MULTIVARIATE DISCRETE PHASE-TYPE DISTRIBUTIONS

By

MATTHEW GOFF

A dissertation submitted in partial fulfillment of

the requirements for the degree of

DOCTOR OF PHILOSOPHY

WASHINGTON STATE UNIVERSITY

Department of Mathematics

May 2005

To the Faculty of Washington State University:

The members of the Committee appointed to examine the dissertation of MATTHEW GOFF find it satisfactory and recommend that it be accepted.

Chair

ACKNOWLEDGMENT

I would like to thank my adviser, Haijun Li, for all of the help he has provided. It was with significant help from him that I was able to find a suitable topic. With his guidance the process of researching, writing, and defending this dissertation went much more smoothly than I had any reason to expect and for that I am very grateful. I also appreciate the financial support I was given by him as a research assistant funded by NSF grant DMI 9812994.

MULTIVARIATE DISCRETE PHASE-TYPE DISTRIBUTIONS

Abstract

by Matthew Goff, Ph.D.
Washington State University
May 2005

Chair: Haijun Li

Many stochastic models involve in one way or another probability distributions of *phase-type*. The family of univariate phase type distributions was introduced by Neuts in [14] as a tool for unifying a variety of stochastic models and for constructing new models that would yield to algorithmic analysis. A class of multivariate phase-type (MPH) distributions was developed by Assaf et al. ([2]). Following Assaf et al., we develop a class of multivariate discrete phase-type (MDPH) distributions.

We demonstrate that MDPH distributions satisfy a number of closure properties and show how they are linked closely with MPH distributions. We develop a number of conditions related to dependency structures of MDPH distributions and show how the link between MPH and MDPH distributions provides a means of analyzing the dependency structure of MPH distributions. A thorough analysis of simple bivariate MDPH distributions is provided and we use the results to find necessary and sufficient conditions for negative dependent bivariate MPH distributions. Two applications of MDPH distributions are given. In the first application, periodic inspection of systems with component lifetimes having a MPH distribution naturally leads to a MDPH distribution. In the final application we derive a simple formula for the mean time to failure of coherent life systems with component lifetimes having an MDPH distribution.

Contents

Chapter 1

Introduction

This dissertation presents a theory of *multivariate discrete phase type* (MDPH) distributions which consists of the representation, structural properties, and some computational methods. Chapter 2 focuses on the representation of MDPH distributions. Chapter 3 establishes useful closure and dependence properties. Chapter 4 provides a complete analysis of bivariate discrete phase type distributions. Finally, Chapter 5 presents some algorithms for computation and applications of MDPH distributions. Code implementing algorithms described and used in Chapter 5 is given in the Appendix. In this chapter we discuss the motivation for this work and summarize the main results.

Many stochastic models involve in one way or another probability distributions of *phase-type*. The family of univariate phase type distributions was introduced by Neuts in [14] as a tool for unifying a variety of stochastic models and for constructing new models that would yield to algorithmic analysis. A non-negative random variable T (or its distribution function) is said to be of phase-type (PH) if T is the time until absorption in a finite-state continuous-time Markov chain $\{X(t), t \geq 0\}$ with state space S and an absorbing state Δ. That is, T has a PH distribution if

$$T = \inf\{t : X(t) = \Delta\}.$$

As an example, the Erlang distribution (the distribution of i.i.d. exponentially distributed random variables) is of phase type. Univariate PH distributions (and their densities, Laplace transforms and all moments) can be written in closed form in terms of the parameters of the underlying Markov chains. Furthermore the set of univariate PH distributions is dense in the set of all distributions with support in $[0, \infty)$. The theory of univariate PH distributions has been developed by several researchers and has been found to have various applications in queueing theory and reliability theory (see [14]), and in risk management and finance ([1], [17]).

At least two classes of multivariate phase-type distributions have been introduced and studied in the literature ([2], [9]). Following Assaf et al. in [2], consider a continuous-time and right-continuous Markov chain $X = \{X(t), t \geq 0\}$ on a finite state space S with generator Q. Let Σ_i, $i = 1, \dots, m$, be m nonempty stochastically closed subsets of S such that $\cap_{i=1}^m \Sigma_i$ is a proper subset of $\mathcal{E}$ (A subset of the state space is said to be stochastically closed if once the process X enters it, X never leaves). We assume that absorption into $\cap_{i=1}^m \Sigma_i$ is certain. Since we are interested in the process only until it is absorbed into $\cap_{i=1}^m \Sigma_i$, we may assume, without loss of generality, that $\cap_{i=1}^m \Sigma_i$ consists of one state, which we shall denote by Δ. Let β be an initial probability vector on S such that $\beta(\Delta) = 0$. Then we can write $\beta = (\alpha, 0)$. Define

$$T_i = \inf\{t \geq 0 : X(t) \in \Sigma_i\}, \quad i = 1, \dots, m. \tag{1.1}$$

It is assumed that $P(T_1 > 0, T_2 > 0, \dots, T_m > 0) = 1$. The joint distribution of $(T_1, T_2, \dots, T_m)$ is called a multivariate phase-type (MPH) distribution with underlying Markov chain $(\alpha, Q, \mathcal{E}_i, i = 1, \dots, m)$, and $(T_1, T_2, \dots, T_m)$ is called a (multivariate) phase-type random vector. As in the univariate case, MPH distributions (and their densities, Laplace transforms and moments) can be written in a closed form. The set of m-dimensional MPH distributions is dense in the set of all distributions on $[0, \infty)^m$, and thus any non-negative m-dimensional distribution can be approximated by the MPH distributions. Hence MPH provides a powerful and versatile tool in the study of multivariate stochastic models. Kulkarni in [9] introduced a more general class of multivariate PH distributions (MPH*), based on the total accumulated reward until absorption in a continuous-time Markov chain. However, we did not develop such a class of multivariate discrete phase-type distributions.

An example of MPH distribution is the Marshall-Olkin distribution ([15]); one of the most widely discussed multivariate life distributions in reliability theory (see [3]). The multivariate Marshall-Olkin distribution has also been used recently to model certain correlated queueing systems, such as assemble-to-order systems (see [11]). Due to their complex structure, however, the applications of general MPH distributions have been limited. Indeed, the structural properties of MPH distributions depend not only on the underlying Markov chains, but also the interplay among the overlapping stochastically closed subsets Σ_i, $i = 1, \dots, m$. To the best of our knowledge, [10] is perhaps the only paper in the literature focusing on the positive dependence properties of MPH distributions. Some properties, especially these concerning the dependence structure, are still unknown.

Another problem is with the computation of MPH distributions. As in the univariate case, MPH distributions have closed form solutions to many problems. However, the closed form solutions for stochastic models with multivariate phase-type distributions are in general still computationally intensive. For example, how does one approximate a multivariate non-negative distribution via some "simpler" MPH distribution? In the univariate case, there have been extensive studies on these issues. In contrast, many computational problems for MPH distributions are still open.

To facilitate the analysis of MPH distributions, we introduce a different approach. We define multivariate discrete phase-type (MDPH) distributions and find that MPH distributions can be considered a composition of MDPH and Erlang distributions. We decompose the underlying Markov chain that defines an MPH and separate the underlying event process from state transitions. The nature of this separation suggests different approaches to the analysis of MPH distributions and in the future may allow for the analysis of more general models where the event process is not a Poisson process.

After defining MDPH distributions, we look at the underlying Markov chains and provide a canonical form for the transition matrix. The structure of the canonical form allows for an easier exploration of the properties of a MDPH distribution and its underlying Markov chain. We also obtain sufficient conditions for simplifying an underlying Markov chains for a given MDPH distribution. Some closure properties are straightforward to derive by construction of an appropriate underlying Markov chain. There are results that show marginals, order statistics, and the concatenation of independent distributions are all MDPH distributions if the original distributions were MDPH.

We consider some dependence properties of MDPH distributions with emphasis on bivariate MDPH distributions and show that the positive dependence properties of underlying MDPH distributions are inherited by the overlying MPH distributions. Our result also illustrate differences between MPH and MDPH distributions. Unlike in the continuous case, the discrete Marshall-Olkin distribution may not be positively dependent in some cases. Also unlike the continuous case, that the discrete Freund distribution cannot be positively dependent. We establish a sufficient condition for the discrete Freund distribution to be negatively dependent. This condition settles an open question for the Freund distributions. Li in [10] showed that if the probability that a component i is destroyed in a Freund distribution becomes *larger* when the other components fail first, the Freund distribution is *positively* dependent. Our results show that if the probability that a component i is destroyed becomes *smaller* when the other components fail first, the Freund distribution is *negatively* dependent. Note that, unlike the case with positive dependence, the analysis of negative dependence poses a considerable challenge because of lack of tools such as stochastic monotonicity.

We present a means of directly computing MDPH probabilities over a class of regions. We also develop an algorithm for generating random vectors from MDPH distributions. A straightforward addition to the algorithm also allows for the generation of random vectors from MPH distributions. Implementations of the algorithms in **R** [16], are provided in the Appendix along with the code used in various examples.

In addition to the motivation provided by the relationship of MDPH distributions with MPH distributions, MDPH distributions are of interest in their own right. We illustrate the usefulness of MDPH distributions through two applications from reliability modeling: periodic inspections of systems with MPH distributed component lifetimes and coherent systems with dependent component lifetimes having a MDPH distribution.

Throughout this dissertation, the term 'increasing' and 'decreasing' mean 'non-decreasing' and 'non-increasing' respectively, and the measurability of sets and functions as well as the existence of expectations are often assumed without explicit mention. We also assume that all the states of a Markov chain are reachable.

Chapter 2

Multivariate Discrete Phase Type Distribution and its Representation

In this chapter we introduce the multivariate discrete phase type (MDPH) distributions and related notations for their representation. We study some basic properties in Section 2.2, and in Section 2.3 also discuss two examples of MDPH distributions from reliability modeling.

A discrete phase type (DPH) random variable is defined as the number of transitions it takes for a discrete-time Markov chain to enter an absorbing subset of the state space. As such, the properties and dependence structure of a MDPH distribution depend on its underlying Markov chain and the absorbing sets. Since the Markov chain underlying a MDPH distribution is not unique, we focus on the equivalence and reducibility of these underlying Markov chains. We obtain a sufficient condition under which a Markov chain representation can be reduced to one with simpler structure, and also show that a simple Markov chain representation for an MDPH must be unique.

2.1 Definitions and Notation

Let $\{\mathbb{S}^n, n \geq 0\}$ be a discrete-time Markov chain with a finite state space S. Let $\mathbf{\Sigma} = \{\Sigma_1, \ldots, \Sigma_m\}$ be a set of non-empty stochastically closed subsets such that $\Sigma_\Delta = \cap_{k=1}^m \Sigma_k \neq \emptyset$. A subset of the state space is called stochastically closed if when the Markov chain $\{\mathbb{S}^n, n \geq 0\}$ enters that subset, it subsequently never leaves. We also define $\Sigma_0 = \bigcap\limits_{i=1}^{m} \Sigma_i^c$. We assume that absorption in Σ_Δ is certain. Let $\mathfrak{S} = [\mathfrak{s}_{ij}]$ be the transition matrix of the process. Hence, $\mathfrak{s}_{ij} = 0$ if $i \in \Sigma_k$ and $j \in \Sigma_k^c$ for some k. Also let σ be an initial probability vector on S with $\sigma(\Sigma_\Delta) = 0$ then $\{\mathbb{S}^n, n \geq 0\}$ will also be written as $\mathbb{S} = \{S, \mathbf{\Sigma}, \mathfrak{S}, \sigma\}$ in order to make explicit the algebraic structures underlying the Markov chain.

Definition 1 (Multivariate Discrete Phase Type Distribution) *Let*

$$\mathrm{S}_j = \min\{n \mid \mathbb{S}^n \in \Sigma_j\}, j = 1, \ldots, m.$$

Then the joint distribution of $\mathbf{S} = (\mathrm{S}_1, \ldots, \mathrm{S}_m)$,

$$\mathbf{P}\{\mathrm{S}_1 \leq n_1, \ldots, \mathrm{S}_m \leq n_m\} = \mathbf{P}\{\min\{n \mid \mathbb{S}^n \in \Sigma_1\} \leq n_1, \ldots, \min\{n \mid \mathbb{S}^n \in \Sigma_m\} \leq n_m\},$$

is called a multivariate discrete phase-type distribution (MDPH).

Theorem 1 (Calculating Survival Probabilities) *For a vector of non-negative integers* $(n_1, \ldots, n_m)$, *let* $\mathbf{o}$ *be a permutation of* $1, 2, \ldots, m$ *such that* $n_{o_1} \leq n_{o_2} \leq \ldots \leq n_{o_m}$. *Taking*

$$\Gamma_i = \bigcap_{j=i}^{m} \Sigma_{o_j}^c, \; i = 1, \ldots, m,$$

(note: $\Gamma_1 = \Sigma_0$*) and, for* $\bar{S} \subseteq S$, $\mathbf{I}_{\bar{S}} = [\mathbf{I}_{ij}]$ *an* $|S| \times |S|$ *diagonal matrix where*

$$\mathbf{I}_{ii} = \begin{cases} 0 & \text{if } s_i \notin \bar{S} \\ 1 & \text{if } s_i \in \bar{S}, \end{cases}$$

then the joint survival function of an MDPH with the underlying Markov chain $\mathbb{S} = \{S, \mathbf{\Sigma}, \mathfrak{S}, \sigma\}$ *is given by*

$$\mathbf{P}\{\mathrm{S}_1 > n_1, \ldots, \mathrm{S}_m > n_m\} = \sigma \cdot \mathfrak{S}^{n_{o_1}} \cdot \mathbf{I}_{\Gamma_1} \cdot \prod_{i=2}^{m} \left(\mathfrak{S}^{n_{o_i} - n_{o_{i-1}}} \cdot \mathbf{I}_{\Gamma_i}\right) \cdot \underline{\mathbf{1}}$$

where $\underline{\mathbf{1}}$ *is a column vector of 1's.*

Proof:

Obviously, we have

$$\mathbf{P}\{\mathrm{S}_1 > n_1, \ldots, \mathrm{S}_m > n_m\} = \mathbf{P}\{\mathbb{S}^{n_1} \in \Sigma_1^c, \ldots, \mathbb{S}^{n_m} \in \Sigma_m^c\}.$$

Note that $\sigma \cdot \mathfrak{S}^{n_{o_1}}$ gives the probability distribution after n_{o_1} transitions. Multiplying by $\mathbf{I}_{\Gamma_1}$ restricts the possibilities to only those that satisfy $\mathrm{S}_{o_j} > n_{o_1}, j = 1, 2, \ldots, m$. Similarly each further pair of matrix factors $\mathfrak{S}^{n_{o_i} - n_{o_{i-1}}} \cdot \mathbf{I}_{\Gamma_i}$ gives the probability distribution of states after n_{o_i} total transitions restricted to those which satisfy $\mathrm{S}_{o_j} > n_{o_i}, j = i, i+1, \ldots, m$. Finally, the last probability distribution is multiplied by a column vector of 1's to give the overall survival probability for the MDPH distribution. □

The underlying Markov chain for an MDPH may not be unique, and so a natural question is how to find an underlying Markov chain with simpler structure.

Definition 2 (Equivalence) *If* $\mathbb{S} = \{S, \mathbf{\Sigma}, \mathfrak{S}, \sigma\}$ *and* $\mathbb{S}^* = \{S^*, \mathbf{\Sigma}^*, \mathfrak{S}^*, \sigma^*\}$ *are two Markov chains underlying the same phase type distribution,* $\mathbb{S}$ *and* $\mathbb{S}^*$ *are said to be equivalent.*

Let $\mathbf{\Psi}$ be a vector valued function having domain S and range $\{0, 1\}^m$, with

$$\mathbf{\Psi}_i(s) = \begin{cases} 0 & \text{if } s \notin \Sigma_i \\ 1 & \text{if } s \in \Sigma_i, \end{cases}$$

for $1 \leq i \leq m$. Let

$$S_k = \{s \mid \sum_{i=1}^{m} \mathbf{\Psi}_i(s) 2^{m-i} = k\},\ k = 0, 1, \ldots, 2^m - 1,$$

then $\{S_k\}$ is a partition of S.

Definition 3 (Simple) *An underlying Markov chain for an m-dimensional MDPH is called simple if* $|S_k| = 1$ *for* $k = 0, 1, \ldots, 2^m - 1$.

Definition 4 (Order) *If a random vector* $\mathbf{S}$ *has an MDPH distribution with underlying Markov chain* $\mathbb{S} = \{S, \mathbf{\Sigma}, \mathfrak{S}, \sigma\}$ *and* $|S| = N$*, then* $\mathbf{S}$ *is said to be of order* N.

Definition 5 (Minimal) *An underlying Markov chain* $\mathbb{S} = \{S, \mathbf{\Sigma}, \mathfrak{S}, \sigma\}$ *for* $\mathbf{S}$ *with an MDPH distribution is called minimal if for any equivalent Markov chain* $\mathbb{S}^* = \{S^*, \mathbf{\Sigma}^*, \mathfrak{S}^*, \sigma^*\}$, $|S| \leq |S^*|$.

Definition 6 (Reducible) *An underlying Markov chain* $\mathbb{S} = \{S, \mathbf{\Sigma}, \mathfrak{S}, \sigma\}$ *for an MDPH distributed random vector* $\mathbf{S}$ *is said to be reducible if one or more states in* S *can be combined to form a state space* S^* *for an equivalent Markov chain* $\mathbb{S}^* = \{S^*, \mathbf{\Sigma}^*, \mathfrak{S}^*, \sigma^*\}$.

Example 1 If $\mathbf{S}$ has an m-dimensional MDPH distribution with underlying Markov chain $\mathbb{S} = \{S, \mathbf{\Sigma}, \mathfrak{S}, \sigma\}$ and $|S_{2^m-1}| > 1$, then $\mathbb{S} = \{S, \mathbf{\Sigma}, \mathfrak{S}, \sigma\}$ is reducible. When the Markov chain enters S_{2^m-1}, $\mathbf{S}$ has been completely realized and whatever takes place inside S_{2^m-1} from this point onward does not change the distribution of $\mathbf{S}$. Therefore, from the standpoint of $\mathbf{S}$, there needs to be only one state in S_{2^m-1}, and $\mathbb{S} = \{S, \mathbf{\Sigma}, \mathfrak{S}, \sigma\}$ can be reduced accordingly.

Definition 7 (Multivariate Geometric) *An MDPH distribution is said to be a multivariate geometric distribution if each univariate marginal distribution has an underlying Markov chain which is simple (and therefore, geometric).*

2.2 Properties and Algebra of Underlying Markov Chains

2.2.1 The Transition Matrix

The transition matrix $\mathfrak{S}$ is the key to studying and establishing many properties of MDPH distributions. As such, it is helpful to investigate its structure. The transition matrix $\mathfrak{S}$ can be partitioned and written as

$$\mathfrak{S} = \begin{pmatrix} \mathfrak{S}_{0,0} & \mathfrak{S}_{0,1} & \dots & \mathfrak{S}_{0,2^m-1} \\ \mathfrak{S}_{1,0} & \mathfrak{S}_{1,1} & \dots & \mathfrak{S}_{1,2^m-1} \\ \vdots & \vdots & \ddots & \vdots \\ \mathfrak{S}_{2^m-1,0} & \mathfrak{S}_{2^m-1,1} & \dots & \mathfrak{S}_{2^m-1,2^m-1} \end{pmatrix},$$

where $\mathfrak{S}_{i,j}$ corresponds to the transition probabilities from states in S_i to states in S_j.

Theorem 2 (Cannonical Form of Transition Matrix) *The transition matrix of an underlying Markov chain for an m-dimensional MDPH distribution can always be written in a recursive block upper triangular form with $m-1$ levels.*

Proof:
Consider $\mathfrak{S}_{i,j}$ which is the transition probabilities from states in S_i to states in S_j. The states in S_i are all mapped to a vector by $\mathbf{\Psi}$ and this vector can be seen as the binary representation of i. However, this vector is also an indicator vector for the m stochastically closed classes of S. Therefore, if i has a 1 in the binary representation where j does not, then $\mathfrak{S}_{ij} = 0$. Using this observation, we can construct a general transition matrix which includes the 0's induced by the stochastically closed classes.

Taking $\mathfrak{S}^{(m)}$ to be the structure of the transition matrix for an underlying Markov chain of an m-dimensional MDPH distribution, in univariate case,

$$\mathfrak{S}^{(1)} = \begin{pmatrix} X & X \\ 0 & X \end{pmatrix}$$

where X denotes a non-zero block (which may not be the same for each position). In the bivariate case,

$$\mathfrak{S}^{(2)} = \left(\begin{array}{cc|cc} X & X & X & X \\ 0 & X & 0 & X \\ \hline 0 & 0 & X & X \\ 0 & 0 & 0 & X \end{array}\right) = \left(\begin{array}{c|c} \mathfrak{S}^{(1)} & \mathfrak{S}^{(1)} \\ \hline 0 & \mathfrak{S}^{(1)} \end{array}\right).$$

The general relationship is given by,

$$\mathfrak{S}^{(m)} = \left(\begin{array}{c|c} \mathfrak{S}^{(m-1)} & \mathfrak{S}^{(m-1)} \\ \hline 0 & \mathfrak{S}^{(m-1)} \end{array} \right).$$

Note, again, that the blocks in the above transition matrices are generic and may not be the same.

To better understand why this is the case, consider what each block represents. The lower left block corresponds to transitions from Σ_m to Σ_m^c so it clearly must be 0. The upper left block corresponds to transitions between states in Σ_m^c. The possible transitions are exactly those which can occur in the $m-1$ dimensional case, that is transitions into the other $m-1$ stochastically closed classes. The upper right block corresponds to the same transitions, but here they are accompanied by a transition from Σ_m^c to Σ_m. Finally, the lower right block corresponds to transitions between states in Σ_m and they are again precisely those which can occur in the $m-1$ dimensional case. □

2.2.2 Reducibility and Equivalence

Let $\mathbf{S}$ have an m-dimensional MDPH with $\mathbb{S} = \{S, \mathbf{\Sigma}, \mathfrak{S}, \sigma\}$. Suppose that h states $j_1, j_2, \ldots, j_h \in S_k$ for some k satisfy that $\forall i \neq j_1, j_2, \ldots, j_h$, $\mathfrak{s}_{j_1 i} = \mathfrak{s}_{j_2 i} = \cdots = \mathfrak{s}_{j_h i}$, where $\mathfrak{s}_{ji}$ is the transition probability from states j to i. Note that for states $j_1, j_2, \ldots, j_h$,

$$\sum_{l=j_1}^{j_h} \mathfrak{s}_{j_1 l} = \cdots = \sum_{l=j_1}^{j_h} \mathfrak{s}_{j_h l}.$$

Let $\mathbb{S}^* = \{S^*, \mathbf{\Sigma}^*, \mathfrak{S}^*, \sigma^*\} = \{\mathbb{S}^{*n}, n \geq 0\}$ be a new Markov chain, where $S^* = S \cup \{j^*\} - \{j_1, \ldots, j_h\}$ with a new state j^* added into $S_k^* = S_k \cup \{j^*\} - \{j_1, \ldots, j_h\}$. Σ^* is the same as Σ with the obvious adjustment to accommodate the reduction. $\mathfrak{S}^*$ and σ^* are the same as $\mathfrak{S}$ and σ with the column vectors corresponding to $j_1, \ldots, j_h$ replaced by the vector of their entry-wise sums, which is the column corresponding to the new state j^*. The summing of the columns results in h identical rows, corresponding to $j_1, \ldots, j_h$. The $h-1$ rows are removed and the other corresponds to j^*. Note that

$$\begin{aligned} \mathbf{P}\{\mathbb{S}^{*n} = j^* \mid \mathbb{S}^{*n-1} = j^*\} &= \sum_{l=j_1}^{j_h} \mathfrak{s}_{j_1 l} \\ \mathbf{P}\{\mathbb{S}^{*n} = i \mid \mathbb{S}^{*n-1} = j^*\} &= \mathfrak{s}_{j_1 i}, \end{aligned}$$

$\forall i \neq j_1, j_2, \ldots, j_h$. Such new state j^* is called a condensation of $\{j_1, \ldots, j_h\}$.

Theorem 3 (Sufficient Condition for Reducibility) *Let* $\mathbf{S}$ *have an m-dimensional MDPH with* $\mathbb{S} = \{S, \mathbf{\Sigma}, \mathfrak{S}, \sigma\}$*, and* $\mathbb{S}^* = \{S^*, \mathbf{\Sigma}^*, \mathfrak{S}^*, \sigma^*\}$ *be the Markov chain described above. Then* $\mathbb{S} = \{S, \mathbf{\Sigma}, \mathfrak{S}, \sigma\}$ *is reducible to* $\mathbb{S}^* = \{S^*, \mathbf{\Sigma}^*, \mathfrak{S}^*, \sigma^*\}$*.*

Proof:

Note that the only difference between $\mathbb{S}$ and $\mathbb{S}^*$ is that states $j_1, \ldots, j_h$ in S are condensed into one state j^* in S^* in the sense that

1. the transition probability from any state to j^* is the sum of the transition probabilities from that state into subset $\{j_1, \ldots, j_h\}$, and
2. the transition probability from j^* to any other state is the transition probability from j_l $(1 \leq l \leq h)$ to that state.

It suffices to show that the exit probability of $\mathbb{S}$ from $\{j_1, \ldots, j_h\}$ to state i after n transitions within $\{j_1, \ldots, j_h\}$ equals the exit probability of $\mathbb{S}^*$ from j^* to state i after n transitions at j^*. For this, let $a_{j_1}, \ldots, a_{j_h}$ be non-negative real numbers with $\sum_{l=1}^{h} a_{j_l} = a \leq 1$. We then have, for any $n \geq 0$ and any $i \neq j_1, \ldots, j_h$,

$$
\begin{aligned}
& (a_{j_1}, \ldots, a_{j_h}) \begin{pmatrix} \mathfrak{s}_{j_1 j_1} & \mathfrak{s}_{j_1 j_2} & \cdots & \mathfrak{s}_{j_1 j_h} \\ \mathfrak{s}_{j_2 j_1} & \mathfrak{s}_{j_2 j_2} & \cdots & \mathfrak{s}_{j_2 j_h} \\ \cdots & \cdots & \cdots & \cdots \\ \mathfrak{s}_{j_h j_1} & \mathfrak{s}_{j_h j_2} & \cdots & \mathfrak{s}_{j_h j_h} \end{pmatrix}^n \begin{pmatrix} \mathfrak{s}_{j_1 i} \\ \mathfrak{s}_{j_2 i} \\ \cdots \\ \mathfrak{s}_{j_h i} \end{pmatrix} \\
= \; & (a_{j_1}, \ldots, a_{j_h}) \begin{pmatrix} \mathfrak{s}_{j_1 j_1} & \mathfrak{s}_{j_1 j_2} & \cdots & \mathfrak{s}_{j_1 j_h} \\ \mathfrak{s}_{j_2 j_1} & \mathfrak{s}_{j_2 j_2} & \cdots & \mathfrak{s}_{j_2 j_h} \\ \cdots & \cdots & \cdots & \cdots \\ \mathfrak{s}_{j_h j_1} & \mathfrak{s}_{j_h j_2} & \cdots & \mathfrak{s}_{j_h j_h} \end{pmatrix}^n \underline{\mathbf{1}} \mathfrak{s}_{j_1 i} \\
= \; & (a_{j_1}, \ldots, a_{j_h}) \begin{pmatrix} \mathfrak{s}_{j_1 j_1} & \mathfrak{s}_{j_1 j_2} & \cdots & \mathfrak{s}_{j_1 j_h} \\ \mathfrak{s}_{j_2 j_1} & \mathfrak{s}_{j_2 j_2} & \cdots & \mathfrak{s}_{j_2 j_h} \\ \cdots & \cdots & \cdots & \cdots \\ \mathfrak{s}_{j_h j_1} & \mathfrak{s}_{j_h j_2} & \cdots & \mathfrak{s}_{j_h j_h} \end{pmatrix}^{n-1} \underline{\mathbf{1}} (\sum_{l=j_1}^{j_h} \mathfrak{s}_{j_1 l}) \mathfrak{s}_{j_1 i} \\
= \; & (a_{j_1}, \ldots, a_{j_h}) \underline{\mathbf{1}} (\sum_{l=j_1}^{j_h} \mathfrak{s}_{j_1 l})^n \mathfrak{s}_{j_1 i} \\
= \; & (\sum_{l=1}^{h} a_{j_l})(\sum_{l=j_1}^{j_h} \mathfrak{s}_{j_1 l})^n \mathfrak{s}_{j_1 i},
\end{aligned}
$$

where $\underline{\mathbf{1}}$ is the column vector of 1's. Thus, from the standpoint of $\mathbf{S}$, $\{j_1, \ldots, j_h\}$ can be considered as a single state and the state space and transition matrix can be modified accordingly to form an equivalent underlying Markov chain for $\mathbf{S}$. $\square$

As discussed in Example 1, Σ_Δ can always be considered a single state.

