Statistica Sinica 4(1994), 567-579

DISCRETE PROBABILISTIC ORDERINGS IN RELIABILITY THEORY

Moshe Shaked, J. George Shanthikumar and Jose Benigno Valdez-Torres

University of Arizona, University of California and Universidad Autonoma de Sinaloa

Abstract: Discrete failure rates arise in several common situations in reliability theory where clock time is not the best scale on which to describe lifetime. A new hazard rate ordering of vectors of discrete dependent random lifetimes is introduced and illustrated in this paper. The relationship between this new ordering and the usual stochastic ordering and the likelihood ratio ordering is discussed.

Key words and phrases: Stochastic ordering, likelihood ratio ordering, hazard rate ordering, discrete dynamic construction, history, simulation, construction on the same probability space, discrete Freund model.

1. Introduction

Discrete failure rates arise in several common situations in reliability theory where clock time is not the best scale on which to describe lifetime. For example, in weapons reliability, the number of rounds fired until failure is more important than age in failure. This is the case also when a piece of equipment operates in cycles and the observation is the number of cycles successfully completed prior to failure. In other situations a device is monitored only once per time period and the observation then is the number of time periods successfully completed prior to the failure of the device.

In this article we continue the study which was started in Shaked, Shanthikumar and Valdez-Torres (1994). In that paper we introduced, among other things, a definition of discrete multivariate conditional hazard rate functions. We showed there the usefulness of these functions for modeling imperfect repair in the discrete multivariate setting and for characterizing aging in the discrete univariate setting. In the present paper we study several notions of probabilistic ordering among vectors of discrete random lifetimes. We discuss the well known stochastic ordering relation and the likelihood ratio ordering relation among such random vectors. We also introduce a new hazard rate ordering relation among such random vectors and we study the relationships among these probabilistic orderings.

Some of the notions and definitions which were introduced in Shaked, Shanthikumar and Valdez-Torres (1994) are used here. They are reproduced in Section 2. In Section 4 we give the definitions of the probabilistic orderings which are studied later in the paper.

In Section 3 we introduce an algorithm (called the *discrete dynamic construction*) which can construct dynamically, using the discrete multivariate conditional hazard rate functions, a random vector having a desirable distribution. This algorithm may be used for simulation purposes, but here we use it as a technical tool for proving one of our main results. This result, which states that the discrete multivariate hazard rate ordering implies stochastic ordering, is proved in Section 5.

In Section 6 we study the relationship between the discrete likelihood ordering and the discrete hazard rate ordering.

The results of the present paper can be looked at as a discrete parallel development of the absolute continuous case study of Shaked and Shanthikumar (1990). However, in the discrete case there are some technical problems which do not appear in the absolute continuous case. These require the different methodology which is used in the present paper.

2. Preliminary: Discrete Multivariate Conditional Hazard Rate Functions

Consider a random vector $\boldsymbol{T} = (T_1, T_2, \ldots, T_n)$ which takes on values in $\{1, 2, \ldots\}^n \equiv \mathbb{N}_{++}^n$. For our purposes it is intuitive to think of T_i as the failure time of component i, $i = 1, 2, \ldots, n$.

The following notation will be used. Let $\boldsymbol{z} = (z_1, z_2, \ldots, z_n) \in \mathbb{N}_{++}^n$ and $I = \{i_1, i_2, \ldots, i_k\} \subset \{1, 2, \ldots, n\}$. Then $\boldsymbol{z}_I$ will denote $(z_{i_1}, z_{i_2}, \ldots, z_{i_k})$. The complement of I will be denoted by $\bar{I} = \{1, 2, \ldots, n\} - I$. We will also denote $\boldsymbol{e} = (1, 1, \ldots, 1)$; the length of $\boldsymbol{e}$ will vary according to the expression in which $\boldsymbol{e}$ appears.

Suppose that all the components start to live at time 0 and are new then. As time progresses the components fail one by one (we do not rule out the possibility of multiple failures). Thus, at time $t \in \mathbb{N}_{++}$, the information which has been gained by observing the components is an event of the form $\{\boldsymbol{T}_I = \boldsymbol{t}_I, \boldsymbol{T}_{\bar{I}} \geq t\boldsymbol{e}\}$ for some $I \subset \{1, 2, \ldots, n\}$ and $\boldsymbol{t}_I < t\boldsymbol{e}$. The *multivariate conditional hazard rate functions* of $\boldsymbol{T}$ are conditioned on such events. They are defined as

$$\lambda_{J|I}(t|\boldsymbol{t}_I) = P\{\boldsymbol{T}_J = t\boldsymbol{e}, \boldsymbol{T}_{\bar{I}-J} > t\boldsymbol{e} | \boldsymbol{T}_I = \boldsymbol{t}_I, \boldsymbol{T}_{\bar{I}} \geq t\boldsymbol{e}\} \tag{2.1}$$

for some $J \subset \bar{I} \subset \{1, 2, \ldots, n\}$ and $\boldsymbol{t}_I < t\boldsymbol{e}$. If in (2.1) the probability of $\{\boldsymbol{T}_I = \boldsymbol{t}_I, \boldsymbol{T}_{\bar{I}} \geq t\boldsymbol{e}\}$ is zero, then $\lambda_{J|I}(t|\boldsymbol{t}_I)$ is defined as 1. Note that in (2.1) it is possible that $J = \emptyset$. In that case we have

$$\lambda_{\emptyset|I}(t|\boldsymbol{t}_I) = P\{\boldsymbol{T}_{\bar{I}} > t\boldsymbol{e}|\boldsymbol{T}_I = \boldsymbol{t}_I, \boldsymbol{T}_{\bar{I}} \geq t\boldsymbol{e}\}.$$

If $I = \emptyset$ in (2.1) then we abbreviate $\lambda_{J|I}(t|\boldsymbol{t}_I)$ by $\lambda_J(t)$. These hazard rates can be called *initial* because they describe the hazard rates of the components before having had any failures.

Clearly, the hazard rate functions are determined by the probability function of $\boldsymbol{T}$. But also the converse is true. It is possible to express explicitly the joint probability function of $\boldsymbol{T}$ by means of the hazard rate functions (2.1); see Shaked, Shanthikumar and Valdez-Torres (1994). It follows that in order to describe the life distribution of $\boldsymbol{T}$ it is enough to postulate the hazard rate functions (2.1). This is a useful fact because in the setting of reliability theory the hazard rate functions have more intuitive meaning than the joint probability function. In this paper we use these functions to characterize various probabilistic orderings of discrete multivariate vectors of random lifetimes.

