Probability in the Engineering and Informational Sciences, **15**, 2001, 369–381. Printed in the U.S.A.

AVAILABILITY OF CONTINUOUS SERVICE AND COMPUTING LONG-RUN MTBF AND RELIABILITY FOR MARKOV SYSTEMS

John E. Angus
Claremont Graduate University
Claremont, California 91711
E-mail: john.angus@cgu.edu

Steady-state availability has long been a popular descriptor of effectiveness for repairable systems because it captures both the operability and repairability aspects of the system. A related measure of effectiveness is the availability of continuous service, which is particularly relevant for safety critical applications. In this article, two different measures of this quantity are described for a repairable system whose state is described by an ergodic finite-state-space continuous-time Markov chain. Using these ideas, formulas for computing system long-run mean time between failures and the long-run system reliability function are derived.

1. INTRODUCTION AND BACKGROUND

Consider a repairable system that may be in any one of the states in $S = \{0,1,\ldots,J\}$, J a nonnegative integer, and suppose that the system state at time t is given by an irreducible, homogeneous, continuous-time ergodic Markov process $\{X(t), t \geq 0\}$ with state space S. This process begins in an initial state $i \in S$, remains there for a random amount of time that is exponentially distributed with mean $1/q_i \in (0,\infty)$, then transitions independently to state j with probability $Q(i,j)$, and continues in this fashion. The rows of the matrix $Q = (Q(i,j))$ define probability distributions on S, and the diagonal elements of Q are all 0. Denote the infinitesimal parameters (i.e., one-step transition rates) for this process by

$$\begin{aligned} r_{ij} &= q_i Q(i,j), \qquad i \neq j, \\ r_{ii} &= -q_i \end{aligned} \tag{1.1}$$

 0269-9648/01 $12.50

https://doi.org/10.1017/S0269964801153064 Published online by Cambridge University Press

and the transition probabilities by $p_{ij}(t) = P\{X(t+s) = j \mid X(s) = i\}$. Denote by $\Lambda = (r_{ij})$ the matrix of infinitesimal parameters defined in Eq. (1.1), and by $P(t) = (p_{ij}(t))$ the matrix of transition probabilities. It is well known (see Ross [13] or Cinlar [5]) that the transition probabilities satisfy the Kolmogorov backward and forward equations:

$$p'_{ij}(t) = \sum_{k\in S} r_{ik} p_{kj}(t) = \sum_{k\in S} p_{ik}(t) r_{kj}, \tag{1.2}$$

which may be written compactly in matrix form as

$$P'(t) = \Lambda P(t) = P(t)\Lambda \tag{1.3}$$

with initial condition $P(0) = I$. It follows that the transition probabilities may be expressed by the matrix exponential

$$P(t) = \exp(\Lambda t) \tag{1.4}$$

and that, because the process is assumed irreducible and ergodic, there exists a stationary distribution π on S such that for every $j \in S$,

$$\lim_{t\to\infty} p_{ij}(t) = \pi_j > 0; \tag{1.5}$$

that is, π represents the "steady-state" or limiting state-occupancy probabilities. It follows under these assumptions that $\pi = (\pi_0, \pi_1, \ldots, \pi_J)$ is the unique solution to

$$\pi\Lambda = 0,$$
$$\sum_{j\in S} \pi_j = 1. \tag{1.6}$$

Efficient methods for approximating Eq. (1.4) are described in Ross [11] and Angus [1]. Finding π in Eq. (1.5) by solving Eq. (1.6) can often be carried out analytically to yield a closed-form solution for many systems of interest (e.g., parallel, k-out-of-n, and series systems under various standby and repair assumptions). When a closed-form solution is not possible, efficient and accurate numerical algorithms are known (see Stewart [14], for example).

The state space S can be partitioned into mutually exclusive sets O and F, the states in O representing operational states in which the system is providing a nominal level (or better) of service (i.e., "up" states) and the states in F representing failure states in which the system has failed to provide service (i.e., "down" states). To avoid trivial cases, it is assumed that both O and S are nonempty. The steady-state availability of the system is then given by

$$A \equiv \sum_{j\in O} \pi_j \tag{1.7}$$

and is interpretable both as the limiting probability of finding the system in one of the states in O and as the long-run fraction of time that the system will be in one of

https://doi.org/10.1017/S0269964801153064 Published online by Cambridge University Press

the states in O. To see that A may be interpreted both ways, it is instructive to recall the following result, whose proof may be found in Norris [10, p. 126].