Corollary 1 *If* $\mathbf{S}$ *is of order* N*, then* $\mathbf{S}$ *is also of order* $N+1$*.*

Proof:

This follows immediately from the reduction theorem. Simply split one state into a pair of states to create a new underlying Markov chain which can be reduced to the original. $\square$

Theorem 4 (Simple underlying Markov chain is unique) *If* $\mathbf{S}$ *has an MDPH distribution with a simple underlying Markov chain* $\mathbb{S} = \{S, \mathbf{\Sigma}, \mathfrak{S}, \sigma\}$*,* $\mathbb{S}$ *is unique in the sense that all the simple underlying Markov chains for* $\mathbf{S}$ *are stochastically identical.*

Proof:

Suppose $\mathbb{S} = \{S, \mathbf{\Sigma}, \mathfrak{S}, \sigma\}$ and $\mathbb{S}' = \{S, \mathbf{\Sigma}, \mathfrak{S}', \sigma'\}$ are two simple underlying Markov chains for $\mathbf{S} = (\mathrm{S}_1, \ldots, \mathrm{S}_m)$. Since they are simple, S_k, where $0 \le k \le 2^m - 1$, contains exactly one state, which we also denote by S_k. Thus, without loss of generality we may assume that the state space and stochastically closed classes are the same for both Markov chains. Due to the way that S_k was defined, each S_k is in a unique set of Σ_i, $i = 1, \cdots, m$. So there is a natural association between each S_k, $0 \le k \le 2^m - 1$ and the sets I of the power set of $\{1, \ldots, m\}$, with $\emptyset = S_0$ and $\{1, \ldots, m\} = S_{2^m-1}$. We will use the following interpretation: the Markov chain is in state $I \subseteq \{1, \ldots, m\}$ at time n if and only if $\mathrm{S}_i \le n$ for all and only $i \in I$.

Clearly $\sigma^* = \sigma'$. To show the transition matrices $\mathfrak{S}$ and $\mathfrak{S}'$ to be the same, we have, for any $I, J \subseteq \{1, \ldots, m\}$,

$$\begin{aligned} \mathfrak{s}_{I,J} &= \mathbf{P}\{\mathbb{S}^1 = J \mid \mathbb{S}^0 = I\} \\ &= \mathbf{P}\{\mathrm{S}_i \le 1 \text{ for all and only } i \in J \mid \mathrm{S}_i \le 0, \text{ for all and only } i \in I\} \\ &= \mathbf{P}\{\mathbb{S}'^1 = J \mid \mathbb{S}'^0 = I\} = \mathfrak{s}'_{I,J}. \end{aligned}$$

Hence $\mathbb{S} = \{S, \mathbf{\Sigma}, \mathfrak{S}, \sigma\}$ and $\mathbb{S}' = \{S, \mathbf{\Sigma}, \mathfrak{S}', \sigma'\}$ must be the same. $\square$

Note that the above theorem only says that simple underlying Markov chain of an MDPH is unique. It follows from Theorem 3 that a simple MDPH can have different, non-simple underlying Markov chains.

Example 2 If $\mathbf{S}$ does not have a simple underlying Markov chain, the underlying Markov chain needs not be the only one of the same order. Consider the following two Markov chains $\mathbb{S} = \{S, \mathbf{\Sigma}, \mathfrak{S}, \sigma\}$ and

$\hat{\mathbb{S}} = \{S, \mathbf{\Sigma}, \hat{\mathfrak{S}}, \hat{\sigma}\}$ with

$$S = \{s_1, s_2, s_3, s_4, s_5, s_6\},$$

$$\Sigma_1 = \{s_3, s_6\}, \Sigma_2 = \{s_4, s_5, s_6\},$$

$$\mathfrak{S} = \begin{pmatrix} \frac{9}{50} & \frac{9}{50} & \frac{12}{50} & \frac{6}{50} & \frac{6}{50} & \frac{8}{50} \\ \frac{3}{50} & \frac{18}{50} & \frac{9}{50} & \frac{2}{50} & \frac{12}{50} & \frac{6}{50} \\ 0 & 0 & \frac{3}{5} & 0 & 0 & \frac{2}{5} \\ 0 & 0 & 0 & \frac{3}{10} & \frac{3}{10} & \frac{4}{10} \\ 0 & 0 & 0 & \frac{1}{10} & \frac{6}{10} & \frac{3}{10} \\ 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix},$$

$$\hat{\mathfrak{S}} = \begin{pmatrix} \frac{315}{1000} & \frac{225}{1000} & \frac{60}{1000} & \frac{354}{1000} & \frac{6}{1000} & \frac{40}{1000} \\ \frac{3}{40} & \frac{9}{40} & \frac{12}{40} & \frac{2}{40} & \frac{6}{40} & \frac{8}{40} \\ 0 & 0 & \frac{3}{5} & 0 & 0 & \frac{2}{5} \\ 0 & 0 & 0 & \frac{1263}{2200} & \frac{357}{2200} & \frac{580}{2200} \\ 0 & 0 & 0 & \frac{503}{2200} & \frac{717}{2200} & \frac{980}{2200} \\ 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix},$$

$$\sigma = \begin{bmatrix} \frac{1}{2} & \frac{1}{2} & 0 & 0 & 0 & 0 \end{bmatrix}, \text{ and } \hat{\sigma} = \begin{bmatrix} \frac{3}{8} & \frac{5}{8} & 0 & 0 & 0 & 0 \end{bmatrix}.$$

Note that $\hat{\mathfrak{S}} = P^{-1}\mathfrak{S}P$, where

$$P = \begin{pmatrix} \frac{1}{4} & \frac{3}{4} & 0 & 0 & 0 & 0 \\ \frac{1}{2} & \frac{1}{2} & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & \frac{1}{4} & \frac{3}{4} & 0 \\ 0 & 0 & 0 & \frac{4}{5} & \frac{1}{5} & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}.$$

Furthermore, because the blocks in P match up with the blocks in $\mathfrak{S}$ (which is to say, the blocks induced by $\mathbf{\Sigma}$),

$$P^{-1}\mathbf{I}_{\Gamma_i}P = \mathbf{I}_{\Gamma_i}, i = 1, 2.$$

If $\mathbb{S} = \{S, \mathbf{\Sigma}, \mathfrak{S}, \sigma\}$ is an underlying Markov chain for $\mathbf{S}$ then

$$\begin{aligned}\mathbf{P}\{\mathrm{S}_1 > n_1, \mathrm{S}_2 > n_2\} &= \sigma \cdot \mathfrak{S}^{n_{o_1}} \cdot \mathbf{I}_{\Gamma_1} \cdot \mathfrak{S}^{n_{o_2} - n_{o_1}} \cdot \mathbf{I}_{\Gamma_2} \cdot \underline{\mathbf{1}} \\ &= \sigma P \cdot \left(P^{-1} \mathfrak{S} P\right)^{n_{o_1}} \cdot \mathbf{I}_{\Gamma_1} \cdot \left(P^{-1} \mathfrak{S} P\right)^{n_{o_2} - n_{o_1}} \cdot \mathbf{I}_{\Gamma_2} P^{-1} \cdot \underline{\mathbf{1}} \\ &= \hat{\sigma} \cdot \hat{\mathfrak{S}}^{n_{o_1}} \cdot \mathbf{I}_{\Gamma_1} \cdot \hat{\mathfrak{S}}^{n_{o_2} - n_{o_1}} \cdot \mathbf{I}_{\Gamma_2} \cdot \underline{\mathbf{1}}\end{aligned}$$

which implies that $\hat{\mathbb{S}} = \{S, \mathbf{\Sigma}, \hat{\mathfrak{S}}, \hat{\sigma}\}$ is also an underlying Markov chain for $\mathbf{S}$ and so is equivalent to $\mathbb{S} = \{S, \mathbf{\Sigma}, \mathfrak{S}, \sigma\}$.

2.3 Examples from Reliability Theory

2.3.1 Discrete Marshall-Olkin Distribution

Consider a system with m components subject to multiple types of fatal shocks. Let $\{K_0, K_1, \ldots, K_{2^m-1}\}$ be the power set of $\{1, \ldots, m\}$ where K_i consists of the components j such that $\sum_{j|j \in K_i} 2^{m-j} = i$. Note that $K_0 = \emptyset$ and $K_{2^m-1} = \{1, \ldots, m\}$. Let p_i, $i = 0, \ldots, 2^{m-1}$, be real numbers such that $0 \leq p_i \leq 1$ and $\sum_{i=0}^{2^{m-1}} p_i = 1$. Any arriving shock, with probability p_i, destroys all the components in K_i simultaneously. Let S_j, $j = 1, \ldots, m$, be the number of shocks needed to destroy component j. The joint distribution of $(\mathrm{S}_1, \ldots, \mathrm{S}_m)$ is called a multivariate discrete Marshall-Olkin distribution.

To see that this distribution is a special case of MDPH type distributions, we let $\{X_n, n \geq 0\}$ be a Markov chain with state space $\{K_i, i = 0, \ldots, 2^m - 1\}$, and starting at $X_0 = K_0$ almost surely. State K_i means that all the components in K_i have failed and the others are still operational. The transition probabilities of the chain from state K_i to K_j are given by

$$\mathfrak{s}_{ij} = \sum_{k|K_k \subseteq K_j, K_i \cup K_k = K_j} p_k, \quad K_i \subseteq K_j.$$

Let $\Sigma_j = \{K_i \mid K_i \ni j\}$, $j = 1, \ldots, m$, be the set of failure states of component j. Clearly,

$$\mathrm{S}_j = \inf\{n : X_n \in \Sigma_j\}, j = 1, \ldots, m.$$

Thus, the joint distribution of $(S_1, \ldots, S_m)$ is an MDPH distribution.

Example 3 (Markov Chain for 2D Case) If $m = 2$ then the state space is $\{K_0, K_1, K_2, K_3\}$ with $\Sigma_1 =$

$\{K_1, K_3\}$ and $\Sigma_2 = \{K_2, K_3\}$. The transition matrix will be:

$$\begin{pmatrix} p_0 & p_1 & p_2 & p_3 \\ 0 & p_0 + p_1 & 0 & p_2 + p_3 \\ 0 & 0 & p_0 + p_2 & p_1 + p_3 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

Example 4 (Markov Chain for 3D Case) If $m = 3$ then the state space is

$$\{K_0, K_1, K_2, K_3, K_4, K_5, K_6, K_7\}$$

with $\Sigma_1 = \{K_1, K_3, K_5, K_7\}$, $\Sigma_2 = \{K_2, K_3, K_6, K_7\}$, and $\Sigma_3 = \{K_4, K_5, K_6, K_7\}$. The transition matrix is given by,

$$\begin{pmatrix} p_0 & p_1 & p_2 & p_3 & p_4 & p_5 & p_6 & p_7 \\ 0 & p_0 + p_1 & 0 & p_2 + p_3 & 0 & p_4 + p_5 & 0 & p_6 + p_7 \\ 0 & 0 & p_0 + p_2 & p_1 + p_3 & 0 & 0 & p_4 + p_6 & p_5 + p_7 \\ 0 & 0 & 0 & p_0 + p_1 + p_2 + p_3 & 0 & 0 & 0 & p_4 + p_5 + p_6 + p_7 \\ 0 & 0 & 0 & 0 & p_0 + p_4 & p_1 + p_5 & p_2 + p_6 & p_3 + p_7 \\ 0 & 0 & 0 & 0 & 0 & p_0 + p_1 + p_4 + p_5 & 0 & p_2 + p_3 + p_6 + p_7 \\ 0 & 0 & 0 & 0 & 0 & 0 & p_0 + p_2 + p_4 + p_6 & p_1 + p_3 + p_5 + p_7 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}$$

2.3.2 Discrete Freund Distribution

Consider a system of m components operating in a random environment. A component fails when it receives a fatal shock from the random environment. As long as all the components are functioning, an arriving shock destroys component i with probability p_i, $i = 1, \ldots, m$. With probability p_0, no components are destroyed. Clearly, $0 \leq p_i \leq 1$ and $\sum_{i=0}^{m} p_i = 1$. There is no simultaneous failure of components. Suppose component index i_j corresponds to the jth component failure. Upon the lth component failure, with the order of failures given by $i_1 \ldots i_l$, an arriving shock destroys component i with probability $p_{i|i_1 \ldots i_l}$, $i \neq i_j, j = 1, \ldots, l$ and with probability $p_{0|i_1 \ldots i_l}$, no additional failures occur. Let S_j, $j = 1, \ldots, m$, be the number of shocks needed to destroy component j. Clearly, the discrete lifetime vector $(\mathrm{S}_1, \ldots, \mathrm{S}_m)$ has dependent components. When $m = 2$, the joint distribution of the lifetime vector is a discrete version of the bivariate extension of the exponential distribution introduced by Freund in [6].

To see that this model is an MDPH type distribution, let the state space contain all the permutations of each subset of $\{1, \ldots, m\}$. The initial state of the Markov chain $\{X_n, n \geq 0\}$ is $\emptyset$. Its transition probability matrix is given as follows, for any state $K = i_1 \cdots i_l$, $\mathfrak{s}_{KK} = p_{0|i_1 \ldots i_l}$ and $\mathfrak{s}_{KL} = p_{i|i_1 \ldots i_l}$, if $L = i_1 \ldots i_l i$ and zero otherwise.

Let the set of failure states of component i be $\Sigma_i = \{K : i \in K\}$. Clearly

$$\mathrm{S}_i = \min\{n \mid X_n \in \Sigma_i\}, i = 1, \ldots, m.$$

So the joint distribution of $(S_1, \ldots, S_m)$ is an MDPH distribution.

Example 5 (Markov Chain for 2D Case) If $m = 2$ then the state space is $\{\emptyset, \{1\}, \{2\}, \{12\}\}$ with $\Sigma_1 = \{\{1\}, \{12\}\}$ and $\Sigma_2 = \{\{2\}, \{12\}\}$. The transition matrix is given by,

$$\begin{pmatrix} p_0 & p_1 & p_2 & 0 \\ 0 & p_{0|1} & 0 & p_{2|1} \\ 0 & 0 & p_{0|2} & p_{1|2} \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

Example 6 (Markov Chain for General 3D Case) If $m = 3$ the state space is:

$$\{\emptyset, \{1\}, \{2\}, \{3\}, \{12\}, \{21\}, \{13\}, \{31\}, \{23\}, \{32\}, \{123\}\}$$

with failure sets

$$\Sigma_1 = \{\{1\}, \{12\}, \{21\}, \{13\}, \{31\}, \{123\}\},\ \Sigma_2 = \{\{2\}, \{21\}, \{12\}, \{23\}, \{32\}, \{123\}\},$$

$$\Sigma_3 = \{\{3\}, \{31\}, \{13\}, \{32\}, \{23\}, \{123\}\}.$$

The transition matrix is given by,

$$\begin{pmatrix} p_0 & p_1 & p_2 & p_3 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & p_{0|1} & 0 & 0 & p_{2|1} & p_{3|1} & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & p_{0|2} & 0 & 0 & 0 & p_{1|2} & p_{3|2} & 0 & 0 & 0 \\ 0 & 0 & 0 & p_{0|3} & 0 & 0 & 0 & 0 & p_{1|3} & p_{2|3} & 0 \\ 0 & 0 & 0 & 0 & p_{0|12} & 0 & 0 & 0 & 0 & 0 & p_{3|12} \\ 0 & 0 & 0 & 0 & 0 & p_{0|13} & 0 & 0 & 0 & 0 & p_{2|13} \\ 0 & 0 & 0 & 0 & 0 & 0 & p_{0|21} & 0 & 0 & 0 & p_{3|21} \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & p_{0|23} & 0 & 0 & p_{1|23} \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & p_{0|31} & 0 & p_{2|31} \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & p_{0|32} & p_{1|32} \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}.$$

Example 7 (Markov Chain for Simple 3D Case) In the $m = 3$ case where $p_{i|jk} = p_{i|kj}$ for $i, j, k = 1, 2, 3$ the Markov chain simplifies to have the state space:

$$\{\emptyset, \{1\}, \{2\}, \{3\}, \{12\}, \{13\}, \{23\}, \{123\}\}$$

with absorbing sets

$$\Sigma_1 = \{\{1\}, \{12\}, \{13\}, \{123\}\},$$
$$\Sigma_2 = \{\{2\}, \{12\}, \{23\}, \{123\}\},$$
$$\Sigma_3 = \{\{3\}, \{13\}, \{23\}, \{123\}\}.$$

The transition matrix is given by

$$\begin{pmatrix} p_0 & p_1 & p_2 & 0 & p_3 & 0 & 0 & 0 \\ 0 & p_{0|1} & 0 & p_{2|1} & 0 & p_{3|1} & 0 & 0 \\ 0 & 0 & p_{0|2} & p_{1|2} & 0 & 0 & p_{3|2} & 0 \\ 0 & 0 & 0 & p_{0|12} & 0 & 0 & 0 & p_{3|12} \\ 0 & 0 & 0 & 0 & p_{0|3} & p_{1|3} & p_{2|3} & 0 \\ 0 & 0 & 0 & 0 & 0 & p_{0|13} & 0 & p_{2|13} \\ 0 & 0 & 0 & 0 & 0 & 0 & p_{0|23} & p_{1|23} \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}.$$

Chapter 3

Basic Theory and General Structural Properties

This chapter presents some distributional properties for MDPH distributions. Section 3.1 provides a number of closure properties of MDPH. In section 3.2, we investigate dependence structures of MDPH distributions. The relationship between MDPH distributions and MPH distributions is presented in Section 3.3.

Most closure properties are obtained via direct constructions of underlying Markov chains that generate a random vector with a MDPH distribution and the desired characteristics. Since discrete phase type random variables are monotone functionals of the underlying Markov chain, many dependence properties of a MDPH distribution can be derived from the dependence properties of discrete-time Markov chains. The properties we obtain here not only form a useful set of tools for applications of MDPH distributions (See Chapter 5), but they also provide a new way of understanding dependencies in MPH distributions. In fact, we show that positive dependence of multivariate continuous phase type distributions may be inherited from positive dependence in an underlying MDPH distribution.

3.1 Closure Properties

Theorem 5 (Marginal Distributions) *If* $\mathbf{S}$ *has an* m*-dimensional MDPH distribution with underlying Markov chain* $\mathbb{S} = \{S, \boldsymbol{\Sigma}, \mathfrak{S}, \sigma\}$ *and* $\mathbf{S_J} = (\mathrm{S}_{j_1}, \mathrm{S}_{j_2}, \ldots, \mathrm{S}_{j_k})$ *is a* k*-dimensional marginal of* $\mathbf{S}$*, then* $\mathbf{S_J}$ *has a* k*-dimensional MDPH distribution with underlying Markov chain* $\mathbb{S} = \{S, \boldsymbol{\Sigma}_{\mathbf{J}}, \mathfrak{S}, \sigma\}$*, where* $\boldsymbol{\Sigma}_{\mathbf{J}} = \{\Sigma_{j_1}, \Sigma_{j_2}, \ldots, \Sigma_{j_k}\}$.

Proof:

For $i = 1, \ldots, m$, $\mathrm{S}_i = \min\{n \mid \mathbb{S}^n \in \Sigma_i\}$, so, for $i = 1, \ldots, k$, $\mathrm{S}_{j_i} = \min\{n \mid \mathbb{S}^n \in \Sigma_{j_i}\}$. This implies that $\mathbf{S_J}$ has a k-dimensional MDPH distribution with underlying Markov chain $\mathbb{S} = \{S, \mathbf{\Sigma_J}, \mathfrak{S}, \sigma\}$ as claimed. □

Proposition 1 (Finite Mixtures) *If* $\mathbf{S}^*$ *and* $\mathbf{S}'$ *have m-dimensional MDPH distributions, then for* $0 \leq \alpha \leq 1$, $\mathbf{S} = \alpha\mathbf{S}' + (1-\alpha)\mathbf{S}^*$ *also has an m-dimensional MDPH distribution.*

Proof:

Suppose $\mathbf{S}^*$ and $\mathbf{S}'$ have underlying Markov chains $\mathbb{S}^* = \{S^*, \mathbf{\Sigma}^*, \mathfrak{S}^*, \sigma^*\}$ and $\mathbb{S}' = \{S', \mathbf{\Sigma}', \mathfrak{S}', \sigma'\}$ respectively. Let $\mathbf{S}$ have an underlying Markov chain $\mathbb{S} = \{S, \mathbf{\Sigma}, \mathfrak{S}, \sigma\}$ where

$$S = S' \cup S^*$$

$$\Sigma_i = \Sigma_i' \cup \Sigma_i^*$$

$$\mathfrak{S} = \begin{pmatrix} \mathfrak{S}' & 0 \\ 0 & \mathfrak{S}^* \end{pmatrix}$$

$$\sigma = [\ \alpha \cdot \sigma', \quad (1-\alpha) \cdot \sigma^*\ ].$$

Then,

$$\begin{aligned}
\mathbf{P}\{\mathrm{S}_1 > n_1, \ldots, \mathrm{S}_m > n_m\} &= \sigma \cdot \mathfrak{S}^{n_{o_1}} \cdot \mathbf{I}_{\Gamma_1} \cdot \prod_{i=2}^{m} \left(\mathfrak{S}^{n_{o_i} - n_{o_{i-1}}} \cdot \mathbf{I}_{\Gamma_i}\right) \cdot \underline{\mathbf{1}} \\
&= \alpha \cdot \sigma' \cdot \mathfrak{S}'^{n_{o_1}} \cdot \mathbf{I}_{\Gamma_1'} \cdot \prod_{i=2}^{m} \left(\mathfrak{S}'^{(n_{o_i} - n_{o_{i-1}})} \cdot \mathbf{I}_{\Gamma_i'}\right) \cdot \underline{\mathbf{1}} \\
&\quad + (1-\alpha) \cdot \sigma^* \cdot \mathfrak{S}^{*n_{o_1}} \cdot \mathbf{I}_{\Gamma_1^*} \cdot \prod_{i=2}^{m} \left(\mathfrak{S}^{*(n_{o_i} - n_{o_{i-1}})} \cdot \mathbf{I}_{\Gamma_i^*}\right) \cdot \underline{\mathbf{1}} \\
&= \alpha \cdot \mathbf{P}\{\mathbf{S}'_1 > n_1, \ldots, \mathbf{S}'_m > n_m\} + (1-\alpha) \cdot \mathbf{P}\{\mathbf{S}^*_1 > n_1, \ldots, \mathbf{S}^*_m > n_m\}
\end{aligned}$$

and $\mathbf{S}$ is a mixture of $\mathbf{S}^*$ and $\mathbf{S}'$. □

Proposition 2 (Order Statistics) *If random vectors* $\mathbf{S}$ *and* $\mathbf{O}$ *have m-dimensional MDPH distributions with underlying Markov chains* $\mathbb{S} = \{S, \mathbf{\Sigma}, \mathfrak{S}, \sigma\}$ *and* $\mathbb{O} = \{S, \mathbf{\Omega}, \mathfrak{S}, \sigma\}$ *respectively, where*

$$\mathbf{\Omega} = \{\Omega_1, \cdots, \Omega_m\},$$

$$\Omega_i = \bigcup_{k=i}^{m} O_k,$$

and

$$O_k = \{s_j | \sum_{i=1}^{m} \Psi_i(s_j) = k\},$$

then the order statistics of $\mathbf{S}$ *have the same distribution as* $\mathbf{O}$.

Proof:

It is easy to see that $\mathbf{\Omega} = \{\Omega_1, \cdots, \Omega_m\}$ is a set of stochastically closed classes that satisfy the conditions in the definition of MDPH distributions with $\Omega_0 = \Sigma_0$ and $\Omega_\Delta = \Sigma_\Delta$. Also note that Ω_i consists of the states s_j that are in at least i classes in $\mathbf{\Sigma}$. It is not difficult to see that $\mathrm{O}_i \le n_i$ if and only if $\min_i\{\mathrm{S}_1, \cdots, \mathrm{S}_m\} \le n_i$ where $\min_i$ is the i^{th} order statistic of $\mathbf{S}$. That is, $\mathrm{O}_1 < n_1$ if and only if there is an i such that $\mathrm{S}_i < n_1$, $\mathrm{O}_2 < n_2$ if and only if there is an i, j such that $\mathrm{S}_i < n_2$ and $\mathrm{S}_j < n_2$. So,

$$\mathbf{P}\{\mathrm{O}_1 \le n_1, \cdots, \mathrm{O}_m \le n_m\} = \mathbf{P}\{\min_1\{\mathrm{S}_1, \cdots, \mathrm{S}_m\} \le n_1, \cdots, \min_m\{\mathrm{S}_1, \cdots, \mathrm{S}_m\} \le n_m\}.$$

Thus, the vector of order statistics of $\mathbf{S}$ has an MDPH with the underlying Markov chain $\mathbb{O} = \{S, \mathbf{\Omega}, \mathfrak{S}, \sigma\}$.
□

Let $(\mathrm{S}_1, \ldots, \mathrm{S}_m)$ be a vector of discrete lifetimes, and $\mathrm{S}_{(1)} = \min\{\mathrm{S}_1, \ldots, \mathrm{S}_m\}$. Given that all the items have survived up to time k, that is, $\mathrm{S}_{(1)} \ge k$, $\mathrm{S}_i - k$, $1 \le i \le m$, are the (discrete) residual lifetimes.

