

Repairable Redundant Systems and the Markov Fallacy

by **W G Gulland (4-sight Consulting)**

Summary

Markov analysis is presented in reliability text books and incorporated in a number of computer programs for reliability calculations as a way of dealing with repairable systems. This paper demonstrates that the application of Markov analysis to this problem is flawed, and leads to erroneous results. It proposes an alternative method of dealing with such systems based on the application of standard probability theory and a consideration of the practical situation. It is recognised that the argument is largely academic because common cause failures of such systems are liable to dominate independent failures in almost all real cases.

1 Application of Markov Analysis to Repairable Systems

A Markov process is defined as:

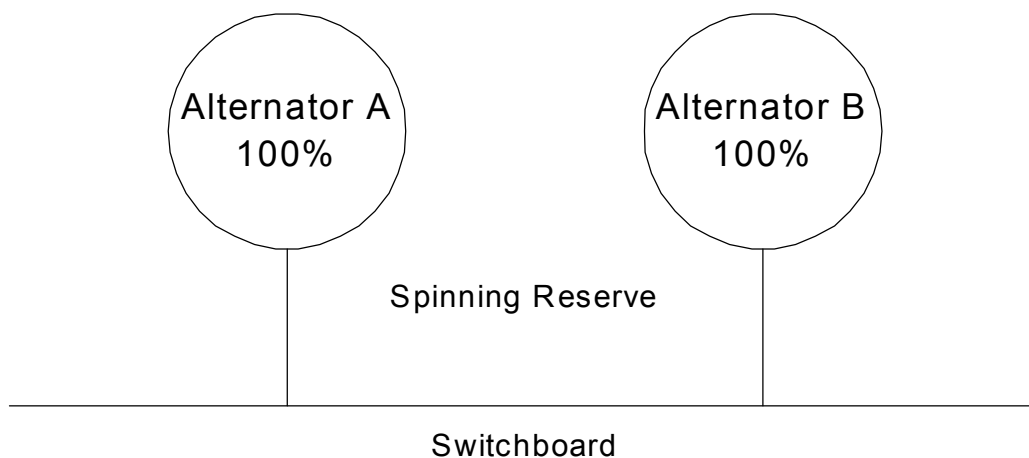
A sequence of possibly dependent random variables (x_1, x_2, x_3, \dots) - identified by increasing values of a parameter, commonly time - with the property that any prediction of the value of x_n , knowing x_1, x_2, \dots, x_{n-1} , may be based on x_{n-1} alone. That is, the future value of the variable depends only upon the present value and not on the sequence of past values.

(Encyclopaedia Britannica)

Nuclear decay, for example, is a Markov process. It is a random process, and the probability that a nucleus decays over a period δt is determined only by its state at time t (un-decayed or decayed).

1.1 Duplicate System with Single Repair Crew

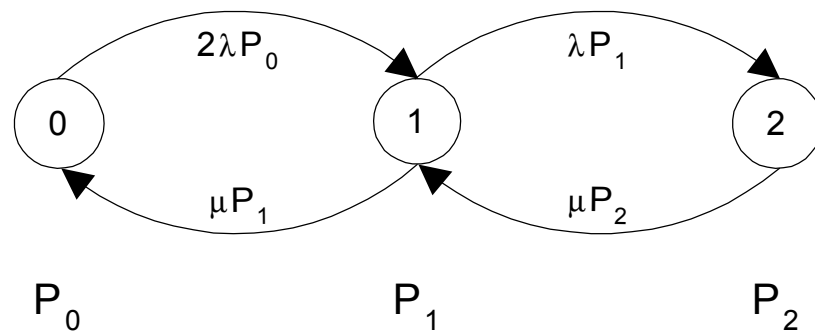
Markov analysis is widely used as a method of analysing the reliability parameters of repairable systems with constant failure and repair rates. Consider an example from power generation:



The system may be in any one of 3 states:

