Journal of Multivariate Analysis **80**, 344–357 (2002)
doi:10.1006/jmva.2000.1982, available online at http://www.idealibrary.com on IDEAL®

On Stochastic Orders for Sums of Independent Random Variables

Ramesh M. Korwar

University of Massachusetts

Received September 7, 1999; published online July 19, 2001

In this paper, it is shown that a convolution of uniform distributions (a) is more dispersed and (b) has a smaller hazard rate when the scale parameters of the uniform distributions are more dispersed in the sense of majorization. It is also shown that a convolution of gamma distributions with a common shape parameter greater than 1 is larger in (a) likelihood ratio order and (b) dispersive order when the scale parameters are more dispersed in the sense of majorization. © 2001 Elsevier Science

AMS 1991 subject classifications: 62G30; 60E15; 62N05; 62D05.

Key words and phrases: convolutions; dispersive order; hazard rate order; likelihood ratio order; majorization; gamma distribution; exponential distribution; uniform distribution; increasing failure rate distribution; totally positive of order two function.

1. INTRODUCTION

The uniform distribution, because of its simplicity, serves as an important model in many applications in statistics. The uniform distribution is also of great theoretical importance because of its connection to the integral transform and unimodal distributions. The gamma distribution is important in reliability and engineering applications. In this paper various stochastic orders are considered for convolutions of gamma as well as convolutions of uniform distributions.

Let X and Y be random variables with distribution functions F and G, survival functions $\bar{F}=1-F$ and $\bar{G}=1-G$, with right continuous inverse distribution functions F^{-1} and G^{-1} respectively.

DEFINITION 1.1. X is smaller than Y in the dispersive order (written $X \leqslant_{disp} Y$) if $F^{-1}(v)-F^{-1}(u) \leqslant G^{-1}(v)-G^{-1}(u)$ whenever $0<u \leqslant v<1$.

0047-259X/01 $35.00

DEFINITION 1.2. X is said to be smaller then Y in the hazard rate order (written $X \leqslant_{hr} Y$) if $\bar{G}(x)/\bar{F}(x)$ is increasing in x. When X and Y have probability density functions f and g respectively, then $X \leqslant_{hr} Y$ if, and only if, $f(x)/\bar{F}(x) \geqslant g(x)/\bar{G}(x)$ for all real x.

DEFINITION 1.3. Let X and Y have density or mass functions f and g respectively. Then, X is said to be smaller than Y in the likelihood ratio order (written $X \leqslant_{\ell r} Y$) if

$$f(u)\, g(v) \geqslant f(v)\, g(u) \qquad \text{for all} \quad u \leqslant v.$$

In the remainder of this paper, the abbreviations r.v., d.f., and p.d.f. will be used for respectively random variable, distibution function, and density function.

DEFINITION 1.4. Let $\mathbf{x} = (x_1, ..., x_n)$ and $\mathbf{y} = (y_1, ..., y_n)$ be two real vectors. Let $x_{[1]} \geqslant \cdots \geqslant x_{[n]}$ be the x_i's ordered from the largest to the smallest. Then $\mathbf{y}$ is said to majorize $\mathbf{x}$ (written $\mathbf{x} \overset{m}{\leqslant} \mathbf{y}$) if $\sum_{i=1}^{j} x_{[i]} \leqslant \sum_{i=1}^{j} y_{[i]}$, $j = 1, ..., n-1$ and $\sum_{i=1}^{n} x_{[i]} = \sum_{i=1}^{n} y_{[i]}$.

DEFINITION 1.5. Let $A \subseteq R^n$, where R^n is the n-dimensional real space. A function ϕ: $A \to R$ is said to be Schur convex (concave) if

$$\mathbf{x} \overset{m}{\leqslant} \mathbf{y} \qquad \text{implies} \quad \phi(\mathbf{x}) \leqslant (\geqslant)\, \phi(\mathbf{y}), \qquad \text{for} \quad \mathbf{x}, \mathbf{y} \in A.$$

DEFINITION 1.6. Let A and B be subsets of the real line. A function $K(x, y)$ on $A \times B$ is said to be totally positive of order 2 (TP_2) if

$$K(x_1, y_1)\, K(x_2, y_2) - K(x_1, y_2)\, K(x_2, y_1) \geqslant 0$$

for all $x_1 < x_2$ in A and all $y_1 < y_2$ in B.

Definition 1.1 makes it clear that the $\leqslant_{disp}$ order compares X and Y by variability and that it is location free. When variances exist, $X \leqslant_{disp} Y$ implies variance $(X) \leqslant$ variance (Y). There are no implications in general among the three stochastic orders except that the likelihood ratio order implies the hazard rate order. Letting $K(1, y) = f(y)$ and $K(2, y) = g(y)$ in Definition 1.6, it can be seen that establishing $X \leqslant_{\ell r} Y$ is equivalent to showing that K is TP_2 on $\{1, 2\} \times (-\infty, \infty)$. See the proof of Theorem 3.3 for an exploitation of this relationship. For more details on these orders, see Shaked and Shantikumar (1994).

Majorization makes precise the vague notion that the components of a vector are "less dispersed" than those of another vector (of the same dimension). Majorization is useful and very powerful in deriving certain types of inequalities. In this paper, majorization is used to partially order the vectors of scale parameters of gamma (as well as uniform) distributions in convolutions to obtain stochastic orderings. For more details on majorization and its applications, see Marshall and Olkin (1979).

A r.v. X is said to have a gamma distribution with shape parameter $a(>0)$ and scale parameter $\lambda(>0)$ if its p.d.f. $f(x; a, \lambda)$ is given by $f(x; a, \lambda) = \{1/\Gamma(a)\}\, \lambda^a x^{a-1} \exp(-\lambda x)$, $x > 0$. The main results of this paper are:

(A) A convolution of gamma distributions with a common shape parameter greater than or equal to 1 is more dispersed (in the dispersive order) if the vector of scale parameters of the gamma distributions is more dispersed (in the sense of majorization) (Theorem 3.6).