Proposition 3 *If* $\mathbf{S} = (\mathrm{S}_1, \ldots, \mathrm{S}_m)$ *has an m-dimensional MDPH distribution with underlying Markov chain* $\mathbb{S} = \{S, \mathbf{\Sigma}, \mathfrak{S}, \sigma\}$, *then* $[(\mathrm{S}_1 - k, \ldots, \mathrm{S}_m - k) \mid \mathrm{S}_{(1)} \ge k]$ *has an MDPH distribution with the underlying Markov chain* $\mathbb{S} = \{S, \mathbf{\Sigma}, \mathfrak{S}, \sigma_k\}$ *where* $\sigma_k = \sigma\mathfrak{S}^k$. *That is, the vector of residual lifetimes has an underlying Markov chain that has the same state space, stochastically closed sets, transition matrix as those of* $\mathbf{S}$, *but the initial distribution* $\sigma_k = \sigma\mathfrak{S}^k$.

Proof:

It is easy to see that

$$[(\mathrm{S}_1, \ldots, \mathrm{S}_m) \mid \mathrm{S}_{(1)} \ge k] = (k, \ldots, k) + (\min\{n | \mathbb{S}^{k+n} \in \Sigma_1\}, \ldots, \min\{n | \mathbb{S}^{k+n} \in \Sigma_m\}).$$

From the Markovian property, $(\min\{n | \mathbb{S}^{k+n} \in \Sigma_1\}, \ldots, \min\{n | \mathbb{S}^{k+n} \in \Sigma_m\})$ has the same distribution as that of

$$(\min\{n | \mathbb{S}^{*n} \in \Sigma_1\}, \ldots, \min\{n | \mathbb{S}^{*n} \in \Sigma_m\}),$$

where $\mathbb{S} = \{S, \mathbf{\Sigma}, \mathfrak{S}, \sigma_k\}$ where $\sigma_k = \sigma\mathfrak{S}^k$. So the distribution of the joint residual lifetimes,

$$[(\mathrm{S}_1 - k, \ldots, \mathrm{S}_m - k) \mid \mathrm{S}_{(1)} \ge k],$$

has an MDPH with underlying Markov chain $\mathbb{S} = \{S, \mathbf{\Sigma}, \mathfrak{S}, \sigma_k\}$. □

To study the concatenation of phase type distributed random vectors, we need the following concepts.

Definition 8 (Direct Product) *Let $\mathbb{S}^* = \{S^*, \mathbf{\Sigma}^*, \mathfrak{S}^*, \sigma^*\}$ and $\mathbb{S}' = \{S', \mathbf{\Sigma}', \mathfrak{S}', \sigma'\}$ be two Markov chains. $\mathbb{S} = \{S, \mathbf{\Sigma}, \mathfrak{S}, \sigma\}$ is said to be the direct product of $\mathbb{S}^*$ and $\mathbb{S}'$ (written $\mathbb{S} = \mathbb{S}^* \otimes \mathbb{S}'$) if and only if*

$$S = S^* \otimes S'$$

$$\mathbf{\Sigma} = \{\Sigma_i^* \otimes S' \mid \Sigma_i^* \in \mathbf{\Sigma}^*\} \cup \{S^* \otimes \Sigma_j' \mid \Sigma_j' \in \mathbf{\Sigma}'\},$$

$$\mathfrak{S} = \mathfrak{S}^* \otimes \mathfrak{S}',$$

$$\text{and } \sigma = \sigma^* \otimes \sigma',$$

where $A \otimes B$ is understood to be the direct (or Kronecker) product for matrices when A and B are matrices and the direct (or Cartesian) product for sets when A and B are sets.

The transition probabilities of a direct product of two Markov chains can be explicitly calculated as follows:

$$\begin{aligned}\mathbf{P}\{\mathbb{S}^0 = (s_i^*, s_j')\} &= \mathbf{P}\{\mathbb{S}^{*0} = s_i^*\} \cdot \mathbf{P}\{\mathbb{S}'^0 = s_j'\} \\ &= \sigma_i^* \cdot \sigma_j'\end{aligned}$$

and

$$\begin{aligned}\mathbf{P}\{\mathbb{S}^{n+1} = (s_{i_2}^*, s_{j_2}') \mid \mathbb{S}^n = (s_{i_1}^*, s_{j_1}')\} &= \mathbf{P}\{\mathbb{S}^{*n+1} = s_{i_2}^* \mid \mathbb{S}^{*n} = s_{i_1}^*\} \cdot \mathbf{P}\{\mathbb{S}'^{n+1} = s_{j_2}' \mid \mathbb{S}'^n = s_{j_1}'\} \\ &= \mathfrak{s}_{i_1 i_2}^* \cdot \mathfrak{s}_{j_1 j_2}'.\end{aligned}$$

Lemma 1 *If $\mathbb{S}^*$ and $\mathbb{S}'$ are independent Markov chains underlying k and m dimensional MDPH distributions $\mathbf{S}^*$ and $\mathbf{S}'$ respectively, then $\mathbb{S} = \mathbb{S}^* \otimes \mathbb{S}'$ is an underlying Markov chain for a $k+m$ dimensional MDPH distribution $\mathbf{S} = (\mathbf{S}^*, \mathbf{S}')$.*

Proof:

Clearly, $\Sigma_0 = \Sigma_0^* \otimes \Sigma_0'$ and $\Sigma_\Delta = \Sigma_\Delta^* \otimes \Sigma_\Delta'$.

$$\begin{aligned}&\mathbf{P}\{\mathbb{S}^n \in \Sigma_\Delta\} = \mathbf{P}\{\mathbb{S}^n = (s_\Delta^*, s_\Delta')\} \\ &= \sum_i \sum_j \mathbf{P}\{\mathbb{S}^n = (s_\Delta^*, s_\Delta') \mid \mathbb{S}^{n-1} = (s_i^*, s_j')\}\mathbf{P}\{\mathbb{S}^{n-1} = (s_i^*, s_j')\} \\ &= \sum_i \sum_j \mathbf{P}\{\mathbb{S}^{*n} = s_\Delta^* \mid \mathbb{S}^{*n-1} = s_i^*\}\mathbf{P}\{\mathbb{S}^{*n-1} = s_i^*\} \cdot \mathbf{P}\{\mathbb{S}'^n = s_\Delta' \mid \mathbb{S}'^{n-1} = s_j'\}\mathbf{P}\{\mathbb{S}'^{n-1} = s_j'\} \\ &= \sum_i \mathbf{P}\{\mathbb{S}^{*n} = s_\Delta^* \mid \mathbb{S}^{*n-1} = s_i^*\}\mathbf{P}\{\mathbb{S}^{*n-1} = s_i^*\} \sum_j \mathbf{P}\{\mathbb{S}'^n = s_\Delta' \mid \mathbb{S}'^{n-1} = s_j'\}\mathbf{P}\{\mathbb{S}'^{n-1} = s_j'\} \\ &= \mathbf{P}\{\mathbb{S}^{*n} = s_\Delta^*\} \cdot \mathbf{P}\{\mathbb{S}'^n = s_\Delta'\}.\end{aligned}$$

Since $\mathbb{S}^*$ and $\mathbb{S}'$ are underlying Markov chains, taking the limit we have:

$$\lim_{n\to\infty} \mathbf{P}\{\mathbb{S}^n \in \Sigma_\Delta\} = \lim_{n\to\infty} \mathbf{P}\{\mathbb{S}^{*n} = s^*_\Delta\} \cdot \mathbf{P}\{\mathbb{S}'^n = s'_\Delta\}$$
$$= 1.$$

Furthermore, $\forall i, j$, $\Sigma^*_i \subsetneq S^*$ and $\Sigma'_j \subsetneq S'$. This implies

$$\{\Sigma^*_i \otimes S' \mid \Sigma^*_i \in \mathbf{\Sigma}^*\} \cap \{S^* \otimes \Sigma'_j \mid \Sigma'_j \in \mathbf{\Sigma}'\} = \emptyset,$$

so,

$$|\mathbf{\Sigma}| = |\mathbf{\Sigma}^*| + |\mathbf{\Sigma}'| = k + m.$$

Since there are $k+m$ stochastically closed classes, $\mathbb{S}$ is an underlying Markov chain for the $k+m$ dimensional MDPH distribution of $\mathbf{S}$. □

Lemma 2 (Concatenation of Simple Distributions) *If $\mathbb{S}^*$ and $\mathbb{S}'$ are simple, then so is $\mathbb{S}^* \otimes \mathbb{S}'$.*

Proof:

$\mathbb{S}^*$ and $\mathbb{S}'$ being simple implies that $\forall i, j$, $|S^*_i| = |S'_j| = 1$ so $|S^*_i \otimes S'_j| = 1$. Therefore, $\mathbb{S}^* \otimes \mathbb{S}'$ is simple. □

3.2 Dependence Properties

3.2.1 Independence Results

Proposition 4 *Suppose $\mathbf{S} = (\mathbf{S}^*, \mathbf{S}')$, where $\mathbf{S}^*$ and $\mathbf{S}'$ have MDPH distributions with underlying Markov chains $\mathbb{S}^* = \{S^*, \mathbf{\Sigma}^*, \mathfrak{S}^*, \sigma^*\}$ and $\mathbb{S}' = \{S', \mathbf{\Sigma}', \mathfrak{S}', \sigma'\}$, respectively. If $\mathbb{S}$ is an underlying Markov chain of $\mathbf{S}$, with $\mathbb{S} = \mathbb{S}^* \otimes \mathbb{S}'$, then $\mathbf{S}^*$ and $\mathbf{S}'$ are independent.*

Proof:

Let $\mathbf{s_k} = (s_{k_0}, s_{k_1}, s_{k_2}, \dots)$ be a realization of $\mathbb{S}$. $\mathbb{S} = \mathbb{S}^* \otimes \mathbb{S}'$ implies

$$\mathbf{s_k} = ((s^*_{i_0}, s'_{j_0}), (s^*_{i_1}, s'_{j_1}), \dots)$$
$$= (\mathbf{s_i}^*, \mathbf{s_j}'),$$

where $\mathbf{s_i}^* = (s^*_{i_0}, s^*_{i_1}, s^*_{i_2}, \ldots)$, and $\mathbf{s_j}' = (s'_{j_0}, s'_{j_1}, s'_{j_2}, \ldots)$ are sample paths of $\mathbb{S}^*$ and $\mathbb{S}'$ respectively. So,

$$\begin{aligned}\mathbf{P}\{\mathbf{s_k}\} &= \mathbf{P}\{(\mathbf{s_i}^*, \mathbf{s_j}')\} \\ &= \sigma^*_{i_0}\sigma'_{j_0}\prod_{m=0}^{\infty}\mathfrak{s}^*_{i_m i_{m+1}}\mathfrak{s}'_{j_m j_{m+1}} \\ &= (\sigma^*_{i_0}\prod_{m=0}^{\infty}\mathfrak{s}^*_{i_m i_{m+1}})(\sigma'_{j_0}\prod_{m=0}^{\infty}\mathfrak{s}'_{j_m j_{m+1}}) \\ &= \mathbf{P}\{\mathbf{s_i}^*\}\mathbf{P}\{\mathbf{s_j}'\}.\end{aligned}$$

Taking $\mathbf{I}_{\mathbf{n}^*} = \{\mathbf{i} \mid \mathbf{s}^*_{\mathbf{i}} \text{ satisfies } \mathbf{S}^* = \mathbf{n}^*\}$ and $\mathbf{J}_{\mathbf{n}'} = \{\mathbf{j} \mid \mathbf{s}'_{\mathbf{j}} \text{ satisfies } \mathbf{S}' = \mathbf{n}'\}$, we have,

$$\begin{aligned}\mathbf{P}\{\mathbf{S}^* = \mathbf{n}^*, \mathbf{S}' = \mathbf{n}'\} &= \sum_{\mathbf{i}\in\mathbf{I}_{\mathbf{n}^*}}\sum_{\mathbf{j}\in\mathbf{J}_{\mathbf{n}'}}\mathbf{P}\{(\mathbf{s}^*_{\mathbf{i}}, \mathbf{s}'_{\mathbf{j}})\} \\ &= \sum_{\mathbf{i}\in\mathbf{I}_{\mathbf{n}^*}}\sum_{\mathbf{j}\in\mathbf{J}_{\mathbf{n}'}}\mathbf{P}\{\mathbf{s}^*_{\mathbf{i}}\}\mathbf{P}\{\mathbf{s}'_{\mathbf{j}}\} \\ &= \sum_{\mathbf{i}\in\mathbf{I}_{\mathbf{n}^*}}\mathbf{P}\{\mathbf{s}^*_{\mathbf{i}}\}\sum_{\mathbf{j}\in\mathbf{J}_{\mathbf{n}'}}\mathbf{P}\{\mathbf{s}'_{\mathbf{j}}\} \\ &= \mathbf{P}\{\mathbf{S}^* = \mathbf{n}^*\}\mathbf{P}\{\mathbf{S}' = \mathbf{n}'\}.\end{aligned}$$

□

Note that the underlying Markov chain may not be unique and the converse of the above result is not generally true, as is shown in the following example.

Example 8 Consider $\hat{\mathfrak{S}}$ from Example 2. $\mathbf{S}$ is the vector of two independent random variables whose marginal distributions have transition matrices

$$\mathfrak{S}_1 = \begin{pmatrix} \frac{3}{10} & \frac{3}{10} & \frac{4}{10} \\ \frac{1}{10} & \frac{6}{10} & \frac{3}{10} \\ 0 & 0 & 1 \end{pmatrix}.$$

and

$$\mathfrak{S}_2 = \begin{pmatrix} \frac{3}{5} & \frac{2}{5} \\ 0 & 1 \end{pmatrix}$$

It is easy to see that $\mathfrak{S} = \mathfrak{S}_1 \otimes \mathfrak{S}_2$, but, $\hat{\mathfrak{S}} \neq \mathfrak{S}_1 \otimes \mathfrak{S}_2$.

3.2.2 Dependence Structures

Dependence structures of multivariate phase-type distributions can be examined using stochastic comparison methods and the numerical methods to be discussed in Chapter 5. Numerical experiments with MPH distributions may provide new insight into their dependence structures beyond the properties that can be established through various comparison techniques.

Many different notions of dependence have been introduced and studied extensively in the literature (see, for example, [20] and [8]). Here we only discuss notions of dependence that are most relevant to this research. These include dependence orders *between* two random vectors and dependence *among* the components of a random vector.

To express the nature of dependence among the components of a random vector, one can of course use the covariance matrix. $\mathbf{X}$ is said to be positively (negatively) pairwise-correlated if the covariance matrix $\mathrm{Cov}(X_i, X_j) \geq (\leq)\ 0$, for all $i \neq j$. However, the following notions of dependence are stronger and frequently used in the literature.

Definition 9 Let $\mathbf{X} = (X_1, \ldots, X_m)$ and $\mathbf{Y} = (Y_1, \ldots, Y_m)$ be two $\mathcal{R}^m$-valued random vectors.

1. $\mathbf{X}$ is said to be larger (smaller) than $\mathbf{Y}$ in the upper (lower) orthant order, denoted by $\mathbf{X} \geq_{uo} (\leq_{lo})\ \mathbf{Y}$, if $P(\mathbf{X} > (\leq)\ \mathbf{x}) \geq P(\mathbf{Y} > (\leq)\ \mathbf{x})$, for all $\mathbf{x} \in \mathcal{R}^m$.

2. $\mathbf{X}$ is said to be more upper (lower) orthant dependent than $\mathbf{Y}$ if $\mathbf{X} \geq_{uo} (\leq_{lo})\ \mathbf{Y}$, and $X_i =_{st} Y_i$ for all i. ('$=_{st}$' denotes the equality in distribution).

Note that the orthant orders, coupled with identical marginals, emphasize the comparisons of dependence strengths of the two vectors by separating the marginals from consideration.

Definition 10 Let $\mathbf{X} = (X_1, \ldots, X_m)$ be an $\mathcal{R}^m$-valued random vector. Let $\mathbf{X}^I = (X_1^I, \ldots, X_m^I)$ denote a vector of real random variables such that $X_j =_{st} X_j^I$ for each j and $X_1^I, \ldots, X_m^I$ are independent.

1. $\mathbf{X}$ is said to be positively upper (lower) orthant dependent (PUOD, PLOD) if $\mathbf{X} \geq_{uo}\ (\leq_{lo})\ \mathbf{X}^I$. $\mathbf{X}$ is said to be negatively upper (lower) orthant dependent (NUOD, NLOD) if $\mathbf{X} \leq_{uo}\ (\geq_{lo})\ \mathbf{X}^I$.

2. $\mathbf{X}$ is said to be associated if $\mathrm{Cov}(f(\mathbf{X}), g(\mathbf{X})) \geq 0$ whenever f and g are non-decreasing. $\mathbf{X}$ is said to be negatively associated if for every subset $K \subseteq \{1, \ldots, m\}$, $\mathrm{Cov}(f(X_i, i \in K), g(X_j, j \in K^c)) \leq 0$ whenever f and g are non-decreasing.

It is known ([4], [19]) that

$$\mathbf{X} \text{ is (positively) associated} \implies \mathbf{X} \text{ is PUOD and PLOD}, \tag{3.1}$$

$$\mathbf{X} \text{ is negatively associated} \implies \mathbf{X} \text{ is NUOD and NLOD}. \tag{3.2}$$

The PLOD (PUOD, NLOD, NUOD)-property of a random vector means that its joint distribution or survival function can be bounded below or above by the products of its marginal distributions or survival functions. Various properties of these classifications have been discussed in [20] and [19].

We first discuss a fundamental property of positive association, from which many useful properties follow.

Theorem 6 *Let* $\mathbf{X}$ *be an* $\mathcal{R}^m$*-valued random variable and* $\mathbf{Y}$ *be an* $\mathcal{R}^s$*-valued random variable. If*

1. $\mathbf{X}$ *is associated,*
2. $(\mathbf{Y} \mid \mathbf{X} = \mathbf{x})$ *is associated for all* $\mathbf{x}$*, and*
3. $E[f(\mathbf{Y}) \mid \mathbf{X} = \mathbf{x}]$ *is increasing in* $\mathbf{x}$ *for all increasing function* f*,*

then $(\mathbf{X}, \mathbf{Y})$ *is associated.*

Proof:
Let E_Z denote the expectation with respect to the distribution of random variable Z. For any two increasing functions f and g, we have

$$E[f(\mathbf{X},\mathbf{Y})g(\mathbf{X},\mathbf{Y})] = E_{\mathbf{X}}[E_{\mathbf{Y}|\mathbf{X}}[f(\mathbf{X},\mathbf{Y})g(\mathbf{X},\mathbf{Y}) \mid \mathbf{X}]]$$

$$\geq E_{\mathbf{X}}[E_{\mathbf{Y}|\mathbf{X}}f(\mathbf{X},\mathbf{Y})E_{\mathbf{Y}|\mathbf{X}}g(\mathbf{X},\mathbf{Y}) \mid \mathbf{X}]$$

$$\geq E_{\mathbf{X}}[E_{\mathbf{Y}|\mathbf{X}}f(\mathbf{X},\mathbf{Y}) \mid \mathbf{X}]E[E_{\mathbf{Y}|\mathbf{X}}g(\mathbf{X},\mathbf{Y}) \mid \mathbf{X}] = Ef(\mathbf{X},\mathbf{Y})Eg(\mathbf{X},\mathbf{Y}),$$

where the first inequality follows from the association property of $\mathbf{Y}|\mathbf{X} = \mathbf{x}$, and the second inequality follows from the association property of $\mathbf{X}$ and the stochastic monotonicity property (3). □

The following properties of association can be easily verified from either Definition 10 or Theorem 6.

Theorem 7 1. *Any real random variable is associated.*

2. *If a* $\mathcal{R}^m$*-valued random vector* $\mathbf{X}$ *is associated and* $f : \mathcal{R}^m \to \mathcal{R}^s$ *is increasing (or decreasing), then* $f(\mathbf{X})$ *is associated.*
3. *Assume that a* $\mathcal{R}^m$*-valued random vector* $\mathbf{X}$ *is associated and a* $\mathcal{R}^s$*-valued random vector* $\mathbf{Y}$ *is associated. If* $\mathbf{X}$ *and* $\mathbf{Y}$ *are independent, then* $(\mathbf{X}, \mathbf{Y})$ *is associated.*
4. *Suppose that random vector* $\mathbf{X}^{(n)}$ *is associated for any* $n \geq 1$*. If* $\mathbf{X}^{(n)}$ *converges weakly to* $\mathbf{X}$*, then* $\mathbf{X}$ *is also associated.*

The following properties of orthant dependence can be easily verified from Definition 10.

Theorem 8 *Let* $f_i : \mathcal{R} \to \mathcal{R}$ *be a real valued function,* $1 \le i \le m$.

1. *If* $(X_1, \ldots, X_m)$ *is PUOD (PLOD) then* $(f_1(X_1), \ldots, f_m(X_m))$ *is also PUOD (PLOD) for all increasing functions* f_i, $1 \le i \le m$.

2. *If* $(X_1, \ldots, X_m)$ *is PUOD (PLOD) then* $(f_1(X_1), \ldots, f_m(X_m))$ *is PLOD (PUOD) for all decreasing functions* f_i, $1 \le i \le m$.

3. *If* $(X_1, \ldots, X_m)$ *is NUOD (NLOD) then* $(f_1(X_1), \ldots, f_m(X_m))$ *is also NUOD (NLOD) for all increasing functions* f_i, $1 \le i \le m$.

4. *If* $(X_1, \ldots, X_m)$ *is NUOD (NLOD) then* $(f_1(X_1), \ldots, f_m(X_m))$ *is NLOD (NUOD) for all decreasing functions* f_i, $1 \le i \le m$.

5. *Suppose that random vector* $X^{(n)}$ *is PUOD (PLOD, NUOD, NLOD) for any* $n \ge 1$. *If* $X^{(n)}$ *converges weakly to* X, *then* X *is also PUOD (PLOD, NUOD, NLOD).*

We can use Theorems 6 and 7 to obtain a sufficient condition for the association of MDPH.

Proposition 5 *Suppose that* $(\mathrm{S}_1, \ldots, \mathrm{S}_m)$ *has an MDPH distribution with an underlying Markov chain* $\{\mathbb{S}^n, n \ge 0\}$ *with state space* $S \subseteq \mathcal{R}^h$ *such that*

1. $\mathbb{S}^0$ *is associated,*

2. $(\mathbb{S}^n \mid \mathbb{S}^{n-1} = s)$ *is associated, and*

3. $Ef(\mathbb{S}^n \mid \mathbb{S}^{n-1} = s)$ *is increasing in* s *for any increasing function* f.

Then $(\mathrm{S}_1, \ldots, \mathrm{S}_m)$ *is associated.*

Proof:

We first show that $(\mathbb{S}^0, \ldots, \mathbb{S}^k)$ is associated for any $k \ge 0$ via induction. The claim is true for $k = 0$ because of (1). Suppose that $(\mathbb{S}^0, \ldots, \mathbb{S}^{k-1})$ is associated. It follows from the Markovian property that

$$(\mathbb{S}^k \mid \mathbb{S}^{k-1} = s_{k-1}, \ldots, \mathbb{S}^0 = s_0), \text{ and } (\mathbb{S}^k \mid \mathbb{S}^{k-1} = s_{k-1})$$

have the same distribution. Thus, by (2), $(\mathbb{S}^k \mid \mathbb{S}^{k-1} = s_{k-1}, \ldots, \mathbb{S}^0 = s_0)$ is associated, and by (3),

$$Ef(\mathbb{S}^k \mid \mathbb{S}^{k-1} = s_{k-1}, \ldots, \mathbb{S}^0 = s_0) = Ef(\mathbb{S}^k \mid \mathbb{S}^{k-1} = s_{k-1})$$

is increasing for any increasing function f. From Theorem 6, we have that $(\mathbb{S}^0, \ldots, \mathbb{S}^k)$ is associated. Consider

$$\mathrm{S}_j^k = \min\{\mathrm{S}_j, k\} = \min\{n \leq k : \mathbb{S}^n \in \Sigma_j\},\ 1 \leq j \leq m.$$

Obviously, S_j^k, $1 \leq j \leq m$, are non-increasing functions of $(\mathbb{S}^0, \ldots, \mathbb{S}^k)$. It follows from Theorem 7 (2) that $(\mathrm{S}_1^k, \ldots, \mathrm{S}_m^k)$ is associated. As $k \to \infty$, Theorem 7 (4) implies that $(\mathrm{S}_1, \ldots, \mathrm{S}_m)$ is associated. □

The condition (3) in Proposition 5 is known as the stochastic monotonicity of a Markov chain, and has been widely studied in the literature (see, for example, [12], [13]).

3.3 Relationship with Multivariate Phase Type Distributions

In this section, we consider the class of multivariate phase-type distributions studied by Assaf, et al. ([2]). Following that paper, consider a continuous-time, right-continuous Markov chain, $X = \{X(t), t \geq 0\}$ on a finite state space S with generator Q. Let Σ_i, $i = 1, \ldots, m$, be m nonempty stochastically closed subsets of S such that $\cap_{i=1}^m \Sigma_i$ is a proper subset of S. It is assumed that absorption into $\cap_{i=1}^m \Sigma_i$ is certain. Since we are interested in the process only until it is absorbed into $\cap_{i=1}^m \Sigma_i$, we may assume, without loss of generality, that $\cap_{i=1}^m \Sigma_i$ consists of one state, which we shall denote by S_Δ. Thus, without loss of generality, we may write $S = \Sigma_0 \cup (\cup_{i=1}^m \Sigma_i)$ for some $\Sigma_0 \subset S$. Let β be an initial probability vector on S such that $\beta(S_\Delta) = 0$. Then we can write $\beta = (\alpha, 0)$. Define

$$T_i = \inf\{t \geq 0 : X(t) \in \Sigma_i\}, \quad i = 1, \ldots, m. \tag{3.3}$$

As in [2], for simplicity, we shall assume that $P(T_1 > 0, T_2 > 0, \ldots, T_m > 0) = 1$. The joint distribution of $(T_1, T_2, \ldots, T_m)$ is called a multivariate phase-type distribution (MPH) with representation $(\alpha, Q, \Sigma_i, i = 1, \ldots, m)$, and $(T_1, T_2, \ldots, T_m)$ is called a phase-type random vector. The class of MPH distributions includes the well-known Marshall-Olkin distribution (Marshall and Olkin 1967), which is one of the most widely discussed multivariate life distributions in reliability theory (see, for example, [3]).

The MPH distributions defined by (3.3), their properties, and some related applications in reliability theory are discussed in [2]. As in the univariate case, those MPH distributions, their densities, Laplace transforms, and moments can be written in a closed form. The set of m-dimensional MPH distributions is dense in the set of all distributions on $[0, \infty)^m$.

The relation between MPH distributions and MDPH distributions is given below.

