

Availability

Table of Contents

- Introduction
- Statistical Models for Simple Systems (Up/Down) and Interpretation
- Markov Models for Simple Systems (Up/Down) and Interpretation
- A Markov Model for a Simple Redundant System
- Model Extensions and Comparisons
- For Further Study
- About the Author
- Other START Sheets Available

Introduction

There is a fundamental difference between the approaches used to perform statistical and reliability analyses of non-maintained (some times called “one-shot”) and those used for maintained systems. Non-maintained systems either fulfill their missions (by surviving beyond mission time) or fail it (by perishing before the mission is completed). In contrast, maintained systems can be repaired (maintained) and put back into operation. During “maintenance,” however, the system is “down” and unavailable for its intended use. This situation changes the analysis approach because maintenance introduces the related and new concept of “availability” (that the system will be “up” and “available” for use, when needed, in lieu of undergoing maintenance). The main objective of this START sheet is to help engineers better understand the meaning and implications of the statistical methods used to develop performance measures (PM) for assessing system “availability” (A).

We start by reviewing some relevant definitions. RAC’s Reliability Toolkit defines availability as “a measure of the degree to which an item is in an operable state at any time.” It defines “maintainability” as “a measure of the ability of an item to be retained in, or restored to, a specified condition, when maintenance is performed using prescribed procedures and technician skill levels.”

From these definitions, we can deduce that system availability is a probabilistic concept. It is based on the system life (X), a random variable (RV). Since the system can fail at any random time, availability is also based on a second RV: the

maintenance time (Y). We calculate their “long run averages” (i.e., results obtained as time goes to infinity, $t \rightarrow \infty$) or “expected values” and denote them respectively $E(X)$ and $E(Y)$. $E(X)$ is the expected system life or the up time, often measured as “Mean Time Between Failures” (MTBF). $E(Y)$ is the expected maintenance time (which includes the activities of fault isolation, repair or removal and function check), often measured as “Mean Time To Repair” (MTTR).

The random nature of system life (X) and maintenance times (Y) demands the use of statistics for obtaining system PM. Therefore, we need to redefine Availability in statistical terms. Hoyland et al (Reference 1), for example, define “availability at time t”, $A(t)$, as “the probability that the system is functioning at time t”. If we call $X(t)$ the “state” of a system at time “t” (which can be either “up” and running [$X(t) = 1$], or “down” and failed [$X(t) = 0$]), then this definition of $A(t)$ can be written as:

$$A(t) = P\{X(T) = 1\}; t > 0$$

The availability concept becomes even more complex when we realize that it is divided into several classes. For example, Blanchard (Reference 2) states that “availability may be expressed differently, depending on the system and its mission” and defines three types:

1. Inherent availability (A_i) is the probability that a system, when used under stated conditions ... will operate satisfactorily at any point in time, as required. A_i excludes preventive maintenance, logistics, and administrative delays, etc.

$$A_i = \frac{MTBF}{MTBF + MTTR}$$

2. Achieved availability (A_a) is the probability that a system when used under stated conditions ... will operate satisfactorily at any point in time. A_a includes preventive maintenance but excludes logistics and administrative delays, etc.

3. Operational availability (A_o) is the probability that a system when used under stated conditions ... will operate satisfactorily when called upon. A_o includes all the fac-

tors that contribute to system downtime for all reasons (i.e., maintenance actions and delays, access, diagnostics, active repair, supply delays, etc.). We call it Mean Down Time (MDT).

$$A_0 = \frac{MTBM}{MTBM + MDT}$$

Notice that, as opposed to $A(t)$, these formulae do not include any reference to a random time “ t ”. The reason is that they are based on the “long run averages” of the system life (X) and maintenance times (Y) and, hence, about the “long run average” (or expected) Availability. To better understand this statistical concept, consider the whole cycle of “up-time plus maintenance” (i.e., $X + Y$). The cycle repeats itself again and again, throughout the entire system life, and constitutes the “totality” of possibilities for the RV system state $X(t)$ within a single cycle. Now consider “system up-time” ($X(t) = 1$) as our “event” of interest. Since the probability of any “event” is defined as the ratio of its “favorable” to its “total” possibilities, we can state that the “long run availability” A is:

$$A = P(\text{System Up}) = \frac{\text{Favorable Cases}}{\text{Total Cases}} = \frac{\text{Up Time}}{\text{Cycle Time}}$$

$$\cong P(X(t) = 1); \text{ as } t \rightarrow \infty$$

Since we are not interested in the “up-time” during a specific cycle but in the long run ($t \rightarrow \infty$), we substitute Up (X) or Down (Y) times by their respective long run averages, $E(X)$ and $E(Y)$. For example, if a system has MTBF = 500 hours and MTTR = 30 hours we obtain:

$$A = A(\infty) = P\{X(\infty) = 1\} = \frac{E(X)}{E(X) + E(Y)}$$

$$= \frac{MTBF}{MTBF + MTTR} = \frac{500}{500 + 30} = 0.9433$$

Finally, we must select which Availability definition we want to discuss. The three classes of “availability” differ only in their respective definition scope. For example, in A_i , Inherent Availability, the average up-time (MTBF) includes only design and manufacturing failures, and the average maintenance (MTTR) only includes active repair time. In Operational availability, average up-time (MTBM) includes all downing events, whatever the cause (e.g., design or manufacturing failure, induced failure, preventive maintenance event, etc.), and average maintenance (MDT) includes all possible downtime (Reference 3).

Without diminishing the practical importance and impact of the differences, the conceptual treatment of availability, from a strictly statistical point of view, is similar in all these cases. Their parameters, “expected values” (and perhaps even the dis-

tributions of the variables involved) may change, and with them, also the interpretation and form of their results. However, such differences will not modify the basic statistical philosophy used for obtaining them, nor the concepts on which their approaches rest. Since the main objective of this START sheet is helping the engineer better understand such statistical philosophy and approaches, we will consider all three cases as a single one and adopt the nomenclature given previously for A_i when referring and dealing with availability.

In the remainder of this START sheet, we first overview and give numerical examples of the statistical treatment of the availability of a simple, repairable system in discrete times. Then, we will consider the case of a simple, repairable, parallel redundant system, comparing two different statistical modeling and analysis approaches. We will illustrate these approaches by developing more numerical examples, performing systems analyses, and comparing their results. Finally, we will mention several ways to improve system availability and provide additional bibliography for further study of these topics.

Statistical Models for Simple Systems (Up/Down) and Interpretation

In the Introduction, the “long run average” Availability was obtained as the ratio of the “long run averages” of Up-Time to Cycle Time. However, Availability is a (cycle) RV itself. Hence, like any other RV, it has its own distribution, density function, etc. In this section, we overview a statistical model that describes Availability (A) as the RV resulting of the algebraic combination of the RV “time between failures” (X) and the RV “time to repair” (Y), at every cycle. In the following section, we describe Availability as a two-state, discrete time Markov Chain. The objective for presenting two contrasting models is to enhance the understanding of different statistical approaches, so engineers can better use them and get meaningful results from their implementation. We will briefly illustrate their mathematical derivations using a practical example. In For Further Study, we reference documents that provide more in-depth information on the subject.

We start by defining random variable (cycle) Availability as the following ratio:

$$A_i = \frac{X_i}{X_i + Y_i}; X_i, Y_i > 0, i = 1, \dots, n$$

The problem of obtaining the “density function” or statistical description of A is resolved using a variable transformation of the joint distribution of the Availability (A) and of some other convenient function (denoted B), of time between failures X and time to repair Y , such as $B(X, Y) = X + Y$.

Assume that system times between failures and to repair (X and Y) are independent of each other and Exponentially distributed,

with mean $\mu = \text{MTBF} = \text{MTTR} = 1$ hour. Hence, their individual density functions are $f_1(X) = \text{Exp}(-X)$ and $f_2(Y) = \text{Exp}(-Y)$. Their joint density function, denoted $f(X, Y)$, is just the product of the two individual Exponential densities, since both (failure and repair) times X and Y are independent of each other. Therefore:

$$f(X, Y) = f_1(X) \times f_2(Y) = \text{Exp}\{-X\} \times \text{Exp}\{-Y\} = \text{Exp}\{-(X+Y)\}$$

Define the (cycle) Availability function $A(X, Y) = X/(X + Y)$. Define auxiliary function $B(X, Y) = X + Y$. Then, their inverse functions are $X = W(A, B) = AB$ and $Y = Z(A, B) = B(1 - A)$. Finally, obtain the matrix of the partial derivatives of the inverses $W(A, B)$ and $Z(A, B)$ and denote it $J(W, Z)$. To derive the joint distribution $g(A, B)$ of A and B , just substitute the values of X and Y in the original joint distribution function $f(X, Y)$, with their inverses [$X = W(A, B) = AB$ and $Y = Z(A, B) = B(1 - A)$] and multiply this by the absolute value of the matrix of the partial derivatives $|J(W, Z)| = |B|$. That is:

$$g(A, B) = f(AB; B(1 - A)) \times |J(W, Z)| \\ = \text{Exp}\{-(AB+B(1 - A))\} \times |B| = B \text{Exp}\{-B\}$$

This variable transformation yields the density of Availability. For, density function $g_1(A)$ of the resulting system Availability ($A = X/(X + Y)$) is just the “marginal distribution” of the above derived, bivariate density function $g(A, B)$.

