POSITIVE DEPENDENCE FROM AN OPERATIONAL POINT OF VIEW

Arne Bang Huseby
University of Oslo

Abstract

In the talk we will review different notions of positive dependence. We argue that dependence should be modeled in terms of operationally well-defined quantities rather than abstract parameters such as covariance or correlation coefficients. In particular we suggest an operational alternative to the concept of association. The approach is illustrated by considering some specialized examples.

Keywords: DEPENDENCE; ASSOCIATION; OPERATIONAL METHODS; SHOCK MODELS; MAINTENANCE MODELS

1. Introduction

Multivariate models based on some sort of positive dependence among the variables have found extensive application in reliability theory as well as other branches of statistics. A classical way of modeling positive dependence, especially in reliability theory, is by introducing the concept of association (Esary, Proschan and Walkup (1967)). This is defined as follows.

Definition 1. The random variables $X_1, \ldots, X_n$ are said to be associated random variables if:

$$(1) \qquad \text{Cov}[f(X_1, \ldots, X_n), g(X_1, \ldots, X_n)] \geq 0$$

for all functions f, g which are nondecreasing in each argument and such that the covariance exists. ❑

Based on the above notion many useful inequalities can be deduced. For details on this we refer to Barlow and Proschan (1981). See also Tong (1980).

If only the joint distribution of $X_1, \ldots, X_n$ is given, it is often very difficult to check the condition (1) directly. The usual approaches are to either use theorems specific to a particular family of distributions, (see e.g. Pitt (1982)), or to use sufficient conditions, (see Barlow and Proschan (1981)). This problem is even more difficult when the distribution is described numerically, such as observed contingency tables. For a discussion of this we refer to Sampson and Whitaker (1989).

From an operational point of view the concept of association is unsatisfactory, since the condition (1) is expressed in terms of abstract parameters, i.e., covariances, rather than operationally well-defined quantities. Furthermore association does not utilize specific knowledge about the physical nature of the objects we are modeling. Consequently, association is often just assessed without any further justification.

Most papers based on an operational approach to modeling, has only considered situations with replications of univariate variables. Although in the finite sample cases this approach leads to models with dependencies, this is *not* the kind of physical dependence on would like to incorporate. Thus, there is a need for some other type of methodology addressing this issue directly. The following definition is motivated by this:

Definition 2. The random variables $X_1,\ldots, X_n$ are said to be increasing (nondecreasing) functions of independent random variables (IFIV) if:

$$(2) \qquad (X_1,\ldots,X_n) \overset{D}{=} (f_1(\mathbf{U}),\ldots,f_n(\mathbf{U}))$$

for some suitable nondecreasing functions $f_1,\ldots,f_n$ and independent random variables $U_1,\ldots, U_N$. [$\mathbf{U} = (U_1,\ldots, U_N)$.] ❑

In a given practical situation, (2) may be a *strict functional relation* derived from some sort of physical insight, not just a way of representing the joint distribution. An IFIV model is considered *operational* if all the U_i's are physically meaningful quantities. We do not, however, require the U_i's to be *observable* in a strict sense.

An IFIV-model can be derived in a natural way from operationally well-defined quantities, i.e., $U_1,\ldots, U_N$. The usage of such models encourages use of physical knowledge, and can be applied together with indifference to derive multivariate distributions.

As we shall see in the next section the IFIV-class possesses more or less all the nice properties as the corresponding class based on association. Thus, the IFIV-class represents an attractive alternative to the classical theory.

Although Definition 2 provides the fundamental basis for our theory, it does not explain how to use this methodology in practice. As already stated, this methodology is based on incorporating physical knowledge. Thus, we will need to consider specific examples to illustrate its practical usefulness. Most of this paper will focus on one such example: a binary shock model. In particular we will explore how the physical knowledge built into this model, imposes restrictions on the model. Although the model has got a lot of flexibility, there are certain limitations one needs to take into account and understand. We would like to stress that when building a probability model in such a fashion, it is extremely important to understand such physical limitations.

2. Some Properties of the IFIV-class

Although the IFIV-class has not been studied as a "stand-alone" concept, the following result is well-known. (Esary, Proschan and Walkup (1967))

Theorem 3. If $X_1, \ldots, X_n$ are IFIV, then they are associated as well. ❑

However, it appears to be unknown whether or not this implication is *strict*. In other words, does there exist associated random variables $X_1, \ldots, X_n$ which do not admit an IFIV-representation? Do we actually *lose* anything (important) by considering the possibly smaller IFIV-class?

By Theorem 3, all results valid for associated random variables, will automatically hold for the IFIV-class as well. In fact many of these results are much easier to prove directly for the IFIV-class. In particular we have:

Theorem 4. The following statements are true:

1) Any subset of a set of IFIV-variables are a set of IFIV-variables.
2) The set consisting of a single random variable is IFIV.
3) If two sets of IFIV-variables are independent, then their union is a set of IFIV-variables.
4) Independent variables are IFIV.
5) Nondecreasing functions of IFIV-variables are IFIV-variables.

Proof: The proofs of the different statements in Theorem 6 are completely trivial. Some of the corresponding proofs for the class of associated random variables are far more complicated. ❑

The following result is only known to hold in the bivariate case for associated random variables: [See Barlow and Proschan (1981).]

Theorem 5. Assume that $X_1, \ldots, X_n$ are IFIV, and that $\text{Cov}(X_i, X_j)$ exists for all $i \neq j$. Then $X_1, \ldots, X_n$ are independent if and only if:

(3) $\qquad \text{Cov}(X_i, X_j) = 0, \;$ for all $i \neq j$.

Proof: It is obviously sufficient to prove that if $X_1, \ldots, X_n$ are dependent IFIV-variables, then there exist i and j such that $\text{Cov}(X_i, X_j) > 0$.

By the IFIV-property, we may write $X_i = f_i(\mathbf{U})$, $i = 1, \ldots, n$, where $\mathbf{U} = (U_1, \ldots, U_N)$, and $f_1, \ldots, f_n$ are nondecreasing functions. If $X_1, \ldots, X_n$ are dependent, then there must exist i, j and k such that X_i and X_j depend on U_k. From this it can be shown by conditioning on U_k and using Kimball's inequality [see Kimball (1951)] that $\mathrm{Cov}(X_i, X_j) > 0$. ❑

3. A binary shock model

In this section we shall focus on a model for positive dependence introduced by Boyles and Samaniego (1984). This model is restricted to binary variables and may be viewed as a discrete analog to the well-known model of Marshall and Olkin (1967). In Egeland and Huseby (1991) this model was used to study errors in reliability computation due to false independence assumptions. More recently, the model is used in Gåsemyr and Natvig (1994) to obtain improved bounds for system reliability.

As we shall see, in its most general form, the model contains a large number of parameters, and hence provides great modeling flexibility. However, the subtle physical knowledge being built into the model, implies that not *all* multivariate joint distributions can be represented in such a fashion. In fact the model class is a proper subset of the class models derived from association. In order to fully understand the limitations, we shall answer the following two questions:

1. When can a multivariate distribution of binary variables be represented by this model?
2. How should the parameters of this model be chosen in order to obtain such a representation?