THEOREM 1.1: *Let $\{X(t), t \geq 0\}$ be irreducible and positive recurrent with invariant distribution π, and let f be a bounded function defined on S. Then,*

$$P\left\{\lim_{t\to\infty} \frac{1}{t}\int_0^t f(X(u))\,du = \bar{f}\right\} = 1, \tag{1.8}$$

where

$$\bar{f} = \sum_{i\in S} \pi_i f(i). \tag{1.9}$$

By applying the theorem with $f(i) = 1$ if $i \in O$ and 0 otherwise, it follows that

$$\lim_{t\to\infty} \frac{1}{t}\int_0^t 1_{\{X(u)\in O\}}\,du = A \tag{1.10}$$

with probability 1; that is, A is interpretable as the long-run fraction of time that the system will be in O.

Alternatively, let $U_1, U_2, \ldots$ denote the successive system sojourn times in O and $D_1, D_2, \ldots$ denote the successive system sojourn times in F; that is, assuming the system starts in an operational state initially, the system spends U_1 time units operational, followed by D_1 time units failed, followed by U_2 time units operational, followed by D_2 time units failed, and so on throughout time.

Because of the strong Markov property that $\{X(t), t \geq 0\}$ enjoys, $U_1, U_2, \ldots$ are independent and $D_1, D_2, \ldots$ are independent, although neither sequence need be identically distributed. However, each sequence has a limit distribution with finite mean (i.e., a steady-state distribution) as demonstrated by following problem 5.34 of Ross [12]. Moreover, in both cases, the sequences are stochastically bounded by random variables with finite means (i.e., their survivor functions are uniformly bounded above by the survivor function of the respective bounding random variable). To see this, consider first U_n. By considering the process to start in state $i \in O$ at time 0, we can compute the distribution of the first passage time out of the set O by making the set F absorbing. The survivor function of this distribution, $R_i(\cdot)$, is the reliability function starting in state $i \in O$, and methods for computing it and its mean will be given in Eqs. (1.16)–(1.21). Regardless of where the process entered O for the nth time, the time the process could spend in O is always stochastically smaller than the random variable that has a distribution function which is the convolution of all the distributions corresponding to the reliability functions $R_i(\cdot)$, $i \in O$, and the mean of this stochastically bounding random variable is the sum of the means $\theta_i = \int_0^\infty R_i(t)\,dt$ over $i \in O$. By interchanging the role of up states and down states, this argument exhibits a stochastically bounding random variable for D_n having finite mean. It therefore follows

https://doi.org/10.1017/S0269964801153064 Published online by Cambridge University Press

using the dominated convergence theorem as in the proof of Theorem 1 of Mi [9] and Lemma 1 of Mi [8] that the limits

$$\theta = \lim_{n\to\infty} \frac{1}{n} \sum_{i=1}^{n} E(U_i) = \lim_{n\to\infty} \frac{1}{n} \sum_{i=1}^{n} U_i,$$

$$\tau = \lim_{n\to\infty} \frac{1}{n} \sum_{i=1}^{n} E(D_i) = \lim_{n\to\infty} \frac{1}{n} \sum_{i=1}^{n} D_i \tag{1.11}$$

exist with probability 1 and are finite and that θ and τ are the means of the limiting distributions of U_n and D_n, respectively. Explicit formulas for θ and τ will be derived later, but for now, these can be related to the steady-state availability via

$$A = \frac{\theta}{\theta + \tau}. \tag{1.12}$$

To see why this is true, for $t \geq 0$ and $n \geq 0$, let $N(t) = n$ on the set where $\sum_{i=1}^{n}(U_i + D_i) \leq t < \sum_{i=1}^{n+1}(U_i + D_i)$, where a sum with an upper limit of 0 is defined to be 0. Thus, assuming that the process starts in a state in O at time 0, $N(t)$ is the number of complete up–down cycles during the time period $[0, t]$. Obviously, $N(t) \to \infty$ as $t \to \infty$ with probability 1 because the process is ergodic, and moreover,

$$\frac{\sum_{i=1}^{N(t)} (U_i + D_i)}{N(t)} \leq \frac{t}{N(t)} \leq \frac{\sum_{i=1}^{N(t)+1} (U_i + D_i)}{N(t)}. \tag{1.13}$$