- State (0) Both operating, with probability P_0
- State (1) One unit failed, with probability P_1
- State (2) Both units failed, i.e. the system is unavailable, with probability P_2

Markov analysis is applied to the state transition diagram for this system to derive the probabilities of each of the states, as follows:



where:

λ = failure rate = 1 / MTBF (mean time between failures) for constant failure rate.

and μ = repair rate = 1 / MDT (mean down time)

The rate of leaving state 0 is $2\lambda P_0$, i.e. it is determined by the probability of being in state 0 and the failure rate of one generator. Similarly, the rate of arriving in state 0 is μP_1 , the rate at which items in state 1 are repaired. This leads to a set of equations:

$$dP_0/dt = -2\lambda P_0 + \mu P_1$$

$$dP_1/dt = 2\lambda P_0 - \mu P_1 - \lambda P_1 + \mu P_2$$

$$dP_2/dt = \lambda P_1 - \mu P_2$$

In the steady state the rates of change of probabilities are zero, and these equations reduce to:

$$-2\lambda P_0 + \mu P_1 = 0$$

$$+2\lambda P_0 - (\lambda + \mu)P_1 + \mu P_2 = 0$$

$$+\lambda P_1 - \mu P_2 = 0$$

to which we can add:

$$P_0 + P_1 + P_2 = 1$$

because the probability of being in one or other of these 3 states is unity.

Solving for P_2 (the unavailability of the system) we obtain:

$$P_2 = 2\lambda^2 / (2\lambda^2 + 2\lambda\mu + \mu^2)$$

which for $\mu \gg \lambda$ (as would normally be the case) reduces to:

$$P_2 \approx 2\lambda^2 / \mu^2 = 2\lambda^2 \cdot \text{MDT}^2 \text{ for } \mu \gg \lambda$$

But, **this is the wrong result.** We know that the unavailability of single generator is given by:

$$Q_a = Q_b = \text{MDT} / (\text{MTBF} + \text{MDT}) = \lambda / (\lambda + \mu)$$

which for $\text{MTBF} \gg \text{MDT}$ or (the corollary) $\mu \gg \lambda$ reduces to

$$Q_a = Q_b = \text{MDT} / \text{MTBF} = \lambda / \mu$$

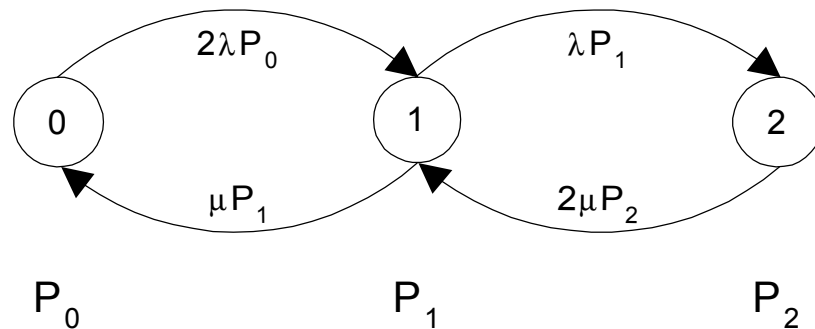
By the multiplication of probabilities rule the unavailability of the system is given by:

$$Q_{\text{system}} = P_2 = Q_a \cdot Q_b = \lambda^2 / \mu^2 = \lambda^2 \cdot \text{MDT}^2$$

i.e. exactly half the value obtained from Markov analysis.

1.2 Duplicate System with 2 Repair Crews

Consider now the application of Markov analysis to the state transition diagram of the same example with 2 repair crews:



