Probability in the Engineering and Informational Sciences, **14**, 2000, 203–218. Printed in the U.S.A.

SUPPLEMENTARY VARIABLE TECHNIQUE IN STOCHASTIC MODELS

Attahiru Sule Alfa*
Department of Industrial and Manufacturing Systems Engineering
The University of Windsor
Windsor, Ontario N9B 3P4, Canada
E-mail: alfa@uwindsor.ca

T. S. S. Srinivasa Rao*
Institute for Systems Studies and Analyses
Defence R & D Organisation
Metcalfe House
Delhi-110 054, India
tssr@mailcity.com

In this article, we formalize the framework for the supplementary variable technique in stochastic models. Specifically, we show that the use of remaining or elapsed times (or any metric) as supplementary variables leads to the notion of forward or backward Chapman–Kolmogorov equations, respectively. We further show that for a class of queueing systems, using remaining time as the supplementary variable makes analysis simpler.

1. INTRODUCTION

Queueing theory forms the classical framework for modeling stochastic service processes which have strong applications in computers and communication networks, and telecommunication systems. Queueing systems are abound in society. They are rife in commercial service systems, where examples include store checkout stations, gas stations, and bank teller stations. Many social service systems and community

*This research work was carried out when the first author was with the University of Manitoba and the second author was a Post-Doctoral Fellow at the same university.

 0269-9648/00 $12.50

https://doi.org/10.1017/S0269964800142068 Published online by Cambridge University Press

emergency systems (fire, rescue, police) have elements of queueing systems. In industry, tool cribs, material-handling systems, maintenance systems, and so forth have been analyzed in a queueing system context.

The first step in analyzing a queueing system is to set it up as a Markov process. In most practical queueing systems, a supplementary variable is usually needed to achieve this. The alternative to that is an embedded Markov chain method. In the queueing literature, we have two kinds of supplementary variable, in general. One is the elapsed service (arrival) time and the other is the remaining service (arrival) time, and for both cases, the approaches of deriving the queueing characteristics are different. The main reason for adding supplementary variables to a stochastic process variable is to make the system Markovian. The use of the supplementary variable technique (SVT) in queueing dates back to 1942 when it was introduced by Kosten [16]. Later, the technique became popular for most stochastic models. This was the result of the article by Cox [4], in which he used the supplementary variable by considering elapsed service time to study the M/G/1 queueing system. Keilson and Kooharian [13] gave a solution for time-dependent queueing processes for the M/G/1 queue by considering elapsed service time as the supplementary variable. In his work on priority queues, Jaiswal [11] makes heavy use of supplementary variables by considering elapsed service time. Knessl et al. [15] applied an integral equation approach to study the stationary distribution of the number of customers in M/G/2 queueing systems by introducing the elapsed service times as supplementary variables. Elapsed time as the supplementary variable has also been used in reliability (Li and Cao [20] and Li, Shi, and Chao [21]), and in scheduling (Li, Braun, and Zhao [19]). For more details about the use of elapsed service time as the supplementary variable in various continuous and discrete queueing systems, see Cohen [2], Cooper [3], and Takagi [26–28]. The technique of remaining service time as the supplementary variable was first proposed by Henderson [9] and later used by Hokstad [10]. The transient behavior of the M/G/1 queue was studied in [10], using remaining service time as the supplementary variable. In a series of articles, Dafermos and Neuts [5], Klimko and Neuts [14], and Heimann and Neuts [8] used remaining service time as supplementary variables in discrete-time analysis of queues. A similar idea was used by Minh [22] and Alfa [1] for discrete-time analysis of time-inhomogeneous queues. In fact, Minh [22] went as far as adding another variable, which he called the "surplus" variable, that enabled him to study the departure process. The works of Minh [22] and Alfa [1] are for queues with time-varying parameters in which the embedded Markov chain approach is not the best choice. The results for the steady state system length distributions of the discrete-time GI/G/1 queue using the residual service time as the supplementary variable were obtained by Yang and Chaudhry [30]. The concept of remaining service time has also been used for analyzing finite-capacity queues; Lee [18] analyzed the M/G/1/N queue with vacation time and exhaustive service discipline by using a combination of supplementary variable and biasing techniques.

Little exists in the literature by way of discussions of the better choice of supplementary variables, whether remaining time or elapsed time. In this article, we attempt to address this issue. When we use the elapsed service (arrival) time as the supplementary variable, the hazard rate function, the integral boundary conditions, and initial conditions are needed. However, when we use remaining service (arrival) time as the supplementary variable, all of these conditions are not needed and we have much simpler calculations. From the above observations, one can assert that the technique of remaining time as the supplementary variable is simple and elegant. The main purpose of this article is to describe the importance of remaining time as the supplementary variable. The technique of remaining time can also easily be applied to study various problems in other areas such as inventory theory, scheduling theory, manufacturing systems, transportation systems, computer-communication systems, and telecommunication systems.

The article is organized as follows. In Section 2, a generalized stochastic model is presented. Section 3 gives the special cases of GI/G/1 queueing systems. In Section 4, observations for considering remaining time and elapsed time as supplementary variables are given. Sections 5 and 6 deal with special cases for remaining time and elapsed time, respectively. Special cases for finite buffer GI/G/1 queueing systems are also included in Section 7. Section 8 gives the conclusion. The Appendix gives some steady state results of the GI/G/1 system of Section 3.1.