Hence, the desired marginal (Availability density) is obtained by integrating function $g(A, B)$ out on B :

$$g_1(A) = g_1\left(\frac{X}{X+Y}\right) = \int g(A, B) dB = \int_0^{\infty} B \text{Exp}\{-B\} dB = 1;$$

$$\text{for } 0 < A < 1$$

Hence, $g_1(A) = 1$ is the theoretical density of Availability and corresponds to the density of the Uniform (0, 1). As a result, all performance measures (PM) of interest, such as the expected value, variance, percentiles, probabilities, etc., are now obtained from the theoretical Uniform (0, 1) distribution parameters. For the simple example given, where both times between failure (X) and to repair (Y) are distributed Exponentially with mean $\mu = \text{MTBF} = \text{MTTR} = 1$, we obtain:

1. Expected Availability = $E\{\text{Uniform}(0,1)\} = 1/2 = 0.5$.
2. Variance of Availability = $\text{Var}\{\text{Uniform}(0,1)\} = 1/12 = 0.083$.
3. L_{10} = Percentile of 10% of Availabilities = $P\{A < 0.1\} = 0.1$.
4. First and Third Quartiles of Availability = 0.25 and 0.75.

The Uniform distribution is a special case of the Beta, which is the general distribution of Availability when the failure and repair times are exponentially distributed. Such theoretical results allow us to obtain empirically the Availability distribution via Monte Carlo (MC) simulation. To verify this for the results just given, we generate $n = 5,000$ Exponentially distributed random failure and repair times, X_i and Y_i , $i = 1, \dots, n$, with $\mu = \text{MTBF} = \text{MTTR} = 1$. We then obtain the corresponding Availabilities $A_i = X_i/(X_i + Y_i)$, sort them, and calculate the $n = 5,000$ results numerically, via MC.

The Expected Value is obtained from the sample average (0.5067); the Variance, from the sample variance (0.0826). Percentile L_{10} (Availability achieved 90% of the times) and all other probabilities are obtained by manipulating the sorted ranks of the total number of data points “ n ”, of the MC generated values. For example, L_{10} corresponds to the 500th sorted rank (10% of the $n = 5,000$ MC data points) and yields a value 0.1048. The quartiles are 0.2558 and 0.7559. Empirical and theoretical results agree closely.

For Exponential means different than unit ($\mu = \text{MTBF} = \text{MTTR} \neq 1$) the mathematical treatment is more difficult and we use the Beta distribution, directly. For a more realistic example, we reuse the example of times between failure (X) with $\mu = \text{MTBF} = 500$ hours, and to repair (Y) with $\mu = \text{MTTR} = 30$ hours. We generate $n = 5,000$ random Beta values with parameters corresponding to the said failure and repair times and obtain the (cycle) MC Availabilities, A_i . Results are given in Figure 1 and Table 1.

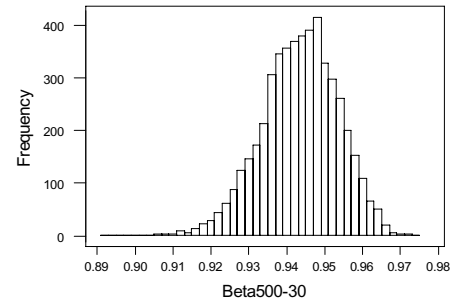


Figure 1. Histogram of Realistic Example

Table 1: Realistic Availability Results – MC Results for Beta (500,30) Example:

Average Availability	=	0.9435
Variance of Availability	=	9.92×10^{-5}
Life L_{10}	=	0.9305
Quartiles	=	0.9370 and 0.9505
$P\{A > 0.95\}$	=	0.2694

For example, the probability that the Availability is greater than value 0.95, $P\{A > 0.95\}$, is obtained by looking at the sorted rank corresponding to a MC Availability closer to 0.95: ($A =$

0.9499 → Rk = 3,653). Then, we divide this Rank by n = 5,000 and subtract it from unit:

$$P\{A > 0.95\} = 1 - P\{A \leq 0.95\} \approx 1 - \frac{3,653}{5,000} = 1 - 0.7306 = 0.2694$$

Markov Models for Simple Systems (Up/Down) and Interpretation

Now, consider the previous problem, approached as a two-state Markov Chain (References 4, 5, 6, and 7). Here we monitor the “status” of the system at time T, denoted X(T), instead of its “availability” A(T). Denote State 0 (Down) and State 1 (Up), and assess the status X(T) of your system S every hour (T = 0, 1, ...). Hence, X(T) = 0 means that system S was Down at time T and X(T) = 1, that system S was Up at time T. We are interested in studying how the System S develops (transitions) over time. That is, we want to know what is the probability q (or p) that system S is Up (or Down) at time T, given that it was Down (or Up) an hour earlier (at time T - 1). We represent this problem using the Markov Chain state diagram shown in Figure 2.

