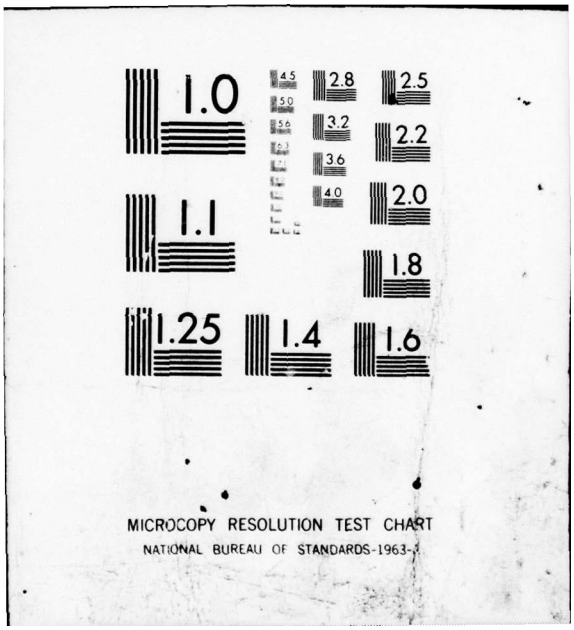
AD-A072 338 NAVAL RESEARCH LAB WASHINGTON DC F/G 12/1
INFORMATION THEORETIC APPROXIMATIONS FOR M/G/1 AND G/G/1 QUEUIN--ETC(U)
JUL 79 J E SHORE

UNCLASSIFIED NRL-MR-4047

| OF |
AD
A072338

END
DATE
FILMED
9-79
DDC


1.0
1.1
1.25
1.4
1.6
1.8
2.0
2.2
2.5
2.8
3.2
3.6
4.0
4.5
5.0
5.6
MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS-1963-A

NRL Memorandum Report 4047

Information Theoretic Approximations for M/G/1 and G/G/1 Queuing Systems

JOHN E. SHORE

Information Systems Staff
Communications Sciences Division

July 18, 1979

NAVAL RESEARCH LABORATORY
Washington, D.C.

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE

READ INSTRUCTIONS BEFORE COMPLETING FORM

1. REPORT NUMBER: NRL Memorandum Report 4047
2. GOVT ACCESSION NO.:
3. RECIPIENT'S CATALOG NUMBER:

4. TITLE (and Subtitle): INFORMATION THEORETIC APPROXIMATIONS FOR M/G/1 AND G/G/1 QUEUING SYSTEMS
5. TYPE OF REPORT & PERIOD COVERED: Final Report
6. PERFORMING ORG. REPORT NUMBER:

7. AUTHOR(s): John E. Shore
8. CONTRACT OR GRANT NUMBER(s):

9. PERFORMING ORGANIZATION NAME AND ADDRESS: Naval Research Laboratory, Washington, DC 20375
10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS: NRL Problem B02-35; 61153N, RR014-09-41

11. CONTROLLING OFFICE NAME AND ADDRESS: 18 Jul 79
12. REPORT DATE: July 18, 1979
13. NUMBER OF PAGES: 39

14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office): 40 pp
15. SECURITY CLASS. (of this report): UNCLASSIFIED
15a. DECLASSIFICATION/DOWNGRADING SCHEDULE:

16. DISTRIBUTION STATEMENT (of this Report)

Approved for public release; distribution unlimited. NRL-MR-4047

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

RR01409 RR014094 1

18. SUPPLEMENTARY NOTES

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

Queue approximations
Maximum entropy
Information theory
M/G/1 systems
G/G/1 systems

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

The behavior of single server queuing systems is characterized by various "performance distributions", including distributions of queue length, waiting time, residence time, busy period, number served in a busy period, etc. In principle, if the arrival and service time distributions are known exactly, then these performance distributions can be computed using standard techniques. We consider the problem of estimating the distributions when only the first few moments of the service time distribution are known. Our approach uses standard relations to compute moments
→ next page (Continues)

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE
S/N 0102-014-6601

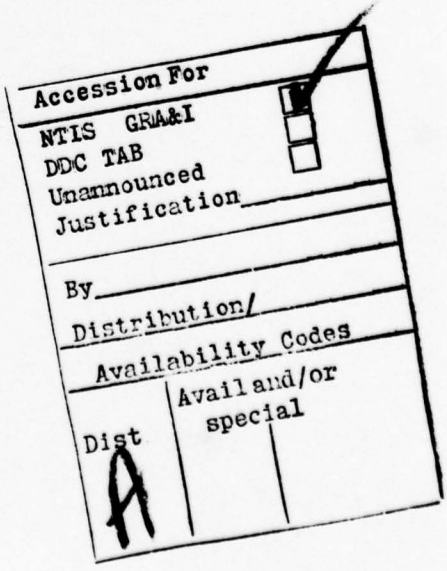
SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

251 950

20. Abstract (Continued)

of the performance distributions from the known moments of the service time distribution and the principles of maximum entropy and minimum cross-entropy to estimate the performance distributions themselves. For M/G/1 systems with known average arrival rates, we derive analytic results for cases when one or two moments of the service time distribution are known, and we show how one can compute results using as many moments of the service time distribution as are available. For G/G/1 systems, our results are limited to the case in which only the average arrival and service rates are known. Among the results obtained with this information theoretic approach is a new light-load approximation for the M/M/1 busy period probability density. Throughout the paper, we illustrate our approach using M/M/1, M/H_2/1, and M/D/1 examples.

CONTENTS



I. INTRODUCTION

For single server queuing systems, we consider the problem of estimating various interesting probability densities and distributions when only partial knowledge of the service time distribution is available. In particular, we consider M/G/1 systems: customers arrive with independent, exponentially distributed interarrival times from an infinite customer pool, wait in an infinite capacity queue, are served independently by a single server with a general service time distribution, and return to the customer pool. The performance of such systems depends on the details of the service time distribution and is characterized by various interesting probability distributions and densities, including queue length, busy period length, number served during a busy period, waiting time, etc. We refer to these as the "performance distributions." In principle, given the service time probability density s(t), one can compute the performance distributions using standard techniques [1]-[3]. But suppose, instead of s(t), one knows only its first n moments

$$s_m = \int dt\ t^m s(t) \quad . \tag{1}$$

What is the best way to use this information in estimating the performance distributions?

