Markov models for periodically tested components[1]

Bo Henry Lindqvist and Henriette Amundrustad
Department of Mathematical Sciences, Norwegian University of Science and Technology, Trondheim

ABSTRACT: Simple Markov models for the failure mechanisms of periodically tested standby components were studied in Hokstad & Frøvig (1996). In this paper we reconsider and generalize their models, presenting a systematic and flexible modeling approach. In addition we indicate how computer packages for symbolic mathematical computations, like Maple and Mathematica, are well suited for solving the models, yielding, for example, general expressions for interesting reliability measures.

1 INTRODUCTION

Component failures can often be classified into se-verity classes, for example critical or degraded failures. Simple Markov models for the failure mechanism of a component, suitable for making statistical inferences based on reliability data bases, were suggested and studied in Hokstad and Frøvig (1996). Their study involved dormant (hidden) failures of periodically tested standby components. In this paper we reconsider and generalize their models, presenting a systematic and flexible modeling approach which enables easy computation of key figures like mean time between failures and mean fractional deadtime, here called critical safety unavailability. For simplicity we shall in the sequel refer to the paper Hokstad & Frøvig (1996) as HF.

Our approach involves a continuous time Markov model for the component state when time runs *between* testing epochs, and in addition two discrete time Markov chains for the states of the component reported immediately before and after each test, respectively. As will be seen, the given framework also allows in an easy manner the potentially useful extension to modeling of incomplete repairs or maintenance actions.

It will be briefly indicated how computer packages for symbolic mathematical computations, like Maple V (Waterloo Maple, Inc., Ontario) and Mathematica (Wolfram Research, Inc., Champaign, Illinois) easily can be used for solving the models. This aspect of the paper is much inspired by papers by Aalen (1993, 1995), who studied models much similar to the ones considered here, but with emphasis on applications in medicine. In the present paper we have used Maple for computations and for making graphical plots.

A recent reference summarizing some of the previous papers on the subject, and suggesting some interesting approximations, is Hokstad & Jensen (1998). A case study using the models of HF is presented in Haugen et al. (1997). Extension to models involving failures other than dormant failures, as considered in Haugen et al. (1996), can be obtained by an appropriate modification of the models considered in the present paper.

2 GENERAL MODEL

We consider a standby component observed from time 0, with testing and preventive maintenance, abbreviated PM, performed periodically at times $\tau, 2\tau, 3\tau, \ldots$, called PM epochs. Here $0 < \tau \leq \infty$ is the length of what we shall call the PM interval.

The general model is presented next, while the models considered in HF are studied in some detail in Section 5. It may be helpful to consult that section while reading the present general description.

Let $X(t) \in S$ denote the state of the component at time t, where the set S of possible states

[1]To appear in proceedings of European Safety and Reliability Conference, 17-19 June 1998 (ESREL '98), Trondheim, Norway.

is finite. It is assumed that $X(t)$ behaves like a time homogeneous Markov chain as long as time runs inside PM intervals, i.e. inside time intervals $n\tau \leq t < (n+1)\tau$ for $n = 0, 1, \ldots$. This Markov chain is governed by an infinitesimal intensity matrix A, where the entry a_{jk} of A for $j \neq k$ is the transition intensitiy from state j to state k, see Taylor & Karlin (1984, p. 254). An example of an intensity matrix A is given by (3), an illustration of which is provided by the state diagram in Figure 2. Let

$$P_{jk}(t) = P(X(t) = k | X(0) = j);\ j, k \in S,\ t > 0$$

denote transition probabilities for the Markov chain governed by A and let

$$P(t) = (P_{jk}(t); j, k \in S)$$

be the corresponding transition matrix.

In order to specify the effect of maintenance and repair at PM epochs, we next introduce for $n = 1, 2, \ldots$,

$$Y_n = X(n\tau-) \equiv \lim_{t \uparrow n\tau} X(t)$$

which is the state of the component *immediately before* the nth PM epoch. The effect of PM at time $n\tau$ is to change the state of the component from Y_n to Z_n according to a transition matrix $R = (R_{jk})$, where

$$P(Z_n = k | Y_n = j) = R_{jk}\ ;\ j, k \in S.$$

Moreover, given Y_n it is assumed that Z_n is independent of all transitions of the component state before time $n\tau$.