Solving the equations in the same way as above yields:

$$P_2 = 2\lambda^2 / (2\lambda^2 + 4\lambda\mu + 2\mu^2)$$

which for $\mu \gg \lambda$ (as would normally be the case) reduces to:

$$P_2 \approx \lambda^2 / \mu^2 = \lambda^2 \cdot \text{MDT}^2$$

i.e. Markov analysis indicates that the unavailability is halved by introducing a second repair crew. **A little thought, however, indicates that this result is nonsense.** If both generators are broken down it is only necessary to restore one generator to restore the system. If it takes $\text{MDT}_{\text{system}}$ on average to restore the system with one repair crew, it will also take $\text{MDT}_{\text{system}}$ on average to restore the system, even if there are 2 repair crews, because the first crew is already repairing the first failure when the second occurs; the system downtime is not reduced by having 2 crews. If $\mu \gg \lambda$, the probability of the first failed and repaired generator failing again before the second machine is repaired is tiny, even if the repairs are carried out sequentially by a single repair crew, so that the reduction

in unavailability is negligible. Thus the benefit of having 2 repair crews is marginal to non-existent!

1.3 Source of the Fallacy in Applying a Markov Model

So what is the reason for the discrepancies between the results which are obtained from reasoning based on combination of probabilities and consideration of practical situations and those obtained by Markov analysis? **The answer in a nutshell is that, while random failure is a Markov process, repair is not a Markov process.** Whilst failures are events which occur at random points in time, repairs occur at anything but random points in time. Firstly, they occur deterministically immediately after failure, or at fixed intervals, depending upon the maintenance policy; secondly, and unlike failures, they are not events which occur at instants in time but processes which occur over periods of time. The consequence is that Markov analysis is not applicable to the process of failure and repair, and results obtained by applying it are fundamentally flawed and cannot be expected to be correct.

Part of the reason for the apparent plausibility of Markov analysis of the problem of failure and repair is that failure rate and repair rate are deceptively similar quantities. They are represented by Greek letters (λ for failure rate, μ for repair rate) which are consecutive in the Greek alphabet; both are terms which include the word “rate”, both are the reciprocal of “mean time”, and both have the dimensions of t^{-1} . These similarities disguise the reality that they are fundamentally different quantities:

- Failure rate (λ) is the rate at which random events occur on average. Calculations of reliability and availability based on λ are only valid for random failures with constant λ , for which failure times are exponentially distributed.
- Repair rate (μ) is the reciprocal of the average period of time it takes to repair the relevant item of equipment, i.e. $1 / \text{MDT}$. The period will probably be subject to some randomness, but even if it is constant, so that the down time is always the same, the results for availability remain valid, i.e. there is no need to assume any particular distribution of repair time. (Indeed, it may be wise to abandon the misleading term “repair rate” and the misleading symbol “ μ ”, because items are not repaired at a constant rate μ , they are repaired in brief periods of time which have an average duration = MDT.)

2 Alternative Models

2.1 Active Redundancy

2.1.1 Combination of Probabilities

We can derive the unavailability of an active redundant system, where n out of m units are required to operate and all m units are normally on-line, and where $\text{MTBF}_{\text{unit}} \gg \text{MDT}_{\text{unit}}$ or (the corollary) $\lambda \cdot \text{MDT}_{\text{unit}} \ll 1$ (which would normally be the case) by applying the multiplication of probabilities rule, as:

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

$$= [m! / \{(m-n+1)! \cdot (n-1)!\}] \cdot (\lambda \cdot \text{MDT}_{\text{unit}})^{m-n+1}$$

where ${}^m C_{m-n+1}$ is the number of combinations of $m-n+1$ out of m , since $m-n+1$ must fail for the system to be unavailable, e.g. for a 2 out of 4 system, 3 units must fail before the system fails, $m = 4$, $n = 2$ and $m-n+1 = 3$. This produces the following results:

Table 1 Unavailability by Combination of Probabilities for Active Redundancy (Independent of Number of Repair Crews)

Total number of units (m)	1	$\lambda \cdot \text{MDT}_{\text{unit}}$			
	2	$(\lambda \cdot \text{MDT}_{\text{unit}})^2$	$2\lambda \cdot \text{MDT}_{\text{unit}}$		
	3	$(\lambda \cdot \text{MDT}_{\text{unit}})^3$	$3(\lambda \cdot \text{MDT}_{\text{unit}})^2$	$3\lambda \cdot \text{MDT}_{\text{unit}}$	
	4	$(\lambda \cdot \text{MDT}_{\text{unit}})^4$	$4(\lambda \cdot \text{MDT}_{\text{unit}})^3$	$6(\lambda \cdot \text{MDT}_{\text{unit}})^2$	$4\lambda \cdot \text{MDT}_{\text{unit}}$
		1	2	3	4