Our approach exploits the fact that moments of the performance distributions are themselves determined by the service time moments (1) and the average arrival time (a sufficient statistic of the exponential interarrival time density). For example, the first n moments of the distribution of the number of customers served during a busy period can be expressed in terms of the the average arrival time and the first n moments of

Note: Manuscript submitted May 29, 1979.

s(t). Thus, the information (1) giving moments of s(t) is equivalent to information giving moments of the performance distributions. Given this information, we use the principle of maximum entropy to estimate the performance distributions themselves. Because entropy maximization has been shown to be a uniquely correct, self-consistent method of inference about probability distributions when new information is in the form of expected values [4], [5], we refer to the resulting estimates of the performance distributions as information theoretic approximations.

The remainder of this paper is organized as follows: Section II summarizes the principles of maximum entropy and minimum cross-entropy (a generalization), and discusses informally the sense in which these principles provide correct, general methods of inductive inference. Information theoretic approximations for M/G/1 queue length, number served in busy period, busy period length, residence time, and waiting time are discussed, with examples, in Sections III-VI. In these applications we assume uniform distributions for estimates of the performance distributions available prior to learning the service time moments (1). Additional applications involving the use of non-uniform prior estimates of the performance distributions are suggested in Section VII. Some results for G/G/1 systems are derived in Section VIII. Discussion follows in Section IX.

II. ENTROPY MAXIMIZATION AND CROSS-ENTROPY MINIMIZATION

A. The Maximum Entropy Principle and the Minimum Cross-entropy Principle

Suppose you know that a system has a set of possible states x_i with unknown probabilities $q^\dagger(x_i)$, and you then learn constraints on the

distribution $q^\dagger$: either values of certain expectations $\sum_i q(x_i) f_k(x_i)$ or bounds on these values. Suppose you need to choose a distribution q that is in some sense the best estimate of $q^\dagger$ given what you know. Usually, there remains an infinite set of distributions that are not ruled out by the constraints. Which one should you choose?

The principle of maximum entropy states that, of all the distributions q that satisfy the constraints, you should choose the one with the largest entropy $-\sum_i q(x_i)\log(q(x_i))$. Entropy maximization was first proposed as a general inference procedure by Jaynes [6]. Since then, it has been applied successfully in a remarkable variety of fields, including traffic networks [7], and queuing theory [8]. For a lengthy list of applications and references, see [5].

The principle of minimum cross-entropy is a generalization that applies in cases when a prior distribution p that estimates $q^\dagger$ is known in addition to the constraints. The principle states that, of the distributions q that satisfy the constraints, you should choose the one with the least cross-entropy $\sum_i q(x_i)\log(q(x_i)/p(x_i))$. Minimizing cross-entropy is equivalent to maximizing entropy when the prior is a uniform distribution. Unlike entropy maximization, cross-entropy minimization generalizes correctly for continuous probability densities. One then minimizes the functional

$$\int dx\, q(x)\log(q(x)/p(x)) \quad . \tag{2}$$

The name cross-entropy is due to Good [9]. Other names include expected weight of evidence [10, p. 72], directed divergence [11, p. 7], and discrimination information [11, p. 37]. First proposed by Kullback [11, p. 37], the principle of minimum cross-entropy has been advocated in

various forms by others [9], [12], [13], including Jaynes [14], who obtained (2) with an "invariant measure" playing the role of the prior density. Like entropy maximization, cross-entropy minimization has been applied in many fields (see [5]).

B. Justifying the Principles as General Methods of Inference

Until recently, entropy maximization was justified best on the basis of entropy's unique properties as an uncertainty measure. That entropy has such properties is undisputed: one can prove, up to a constant factor, that entropy is the only function satisfying axioms that are accepted as requirements for an uncertainty measure [15]. Intuitively, the maximum entropy principle follows quite naturally from such axiomatic characterizations. Jaynes states that the maximum entropy distribution "is uniquely determined as the one which is maximally noncommittal with regard to missing information" [6, p. 623], and that it "agrees with what is known, but expresses 'maximum uncertainty' with respect to all other matters, and thus leaves a maximum possible freedom for our final decisions to be influenced by the subsequent sample data" [14, p. 231]. Somewhat whimsically, Benes justified his use of entropy maximization as "a reasonable and systematic way of throwing up our hands" [7, p. 234].

Similar justifications can be advanced for cross-entropy minimization. Like entropy, cross-entropy has various properties that are desirable for an information measure [12],[13], and one can argue [16] that cross-entropy measures the amount of information necessary to change a prior p into a posterior q. The principle of minimum cross-entropy then follows intuitively much like entropy maximization.

To some, entropy's unique properties make it obvious that entropy maximization is the correct way to account for constraint information. To

others, such an informal and intuitive justification yields plausibility but not proof --- why maximize entropy; why not some other function? As a result, entropy maximization has remained controversial despite its success.

Recently, we have obtained a new, formal justification for entropy maximization using a different approach [5]. This approach is based on the observation that previous justifications are weak, not only because they rely on informal, intuitive arguments, but also because they are indirect --- they are based on a formal description of what is required of an information measure rather than on a formal description of what is required of a method for taking new information into account.

Our approach in [5] was to formalize the requirements of inductive inference directly in terms of a set of consistency axioms that make no reference to information measures or properties of information measures. All of the axioms are based on a single fundamental principle: If a problem can be solved in more than one way, the results should be consistent. Informally, the axioms may be phrased as follows:

1) Uniqueness. The results of taking new information into account should be unique.
2) Invariance. It shouldn't matter in which coordinate system we account for new information.
3) System independence. It shouldn't matter whether we account for independent information about independent systems separately in terms of different probability densities or together in terms of a joint density.
4) Subset Independence. It shouldn't matter whether we account for information about an independent subset of system states in terms of a separate conditional density or in terms of the full system density.