(B) A convolution of gamma distributions with a common shape parameter greater than or equal to 1 is larger in the likelihood ratio order if the vector of scale parameters of the gamma distributions is more dispersed (in the sense of majorization) (Theorem 3.4).

(C) Result (A) holds for a convolution of uniform distributions. Result (B) where hazard rate ordering replaces likelihood rates ordering is valid (Theorem 2.3, Remark 2.2, and Theorem 2.5. respectively).

Recently Kochar and Ma (1999 a, b) proved result (A) for the case of an integral common shape parameter. Result (B) for a common integral shape parameter can be proved by using the likelihood ratio order for convolutions of exponential distributions established by Boland *et al.* (1994). However, the extensions to the case of a nonintegral shape parameter considered here are nontrivial.

2. STOCHASTIC ORDERS FOR CONVOLUTIONS OF UNIFORM DISTRIBUTIONS

This section deals with the dispersive order and hazard rate order for convolutions of uniform distributions by using majorization.

Proposition 2.1 gives the distribution of a linear combination (with positive coefficients) of two independent uniform r.v.'s on $[0, 1]$. This result is then used to prove that a convolution of two uniform distributions is more dispersed (in the dispersive order) when the scale parameters of the uniform distributions are more dispersed in the sense of majorization. This latter result is needed in proving its generalization (Theorem 2.3).

PROPOSITION 2.1. *Let Y_1 and Y_2 be two independent r.v.'s each with a uniform distribution on $[0, 1]$. Furthermore, let $0 < \lambda_2 \leqslant \lambda_1$ be any constants. Then, the p.d.f. f, and the d.f. F, of $Y = Y_1/\lambda_1 + Y_2/\lambda_2$ are given by*

$$f(y) = \begin{cases} \lambda_1 \lambda_2 y, & 0 \leqslant y \leqslant 1/\lambda_1 \\ \lambda_2, & 1/\lambda_1 \leqslant y \leqslant 1/\lambda_2 \\ \lambda_1 + \lambda_2 - \lambda_1 \lambda_2 y, & 1/\lambda_2 \leqslant y \leqslant 1/\lambda_1 + 1/\lambda_2 \\ 0, & \textit{otherwise} \end{cases} \tag{2.1}$$

and

$$F(y) = \begin{cases} 0, & y \leqslant 0 \\ \lambda_1 \lambda_2 y^2/2, & 0 \leqslant y < 1/\lambda_1 \\ \lambda_2 y - \lambda_2/2\lambda_1, & 1/\lambda_1 \leqslant y \leqslant 1/\lambda_2 \\ 1 - \{\lambda_1 + \lambda_2 - \lambda_1 \lambda_2 y\}^2/2\lambda_1 \lambda_2, & 1/\lambda_2 \leqslant y < 1/\lambda_1 + 1/\lambda_2 \\ 1, & y \geqslant 1/\lambda_1 + 1/\lambda_2. \end{cases} \tag{2.2}$$

Proof. Let $Z = Y_2/\lambda_2$. Then, routine calculations show that the joint p.d.f. $g(y, z)$ of Y and Z is given by $g(y, z) = \lambda_1 \lambda_2, 0 \leqslant (y - z) \leqslant 1/\lambda_1$, $0 \leqslant z \leqslant 1/\lambda_2$; 0, otherwise. From this it follows that

$$\begin{aligned} f(y) &= \lambda_1 \lambda_2 \{\min(1/\lambda_2,\, y) - \max(0,\, y - 1/\lambda_1)\}, && 0 \leqslant y \leqslant 1/\lambda_1 + 1/\lambda_2 \\ &= 0, && \text{otherwise.} \end{aligned}$$

Now, (2.1) and (2.2) follow from this last result. ∎

Theorem 2.2 below shows that the convolution F defined in the proof of Proposition 2.1 is more dispersed when the parameters λ_1 and λ_2 are more dispersed.

The proof of Theorem 2.2 uses a result due to Saunders and Moran (1978).

THEOREM 2.1 (Saunders and Moran, 1978). *Let $\{F_b, b \in R\}$ be a collection of d.f.'s such that F_b is supported on some interval $(x_-^{(b)}, x_+^{(b)}) \subseteq (0, \infty)$ and has p.d.f. f_b which does not vanish on any subinterval of $(x_-^{(b)}, x_+^{(b)})$. Then,*

$$F_b \leqslant_{disp} F_{b^*}, \qquad b, b^* \in R, \qquad b \leqslant b^*$$

if and only if $F'_b(y)/f_b(y)$ is decreasing in y, where F'_b is the derivative of F_b with respect to b.

THEOREM 2.2. *Let* $c>0$ *be any arbitrary number. For each* $\lambda \in [c/2, c)$, *let* $F(\cdot;\lambda)$ *be the d.f. of the convolution* $Y_1/\lambda + Y_2/(c-\lambda)$, *where* Y_1 *and* Y_2 *are independent and identically distributed as a uniform r.v. on* $[0,1]$. *Then,*

$$F(\cdot;\lambda) \leqslant_{disp} F(\cdot;\lambda^*), \qquad \lambda, \lambda^* \in \left[\frac{c}{2}, c\right), \quad \lambda \leqslant \lambda^*.$$

Proof. Using (2.1) and (2.2) one computes that

$$F'(y;\lambda)/f(y;\lambda)$$
$$= \begin{cases} (c-2\lambda)\, y/2\lambda(c-\lambda), & 0 \leqslant y \leqslant 1/\lambda \\ -y/(c-\lambda) + (c-2\lambda)/2\lambda^2(c-\lambda), & 1/\lambda \leqslant y \leqslant 1/(c-\lambda) \\ (c-2\lambda)\{(c+\lambda(c-\lambda)\, y\}/2\lambda^2(c-\lambda)^2, & 1/(c-\lambda) \leqslant y \leqslant 1/\lambda + 1/(c-\lambda) \end{cases}$$

which is clearly a decreasing function of y since $c-2\lambda \leqslant 0$. The result now follows from Theorem 2.1. ∎

Theorem 2.3 generalizes Theorem 2.2. To do this, the following definition of a dispersive distribution is needed.