3. Preliminary: The Discrete Dynamic Construction

Let $\boldsymbol{T} = (T_1, T_2, \ldots, T_n)$ be a discrete random vector taking on values in $\mathbb{N}^n_{++}$. Let $\lambda_{\cdot|\cdot}(\cdot|\cdot)$ be its discrete multivariate conditional hazard rate functions as described in (2.1). We describe now an algorithm, called the *discrete dynamic construction*, which, using the functions $\lambda_{\cdot|\cdot}(\cdot|\cdot)$, constructs a random vector $\hat{\boldsymbol{T}} = (\hat{T}_1, \hat{T}_2, \ldots, \hat{T}_n)$ such that

$$\hat{\boldsymbol{T}} =_{\text{st}} \boldsymbol{T} \tag{3.1}$$

(here '$=_{\text{st}}$' means equality in law).

The algorithm is similar to, but different than, the dynamic construction described in Shaked and Shanthikumar (1991b). The latter construction applies to vectors of random lifetimes with absolutely continuous joint distributions. In such a case, no two components can fail at the same time epoch. Here, in the discrete case, it is possible. Therefore, the discrete construction is different in nature than the one in Shaked and Shanthikumar (1991b) — it has to allow multiple failures at some time epochs. The discrete dynamic construction is described below by induction on $t \in \mathbb{N}_{++}$ — the countable number of time epochs in which components may fail. It is unlike the continuous construction of Shaked and Shanthikumar (1991b) in which the induction was over the ordered failure times.

We describe now the steps of the discrete dynamic construction. As we mentioned above, they are indexed by $t \in \mathbb{N}_{++}$. In general, Step t describes which components failed at time t, if any. These failure times are the $\hat{T}_i$'s.

Step 1. The algorithm enters this step when all the components are alive. The algorithm now chooses a set $J \subset \{1, 2, \dots, n\}$ with probability $\lambda_J(1)$ [J may be empty], and defines (if $J \neq \emptyset$)

$$\hat{\boldsymbol{T}}_J = \boldsymbol{e}. \tag{3.2}$$

For $i \in \bar{J}$ the algorithm does not define $\hat{T}_i$ in this step; these $\hat{T}_i$'s will be defined in a later step. Upon determination of J and $\hat{\boldsymbol{T}}_J$ the algorithm sets $t = 2$ and then proceeds to Step t.

Thus, upon exit from Step 1, some of the $\hat{T}_i$'s (if any) have been determined already as described in (3.2), and the other $\hat{T}_i$'s (i.e., for $i \in \bar{J}$) are still to be determined. Therefore $\hat{\boldsymbol{T}}_{\bar{J}} > \boldsymbol{e}$. (If $J = \emptyset$ then after Step 1 one has $\hat{\boldsymbol{T}} > \boldsymbol{e}$.)

Step t. Upon entrance to this step some of the $\hat{T}_i$'s (if any) have already been determined. Suppose that the algorithm has already determined the $\hat{T}_i$'s with $i \in I$ for some set $I \subset \{1, 2, \dots, n\}$. More explicitly, suppose that upon entrance to this step we already know that $\boldsymbol{T}_I = \boldsymbol{t}_I$ (where, of course, $\boldsymbol{t}_I < t\boldsymbol{e}$) and that $\boldsymbol{T}_{\bar{I}} \geq t\boldsymbol{e}$. The algorithm now chooses a set $J \subset \bar{I}$ with probability $\lambda_{J|I}(t|\boldsymbol{t}_I)$ and defines (if $J \neq \emptyset$)

$$\hat{\boldsymbol{T}}_J = t\boldsymbol{e}.$$

For $i \in \overline{I \cup J}$ the algorithm does not define $\hat{T}_i$ in this step; these $\hat{T}_i$'s (if any) will be determined in a later step. From Step t the algorithm proceeds to Step $t+1$ provided $\overline{I \cup J} \neq \emptyset$. Otherwise the construction is complete.

Thus, upon exit from Step t, the $\hat{T}_i$'s with $i \in I \cup J$ have been determined already. The other $\hat{T}_i$'s (if any) are still to be determined, that is, $\boldsymbol{T}_{\overline{I \cup J}} > t\boldsymbol{e}$. Upon entrance to Step $t+1$ (if ever) we already know the values of $\hat{T}_i$ for $i \in I \cup J$.

The algorithm performs the steps in sequence until all the $\hat{T}_i$'s have been determined. With probability one this will happen in a finite number of steps whenever $P\{T_i < \infty, \ i = 1, 2, \dots, n\} = 1$.

From the construction it is clear that $\hat{\boldsymbol{T}}$ has the discrete multivariate conditional hazard rate functions of $\boldsymbol{T}$. Since the discrete multivariate conditional hazard rate functions uniquely determine the probability function, it follows that $\hat{\boldsymbol{T}} =_{\text{st}} \boldsymbol{T}$.

The discrete dynamic construction can be used to simulate discrete dependent lifetimes. This can be done by generating a sequence of independent uniform random variables $\{U_t, t \in \mathbb{N}_{++}\}$ and using U_t in order to generate the required probabilities in Step t, $t \in \mathbb{N}_{++}$. In this paper, however, we use the discrete dynamic construction as a technical tool for proving Theorem 5.1 in Section 5.

4. Discrete Probabilistic Ordering

Let $\boldsymbol{X} = (X_1, X_2, \ldots, X_n)$ and $\boldsymbol{Y} = (Y_1, Y_2, \ldots, Y_n)$ be two discrete random vectors taking on values in $\{\ldots, -1, 0, 1, \ldots\}^n = \mathbb{Z}^n$. The random vector $\boldsymbol{X}$ is said to be *stochastically smaller* than the random vector $\boldsymbol{Y}$ (denoted $\boldsymbol{X} \leq_{\text{st}} \boldsymbol{Y}$) if

$$E\phi(\boldsymbol{X}) \leq E\phi(\boldsymbol{Y}) \tag{4.1}$$

for every real function ϕ, with domain in $\mathbb{Z}^n$, which is increasing with respect to the componentwise partial ordering in $\mathbb{Z}^n$ (and for which the expectations in (4.1) exist). In this paper 'increasing' means 'nondecreasing' and 'decreasing' means 'nonincreasing'. If Q denotes the probability measure of $\boldsymbol{X}$ and R denotes the probability measure of $\boldsymbol{Y}$ then we sometimes write $Q \leq_{\text{st}} R$ to denote $\boldsymbol{X} \leq_{\text{st}} \boldsymbol{Y}$.