The definitions of the Y_n and Z_n are illustrated in Figure 1.

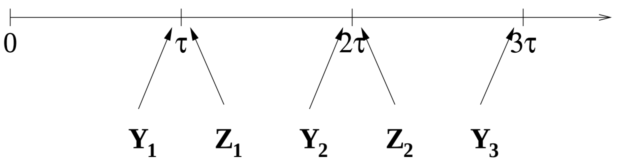

Figure 1: The definition of Y_n and Z_n.

The model description is completed by defining the initial state of the Markov chain $X(t)$ running inside the PM interval $[n\tau, (n+1)\tau)$ to be $X(n\tau) \equiv Z_n$ $(n = 0, 1, \ldots)$, where Z_0 is the intital state of the component, usually the perfect state in S. It is, furthermore, assumed that the Markov chain $X(t)$ on $[n\tau, (n+1)\tau)$, given its initial state Z_n, is independent of all transitions occurring before time $n\tau$.

Let the distribution of $Z_0 \equiv X(0)$ be denoted $\rho = (\rho_j; j \in S)$, where $\rho_j = P(Z_0 = j)$.

Then for any $k \in S$,

$$P(Y_1 = k) = P(X(\tau-) = k)$$

$$= \sum_{j \in S} P(X(\tau-) = k | X(0) = j) P(X(0) = j)$$

$$= \sum_{j \in S} \rho_j P_{jk}(\tau) = [\rho P(\tau)]_k$$

Thus the distribution of Y_1 is given by the vector-matrix product $\rho P(\tau)$. Further, for $n \geq 1$,

$$P(Y_{n+1} = k | Y_n = j)$$

$$= \sum_{\ell \in S} P(Y_{n+1} = k | Z_n = \ell, Y_n = j) P(Z_n = \ell | Y_n = j)$$

$$= \sum_{\ell \in S} P_{\ell k}(\tau) R_{j\ell} = [RP(\tau)]_{jk}.$$

It follows that $Y_1, Y_2, \ldots$ is a discrete time Markov chain on S with transition matrix

$$Q = RP(\tau).$$

On the other hand,

$$P(Z_{n+1} = k | Z_n = j)$$

$$= \sum_{\ell \in S} P(Z_{n+1} = k | Y_{n+1} = \ell, Z_n = j)$$

$$\times P(Y_{n+1} = \ell | Z_n = j)$$

$$= \sum_{\ell \in S} P_{j\ell}(\tau) R_{\ell k} = [P(\tau)R]_{jk}.$$

Thus, our assumptions imply that $Z_0, Z_1, \ldots$ is a discrete time Markov chain on S with transition matrix

$$T = P(\tau)R.$$

3 RELIABILITY MEASURES

3.1 *Average rate of critical failures*

Let $\pi = (\pi_j, j \in S)$ be the stationary distribution of the Markov chain $Y_1, Y_2, \ldots$, i.e. π is the unique probability vector satisfying the equation

$$\pi Q \equiv \pi RP(\tau) = \pi.$$

For any subset $G \subset S$, define $\pi_G = \sum_{j \in G} \pi_j$. This is the expected relative number of PM epochs, in the long run, where the component is found to be in G. Moreover, $1/\pi_G$ is the mean time, in the long run, between visits to G (measured with time unit τ). These facts are well known from the theory of Markov chains (Taylor & Karlin 1984).

Let in the following G be the subset of S defining the *critical failure states* of the component. Then as in HF we define the *mean time between critical failures* to be

$$MTBF_{crit} = \tau/\pi_G$$

and the *average rate of critical failures* to be

$$\lambda_{crit} = 1/MTBF_{crit} = \pi_G/\tau. \tag{1}$$

3.2 *Critical safety unavailability*

Consider a PM interval $[n\tau, (n+1)\tau)$. The expected relative amount of time in this interval that the component is in a critical state, i.e. in G, is

$$U_n = \frac{1}{\tau}\int_{n\tau}^{(n+1)\tau} P(X(t) \in G)dt$$

By our assumptions, $X(t)$ behaves in the interval $[n\tau, (n+1)\tau)$ in the same manner as if it was run in the interval $[0, \tau)$ and started in state Z_n. Thus

$$U_n = \frac{1}{\tau}\int_0^{\tau} \sum_{j \in S} P_{jG}(t)P(Z_n = j)dt$$

where $P_{jG}(t) = \sum_{k \in G} P_{jk}(t)$.