We were then able to prove that the principle of maximum entropy is correct in the following sense: Given information in the form of constraints on expected values, there is only one distribution satisfying these constraints that can be chosen in a manner that satisfies the axioms; this unique distribution can be obtained by maximizing entropy. This result for entropy maximization was obtained both directly and as a special case (uniform priors) of an analogous, more general result for the principle of minimum cross-entropy.

C. Mathematics of Entropy Maximization and Cross-entropy Minimization

We treat entropy maximization as a special case of cross-entropy minimization. Let $\underset{\sim}{x}$ denote a single state of some system that has a set $\underset{\sim}{D}$ of possible system states and a probability density $q^{\dagger}(\underset{\sim}{x})$ of states. We assume that the existence of $q^{\dagger}$ is known but that $q^{\dagger}$ itself is unknown. Given $p(\underset{\sim}{x})$, a prior density that is our current estimate of $q^{\dagger}$, we wish to select a posterior $q(\underset{\sim}{x})$ based on new information that $q^{\dagger}$ (and therefore q) must satisfy the expected value constraints

$$\int_{\underset{\sim}{D}} d\underset{\sim}{x}\; q(\underset{\sim}{x}) g_r(\underset{\sim}{x}) \equiv \langle g_r \rangle = \overline{g}_r \;, \tag{3}$$

for a known set of bounded functions $g_r(\underset{\sim}{x})$ and numbers $\overline{g}_r$, $r = 1,\ldots,n$. The solution to this inference problem is obtained by varying $q(\underset{\sim}{x})$ so that the cross-entropy

$$H(q,p) = \int_{\underset{\sim}{D}} d\underset{\sim}{x}\; q(\underset{\sim}{x}) \log(q(\underset{\sim}{x})/p(\underset{\sim}{x})) \tag{4}$$

is minimized subject to the constraints (3) and the normalization constraint

$$\int_{\underset{\sim}{D}} d\underset{\sim}{x}\; q(\underset{\sim}{x}) = 1 \quad . \tag{5}$$

Mathematically, the solution is obtained using the method of Lagrangian multipliers and standard techniques from the calculus of variations. The minimization condition is

$$\log(q(\underset{\sim}{x})/p(\underset{\sim}{x})) + 1 + \alpha_0 + \sum_r \beta_r g_r(\underset{\sim}{x}) = 0 , \tag{6}$$

where the β_r and α_0 are Lagrangian multipliers corresponding to the constraints (3) and (5). The solution of (6) is

$$q(\underset{\sim}{x}) = p(\underset{\sim}{x})\exp(-\alpha - \sum_r \beta_r g_r(\underset{\sim}{x})), \tag{7}$$

where $\alpha = \alpha_0 + 1$. It is convenient to write (7) in the form

$$q(\underset{\sim}{x}) = Z^{-1}p(\underset{\sim}{x})\exp(-\sum_r \beta_r g_r(\underset{\sim}{x})) , \tag{8}$$

where Z is the "partition function",

$$Z = \exp(\alpha) = \int_{\underset{\sim}{D}} d\underset{\sim}{x}\; p(\underset{\sim}{x})\exp(-\sum_r \beta_r g_r(\underset{\sim}{x})) . \tag{9}$$

The values of the multipliers β_r are determined by the known expectation values $\bar{g}_r$ in (3). One can express the posterior q directly in terms of the values $\bar{g}_r$ by solving the equations

$$\bar{g}_r = -Z^{-1}\frac{\partial Z}{\partial \beta_r}$$

$$= -\frac{\partial}{\partial \beta_r}\log(Z) \tag{10}$$

for the β_r, or by substituting (8) into the constraint equations (3) and solving for the β_r. Such solutions are often difficult or impossible to obtain analytically, but one can obtain them computationally in general [4, Appendix B], [17].

The principle of minimum cross-entropy also applies when, in addition to equality constraints (3), we gain new information in the form of a bound on an expected value,

$$\int_{\underset{\sim}{D}} d\underset{\sim}{x}\, q(\underset{\sim}{x})g(\underset{\sim}{x}) \equiv \langle g \rangle \geq \overline{g} \quad . \tag{11}$$

Such an inequality constraint is handled as follows: First one solves for the minimum cross-entropy density given only the equality constraints (3). If the resulting density happens to satisfy (11), then this density is the overall solution. If (11) is not satisfied, then the overall solution is the minimum cross-entropy density given (3) and the additional equality constraint $\langle g \rangle = \overline{g}$.

The principle of maximum entropy applies when one has constraint information (3) but no prior. In this case one selects the posterior by maximizing the posterior entropy

$$H(q) = \int_{\underset{\sim}{D}} d\underset{\sim}{x}\, q(\underset{\sim}{x}) \log(q(\underset{\sim}{x}))$$

subject to the constraints (3) and (5). The solution is the same as (6)-(9) with $p(\underset{\sim}{x})$ deleted. In general, selecting a posterior by maximizing entropy is equivalent to assuming a uniform prior and minimizing cross-entropy [5].

III. M/G/1 QUEUE LENGTH DISTRIBUTION

We consider an M/G/1 queuing system with an average arrival rate λ and a service time probability density s(t) with moments s_i (see (1)). Let $q_c(k)$ be the probability that k customers are in the system (queued or being served), and let c_m be the moments

$$c_m = \sum_{k=0}^{\infty} k^m q_c(k) \quad .$$

The first moment c_1 is just the expected number of customers in the system. In this Section, we use the Pollaczeck-Khinchin formula to express c_1 in terms of the first two service time moments s_1, s_2, and we derive a maximum entropy estimate of $q_c(k)$ given c_1. We then derive a formula expressing c_2 in terms of s_1, s_2, and s_3, and we compute maximum entropy estimates of q_c given c_1 and c_2. As examples, we consider M/M/1, $M/H_2/1$, and M/D/1 systems.