Now, we turn to the model itself. Let $E = \{1,\ldots, n\}$ be a set of components in a reliability system. Each component is either functioning or failed, and we let $X_1,\ldots, X_n$ represent the states of these components. That is, we let:

$$(4) \qquad X_i = \begin{cases} 1 & \text{if component } i \text{ is functioning} \\ 0 & \text{otherwise} \end{cases}, \quad i = 1,\ldots, n.$$

We assume that the failures of the components are caused by different types of "*shocks*" striking single components or groups of components. More precisely, we assume that for each nonempty set $A \subseteq E$, there exists a possible shock which, if it occurs, kills all the components in the set A and these *alone*.

In order to describe the "shock status" of a set $A \subseteq E$ we introduce:

$$(5) \qquad Y_A = \begin{cases} 1 & \text{if the shock striking the set } A \text{ has not yet occured} \\ 0 & \text{otherwise} \end{cases}$$

In terms of the Y_A's we may now express the X_i's as follows:

(6) $$X_i = \prod_{A:i\in A} Y_A, \; i = 1,\dots, n.$$

We assume that the Y_A's are independent and that $\Pr(Y_A = 1) = \pi(A)$. Furthermore, let $P(\varnothing) = 1$ and:

(7) $$P(A) = \Pr\left(\prod_{i\in A} X_i = 1\right), \quad \varnothing \subset A \subseteq E.$$

By Möbius inversion (see e.g., Welsh (1976)) it is easily obtained that:

(8) $$\Pr(X_1 = x_1 \cap \cdots \cap X_n = x_n) = \sum_{B\subseteq E\setminus A(\mathbf{x})} (-1)^{|B|} P(A(\mathbf{x}) \cup B)$$

where $\mathbf{x} = (x_1,\dots, x_n)$ and $A(\mathbf{x}) = \{i : x_i = 1\}$.

Hence, we see that the simultaneous distribution of the X_i's is determined by the $P(A)$'s. Moreover, we have the following relation between the $\pi(A)$'s and the $P(A)$'s:

(9) $$P(A) = \prod_{B\,:\,(B\cap A)\neq\varnothing} \pi(B), \qquad \varnothing \subset A \subseteq E.$$

Using Möbius inversion methods, one can prove the following result:

Theorem 6. Assume that (9) holds. Then the $\pi(A)$'s are given by:

(10) $$\pi(A) = \exp\left\{-\sum_{D\subseteq A} (-1)^{|D|} L((E\setminus A)\cup D)\right\}, \varnothing \subset A \subseteq E.$$

where $L(A) = \ln P(A)$, $\varnothing \subset A \subseteq E$.

Proof: We skip the details here. ❑

Note: To avoid trivial cases we assume that $\Pr(X_i = 1) > 0$, $i = 1,\dots, n$. This implies that $\pi(A) > 0$, for all A, and hence also $P(A) > 0$, for all A. Thus, $L(A)$ exists for all A.

Using Theorem 6 we get the following necessary and sufficient condition for a joint distribution to be representable by a shock model:

Theorem 7. The simultaneous distribution of the binary variables $X_1,\dots, X_n$ may be represented by a shock model if and only if:

(11) $$\sum_{D \subseteq A} (-1)^{|D|} L((E \setminus A) \cup D) \geq 0, \varnothing \subset A \subseteq E.$$

Proof: The result is a direct consequence of (10). ❑

Example 8. Let X_1 and X_2 be two binary variables. By Theorem 7 we get the following three conditions for when their joint distribution may be represented by a shock model:

(12) $L(\{1\}) - L(\{1, 2\}) \geq 0$ (trivially true)
(13) $L(\{2\}) - L(\{1, 2\}) \geq 0$ (trivially true)
(14) $L(\{1, 2\}) - L(\{1\}) - L(\{2\}) \geq 0$

The only nontrivial condition is equivalent to:

(15) $\text{Cov}(X_1, X_2) \geq 0$

Thus, in the bivariate case everything is reduced to a simple condition on the covariance between the two variables. ❑

Example 9. Let X_1, X_2 and X_3 be three binary variables. The nontrivial conditions for the existence of a shock model representation, as derived from Theorem 7 are the following:

(16) $L(\{1\}) - L(\{1, 2\}) - L(\{1, 3\}) + L(\{1, 2, 3\}) \geq 0$

(17) $L(\{2\}) - L(\{1, 2\}) - L(\{2, 3\}) + L(\{1, 2, 3\}) \geq 0$

(18) $L(\{3\}) - L(\{1, 3\}) - L(\{2, 3\}) + L(\{1, 2, 3\}) \geq 0$

(19) $$-L(\{1\}) - L(\{2\}) - L(\{3\}) + L(\{1, 2\}) + L(\{1, 3\}) + L(\{2, 3\}) - L(\{1, 2, 3\}) \geq 0$$

We may interpret the first three conditions, (16-18) in terms of conditional probabilities e.g. as follows:

(20) $\Pr(X_k = 1 \mid X_i = 1, X_j = 1) \geq \Pr(X_k = 1 \mid X_i = 1),$

where $\{i, j, k\} = \{1, 2, 3\}$. Similarly, (19) may be interpreted as:

(21) $$\Pr(X_k = 1 \mid X_i = 1) \cdot \Pr(X_k = 1 \mid X_j = 1) \geq \Pr(X_k = 1 \mid X_i = 1, X_j = 1) \cdot \Pr(X_k = 1),$$

where again $\{i, j, k\} = \{1, 2, 3\}$.

Note: Letting $\{i, j, k\}$ run through all permutations of the set $\{1, 2, 3\}$, 12 apparently different inequalities can be derived from (20) and (21). However, these relations will contain a lot of redundancy. Especially the 6 inequalities derived from (21) will all be equivalent. ❑

4. Exchangeable variables.

Of particular interest is situations where the binary variables, $X_1,\ldots, X_n$ are exchangeable, i.e., when the distribution of $X_1,\ldots, X_n$ is invariant with respect to permutations. In this case $P(A)$ depends only on $|A|$, for each $A \subseteq E$. Thus, we may introduce the following simplified notation:

(22) $\quad P_i = P(\{1,\ldots, i\})\,,\ L_i = \ln(P_i)\,,\ i = 1,\ldots, n.$

In particular we define $P_0 = P(\varnothing) = 1$ and thus $L_0 = 0$. Since $P(A)$ depends only on $|A|$, we of course get that:

(23) $\quad P(A) = P_{|A|}$ and $L(A) = L_{|A|}$, for all $A \subseteq E$.

It is easily seen that when $X_1,\ldots, X_n$ are exchangeable, also $\pi(A)$ depends only on $|A|$, for each $A \subseteq E$. We may thus introduce:

(24) $\quad \pi_i = \pi(\{1,\ldots, i\}), i = 1,\ldots, n.$

and as above we get that:

(25) $\quad \pi(A) = \pi_{|A|}$, for all $A \subseteq E$.

