

## STOCHASTIC ORDERING OF EXPONENTIAL FAMILY DISTRIBUTIONS AND THEIR MIXTURES

YAMING YU,\* *University of California, Irvine*

### Abstract

We investigate stochastic comparisons between exponential family distributions and their mixtures with respect to the usual stochastic order, the hazard rate order, the reversed hazard rate order, and the likelihood ratio order. A general theorem based on the notion of *relative log-concavity* is shown to unify various specific results for the Poisson, binomial, negative binomial, and gamma distributions in recent literature. By expressing a convolution of gamma distributions with arbitrary scale and shape parameters as a scale mixture of gamma distributions, we obtain comparison theorems concerning such convolutions that generalize some known results. Analogous results on convolutions of negative binomial distributions are also discussed.

*Keywords:* Binomial mixture; convolution; gamma mixture; hazard rate order; likelihood ratio order; log-concavity; negative binomial mixture; Poisson mixture; stochastic comparison; stochastic orders

2000 Mathematics Subject Classification: Primary 60E15

Secondary 60E05

### 1. Stochastic orders and some general observations

The study of stochastic orders has received attention in diverse areas including economics, operations research, reliability, and statistics (e.g. survival analysis). For book-length treatments of both theory and applications, see Shaked and Shanthikumar (1994), (2007). This paper is mainly concerned with four orders, namely, the usual stochastic order (denoted by ' $\leq_{st}$ '), the hazard rate order (denoted by ' $\leq_{hr}$ '), the reversed hazard rate order (denoted by ' $\leq_{rh}$ '), and the likelihood ratio order (denoted by ' $\leq_{lr}$ '). We recall the familiar definitions.

**Definition 1.** Let  $X$  and  $Y$  be continuous random variables on  $\mathbb{R}$  with probability density functions (PDFs) or discrete random variables on  $\mathbb{Z}$  with probability mass functions (PMFs)  $f(x)$  and  $g(x)$ , respectively. Denote their respective cumulative distribution functions (CDFs) by  $F(x)$  and  $G(x)$ .

- $X$  is said to be smaller than  $Y$  in the usual *stochastic order* (written  $X \leq_{st} Y$ ) if  $\bar{F}(x) \leq \bar{G}(x)$  for all  $x$ , where  $\bar{F}(x) = 1 - F(x)$  and  $\bar{G}(x) = 1 - G(x)$ .
- $X$  is said to be smaller than  $Y$  in the *hazard rate order* (written  $X \leq_{hr} Y$ ) if  $f(x)/\bar{F}(x) \geq g(x)/\bar{G}(x)$  for all  $x$ .
- $X$  is said to be smaller than  $Y$  in the *reversed hazard rate order* (written  $X \leq_{rh} Y$ ) if  $f(x)/F(x) \leq g(x)/G(x)$  for all  $x$ .

Received 4 April 2008; revision received 27 January 2009.

\* Postal address: Department of Statistics, University of California, Irvine, CA 92697-1250, USA.

Email address: yamingy@uci.edu

- $X$  is said to be smaller than  $Y$  in the *likelihood ratio order* (written  $X \leq_{lr} Y$ ) if the likelihood ratio  $f(x)/g(x)$  is a monotone decreasing function on the set  $\{x : f(x) > 0 \text{ or } g(x) > 0\}$ . By convention,  $a/0 = \infty$  whenever  $a > 0$ .

As is well known,  $X \leq_{lr} Y$  implies that  $X \leq_{hr} Y$  and  $X \leq_{rh} Y$ , either of which in turn implies that  $X \leq_{st} Y$ . Further basic properties of these orders can be found in Shaked and Shanthikumar (1994).

Despite their importance, it can be nontrivial to verify the relations ‘ $\leq_{st}$ ’, ‘ $\leq_{hr}$ ’, ‘ $\leq_{rh}$ ’, or ‘ $\leq_{lr}$ ’, e.g. when the relevant distributions are not in closed form. This work provides some simple conditions that unify and generalize many results for specific distributions in recent literature. The following *relative log-concavity order*, introduced in Whitt (1985) (see also Yu (2008)), plays a critical role in the development.

**Definition 2.** Let  $X$  and  $Y$  be continuous or discrete random variables with PDFs or, respectively, PMFs  $f(x)$  and  $g(x)$ , respectively. We say that  $X$  is *log-concave relative to  $Y$*  (written  $X \leq_{lc} Y$ ) if

1. the support of  $X$ ,  $\text{supp}(X) = \{x : f(x) > 0\}$ , and the support of  $Y$ ,  $\text{supp}(Y) = \{x : g(x) > 0\}$ , are both intervals on  $\mathbb{R} (\mathbb{Z})$ ;
2.  $\text{supp}(X) \subset \text{supp}(Y)$ ;
3.  $\log(f(x)/g(x))$  is a concave function on  $\text{supp}(X)$ .