Letting n tend to infinity, the $P(Z_n = j)$ tend to the limiting values γ_j defined from the stationary distribution $\gamma = (\gamma_j)$ of the Markov chain $Z_0, Z_1, \ldots$. This distribution is found by solving the equations

$$\gamma T \equiv \gamma P(\tau)R = \gamma.$$

Following the terminology of HF we shall define the *Critical Safety Unavailability* (CSU) of the component by

$$CSU = \lim_{n \to \infty} U_n$$

$$= \frac{1}{\tau}\int_0^{\tau} \sum_{j \in S} P_{jG}(t)\gamma_j dt = \sum_{j \in S} \gamma_j Q_j \tag{2}$$

where

$$Q_j = \frac{1}{\tau}\int_0^{\tau} P_{jG}(t)dt$$

is the critical safety unavailability given that the component state is j at the beginning of the PM interval.

4 USING MAPLE

The formulae and function graphs of the present paper have been obtained using the mathematical computer package Maple V Release 4. Maple, as well as the competing package called Mathematica, are in particular able to perform symbolic manipulations and are thus well suited for treating models like the ones considered here.

For example, it is well known (Taylor and Karlin 1984, p. 256) that the transition matrix $P(t)$ can be found from the infinitesimal transition matrix A by using a matrix exponential function, $P(t) = e^{tA}$. Using Maple's package for linear algebra we are able to get symbolic expressions for the entries of $P(t)$ by the command `exponential(t*A)`. This generally works, however, only when A, possibly after a renumbering of the states, can be written as an upper or lower triangular matrix. Fortunately, this is the case for the type of models considered in this paper, where no repair is done between PM epochs. Otherwise, the computation of $P(t)$ with Maple is more involved and may be possible only numerically (see Amundrustad 1997 for general recipes using Maple, and Aalen 1993 for using Mathematica).

5 HOKSTAD & FRØVIG'S FAILURE MODEL

As an illustration we shall reconsider the most general failure model of HF, namely their Failure Mechanism III. Here the state space is

$$S = \{O, D, K^I, K^{II}\},$$

where $O =$ the component is as good as new, $D =$ the component has a failure classified as *degraded* (noncritical), $K^I =$ the component has a failure classified as *critical*, caused by a *sudden shock*, $K^{II} =$ the component has a failure classified as *critical*, caused by the *degradation process.*

It is assumed that the Markov chain $X(t)$ is defined by the state diagram of Figure 2, and thus has infinitesimal transition matrix

$$A = \begin{bmatrix} -\lambda_d - \lambda_k & \lambda_d & \lambda_k & 0 \\ 0 & -\lambda_k - \lambda_{dk} & \lambda_k & \lambda_{dk} \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \tag{3}$$

Note that both K^I and K^{II} are absorbing states.

The model assumes that no repairs are done in the time intervals between PM epochs. Moreover, since A is upper triangular, we can obtain

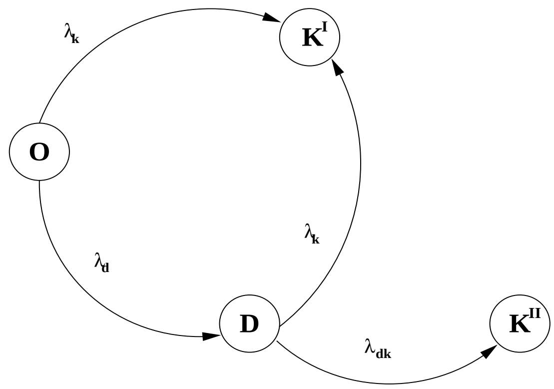

Figure 2: State diagram for the failure mechanism of HF.