Taking $t \to \infty$ and applying Eqs. (1.11) shows that, with probability 1,

$$\lim_{t\to\infty} \frac{N(t)}{t} = \frac{1}{\theta + \tau}. \tag{1.14}$$

Next, for t sufficiently large that $N(t) > 0$, the inequalities

$$\frac{\sum_{i=1}^{N(t)} U_i}{N(t)} \frac{N(t)}{t} \leq \frac{1}{t} \int_0^t 1_{\{X(u)\in O\}}\, du \leq \frac{\sum_{i=1}^{N(t)+1} U_i}{N(t)+1} \frac{N(t)+1}{t} \tag{1.15}$$

hold, so, again, letting $t \to \infty$ and applying Eqs. (1.10) and (1.11), the result [Eq. (1.12)] follows. See Ross [13] or Barlow [2], for example, for further discussions along these lines in the reliability context.

In reliability characterization terms, both A and θ are common measures of a system's effectiveness. The parameter θ is the long-run (or steady state) mean time

https://doi.org/10.1017/S0269964801153064 Published online by Cambridge University Press

between failures (MTBF) for the system, and τ is the long-run mean time to repair (MTTR) the system. Typically, the qualifier "long run" is dropped, but it is important to remember that, in general, the sequence $U_1, U_2, \ldots$ is not independent and identically distributed (and similarly for the sequence $D_1, D_2, \ldots$). Cases where the sequence $\{U_i\}$ is indeed independent and identically distributed (at least after $i = 1$) occur when the instantaneous transition out of a state of failure into an operating state always leaves the system in the same operating state. In this case, the reentry into operation from failure constitutes a regeneration point for the process. Many structures that fit into this category have been analyzed and closed-form expressions for A and θ have been worked out. See Kozlov and Ushakov [7], for example, for an encyclopedia of series, parallel, and k-out-of-n structures in the Markov case. When the system is not necessarily Markovian but the individual components are statistically independent with respect to failures and repairs, Ross [13] or Barlow [2] give general formulas for A and θ in terms of the system reliability function (i.e., the h function) and the individual component mean lives and mean repair times.

The parameter θ is often confused with the mean time to (first) failure (MTTF), the latter usually signifying the mean time to failure after starting at time 0 in some fixed initial state $j \in O$. Generally, MTTF is not equal to MTBF, but in the case just described, where instantaneous reentry into operation from failure always leaves the system in the same operating state, the two would be equal if the initial system state is chosen to be that operating state. This is, in fact, how MTBF is computed for those special cases (which include those treated in Kozlov and Ushakov [7]). When reentry to the operating states O can occur at more than one state, the computation of θ is not as straightforward. In any case, having expressions for A and θ immediately yields τ from Eq. (1.12).

Starting at time 0 in state $j \in O$, let $T = \inf\{t \geq 0 : X(t) \in F\}$. Of course, the MTTF starting in state j is

$$\theta_j = E(T \mid X(0) = j). \tag{1.16}$$

Also useful is the reliability of the system starting in j at time 0. This function is defined by

$$R_j(t) = P\{T > t \mid X(0) = j\}. \tag{1.17}$$

The functions $R_j, j \in O$, are not difficult to find in principle. One simply replaces r_{ik} by 0 in Λ for all $k \in S$ for each $i \in F$, solves either Eq. (1.2), (1.3), or (1.4) for $p_{jk}(t)$ for all $j, k \in O$, and then sets $R_j(t) = \sum_{k \in O} p_{jk}(t)$. By doing this, we have effectively lumped all of the down states into one absorbing down state (i.e., once the process enters it, it cannot get out) and then computed the probability that, starting in an up state, the process is in O at time t. Since entering the down state prior to t would render it impossible to be in an up state at time t, this is then the probability that the process transitions entirely among up states throughout the time period $[0, t]$, which, in turn, represents $P\{T > t \mid X(0) = j\}$. Following this recipe, it is not difficult to show that the functions $R_j, j \in O$, are the unique solution to the system of first-order linear differential equations (analogous to the Kolmogorov backward equations):