The order ‘ $\leq_{lc}$ ’ provides a way of deriving conditions that imply the four ground-level orders ‘ $\leq_{st}$ ’, ‘ $\leq_{hr}$ ’, ‘ $\leq_{rh}$ ’, and ‘ $\leq_{lr}$ ’. This is analogous to gaining an understanding of the monotonicity properties of a function by studying its second derivative. We summarize some general observations below.

**Theorem 1.** Let the random variables  $X$  and  $Y$  have PDFs  $f(x)$  and  $g(x)$ , respectively, both supported on  $(0, \infty)$ . Assume that the log density ratio  $l(x) = \log(f(x)/g(x))$  is continuous and, moreover, concave, i.e.  $X \leq_{lc} Y$ . Then,

1.  $X \leq_{st} Y$  and  $X \leq_{hr} Y$  are equivalent, and each holds if and only if  $\lim_{x \downarrow 0} l(x) \geq 0$ ;
2. assuming that  $l(x)$  is continuously differentiable,  $X \leq_{lr} Y$  and  $X \leq_{rh} Y$  are equivalent, and each holds if and only if  $\lim_{x \downarrow 0} l'(x) \leq 0$ .

*Proof.* 1. Let  $A = \{x : l(x) \geq 0\} = \{x : f(x) \geq g(x), x > 0\}$ . Because  $l(x)$  is concave,  $A$  is an interval. We first show that  $X \leq_{st} Y$  is equivalent to  $\lim_{x \downarrow 0} l(x) \geq 0$ . If  $\lim_{x \downarrow 0} l(x) \geq 0$  then it is easy to see that the left endpoint of  $A$  is 0. That is,  $f(x) - g(x)$  changes sign at most once from + to - as  $x$  increases from 0 to  $\infty$ ; it follows that  $F(x) - G(x) = \int_0^x (f(u) - g(u)) du$  does not change sign at all, i.e.  $F(x) \geq G(x)$  for all  $x$ , and that  $X \leq_{st} Y$ , by definition. Conversely, if  $X \leq_{st} Y$  then  $\int_0^x (f(u) - g(u)) du \geq 0$  for all  $x$ , forcing the left endpoint of  $A$  to 0, which implies that  $\lim_{x \downarrow 0} l(x) \geq 0$ . Note that this limit exists by the concavity of  $l(x)$ .

Concerning the hazard rate order, we only need to show that

$$X \leq_{st} Y \implies X \leq_{hr} Y,$$

since the implication

$$X \leq_{hr} Y \implies X \leq_{st} Y$$

is well known. By definition, if  $X \leq_{st} Y$  then  $\bar{F}(x) \leq \bar{G}(x)$  for all  $x$ . Given  $x_0 > 0$ , if  $f(x_0) \geq g(x_0)$  then  $f(x_0)/\bar{F}(x_0) \geq g(x_0)/\bar{G}(x_0)$ . Otherwise,  $f(x_0) < g(x_0)$ , i.e.  $x_0 \notin A$ .

As before, since  $X \leq_{st} Y$ , the left endpoint of  $A$  must be 0. Hence,  $l(x) < 0$  for all  $x \geq x_0$ . If there exist some  $x_2 > x_1 \geq x_0$  such that  $l(x_2) > l(x_1)$  then, by the concavity of  $l(x)$ , for all  $x \leq x_1$ , we have

$$l(x) \leq l(x_1) + (x - x_1) \frac{l(x_2) - l(x_1)}{x_2 - x_1} < 0,$$

i.e.  $l(x) < 0$  for all  $x$ , a contradiction. Thus,  $l(x)$  (or  $f(x)/g(x)$ ) decreases on  $[x_0, \infty)$ , and, consequently,

$$\frac{f(x_0)}{\bar{F}(x_0)} = \frac{f(x_0)}{\int_{x_0}^{\infty} f(u) \, du} \geq \frac{f(x_0)}{\int_{x_0}^{\infty} g(u) f(x_0)/g(x_0) \, du} = \frac{g(x_0)}{\bar{G}(x_0)}.$$

That is, the hazard rate of  $X$  is always greater than or equal to that of  $Y$ .

2. Note that  $l'(x)$  decreases in  $x$  since  $l(x)$  is concave; therefore, to ensure a monotone density ratio, or  $l'(x) \leq 0$  for all  $x$ , we need only  $\lim_{x \downarrow 0} l'(x) \leq 0$ . That is,

$$X \leq_{lr} Y \iff \lim_{x \downarrow 0} l'(x) \leq 0.$$

Concerning the reversed hazard rate order, we only need to show that

$$X \leq_{rh} Y \implies X \leq_{lr} Y,$$

since the implication

$$X \leq_{lr} Y \implies X \leq_{rh} Y$$

is known. Assume the contrary, i.e.  $X \leq_{rh} Y$  but  $X \not\leq_{lr} Y$ . Then, by the discussion above,  $\lim_{x \downarrow 0} l'(x) > 0$ , and, by continuity, there exists  $\varepsilon > 0$  such that  $l'(x) > 0$  for all  $x \in (0, \varepsilon]$ . That is,  $f(x)/g(x)$  strictly increases on  $x \in (0, \varepsilon]$ . Thus,

$$\frac{f(\varepsilon)}{F(\varepsilon)} = \frac{f(\varepsilon)}{\int_0^\varepsilon f(u) \, du} > \frac{f(\varepsilon)}{\int_0^\varepsilon g(u) f(\varepsilon)/g(\varepsilon) \, du} = \frac{g(\varepsilon)}{G(\varepsilon)},$$

which contradicts the definition of  $X \leq_{rh} Y$ .

