Probability in the Engineering and Informational Sciences, **36**, 2022, 461–481.
doi:10.1017/S0269964820000467

ORDERINGS OF FINITE MIXTURE MODELS WITH LOCATION-SCALE DISTRIBUTED COMPONENTS

GHOBAD BARMALZAN
Department of Statistics, University of Zabol, Sistan and Baluchestan, Iran
E-mail: ghobad.barmalzan@gmail.com

SAJAD KOSARI
Department of Mathematics, University of Zabol, Sistan and Baluchestan, Iran

NARAYANASWAMY BALAKRISHNAN
Department of Mathematics and Statistics, McMaster University, Hamilton, Ontario L8S 4L8, Canada

In this paper, we consider finite mixture models with components having distributions from the location-scale family. We then discuss the usual stochastic order and the reversed hazard rate order of such finite mixture models under some majorization conditions on location, scale and mixing probabilities as model parameters.

Keywords: finite mixture models, location parameters, location-scale distribution, multivariate chain majorization order, reversed hazard rate order, scale parameters, usual stochastic order

1. INTRODUCTION

Finite mixtures have found applications in many areas, such as reliability theory, actuarial science, medicine, and economics, for modeling data arising from a finite number of heterogeneous populations; see, for example, Everitt and Hand [5], Titterington *et al.* [19], Finkelstein and Esaulova [7], Franco *et al.* [8], and Amini-Seresht and Zhang [1]. For instance, manufactured units are often heterogeneous for different reasons such as quality of resources and components used in the production process, operational history and human errors; see Marshall and Olkin [14], Cha and Finkelstein [4] and Finkelstein [6].

For an absolutely continuous random variable X, denote its probability density function (pdf) by $f_X(\cdot)$, cumulative distribution function (cdf) by $F_X(\cdot)$, survival (reliability) function (sf) by $\bar{F}_X(\cdot)$, hazard rate function by $r_X(\cdot)$, and reversed hazard rate function by $\tilde{r}_X(\cdot)$. Formally, these functions describe a homogeneous, infinite population of units meaning that if we draw a unit at random from a population, its lifetime can be characterized by any of these functions. However, if we have different infinite populations (called subpopulations) and draw a unit with certain probabilities from each subpopulation, this unit will naturally be described by the corresponding mixed characteristics and the whole population of units will then be heterogeneous.

© The Author(s), 2020. Published by Cambridge University Press 0269-9648/19 $25.00

https://doi.org/10.1017/S0269964820000467 Published online by Cambridge University Press

Comparison of important lifetime characteristics of technical systems is of great importance in reliability theory. A convenient tool for this purpose is the theory of stochastic orderings.

Suppose we have n homogeneous infinite subpopulations of units with X_i denoting the lifetime of a unit in the ith subpopulation, $i = 1, 2, \ldots, n$. Then, let $M(\boldsymbol{X}, \boldsymbol{p})$ be a random variable representing the mixture of units that are drawn from these n subpopulations. Evidently, its distribution, survival and density functions are given by

$$F_{M(\boldsymbol{X},\boldsymbol{p})}(x) = \sum_{i=1}^{n} p_i F_{X_i}(x),$$

$$\bar{F}_{M(\boldsymbol{X},\boldsymbol{p})}(x) = \sum_{i=1}^{n} p_i \bar{F}_{X_i}(x),$$

$$f_{M(\boldsymbol{X},\boldsymbol{p})}(x) = \sum_{i=1}^{n} p_i f_{X_i}(x),$$

respectively, where p_i (>0) are the mixture proportions (weights) with $\sum_{i=1}^{n} p_i = 1$. As the failure rate is a conditional characteristic, the corresponding mixture failure rate is defined through modified conditional weights (on the condition of survival in $[0, t)$); see, for example, Navarro and Hernandez [17], Finkelstein [6], and Cha and Finkelstein [4].

Hazra and Finkelstein [9] recently established some results concerning stochastic comparisons of two finite mixtures when each of them is from one of the following semi-parametric families: proportional hazards, accelerated lifetime, and proportional reversed hazards families.

A random variable X is said to follow the location-scale family (denoted by $X \sim LS(\lambda, \mu)$) if its distribution function is given by

$$F(x; \lambda, \mu) = F\left(\frac{x-\lambda}{\mu}\right), \quad x > \lambda,\ \mu > 0,$$

where λ ($\in \mathbb{R}$) and μ are the location and scale parameters, respectively, and $F(\cdot)$ is the baseline distribution function (cf. [14]). Some well-known distributions that could be used as a baseline distribution are generalized Pareto, Burr, exponentiated Weibull, Pareto, power-generalized Weibull, generalized gamma, and exponentiated Weibull-geometric models. These models have all been used extensively in the analysis of lifetime data arising from reliability and survival studies; see, for example, Johnson *et al.* [11,12] and Marshall and Olkin [14].

The lifetime models mentioned above have been utilized in the analysis of lifetime data from reliability experiments when the location parameter is both present and absent. The location parameter, when considered, in these models are often interpreted to be a lower threshold value, and such models have been used extensively in hydrological, environmental, and climate applications in which the location parameter corresponds to the minimum threshold value of the observed characteristic; see, for example, Hosking and Wallis [10] and Castillo *et al.* [3] for many practical data sets which have been analyzed with the use of such threshold models comprising a location (threshold) parameter. Thus, the stochastic ordering results established here will not only have applications and interpretations in reliability theory but also in some applied fields such as hydrology, climatology, and environmental science.

In this paper, we discuss some stochastic comparisons of two finite mixture models with the components having location-scale distributions. Section 2 reviews some basic concepts

https://doi.org/10.1017/S0269964820000467 Published online by Cambridge University Press

that will be used in the sequel. Section 3 is divided into two parts, with the first part dealing with stochastic comparison of location-scale mixture models when the mixture proportions and threshold (location) parameters are different, while the second part deals with location-scale mixture models when the mixture proportions and scale parameters are different. Section 4 discusses the reversed hazard rate order of two finite mixture models with location-scale distributed components. Several examples are also presented to illustrate all the established results. Finally, some concluding remarks are made in Section 5.

2. PRELIMINARIES

In this section, we present some basic definitions and lemmas that are essential for subsequent developments.

DEFINITION 1: *Suppose X and Y are two non-negative continuous random variables with density functions f_X and f_Y, distribution functions F_X and F_Y, survival functions $\bar{F}_X$ and $\bar{F}_Y$, and reversed hazard rate functions $\tilde{r}_X$ and $\tilde{r}_Y$, respectively. We assume that all expectations exist wherever they are given. Then:*

(i) X is said to be larger than Y in the usual stochastic order (denoted by $X \geq_{st} Y$) if $\bar{F}_X(t) \geq \bar{F}_Y(t)$, for all $t \in \mathbb{R}_+$. This is equivalent to saying that $\mathbb{E}(\phi(X)) \geq \mathbb{E}(\phi(Y))$ for all increasing functions $\phi : \mathbb{R}_+ \to \mathbb{R}$;

(ii) X is said to be larger than Y in the reversed hazard rate order (denoted by $X \geq_{\mathrm{rh}} Y$) if and only if $F_X(t)/F_Y(t)$ is increasing in $t \in \mathbb{R}_+$. This is equivalent to saying that $\tilde{r}_Y(t) \geq \tilde{r}_X(t)$ for all $t \in \mathbb{R}_+$.

It is clear that the reversed hazard rate order includes the usual stochastic order as a special case. Interested readers may refer to Müller and Stoyan [16] and Shaked and Shanthikumar [18] for a comprehensive discussion on various stochastic orderings and inter relationships between them.

DEFINITION 2: *Consider two vectors $\boldsymbol{a} = (a_1, \ldots, a_n)$ and $\boldsymbol{b} = (b_1, \ldots, b_n)$, with increasing arrangements $a_{(1)} \leq \cdots \leq a_{(n)}$ and $b_{(1)} \leq \cdots \leq b_{(n)}$, respectively. Then, $\boldsymbol{a}$ is said to majorize $\boldsymbol{b}$, denoted by $\boldsymbol{a} \overset{m}{\succ} \boldsymbol{b}$, if $\sum_{j=1}^{i} a_{(j)} \leq \sum_{j=1}^{i} b_{(j)}$ for $i = 1, \ldots, n-1$, and $\sum_{j=1}^{n} a_{(j)} = \sum_{j=1}^{n} b_{(j)}$.*

DEFINITION 3: *We say that a real-valued function ϕ, defined on a set $\mathbb{A} \subseteq \mathbb{R}^n$, is Schur-convex on $\mathbb{A}$ if*

$$\boldsymbol{u} \overset{m}{\succeq} \boldsymbol{v} \Rightarrow \phi(\boldsymbol{u}) \geq \phi(\boldsymbol{v}) \quad \text{for any } \boldsymbol{u}, \boldsymbol{v} \in \mathbb{A}.$$

ϕ is said to be Schur-concave function on $\mathbb{A}$ if $-\phi$ is Schur-convex on $\mathbb{A}$.

A square matrix Π_n, of order n, is said to be a permutation matrix if each row and column has a single entry as 1, and all other entries as zero. Clearly, such matrices are obtained by interchanging rows (or columns) of the identity matrix I_n. The T-transform matrix has the form $T = wI_n + (1-w)\Pi_n$, where $0 \leq w \leq 1$, and Π_n is a permutation matrix. We say that two T-transform matrices $T_1 = w_1 I_n + (1-w_1)\Pi_1$ and $T_2 = w_2 I_n + (1-w_2)\Pi_2$ have the same structure if $\Pi_1 = \Pi_2$. It is well-known that a finite product of T-transform matrices with the same structure is also a T-transform matrix, while this

https://doi.org/10.1017/S0269964820000467 Published online by Cambridge University Press

product may not be a T-transform matrix if its elements do not have the same structure; see Balakrishnan *et al.* [2] for more details.

The following definition presents two generalizations of the vector majorization to the case of matrices.

DEFINITION 4: *Consider the* $m \times n$ *matrices* $U = \{u_{ij}\}$ *and* $V = \{v_{ij}\}$ *with rows* $\boldsymbol{u}_1, \ldots, \boldsymbol{u}_m$ *and* $\boldsymbol{v}_1, \ldots, \boldsymbol{v}_m$, *respectively. Then:*

(i) If $\boldsymbol{u}_i \overset{m}{\succeq} \boldsymbol{v}_i$, *for* $i = 1, \ldots, m$, *then we say that the matrix* U *is larger than the matrix* V *in row majorization and is denoted by* $U >^{row} V$;

(ii) If there exists a finite set of $n \times n$ T*-transform matrices* $T_1, \ldots, T_k$ *such that* $V = UT_1T_2 \ldots T_k$, *then we say that the matrix* U *is larger than matrix* V *in chain majorization and is denoted by* $U \gg V$.

It is well-known that chain majorization implies row majorization. Interested readers may refer to Marshall *et al.* [15] for an elaborate discussion on the theory of vector and matrix majorizations and their applications.

Let us set

$$\mathcal{S}_n = \left\{ (\boldsymbol{a}, \boldsymbol{b}) = \begin{pmatrix} a_1 \ldots a_n \\ b_1 \ldots b_n \end{pmatrix} : a_i, b_j > 0 \text{ and } (a_i - a_j)(b_i - b_j) \leq 0,\ i, j = 1, \ldots, n \right\}$$

and

$$\mathcal{H}_n = \left\{ (\boldsymbol{a}, \boldsymbol{b}) = \begin{pmatrix} a_1 \ldots a_n \\ b_1 \ldots b_n \end{pmatrix} : a_i, b_j > 0 \text{ and } (a_i - a_j)(b_i - b_j) \geq 0,\ i, j = 1, \ldots, n \right\}.$$

The proof of the two next lemmas that will be useful in establishing the main results in the subsequent sections are presented in the Appendix.