2. GENERALIZED STOCHASTIC MODEL

Consider a stochastic process $\{X(t); t \geq 0\}$ which assumes discrete positive values $0,1,2,\ldots$. Further, define n continuous random variables parameterized by t, $\mathbf{W}(t) = \{W_1(t), W_2(t), \ldots, W_n(t)\}$ and $\mathbf{Z}(t) = \{Z_1(t), Z_2(t), \ldots, Z_n(t)\}$, which assume values in $[0,\infty)$. Let the process $[X(t), \mathbf{W}(t); t \geq 0]$ be a Markov process. If we define

$$P_{ij}(t,\mathbf{w},\mathbf{z}) = \Pr\{X(t) = j,\, \mathbf{z} \leq \mathbf{Z}(t) \leq \mathbf{z} + \delta\mathbf{z} | X(0) = i,\, \mathbf{w} \leq \mathbf{W}(0) \leq \mathbf{w} + \delta\mathbf{w}\},$$

where $\mathbf{W}(t), \mathbf{Z}(t) \in \Re^n$, then we have

$$P_{ij}(t+s,\mathbf{w},\mathbf{z}) = \sum_k \int_{\mathbf{E}_n} P_{ik}(s,\mathbf{w},\mathbf{E}_n) P_{kj}(t,\mathbf{E}_n,\mathbf{z})\, d\mathbf{E}_n,$$

where $\mathbf{E}_n \in \Re^n$. This leads to the Chapman–Kolmogorov backward difference equation. The Chapman–Kolmogorov forward difference equation is obtained from

$$P_{ij}(t+s,\mathbf{w},\mathbf{z}) = \sum_k \int_{\mathbf{E}_n} P_{ik}(t,\mathbf{w},\mathbf{E}_n) P_{kj}(s,\mathbf{E}_n,\mathbf{z})\, d\mathbf{E}_n.$$

For more details, see Ross [24]. In what follows, we cast the GI/G/1 queueing model as backward and also as forward Chapman–Kolmogorov equations. In the backward system, we use elapsed times as the supplementary variables, and in the forward system, we use remaining times. In order to use elapsed time we need to consider the hazard rate function. This is necessary in order to keep the system Markovian.

3. SPECIAL CASES

3.1. A Queueing Model for the GI/G/1 System: Remaining Time

In this section, a queueing model for the GI/G/1 system is developed by considering remaining times as the supplementary variables.

Consider a single-server queue wherein the interarrival times and service times are independent and identically distributed (i.i.d.) random variables having density functions $a(x)$ $(x \geq 0)$ and $b(x)$ $(x \geq 0)$, respectively, and $A(x) = \int_0^x a(y)\,dy$ and $B(x) = \int_0^x b(y)\,dy$. It is perhaps possible to obtain the results without assuming interarrival and service times to have a density functions (see Gnedenko and Kovalenko [6, pp. 157–160]), but this assumption seems to simplify the argument. Further, we know that any probability distribution is the limit of a sequence of continuous distributions. Thus, our assumption could be eliminated by a continuity argument (see Stoyan [25] and Whitt [29]) and does not seem to infer any loss of generality in the results.

Let $N(t)$ denote the number of units present in the system at time t, $V(t)$ denote the remaining interarrival time at time t (i.e., the time until the next arrival after time t), and $U(t)$ denote the remaining service time for the unit in service at time t. The state of the system at time t is defined by $[N(t), V(t), U(t); t \geq 0]$, which is Markovian in continuous time.

Define

$$P_0(v,t)\Delta v = \text{Prob}\{(N(t)=0),(v < V(t) \leq v + \Delta v)\}, \qquad v \geq 0,$$

$$P_n(v,u,t)\Delta v\Delta u = \text{Prob}\{(N(t)=n),(v < V(t) \leq v + \Delta v)\ \&\ (u < U(t) \leq u + \Delta u)\},$$
$$v \geq 0,\ u \geq 0,\ n \geq 1.$$

From the above equations, it follows that

$$P_0(t) = \text{Prob}(N(t)=0) = \int_0^\infty P_0(v,t)\,dv,$$

$$P_n(t) = \text{Prob}(N(t)=n) = \int_0^\infty \int_0^\infty P_n(v,u,t)\,dv\,du, \qquad n \geq 1.$$

By relating the states of the system at time t and $t + \Delta t$, we obtain the following Chapman–Kolmogorov forward equations:

$$P_0(v - \Delta t, t + \Delta t) = P_0(v,t) + P_1(v,0,t)\Delta t + O(\Delta t), \tag{1}$$

$$P_1(v - \Delta t, u - \Delta t, t + \Delta t) = P_1(v,u,t) + P_0(0,t)b(u)a(v)\Delta t + P_2(v,0,t)b(u)\Delta t + O(\Delta t), \tag{2}$$

$$P_n(v - \Delta t, u - \Delta t, t + \Delta t) = P_n(v,u,t) + P_{n-1}(0,u,t)a(v)\Delta t + P_{n+1}(v,0,t)b(u)\Delta t + O(\Delta t), \qquad n \geq 2. \tag{3}$$

