

Further Thoughts on Markov Analysis in Reliability Modelling

An Addendum to “Repairable Redundant Systems and the Markov Fallacy”

by [W G Gulland \(4-sight Consulting\)](#)

1 Why Repair is NOT a Markov Process

The basic contention of my earlier paper ([Ref. 1](#)) and of this article is that repair is not and never can be a Markov process. One of the essential attributes of a Markov process is "that the probability of being in a state can be determined solely from knowledge of the previous state", and is therefore independent of all states prior to that immediately preceding state (the system has no memory). For example, in nuclear decay, which is a Markov process *par excellence*, the probability that a nucleus is species Z is a result of its immediately previous state being species Y, not of its having been species X before it became species Y.

In contrast, if I need to replace the clutch to repair my car, I must first remove the gearbox. The gearbox remains out of my car while I remove and replace the clutch, and until I put it back in. The state of my "car-under-repair" system remains without a gearbox, and dependent on the past state "gearbox removed", while it goes through the states of "clutch removed" and "clutch replaced", until I put the gearbox back. The "car-under-repair" system has memory (fortunately it "remembers" that the gearbox has been removed until it is deliberately replaced!), and, by definition, repairing it cannot be a Markov process. If I incorporate this non-Markov process into the larger process of the "operate / fail / repair / operate" cycle, the larger process does not and, indeed, cannot become a Markov process.

Markov developed his theory to model Brownian motion of gases, where particles suddenly change direction. Suddenness / instantaneousness is an essential characteristic of Markov processes which follows from the basic definition, since any process which does not occur instantaneously in a continuous-time system, or between one time point and the next in a discrete-time system, must effectively be going through a series of states, with the state at any instant dependent not only on the immediately preceding state but on other earlier states. If this were not so, the process could be broken down into a series of sub-processes which would themselves be Markov processes. The mental model implicit in various texts on the subject of Markov analysis is that items jump from the failed state to the repaired state. For example:

- “*Transition matrix method.* This method is applicable to systems with exponential component failure and repair time distributions.” (Ref. 2)
- “It is important to remember one rule with Markov analysis, namely, that the probabilities of changing state are dependent only on the state itself. In other words, the probability of failure or repair is not dependent on the past history of the system.” (Ref. 3)
- “The 2λ and 2μ terms indicate that two components are available for failure or repair respectively in the next increment of time and that one of the two can fail or be repaired, but not both in that interval.” (Ref. 4)

It will be observed that the first reference even requires that item downtimes are exponentially distributed. The mental model involved here must be a cycle of "operate / fail

(suddenly and at random, but at a constant rate in a sample of infinite size) / not-operate / recover (suddenly and at random, but at a constant rate in a sample of infinite size) / operate". This would indeed be a Markov system, but it is not even close to being an accurate description of the cycle in real repairable systems. Also, as stated in Ref. 1, analysis of these systems based on probability theory demonstrates that the only restriction on the distribution of item downtimes is that it does not vary with time, so that the mean value (MDT) is constant. Apart from that restriction, the distribution of downtimes can have any feasible shape. Typically, downtimes may be grouped in a few bands, corresponding to the times to replace major components of the item. Certainly no maintenance manager worth his salt would tolerate a situation where downtimes could in rare cases extend to infinity, and are predictable only to the extent of saying that they are exponentially distributed. Nor would he accept recoveries happening at random; he would expect them to happen within X hours of the failure.

The conclusion must be that repair fails the conditions for being a Markov process on a number of counts.

2 Practical Significance

The practical significance of this issue is limited to:

- Identifying that in most cases almost the full benefit of redundancy can be obtained by employing a single repair crew (this result only starts to fail once $\lambda \cdot \text{MDT} > \sim 0.03$, and the error is only significant for $\lambda \cdot \text{MDT} > \sim 0.3$ - or $\text{MDT} > \sim 0.3 \times \text{MTBF}$, see formula for error developed below). In practice, it is doubtful whether anyone actually employs extra personnel on the basis of Markov theory, but, if they do, they could save themselves some money.
- Saving authors writing textbooks and programmers writing software packages based on erroneous theory, and reliability practitioners producing erroneous answers to problems. However, other sources of error, such as common cause failure rates, are in practice much more significant.
- Providing an object lesson in how easy it is to get the logic of a probability based problem wrong, and the importance whenever possible of attacking such a problem with more than one approach, and ideally by more than one analyst (though the conventional Markov analysis seems to have been accepted without question by a remarkable number of people for a remarkably long time).

3 What Remains of Markov Analysis

State transition matrices and diagrams remain useful tools for analysing repairable redundant systems with immediately detected failures, provided that the correct terms are used in these models. The correct terms are those derived from probability theory, and not those derived from the conventional Markov analysis. The development in Appendix 3 of Ref. 6 of these tools for analysing proof-tested systems, again on the basis of probability theory, is also very useful.