LEMMA 1: *Consider a differentiable function* $\phi : R_+^4 \longrightarrow R_+$. *Then,*

$$\phi(U) \geq (\leq) \phi(V) \quad \text{for all } U, V \text{ such that } U \in \mathcal{S}_2(\mathcal{H}_2), \text{ and } U \gg V \tag{1}$$

if and only if

(i) $\phi(U) = \phi(U\Pi)$ *for all permutation matrices* Π, *and for all* $U \in \mathcal{S}_2(\mathcal{H}_2)$;

(ii) $\sum_{i=1}^{2} (u_{ik} - u_{ij})[\phi_{ik}(U) - \phi_{ij}(U)] \geq (\leq) 0$ *for all* $j, k = 1, 2$, *and for all* $U \in \mathcal{S}_2(\mathcal{H}_2)$, *where* $\phi_{ij}(U) = \partial \phi(U) / \partial u_{ij}$.

LEMMA 2: *Consider a differentiable function* $\varphi : R_+^2 \to R_+$, *and let the function* $\phi_n : R_+^{2n} \to R_+$ *be defined as* $\phi_n(U) = \sum_{i=1}^{n} \varphi(u_{1i}, u_{2i})$. *If* ϕ_2 *satisfies (1), then* $\phi_n(U) \geq \phi_n(V)$, *where* $U \in \mathcal{S}_n(\mathcal{H}_n)$ *and* $V = UT$.

3. USUAL STOCHASTIC ORDER

Let $X_1, \ldots, X_n$ be a set of random variables with $X_i \sim LS(\mu_i, \lambda_i)$, $i = 1, \ldots, n$, and $V_n(\boldsymbol{p}, \boldsymbol{\mu}, \boldsymbol{\lambda})$ be a random variable representing a finite mixture of location-scale distributions

https://doi.org/10.1017/S0269964820000467 Published online by Cambridge University Press

of X_i's. Then, the survival function of $V_n(\boldsymbol{p}, \boldsymbol{\mu}, \boldsymbol{\lambda})$ are given by

$$\bar{F}_{V_n(\boldsymbol{p},\boldsymbol{\mu},\boldsymbol{\lambda})}(x) = \sum_{i=1}^{n} p_i \left[1 - F\left(\frac{x-\mu_i}{\lambda_i}\right)\right], \quad x > \max\{\mu_1, \ldots, \mu_n\},$$

where $0 < p_i < 1$ are the mixture proportions such that $\sum_{i=1}^{n} p_i = 1$, and $\mu_i > 0$, $\lambda_i > 0$, $i = 1, \ldots, n$. This section is divided into two subsections, with the first one dealing with stochastic comparison of location-scale mixture models when mixture proportions and threshold (location) parameters are different, and the second one dealing with the usual stochastic order of location-scale mixture models when the mixture proportions and scale parameters are different.

3.1. Case of Mixture Proportions and Threshold Parameters Being Different

In the following theorem, we show that if the matrix of parameters of one mixture model chain majorizes that of another mixture, then the first mixture is smaller than the second with respect to the usual stochastic order.

THEOREM 1: *Let* $\bar{F}_{V_2(\boldsymbol{p},\boldsymbol{\mu})}(x) = \sum_{i=1}^{2} p_i[1 - F((x-\mu_i)/\lambda)]$ *and* $\bar{F}_{W_2(\boldsymbol{q},\boldsymbol{\theta})}(x) = \sum_{i=1}^{2} q_i[1 - F((x-\theta_i)/\lambda)]$ *be survival functions of two mixtures corresponding to* $V_2(\boldsymbol{p}, \boldsymbol{\mu})$ *and* $W_2(\boldsymbol{q}, \boldsymbol{\theta})$, *respectively. Further, suppose* $f(t)$ *is a decreasing function of* t. *Then, for* $(\boldsymbol{p}, \boldsymbol{\mu}) \in \mathcal{H}_2$, *we have*

$$\begin{pmatrix} p_1 & p_2 \\ \mu_1 & \mu_2 \end{pmatrix} \gg \begin{pmatrix} q_1 & q_2 \\ \theta_1 & \theta_2 \end{pmatrix} \implies V_2(\boldsymbol{p}, \boldsymbol{\mu}) \geq_{st} W_2(\boldsymbol{q}, \boldsymbol{\theta}).$$

PROOF: Assume that

$$\bar{F}_{V_2(\boldsymbol{p},\boldsymbol{\mu})}(x) = \sum_{i=1}^{2} p_i \left[1 - F\left(\frac{x-\mu_i}{\lambda}\right)\right], \quad x \geq \max\{\mu_1, \mu_2\}.$$

To establish the desired result, we have to check Conditions (i) and (ii) of Lemma 1. Clearly, $V_2(\boldsymbol{p}, \boldsymbol{\mu})$ is permutation invariant on $\mathcal{H}_2$, for fixed $x > 0$, which confirms Condition (i). Now, for fixed $x > 0$ and $i \neq j$, consider the function φ

$$\begin{aligned}\varphi(\boldsymbol{p}, \boldsymbol{\mu}) = (p_1 - p_2) &\left(\frac{\partial \bar{F}_{V_2(\boldsymbol{p},\boldsymbol{\mu})}(x)}{\partial p_1} - \frac{\partial \bar{F}_{V_2(\boldsymbol{p},\boldsymbol{\mu})}(x)}{\partial p_2}\right) \\ &+ (\mu_1 - \mu_2)\left(\frac{\partial \bar{F}_{V_2(\boldsymbol{p},\boldsymbol{\mu})}(x)}{\partial \mu_1} - \frac{\partial \bar{F}_{V_2(\boldsymbol{p},\boldsymbol{\mu})}(x)}{\partial \mu_2}\right). \qquad (2)\end{aligned}$$

The partial derivative of $\bar{F}_{V_2(\boldsymbol{p},\boldsymbol{\mu})}(x)$ with respect to p_i and μ_i are

$$\frac{\partial \bar{F}_{V_2(\boldsymbol{p},\boldsymbol{\mu})}(x)}{\partial p_i} = 1 - F\left(\frac{x-\mu_i}{\lambda}\right) \quad \text{and} \quad \frac{\partial \bar{F}_{V_2(\boldsymbol{p},\boldsymbol{\mu})}(x)}{\partial \mu_i} = \frac{p_i}{\lambda} f\left(\frac{x-\mu_i}{\lambda}\right),$$

respectively. Now, upon substituting these expressions in (2), we obtain

$$\begin{aligned}\varphi(\boldsymbol{p}, \boldsymbol{\mu}) = (p_1 - p_2) &\left[F\left(\frac{x-\mu_2}{\lambda}\right) - F\left(\frac{x-\mu_1}{\lambda}\right)\right] \\ &+ \frac{1}{\lambda}(\mu_1 - \mu_2)\left[p_1 f\left(\frac{x-\mu_1}{\lambda}\right) - p_2 f\left(\frac{x-\mu_2}{\lambda}\right)\right]. \qquad (3)\end{aligned}$$

The assumption that $(\boldsymbol{p}, \boldsymbol{\mu}) \in \mathcal{H}_2$ implies that $(p_1 - p_2)(\mu_1 - \mu_2) \geq 0$, which means that either $p_1 \geq p_2$ and $\mu_1 \geq \mu_2$, or $p_2 \geq p_1$ and $\mu_2 \geq \mu_1$. We present the proof only for the case

https://doi.org/10.1017/S0269964820000467 Published online by Cambridge University Press

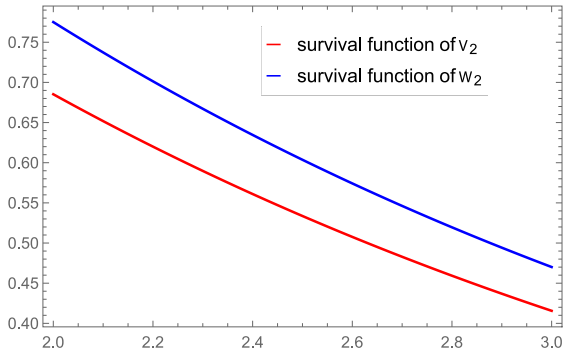


FIGURE **1.** Plots of survival functions of two finite mixture models when $\lambda = 2$.

when $p_1 \geq p_2$ and $\mu_1 \geq \mu_2$, since the proof for the other case is quite similar. Because $F(t)$ is an increasing function with respect to t, and $\mu_1 \geq \mu_2$, we readily observe that the first term on the right-hand side of (3) is positive. Moreover, according to the assumption that f is a decreasing function, the second term on the right-hand side of (3) is also positive. Upon combining these observations, we have $\varphi(\boldsymbol{p}, \boldsymbol{\mu}) \geq 0$, and so Condition (ii) of Lemma 1 is also satisfied. This completes the proof of the theorem. ■

We may question whether the result of Theorem 1 could hold if $(\boldsymbol{p}, \boldsymbol{\mu}) \notin \mathcal{H}_2$? The following discussion provides a counterexample. Let $(\boldsymbol{p}, \boldsymbol{\mu}) \in \mathcal{S}_2$. Then, we have $(p_1 - p_2)(\mu_1 - \mu_2) \leq 0$, which means that either $p_1 \geq p_2$ and $\mu_2 \geq \mu_1$, or $p_2 \geq p_1$ and $\mu_1 \geq \mu_2$. For example, consider the case when $p_1 \geq p_2$ and $\mu_2 \geq \mu_1$. In this case, the first term on the right-hand side of (3) is non-positive. Also, there exist two possibilities for the sign of $p_1 f((x-\mu_1)/\lambda) - p_2 f((x-\mu_2)/\lambda)$. If $p_1 f((x-\mu_1)/\lambda) - p_2 f((x-\mu_2)/\lambda) \leq 0$, then the second term on the right-hand side of (3) is non-negative. In this case, if λ is sufficiently small, then the magnitude of the second term will be more than the first term and so $V_2(\boldsymbol{p}, \boldsymbol{\mu}) \geq_{\mathrm{st}} W_2(\boldsymbol{q}, \boldsymbol{\theta})$. In the other case, if $p_1 f((x-\mu_1)/\lambda) - p_2 f((x-\mu_2)/\lambda) \geq 0$, then the second term on the right-hand side of (3) is non-positive and $\varphi(\boldsymbol{p}, \boldsymbol{\mu}) \leq 0$, which implies that $V_2(\boldsymbol{p}, \boldsymbol{\mu}) \leq_{\mathrm{st}} W_2(\boldsymbol{q}, \boldsymbol{\theta})$.

COUNTEREXAMPLE 1: *Let us consider the standard exponential as the baseline distribution. Set* $(\mu_1, \mu_2) = (1, 2)$, $(p_1, p_2) = (0.8, 0.2)$, $(\theta_1, \theta_2) = (1.4, 1.6)$, *and* $(q_1, q_2) = (0.56, 0.44)$. *We can easily observe that* $(\boldsymbol{q}, \boldsymbol{\theta}) = (\boldsymbol{p}, \boldsymbol{\mu}) T_{0.6}$ *which, according to Part (ii) of Definition 4, implies* $(\boldsymbol{p}, \boldsymbol{\mu}) \gg (\boldsymbol{q}, \boldsymbol{\theta})$. *The survival functions of the two mixture models when* $\lambda = 2$ *are plotted in Figure 1. It can be seen in Figure 1 that* $V_2(\boldsymbol{p}, \boldsymbol{\alpha}) \geq_{st} W_2(\boldsymbol{q}, \boldsymbol{\lambda})$. *We can also see the survival functions of the two mixture models when* $\lambda = 0.2$ *in Figure 2 wherein it can be seen that* $W_2(\boldsymbol{p}, \boldsymbol{\alpha}) \geq_{st} V_2(\boldsymbol{q}, \boldsymbol{\theta})$. *Thus, in general, the result of Theorem 1 could hold if* $(\boldsymbol{p}, \boldsymbol{\mu}) \notin \mathcal{H}_2$, *but it would depend on the value of* λ.

Some generalizations of the result in Theorem 1 to the case when the number of underlying mixing variables is arbitrary are presented below.