A discrete version of Theorem 1 is stated as follows.

**Theorem 2.** *Let the random variables  $X$  and  $Y$  have PMFs  $f(x)$  and  $g(x)$ , respectively, both supported on the same  $\mathbb{Z}_+ = \{0, 1, \dots\}$  (or  $\{0, 1, \dots, n\}$  for some  $n > 0$ ). Assume that  $X \leq_{lc} Y$ . Then*

1.  $X \leq_{st} Y$  and  $X \leq_{hr} Y$  are equivalent, and each holds if and only if  $f(0)/g(0) \geq 1$ ;
2.  $X \leq_{lr} Y$  and  $X \leq_{rh} Y$  are equivalent, and each holds if and only if  $f(1)/g(1) \leq f(0)/g(0)$ .

Basically, if  $X \leq_{lc} Y$  then  $X \leq_{st} Y$ ,  $X \leq_{hr} Y$ ,  $X \leq_{rh} Y$ , and  $X \leq_{lr} Y$  are all determined by the behavior of  $\Pr(X = x)/\Pr(Y = x)$  near the left endpoint  $x = 0$ .

**Example 1.** Let  $Y \sim \text{Bin}(n, p)$ ,  $p \in (0, 1)$ , and let  $X = \sum_{i=1}^n B_i$ , where the  $B_i$  are independent Bernoulli random variables, i.e.  $\Pr(B_i = 1) = 1 - \Pr(B_i = 0) = p_i$ ,  $i = 1, \dots, n$ . In the context of software testing, Boland *et al.* (2002) considered comparisons between  $X$  and  $Y$  with respect to several stochastic orders. We note that Theorem 2 gives an alternative, somewhat faster, derivation of some of their results. Our starting point is the well-known relation  $X \leq_{lc} Y$ , which is equivalent to Newton’s inequalities (see Hardy *et al.* (1964, p. 52)). Thus, Theorem 2

and simple calculations yield

1.  $X \leq_{st} Y$  ( $X \leq_{hr} Y$ ) if and only if  $p \geq 1 - (\prod_{i=1}^n (1 - p_i))^{1/n}$ ;
2.  $X \leq_{lr} Y$  ( $X \leq_{rh} Y$ ) if and only if  $p \geq 1 - n / (\sum_{i=1}^n (1 - p_i)^{-1})$ .

If we let  $X' = n - X$  and  $Y' = n - Y$ , then, obviously,  $X' \leq_{lc} Y'$  and

$$\begin{aligned} X' \leq_{st} Y' &\iff Y \leq_{st} X, & X' \leq_{hr} Y' &\iff Y \leq_{rh} X, \\ X' \leq_{lr} Y' &\iff Y \leq_{lr} X, & X' \leq_{rh} Y' &\iff Y \leq_{hr} X. \end{aligned}$$

Applying Theorem 2 to  $X'$  and  $Y'$ , we obtain

1.  $Y \leq_{st} X$  ( $Y \leq_{rh} X$ ) if and only if  $p \leq (\prod_{i=1}^n p_i)^{1/n}$ ;
2.  $Y \leq_{lr} X$  ( $Y \leq_{hr} X$ ) if and only if  $p \leq n / (\sum_{i=1}^n p_i^{-1})$ .

Our result,

$$X \leq_{hr} Y \iff p \geq 1 - \left(\prod (1 - p_i)\right)^{1/n}, \tag{1}$$

corrects a slight oversight in Boland *et al.* (2002, Theorem 1(iv)(b)). Basically, Boland *et al.* (2002) found the correct criterion for  $Y' \leq_{hr} X'$  and claimed that the same criterion holds for  $X \leq_{hr} Y$ . However,  $Y' \leq_{hr} X'$  is equivalent to  $X \leq_{rh} Y$ , not  $X \leq_{hr} Y$ . This explains the discrepancy between (1) and Theorem 1(iv)(b) of Boland *et al.* (2002).

Theorems 1 and 2 are particularly useful for comparing exponential family distributions with their mixtures, as will be illustrated in Section 2, where various specific results concerning Poisson, binomial, negative binomial, and gamma distributions are unified and generalized. In Section 3 we apply the results of Section 2 to convolutions of gamma distributions, which are useful in modeling, for example, the lifetime of a redundant standby system without repairing (see Bon and Păltănea (1999)). It is shown that, if  $S = \sum_{i=1}^n \beta_i S_i$ , where  $S_i \sim \text{gamma}(\alpha_i, 1)$  independently,  $\alpha_i, \beta_i > 0$ , and  $T = \beta \sum_{i=1}^n S_i$ ,  $\beta > 0$ , then

$$T \leq_{st} S \iff T \leq_{hr} S \iff \beta \leq \left(\prod_{i=1}^n \beta_i^{\alpha_i}\right)^{1/\alpha_+},$$

where  $\alpha_+ = \sum_{i=1}^n \alpha_i$ . Moreover,

$$T \leq_{lr} S \iff T \leq_{rh} S \iff \beta \leq \alpha_+ \left(\sum_{i=1}^n \frac{\alpha_i}{\beta_i}\right)^{-1}.$$

In Section 4, convolutions of negative binomial distributions are considered and results analogous to those of Section 3 are obtained.