The establishment of the relationship $\boldsymbol{X} \leq_{\text{st}} \boldsymbol{Y}$ is of importance in various applications. One can view Theorem 5.1 in Section 5 as a set of sufficient conditions on the discrete multivariate conditional hazard rate functions which assure the stochastic ordering relation between two vectors of discrete random vectors.

In order to define the next ordering (the one we call the hazard rate ordering) we need to introduce some notation. This ordering will be used only in order to compare vectors of discrete random lifetimes. Therefore, we assume now that $\boldsymbol{X}$ and $\boldsymbol{Y}$ can take on values only in $\mathbb{N}_{++}$.

For $t \in \mathbb{N}_{++}$ let h_t denote a realization of the failure times of n components up to time t, *exclusive*. That is, if $X_1, X_2, \ldots, X_n$ are the discrete random lifetimes of the components, then h_t is an event of the form $\{\boldsymbol{X}_I = \boldsymbol{x}_I, \boldsymbol{X}_{\bar{I}} \geq t\boldsymbol{e}\}$ for some $I \subset \{1, 2, \ldots, n\}$ and $\boldsymbol{x}_I < t\boldsymbol{e}$. On such events we condition the probabilities in the definition (2.1) of the discrete multivariate conditional hazard rate functions. Such an event will be called a *history*.

Fix a $t \in \mathbb{N}_{++}$. If h_t and h'_t are two histories such that in h_t there are more failures than in h'_t and every component which failed in h'_t also failed in h_t, and, for components which failed in both histories, the failures in h_t are earlier than the failures in h'_t, then we say that $h_t \leq h'_t$. More explicitly, if h_t is a history associated with $\boldsymbol{X}$ of the form $\{\boldsymbol{X}_I = \boldsymbol{x}_I, \boldsymbol{X}_{\bar{I}} \geq t\boldsymbol{e}\}$ and h'_t is a history associated with $\boldsymbol{Y}$ of the form $\{\boldsymbol{Y}_A = \boldsymbol{y}_A, \boldsymbol{Y}_{\bar{A}} \geq t\boldsymbol{e}\}$ then $h_t \leq h'_t$ if, and only if, $A \subset I$ and $\boldsymbol{x}_A \leq \boldsymbol{y}_A$ (of course, we also have $\boldsymbol{x}_{I-A} < t\boldsymbol{e}$ and $\boldsymbol{y}_A < t\boldsymbol{e}$).

Remark 4.1. Before proceeding, we note a 1-1 association between $\{0,1\}^n$ and the set of subsets of $\{1, 2, \ldots, n\}$. For each point $\boldsymbol{u} \in \{0,1\}^n$ let $A(\boldsymbol{u}) \subset \{1, 2, \ldots, n\}$ be the set of the coordinates of $\boldsymbol{u}$ which are 1's. Conversely, for each set $A = \{i_1, i_2, \ldots, i_k\} \subset \{1, 2, \ldots, n\}$ let $\boldsymbol{u}(A) \in \{0,1\}^n$ be the vector which has 1's in places $i_1, i_2, \ldots, i_k$ and 0's elsewhere.

Let $\mu_{\cdot|\cdot}(\cdot|\cdot)$ denote the discrete multivariate conditional hazard rate function of $\boldsymbol{X}$ (as defined in (2.1)). Similarly let $\eta_{\cdot|\cdot}(\cdot|\cdot)$ be the hazard rate functions of $\boldsymbol{Y}$.

Given a history h_t, associated with $\boldsymbol{X}$, of the form $\{\boldsymbol{X}_I = \boldsymbol{x}_I, \boldsymbol{X}_{\bar{I}} \geq t\boldsymbol{e}\}$, we define now a probability measure Q_{h_t} on $\{0,1\}^n$ as follows. For $A \subset \bar{I}$ set

$$Q_{h_t}\{\boldsymbol{u}(I \cup A)\} = \mu_{A|I}(t|\boldsymbol{x}_I), \quad A \subset \bar{I}, \tag{4.2}$$

and let the mass of Q_{h_t} on all other points of $\{0,1\}^n$ be 0. It is obvious that Q_{h_t} is a proper probability measure; it corresponds to the indicators of the components that have failed by time t, *inclusive*. We call $Q_.$ the *discrete multivariate conditional hazard rate measure* of $\boldsymbol{X}$.

Similarly, given a history h'_t, associated with $\boldsymbol{Y}$, one can define, as in (4.2), the discrete multivariate conditional hazard rate measure of $\boldsymbol{Y}$. It is denoted by $R_.$.

Definition 4.2. Let $\boldsymbol{X}$ and $\boldsymbol{Y}$ be random vectors which take on values in $\mathbb{N}^n_{++}$. The random vector $\boldsymbol{X}$ is said to be smaller than $\boldsymbol{Y}$ in the *discrete hazard rate ordering* (denoted $\boldsymbol{X} \leq_{\text{h}} \boldsymbol{Y}$) if

$$Q_{h_t} \geq_{\text{st}} R_{h'_t} \quad \text{whenever} \quad h_t \leq h'_t. \tag{4.3}$$