The next result expresses the π_i's as well the condition (11) in terms of the L_i's:

Theorem 10. The joint distribution of the exchangeable binary variables $X_1,\ldots, X_n$ may be represented by a shock model if and only if:

(26) $$\sum_{j=0}^{k} (-1)^j \binom{k}{j} L_{n-k+j} \geq 0\,,\ 0 < k \leq n.$$

Moreover:

(27) $$\pi_k = \exp\left\{-\sum_{j=0}^{k}(-1)^j \binom{k}{j} L_{n-k+j}\right\}, \quad 0 < k \leq n.$$

Proof: The result is a direct consequence of (11). ❑

It is a well-known fact that if $X_1,\ldots, X_n$ can be embedded in an *infinite* sequence of exchangeable binary variables, then the distribution of $X_1,\ldots, X_n$ admits the following representation:

(28) $$\Pr\left(\bigcap_{i=1}^{n} X_i = x_i\right) = \int_0^1 \theta^{\sum_{i=1}^n x_i}(1-\theta)^{n-\sum_{i=1}^n x_i} F(d\theta)$$

for some suitable distribution F on [0, 1]. That is, the X_i's are conditionally independent given a "parameter" θ. The parameter θ can be interpreted as the limiting frequency of 1's in the infinite sequence.

For general random variables exchangeability does not imply that the variables are IFIV or associated. The following example shows this:

Example 11. Let $Y_1,\ldots, Y_n$ be independent and identically distributed with zero mean and positive variance. Let U be independent of the Y_i's and have positive variance, and define:

(29) $$X_i = Y_i \cdot U, \qquad i = 1,\ldots, n.$$

Then it is easy to see that $X_1,\ldots, X_n$ can be embedded into an infinite sequence of exchangeable variables (simply by extending the sequence of Y_i's). Since $Y_1,\ldots, Y_n$ and U are nondegenerated, it is clear that the X_i's are dependent. However $\text{Cov}(X_i, X_j) = 0$, for all $i \neq j$, so by Theorem 5, $X_1,\ldots, X_n$ cannot be IFIV. In fact they are not even associated. ❑

However, in the binary case we get the following nice result:

Theorem 12. Let $X_1,\ldots, X_n$ be binary variables which can be embedded in an infinite sequence of exchangeable variables. Then $X_1,\ldots, X_n$ are IFIV.

Proof: Assume that $X_1,\ldots, X_n$ are binary variables which can be embedded in an infinite sequence of exchangeable variables. Thus, the joint distribution of the X_i's admit the representation (28) for some distribution F on [0, 1]. Now let $Y_1,\ldots, Y_n$ be independent and uniformly distributed on [0, 1], and let U have the distribution F and be independent of the Y_i's. Then it is easily seen that $X_1,\ldots, X_n$ satisfy:

$$(30) \qquad (X_1,\ldots,X_n) \stackrel{D}{=} \big(\mathrm{I}(Y_1+U>1),\ldots,\mathrm{I}(Y_n+U>1)\big),$$

and hence it is follows that $X_1,\ldots, X_n$ are IFIV. ❑

Now, if any IFIV-distribution could be represented by a shock model, then clearly Theorem 12 would imply that infinite exchangeability was a sufficient condition for shock model representability. However, as we shall see, this is not true. Before we show this, we shall restate Theorem 10 in terms of the moments of the θ-parameter.

From (28) it follows that under the infinite exchangeability assumption, we may write:

$$(31) \qquad P_i = \mathrm{E}(\theta^i)\,,\ \ i = 0, 1,\ldots, n,$$

where we always define $\mathrm{E}(\theta^0) = P_0 = 1$. Thus, we have arrived at the following result:

Theorem 13. Let $X_1,\ldots, X_n$ be binary variables which can be embedded in an infinite sequence of exchangeable variables, and let θ denote the limiting frequency of 1's in the sequence. Then the joint distribution of $X_1,\ldots, X_n$ may be represented by a shock model if and only if:

$$(32) \qquad \sum_{j=0}^{k}(-1)^j\binom{k}{j}\ln\left[\mathrm{E}\left(\theta^{n-k+j}\right)\right] \geq 0\,,\ \ 0 < k \leq n.$$

Moreover,

$$(33) \qquad \pi_k = \exp\left\{-\sum_{j=0}^{k}(-1)^j\binom{k}{j}\ln\left[\mathrm{E}\left(\theta^{n-k+j}\right)\right]\right\},\ \ 0 < k \leq n.$$

Proof: The result follows by the above discussion ❑

Before illustrating this result with some examples, we note that for $k = 1$ and 2 the inequality in (32) can be written respectively as:

$$(34) \qquad \mathrm{E}\left(\theta^{n-1}\right) \geq \mathrm{E}\left(\theta^{n}\right)$$

$$(35) \qquad \mathrm{E}\left(\theta^{n}\right)\mathrm{E}\left(\theta^{n-2}\right) \geq \left[\mathrm{E}\left(\theta^{n-1}\right)\right]^2$$

Since $\theta \in [0, 1]$, (34) is trivially true. Moreover, by a moment inequality for nonnegative random variables given in Tong (1977) and Proschan and Sethuraman (1974), it follows that (35) is true also. However, as we shall see, for $k > 2$ the conditions are not trivial.

Example 14. Let $X_1, \ldots, X_n$ be binary variables which can be embedded in an infinite sequence of exchangeable variables, and let θ denote the limiting frequency of 1's in the sequence. Assume that the uncertainty about θ can be described by a beta-distribution with parameters α and β. That is, θ has the following density:

$$(36) \qquad f(\theta) = \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} \theta^{\alpha-1}(1-\theta)^{\beta-1}, \quad 0 \leq \theta \leq 1.$$

where α and β are some suitable positive numbers. In this case one easily calculates that:

$$(37) \qquad \mathrm{E}\left(\theta^i\right) = \frac{\Gamma(\alpha+\beta)\Gamma(\alpha+i)}{\Gamma(\alpha+\beta+i)\Gamma(\alpha)}, \quad i = 0, 1, \ldots, n.$$

Hence, the conditions (32) can be written in terms of α and β as:

$$(38) \qquad \sum_{j=0}^{k} (-1)^j \binom{k}{j} \ln\left[\frac{\Gamma(\alpha+\beta)\Gamma(\alpha+n-k+j)}{\Gamma(\alpha)\Gamma(\alpha+\beta+n-k+j)}\right] \geq 0, \quad 0 < k \leq n.$$

Applying the following identity:

$$(39) \qquad \sum_{j=0}^{k} (-1)^j \binom{k}{j} = 0, \quad \text{for } k = 1, 2, \ldots$$

it follows that the two constant factors $\Gamma(\alpha)$ and $\Gamma(\alpha+\beta)$ vanish, and thus (38) can be rewritten as:

$$(40) \qquad \sum_{j=0}^{k} (-1)^j \binom{k}{j} \left[\ln[\Gamma(\alpha+n-k+j)] - \ln[\Gamma(\alpha+\beta+n-k+j)]\right] \geq 0, \quad 0<k\leq n.$$

More compactly we can rewrite (40) as:

$$(41) \qquad \Psi_k(\alpha + n - k) \geq \Psi_k(\alpha + \beta + n - k), \quad 0 < k \leq n.$$

where we have introduced:

$$(42) \qquad \Psi_k(a) = \sum_{j=0}^{k} \binom{k}{j} (-1)^j \ln[\Gamma(a+j)], \quad k = 1, 2, \ldots$$

By using the well-known fact that $\Gamma(s+1) = s\,\Gamma(s)$, it is easily established that Ψ_k can be simplified to the following function:

$$(43) \qquad \Psi_k(a) = \sum_{j=0}^{k-1} \binom{k-1}{j} (-1)^{j+1} \ln(a+j) \ , \ k = 1, 2, \ldots$$

where we define $\binom{0}{0} = 1$.

