The association in time of a Markov process
with application to multistate reliability theory

N.L. HJORT, B. NATVIG and E. FUNNEMARK

In traditional reliability theory the system and the components are described just as functioning or failed. For this case a series of bounds for the availability and unavailability in a fixed time interval, I , for a system of maintained, interdependent components are given in Natvig (1980). For the special case of independent components the only assumption needed is that the marginal performance process of each component is associated in I . When these processes are Markovian a sufficient condition for this to hold is given by Esary and Proschan (1970).

By now the traditional binary theory is being replaced by a theory for multistate systems of multistate components. Here the states represent successive levels of performance ranging from a perfect functioning level down to a complete failure level. In Funnemark and Natvig (1982) the work of Natvig (1980) is generalized to the multistate case. In the present paper we generalize the sufficient condition given by Esary and Proschan (1970) and give an equivalent and much more convenient condition in terms of the transition intensities of the Markov process.

KEY WORDS: Association in time; Markov processes; Multistate reliability theory.

1. INTRODUCTION

In reliability theory a key problem is to find out how the performance of a complex system can be determined from knowledge of the performance of its components. One inherent weakness of the traditional theory in this field is that the system and the components are always described just as functioning or failed. This approach represents an oversimplification in many real-life situations where the systems and their components are capable of assuming a whole range of levels of performance, varying from perfect functioning to complete failure.

Fortunately, by now the traditional binary theory is being replaced by a theory for multistate systems of multistate components. Some recent references are Natvig (1982) and Block and Savits (1982). Let the set of states of the system be $S = \{0, 1, \dots, M\}$. The $M+1$ states represent successive levels of performance ranging from the perfect functioning level M down to the complete failure level 0 . Furthermore, let the set of components be $C = \{1, 2, \dots, n\}$ and the set of states of the i -th component S_i ($i=1, \dots, n$) where $\{0, M\} \subseteq S_i \subseteq S$. If x_i ($i=1, \dots, n$) denotes the state or performance level of the i -th component and $\underline{x} = (x_1, \dots, x_n)$, it is assumed that the state ϕ of the system is a deterministic function of $\underline{x}$; i.e. $\phi = \phi(\underline{x})$. Here $\underline{x}$ takes values in $S_1 \times S_2 \times \dots \times S_n$ and ϕ takes values in S . The function ϕ is called the structure function of the system.

Definition 1.1 A system is a multistate monotone system (MMS) iff

- i) $\phi(\underline{x})$ is non-decreasing in each argument.
- ii) $\phi(\underline{0}) = 0$ and $\phi(\underline{M}) = M$ ($\underline{0}=(0, \dots, 0), \underline{M}=(M, \dots, M)$).

As a simple example of an MMS consider the network of Figure A.

Figure A. Example of an MMS.

Here component 1 (2) is the parallel module of the branches a_1 and b_1 (a_2 and b_2).

Let ($i=1, 2$)

$$\begin{aligned} x_i &= 3 && \text{if two branches work,} \\ &= 1 && \text{if one branch works,} \\ &= 0 && \text{if no branch works.} \end{aligned}$$

The state of the system is given in Table 1.

Table 1. State of network in Figure A.

	3	0	2	3
Component 2	1	0	1	2
	0	0	0	0
	0	1	3	
Component 1				

Note for instance that the state 1 is a critical one both for each component and the system as a whole in the sense that the failing of a branch leads to the 0 state. In binary theory the functioning state comprises the states $\{1,2,3\}$ and hence just a rough description of the system's performance is possible.

We now give some basic definitions.

Definition 1.2 The performance process of the i -th component is a stochastic process $\{X_i(t), t \in \tau\}$ where for each fixed $t \in \tau, X_i(t) \in S_i$ denotes the state of the component at time t . The joint performance process of the components is given by

$$\{\underline{X}(t), t \in \tau\} = \{X_1(t), \dots, X_n(t), t \in \tau\}.$$

The time domain τ is contained in $[0, \infty)$. We assume that the sample functions $X_i(t), t \in \tau, i=1, \dots, n$ are continuous from the right on τ . The performance process of the system is now given by $\{\phi(\underline{X}(t)), t \in \tau\}$.