For example suppose that $n = 2$. Then (4.3) is equivalent to

$$\begin{aligned}
\mu_{\{1,2\}}(t) &\geq \eta_{\{1,2\}}(t), \quad t \in \mathbb{N}_{++}, && (4.4)\\
\mu_{\{1\}}(t) + \mu_{\{1,2\}}(t) &\geq \eta_{\{1\}}(t) + \eta_{\{1,2\}}(t), \quad t \in \mathbb{N}_{++}, && (4.5)\\
\mu_{\{2\}}(t) + \mu_{\{1,2\}}(t) &\geq \eta_{\{2\}}(t) + \eta_{\{1,2\}}(t), \quad t \in \mathbb{N}_{++}, && (4.6)\\
\mu_{\{1\}}(t) + \mu_{\{2\}}(t) + \mu_{\{1,2\}}(t) &\geq \eta_{\{1\}}(t) + \eta_{\{2\}}(t) + \eta_{\{1,2\}}(t), \quad t \in \mathbb{N}_{++}, && (4.7)\\
\mu_{\{2\}|\{1\}}(t|x_1) &\geq \eta_{\{2\}}(t) + \eta_{\{1,2\}}(t), \quad t > x_1 \geq 1, &&\\
\mu_{\{1\}|\{2\}}(t|x_2) &\geq \eta_{\{1\}}(t) + \eta_{\{1,2\}}(t), \quad t > x_2 \geq 1, &&\\
\mu_{\{2\}|\{1\}}(t|x_1) &\geq \eta_{\{2\}|\{1\}}(t|y_1), \quad t > y_1 \geq x_1 \geq 1, \quad \text{and} &&\\
\mu_{\{1\}|\{2\}}(t|x_2) &\geq \eta_{\{1\}|\{2\}}(t|y_2), \quad t > y_2 \geq x_2 \geq 1. &&
\end{aligned}$$

If in the case $n = 2$ there cannot be simultaneous failures, that is, if $P\{X_1 = X_2\} = P\{Y_1 = Y_2\} = 0$, then $\mu_{\{1,2\}}(t) = \eta_{\{1,2\}}(t) = 0, t \in \mathbb{N}_{++}$, and (4.7) is superfluous because it follows from (4.5) and (4.6). Also (4.4) then obviously holds. The remaining conditions can then be written as

$$\begin{aligned}
\mu_{\{k\}|I\cup J}(t|\boldsymbol{x}_I, \boldsymbol{x}_J) \geq &\eta_{\{k\}|I}(\boldsymbol{y}_I), \quad \boldsymbol{x}_I \leq \boldsymbol{y}_I < t\boldsymbol{e},\\
&\boldsymbol{x}_J < t\boldsymbol{e}, \ I \subset \{1,2\}, \ J \subset \{1,2\}, \ I \cap J = \emptyset, \ k \in \overline{I \cup J}. \quad (4.8)
\end{aligned}$$

In fact, for a general n, if no two or more simultaneous failures can occur for a collection of components with lifetimes $X_1, X_2, \ldots, X_n$, and for a collection of components with lifetimes $Y_1, Y_2, \ldots, Y_n$, then (4.3) reduces to (4.8) (with

$\{1,2\}$ replaced there by $\{1,2,\ldots,n\}$). Condition (4.8) is similar to the condition of Shaked and Shanthikumar (1990) which defines the hazard rate ordering for vectors of random lifetimes with absolutely continuous distributions. But in Definition 4.2 we need Condition (4.3) rather than (4.8) because of the positive probability of multiple failures when one deals with discrete failure times. One can see now the additional complexity which is involved when one studies components with discrete random lifetimes which may have multiple failures, as opposed to the case of random lifetimes with absolutely continuous distributions.

Example 4.3. Consider the following discrete analogue of a model of Ross (1984) and Freund (1961). Suppose n components start to live at time 0. The discrete failure rate of each of them at time $t = 1$ is p_n and they may fail at time 1 independently of each other. At any time $t \in \mathbb{N}_{++}$, the failure rate of each of the surviving components is independent of t. It depends only on the number of surviving components, and the surviving components may fail at time t independently of each other. More formally, the $\lambda_{J|I}(t|\boldsymbol{t}_I)$ of (2.1) is now a function of $|\bar{I}|$ (the cardinality of $\bar{I}$) and of $|J|$ only. If $p_{|\bar{I}|}$ is the failure rate of any of the surviving components then

$$\lambda_{J|I}(t|\boldsymbol{t}_I) = p_{|\bar{I}|}^{|J|}(1-p_{|\bar{I}|})^{|\bar{I}|-|J|}, \quad J \subset \bar{I} \subset \{1,2,\ldots,n\}.$$

Let $\boldsymbol{X}$ be a vector of lifetimes having the above distribution. That is, suppose that $\boldsymbol{X}$ has the discrete multivariate conditional hazard rate functions

$$\mu_{J|I}(t|\boldsymbol{x}_I) = p_{|\bar{I}|}^{|J|}(1-p_{|\bar{I}|})^{|\bar{I}|-|J|}, \quad J \subset \bar{I} \subset \{1,2,\ldots,n\},\ t \in \mathbb{N}_{++}.$$

Let $\boldsymbol{Y}$ have the same distribution but with parameters $q_n, q_{n-1}, \ldots, q_1$, rather than $p_n, p_{n-1}, \ldots, p_1$. That is, suppose that the discrete multivariate conditional hazard rate functions of $\boldsymbol{Y}$ are

$$\eta_{J|I}(t|\boldsymbol{x}_I) = q_{|\bar{I}|}^{|J|}(1-q_{|\bar{I}|})^{|\bar{I}|-|J|}, \quad J \subset \bar{I} \subset \{1,2,\ldots,n\},\ t \in \mathbb{N}_{++}.$$

Then it can be verified (using coupling arguments) that if $p_i \geq q_j$, $j \geq i$, $i = 1,2,\ldots,n$, then

$$Q_{h_t} \geq_{\text{st}} R_{h'_t} \quad \text{whenever} \quad h_t \leq h'_t, \quad t \in \mathbb{N}_{++},$$

where $Q.$ and $R.$ are as described in (4.2) and (4.3). Therefore $\boldsymbol{X} \leq_{\text{h}} \boldsymbol{Y}$.