The result of maximizing the entropy of q_c subject to the single known constraint c_1 and the normalization constraint $\sum_k q_c(k) = 1$ is

$$q_c(k) = Z^{-1} e^{-\beta k} \quad , \tag{12}$$

where

$$Z = \sum_{k=0}^{\infty} e^{-\beta k} = (1 - e^{-\beta})^{-1} \tag{13}$$

(see (8)-(9)). We apply (10) in order to express the multiplier β in terms of the constraint c_1,

$$c_1 = -\frac{\partial}{\partial \beta} \log(Z) = (e^{\beta} - 1)^{-1} \quad .$$

This yields $e^{-\beta} = c_1/(1 + c_1)$, which we use (with (13)) to eliminate β from (12):

$$q_c(k) = \frac{1}{1 + c_1}\left(\frac{c_1}{1 + c_1}\right)^k \tag{14}$$

This expression gives the maximum entropy estimate of q_c directly in terms of the known information c_1.

Now, knowledge of s_1 and s_2 yields knowledge of c_1 by the Pollaczeck-Khinchen formula [3, p. 187]

$$c_1 = \rho + \rho^2 \frac{(1 + C^2)}{2(1 - \rho)} ,$$

$$= \rho + \frac{\lambda^2 s_2}{2(1 - \rho)} , \tag{15}$$

where $\rho = \lambda s_1$, and C is the coefficient of variation $C = (s_2 - s_1^2)^{\frac{1}{2}}/s_1$. Thus, (14) and (15) provide an information theoretic approximation to q_c for an M/G/1 system given the average arrival rate and the first two moments of the service time density.

As an example application, we consider an $M/H_2/1$ system solved exactly by Kleinrock [3, pp. 195-96]. The service time distribution is

$$s(t) = \frac{1}{4}\lambda e^{-\lambda t} + \frac{3}{4}(2\lambda)e^{-2\lambda t} , \tag{16}$$

for which $\rho = \lambda s_1 = 5/8$ and $C^2 = 31/25$. Substituting these values into (15) yields $c_1 = 1.79166$. The information theoretic approximation (14) then becomes

$$q_c(k) = .358209(.641791)^k . \tag{17}$$

The exact solution for q_c is [3, p. 196]

$$q_c(k) = \frac{3}{32}\left(\frac{2}{5}\right)^k + \frac{9}{32}\left(\frac{2}{3}\right)^k \quad . \tag{18}$$

We compare the one-moment approximation (17) with the exact solution (18) in the first three columns of Table I.

The extremely close agreement arises because the exact solution (18) is the sum of two similar geometric terms which can be approximated closely by a single geometric term (17). In general, the single-moment result (14) can be thought of as providing the geometric distribution that is the best information theoretic approximation to q_c.

If the exact solution itself happens to be geometric, then the approximation (14) will be the same as the exact solution. For example, suppose that the service time distribution is exponential $s(t) = \mu e^{-\mu t}$. Then (15) reduces to $c_1 = \rho/(1-\rho)$, with $\rho = \lambda s_1 = \lambda/\mu$, and the approximation (14) becomes $q_c(k) = (1-\rho)\rho^k$, which is the well-known exact solution for the M/M/1 system [3, p. 96].

If other moments besides c_1 are known, the maximum entropy estimate of q_c will no longer in general be geometric. In order to illustrate multi-moment approximations, we begin by deriving an expression for c_2 in terms of the service time moments s_m. Our starting point is the relation

$$c_2 = c_1 + \lambda^2 r_2 \quad , \tag{19}$$

where r_2 is the second moment of the system residence time probability density [3, p. 240]. The moments r_m are related to the s_m and to the moments w_m of the waiting time probability density by

$$r_k = \sum_{i=0}^{k} \binom{k}{i} w_{k-i} s_i \quad , \tag{20}$$

where $w_0 = s_0 \equiv 1$ [3, p. 202], and the w_m are in turn related to the s_m by the Takacs recurrance formula [3, p. 201]

$$w_k = \frac{\lambda}{1 - \rho} \sum_{i=1}^{k} \binom{k}{i} \frac{s_{i+1}}{i + 1} w_{k-i} \quad . \tag{21}$$

By combining (15), (20), and (21) with (19), we obtain

$$c_2 = \rho + \frac{\lambda^2 s_2}{2(1 - \rho)} + \frac{\lambda^4 s_2^2}{2(1 - \rho)^2} + \frac{\lambda^3 s_3}{3(1 - \rho)} + \frac{\lambda^3 s_1 s_2}{(1 - \rho)} + \lambda^2 s_2 \tag{22}$$

Now the maximum entropy solution for $q_c(k)$ given c_1 and c_2 cannot be expressed analytically in terms of the moments s_m, so we resort to numerical techniques. We use an APL function written by Johnson [18] that computes maximum entropy distributions given arbitrary expected value constraints. This function requires that the constraints (3) be written in the form

$$\sum_k g_r(k) q(k) = 0 \quad .$$

The APL function accepts as input the matrix $M_{rk} = g_r(k)$ and uses the Newton-Raphson method to find the maximum entropy solution for q(k). When the known expected values are moments c_m, the input matrix becomes

$$M_{mk} = g_m(k) = k^m - c_m \quad .$$

For the $M/H_2/1$ example, we have $c_1 = 1.79166$ from before. The moment s_3 is easily obtained from (16), and $c_2 = 8.68055$ follows from (22). Using the APL function to find the maximum entropy appproximations for q_c given c_1 and c_2, we obtain the results shown in the fourth column in Table I. This approximation, which was computed for 50 points, required 1.5 CPU seconds on a DEC PDP-10 KI processor. It is worth noting that single-moment

results from the APL function agreed with the analytic expression (14) up to eight digits.

