)}{\overline{G}(t''|\theta)} d\Pi(\theta|D), \forall(t', t'') \in A \tag{11}$$

Beside allowing us to see what are, in principle, the differences and the relations between (5) and (6), (10) and (11) can also be exploited to obtain some results for specific cases. By extending the analysis worked out in Finkelstein and Esaulova (2001), for instance, results of interest might be obtained for the two special cases of *Additive Models* and *Multiplicative Models*, respectively defined by the positions $\rho(t|\theta) = \alpha(t) + \theta$ and $\rho(t|\theta) = \theta \cdot \alpha(t)$.

We now conclude this note by mentioning that the same problems considered so far, for joint models of the type in (3), can be posed for the more general models defined as follows. Consider lifetimes $T_1, ..., T_n$ of n different components, where the distribution of each lifetime T_i $(i = 1, ..., n)$ is affected by Θ_i, the frailty of the component i. We assess the conditional distribution of T_i given Θ_i according to the position

$$P\{T_i > t|\Theta_i = \theta\} = \overline{G}(t|\theta), \tag{12}$$

so that, if Θ_i is distributed according to $\Pi^{(1)}$, the (predictive) one-dimensional survival function of T_i is then again the mixture given in (1).

We are interested in modelling, at a time, situations of both heterogeneity and possible (symmetric) dependence among $T_1, ..., T_n$ and then consider the following further assumptions:

i) The distribution of T_i conditional on $\Theta_1, ..., \Theta_n$ is given by

$$P\{T_i > t|\Theta_1 = \theta_1, ..., \Theta_n = \theta_n\} = P\{T_i > t|\Theta_i = \theta_i\} = \overline{G}(t|\theta_i), \tag{13}$$

ii) T_i and $T_1, ..., T_{i-1}, T_{i+1}, ..., T_n$ are conditionally independent given the vector $\Theta_1, ..., \Theta_n$

iii) $\Theta_1, ..., \Theta_m$ are *exchangeable*, and we denote by $\Pi^{(n)}$ their joint distribution. They in particular are identically distributed, and we assume their common marginal prior distribution is $\Pi^{(1)}$.

It can easily be shown that, under (12), i), ii), and iii), $T_1, ..., T_n$ are exchangeable; in particular the unconditional joint survival function is

$$\overline{F}^{(n)}(t_1, ..., t_n) = \int_0^\infty ... \int_0^\infty \prod_{i=1}^n \overline{G}(t_i|\theta_i) d\Pi^{(n)}(\theta_1, ..., \theta_n); \tag{14}$$

The marginal survival functions clearly coincide with (1). Even if the position (14) describes very simple and special models accounting both for frailties and dependence among lifetimes, they can be useful for some clarification of ideas. In particular, two very special cases of the above models are obtained by assuming

α) $\Theta_1, ..., \Theta_n$ i.i.d, in which case $T_1, ..., T_n$ are also i.i.d

β) $\Theta_1, ..., \Theta_n$ are such that $P\{\Theta_1 = ... = \Theta_n\} = 1$, in which case $T_1, ..., T_n$ are conditionally i.i.d, giving rise to the usual case of Bayesian statistics.

Several other cases of interest for the (exchangeable) joint distribution of $\Theta_1, ..., \Theta_n$ are possible, by modifying the joint distribution $\Pi^{(n)}$. The model in β), which in a sense eliminates heterogeneity, does of course coincide with (3). One can generally expect heuristically that heterogeneity among components force the model to present situations of infant mortality. In order to extend to such models some of the analyses above notice that, in such a case, the formula (11) is to be replaced, for the case $n = 2$, by

$$\frac{\overline{G}(t' + \tau|\theta)}{\overline{G}(t'|\theta)} d\Pi(\theta_1|D) \geq_* \int_0^\infty \frac{\overline{G}(t'' + \tau|\theta)}{\overline{G}(t''|\theta)} d\Pi(\theta_2|D), \forall (t', t'') \in A$$

This can be used to show that conditions which guarantee (5) for the case in β) are not generally sufficient to achieve a similar result for the case in (14). Multivariate negative aging properties for these models have been studied in Spizzichino and Torrisi (2001).

References

[1] Barlow R. and Proschan F. (1981), *Statistical Theory of Reliability and Life Testing.* To Begin With: Silver Spring, Maryland.

[2] Bassan B., Kochar S. and Spizzichino F. (2002), Some bivariate notions of IFR and DMRL and related properties. Preprint Department of Mathematics, University “La Sapienza”.

[3] Bassan B. and Spizzichino F. (1999), Stochastic comparison for residual lifetimes and Bayesian notions of multivariate aging. *Advances in Applied Probability*, **31**, 1078–1094.

[4] Finkelstein M.S. and Esaulova V. (2001), Modeling a failure rate for a mixture of distribution functions. *Probability in Engineering and Informational Sciences*, **15**, 383–400.

[5] Lynn, N, J and Singpurwalla, N. D. (1997). Comment: “Burn-in” makes us feel good. *Statistical Science*, **12**, 13-19.

[6] Lynch J.D. (1999), On conditions for mixtures of increasing failure rate distributions to have an increasing failure rate. *Probab. Engrg. Inform. Sci.*, **13**, 33–36.

[7] Shaked, M. and Shanthikumar, J.G. (1994). *Stochastic Orders and their Applications.* Academic Press, San Diego, CA.

[8] Shaked, M. and Spizzichino F. (2001), Mixtures and Monotonicity of failure Rate Functions, in *Advances in Reliability* (N. Balakrishnan and C.R. Rao Eds.) Handbook of Statistics, Vol. 20, Elsevier Science B.V.

[9] Spizzichino, F. (1992). Reliability Decision Problems under Conditions of Ageing. In *Bayesian Statistic 4*, J. Bernardo, J. Berger, A.P. Dawid, and A.F.M. Smith Eds., Clarendon Press, Oxford, 803-811.

[10] Spizzichino F. (2001), *Subjective probability models for lifetimes.* Chapman & Hall / CRC, Boca Raton, FL.

[11] Spizzichino, F. and Torrisi, G. L. (2001) Multivariate negative aging in an exchangeable model of heterogeneity. *Statist. Probab. Lett.* **55** no. 1, 71–82.

[12] Vaupel, J.V. and Yashin, A.I. (1985). Some surprising effects of selection on population dynamics. *The Am. Statist.*, 39, 176-185.