Let $\boldsymbol{X}$ and $\boldsymbol{Y}$ take on values in $\mathbb{Z}^n$. Let f denote the discrete probability density of $\boldsymbol{X}$, that is,

$$f(x_1,x_2,\ldots,x_n) = P\{X_1 = x_1, X_2 = x_2, \ldots, X_n = x_n\}, \quad \boldsymbol{x} \in \mathbb{Z}^n.$$

Similarly, let g denote the discrete probability density of $\boldsymbol{Y}$. We say that $\boldsymbol{X}$ is smaller than $\boldsymbol{Y}$ in the *likelihood ratio ordering* (denoted $\boldsymbol{X} \leq_{\rm lr} \boldsymbol{Y}$) if

$$f(\boldsymbol{x})g(\boldsymbol{y}) \leq f(\boldsymbol{x} \wedge \boldsymbol{y})g(\boldsymbol{x} \vee \boldsymbol{y}), \quad \boldsymbol{x} \in \mathbb{Z}^n, \ \boldsymbol{y} \in \mathbb{Z}^n,$$

where $\boldsymbol{x} \wedge \boldsymbol{y}$ denotes $(x_1 \wedge y_1, x_2 \wedge y_2, \ldots, x_n \wedge y_n)$ and $\boldsymbol{x} \vee \boldsymbol{y}$ denotes $(x_1 \vee y_1, x_2 \vee y_2, \ldots, x_n \vee y_n)$; see Karlin and Rinott (1980) and Whitt (1982), where examples of random vectors which are ordered by the likelihood ratio ordering can be found.

It should be noted that $\leq_{\rm h}$ and $\leq_{\rm lr}$ are not orderings in the usual sense because they are not necessarily reflexive.

5. The Relationship Between the Hazard Rate Ordering and the Usual Stochastic Ordering

In this section we prove the following result.

Theorem 5.1. *Let* $\boldsymbol{X} = (X_1, X_2, \ldots, X_n)$ *and* $\boldsymbol{Y} = (Y_1, Y_2, \ldots, Y_n)$ *be two random vectors which can take on values in* $\mathbb{N}^n_{++}$. *If* $\boldsymbol{X} \leq_{\rm h} \boldsymbol{Y}$ *then*

$$\boldsymbol{X} \leq_{\rm st} \boldsymbol{Y}. \tag{5.1}$$

Proof. The proof will be done by constructing, on the same probability space, two random vectors $(\hat{X}_1, \hat{X}_2, \ldots, \hat{X}_n)$ and $(\hat{Y}_1, \hat{Y}_2, \ldots, \hat{Y}_n)$ such that

$$\hat{\boldsymbol{X}} =_{\rm st} \boldsymbol{X}, \tag{5.2}$$

$$\hat{\boldsymbol{Y}} =_{\rm st} \boldsymbol{Y}, \quad \text{and} \tag{5.3}$$

$$\hat{\boldsymbol{X}} \leq \hat{\boldsymbol{Y}} \quad \text{a.s..} \tag{5.4}$$

From (5.2), (5.3) and (5.4) one obtains (5.1).

Denote the discrete multivariate conditional hazard rate functions of $\boldsymbol{X}$ by $\mu_{\cdot|\cdot}(\cdot|\cdot)$ and of $\boldsymbol{Y}$ by $\eta_{\cdot|\cdot}(\cdot|\cdot)$.

The construction of $\hat{\boldsymbol{X}}$ and $\hat{\boldsymbol{Y}}$ will be done in steps indexed by $t \in \mathbb{N}_{++}$. Here, as in the discrete dynamic construction, we describe an algorithm in which t is to be thought of as a value of discrete time. In Step t it is determined which $\hat{X}_i$'s (if any) and which $\hat{Y}_i$'s (if any) are equal to t.

Step 1. The algorithm enters this step with the obvious information that $\hat{\boldsymbol{X}} \geq \boldsymbol{e}$ and $\hat{\boldsymbol{Y}} \geq \boldsymbol{e}$. Consider Q_{h_1} as in (4.3) with $t = 1$ and $I = \emptyset$ (because $h_1 = \{\boldsymbol{X} \geq \boldsymbol{e}\}$). Consider R_{h_1} as in (4.3) with $t = 1$ and $I = \emptyset$ except that here η replaces μ. From (4.3) it follows that $Q_{h_1} \geq_{\rm st} R_{h_1}$. Therefore random vectors $\boldsymbol{U}_1$ and $\boldsymbol{V}_1$, which can take on values in $\{0,1\}^n$, can be defined on the same probability space such that $\boldsymbol{U}_1$ has the probability measure Q_{h_1}, $\boldsymbol{V}_1$ has the probability measure R_{h_1} and $\boldsymbol{U}_1 \geq \boldsymbol{V}_1$ with probability one (see, e.g., Kamae, Krengel and O'Brien (1977)). Let S_1 be the joint probability measure of $(\boldsymbol{U}_1, \boldsymbol{V}_1)$. The algorithm now chooses a realization $(\boldsymbol{u}_1, \boldsymbol{v}_1)$ according to S_1.

Let $A \subset \{1, 2, \ldots, n\}$ be the set associated with $\boldsymbol{u}_1$ as described in Remark 4.1. Similarly let $A' \subset \{1, 2, \ldots, n\}$ be the set associated with $\boldsymbol{v}_1$. Since $\boldsymbol{u}_1 \geq \boldsymbol{v}_1$ it follows that $A \supset A'$. Of course A' or A may be the empty sets. Define

$$\hat{\boldsymbol{X}}_A = \boldsymbol{e}, \quad \hat{\boldsymbol{Y}}_{A'} = \boldsymbol{e},$$

set $t = 2$ and proceed to Step t.

Upon exit from Step 1 some of the $\hat{X}_i$'s and some of the $\hat{Y}_j$'s (if any) have been determined and it is known, then, that $\hat{\boldsymbol{X}}_{\bar{A}} > \boldsymbol{e}$ and $\boldsymbol{Y}_{\bar{A}'} > \boldsymbol{e}$. It follows that we already have

$$\hat{\boldsymbol{X}}_A \leq \hat{\boldsymbol{Y}}_A \quad \text{with probability one.}$$

Notice that not all the $\hat{Y}_i$'s with $i \in A$ have been already determined. Some of the $\hat{Y}_i$'s (those with $i \in A - A'$) still have not been determined, but they must satisfy $\hat{Y}_i > 1$.