Because these tools are so closely associated with conventional Markov analysis, using them could perhaps be described as "Pseudo-Markov analysis" or "Quasi-Markov analysis" to recognise the similarity to Markov analysis, but also to emphasise that unmodified Markov analysis is not valid.

Personally I find the state transition diagram particularly useful for visualising how the probabilities of the various possible states of a system are related to each other. As I attempted to demonstrate in Ref. 1, this is even more apparent in the case where $\lambda \ll 1/MDT$ (or $\lambda \cdot \mu \ll 1$), where the results for all the states can be developed in cascade starting with the assumption that the probability of State 0 (zero items failed) = ~ 1 .

4 Derivation of Formulae for Availability / Probability of Failure on Demand

For proof-tested systems, multiple failures are on average equally spaced through the proof-test interval; for immediately detected failures, subsequent failures are on average equally spaced through the downtime of the first item to fail. This is essentially the explanation given in Ref. 3 for proof-tested systems and used there to derive the results for availability / probability of failure on demand, etc. However, both sets of results for unavailability (Q_{system}) can be derived independently of these explanations, by integrating the unavailability function for proof-tested systems, by applying the standard probability theory for coincident independent events (the multiplication rule) to immediately detected failures, and then applying combination theory in both cases. The results are:

- For proof-tested systems, n out of m required to operate,

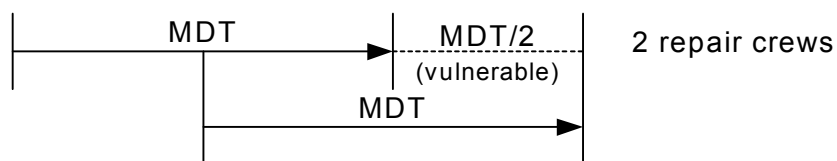
$$Q_{system} = {}^m C_{m-n+1} \cdot (\lambda \cdot \tau)^{m-n+1} / (m-n+2)$$
- For immediately detected failures, n out of m required to operate,

$$Q_{system} = {}^m C_{m-n+1} \cdot (\lambda \cdot MDT_{unit})^{m-n+1}$$

The theorems on average spacing of failures can then be derived as a corollary of those proofs, to provide a physical interpretation of the results. I know of no other separate proof of these theorems on average spacing of failures, though the cases of a single failure in a proof-tested system, and of the second failure in a system where failures are immediately revealed, are essentially self-evident, and the general results seem intuitively correct.

5 Error in Formulae / Incremental Unavailability if only a Single Repair Crew

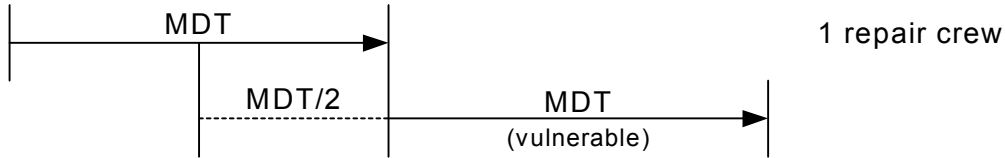
5.1 Given a 1oo2 system and 2 repair crews



- If there are 2 repair crews the average period of vulnerability to re-failure of the first failed-and-repaired item while the second failure is being repaired is $MDT/2$.
- Therefore the probability of a coincident re-failure given that there has already been a coincident failure is $\lambda MDT/2$.
- The frequency of first coincident failures is $\lambda^2 MDT$.
- Therefore the frequency of coincident re-failures is $\lambda^3 MDT^2/2$.
- On average the coincident re-failure will occur half way through the period of vulnerability, i.e. at $MDT/4$ after the first repair is completed, so that the

unavailability attributable to re-failures is $\lambda^3\text{MDT}^3/8$ (though this is of course already included in the basic unavailability of $\lambda^2\text{MDT}^2$, derived by probability theory, for a system with 2 repair crews where repair of a failed item always starts immediately it fails).

5.2 Given a 1oo2 system and a single repair crew



- If there is a single repair crew the average period of vulnerability to re-failure of the first failed-and-repaired item while the second failure is being repaired is MDT.
- Therefore the probability of a coincident re-failure given that there has already been a coincident failure is λMDT
- The frequency of first coincident failures is $\lambda^2\text{MDT}$.
- Therefore the frequency of coincident re-failures is $\lambda^3\text{MDT}^2$.
- On average the coincident re-failure will occur half way through the period of vulnerability, i.e. at $\text{MDT}/2$ after the first repair is completed, so that the unavailability attributable to re-failures is $\lambda^3\text{MDT}^3/2$.