THEOREM 2: *Let* $\bar{F}_{V_n(\boldsymbol{p},\boldsymbol{\mu})}(x) = \sum_{i=1}^{n} p_i[1 - F((x-\mu_i)/\lambda)]$ *and* $\bar{F}_{W_n(\boldsymbol{q},\boldsymbol{\theta})}(x) = \sum_{i=1}^{n} q_i[1 - F((x-\theta_i)/\lambda)]$ *be survival functions of two finite mixtures corresponding to* $V_n(\boldsymbol{p}, \boldsymbol{\mu})$ *and* $W_n(\boldsymbol{q}, \boldsymbol{\theta})$, *respectively. Further, suppose* $f(t)$ *is a decreasing function of* t. *Then, for* $(\boldsymbol{p}, \boldsymbol{\mu}) \in$

https://doi.org/10.1017/S0269964820000467 Published online by Cambridge University Press

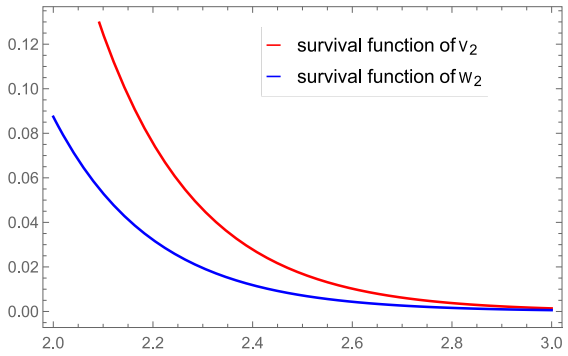


FIGURE **2.** Plots of survival functions of two finite mixture models when $\lambda = 0.2$.

$\mathcal{H}_n$, *we have*

$$\begin{pmatrix} q_1 & \dots & q_n \\ \theta_1 & \dots & \theta_n \end{pmatrix} = \begin{pmatrix} p_1, & \dots & p_n \\ \mu_1 & \dots & \mu_n \end{pmatrix} T \implies V_n(\boldsymbol{p}, \boldsymbol{\mu}) \geq_{st} W_n(\boldsymbol{q}, \boldsymbol{\theta}).$$

PROOF: Setting $\phi_n(\boldsymbol{p}, \boldsymbol{\mu}) = \bar{F}_{V_n(\boldsymbol{p},\boldsymbol{\mu})}(x)$ and $\varphi(p, \mu) = p(1 - F((x - \mu)/\lambda))$, for fixed $x > \mu$, we then have $\phi_n(\boldsymbol{p}, \boldsymbol{\mu}) = \sum_{i=1}^{n} \varphi(p_i, \mu_i) = \sum_{i=1}^{n} p_i(1 - F((x - \mu_i)/\lambda))$. According to Theorem 1, ϕ_2 satisfies (1). Now, the required result follows immediately from Lemma 2. ■

Theorem 2 shows that higher heterogeneity of matrix of parameters implies the survival function of the corresponding mixture model to be higher.

The following corollary is a direct consequence of Theorem 1 and the fact that a finite product of T-transform matrices with the same structure is also a T-transform matrix.

COROLLARY 1: *Let* $\bar{F}_{V_n(\boldsymbol{p},\boldsymbol{\mu})}(x) = \sum_{i=1}^{n} p_i[1 - F((x - \mu_i)/\lambda)]$ *and* $\bar{F}_{W_n(\boldsymbol{q},\boldsymbol{\theta})}(x) = \sum_{i=1}^{n} q_i[1 - F((x - \theta_i)/\lambda)]$ *be survival functions of two finite mixtures corresponding to* $V_n(\boldsymbol{p}, \boldsymbol{\mu})$ *and* $W_n(\boldsymbol{q}, \boldsymbol{\theta})$, *respectively. Further, suppose* $f(t)$ *is a decreasing function with respect to* t. *If the* T*-transform matrices* $T_1, \dots, T_k$ *have the same structure, then for* $(\boldsymbol{p}, \boldsymbol{\mu}) \in \mathcal{H}_n$, *we have*

$$\begin{pmatrix} q_1 & \dots & q_n \\ \theta_1 & \dots & \theta_n \end{pmatrix} = \begin{pmatrix} p_1 & \dots & p_n \\ \mu_1 & \dots & \mu_n \end{pmatrix} T_1 \dots T_k \implies V_n(\boldsymbol{p}, \boldsymbol{\mu}) \geq_{st} W_n(\boldsymbol{q}, \boldsymbol{\theta}).$$

The following example provides an illustration of the result in Corollary 1.

EXAMPLE 1: *Suppose* $(\mu_1, \mu_2, \mu_3) = (5, 4, 3)$, $(\theta_1, \theta_2, \theta_3) = (3.6, 4.04, 4.36)$, $(p_1, p_2, p_3) = (0.1, 0.4, 0.5)$, *and* $(q_1, q_2, q_3) = (0.42, 0.3, 0.28)$. *Consider* T*-transform matrices* T_1 *and* T_2 *(with the same structure) as follows:*

$$T_1 = \begin{pmatrix} 0.2 & 0 & 0.8 \\ 0.8 & 0.2 & 0 \\ 0 & 0.8 & 0.2 \end{pmatrix} \quad \textit{and} \quad T_2 = \begin{pmatrix} 0.4 & 0 & 0.6 \\ 0.6 & 0.4 & 0 \\ 0 & 0.6 & 0.4 \end{pmatrix}.$$

Then, it is easy to observe that $(\boldsymbol{q}, \boldsymbol{\theta}) = (\boldsymbol{p}, \boldsymbol{\mu})T_1T_2$ *which, based on Corollary 1, readily yields* $V_3(\boldsymbol{p}, \boldsymbol{\mu}) \geq_{st} W_3(\boldsymbol{q}, \boldsymbol{\theta})$.

It will be of interest to know whether the results of Corollary 1 will still hold if the matrices T_i, $i = 1, \dots, k$, do not have the same structure. The following corollary gives an

https://doi.org/10.1017/S0269964820000467 Published online by Cambridge University Press

answer to this question, by repeating the result of Theorem 2 for the matrices $(\boldsymbol{p},\boldsymbol{\mu})T_1\dots T_i$, $i=1,\dots,k-1$.

COROLLARY 2: *Let* $\bar{F}_{V_n(\boldsymbol{p},\boldsymbol{\mu})}(x)=\sum_{i=1}^n p_i[1-F((x-\mu_i)/\lambda)]$ *and* $\bar{F}_{W_n(\boldsymbol{q},\boldsymbol{\theta})}(x)=\sum_{i=1}^n q_i[1-F((x-\theta_i)/\lambda)]$ *be survival functions of two finite mixtures corresponding to* $V_n(\boldsymbol{p},\boldsymbol{\mu})$ *and* $W_n(\boldsymbol{q},\boldsymbol{\theta})$, *respectively. Further, suppose* $f(t)$ *is a decreasing function of* t. *If* $(\boldsymbol{p},\boldsymbol{\mu})\in\mathcal{H}_n$ *and* $(\boldsymbol{p},\boldsymbol{\mu})T_1\cdots T_i\in\mathcal{H}_n$ *for* $i=1,\dots,k-1$, *where* $k\geq 2$, *we have*

$$\begin{pmatrix} q_1 & \dots & q_n \\ \theta_1 & \dots & \theta_n \end{pmatrix} = \begin{pmatrix} p_1 & \dots & p_n \\ \mu_1 & \dots & \mu_n \end{pmatrix} T_1\dots T_k \implies V_n(\boldsymbol{p},\boldsymbol{\mu}) \geq_{st} W_n(\boldsymbol{q},\boldsymbol{\theta}).$$

The following example provides an illustration of the result in Corollary 2.

EXAMPLE 2: *Suppose* $(\mu_1,\mu_2,\mu_3)=(6,4,2)$, $(\theta_1,\theta_2,\theta_3)=(0.344,8.24,3.416)$, $(p_1,p_2,p_3)=(0.1,0.3,0.6)$, *and* $(q_1,q_2,q_3)=(0.0404,0.554,0.4056)$. *Consider the* T*-transform matrices* T_1, T_2, *and* T_3 *as follows:*

$$T_1=\begin{pmatrix} 1 & 0 & 0 \\ 0 & 0.4 & 0.6 \\ 0 & 0.6 & 0.4 \end{pmatrix},\quad T_2=\begin{pmatrix} 0.2 & 0.8 & 0 \\ 0.8 & 0.2 & 0 \\ 0 & 0 & 1 \end{pmatrix},\quad \text{and}\quad T_3=\begin{pmatrix} 0.1 & 0 & 0.9 \\ 0 & 1 & 0 \\ 0 & 0.9 & 0.1 \end{pmatrix}.$$

It is then easy to observe that $(\boldsymbol{p},\boldsymbol{\mu})$, $(\boldsymbol{p},\boldsymbol{\mu})T_1$, *and* $(\boldsymbol{p},\boldsymbol{\mu})T_1T_2$ *are all in* $\mathcal{H}_n$ *and* $(\boldsymbol{q},\boldsymbol{\theta})=(\boldsymbol{p},\boldsymbol{\mu})T_1T_2T_3$ *which, based on Corollary 2, readily yields* $V_3(\boldsymbol{p},\boldsymbol{\mu})\geq_{st} W_3(\boldsymbol{q},\boldsymbol{\theta})$.

3.2. Case of Mixture Proportions and Scale Parameters Being Different

THEOREM 3: *Let* $\bar{F}_{Z_2(\boldsymbol{p},\boldsymbol{\lambda})}(x)=\sum_{i=1}^2 p_i[1-F((x-\mu)/\lambda_i)]$ *and* $\bar{F}_{Y_2(\boldsymbol{q},\boldsymbol{\beta})}(x)=\sum_{i=1}^2 q_i[1-F((x-\mu)/\beta_i)]$ *be survival functions of two finite mixtures corresponding to* $Z_2(\boldsymbol{p},\boldsymbol{\lambda})$ *and* $Y_2(\boldsymbol{q},\boldsymbol{\beta})$, *respectively. Further, suppose* $f(t)$ *is a decreasing function of* t. *Then, for* $(\boldsymbol{p},\boldsymbol{\lambda}^{-1})\in\mathcal{S}_2$, *we have*

$$\begin{pmatrix} p_1 & p_2 \\ \lambda_1^{-1} & \lambda_2^{-1} \end{pmatrix} \gg \begin{pmatrix} q_1 & q_2 \\ \beta_1^{-1} & \beta_2^{-1} \end{pmatrix} \implies Z_2(\boldsymbol{p},\boldsymbol{\lambda}) \geq_{st} Y_2(\boldsymbol{q},\boldsymbol{\beta}).$$

PROOF: Assume that $\alpha_i=\lambda_i^{-1}$, for $i=1,2$. Then, we can rewrite $\bar{F}_{Z_2(\boldsymbol{p},\boldsymbol{\lambda})}(x)$ as follows:

$$\bar{F}_{Z_2(\boldsymbol{p},\boldsymbol{\alpha})}(x)=\sum_{i=1}^2 p_i[1-F(\alpha_i(x-\mu))],\quad x\geq\mu.$$

To establish the desired result, we have to check Conditions (i) and (ii) of Lemma 1. Clearly, $Z_2(\boldsymbol{p},\boldsymbol{\alpha})$ is permutation invariant on $\mathcal{S}_2$, for fixed $x>0$, which confirms Condition (i). Now, for fixed $x>0$ and $i\neq j$, consider the function

$$\begin{aligned}\Phi(\boldsymbol{p},\boldsymbol{\alpha}) &= (p_1-p_2)\left(\frac{\partial\bar{F}_{Z_2(\boldsymbol{p},\boldsymbol{\alpha})}(x)}{\partial p_1}-\frac{\partial\bar{F}_{Z_2(\boldsymbol{p},\boldsymbol{\alpha})}(x)}{\partial p_2}\right)\\ &\quad+(\alpha_1-\alpha_2)\left(\frac{\partial\bar{F}_{Z_2(\boldsymbol{p},\boldsymbol{\alpha})}(x)}{\partial\alpha_1}-\frac{\partial\bar{F}_{Z_2(\boldsymbol{p},\boldsymbol{\alpha})}(x)}{\partial\alpha_2}\right).\end{aligned}\tag{4}$$