$P(t) = e^{tA}$ by using Maple (Section 4). It is clear that $P(t)$ can be written

$$\begin{bmatrix} P_{OO}(t) & P_{OD}(t) & P_{OK^I(t)} & P_{OK^{II}}(t) \\ 0 & P_{DD}(t) & P_{DK^I}(t) & P_{DK^{II}}(t) \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \tag{4}$$

while the expressions for the entries obtained by Maple are

$$P_{OO}(t) = e^{-(\lambda_k+\lambda_d)t}$$

$$P_{OD}(t) = \frac{\lambda_d}{\lambda_d - \lambda_{dk}} \left(e^{-(\lambda_k+\lambda_{dk})t} - e^{-(\lambda_k+\lambda_d)t}\right)$$

$$P_{OK^I}(t) = \frac{\lambda_k(\lambda_d + \lambda_k + \lambda_{dk})}{(\lambda_d + \lambda_k)(\lambda_k + \lambda_{dk})}$$

$$+ \frac{\lambda_k\lambda_{dk}}{(\lambda_d - \lambda_{dk})(\lambda_d + \lambda_k)} e^{-(\lambda_k+\lambda_d)t}$$

$$+ \frac{\lambda_k\lambda_d}{(\lambda_{dk} - \lambda_d)(\lambda_k + \lambda_{dk})} e^{-(\lambda_k+\lambda_{dk})t}$$

$$P_{OK^{II}}(t) = \frac{\lambda_d\lambda_{dk}}{(\lambda_d + \lambda_k)(\lambda_k + \lambda_{dk})}$$

$$+ \frac{\lambda_d\lambda_{dk}}{(\lambda_d - \lambda_{dk})(\lambda_d + \lambda_k)} e^{-(\lambda_k+\lambda_d)t}$$

$$+ \frac{\lambda_d\lambda_{dk}}{(\lambda_{dk} - \lambda_d)(\lambda_k + \lambda_{dk})} e^{-(\lambda_k+\lambda_{dk})t}$$

$$P_{DD}(t) = e^{-(\lambda_k+\lambda_{dk})t}$$

$$P_{DK^I}(t) = \frac{\lambda_k}{\lambda_k + \lambda_{dk}} \left(1 - e^{-(\lambda_k+\lambda_{dk})t}\right)$$

$$P_{DK^{II}}(t) = \frac{\lambda_{dk}}{\lambda_k + \lambda_{dk}} \left(1 - e^{-(\lambda_k+\lambda_{dk})t}\right)$$

In practice it is of great interest to quantify the effect of doing preventive maintenance. HF in particular compare the cases where (i) all failures are repaired at PM epochs and (ii) where only critical failiures are repaired at PM epochs. These cases are reconsidered below in the framework of the present paper. Generalizations to certain kinds of *imperfect repair* are also considered.

5.1 *All failures are repaired at PM epochs*

We shall later call this the standard model. Because the PM always returns the component back to state O, we have

$$R = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{bmatrix} \tag{5}$$

This implies that the transition matrix $Q = RP(\tau)$ of $Y_1, Y_2, \ldots$ has all rows equal to the first row of (4), with t replaced by τ. It is clear that this row is also the stationary distribution π for this case. Thus by (1),

$$\lambda_{crit} = \frac{P_{OK^I}(\tau) + P_{OK^{II}}(\tau)}{\tau}$$

$$= \frac{\lambda_{dk} - \lambda_d - \lambda_{dk}e^{-(\lambda_d+\lambda_k)t} + \lambda_d e^{-(\lambda_k+\lambda_{dk})t}}{\tau(\lambda_{dk} - \lambda_d)}$$

which was obtained in HF by using the geometric distribution. The Markov chain approach of the present paper is, however, most useful when studying repair strategies leading to less trivial matrices R than the one in (5).

The critical safety unavailability, CSU, is particularly easy to obtain in the present case, because $P(Z_n = O) = 1$ for all i and hence $\gamma = (1, 0, 0, 0)$. It follows from the general expression (2) that

$$CSU = Q_O = \int_0^\tau (P_{OK^I}(t) + P_{OK^{II}}(t))dt$$

which was computed by HF and is found from their formula (12).