From Barlow and Proschan (1975) we have

Definition 1.3 The r.v.'s $T_1, \dots, T_n$ are associated iff $\text{cov}[\Gamma(\underline{T}), \Delta(\underline{T})] > 0$ for all pairs of non-decreasing binary functions Γ, Δ .

Consider a time interval $I = [t_A, t_B] \subset [0, \infty)$ and let $\tau(I) = \tau \cap I$.

Definition 1.4 The joint performance process $\{\underline{X}(t), t \in \tau\}$ of the components is associated in the time interval I iff, for any integer m and $\{t_1, \dots, t_m\} \subset \tau(I)$, the r.v.'s in the array

$$\begin{array}{c} X_1(t_1), \dots, X_1(t_m) \\ \vdots \\ X_n(t_1), \dots, X_n(t_m) \end{array}$$

are associated.

This definition obviously applies to a marginal performance process too.

Definition 1.5 The availability, $h_\phi^{j(I)}$, and the unavailability, $g_\phi^{j(I)}$, to the level j in the time interval I for an MMS with structure function ϕ are given by ($j=1, \dots, M$):

$$h_\phi^{j(I)} = \Pr[\phi(\underline{X}(s)) \geq j \quad \forall s \in \tau(I)],$$

$$g_\phi^{j(I)} = \Pr[\phi(\underline{X}(s)) < j \quad \forall s \in \tau(I)].$$

In Funnemark and Natvig (1982) a series of bounds for $h_{\phi}^j(I)$ and $g_{\phi}^j(I)$ are given in the case of maintained, interdependent components, thus generalizing the work of Natvig (1980) treating the binary case ($M=1$). These bounds are of great interest when trying to predict the performance process of the system. For the special case where the performance processes of the components are independent, the only assumption needed is that the marginal performance process of each component is associated in I . This ensures for instance that the joint performance process of the components is associated in I .

When these marginal performance processes are Markovian, we present and prove in the next section a theorem providing a sufficient condition for each of them to be associated in I , which is what is needed, thus generalizing a result by Esary and Proschan (1970). Imbedded in our theorem is an equivalent and much more convenient condition in terms of the transition intensities of the Markov process.

Before concentrating on this theorem it should be mentioned that concepts of positive dependence of sets of r.v.'s have recently received a lot of attention; see Block and Ting (1981), Bølviken (1982). Note also that the obvious guess that a normal vector is associated iff all simple correlation coefficients are nonnegative, has just recently been confirmed by Pitt (1982).

2. THE MAIN THEOREM

Let $X = \{X(t), t \in \tau\}$ be a Markov process with state space $\{0, 1, \dots, k\}$. (Note that when considering the i -th component of Section 1, we renumber the elements of S_i ; $k+1$ is then the number of elements of S_i .) Denote the corresponding transition probabilities

$$P_{i,j}(s,t) = \Pr[X(t)=j | X(s)=i], \quad s < t, \quad (2.1)$$

and let $\tilde{P}(s,t) = \{P_{i,j}(s,t)\}_{\substack{i=0,1,\dots,k \\ j=0,1,\dots,k}}$. Consider $\tau(I) = [0, \infty)$

and assume the existence of the transition intensities

$$\mu_{i,j}(s) = \lim_{h \rightarrow 0^+} P_{i,j}(s, s+h)/h, \quad i \neq j. \quad (2.2)$$

The following notation is needed:

$$P_{i,>j}(s,t) = \sum_{v=j}^k P_{i,v}(s,t) = \Pr[X(t) > j | X(s)=i],$$

$$\mu_{i,>j}(s) = \sum_{v=j}^k \mu_{i,v}(s), \quad i < j,$$

$$\mu_{i,<j}(s) = \sum_{v=0}^{j-1} \mu_{i,v}(s), \quad i > j.$$

Theorem 2.1

Let X be a continuous time Markov process with state space $\{0,1,\dots,k\}$ and matrix of transition probabilities $P(s,t)$. Assume the transition intensities to be continuous. Consider the following statements about X :

- (i) X is associated in time.
- (ii) X is conditionally, stochastically, non-decreasing in time; i.e.

$$\Pr[X(t) > j | X(s_1)=i_1, \dots, X(s_n)=i_n]$$

is non-decreasing in $i_1, \dots, i_n$ for each j and for each choice of $s_1 < s_2 < \dots < s_n < t, n \geq 1$.