In particular we get that:

$$(44) \qquad \lim_{a \to 0^+} \Psi_k(a) = +\infty, \ \lim_{a \to \infty} \Psi_k(a) = 0 \ , \ k = 1, 2, \ldots ,$$

By expanding the expression for Ψ_{k+1} and using the well-known identity:

$$(45) \qquad \binom{k}{j} = \binom{k-1}{j} + \binom{k-1}{j-1}, \qquad \text{for } 0 < j < k,$$

we get that:

$$(46) \qquad \Psi_{k+1}(a) = \left[\sum_{j=0}^{k-1} \binom{k-1}{j} (-1)^{j+1} \ln(a+j) \right] + \left[\sum_{j=1}^{k} \binom{k-1}{j-1} (-1)^{j+1} \ln(a+j) \right]$$

By substituting $(j\text{-}1)$ with j in the last of the two sums in (46), we obtain the following nice recursion property:

$$(47) \qquad \Psi_{k+1}(a) = \Psi_k(a) - \Psi_k(a+1).$$

The derivative of Ψ_k is easily calculated to be:

$$(48) \qquad \tfrac{\partial}{\partial a} \Psi_k(a) = \sum_{j=0}^{k-1} \binom{k-1}{j} (-1)^{j+1} \frac{1}{a+j}, \qquad k = 1, 2, \ldots$$

Using (47) and induction it is easily established that the expression (48) may be re-written as:

$$(49) \qquad \tfrac{\partial}{\partial a} \Psi_k(a) = -(k-1)! \left[\prod_{j=0}^{k-1} \frac{1}{a+j} \right], \ k = 1, 2, \ldots$$

Hence, in particular this derivative is negative on $\langle 0, \infty\rangle$, i.e., the function Ψ_k is decreasing for all $a > 0$. Thus, since $\beta > 0$, it follows that (41) is in fact always true. That is, if θ is beta-distributed, then the joint distribution of $X_1,\dots, X_n$ admits a shock model representation. ❑

Example 15. Let $X_1,\dots, X_n$ be binary variables which, as in the previous example, can be embedded in an infinite sequence of exchangeable variables with θ as the limiting frequency of 1's in the sequence. Assume that in this case the uncertainty about θ can be described by the following distribution:

$$(50) \qquad f(\theta) = \frac{\alpha^{\gamma}}{\Gamma(\gamma)}\theta^{\alpha-1}[-\ln(\theta)]^{\gamma-1}, \qquad 0 \le \theta \le 1.$$

where α and γ are positive real numbers.

This distribution is related to the well-known gamma-distribution such that if θ has the distribution (50), then $-\ln(\theta)$ is gamma-distributed with parameters γ and α. Note that if in particular $\gamma = 1$, then θ is beta-distributed with parameters α and 1.

The moments of the distribution (50) is given by:

$$(51) \qquad \mathrm{E}\left(\theta^{p}\right) = \left(\frac{\alpha}{\alpha+p}\right)^{\gamma} \qquad p > -\alpha.$$

Inserting this into the conditions (32) we get the following set of conditions for shock model representability:

$$(52) \qquad \sum_{j=0}^{k}\binom{k}{j}(-1)^{j}\ln\left[\left(\frac{\alpha}{\alpha+n-k+j}\right)^{\gamma}\right] \ge 0, \quad 0 < k \le n.$$

Since $\gamma > 0$, we may get rid of this constant simply by dividing each side of the inequality with γ. Moreover, using (39) once again we obtain the following simplified set of conditions:

$$(53) \qquad \sum_{j=0}^{k}\binom{k}{j}(-1)^{j+1}\ln(\alpha+n-k+j) \ge 0, \qquad 0 < k \le n.$$

By (43) the left-hand side of (53) is equal to $\Psi_{k+1}(\alpha+n-k)$. Thus, by the previous example, it follows that (53) is always satisfied. ❑

Two other types of examples where the conditions (32) are always satisfied, can be obtained as follows:

Assume that θ has one of the following two distributions:

(54) $\Pr(\theta = \theta_0) = 1$, where $\theta_0 \in \langle 0, 1]$.

(55) $\Pr(\theta = 1) = 1\text{-}\Pr(\theta = 0) = \alpha$, where $\alpha \in \langle 0, 1]$.

In the case of (54) there is no uncertainty about θ, so $X_1,\dots, X_n$ are independent. In this case it is obvious that the joint distribution admits a shock model representation. In particular (32) can be written in the following form:

$$(56) \qquad \ln(\theta_0)\sum_{j=0}^{k}(-1)^j\binom{k}{j}(n-k+j) \geq 0 \ , \ 0 < k \leq n.$$

Using (39) and the following related identity:

$$(57) \qquad \sum_{j=0}^{k}(-1)^j\binom{k}{j}j = \begin{cases} -1 & k=1 \\ 0 & k=2,3,\dots \end{cases}$$

we see that the left-hand side of (56) is $-\ln(\theta_0)$ for $k = 1$, and zero otherwise. Thus, since $\theta_0 \in \langle 0, 1]$, it follows that (56) is always true.

To see that (32) holds for the distribution (55), we note that since in this case θ is either 0 or 1, it follows that $E(\theta^x) = \alpha$, for all $x > 0$, while of course $E(\theta^0) = 1$ by definition. In particular $\ln[E(\theta^{n-k+j})] = \ln(\alpha) \leq 0$, for $0 \leq j \leq k < n$, and zero if $j = 0$ and $k = n$. Thus, (32) can be written as:

$$(58) \qquad \ln(\alpha)\left[\sum_{j=0}^{k}(-1)^j\binom{k}{j} - I(k=n)\right] \geq 0 \ , \ 0 < k \leq n,$$

and by (39) this is again always true.

Extensions of the above examples can be obtained using the following two result:

Theorem 16. Assume that $\theta = \prod_{i=1}^{m}\theta_i$, where $\theta_1,\dots, \theta_m$ are independent and each θ_i has a distribution satisfying the conditions (32). Then the distribution of θ satisfies (32) as well. In particular if $\theta = \theta_1\, \theta_2$, where $\theta_1 \in \langle 0, 1]$ is a constant, then the distribution of θ satisfies (32) if and only if the distribution of θ_2 satisfies (32).