Number of units required to operate (n)

This table can be compared with the results obtained by Markov analysis, which are:

Table 2 Unavailability by Markov Analysis for Active Redundancy with a Single Repair Crew

Total number of units (m)	1	$\lambda \cdot \text{MDT}_{\text{unit}}$			
	2	$2(\lambda \cdot \text{MDT}_{\text{unit}})^2$	$2\lambda \cdot \text{MDT}_{\text{unit}}$		
	3	$6(\lambda \cdot \text{MDT}_{\text{unit}})^3$	$6(\lambda \cdot \text{MDT}_{\text{unit}})^2$	$3\lambda \cdot \text{MDT}_{\text{unit}}$	
	4	$24(\lambda \cdot \text{MDT}_{\text{unit}})^4$	$24(\lambda \cdot \text{MDT}_{\text{unit}})^3$	$12(\lambda \cdot \text{MDT}_{\text{unit}})^2$	$4\lambda \cdot \text{MDT}_{\text{unit}}$
		1	2	3	4

Number of units required to operate (n)

For $\lambda \cdot \text{MDT}_{\text{unit}} \ll 1$, the results for any system with spares in excess of one, are of course essentially academic, and even for systems with a single spare ($n = m-1$), unavailability due to common cause failures of the order of $\lambda \cdot \text{MDT}_{\text{unit}}$ are likely to dominate unavailability due to independent failures of the order of $(\lambda \cdot \text{MDT}_{\text{unit}})^2$.

2.1.2 Consideration of Practical Failure and Repair Process

Another way of deriving the failure rates and unavailabilities for single spare ($m-1$ out of m) active redundancy systems is to consider the basic failure rate, i.e. the mean rate at which a single unit fails = $m \cdot \lambda$, since all m units are on-line. When a single unit fails, it is repaired with mean down time = MDT_{unit} . The probability of a second unit failing during that period is $(m-1) \cdot \lambda \cdot \text{MDT}_{\text{unit}}$, since $(m-1)$ units remain on-line, so that the system failure rate is $m \cdot (m-1) \cdot \lambda^2 \cdot \text{MDT}_{\text{unit}}$.

Failure rate	$2\lambda^2 \cdot \text{MDT}_{\text{unit}}$	$6\lambda^2 \cdot \text{MDT}_{\text{unit}}$	$12\lambda^2 \cdot \text{MDT}_{\text{unit}}$
	1 out of 2	2 out of 3	3 out of 4

Number of units required to operate (m-1 out of m, Active Redundancy)

These are the same as the results obtained by Markov analysis, but this does not mean that Markov results can be generalised to the n out of m case, and it does not give the correct result for unavailability.

On average, if a second unit fails, it will fail half way through the down period of the first unit, i.e. at $MDT_{unit} / 2$, so that:

$$MDT_{system} = MDT_{unit} / 2$$

and the system unavailability is given by:

$$\begin{aligned} Q_{system} &= MDT_{system} / (MTBF_{system} + MDT_{system}) \\ &= MDT_{system} / MTBF_{system} \text{ if } MTBF_{system} \gg MDT_{system} \\ &= (MDT_{unit} / 2) \cdot \{m \cdot (m-1) \cdot \lambda^2 \cdot MDT_{unit}\} \\ &= m \cdot (m-1) \cdot (\lambda \cdot MDT_{unit})^2 / 2 \end{aligned}$$

which is the same result as obtained by applying the multiplication of probabilities rule.