- (iii) $P_{i,>j}(s,t)$ is non-decreasing in i for each j and for each choice of $s < t$.

- (iv) For each j and s

$$\mu_{i,>j}(s) \text{ is non-decreasing in } i \in \{0,1,\dots,j-1\},$$

$$\mu_{i,<j}(s) \text{ is non-increasing in } i \in \{j,j+1,\dots,k\}.$$

Then (ii), (iii), (iv) are equivalent and each of them implies (i).

Proof of Theorem

The crucial implication (ii) $\Rightarrow$ (i) is an easy consequence of Theorem 4.7, p.146 in Barlow and Proschan (1975), a result dating back to Esary and Proschan (1968). The equivalence of (ii) and (iii) follows directly from the Markov property of X . Choosing $s < t$ (iii) is equivalent to

$$P_{i,>j}(s,t) \text{ is non-decreasing in } i \in \{0,1,\dots,j-1\},$$

$$P_{i,<j}(s,t) \text{ is non-increasing in } i \in \{j,j+1,\dots,k\}, \quad (2.4)$$

$$1 \geq P_{j,<j}(s,t) + P_{j-1,>j}(s,t).$$

By setting $t = s+h$, dividing by h and letting $h \rightarrow 0^+$, (iv) follows. We may hence think of (iv) as the local version of (iii). What remains to be shown is that (iv) implies (iii); i.e. that a process possessing the (iii) property locally, also possesses it globally.

Let $\mathcal{M}$ denote the class of all transition probability matrices $\tilde{P} = \{P_{ij}\}_{i=0,1,\dots,k}^{j=0,1,\dots,k}$ where $P_{i,j}$ is non-decreasing in

i for each j . Saying that X has property (iii) amounts to saying that $\tilde{P}(s,t) \in \mathcal{M}$ for each choice of $s < t$. Let now $\tilde{Q}(u) = \{q_{i,j}(u)\}_{i=0,1,\dots,k}^{j=0,1,\dots,k}$ be the intensity matrix for the process X ; i.e.

$$\begin{aligned} q_{i,j}(u) &= \mu_{i,j}(u), \quad i \neq j, \\ q_{i,i}(u) &= - \sum_{j \neq i} \mu_{i,j}(u). \end{aligned} \quad (2.5)$$

We then have

$$\tilde{P}(s,t) = \exp\left(\int_{[s,t]} \tilde{Q}(u) du\right) \quad (2.6)$$

(This formula is perhaps best known in the time-homogeneous case, where $\tilde{Q}$ becomes a constant matrix. However, the more general case we treat here, with continuous $\tilde{Q}(u)$, can be worked out using the same methods; see e.g. Karlin and Taylor (1975, p.152). It is in fact possible to prove an even more general version, avoiding continuity of the $\mu_{i,j}$'s, using the product integral, see Johansen (1977).) From (2.6) we get the following representation

$$\tilde{P}(s,t) = \lim_{n \rightarrow \infty} \prod_{j=0}^{n-1} [I + \tilde{Q}(s + (j/n)(t-s))(t-s)/n], \quad (2.7)$$

where I is the identity matrix.

If now (iv) is true, it readily follows from (2.4) and (2.5) that for all small enough h , say for $h \in [0, h_0]$,

$$I + \tilde{Q}(u)h \in \mathcal{M}.$$

Since the intensities are bounded on $[s,t]$ we may choose h_0 independent of u in $[s,t]$. If we can show that $\mathcal{M}$ is closed under multiplication, it follows from (2.7) that

$$\tilde{P}(s,t) = \lim_{n \rightarrow \infty} \tilde{M}_n,$$

where $\tilde{M}_n \in \mathcal{M}$ for all large enough n , say for $n \geq n_0$, and we can conclude, since $\mathcal{M}$ is also closed under pointwise limits, that $\tilde{P}(s,t) \in \mathcal{M}$; i.e. (iii) is true.