Step t. Upon entrance to this step some of the $\hat{X}_i$'s and some of the $\hat{Y}_i$'s (if any) have already been determined. Suppose that the $\hat{X}_i$'s have been determined for all $i \in A$ for some set $A \subset \{1, 2, \ldots, n\}$. More explicitly suppose that $\hat{\boldsymbol{X}}_A = \boldsymbol{x}_A$, $\hat{\boldsymbol{X}}_{\bar{A}} \geq t\boldsymbol{e}$. Suppose, also, that the $\hat{Y}_i$'s have been determined for $i \in A'$ for some set $A' \subset \{1, 2, \ldots, n\}$. More explicitly, suppose $\hat{\boldsymbol{Y}}_{A'} = \boldsymbol{y}_{A'}$, $\hat{\boldsymbol{Y}}_{\bar{A}'} \geq t\boldsymbol{e}$. By the induction hypothesis, $A \supset A'$, $\boldsymbol{x}_A < t\boldsymbol{e}$, $\boldsymbol{x}_{A'} \leq \boldsymbol{y}_{A'} < t\boldsymbol{e}$. Therefore, if we define $h_t = \{\boldsymbol{X}_{A'} = \boldsymbol{x}_{A'}, \boldsymbol{X}_{A-A'} = \boldsymbol{x}_{A-A'}, \boldsymbol{X}_{\bar{A}} \geq t\boldsymbol{e}\}$ and $h'_t = \{\boldsymbol{Y}_{A'} = \boldsymbol{y}_{A'}, \boldsymbol{Y}_{\bar{A}'} \geq t\boldsymbol{e}\}$ we have $h_t \leq h'_t$. Consider now Q_{h_t} and $R_{h'_t}$ as defined in Section 4. From (4.3) it follows that $Q_{h_t} \geq_{\text{st}} R_{h'_t}$. Therefore, random vectors $\boldsymbol{U}_t$ and $\boldsymbol{V}_t$, taking on values in $\{0, 1\}^n$, can be defined, on the same probability space, such that $\boldsymbol{U}_t$ is distributed according to Q_{h_t}, $\boldsymbol{V}_t$ is distributed according to $R_{h'_t}$, and $\boldsymbol{U}_t \geq \boldsymbol{V}_t$ with probability one. Let S_t be the joint probability measure of $(\boldsymbol{U}_t, \boldsymbol{V}_t)$. The algorithm now chooses a realization $(\boldsymbol{u}_t, \boldsymbol{v}_t)$ according to S_t.

Let $B \subset \{1, 2, \ldots, n\}$ be the set associated with $\boldsymbol{u}_t$ as described in Remark 4.1 and let $B' \subset \{1, 2, \ldots, n\}$ be the set similarly associated with $\boldsymbol{v}_t$. From the definition of Q_{h_t} is clear that $B \supset A$. Similarly from the definition of $R_{h'_t}$ it is seen that $B' \supset A'$. Also, since $\boldsymbol{u}_t \geq \boldsymbol{v}_t$ it follows that $B \supset B'$. Define

$$\hat{\boldsymbol{X}}_{B-A} = t\boldsymbol{e}, \quad \hat{\boldsymbol{Y}}_{B'-A'} = t\boldsymbol{e}$$

and proceed to Step $t + 1$.

Upon exit from Step t some of the $\hat{X}_i$'s and some of the $\hat{Y}_i$'s (if any) have been determined and it is known that $\hat{\boldsymbol{X}}_{\bar{B}} > t\boldsymbol{e}$ and $\hat{\boldsymbol{Y}}_{\bar{B}'} > t\boldsymbol{e}$. Also, since $B \supset B'$, it follows (using the induction hypothesis $\hat{\boldsymbol{X}}_A \le \hat{\boldsymbol{Y}}_A$ a.s.) that

$$\hat{\boldsymbol{X}}_B \le \hat{\boldsymbol{Y}}_B \quad \text{a.s..}$$

Notice that not necessarily all the $\hat{Y}_i$'s with $i \in B$ have been determined by Step t. The $\hat{Y}_i$'s with $i \in B - B'$ have not been determined yet, but they must satisfy $\hat{Y}_i > t$.

Performing the steps of this procedure in sequence the algorithm finally determines all the $\hat{X}_i$'s and $\hat{Y}_i$'s using a construction for all h_t and h'_t which are realized. The resulting $\hat{\boldsymbol{X}}$ and $\hat{\boldsymbol{Y}}$ must satisfy (5.4). The $\hat{\boldsymbol{X}}$ satisfies (5.2) because it is marginally constructed as in the discrete dynamic construction. Similarly $\hat{\boldsymbol{Y}}$ satisfies (5.3).

As an example for the use of Theorem 5.1 consider the $\boldsymbol{X}$ and the $\boldsymbol{Y}$ defined in Example 4.3. It has been shown in Example 4.3 that $\boldsymbol{X} \le_{\mathrm{h}} \boldsymbol{Y}$. It follows from Theorem 5.1 that $\boldsymbol{X} \le_{\mathrm{st}} \boldsymbol{Y}$.

6. The Relationship Between the Likelihood Ratio Ordering and the Hazard Rate Ordering

The following notation is used in this section: Let Z be a random variable (or vector) and let E be an event. Then $[Z|E]$ denotes any random variable (or vector) whose distribution is the conditional distribution of Z given E.

In this section we prove the following result.

Theorem 6.1. *Let* $\boldsymbol{X} = (X_1, X_2, \ldots, X_n)$ *and* $\boldsymbol{Y} = (Y_1, Y_2, \ldots, Y_n)$ *be two random vectors taking on values in* $\mathbb{N}^n_{++}$. *If* $\boldsymbol{X} \le_{\mathrm{lr}} \boldsymbol{Y}$ *then*

$$\boldsymbol{X} \le_{\mathrm{h}} \boldsymbol{Y}. \tag{6.1}$$

Proof. Denote the discrete density of $\boldsymbol{X}$ by f and of $\boldsymbol{Y}$ by g.