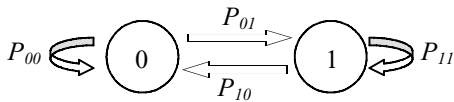


Figure 2. State Diagram for System S

The transition probabilities are:

$$p_{01} = P\{X(T) = 1 | X(T-1) = 0\} = q$$

$$p_{10} = P\{X(T) = 0 | X(T-1) = 1\} = p$$

$$p_{00} = 1 - q$$

$$p_{11} = 1 - p$$

For example, let system S be Down at time T - 1. Then, either the system is Up at time T, with probability q, or it will be Down with probability 1 - q. If the system was Up at T - 1 then it is either Down at T with probability p, or it is Up at T, with probability 1 - p. This occurs because there are only two possibilities (Up or Down) for S at any given time T. The two state probabilities have to add up to Unit. Let’s analyze this situation further.

Each time unit (hour) T, transitioned by system S, can be considered as an independent trial, and the probability p_{ij} of moving from i into the other state j, as the probability of “success”. For example, let system S be in state Up. Then, moving to state Down by one step with probability $p_{10} = p = 0.002$ yields a Geometric distribution with Mean $\mu = 1/p = 500$ hours. If, instead, system S is in state Down then, moving to state Up by one-step (hour), with probability $p_{01} = q = 0.033$, yields a Geometric, with Mean $\mu = 1/q = 30$ hours. These are the same parameters of the “realistic” example of the previous section.

The Geometric distribution is the discrete counterpart of the continuous Exponential and, as the units of time T become smaller (hours to minutes, seconds, etc.), the two distributions converge. Therefore, this numerical example is equivalent to the one given in the previous section, which used similar time parameters, and will serve as a vehicle for comparison and contrast.

One important property of Markov Chains is their “lack of Memory.” This means that only the system status at the immediately previous time has any bearing on the status at the current time, and every other past history goes into oblivion. In addition, these Markov Chains are time homogeneous (the transition probabilities p_{ij} do not change over time). Hence, it is enough to know the “one-step state transition probabilities” or the probabilities p_{ij} of going from any state “i” to any other state “j” in one step, to resolve the problems.

Markov Chains can be represented by a “Transition Probability Matrix” P, where rows represent every system state we can be in at time T, and columns represent every other state we can go to, in one step (i.e., where we will be, at T + 1). Entries of Matrix P (p_{ij}) correspond to the Markov Chain’s one-step transition probabilities and must add up to unit, on every matrix row. For our numerical example, the Transition Probability matrix P is:

States	0	1	States	0	1
0	(1 - q)	q	0	(0.967)	(0.033)
1	p	(1 - p)	1	(0.002)	(0.998)

If we need the probabilities of moving from one state to any other, in two steps, we raise matrix P to the second power. For example, moving from states Up to Down in two steps, entails either moving from Up to Down in first step, and remaining in Down state another step. Or it may entail first remaining Up for one step, before moving from states Up to Down in the second step. In matrix language, this is expressed in the following way:

$$P^2 = \begin{bmatrix} 1-q & q \\ p & 1-p \end{bmatrix}^2 = \begin{bmatrix} 1-q & q \\ p & 1-p \end{bmatrix} \times \begin{bmatrix} 1-q & q \\ p & 1-p \end{bmatrix} = \begin{bmatrix} (1-q)^2 + pq & q(1-q) + q(1-p) \\ p(1-q) + p(1-p) & pq + (1-p)^2 \end{bmatrix} = \begin{bmatrix} P_{00}^{(2)} & P_{01}^{(2)} \\ P_{10}^{(2)} & P_{11}^{(2)} \end{bmatrix}$$

In our example, the $p_{10}^{(2)}$ result provides the probability that system S is Down, if it was initially Up, after operating for two hours (T = 2): $p_{10}^{(2)} = p(1 - q) + (1 - p)p = 0.003$. The $p_{11}^{(2)}$ result (probability that S is Up after 2 hours, given that it started Up) can be obtained as one minus the probability that S is down, after 2 hours: $p_{11}^{(2)} = 1 - p_{10}^{(2)}$. We can then interpret $p_{11}^{(2)} = A(T) = 0.997$ as the system Availability, after T = 2 hours of operation, if it started in state Up (at T = 0).