5.3 Incremental Unavailability

Therefore the incremental unavailability is given by:

$$Q_i = \lambda^3\text{MDT}^3/2 - \lambda^3\text{MDT}^3/8 = 3\lambda^3\text{MDT}^3/8 = 0.375\lambda^3\text{MDT}^3$$

Basic unavailability, Q , calculated assuming 2 repair crews is $\lambda^2\text{MDT}^2$, so:

$$Q_i/Q = 0.375\lambda\text{MDT}$$

and:

$$\begin{aligned} Q_i/Q &< 0.1 \text{ (10\%)} \text{ if } \lambda\text{MDT} < 0.267 \\ Q_i/Q &< 0.01 \text{ (1\%)} \text{ if } \lambda\text{MDT} < 0.0267 \end{aligned}$$

Typically $\lambda\text{MDT} < 0.01$, and:

$$Q_i/Q < 0.00375 \text{ (0.375\%)}$$

6 Range of Validity of the Results

It follows from Ref. 1 and Section 5 above that, no matter how many repair crews there actually are, provided there is at least one and $\lambda\text{MDT} < \sim 0.3$, correct or nearly correct results for immediately detected failures can be obtained from the formulae derived with the conventional Markov analysis by setting the number of repair crews in that analysis to the number of failures required to render the system unavailable, e.g. the number of repair crews should be set to 3 for a 2oo4 system.

7 Reflection on the Use of Computer Programs

It is dangerous to use computer programmes to unthinkingly model complicated systems. Markov analysis has been used in this way. This danger is common to all reliability calculations using a computer, which will produce the wrong answer if the underlying model is wrong, and it is important for the user to be able to perform some simple manual checks. The user must have insight into the build-up and main contributors to the final answer, and the ability to perform sensitivity checks on the solution. The use of reliability block diagrams (RBDs) or fault trees is not necessarily a completely satisfactory alternative to Markov analysis because at least some fault tree / RBD programs incorporate the Markov results in "super-blocks".

8 Additional Comment

It is instructive to add a further column to Table 2 of Ref. 6 for the formulae based on standard probability theory for coincident independent events (the multiplication rule) and combination theory applied to faults found at proof test, assuming random proof testing of redundant elements (Ref. 5). In this case the result for unavailability are given by the formula:

- $Q_{\text{system}} = {}^m C_{m-n+1} \cdot (\lambda \cdot \tau/2)^{m-n+1}$

These results are shown in Column 2A of the table below:

1 Configuration	2 PFD Using Probability Analysis (Appendix 1)	2A PFD based on probabilities if proof testing is random (MDT _{unit} = T/2)	3 PFD Using Markov Standard Formula (Appendix 3)	4 PFD Using Markov Modified Formula (Appendix 3)	5 Ratio Factor For Markov Standard Results (column 3) to Correct Formulae (column 2)
1oo2	$\lambda^2 T^2/3$	$\lambda^2 T^2/4$	$\lambda^2 T^2/2$	$\lambda^2 T^2/3$	1.5
2oo2	λT	λT	λT	λT	1
1oo3	$\lambda^3 T^3/4$	$\lambda^3 T^3/8$	$3\lambda^3 T^3/4$	$\lambda^3 T^3/4$	3
2oo3	$\lambda^2 T^2$	$3\lambda^2 T^2/4$	$3\lambda^2 T^2/2$	$\lambda^2 T^2$	1.5
3oo3	$3\lambda T/2$	$3\lambda T/2$	$3\lambda T/2$	$3\lambda T/2$	1
1oo4	$\lambda^4 T^4/5$	$\lambda^4 T^4/16$	$3\lambda^4 T^4/2$	$\lambda^4 T^4/5$	7.5
2oo4	$\lambda^3 T^3$	$\lambda^3 T^3/2$	$3\lambda^3 T^3$	$\lambda^3 T^3$	3
3oo4	$2\lambda^2 T^2$	$3\lambda^2 T^2/2$	$3\lambda^2 T^2$	$2\lambda^2 T^2$	1.5
4oo4	$2\lambda T$	$2\lambda T$	$2\lambda T$	$2\lambda T$	1

Clearly the results obtained by application of standard Markov theory bear very little similarity to the results based on probability analysis, either for co-ordinated or for random proof testing.

9 Acknowledgements

The author acknowledges the help given to him in preparing these comments by his colleague Dr. Saeed Fararoy, who performed a literature search which identified several of the references quoted, and who also made valuable comments on early drafts of my original

paper; also his colleague Dr. Tony Foord who reviewed not only the original paper (Ref. 1) but also these comments.

10 References

1. ["Repairable Redundant Systems and the Markov Fallacy"](#)
2. BS5760 : Part 2, Section 8.1.8.1.
3. "Reliability, Maintainability and Risk", David J. Smith, ISBN 0-7506-0854-4.
4. "Reliability Evaluation of Engineering Systems: Concepts & Techniques", Roy Billington & Ronald N. Allan, ISBN 0-306-44063-6
5. "Effect of Proof Testing", David J. Smith, The Journal of the Safety & Reliability Society, Volume 21, No. 3
6. "Reliability Assessments of Repairable Systems – is Markov Modelling Correct?", K G L Simpson and M Kelly (Silvertech Safety Consultancy Ltd)