From (1)–(3), we have the following partial differential equations of the system:

$$\left(\frac{\partial}{\partial t} - \frac{\partial}{\partial v}\right) P_0(v,t) = P_1(v,0,t), \tag{4}$$

$$\left(\frac{\partial}{\partial t} - \frac{\partial}{\partial v} - \frac{\partial}{\partial u}\right) P_1(v,u,t) = P_0(0,t)b(u)a(v) + P_2(v,0,t)b(u), \tag{5}$$

$$\left(\frac{\partial}{\partial t} - \frac{\partial}{\partial v} - \frac{\partial}{\partial u}\right) P_n(v,u,t) = P_{n-1}(0,u,t)a(v) + P_{n+1}(v,0,t)b(u), \qquad n \geq 2. \tag{6}$$

3.2. A Queueing Model for the GI/G/1 System: Elapsed Time

In this section, a queueing model for the GI/G/1 system is developed by considering elapsed times as the supplementary variables.

Consider the queueing system as in Section 3.1. Let $N(t)$ denote the number of customers present in the system at time t, $Y(t)$ denote the elapsed interarrival time at time t (i.e., the time since the last arrival at time t), and $X(t)$ denote the elapsed service time for the customer currently in service at time t. The state of the system at time t is defined by $[N(t), Y(t), X(t); t \geq 0]$, which is Markovian in continuous time.

Define

$$P_0(y,t)\Delta y = \text{Prob}\{(N(t)=0),(y < Y(t) \leq y + \Delta y)\}, \qquad y \geq 0,$$

$$P_n(y,x,t)\Delta y \Delta x = \text{Prob}\{(N(t)=n),(y < Y(t) \leq y + \Delta y)\ \&\ (x < X(t) \leq x + \Delta x)\},$$
$$y \geq 0, x \geq 0, n \geq 1.$$

From the above equations, it follows that

$$P_0(t) = \text{Prob}(N(t)=0) = \int_0^\infty P_0(y,t)\,dy,$$

$$P_n(t) = \text{Prob}(N(t)=n) = \int_0^\infty \int_0^\infty P_n(y,x,t)\,dy\,dx, \qquad n \geq 1.$$

Further define the two hazard rate functions

$$a(y) = \lambda(y) \exp\left[-\int_0^y \lambda(y_1)\,dy_1\right], \qquad y \geq 0,$$

where

$$\lambda(y) = \frac{a(y)}{1 - A(y)}, \qquad y \geq 0,$$

and

$$b(x) = \mu(x) \exp\left[-\int_0^x \mu(x_1)\,dx_1\right], \qquad x \geq 0,$$

https://doi.org/10.1017/S0269964800142068 Published online by Cambridge University Press

where

$$\mu(x) = \frac{b(x)}{1 - B(x)}, \qquad x \geq 0.$$

By relating the states of the system at time t and $t + \Delta t$, we obtain the following Chapman–Kolmogorov backward equations:

$$P_0(y + \Delta t, t + \Delta t) = [1 - \lambda(y)\Delta t]P_0(y,t) + \int_0^\infty P_1(y,x,t)\mu(x)\Delta t\, dx + O(\Delta t),$$

$$P_n(y + \Delta t, x + \Delta t, t + \Delta t) = [1 - \lambda(y)\Delta t][1 - \mu(x)\Delta t]P_n(y,x,t) + O(\Delta t), \qquad n \geq 1.$$

The above equations can be written

$$\left(\frac{\partial}{\partial t} + \frac{\partial}{\partial y}\right)P_0(y,t) = -\lambda(y)P_0(y,t) + \int_0^\infty P_1(y,x,t)\mu(x)\, dx, \tag{7}$$

$$\left(\frac{\partial}{\partial t} + \frac{\partial}{\partial y} + \frac{\partial}{\partial x}\right)P_n(y,x,t) = -[\lambda(y) + \mu(x)]P_n(y,x,t), \qquad n \geq 1. \tag{8}$$

The integral boundary conditions are

$$P_n(y,0,t) = \int_0^\infty P_{n+1}(y,x,t)\mu(x)\, dx, \qquad n \geq 1, \tag{9}$$

$$P_0(0,x,t) = 0, \tag{10}$$

$$P_n(0,x,t) = \int_0^\infty P_{n-1}(y,x,t)\lambda(y)\, dy, \qquad n \geq 1. \tag{11}$$

The initial conditions are

$$P_0(y,0) = \delta(y - y_0), \tag{12}$$

$$P_n(y,x,0) = \delta_{ni}\delta(y - y_0, x - x_0), \qquad n \geq 0, \tag{13}$$

where δ_{ni} is the Kronecker delta and $\delta(y - y_0, x - x_0)$ is the Dirac delta function. In words, (13) corresponds to starting the system with a specific queue length i, and the elapsed arrival time and elapsed service time are y_0 and x_0, respectively.

Equations (7)–(11) are also known as partial differential–integral equations of the system.