Proof: The result follows directly by observing the following:

(59) $$\ln\left[E\left(\theta^{n-k+j}\right)\right] = \sum_{i=1}^{m} \ln\left[E\left(\theta_i^{n-k+j}\right)\right]$$

and inserting this into (32). In the special case where $\theta = \theta_1 \theta_2$, where θ_1 is a constant, the equivalence follows since by (57) the moments of the constant θ_1 does not contribute to the sum in (59) except in the trivial case where $k = 1$ (see (34)). ❑

Thus, e.g., products of independent beta-distributed limiting frequencies will satisfy (32). Similarly, if $\theta = \theta_1 \cdot \theta_2$ where θ_1 is a known constant between 0 and 1 and θ_2 is beta-distributed, then θ will satisfy (32).

Theorem 17. Let $\{\theta_i\}$ be a sequence of variables such that each θ_i has a distribution satisfying the conditions (32). Assume that $\theta_i \xrightarrow{D} \theta$, where $\Pr(\theta > 0) > 0$. Then the distribution of θ satisfies (32) as well.

Proof: Since $\theta, \theta_1, \theta_2, \ldots \in [0,1]$ with probability one and $f(x) = x^a$ is continuous and bounded on [0,1] for all $a \geq 0$, it follows by standard theory that $\theta_i \xrightarrow{D} \theta$ implies that:

(60) $$E\left(\theta_i^a\right) \to E\left(\theta^a\right), \text{ for all } a \geq 0.$$

Moreover, since $\Pr(\theta > 0) > 0$, $E(\theta^a) > 0$ for all a, so $\ln[E(\theta^a)]$ is always well-defined. By inserting this into (32) the result follows. ❑

Note that if the limiting θ is concentrated in 0, then all the X_i's will be concentrated in 0 as well. In this case of course the joint distribution of $X_1, \ldots, X_n$ trivially admits a shock model representation. However, the conditions (32) do not apply since in this situation $\ln[E(\theta^a)]$ is not well-defined. In fact the conditions (32) are valid only under the assumption that $\Pr(X_i = 1) > 0$, $i = 1, \ldots, n$.

So far we have not seen any example where the conditions (32) do *not* hold. Thus, one may think that (32) is true for any distribution on the limiting frequency θ. The next example, however, shows that this is not so.

Example 18. Let X_1, X_2, X_3 be binary variables which can be embedded in an infinite sequence of exchangeable variables, and let θ denote the limiting frequency of 1's in the sequence. Assume that θ has the following distribution:

(61) $$\Pr(\theta = 1) = 1 - \Pr(\theta = \theta_0) = \alpha$$

where α, $\theta_0 \in [0, 1]$, and $\max(\alpha, \theta_0) > 0$ (avoiding distributions concentrated at zero). Note that if $\theta_0 > 0$ and $\alpha = 0$, we get the model (54). Similarly if $\theta_0 = 0$ and $\alpha > 0$, we get the model (55).

To study the conditions (32) for this model, we introduce the following functions corresponding to the left-hand sides of (32):

$$\text{(62)} \qquad \phi_k(\alpha,\theta_0) = \sum_{j=0}^{k} (-1)^j \binom{k}{j} \ln\left[(1-\alpha)\theta_0^{3-k+j} + \alpha\right], \quad k = 1, 2, 3.$$

By the remark to Theorem 13 (see (34) and (35)) the only nontrivial condition can be written as:

$$\text{(63)} \qquad \phi_3(\alpha,\theta_0) = \sum_{j=0}^{3} (-1)^j \binom{3}{j} \ln\left[(1-\alpha)\theta_0^{j} + \alpha\right] \geq 0.$$

It is easy to verify that:

$$\text{(64)} \qquad \phi_3(\alpha, 0) = -\ln(\alpha), \ \phi_3(\alpha, 1) = 0, \quad \text{for all } \alpha \in \langle 0, 1].$$

Thus, if we could show that for a given $\alpha > 0$, $\phi_3(\alpha, \theta_0)$ is decreasing in θ_0, then it would follow that $\phi_3(\alpha, \theta_0) \geq 0$ for all $\theta_0 \in [0, 1]$. We will show that this is true only for large values of α. To see this, we calculate the partial derivative of $\phi_3(\alpha, \theta_0)$ with respect to θ_0:

$$\text{(65)} \qquad \frac{\partial}{\partial\theta_0}\phi_3(\alpha,\theta_0) = -\frac{3(1-\alpha)}{(1-\alpha)\theta_0 + \alpha} + \frac{6(1-\alpha)\theta_0}{(1-\alpha)\theta_0^2 + \alpha} - \frac{3(1-\alpha)\theta_0^2}{(1-\alpha)\theta_0^3 + \alpha}.$$

To study the sign of this we introduce:

$$\text{(66)} \qquad \delta(y) = \frac{\theta_0^y}{(1-\alpha)\theta_0^{y+1} + \alpha}$$

By (65) it follows that the partial derivative of $\phi_3(\alpha, \theta_0)$ with respect to θ_0 is nonpositive if and only if:

$$\text{(67)} \qquad \delta(1) \leq \tfrac{1}{2}\delta(0) + \tfrac{1}{2}\delta(2)$$

By Jensen's inequality we get that a sufficient condition for (67) to hold, is that δ is a convex function. By elementary calculus, this can be shown to hold if $\alpha \geq \frac{1}{2}$. However, when α is small but positive, $\phi_3(\alpha, \theta_0)$ will not be monotone, and one can find θ_0-values where the condition (63) does not hold. E.g., $\phi_3(0.05, 0.25) = -0.16 < 0$.

In Figure 1 we have plotted $\phi_3(\alpha, \theta_0)$.

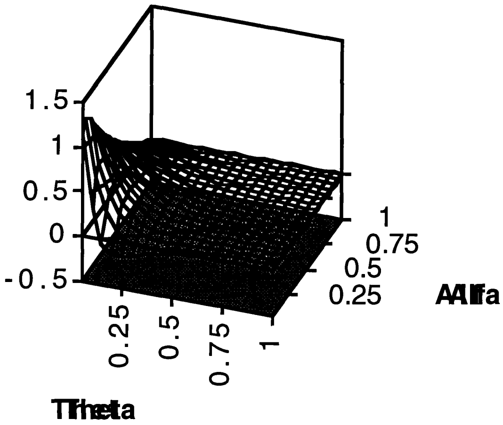

Figure 1.

In this plot one clearly sees that there is a region where the condition (63) does not hold. ❑

Although the last example is rather specialized, it is possible to use this to derive other related examples as well. First we observe that if θ has the distribution (61) with α and θ_0 in the region where (32) does not hold, then the same will also be true for $a\theta$, for any $a \in \langle 0, 1]$. This follows by applying the same argument as in the proof of Theorem 16, and using that by the identity (57) we have that:

$$(68) \qquad \sum_{j=0}^{3} (-1)^j \binom{3}{j} \ln\left(a^j\right) = 0$$

The variable $a\theta$ will have a distribution concentrated on the set $\{a\theta_0, a\}$.