$$R_j'(t) = \sum_{k \in O} r_{jk} R_k(t), \qquad j \in O,$$
$$R_j(0) = 1, \qquad j \in O. \tag{1.18}$$

By rearranging the state labels, we can assume without loss of generality that $O = \{0,1,\ldots,k\}$ and $F = \{k+1,\ldots,J\}$ and that the matrix Λ is partitioned as

$$\Lambda = \begin{pmatrix} \Lambda_{oo} & \Lambda_{of} \\ \Lambda_{fo} & \Lambda_{ff} \end{pmatrix}, \tag{1.19}$$

where Λ_{oo} represents transitions from operating states to operating states, Λ_{of} represents transitions from operating states to failed states, Λ_{fo} represents transitions from failed states to operating states, and Λ_{ff} represents transitions from failed states to failed states. Then, viewing R as the column vector of the functions R_j, $j \in O$, Eq. (1.18) can be expressed as

$$R'(t) = \Lambda_{oo} R(t), \qquad R(0) = (1,1,\ldots,1)^T. \tag{1.20}$$

Integrating Eq. (1.20) and setting $\Theta = (\theta_0, \theta_1, \ldots, \theta_k)^T$, it is seen that the column vector of mean times to failure solves

$$\Lambda_{oo}\Theta = -(1,1,\ldots,1)^T. \tag{1.21}$$

2. AVAILABILITY OF CONTINUOUS SERVICE

In modern safety critical systems, it is necessary to know how frequently (or with what probability) the system is in a state which is not only a nominal “up” state, but for which there is some assurance that it will remain in an up state for a “comfortable” period of time. For example, in air traffic control or civil aviation navigation systems, the critical phase of a so-called “Category I” precision approach lasts for 150 s, during which highly trusted and accurate navigation must prevail continuously with high probability; that is, for applications like this, some measure of the “availability of continuous service” is needed.

The concept of the probability of having assurance of continuous service is not new. One version has been variously described as the “interval reliability” (see Barlow and Proschan [4]) and “interval availability” (see Mi [9]), although the latter term is normally reserved to describe the fraction of time during an interval that the system is in one of its up states. In either case, it is desirable to know the probability of finding the system up at time t, and that it continues to remain up without entering a down state throughout the ensuing Δt time units. In the context of the model used here, assuming the system starts in state i, this would be

$$P\{X(s) \in O, t \le s \le t + \Delta t | X(0) = i\} \tag{2.1}$$

https://doi.org/10.1017/S0269964801153064 Published online by Cambridge University Press

and it follows from the Markov property that this is given by

$$P\{X(s) \in O, t \le s \le t + \Delta t | X(0) = i\}$$
$$= \sum_{j \in O} p_{ij}(t) P\{X(s) \in O, 0 \le s \le \Delta t | X(0) = j\}. \quad \textbf{(2.2)}$$

Typically, the interest is in the steady state as $t \to \infty$, and in this limit, Eq. (2.2) becomes

$$A_1(\Delta t) \equiv \sum_{j \in O} \pi_j R_j(\Delta t), \quad \textbf{(2.3)}$$

where R_j is the reliability of the system when it starts in state j at time 0.

Because the Markov process is ergodic, the measure $A_1(\Delta t)$ is the steady-state probability of being operational at a given point in time and remaining operational for the ensuing Δt time units and it also represents the long-run fraction of total time that the system spends in O for which the exit time from O is at least Δt time units. That the latter interpretation is valid is made clear later when Eq. (2.3) is related to the results of Mi [9]. Thus, this is one measure that fulfills the aforementioned need in safety critical systems.