## 2. Comparing exponential family distributions with their mixtures

Consider the density of an exponential family,

$$f(x; \theta) = f_0(x)e^{b(\theta)x}h(\theta), \tag{2}$$

where  $\theta$  is a parameter, and, for simplicity, assume that the support of  $f(x; \theta)$  is the interval  $(0, \infty)$  (regardless of the value of  $\theta$ ). Let  $g(x) = \int f(x; t) d\mu(t)$  be the mixture of  $f(x; \theta)$

with respect to a probability distribution  $\mu$  on  $\theta$ . Shaked (1980) considered the comparison between  $g(x)$  and  $f(x; \theta)$  with a fixed  $\theta$ , focusing on the case when the two distributions have the same mean. Our comparisons here are in terms of ‘ $\leq_{st}$ ’, ‘ $\leq_{hr}$ ’, ‘ $\leq_{rh}$ ’, and ‘ $\leq_{lr}$ ’. As noted in Whitt (1985),

$$\log\left(\frac{g(x)}{f(x; \theta)}\right) = \log\left(\int \frac{e^{(b(t)-b(\theta))x} h(t)}{h(\theta)} d\mu(t)\right)$$

is a convex function of  $x$ , i.e.  $l(x) = \log(f(x; \theta)/g(x))$  is concave. (This holds because log-convexity is closed under mixture.) We may compute

$$\lim_{x \downarrow 0} l(x) = -\log\left(\int \frac{h(t)}{h(\theta)} d\mu(t)\right)$$

and

$$\lim_{x \downarrow 0} l'(x) = \frac{\int (b(\theta) - b(t))h(t) d\mu(t)}{\int h(t) d\mu(t)},$$

provided that the interchange of limit (differentiation) and integration is valid. Thus, if the random variables  $X$  and  $Y$  have densities  $f(x; \theta)$  and  $g(x)$ , respectively, then, by Theorem 1,

1.  $X \leq_{st} Y$  ( $X \leq_{hr} Y$ ) if and only if

$$\int h(t) d\mu(t) \leq h(\theta); \tag{3}$$

2.  $X \leq_{lr} Y$  ( $X \leq_{rh} Y$ ) if and only if

$$b(\theta) \leq \frac{\int b(t)h(t) d\mu(t)}{\int h(t) d\mu(t)}. \tag{4}$$

If  $f(x; \theta)$  is a discrete PMF on  $\mathbb{Z}_+$  then, by Theorem 2,

1.  $X \leq_{st} Y$  ( $X \leq_{hr} Y$ ) if and only if

$$\int h(t) d\mu(t) \leq h(\theta); \tag{5}$$

2.  $X \leq_{lr} Y$  ( $X \leq_{rh} Y$ ) if and only if

$$e^{b(\theta)} \leq \frac{\int h(t)e^{b(t)} d\mu(t)}{\int h(t) d\mu(t)}. \tag{6}$$

Let us illustrate (5) and (6) with some discrete examples. In Examples 2 and 3, below, certain results given in Misra *et al.* (2003) are recovered concerning the comparisons of Poisson and binomial distributions with their mixtures; in Example 4 we consider the negative binomial and recover analogous results given in Alamatsaz and Abbasi (2008). In addition to ‘ $\leq_{st}$ ’ and ‘ $\leq_{lr}$ ’ studied in Misra *et al.* (2003) and Alamatsaz and Abbasi (2008), comparisons in terms of ‘ $\leq_{hr}$ ’ and ‘ $\leq_{rh}$ ’ are also included.

**Example 2.** Let  $X$  have a Poisson distribution  $Po(\lambda)$ ,  $\lambda > 0$ , whose PMF is

$$f(x; \lambda) = \frac{1}{x!} \lambda^x e^{-\lambda}, \quad x = 0, 1, \dots,$$

or, in the form of (2),

$$f(x; \lambda) = \frac{1}{x!} e^{xb(\lambda)} h(\lambda)$$

with  $b(\lambda) = \log(\lambda)$  and  $h(\lambda) = e^{-\lambda}$ . Suppose that  $Y$  is a mixture of  $Po(t)$  with respect to a distribution  $\mu(t)$  on  $t \in (0, \infty)$ . Then, by (5) and (6), we have

1.  $X \leq_{st} Y$  ( $X \leq_{hr} Y$ ) if and only if

$$\int e^{-t} d\mu(t) \leq e^{-\lambda};$$

2.  $X \leq_{lr} Y$  ( $X \leq_{rh} Y$ ) if and only if

$$\lambda \leq \frac{\int t e^{-t} d\mu(t)}{\int e^{-t} d\mu(t)}.$$

**Example 3.** Let  $X$  have a binomial distribution with parameters  $(n, p)$ , where  $0 < p < 1$  and  $n$  is a positive integer. The PMF of  $X$  is

$$f(x; p) = \binom{n}{x} p^x (1-p)^{n-x}, \quad x = 0, \dots, n,$$

or, in the form of (2),

$$f(x; p) = \binom{n}{x} e^{xb(p)} h(p)$$

with  $b(p) = \log(p/(1-p))$  and  $h(p) = (1-p)^n$ . Suppose that  $Y$  is a mixture of  $Bin(n, t)$  with respect to a distribution  $\mu(t)$  on  $t \in (0, 1)$ . Then, after simple algebra, (5) and (6) give

1.  $X \leq_{st} Y$  ( $X \leq_{hr} Y$ ) if and only if

$$\int (1-t)^n d\mu(t) \leq (1-p)^n;$$

2.  $X \leq_{lr} Y$  ( $X \leq_{rh} Y$ ) if and only if

$$p \leq \frac{\int t(1-t)^{n-1} d\mu(t)}{\int (1-t)^{n-1} d\mu(t)}.$$

By considering  $X' = n - X$  and  $Y' = n - Y$ , we obtain

1.  $Y \leq_{st} X$  ( $Y \leq_{rh} X$ ) if and only if

$$\int t^n d\mu(t) \leq p^n;$$

2.  $Y \leq_{lr} X$  ( $Y \leq_{hr} X$ ) if and only if

$$p \geq \frac{\int t^n d\mu(t)}{\int t^{n-1} d\mu(t)}.$$

**Example 4.** Let  $X$  have a negative binomial distribution, denoted by  $NB(k, p)$ , where  $k$  (not necessarily an integer) is positive and  $0 < p < 1$ . The PMF of  $X$  is

$$f(x; p) = \binom{k+x-1}{x} p^k (1-p)^x, \quad x = 0, 1, \dots,$$

or, in the form of (2),

$$f(x; p) = \binom{k+x-1}{x} e^{xb(p)} h(p)$$

with  $b(p) = \log(1-p)$  and  $h(p) = p^k$ . Suppose that  $Y$  is a mixture of  $NB(k, t)$  with respect to a distribution  $\mu(t)$  on  $t \in (0, 1)$ . Then (5) and (6) give

1.  $X \leq_{st} Y$  ( $X \leq_{hr} Y$ ) if and only if

$$\int t^k d\mu(t) \leq p^k; \tag{7}$$

2.  $X \leq_{lr} Y$  ( $X \leq_{rh} Y$ ) if and only if

$$p \geq \frac{\int t^{k+1} d\mu(t)}{\int t^k d\mu(t)}. \tag{8}$$

Let us illustrate (3) and (4) with a continuous example.

**Example 5.** Let  $X$  have a gamma distribution,  $\text{gamma}(\alpha, \beta)$ ,  $\alpha > 0$ ,  $\beta > 0$ , which is parameterized so that the PDF is

$$f(x; \beta) = \Gamma(\alpha)^{-1} \beta^{-\alpha} x^{\alpha-1} e^{-x/\beta}, \quad x > 0,$$

or, in the form of (2),

$$f(x; \beta) = \Gamma(\alpha)^{-1} x^{\alpha-1} e^{xb(\beta)} h(\beta)$$

with  $b(\beta) = -\beta^{-1}$  and  $h(\beta) = \beta^{-\alpha}$ . Suppose that  $Y$  is a mixture of  $\text{gamma}(\alpha, t)$  with respect to a distribution  $\mu(t)$  on  $t \in (0, \infty)$ . Then (3) and (4) give

1.  $X \leq_{st} Y$  ( $X \leq_{hr} Y$ ) if and only if

$$\int t^{-\alpha} d\mu(t) \leq \beta^{-\alpha}; \tag{9}$$

2.  $X \leq_{lr} Y$  ( $X \leq_{rh} Y$ ) if and only if

$$\beta \int t^{-\alpha-1} d\mu(t) \leq \int t^{-\alpha} d\mu(t) < \infty. \tag{10}$$

Note that, unlike previous examples, this is a continuous case and the regularity conditions (interchange of limit (differentiation) and integration) required in the derivation of (9) and (10) need to be verified. For example, to establish (9), we note that

$$\begin{aligned} \lim_{x \downarrow 0} \frac{f(x; \beta)}{\int f(x; t) d\mu(t)} &= \lim_{x \downarrow 0} \frac{\beta^{-\alpha} e^{-x/\beta}}{\int t^{-\alpha} e^{-x/t} d\mu(t)} \\ &= \frac{\beta^{-\alpha}}{\lim_{x \downarrow 0} \int t^{-\alpha} e^{-x/t} d\mu(t)} \\ &= \frac{\beta^{-\alpha}}{\int t^{-\alpha} d\mu(t)}, \end{aligned}$$

where we appeal to the monotone convergence theorem for the last equality.

### 3. Convolutions of gamma distributions

Example 5 in Section 2 enables us to compare a sum of independent gamma random variables with a particular gamma variate. To achieve this, we exploit a connection between such a convolution of gamma distributions and a mixture of gamma distributions. Specifically, let  $S = \sum_{i=1}^n \beta_i S_i$ , where  $S_i \sim \text{gamma}(\alpha_i, 1)$  independently and  $\beta_i > 0, i = 1, \dots, n$ . Let  $T \sim \text{gamma}(\sum_{i=1}^n \alpha_i, \beta), \beta > 0$ . We are interested in conditions on  $\beta$  that ensure that  $T \leq_{\text{st}} S, T \leq_{\text{hr}} S, T \leq_{\text{rh}} S$ , or  $T \leq_{\text{lr}} S$ . Relevant works on this problem include Boland *et al.* (1994), Bon and Păltănea (1999), Kochar and Ma (1999), Korwar (2002), Khaledi and Kochar (2004). In particular, using *majorization* techniques (see Marshall and Olkin (1979)), Boland *et al.* (1994) showed that, in the case where  $\alpha_i = 1$ , i.e. when  $S$  is a sum of independent exponential variables with possibly different scales, we have

$$\beta \leq \frac{n}{\sum_{i=1}^n \beta_i^{-1}} \implies T \leq_{\text{lr}} S.$$

Bon and Păltănea (1999) extended this to (still with  $\alpha_i = 1$ )

$$T \leq_{\text{st}} S \iff T \leq_{\text{hr}} S \iff \beta \leq \left( \prod_{i=1}^n \beta_i \right)^{1/n}, \tag{11}$$

$$T \leq_{\text{lr}} S \iff \beta \leq \frac{n}{\sum_{i=1}^n \beta_i^{-1}}. \tag{12}$$

The results of Korwar (2002) and Khaledi and Kochar (2004) imply that the ‘ $\Leftarrow$ ’ parts of (11) and (12) hold when all the  $\alpha_i$  are equal and their common value  $\alpha \geq 1$ . As an application of the calculations in Sections 1 and 2, we give a further extension for general  $\alpha_i > 0$ . Such results are of interest in reliability theory as they provide convenient bounds (for example) on the hazard rate function of  $S$  through the simpler hazard rate function of  $T$  (see Bon and Păltănea (1999)).

**Theorem 3.** Assume that  $\alpha_i > 0$ , and let  $\alpha_+ = \sum_{i=1}^n \alpha_i$ . Then

1.  $T \leq_{\text{st}} S$  ( $T \leq_{\text{hr}} S$ ) if and only if  $\beta \leq (\prod_{i=1}^n \beta_i^{\alpha_i})^{1/\alpha_+}$ ;
2.  $T \leq_{\text{lr}} S$  ( $T \leq_{\text{rh}} S$ ) if and only if  $\beta \leq \alpha_+ / (\sum_{i=1}^n \alpha_i / \beta_i)$ .

*Proof.* Let  $T_0 = \sum_{i=1}^n S_i$ . According to a classical property of the gamma distribution,  $(S_1/T_0, \dots, S_n/T_0)$  is independent of  $T_0$ ; consequently,  $S/T_0 = \sum \beta_i S_i/T_0$  is independent of  $T_0$ . Denote the distribution of  $S/T_0$  by  $\mu$ . Then  $S = (S/T_0)T_0$  has the distribution of a mixture of  $\text{gamma}(\alpha_+, \gamma)$  with respect to  $\mu(\gamma)$  on  $\gamma \in (0, \infty)$ , whereas  $T \sim \text{gamma}(\alpha_+, \beta)$ . Thus, the results of Example 5, i.e. (9) and (10), are directly applicable. We only need to calculate

$$\int \gamma^{-\alpha_+} d\mu(\gamma) = \mathbb{E} \left[ \left( \frac{S}{T_0} \right)^{-\alpha_+} \right]$$

and

$$\int \gamma^{-\alpha_+-1} d\mu(\gamma) = \mathbb{E} \left[ \left( \frac{S}{T_0} \right)^{-\alpha_+-1} \right].$$



It can be shown that

$$E \left[ \left( \frac{S}{T_0} \right)^{-\alpha_+} \right] = \prod_{i=1}^n \beta_i^{-\alpha_i} \tag{13}$$

$$\text{and } E \left[ \left( \frac{S}{T_0} \right)^{-\alpha_+-1} \right] = \frac{\sum_{i=1}^n \alpha_i / \beta_i}{\alpha_+} \prod_{i=1}^n \beta_i^{-\alpha_i}. \tag{14}$$