As a more advanced example, we can show that the class of distributions satisfying (32) is not closed under mixtures. In particular we can show that mixtures of beta-distributed limiting frequencies may not satisfy (32). To see this we let $\theta_{i,n}$ be beta-distributed with parameters $\alpha_{i,n}$ and $\beta_{i,n}$, $n = 1, 2, \ldots$ and $i = 1, 2$. Assume that:

$$(69) \qquad \lim_{n\to\infty} \alpha_{i,n} = \lim_{n\to\infty} \beta_{i,n} = \infty, \ \lim_{n\to\infty} \frac{\alpha_{1,n}}{\alpha_{1,n} + \beta_{1,n}} = 1 \text{ and } \lim_{n\to\infty} \frac{\alpha_{2,n}}{\alpha_{2,n} + \beta_{2,n}} = \theta_0$$

We then assume that θ_n has a distribution that is a mixture of the distributions of $\theta_{1,n}$ and $\theta_{2,n}$. That is, we assume that:

$$(70) \qquad \Pr(\theta_n = \theta_{1,n}) = 1 - \Pr(\theta_n = \theta_{2,n}) = \alpha.$$

If θ has the distribution (61), then it is easily seen that:

$$(71) \qquad \theta_n \xrightarrow{D} \theta.$$

If the distribution of θ_n satisfies (32) for all n, then by Theorem 17, so would the distribution of θ. However, as we have seen above, we may choose α and θ_0 such that this is not true, implying that for sufficiently large n, θ_n will have a distribution not satisfying (32).

We close this section by presenting yet another interesting consequence of Example 18. This is related to a situation with non-fatal shocks, and is described as follows:

Example 19. Let $X_1, \ldots, X_n$ be exchangeable binary variables representing the states of n components living in a common environment. The components are exposed to certain non-fatal shocks which affect the reliabilities of the components. For simplicity we only consider shocks striking all components here. Given the number of shocks that have occurred, the X_i's are assumed to be independent. Specifically we assume that:

$$(72) \qquad \Pr(X_i = 1 \mid k \text{ shocks have occurred}) = \theta_k, \qquad i = 1, \ldots, n, \;\; k = 1, 2, \ldots$$

where $\theta_0 \geq \theta_1 \geq \cdots$. That is, the component reliability is decreasing as the number of shocks grows. We denote the number of shocks by K. In the case where $\Pr(K = 0) = \alpha = 1 - \Pr(K = 1)$, and $n = 3$, we essentially have the same model as in Example 18. As we already know, this type of model may not have a shock model representation. Clearly the same thing may happen for other distributions of K.

By extending the concept of non-fatal shocks to more general situations with shock striking different subsets of the component set in the same fashion as above we get a new class of probability models containing the class of fatal shock models as limiting cases. It follows from the above discussion that this extended class is larger than the class where only fatal shocks can occur. It is interesting to compare this with the continuous time shock model of Marshall and Olkin (1967). In their case it turns out that the two corresponding classes are equivalent, in the sense that they both lead to the same kind of multivariate distribution. ❑

4. A dual shock model.

In the previous section we investigated a situation where component failures could occur as a consequence of an “external” shock. Thus, the external environment where the components operated, could never have a positive effect on the components. In many practical situations, however, one would like to include positive environmental effects as well. If e.g. the components are inspected and maintained according to a certain strategy, the performance of the components may be improved. If the same strategy is applied to groups of components, weakness in this strategy could affect all components in such a group. As a result, the component states would be dependent. Assuming that the maintenance strategy never causes component failures by itself, it does not seem appropriate to use a shock model to represent this kind of dependence. Instead one may do as follows:

As usual we let $E = \{1,\ldots n\}$ be the component set of a reliability system, and let the component states be represented by the binary variables $X_1,\ldots, X_n$. For each non-empty subset A of E there exists a maintenance strategy. If the strategy “works”, all components in the set survive. To describe the “maintenance status” of a particular set, we introduce:

(73) $$Y_A = \begin{cases} 1 & \text{if maintenance of } A \text{ has worked} \\ 0 & \text{otherwise} \end{cases}, \emptyset \subset A \subseteq E.$$

In the same way as we did with the shock model, we may now express the X_i's in terms of the Y_A's:

(74) $$X_i = \coprod_{A:i\in A} Y_A,\ i = 1,\ldots, n.$$

As for the shock model, we assume that the Y_A's are independent.

We observe that the model (74) is dual to the shock model (6). Thus, we may refer to (74) as a dual shock model. This model was suggested and analyzed by Egeland and Huseby (1991).

Because of the duality between the two models, almost all results we present in this section, are derived as simple consequences of the corresponding shock model results. To show this, we shall find it convenient to introduce the following dual variables:

(75) $$X_i^d = 1 - X_i,\ i = 1,\ldots, n.$$

(76) $$Y_A^d = 1 - Y_A,\ \emptyset \subset A \subseteq E.$$

The model (74) can then be stated in terms of the dual variables as:

(77) $$X_i^d = \prod_{A:i\in A} Y_A^d,\ i = 1,\ldots, n.$$

We also introduce:

(77) $$P^d(A) = \Pr\left(\prod_{i\in A} X_i^d = 1\right) = \Pr\left(\coprod_{i\in A} X_i = 0\right), \varnothing \subset A \subseteq E.$$

(78) $$\pi^d(A) = \Pr\left(Y_A^d = 1\right) = \Pr(Y_A = 0), \varnothing \subset A \subseteq E.$$

(79) $$L^d(A) = \ln P^d(A) \text{ and } \lambda^d(A) = \ln \pi^d(A), \varnothing \subset A \subseteq E.$$

As usual we also define $P^d(\varnothing) = 1$ and $L^d(\varnothing) = 0$. To avoid trivial cases we assume that $\Pr(X_i = 0) > 0$. [Compare the assumptions underlying Theorem 6.] This implies that all P^d's and π^d's are positive, and thus their logarithms are well-defined.

The fundamental relation between the shock model and its dual can now be stated as follows:

Theorem 20. Let $X_1,\ldots, X_n$ be binary variables, and let $X_1^d,\ldots,X_n^d$ denote the corresponding duals. Then the joint distribution of $X_1,\ldots, X_n$ can be represented by a dual shock model if and only if the joint distribution of $X_1^d,\ldots,X_n^d$ can be represented by a shock model .

Proof: The result follows immediately from (77). ❑

Theorem 20 allows us to simply transform the results from the previous section to obtain corresponding results for the dual shock model. No new proofs are needed. In particular we have:

Theorem 21. The joint distribution of the binary variables $X_1,\ldots, X_n$ may be represented by a dual shock model if and only if:

(80) $$\sum_{D\subseteq A}(-1)^{|D|} L^d((E \setminus A) \cup D) \geq 0,\ \varnothing \subset A \subseteq E.$$

Proof: The result is a direct consequence of Theorem 20 and Theorem 7. ❑

By translating Example 8 we get that the joint distribution of two binary variables X_1 and X_2 can be represented by a dual shock model if and only if $\mathrm{Cov}\left(X_1^d, X_2^d\right) \geq 0$. However, by elementary properties of the covariance, this condition is equivalent to

$\text{Cov}(X_1, X_2) \geq 0$. Hence, we observe that in the bivariate case the two model classes are in fact equal.