One practical disadvantage of the approach leading to the measure $A_1(\Delta t)$ is that one cannot at any given instant in time look at the system and instantaneously determine if the system is operational *and* will remain so for the next Δt time units. An alternative approach to alleviate this difficulty is to define the availability of reliable service. This measure is defined like the usual steady-state availability, except that only certain operational states are included in the sum [Eq. (1.7)]; specifically, only operational states $j \in O$ for which $R_j(\Delta t) \ge c$ are included, where c is a prespecified (usually high) probability. Mathematically, the availability of reliable service is defined as

$$A_c(\Delta t) = \sum_{j \in O} \pi_j 1_{\{R_j(\Delta t) \ge c\}}. \quad \textbf{(2.4)}$$

This is just steady-state availability, only with operational states redefined to require not only that $j \in O$ but also that $R_j(\Delta t) \ge c$. When c is high, this usually has the effect of limiting the operational states to only those in which there remains some level of component redundancy (i.e., in which a single component failure will not cause system failure), and this makes this measure particularly attractive for specifying performance in safety critical systems. This measure also overcomes the disadvantage of the measure $A_1(\Delta t)$ because it can be determined a priori, which states that $j \in O$ satisfy $R_j(\Delta t) \ge c$.

When Δt is small, Eqs. (2.3) and (2.4) can often be adequately approximated by substituting

$$R_j(\Delta t) \approx 1 + R_j'(0+)\Delta t = 1 - \sum_{k \in F} r_{jk} \Delta t \quad \textbf{(2.5)}$$

by Eq. (1.18) (since the rows of Λ sum to 0). In this case,

$$A_c(\Delta t) \approx \sum_{j\in O} \pi_j 1_{\{\sum_{k\in F} r_{jk}\Delta t \le 1-c\}} \tag{2.6}$$

and

$$A_1(\Delta t) \approx A - \sum_{j\in O} \pi_j \sum_{k\in F} r_{jk}\,\Delta t. \tag{2.7}$$

In fact, it will be seen later (by combining Eqs. (1.12) and (3.6)) that the right-hand side of Eq. (2.7) is actually $A(1 - \Delta t/\theta)$.

The measure $A_c(\Delta t)$ is specified, for example, in the Federal Aviation Administration's Wide Area Augmentation System (WAAS) specification (see [15]), and is sometimes referred to in engineering jargon as the availability of continuity-of-function (or continuity-of-service). The WAAS system is designed to augment the Global Positioning System for use in providing safe and accurate navigation in civil aviation. In that specification (for the first phase of the system), the number $c = 1 - 10^{-5}$ and $\Delta t = 1$ h for en route and nonprecision approach modes of flight, and $c = 1 - (5.5 \times 10^{-5})$ and $\Delta t = 150$ s for the Category I precision approach. Using these numbers, that specification requires that the availability of reliable service be 0.999 for en route and nonprecision approach modes of flight, and 0.95 for precision approach.

To illustrate and compare these measures, consider a simple parallel system with two statistically independent and identical units (or components), each having a constant failure rate λ and each with constant repair rate μ. We will assume that units are switched on (when not failed), repair is unlimited (i.e., that there are enough repair resources to simultaneously work on repairing both units if need be), and that this system is nominally operational as long as at least one unit is functioning. We shall denote the nominally operating states as $O = \{0,1\}$ and the failure state as $F = \{2\}$. The transition rate matrix for this system is

$$\Lambda = \begin{pmatrix} -2\lambda & 2\lambda & 0 \\ \mu & -(\lambda+\mu) & \lambda \\ 0 & 2\mu & -2\mu \end{pmatrix} \tag{2.8}$$

and the steady-state probabilities are

$$\pi_0 = \frac{\mu^2}{(\lambda+\mu)^2}, \qquad \pi_1 = \frac{2\lambda\mu}{(\lambda+\mu)^2}, \qquad \pi_2 = \frac{\lambda^2}{(\lambda+\mu)^2}, \tag{2.9}$$

so that the steady-state availability is

$$A = \frac{\mu^2 + 2\lambda\mu}{(\lambda+\mu)^2}. \tag{2.10}$$

https://doi.org/10.1017/S0269964801153064 Published online by Cambridge University Press