DEFINITION 2.1. A r.v. Z (and its d.f. F) is said to be dispersive if $X+Z \leqslant_{disp} Y+Z$ whenever Z is independent of X and Y and $X \leqslant_{disp} Y$.

Examples of dispersive distributions are: exponential, normal, and uniform.

THEOREM 2.3. *Let* $Y_1, ..., Y_n$ *be independent and identically distributed as a uniform r.v. on* $[0,1]$. *Let, further,* λ_i*'s* (λ_i^**'s*) *be any positive numbers such that* $(\lambda_1, ..., \lambda_n) \overset{m}{\leqslant} (\lambda_1^*, ..., \lambda_n^*)$. *Set* $Y = Y_1/\lambda_1 + \cdots + Y_n/\lambda_n$ *and* $Y^* = Y_1/\lambda_1^* + \cdots + Y_n/\lambda_n^*$. *Then,* $Y \leqslant_{disp} Y^*$.

Proof. It is sufficient to prove the Theorem for the case $(\lambda_1, \lambda_2) \overset{m}{\leqslant} (\lambda_1^*, \lambda_2^*)$ and $\lambda_i = \lambda_i^*$, $i = 3, ..., n$. See Lemma 2B.1, p. 21, in Marshall and Olkin (1979). Now, by Theorem 2.2,

$Y_1/\lambda_1 + Y_2/\lambda_2 \leqslant_{disp} Y_1/\lambda_1^* + Y_2/\lambda_2^*$. It is known that a convolution of dispersive r.v.'s is dispersive (Theorem 5(2) of Lewis and Thompson (1981)). Hence $Y_3/\lambda_3 + \cdots + Y_n/\lambda_n$ $(= Y_3/\lambda_3^* + \cdots + Y_n/\lambda_n^*)$ is dispersive since each of Y_i/λ_i, being a uniform r.v., is dispersive. The proof is now completed by using Definition 2.1. ∎

COROLLARY 2.1. *Let* Y_i*'s be as in Theorem 2.3,* $\lambda_1 \leqslant \cdots \leqslant \lambda_n$, $\lambda_1^* \leqslant \cdots \leqslant \lambda_n^*$, *and* $(\lambda_1, ..., \lambda_n) \overset{m}{\leqslant} (\lambda_1^*, ..., \lambda_n^*)$. *Then, for each* $k = 1, ..., n$, $\sum_{i=1}^k Y_i/\lambda_i \leqslant_{disp} \sum_{i=1}^k Y_i/\lambda_i^*$.

The proof is similar to that of Corallary 3.2 *and will be omitted here.*

The generalization of Theorem 2.3 to a convolution of any uniform distributions is achieved by using the fact that the order $\leqslant_{disp}$ is location free and Theorem 2.3.

THEOREM 2.4. *Let* X_{λ_i} $(X_{\lambda_i^*})$, $i=1, ..., n$ *be independent r.v.'s, with* X_{λ_i} $(X_{\lambda_i^*})$ *having a uniform distribution on* $[a_i, b_i]$ $([a_i^*, b_i^*])$, $\lambda_i = 1/(b_i - a_i)$ $(\lambda_i^* = 1/(b_i^* - a_i^*))$. *Let, further, that* $(\lambda_1, ..., \lambda_n) \overset{m}{\leqslant} (\lambda_1^*, ..., \lambda_n^*)$. *Set* $X = X_{\lambda_1} + \cdots + X_{\lambda_n}$ *and* $X^* = X_{\lambda_1^*} + \cdots + X_{\lambda_n^*}$. *Then,* $X \leqslant_{disp} X^*$.

Proof. Write $X_{\lambda_i} = Y_{\lambda_i}/\lambda_i + a_i$ $(X_{\lambda_i^*} = Y_{\lambda_i^*}/\lambda_i^* + a_i^*)$, where $Y_{\lambda_i} = (X_{\lambda_i} - a_i)/(b_i - a_i)$ $(Y_{\lambda_i^*} = (X_{\lambda_i^*} - a_i^*)/(b_i^* - a_i^*))$. Thus, Y_{λ_i}'s ($Y_{\lambda_i^*}$'s) are independent and identically distributed, each with a uniform distribution on $[0, 1]$. By Theorem 2.3, $Y \leqslant_{disp} Y^*$, where $Y = Y_{\lambda_1}/\lambda_1 + \cdots + Y_{\lambda_n}/\lambda_n$ and $Y^* = Y_{\lambda_1^*}/\lambda_1^* + \cdots + Y_{\lambda_n^*}/\lambda_n^*$. Now, $Y + a \leqslant_{disp} Y \leqslant_{disp} Y^* \leqslant_{disp} Y^* + a^*$, where $a = \sum_{i=1}^n a_i$ and $a^* = \sum_{i=1}^n a_i^*$. But $Y + a = X$ and $Y^* + a^* = X^*$. Thus, $X \leqslant_{disp} X^*$. ∎

The definition of an increasing failure rate (IFR) distibution is needed in proving the second main result.

DEFINITION 2.2. A. d.f. F is an increasing failure rate (IFR) distribution if $\bar{F}(t + x)/\bar{F}(t)$ is decreasing in $-\infty < t < \infty$ for each $x \geqslant 0$.

Now for the final results.

THEOREM 2.5. *Under the assumptions of Theorem* 2.4, $X \leqslant_{hr} X^*$.