5.2 *All critical failures are repaired at PM epochs*

If degraded failures are not repaired, then the appropriate R matrix is

$$R = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{bmatrix}$$

More generally we shall consider an extension of the models in HF by assuming that all critical failures are repaired, while degraded failures are repaired with probability $1-r$ and remain unrepaired with probability r, $0 \le r \le 1$. The repair strategy is thus determined by the parameter r.

More precisely, we shall consider

$$R = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 1-r & r & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{bmatrix}$$

Computing $Q = RP(\tau)$, solving the equations $\pi Q = \pi$ and using (1), we get the average rate of critical failures,

$$\begin{aligned} \lambda_{crit}(\tau, r) &= \frac{\pi_{K^I} + \pi_{K^{II}}}{\tau} \\ &= \frac{P_{OK^I} + P_{OK^{II}}}{\tau} \qquad (6) \\ &- \frac{rP_{OD}(P_{OK^I} - P_{DK^I} + P_{OK^{II}} - P_{DK^{II}})}{\tau(1 + rP_{OD} - rP_{DD})} \end{aligned}$$

where for simplicity of writing we have put $P_{jk} \equiv P_{jk}(\tau)$, $j, k \in S$.

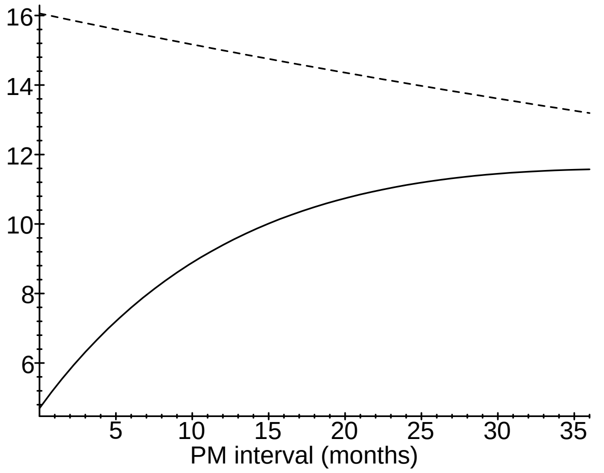

Figure 3: Average rate of critical failures (failures per 10^6 hours) for $r = 0$ (solid curve) and $r = 1$ (dashed curve), plotted as a function of the length τ of the PM interval (months).

HF study the effect of preventive maintenance by comparing $\lambda_{crit}(\tau, 0)$, which is the average rate of failures for the standard model, and the so called *naked failure rate*, $\lambda_{crit}(\tau, 1)$, which is the failure rate obtained when degraded failures are not repaired. The difference is illustrated in Figure 3, where the corresponding average rates of failures (6) are plotted as functions of the length of the PM intervals, τ (months), for a numerical example from HF involving a gas detector, where $\lambda_d = 12.9$, $\lambda_{dk} = 130$, $\lambda_k = 4.7$, all with unit *failures per* 10^6 *hours*. (Note that the naked failure rate is decreasing with τ. This may seem unreasonable, but is caused by the fact that failures are reported only at the end of each PM interval. Thus for long PM intervals the rate of failures will appear too low, because the component can experience at most one failure per interval).

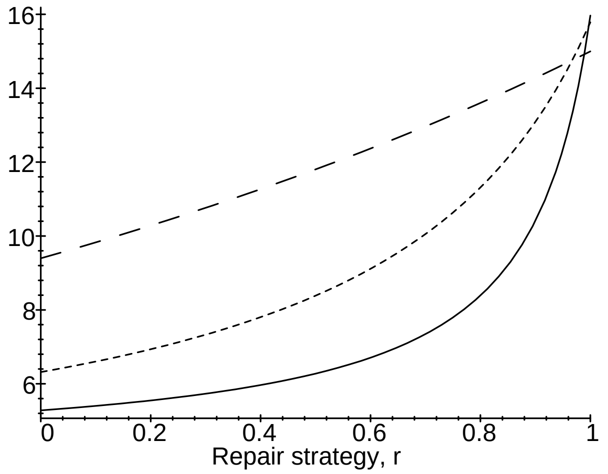

Figure 4: Average rate of critical failures for PM-intervals of 1 month (solid curve), 3 months (short dashes) and 12 months (long dashes), plotted as a function of repair strategy r.