The matrix Λ_{oo} is

$$\Lambda_{oo} = \begin{pmatrix} -2\lambda & 2\lambda \\ \mu & -(\lambda+\mu) \end{pmatrix} \tag{2.11}$$

and the solution to $R'(t) = \Lambda_{oo} R(t)$ gives the reliability functions

$$R_1(t) = c_1 e^{\alpha_1 t} + c_2 e^{\alpha_2 t},$$
$$R_0(t) = d_1 e^{\alpha_1 t} + d_2 e^{\alpha_2 t}, \tag{2.12}$$

where

$$\alpha_1 = \frac{-(3\lambda+\mu) + \sqrt{\lambda^2 + 6\lambda\mu + \mu^2}}{2}, \tag{2.13}$$

$$\alpha_2 = \frac{-(3\lambda+\mu) - \sqrt{\lambda^2 + 6\lambda\mu + \mu^2}}{2}, \tag{2.14}$$

$$c_1 = \frac{\lambda + \alpha_2}{\alpha_2 - \alpha_1}, \qquad c_2 = \frac{-\alpha_1 - \lambda}{\alpha_2 - \alpha_1}, \tag{2.15}$$

$$d_1 = \frac{\alpha_2}{\alpha_2 - \alpha_1}, \quad d_2 = \frac{-\alpha_1}{\alpha_2 - \alpha_1}. \tag{2.16}$$

Typical values for the parameters are $\lambda = 0.001$ per hour and $\mu = 1$ per hour, yielding the value for availability $A = 1 - 10^{-6}$. Using $c = 1 - 10^{-5}$ and $\Delta t = 1$ h, it turns out that $A_1(\Delta t) = 0.999997$, not far below the availability. Applying the above formulas for the reliabilities, $R_0(1) = 0.999999265$ and $R_1(1) = 0.9993679$, and so $A_c(\Delta t) = \pi_0 = 0.998$. Note that this measure is considerably more conservative than the others: There is no credit for the nominal operating state 1 because starting in that state, the probability of surviving 1 h is too low. In order to boost $A_c(\Delta t)$ above 0.999, an additional redundant unit would be needed.

3. LONG-RUN MTBF, MTTR, AND RELIABILITY

Despite the importance and usefulness of the long-run MTBF and MTTR (θ and τ) for describing system effectiveness, some of the most popular mathematical reliability sources (e.g., Hoyland and Rausand [6], Barlow and Proschan [3,4], Kozlov and Ushakov [7]) do not present formulas for computing them for the case where the system state follows a continuous-time Markov process as discussed here (arguably the most commonly assumed case in industry). Expressions for these, which do not seem to be widely known among reliability engineers, are derived in this section. It will be seen shortly that the computations of long-run MTBF, MTTR, and system reliability can be accomplished surprisingly easily by making a connection with the steady-state availability of continuous service [Eq. (2.3)], a connection facilitated by the work of Mi [9], who gives a fairly general treatment of interval availability.

Let R_S denote the long-run system reliability function (i.e., the limiting distribution of U_n as $n \to \infty$). By Eq. (1.11) and the justifying discussions in the paragraph preceding it, conditions (i) through (iv) of Theorem 1 of Mi hold true for the sequences $U_1, U_2, \ldots$ and $D_1, D_2, \ldots$. It then follows from Theorem 2 of Mi [9] that

$$\frac{1}{T}\int_0^T 1_{\{X(s)\in O,\, u\le s\le u+t\}}\,du \to \frac{\displaystyle\int_t^\infty R_S(u)\,du}{\theta+\tau} \tag{3.1}$$

with probability 1 and

$$E\left[\left(\frac{1}{T}\int_0^T 1_{\{X(s)\in O,\, u\le s\le u+t\}}\,du\right)\middle| X(0)=i\right] \to \frac{\displaystyle\int_t^\infty R_S(u)\,du}{\theta+\tau} \tag{3.2}$$

as $T \to \infty$. However, since the integrand is everywhere nonnegative, Fubini's theorem gives