To obtain the probability of moving from one state to another in “n” steps, we raise the matrix P to the “nth” power (P^n). For example, the probability that S is Down after T = 10 hours

(steps) if it was initially Up, is $p_{10}^{(10)} = 0.017$ (this includes that S could have gone Down or Up, then restored, and this may have occurred more than once during the $T = 10$).

For a sufficiently large “n” matrix P^n yields quasi identical rows. Results are interpreted as “long run averages” or limiting probabilities “ p_i ” of S being in the state corresponding to column “i”. These results are similar to the ones obtained using the Expected Availability and Unavailability and the variable transformations approaches. To obtain these limiting probabilities (i.e., to calculate P^n) we need a practical result. For any two-state (e.g., Up, Down) system, as the one described above, this practical method is as follows.

$$P^n = \begin{bmatrix} 1-q & q \\ p & 1-p \end{bmatrix}^n = \frac{1}{p+q} \begin{bmatrix} p & q \\ p & q \end{bmatrix} + \frac{(1-p-q)^n}{p+q} \begin{bmatrix} q & -q \\ -p & p \end{bmatrix}$$

Then, for a sufficiently large “n”, the second term goes to zero and the matrix P^n reduces to:

$$\begin{aligned} \text{Limit}_{n \rightarrow \infty} P^n &= \text{Limit}_{n \rightarrow \infty} \left\{ \frac{1}{p+q} \begin{bmatrix} p & q \\ p & q \end{bmatrix} + \frac{(1-p-q)^n}{p+q} \begin{bmatrix} q & -q \\ -p & p \end{bmatrix} \right\} \\ &= \begin{bmatrix} p/(p+q) & q/(p+q) \\ p/(p+q) & q/(p+q) \end{bmatrix} \end{aligned}$$

Verify, for our given numerical example, that the probability of being in state Up at any arbitrary time T is $q/(p+q) = 0.943$, and the probability of being in state Down, is $p/(p+q) = 0.057$. These two “state occupancy rates”, $E(X)$ and $E(Y)$, can also be interpreted as the percent of the time that the system S will spend in states Up and Down.

A Markov Model for a Simple Redundant System

In Reference 8, we developed a statistical model for a non-maintained, simple redundant system, composed of two identical devices in parallel. The approach was based on the two RV, corresponding to the two device lives. In this section we also analyze a simple redundant system composed of two identical devices in parallel. The differences now are that we use a Markov Chain approach, and that system S is maintained and can function at a degraded level with only one unit. The advantages of Markov modeling of system Availability, as will become apparent from the numerical example that follows, increase as the system becomes more complex (as also do the mathematics behind the analyses involved).

Let, as before, $X(T)$ be the state of the system at time $T (= 0,1,2, \dots \text{ hours})$. Let State 0 be the Down state, where both devices have failed and one of them is being repaired. Let State 1 be the Degraded state, where one device has failed and is being repaired and the second is working (and the system is operable but with lesser capabilities). Finally, let State 2 be the Up state,

where both units are operating and the system is working at full capacity. The state diagram for this model is shown in Figure 3.

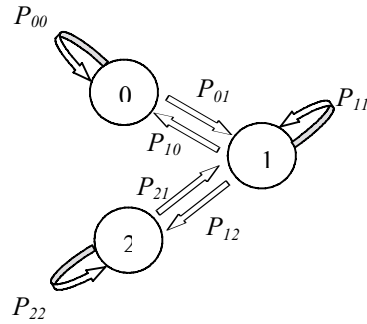


Figure 3. Markov Chain for Redundant System

The state equations are:

$$\begin{aligned} p_{01} &= P\{X(T)=1|X(T-1)=0\} = q \\ p_{10} &= P\{X(T)=0|X(T-1)=1\} = p \\ p_{12} &= P\{X(T)=2|X(T-1)=1\} = q \\ p_{21} &= P\{X(T)=1|X(T-1)=2\} = 2p \\ p_{ii} &= P\{X(T)=i|X(T-1)=i\} = 1 - \sum_{j \neq i} p_{ij} \end{aligned}$$

As before, we can consider every step (hour) T as an independent trial, having probability of success p_{ij} corresponding to the feasible transitions from our current state “i” into state $j = 0,1,2$. Hence, we can again think of the distribution of every change of state (produced by the occurrence of a failure or a repair) as being geometric, the discrete counterpart of the Exponential. It will have “probability of success” $p = p_{ij}$ (corresponding to the change into that state) and a mean time to accomplishing such change of $\mu = 1/p_{ij}$.