For the same numerical values of the parameters we have plotted, Figure 4, the average rate of failures (6), as function of the repair strategy r, for three different lengths of the PM interval. Especially for the short PM intervals it is seen that the increase of critical failure rate is only moderate when r is increased from 0 to, say, 0.5, but that the increase is much higher when r is increased from 0.5 to 1.

In order to obtain an expression for the critical safety unavailability, CSU, we need to compute the matrix $T = P(\tau)R$ and solve for the stationary distribution γ. We get (obviously) that $\gamma_{K^I} = \gamma_{K^{II}} = 0$, and furthermore,

$$\gamma_O = \frac{1 - rP_{DD}(\tau)}{1 + rP_{OD}(\tau) - rP_{DD}(\tau)}$$

$$\gamma_D = \frac{rP_{OD}(\tau)}{1 + rP_{OD}(\tau) - rP_{DD}(\tau)}$$

While Q_O, as mentioned in the previous subsection, is found in HF, we need in addition to compute

$$Q_D = \frac{1}{\tau}\int_0^\tau (P_{DK^I}(t) + P_{DK^{II}}(t))dt$$

$$= 1 - \frac{1}{\tau}\int_0^\tau P_{DD}(t)dt$$

$$= 1 - \frac{1 - e^{-(\lambda_k+\lambda_{dk})\tau}}{\tau(\lambda_k + \lambda_{dk})}$$

which can be obtained very easily from Maple.

Now we are able to compute CSU from (2) for varying τ and r, using

$$CSU(\tau, r) = \gamma_O Q_O + \gamma_D Q_D \tag{7}$$

Figure 5 shows the effect of PM by plotting CSU from (7), in per cent of time, as a function of the repair strategy r for two different PM intervals, 1 month and 3 months, with the same parameter values as before. It is remarkable that the curves increase only moderately until r is sufficiently near 1, and then rise quite rapidly. Thus it seems that any PM strategy with r less than, say, 0.5, is much better than a strategy where degraded failures are not repaired at all ($r = 1$), and not too much worse than one where all failures are repaired at PM epochs ($r = 0$).

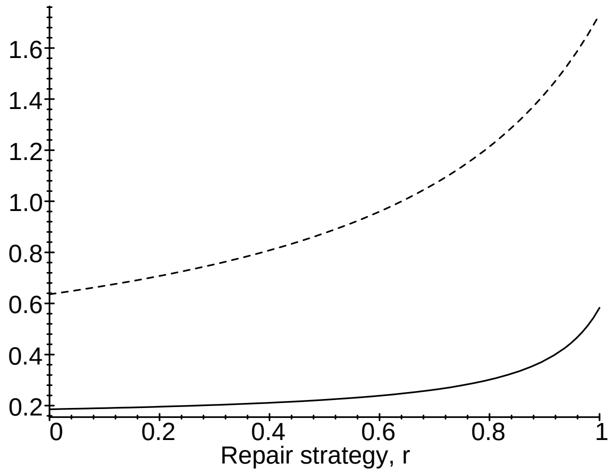

Figure 5: Critical safety unavailability, CSU (per cent), plotted as a function of repair strategy r, for PM intervals of length 1 month (solid curve) and 3 months (dashed curve).

HF suggest a measure of the effect of PM given by $I_{PM}(\tau) = 1 - \lambda_{crit}(\tau, 0)/\lambda_{crit}(\tau, 1)$. Here we suggest a different, but reasonable measure,

$$E_{PM}(\tau) = \frac{CSU(\tau, 1)}{CSU(\tau, 0)} \tag{8}$$

which expresses the relative increase in critical safety unavailability when going from the standard PM strategy to no PM. A plot of the measure (8) as a function of τ (months) is given in Figure 6.