4. OBSERVATIONS FOR REMAINING TIME AND ELAPSED TIME

In this section, we give the observations for considering the supplementary variable as the remaining time and the elapsed time.

https://doi.org/10.1017/S0269964800142068 Published online by Cambridge University Press

4.1. Remaining Time

1. The supplementary variable technique analysis for the queueing problems by considering supplementary variable as the remaining time involves the following:
   i. probability density function
   ii. partial differential equations.
2. The algorithm for considering the supplementary variable as remaining time works efficiently for any service (arrival) time distribution including both phase as well as nonphase types.
3. The analysis using the remaining time as the supplementary variable provides a simple procedure for deriving relations among state probabilities at various epochs (arbitrary, departure, and prearrival).
4. The only input required for efficient evaluation of state probabilities is the Laplace transform (LT) of the service (arrival) time distribution.
5. The state-dependent service concept can easily be introduced to study the queueing problem by considering remaining service time as the supplementary variable.
6. The more general model (i.e., state-dependent arrival and state-dependent service) can also be studied using remaining service time as the supplementary variable.
7. Features 5 and 6 are more useful from a practical point of view.

4.2. Elapsed Time

1. The supplementary variable technique analysis for the queueing problems by considering elapsed time as the supplementary variable involves the following:
   i. hazard rate function
   ii. partial differential integral equations
   iii. integral boundary conditions
   iv. initial conditions.
2. In the case of elapsed time, the solution procedure is not as simple compared to remaining time.
3. The state-dependent service concept cannot easily be introduced to study the queueing problem by considering the elapsed service time as the supplementary variable. This is because one has to define the hazard rate function in terms of state-dependent service. It should also be pointed out that the definition of hazard rate function involves density function.
4. The more general model (i.e., state-dependent arrival and state-dependent service) is also not possible to study using elapsed service time as the supplementary variable.
5. The elapsed time as the supplementary variable will not give results in a unified way compared to remaining time as the supplementary variable.

6. It is our observation that not much work has been done by considering elapsed interarrival time as the supplementary variable.
7. It should be pointed out that elapsed time is a more practical manner in which to observe a system, even though it is more difficult to model a system in that form.

It should be noted here that from the above observations [cf. Sections 4.1 and 4.2] one can assert that the concept of using remaining time as a supplementary variable gives a very simple and elegant solution; see Gupta and Srinivasa Rao [7], Lee and Ahn [17], and Niu and Takahashi [23].

In the study of the GI/G/1 system, we have two supplementary variables. In the next section, we show that our observation regarding remaining time and elapsed time still hold even when there is only one supplementary variable involved.

5. SPECIAL CASES FOR REMAINING TIME

Case 1: When we consider the interarrival time distribution is exponential, the resulting system is the well-known M/G/1 queue. We let the arrival rate be λ.

Let $N(t)$ denote the number of units present in the system at time t and let $U(t)$ denote the remaining service of the unit in service at time t. The state of the system at time t is defined by $[N(t), U(t); t \geq 0]$, which is Markovian in continuous time. Further, let $P_0(t)$ be defined as the probability that at time t, the system is empty (i.e., idle), and let $P_n(u,t)\Delta u$ $(n \geq 1)$ be defined as the joint probability that at time t, the number of units in the system is n and the remaining service time of the unit in service lies in the interval $(u, u+\Delta u)$; that is,

$$P_0(t) = \text{Prob}(N(t) = 0),$$

$$P_n(u,t)\Delta u = \text{Prob}\{N(t) = n, u < U(t) \leq u + \Delta u\}, \qquad u \geq 0, n \geq 1.$$

From the above equations, it follows that

$$P_n(t) = \text{Prob}(N(t) = n) = \int_0^\infty P_n(u,t)\,du, \qquad n \geq 1.$$

Relating the states of the system at time t and $t + \Delta t$, we obtain the following partial differential equations:

$$\frac{d}{dt} P_0(t) = -\lambda P_0(t) + P_1(0,t), \tag{14}$$

$$\left(\frac{\partial}{\partial t} - \frac{\partial}{\partial u}\right) P_1(u,t) = -\lambda P_1(u,t) + \lambda P_0(t) b(u) + P_2(0,t) b(u), \tag{15}$$

$$\left(\frac{\partial}{\partial t} - \frac{\partial}{\partial u}\right) P_n(u,t) = -\lambda P_n(u,t) + \lambda P_{n-1}(u,t) + P_{n+1}(0,t) b(u), \qquad n \geq 2. \tag{16}$$

Case 2: When we consider the service time distribution to be exponential, the resulting model is the well-known GI/M/1 queue. We let the service rate be μ.

Let $N(t)$ denote the number of units present in the system at time t and let $V(t)$ denote the remaining interarrival time at time t. The state of the system at time t is defined by $[N(t), V(t); t \geq 0]$, which is Markovian in continuous time. Further, let $P_n(v,t)\Delta v$ $(n \geq 0)$ be defined as the joint probability that at time t, the number in the system is n and the time until the next unit arrival lies in the interval $(v, v+\Delta v)$; that is,

$$P_n(v,t)\Delta v = \text{Prob}\{N(t)=n, v < V(t) \leq v + \Delta v\}, \qquad v \geq 0, n \geq 0.$$

From the above equations, it follows that

$$P_n(t) = \text{Prob}(N(t)=n) = \int_0^\infty P_n(v,t)\,dv, \qquad n \geq 0.$$

Relating the states of the system at time t and $t+\Delta t$, we obtain the following partial differential equations:

$$\left(\frac{\partial}{\partial t} - \frac{\partial}{\partial v}\right) P_0(v,t) = \mu P_1(v,t), \tag{17}$$

$$\left(\frac{\partial}{\partial t} - \frac{\partial}{\partial v}\right) P_n(v,t) = -\mu P_n(v,t) + P_{n-1}(0,t)a(v) + \mu P_{n+1}(v,t), \qquad n \geq 1. \tag{18}$$

6. SPECIAL CASES FOR ELAPSED TIME

Case 1: When we consider the interarrival time distribution to be exponential, the resulting queueing system is the well-known M/G/1 queue. We let the arrival rate be λ.