As an additional example, we consider a system with constant ("deterministic") service time $1/\mu$ --- i.e., M/D/1. The service time probability density is $s(t) = \delta(t - 1/\mu)$, with moments

$$s_m = 1/\mu^m \quad . \tag{23}$$

We use (23), (15), and (14) to obtain a single-moment approximation for q_c, and we use (23), (22), (15), and the APL function to obtain a two-moment approximation. For $\lambda = 1$ and $\mu = 2$, the results are shown in Table II together with simulation results. The simulation result $q_c(k)$ is the relative amount of time the system had k customers present during an overall period covering 5000 arrivals. The two-moment approximation in Table II required 1.6 CPU seconds.

Approximations *involving more moments* can be computed similarly since c_m can in general be expressed as a function of $s_1,\ldots,s_{m+1}$ --- one method is to differentiate the Pollaczek-Khinchin transform equation [3, p. 194]. But the accuracy of the two-moment approximation for the $M/H_2/1$ and M/D/1 examples, which have radically different service time densities, and the reduction of the one-moment approximation to the exact result in the M/M/1 case, together suggest that the two-moment approximation will in general be quite accurate for M/G/1 systems. This is only a conjecture, however, and more detailed studies are needed.

IV. NUMBER SERVED IN A M/G/1 BUSY PERIOD

If the system is empty and a customer arrives at time t_1, and if t_2 is the next time at which the system is empty, then the period between t_1 and t_2 is called a busy period. Let $q_n(k)$ be the probability that the number of customers served in a busy period is k, and let n_m be the moments of q_n. As before, we assume that the mean arrival rate λ and some moments of the service time density s(t) are known.

Now the first four moments of q_n can be expressed in terms of the first four moments of s as follows [2, p. 158]:

$$
\begin{aligned}
n_1 &= \frac{1}{1-\rho} \\
n_2 &= \frac{1+\lambda^2 K_2}{(1-\rho)^3} \\
n_3 &= \frac{3(1+\lambda^2 K_2)^2}{(1-\rho)^5} - \frac{2-\lambda^3 K_3}{(1-\rho)^4} \\
n_4 &= \frac{6+\lambda^4 K_4}{(1-\rho)^5} - \frac{10(1+\lambda^2 K_2)(2-\lambda^3 K_3)}{(1-\rho)^6} + \frac{15(1+\lambda^2 K_2)^3}{(1-\rho)^7}
\end{aligned}
\tag{24}
$$

where $\rho = \lambda s_1 = K_1$ and the K_m are the cumulants (semi-invariants)

$$
\begin{aligned}
K_1 &= s_1 \\
K_2 &= s_2 - s_1^2 \\
K_3 &= s_3 - 3s_2 s_1 + 2s_1^3 \\
K_4 &= s_4 - 4s_3 s_1 - 3s_2^2 + 12 s_1^2 s_2 - 3s_1^4
\end{aligned}
\tag{25}
$$

Thus, for example, knowing the mean service time (s_1) is equivalent to knowing the mean number served during a busy period (n_1). The maximum entropy distribution in this case is

$$q_n(k) = Z^{-1} e^{-\beta k} ,$$

where

$$Z = \sum_{k=1}^{\infty} e^{-\beta k} = (e^{\beta k} - 1)^{-1} .$$

From (10) we have

$$n_1 = - \frac{\partial}{\partial \beta} \log(Z) = (e^{\beta} - 1)^{-1} ,$$

which we use to express q_n directly in terms of the known constraint n_1:

$$q_n(k) = \frac{1}{n_1 - 1} \left(\frac{n_1 - 1}{n_1} \right)^k \tag{26}$$

This result differs from the previous single moment result (14) because the domain of $q_n(k)$ is $k = 1, \ldots, \infty$ instead of $k = 0, \ldots, \infty$. Using the relation $n_1 = 1/(1 - \rho)$ (see (24)), (26) becomes

$$q_n(k) = (1 - \rho) \rho^{k-1} \tag{27}$$

where $\rho = \lambda s_1$. Eq. (27) provides an information theoretic approximation to the number served in a busy period for an M/G/1 system given the mean arrival rate and the mean service time.

Now, unlike the case for the distribution q_c, the distribution q_n for an M/M/1 system is not geometric. In fact, the exact result is [3, p. 218]

$$q_n(k) = \frac{1}{k} \binom{2k - 2}{k - 1} \rho^{k-1} (1 + \rho)^{1-2k} \tag{28}$$

This gives an opportunity to show how knowledge of higher moments than s_1 can be used to provide better approximations than (27). Now, for an M/M/1 system with $s(t) = \mu e^{-\mu t}$, the moments s_m are

$$s_m = m!/\mu^m \quad . \tag{29}$$

For a given λ and μ, we use (29), (25), and (24) to compute the moments n_m, and we use the APL function to compute the maximum entropy distribution $q_n(k)$ given the n_m. In Table III, for $\lambda = 2$ and $\mu = 8$, we compare the exact solution (28) with the single moment approximation (27) and the four moment approximation computed by the APL function. (As should be expected, approximations based on two and three moments fall between the approximations shown.) In Table IV we present the same comparison for $\lambda = 1$ and $\mu = 2$.

As another example, we again consider the M/D/1 system. As in the M/M/1 case, the exact result for q_n is known, namely [3, p. 219]

$$q_n(k) = \frac{(k\rho)^{k-1}}{k!} e^{-k\rho} \tag{30}$$

For a given λ and μ, we use (23), (25), and (24) to compute the n_m and then the APL function to compute maximum entropy approximations to q_n. Results comparing the exact solution (30) with one- and four-moment approximations are given in Table V and VI. (The values for λ and μ are the same as those for the M/M/1 examples in Table III and IV.) The four-moment approximations in Tables V and VI required about 1.5 CPU seconds each.