The transition probability matrix P for this model is given by:

	States	0	1	2	States	0	1	2
	0	p_{00}	p_{01}	p_{02}	0	$1-q$	q	0
$P =$	1	p_{10}	p_{11}	p_{12}	1	p	$1-p-q$	q
	2	p_{20}	p_{21}	p_{22}	2	0	$2p$	$1-2p$

Rows must add to one (probability is unity because the system is always in one of its three states). And, if we want to know the probability $p_{ij}^{(n)}$ of being in some state “j” after “n” steps, given that we started in some state “i” of the system, we raise matrix P to the power “n” as we did before, and look at entry p_{ij} of the resulting matrix P^n . With the advent of modern computers and math software, these operations are no longer tedious or difficult.

Modify the numerical example of previous section, now using two units instead of one. The probability p of either unit failing

in the next hour is 0.002. The probability q of the repair crew completing a maintenance job in the next hour is 0.033. Only one failure is allowed in each unit time period, and only one repair can be undertaken at a time.

With these new conditions, the probability that a degraded system (State 1) remains degraded after two hours is the sum of the probabilities corresponding to three events. First, that system status has never changed. Second, that one unit is first repaired and then another unit fails during the second hour. Third, that remaining unit fails in the first hour (the entire system goes down) then, a repair is completed in the second hour (system goes up, at degraded level):

$$\begin{aligned} P_{11}^2 &= [P \times P]_{11} = p_{11}^{(2)} = p_{10}p_{01} + p_{11}p_{11} + p_{12}p_{21} \\ &= pq + (1-p-q)^2 + 2pq \\ &= 0.002 \times 0.033 + (1-0.035)^2 + 2 \times 0.002 \times 0.033 \\ &= 0.9314 \end{aligned}$$

We are also interested in the mean time that the system spends in any given state. For example, System S can change to Up or Down, from state Degraded, in one step, with probabilities p and q . Hence, S will remain in the state Degraded with probability $1-p-q$. Then, on average, S will spend a “sejour” of length $1/\{1-(1-p-q)\} = 1/0.035 = 28.57$ consecutive hours in the Degraded state, before moving out to either Up or Down states.

Let’s now analyze “Availability at time T” = $A(T) = P\{S \text{ is Available at } T\}$. But this just means that system S is not Down at time “T” (it can be Up or Degraded). In addition, S could have initially been Up, Down or Degraded. Hence, $A(T)$ depends on the initial state of S (States 0,1,2), actual system availability level (States 1,2) and time (T). Assume we are interested in S being “Degraded Available” at T, given it was Degraded at T = 0: $p_{11}^{(T)}$. Since for matrix P^T every row has to add to unit, we can obtain such Availability via:

$$\begin{aligned} p_{10}^{(T)} + p_{11}^{(T)} + p_{12}^{(T)} &= 1 \Rightarrow p_{11}^{(T)} = 1 - p_{10}^{(T)} - p_{12}^{(T)} \\ A(T) = P\{X(T)=1 \mid X(0)=1\} &= p_{11}^{(T)} = 1 - p_{10}^{(T)} - p_{12}^{(T)} \end{aligned}$$

We may instead be interested in “long run averages” or “state occupancies”. These are the asymptotic probabilities of system S being in each one of its possible states at any time T, or the percent time spent in these states, irrespective of the state they were in, initially. These results are obtained by considering the Vector (denoted Π) of “long run” probabilities:

$$\begin{aligned} \Pi &= \lim_{T \rightarrow \infty} [\text{Prob}\{X(T)=0\}, \text{Prob}\{X(T)=1\}, \text{Prob}\{X(T)=2\}] \\ &= (\Pi_0; \Pi_1; \Pi_2) \end{aligned}$$

Vector Π fulfills two important properties that allow the calculation of such values:

$$(1) : \Pi \times P = \Pi; (2) \sum \Pi_i = 1; \text{ with } : \Pi_i = \lim_{T \rightarrow \infty} \text{Prob}\{X(T)=i\}$$

In plain English, $\Pi \times P = \Pi$ (Vector Π times the matrix P equals Π) defines a system of linear equations, that are “normalized” by the second property (that probabilities in the components of Vector Π add to Unit). For our example, we have the following.