2.2 Standby Redundancy

2.2.1 Consideration of Practical Failure and Repair Process

The same argument can be applied to single spare standby redundancy systems. In this case the basic failure rate is $(m-1) \cdot \lambda$, since only $(m-1)$ units are normally on-line, and the probability of a second unit failing during the period when one unit has already failed is $(m-1) \cdot \lambda \cdot MDT_{unit}$, since $(m-1)$ units remain on-line, so that the system failure rate is $(m-1)^2 \cdot \lambda^2 \cdot MDT_{unit}$.

Failure rate	$\lambda^2 \cdot MDT_{unit}$	$4\lambda^2 \cdot MDT_{unit}$	$9\lambda^2 \cdot MDT_{unit}$
	1 out of 2	2 out of 3	3 out of 4

**Number of units required to operate
(m-1 out of m, Standby Redundancy)**

Applying the same argument as above about the average time when the second unit will fail gives the following results for unavailability:

Table 3 Unavailability by Consideration of Practical Failure and Repair Process for Standby Redundancy (Independent of Number of Repair Crews)

Total number of units (m)	1	$\lambda \cdot \text{MDT}_{\text{unit}}$			
	2	$(\lambda \cdot \text{MDT}_{\text{unit}})^2 / 2$	$2\lambda \cdot \text{MDT}_{\text{unit}}$		
	3	?	$2(\lambda \cdot \text{MDT}_{\text{unit}})^2$	$3\lambda \cdot \text{MDT}_{\text{unit}}$	
	4	?	?	$9(\lambda \cdot \text{MDT}_{\text{unit}})^2 / 2$	$4\lambda \cdot \text{MDT}_{\text{unit}}$
		1		2	3

Number of units required to operate (n)

This table also can be compared with the results obtained by Markov analysis (with a single repair crew), which are:

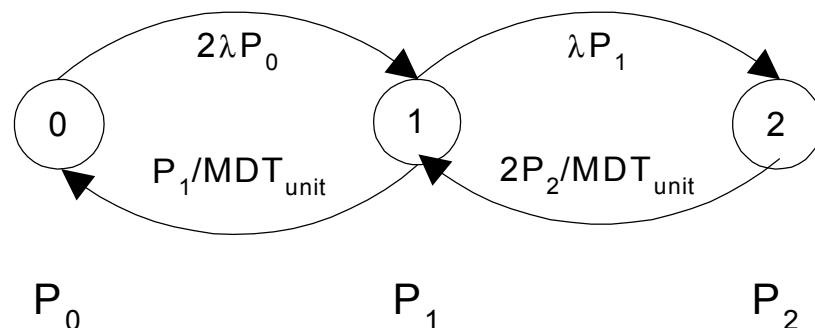
Table 4 Unavailability by Markov Analysis for Standby Redundancy with a Single Repair Crew

Total number of units (m)	1	$\lambda \cdot \text{MDT}_{\text{unit}}$			
	2	$(\lambda \cdot \text{MDT}_{\text{unit}})^2$	$2\lambda \cdot \text{MDT}_{\text{unit}}$		
	3	$(\lambda \cdot \text{MDT}_{\text{unit}})^3$	$4(\lambda \cdot \text{MDT}_{\text{unit}})^2$	$3\lambda \cdot \text{MDT}_{\text{unit}}$	
	4	$(\lambda \cdot \text{MDT}_{\text{unit}})^4$	$8(\lambda \cdot \text{MDT}_{\text{unit}})^3$	$9(\lambda \cdot \text{MDT}_{\text{unit}})^2$	$4\lambda \cdot \text{MDT}_{\text{unit}}$
		1	2	3	4

Number of units required to operate (n)