Let $N(t)$ denote the number of units present in the system at time t and let $X(t)$ denote the elapsed service of the unit currently in service at time t. The state of the system at time t is defined by $[N(t), X(t); t \geq 0]$, which is Markovian in continuous time. Further, let $P_0(t)$ be defined as the probability that at time t, the system is empty (i.e., idle) and let $P_n(x,t)\Delta x$ $(n \geq 1)$ be defined as the joint probability that at time t, the number in the system is n and the elapsed service time of the unit currently in service lies in the interval $(x, x+\Delta x)$; that is,

$$P_0(t) = \text{Prob}(N(t)=0),$$

$$P_n(x,t)\Delta x = \text{Prob}\{N(t)=n, x < X(t) \leq x + \Delta x\}, \qquad x \geq 0, n \geq 1.$$

From the above equations, it follows that

$$P_n(t) = \text{Prob}(N(t)=n) = \int_0^\infty P_n(x,t)\,dx, \qquad n \geq 1.$$

Relating the states of the system at time t and $t + \Delta t$, we obtain the following partial differential–integral equations:

$$\frac{d}{dt} P_0(t) = -\lambda P_0(t) + \int_0^\infty P_1(x,t)\mu(x)\,dx, \tag{19}$$

$$\left(\frac{\partial}{\partial t} + \frac{\partial}{\partial x}\right) P_1(x,t) = -[\lambda + \mu(x)]P_1(x,t), \tag{20}$$

$$\left(\frac{\partial}{\partial t} + \frac{\partial}{\partial x}\right) P_n(x,t) = -[\lambda + \mu(x)]P_n(x,t) + \lambda P_{n-1}(x,t), \qquad n \geq 2. \tag{21}$$

The boundary conditions are given by

$$P_1(0,t) = \lambda P_0(t) + \int_0^\infty P_2(x,t)\mu(x)\,dx, \tag{22}$$

$$P_n(0,t) = \int_0^\infty P_{n+1}(x,t)\mu(x)\,dx, \qquad n \geq 2. \tag{23}$$

The initial conditions are given by

$$P_n(x,0) = \delta_{ni}\delta(x - x_0), \qquad n \geq 0, \tag{24}$$

where δ_{ni} is the Kronecker delta and $\delta(x - x_0)$ is the Dirac delta function. In words, (24) corresponds to starting the system with specific queue length i, and the elapsed service time is x_0.

Case 2: When we consider the service time distribution to be exponential, the resulting queueing system is the well-known GI/M/1 queue. We let the service rate be μ.

Let $N(t)$ denote the number of units present in the system at time t and let $Y(t)$ denote the elapsed interarrival time at time t. The state of the system at time t is defined by $[N(t), Y(t); t \geq 0]$, which is Markovian in continuous time. Further, let $P_n(y,t)\Delta y$ $(n \geq 0)$ be defined as the joint probability that at time t, the number of units in the system is n and the time since the last unit arrival lies in the interval $(y, y + \Delta y)$; that is,

$$P_n(y,t)\Delta y = \text{Prob}\{N(t) = n, y < Y(t) \leq y + \Delta y\}, \qquad y \geq 0, n \geq 0.$$

From the above equations, we have

$$P_n(t) = \text{Prob}(N(t) = n) = \int_0^\infty P_n(y,t)\,dy, \qquad n \geq 0.$$

Relating the states of the system at time t and $t + \Delta t$, we have the following partial differential–integral equations:

$$\left(\frac{\partial}{\partial t} + \frac{\partial}{\partial y}\right) P_0(y,t) = -\lambda(y) P_0(y,t) + \mu P_1(y,t), \tag{25}$$

$$\left(\frac{\partial}{\partial t} + \frac{\partial}{\partial y}\right) P_n(y,t) = -[\lambda(y) + \mu]\, P_n(y,t) + \mu P_{n+1}(y,t), \qquad n \geq 1. \tag{26}$$

The boundary conditions are given by

$$P_0(0,t) = 0, \tag{27}$$

$$P_n(0,t) = \int_0^\infty P_{n-1}(y,t)\lambda(y)\, dy, \qquad n \geq 1. \tag{28}$$

The initial conditions are given by

$$P_n(y,0) = \delta_{ni}\,\delta(y - y_0), \qquad n \geq 0, \tag{29}$$

where δ_{ni} is the Kronecker delta and $\delta(y - y_0)$ is the Dirac delta function.

7. SPECIAL CASES FOR FINITE BUFFER QUEUE

In this section, we consider the GI/G/1/N system and show that our observations regarding the remaining time and elapsed time is even more obvious when we have two reflecting boundaries 0 and N.