Proof. By Theorem 2.4, $X \leqslant_{disp} X^*$. X has an IFR distribution since a convolution of IFR distributions is IFR (Barlow, Marshall and Proschan (1963)). The result now follows from the fact that if $X \leqslant_{disp} Y$ and X or Y is IFR, then $X \leqslant_{hr} Y$ (Bagai and Kochar, 1986). ∎

Remark 2.1. Theorem 2.5 can be proved more directly as follows. Using Proposition 2.1, one can prove that $Y_1/\lambda_1 + Y_2/\lambda_2 \leqslant_{hr} Y_1/\lambda_1^* + Y_2/\lambda_2^*$ if $(\lambda_1, \lambda_2) \overset{m}{\leqslant} (\lambda_1^*, \lambda_2^*)$. Now, a sum of independent uniform r.v.'s, being a sum of independent r.v's, each with an IFR d.f. is IFR. It then follows from Lemma 1.B.5 in Shaked and Shantikumar (1994) that $Y \leqslant_{hr} Y^*$.

Remark 2.2. The hazard rate order $X \leqslant_{hr} X^*$ in Theorem 2.5 cannot be replaced by the stronger order $X \leqslant_{\ell r} X^*$. Consider $n = 2$. Let $a_1 = a_2 =$

$a_1^* = a_2^* = 0$ and let $(\lambda_1, \lambda_2) \overset{m}{\leqslant} (\lambda_1^*, \lambda_2^*)$ and let f and f^* be the p.d.f's of X and X^* respectively. Then, from Proposition 2.1 it follows that

$$\begin{aligned} f^*(x)/f(x) &= \lambda_1^* \lambda_2^* / \lambda_1 \lambda_2, \qquad & x \in [0, 1/\lambda_1^*) \\ &= \lambda_2^* / \lambda_1 \lambda_2 x, \qquad & x \in [1/\lambda_1^*, 1/\lambda_1). \end{aligned}$$

Thus, while $f^*(x)/f(x)$ is increasing in $x \in [10, 1/\lambda_1^*)$, it is decreasing in $x \in [1/\lambda_1^* .1/\lambda_1)$. Hence,

$$(\lambda_1, \lambda_2) \overset{m}{\leqslant} (\lambda_1^*, \lambda_2^*) \qquad \text{does not imply that} \quad X \leqslant_{\ell r} X^*. \quad \blacksquare$$

3. LIKELIHOOD RATIO AND DISPERSIVE ORDERS FOR CONVOLUTIONS OF GAMMA DISTRIBUTIONS

In this section, stochastic orders for convolutions of gamma distributions will be considered. Specifically, it will be shown that a convolution of gamma distributions with a common shape parameter larger than or equal to 1 is larger (a) in likelihood ratio order (Theorem 3.4) and (b) dispersive order (Theorem 3.6) when the scale parameters of the gamma distributions are more dispersed in the sense of majorization.

The following results comparing two gamma distributions with either a common shape parameter or a common scale parameter using the dispersive order will be useful in proving Corollaries 3.1, 3.2, and Remark 3.2. Remark 3.2 and Theorem 3.4 respectively extend Theorem 3.1 and Theorem 3.2 to convolutions of gamma distributions.

THEOREM 3.1 (Saunders and Moran, 1978). *Let $F^{-1}(\alpha)$ denote the αth quantile of a gamma distribution with p.d.f. $f(y, a, \lambda) = \{\lambda^a y^{a-1}/\Gamma_{(a)}\}$* exp *$(-\lambda y)$, $y > 0$. Then, $F^{-1}(\beta) - F^{-1}(\alpha)$ is increasing with a while $F^{-1}(\beta)/F^{-1}(\alpha)$ is decreasing with a, for $\beta > \alpha$.*

The following companion result may be of independent interest.

THEOREM 3.2. *Under the assumptions of Theorem* 3.1, *$F^{-1}(\beta) - F^{-1}(\alpha)$ decreases with λ, for $\beta > \alpha$.*

The proof of Theorem 3.2 follows from a straightforward application of Theorem 2.1, and it will be omitted here.

Proposition 3.1 giving the distribution of a convolution of two gamma distributions with a common shape parameter greater then or equal to 1 leads to Theorem 3.3 establishing the likelihood ratio order for convolutions of two such gamma distributions.

PROPOSITION 3.1. *Let $Y_i, i=1,2$ be two independent gamma r.v.'s with Y_i having the p.d.f. $f(y; a, \lambda_i) = \{y^{a-1}\lambda_i^a/\Gamma(a)\} \exp(-\lambda_i y)$, $y>0$; 0, otherwise. Assume, without loss of generality, that $\lambda_1 > \lambda_2$. Then, the p.d.f. $h(y; a, \lambda_1, \lambda_2)$ of $Y_1 + Y_2$ is given by*

$$h(y; a, \lambda_1, \lambda_2) = \int_0^\theta g(w, y; a, \lambda_1, \lambda_2)\, k(w; a, \lambda_1, \lambda_2)\, dw, \tag{3.1}$$

where

$$k(w; a, \lambda_1, \lambda_2) = (1 - w^2/\theta^2)^{a-1}/\{B(a, a)\, 2^{2(a-1)}\theta\} \tag{3.2}$$

$$g(w, y; a, \lambda_1, \lambda_2) = \{1/\Gamma(2a)\}(\lambda_1\lambda_2)^a\, y^{2a-1} \cosh(wy) \exp(-cy/2),$$
$$y > 0, \qquad 0 < w < \theta \tag{3.3}$$

and where $c = (\lambda_1 + \lambda_2)$, $\theta = (\lambda_1 - \lambda_2)/2$ and $B(a, b)$ is the beta function with parameters $a > 0, b > 0$.