To show that M is closed under multiplication let $\underline{M} \in M$ have elements $M_{i,j}$, $i, j = 0, 1, \dots, k$. We start by giving the $M_{i,j}$'s a special representation. Consider first the k -th column. Since $M_{i,k}$ is non-decreasing in i , there exist nonnegative

numbers $\varepsilon_i^{(k)}$, $i = 0, \dots, k$ with corresponding sums

$$\delta_i^{(k)} = \sum_{j=0}^i \varepsilon_j^{(k)}, \quad i = 0, \dots, k$$

(setting $\varepsilon_0^{(k)} = \delta_0^{(k)} = 0$), such that

$$M_{i,k} = m_k + \delta_i^{(k)}, \quad i = 0, 1, \dots, k.$$

Next look at the $(k-1)$ -th column. Since $M_{i, >k-1}$ is non-decreasing in i , we have

$$\begin{aligned} M_{i, >k-1} &= m_{>k-1} + \varepsilon_0^{(k-1)} + \varepsilon_1^{(k-1)} + \dots + \varepsilon_i^{(k-1)} \\ &= m_{k-1} + m_k + \delta_i^{(k-1)}, \quad i = 0, 1, \dots, k \end{aligned}$$

for suitable nonnegative $\varepsilon_i^{(k-1)}$'s with corresponding sums $\delta_i^{(k-1)}$ ($\varepsilon_0^{(k-1)} = \delta_0^{(k-1)} = 0$).

We continue this process and end up with $\varepsilon_i^{(j)}$, $i, j = 0, \dots, k$, all being nonnegative such that

$$M_{i, >j} = m_j + \dots + m_k + \varepsilon_0^{(j)} + \varepsilon_1^{(j)} + \dots + \varepsilon_i^{(j)} = m_j + \dots + m_k + \delta_i^{(j)},$$

where $\varepsilon_0^{(j)} = \delta_0^{(j)} = 0$, $\varepsilon_i^{(0)} = \delta_i^{(0)} = 0$ and $\sum_{j=0}^k m_j = 1$. This gives

by also setting $\varepsilon_i^{(k+1)} = \delta_i^{(k+1)} = 0$,

$$M_{i,j} = M_{i, >j} - M_{i, >j+1} = m_j + \delta_i^{(j)} - \delta_i^{(j+1)}, \quad i, j = 0, 1, \dots, k.$$

Now let $\underline{M} = \{M_{i,j}\}$ and $\bar{\underline{M}} = \{\bar{M}_{i,j}\}$ be matrices in M represented as above, where $\bar{m}_j$, $\bar{\varepsilon}_i^{(j)}$, $\bar{\delta}_i^{(j)}$ $i = 0, \dots, k$; $j = 0, \dots, k+1$ have the obvious interpretation. We study $\underline{N} = \underline{M}\bar{\underline{M}}$, with elements

$$N_{i,j} = \sum_{v=0}^k M_{i,v} \bar{M}_{v,j}.$$

Obviously

$$N_{i, >j} = \sum_{v=0}^k M_{i,v} \bar{M}_{v, >j}.$$

This gives for $i = 0, \dots, k-1$; $j = 0, \dots, k$ by using the represen-

tations introduced above:

$$\begin{aligned}
 N_{i+1, > j} - N_{i, > j} &= \sum_{v=0}^k (M_{i+1, v} - M_{i, v}) \bar{M}_{v, > j} \\
 &= \sum_{v=0}^k (\delta_{i+1}^{(v)} - \delta_{i+1}^{(v+1)} - \delta_i^{(v)} + \delta_i^{(v+1)}) \left(\sum_{r=j}^k \bar{m}_r + \bar{\delta}_v^{(j)} \right) \\
 &= \sum_{v=0}^k (\varepsilon_{i+1}^{(v)} - \varepsilon_{i+1}^{(v+1)}) \left(\sum_{r=j}^k \bar{m}_r + \bar{\delta}_v^{(j)} \right) \\
 &= \sum_{v=0}^k \varepsilon_{i+1}^{(v)} \bar{\delta}_v^{(j)} - \sum_{v=1}^k \varepsilon_{i+1}^{(v)} \bar{\delta}_{v-1}^{(j)} = \sum_{v=1}^k \varepsilon_{i+1}^{(v)} \bar{\varepsilon}_v^{(j)} \geq 0,
 \end{aligned}$$

and we have proved that M is closed under multiplication. This ends the proof of our theorem.