The partial derivatives of $\bar{F}_{Z_2(\boldsymbol{p},\boldsymbol{\alpha})}(x)$ with respect to p_i and α_i are

$$\frac{\partial\bar{F}_{Z_2(\boldsymbol{p},\boldsymbol{\alpha})}(x)}{\partial p_i}=1-F(\alpha_i(x-\mu))\quad\text{and}\quad\frac{\partial\bar{F}_{Z_2(\boldsymbol{p},\boldsymbol{\alpha})}(x)}{\partial\alpha_i}=-(x-\mu)p_i\,f(\alpha_i(x-\mu)),$$

https://doi.org/10.1017/S0269964820000467 Published online by Cambridge University Press

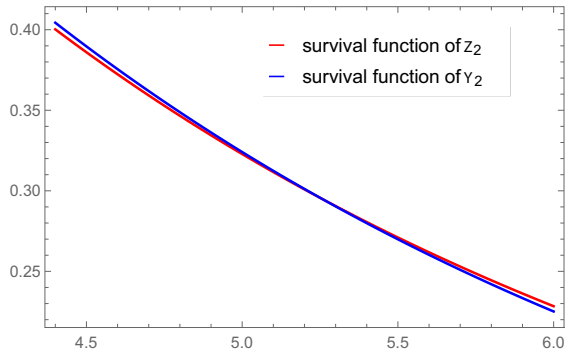


FIGURE **3.** Plots of survival functions of $Z_2(\boldsymbol{p}, \boldsymbol{\lambda})$ and $Y_2(\boldsymbol{q}, \boldsymbol{\beta})$ when $\mu = 2$.

respectively. Now, upon substituting these expressions in (4), we obtain

$$\begin{aligned}\Phi(\boldsymbol{p}, \boldsymbol{\alpha}) = (p_1 - p_2)[F(\alpha_2(x-\mu)) - F(\alpha_1(x-\mu))] \\ + (x-\mu)(\alpha_1 - \alpha_2)[p_2 f(\alpha_2(x-\mu)) - p_1 f(\alpha_1(x-\mu))]. \end{aligned} \tag{5}$$

The assumption that $(\boldsymbol{p}, \boldsymbol{\alpha}) \in \mathcal{S}_2$ implies that either $(p_1 - p_2)(\alpha_1 - \alpha_2) \leq 0$, which means that $p_1 \geq p_2$ and $\alpha_1 \leq \alpha_2$, or $p_2 \geq p_1$ and $\alpha_2 \leq \alpha_1$. We present the proof only for the case when $p_1 \geq p_2$ and $\alpha_1 \leq \alpha_2$, since the other case can be proved similarly. Because $F(t)$ is an increasing function of t, and $\alpha_1 \leq \alpha_2$, we readily observe that the first term on the right-hand side of (5) is positive. Furthermore, according to the assumption that f is a decreasing function, the second term on the right-hand side of (5) is also positive. Upon combining these two observations, we have $\Phi(\boldsymbol{p}, \boldsymbol{\alpha}) \geq 0$, and so Condition (ii) is also satisfied. This completes the proof of the theorem. ■

The following counterexample shows that the result in Theorem 3 is not necessarily true if $(\boldsymbol{p}, \boldsymbol{\lambda}^{-1}) \notin \mathcal{S}_2$.

COUNTEREXAMPLE 2: *Let us consider the standard exponential distribution as the baseline distribution for the location-scale model. Set* $(p_1, p_2) = (0.4, 0.6)$, $(q_1, q_2) = (0.44, 56)$, $(\lambda_1^{-1}, \lambda_2^{-1}) = (0.25, 0.5)$, *and* $(\beta_1^{-1}, \beta_2^{-1}) = (0.3, 0.45)$. *It is then easy to check that* $(\boldsymbol{q}, \boldsymbol{\beta}^{-1}) = (\boldsymbol{p}, \boldsymbol{\lambda}^{-1})T_{0.8}$ *which, according to Part (ii) of Definition 4, implies* $(\boldsymbol{p}, \boldsymbol{\lambda}^{-1}) \gg (\boldsymbol{q}, \boldsymbol{\beta}^{-1})$. *The survival functions of the two mixture models when* $\mu = 2$ *are plotted in Figure 3. It can be seen in Figure 3 that these survival functions do cross, which means that* $Z_2(\boldsymbol{p}, \boldsymbol{\lambda}) \not\geq_{st} Y_2(\boldsymbol{q}, \boldsymbol{\beta})$.

Some generalizations of the result in Theorem 3 to the case when the number of underlying mixing variables is arbitrary are presented below.

THEOREM 4: *Let* $\bar{F}_{Z_n(\boldsymbol{p}, \boldsymbol{\lambda})}(x) = \sum_{i=1}^n p_i[1 - F((x-\mu)/\lambda_i)]$ *and* $\bar{F}_{Y_n(\boldsymbol{q}, \boldsymbol{\beta})}(x) = \sum_{i=1}^n q_i[1 - F((x-\mu)/\beta_i)]$ *be survival functions of two finite mixtures corresponding to* $Z_n(\boldsymbol{p}, \boldsymbol{\lambda})$ *and* $Y_n(\boldsymbol{q}, \boldsymbol{\beta})$, *respectively. Suppose* $f(t)$ *is a decreasing function of* t. *Then, for* $(\boldsymbol{p}, \boldsymbol{\lambda}^{-1}) \in \mathcal{S}_n$, *we have*

$$\begin{pmatrix} q_1 & \cdots & q_n \\ \beta_1^{-1} & \cdots & \beta_n^{-1} \end{pmatrix} = \begin{pmatrix} p_1 & \cdots & p_n \\ \lambda_1^{-1} & \cdots & \lambda_n^{-1} \end{pmatrix} T \implies Z_n(\boldsymbol{p}, \boldsymbol{\lambda}) \geq_{st} Y_n(\boldsymbol{q}, \boldsymbol{\beta}).$$

https://doi.org/10.1017/S0269964820000467 Published online by Cambridge University Press

PROOF: The proof is similar to that of Theorem 2 and is, therefore, omitted here for the sake of brevity. ■

Theorem 4 shows that higher heterogeneity of matrix of parameters implies the survival function of the corresponding mixture model to be higher.

The following corollary is a direct consequence of Theorem 4 and the fact that a finite product of T-transform matrices with the same structure is also a T-transform matrix.

COROLLARY 3: *Let* $\bar{F}_{Z_n(\boldsymbol{p},\boldsymbol{\lambda})}(x) = \sum_{i=1}^{n} p_i[1 - F((x-\mu)/\lambda_i)]$ *and* $\bar{F}_{Y_n(\boldsymbol{q},\boldsymbol{\beta})}(x) = \sum_{i=1}^{n} q_i[1 - F((x-\mu)/\beta_i)]$ *be survival functions of two finite mixtures corresponding to* $Z_n(\boldsymbol{p},\boldsymbol{\lambda})$ *and* $Y_n(\boldsymbol{q},\boldsymbol{\beta})$*, respectively. Further, suppose* $f(t)$ *is a decreasing function of* t*. If the* T*-transform matrices* $T_1, \ldots, T_k$ *have the same structure, then for* $(\boldsymbol{p}, \boldsymbol{\lambda}^{-1}) \in \mathcal{S}_n$*, we have*

$$\begin{pmatrix} q_1 & \cdots & q_n \\ \beta_1^{-1} & \cdots & \beta_n^{-1} \end{pmatrix} = \begin{pmatrix} p_1 & \cdots & p_n \\ \lambda_1^{-1} & \cdots & \lambda_n^{-1} \end{pmatrix} T_1 \ldots T_k \implies Z_n(\boldsymbol{p},\boldsymbol{\lambda}) \geq_{st} Y_n(\boldsymbol{q},\boldsymbol{\beta}).$$

It will be of interest to know whether the results of Corollary 3 will still hold if the matrices T_i, $i = 1, \ldots, k$, do not have the same structure. The following corollary gives an answer to this question.

COROLLARY 4: *Let* $\bar{F}_{Z_n(\boldsymbol{p},\boldsymbol{\lambda})}(x) = \sum_{i=1}^{n} p_i[1 - F((x-\mu)/\lambda_i)]$ *and* $\bar{F}_{Y_n(\boldsymbol{q},\boldsymbol{\beta})}(x) = \sum_{i=1}^{n} q_i[1 - F((x-\mu)/\beta_i)]$ *be survival functions of two finite mixtures corresponding to* $Z_n(\boldsymbol{p},\boldsymbol{\lambda})$ *and* $Y_n(\boldsymbol{q},\boldsymbol{\beta})$*, respectively. Further, suppose* $f(t)$ *is a decreasing function of* t*. If* $(\boldsymbol{p}, \boldsymbol{\lambda}^{-1}) \in \mathcal{S}_n$ *and* $(\boldsymbol{p}, \boldsymbol{\lambda}^{-1})T_1 \cdots T_i \in \mathcal{S}_n$ *for* $i = 1, \ldots, k-1$*, where* $k \geq 2$*, we have*

$$\begin{pmatrix} q_1 & \cdots & q_n \\ \beta_1^{-1} & \cdots & \beta_n^{-1} \end{pmatrix} = \begin{pmatrix} p_1 & \cdots & p_n \\ \lambda_1^{-1} & \cdots & \lambda_n^{-1} \end{pmatrix} T_1 \ldots T_k \implies Z_n(\boldsymbol{p},\boldsymbol{\lambda}) \geq_{st} Y_n(\boldsymbol{q},\boldsymbol{\beta}).$$

The following example provides an illustration of the result in Corollary 4.

EXAMPLE 3: *Suppose* $(\lambda_1^{-1}, \lambda_2^{-1}, \lambda_3^{-1}) = (6, 4, 2)$, $(\beta_1^{-1}, \beta_2^{-1}, \beta_3^{-1}) = (0.344, 8.24, 3.416)$, $(p_1, p_2, p_3) = (0.1, 0.3, 0.6)$, *and* $(q_1, q_2, q_3) = (0.0404, 0.554, 0.4056)$. *Consider the* T*-transform matrices* T_1*,* T_2*, and* T_3 *as follows:*

$$T_1 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0.4 & 0.6 \\ 0 & 0.6 & 0.4 \end{pmatrix}, \quad T_2 = \begin{pmatrix} 0.2 & 0.8 & 0 \\ 0.8 & 0.2 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \quad \text{and} \quad T_3 = \begin{pmatrix} 0.1 & 0 & 0.9 \\ 0 & 1 & ,0 \\ 0 & 0.9 & 0.1 \end{pmatrix}.$$

It is then easy to observe that $(\boldsymbol{p}, \boldsymbol{\lambda}^{-1})$*,* $(\boldsymbol{p}, \boldsymbol{\lambda}^{-1})T_1$*, and* $(\boldsymbol{p}, \boldsymbol{\lambda}^{-1})T_1T_2$ *are all in* $\mathcal{S}_3$ *and* $(\boldsymbol{q}, \boldsymbol{\lambda}) = (\boldsymbol{p}, \boldsymbol{\lambda}^{-1})T_1T_2T_3$ *which, based on Corollary 4, readily yields* $Z_3(\boldsymbol{p},\boldsymbol{\lambda}) \geq_{st} Y_3(\boldsymbol{q},\boldsymbol{\beta})$.