Similarly, by using Example 9 we get that the joint distribution of three binary variables X_1, X_2 and X_3 can be represented by a dual shock model if and only if the following two sets of conditions hold:

(81) $$\Pr\left(X_k = 0 \middle| X_i = 0,\ X_j = 0\right) \geq \Pr\left(X_k = 0 \middle| X_i = 0\right)$$

(82) $$\Pr\left(X_k = 0 \middle| X_i = 0\right) \cdot \Pr\left(X_k = 0 \middle| X_j = 0\right)$$
$$\geq \Pr\left(X_k = 0 \middle| X_i = 0,\ X_j = 0\right) \cdot \Pr\left(X_k = 0\right)$$

where $\{i, j, k\} = \{1, 2, 3\}$.

Note that in the trivariate case we cannot obtain equivalent conditions for the two models. Although (81) and (82) look similar to (20) and (21), the conditions are not equivalent, as we shall see later on.

We now turn to the case where $X_1, \ldots, X_n$ are exchangeable. Analogous to what we did in the shock model case, we introduce the simplified notation:

(83) $$P_i^d = P^d(\{1, \ldots, i\}),\ L_i^d = \ln(P_i),\ i = 1, \ldots, n,$$

and let $P_0^d = P^d(\emptyset) = 1$, and $L_0^d = 0$.

Using this notation, we get the following result corresponding to Theorem 10:

Theorem 22. The joint distribution of the exchangeable binary variables $X_1, \ldots, X_n$ may be represented by a dual shock model if and only if:

(84) $$\sum_{j=0}^{k} (-1)^j \binom{k}{j} L_{n-k+j}^d \geq 0,\ 0 < k \leq n.$$

Moreover:

(85) $$\pi_k^d = \exp\left\{-\sum_{j=0}^{k} (-1)^j \binom{k}{j} L_{n-k+j}^d\right\},\ 0 < k \leq n.$$

Proof: The result is a direct consequence of Theorem 20 and Theorem 10. ❑

Assume then that $X_1,\ldots, X_n$ can be embedded in an *infinite* sequence of exchangeable variables, i.e., the joint distribution of $X_1,\ldots, X_n$ admits the representation (28). In this case it follows that:

$$(86) \qquad P_i^d = \mathrm{E}\left[(1-\theta)^i\right], \; i = 0, 1,\ldots, n,$$

where we always define $\mathrm{E}[(1\text{-}\theta)^0] = P_0^d = 1$. By inserting this into (84), we obtain:

Theorem 23. Let $X_1,\ldots, X_n$ be binary variables which can be embedded in an infinite sequence of exchangeable variables, and let θ denote the limiting frequency of 1's in the sequence. Then the joint distribution of $X_1,\ldots, X_n$ may be represented by a dual shock model if and only if:

$$(87) \qquad \sum_{j=0}^{k}(-1)^j\binom{k}{j}\ln\left[\mathrm{E}\left((1-\theta)^{n-k+j}\right)\right] \geq 0, \; 0 < k \leq n.$$

Moreover,

$$(88) \qquad \pi_k^d = \exp\left\{-\sum_{j=0}^{k}(-1)^j\binom{k}{j}\ln\left[\mathrm{E}\left((1-\theta)^{n-k+j}\right)\right]\right\}, \; 0 < k \leq n.$$

Proof: The result is a direct consequence of Theorem 20, Theorem 13 and the above discussion. ❑

Corollary 24. Let θ denote the limiting frequency of 1's in an infinite sequence of binary variables. Then θ satisfies (87) if and only if $(1\text{-}\theta)$ satisfies (32).❑

We now consider some examples:

Example 25. Let $X_1,\ldots, X_n$ be binary variables which can be embedded in an infinite sequence of exchangeable variables, let θ denote the limiting frequency of 1's in the sequence, and assume that θ is beta-distributed with parameters α and β, defined in (36). Then $(1\text{-}\theta)$ is beta-distributed with parameters β and α. Hence by Example 14, $(1\text{-}\theta)$ satisfies (32). By Corollary 24 this implies that θ satisfies the condition (87), so the joint distribution of $X_1,\ldots, X_n$ admits a dual shock model representation (as well as a regular shock model representation). ❑

Example 26. Let $X_1,\ldots, X_n$ and θ be defined as in the previous example, and assume that θ is distributed according to the following density:

(89) $$f(\theta)=\frac{\alpha^{\gamma}}{\Gamma(\gamma)}(1-\theta)^{\alpha-1}[-\ln(1-\theta)]^{\gamma-1}, \quad 0\le\theta\le 1.$$

where α and γ are positive real numbers.

It is easy to see that this implies that (1-θ) has the distribution (50). By Example 15 and Corollary 24 this implies that θ satisfies the condition (87), so the joint distribution of $X_1,\dots, X_n$ admits a dual shock model representation. We observe however, that in this case, (1-θ) does *not* have a distribution in the same class as θ. As a consequence, we cannot immediately apply Corollary 24 to establish that the joint distribution of $X_1,\dots, X_n$ admits a regular shock model representation as well. Alternatively we cannot conclude so far that the distribution (50) satisfies (87). In order to investigate this issue further, we compute the moments of the distribution (89). Applying (51) and induction, we get:

(90) $$\mathrm{E}\left(\theta^{p}\right)=\sum_{i=0}^{p}\binom{p}{i}(-1)^{i}\left(\frac{\alpha}{\alpha+i}\right)^{\gamma}, \quad p=1,2,\dots$$

Inserting this into (32), we get the following condition for shock representability:

(91) $$\sum_{j=0}^{k}(-1)^{j}\binom{k}{j}\ln\left[\sum_{i=0}^{n-k+j}\binom{n-k+j}{i}(-1)^{i}\left(\frac{\alpha}{\alpha+i}\right)^{\gamma}\right]\ge 0, \; 0<k\le n.$$

If $\gamma = 1$, it is easy to see that (89) is reduced to a beta-distribution with parameters 1 and α. Thus, by the previous examples, we now that (91) holds for this particular γ-value (for all $\alpha > 0$). Furthermore, it is easy to see that the left-hand side of (91) goes to zero as γ goes to infinity. Thus, if we could establish that the left-hand side of (91) is non-increasing in γ, it would follow that (91) is true for all $\gamma > 0$. Due to the complexity of the expression (91) we have not been able to establish this analytically. However, numerical calculations indicates that this in fact is true. Thus, we conjecture that the joint distribution of $X_1,\dots, X_n$ admits a regular shock model representation. ❑

By Corollary 24 and (54-55) it also follows that the following two distributions satisfy the condition (87):

(92) $\Pr(\theta = \theta_0) = 1$, where $\theta_0 \in [0, 1\rangle$.