3. SOME CONCLUDING REMARKS

For the binary case ($k=1$) it is easily seen that statement (iii) of our theorem is equivalent to

$$P_{1,1}(s,t) + P_{0,0}(s,t) \geq 1, \text{ for each } s < t,$$

which is just the sufficient condition given by Esary and Proschan (1970). It should, however, be noted that their proof is cumbersome compared to the one given here. (Originally, for the case $k=2$, the second present author suggested a proof along the same lines, however employing Theorem 4.7, p.146 in Barlow and Proschan (1975) indirectly. This did not lead to nice sufficient conditions at all. The direct use of this theorem along with the Markov property in our proof was suggested by the first present author who is also mainly responsible for the rest of the material in Section 2.) Furthermore, for the binary case, note that when $\mu_{1,0}(s)$ and $\mu_{0,1}(s)$ are continuous, statement (iv) of our theorem is always satisfied and hence the corresponding Markov process is always associated in time. This was just noted for constant intensities in Esary and Proschan (1970).

Now let us turn to the case $k=2$ covering our example from Section 1. Then it is easily seen that (iii) is equivalent to

$$P_{0,2}(s,t) \leq P_{1,2}(s,t) \leq P_{2,2}(s,t),$$

$$P_{2,0}(s,t) \leq P_{1,0}(s,t) \leq P_{0,0}(s,t)$$

(for each $s < t$), and (iv) is equivalent to

$$\mu_{0,2}(s) \leq \mu_{1,2}(s)$$

$$\mu_{2,0}(s) \leq \mu_{1,0}(s)$$

(for all s). We emphasize that the latter condition is usually very simple to check. If in our example the two branches of each component can never be repaired/replaced simultaneously and can furthermore never fail simultaneously, i.e. we have $\mu_{0,2}(s) = \mu_{2,0}(s) = 0$ for all s , (iv) is always satisfied.

Let more generally a component consist of k branches in parallel and let its state be the number of functioning branches. Assume that the branches fail and are repaired/replaced independently of each other, all having the same instantaneous failure rate $\lambda(s)$ and repair/replacement rate $\mu(s)$. Then

$$\begin{aligned} \mu_{i,>j}(s) &= (k-j+1)\mu(s) & \text{for } i = j-1, \\ &= 0 & \text{for } i = 0, 1, \dots, j-2, \end{aligned}$$

and

$$\begin{aligned} \mu_{i,<j} &= j\lambda(s) & \text{for } i = j, \\ &= 0 & \text{for } i = j+1, \dots, k. \end{aligned}$$

Hence (iv) is satisfied.

Let in this general case the state of the component have no specific interpretation and let for all s $\mu_{j,k}(s) = 0$, $j = 1, \dots, k-1$ whereas $\mu_{0,k}(s) > 0$. This means that the component can be repaired/replaced to perfect functioning only when it has failed completely. Now (iv) does not hold. It should, however, be mentioned that this repair/replacement strategy does not seem very sensible.

Let us end up by trying to answer the question of how much stronger the statements (ii), (iii), (iv) of our theorem are than statement (i). From (i) and Definitions 1.3, 1.4 we have

$$A_{i,j} = \text{Cov}[I\{X(s) > i\}, I\{X(t) > j\}] > 0,$$

where

$$\begin{aligned} I\{X(s) > i\} &= 1 & \text{for } X(s) > i, \\ &= 0 & \text{for } X(s) < i. \end{aligned}$$

Introduce $\pi_u(s) = \Pr[X(s)=u]$. We now have

$$\begin{aligned} A_{i,j} &= E[I\{X(s) > i\} \cdot I\{X(t) > j\}] - E[I\{X(s) > i\}] E[I\{X(t) > j\}] \\ &= \sum_{u>i} \pi_u(s) P_{u,>j}(s,t) - \sum_{u>i} \pi_u(s) \sum_{v=0}^k \pi_v(s) P_{v,>j}(s,t) = \end{aligned}$$

$$\begin{aligned}
 &= \sum_{u \geq i} \sum_{v=0}^k \pi_u(s) \pi_v(s) [P_{u, \geq j}(s, t) - P_{v, \geq j}(s, t)] \\
 &= \sum_{u \geq i} \sum_{v < i} \pi_u(s) \pi_v(s) [P_{u, \geq j}(s, t) - P_{v, \geq j}(s, t)]
 \end{aligned}$$

If we claim $A_{ij} > 0$ for all entrance distributions $\{\pi_u(s)\}_{u=0}^k$ for $X(s)$, then (iii) follows. This shows that we cannot obtain any better criterion than (iii) formulated in terms of the transition probabilities only, to ensure X to be associated in time.