4. REVERSED HAZARD RATE AND HAZARD RATE ORDERS

In this section, we strengthen the result in Theorem 1 from the usual stochastic order to the reversed hazard rate order. The concept of chain majorization is a way of comparing two matrices of same dimension in terms of dispersion of their components. Theorem 5 below shows that if the matrix of parameters $(\boldsymbol{p}, \boldsymbol{\mu})$ is more dispersive than the matrix of parameters $(\boldsymbol{q}, \boldsymbol{\theta})$, then the first mixture model is larger than the second mixture model with respect to the reversed hazard rate order.

https://doi.org/10.1017/S0269964820000467 Published online by Cambridge University Press

THEOREM 5: *Let* $F_{V_2(\boldsymbol{p},\boldsymbol{\mu})}(x) = \sum_{i=1}^{2} p_i F((x-\mu_i)/\lambda)$ *and* $F_{W_2(\boldsymbol{q},\boldsymbol{\theta})}(x) = \sum_{i=1}^{2} q_i F((x-\theta_i)/\lambda)$ *be distribution functions of two finite mixtures corresponding to* $V_2(\boldsymbol{p},\boldsymbol{\mu})$ *and* $W_2(\boldsymbol{q},\boldsymbol{\theta})$, *respectively. Further, suppose* $f(t)$ *is a decreasing and convex function of* t. *Then, for* $(\boldsymbol{p},\boldsymbol{\mu}) \in \mathcal{H}_2$, *we have*

$$\begin{pmatrix} p_1 & p_2 \\ \mu_1 & \mu_2 \end{pmatrix} \gg \begin{pmatrix} q_1 & q_2 \\ \theta_1 & \theta_2 \end{pmatrix} \implies V_2(\boldsymbol{p},\boldsymbol{\mu}) \geq_{\text{rh}} W_2(\boldsymbol{q},\boldsymbol{\theta}).$$

PROOF: The reversed hazard rate function of $V_2(\boldsymbol{p},\boldsymbol{\mu})$ is given by

$$\tilde{r}(x;\boldsymbol{p},\boldsymbol{\mu}) = \frac{1}{\lambda} \frac{p_1 f\left(\frac{x-\mu_1}{\lambda}\right) + p_2 f\left(\frac{x-\mu_2}{\lambda}\right)}{p_1 F\left(\frac{x-\mu_1}{\lambda}\right) + p_2 F\left(\frac{x-\mu_2}{\lambda}\right)}.$$

To establish the desired result, we have to check Conditions (i) and (ii) of Lemma 1. Clearly, $r(x;\boldsymbol{p},\boldsymbol{\mu})$ is permutation invariant on $\mathcal{H}_2$, for fixed $x > 0$, which confirms Condition (i). Next, for fixed $x > 0$ and $i \neq j$, consider the function

$$\begin{aligned} \psi(\boldsymbol{p},\boldsymbol{\mu}) = (p_1 - p_2) &\left(\frac{\partial \tilde{r}(x;\boldsymbol{p},\boldsymbol{\mu})}{\partial p_1} - \frac{\partial \tilde{r}(x;\boldsymbol{p},\boldsymbol{\mu})}{\partial p_2} \right) \\ &+ (\mu_1 - \mu_2) \left(\frac{\partial \tilde{r}(x;\boldsymbol{p},\boldsymbol{\mu})}{\partial \mu_1} - \frac{\partial \tilde{r}(x;\boldsymbol{p},\boldsymbol{\mu})}{\partial \mu_2} \right). \end{aligned} \tag{6}$$

The partial derivatives of $\tilde{r}(x;\boldsymbol{p},\boldsymbol{\mu})$ with respect to p_1, p_2, μ_1, and μ_2 are given by

$$\begin{aligned} \frac{\partial \tilde{r}(x;\boldsymbol{p},\boldsymbol{\mu})}{\partial p_1} &\overset{\text{sgn}}{=} p_2 f\left(\frac{x-\mu_1}{\lambda}\right) F\left(\frac{x-\mu_2}{\lambda}\right) - p_2 f\left(\frac{x-\mu_2}{\lambda}\right) F\left(\frac{x-\mu_1}{\lambda}\right), \\ \frac{\partial \tilde{r}(x;\boldsymbol{p},\boldsymbol{\mu})}{\partial p_2} &\overset{\text{sgn}}{=} p_1 f\left(\frac{x-\mu_2}{\lambda}\right) F\left(\frac{x-\mu_1}{\lambda}\right) - p_1 f\left(\frac{x-\mu_1}{\lambda}\right) F\left(\frac{x-\mu_2}{\lambda}\right), \end{aligned}$$

and

$$\begin{aligned} \frac{\partial \tilde{r}(x;\boldsymbol{p},\boldsymbol{\mu})}{\partial \mu_1} &\overset{\text{sgn}}{=} p_1^2 f^2\left(\frac{x-\mu_1}{\lambda}\right) + p_1 p_2 f\left(\frac{x-\mu_1}{\lambda}\right) f\left(\frac{x-\mu_2}{\lambda}\right) \\ &\quad - p_1^2 f'\left(\frac{x-\mu_1}{\lambda}\right) F\left(\frac{x-\mu_1}{\lambda}\right) - p_1 p_2 f'\left(\frac{x-\mu_1}{\lambda}\right) F\left(\frac{x-\mu_2}{\lambda}\right), \\ \frac{\partial \tilde{r}(x;\boldsymbol{p},\boldsymbol{\mu})}{\partial \mu_2} &\overset{\text{sgn}}{=} p_2^2 f^2\left(\frac{x-\mu_2}{\lambda}\right) + p_1 p_2 f\left(\frac{x-\mu_2}{\lambda}\right) f\left(\frac{x-\mu_1}{\lambda}\right) \\ &\quad - p_2^2 f'\left(\frac{x-\mu_2}{\lambda}\right) F\left(\frac{x-\mu_2}{\lambda}\right) - p_1 p_2 f'\left(\frac{x-\mu_2}{\lambda}\right) F\left(\frac{x-\mu_1}{\lambda}\right), \end{aligned}$$

where $a \overset{\text{sgn}}{=} b$ means that a and b have same sign. The assumption that $(\boldsymbol{p},\boldsymbol{\mu}) \in \mathcal{H}_2$ implies that either $(p_1 - p_2)(\mu_1 - \mu_2) \geq 0$, which means that $p_1 \geq p_2$ and $\mu_1 \geq \mu_2$, or $p_2 \geq p_1$ and $\mu_2 \geq \mu_1$. We present the proof only for the case when $p_1 \geq p_2$ and $\mu_1 \geq \mu_2$, since the proof

https://doi.org/10.1017/S0269964820000467 Published online by Cambridge University Press

for the other case is quite similar. Let us set $z_1 = (x-\mu_1)/\lambda$ and $z_2 = (x-\mu_2)/\lambda$. Then,

$$\frac{\partial \tilde{r}(x;\boldsymbol{p},\boldsymbol{\mu})}{\partial p_1} - \frac{\partial \tilde{r}(x;\boldsymbol{p},\boldsymbol{\mu})}{\partial p_2} \overset{\text{sgn}}{=} (f(z_1)-f(z_2))(p_1F(z_1)+p_2F(z_2)) + (F(z_2)-F(z_1))(p_1f(z_1)+p_2f(z_2))\}.$$

Since $\mu_1 \geq \mu_2$, we have $z_1 \leq z_2$ and so $F(z_1) \leq F(z_2)$. The decreasing property of f implies that $f(z_1) \geq f(z_2)$. Upon combining these observations, we can conclude that

$$\frac{\partial \tilde{r}(x;\boldsymbol{p},\boldsymbol{\mu})}{\partial p_1} - \frac{\partial \tilde{r}(x;\boldsymbol{p},\boldsymbol{\mu})}{\partial p_2} \geq 0.$$

Now, the assumption $p_1 \geq p_2$ implies that the first term on the right-hand side of (6) is non-negative. Moreover, we have

$$\frac{\partial \tilde{r}(x;\boldsymbol{p},\boldsymbol{\mu})}{\partial \mu_1} - \frac{\partial \tilde{r}(x;\boldsymbol{p},\boldsymbol{\mu})}{\partial \mu_2} \overset{\text{sgn}}{=} (p_1F(z_1)+p_2F(z_2))(p_2f'(z_2)-p_1f'(z_1)) + (p_1f(z_1)+p_2f(z_2))(p_1f(z_1)-p_2f(z_2)).$$

If f is decreasing, then $p_1f(z_1) \geq p_2f(z_2)$ and if f' is decreasing and convex, then we have $p_1f'(z_1) \leq p_2f'(z_2)$, which imply that

$$\frac{\partial \tilde{r}(x;\boldsymbol{p},\boldsymbol{\mu})}{\partial \mu_1} - \frac{\partial \tilde{r}(x;\boldsymbol{p},\boldsymbol{\mu})}{\partial \mu_2} \geq 0.$$

Now, the assumption $\mu_1 \geq \mu_2$ implies that the second term on the right-hand side of (6) is non-negative. We can, thus, conclude that $\psi(\boldsymbol{p},\boldsymbol{\mu}) \geq 0$, and the proof of the theorem, thus, gets completed. ■

The following example provides an illustration of the result in Theorem 5, which shows that the curve of the reversed hazard rate function of $V_2(\boldsymbol{p},\boldsymbol{\mu})$ lies above that of $W_2(\boldsymbol{q},\boldsymbol{\theta})$.

EXAMPLE 4: *Let us consider the standard exponential as the baseline distribution. Set* $(\mu_1,\mu_2)=(2,5)$, $(p_1,p_2)=(0.3,0.7)$, $(\theta_1,\theta_2)=(3.8,3.2)$, $(q_1,q_2)=(0.54,0.46)$, *and* $\lambda=2$. *We can easily observe that* $(\boldsymbol{p},\boldsymbol{\mu}) \in \mathcal{H}_2$ *and* $(\boldsymbol{q},\boldsymbol{\theta}) = (\boldsymbol{p},\boldsymbol{\mu})T_{0.4}$ *which, according to Part (ii) of Definition 4, implies* $(\boldsymbol{p},\boldsymbol{\mu}) \gg (\boldsymbol{q},\boldsymbol{\theta})$. *The reversed hazard rate functions of these two mixture models are plotted in Figure 4. It can be seen from Figure 4 that* $V_2(\boldsymbol{p},\boldsymbol{\mu}) \geq_{\text{hr}} W_2(\boldsymbol{q},\boldsymbol{\theta})$.

REMARK 1: *It should be mentioned that the condition "*$f(x)$ *is a decreasing and convex function of* x*" in Theorem 5 is quite general and is easy to verify for many family of distributions. For example, we consider the following cases.*

(i) Truncated normal distribution: The density function of the truncated standard normal distribution is

$$f(x) = \frac{1}{1-\Phi(1)+\Phi(-1)}\frac{1}{\sqrt{2\pi}}e^{-x^2/2}, \quad x \in \mathbb{R}-[-1,1],$$

where Φ *is the cumulative distribution function of standard normal. The first two derivatives of* $f(x)$ *with respect to* x, *respectively, are*

$$f'(x) \overset{sgn}{=} -xe^{-x^2/2} \quad \text{and} \quad f''(x) \overset{sgn}{=} -(1-x^2)\,e^{-x^2/2},$$

which confirms that $f(x)$ *is a decreasing and convex function of* x.

https://doi.org/10.1017/S0269964820000467 Published online by Cambridge University Press

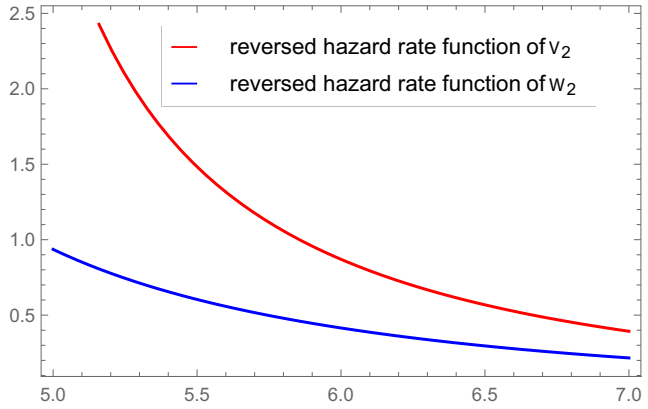


FIGURE **4.** Plots of reversed hazard rate functions of two finite mixture models when $\lambda = 2$.