Theorem 9 *Let $(T_1, \ldots, T_m)$ be a random vector from a MPH distribution as defined in (3.3). Then we* *have,*

$$T_j = \sum_{i=1}^{\mathrm{S}_j} E_i,\ j = 1, 2, \ldots, m, \tag{3.4}$$

where $(\mathrm{S}_1, \ldots, \mathrm{S}_m)$ *is a random vector from an MDPH distribution and* $E_i, i = 1, 2, \ldots,$ *have independent and identically distributed exponential distributions.*

Proof:
Let $\{X(t), t \geq 0\}$ be the underlying Markov chain for $\mathbf{T} = (T_1, \ldots, T_m)$ with representation $(\alpha, Q, \Sigma_i, i = 1, \ldots, m)$. Since the state space is finite, we can use uniformization (see, for example, [18]) to rewrite $X(t)$ as

$$X(t) = X_{N(t)},\ t \geq 0,$$

where $\{X_n, n \geq 0\}$ is a discrete-time (embedded) Markov chain and $\{N(t), t \geq 0\}$ is a Poisson process with rate λ which is independent of $\{X_n, n \geq 0\}$. Define

$$\mathrm{S}_j = \min\{n | X_n \in \Sigma_j\},\ j = 1, \ldots, m.$$

So S_j is the number of transitions of Markov chain $\{X(t), t \geq 0\}$ prior to absorption to Σ_j. Obviously, $\mathbf{S} = (\mathrm{S}_1, \ldots, \mathrm{S}_m)$ has a distribution that is MDPH. Suppose $E_i, i = 1, 2, \ldots,$ have independent and identically distributed exponential distributions with parameter λ. Here each E_i denotes a sojourn time between two transitions of $\{X(t), t \geq 0\}$. Then we have $T_j = \sum_{i=1}^{\mathrm{S}_j} E_i,\ j = 1, 2, \ldots, m.$ □

Hereafter, $\mathbf{S}$ is said to be the underlying MDPH for $\mathbf{T}$ associated by parameter λ.

Corollary 2 *The transition function* $P(t)$ *of* $\{X(t), t \geq 0\}$ *is given by:*

$$P(t) = e^{-\lambda t} e^{\lambda t \mathfrak{S}} = \sum_{k=0}^{\infty} \frac{e^{-\lambda t}(\lambda t)^k}{k!} \mathfrak{S}^k.$$

Furthermore if $\mathfrak{S}$ *is given, the generator* $Q = \lambda\mathfrak{S} - \lambda I$*, or, if* Q *is given,* $\mathfrak{S} = I + \frac{1}{\lambda}Q$ *for some constant* $\lambda > 0$*.*

From (3.4), we can now show that MPH distributions inherit some positive dependence properties from underlying MDPH distributions.

Theorem 10 *Let* $(T_1, \ldots, T_m)$ *be a random vector with the multivariate phase type distribution as described in (3.4).*

1. *If* $(\mathrm{S}_1, \ldots, \mathrm{S}_m)$ *is PUOD (PLOD), then* $(T_1, \ldots, T_m)$ *is also PUOD (PLOD).*

2. If $(\mathrm{S}_1, \ldots, \mathrm{S}_m)$ *is positively associated, then* $(T_1, \ldots, T_m)$ *is also positively associated.*

Proof:

From (3.4), we have

$$(T_1, \ldots, T_m) \mid \{\mathrm{S}_1 = n_1, \ldots, \mathrm{S}_m = n_m\} = (\sum_{i=1}^{n_1} E_i, \ldots, \sum_{i=1}^{n_m} E_i). \tag{3.5}$$

It follows from Theorem 7 (1), (2) and (3) that $(T_1, \ldots, T_m) \mid \{\mathrm{S}_1 = n_1, \ldots, \mathrm{S}_m = n_m\}$ is associated.

(1) We only prove the PUOD case, and the other case is similar. Since $(T_1, \ldots, T_m) \mid \{\mathrm{S}_1 = n_1, \ldots, \mathrm{S}_m = n_m\}$ is also PUOD, then we have

$$E[f_1(T_1) \ldots f_m(T_m) \mid \{\mathrm{S}_1 = n_1, \ldots, \mathrm{S}_m = n_m\}] \geq \prod_{i=1}^{m} E[f_i(T_i) \mid \{\mathrm{S}_1 = n_1, \ldots, \mathrm{S}_m = n_m\}].$$

for all non-negative, increasing functions $f_1, \ldots, f_m$. Unconditioning on $S_1, \ldots, S_m$ yields that

$$E[\prod_{i=1}^{m} f_i(T_i) \mid \mathrm{S}_1, \ldots, \mathrm{S}_m] \geq \prod_{i=1}^{m} E[f_i(T_i) \mid \mathrm{S}_1, \ldots, \mathrm{S}_m].$$

Obviously,

$$g_i(n_i) = E[f_i(T_i) \mid \{\mathrm{S}_1 = n_1, \ldots, \mathrm{S}_m = n_m\}] = Ef_i(\sum_{j=1}^{n_i} E_j)$$

is increasing in n_i, $i = 1, \ldots m$. Since $(\mathrm{S}_1, \ldots, \mathrm{S}_m)$ is PUOD, then it follows from Theorem 8 (1) that $(g_1(\mathrm{S}_1), \ldots, g_m(\mathrm{S}_m))$ is also PUOD. That is,

$$E\prod_{i=1}^{m} f_i(T_i) = E(E[\prod_{i=1}^{m} f_i(T_i) \mid \mathrm{S}_1, \ldots, \mathrm{S}_m])$$

$$\geq E\prod_{i=1}^{m} g_i(\mathrm{S}_i) \geq \prod_{i=1}^{m} Eg_i(\mathrm{S}_i) = \prod_{i=1}^{m} Ef_i(T_i).$$

Thus, $(T_1, \ldots, T_m)$ is PUOD.

(2) To show that $(T_1, \ldots, T_m)$ is associated, we need Theorem 6. Consider the following three facts.

1. $(\mathrm{S}_1, \ldots, \mathrm{S}_m)$ is associated.

2. $(T_1, \ldots, T_m) \mid \{\mathrm{S}_1 = n_1, \ldots, \mathrm{S}_m = n_m\}$ is associated.

3. Since $E_i, i = 1, 2, \ldots,$ defined in (3.5), are independent and identically distributed, we have that $E[f(T_1, \ldots, T_m) \mid \{S_1 = n_1, \ldots, S_m = n_m\}] = Ef(\sum_{i=1}^{n_1} E_i, \ldots, \sum_{i=1}^{n_m} E_i)$ is increasing in $(n_1, \ldots, n_m)$, for any increasing function f defined on $\mathcal{R}^m$.

From Theorem 6, we obtain that $(T_1, \ldots, T_m)$ is associated. $\square$

As illustrated in Theorem 10, the positive dependence properties of MPH distributions inherit from similar properties of an underlying MDPH distributions. Therefore the dependence analysis of MDPH distributions provides a tool for understanding the dependence structure of MPH distributions.

Chapter 4

Bivariate Discrete Phase-type Distributions

This chapter provides an analysis of bivariate MDPH distributions. In Section 4.1, an explicit expression for the joint survival function of bivariate DPH distributions is given. Section 4.2 focuses on the simple bivariate DPH, and also highlights a sufficient condition for positive and negative dependence in simple bivariate DPH distributions. In Sections 4.3 and 4.4, we discuss the bivariate discrete Marshall-Olkin and Freund distributions. We show that the discrete Marshall-Olkin has univariate marginals with geometric distributions and obtain sufficient conditions for a discrete Marshall-Olkin distribution to be PQD or NQD. This shows that, unlike the continuous case, the discrete Marshall-Olkin distribution is not always positively dependent. We also show that, unlike the continuous case, that the discrete Freund distribution cannot be positively dependent, and establish a sufficient condition for the discrete Freund distribution to be NQD.

In Section 4.5 we look at the relationship between bivariate MDPH distributions and bivariate MPH distributions with particular focus on the Marshall-Olkin and Freund distributions. We show that all continuous Marshall-Olkin and Freund distributions have underlying MDPH distribution that are Marshall-Olkin and Freund, respectively. Li in [10] showed that if the probability that destroys component i in a Freund distribution becomes *larger* when the other component fails first, the Freund distribution is *positively* dependent. We settle an open question by showing that if the probability that destroys component i in a Freund distribution becomes *smaller* when the other components fail first, the Freund distribution is *negatively* dependent. Note that, unlike for positive dependence, the analysis of negative dependence poses a considerable challenge because of lack of the tools such as stochastic monotonicity.

4.1 General Form of Bivariate Discrete Phase-type Distribution

If $\mathbf{S} = (\mathrm{S}_1, \mathrm{S}_2)$ has a bivariate DPH distribution, the underlying Markov chain $\mathbb{S} = \{S, \mathbf{\Sigma}, \mathfrak{S}, \sigma\}$ is of the form: state space $S = S_0 \bigcup S_1 \bigcup S_2 \bigcup S_3$, $|S_i| \geq 1$ and $|S_3| = 1$, $\Sigma_1 = S_1 \bigcup S_3$, $\Sigma_2 = S_2 \bigcup S_3$ and the transition matrix can be partitioned as follows,

$$\mathfrak{S} = \begin{pmatrix} \mathfrak{S}_{00} & \mathfrak{S}_{01} & \mathfrak{S}_{02} & \mathfrak{S}_{03} \\ 0 & \mathfrak{S}_{11} & 0 & \mathfrak{S}_{13} \\ 0 & 0 & \mathfrak{S}_{22} & \mathfrak{S}_{23} \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

We assume that $\mathbf{P}\{\mathrm{S}_1 > 0, \mathrm{S}_2 > 0\} = 1$, that is, the underlying Markov chain starts in S_0 almost surely. The marginal distributions of S_1 and S_2 have underlying Markov chains $\mathbb{S}_1$ and $\mathbb{S}_2$, respectively. $\mathbb{S}_1 = \{S_{(1)}, \Sigma_1, \mathfrak{S}_{(1)}, \sigma\}$ where $S_{(1)} = \{S_0, S_2, (S_1, S_3)\}$, with (S_1, S_3) collapsed to a single state, and $\Sigma_1 = (S_1, S_3)$ (See Section 2.2.), and

$$\mathfrak{S}_{(1)} = \begin{pmatrix} \mathfrak{S}_{00} & \mathfrak{S}_{02} & \mathfrak{S}_{01}\underline{\mathbf{1}} + \mathfrak{S}_{03} \\ 0 & \mathfrak{S}_{22} & \mathfrak{S}_{23} \\ 0 & 0 & 1 \end{pmatrix}$$

Similarly, $\mathbb{S}_2 = \{S_{(2)}, \Sigma_2, \mathfrak{S}_{(2)}, \sigma\}$ where $S_{(2)} = \{S_0, S_1, (S_2, S_3)\}$, $\Sigma_2 = \{(S_2, S_3)\}$ and

$$\mathfrak{S}_{(2)} = \begin{pmatrix} \mathfrak{S}_{00} & \mathfrak{S}_{01} & \mathfrak{S}_{02}\underline{\mathbf{1}} + \mathfrak{S}_{03} \\ 0 & \mathfrak{S}_{11} & \mathfrak{S}_{13} \\ 0 & 0 & 1 \end{pmatrix}.$$

The joint and marginal survival functions can be calculated explicitly. For any probability vector α and any subset $\bar{S} \subseteq S$, we denote $\alpha_{\bar{S}}$ the sub-vector of α by removing its i-th entry for all $i \notin \bar{S}$.

$$\mathbf{P}\{\mathrm{S}_1 > n_1\} = \sigma \mathfrak{S}_{(1)}{}^{n_1} \begin{bmatrix} \underline{\mathbf{1}} \\ \underline{\mathbf{1}} \\ 0 \end{bmatrix} = \sigma_{S_0} \mathfrak{S}_{00}^{n_1} \underline{\mathbf{1}} + \sigma_{S_0} \sum_{i=1}^{n_1} \mathfrak{S}_{00}^{n_1 - i} \mathfrak{S}_{02} \mathfrak{S}_{22}^{i-1} \underline{\mathbf{1}}$$

$$\mathbf{P}\{\mathrm{S}_2 > n_2\} = \sigma \mathfrak{S}_{(2)}{}^{n_2} \begin{bmatrix} \underline{\mathbf{1}} \\ \underline{\mathbf{1}} \\ 0 \end{bmatrix} = \sigma_{S_0} \mathfrak{S}_{00}^{n_2} \underline{\mathbf{1}} + \sigma_{S_0} \sum_{i=1}^{n_2} \mathfrak{S}_{00}^{n_2 - i} \mathfrak{S}_{01} \mathfrak{S}_{11}^{i-1} \underline{\mathbf{1}}$$

and, without loss of generality, taking $n_1 \le n_2$, we have,

$$\mathbf{P}\{\mathrm{S}_1 > n_1, \mathrm{S}_2 > n_2\} = \sigma \mathfrak{S}^{n_1} \mathbf{I}_{\Sigma_0} \mathfrak{S}^{n_2-n_1} \begin{bmatrix} \underline{\mathbf{1}} \\ \underline{\mathbf{1}} \\ \underline{\mathbf{0}} \\ 0 \end{bmatrix}$$

$$= \sigma_{S_0} \mathfrak{S}_{00}^{n_1} [\mathfrak{S}_{00}^{n_2-n_1} \underline{\mathbf{1}} + \sum_{i=1}^{n_2-n_1} \mathfrak{S}_{00}^{n_2-n_1-i} \mathfrak{S}_{01} \mathfrak{S}_{11}^{i-1} \underline{\mathbf{1}}].$$

4.2 Simple Bivariate Discrete Phase-type Distributions

Let $\mathbf{S} = (\mathrm{S}_1, \mathrm{S}_2)$ be a simple bivariate DPH random vector with underlying Markov chain $\mathbb{S} = \{S, \mathbf{\Sigma}, \mathfrak{S}, \mathbf{e_1}\}$ where $S = \{s_0, s_1, s_2, s_3\}$, $\Sigma_1 = \{s_1, s_3\}$, $\Sigma_2 = \{s_2, s_3\}$ and

$$\mathfrak{S} = \begin{pmatrix} \mathfrak{s}_{00} & \mathfrak{s}_{01} & \mathfrak{s}_{02} & \mathfrak{s}_{03} \\ 0 & \mathfrak{s}_{11} & 0 & \mathfrak{s}_{13} \\ 0 & 0 & \mathfrak{s}_{22} & \mathfrak{s}_{23} \\ 0 & 0 & 0 & 1 \end{pmatrix},$$

with $\mathfrak{s}_{ij} \ge 0$, and $\mathbf{e}_1$ a $\{1, 0\}$-vector with 1 in the first entry and 0's elsewhere. The marginals S_1 and S_2 have the following underlying Markov chains $\mathbb{S}_1$ and $\mathbb{S}_2$, respectively.

$\mathbb{S}_1 = \{S_{(1)}, \Sigma_1, \mathfrak{S}_{(1)}, \mathbf{e}_1\}$ where $S_{(1)} = \{s_0, s_2, (s_1, s_3)\}$ with (s_1, s_3) considered a single state. $\Sigma_1 = \{(s_1, s_3)\}$.

$$\mathfrak{S}_{(1)} = \begin{pmatrix} \mathfrak{s}_{00} & \mathfrak{s}_{02} & \mathfrak{s}_{01} + \mathfrak{s}_{03} \\ 0 & \mathfrak{s}_{22} & \mathfrak{s}_{23} \\ 0 & 0 & 1 \end{pmatrix}$$

Similarly, $\mathbb{S}_2 = \{S_{(2)}, \Sigma_2, \mathfrak{S}_{(2)}, \mathbf{e}_1\}$ where $S_{(2)} = \{s_0, s_1, (s_2, s_3)\}$, $\Sigma_2 = \{(s_2, s_3)\}$ and

$$\mathfrak{S}_{(2)} = \begin{pmatrix} \mathfrak{s}_{00} & \mathfrak{s}_{01} & \mathfrak{s}_{02} + \mathfrak{s}_{03} \\ 0 & \mathfrak{s}_{11} & \mathfrak{s}_{13} \\ 0 & 0 & 1 \end{pmatrix}.$$

So,

$$\mathbf{P}\{\mathrm{S}_1 > n_1\} = \mathbf{e}_1 {\mathfrak{S}_{(1)}}^{n_1} \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}$$
$$= \mathfrak{s}_{00}^{n_1} + \sum_{i=1}^{n_1} \mathfrak{s}_{00}^{n_1-i} \mathfrak{s}_{02} \mathfrak{s}_{22}^{i-1}$$
$$\mathbf{P}\{\mathrm{S}_2 > n_2\} = \mathbf{e}_1 {\mathfrak{S}_{(2)}}^{n_2} \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}$$
$$= \mathfrak{s}_{00}^{n_2} + \sum_{i=1}^{n_2} \mathfrak{s}_{00}^{n_2-i} \mathfrak{s}_{01} \mathfrak{s}_{11}^{i-1}$$

and, without loss of generality, taking $n_1 \leq n_2$,

$$\mathbf{P}\{\mathrm{S}_1 > n_1, \mathrm{S}_2 > n_2\} = \mathbf{e}_1 \mathfrak{S}^{n_1} \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \mathfrak{S}^{n_2-n_1} \begin{bmatrix} 1 \\ 1 \\ 0 \\ 0 \end{bmatrix}$$
$$= \mathfrak{s}_{00}^{n_1} \left(\mathfrak{s}_{00}^{n_2-n_1} + \sum_{i=1}^{n_2-n_1} \mathfrak{s}_{00}^{n_2-n_1-i} \mathfrak{s}_{01} \mathfrak{s}_{11}^{i-1} \right) \tag{4.1}$$

If $\mathfrak{s}_{00} \neq \mathfrak{s}_{22}$ and $\mathfrak{s}_{00} \neq \mathfrak{s}_{11}$, then these simplify to

$$\mathfrak{s}_{00}^{n_1} + \mathfrak{s}_{02} \frac{\mathfrak{s}_{00}^{n_1} - \mathfrak{s}_{22}^{n_1}}{\mathfrak{s}_{00} - \mathfrak{s}_{22}}, \mathfrak{s}_{00}^{n_2} + \mathfrak{s}_{01} \frac{\mathfrak{s}_{00}^{n_2} - \mathfrak{s}_{11}^{n_2}}{\mathfrak{s}_{00} - \mathfrak{s}_{11}}, \text{ and } \mathfrak{s}_{00}^{n_1} \left(\mathfrak{s}_{00}^{n_2-n_1} + \mathfrak{s}_{01} \frac{\mathfrak{s}_{00}^{n_2-n_1} - \mathfrak{s}_{11}^{n_2-n_1}}{\mathfrak{s}_{00} - \mathfrak{s}_{11}} \right),$$

respectively. Furthermore, the following results will typically assume without loss of generality that $\mathfrak{s}_{00} \neq \mathfrak{s}_{22}$ and $\mathfrak{s}_{00} \neq \mathfrak{s}_{11}$, since it suffices to take limits to get similar results in the case of equality.

The joint probability mass function for a simple bivariate DPH vector is given by

$$\mathbf{P}\{\mathrm{S}_1 = n_1, \mathrm{S}_2 = n_2\} = \begin{cases} \mathfrak{s}_{00}^{n_2-1} \mathfrak{s}_{02} \mathfrak{s}_{22}^{n_1-n_2-1} \mathfrak{s}_{23} & n_1 > n_2 \\ \mathfrak{s}_{00}^{n_2-1} \mathfrak{s}_{03} & n_1 = n_2 \\ \mathfrak{s}_{00}^{n_1-1} \mathfrak{s}_{01} \mathfrak{s}_{11}^{n_2-n_1-1} \mathfrak{s}_{13} & n_1 < n_2. \end{cases}$$

Although the algebra is cumbersome, it is straightforward to verify with a computer algebra system the following formula for the moment generating function:

$$\begin{aligned}
M(t_1,t_2) &= \sum_{n_1=1}^{\infty}\sum_{n_2=1}^{\infty} \mathbf{P}\{\mathrm{S}_1=n_1,\mathrm{S}_2=n_2\} e^{t_1n_1+t_2n_2} \\
&= \sum_{n_1=1}^{\infty}\sum_{n_2=n_1+1}^{\infty} \mathfrak{s}_{00}^{n_1-1}\mathfrak{s}_{01}\mathfrak{s}_{11}^{n_2-n_1-1}\mathfrak{s}_{13}e^{t_1n_1+t_2n_2} + \sum_{n_2=1}^{\infty}\sum_{n_1=n_2+1}^{\infty} \mathfrak{s}_{00}^{n_2-1}\mathfrak{s}_{02}\mathfrak{s}_{22}^{n_1-n_2-1}\mathfrak{s}_{23}e^{t_1n_1+t_2n_2} \\
&\quad + \sum_{n_2=1}^{\infty} \mathfrak{s}_{00}^{n_2-1}\mathfrak{s}_{03}e^{n_2(t_1+t_2)} \\
&= \frac{1}{e^{-t_1-t_2}-\mathfrak{s}_{00}}\left(\frac{\mathfrak{s}_{01}\mathfrak{s}_{13}}{e^{-t_2}-\mathfrak{s}_{11}} + \frac{\mathfrak{s}_{02}\mathfrak{s}_{23}}{e^{-t_1}-\mathfrak{s}_{22}} + \mathfrak{s}_{03}\right).
\end{aligned}$$

To discuss the properties of bivariate DPH, we need to avoid some trivial cases. One of them is the following.

Definition 11 (Degenerate) *If* $\mathfrak{s}_{00}+\mathfrak{s}_{01}=0$ *or* $\mathfrak{s}_{00}+\mathfrak{s}_{02}=0$*, then* $\mathbf{S}$ *is said to be degenerate.*

Observe that if $\mathfrak{s}_{00}+\mathfrak{s}_{01}=0$, then $\mathbf{P}\{\mathrm{S}_2=1\}=\mathfrak{s}_{02}+\mathfrak{s}_{03}=1$. That is, S_2 is degenerate at the constant 1. Similarly, if $\mathfrak{s}_{00}+\mathfrak{s}_{02}=0$, then S_1 is degenerate at the constant 1. In these cases, S_1 and S_2 are trivially independent. Further results will assume that $\mathbf{S}$ is not degenerate.

Theorem 11 $\mathbf{S}$ *has a bivariate geometric distribution* $\iff$ $\mathfrak{s}_{00}+\mathfrak{s}_{01}=\mathfrak{s}_{11}$ *and* $\mathfrak{s}_{00}+\mathfrak{s}_{02}=\mathfrak{s}_{22}$.

Proof:

($\Leftarrow$)

Assume $\mathfrak{s}_{00}+\mathfrak{s}_{01}=\mathfrak{s}_{11}$; so

$$\begin{aligned}
\mathbf{P}\{\mathrm{S}_2>1\} &= (\mathfrak{s}_{00}+\mathfrak{s}_{01}) \\
&= \mathfrak{s}_{11}.
\end{aligned} \tag{4.2}$$

Since

$$\begin{aligned}
\mathbf{P}\{\mathrm{S}_2>n_2\} &= \mathfrak{s}_{00}^{n_2} + \sum_{i=1}^{n_2}\mathfrak{s}_{00}^{n_2-i}\mathfrak{s}_{01}\mathfrak{s}_{11}^{i-1} \\
&= \mathfrak{s}_{00}(\mathfrak{s}_{00}^{n_2-1} + \sum_{i=1}^{n_2-1}\mathfrak{s}_{00}^{n_2-1-i}\mathfrak{s}_{01}\mathfrak{s}_{11}^{i-1}) + \mathfrak{s}_{01}\mathfrak{s}_{11}^{n_2-1} \\
&= \mathfrak{s}_{00}\mathbf{P}\{\mathrm{S}_2>n_2-1\} + \mathfrak{s}_{01}\mathfrak{s}_{11}^{n_2-1},
\end{aligned} \tag{4.3}$$

a simple induction argument based upon (4.2) and (4.3) can be used to show that

$$\begin{aligned}\mathbf{P}\{\mathrm{S}_2 > n_2\} &= \mathfrak{s}_{11}^{n_2} \\ &= \mathbf{e}_1 \cdot \begin{pmatrix} \mathfrak{s}_{11} & \mathfrak{s}_{13} \\ 0 & 1 \end{pmatrix}^{n_2} \cdot \begin{bmatrix} 1 \\ 0 \end{bmatrix}\end{aligned} \tag{4.4}$$

which implies that S_2 has a simple underlying Markov chain $\mathbb{S}_2 = \{\{s_1, s_3\}, \{s_3\}, \mathfrak{S}_{(2^*)}, \mathbf{e}_1\}$ where

$$\mathfrak{S}_{(2^*)} = \begin{pmatrix} \mathfrak{s}_{11} & \mathfrak{s}_{13} \\ 0 & 1 \end{pmatrix}.$$

Similarly, for $\mathfrak{s}_{00} + \mathfrak{s}_{02} = \mathfrak{s}_{22}$, S_1 has a simple underlying Markov chain $\mathbb{S}_1 = \{\{s_2, s_3\}, \{s_3\}, \mathfrak{S}_{(1^*)}, \mathbf{e}_1\}$ where

$$\mathfrak{S}_{(1^*)} = \begin{pmatrix} \mathfrak{s}_{22} & \mathfrak{s}_{23} \\ 0 & 1 \end{pmatrix}.$$

($\Rightarrow$) Suppose that in addition to the underlying Markov chain inherited from the joint distribution, S_2 has a simple underlying Markov chain with transition matrix

$$\mathfrak{S}_{(2^*)} = \begin{pmatrix} \mathfrak{s}'_{00} & \mathfrak{s}'_{01} \\ 0 & 1 \end{pmatrix}.$$

Then,

$$\begin{aligned}\mathbf{P}\{\mathrm{S}_2 > 1\} &= \mathfrak{s}'_{00} \\ &= \mathfrak{s}_{00} + \mathfrak{s}_{01}\end{aligned}$$

and

$$\begin{aligned}\mathbf{P}\{\mathrm{S}_2 > 2\} &= \mathfrak{s}'^2_{00} \\ &= \mathfrak{s}_{00}(\mathfrak{s}_{00} + \mathfrak{s}_{01}) + \mathfrak{s}_{01}\mathfrak{s}_{11} \\ &= \mathfrak{s}_{00}\mathfrak{s}'_{00} + \mathfrak{s}_{01}\mathfrak{s}_{11}\end{aligned}$$

which implies

$$\mathfrak{s}'_{00}(\mathfrak{s}'_{00} - \mathfrak{s}_{00}) = \mathfrak{s}_{01}\mathfrak{s}_{11}.$$

Because $\mathfrak{s}_{01} = \mathfrak{s}'_{00} - \mathfrak{s}_{00}$, we have if $\mathfrak{s}_{01} > 0$,

$$\begin{aligned}\mathfrak{s}'_{00}\mathfrak{s}_{01} &= \mathfrak{s}_{01}\mathfrak{s}_{11} \\ \mathfrak{s}'_{00} &= \mathfrak{s}_{11} \\ \mathfrak{s}_{00} + \mathfrak{s}_{01} &= \mathfrak{s}_{11}.\end{aligned}$$

By a parallel argument, it is easy to see that if $\mathfrak{s}_{02} > 0$, then $\mathfrak{s}_{00} + \mathfrak{s}_{02} = \mathfrak{s}_{22}$.

Finally consider the case that $\mathfrak{s}_{01} = 0$ ($\mathfrak{s}_{02} = 0$). Since the Markov chain starts in s_0 with probability 1, s_1 (s_2) is unreachable so we can change $\mathfrak{s}_{11}$ ($\mathfrak{s}_{22}$) so that an equivalent underlying Markov chain has the property that $\mathfrak{s}'_{00} = \mathfrak{s}_{11}$ ($\mathfrak{s}'_{00} = \mathfrak{s}_{22}$). □

Note that the condition in Theorem 11 is not surprising. Since $\mathfrak{s}_{00} + \mathfrak{s}_{01} = \mathfrak{s}_{11}$ and $\mathfrak{s}_{00} + \mathfrak{s}_{02} = \mathfrak{s}_{22}$ are equivalent to $\mathfrak{s}_{02} + \mathfrak{s}_{03} = \mathfrak{s}_{13}$ and $\mathfrak{s}_{01} + \mathfrak{s}_{03} = \mathfrak{s}_{23}$ respectively, it follows from Theorem 3 that the marginal distributions of S_1 and S_2 have simple underlying Markov chains.