V. M/G/1 BUSY PERIOD LENGTH

We now consider the probability density $q_b(t)$ for the length of the busy

period. The first four moments of $q_b(t)$

$$b_m = \int_0^\infty dt\ t^m q_b(t) \quad ,$$

can be expressed in terms of λ and the service time moments s_m as follows:

$$\begin{aligned} b_1 &= \frac{s_1}{1-\rho} \\ b_2 &= \frac{s_2}{(1-\rho)^3} \\ b_3 &= \frac{s_3}{(1-\rho)^4} + \frac{3\lambda s_2^2}{(1-\rho)^5} \\ b_4 &= \frac{s_4}{(1-\rho)^5} + \frac{10\lambda s_2 s_3}{(1-\rho)^6} + \frac{15\lambda^2 s_2^3}{(1-\rho)^7} \end{aligned} \tag{31}$$

where, as usual, $\rho = \lambda s_1$ [3, pp. 214-5].

If only s_1 is known, then only b_1 is determined. The resulting maximum entropy solution for q_b is $q_b(t) = (1/b_1)\exp(-t/b_1)$. (We omit the standard derivation, which is just the continuous analog of the derivation of (14).) Combining this solution with the expression for b_1 from (31) yields

$$q_b(t) = (\mu' - \lambda)e^{-(\mu' - \lambda)t} \tag{32}$$

where $\mu' = 1/s_1$. Eq. (32) provides an information theoretic approximation to the busy period probability density for an M/G/1 system given the mean arrival rate and the mean service time.

If higher moments than s_1 are known, then better approximations can be obtained using (31) and the numerical techniques described in Section IV. These techniques must be modified slightly since here we are dealing with a continuous probability density. To put the problem into a discrete form, we

approximate the moment integrals by

$$b_m = \int_0^\infty dt\ t^m q_b(t) \approx \int_0^T dt\ t^m q_b(t)$$

$$\approx \sum_k \Delta_k t_k^m q_b(t_k)$$

for some sufficiently large T, where the Δ_k are widths of intervals surrounding the points t_k. In these terms, the normalization constraint is

$$1 = \int_0^\infty dt\ q_b(t) \approx \sum_k \Delta_k q_b(t_k) \ .$$

We can write the known constraints as

$$\sum_k q(k) = 1 \tag{33}$$

$$\sum_k (t_k^m - b_m) q(k) = 0 \ , \tag{34}$$

where q(k) is a discrete distribution defined by $q(k) = \Delta_k q_b(t_k)$. From (33)-(34) we can use the APL function to compute the maximum entropy distribution q(k) given the known moments b_m. The result yields an approximate solution for q_b at points t_k, since $q_b(t_k) = q(k)/\Delta_k$.

As in the previous section, the exact solution for an M/M/1 system is known, namely [3, p. 215]

$$q_b(t) = \frac{1}{t\sqrt{\lambda/\mu}}\, e^{-(\lambda+\mu)t}\, I_1(2t\sqrt{\lambda\mu}) \ , \tag{35}$$

where I_1 is the modified Bessel function of the first kind (order one). We therefore illustrate the foregoing by assuming s(t) to be exponential, computing various approximations based on (31) and (29), and comparing the results with (35). Results for the case $\lambda = 5$ and $\mu = 10$ are shown in Fig. 1. Results for the case $\lambda = 1$ and $\mu = 10$ are shown in Fig. 2 for the

one-moment approximation and in Table VII for the four-moment approximation at selected points. The single moment approximations, which were computed by the APL function, agree in both cases with (32).

The results in Fig. 2 suggest that (32) might be a good light-load approximation for the M/M/1 busy period density (35). Although systematic studies are needed to support this conjecure, it appears from a few additional runs that, for $\rho \leq .1$, (32) is accurate to within 5-10% in the range where the cumulative probability distribution of $q_b(t)$ is as large as about .95. The conjecture is supported further by the following argument, which is due to A. E. Ephremides [18]: Equation (2) is identical to the exact M/M/1 residence time probability density [3, p. 202]. Since most busy periods will consist of single customer residences under light load conditions, it makes sense that the busy period should tend to (2).

VI. M/G/1 RESIDENCE TIME AND WAITING TIME

Residence time is the total time a customer spends in the system. Waiting time is the interval between the arrival time and the time at which service begins. Moments r_m of the residence time probability density $q_r(t)$ can be expressed in terms of the service time moments s_m by using (20) and (21). For example, we have

$$\begin{aligned} r_1 &= \frac{\lambda s_2}{2(1-\rho)} + s_1 \\ &= \frac{\lambda s_2}{2(1-\rho)} + \frac{\rho}{\lambda} \end{aligned} \tag{36}$$

where $\rho = \lambda s_1$. This is related to (15) by Little's result $\lambda r_1 = c_1$. The maximum entropy density $q_r(t)$ given r_1 is just

$$q_r(t) = (1/r_1)\exp(-t/r_1) \quad . \tag{37}$$

Eqs. (36)-(37) provide an information theoretic approximation to the residence time probability density for an M/G/1 system given the mean arrival rate and the first two moments of the service time density. If higher moments than s_2 are known, then better approximations for q_r can be obtained by using (20), (21), and the computational methods discussed earlier.

For an M/M/1 system, (36) reduces to $r_1 = \rho/\lambda(1 - \rho)$ and (37 becomes $g_r(t) = \mu(1 - \rho)\exp(-\mu(1 - \rho)t)$, where $\mu = 1/s_1$, which is the exact M/M/1 solution [3, p. 202]. This behavior is similar to that of the one-moment approximation for $q_c(k)$ discussed in Section III. The similarity arises from (37) being the continuous analog of (14) and from Little's result.