$$\begin{aligned} \Pi \times P &= (\Pi_0, \Pi_1, \Pi_2) \times \begin{bmatrix} 0.967 & 0.033 & 0 \\ 0.002 & 0.965 & 0.033 \\ 0 & 0.004 & 0.996 \end{bmatrix} = (\Pi_0, \Pi_1, \Pi_2) \\ \Rightarrow &\begin{cases} 0.967\Pi_0 + 0.002\Pi_1 = \Pi_0 \\ 0.033\Pi_0 + 0.965\Pi_1 + 0.004\Pi_2 = \Pi_1 \\ 0.033\Pi_1 + 0.996\Pi_2 = \Pi_2 \end{cases} \\ &\text{with } : \sum_i \Pi_i = \Pi_0 + \Pi_1 + \Pi_2 = 1 \end{aligned}$$

The solution of this linear system of equations yields the long run or asymptotic occupancy rates:

$$\Pi = (\Pi_0, \Pi_1, \Pi_2) = (0.0065, 0.1074, 0.8861)$$

A $\Pi_2 = 0.8861$ indicates that the system S is operating at full capacity 88% of the time. A $\Pi_1 = 0.1074$ means that S is operating at a Degraded capacity 10% of the time. Only Π_0 , the probability corresponding to State 0 (Down state), is associated with the system being Unavailable. The “long run” system Availability is then: $1 - \Pi_0 = 1 - 0.0065 = 0.9935$.

Finally, we are also interested in the expected times for System S to go Down if initially S was in State Up (denoted V_1) or Degraded (V_2), or in the average time S spent in each of these states before going “Down”. We obtain them by assuming Down is an “absorbing” state (one that, once entered, can never be left) and solving the linear system of equations leading to all such possible situations. That is, one step is taken at minimum (when the system goes Down, directly). If S is not absorbed in one step, then it will necessarily move on to any of other, non-absorbing (Up or Degraded) states, with the corresponding probability, and the process restarts.

$$V_1 = 1 + p_{11}V_1 + p_{12}V_2 = 1 + 0.965V_1 + 0.033V_2$$

$$V_2 = 1 + p_{21}V_1 + p_{22}V_2 = 1 + 0.004V_1 + 0.996V_2$$

Average times until System S goes down yield $V_1 = 4,625$ hours (starting in state Degraded) and $V_2 = 4,875$ (starting Up). For comparison, the non maintained system version referred to initially, would work an Expected $3/2\lambda = 3/0.004 = 750$ hours in Up state, before going Down (Reference 8). The fact that main-

tenance is now possible, while S continues operating in a Degraded state (with a single unit), results in an increase of $\mu/2\lambda^2 = 0.033/2 \times 0.002^2 = 4,125$ hours in its Expected Time to go Down (from Up). Verify that the new Expected Time is due to the sum of Expected times to failures, plus *maintenance*: $V_2 = 3/2\lambda + \mu/2\lambda^2 = 750 + 4,125 = 4,875$.

Model Extensions and Comparisons

We have seen how a stochastic process is just a R.V. $X(T)$, indexed in some parameter T called “time.” The processes overviewed here are collectively known as “discrete time parameter” Markov Chains, because transitions only occur at regular intervals (in our examples, every hour). Hourly time intervals can be shortened (to minutes, seconds, etc.) and $X(T)$ approximates a “continuous time parameter” Markov Chain (also known as a Markov Processes) just like a Riemann sum approximates an Integral.

We have not dealt with continuous time parameter Markov Chains in this START sheet, because their mathematical treatment requires using differential equations, Laplace Transforms and other tools of advanced calculus and mathematics. The objective of this START sheet is not to discuss mathematics, but to convey important statistical principles to the engineer who uses software and tools that implement them. The reader interested in learning more about these advanced methods is referred to the sources in the Bibliography.

Reference (6), is a START sheet that discusses the mathematical derivation of a simple continuous time parameter Markov Chain and some uses. It is available on our web site at: <http://rac.alionscience.com/pdf/MARKOV.pdf>. Reference 7, also discusses these models in more detail. Reference 5, Chapter 10, is an older but classic reliability book that treats this problem at introductory level. Reference 1, Chapter 6, is a recent textbook that treats the subject extensively and in a more mathematically advanced way. Reference 4, is a mathematics book about stochastic modeling with a clear approach to the topic. Finally, Reference 8 develops the system example used for comparison here.

We have discussed extensively, however, the understanding and use of several important statistical models. Among them, Availability via RV transformation and via defining a Markov Chain that represents the system as it moves through time. In doing so, we have shown how different but complementary statistical modeling approaches provide different answers to different types of problems and questions.

For example, if the problem is one of characterizing the RV Availability (A) via finding a confidence interval, a percentile (Life L_{10}) or the specific probability of some events (say, $A > 0.9$) then we may want to derive the distribution of A, directly. Obtaining such theoretical distributions may not always be easy.

But then, one can resort to MC methods, which will provide working approximations to the exact but unavailable solutions.