7.1. Remaining Time

In this section, a finite buffer GI/G/1 model is developed by considering supplementary variables as the remaining times. We assume that the number of waiting places is $N - 1$ (i.e., the maximum number of units allowed in the system is N).

Define

$$P_0(v,t)\Delta v = \text{Prob}\{(N(t) = 0),(v < V(t) \leq v + \Delta v)\}, \qquad v \geq 0,$$

$$P_n(v,u,t)\Delta v \Delta u = \text{Prob}\{(N(t) = n),(v < V(t) \leq v + \Delta v)\ \&\ (u < U(t) \leq u + \Delta u)\},$$
$$v \geq 0,\ u \geq 0,\ 1 \leq n \leq N.$$

From the above equations, it follows that

$$P_0(t) = \text{Prob}(N(t) = 0) = \int_0^\infty P_0(v,t)\, dv,$$

$$P_n(t) = \text{Prob}(N(t) = n) = \int_0^\infty \int_0^\infty P_n(v,u,t)\, dv\, du, \qquad 1 \leq n \leq N.$$

https://doi.org/10.1017/S0269964800142068 Published online by Cambridge University Press

By relating the states of the system at time t and $t + \Delta t$, we obtain the following partial differential equations of the system:

$$\left(\frac{\partial}{\partial t} - \frac{\partial}{\partial v}\right) P_0(v, t) = P_1(v, 0, t), \tag{30}$$

$$\left(\frac{\partial}{\partial t} - \frac{\partial}{\partial v} - \frac{\partial}{\partial u}\right) P_1(v, u, t) = P_0(0, t) b(u) a(v) + P_2(v, 0, t) b(u), \tag{31}$$

$$\left(\frac{\partial}{\partial t} - \frac{\partial}{\partial v} - \frac{\partial}{\partial u}\right) P_n(v, u, t) = P_{n-1}(0, u, t) a(v) + P_{n+1}(v, 0, t) b(u), \quad 2 \le n \le N - 1 \tag{32}$$

$$\left(\frac{\partial}{\partial t} - \frac{\partial}{\partial v} - \frac{\partial}{\partial u}\right) P_N(v, u, t) = P_{N-1}(0, u, t) a(v) + P_N(0, u, t) a(v). \tag{33}$$

7.2. Elapsed Time

In this section, a finite buffer GI/G/1 model has been developed by considering the supplementary variable as the elapsed time. If $a(y)$ and $b(x)$, and $\lambda(y)$ and $\mu(x)$ are the density functions and hazard rate functions for the interarrival and service time distributions, respectively, the system capacity is N (i.e., the number of waiting places is $N - 1$).

Define

$$P_0(y, t)\Delta y = \text{Prob}\{(N(t) = 0), (y < Y(t) \le y + \Delta y)\}, \qquad y \ge 0,$$

$$P_n(y, x, t)\Delta y \Delta x = \text{Prob}\{(N(t) = n), (y < Y(t) \le y + \Delta y) \;\&\; (x < X(t) \le x + \Delta x)\},$$
$$y \ge 0,\ x \ge 0,\ 1 \le n \le N.$$

From the above equations, it follows that

$$P_0(t) = \text{Prob}(N(t) = 0) = \int_0^\infty P_0(y, t)\, dy,$$

$$P_n(t) = \text{Prob}(N(t) = n) = \int_0^\infty \int_0^\infty P_n(y, x, t)\, dy\, dx, \qquad 1 \le n \le N.$$

By relating the states of the system at time t and $t + \Delta t$, we obtain the following partial differential–integral equations:

$$\left(\frac{\partial}{\partial t} + \frac{\partial}{\partial y}\right) P_0(y, t) = -\lambda(y) P_0(y, t) + \int_0^\infty P_1(y, x, t) \mu(x)\, dx, \tag{34}$$

$$\left(\frac{\partial}{\partial t} + \frac{\partial}{\partial y} + \frac{\partial}{\partial x}\right) P_n(y, x, t) = -[\lambda(y) + \mu(x)] P_n(x, t), \qquad 1 \le n \le N. \tag{35}$$

https://doi.org/10.1017/S0269964800142068 Published online by Cambridge University Press

The boundary conditions are

$$P_n(y,0,t) = \int_0^\infty P_{n+1}(y,x,t)\mu(x)\,dx, \qquad 1 \le n \le N-1, \tag{36}$$

$$P_N(y,0,t) = 0, \tag{37}$$

$$P_0(0,x,t) = 0, \tag{38}$$

$$P_n(0,x,t) = \int_0^\infty P_{n-1}(y,x,t)\lambda(y)\,dy, \qquad 1 \le n \le N-1, \tag{39}$$

$$P_N(0,x,t) = \int_0^\infty P_{N-1}(y,x,t)\lambda(y)\,dy + \int_0^\infty P_N(y,x,t)\lambda(y)\,dy. \tag{40}$$

The initial conditions are

$$P_0(y,0) = \delta(y - y_0), \tag{41}$$

$$P_n(y,x,0) = \delta_{ni}\delta(y - y_0, x - x_0), \qquad 0 \le n \le N, \tag{42}$$

where δ_{ni} is the Kronecker delta and $\delta(y - y_0, x - x_0)$ is the Dirac delta function. In words, (42) corresponds to starting the system with a specific queue length i, and the elapsed arrival time and elapsed service time are y_0 and x_0, respectively.