Theorem 12 S_1 *and* S_2 *are independent* $\iff$ $\mathfrak{s}_{00} + \mathfrak{s}_{01} = \mathfrak{s}_{11}$, $\mathfrak{s}_{00} + \mathfrak{s}_{02} = \mathfrak{s}_{22}$ *and* $\mathfrak{s}_{11}\mathfrak{s}_{22} = \mathfrak{s}_{00}$.

Proof:
($\Rightarrow$) Suppose S_1 and S_2 are independent. First consider the case when $\mathfrak{s}_{01} = 0$. This implies that $\mathrm{S}_1 \geq \mathrm{S}_2$ almost surely so S_1 and S_2 are not independent. Therefore, $\mathfrak{s}_{01} > 0$ and by a similar argument, $\mathfrak{s}_{02} > 0$.

If S_1 and S_2 are independent, then

$$\mathbf{P}\{\mathrm{S}_1 > n_1, \mathrm{S}_2 > n_2\} = \mathbf{P}\{\mathrm{S}_1 > n_1\}\mathbf{P}\{\mathrm{S}_2 > n_2\}.$$

For $n_1 = 1$ and $n_2 = 1$;

$$\mathfrak{s}_{00} = (\mathfrak{s}_{00} + \mathfrak{s}_{02})(\mathfrak{s}_{00} + \mathfrak{s}_{01}) \tag{4.5}$$

and $n_1 = 1$ and $n_2 = 2$;

$$\mathfrak{s}_{00}(\mathfrak{s}_{00} + \mathfrak{s}_{01}) = (\mathfrak{s}_{00} + \mathfrak{s}_{02})(\mathfrak{s}_{00}(\mathfrak{s}_{00} + \mathfrak{s}_{01}) + \mathfrak{s}_{01}\mathfrak{s}_{11}). \tag{4.6}$$

Substituting (4.5) into the left side of (4.6) gives:

$$\begin{aligned}
(\mathfrak{s}_{00} + \mathfrak{s}_{02})(\mathfrak{s}_{00} + \mathfrak{s}_{01})^2 &= (\mathfrak{s}_{00} + \mathfrak{s}_{02})(\mathfrak{s}_{00}(\mathfrak{s}_{00} + \mathfrak{s}_{01}) + \mathfrak{s}_{01}\mathfrak{s}_{11}) \\
(\mathfrak{s}_{00} + \mathfrak{s}_{01})^2 - \mathfrak{s}_{00}(\mathfrak{s}_{00} + \mathfrak{s}_{01}) &= \mathfrak{s}_{01}\mathfrak{s}_{11} \\
(\mathfrak{s}_{00} + \mathfrak{s}_{01})(\mathfrak{s}_{00} + \mathfrak{s}_{01} - \mathfrak{s}_{00}) &= \mathfrak{s}_{01}\mathfrak{s}_{11} \\
\mathfrak{s}_{01}(\mathfrak{s}_{00} + \mathfrak{s}_{01}) &= \mathfrak{s}_{01}\mathfrak{s}_{11} \\
\mathfrak{s}_{11} &= \mathfrak{s}_{00} + \mathfrak{s}_{01}.
\end{aligned} \tag{4.7}$$

Similarly, for $n_1 = 2$ and $n_2 = 1$

$$\mathfrak{s}_{22} = \mathfrak{s}_{00} + \mathfrak{s}_{02}. \tag{4.8}$$

Also (4.5), (4.7) and (4.8) imply

$$\mathfrak{s}_{00} = \mathfrak{s}_{11}\mathfrak{s}_{22}.$$

($\Leftarrow$)

Suppose $\mathfrak{s}_{00}+\mathfrak{s}_{01}=\mathfrak{s}_{11}$, $\mathfrak{s}_{00}+\mathfrak{s}_{02}=\mathfrak{s}_{22}$ and $\mathfrak{s}_{00}=\mathfrak{s}_{11}\mathfrak{s}_{22}$. (4.4) implies that

$$\mathbf{P}\{S_2 > n_2\} = \mathfrak{s}_{11}^{n_2}$$

and

$$\mathbf{P}\{S_1 > n_1\} = \mathfrak{s}_{22}^{n_1},$$

so for any $n_1 \leq n_2$,

$$\begin{aligned}\mathbf{P}\{S_1 > n_1, S_2 > n_2\} &= \mathfrak{s}_{00}^{n_1}(\mathfrak{s}_{00}^{n_2-n_1} + \sum_{i=1}^{n_2-n_1} \mathfrak{s}_{00}^{n_2-n_1-i}\mathfrak{s}_{01}\mathfrak{s}_{11}^{i-1}) \\ &= \mathfrak{s}_{00}^{n_1}\mathbf{P}\{S_2 > n_2 - n_1\} \\ &= \mathfrak{s}_{00}^{n_1}\mathfrak{s}_{11}^{n_2-n_1} \\ &= (\mathfrak{s}_{22}\mathfrak{s}_{11})^{n_1}\mathfrak{s}_{11}^{n_2-n_1} \\ &= \mathfrak{s}_{22}^{n_1}\mathfrak{s}_{11}^{n_2} \\ &= \mathbf{P}\{S_1 > n_1\}\mathbf{P}\{S_2 > n_2\}.\end{aligned} \tag{4.9}$$

□

For example, let $\mathfrak{s}_{00} = \mathfrak{s}_{01} = \mathfrak{s}_{02} = \mathfrak{s}_{03} = 1/4$, and $\mathfrak{s}_{11} = \mathfrak{s}_{13} = \mathfrak{s}_{22} = \mathfrak{s}_{23} = 1/2$, then the conditions in Theorem 12 are satisfied, and the corresponding S_1 and S_2 are independent.

Lemma 3 *If* $\mathbf{S}$ *has a bivariate distribution which is PUOD (NUOD) then* $\mathbf{S}$ *is PQD (NQD).*

Proof:

If $\mathbf{S}$ is PUOD, then $\mathbf{P}\{S_1 > n_1, S_2 > n_2\} \geq \mathbf{P}\{S_1 > n_1\}\mathbf{P}\{S_2 > n_2\}$, and the following are all equivalent:

$$\mathbf{P}\{S_1 > n_1, S_2 > n_2\} \geq \mathbf{P}\{S_1 > n_1\}\mathbf{P}\{S_2 > n_2\}$$

$$\mathbf{P}\{S_1 \leq n_1, S_2 \leq n_2\} - \mathbf{P}\{S_1 \leq n_1\} - \mathbf{P}\{S_2 \leq n_2\} + 1 \geq (1 - \mathbf{P}\{S_1 \leq n_1\})(1 - \mathbf{P}\{S_2 \leq n_2\})$$

$$\mathbf{P}\{S_1 \leq n_1, S_2 \leq n_2\} \geq \mathbf{P}\{S_1 \leq n_1\})\mathbf{P}\{S_2 \leq n_2\}.$$

Therefore, $\mathbf{S}$ is PLOD and hence, PQD. For NUOD and NQD, reverse the inequality. $\square$

Corollary 3 $\mathfrak{s}_{00} + \mathfrak{s}_{01} = \mathfrak{s}_{11}$, $\mathfrak{s}_{00} + \mathfrak{s}_{02} = \mathfrak{s}_{22}$ *and* $\mathfrak{s}_{00} \geq (\leq)\mathfrak{s}_{11}\mathfrak{s}_{22}$ *are sufficient conditions for* $\mathbf{S}$ *to be PQD (NQD).*

Proof:

Following the $(\Leftarrow)$ portion of the proof of Theorem 12, (4.9) becomes, for any $n_1 \leq n_2$,

$$\begin{aligned}
\mathbf{P}\{\mathrm{S}_1 > n_1, \mathrm{S}_2 > n_2\} &= \mathfrak{s}_{00}^{n_1}\mathfrak{s}_{11}^{n_2-n_1} \\
&\geq (\leq)(\mathfrak{s}_{22}\mathfrak{s}_{11})^{n_1}\mathfrak{s}_{11}^{n_2-n_1} \\
&= \mathfrak{s}_{22}^{n_1}\mathfrak{s}_{11}^{n_2} \\
&= \mathbf{P}\{\mathrm{S}_1 > n_1\}\mathbf{P}\{\mathrm{S}_2 > n_2\}.
\end{aligned}$$

This implies $\mathbf{S}$ is PUOD (NUOD) and by Lemma 3, PQD (NQD). $\square$

4.3 Discrete Bivariate Marshall-Olkin Distributions

Recall that if $\mathbf{S} = (\mathrm{S}_1, \mathrm{S}_2)$ is a bivariate simple DPH with underlying Markov chain $\mathbb{S} = \{S, \mathbf{\Sigma}, \mathfrak{S}, \mathbf{e_1}\}$, then,

$$\mathfrak{S} = \begin{pmatrix} \mathfrak{s}_{00} & \mathfrak{s}_{01} & \mathfrak{s}_{02} & \mathfrak{s}_{03} \\ 0 & \mathfrak{s}_{11} & 0 & \mathfrak{s}_{13} \\ 0 & 0 & \mathfrak{s}_{22} & \mathfrak{s}_{23} \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

Lemma 4 (Discrete Bivariate Marshall-Olkin Distributions) $\mathbf{S}$ *has a discrete Marshall-Olkin distribution if and only if* $\mathbf{S}$ *has a bivariate geometric distribution.*

Proof:

If $\mathbf{S}$ is a discrete bivariate Marshall-Olkin distribution it follows from Example 3 that $\mathfrak{s}_{11} = \mathfrak{s}_{00} + \mathfrak{s}_{01}$ and $\mathfrak{s}_{22} = \mathfrak{s}_{00} + \mathfrak{s}_{02}$. By Theorem 11, $\mathbf{S}$ is a bivariate geometric distribution.

If $\mathbf{S}$ is a bivariate geometric distribution then, S_1 and S_2 are simple if and only if $\mathfrak{s}_{11} = \mathfrak{s}_{00} + \mathfrak{s}_{01}$ and $\mathfrak{s}_{22} = \mathfrak{s}_{00} + \mathfrak{s}_{02}$. These equations are equivalent to $\mathfrak{s}_{13} = \mathfrak{s}_{02} + \mathfrak{s}_{03}$ and $\mathfrak{s}_{23} = \mathfrak{s}_{01} + \mathfrak{s}_{03}$. So,

$$\mathfrak{S} = \begin{pmatrix} \mathfrak{s}_{00} & \mathfrak{s}_{01} & \mathfrak{s}_{02} & \mathfrak{s}_{03} \\ 0 & \mathfrak{s}_{00} + \mathfrak{s}_{01} & 0 & \mathfrak{s}_{02} + \mathfrak{s}_{03} \\ 0 & 0 & \mathfrak{s}_{00} + \mathfrak{s}_{02} & \mathfrak{s}_{01} + \mathfrak{s}_{03} \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

$\mathfrak{S}$ is identical in form to the one given in Example 3 for the discrete bivariate Marhsall-Olkin distribution so $\mathbf{S}$ must be a discrete bivariate Marshall-Olkin distribution. □

It follows from Lemma 4 that the marginal distributions of a discrete bivariate Marhsall-Olkin distribution are geometric, which is a discrete analog of the fact that the continuous Marshall-Olkin distribution has exponential marginals.

Using formula (4.1) for the survival probabilities of a simple bivariate DPH and the relationships $\mathfrak{s}_{11} = \mathfrak{s}_{00} + \mathfrak{s}_{01}$ and $\mathfrak{s}_{22} = \mathfrak{s}_{00} + \mathfrak{s}_{02}$ defining a bivariate discrete Marshall-Olkin distribution, we have:

$$\mathbf{P}\{\mathrm{S}_1 > n_1, \mathrm{S}_2 > n_2\} = \begin{cases} \mathfrak{s}_{00}^{n_1}\mathfrak{s}_{11}^{n_2-n_1} & n_2 \geq n_1 \\ \\ \mathfrak{s}_{00}^{n_2}\mathfrak{s}_{22}^{n_1-n_2} & n_1 \geq n_2. \end{cases}$$

Theorem 13 (Orthant Comparisons) *Let $\mathbf{S}^*$ and $\mathbf{S}'$ be bivariate discrete Marshall-Olkin distributions with transition matrices*

$$\mathfrak{S}' = \begin{pmatrix} \mathfrak{s}'_{00} & \mathfrak{s}'_{01} & \mathfrak{s}'_{02} & \mathfrak{s}'_{03} \\ 0 & \mathfrak{s}'_{11} & 0 & \mathfrak{s}'_{13} \\ 0 & 0 & \mathfrak{s}'_{22} & \mathfrak{s}'_{23} \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

and

$$\mathfrak{S}^* = \begin{pmatrix} \mathfrak{s}^*_{00} & \mathfrak{s}^*_{01} & \mathfrak{s}^*_{02} & \mathfrak{s}^*_{03} \\ 0 & \mathfrak{s}^*_{11} & 0 & \mathfrak{s}^*_{13} \\ 0 & 0 & \mathfrak{s}^*_{22} & \mathfrak{s}^*_{23} \\ 0 & 0 & 0 & 1 \end{pmatrix},$$

respectively. Then $\mathbf{P}\{\mathrm{S}'_1 > n_1, \mathrm{S}'_2 > n_2\} \geq \mathbf{P}\{\mathrm{S}^*_1 > n_1, \mathrm{S}^*_2 > n_2\}$ if and only if $\mathfrak{s}'_{00} \geq \mathfrak{s}^*_{00}$, $\mathfrak{s}'_{11} \geq \mathfrak{s}^*_{11}$, and $\mathfrak{s}'_{22} \geq \mathfrak{s}^*_{22}$.

Proof:

Suppose $\mathfrak{s}'_{00} \geq \mathfrak{s}^*_{00}$, $\mathfrak{s}'_{11} \geq \mathfrak{s}^*_{11}$, and $\mathfrak{s}'_{22} \geq \mathfrak{s}^*_{22}$. Then for every $n_2 \geq n_1 \geq 0$,

$$\left(\frac{\mathfrak{s}'_{00}}{\mathfrak{s}^*_{00}}\right)^{n_1} \geq 1 \geq \left(\frac{\mathfrak{s}^*_{11}}{\mathfrak{s}'_{11}}\right)^{n_2-n_1},$$

$$\mathfrak{s}'^{n_1}_{00}\mathfrak{s}'^{n_2-n_1}_{11} \geq \mathfrak{s}^{*n_1}_{00}\mathfrak{s}^{*n_2-n_1}_{11}$$

$$\mathbf{P}\{\mathrm{S}'_1 > n_1, \mathrm{S}'_2 > n_2\} \geq \mathbf{P}\{\mathrm{S}^*_1 > n_1, \mathrm{S}^*_2 > n_2\}.$$

A parallel argument works for $n_1 \geq n_2 \geq 0$.

If $\mathbf{P}\{\mathrm{S}_1' > n_1, \mathrm{S}_2' > n_2\} \geq \mathbf{P}\{\mathrm{S}_1^* > n_1, \mathrm{S}_2^* > n_2\}$, then for every $n_2 \geq n_1$,

$$\mathfrak{s}_{00}^{\prime n_1} \mathfrak{s}_{11}^{\prime n_2 - n_1} \geq \mathfrak{s}_{00}^{* n_1} \mathfrak{s}_{11}^{* n_2 - n_1}.$$

This implies

$$\left(\frac{\mathfrak{s}_{00}'}{\mathfrak{s}_{00}^*} \right)^{n_1} \geq \left(\frac{\mathfrak{s}_{11}^*}{\mathfrak{s}_{11}'} \right)^{n_2 - n_1}.$$

For $n_2 = n_1 = 1$,

$$\frac{\mathfrak{s}_{00}'}{\mathfrak{s}_{00}^*} \geq 1$$

and

$$\mathfrak{s}_{00}' \geq \mathfrak{s}_{00}^*.$$

Taking n_1 fixed and allowing n_2 to be arbitrarily large, we see that

$$\frac{\mathfrak{s}_{11}^*}{\mathfrak{s}_{11}'} \leq 1$$

to satisfy the inequality for all $n_2 > n_1$. This implies $\mathfrak{s}_{11}' \geq \mathfrak{s}_{11}^*$. A parallel argument for $n_1 \geq n_2$ shows that $\mathfrak{s}_{22}' \geq \mathfrak{s}_{22}^*$. □

Corollary 4 *If $\mathbf{S}^*$ and $\mathbf{S}'$ have identical marginals, then $\mathbf{P}\{\mathrm{S}_1' > n_1, \mathrm{S}_2' > n_2\} \geq \mathbf{P}\{\mathrm{S}_1^* > n_1, \mathrm{S}_2^* > n_2\}$ if and only if $\mathfrak{s}_{00}' \geq \mathfrak{s}_{00}^*$.*

Proof:

If $\mathbf{S}^*$ and $\mathbf{S}'$ have identical marginals, $\mathfrak{s}_{11}' = \mathfrak{s}_{11}^*$ and $\mathfrak{s}_{22}' = \mathfrak{s}_{22}^*$. Therefore $\mathfrak{s}_{00}' \geq \mathfrak{s}_{00}^*$ is all that is needed to satisfy the conditions of the theorem. □

4.4 Discrete Bivariate Freund Distributions

Let $\mathbf{S} = (\mathrm{S}_1, \mathrm{S}_2)$ be a bivariate simple DPH with underlying Markov chain $\mathbb{S} = \{S, \mathbf{\Sigma}, \mathfrak{S}, \mathbf{e_1}\}$. Where

$$\mathfrak{S} = \begin{pmatrix} \mathfrak{s}_{00} & \mathfrak{s}_{01} & \mathfrak{s}_{02} & 0 \\ 0 & \mathfrak{s}_{11} & 0 & \mathfrak{s}_{13} \\ 0 & 0 & \mathfrak{s}_{22} & \mathfrak{s}_{23} \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

Then it is clear from Example 5 that $\mathbf{S}$ has a bivariate discrete Freund distribution.

Lemma 5 S_1 *and* S_2 *cannot be PUOD except when the distribution is degenerate in which case PUOD holds trivially.*

Proof:

Suppose S_1 and S_2 are positively dependent. Then $\forall n_1, n_2$,

$$\mathbf{P}\{\mathrm{S}_1 \leq n_1, \mathrm{S}_2 \leq n_2\} \geq \mathbf{P}\{\mathrm{S}_1 \leq n_1\}\mathbf{P}\{\mathrm{S}_2 \leq n_2\}$$

and in particular,

$$\mathbf{P}\{\mathrm{S}_1 = 1, \mathrm{S}_2 = 1\} \geq \mathbf{P}\{\mathrm{S}_1 = 1\}\mathbf{P}\{\mathrm{S}_2 = 1\}$$

but $\mathbf{P}\{\mathrm{S}_1 = 1, \mathrm{S}_2 = 1\} = 0$, since there can be no simultaneous failure. The only way to satisfy this inequality is if $\mathbf{P}\{\mathrm{S}_1 = 1\} = 0$ or $\mathbf{P}\{\mathrm{S}_2 = 1\} = 0$, but not both. Without loss of generality, suppose $\mathbf{P}\{\mathrm{S}_2 = 1\} = 0$. So $\mathfrak{s}_{02} = 0$ and $\mathbf{P}\{\mathrm{S}_1 = n_1, \mathrm{S}_2 = n_2\} = 0$ if $n_1 > n_2$. Since $\mathfrak{s}_{00} = \mathfrak{s}_{02} = 0$ implies the distribution is degenerate and therefore trivially independent, assume that $\mathfrak{s}_{00} > 0$. Then

$$\begin{aligned} \mathbf{P}\{\mathrm{S}_1 = 3, \mathrm{S}_2 = 2\} &= 0 \\ &\ngeq (\mathfrak{s}_{00}^2\mathfrak{s}_{01})(\mathfrak{s}_{01}\mathfrak{s}_{13}) \qquad (4.10) \\ &= \mathbf{P}\{\mathrm{S}_1 = 3\}\mathbf{P}\{\mathrm{S}_2 = 2\} \end{aligned}$$

so S_1 and S_2 cannot be positively dependent. □

Corollary 5 S_1 *and* S_2 *cannot be independent.*

Proof:

The proof is the same as above with the inequalities replaced by equality. □

Sufficient conditions for NUOD are given in Corollary 3, while the following example shows that bivariate discrete Freund distributions need not be NUOD.

Example 9 *Suppose* $\mathbf{S} = (\mathrm{S}_1, \mathrm{S}_2)$ *is a bivariate discrete Freund distribution with underlying Markov chain* $\mathbb{S} = \{S, \mathbf{\Sigma}, \mathfrak{S}, \mathbf{e_1}\}$. *Where*

$$\mathfrak{S} = \begin{pmatrix} .98 & .01 & .01 & 0 \\ 0 & .01 & 0 & .99 \\ 0 & 0 & .01 & .99 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

Then

$$\begin{aligned}
\mathbf{P}\{\mathrm{S}_1 > 1, \mathrm{S}_2 > 2\} &= \mathfrak{s}_{00}(\mathfrak{s}_{00} + \mathfrak{s}_{01}) \\
&= .98(.98 + .01) \\
&= .9702 \\
&> .9606 \\
&= (.98 + .01)(.98^2 + .01(.98 + .01)) \\
&= (\mathfrak{s}_{00} + \mathfrak{s}_{01})(\mathfrak{s}_{00}^2 + \mathfrak{s}_{01}(\mathfrak{s}_{00} + \mathfrak{s}_{11})) \\
&= \mathbf{P}\{\mathrm{S}_1 > 1\}\mathbf{P}\{\mathrm{S}_2 > 2\}.
\end{aligned}$$

4.5 Bivariate MDPH and Bivariate MPH Distributions

4.5.1 Relationship of Bivariate Simple MDPH Distributions to MPH distributions

Following Section 3.3, let $\mathbf{S} = (\mathrm{S}_1, \mathrm{S}_2)$ be a simple bivariate DPH distribution with underlying Markov chain $\mathbb{S} = \{S, \mathbf{\Sigma}, \mathfrak{S}, \mathbf{e_1}\}$ where $S = \{s_0, s_1, s_2, s_3\}$, $\Sigma_1 = \{s_1, s_3\}$, $\Sigma_2 = \{s_2, s_3\}$ and

$$\mathfrak{S} = \begin{pmatrix} \mathfrak{s}_{00} & \mathfrak{s}_{01} & \mathfrak{s}_{02} & \mathfrak{s}_{03} \\ 0 & \mathfrak{s}_{11} & 0 & \mathfrak{s}_{13} \\ 0 & 0 & \mathfrak{s}_{22} & \mathfrak{s}_{23} \\ 0 & 0 & 0 & 1 \end{pmatrix},$$

with $\mathfrak{s}_{ij} \geq 0$, and $\mathbf{e_1}$ is a $\{1, 0\}$-vector with 1 in the first element and 0's elsewhere.

As before, we assume $\mathfrak{s}_{00} \neq \mathfrak{s}_{11}$ and $\mathfrak{s}_{00} \neq \mathfrak{s}_{22}$, so

$$\mathbf{P}\{\mathrm{S}_1 > n_1\} = \mathfrak{s}_{00}^{n_1} + \mathfrak{s}_{02}\frac{\mathfrak{s}_{22}^{n_1} - \mathfrak{s}_{00}^{n_1}}{\mathfrak{s}_{22} - \mathfrak{s}_{00}},$$
$$\mathbf{P}\{\mathrm{S}_2 > n_2\} = \mathfrak{s}_{00}^{n_2} + \mathfrak{s}_{01}\frac{\mathfrak{s}_{11}^{n_1} - \mathfrak{s}_{00}^{n_1}}{\mathfrak{s}_{11} - \mathfrak{s}_{00}}$$

and, without loss of generality, taking $n_1 \leq n_2$,

$$\mathbf{P}\{\mathrm{S}_1 > n_1, \mathrm{S}_2 > n_2\} = \mathfrak{s}_{00}^{n_1}[\mathfrak{s}_{00}^{n_2-n_1} + \mathfrak{s}_{01}\frac{\mathfrak{s}_{11}^{n_2-n_1} - \mathfrak{s}_{00}^{n_2-n_1}}{\mathfrak{s}_{11} - \mathfrak{s}_{00}}].$$

Let $\mathbf{T} = (\mathrm{T}_1, \mathrm{T}_2)$ have an MPH distribution as described in (3.4) with the DPH random vector $(\mathrm{S}_1, \mathrm{S}_2)$ and exponential distribution of rate λ. Let $\{N(t), t \geq 0\}$ denote a Poisson process with rate λ. Then for

$0 < x < y$,

$$\begin{aligned}
\mathbf{P}\{\mathrm{T}_1 > x, \mathrm{T}_2 > y\} &= \sum_{i=0}^{\infty}\sum_{j=0}^{\infty} \mathbf{P}\{\mathrm{S}_1 > i, \mathrm{S}_2 > i+j\} \cdot \mathbf{P}\{N(x) = i\}\mathbf{P}\{N(y-x) = j\} \\
&= \sum_{i=0}^{\infty}\sum_{j=0}^{\infty} \mathfrak{s}_{00}^i(\mathfrak{s}_{00}^j + \mathfrak{s}_{01}\frac{\mathfrak{s}_{11}^j - \mathfrak{s}_{00}^j}{\mathfrak{s}_{11} - \mathfrak{s}_{00}}) \cdot \frac{(\lambda x)^i}{i!}e^{-\lambda x} \cdot \frac{(\lambda(y-x))^j}{j!} \cdot e^{-\lambda(y-x)} \\
&= \sum_{i=0}^{\infty} \mathfrak{s}_{00}^i \cdot \frac{(\lambda x)^i}{i!}e^{-\lambda x} \sum_{j=0}^{\infty}(\mathfrak{s}_{00}^j + \mathfrak{s}_{01}\frac{\mathfrak{s}_{11}^j - \mathfrak{s}_{00}^j}{\mathfrak{s}_{11} - \mathfrak{s}_{00}}) \cdot \frac{(\lambda(y-x))^j}{j!} \cdot e^{-\lambda(y-x)} \\
&= e^{\lambda x \mathfrak{s}_{00}} e^{-\lambda x} \cdot (e^{\lambda(y-x)\mathfrak{s}_{00}} + \frac{\mathfrak{s}_{01}}{\mathfrak{s}_{11} - \mathfrak{s}_{00}}(e^{\lambda(y-x)\mathfrak{s}_{11}} - e^{\lambda(y-x)\mathfrak{s}_{00}})) \cdot e^{-\lambda(y-x)} \\
&= \frac{\mathfrak{s}_{01}}{\mathfrak{s}_{11} - \mathfrak{s}_{00}} e^{\lambda x(\mathfrak{s}_{00} - \mathfrak{s}_{11})} e^{\lambda y(\mathfrak{s}_{11} - 1)} + (1 - \frac{\mathfrak{s}_{01}}{\mathfrak{s}_{11} - \mathfrak{s}_{00}}) e^{\lambda y(\mathfrak{s}_{00} - 1)}.
\end{aligned}$$

Following parallel steps for $0 < y < x$,

$$\begin{aligned}
\mathbf{P}\{\mathrm{T}_1 > x, \mathrm{T}_2 > y\} &= \frac{\mathfrak{s}_{02}}{\mathfrak{s}_{22} - \mathfrak{s}_{00}} e^{\lambda y(\mathfrak{s}_{00} - \mathfrak{s}_{22})} e^{\lambda x(\mathfrak{s}_{22} - 1)} + (1 - \frac{\mathfrak{s}_{02}}{\mathfrak{s}_{22} - \mathfrak{s}_{00}}) e^{\lambda x(\mathfrak{s}_{00} - 1)}, \\
\mathbf{P}\{\mathrm{T}_1 > x, \mathrm{T}_2 > x\} &= e^{\lambda x(\mathfrak{s}_{00} - 1)}.
\end{aligned}$$

The marginals are given by,

$$\begin{aligned}
\mathbf{P}\{\mathrm{T}_1 > x\} &= \sum_{i=0}^{\infty} \mathbf{P}\{\mathrm{S}_1 > i\} \cdot \mathbf{P}\{N(x) = i\} \\
&= \sum_{i=0}^{\infty}(\mathfrak{s}_{00}^i + \mathfrak{s}_{02}\frac{\mathfrak{s}_{22}^i - \mathfrak{s}_{00}^i}{\mathfrak{s}_{22} - \mathfrak{s}_{00}}) \cdot \frac{(\lambda x)^i}{i!}e^{-\lambda x} \\
&= (e^{\lambda x \mathfrak{s}_{00}} + \frac{\mathfrak{s}_{02}}{\mathfrak{s}_{22} - \mathfrak{s}_{00}}(e^{\lambda x \mathfrak{s}_{22}} - e^{\lambda x \mathfrak{s}_{00}})) \cdot e^{-\lambda x} \\
&= \frac{\mathfrak{s}_{02}}{\mathfrak{s}_{22} - \mathfrak{s}_{00}} e^{\lambda x(\mathfrak{s}_{22} - 1)} + (1 - \frac{\mathfrak{s}_{02}}{\mathfrak{s}_{22} - \mathfrak{s}_{00}}) e^{\lambda x(\mathfrak{s}_{00} - 1)},
\end{aligned}$$

and similarly,

$$\mathbf{P}\{\mathrm{T}_2 > y\} = \frac{\mathfrak{s}_{01}}{\mathfrak{s}_{11} - \mathfrak{s}_{00}} e^{\lambda y(\mathfrak{s}_{11} - 1)} + (1 - \frac{\mathfrak{s}_{01}}{\mathfrak{s}_{11} - \mathfrak{s}_{00}}) e^{\lambda y(\mathfrak{s}_{00} - 1)}.$$

4.5.2 Bivariate Marshall-Olkin Distributions

Suppose $\mathbf{S}$ has a discrete Marshall-Olkin distribution and $\mathbf{T} = (\mathrm{T}_1, \mathrm{T}_2)$ is a corresponding continuous phase-type random vector associated with $\mathbf{S}$.