If the system is more complex, involving redundancy, degradation, etc. and one is more interested in asymptotic or steady state results, we may want to implement a Markov model. PM such as “long run” Availability (state occupancies), Expected time to failure, etc. can also be obtained as the system $X(T)$ moves through time. Markov Model assumptions (e.g., that distributions of times to failure, to repair, etc. should be Exponential) are some times unrealistic. But here too, one can resort to Monte Carlo methods.

Some software packages (e.g., BlockSim) implement some of these models and methods. Knowledge of the mathematics involved in model development is no longer necessary for the engineer. But a better understanding of the nature and implications of the methods they implement provides a safer use of such software packages.

Finally, it becomes clear that there are two ways of improving Availability: either extending the system life or improving its Maintainability. Logistics deals directly with the latter issue and its implications. Due to its complexity, Logistics will be the topic of a forthcoming START sheets.

For Further Study

1. System Reliability Theory: Models and Statistical Methods, Hoyland, A. and M. Rausand, Wiley, NY, 1994.
2. Logistics Engineering and Management, Blanchard, B.S., Prentice Hall, NJ, 1998.
3. Criscimagna, N., RAC, Personal communication.
4. An Introduction to Stochastic Modeling, Taylor, H. and S. Karlin, Academic Press, NY, 1993.
5. Methods for Statistical Analysis of Reliability and Life Data, Mann, N., R. Schafer, and N. Singpurwalla, John Wiley, NY, 1974.
6. Applicability of Markov Analysis Methods to Availability, Maintainability and Safety, Fuqua, N., RAC START Sheet, Volume 10, No. 2, <http://rac.alionscience.com/pdf/MARKOV.pdf>.
7. Appendix C of the Operational Availability Handbook (OPAH), Manary, J. RAC.
8. Understanding Series and Parallel Systems Reliability, Romeu, J.L., RAC START Sheet, Volume 11, No. 5 <http://rac.alionscience.com/pdf/S&PSYSREL.pdf>.

About the Author

Dr. Jorge Luis Romeu has over thirty years of statistical and operations research experience in consulting, research, and teaching. He was a consultant for the petrochemical, construction, and agricultural industries. Dr. Romeu has also worked in statistical and simulation modeling and in data analysis of software and hardware reliability, software engineering, and ecological problems.

Dr. Romeu has taught undergraduate and graduate statistics, operations research, and computer science in several American and foreign universities. He teaches short, intensive professional training courses. He is currently an Adjunct Professor of Statistics and Operations Research for Syracuse University and a Practicing Faculty of that school's Institute for Manufacturing Enterprises.

For his work in education and research and for his publications and presentations, Dr. Romeu has been elected Chartered Statistician Fellow of the Royal Statistical Society, Full Member of the Operations Research Society of America, and Fellow of the Institute of Statisticians.

Romeu has received several international grants and awards, including a Fulbright Senior Lectureship and a Speaker Specialist Grant from the Department of State, in Mexico. He has extensive experience in international assignments in Spain and Latin America and is fluent in Spanish, English, and French.

Romeu is a senior technical advisor for reliability and advanced information technology research with Alion Science and Technology previously IIT Research Institute (IITRI). Since rejoining Alion in 1998, Romeu has provided consulting for several statistical and operations research projects. He has written a

State of the Art Report on Statistical Analysis of Materials Data, designed and taught a three-day intensive statistics course for practicing engineers, and written a series of articles on statistics and data analysis for the AMPTIAC Newsletter and RAC Journal.

Other START Sheets Available

Many Selected Topics in Assurance Related Technologies (START) sheets have been published on subjects of interest in reliability, maintainability, quality, and supportability. START sheets are available on-line in their entirety at <http://rac.alionscience.com/rac/jsp/start/startsheet.jsp>.

For further information on RAC START Sheets contact the:

Reliability Analysis Center
201 Mill Street
Rome, NY 13440-6916
Toll Free: (888) RAC-USER
Fax: (315) 337-9932

or visit our web site at:

<http://rac.alionscience.com>



About the Reliability Analysis Center

The Reliability Analysis Center is a world-wide focal point for efforts to improve the reliability, maintainability, supportability and quality of manufactured components and systems. To this end, RAC collects, analyzes, archives in computerized databases, and publishes data concerning the quality and reliability of equipments and systems, as well as the microcircuit, discrete semiconductor, electronics, and electromechanical and mechanical components that comprise them. RAC also evaluates and publishes information on engineering techniques and methods. Information is distributed through data compilations, application guides, data products and programs on computer media, public and private training courses, and consulting services. Alion, and its predecessor company IIT Research Institute, have operated the RAC continuously since its creation in 1968.