8. CONCLUSIONS

In this article, we have tried to give an overview on supplementary variable technique (SVT). We have also demonstrated the importance of SVT by considering remaining time as the supplementary variable. It is hoped that the basic idea of writing this article on SVT in stochastic models will help the researchers, practitioners, and engineers apply this technique in various systems.

In this article, we have pointed out that the use of remaining time as the supplementary variable is much easier to work with. This technique will give results in a unified way (e.g., one can obtain the distributions of units in the system at departure, prearrival, and arbitrary epochs). Further, it is easier to develop relations among state probabilities at various epochs where, otherwise, one has to use level crossing analysis (cf. Jeyachandra and Shanthikumar [12]) to derive these relations.

Acknowledgments

This research has been supported in part by Grant OGP0006584 from the Natural Sciences and Engineering Research Council of Canada (to A.S. Alfa).

References

1. Alfa, A.S. (1982). Time-inhomogeneous bulk server queue in discrete time: A transportation type problem. *Operations Research* 30: 650–658.
2. Cohen, J.W. (1982). *The single server queue*, 2nd ed. Amsterdam: North-Holland.
3. Cooper, R.B. (1990). *Introduction to queueing theory*, 3rd ed. New York: North-Holland.

4. Cox, D.R. (1955). A use of complex probabilities in the theory of stochastic processes. *Proceedings of the Cambridge Philosophical Society* 51: 313–319.
5. Dafermos, S.C. & Neuts, M.F. (1971). A single server queue in discrete time. *Cahiers du Centre de Recherche Opérationnelle* 13: 23–40.
6. Gnedenko, B.V. & Kovalenko, J.N. (1968). *Introduction to queueing theory.* Jerusalem: Israel Program for Scientific Translation.
7. Gupta, U.C. & Srinivasa Rao, T.S.S. (1998). On the analysis of single server finite queue with state dependent arrival and service processes: $M(n)/G(n)/1/K$. *Operations Research Spektrum* 20: 83–89.
8. Heimann, D. & Neuts, M.F. (1973). The single server queue in discrete time—Numerical analysis IV. *Naval Research Logistics Quarterly* 20: 753–766.
9. Henderson, W. (1972). Alternative approaches to the analysis of the $M/G/1$ and $GI/M/1$ queues. *Journal of the Operations Research Society of Japan* 12: 92–101.
10. Hokstad, P. (1975). A supplementary variable technique applied to the $M/G/1$ queue. *Scandinavian Journal of Statistics* 2: 95–98.
11. Jaiswal, N.K. (1968). *Priority queues.* New York: Academic Press.
12. Jeyachandra, M.J. & Shanthikumar, J.G. (1983). On a machine interference problem with several types of machines attended by a single repairman. *International Journal of Production Research* 7: 529–541.
13. Keilson, J. & Kooharian, A. (1960). On the time dependent queueing process. *Annals of Mathematical Statistics* 31: 104–112.
14. Klimko, F. & Neuts, M.F. (1973). The single server queue in discrete time—Numerical analysis II. *Naval Research Logistics Quarterly* 20: 305–319.
15. Knessl, C., Matkowsky, B.J., Schuss, Z., & Tier, C. (1990). An integral equation approach to the M/G/2 queue. *Operations Research* 38: 506–518.
16. Kosten, L. (1973). *Stochastic theory of service systems.* Oxford: Pergamon Press.
17. Lee, H.W. & Ahn, B.Y. (1998). A queueing model for optimal control of partial buffer sharing in ATM. *Computers and Operations Research* 25: 113–126.
18. Lee, T.T. (1984). $M/G/1/N$ queue with vacation time and exhaustive service discipline. *Operations Research* 32: 774–784.
19. Li, W., Braun, J., & Zhao, Y.-Q. (1998). Stochastic scheduling on a repairable machine with Erlang uptime. *Advances in Applied Probability* 30: 1073–1088.
20. Li, W. & Cao, J. (1993). Limiting distribution of residual lifetimes of Markov repairable system. *Microelectronics & Reliability* 33: 1069–1072.
21. Li, W., Shi, D.-H., & Chao, X. (1997). Reliability analysis of $M/G/1$ queueing systems with server breakdowns and vacations. *Journal of Applied Probability* 34: 546–555.
22. Minh, D.L. (1978). The discrete-time, single-server queue with time-inhomogeneous compound Poisson input and general service time distribution. *Journal of Applied Probability* 15: 590–601.
23. Niu, Z. & Takahashi, Y. (1999). A finite-capacity queue with exhaustive vacation/close-down/setup times and Markovian arrival processes. *Queueing Systems* 31: 1–23.
24. Ross, S.M. (1983). *Stochastic processes.* New York: Wiley.
25. Stoyan, D. (1973). A continuity theorem for queue size. *Bulletin of Academic Polond Science* 21: 1143–1146.
26. Takagi, H. (1991). *Queueing analysis: A foundation of performance evaluation.* Vol. 1: *Vacation and priority systems.* New York: North-Holland.
27. Takagi, H. (1993). *Queueing analysis: A foundation of performance evaluation.* Vol. 2: *Finite systems.* New York: North-Holland.
28. Takagi, H. (1993). *Queueing analysis: A foundation of performance evaluation.* Vol. 3: *Discrete-time systems.* New York: North-Holland.
29. Whitt, W. (1974). The continuity of queues. *Advances in Applied Probability* 6: 175–183.
30. Yang, T. & Chaudhry, M.L. (1996). On steady-state queue size distributions of the discrete-time $GI/G/1$ queue. *Advances in Applied Probability* 28: 1177–1200.