2.3 Application of Transition Diagrams

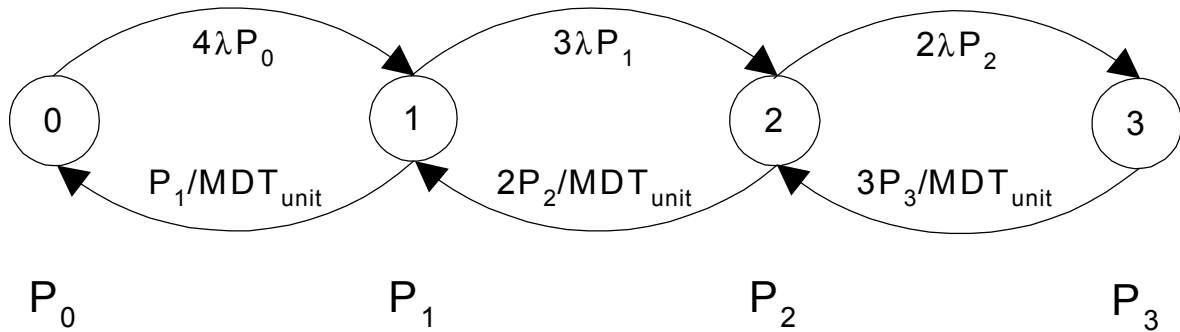
The same results can be obtained by the state transition diagram method. For example, the diagram for the 1 out of 2 active redundancy case is:



which yields:

$$P_2 = Q_{\text{system}} = (\lambda \cdot \text{MDT}_{\text{unit}})^2 \text{ for } \lambda \cdot \text{MDT}_{\text{unit}} \ll 1$$

and for the 2 out of 4 active redundancy case is:



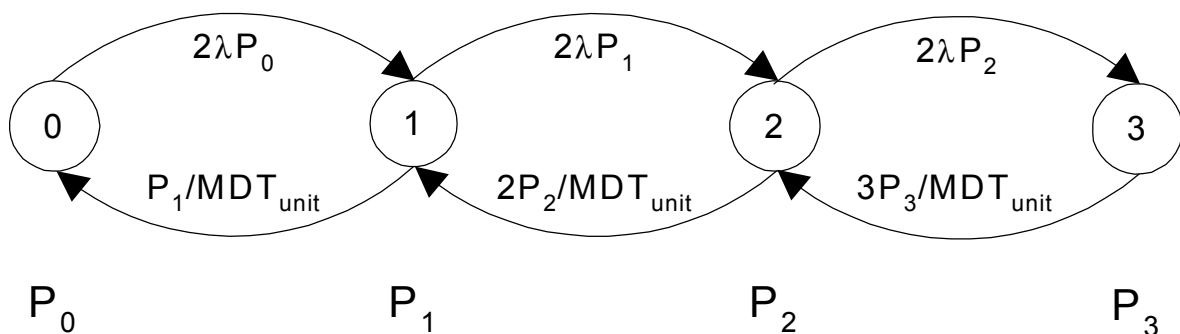
which yields:

$$P_3 = Q_{system} = 4(\lambda \cdot MDT_{unit})^3 \text{ for } \lambda \cdot MDT_{unit} \ll 1$$

In general, it can be seen that the same results are obtained from state transition diagrams as from combination of probabilities, and from consideration of the practical failure and repair process, if the restoration rate from State (k) (i.e. k units failed) is set to k/MDT_{unit} , or rather that the restoration period is set to MDT_{unit}/k . Physically, the explanation follows a similar argument to that developed above for the 1 out of 2 case. On average, if k units fail, the second unit fails after the first unit has been failed for MDT_{unit}/k , the third unit after $2MDT_{unit}/k$, the k^{th} unit after $(k-1) \cdot MDT_{unit}/k$, so that there is only a period MDT_{unit}/k remaining before the first failed unit is restored, and it is only for this period that the system remains in State (k) before it is returned to State (k-1). Thus for the 2 out of 4 example above the third failure occurs on average $\frac{2}{3}^{rd}$ of the way through the down time of the first unit to fail. (Note that, for $\lambda \cdot MDT_{unit} \ll 1$, $P_0 \approx 1$, $P_0 \gg P_1 \gg P_2$ and $P_1 \approx 4\lambda \cdot MDT_{unit}$, etc.)