In the binary case we have

$$A_{11} = \pi_1(s) \pi_0(s) [P_{1,1}(s, t) - P_{0,1}(s, t)] > 0.$$

Hence (i) implies (iii) unless $X(s) = 0$ a.s. or $X(s) = 1$ a.s.

Finally, it should be mentioned that a discrete time version of our theorem is proved along the same lines. In particular, if $\{X(t), t=0,1,2,\dots\}$ is a Markov chain with state space $\{0,1,\dots,k\}$ and matrix of transition probabilities $P(s,t)$, then X is associated in time if only $P(s,s+1) \in M$ for all $s \geq 0$.

Acknowledgement

The authors would like to thank Arne Bang Huseby for assistance on the linear programming problems that turned up when trying to generalize the proof of Esary and Proschan (1970).

REFERENCES

- BARLOW, R.E., and PROSCHAN, F. (1975), Statistical Theory of Reliability and Life Testing. Probability Models, New York: Holt, Rinehart and Winston.
- BLOCK, H.W., and SAVITS, T.S. (1982), "A Decomposition for Multistate Monotone Systems", Journal of Applied Probability, 19, 391-402.
- _____, and TING, M.-L. (1981), "Some Concepts of Multivariate Dependence", Communications in Statistics, A 10, 749-762.
- BØLVIKEN, E. (1982), "Probability Inequalities for the Multivariate Normal with Non-negative Partial Correlations", Scandinavian Journal of Statistics, 9, 49-58.
- ESARY, J.D., and PROSCHAN, F. (1968), "Generating Associated Random Variables", Boeing Scientific Research Laboratories Document D1-82-0696. Seattle, Washington.

_____, and PROSCHAN, F. (1970), "A Reliability Bound for Systems of Maintained, Interdependent Components", Journal of the American Statistical Association, 65, 329-338.

FUNNEMARK, E. and NATVIG, B. (1982), "Bounds for the Availability and Unavailability in a Fixed Time Interval for a Multistate System of Multistate, Maintained, Interdependent Components", to appear as a Research Report, University of Oslo, Dept. of Statistics.

JOHANSEN, S. (1977), "Product Integrals and Markov Processes", Preprint No.3, University of Copenhagen, Institute of Mathematical Statistics.

KARLIN, S., and TAYLOR, H.M. (1975), A First Course in Stochastic Processes, New York: Academic Press.

NATVIG, B. (1980), "Improved Bounds for the Availability and Unavailability in a Fixed Time Interval for Systems of Maintained, Interdependent Components", Advances in Applied Probability, 12, 200-221.

_____, (1982), "Two Suggestions of How to Define a Multistate Coherent System", Advances in Applied Probability, 14, 434-455.

PITT, L.D. (1982), "Positively Correlated Normal Variables are Associated", The Annals of Probability, 10, 496-499.

CORRECTION NOTE:

Odd Aalen has kindly pointed out that formula (2.6) is not generally true in its stated form. However, (2.7) remains true, so that our proof is not affected.

Replace lines 14-22 on page 6 by:

We then have

$$\tilde{P}(s,t) = \prod_{[s,t]} (I + Q(u)du). \quad (2.6)$$

The expression on the right hand side is a product integral. A general discussion of product integrals, containing the proof of (2.6), can be found in Johansen (1977). In the time-homogeneous case, where Q becomes a constant matrix, (2.6) reduces to the better known formula $\tilde{P}(s,t) = \exp((t-s)Q)$, see e.g. Karlin and Taylor (1975, p. 152). The following representation follows from Johansen's Theorem 2.5, utilizing that all the intensities are uniformly continuous on $[s,t]$: