A Positive Dependence Paradox and Aging Distributions

Kumar Joag-dev
University of Illinois, Urbana

Subhash Kochar
Indian Statistical Institute, New Delhi.

Abstract

Let $X_1, X_2, ...$ be a sequence of independent random variables and let N be an integer valued stopping variable independent of the X_i's. We study the dependence between N and $S_N = \sum_{i=1}^{N} X_i$. It is shown that when X_i's have logconcave densities, N and S_N are TP_2 dependent. Where as $P[S_N \geq x | N \geq n]$ is always increasing in n, $P[N \geq n | S_N \geq x]$ may not be increasing in x, in some cases. We prove in this note that a sufficient condition for (1.1) to hold is that X_i's have increasing failure rate distributions.

Key words : Logconcave density, IFR distribution, positive dependence, TP_2 dependence, RTI property, shock models.

1 Introduction

In reliability theory shock models have been used extensively. It is assumed that the shocks arrive at a certain rate. Each shock inflicts certain damage to system. The amounts of damage, X_i say, form a sequence of independent (but *not necessarily* identically distributed) random variables and it is further assumed that they are independent of the number of shocks which itself is a random variable. Assuming that the shocks are cumulative one can visualize situations where the failure of the system indicates that the cumulative damage has exceeded certain threshold. We want to make certain inferences regarding N, the number of shocks received by the system at the time of its failure. In particular the following question is of primary interest.

Is

$$P[N \geq n \quad | \sum_{1}^{N} X_i \geq w] \quad \text{increasing in } w \text{, for every } n? \tag{1.1}$$

Surprisingly, the answer to the above question is not always in the affirmative even when the X_i's are identically distributed. We prove in this note that a sufficient condition for (1.1) to hold is that X_i's have IFR (*increasing failure rate*) distributions. If (1.1) holds, we say that N is *right tail increasing in* $S_N = \sum_1^N X_i$ (written as $RTI\,(N|S_N)$).

There are other situations where the same question has been considered. A particular application related to the modeling of a single ion channel was recently studied by Milne, Edeson and Madsen (1986) in which the random variables X_i were exponentially distributed while N was a geometric random variable. One of the objects of their study was to examine the dependence between N and S_N. Typically, in applications, the information about S_N is more readily available while that about N is not. Consequently the nature of the dependence relation between these two variables becomes quite important for estimating the unknown N. By using explicit knowledge about the distributions, Milne, Edeson and Madsen (1986) showed that $E[N|S_N]$ is monotone increasing in S_N. We will show that such monotonicity holds under more general conditions on the distributions of X_i (exponential being a special case).

2 Preliminaries and Summary

Since most of the applications are typically for nonnegative random variables we will make that assumption in what follows.

It may not be much of an exaggeration if one says that for every monotonicity result about a family of distributions there is a TP_2 function lurking around. A nonnegative function f, defined on the plane is said to be a TP_2 function if for every $x_1 < x_2$ and $y_1 < y_2$, one has the inequality $f(x_1, y_1)f(x_2, y_2) \geq f(x_2, y_1)f(x_1, y_2)$.

If a nonnegative function f is logconcave, then it is easy to see that $f(x-y)$, considered as a function of x and y is TP_2. A family of distributions, usually parametrized, whose members possess logconcave probability density functions, is referred to as PF_2, or *Pólya Family* of order 2. It is well-known that convolution of logconcave densities is logconcave.

Let F denote a distribution function and let f be its density (with respect to Lebesgue measure). The symbol $\overline{F} = 1 - F$ is called the survival function corresponding to F.

The condition that $ln\,\overline{F}$ be concave is weaker than requiring $ln\,f$ to be concave. Note that this weaker condition is equivalent to

$$\frac{f(x)}{\overline{F}(x)} \quad \text{is monotone increasing in } x, \tag{2.1}$$

or equivalently, for every $t > 0$,

$$P[X \geq x | X \geq x - t] = \frac{\overline{F}(x)}{\overline{F}(x-t)} \quad \text{is monotone decreasing in } x. \tag{2.2}$$

The class of distributions satisfying (2.1) plays a very important role and is known as the *increasing failure rate* (IFR) class. The condition (2.2) is useful when the density does not exist.

Similarly, the aging of a component may be characterized by its *survival rate*. Given that a component has failed at time x, the conditional probability that it might have survived upto time $x - t$, is given by

$$P[X \geq x - t | X \leq x] = \frac{F(x) - F(x - t)}{F(x)}.$$

If this conditional probability is monotone decreasing in x for every $t > 0$, then it is equivalent to saying that $f(x)/F(x)$ is monotone decreasing in x, or F is *decreasing survival rate* (DSR) distribution. It is clear that this in turn is equivalent to having $ln\, F$ concave.

It is well-known that the class of IFR distributions is closed under the operation of convolution. The fact that the same is true for the class of DSR distributions follows from a proof similar to that for the IFR class (see for example, Barlow and Proschan (1981), Section 4.2, Chapter 4). Keilson and Sumita (1982) consider orderings of distributions based on failure and survival rates. It is shown there that a distribution with a logconcave probability density possesses IFR as well as DSR properties.

There are several notions of positive dependence between random variables and there is a vast literature on this topic with important contributions by Lehmann (1966), Esary and Proschan (1972), Barlow and Proschan (1981) and Shaked (1977), amongst others. Perhaps the strongest notion of positive dependence between two random variables X and Y is TP_2 dependence. Two random variables X and Y are TP_2 *dependent* if their joint density $f(x, y)$ is totally positive of order 2 in x and y. Alternatively, it can be shown that X and Y are TP_2 dependent if and only if

$$\frac{h(x|Y = y_2)}{h(x|Y = y_1)} \quad \text{is monotone increasing in } x, \tag{2.3}$$

for every pair $y_2 > y_1$, where $h(x|Y = y)$ denotes the conditional density of the random variable X given $Y = y$. Note that this dependence relation between the random variables is symmetric in X and Y. Another interesting notion of positive dependence is that of *right tail increasing*. We say that X is *right tail increasing* in Y if $P[X > x|Y > y]$ is nondecreasing in y for all x and denote this relationship by $RTI\ (X|Y)$. We say that X is *left tail decreasing* in Y if $P[X < x|Y < y]$ is nonincreasing in x for all y, and denote this by $LTD\,(X|Y)$. It is well known that if X and Y are TP_2 dependent, then $RTI\,(X|Y)$ as well as $LTD\,(X|Y)$ (cf. Barlow and Proschan (1981)). Also note that the later two notions of positive dependence are *not* symmetric in X and Y.

It is easy to see that $P[S_N \geq x|N \geq n]$ is *always* nondecreasing in n for all x, proving thereby that $RTI\,(S_N|N)$. However, we shall see in the next section that $RTI\,(N|S_N)$ does not hold in general. We prove in the next section that if X_i's have logconcave densities, then N and S_N are TP_2 dependent and that a sufficient condition for $RTI\,(N|S_N)$ is that X_i's have IFR distributions.

3 Dependence Results

As before, let $\{X_i\}$ be a sequence of independent random variables and N be an integer valued nonnegative random variable independent of this sequence.

Theorem **3.1** *Suppose that $X_1, X_2, \ldots$ are independent random variables with logconcave densities. Then N and S_N are TP_2 dependent.*

Proof. It suffices to prove that

$$\frac{h(x|N=n+1)}{h(x|N=n)} \quad \text{is monotone increasing in } x$$

for every positive integer n, where h denotes the conditional density of S_N given N. Due to the independence of the X_i and N the above monotonicity is equivalent to showing

$$\frac{g^{(n+1)}(x)}{g^{(n)}(x)} \quad \text{is monotone increasing in } x,$$

where $g^{(n)}(x)$ represents the density of the convolution $X_1 + X_2 + \cdots + X_n$. It is well known that the convolution of two logconcave densities is logconcave. Since we have assumed that each X_i has logconcave density, the required result follows from Karlin and Proschan (1960). Also see Theorem 1.C.5 of Shaked and Shanthikumar (1994) in this connection.

■

The positive dependence described above is stronger than the one described by either of the following two conditions .

$$\frac{P[N = n+1|S_N \geq x]}{P[N = n|S_N \geq x]} \quad \text{is monotone increasing in } x, \tag{3.1}$$

and

$$\frac{P[N = n+1|S_N \leq x]}{P[N = n|S_N \leq x]} \quad \text{is monotone increasing in } x. \tag{3.2}$$

The relations (3.1) and (3.2) state that the conditioning events $\{S_N \geq x\}$ and $\{S_N \leq x\}$ create MLR orderings for the conditional distributions of N in x. Shaked (1977) calls such orderings as $DTP\,(1,0)$ orderings. He has shown that (3.1) implies $RTI\,(N|S_N)$ and similarly (3.2) implies $LTD\,(N|S_N)$.