Proof. Routine calculations show that the p.d.f. $h(y; a, \lambda_1, \lambda_2)$ of $Y_1 + Y_2$ is given by $h(y; a, \lambda_1, \lambda_2) = \{y^{2a-1}/\Gamma(2a)\}[\int_0^1 \{x(1-x)\}^{a-1} (\lambda_1\lambda_2)^a \exp[-\{\lambda_1 x + \lambda_2(1-x)\}\, y]/B(a, a)]\, dx$, $y > 0$. Now, change the variable x of integration to $u = (2x-1)\,\theta$ to obtain

$$h(y; a, \lambda_1, \lambda_2) = \{y^{2a-1}/\Gamma(2a)\}(\lambda_1\lambda_2)^a \exp(-cy/2)$$
$$\times \int_{-\theta}^{\theta} (1 - u^2/\theta^2)^{a-1} \exp(-yu)/\{B(a, a)\, 2^{2(a-1)}\theta\}\, du.$$

To complete the proof, split the interval $(-\theta, \theta)$ of integration into $(-\theta, 0)$ and $[0, \theta)$, and to evaluate the integral over $(-\theta, 0)$, make the change of variable $u = -w$. ▮

THEOREM 3.3. *Let $Y_{\lambda_i}(Y_{\lambda_i^*})$, $i = 1, 2$ be independent gamma r.v.'s, $Y_{\lambda_i}(Y_{\lambda_i^*})$ having p.d.f. $f(y; a, \lambda_i) = \{y^{a-1}\lambda_i^a/\Gamma(a)\} \exp(-\lambda_i y)$, $y > 0$; 0, otherwise $f((y; a, \lambda_i^*))$. Then, $Y_{\lambda_i} + Y_{\lambda_2} \leqslant_{\ell r} Y_{\lambda_1^*} + Y_{\lambda_2^*}$ if $(\lambda_1, \lambda_2) \overset{m}{\leqslant} (\lambda_1^*, \lambda_2^*)$.*

Proof. Without loss of generality assume that $\lambda_1 \geqslant \lambda_2$ and $\lambda_1^* \geqslant \lambda_2^*$. Because $(\lambda_1, \lambda_2) \overset{m}{\leqslant} (\lambda_1^*, \lambda_2^*)$, it then follows that $\lambda_2^* \leqslant \lambda_2 \leqslant \lambda_1 \leqslant \lambda_1^*$. It is necessary to distinguish three cases: (i) $\lambda_1^* = \lambda_1$, making $(\lambda_1, \lambda_2) = (\lambda_1^*, \lambda_2^*)$, (ii) $\lambda_1^* \neq \lambda_1$, $\lambda_1 \neq \lambda_2$, and (iii) $\lambda_1^* \neq \lambda_1$ and $\lambda_1 = \lambda_2$. In case (i), the theorem is trivially true. Case (iii) will be dealt with at the end of the proof, and case (ii) will be treated first. Let $h(1, y) = h(y; a, \lambda_1, \lambda_2)$, $k(1, w) = k(w; a, \lambda_1, \lambda_2)$, and $g(w, y) = g(w, y; a, \lambda_1, \lambda_2)/(\lambda_1\lambda_2)^a$. Then, the p.d.f. of $Y_{\lambda_1} + Y_{\lambda_2}$, by (3.1), is $h(1, y) = (\lambda_1\lambda_2)^a \int_0^\theta g(w, y)\, k(1, w)\, dw$. Similarly, let the p.d.f. of $Y_{\lambda_1^*} + Y_{\lambda_2^*}$ be given by $h(2, y) = (\lambda_1^*\, \lambda_2^*)^a \int_0^{\theta^*} g(\lambda, y)\, k(2, w)\, dw$, where $h(2, y)=h(y; a, \lambda_1^*, \lambda_2^*)$, $k(2, y)=k(w; a, \lambda_1^*, \lambda_2^*)$ and $\theta^*=(\lambda_1^*-\lambda_2^*)/2$. To prove that $h(2, y)/h(1, y)$ is increasing in y (as required), one way is to appeal to the basic composition formula of Karlin (1968, p. 17) directly for the situation at hand, and show that $h(i, y)$ is TP_2 in $i \in \{1, 2\}$ and $y \in (0, \infty)$. The other one is to use a version of Theorem 1.C.11 in Shaked and Shantikumar (1994). To apply Karlin's result here, note that $\cosh(w_2)/\cosh(w_1)$ is increasing in y whenever $w_1 \leqslant w_2$. This follows from the identity $\cosh(w_2)/\cosh(w_1)=\cosh\{(w_2-w_1)\, y\}+\tanh(w_1 y)\sinh\{(w_2-w_1)\, y\}$ and the fact that all the three functions on the right hand side are increasing in y for $y>0$ and $w_1 \leqslant w_2$. Then, since $(\lambda_1, \lambda_2) \overset{m}{\leqslant} (\lambda_1^*, \lambda_2^*)$ implies $\theta^*>\theta$, it follows that the ratio $(1-w^2/\theta V^2)/(1-w^2/\theta^2)$ is increasing in w whenever $0<w<\theta$. Thus, $g(w, y)$ is TP_2 in $w \in (0, \theta)$ and $y \in (0, \infty)$, and $k(i, w)$ in $i \in \{1, 2\}$ and $y \in (0, \infty)$ (thus proving that $h(2, y)/h(1, y)$ is increasing in $y>0$).

In case (iii), $h(1, y)=\{y^{2a-1}(c/2)^{2a}/\Gamma(2a)\}\exp(-cy/2)$, $y>0$, where $c=\lambda_1^*+\lambda_2^*$. Thus,

$$h(2, y)/h(1, y)=\{(\lambda_1^*\,\lambda_2^*)^a/(c/2)^{2a}\}\int_0^{\theta^*} \times\{(1-x^2/\theta^{*2})^{a-1}/B(a, a)\, 2^{2(a-1)}\theta^*\}\cosh(xy)\, dx$$

which is increasing in y since (i) $\cosh(xy)$ is increasing in $y \in (0, \infty)$ for each $x \geqslant 0$, and (ii) $(1-x^2/\theta^{*2})^{a-1} \geqslant 0$ for $0<x<\theta^*$. ▮