(93) $\Pr(\theta = 1) = 1-\Pr(\theta = 0) = \alpha$, where $\alpha \in [0, 1\rangle$.

We also state the results analogous to Theorem 16 and Theorem 17:

Theorem 27. Assume that $\theta = \coprod_{i=1}^{m} \theta_i$, where $\theta_1, \ldots, \theta_m$ are independent and each θ_i has a distribution satisfying the condition (87). Then the distribution of θ satisfies (87) as well. In particular if $\theta = \theta_1 \coprod \theta_2$, where $\theta_1 \in [0, 1\rangle$ is a constant, then the distribution of θ satisfies (87) if and only if the distribution of θ_2 satisfies (87).

Proof: The result follows directly by Corollary 24. ❑

Theorem 28. Let $\{\theta_i\}$ be a sequence of variables such that each θ_i has a distribution satisfying the condition (87). Assume that $\theta_i \xrightarrow{D} \theta$, where $\Pr(\theta < 1) > 0$. Then the distribution of θ satisfies (87) as well.

Proof: The result follows by the same type of argument as we used to prove Theorem 17 since $f(x) = (1-x)^a$ is continuous and bounded on [0,1] for all $a \geq 0$. ❑

In the same way as we did in the previous section, we can of course use the last two theorems to extend the results in the above examples.

We close this section by considering a situation similar to the one treated in Example 18. In this case, however, we allow θ to have a general two-point distribution.

Example 29. Let X_1, X_2, X_3 be binary variables which can be embedded in an infinite sequence of exchangeable variables, and let θ denote the limiting frequency of 1's in the sequence. Assume that θ has the following distribution:

(94) $$\Pr(\theta = \theta_1) = 1 - \Pr(\theta = \theta_0) = \alpha$$

where $\alpha, \theta_0, \theta_1 \in \langle 0, 1\rangle$ and $\theta_0 < \theta_1$ (avoiding distributions concentrated at zero or one).

We also introduce the two random variables, θ' and θ'', with distributions:

(95) $$\Pr(\theta' = 1) = 1 - \Pr\left(\theta' = \frac{\theta_0}{\theta_1}\right) = \alpha$$

(96) $$\Pr\left(\theta'' = 1 - \frac{1-\theta_1}{1-\theta_0}\right) = 1 - \Pr(\theta'' = 0) = \alpha$$

Note that:

(97) $$\theta \stackrel{D}{=} \theta_1\theta' \stackrel{D}{=} \theta_0 \coprod \theta''$$

Hence, by Theorem 16, θ satisfies (32) if and only if θ' does. Similarly, by Theorem 27, θ satisfies (87) if and only if θ'' does.

We now consider two cases:

Case 1. $\alpha \geq \frac{1}{2}$. In this case we know from Example 18 that θ' always satisfies (32). Hence it follows that the joint distribution of X_1, X_2, X_3 admits a shock model representation.

By Corollary 24 θ'' satisfies (87) if and only if $(1-\theta'')$ satisfies (32). However, since $\alpha \geq \frac{1}{2}$, clearly $(1-\alpha) \leq \frac{1}{2}$. Using Example 18 again, it is easy to see that we can choose θ_0, θ_1 and α such that this is *not* true. Hence it follows that the joint distribution of X_1, X_2, X_3 does *not* necessarily admit a dual shock model representation.

Case 2. $\alpha \leq \frac{1}{2}$. In this case we know from Example 18 that we can choose θ_0, θ_1 and α such that θ' does *not* satisfy (32). Hence it follows that the joint distribution of X_1, X_2, X_3 does *not* necessarily admit a shock model representation.

Since we now have that $(1-\alpha) \geq \frac{1}{2}$, by Example 18 we get that $(1-\theta'')$ always satisfies (32), and hence by Corollary 24, θ'' satisfies (87). Hence it follows that the joint distribution of X_1, X_2, X_3 admits a dual shock model representation.

We can summarize this example by saying that if θ has the distribution (94), then the joint distribution of X_1, X_2, X_3 admits either a shock model representation or a dual shock model representation or both. ❑

In general we conclude that the shock models and dual shock models are two different classes based on quite different physical assumptions. While some distributions belong to both classes, we have shown that there exists distributions belonging to only one of the classes. Thus, the two classes supplement each other nicely. Still we expect to find examples of distributions of exchangeable X_i's belonging to none of the classes.

5. Conclusions

The IFIV class provides a rich class of probability models, all of which can be interpreted in a natural way in terms of operationally well-defined quantities. Thus, from an operational point of view, the IFIV class is more appealing than the class of associated random variables.

In the binary case the IFIV class is contains all PDM models. Generally there is a strong relation between PDM models and IFIV models.

Conditions are given for when a joint distribution of binary variables can be interpreted as a shock model. In particular we have derived conditions for when a binary PDM model admits such an interpretation.

Acknowledgement: This research has been supported by Johan og Mimi Wesmanns Understøttelsesfond and by the Norwegian Council for Research. The author is grateful to these institutions for making this research possible. The author would also like to thank Professor Nozer Singpurwalla, Professor Richard Barlow and Professor Max Mendel for many helpful comments.

5. References:

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

Boyles, R. A. and Samaniego, F. J. (1984) Modeling and Inference for Multivariate Binary Data With Positive Dependence. *J. American Stat. Assoc.* **79**, 188-193.

Egeland, T, and Huseby, A. B. (1991) On Dependence and Reliability Computations. *Networks* **21**, pp 521-545.

Esary, J. D., Proschan, F. and Walkup, D. (1967) Association of Random Variables with Applications. *Ann. Math. Stat.,* **38**, 1466-1474.

Gåsemyr, J. and Natvig, B. (1994) Some Aspects of Reliability Analysis in Shock Models. Submitted to *Scand. J. Statist.*

Kimball, A. W. (1951) On Dependent Tests of Significance in the Analysis of Variance. *Ann. Math. Statist.* **22**, 600-602.

Marshall, A. and Olkin, I. (1967) A Multivariate Exponential Distribution. *J. American Stat. Assoc.* **62**, 30-44.

Pitt, L. (1982) Positively Correlated Random Variables are Associated. *Ann. of Prob.*, **10**, 496-499.

Proschan, F. and Sethuraman, J. (1974) Schur functions in Statistics. I: The preservation theorem. Tech. Rep. No. M254-RR, Dept. of Statistics, Florida State Univ., Tallahassee, Florida.

Sampson, A. R. and Whitaker, L. R. (1989) Computational Aspects of Association for Bivariate Discrete Distributions. In *Contributions to Probability and Statistics, Essays in Honor of Ingram Olkin*, (Gleser, L. J. et. al., eds.), Springer-Verlag.

Tong, Y. L. (1977) An Ordering Theorem for Conditionally Independent and Identically Distributed Random Variables. *Ann. Statist.* **5**, 274-277.

Tong, Y. L. (1980) *Probability Inequalities in Multivariate Distributions.* Ac. Press.

Welsh, D. J. A. (1976) *Matroid Theory*. Ac. Press.