THEOREM **3.2** *Suppose that $X_1, X_2, \ldots$ are independent random variables each having IFR distribution. Then (3.1) holds. If the distribution functions have DSR property, then (3.2) holds.*

Proof. Observe that

$$\begin{aligned} \frac{P[N = n+1|S_N \geq x]}{P[N = n|S_N \geq x]} &= \frac{P[N = n+1, S_N \geq x]}{P[N = n, S_N \geq x]} \\ &= \frac{P[N = n+1]P[S_{n+1} \geq x]}{P[N = n]P[S_n \geq x]}, \end{aligned} \tag{3.3}$$

where the last equality follows from the independence of N and X_i's. Hence the ratio in (3.3), considered as a function of x, is seen to be proportional to

$$\frac{P[S_{n+1} \geq x]}{P[S_n \geq x]},$$

which is nondecreasing in x if each X_i has IFR distribution (cf. Lynch, Mimmack and Proschan (1987)).

The proof for (3.2) is similar.

■

In the next theorem we give a necessary and sufficient condition for N to be right tail increasing in S_N. We shall use the notation $\overline{F}^{(j)}(x)$ to denote the survival function of the convolution $X_1 + \cdots + X_j$.

THEOREM **3.3** *Let $p_j = P[N = j]$ and let $r^{(j)}$ denote the failure rate of S_j. Then a necessary and sufficient condition that $RTI(N|S_N)$ is that*

$$\sum_{j=1}^{n-1} \sum_{i=n}^{\infty} p_i\, p_j [r^{(j)}(x) - r^{(i)}(x)] \geq 0, \qquad \textit{for all } n \geq 1 \textit{ and for all } x \geq 0. \tag{3.4}$$

Proof: By definition, $RTI\,(N|S_N)$ if and only if

$$\begin{aligned}
& P[N \geq n | S_N \geq x] \quad \text{is nondecreasing in } x \\
\iff\ & \frac{P[S_N \geq x,\, N \geq n]}{P[S_N \geq x]} \quad \text{is nondecreasing in } x \\
\iff\ & \frac{\sum_{j=n}^{\infty} P[S_j \geq x]\, P[N = j]}{\sum_{j=1}^{\infty} P[S_j \geq x]\, P[N = j]} \quad \text{is nondecreasing in } x \\
\iff\ & \frac{\sum_{j=n}^{\infty} p_j \overline{F}^{(j)}(x)}{\sum_{i=1}^{\infty} p_i \overline{F}^{(i)}(x)} \quad \text{is nondecreasing in } x
\end{aligned}$$

$$\Longleftrightarrow \quad \frac{\sum_{j=1}^{n-1} p_j \overline{F}^{(j)}(x)}{\sum_{i=n}^{\infty} p_i \overline{F}^{(i)}(x)} \quad \text{is nonincreasing in } x$$

On differentiating and after some simplification, we find that this will hold if and only if

$$\sum_{j=1}^{n-1} \sum_{i=n}^{\infty} p_i \, p_j [r^{(j)}(x) - r^{(i)}(x)] \geq 0, \quad \text{for all } n \geq 1 \text{ and for all } x \geq 0. \tag{3.5}$$

■

It follows from Lynch, Mimmack and Proschan (1987) that if X_i 's are independent with IFR distributions, then for $i \geq j$, the failure rate $r^{(j)}(x)$ of S_j is greater than or equal to that of S_i. Hence we have the following corollary.

COROLLARY 3.1 *A sufficient condition for RTI* $(N|S_N)$ *is that* X_i *'s are independent with IFR distributions.*

Similar results can be obtained in the case of DSR random variables.

It is quite tempting to say that large values of S_N should be associated with large values of N regardless of the distributions of the nonnegative random variables X_i's. The following counter example shows that this is not the case.

Counter Example

To show that, in general,

$$P[N \geq n | X_1 + \cdots + X_N \geq w] \tag{3.6}$$

may not be monotonically increasing in w.

We take X_i's to be identically distributed. First we give an example of a random variable X such that

$$X_1 + X_2 \overset{hr}{\not\geq} X_1,$$

where X_1, X_2 are i.i.d. copies of X.