The situation for waiting times is somewhat more complicated. Let $q_w(t)$ be the waiting time probability density with moments w_m. The w_m can be expressed in terms of the s_m using (21); for example,

$$w_1 = \frac{\lambda s_2}{2(1 - \rho)} \quad , \tag{38}$$

where $\rho = \lambda s_1$. The maximum entropy solution given just w_1 is

$$q_w(t) = (1/w_1)\exp(-t/w_1) \quad . \tag{39}$$

In the M/M/1 case, (38) becomes $w_1 = \rho/\mu(1 - \rho)$ and (39) becomes

$$q_w(t) = (\mu/\rho)(1 - \rho)\exp(-\mu(1 - \rho)t/\rho) \quad , \tag{40}$$

in contrast to the exact M/M/1 result [3, p. 203]

$$q_w(t) = (1 - \rho)\delta(t) + \lambda(1 - \rho)\exp(-\mu(1 - \rho)t) \quad . \tag{41}$$

Eqs. (40) and (41) have the same mean w_1, but (40) lacks the implulse term at $t = 0$ that results from the finite probability $q_c(0)$ that the system is

empty when a customer arrives. We can, however, improve on (40) by noting that s_1 and s_2 provide information about $q_c(0)$. In particular, we have

$$q_c(0) = (1 + c_1)^{-1}$$

$$= (1 + \rho + \lambda w_1)^{-1}$$

from (14). Now the total probability in $q_w(t)$ that is concentrated at $t = 0$ must equal $q_c(0)$. We express this fact as

$$\lim_{\epsilon \to 0} \int dt\, u_\epsilon(t) q_w(t) = q_c(0) = (1 + \rho + \lambda w_1)^{-1} \quad , \tag{42}$$

where

$$u_\epsilon(t) = \begin{cases} 1 , & t \leq \epsilon \\ 0 , & t > \epsilon \end{cases} \quad .$$

But the integral in (42) is just a constraint (3) that we can impose in addition to the moment constraint $\int dt\, t q_w(t) = w_1$. The maximum entropy density that satisfies these constraints is

$$q_w(t) = (\lambda w_1 + \rho + 1)^{-1} \delta(t) + w_1 B^2 \exp(-Bt) \quad , \tag{43}$$

where

$$B = \frac{\rho + \lambda w_1}{w_1(1 + \rho + \lambda w_1)} \quad . \tag{44}$$

Eqs. (43), (44), and (38) provide an information theoretic approximation to the waiting time probability density for an M/G/1 system given λ and s_1, s_2. Unlike (39), (43) reduces to (41) in the M/M/1 case $w_1 = \rho/\mu(1 - \rho)$.

VII. SOME G/G/1 RESULTS

We consider a G/G/1 queue that has a probability density of interarrival

times a(t) with moments a_m and a probability density of service times s(t) with moments s_m. We discuss approximations for the case in which only a_1 and s_1 are known.

Eq. (14) is the maximum entropy distribution of queue length q_c given the first moment c_1. The probability that the system is empty is therefore

$$q_c(0) = (1 + c_1)^{-1}. \tag{45}$$

Now, if the G/G/1 system is in equilibrium,

$$(1 - q_c(0))/s_1 = 1/a_1$$

must hold. Solving for $q_c(0)$ and substituting the result into (45) yields

$$c_1 = \frac{s_1/a_1}{(1 - s_1/a_1)}. \tag{46}$$

Eq. (15) then yields

$$q_c(k) = (1 - \rho)\rho^k \quad , \tag{47}$$

where $\rho = s_1/a_1$. This is an information theoretic approximation for the G/G/1 queue length given the first moments of the arrival and service time densities. As was the case for the M/G/1 approximation (14)-(15), Eq. (47) yields the exact M/M/1 result when a(t) and s(t) are exponential. Stated differently, (47) shows that the M/M/1 result is also the proper information theoretic approximation for G/G/1 systems given only a_1 and s_1.

Next we consider the residence time density q_r. Eq. (46) and Little's result $c_1 = r_1/a_1$ yield

$$r_1 = s_1/(1 - s_1/a_1).$$

The maximum entropy density q_r given r_1 is then

$$q_r(t) = \mu(1 - \rho)\exp(-\mu(1 - \rho)t) \quad . \tag{48}$$

where $\rho = s_1/a_1$ and $\mu = 1/s_1$. This is an information theoretic approximation for the G/G/1 queue length given the first moments of the arrival and service time densities. Like the M/G/1 approximation (36)-(37), (48) yields the exact M/M/1 result when a(t) and s(t) are exponential and also shows that the M/M/1 result is the proper information theoretic approximation for G/G/1 systems given only a_1 and s_1.

Similar arguments based on results from Section VI apply in the case of the waiting time density w_t. In this case, the G/G/1 approximation given a_1 and s_1 is

$$q_w(t) = (1 - \rho)\delta(t) + \lambda(1 - \rho)\exp(-\mu(1 - \rho)t) \quad . \tag{49}$$

where $\rho = s_1/a_1$ and $\mu = 1/s_1$.

VIII. USING NON-UNIFORM PRIORS

Since entropy maximization is equivalent to cross-entropy minimization with a uniform prior (Section II), the information theoretic approximations discussed in Sections III-VII are properly thought of as being based on uniform prior estimates of the performance distributions. If information about the performance distributions in addition to the s_m is available and can be expressed as non-uniform prior approximations, it is likely that better approximations would result. For example, if it is suspected that the service time density s(t) is nearly exponential, it would be reasonable to use M/M/1

performance distributions as prior approximations of M/G/1 distributions. As a specific example, suppose we wish to estimate the busy period density q_b based on measurements of λ, s_1 and s_2. As a prior approximation, we use the exact M/M/1 result (35) with $\mu = s_1$, and we compute the moments b_1 and b_2 from (31). We obtain a posterior approximation by minimizing cross-entropy with respect to the prior subject to the constraints b_1 and b_2. If s_2 happens to satisfy

$$s_2 = 2s_1^2, \tag{50}$$

which would always be the case if s(t) were exponential, then the posterior would be unchanged from the prior since the M/M/1 prior itself satisfies the constraints b_1 and b_2. If (50) is not satisfied, then the M/M/1 prior does not satisfy the constraints b_1 and b_2 and the posterior will be different. In an information theoretic sense, however, it will be the closest distribution that satisfies the constraints. Figure 3 shows an example in which two-moment approximations for q_b were computed using both uniform and M/M/1 priors. The parameters in both cases were $\lambda = 5$, $s_1 = .1$ and $s_2 = .04$. The second moment is larger than it would be if s(t) were exponential --- the coefficient of variation is 1.74 instead of one. Since $\lambda = 5$ and $1/s_1 = 10$, the non-uniform prior used in computing the result in Fig. 3 is the same as the M/M/1 curve shown in Figure 1. The results in Fig. 3 were obtained using an APL function that finds a minimum cross-entropy posterior given an arbitrary prior and an arbitrary constraint matrix [18].