(ii) Feller-Pareto distribution function: In this case, we have

$$f(x) \propto x^{\beta/\gamma-1}[1+x^{1/\gamma}]^{-(\alpha+\beta)}, \quad x>0, \alpha>0, \beta>0, \gamma>0.$$

First, it can be observed that

$$f'(x) \overset{sgn}{=} \left(\frac{\beta}{\gamma}-1\right) x^{\beta/\gamma-2}(1+x^{1/\gamma})^{-(\alpha+\beta)} - \frac{1}{\gamma}(\alpha+\beta)x^{1/\gamma+\beta/\gamma-1}(1+x^{1/\gamma})^{-(\alpha+\beta)-1}$$

is decreasing in x, for $\beta \le \gamma$. Next, after some minor calculations, we find

$$\begin{aligned} f''(x) \overset{sgn}{=} & \left(\frac{\beta}{\gamma}-1\right) x^{\beta/\gamma-3}(1+x^{1/\gamma})^{-(\alpha+\beta)} \left[\left(\frac{\beta}{\gamma}-2\right) - \frac{1}{\gamma}(\alpha+\beta)x^{1/\gamma}[1+x^{1/\gamma}]^{-1}\right] \\ & - \frac{1}{\gamma}(\alpha+\beta)x^{\beta/\gamma+1/\gamma-3}(1+x^{1/\gamma})^{-(\alpha+\beta)-2} \\ & \times \left[\left(\frac{\beta}{\gamma}+\frac{1}{\gamma}-2\right)[1+x^{1/\gamma}]^{-1} - \frac{1}{\gamma^2}(\alpha+\beta+1)x^{1/\gamma}\right]. \end{aligned} \tag{7}$$

So, under the conditions $\beta \le \alpha$ and $\beta+1 \le 2\gamma$, we can see that $f(x)$ is convex. This family of distributions includes the Pareto II ($\beta=\gamma=1$), Pareto III ($\alpha=\beta=1$), and Pareto IV ($\beta=1$) distributions, as well as the Transformed Beta ($\gamma=1/\gamma^*$) and Burr distributions ($\beta=1, \gamma=1/c$).

(iii) Generalized gamma distribution: In this case, we have

$$f(x) \propto x^{q-1}\, e^{-x^p}, \quad x>0,\ p>0,\ q>0.$$

Taking the derivative of $f(x)$ twice with respect to x, we get

$$f'(x) \overset{sgn}{=} (q-1)x^{q-2}\, e^{-x^p} - px^{p+q-2}\, e^{-x^p}$$

and

$$\begin{aligned} f''(x) \overset{sgn}{=} & (q-1)x^{q-2}\, e^{-x^p}[(q-2) - px^{p-1}] \\ & - px^{p+q-3}\, e^{-x^p}[(p+q-2) - px^p]. \end{aligned}$$

Now, it is easy to observe that, for $q \le 1$ and $p+q \le 2$, $f(x)$ is decreasing and convex. The generalized gamma distribution includes the exponential ($p=q=1$), Weibull ($p=q$), and gamma ($p=1$) distributions as special cases.

https://doi.org/10.1017/S0269964820000467 Published online by Cambridge University Press

Some generalizations of the result in Theorem 5 to the case when the number of underlying mixing variables is arbitrary are presented below.

THEOREM 6: *Let*

$$\tilde{r}_{V_n(\boldsymbol{p},\boldsymbol{\mu})}(x) = \frac{1}{\lambda}\frac{\sum_{i=1}^{n} p_i\, f\left(\frac{x-\mu_i}{\lambda}\right)}{\sum_{i=1}^{n} p_i F\left(\frac{x-\mu_i}{\lambda}\right)}, \quad x > \max\{\mu_1,\ldots,\mu_n\},$$

and

$$\tilde{r}_{V_n(\boldsymbol{q},\boldsymbol{\theta})}(x) = \frac{1}{\lambda}\frac{\sum_{i=1}^{n} q_i\, f\left(\frac{x-\theta_i}{\lambda}\right)}{\sum_{i=1}^{n} q_i F\left(\frac{x-\theta_i}{\lambda}\right)}, \quad x > \max\{\theta_1,\ldots,\theta_n\},$$

be the reversed hazard rate functions of two finite mixtures corresponding to $V_n(\boldsymbol{p},\boldsymbol{\mu})$ *and* $W_n(\boldsymbol{q},\boldsymbol{\theta})$*, respectively. Further, suppose* $f(t)$ *is a decreasing and convex function of* t*. Then, for* $(\boldsymbol{p},\boldsymbol{\mu}) \in \mathcal{H}_n$*, we have*

$$\begin{pmatrix} q_1 & \ldots & q_n \\ \theta_1 & \ldots & \theta_n \end{pmatrix} = \begin{pmatrix} p_1 & \ldots & p_n \\ \mu_1 & \ldots & \mu_n \end{pmatrix} T \implies V_n(\boldsymbol{p},\boldsymbol{\mu}) \geq_{\mathrm{rh}} W_n(\boldsymbol{q},\boldsymbol{\theta}).$$

PROOF: The proof is similar to that of Theorem 2 and is, therefore, omitted here for the sake of brevity. ■

Theorem 6 shows that higher heterogeneity of matrix of parameters implies the reversed hazard rate function of the corresponding mixture model will be longer.

The following corollary is a direct consequence of Theorem 5 and the fact that a finite product of T-transform matrices with the same structure is also a T-transform matrix.

COROLLARY 5: *Let*

$$\tilde{r}_{V_n(\boldsymbol{p},\boldsymbol{\mu})}(x) = \frac{1}{\lambda}\frac{\sum_{i=1}^{n} p_i\, f\left(\frac{x-\mu_i}{\lambda}\right)}{\sum_{i=1}^{n} p_i F\left(\frac{x-\mu_i}{\lambda}\right)}, \quad x > \max\{\mu_1,\ldots,\mu_n\},$$

and

$$\tilde{r}_{W_n(\boldsymbol{q},\boldsymbol{\theta})}(x) = \frac{1}{\lambda}\frac{\sum_{i=1}^{n} q_i\, f\left(\frac{x-\theta_i}{\lambda}\right)}{\sum_{i=1}^{n} q_i\, F\left(\frac{x-\theta_i}{\lambda}\right)}, \quad x > \max\{\theta_1,\ldots,\theta_n\},$$

be the reversed hazard rate functions of two finite mixtures corresponding to $V_n(\boldsymbol{p},\boldsymbol{\mu})$ *and* $W_n(\boldsymbol{q},\boldsymbol{\theta})$*, respectively. Further, suppose* $f(t)$ *is a decreasing and convex function of* t*. If* T*-transform matrices* $T_1,\ldots,T_k$ *have the same structure, then for* $(\boldsymbol{p},\boldsymbol{\mu}) \in \mathcal{H}_n$*, we have*

$$\begin{pmatrix} q_1 & \ldots & q_n \\ \theta_1 & \ldots & \theta_n \end{pmatrix} = \begin{pmatrix} p_1 & \ldots & p_n \\ \mu_1 & \ldots & \mu_n \end{pmatrix} T_1 \ldots T_k \implies V_n(\boldsymbol{p},\boldsymbol{\mu}) \geq_{\mathrm{rh}} W_n(\boldsymbol{q},\boldsymbol{\theta}).$$

The following example provides an illustration of the result in Corollary 5.

https://doi.org/10.1017/S0269964820000467 Published online by Cambridge University Press

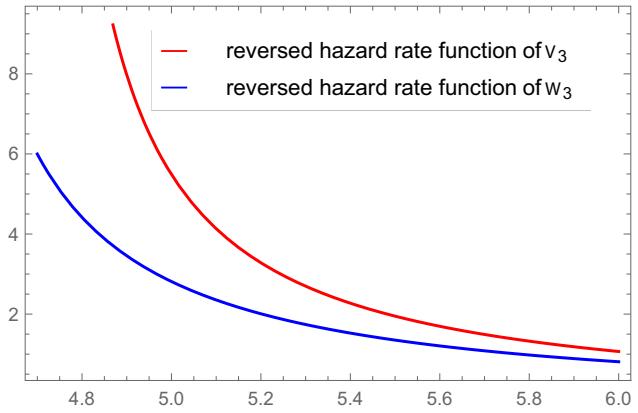


FIGURE **5.** Plots of reversed hazard rate functions of two finite mixture models when $\lambda = 2$.

EXAMPLE 5: *Suppose* $(\mu_1, \mu_2, \mu_3) = (3, 4, 5)$, $(\theta_1, \theta_2, \theta_3) = (3, 4.5, 4.5)$, $(p_1, p_2, p_3) = (0.1, 0.2, 0.7)$, *and* $(q_1, q_2, q_3) = (0.1, 0.45, 0.45)$. *Consider the* T*-transform matrices* T_1 *and* T_2 *with same structure as follows:*

$$T_1 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0.2 & 0.8 \\ 0 & 0.8 & 0.2 \end{pmatrix} \quad \textit{and} \quad T_2 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0.5 & 0.5 \\ 0 & 0.5 & 0.5 \end{pmatrix}.$$

Then, it is easy to observe that $(\boldsymbol{q}, \boldsymbol{\theta}) = (\boldsymbol{p}, \boldsymbol{\mu})T_1T_2$ *which, based on Corollary 5, readily yields* $V_3(\boldsymbol{p}, \boldsymbol{\mu}) \geq_{\text{rh}} W_3(\boldsymbol{q}, \boldsymbol{\theta})$. *The plots of reversed hazard rate functions in this case have has been shown in Figure 5, with exponential as baseline distribution with* $\lambda = 2$.

It will be of interest to know whether the results in Corollary 5 will still hold if the matrices T_i, $i = 1, \ldots, k$, do not have the same structure. The following corollary gives an answer to this question.

COROLLARY 6: *Let*

$$\tilde{r}_{V_n(\boldsymbol{p},\boldsymbol{\mu})}(x) = \frac{1}{\lambda} \frac{\sum_{i=1}^n p_i \, f\left(\frac{x-\mu_i}{\lambda}\right)}{\sum_{i=1}^n p_i F\left(\frac{x-\mu_i}{\lambda}\right)}, \quad x > \max\{\mu_1, \ldots, \mu_n\},$$

and

$$\tilde{r}_{W_n(\boldsymbol{q},\boldsymbol{\theta})}(x) = \frac{1}{\lambda} \frac{\sum_{i=1}^n q_i \, f\left(\frac{x-\theta_i}{\lambda}\right)}{\sum_{i=1}^n q_i F\left(\frac{x-\theta_i}{\lambda}\right)}, \quad x > \max\{\theta_1, \ldots, \theta_n\},$$

be the reversed hazard rate functions of two finite mixtures corresponding to $V_n(\boldsymbol{p}, \boldsymbol{\mu})$ *and* $W_n(\boldsymbol{q}, \boldsymbol{\theta})$, *respectively. Further, suppose* $f(t)$ *is a decreasing and convex function of* t. *If* $(\boldsymbol{p}, \boldsymbol{\mu}) \in \mathcal{H}_n$ *and* $(\boldsymbol{p}, \boldsymbol{\mu})T_1 \ldots T_i \in \mathcal{S}_n$, *for* $i = 1, \ldots, k-1$, *where* $k \geq 2$, *we have*

$$\begin{pmatrix} q_1 & \ldots & q_n \\ \theta_1 & \ldots & \theta_n \end{pmatrix} = \begin{pmatrix} p_1 & \ldots & p_n \\ \mu_1 & \ldots & \mu_n \end{pmatrix} T_1 \ldots T_k \implies V_n(\boldsymbol{p}, \boldsymbol{\mu}) \geq_{\text{rh}} W_n(\boldsymbol{q}, \boldsymbol{\theta}).$$

THEOREM 7: *Let* $\bar{F}_{Z_2(\boldsymbol{p},\boldsymbol{\lambda})}(x) = \sum_{i=1}^2 p_i \bar{F}((x-\mu)/\lambda_i)$ *and* $\bar{F}_{Y_2(\boldsymbol{q},\boldsymbol{\beta})}(x) = \sum_{i=1}^2 q_i \bar{F}((x-\mu)/\beta_i)$ *be the survival functions of two finite mixtures corresponding to* $Z_2(\boldsymbol{p}, \boldsymbol{\lambda})$ *and* $Y_2(\boldsymbol{q}, \boldsymbol{\beta})$,

https://doi.org/10.1017/S0269964820000467 Published online by Cambridge University Press

respectively. Further, suppose $xr(x)$ is an increasing and concave function of x, where $r(x)$ is the hazard rate function of the baseline distribution. Then, for $(\boldsymbol{p}, \boldsymbol{\lambda}^{-1}) \in \mathcal{S}_2$, we have

$$\begin{pmatrix} p_1 & p_2 \\ \lambda_1^{-1} & \lambda_2^{-1} \end{pmatrix} \gg \begin{pmatrix} q_1 & q_2 \\ \beta_1^{-1} & \beta_2^{-1} \end{pmatrix} \implies Z_2(\boldsymbol{p}, \boldsymbol{\lambda}) \geq_{\text{hr}} Y_2(\boldsymbol{q}, \boldsymbol{\beta}).$$