It should be emphasised that reducing the restoration period in this way has nothing whatever to do with deploying multiple repair crews, and certainly nothing to do with Markov analysis.

Developing the state transition diagram in a similar way for 2 out of 4 standby redundancy, we obtain:



which yields the result:

$$P_3 = Q_{system} = \frac{4(\lambda \cdot MDT_{unit})^3}{3} \text{ for } \lambda \cdot MDT_{unit} \ll 1$$

and we can now complete the table of unavailabilities for standby redundancy as follows:

Table 5 Unavailability by Transition Diagrams Based on Combination of Probabilities for Standby Redundancy (Independent of Number of Repair Crews)

Total number of units (m)	1	$\lambda \cdot \text{MDT}_{\text{unit}}$			
	2	$(\lambda \cdot \text{MDT}_{\text{unit}})^2 / 2$	$2\lambda \cdot \text{MDT}_{\text{unit}}$		
	3	$(\lambda \cdot \text{MDT}_{\text{unit}})^3 / 6$	$2(\lambda \cdot \text{MDT}_{\text{unit}})^2$	$3\lambda \cdot \text{MDT}_{\text{unit}}$	
	4	$(\lambda \cdot \text{MDT}_{\text{unit}})^4 / 24$	$4(\lambda \cdot \text{MDT}_{\text{unit}})^3 / 3$	$9(\lambda \cdot \text{MDT}_{\text{unit}})^2 / 2$	$4\lambda \cdot \text{MDT}_{\text{unit}}$
		1	2	3	4

Number of units required to operate (n)

3 Conclusions

1. The correct unavailabilities of active redundant systems (for independent failures) are given by Table 1, i.e.:

Table 1 Unavailability by Combination of Probabilities for Active Redundancy (Independent of Number of Repair Crews)

Total number of units (m)	1	$\lambda \cdot \text{MDT}_{\text{unit}}$			
	2	$(\lambda \cdot \text{MDT}_{\text{unit}})^2$	$2\lambda \cdot \text{MDT}_{\text{unit}}$		
	3	$(\lambda \cdot \text{MDT}_{\text{unit}})^3$	$3(\lambda \cdot \text{MDT}_{\text{unit}})^2$	$3\lambda \cdot \text{MDT}_{\text{unit}}$	
	4	$(\lambda \cdot \text{MDT}_{\text{unit}})^4$	$4(\lambda \cdot \text{MDT}_{\text{unit}})^3$	$6(\lambda \cdot \text{MDT}_{\text{unit}})^2$	$4\lambda \cdot \text{MDT}_{\text{unit}}$
		1	2	3	4

Number of units required to operate (n)

2. The correct unavailabilities of standby redundant systems (for independent failures) are given by Table 5, i.e.:

Table 5 Unavailability by Transition Diagrams Based on Combination of Probabilities for Standby Redundancy (Independent of Number of Repair Crews)

Total number of units (m)	1	$\lambda \cdot \text{MDT}_{\text{unit}}$			
	2	$(\lambda \cdot \text{MDT}_{\text{unit}})^2 / 2$	$2\lambda \cdot \text{MDT}_{\text{unit}}$		
	3	$(\lambda \cdot \text{MDT}_{\text{unit}})^3 / 6$	$2(\lambda \cdot \text{MDT}_{\text{unit}})^2$	$3\lambda \cdot \text{MDT}_{\text{unit}}$	
	4	$(\lambda \cdot \text{MDT}_{\text{unit}})^4 / 24$	$4(\lambda \cdot \text{MDT}_{\text{unit}})^3 / 3$	$9(\lambda \cdot \text{MDT}_{\text{unit}})^2 / 2$	$4\lambda \cdot \text{MDT}_{\text{unit}}$
		1	2	3	4

Number of units required to operate (n)

3. These are NOT the same results as are obtained by Markov analysis.

See also [Further Thoughts on Markov Analysis in Reliability Modelling - An Addendum to "Repairable Redundant Systems and the Markov Fallacy"](#)