5.3 *A general imperfect repair model*

A more general imperfect repair model could be defined by

$$R = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 1-r & r & 0 & 0 \\ 1-r_{k1} & 0 & r_{k1} & 0 \\ 1-r_{k2} & 0 & 0 & r_{k2} \end{bmatrix}$$

Here r has the same meaning as before, while $1 - r_{k1}$ is the probability of successful repair of a K^I failure and $1-r_{k2}$ is the similar for K^{II}. Again we can compute average rates of critical failures by solving $\pi RP(\tau) = \pi$ and using (1), and critical safety unavailability by solving $\gamma P(\tau)R = \gamma$ and using (2).

Note that imperfect repair models are well known in the literature of repairable systems. For example the model of Brown and Proschan (1983) specifies a probability p of the occurrence of a perfect repair and $1 - p$ of a minimal repair.

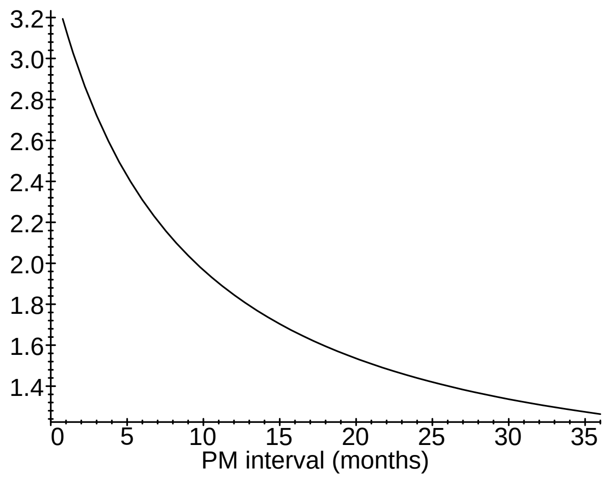

Figure 6: The effect of PM, $E_{PM}(\tau)$, plotted as a function of the length τ of the PM interval (months).

6 CONCLUSIONS

The paper defines a general framework for studying the reliability of components which are tested periodically. The basic parameters are the infinitesimal transition matrix A associated with the failure mechanism, and the one step transition matrix R representing the maintenance strategy at test epochs.

It is shown in particular how expressions for the average rate of critical failures and critical safety unavailability can be obtained for given values of A and R. A possible practical application of the approach is the comparison of different PM strategies, as manifested by the matrix R, for a fixed failure mechanism defined by A.

The approach of the present paper is motivated and inspired by the paper HF. However, while HF in addition to stochastic modeling also included statistical estimation of parameters based on data, we have given no attention to statistical inference here.

REFERENCES

Aalen, O.O. 1993. Phase-type distributions: Computer algebra and a simple mixing model. Section for Medical Statistics, University of Oslo.

Aalen, O.O. 1995. Phase-type distributions in survival analysis. *Scandinavian Journal of Statistics.* 22:447-463.

Amundrustad, H. 1997. Phase-type distributions and Markov modeling in lifetime data analysis. (In Norwegian). Project work. Department of Mathematical Sciences, Norwegian University of Science and Technology.

Brown, M. & F. Proschan 1983. Imperfect repair. *Journal of Applied Probability.* 20:851-859.

Haugen, K., P. Hokstad & H. Langseth 1996. Estimation of the average rate of occurrence of failures. In: Cacciabue, P.C. & I.A. Papazoglou (ed.) *Probabilistic Safety Assessment and Management*, pp. 1452-1457. London: Springer.

Haugen, K., P. Hokstad & H. Sandtorv 1997. The analysis of failure data in the presence of critical and degraded failures. *Reliability Engineering and System Safety.* 58:97-107.

Hokstad, P. & A.T. Frøvig 1996. The modelling of degraded and critical failures for components with dormant failures. *Reliability Engineering and System Safety.* 51:189-199.

Hokstad, P. & R. Jensen 1998. Predicting the failure rate for components that go through a degradation state. To appear in proceedings of ESREL'98.

Taylor, H.M. & S. Karlin 1984. *An introduction to stochastic modeling.* London: Academic Press.