PROOF: Set $\alpha_i = \lambda_i^{-1}$, for $i = 1, 2$. Then, the hazard rate function of $Z_2(\boldsymbol{p}, \boldsymbol{\alpha})$ is given by

$$\begin{aligned} r(x; \boldsymbol{p}, \boldsymbol{\alpha}) &= \frac{\sum_{i=1}^{2} p_i \alpha_i r(\alpha_i(x-\mu)) \bar{F}(\alpha_i(x-\mu))}{\sum_{i=1}^{2} p_i \bar{F}(\alpha_i(x-\mu))} \\ &= \frac{M_1}{M_2}, \quad \text{(say)}. \end{aligned}$$

To establish the desired result, we have to check Conditions (i) and (ii) of Lemma 1. Clearly, $r(x; \boldsymbol{p}, \boldsymbol{\alpha})$ is permutation invariant on $\mathcal{S}_2$, for fixed $x > 0$, which confirms Condition (i). Next, for fixed $x > 0$ and $i \neq j$, consider the function

$$\begin{aligned} \Psi(\boldsymbol{p}, \boldsymbol{\alpha}) = (p_1 - p_2) &\left(\frac{\partial r(x; \boldsymbol{p}, \boldsymbol{\alpha})}{\partial p_1} - \frac{\partial r(x; \boldsymbol{p}, \boldsymbol{\alpha})}{\partial p_2} \right) \\ &+ (\alpha_1 - \alpha_2) \left(\frac{\partial r(x; \boldsymbol{p}, \boldsymbol{\alpha})}{\partial \alpha_1} - \frac{\partial r(x; \boldsymbol{p}, \boldsymbol{\alpha})}{\partial \alpha_2} \right). \end{aligned} \tag{8}$$

The partial derivatives of $r(x; \boldsymbol{p}, \boldsymbol{\alpha})$ with respect to p_1, p_2, α_1, and α_2 are given by

$$\begin{aligned} \frac{\partial r(x; \boldsymbol{p}, \boldsymbol{\alpha})}{\partial p_1} &\overset{\text{sgn}}{=} \bar{F}(\alpha_1(x-\mu))[\alpha_1 r(\alpha_1(x-\mu)) M_2 - M_1], \\ \frac{\partial r(x; \boldsymbol{p}, \boldsymbol{\alpha})}{\partial p_2} &\overset{\text{sgn}}{=} \bar{F}(\alpha_2(x-\mu))[\alpha_2 r(\alpha_2(x-\mu)) M_2 - M_1], \end{aligned}$$

and

$$\begin{aligned} \frac{\partial r(x; \boldsymbol{p}, \boldsymbol{\alpha})}{\partial \alpha_1} &\overset{\text{sgn}}{=} p_1(x-\mu) f(\alpha_1(x-\mu))[M_1 - \alpha_1 r(\alpha_1(x-\mu)) M_2] \\ &\quad + p_1 \bar{F}(\alpha_1(x-\mu))\, M_2 \frac{\partial}{\partial \alpha_1}[\alpha_1 r(\alpha_1(x-\mu))], \\ \frac{\partial r(x; \boldsymbol{p}, \boldsymbol{\alpha})}{\partial \alpha_2} &\overset{\text{sgn}}{=} p_2(x-\mu) f(\alpha_2(x-\mu))[M_1 - \alpha_2 r(\alpha_2(x-\mu)) M_2] \\ &\quad + p_2 \bar{F}(\alpha_2(x-\mu)) M_2 \frac{\partial}{\partial \alpha_2}[\alpha_2 r(\alpha_2(x-\mu))]. \end{aligned}$$

The assumption that $(\boldsymbol{p}, \boldsymbol{\alpha}) \in \mathcal{S}_2$ implies that $(p_1 - p_2)(\alpha_1 - \alpha_2) \leq 0$, which means that either $p_1 \geq p_2$ and $\alpha_1 \leq \alpha_2$, or $p_2 \geq p_1$ and $\alpha_1 \leq \alpha_2$. We present the proof only for the case when $p_1 \geq p_2$ and $\alpha_1 \leq \alpha_2$, since the proof for the other case is quite similar. For $\alpha_1 \leq \alpha_2$,

https://doi.org/10.1017/S0269964820000467 Published online by Cambridge University Press

with the increasing property of $xr(x)$, we have

$$\alpha_1 r(\alpha_1(x-\mu))M_2 - M_1 \le 0 \le \alpha_2 r(\alpha_2(x-\mu))M_2 - M_1, \tag{9}$$

and so

$$\bar{F}(\alpha_1(x-\mu))[\alpha_1 r(\alpha_1(x-\mu))M_2 - M_1] \le 0 \le \bar{F}(\alpha_2(x-\mu))[\alpha_2 r(\alpha_2(x-\mu))M_2 - M_1].$$

Therefore, we conclude that

$$(p_1 - p_2)\left(\frac{\partial r(x;\boldsymbol{p},\boldsymbol{\alpha})}{\partial p_1} - \frac{\partial r(x;\boldsymbol{p},\boldsymbol{\alpha})}{\partial p_2}\right) \le 0. \tag{10}$$

Now, we get

$$\begin{aligned}\frac{\partial r(x;\boldsymbol{p},\boldsymbol{\alpha})}{\partial \alpha_1} - \frac{\partial r(x;\boldsymbol{p},\boldsymbol{\alpha})}{\partial \alpha_2} &\overset{\text{sgn}}{=} (x-\mu)(p_1 f(\alpha_1(x-\mu))[M_1 - \alpha_1 r(\alpha_1(x-\mu))M_2] \\ &\quad - p_2 f(\alpha_2(x-\mu))[M_1 - \alpha_2 r(\alpha_2(x-\mu))M_2]) \\ &\quad + M_2\left(p_1\bar{F}(\alpha_1(x-\mu))\frac{\partial}{\partial\alpha_1}[\alpha_1 r(\alpha_1(x-\mu))] - p_2\bar{F}(\alpha_2(x-\mu))\frac{\partial}{\partial\alpha_2}[\alpha_2 r(\alpha_2(x-\mu))]\right) \\ &= K_1 + K_2,\end{aligned}$$

where

$$\begin{aligned}K_1 = (x-\mu)(&p_1 f(\alpha_1(x-\mu))[M_1 - \alpha_1 r(\alpha_1(x-\mu))M_2] \\ &- p_2 f(\alpha_2(x-\mu))[M_1 - \alpha_2 r(\alpha_2(x-\mu))M_2])\end{aligned}$$

and

$$K_2 = M_2\left(p_1\,\bar{F}(\alpha_1(x-\mu))\frac{\partial}{\partial\alpha_1}[\alpha_1 r(\alpha_1(x-\mu))] - p_2\bar{F}(\alpha_2(x-\mu))\frac{\partial}{\partial\alpha_2}[\alpha_2 r(\alpha_2(x-\mu))]\right).$$

For $p_1 \ge p_2$ and $\alpha_1 \le \alpha_2$, we have $K_1 \ge 0$ by using the inequality in (9). On the other hand, from $p_1 \ge p_2$, $\alpha_1 \ge \alpha_1$, decreasing property of $\bar{F}(\alpha(x-\mu))$ with respect to α and the concavity of $xr(x)$, we also have $K_2 \ge 0$, and so

$$(\alpha_1 - \alpha_2)\left(\frac{\partial r(x;\boldsymbol{p},\boldsymbol{\alpha})}{\partial \alpha_1} - \frac{\partial r(x;\boldsymbol{p},\boldsymbol{\alpha})}{\partial \alpha_2}\right) \le 0.$$

Thus, we can conclude that $\Psi(\boldsymbol{p},\boldsymbol{\alpha}) \le 0$ and therefore $r(x;\boldsymbol{p},\boldsymbol{\lambda}) \le r(x;\boldsymbol{q},\boldsymbol{\beta})$ for all $x \in \mathbb{R}$, which completes the proof of the theorem. ■

REMARK 2: *It should be mentioned that the condition "$xr(x)$ is an increasing and concave function of x" in Theorem 7 is quite general and can be verified for many distributions. For example, consider the following cases.*

(i) Generalized Pareto distribution: The survival function and density function of Generalized Pareto distributions are

$$\bar{F}(x) = (1+\alpha x)^{-1/\alpha} \quad \textit{and} \quad f(x) = (1+\alpha x)^{-(1/\alpha)-1}, \quad x > 0,\ \alpha > 0,$$

respectively. The first two derivatives of $xr(x)$ with respect to x are

$$(xr(x))' = (1+\alpha x)^{-2} \quad \textit{and} \quad (xr(x))' = -2\alpha(1+\alpha x)^{-2},$$

which readily reveal that $xr(x)$ is an increasing and concave function of x.

https://doi.org/10.1017/S0269964820000467 Published online by Cambridge University Press

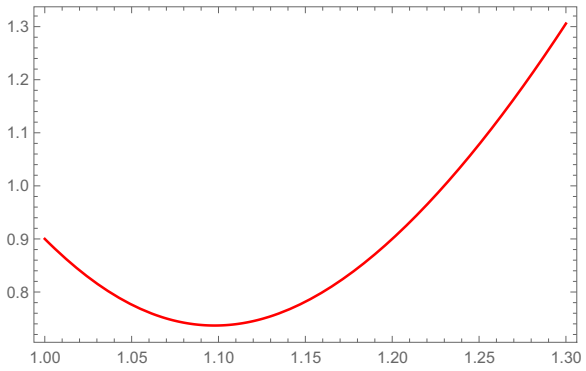


FIGURE **6.** The ratio of density functions of $Z_2(\boldsymbol{p}, \boldsymbol{\lambda})$ and $Y_2(\boldsymbol{q}, \boldsymbol{\beta})$.

(ii) Generalized gamma distribution: In this case, we have

$$f(x) = \frac{p}{\Gamma(q/p)} x^{q-1} e^{-x^p}, \quad x > 0,\ p > 0,\ q > 0.$$

Khaledi et al. [13] have shown that for $p < 1$ and $q < 1$, $xr(x)$ is increasing and concave.

The following counterexample proves that the result in Theorem 7 can not be extended to likelihood ratio ordering.

COUNTEREXAMPLE 3: *Let us consider the standard exponential as a baseline distribution for the location-scale model. Take $(\lambda_1, \lambda_2) = (1/3, 1/20)$, $(\beta_1, \beta_2) = (1/16.6, 1/6.4)$, $(p_1, p_2) = (0.6, 0.4)$, $(q_1, q_2) = (0.44, 0.56)$, and $\mu = 1$. It is then easy to check that $(\boldsymbol{q}, \boldsymbol{\beta}^{-1}) = (\boldsymbol{p}, \boldsymbol{\lambda}^{-1})T_{0.2}$ which, according to Part (ii) of Definition 4, implies $(\boldsymbol{p}, \boldsymbol{\lambda}^{-1}) \gg (\boldsymbol{q}, \boldsymbol{\beta}^{-1})$, and $(\boldsymbol{p}, \boldsymbol{\lambda}^{-1}) \in \mathcal{S}_2$. The ratio of the density functions of $Z_2(\boldsymbol{p}, \boldsymbol{\lambda})$ and $Y_2(\boldsymbol{q}, \boldsymbol{\beta})$ is given by*

$$\frac{f_{Z_2(\boldsymbol{p},\boldsymbol{\lambda})}(x)}{f_{Y_2(\boldsymbol{q},\boldsymbol{\beta})}(x)} = \frac{1.8e^{-3(x-1)} + 0.8e^{-20(x-1)}}{7.304e^{-16.6(x-1)} + 3.584e^{-6.4(x-1)}}, \qquad x \geq 1.$$

As plotted in Figure 6, it can be seen that the above ratio function is not monotone for $x \in [1, \infty)$, which means that $Z_2(\boldsymbol{p}, \boldsymbol{\beta}) \not\geq_{lr} Y_2(\boldsymbol{q}, \boldsymbol{\gamma})$ and $Z_2(\boldsymbol{p}, \boldsymbol{\beta}) \not\leq_{lr} Y_2(\boldsymbol{q}, \boldsymbol{\gamma})$.