Theorem 14 $\mathbf{T}$ *has a continuous Marshall-Olkin distribution.*

Proof:

Recall that $\mathfrak{s}_{00} + \mathfrak{s}_{01} = \mathfrak{s}_{11}$ and $\mathfrak{s}_{00} + \mathfrak{s}_{02} = \mathfrak{s}_{22}$ so, $\frac{\mathfrak{s}_{01}}{\mathfrak{s}_{11} - \mathfrak{s}_{00}} = \frac{\mathfrak{s}_{02}}{\mathfrak{s}_{22} - \mathfrak{s}_{00}} = 1$. Simplifying the general formulas given above:

$$\mathbf{P}\{\mathrm{T}_1 > x, \mathrm{T}_2 > y\} = \begin{cases} e^{-\lambda \mathfrak{s}_{01} x - \lambda(1-\mathfrak{s}_{11}) y} & x < y \\ e^{-\lambda(1-\mathfrak{s}_{00}) x} & x = y \\ e^{-\lambda \mathfrak{s}_{02} y - \lambda(1-\mathfrak{s}_{22}) x} & x > y \end{cases}$$

Both cases where $x \neq y$ simplify to the form given when $x = y$.

With $1 - \mathfrak{s}_{00} = \mathfrak{s}_{01} + \mathfrak{s}_{02} + \mathfrak{s}_{03}$, $1 - \mathfrak{s}_{22} = \mathfrak{s}_{01} + \mathfrak{s}_{03}$, and $1 - \mathfrak{s}_{11} = \mathfrak{s}_{02} + \mathfrak{s}_{03}$ some algebra results in the following simplification:

$$\mathbf{P}\{\mathrm{T}_1 > x, \mathrm{T}_2 > y\} = \begin{cases} e^{-\lambda \mathfrak{s}_{01} x - \lambda \mathfrak{s}_{02} y - \lambda \mathfrak{s}_{03} y} & x < y \\ e^{-\lambda \mathfrak{s}_{01} x - \lambda \mathfrak{s}_{02} y - \lambda \mathfrak{s}_{03} y} & x = y \\ e^{-\lambda \mathfrak{s}_{01} x - \lambda \mathfrak{s}_{02} y - \lambda \mathfrak{s}_{03} x} & x > y \end{cases}$$

or,

$$\mathbf{P}\{\mathrm{T}_1 > x, \mathrm{T}_2 > y\} = e^{-\lambda \mathfrak{s}_{01} x - \lambda \mathfrak{s}_{02} y - \lambda \mathfrak{s}_{03} \max\{x, y\}}.$$

This is the survival function for a bivariate Marshall-Olkin distribution with parameters $\lambda_1 = \lambda \mathfrak{s}_{01}$, $\lambda_2 = \lambda \mathfrak{s}_{02}$, and $\lambda_{12} = \lambda \mathfrak{s}_{03}$ ([15]). □

Corollary 6 *If* $\mathbf{T}$ *is a continuous bivariate Marshall-Olkin distribution with parameters* λ_1, λ_2, *and* λ_{12}, *then there exists a discrete bivariate Marshall-Olkin distribution* $\mathbf{S}$ *with which (3.4) is satisfied.*

Proof:

Let $\lambda > \lambda_1 + \lambda_2 + \lambda_{12}$ with $\mathfrak{s}_{01} = \frac{\lambda_1}{\lambda}$, $\mathfrak{s}_{02} = \frac{\lambda_2}{\lambda}$, and $\mathfrak{s}_{03} = \frac{\lambda_{12}}{\lambda}$. Taking $\mathfrak{s}_{00} = 1 - \mathfrak{s}_{01} - \mathfrak{s}_{02} - \mathfrak{s}_{03}$, $\mathfrak{s}_{11} = \mathfrak{s}_{00} + \mathfrak{s}_{01}$, and $\mathfrak{s}_{22} = \mathfrak{s}_{00} + \mathfrak{s}_{02}$, then

$$\mathfrak{S} = \begin{pmatrix} \mathfrak{s}_{00} & \mathfrak{s}_{01} & \mathfrak{s}_{02} & \mathfrak{s}_{03} \\ 0 & \mathfrak{s}_{11} & 0 & 1 - \mathfrak{s}_{11} \\ 0 & 0 & \mathfrak{s}_{22} & 1 - \mathfrak{s}_{22} \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

is the transition matrix for a bivariate discrete Marshall-Olkin distribution $\mathbf{S}$. Consider $\hat{\mathbf{T}}$, the MPH associated with $\mathbf{S}$ by (3.4) with the exponential distribution with parameter λ. Then

$$\begin{aligned}\mathbf{P}\{\hat{\mathrm{T}}_1 > x, \hat{\mathrm{T}}_2 > y\} &= e^{-\lambda\mathfrak{s}_{01}x-\lambda\mathfrak{s}_{02}y-\lambda\mathfrak{s}_{03}\max\{x,y\}} \\ &= e^{-\lambda\frac{\lambda_1}{\lambda}x-\lambda\frac{\lambda_2}{\lambda}y-\lambda\frac{\lambda_{12}}{\lambda}\max\{x,y\}} \\ &= e^{-\lambda_1 x-\lambda_2 y-\lambda_{12}\max\{x,y\}} \\ &= \mathbf{P}\{\mathrm{T}_1 > x, \mathrm{T}_2 > y\},\end{aligned}$$

so $\mathbf{T}$ is a corresponding MPH random vector with $\mathbf{S}$. □

Corollary 7 *Let* $\mathbf{T}$ *be a bivariate Marshall-Olkin distribution with parameters* λ_1, λ_2, *and* λ_{12}. *Furthermore, let* $\lambda' > \lambda^* > \lambda_1 + \lambda_2 + \lambda_{12}$. *If* $\mathbf{S}'$ *and* $\mathbf{S}^*$ *are two discrete Marshall-Olkin distributions associated with* $\mathbf{T}$ *by* λ' *and* λ^*, *respectively, then for every* $n_1, n_2 > 0$, $\mathbf{P}\{\mathrm{S}'_1 > n_1, \mathrm{S}'_2 > n_2\} \geq \mathbf{P}\{\mathrm{S}^*_1 > n_1, \mathrm{S}^*_2 > n_2\}$.

Proof:

The following statements are equivalent:

$$\begin{gathered}\lambda^* < \lambda' \\ -\lambda^*(\lambda_1+\lambda_2+\lambda_{12}) > -\lambda'(\lambda_1+\lambda_2+\lambda_{12}) \\ \lambda^*\lambda' - \lambda^*(\lambda_1+\lambda_2+\lambda_{12}) > \lambda^*\lambda' - \lambda'(\lambda_1+\lambda_2+\lambda_{12}) \\ \lambda^*(\lambda'-\lambda_1-\lambda_2-\lambda_{12}) > \lambda'(\lambda^*-\lambda_1-\lambda_2-\lambda_{12}) \\ 1-\frac{\lambda_1}{\lambda'}-\frac{\lambda_2}{\lambda'}-\frac{\lambda_{12}}{\lambda'} > 1-\frac{\lambda_1}{\lambda^*}-\frac{\lambda_2}{\lambda^*}-\frac{\lambda_{12}}{\lambda^*} \\ \mathfrak{s}'_{00} > \mathfrak{s}^*_{00}.\end{gathered}$$

Parallel arguments demonstrate that $\mathfrak{s}'_{11} > \mathfrak{s}^*_{11}$ and $\mathfrak{s}'_{22} > \mathfrak{s}^*_{22}$, so by Theorem 13, $\mathbf{P}\{\mathrm{S}'_1 > n_1, \mathrm{S}'_2 > n_2\} \geq \mathbf{P}\{\mathrm{S}^*_1 > n_1, \mathrm{S}^*_2 > n_2\}$. □

The relationship between continuous and discrete bivariate Marshall-Olkin distributions can be used to characterize the dependency relationship of the two components.

Theorem 15 *Any two Marshall-Olkin distributed random variables are PUOD. In the case of no simultaneous failure, they are independent.*

Proof:

The marginal survival functions of T_1 and T_2 are given by:

$$\mathbf{P}\{\mathrm{T}_1 > x\} = e^{-\lambda x(1-\mathfrak{s}_{22})}$$

and

$$\mathbf{P}\{\mathrm{T}_2 > y\} = e^{-\lambda y(1-\mathfrak{s}_{11})}.$$

The following statements are all equivalent for $x > y$ (a similar argument works for $x < y$):

$$\mathfrak{s}_{01} \leq \mathfrak{s}_{01} + \mathfrak{s}_{03}$$

$$\mathfrak{s}_{11} - \mathfrak{s}_{00} \leq \mathfrak{s}_{23}$$

$$\mathfrak{s}_{11} - \mathfrak{s}_{00} \leq 1 - \mathfrak{s}_{22}$$

$$x(\mathfrak{s}_{11} - \mathfrak{s}_{00}) \leq x(1 - \mathfrak{s}_{22})$$

$$x(\mathfrak{s}_{11} - \mathfrak{s}_{00}) + y(1 - \mathfrak{s}_{11}) \leq x(1 - \mathfrak{s}_{22}) + y(1 - \mathfrak{s}_{11})$$

$$-\lambda(x(\mathfrak{s}_{11} - \mathfrak{s}_{00}) + y(1 - \mathfrak{s}_{11})) \geq -\lambda(x(1 - \mathfrak{s}_{22}) + y(1 - \mathfrak{s}_{11}))$$

$$e^{-\lambda(x(\mathfrak{s}_{11}-\mathfrak{s}_{00})+y(1-\mathfrak{s}_{11})} \geq e^{-\lambda(x(1-\mathfrak{s}_{22})+y(1-\mathfrak{s}_{11}))}$$

$$\mathbf{P}\{\mathrm{T}_1 > x, \mathrm{T}_2 > y\} \geq \mathbf{P}\{\mathrm{T}_1 > x\}\mathbf{P}\{\mathrm{T}_2 > y\}.$$

This implies that T_1 and T_2 are strictly PUOD and therefore PQD except when $\mathfrak{s}_{03} = 0$ (no simultaneous failure can occur), in which case they are independent. □

In fact, it follows from (4.11) that for any $(\mathrm{T}_1, \mathrm{T}_2)$ that has a Marshall-Olkin distribution,

$$\mathrm{T}_1 = \min\{E_1, E_{12}\},\ \mathrm{T}_2 = \min\{E_2, E_{12}\},$$

where E_1, E_2, E_{12} are independent, and exponentially distributed with rates $\lambda_1, \lambda_2, \lambda_{12}$ respectively. Thus, by Theorem 7 (2) and (3), $(\mathrm{T}_1, \mathrm{T}_2)$ is associated, and hence positively upper orthant dependent.

4.5.3 Bivariate Freund Distributions

Suppose $\mathbf{S}$ is a discrete Freund distribution and $\mathbf{T} = (\mathrm{T}_1, \mathrm{T}_2)$ is a continuous phase-type distribution associated with $\mathbf{S}$.

Theorem 16 $\mathbf{T}$ *is a continuous Freund distribution.*

Proof:

Using the formulas for the survival function derived for the general bivariate MDPH case we have,

$$\mathbf{P}\{T_1 > x, T_2 > y\} = \begin{cases} \frac{\mathfrak{s}_{01}}{\mathfrak{s}_{11}-\mathfrak{s}_{00}} e^{\lambda x(\mathfrak{s}_{00}-\mathfrak{s}_{11})} e^{\lambda y(\mathfrak{s}_{11}-1)} + (1 - \frac{\mathfrak{s}_{01}}{\mathfrak{s}_{11}-\mathfrak{s}_{00}}) e^{\lambda y(\mathfrak{s}_{00}-1)} & 0 < x < y \\ \\ \frac{\mathfrak{s}_{02}}{\mathfrak{s}_{22}-\mathfrak{s}_{00}} e^{\lambda y(\mathfrak{s}_{00}-\mathfrak{s}_{22})} e^{\lambda x(\mathfrak{s}_{22}-1)} + (1 - \frac{\mathfrak{s}_{02}}{\mathfrak{s}_{22}-\mathfrak{s}_{00}}) e^{\lambda x(\mathfrak{s}_{00}-1)} & 0 < y < x \end{cases}$$

Let $\alpha' = \lambda(1-\mathfrak{s}_{22})$, $\beta' = \lambda(1-\mathfrak{s}_{11})$, $\alpha = \lambda\mathfrak{s}_{01}$, and $\beta = \lambda\mathfrak{s}_{02}$. Recall $\mathfrak{s}_{00}+\mathfrak{s}_{01}+\mathfrak{s}_{02} = 1$, so $\mathfrak{s}_{00} = 1-\frac{\alpha}{\lambda}-\frac{\beta}{\lambda}$. Substituting and simplifying gives:

$$\mathbf{P}\{\mathrm{T}_1 > x, \mathrm{T}_2 > y\} = \begin{cases} \frac{\alpha}{\alpha+\beta-\beta'}e^{-\beta' y}e^{-(\alpha+\beta-\beta')x} + \left(1-\frac{\alpha}{\alpha+\beta-\beta'}\right)e^{-y(\alpha+\beta)} & 0<x<y \\ \\ \frac{\beta}{\alpha+\beta-\alpha'}e^{-\alpha' x}e^{-(\alpha+\beta-\alpha')y} + \left(1-\frac{\beta}{\alpha+\beta-\alpha'}\right)e^{-x(\alpha+\beta)} & 0<y<x \end{cases}$$

which is the survival function for a bivariate Freund distribution with parameters α, β, α', and β' ([6]). $\square$

Corollary 8 *If* $\mathbf{T}$ *is a continuous bivariate Freund distribution with parameters* α, α', β, *and* β', *then there exists a corresponding discrete bivariate Freund distribution* $\mathbf{S}$ *with which* $\mathbf{T}$ *is associated.*

Proof:
Let $\lambda > \max\{\alpha+\beta, \alpha', \beta'\}$ and $\mathfrak{s}_{11} = 1-\frac{\beta'}{\lambda}$, $\mathfrak{s}_{22} = 1-\frac{\alpha'}{\lambda}$, $\mathfrak{s}_{01} = \frac{\alpha}{\lambda}$, $\mathfrak{s}_{02} = \frac{\beta}{\lambda}$, and $\mathfrak{s}_{00} = 1-\frac{\alpha}{\lambda}-\frac{\beta}{\lambda}$. Then

$$\mathfrak{S} = \begin{pmatrix} \mathfrak{s}_{00} & \mathfrak{s}_{01} & \mathfrak{s}_{02} & 0 \\ 0 & \mathfrak{s}_{11} & 0 & 1-\mathfrak{s}_{11} \\ 0 & 0 & \mathfrak{s}_{22} & 1-\mathfrak{s}_{22} \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

is the transition matrix for a bivariate discrete Freund distribution $\mathbf{S}$. Consider $\hat{\mathbf{T}}$, the MPH associated with $\mathbf{S}$ by a Poisson process with parameter λ. Following the work above, it is easy to see that $\hat{\mathbf{T}}$ and $\mathbf{T}$ have the same distribution. $\square$

A Class of NQD Bivariate MPH Distributions

The following results will show that the only NUOD bivariate simple MPH distributions are a class of Freund distributions. Lemma 3 implies that this class must also be the only bivariate simple MPH distributions that are NQD.

Let $R(x,y) = \dfrac{\mathbf{P}\{\mathrm{T}_1 > x\}\mathbf{P}\{\mathrm{T}_2 > y\}}{\mathbf{P}\{\mathrm{T}_1 > x, \mathrm{T}_2 > y\}}$. Clearly $\mathbf{T}$ is NUOD if and only if $\forall x, y$, $R(x,y) \geq 1$.

Without loss of generality, assume $0 \leq x \leq y$. Let $A_1 = \frac{\mathfrak{s}_{01}}{\mathfrak{s}_{11}-\mathfrak{s}_{00}}$, $A_2 = \frac{\mathfrak{s}_{02}}{\mathfrak{s}_{22}-\mathfrak{s}_{00}}$, and $B_x = e^{\lambda x(\mathfrak{s}_{00}-\mathfrak{s}_{11})}$, then

$$\mathbf{P}\{\mathrm{T}_1 > x, \mathrm{T}_2 > y\} = A_1 B_x e^{\lambda y(\mathfrak{s}_{11}-1)} + (1-A_1)e^{\lambda y(\mathfrak{s}_{00}-1)},$$

$$\mathbf{P}\{\mathrm{T}_1 > x\} = (1-A_2)e^{\lambda x(\mathfrak{s}_{00}-1)} + A_2 e^{\lambda x(\mathfrak{s}_{22}-1)},$$

$$\mathbf{P}\{\mathrm{T}_2 > y\} = (1-A_1)e^{\lambda y(\mathfrak{s}_{00}-1)} + A_1 e^{\lambda y(\mathfrak{s}_{11}-1)},$$

and

$$R(x,y) = \frac{\left((1-A_2)e^{\lambda x(\mathfrak{s}_{00}-1)} + A_2 e^{\lambda x(\mathfrak{s}_{22}-1)}\right)\left((1-A_1)e^{\lambda y(\mathfrak{s}_{00}-1)} + A_1 e^{\lambda y(\mathfrak{s}_{11}-1)}\right)}{A_1 B_x e^{\lambda y(\mathfrak{s}_{11}-1)} + (1-A_1)e^{\lambda y(\mathfrak{s}_{00}-1)}}. \tag{4.12}$$

Theorem 17 *A necessary condition for* $\mathbf{T}$ *to be NUOD is* $\mathfrak{s}_{03} = 0$.

Proof:

Consider $R(x,x)$. After simplifying we have

$$R(x,x) = (1-A_2)(1-A_1)e^{\lambda x(\mathfrak{s}_{00}-1)} + (1-A_2)A_1 e^{\lambda x(\mathfrak{s}_{11}-1)} +$$

$$A_2(1-A_1)e^{\lambda x(\mathfrak{s}_{22}-1)} + A_1 A_2 e^{\lambda x(\mathfrak{s}_{22}+\mathfrak{s}_{11}-\mathfrak{s}_{00}-1)},$$

$R(0,0) = 1$, and $\frac{dR}{dx}|_{x=0} = \lambda(\mathfrak{s}_{00} + \mathfrak{s}_{01} + \mathfrak{s}_{02} - 1)$. $\lambda(\mathfrak{s}_{00} + \mathfrak{s}_{01} + \mathfrak{s}_{02} - 1) \leq 0$ with equality if and only if $\mathfrak{s}_{00} + \mathfrak{s}_{01} + \mathfrak{s}_{02} = 1$, but $\mathfrak{s}_{00} + \mathfrak{s}_{01} + \mathfrak{s}_{02} = 1$ if and only if $\mathfrak{s}_{03} = 0$. Suppose $\mathfrak{s}_{03} > 0$, then $\frac{dR}{dx}|_{x=0} < 0$. This implies that there exists some $c > 0$ such that $R(c,c) < R(0,0) = 1$, but $\mathbf{T}$ is NUOD if and only if $R(x,y) \geq 1\ \forall x,y$. Therefore, $\mathfrak{s}_{03} > 0$ implies that $\mathbf{T}$ is not NUOD. □

Theorem 18 *Necessary conditions for* $\mathbf{T}$ *to be NUOD are* $0 < \frac{\mathfrak{s}_{01}}{\mathfrak{s}_{11}-\mathfrak{s}_{00}} \leq 1$, *and* $0 < \frac{\mathfrak{s}_{02}}{\mathfrak{s}_{22}-\mathfrak{s}_{00}} \leq 1$.

Proof:

If $\mathfrak{s}_{11} < \mathfrak{s}_{00}$ then $\lim_{y\to\infty} R(x,y) = \mathbf{P}\{\mathrm{T}_1 > x\} < 1$, so there is an x and y such that $R(x,y) < 1$. This implies that if $\mathfrak{s}_{11} < \mathfrak{s}_{00}$, $\mathbf{T}$ is not NUOD. A parallel argument works for $\mathfrak{s}_{22}$ and $\mathfrak{s}_{00}$. Therefore $\frac{\mathfrak{s}_{01}}{\mathfrak{s}_{11}-\mathfrak{s}_{00}} > 0$ and $\frac{\mathfrak{s}_{02}}{\mathfrak{s}_{22}-\mathfrak{s}_{00}} > 0$ are necessary conditions for $\mathbf{T}$ to be NUOD.

Consider

$$R(x, \tfrac{1}{\lambda}) = \mathbf{P}\{\mathrm{T}_2 > \tfrac{1}{\lambda}\} \frac{(1-A_2)e^{\lambda x(\mathfrak{s}_{00}-1)} + A_2 e^{\lambda x(\mathfrak{s}_{22}-1)}}{A_1 B_x e^{\mathfrak{s}_{11}-1} + (1-A_1)e^{\mathfrak{s}_{00}-1}}.$$

$$\frac{\partial R}{\partial x}|_{x=0} = \mathbf{P}\{\mathrm{T}_2 > \tfrac{1}{\lambda}\} \frac{\lambda \mathfrak{s}_{01}(1-A_1)e^{\mathfrak{s}_{00}-1}(e^{\mathfrak{s}_{11}-\mathfrak{s}_{00}} - 1)}{\left(A_1 e^{\mathfrak{s}_{11}-1} + (1-A_1)e^{\mathfrak{s}_{00}-1}\right)^2}.$$

Since it has already been shown that $\mathfrak{s}_{11} > \mathfrak{s}_{00}$ must be true for $\mathbf{T}$ to be NUOD, all factors except possibly $(1-A_1)$ are positive. If $A_1 > 1$, then $\frac{\partial R}{\partial x}|_{x=0} < 0$. $R(0, \frac{1}{\lambda}) = 1$ so there is some $c > 0$ such that $R(c, \frac{1}{\lambda}) < 1$. Therefore, $A_1 \leq 1$ is a necessary condition for $\mathbf{T}$ to be NUOD. The argument to show that $A_2 \leq 1$ is exactly parallel to this one with $x \geq y$ (and the appropriate changes in $R(x,y)$). □

Theorem 19 $\mathfrak{s}_{03} = 0$, $0 < \frac{\mathfrak{s}_{01}}{\mathfrak{s}_{11}-\mathfrak{s}_{00}} \leq 1$, *and* $0 < \frac{\mathfrak{s}_{02}}{\mathfrak{s}_{22}-\mathfrak{s}_{00}} \leq 1$ *are sufficient conditions for* $\mathbf{T}$ *to be NUOD.*

Proof:

Assuming that $0 < \frac{\mathfrak{s}_{01}}{\mathfrak{s}_{11}-\mathfrak{s}_{00}} \leq 1$ implies that $\mathfrak{s}_{00} < \mathfrak{s}_{11}$, so $1 - B_x > 0$ and $1 - A_1 \geq 0$.

It is straightforward to verify that

$$R(x,y) = \mathbf{P}\{\mathrm{T}_1 > x\}\left(1 + \frac{A_1(1-B_x)}{A_1B_x + (1-A_1)e^{\lambda y(\mathfrak{s}_{00}-\mathfrak{s}_{11})}}\right).$$

Note that for $A_1 = 1$, $R(x,y)$ is constant with respect to y.

Assuming $A_1 < 1$, after some simplification we have,

$$\frac{\partial R}{\partial y} = \frac{\mathbf{P}\{\mathrm{T}_1 > x\}\lambda\mathfrak{s}_{01}(1-A_1)(1-B_x)e^{\lambda y(\mathfrak{s}_{00}-\mathfrak{s}_{11})}}{\left(A_1B_x + (1-A_1)e^{\lambda y(\mathfrak{s}_{00}-\mathfrak{s}_{11})}\right)^2} > 0.$$

That is, for fixed x and $y \geq x$, $R(x,y)$ is increasing in y.

So, for every x, y_1, and y_2 such that $x \leq y_1 \leq y_2$, $0 < \frac{\mathfrak{s}_{01}}{\mathfrak{s}_{11}-\mathfrak{s}_{00}} \leq 1$ implies that $R(x,y_1) \leq R(x,y_2)$. A parallel argument for $0 < \frac{\mathfrak{s}_{02}}{\mathfrak{s}_{22}-\mathfrak{s}_{00}} \leq 1$ implies that for every x_1, x_2, and y such that $y \leq x_1 \leq x_2$ $R(x_1,y) \leq R(x_2,y)$. If it can be shown that for every x, $R(x,x) \geq 1$ then this implies that for every x, y, $R(x,y) \geq 1$ and $\mathbf{T}$ is NUOD.