THEOREM 3.4. *Let $Y_{\lambda_i}(Y_{\lambda_i^*})$, $i=1, ..., n$ be independent gamma r.v.'s with a common shape parameter, $Y_{\lambda_i}(Y_{\lambda_i^*})$ having the p.d.f. $f(y; a, \lambda_i)=\{1/\Gamma(a)\}\lambda_i^a y^{a-1}\exp(-\lambda_i y)$, $y>0$, $a \geqslant 1$ ($f(y; a, \lambda_i^*)$). Let $(\lambda_1, ..., \lambda_n) \overset{m}{\leqslant} (\lambda_1^*, ..., \lambda_n^*)$. Then, $\sum_{i=1}^n Y_{\lambda_i} \leqslant_{\ell r} \sum_{i=1}^n Y_{\lambda_i^*}$.*

Proof. It suffices to prove the theorem for the case that $(\lambda_1, \lambda_2) \overset{m}{\leqslant} (\lambda_1^*, \lambda_2^*)$ and $\lambda_i=\lambda_i^*$, $i=3, ..., n$. It is known (Theorem 2.1 (d) of Keilson and Sumita, 1982) that if $X \leqslant_{\ell r} Y$ and Z is independent of X and Y, and is dispersive, then $X+Z \leqslant_{\ell r} Y+Z$. Take $X=Y_{\lambda_1}+Y_{\lambda_2}$, $Y=Y_{\lambda_1^*}+Y_{\lambda_2^*}$ and $Z=\sum_{i=3}^n Y_{\lambda_i}=\sum_{i=3}^n Y_{\lambda_i^*}$. Then by Theorem 3.3 $X \leqslant_{\ell r}$, Y and Z, being a convolution of dispersive r.v.'s, is dispersive and independent of (X, Y). Now, the result follows from the above result of Keilson and Sumita. ▮

COROLLARY 3.1. *Let the hypotheses be as in Theorem 3.4, where $\lambda_1 \leqslant \cdots \leqslant \lambda_n$, $\lambda_1^* \leqslant \cdots \leqslant \lambda_n^*$ and $(\lambda_1, ..., \lambda_n) \overset{m}{\leqslant} (\lambda_1^*, ..., \lambda_n^*)$. Then. for $k=1, ..., n$, $\sum_{i=1}^k Y_{\lambda_i} \leqslant_{\ell r} \sum_{i=1}^k Y_{\lambda_i^*}$.*

The proof is omitted here since it is similar to that of Corollary 3.2.

Remark 3.1. The result in Theorem 3.4 is not true if the gamma distributions in the convolution have differing shape parameters. Consider the convolution of two gamma distributions, one with shape parameter 2 and scale parameter 3, the other with shape and scale parameters each equal to 1. Then, using Proposition 3.1, one can show that the p.d.f. $f_1(y)$ of the convolution is given by

$$f_1(y) = (9/4)\{\exp(-y) - \exp(-3y) - 2y\exp(-3y)\}, \qquad y > 0.$$

Similarly, the convolution of the two gamma distributions with shape parameters 2 and 1 and equal scale parameters 2 and 2 has p.d.f. $f_2(y)$ given by

$$f_2(y) = 4y^2\exp(-2y), \qquad y > 0.$$

Thus,

$$r(y) = f_1(y)/f_2(y) = (9/16)\{\exp(2y) - 1 - 2y\}/y^2\exp(y).$$

From this it follows that $r(0.5) = 0.98024 > 0.90824 = r(1)$, yet $(3, 1)$ majorizes $(2, 2)$. ∎

Remark 3.2. Results similar to the ones in Theorem 3.4 for convolutions of gamma distributions with a common scale parameter hold as follows. Let $Y_{\lambda_i}(Y_{\lambda_i^*})$, $i = 1, ..., n$ be independent gamma r.v.'s with Y_{λ_i} having the p.d.f. $f(y; a_i, b) = \{1/\Gamma(a_i\}\, b^{a_i}y^{a_i-1}\exp(-by)$, $y > 0$; 0, otherwise $(f(y; a_i^*, b))$, where $a_i \geqslant 1$ and $\lambda_i = 1/a_i$ $(a_i^* \geqslant 1$ and $\lambda_i^* = 1/a_i^*)$. Further, suppose $(\lambda_1, ..., \lambda_n) \overset{m}{\leqslant} (\lambda_1^*, ..., \lambda_n^*)$. Then, $1/\lambda_1^* + \cdots + 1/\lambda_n^* \geqslant 1/\lambda_1 + \cdots + 1/\lambda_n$ (by the Schur convexity of $\phi(\lambda_1, ..., \lambda_n) = \sum_{i=1}^n 1/\lambda_i$) and $\sum_{i=1}^n Y_{\lambda_i}$ $(\sum_{i=1}^n Y_{\lambda_i^*})$ has a gamma distribution with p.d.f. $f.(y; a, b)(f(y;\ a^*, b)$, where $a = \sum_{i=1}^n 1/\lambda_i$ $(a^* = \sum_{i=1}^n 1/\lambda_i^*)$. It follows from Theorem 3.1 that $\sum_{i=1}^n Y_{\lambda_i} \leqslant_{disp} \sum_{i=1}^n Y_{\lambda_i^*}$. Also, $f(y; a^*, b)/f(y; a, b) = \{\Gamma(a)/\Gamma(a^*\}\, y^{a^*-a}$ which is increasing in y. Hence $\sum_{i=1}^n Y_{\lambda_i} \leqslant_{\ell r} \sum_{i=1}^n Y_{\lambda_1^*}$. ∎

The next result shows that a convolution of two gamma distributions with a common shape parameter greater than or equal to 1 is more dispersed when the scale parameters are more dispersed in the sense of majorization.