Let X be a discrete r.v. with pdf

$$f(x) = \begin{cases} .7 \quad \text{for} \quad x = 1 \\ .1 \quad \text{for} \quad x = 2 \\ .2 \quad \text{for} \quad x = 3 \\ 0 \qquad\qquad \text{otherwise} . \end{cases}$$

The survival function of X is

$$\overline{F}(x) = \begin{cases} 0, & 3 < x \\ .2, & 2 < x \leq 3 \\ .3, & 1 < x \leq 2 \\ 1, & x \leq 1 \end{cases}$$

The pdf of $Y = X_1 + X_2$ is

$$g(y) = \begin{cases} .49, & y = 2 \\ .14, & y = 3 \\ .29, & y = 4 \\ .04, & y = 5 \\ .04, & y = 6 \\ 0, & \text{otherwise} \end{cases}$$

with survival function

$$\overline{G}(x) = \begin{cases} 0, & 6 < x \\ .04, & 5 < x \leq 6 \\ .08, & 4 < x \leq 5 \\ .37, & 3 < x \leq 4 \\ .51, & 2 < x \leq 3 \\ 1, & x \leq 2 \end{cases}$$

Now

$$\frac{\overline{F}(x)}{\overline{G}(x)} = \begin{cases} 1, & x \leq 1 \\ .3, & 1 < x \leq 2 \\ \frac{.2}{.51} = .39, & 2 < x \leq 3 \\ 0, & x > 3 \end{cases}$$

is not monotonically decreasing in x.

Let N be a binary random variable taking values $\{1, 2\}$. Then (3.6) can be written as

$$\begin{aligned} & P[N \geq n | X_1 + \cdots + X_N \geq w] \\ = \; & \frac{P[X_1 + \cdots + X_N \geq w, N \geq n]}{P[X_1 + \cdots + X_N \geq w]} \\ = \; & \frac{\sum_{j=n}^{\infty} P[X_1 + \cdots + X_j \geq w] P[N = j]}{\sum_{j=1}^{\infty} P[X_1 + \cdots + X_j \geq w] P[N = j]}, \\ & \text{since } X_i\text{'s and } N \text{ are independent.} \\ = \; & \frac{P[X_1 + X_2 \geq w] p_2}{P[X_1 \geq w] p_1 + P[X_1 + X_2 \geq w] p_2} \\ = \; & \left[1 + \frac{p_1}{p_2} \frac{P[X_1 \geq w]}{P[X_1 + X_2 \geq w]} \right]^{-1}, \\ = \; & \left[1 + \frac{p_1}{p_2} \frac{\overline{F}(w)}{\overline{G}(w)} \right]^{-1} \end{aligned} \tag{3.7}$$

where $p_1 = P[N = 1]$, and $p_2 = P[N = 2]$.

It is easy to see from our above example that (3.7) is not monotonically increasing in w.

References

Barlow, R. E. and Proschan, F. (1981). *Statistical theory of reliability and life testing, probability models.* To Begin With, Silver Spring, MD.

Esary, J.D. and Proschan, F. (1972). Relationships among some notions of bivariate dependence. *Ann. Math. Statist.* **43**, 651-655.

Karlin, S. and Proschan, F. (1960). Polya type distributions of convolutions. *Ann. Math. Statist.* **31**, 761-736.

Keilson, J. and Sumita, U. (1982). Uniform stochastic ordering and related inequalities. *Canad. J. Statist.* **10**, 181–189.

Lehmann, E.L. (1966). Some concepts of dependence. *Ann. Math. Stat.* **37**, 1137-1153.

Lynch, J., Mimmack, G. and Proschan, F. (1987). Uniform stochastic orderings and totao positivity. *Canad. J. Statist.* **63-69**.

Milne, R.K., Edeson, R.O., Madsen, B.W. (1986). Stochastic modelling of a single ion channel: interdependence of burst length and number of openings per burst. *Proc. R. Soc. Lond.* **B 227**, 83-102.

Shaked, M. (1977). A family of concepts for bivariate distributions.*J. Amer. Statist. Assoc.* **72**, 642- 650.

Shaked, M. and Shanthikumar, J.G. (1994). *Stochastic Orders And Their Applications.* Academic Press : SanDiego.

July 30, 2002