Recall that

$$R(x,x) = (1-A_2)(1-A_1)e^{\lambda x(\mathfrak{s}_{00}-1)} + (1-A_2)A_1e^{\lambda x(\mathfrak{s}_{11}-1)} +$$

$$A_2(1-A_1)e^{\lambda x(\mathfrak{s}_{22}-1)} + A_1A_2e^{\lambda x(\mathfrak{s}_{22}+\mathfrak{s}_{11}-\mathfrak{s}_{00}-1)},$$

$R(0,0) = 1$, and if $\mathfrak{s}_{03} = 0$, $\frac{dR}{dx}|_{x=0} = 0$.

$$\frac{d^2R}{dx^2} = \lambda^2(\mathfrak{s}_{00}-1)^2(1-A_2)(1-A_1)e^{\lambda x(\mathfrak{s}_{00}-1)} + \lambda^2(\mathfrak{s}_{11}-1)^2(1-A_2)A_1e^{\lambda x(\mathfrak{s}_{11}-1)}$$
$$+ \lambda^2(\mathfrak{s}_{22}-1)^2A_2(1-A_1)e^{\lambda x(\mathfrak{s}_{22}-1)} + \lambda^2(\mathfrak{s}_{22}+\mathfrak{s}_{11}-\mathfrak{s}_{00}-1)^2A_1A_2e^{\lambda x(\mathfrak{s}_{22}+\mathfrak{s}_{11}-\mathfrak{s}_{00}-1)},$$

Each term is positive, so $R(x,x)$ is concave up with a slope of 0 at $x = 0$. This implies that $R(x,x)$ is increasing for $x \geq 0$, and $R(x,x) \geq R(0,0) = 1$. Therefore $\mathbf{T}$ is NUOD. □

Theorem 19 also leads to the solution of the following open problem for Freund distributions. Consider Example 5 again. The conditions in Theorem 19 can be expressed, in terms of the matrix of Example 5, as follows.

$$1 > p_1 \geq p_{1|2} \geq 0,$$

$$1 > p_2 \geq p_{2|1} \geq 0.$$

Thus, if these inequalities hold, the corresponding bivariate Freund distribution is NUOD. This is reasonable because if the probability that destroys component i $(i = 1,2)$ becomes smaller when the other component fails first, then component i would be more likely to survive upon the failure of the other component, and as such, two components are negatively dependent.

Li showed that if

$$p_1 \leq p_{1|2},\, p_2 \leq p_{2|1},$$

then the Freund distribution is PUOD ([10]). Our result complements Li's result and provides a complete picture for dependence structure of the continuous Freund distributions.

Chapter 5

Computation and Applications

In the first section of this chapter we provide a method for directly calculating probabilities for a general class of conditions. Included in this class is the probability mass function. We also give an algorithm for generating random vectors from a given MDPH distribution and extend this result to random vectors from a given MPH distribution. In the second section we discuss two applications of MDPH distributions. The first application considers a system subject to periodic inspections with component lifetimes that have a MPH distribution. The component lifetimes in terms of the number of inspections have a MDPH distribution. We consider examples where it is of interest to find the inspection interval length that provides the lowest long-term cost given a distribution and a set of expenses. In the second application, we derive a formula for finding the mean time to failure of a coherent reliability system whose component lifetimes have a MDPH distribution.

5.1 Computation

5.1.1 Computing Probabilities

It may be useful to calculate directly the probability of a MDPH vector occurring in a region defined by a set of conditions $\mathrm{S}_{k_i} \leq n_i$, $i = 1, 2, \cdots, N_1$ and $\mathrm{S}_{k_i} > n_i$, $i = N_1+1, N_1+2, \cdots, N$ where the k_i are not necessarily unique. Although it is technically possible to calculate such probabilities using the survival function, the following generalization of the method described in Theorem 1 provides a more convenient approach to this class of regions.

Theorem 20 *Consider a set of conditions* $\mathrm{S}_{k_i} \leq n_i,\ i = 1, 2, \cdots, N_1$ *and* $\mathrm{S}_{k_i} > n_i,\ i = N_1+1, N_1+2, \cdots, N$. *Let* $\Gamma_j = \left(\bigcap_{i \leq j; o_i \leq N_1} \Sigma_{k_{o_i}}\right) \bigcap \left(\bigcap_{i > j; o_i > N_1} \Sigma^c_{k_{o_i}}\right)$. *With* $\mathbf{I}_{\Gamma_j}$ *defined as in Theorem 1, we have:*

$$\mathbf{P}\{\mathrm{S}_{k_i} \leq n_i; i \leq N_1,\ \mathrm{S}_{k_i} > n_i; N_1 < i \leq N\} = \sigma \cdot \mathfrak{S}^{n_{o_1}} \cdot \mathbf{I}_{\Gamma_1} \cdot \prod_{j=2}^{N} \left(\mathfrak{S}^{n_{o_j} - n_{o_{j-1}}} \cdot \mathbf{I}_{\Gamma_j}\right) \cdot \underline{\mathbf{1}}.$$

Proof:

The idea is the same as for Theorem 1. First note that

$$\mathbf{P}\{\mathrm{S}_{k_i} \leq n_i; i \leq N_1,\ \mathrm{S}_{k_i} > n_i; N_1 < i \leq N\} = \mathbf{P}\{\mathbb{S}^{n_i} \in \Sigma_{k_i}; i \leq N_1,\ \mathbb{S}^{n_i} \in \Sigma^c_{k_i}; N_1 < i \leq N\}.$$

As in Theorem 1, Γ_j is constructed so that $\mathbf{I}_{\Gamma_j}$ will restrict the Markov chain to those paths which satisfy the constraints. □

Corollary 9 *Conditions of the form* $\mathrm{S}_i = n_i$, $\mathrm{S}_i \geq n_i$, *and* $\mathrm{S}_i < n_i$ *can be easily converted to the two condition types used above. In particular,* $\mathrm{S}_i = n_i$ *is equivalent to* ($\mathrm{S}_i \leq n_i$, $\mathrm{S}_i > n_i - 1$), $\mathrm{S}_i \geq n_i$ *is equivalent to* $\mathrm{S}_i > n_i - 1$, *and* $\mathrm{S}_i < n_i$ *is equivalent to* $\mathrm{S}_i \leq n_i - 1$.

We have implemented the computational method suggested by this theorem and corollary in the **R** dialect of the **S** programming language [16]. The code is provided in the appendix.

5.1.2 Generating Random Vectors

It may be of interest to generate vectors with a particular MDPH distribution. The approach we have taken is to simulate the underlying Markov chain while keeping track of the time of entry into the stochastically closed classes. Given an m dimensional MDPH random vector $\mathbf{S}$ with underlying Markov chain $\mathbb{S} = \{S, \mathbf{\Sigma}, \mathfrak{S}, \sigma\}$ where $S = \{s_1, \cdots, s_N\}$, the algorithm is as follows:

1. Set $n_1 = \cdots = n_m = 0$
2. Generate the initial state $\mathbb{S}^0 = s_i$ where i is from a multinomial distribution with probabilities given by σ
3. While $\mathbb{S}^n \notin S_\Delta$:
   (a) Generate next state s_j given current state s_i; $\mathbb{S}^{n+1} = s_j \mid \mathbb{S}^n = s_i$ where j is from a multinomial distribution with probabilities given by row i of $\mathfrak{S}$.

(b) For each k such that $s_j \in \Sigma_k^c$, increment n_k by 1

4. $\mathbf{n} = (n_1, \cdots, n_m)$ is an m dimensional random vector from the given MDPH distribution.

Example 10 (Estimating Higher Joint Moments of a Simple MDPH) Consider a simple bivariate MDPH $\mathbf{S}$, with underlying transition matrix:

$$\mathfrak{S} = \begin{pmatrix} 0.2 & 0.35 & 0.35 & 0.1 \\ 0 & 0.6 & 0 & 0.4 \\ 0 & 0 & 0.6 & 0.4 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

Note that the marginal probabilities for S_1 are the same as for S_2, so it suffices to look only at $E[\mathrm{S}_1^n \mathrm{S}_2^m]$ where $n \leq m$. Using the moment generating function derived in Chapter 4, we calculated mixed order moments for $\mathbf{S}$ directly. We also used simulation runs of 500000 repeated for each case to estimate the moments and the standard error of the estimate.

	C*alculated*	E*stimated*
$E[\mathrm{S}_2]$	2.343	2.345 ± 0.005
$E[\mathrm{S}_2^2]$	8.984	8.980 ± 0.049
$E[\mathrm{S}_1\mathrm{S}_2]$	4.609	4.601 ± 0.012
$E[\mathrm{S}_1\mathrm{S}_2^2]$	15.215	15.228 ± 0.093
$E[\mathrm{S}_1^2\mathrm{S}_2^2]$	41.035	40.918 ± 0.400

To simulate an MPH $\mathbf{T}$ we generate a vector from the underlying MDPH. If N is the maximum value of this vector, then generate $(E_1, E_2, \cdots, E_N)$ from an exponential distribution with parameter λ. The value of $T_i = \sum_{j=1}^{n_i} E_i, i = 1, \ldots, m$.

5.2 Applications

5.2.1 Periodic Inspections

Consider an m component system with component lifetimes having a MPH distribution $\mathbf{T} = (\mathrm{T}_1, \cdots, \mathrm{T}_m)$. Suppose $\mathbf{T}$ has an underlying MDPH $\mathbf{S}$ (with underlying Markov chain $\mathbb{S} = \{S, \mathbf{\Sigma}, \mathfrak{S}, \sigma\}$) associated by parameter λ and the system is subject to periodic inspections every i time units. It is assumed that failure of a component in the system can only be known by inspection. $\mathbf{S}^* = (\mathrm{S}_1^*, \cdots, \mathrm{S}_m^*)$ can be considered the discrete distribution of failure times in terms of the number of inspections until failure is observed.

Theorem 21 $\mathbf{S}^*$ *has a MDPH distribution with underlying Markov chain* $\mathbb{S} = \{S, \mathbf{\Sigma}, \mathfrak{S}^*, \sigma\}$, *where* $\mathfrak{S}^* = e^{-\lambda i} e^{\lambda i \mathfrak{S}}$.

Proof:

For inspection intervals of fixed length i, this follows immediately from Corollary 2 by taking $t = i$. □

Example 11 (Bivariate Simple MPH with Periodic Inspection) Let $\mathbf{T}$ be a bivariate simple MPH distribution with underlying MDPH $\mathbf{S}$ associated to $\mathbf{T}$ by parameter λ. If $\mathbb{S} = \{S, \mathbf{\Sigma}, \mathfrak{S}, \sigma\}$ is the underlying Markov chain for $\mathbf{S}$, it is straightforward to verify that, for inspection interval i, the MDPH $\mathbf{S}^*$ induced by the periodic inspection will have the following transition matrix:

$$\mathfrak{S}^* = \begin{pmatrix} \mathfrak{s}^*_{00} & \mathfrak{s}^*_{01} & \mathfrak{s}^*_{02} & 1 - \mathfrak{s}^*_{00} - \mathfrak{s}^*_{01} - \mathfrak{s}^*_{02} \\ 0 & \mathfrak{s}^*_{11} & 0 & 1 - \mathfrak{s}^*_{11} \\ 0 & 0 & \mathfrak{s}^*_{22} & 1 - \mathfrak{s}^*_{22} \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

$$= \begin{pmatrix} e^{\lambda i(\mathfrak{s}_{00}-1)} & \frac{\mathfrak{s}_{01}}{\mathfrak{s}_{11}-\mathfrak{s}_{00}}\left(e^{\lambda i(\mathfrak{s}_{11}-1)} - e^{\lambda i(\mathfrak{s}_{00}-1)}\right) & \frac{\mathfrak{s}_{02}}{\mathfrak{s}_{22}-\mathfrak{s}_{00}}\left(e^{\lambda i(\mathfrak{s}_{22}-1)} - e^{\lambda i(\mathfrak{s}_{00}-1)}\right) & 1 - \mathfrak{s}^*_{00} - \mathfrak{s}^*_{01} - \mathfrak{s}^*_{02} \\ 0 & e^{\lambda i(\mathfrak{s}_{11}-1)} & 0 & 1 - \mathfrak{s}^*_{11} \\ 0 & 0 & e^{\lambda i(\mathfrak{s}_{22}-1)} & 1 - \mathfrak{s}^*_{22} \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

$$= \begin{pmatrix} e^{\lambda i(\mathfrak{s}_{00}-1)} & \frac{\mathfrak{s}_{01}}{\mathfrak{s}_{11}-\mathfrak{s}_{00}}\left(\mathfrak{s}^*_{11} - \mathfrak{s}^*_{00}\right) & \frac{\mathfrak{s}_{02}}{\mathfrak{s}_{22}-\mathfrak{s}_{00}}\left(\mathfrak{s}^*_{22} - \mathfrak{s}^*_{00}\right) & 1 - \mathfrak{s}^*_{00} - \mathfrak{s}^*_{01} - \mathfrak{s}^*_{02} \\ 0 & e^{\lambda i(\mathfrak{s}_{11}-1)} & 0 & 1 - \mathfrak{s}^*_{11} \\ 0 & 0 & e^{\lambda i(\mathfrak{s}_{22}-1)} & 1 - \mathfrak{s}^*_{22} \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

Example 12 (Bivariate Marshall-Olkin with Periodic Inspection) If $\mathbf{T}$ is a Marshall-Olkin distribution then $\mathfrak{s}_{00} + \mathfrak{s}_{01} = \mathfrak{s}_{11}$ and $\mathfrak{s}_{00} + \mathfrak{s}_{02} = \mathfrak{s}_{22}$ so

$$\mathfrak{S}^* = \begin{pmatrix} e^{\lambda i(\mathfrak{s}_{00}-1)} & \mathfrak{s}^*_{11} - \mathfrak{s}^*_{00} & \mathfrak{s}^*_{22} - \mathfrak{s}^*_{00} & 1 + \mathfrak{s}^*_{00} - \mathfrak{s}^*_{11} - \mathfrak{s}^*_{22} \\ 0 & e^{\lambda i(\mathfrak{s}_{11}-1)} & 0 & 1 - \mathfrak{s}^*_{11} \\ 0 & 0 & e^{\lambda i(\mathfrak{s}_{22}-1)} & 1 - \mathfrak{s}^*_{22} \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

Therefore $\mathfrak{s}^*_{00} + \mathfrak{s}^*_{01} = \mathfrak{s}^*_{11}$ and $\mathfrak{s}^*_{00} + \mathfrak{s}^*_{02} = \mathfrak{s}^*_{22}$ which means that $\mathbf{S}^*$ is a Marshall-Olkin distribution also. This implies that if a certain system has a bivariate Marshall-Olkin distribution, the distribution of failures under a periodic inspection regime will have a bivariate discrete Marshall-Olkin distribution. Furthermore, it follows from Theorem 13 that if $\mathbf{S}^*_1$ is induced by i_1 and $\mathbf{S}^*_2$ is induced by i_2, with $i_1 < i_2$, then $\mathbf{S}^*_1 \geq_{uo} \mathbf{S}^*_2$.

Example 13 (Optimization of Inspection Interval for Given Costs) Suppose that units consisting of a two component system with component lifetimes having a Marshall-Olkin distribution are subject to replacement upon failure of at least one component. If an inspection finds that one component has failed, replacement of the unit can occur with minimal system downtime. The cost associated with this replacement will be C_r. If an inspection finds that both components have failed, the system has failed and downtime is incurred along with an associated cost. As a simplification, we assume that the cost due to the downtime is proportional to the length of the inspection interval. Adding the cost due to downtime with cost of replacement we have the cost associated with an observed failure of both components: $i \cdot C_{dt} + C_r$. Finally, each inspection has a certain associated cost given by C_{in}. The lifetime cost of each unit is given by:

$$C(\mathrm{S}_1^*, \mathrm{S}_2^*) = \begin{cases} C_{in} \min\{\mathrm{S}_1^*, \mathrm{S}_2^*\} + C_r & \mathrm{S}_1^* \neq \mathrm{S}_2^* \\ \\ C_{in} \min\{\mathrm{S}_1^*, \mathrm{S}_2^*\} + i \cdot C_{dt} + C_r & \mathrm{S}_1^* = \mathrm{S}_2^* \end{cases}$$

Using the joint pmf given in Section 4.2, it is possible to calculate the expected cost as a function of the given costs and the parameters of the transition matrix $\mathfrak{S}^*$

$$E[C] = \left(\frac{1}{1 - \mathfrak{s}_{00}^*}\right) C_{in} + \frac{\mathfrak{s}_{03}^*}{1 - \mathfrak{s}_{00}^*} \left(i \cdot C_{dt} + C_r\right) + \left(1 - \frac{\mathfrak{s}_{03}^*}{1 - \mathfrak{s}_{00}^*}\right) C_r$$

Making use of the fact that the distribution is Marshall-Olkin, rearranging and substituting to put it in terms of the original underlying MDPH, λ, and i results in:

$$E[C] = \left(\frac{1}{1 - e^{\lambda i(\mathfrak{s}_{00}-1)}}\right) C_{in} + C_r + i \cdot C_{dt} \frac{1 + e^{\lambda i(\mathfrak{s}_{00}-1)} - e^{\lambda i(\mathfrak{s}_{11}-1)} - e^{\lambda i(\mathfrak{s}_{22}-1)}}{1 - e^{\lambda i(\mathfrak{s}_{00}-1)}}.$$

For a given distribution and costs, it may be of interest to find the inspection interval that minimizes the expected per unit cost. Since λ is the expected number of shock events for a fixed unit time interval, without loss of generality, time units can be chosen such that $\lambda = 1$. The costs will also be scaled relative to C_{in} so that $C_{in} = 1$. Taking $\mathfrak{s}_{00} = \frac{9}{10}$, $\mathfrak{s}_{11} = \mathfrak{s}_{22} = \frac{47}{50}$, $C_r = 10$, and $C_{dt} = 10$ gives the expected cost of a unit as a function of i.

$$E[C] = \frac{1}{1 - e^{\frac{-i}{10}}} + 10 + 10i \frac{1 + e^{\frac{-i}{10}} - 2e^{\frac{-3i}{50}}}{1 - e^{\frac{-i}{10}}}.$$

Using a computer algebra system to find the inspection interval period that gives the minimum cost per unit gives 1.86 time units as the approximate optimal inspection interval length.

Clearly not all units will have the same lifespan, so it may be that a question of greater interest is the overall average cost per time of operating many units in succession. In this case it is necessary to know the expected lifetime of a unit. For this example, the lifetime, L, of a particular unit will be considered to end when it is replaced. So,

$$L(\mathrm{S}_1^*, \mathrm{S}_2^*) = i \min\{\mathrm{S}_1^*, \mathrm{S}_2^*\}$$

and the expected lifetime is

$$E[L] = \frac{i}{1 - e^{\frac{-i}{10}}}$$

Finally,

$$\frac{E[C]}{E[L]} = \frac{11 - 10e^{\frac{-i}{10}} + 10i\left(1 + e^{\frac{-i}{10}} - 2e^{\frac{-3i}{50}}\right)}{i}.$$

Using a computer algebra system to minimize this function with respect to i indicates that the optimum inspection interval to minimize the average cost per unit of time is approximately 2.20 time units.

Example 14 Consider a system with two components which is subject to damaging shocks. Each component may exist in one of three states after receiving a shock; no damage, minor damage, or failed. The system fails if both components have failed or experienced shocks which resulted in minor damage. The state space of the system can be described by $S = \{s_a, s_b, s_c, s_1, s_2, s_3\}$ where s_a corresponds to no damage, s_b corresponds to minor damage to the first component and no damage to the second component, s_c corresponds to minor damage to the second component with no damage to the first component, s_1 and s_2 correspond to failure of the first and second components respectively, and s_3 corresponds to the failure of the system. To simplify the example, suppose that the probability of a certain type of shock damage does not depend on the state of the system, but two minor damage shocks to the same component result in the failure of that component and the two components have the same distribution of failures; then

$$\mathfrak{S} = \begin{pmatrix} \mathfrak{s}_{aa} & \mathfrak{s}_{ab} & \mathfrak{s}_{ac} & \mathfrak{s}_{a1} & \mathfrak{s}_{a2} & \mathfrak{s}_{a3} \\ 0 & \mathfrak{s}_{bb} & 0 & \mathfrak{s}_{b1} & 0 & \mathfrak{s}_{b3} \\ 0 & 0 & \mathfrak{s}_{cc} & 0 & \mathfrak{s}_{c2} & \mathfrak{s}_{c3} \\ 0 & 0 & 0 & \mathfrak{s}_{11} & 0 & \mathfrak{s}_{13} \\ 0 & 0 & 0 & 0 & \mathfrak{s}_{22} & \mathfrak{s}_{23} \\ 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}$$

$$= \begin{pmatrix} \mathfrak{s}_{aa} & \mathfrak{s}_{ab} & \mathfrak{s}_{ab} & \mathfrak{s}_{a1} & \mathfrak{s}_{a1} & 1-(\mathfrak{s}_{aa}+2\mathfrak{s}_{ab}+2\mathfrak{s}_{a1}) \\ 0 & \mathfrak{s}_{aa} & 0 & \mathfrak{s}_{ab}+\mathfrak{s}_{a1} & 0 & 1-(\mathfrak{s}_{aa}+\mathfrak{s}_{ab}+\mathfrak{s}_{a1}) \\ 0 & 0 & \mathfrak{s}_{aa} & 0 & \mathfrak{s}_{ab}+\mathfrak{s}_{a1} & 1-(\mathfrak{s}_{aa}+\mathfrak{s}_{ab}+\mathfrak{s}_{a1}) \\ 0 & 0 & 0 & \mathfrak{s}_{aa}+\mathfrak{s}_{ab}+\mathfrak{s}_{a1} & 0 & 1-(\mathfrak{s}_{aa}+\mathfrak{s}_{ab}+\mathfrak{s}_{a1}) \\ 0 & 0 & 0 & 0 & \mathfrak{s}_{aa}+\mathfrak{s}_{ab}+\mathfrak{s}_{a1} & 1-(\mathfrak{s}_{aa}+\mathfrak{s}_{ab}+\mathfrak{s}_{a1}) \\ 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}$$

and, with a simplifying reparameterization

$$= \begin{pmatrix} \mathfrak{s}_{aa} & \mathfrak{s}_{ab} & \mathfrak{s}_{ab} & \mathfrak{s}_{11}-\mathfrak{s}_{aa}-\mathfrak{s}_{ab} & \mathfrak{s}_{11}-\mathfrak{s}_{aa}-\mathfrak{s}_{ab} & 1+\mathfrak{s}_{aa}-2\mathfrak{s}_{11} \\ 0 & \mathfrak{s}_{aa} & 0 & \mathfrak{s}_{11}-\mathfrak{s}_{aa} & 0 & 1-\mathfrak{s}_{11} \\ 0 & 0 & \mathfrak{s}_{aa} & 0 & \mathfrak{s}_{11}-\mathfrak{s}_{aa} & 1-\mathfrak{s}_{11} \\ 0 & 0 & 0 & \mathfrak{s}_{11} & 0 & 1-\mathfrak{s}_{11} \\ 0 & 0 & 0 & 0 & \mathfrak{s}_{11} & 1-\mathfrak{s}_{11} \\ 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}.$$

Further, suppose the shocks arrive as a Poisson process with parameter λ and the system is to be subject to periodic inspection every i time units with failures only observed upon inspection. The transition matrix for the MDPH distribution of failures observed at inspection is given by:

$$\mathfrak{S}^* = \begin{pmatrix} \mathfrak{s}^*_{aa} & \lambda i \mathfrak{s}_{ab}\mathfrak{s}^*_{aa} & \lambda i \mathfrak{s}_{ab}\mathfrak{s}^*_{aa} & \mathfrak{s}^*_{11}-\mathfrak{s}^*_{ab}-\mathfrak{s}^*_{aa} & \mathfrak{s}^*_{11}-\mathfrak{s}^*_{ab}-\mathfrak{s}^*_{aa} & 1+\mathfrak{s}^*_{aa}-2\mathfrak{s}^*_{11} \\ 0 & \mathfrak{s}^*_{aa} & 0 & \mathfrak{s}^*_{11}-\mathfrak{s}^*_{aa} & 0 & 1-\mathfrak{s}^*_{11} \\ 0 & 0 & \mathfrak{s}^*_{aa} & 0 & \mathfrak{s}^*_{11}-\mathfrak{s}^*_{aa} & 1-\mathfrak{s}^*_{11} \\ 0 & 0 & 0 & \mathfrak{s}^*_{11} & 0 & 1-\mathfrak{s}^*_{11} \\ 0 & 0 & 0 & 0 & \mathfrak{s}^*_{11} & 1-\mathfrak{s}^*_{11} \\ 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix},$$

where $\mathfrak{s}_{aa}^* = e^{\lambda i(\mathfrak{s}_{aa}-1)}$, and $\mathfrak{s}_{11}^* = e^{\lambda i(\mathfrak{s}_{aa}+\mathfrak{s}_{ab}+\mathfrak{s}_{a1}-1)}$.

Suppose that a slightly modified version of the system is available with identical characteristics except that a partially damaged component can be repaired for no additional cost as soon as it is observed. This implies that the transitions out of s_b and s_c are the same as those out of s_a. The corresponding change to the transition matrix results in:

$$\begin{pmatrix} \mathfrak{s}_{aa}^* & \lambda i\mathfrak{s}_{ab}\mathfrak{s}_{aa}^* & \lambda i\mathfrak{s}_{ab}\mathfrak{s}_{aa}^* & \mathfrak{s}_{11}^* - \mathfrak{s}_{ab}^* - \mathfrak{s}_{aa}^* & \mathfrak{s}_{11}^* - \mathfrak{s}_{ab}^* - \mathfrak{s}_{aa}^* & 1 + \mathfrak{s}_{aa}^* - 2\mathfrak{s}_{11}^* \\ \mathfrak{s}_{aa}^* & \lambda i\mathfrak{s}_{ab}\mathfrak{s}_{aa}^* & \lambda i\mathfrak{s}_{ab}\mathfrak{s}_{aa}^* & \mathfrak{s}_{11}^* - \mathfrak{s}_{ab}^* - \mathfrak{s}_{aa}^* & \mathfrak{s}_{11}^* - \mathfrak{s}_{ab}^* - \mathfrak{s}_{aa}^* & 1 + \mathfrak{s}_{aa}^* - 2\mathfrak{s}_{11}^* \\ \mathfrak{s}_{aa}^* & \lambda i\mathfrak{s}_{ab}\mathfrak{s}_{aa}^* & \lambda i\mathfrak{s}_{ab}\mathfrak{s}_{aa}^* & \mathfrak{s}_{11}^* - \mathfrak{s}_{ab}^* - \mathfrak{s}_{aa}^* & \mathfrak{s}_{11}^* - \mathfrak{s}_{ab}^* - \mathfrak{s}_{aa}^* & 1 + \mathfrak{s}_{aa}^* - 2\mathfrak{s}_{11}^* \\ 0 & 0 & 0 & \mathfrak{s}_{11}^* & 0 & 1 - \mathfrak{s}_{11}^* \\ 0 & 0 & 0 & 0 & \mathfrak{s}_{11}^* & 1 - \mathfrak{s}_{11}^* \\ 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix},$$

which reduces to

$$\begin{pmatrix} \mathfrak{s}_{aa}^* + 2\lambda i\mathfrak{s}_{ab}\mathfrak{s}_{aa}^* & \mathfrak{s}_{11}^* - \mathfrak{s}_{ab}^* - \mathfrak{s}_{aa}^* & \mathfrak{s}_{11}^* - \mathfrak{s}_{ab}^* - \mathfrak{s}_{aa}^* & 1 + \mathfrak{s}_{aa}^* - 2\mathfrak{s}_{11}^* \\ 0 & \mathfrak{s}_{11}^* & 0 & 1 - \mathfrak{s}_{11}^* \\ 0 & 0 & \mathfrak{s}_{11}^* & 1 - \mathfrak{s}_{11}^* \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

A natural question is which version of the system provides the cheapest long term cost. Suppose that a system is replaced at the first observed failure of a component and system downtime occurs if both components have failed with an associated cost similar to the previous example. The total cost of a unit is given by

$$C(\mathrm{S}_1^*, \mathrm{S}_2^*) = \begin{cases} C_{in} \min\{\mathrm{S}_1^*, \mathrm{S}_2^*\} + C_r & \mathrm{S}_1^* \neq \mathrm{S}_2^* \\ \\ C_{in} \min\{\mathrm{S}_1^*, \mathrm{S}_2^*\} + i \cdot C_{dt} + C_r & \mathrm{S}_1^* = \mathrm{S}_2^* \end{cases}$$

where C_r depends on whether the system is repairable or not.