The claim then follows from (9) and (10). Equation (13) dates back to Mauldon (1959), and the following derivation, which we include for completeness, can be found in Letac *et al.* (2001). For  $t_1, \dots, t_n \in (-\infty, 1)$ , we have, by independence,

$$E \left[ \exp \left( \sum t_i S_i \right) \right] = \prod_{i=1}^n E \left[ \exp(t_i S_i) \right] = \prod_{i=1}^n (1 - t_i)^{-\alpha_i}.$$

On the other hand,

$$\begin{aligned} E \left[ \exp \left( \sum t_i S_i \right) \right] &= E \left[ E \left[ \exp \left( \sum t_i S_i \right) \mid \sum \frac{t_i S_i}{T_0} \right] \right] \\ &= E \left[ \left( 1 - \sum \frac{t_i S_i}{T_0} \right)^{-\alpha_+} \right]. \end{aligned}$$

Thus,

$$E \left[ \left( 1 - \sum \frac{t_i S_i}{T_0} \right)^{-\alpha_+} \right] = \prod_{i=1}^n (1 - t_i)^{-\alpha_i}.$$

Equation (13) is obtained by substituting  $(1 - \beta_i)$  for  $t_i$ ,  $i = 1, \dots, n$ . Moreover, (14) is obtained by differentiating both sides of (13) with respect to  $\beta_i$  and then adding the results for  $i = 1, \dots, n$ .

Actually, Khaledi and Kochar (2004) also compared variables of the form of  $S$  (assuming that the  $\alpha_i$  are equal and that their common value  $\alpha \geq 1$ ) in terms of the *dispersive order* (denoted by ' $\leq_{\text{disp}}$ '). We mention a result comparing  $T$  and  $S$  in terms of ' $\leq_{\text{disp}}$ ' for general  $\alpha_i > 0$ . Let us recall the definitions of ' $\leq_{\text{disp}}$ ' and the related *star order* (denoted by ' $\leq_*$ ').

**Definition 3.** Let  $X$  and  $Y$  be absolutely continuous random variables supported on  $(0, \infty)$  with CDFs  $F$  and  $G$ , respectively, and denote by  $F^{-1}$  and  $G^{-1}$  the inverse functions of  $F$  and  $G$ , respectively.

- We say that  $X$  is smaller than  $Y$  in the *dispersive order* (written  $X \leq_{\text{disp}} Y$ ) if

$$F^{-1}(b) - F^{-1}(a) \leq G^{-1}(b) - G^{-1}(a), \quad 0 < a < b < 1.$$

- We say that  $X$  is smaller than  $Y$  in the *star order* (written  $X \leq_* Y$ ) if  $G^{-1}F(x)/x$  is an increasing function of  $x$ ,  $x > 0$ .

**Theorem 4.** *We have*

$$T \leq_{\text{disp}} S \iff T \leq_{\text{st}} S.$$

*Proof.* The ‘ $\Rightarrow$ ’ part follows from the definitions (see Theorem 3.B.13 of Shaked and Shanthikumar (2007)). To prove the ‘ $\Leftarrow$ ’ part, first we show that  $T \leq_* S$ . The claim that  $T \leq_{\text{disp}} S$  then follows from  $T \leq_{\text{st}} S$  and  $T \leq_* S$  (see Ahmed *et al.* (1986) and Shaked and Shanthikumar (2007, p. 215)). Denote the density functions of  $T$  and  $S$  by  $f(x)$  and  $g(x)$ , respectively. One sufficient condition for  $T \leq_* S$  is that, for all  $a > 0$ ,  $af(ax) - g(x)$  changes sign at most twice as  $x$  increases from 0 to  $\infty$ , the sign sequence being  $-, +, -$  in the case of two changes. This is easily verified by noting that, based on the analysis in Section 2,  $\log(af(ax)/g(x))$  is concave in  $x$ .

### 4. Convolutions of negative binomial distributions

This section contains results for sums of independent negative binomial random variables. The development somewhat parallels that of Section 3.

Let  $N = \sum_{i=1}^n N_i$ , where  $N_i \sim \text{NB}(k_i, p_i)$  independently,  $k_i > 0$ , and  $p_i \in (0, 1)$ ,  $i = 1, \dots, n$ . Let  $M \sim \text{NB}(\sum_{i=1}^n k_i, p)$ ,  $p \in (0, 1)$ . For the special case where  $k_i = 1$ , Boland *et al.* (1994) compared variables of the form of  $N$ , i.e. sums of independent geometric variables with possibly different parameters, with respect to the likelihood ratio order. We have the following result comparing  $M$  and  $N$  for general  $k_i > 0$  (not necessarily integers). Theorem 5 should be compared with Example 1 in Section 1.

**Theorem 5.** Let  $k_+ = \sum_{i=1}^n k_i$ . Then

1.  $M \leq_{\text{st}} N$  ( $M \leq_{\text{hr}} N$ ) if and only if  $p \geq (\prod_{i=1}^n p_i^{k_i})^{1/k_+}$ ;
2.  $M \leq_{\text{lr}} N$  ( $M \leq_{\text{rh}} N$ ) if and only if  $p \geq \sum_{i=1}^n k_i p_i / k_+$ .