Split $\{1, 2, \ldots, n\}$ into three mutually exclusive sets I, J and L (so that $L = \overline{I \cup J}$). Fix $\boldsymbol{x}_I, \boldsymbol{x}_J, \boldsymbol{y}_I$ and $t \in \mathbb{N}_{++}$ such that $\boldsymbol{x}_I \le \boldsymbol{y}_I < t\boldsymbol{e}$ and $\boldsymbol{x}_J < t\boldsymbol{e}$. Let $h_t = \{\boldsymbol{X}_I = \boldsymbol{x}_I, \boldsymbol{X}_J = \boldsymbol{x}_J, \boldsymbol{X}_L \ge t\boldsymbol{e}\}$ and $h'_t = \{\boldsymbol{Y}_I = \boldsymbol{y}_I, \boldsymbol{Y}_{J\cup L} \ge t\boldsymbol{e}\}$. First we show that

$$[(\boldsymbol{X}_I, \boldsymbol{X}_J, \boldsymbol{X}_L)|\boldsymbol{X}_I = \boldsymbol{x}_I, \boldsymbol{X}_J = \boldsymbol{x}_J, \boldsymbol{X}_L \ge t\boldsymbol{e}] \le_{\mathrm{lr}} [(\boldsymbol{Y}_I, \boldsymbol{Y}_J, \boldsymbol{Y}_L)|\boldsymbol{Y}_I = \boldsymbol{y}_I, \boldsymbol{Y}_{J\cup L} \ge t\boldsymbol{e}]. \tag{6.2}$$

Denote the discrete densities of $(\boldsymbol{X}_I, \boldsymbol{X}_J, \boldsymbol{X}_L)$ and of $(\boldsymbol{Y}_I, \boldsymbol{Y}_J, \boldsymbol{Y}_L)$ by $\widetilde{f}$ and $\widetilde{g}$, respectively. The discrete density of $[(\boldsymbol{X}_I, \boldsymbol{X}_J, \boldsymbol{X}_L)|\boldsymbol{X}_I = \boldsymbol{x}_I, \boldsymbol{X}_J = \boldsymbol{x}_J, \boldsymbol{X}_L \ge t\boldsymbol{e}]$ is

$$\tilde{\tilde{f}}(\boldsymbol{a}_I, \boldsymbol{a}_J, \boldsymbol{a}_L) = \frac{\tilde{f}(\boldsymbol{a}_I, \boldsymbol{a}_J, \boldsymbol{a}_L)}{\sum\limits_{\boldsymbol{x}_L \geq t\boldsymbol{e}} \tilde{f}(\boldsymbol{x}_I, \boldsymbol{x}_J, \boldsymbol{x}_L)}$$

provided $\boldsymbol{a}_I = \boldsymbol{x}_I, \boldsymbol{a}_J = \boldsymbol{x}_J, \boldsymbol{a}_L \geq t\boldsymbol{e}$, and is 0 otherwise. The discrete density of $[(\boldsymbol{Y}_I, \boldsymbol{Y}_J, \boldsymbol{Y}_L)|\boldsymbol{Y}_I = \boldsymbol{y}_I, \boldsymbol{Y}_{J\cup L} \geq t\boldsymbol{e}]$ is

$$\tilde{\tilde{g}}(\boldsymbol{b}_I, \boldsymbol{b}_J, \boldsymbol{b}_L) = \frac{\tilde{g}(\boldsymbol{b}_I, \boldsymbol{b}_J, \boldsymbol{b}_L)}{\sum\limits_{\boldsymbol{y}_J \geq t\boldsymbol{e}} \sum\limits_{\boldsymbol{y}_L \geq t\boldsymbol{e}} \tilde{g}(\boldsymbol{y}_I, \boldsymbol{y}_J, \boldsymbol{y}_L)}$$

provided $\boldsymbol{b}_I = \boldsymbol{y}_I, \boldsymbol{b}_J \geq t\boldsymbol{e}, \boldsymbol{b}_L \geq t\boldsymbol{e}$, and is 0 otherwise. In order to prove (6.2) we need to show that

$$\tilde{\tilde{f}}(\boldsymbol{a}_I, \boldsymbol{a}_J, \boldsymbol{a}_L)\tilde{\tilde{g}}(\boldsymbol{b}_I, \boldsymbol{b}_J, \boldsymbol{b}_L) \leq \tilde{\tilde{f}}(\boldsymbol{a}_I \wedge \boldsymbol{b}_I, \boldsymbol{a}_J \wedge \boldsymbol{b}_J, \boldsymbol{a}_L \wedge \boldsymbol{b}_L)\tilde{\tilde{g}}(\boldsymbol{a}_I \vee \boldsymbol{b}_I, \boldsymbol{a}_J \vee \boldsymbol{b}_J, \boldsymbol{a}_L \vee \boldsymbol{b}_L). \tag{6.3}$$

Since $\boldsymbol{x}_I \leq \boldsymbol{y}_I < t\boldsymbol{e}, \boldsymbol{x}_J \leq t\boldsymbol{e}$, it follows that (6.3) holds if

$$\tilde{f}(\boldsymbol{x}_I, \boldsymbol{x}_J, \boldsymbol{a}_L)\tilde{g}(\boldsymbol{y}_I, \boldsymbol{b}_J, \boldsymbol{b}_L) \leq \tilde{f}(\boldsymbol{x}_I, \boldsymbol{x}_J, \boldsymbol{a}_L \wedge \boldsymbol{b}_L)\tilde{g}(\boldsymbol{y}_I, \boldsymbol{b}_J, \boldsymbol{a}_L \vee \boldsymbol{b}_L)$$

for $\boldsymbol{b}_J \geq t\boldsymbol{e}, \boldsymbol{a}_L \geq t\boldsymbol{e}$ and $\boldsymbol{b}_L \geq t\boldsymbol{e}$. But this follows from the assumption that $\boldsymbol{X} \leq_{\text{lr}} \boldsymbol{Y}$. Thus (6.2) holds.

Since $\leq_{\text{lr}} \Longrightarrow \leq_{\text{st}}$ (see, e.g., Karlin and Rinott (1980) or Whitt (1982)) it follows from (6.2) that

$$[(\boldsymbol{X}_I, \boldsymbol{X}_J, \boldsymbol{X}_L)|h_t] \leq_{\text{st}} [(\boldsymbol{Y}_I, \boldsymbol{Y}_J, \boldsymbol{Y}_L)|h'_t]. \tag{6.4}$$

Now define, for $i \in \{1, 2, \ldots, n\}$,

$$W_i = \begin{cases} 1, & \text{if } X_i \leq t, \\ 0, & \text{if } X_i > t, \end{cases}$$

and

$$Z_i = \begin{cases} 1, & \text{if } Y_i \leq t, \\ 0, & \text{if } Y_i > t. \end{cases}$$

From (6.4) it follows that

$$[(W_1, W_2, \ldots, W_n)|h_t] \geq_{\text{st}} [(Z_1, Z_2, \ldots, Z_n)|h'_t]. \tag{6.5}$$

The conditional distribution of $\boldsymbol{W}$ given h_t is determined by the $\mu_{A|I\cup J}\ (t|\boldsymbol{x}_I, \boldsymbol{x}_J)$, $A \subset \overline{I \cup J}$, which are the discrete multivariate conditional hazard rate functions conditioned on h_t. This distribution is the one which is associated with the discrete multivariate conditional hazard rate measure Q_{h_t} of $\boldsymbol{X}$ (see Section 4 for its definition). Similarly, the conditional distribution of $\boldsymbol{Z}$ given h'_t is the one associated with the discrete multivariate conditional hazard rate measure $R_{h'_t}$ of $\boldsymbol{Y}$. And (6.5) is equivalent to

$$Q_{h_t} \geq_{\mathrm{st}} R_{h'_t}. \tag{6.6}$$

Since (6.6) has been shown whenever $h_t \leq h'_t$ one obtains (4.3) and this proves (6.1).