$$\begin{aligned} &E\left[\left(\frac{1}{T}\int_0^T 1_{\{X(s)\in O,\, u\le s\le u+t\}}\,du\right)\middle| X(0)=i\right] \\ &\quad = \frac{1}{T}\int_0^T E\left[1_{\{X(s)\in O,\, u\le s\le u+t\}} \middle| X(0)=i\right] du && (3.3) \\ &\quad = \frac{1}{T}\int_0^T \sum_{j\in O} p_{ij}(u) P\{X(s)\in O, 0\le s\le \Delta t \,|\, X(0)=j\}\,du \\ &\quad \to \sum_{j\in O} \pi_j R_j(t) && (3.4) \end{aligned}$$

as $T \to \infty$. Relation Eq. (3.4) follows by Eqs. (2.2) and (2.3) and the fact that $(1/T)\int_0^T f(u)\,du \to c$ as $T \to \infty$ whenever f is integrable and $f(u) \to c$ as $u \to \infty$. This shows that the measure $A_1(t)$ does, indeed, have the interpretation as the long-run fraction of total time that the system spends in O for which the exit time from O is at least t time units. Accordingly, combining Eqs. (3.2)–(3.4), it follows that

$$A_1(t) \equiv \sum_{j\in O} \pi_j R_j(t) = \frac{\displaystyle\int_t^\infty R_S(u)\,du}{\theta+\tau}. \tag{3.5}$$

Differentiating this at $t = 0+$ gives, using Eq. (1.18),

$$\sum_{i\in O} \pi_i \sum_{k\in O} r_{ik} = -\frac{1}{\theta+\tau} = -\sum_{i\in O} \pi_i \sum_{k\in F} r_{ik}, \tag{3.6}$$

which, after applying Eq. (1.12), yields

$$\theta = -\frac{\sum_{i\in O}\pi_i}{\sum_{i\in O}\pi_i\sum_{k\in O}r_{ik}} = \frac{\sum_{i\in O}\pi_i}{\sum_{i\in O}\pi_i\sum_{k\in F}r_{ik}}. \tag{3.7}$$

The second equality above and the second equality in Eq. (3.6) follows by simply noticing that the limiting probability $\{\pi_i, i\in S\}$ must satisfy Eq. (1.6) and the rows of Λ each sum to 0, so

$$\sum_{i\in O}\pi_i\sum_{k\in F}r_{ik} = -\sum_{i\in O}\pi_i\sum_{k\in O}r_{ik} = \sum_{i\in F}\pi_i\sum_{k\in O}r_{ik}. \tag{3.8}$$

Now, from the relationship (1.12), the long-run MTTR is seen to be

$$\tau = \frac{1-\sum_{i\in O}\pi_i}{\sum_{i\in O}\pi_i\sum_{k\in F}r_{ik}} = \frac{\sum_{i\in F}\pi_i}{\sum_{i\in F}\pi_i\sum_{k\in O}r_{ik}}. \tag{3.9}$$

Also, differentiating Eq. (3.5) at $t+$ and using Eqs. (1.18), (3.7), (3.8), and (3.9) yields

$$R_S(t) = -\frac{\sum_{i\in O}\pi_i\sum_{k\in O}r_{ik}R_k(t)}{\sum_{i\in O}\pi_i\sum_{k\in F}r_{ik}} = \frac{\sum_{i\in F}\pi_i\sum_{k\in O}r_{ik}R_k(t)}{\sum_{i\in F}\pi_i\sum_{k\in O}r_{ik}} \tag{3.10}$$

as the long-run system reliability function (i.e., the steady-state reliability function for the sojourn time in the set O).