THEOREM 3.5. *Let $c > 0$ be any number. For each $\lambda \in [c/2, c)$, let $F(y; a, \lambda)$ be the d.f. of the convolution of two gamma distributions with a common shape parameter a and scale parameters λ and $c - \lambda$. Then,*

$$F(\cdot; a, \lambda) \leqslant_{disp} F(\cdot; a, \lambda^*) \qquad \text{if} \quad \lambda, \lambda^* \in [c/2, c) \qquad \text{and} \qquad \lambda \leqslant \lambda^*.$$

Proof. The result will be proved by using Theorem 2.1. It will be shown first that

$$\begin{aligned}\{\lambda(c-\lambda)/(c-2\lambda)\}\ &F'(y; a, \lambda)\\ &= F(y; a, \lambda) - F(y; a+1, \lambda)\\ &= \{caf(y; a+1, \lambda) + \lambda(c-\lambda) f(y; a, \lambda)\, y\}/\{2a\lambda(c-\lambda)\}, \qquad \lambda \neq c/2\end{aligned} \tag{3.4}$$

and second that

$$f(y; a+1, \lambda)/f(y; a, \lambda) \qquad \text{is increasing in } y \tag{3.5}$$

so that

$$F'(y; a, \lambda)/f(y; a, \lambda) \qquad \text{is decreasing in } y.$$

Note that $F(y; a, \lambda) - F(y; a+1, \lambda) \geqslant 0$ (which follows from (3.5)) and $(c-2\lambda) \leqslant 0$.

From Proposition 3.1, the p.d.f. $f(y; a, \lambda)$ of $F(y; a, \lambda)$ can be written as

$$\begin{aligned}f(y; a, \lambda) = \sqrt{\pi}\, [\{\lambda(c-\lambda)\}^a/\Gamma(a)]\\ \cdot \{y/(2\lambda - c)\}^{a-1/2} \exp(-cy/2)\, I_{a-1/2}((\lambda - c/2)\, y),\end{aligned}$$

where

$$I_{a-1/2}(y) = \{2(y/2)^{a-1/2}/\sqrt{\pi}\, \Gamma(a)\} \int_0^1 (1-t^2)^{a-1} \cosh(ty)\, dt$$

is a modified Bessel function of the first kind. Differentiating $f(y; a, \lambda)$ with respect to λ and simplifying the result using the recurrence formula

$$I'_\nu(z) = I_{\nu+1}(z) + (\nu/z)\, I_\nu(z)$$

(with $z = (\lambda - c/2)\, y$ and $\nu = a - 1/2$) one gets

$$F'(y; a, \lambda) = \{(c-2\lambda)/\lambda(c-\lambda)\}\{F(y; a, \lambda) - F(y; a+1, \lambda)\}. \tag{3.6}$$

The first factor of the right hand side can be simplified by first finding the Laplace transform of $F(y; a, \lambda) - F(y; a+1, \lambda)$ and then inverting it. Towards this end, let the Laplace transform $\mathscr{L}\{G(\cdot)\}$ of $G(\cdot)$ be denoted by $g(s)$. Making use of the formulas

(1) $\mathscr{L}\{f(\cdot; a, \lambda)\} = \{\lambda(c-\lambda)\}^a/\{(s+\lambda)(s+c-\lambda)\}^a$,

(2) $\mathscr{L}\{\int_0^y G(x)\, dx\} = g(s)/s$,

and

(3) $\mathscr{L}\{yG(y)\} = -g'(s)$

one can show that

$$\begin{aligned}\mathscr{L}\{F(\cdot; a, \lambda) - F(\cdot; a+1, \lambda)\} \\ &= (c+s)\{\lambda(c-\lambda)\}^a/\{(\lambda+s)(c-\lambda+s)\}^{a+1} \\ &= \mathscr{L}\{caf(y; a+1, \lambda) + \lambda(c-\lambda)\, yf(y; a, \lambda)\}/\{2a(c-\lambda)\,\lambda\}.\end{aligned}$$

Inverting this Laplace transform and using the result in (3.6), one gets (3.4).

To show (3.5), note that the convolution $F(y; a+1, \lambda)$ is the convolution of $F(y; a, \lambda)$ and two exponentials with scale parameters λ, and $c-\lambda$. Let X, Y, U and V respectively be the r.v.'s associated with these convolutions. Let also W be a r.v. degenerate at 0. Then $W \leqslant_{\ell r} U+V$, and by the Keilson and Sumita (1982) result quoted in the proof of Theorem 3.4,

$$Y \stackrel{d}{=} Y + W \leqslant_{\ell r} Y + U + V \stackrel{d}{=} X,$$

where $S \stackrel{d}{=} T$ denotes that r.v.'s S and T have the same distribution. This proves (3.5). A different proof of (3.5) can be based on Theorem 1.C.5 in Shaked and Shantikumar (1994). ▮

The second major result of this section which generalizes Theorem 3.5 is proved next.

THEOREM 3.6. *Let* $Y_{\lambda_i}(Y_{\lambda_i^*})$, $i=1, ..., n$ *be independent r.v.'s,* $Y_{\lambda_i}(Y_{\lambda_i^*})$ *having a gamma distribution with p.d.f.* $f(y; a, \lambda_i) = \{y^{a-1}\lambda_i^a/\Gamma(a)\}\exp(-\lambda_i y)$, $y>0$ $(f(y; a, \lambda_i^*))$, *where* $a \geqslant 1$. *Let* $(\lambda_1 ..., \lambda_n) \stackrel{m}{\leqslant} (\lambda_1^*, ..., \lambda_n^*)$. *Then,* $\sum_{i=1}^n Y_{\lambda_i} \leqslant_{disp} \sum_{i=1}^n Y_{\lambda_i^*}$.

Proof. It suffices to prove the result for the case $(\lambda_1, \lambda_2) \stackrel{m}{\leqslant} (\lambda_1^*, \lambda_2^*)$ and $\lambda_i = \lambda_i^*$, $i=3, ..., n$. Please see the argument for this given in the proof of Theorem 2.3. By Theorem 3.5,

$$Y_{\lambda_1} + Y_{\lambda_2} \leqslant_{disp} Y_{\lambda_1^*} + Y_{\lambda_2^*}. \tag{3.7}$$

Since a gamma distribution with a shape parameter greater than or equal to 1 is dispersive, it follows from the Theorem S(2) of Lewis and Thomson (1981) (quoted in the proof of Theorem 2.3) that $\sum_{i=3}^n Y_{\lambda_i} = \sum_{i=3}^n Y_{\lambda_i^*}$ is dispersive. This result, (3.7), and the definition of a dispersive distribution now show that $\sum_{i=1}^{n} Y_{\lambda_i} \leqslant_{disp} \sum_{i=1}^{n} Y_{\lambda_i^*}$. ▮

COROLLARY 3.2. *Let the hypotheses be as in Theorem* 3.6, *where* $\lambda_1 \leqslant \cdots \leqslant \lambda_n$, $\lambda_1^* \leqslant \cdots \leqslant \lambda_n^*$ *and* $(\lambda_1, ..., \lambda_n) \overset{m}{\leqslant} (\lambda_1^*, ...\lambda_n^*)$. *Then, for* $k = 1, ..., n$ $\sum_{i=1}^{k} Y_{\lambda^i} \leqslant_{disp} \sum_{i=1}^{k} Y_{\lambda_i^*}$.