It is well known that $\boldsymbol{X} \leq_{\mathrm{lr}} \boldsymbol{Y}$ implies $\boldsymbol{X} \leq_{\mathrm{st}} \boldsymbol{Y}$. Theorem 6.1 gives a stronger result, that is, that $\boldsymbol{X} \leq_{\mathrm{h}} \boldsymbol{Y}$. The order $\leq_{\mathrm{h}}$ enables us to compare the underlying items 'locally' as time progresses, in contrast to the 'global' comparison that the order $\leq_{\mathrm{st}}$ yields. More explicitly, given comparable histories associated with $\boldsymbol{X}$ and $\boldsymbol{Y}$ at time t, the order $\leq_{\mathrm{h}}$ allows us to stochastically compare the predicted behavior of the two underlying systems at the next time point. Such a comparison is not possible by means of the order $\leq_{\mathrm{st}}$ solely.

In Shaked and Shanthikumar (1990) there is an application of the continuous orderings to the area of positive dependence. Several notions of positive dependence, pertaining to the random variables $X_1, X_2, \ldots, X_n$, are obtained in Shaked and Shanthikumar (1990) by requiring, for example, that $\boldsymbol{X} \leq_{\mathrm{h}} \boldsymbol{X}$ or that $\boldsymbol{X} \leq_{\mathrm{lr}} \boldsymbol{X}$. The relationships among the continuous orderings enable one to study the relationships among the various resulting positive dependence notions (see Shaked and Shanthikumar (1990)). These notions were also compared there to other well known positive dependence notions such as the positive association notion of Esary, Proschan and Walkup (1967). In the present paper we have not studied the corresponding analogous discrete positive dependence notions. However we believe that no essential new technical difficulties arise when one tries to study them. One use of Theorem 6.1 is to show that the positive dependence notion defined by $\boldsymbol{X} \leq_{\mathrm{lr}} \boldsymbol{X}$ implies the positive dependence notion defined by $\boldsymbol{X} \leq_{\mathrm{h}} \boldsymbol{X}$.

7. Conclusions and Some Remarks

In this paper we have introduced some discrete probabilistic orderings and have studied the relationships among them. These orderings are discrete analogues of the continuous orderings of Shaked and Shanthikumar (1990), but the technical difficulties which are encountered while studying the discrete orderings are different from those involved with the continuous orderings of Shaked and Shanthikumar (1990).

In Shaked and Shanthikumar (1990) an ordering relation, called the cumulative hazard ordering, denoted by $\leq_{\mathrm{ch}}$, is also studied. An analogue of this ordering is not studied here because, as of yet, we do not know what a "correct" discrete analogue of $\leq_{\mathrm{ch}}$ should look like; see Valdez-Torres (1989).

Shaked and Shanthikumar (1991a) used the orderings of Shaked and Shanthikumar (1990) in order to define several multivariate aging notions for continuous dependent random lifetimes such as MIFR (multivariate increasing failure rate) and a kind of multivariate logconcavity which was called MPF_2 (multivariate Polya frequency of order 2). Similar discrete analogues can be developed using the discrete multivariate orderings of the present paper. We may do it elsewhere.

Acknowledgement

The research of the first author was supported by NSF Grant DMS 9303891. The research of the second author was supported by NSF Grant DMS 9308149.

References

Esary, J. D., Proschan, F. and Walkup, D. W. (1967). Association of random variables, with applications. *Ann. Math. Statist.* **38**, 1466-1474.

Freund, J. E. (1961). A bivariate extension of the exponential distribution. *J. Amer. Statist. Assoc.* **56**, 971-977.

Kamae, T., Krengel, U. and O'Brien, G. L. (1977). Stochastic inequalities on partially ordered spaces. *Ann. Probab.* **5**, 899-912.

Karlin, S. and Rinott, Y. (1980). Classes of orderings of measures and related correlation inequalities. I. Multivariate totally positive distributions. *J. Multivariate Anal.* **10**, 467-498.

Ross, S. M. (1984). A model in which component failure rates depend on the working set. *Naval Research Logistics Quarterly* **31**, 297-300.

Shaked, M. and Shanthikumar, J. G. (1990). Multivariate stochastic orderings and positive dependence in reliability theory. *Math. Oper. Res.* **15**, 545-552.

Shaked, M. and Shanthikumar, J. G. (1991a). Dynamic multivariate aging notions in reliability theory. *Stochastic Process. Appl.* **38**, 85-97.

Shaked, M. and Shanthikumar, J. G. (1991b). Dynamic construction and simulation of random vectors. *Topics in Statistical Dependence* (Edited by H. W. Block, A. Sampson and T. H. Savits), 415-433, IMS Lecture Notes.

Shaked, M., Shanthikumar, J. G. and Valdez-Torres, J. B. (1994). Discrete hazard rate functions. *Comput. Oper. Res.*, to appear.

Valdez-Torres, J. B. (1989). Multivariate discrete failure rates with some applications. Doctoral Dissertation, University of Arizona.

Whitt, W. (1982). Multivariate monotone likelihood ratio and uniform conditional stochastic order. *J. Appl. Probab.* **19**, 695-701.

Department of Mathematics, University of Arizona, Tucson, AR 85721, U.S.A.
School of Business Administration, University of California, Berkeley, CA 94720, U.S.A.
Escuela de Ciencias Quimicas, Universidad Autonoma de Sinaloa, Culiacan, Sinaloa, Mexico.

(Received March 1993; accepted January 1994)