If T_O represents the sojourn time in the set O at steady state, then the previous analysis shows that $R_S(t) = P\{T_O > t\}$ and $E(T_O) = \theta$. A related quantity is the exit time from the set O, given the process is somewhere in the set O. If at steady state, this quantity is denoted by T_x, then

$$P\{T_x > t\} = \lim_{u\to\infty} P\{X(s)\in O, u\le s\le u+t \mid X(u)\in O\} \tag{3.11}$$

and it follows from the relation

$$\begin{aligned} &P\{X(s)\in O, u\le s\le u+t\} \\ &\quad = P\{X(s)\in O, u\le s\le u+t \mid X(u)\in O\}P\{X(u)\in O\} \end{aligned} \tag{3.12}$$

along with Eqs. (3.11), (2.2), and (2.3) that

$$\sum_{j\in O}\pi_j R_j(t) = AP\{T_x > t\}. \tag{3.13}$$

https://doi.org/10.1017/S0269964801153064 Published online by Cambridge University Press

Using this in Eq. (3.5) and applying Eq. (1.12) gives

$$P\{T_x > t\} = \frac{\int_t^{\infty} R_S(u)\,du}{\theta}, \tag{3.14}$$

and integrating Eqs. (3.13) and (3.14) gives

$$E(T_x) = \frac{E(T_O^2)}{2E(T_O)} = A^{-1} \sum_{j \in O} \pi_j \theta_j, \tag{3.15}$$

where θ_j is the MTTF starting in state $j \in O$, given in Eq. (1.21). This also serves to provide a formula for the second moment of T_O and to show that T_x has the same relation to T_O as the steady-state excess or residual life has to the time between renewals in a renewal process. Ross [12, Prob. 5.34] also points out this structural similarity and gives an outline of a (different) multistep approach to find the steady-state distribution and mean of the sojourn time in an arbitrary set of states $B \subset S$ (i.e., the time spent in B during a visit). Taking $B = O$ gives these measures the interpretation of steady-state reliability and MTBF, respectively, and those results then match Eqs. (3.7) and (3.10).

4. SUMMARY AND FINAL REMARKS

Two availability measures appropriate for safety critical applications have been discussed and illustrated for systems whose state evolves according to a continuous Markov process: the availability of continuous service and the availability of reliable service. Both measures are computationally straightforward and are particularly relevant to safety critical applications. At the same time, general expressions for the long-run values of MTBF, MTTR, and system reliability have been derived as easy consequences of applying the concept of steady-state availability of continuous service. These formulas are not widely exposed in popular reliability methodology sources, and this discussion serves to make them accessible to practitioners.

References

1. Angus, J.E. (1992). Some bounds on the error in approximating transition probabilities in continuous-time Markov processes. *SIAM Review* 34(1): 110–113.
2. Barlow, R.E. (1998). *Engineering reliability.* Philadelphia: SIAM.
3. Barlow, R.E. & Proschan, F. (1966). *Statistical theory of reliability and life testing.* Silver Spring, MD: To Begin With.
4. Barlow, R.E. & Proschan, F. (1996). *Mathematical theory of reliability.* Philadelphia: SIAM.
5. Cinlar, E. (1975). *Introduction to stochastic processes.* Englewood Cliffs, NJ: Prentice-Hall.
6. Hoyland, A. & Rausand, M. (1994). *System reliability theory: Models and statistical methods.* New York: Wiley.
7. Kozlov, B.A. & Ushakov, I.A. (1970). *Reliability handbook.* New York: Holt, Rinehart and Winston.
8. Mi, J. (1995). Limiting behavior of some measures of system availability. *Journal of Applied Probability* 32: 482–493.

https://doi.org/10.1017/S0269964801153064 Published online by Cambridge University Press

9. Mi, J. (1999). On measure of system interval availability. *Probability in the Engineering and Informational Sciences* 13: 359–375.
10. Norris, J.R. (1999). *Markov chains.* Cambridge: Cambridge University Press.
11. Ross, S.M. (1987). Approximating the transition probabilities and mean occupancy times in continuous time Markov chains. *Probability in the Engineering and Informational Sciences* 1: 251–264.
12. Ross, S.M. (1996). *Stochastic processes*, 2nd ed. New York: Wiley.
13. Ross, S.M. (1997). *Introduction to probability models*, 6th ed. San Diego: Academic Press.
14. Stewart, W.J. (1994). *Introduction to the numerical solution of Markov chains.* Princeton, NJ: Princeton University Press.
15. U.S. Department of Transportation Federal Aviation Administration Specification: Wide Area Augmentation System (WAAS). FAA-2892B. March 10, 1997.