APPENDIX

Steady State Analysis for the GI/G/1 Queueing System Using Remaining Times as the Supplementary Variables

Let us discuss the system of equations in steady state; we let the limit of t in (4)–(6) approach infinity and define

$$P_0(v) = \lim_{t\to\infty} P_0(v,t),$$

$$P_n(v,u) = \lim_{t\to\infty} P_n(v,u,t), \qquad n \geq 1.$$

The steady state equations are

$$-\frac{\partial}{\partial v} P_0(v) = P_1(v,0), \tag{43}$$

$$\left(-\frac{\partial}{\partial v} - \frac{\partial}{\partial u}\right) P_1(v,u) = P_0(0)b(u)a(v) + P_2(v,0)b(u), \tag{44}$$

$$\left(-\frac{\partial}{\partial v} - \frac{\partial}{\partial u}\right) P_n(v,u) = P_{n-1}(0,u)a(v) + P_{n+1}(v,0)b(u), \qquad n \geq 2. \tag{45}$$

Further, define

$$P(z,v,u) = \sum_{n=1}^{\infty} P_n(v,u)z^n \qquad \text{mod } z \leq 1,\ v \geq 0,\ u \geq 0,$$

$$P(z,0,u) = \sum_{n=1}^{\infty} P_n(0,u)z^n \qquad \text{mod } z \leq 1,\ u \geq 0,$$

$$P(z,v,0) = \sum_{n=1}^{\infty} P_n(v,0)z^n \qquad \text{mod } z \leq 1,\ v \geq 0,$$

$$P^*(z,r,s) = \int_0^\infty \int_0^\infty e^{-rv}e^{-su}P(z,v,u)\,dv\,du,$$

$$Q^*(z,s) = \int_0^\infty e^{-su}P(z,0,u)\,du,$$

$$R^*(z,r) = \int_0^\infty e^{-rv}P(z,v,0)\,dv.$$

Multiplying (44) by z and (45) by z^n ($n \geq 2$) and then adding the terms, and also using (43), we get

$$-\frac{\partial}{\partial v} P(z,v,u) - \frac{\partial}{\partial u} P(z,v,u)$$
$$= zP_0(0)a(v)b(u) + za(v)P(z,0,u) + \frac{1}{z} b(u)P(z,v,0) + b(u)\frac{\partial}{\partial u} P_0(v). \tag{46}$$

https://doi.org/10.1017/S0269964800142068 Published online by Cambridge University Press

Again multiplying (46) by e^{-rv} and e^{-su} and then double integrating, we get

$$(r+s)P^*(z,r,s) = P_0(0)B^*(s)[1-zA^*(r)] + Q^*(z,s)[1-zA^*(r)] + R^*(z,r)\left[1-\frac{1}{z}B^*(s)\right] - rP_0^*(r)B^*(s). \quad (47)$$

The unknown quantities in (47) are $P^*(z,r,s)$, $Q^*(z,s)$, $R^*(z,r)$, and $P_0^*(r)$. Letting $r \to 0$ and $s \to 0$ in (47), we have

$$sP^*(z,0,s) = P_0(0)B^*(s)[1-z] + Q^*(z,s)[1-z] + R^*(z,0)\left[1-\frac{1}{z}B^*(s)\right], \quad (48)$$

$$rP^*(z,r,0) = P_0(0)[1-zA^*(r)] + Q^*(z,0)[1-zA^*(r)] + R^*(z,r)\left[1-\frac{1}{z}\right] - rP_0^*(r). \quad (49)$$

Differentiating (47) once w.r.t. r and letting $r,s \to 0$, we obtain

$$P^*(z,0,0) = a_1 z[P_0(0) + Q^*(z,0)] + R^{*(1)}(z,0)\left[1-\frac{1}{z}\right] - P_0^*(0), \quad (50)$$

where $a_1 = -A^{*(1)}(0)$ is the mean interarrival time.

Again differentiating (47) once w.r.t. s and letting $r,s \to 0$, we get

$$P^*(z,0,0) = b_1\left[\frac{1}{z}R^*(z,0) - P_0(0)(1-z)\right] + Q^{*(1)}(z,0)(1-z), \quad (51)$$

where $b_1 = -B^{*(1)}(0)$ is the mean service time. Letting $z \to 1$ in (50) and (51) we obtain

$$P^*(1,0,0) = a_1[P_0(0) + Q^*(1,0)] - P_0^*(0) \quad (52)$$

and

$$P^*(1,0,0) = b_1 R^*(0), \quad (53)$$

respectively. From (52) and (53), we get

$$P_0(0) = \frac{1}{a_1}[P_0^*(0) + b_1 R^*(1,0) - a_1 Q^*(1,0)]. \quad (54)$$

https://doi.org/10.1017/S0269964800142068 Published online by Cambridge University Press