5. CONCLUDING REMARKS

Finite mixtures have found key applications in many different areas including reliability theory, actuarial science, medicine, and economics. In this paper, under certain conditions on the model parameters, we have discussed stochastic comparisons of two finite mixture models with location-scale distributed components in the sense of the usual stochastic order and the reversed hazard rate order. Several examples have also been presented for illustrating for the established results.

As a main tool for proving the results, we have used the concept of multivariate majorization. Our findings in this paper are theoretical, but various applications in reliability practice can be readily seen. For example, let us consider the following scenario:

Suppose we have two systems produced by different manufacturers resulting in different reliability characteristics for the systems. Each system can operate in n operational regimes

https://doi.org/10.1017/S0269964820000467 Published online by Cambridge University Press

with corresponding (because they are not statistically identical) mixing probabilities p_i and q_i. Let their lifetimes in each regime be described by different distributions, say, $LS(\mu_i, \lambda)$ and $LS(\theta_i, \lambda)$, respectively. Then, a question of interest is which of the two systems will perform better in a certain stochastic sense? By the use of Theorem 1, we can conclude that if $(\boldsymbol{p}, \boldsymbol{\mu})$ majorizes $(\boldsymbol{q}, \boldsymbol{\theta})$, then the first system will perform better than the second in the sense of the usual stochastic order.

In the present work, we have established results for the usual stochastic order, the hazard rate order, and the reversed hazard rate order. Dispersion and variability orders in this general setup will be of great interest, and we are currently working on this problem and hope to report the findings in a future paper.

One final comment we wish to make is that we have established in this work stochastic orderings for finite mixtures of location-scale models when either the mixture proportions and threshold (location) parameters are different, or the mixture proportions and scale parameters are different. It will naturally be of interest to consider the general situation when all proportion, location and scale parameters are different!

ACKNOWLEDGMENTS

Thanks are due to the Editor and anonymous Reviewers for their useful comments and suggestions on an earlier version of this manuscript which led to this improved version. This study has been conducted by the University of Zabol, Grant No. UOZ-GR-9618-14.

References

1. Amini-Seresht, E. & Zhang, Y. (2017). Stochastic comparisons on two finite mixture models. *Operations Research Letters* 45: 475–480.
2. Balakrishnan, N., Haidari, A. & Masoumifard, K. (2015). Stochastic comparisons of series and parallel-systems with generalized exponential components. *IEEE Transactions on Reliability* 64: 333–348.
3. Castillo, E., Hadi, A.S., Balakrishnan, N. & Sarabia, J.M. (2005). *Extreme value and related models with applications in engineering and science*. Hoboken, NJ: John Wiley & Sons.
4. Cha, J.H. & Finkelstein, M. (2013). The failure rate dynamics in heterogeneous populations. *Reliability Engineering and System Safety* 112: 120–128.
5. Everitt, B.S. & Hand, D.J. (1981). *Finite mixture distributions*. London: Chapman and Hall.
6. Finkelstein, M. (2008). *Failure rate modeling for reliability and risk*. London: Springer.
7. Finkelstein, M. & Esaulova, V. (2006). On mixture failure rate ordering. *Communications in Statistics – Theory and Methods* 35: 1943–1955.
8. Franco, M., Balakrishnan, N., Kundu, D., & Vivo, J.M. (2014). Generalized mixtures of Weibull components. *Test* 23: 515–535.
9. Hazra, N.K. & Finkelstein, M. (2018). On stochastic comparisons of finite mixtures for some semiparametric families of distributions. *Test* 27: 988–1006.
10. Hosking, J.R.M. & Wallis, J.R. (1997). *Regional frequency analysis: An approach based on L-moment*. Cambridge, UK: Cambridge University Press.
11. Johnson, N.L., Kotz, S., & Balakrishnan, N. (1994). *Continuous univariate distributions*, 2nd ed., vol. 1. New York: John Wiley & Sons.
12. Johnson, N.L., Kotz, S., & Balakrishnan, N. (1995). *Continuous univariate distributions*, 2nd ed., vol. 2. New York: John Wiley & Sons.
13. Khaledi, B.E, Farsinezhad, S., & Kochar, S.C. (2011). Stochastic comparisons of order statistics in the scale model. *Journal of Statistical Planning and Inference* 141: 276–286.
14. Marshall, A.W. & Olkin, I. (2007). *Life Distributions*. New York: Springer.
15. Marshall, A.W., Olkin, I., & Arnold, B.C. (2011). *Inequalities: theory of majorization and its applications*, 2nd ed. New York: Springer.
16. Müller, A. & Stoyan, D. (2002). *Comparison methods for stochastic models and risks*. New York: John Wiley & Sons.
17. Navarro, J. & Hernandez, P.J. (2004). How to obtain bathtub-shaped failure rate models from normal mixtures. *Probability in the Engineering and Informational Sciences* 18: 511–531.
18. Shaked, M. & Shanthikumar, J.G. (2007). *Stochastic orders*. New York: Springer.

https://doi.org/10.1017/S0269964820000467 Published online by Cambridge University Press

19. Titterington, D.M., Smith, A.F.M., & Makov, U.E. (1985). *Statistical analysis of finite mixture distributions*. Chichester: John Wiley & Sons.

APPENDIX

Proof of Lemma 1: Let us assume (1) holds. Under $U \in \mathcal{S}_2$, (i) is obtained from (1), and the facts that $U \gg U\Pi$ and $U\Pi \gg U$. For obtaining (ii), consider the following T-transform matrix:

$$T_\eta = \eta \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + (1-\eta) \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} \eta & 1-\eta \\ 1-\eta & \eta \end{pmatrix}. \tag{A.1}$$

Now, let us assume that $U \in \mathcal{S}_2$. Because $U \gg UT_\eta$, then from (1), we get

$$\frac{\phi(U) - \phi(UT_\eta)}{1-\eta} \geq 0 \quad \text{for all } 0 \leq \eta < 1.$$

Upon using L'Hopital's rule, we obtain

$$\begin{aligned} 0 \leq \lim_{\eta \to 1} \frac{\phi(U) - \phi(UT_\eta)}{1-\eta} &= (u_1^C - u_2^C)' \begin{pmatrix} \phi_{11}(U) - \phi_{12}(U) \\ \phi_{21}(U) - \phi_{22}(U) \end{pmatrix} \\ &= \sum_{i=1}^{2} (u_{i1} - u_{i2})(\phi_{i1}(U) - \phi_{i2}(U)), \end{aligned}$$

which implies (ii).

Conversely, suppose (i) and (ii) hold. To obtain (1), assume that $U \in \mathcal{S}_2$ and $V_1 = UT_{\eta_1}$. Consider the function $g : [0,1] \to R^+$ as follows:

$$\begin{aligned} g(\eta_1) &= \phi(UT_{\eta_1}) \\ &= \phi(v_1^C, v_2^C), \end{aligned}$$

where $v_1^C = \eta_1 u_1^C + (1-\eta_1)u_2^C$ and $v_2^C = (1-\eta_1)u_1^C + \eta_1 u_2^C$. By taking the derivative of $g(\eta_1)$ with respect to η_1, for $\eta_1 > 1/2$, we have

$$\begin{aligned} g'(\eta_1) &= (u_1^C - u_2^C)' \begin{pmatrix} \phi_{11}(v_1^C, v_2^C) - \phi_{12}(v_1^C, v_2^C) \\ \phi_{21}(v_1^C, v_2^C) - \phi_{22}(v_1^C, v_2^C) \end{pmatrix} \\ &= \sum_{i=1}^{2} (u_{i1} - u_{i2})(\phi_{i1}(v_1^C, v_2^C) - \phi_{i2}(v_1^C, v_2^C) \\ &= \sum_{i=1}^{2} \frac{v_{i1} - v_{i2}}{2\eta_1 - 1} (\phi_{i1}(v_1^C, v_2^C) - \phi_{i2}(v_1^C, v_2^C) \quad (\text{because } v_1^C - v_2^C = (2\eta_1 - 1))(u_1^C - u_2^C)) \\ &\geq 0 \quad (\text{from (ii)}). \end{aligned}$$

So, $g(\eta_1)$ is non-decreasing in $\eta_1 \in (1/2, 1]$. On the other hand, from (i), it readily follows that $g(1/2 + \epsilon) = g(1/2 - \epsilon)$ for $0 < \epsilon < 1/2$. These observations mean that $g(\eta_1)$ is symmetric about $1/2$, which along with the non-decreasing property of $g(\eta_1)$ in $\eta_1 \in (1/2, 1]$, implies that $g(\eta_1) \leq g(1)$ for all $0 \leq \eta_1 \leq 1$. So, $\phi(U) \geq \phi(UT_{\eta_1})$. Now, suppose $U \in \mathcal{S}_2$, and $V_k = UT_{\eta_1}T_{\eta_2} \cdots T_{\eta_k}$, where $k \geq 2$ is a finite integer. Because $U \in \mathcal{S}_2$, then $V_i = UT_{\eta_1} \ldots T_{\eta_i} \in \mathcal{S}_2$, $i = 2, \ldots, k$. But, we easily observe that $U \gg V_2 \gg \cdots \gg V_i \gg \cdots \gg V_k$ which, from the above discussion, implies that $\phi(U) \geq \phi(V_2) \geq \cdots \geq \phi(V_i) \geq \cdots \geq \phi(V_k)$. This result completes the proof of Lemma 1. ■

https://doi.org/10.1017/S0269964820000467 Published online by Cambridge University Press

PROOF OF LEMMA 2: Based on the definition of a T-transform matrix, we can observe that the matrices U and V have $(n-2)$ columns to be the same and are possibly different in two columns i and j. Thus, we have

$$\begin{aligned} &v_{ri} = \eta u_{ri} + (1-\eta)u_{rj} \quad \text{and} \quad v_{rj} = (1-\eta)u_{ri} + \eta u_{rj}, \quad \text{for } r = 1, 2, \\ &v_{rs} = u_{rs}, \quad \text{for } r = 1, 2, \quad s = 1, \ldots, n, \text{ and } s \neq i, j. \end{aligned} \tag{A.2}$$

Let us set

$$U_0 = \begin{pmatrix} u_{1i} & u_{1j} \\ u_{2i} & u_{2j} \end{pmatrix}, \quad V_0 = \begin{pmatrix} v_{1i} & v_{1j} \\ v_{2i} & v_{2j} \end{pmatrix}, \quad \text{and} \quad T_0 = \begin{pmatrix} \eta & 1-\eta \\ 1-\eta & \eta \end{pmatrix}.$$

Indeed, keeping the ith and jth columns of the matrices U, V, and T_η and deleting all other columns, we immediately obtain the matrices U_0, V_0, and T_0, respectively. From the fact that $U \in \mathcal{S}_n$, we readily conclude that $U_0 \in \mathcal{S}_2$. On the other hand, we have $V_0 = U_0 T_0$, i.e., $U_0 \gg V_0$. Now, upon using the above observations, and the fact that ϕ_2 satisfies (1), we obtain from Lemma 1 that

$$\phi_2(U_0) \geq \phi_2(V_0). \tag{A.3}$$

Therefore, we obtain

$$\begin{aligned} \phi_n(U) &= \sum_{l=1}^{n} \varphi(u_{1l}, u_{2l}) \\ &= \phi_2(U_0) + \left(\sum_{l=1, l\neq i,j}^{n} \varphi(u_{1l}, u_{2l}) \right) \geq \phi_2(V_0) + \left(\sum_{l=1, l\neq i,j}^{n} \varphi(u_{1l}, u_{2l}) \right) \quad \text{(according to (A.2))} \\ &= \phi_2(V_0) + \left(\sum_{l=1, l\neq i,j}^{n} \varphi(v_{1l}, v_{2l}) \right) \quad \text{(according to (A.3))} \\ &= \phi_n(V), \end{aligned}$$

which completes the proof of Lemma 2. ■

https://doi.org/10.1017/S0269964820000467 Published online by Cambridge University Press