Since the repairable system has a simple lifetime distribution, the expected cost and expected lifetime are similar to the previous example with (taking $\lambda = 1$)

$$\begin{aligned} E[C] &= \left(\frac{1}{1 - \mathfrak{s}_{00}^*}\right) C_{in} + \frac{\mathfrak{s}_{03}^*}{1 - \mathfrak{s}_{00}^*} (i \cdot C_{dt} + C_r) + \left(1 - \frac{\mathfrak{s}_{03}^*}{1 - \mathfrak{s}_{00}^*}\right) C_r \\ &= \frac{C_{in}}{1 - (\mathfrak{s}_{aa}^* + 2i\mathfrak{s}_{ab}\mathfrak{s}_{aa}^*)} + i \cdot C_{dt} \frac{1 + \mathfrak{s}_{aa}^* - 2\mathfrak{s}_{11}^*}{1 - (\mathfrak{s}_{aa}^* + 2i\mathfrak{s}_{ab}\mathfrak{s}_{aa}^*)} + C_r. \end{aligned}$$

Likewise, the lifetime and expected lifetime of a repairable system are

$$L(\mathrm{S}_1^*, \mathrm{S}_2^*) = i \min\{\mathrm{S}_1^*, \mathrm{S}_2^*\}$$

and

$$\begin{aligned} E[L] &= \frac{i}{1 - \mathfrak{s}_{00}^*} \\ &= \frac{i}{1 - (\mathfrak{s}_{aa}^* + 2i\mathfrak{s}_{ab}\mathfrak{s}_{aa}^*)}. \end{aligned}$$

So

$$\frac{E[C]}{E[L]} = \frac{C_{in} + i \cdot C_{dt}\left(1 + \mathfrak{s}_{aa}^* - 2\mathfrak{s}_{11}^*\right) + C_r\left(1 - (\mathfrak{s}_{aa}^* + 2i\mathfrak{s}_{ab}\mathfrak{s}_{aa}^*)\right)}{i}.$$

Taking

$$\mathfrak{S} = \begin{pmatrix} \frac{9}{10} & \frac{3}{100} & \frac{3}{100} & \frac{3}{200} & \frac{3}{200} & \frac{1}{100} \\ 0 & \frac{9}{10} & 0 & \frac{9}{200} & 0 & \frac{11}{200} \\ 0 & 0 & \frac{9}{10} & 0 & \frac{9}{200} & \frac{11}{200} \\ 0 & 0 & 0 & \frac{189}{200} & 0 & \frac{11}{200} \\ 0 & 0 & 0 & 0 & \frac{189}{200} & \frac{11}{200} \\ 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix},$$

and $C_{in} = 1$, $C_{dt} = 10$, and $C_r = 100$ for the repairable system, the minimum average cost per unit time occurs with inspection internal $i \approx 2.13$ and the average cost per unit time approximately 4.92.

Using the same parameters as above with the exception of a lower replacement cost for the non-repairable system of $C_r = 50$. The transition matrix for this MDPH with inspection interval i is

$$\begin{pmatrix} e^{-\frac{i}{10}} & \frac{3i}{100}e^{-\frac{i}{10}} & \frac{3i}{100}e^{-\frac{i}{10}} & e^{-\frac{11i}{200}} - \frac{3i}{100}e^{-\frac{i}{10}} - e^{-\frac{i}{10}} & e^{-\frac{11i}{200}} - \frac{3i}{100}e^{-\frac{i}{10}} - e^{-\frac{i}{10}} & 1 - 2e^{-\frac{11i}{200}} + e^{-\frac{i}{10}} \\ 0 & e^{-\frac{i}{10}} & 0 & e^{-\frac{11i}{200}} - e^{\frac{i}{10}} & 0 & 1 - e^{-\frac{11i}{200}} \\ 0 & 0 & e^{-\frac{i}{10}} & 0 & e^{-\frac{11i}{200}} - e^{\frac{i}{10}} & 1 - e^{-\frac{11i}{200}} \\ 0 & 0 & 0 & e^{-\frac{11i}{200}} & 0 & 1 - e^{-\frac{11i}{200}} \\ 0 & 0 & 0 & 0 & e^{-\frac{11i}{200}} & 1 - e^{-\frac{11i}{200}} \\ 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}.$$

This system is not as straightforward to analyze in closed form, so we ran simulations to estimate the optimum inspection period and the average cost per unit time. First several simulations were run to find a range of values for inspection period that give the lowest expected cost per unit time. After finding the optimum period to be between 2.0 and 3.0, we performed simulation runs of two million systems for the eleven inspection intervals $i = 2 + .1j$, $j = 0, \cdots, 10$. We estimated average cost per unit time to be to be approximately 3.9. Therefore, at the given prices the non-repairable system is cheaper to own and operate than the repairable system.

5.2.2 Coherent Reliability Systems

Many realistic engineering applications can be described using coherent systems ([3] 1981). In this section, we show that any coherent system with components whose lifetime vector is distributed MDPH has a lifetime that is distributed DPH. We also obtain an explicit expression for the reliability of coherent systems with dependent components.

Consider a system comprising m components. We assign a binary variable x_i to component i:

$$x_i = \begin{cases} 1 & \text{if component } i \text{ is functioning} \\ 0 & \text{if component } i \text{ has failed.} \end{cases}$$

Similarly, the binary variable Φ indicates the state of the system:

$$\Phi = \begin{cases} 1 & \text{if the system is functioning} \\ 0 & \text{if the system has failed.} \end{cases}$$

We assume that

$$\Phi = \Phi(\mathbf{x}),$$

where $\mathbf{x} = (x_1, x_2, \ldots, x_m)$. That is, the state of the system is completely determined by the states of the components. We refer to the function $\Phi(\mathbf{x})$ as the structure function of the system.

Definition 12 *A system is said to be coherent if following conditions hold.*

1. *The structure function Φ is non-decreasing in each argument.*
2. *Each component is relevant, i.e., there exists at least one vector $(\cdot_i, \mathbf{x})$ such that $\Phi(1_i, \mathbf{x}) = 1$ and $\Phi(0_i, \mathbf{x}) = 0$.*

Example 15 A system that is functioning if and only if each component is functioning is called a series system. The structure function for this system is given by

$$\Phi(\mathbf{x}) = \prod_{i=1}^{m} x_i.$$

A system that is functioning if and only if at least one component is functioning is called a parallel system. The structure function for this system is given by

$$\Phi(\mathbf{x}) = 1 - \prod_{i=1}^{m}(1 - x_i).$$

In general, a system that is functioning if and only if at least k out of n components are functioning is called a k-out-of-n system. The structure function for this system is given by

$$\Phi(\mathbf{x}) = \begin{cases} 1 & \text{if } \sum_{i=1}^{m} x_i \geq k \\ 0 & \text{if } \sum_{i=1}^{m} x_i < k. \end{cases}$$

A system that is functioning if and only if at least k consecutive components are functioning is called a consecutive k-out-of-n system. The structure function for this system is given by

$$\Phi(\mathbf{x}) = \begin{cases} 1 & \text{if there exists a } l \text{ such that } \sum_{i=l+1}^{l+k} x_i \geq k \\ 0 & \text{otherwise.} \end{cases}$$

All of these systems are coherent systems.

Let $\mathbf{S}_\Phi$ be a system with m components and structure function Φ. Suppose $(\mathrm{S}_1, \ldots, \mathrm{S}_m)$ is a random vector of component lifetimes that has has a MDPH distribution with underlying Markov chain $\mathbb{S} = \{S, \mathbf{\Sigma}, \mathfrak{S}, \sigma\}$ $= \{\mathbb{S}^n, n \geq 0\}$. Define

$$X_i^n = I_{\Sigma_i^c}(\mathbb{S}^n),\ i = 1, \ldots, m,$$

where I_A denotes the indicator function of the set A. The reliability of $\mathbf{S}_\Phi$ at time n is given by

$$R_\Phi(n) = \mathbf{P}\{\Phi(X_1^n, \ldots, X_m^n) = 1\}.$$

Let $\tau_\Phi(\mathrm{S}_1, \ldots, \mathrm{S}_m)$ be the lifetime of the system $\mathbf{S}_\Phi$. Then τ_Φ is called the life function associated with the structure function Φ. It follows that

$$R_\Phi(n) = \mathbf{P}\{\Phi(X_1^n, \ldots, X_m^n) = 1\} = \mathbf{P}\{\tau_\Phi(\mathrm{S}_1, \ldots, \mathrm{S}_m) \geq n\}.$$

The system mean time to failure is given by $E[\tau_\Phi(\mathrm{S}_1, \ldots, \mathrm{S}_m)]$. The following result provides a useful tool for calculating the reliability and mean time to failure of coherent systems.

Theorem 22 *Let $\mathbf{S}_{\Phi_1}, \cdots, \mathbf{S}_{\Phi_k}$ be coherent systems each having identically MDPH distributed component lifetime vectors $\mathbf{S} = (\mathrm{S}_1, \ldots, \mathrm{S}_m)$. If $\tau_1, \ldots, \tau_k$ are the life functions associated with structure functions $\Phi_1, \ldots, \Phi_k$, respectively, then $(\tau_1(\mathbf{S}), \ldots, \tau_k(\mathbf{S}))$ is a k-dimensional MDPH random vector.*

Proof:

Let $\mathbb{S} = \{S, \mathbf{\Sigma}, \mathfrak{S}, \sigma\} = \{\mathbb{S}^n, n \geq 0\}$ be the underlying Markov chain for $\mathbf{S}$. Define

$$\Lambda_j = \{s \mid \Phi_j(I_{\Sigma_1^c}(s), \ldots, I_{\Sigma_m^c}(s)) = 0\},\ j = 1, \ldots, k.$$

Since $\Sigma_1, \ldots, \Sigma_m$ are all stochastically closed and Φ is non-decreasing in each argument, the sets $\Lambda_1, \ldots, \Lambda_k$ are k stochastically closed subsets of S for the same Markov chain $\{\mathbb{S}^n, n \geq 0\}$, and

$$\tau_j(\mathbf{S}) = \min\{n \mid \mathbb{S}^n \in \Lambda_j\},\ j = 1, \ldots, k.$$

Thus, $(\tau_1(\mathbf{S}), \ldots, \tau_k(\mathbf{S}))$ is a k-dimensional MDPH random vector. □

Corollary 10 *Let $\mathbf{S}_\Phi$ be a coherent system with m components having lifetime vector $(\mathrm{S}_1, \ldots, \mathrm{S}_m)$ with a MDPH distribution with underlying Markov chain $\mathbb{S} = \{S, \mathbf{\Sigma}, \mathfrak{S}, \sigma\}$. If τ_Φ is the life function associated with the structure function Φ, then $\tau_\Phi(\mathrm{S}_1, \ldots, \mathrm{S}_m)$ has a univariate DPH distribution with underlying Markov chain $\{S, \Lambda, \mathfrak{S}, \sigma\}$, where $\Lambda = \{s \mid \Phi(I_{\Sigma_1^c}(s), \ldots, I_{\Sigma_m^c}(s)) = 0\} \subseteq S$.*

Thus the reliability of the coherent system $\mathbf{S}_\Phi$ with component lifetimes $(\mathrm{S}_1, \ldots, \mathrm{S}_m)$ at time n is given by

$$R_\Phi(n) = \sigma \mathfrak{S}^n \mathbf{I}_{S-\Lambda} \underline{\mathbf{1}}.$$

The mean time to failure of the coherent system $\mathbf{S}_\Phi$ is given by

$$E[\tau_\Phi(\mathrm{S}_1, \ldots, \mathrm{S}_m)] = \sigma (I - \mathfrak{S})^{-1} \mathbf{I}_{S-\Lambda} \underline{\mathbf{1}}.$$

For a given coherent system with structure function Φ and a component lifetime vector $\mathbf{S}$ that has a MDPH distribution, the reliability and mean time to failure can be easily calculated using these formulas. This is a substantial improvement over the formulas used in engineering applications where only the independent case has been considered.

Appendix A

Computer Code

A.1 MDPH Data Structures

Each MDPH is represented as a list with four elements; the transition matrix, dimension, absorbing classes (a set of logical indicator vectors), and the initial probability vector.

A.1.1 Bivariate MDPH Distribution

```
mdph2d<-list(trmat=c(),dim=NA, classes=c(), init=c())
mdph2d$trmat<-matrix(c(
                        c(0.2, 0.30, 0.45, 0.05),
                        c( 0 , 0.35,   0 , 0.65),
                        c( 0 ,  0  , 0.40, 0.60),
                        c( 0 ,  0  ,   0 ,  1  )
                        ), nrow=4, ncol=4, byrow=T)
mdph2d$dim<-2
mdph2d$init<-matrix(c(1,0,0,0), nrow=1, ncol=4, byrow=T)
mdph2d$classes<-array(c(F,T, F, T, F,F,T,T), dim=c(4,2))
```

A.1.2 Trivariate MDPH Distribution

```
mdph3d<-list(trmat=c(),dim=NA, classes=c(), init=c())
```

```
mdph3d$trmat<-matrix(c(
                        c(0.15, 0.05, 0.30, 0.05, 0.10, 0.20, 0.10, 0.05),
                        c(  0 , 0.30,   0 , 0.25,   0 , 0.20,   0 , 0.25),
                        c(  0,    0 , 0.25, 0.25,   0 ,   0 , 0.25, 0.25),
                        c(  0 ,   0 ,   0 , 0.75,   0 ,   0 ,   0 , 0.25),
                        c(  0 ,   0 ,   0 ,   0 , 0.15, 0.25, 0.15, 0.45),
                        c(  0 ,   0 ,   0 ,   0 ,   0 , 0.75,   0 , 0.25),
                        c(  0 ,   0 ,   0 ,   0 ,   0 ,   0 , 0.40, 0.60),
                        c(  0 ,   0 ,   0 ,   0 ,   0 ,   0 ,   0 ,   1 ),
                        ),
            nrow=8, ncol=8, byrow=T)
mdph3d$dim<-3
mdph3d$init<-matrix(c(1,0,0,0,0,0,0,0), nrow=1, ncol=8, byrow=T)
mdph3d$classes<-array(c(
                        c(F,T,F,T,F,T,F,T),
                        c(F,F,T,T,F,F,T,T),
                        c(F,F,F,F,T,T,T,T)
                        ),
                    dim=c(8,3))
```

A.2 Functions

A.2.1 MDPH Probabilities

This function implements the ideas in Theorem 20 to calculate the probability of a MDPH being in a region defined by a given set of conditions.

```
# mdph.prob(distribution, event.times, conditions)
# distribution: list with components for MDPH
# event.times: the n's for the conditions in order of variables
# cond: The conditions (<, <=, =, >=, >) associated with event.times
```

```
mdph.prob<-function(distribution=NA, variables=c(), event.times=c(), cond=c("")) {
#  Verify that transition matrix is valid
        if(length(which(rowSums(distribution$trmat)!=1))) {
              return("Error! Rows of Transition Matrix Do Not Sum to 1")
        }
#  Verify validity of conditions
        if(length(which(cond!="<" & cond!="<=" & cond!=">=" & cond!="=" & cond!=">"))) {
              return("Error! Invalid Condition")
        }
#  Verify validity of event.times
        if(length(which(event.times<0))) { return("Error! Negative Event Time") }

#  Fix conditions and event.times so all conditions are <= or >
#               <n becomes <= n-1
#               >=n becomes > n-1
#               =n becomes >n-1 & <= n
    j=1
    orig.cond=cond
    new.cond=c()
    new.variables=c()
    new.event.times=c()
    for(i in 1:length(cond)) {
        if(cond[i] == "<") {cond[i]<-"<="; event.times[i]<-event.times[i]-1;}
        if(cond[i] == ">=") {cond[i]<-">"; event.times[i]<-event.times[i]-1;}
        if(cond[i] == "=") {
            cond[i]<-"<="; event.times[i]=event.times[i];
            new.variables[j]=variables[i]
            new.cond[j]=">";
            new.event.times[j] <- event.times[i]-1;
            j<-j+1
        }
```

```
    }
    if(j>1) {
        cond<-c(cond,new.cond)
        event.times<-c(event.times, new.event.times)
        variables<-c(variables, new.variables)
    }

#  <=n becomes < n for purposes of constructing diagonal
    for(i in 1:length(cond)) { if(cond[i]=="<=") cond[i]<-"<"}

#  Order n's
    order.times<-rank(event.times)

#  Loop over Product
#   Initialize loop variables
    temp.dist<-distribution$init
    N.current<-0
    for(i in unique(sort(order.times))) {
        N.old<-N.current
#     Find the event time
        N.current<-unique(event.times[order.times==i])
#     Multiply to event
        temp.dist<-temp.dist %*% matPower(distribution$trmat,N.current-N.old)
#     Build diagonal and multiply
        temp.dist<-temp.dist %*% build_diag(N.current, event.times, cond, variables,
                                            distribution$dim, distribution$classes)
    }
    return(sum(temp.dist))
}
```

A.2.2 Construct Diagonal Matrix

This function constructs the diagonal matrix $\mathbf{I}_{\Gamma_i}$ from Theorem 20

```
# N: time of the current event
# event.times: the n's for the conditions in order of variables
# cond: conditions (<, <=, =, >=, >) associated with event.times
# variables: variables associated with conditions and event.times
# dim: dimension of MDPH distribution of interest
# classvectors: indicator vectors for the absorbing sets in the MDPH distribution
build_diag<-function(N=c(),event.times=c(), cond=c(""), variables=c(), dim=c(),
                                                                  classvectors=NA)
{
#Construct logical vector by &'ing all conditions with current event time
        cond.vect<-sapply(N, paste, cond, event.times)
        truth.vect<-sapply(cond.vect, function(x) eval(parse(text=x)), USE.NAMES=F)
#Construct diagonal vector based on which conditions are satisfied
        diagonal<-c(rep(T,length(classvectors[,1])))
        for ( i in 1:length(cond)) {
            if ( !truth.vect[i] ) {
                if(cond[i]=="<") {
                    diagonal<-diagonal&classvectors[,variables[i]]
                }else {
                    diagonal<-diagonal&!classvectors[,variables[i]]
                }
            }
        }
        return(diag(diagonal))
}
```

A.2.3 Generate MDPH Random Vector

This function generates a random vector from a given MDPH distribution using the algorithm given in Section 5.1.2.

```
mdph.randgen<- function(distribution=NA) {
#  Initialize n vector to 0
    n<-c(rep(0,distribution$dim))
#  Sample from initial distribution multinomial
    i<-sample(length(distribution$init), 1, prob=distribution$init)
    n[!distribution$classes[i,]]<-n[!distribution$classes[i,]]+1
#  Loop until current value is in Sigma_delta
    while (i!=length(distribution$init)) {
#  Use value from previous sample to choose row for current sample
        i<-sample(length(distribution$init), 1, prob=distribution$trmat[i,])
#  Unless value of current sample is in Sigma_i add n_i<-n_i+1
        n[!distribution$classes[i,]]<-n[!distribution$classes[i,]]+1
    }
    mdphrv<-n
    return(mdphrv)
}
```

A.2.4 Generate MPH Random Vector

This function generates a random vector from an MPH vector given an underlying MDPH distribution and the rate parameter (λ) associating them.

```
mph.randgen<-function(distribution=NA, rate=1) {
        discrete.rv<-mdph.randgen(distribution)$n
        exp.rv<-rexp(max(discrete.rv), rate=rate)
        mph.rv<-sapply(discrete.rv, function(i) sum(exp.rv[1:i]))
        return(mph.rv)
}
```

A.3 Examples

A.3.1 Estimating Higher Joint Moments

This function is used in Example 10 to estimate the higher joint moments.

```
# distribution: list with elements of MDPH distribution
# order: A vector of orders for each variable
# runs: number of runs for Monte Carlo estimate of moment
mdph.moment.est<-function(distribution=NA, order=c(), runs=1000) {
if ( length(order)!=distribution$dim) { print("Error! Order vector not correct size") }
generatedmoments<-rep(0, runs)
for( i in 1:runs) {
    rv<-mdph.randgen(distribution)
    generatedmoments[i]<-prod(rv^order)
}
return(c(mean(generatedmoments), sd(generatedmoments)/sqrt(runs)))

}
```

A.3.2 Simulation

The following code was used to get Monte Carlo estimates for the long run average cost per unit time of the non-repairable system in Example 14.

```
# Define S* the bivariate inspection interval derived MDPH
  mdphStar<-list(trmat=c(),dim=NA, classes=c(), init=c())
  mdphStar$trmat<-build_trmat(1)
  mdphStar$dim<-2
  mdphStar$init<-matrix(c(1,0,0,0,0,0), nrow=1, ncol=6, byrow=T)
  mdphStar$classes<-array(c(F,F,F,T,F,T,  F,F,F,F,T,T), dim=c(6,2))

# Function to build the transition matrix for a given inspection interval length
# The rate parameter lambda is assumed to be 1.
```

```
  build_trmatstar<-function(i=1) {
          saa=exp(-i/10)
          sbb=exp(-i/10)
          scc=exp(-i/10)
          s11=exp(-11*i/200)
          s22=exp(-11*i/200)
          sab=(3/100)*i*exp(-i/10)
          sac=(3/100)*i*exp(-i/10)
          sa1=exp(-i*11/200) - exp(-i/10) * (3*i/100 + 1)
          sa2=exp(-i*11/200) - exp(-i/10) * (3*i/100 + 1)
          sb1=exp(-i*11/200) - exp(-i/10)
          sc2=exp(-i*11/200) - exp(-i/10)
          matstar<-matrix(c(
                          c(saa, sab, sac, sa1, sa2, 1-(saa+sab+sac+sa1+sa2)),
                          c(0,   sbb,   0, sb1,   0, 1-(sbb+sb1)            ),
                          c(0,   0,   scc,   0, sc2, 1-(scc+sc2)            ),
                          c(  0,   0,   0, s11,   0, 1-s11                  ),
                          c(  0,   0,   0,   0, s22, 1-s22                  ),
                          c(  0,   0,   0,   0,   0, 1                      )
                          ), nrow=6, ncol=6, byrow=T)
          return(matstar);
  }

# Define Cost Parameters
  Ci<-1; Cr<-50;  Cdt<-10;
# Set the number of units to simulate
  runs<-2000000
# Values of inspection interval to simulate
  intvalues<-seq(2.0,3, by=.1)
# Initialize vector where average costs will be saved for each interval length
  avecost<-c()
```

```
# Loop over Interval lengths
  for (interval in intvalues) {
#  Reset cost and time totals
      cost<-0
      time<-0
#  Build transition matrix for current interval length
      mdphStar$trmat<-build_trmatstar(interval)
#  Simulation loop
      for(j in 1:runs) {
              rv<-mdph.randgen(mdphStar)
              if( rv[2]==rv[1] ) {
                  cost<-cost+ Ci*min(rv) + interval*Cdt + Cr
              }else{
                  cost<-cost+ Ci*min(rv) + Cr
              }
              time<- time + interval*min(rv)
          }
#  Average cost estimate is total cost of all units over total lifetime of all units
      avecost<-c(avecost, cost/time)
  }

# Plot estimated average cost for each interval length
  x11()
  plot(intvalues,avecost, xlim=c(min(intvalues),max(intvalues)),
                          ylim=c(min(avecost)-.05,max(avecost)+.05),
                          col="black", type="p")
```

Bibliography

[1] Asmussen, S. (2000) *Ruin Probabilities.* World Scientific, Singapore.

[2] Assaf, D., Langberg, N., Savits, T., and Shaked, M. (1984), Multivariate phase-type distributions, *Operations Research* **32**, 688–701.

[3] Barlow, R. E. and Proschan, F. (1981), *Statistical Theory of Reliability and Life Testing*, To Begin With, Silver Spring, MD.

[4] Esary, J. D. and Proschan, F. (1970), A reliability bound for systems of maintained, independent components, *Journal of American Statistical Association*, **65**, 329-338.

[5] Erlang, A. K. (1917/1918), Solution of some problems in the theory of probabilities of significance in automatic telephone exchanges. *The Post Office Engineer's Journal* **10**, 189-197.

[6] Freund, J. E. (1961), A bivariate extension of the exponential distribution, *J. Am. Statist. Asso.* **56**, 971–977.

[7] Harris, T. E. (1977), A correlation inequality for Markov processes in partially ordered state spaces. *Ann. Prob.* **5**, 451-454.

[8] Joe, H. (1998), *Multivariate Models and Dependence Concepts*, Chapman and Hall/CRC.

[9] Kulkarni, V. G. (1989), A new class of multivariate phase type distributions, *Operations Research* **37**, 151–158.

[10] Li, H. (2003), Association of multivariate phase-type distributions with applications to shock models. *Statistics and Probability Letters*, **64**, 381-392.

[11] Li, H. and Xu, S. H. (2000), On the dependence structure and bounds of correlated parallel queues and their applications to synchronized stochastic system. *Journal of Applied Probability*, *37*, 1020-1043.

[12] Lindqvist, H. (1987), Monotone and associated Markov chains, with applications to reliability theory. *Journal of Applied Probability* **24**, 679-695.

[13] Lindqvist, H. (1988), Association of probability measures on partially ordered spaces, *Journal of Multivariate Analysis*, **26**, 111-131.

[14] Neuts, M. F. (1981), *Matrix-Geometric Solutions in Stochastic Models: An Algorithmic Approach*, The Johns Hopkins University Press, Baltimore.

[15] Marshall, A. W. and Olkin, I. (1967). A multivariate exponential distribution. *J. Amer. Statist. Assoc.* **62**, 30-44.

[16] R Development Core Team (2004), R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL http://www.R-project.org.

[17] Rolski, T., Schmidli, H., Schmidt, V., Teugels, J. (1999), *Stochastic Processes for Finance and Insurance.* Wiley, New York.

[18] Ross, S. M. (1983), *Stochastic Processes*, John Wiley & Sons.

[19] Szekli, R. (1995), *Stochastic Ordering and Dependence in Applied Probability*, Springer-Verlag.

[20] Tong, Y.L. (1980), *Probability Inequalities in Multivariate Distributions*, Academic Press, New York.