Proof. Let $\theta_k = \sum_{i=1}^{k} \lambda_i - \sum_{i=1}^{k} \lambda_i^*$. Then $(\lambda_1, ..., \lambda_k) \overset{m}{\leqslant} (\lambda_1^*, ..., \lambda_k^* + \theta_k)$ since $(\lambda_1, ..., \lambda_n) \overset{m}{\leqslant} (\lambda_1^*, ..., \lambda_n^*)$. Let $Y_{\lambda_k^* + \theta_k}$ be a gamma r.v. with shape parameter a and scale parameter $\lambda_k^* + \theta_k$. It then follows from Theorem 3.6 that $\sum_{i=1}^{k} Y_{\lambda_i} \leqslant_{disp} \sum_{i=1}^{k-1} Y_{\lambda_i^*} + Y_{\lambda_k^* + \theta_k}$, $k \geqslant 2$. Finally, $\sum_{i=1}^{k-1} Y_{\lambda_i^*} + Y_{\lambda_k^* + \theta_k}$ $\leqslant_{disp} \sum_{i=1}^{k} Y_{\lambda_i^*}$, since by Theorem 3.2, $Y_{\lambda_k^* + \theta_k} \leqslant_{disp} Y_{\lambda_k^*}$ and $\sum_{i=1}^{k-1} Y_{\lambda_i^*}$ is dispersive. For the case $k = 1$, the result follows since $\lambda_1 = \lambda_1^* + \theta_1$ and Y_{λ_1} $= Y\lambda_1^* + \theta_1 \leqslant_{disp} Y_{\lambda_1^*}$. ▮

Remark 3.3. Theorem 3.6 is not true for a convolution of gamma distributions with differing shape parameters. Let $Y_{a,\lambda}$ denote a gamma r.v. with a shape parameter a and scale parameter λ. Consider the two r.v's $U = Y_{100,5} + Y_{1,1}$ and $V = Y_{100,4} + Y_{1,2}$ where the r.v.'s in each sum are independent. Then, $(4, 2) \overset{m}{\leqslant} (5.1)$ and yet Variance $(U) = 100/25 + 1/1 = 5$ $< 6.5 = 100/16 + 1/4 =$ Variance (V). Now, $V \leqslant_{disp} U$ implies that Variance $(V) \leqslant$ Variance (U). Thus, U cannot be greater than V in dispersive order.

Consider a convolution of an arbitrary number n of gamma distributions with a common shape parameter $a \geqslant 1$ and scale parameters λ_i's. Then, as applications of Theorems 3.4 and 3.6, it is possible to construct lower bounds for (a) the hazard function and (b) the difference between any two quantiles of the above convolution in terms of corresponding quantities of a gamma distribuion with shape parameter a and scale parameter $\bar{\lambda}$ as follows, where $\bar{\lambda} = \sum_{i=1}^{n} \lambda_i / n$. Since the likelihood ratio order implies the hazard rate order, and $(\bar{\lambda}, ..., \bar{\lambda} \overset{m}{\leqslant} (\lambda_1, ..., \lambda_2)$, Theorem 3.4 implies that a lower bound for the hazard function of the convolution at a point x is given by the hazard function of the gamma distribution at x. Similarly, Theorem 3.6 implies that the difference between any two quantiles of the convolution is larger than the difference between the correspoinding quantiles of the gamma distribution.

ACKNOWLEDGMENTS

The author thanks two referees for their helpful comments and suggestions which greatly improved the paper.

REFERENCES

1. I. Bagai and S. C. Kochar, On tail-ordering and comparisions of failure rates, *Commun. Statist. Theory Methods* **15** (1986), 1377–1388.
2. R. E. Barlow, A. W. Marshall, and F. Proschan, Properties of probability distributions with monotone hazard rates, *Ann. Math. Statist.* **34** (1963), 375–399.
3. P. J. Boland, E. El-Neweihi, and F. Proschan, Schur properties of convolutions of exponential and geometric random variables, *J. Multivariate Anal.* **48** (1994), 157–167.
4. S. Karlin, "Total Positivity," Vol. 1, Stanford Univ. Press, Stanford, CA, 1968.
5. J. Keilson and U. Sumita, Uniform stochastic ordering and related inequalities, *Canad. J. Statist.* **10** (1982), 181–198.
6. S. C. Kochar and C. Ma, Dispersive ordering of convolutions of exponential random variables, *Statist. Probab. Lett.* **43** (1999a), 321–324.
7. S. C. Kochar and C. Ma, Erratum to [6], *Statist. Probab. Lett.* **45** (1999b), 283.
8. T. Lewis and J. W. Thompson, Dispersive distributions and connections between dispersivity and strong unimodality, *J. Appl. Probab.* **18** (1981), 76–90.
9. A. W. Marshall and I. Olkin, "Inequalities: Theory of Majorization and Its Applications," Academic Press, New York, 1979.
10. I. W. Saunders and P. A. P. Moran, On the quantiles of the gamma and *F* distributions, *J. Appl. Probab.* **15** (1978), 426–432.
11. M. Shaked and J. G. Shantikumar, "Stochastic Orders and Their Applications," Academic Press, New York, 1994.