*Proof.* The negative binomial  $\text{NB}(k, t)$  is a mixture of  $\text{Po}(\lambda(1 - t)/t)$ , where the mixing distribution is  $\lambda \sim \text{gamma}(k, 1)$ . It follows that the distribution of  $N = \sum_{i=1}^n N_i$  is given by

$$N \mid (\lambda_1, \dots, \lambda_n) \sim \text{Po}\left(\sum_{i=1}^n \frac{\lambda_i(1 - p_i)}{p_i}\right),$$

$$\lambda_i \sim \text{gamma}(k_i, 1) \text{ independently.}$$

In this setup, let  $L = \sum_{i=1}^n \lambda_i(1 - p_i)/p_i$  and  $\lambda_+ = \sum_{i=1}^n \lambda_i$ . As in Section 3,  $L = (L/\lambda_+)\lambda_+$  is a scale mixture  $\text{gamma}(k_+, \gamma)$ , where the distribution of  $\gamma$  is that of  $L/\lambda_+$ . It is clear that  $N$  can be expressed as a mixture of negative binomial variates:

$$N \mid \gamma \sim \text{NB}(k_+, (1 + \gamma)^{-1}),$$

where again  $\gamma$  has the distribution of  $L/\lambda_+$ . We may apply the results of Example 4 in Section 2, namely (7) and (8). However, as pointed out by an anonymous reviewer, it is simpler to appeal to Theorem 2 directly. By the mixture representation of  $N$  above we have  $M \leq_{\text{lc}} N$ . A quick calculation yields

$$\frac{\Pr(M = 0)}{\Pr(N = 0)} = \frac{p^{k_+}}{\prod_{i=1}^n p_i^{k_i}}$$

and

$$\frac{\Pr(M = 1)}{\Pr(N = 1)} = \frac{k_+ p^{k_+} (1 - p)}{(\prod_{i=1}^n p_i^{k_i}) \sum_{i=1}^n k_i (1 - p_i)}.$$

The claims then follow from Theorem 2.

## Acknowledgements

The author would like to thank the Editor and an anonymous reviewer for their helpful comments.

## References

- AHMED, A. N., ALZAID, A., BARTOSZEWICZ, J. AND KOCHAR, S. C. (1986). Dispersive and superadditive ordering. *Adv. Appl. Prob.* **18**, 1019–1022.
- ALAMATSAZ, M. H. AND ABBASI, S. (2008). Ordering comparison of negative binomial random variables with their mixtures. *Statist. Prob. Lett.* **78**, 2234–2239.
- BOLAND, P. J., EL-NEWEIHI, E. AND PROSCHAN, F. (1994). Schur properties of convolutions of exponential and geometric random variables. *J. Multivariate Anal.* **48**, 157–167.
- BOLAND, P. J., SINGH, H. AND CUKIC, B. (2002). Stochastic orders in partition and random testing of software. *J. Appl. Prob.* **39**, 555–565.
- BON, J.-L. AND PĂLTĂNEA, E. (1999). Ordering properties of convolutions of exponential random variables. *Lifetime Data Anal.* **5**, 185–192.
- HARDY, G. H., LITTLEWOOD, J. E. AND POLYA, G. (1964). *Inequalities*. Cambridge University Press.
- KHALEDI, B.-E. AND KOCHAR, S. C. (2004). Ordering convolutions of gamma random variables. *Sankhyā* **66**, 466–473.
- KOCHAR, S. C. AND MA, C. (1999). Dispersive ordering of convolutions of exponential random variables. *Statist. Prob. Lett.* **43**, 321–324. (Erratum: **45** (1999), 283.)
- KORWAR, R. M. (2002). On stochastic orders for sums of independent random variables. *J. Multivariate Anal.* **80**, 344–357.
- LETAC, G., MASSAM, H. AND RICHARDS, D. (2001). An expectation formula for the multivariate Dirichlet distribution. *J. Multivariate Anal.* **77**, 117–137.
- MARSHALL, A. W. AND OLKIN, I. (1979). *Inequalities: Theory of Majorization and Its Applications*. Academic Press, New York.
- MAULDON, J. G. (1959). A generalization of the beta-distribution. *Ann. Math. Statist.* **30**, 509–520.
- MISRA, N., SINGH, H. AND HARNER, E. J. (2003). Stochastic comparisons of poisson and binomial random variables with their mixtures. *Statist. Prob. Lett.* **65**, 279–290.
- SHAKED, M. (1980). On mixtures from exponential families. *J. R. Statist. Soc. B* **42**, 192–198.
- SHAKED, M. AND SHANTIKUMAR, J. G. (1994). *Stochastic Orders and Their Applications*. Academic Press, Boston, MA.
- SHAKED, M. AND SHANTIKUMAR, J. G. (2007). *Stochastic Orders*. Springer, New York.
- WHITT, W. (1985). Uniform conditional variability ordering of probability distributions. *J. Appl. Prob.* **22**, 619–633.
- YU, Y. (2008). Relative log-concavity and a pair of triangle inequalities. Tech. Rep. Department of Statistics, University of California, Irvine.