IX. DISCUSSION

There are at least three possible uses for the results in this paper. First, the techniques presented could be used as a general method of computing the performance distributions in cases where all of the service density moments are available, i.e., when the density s(t) is known exactly. Second, the analytic approximations --- (14) and (15), (27), (32), (36) and (37), (38) and (43), (47)-(49) --- could be useful in various studies whenever explicit forms for the performance distributions are required. Third, and probably best, the techniques provide a means of estimating the performance distributions when only the first few moments of s(t) are known and s(t) itself is not known.

How accurate are these information theoretic approximations? Unfortunately, about all that can be said in general is that the approximations are the least-biased choices given the information available. To use the language of statistics [11], the approximations are the hypotheses that are best supported by the information available. Depending on the actual performance distribution and the number of moments considered, an information theoretic approximation may or may not be a good approximation in the mean-squared-error sense, although it is true that the mean-squared-error can always be made sufficiently small by taking sufficiently many moments into account. On the other hand, it is not generally known what kind of error measure is best for judging the accuracy of performance distribution approximations. It may well be that measures such as mean-squared-error are less important than information measures such as cross-entropy.

More can be said about the queue length distribution q_c and the busy period density q_b, because, although an explicit proof is lacking, it seems clear that these must be monotonically decreasing functions for a wide class of M/G/1 systems. If so, then q_c and q_b don't have basic stucture that would be seen in approximations based on many moments but not seen in approximations based on only a few moments. This in turn means that the basic shape will be revealed by approximations based on the first few moments, and suggests that a large number of moments will not in general be required in order to achieve low mean-squared-error. In the case of the queue length distribution, the diverse examples discussed in Section III suggest that a two-moment approximation may in general be quite good. Assuming that both q_c and q_b are monotonic, it seems reasonable to conjecture that both the mean-square-error and its rate of change will decrease monotonically with the number of moments used. If true, this would help in judging how close the approximation is to the unknown true distribution.

Queuing models, particularly ones with Poisson arrivals and exponentially distributed service times, have been used with remarkable success in the performance modeling and analysis of computer systems. Because computer systems do not satisfy many assumptions made by the stochastic process models that are used, this success has been somewhat puzzling. The results presented in this paper show that the information theory viewpoint may be the best one from which to understand this success. For example, Section VII showed that various M/M/1 formulas are also information theoretic approximations for G/G/1

systems. That is, the M/M/1 formulas are the best hypotheses about G/G/1 behavior given only the mean arrival and service rates. This fact has nothing at all to do with the various assumptions that must be debated when considering the applicability of stochastic models.

ACKNOWLEDGMENTS

I thank R. W. Johnson for H. Vantilborgh for helpful discussions, H. Vantilborgh for suggesting that (32) might be a good light-load approximation, and D. Baker and A. E. Ephemides for their comments on an earlier version of this paper. I also thank J. Stroup for typing the equations and for other help in preparing the manuscript.

Table I

Comparison of exact and approximate solutions for $M/H_2/1$ queue length distribution

k	$q_c(k)$ (exact)	$q_c(k)$ (1 moment approx.)	$q_c(k)$ (2 moment approx.)
0	.375	.358	.367
1	.225	.230	.229
2	.140	.148	.144
3	.0893	.0947	.0914
4	.0580	.0608	.0583
5	.0380	.0390	.0375
6	.0251	.0250	.0243
7	.0166	.0161	.0158
8	.0110	.0103	.0104
9	.00734	.00662	.00688
10	.00489	.00425	.00458

Table II

Comparison of information theoretic approximations and simulation results for *M/D/1 queue length distribution.* ($\lambda = 1$ and $\mu = 2$)

k	$q_c(k)$ (simulation)	$q_c(k)$ (1 moment approx.)	$q_c(k)$ (2 moment approx.)
0	.50	.57	.51
1	.33	.24	.30
2	.12	.10	.13
3	.038	.045	.044
4	.0093	.019	.011
5	.0025	.0083	.0022
6	.00047	.0035	.00033
7	.0000081	.0015	.000037
8	0	.00065	.0000032

Table III

Comparison of exact and approximate solutions for distribution of number served in an M/M/1 busy period. ($\lambda = 2$ and $\mu = 8$)

k	$q_n(k)$ (exact)	$q_n(k)$ (1 moment approx.)	$q_n(k)$ (4 moment approx.)
1	.800	.750	.793
2	.128	.187	.142
3	.0410	.0469	.0372
4	.0164	.0117	.0133
5	.00734	.00293	.00611
6	.00352	.000732	.00334
7	.00177	.000183	.00205
8	.000921	.0000458	.00134
9	.000491	.0000114	.000888
10	.000267	.00000286	.000567

Table IV

Comparison of exact and approximate solutions for distribution of number served in an M/M/1 busy period. ($\lambda = 1$ and $\mu = 2$)

k	$q_n(k)$ (exact)	$q_n(k)$ (1 moment approx.)	$q_n(k)$ (4 moment approx.)
1	.666	.500	.629
2	.148	.250	.195
3	.0658	.125	.0737
4	.0365	.0625	.0332
5	.0227	.0312	.0174
6	.0152	.0156	.0104
7	.0106	.00781	.00696
8	.00765	.00391	.00511
9	.00567	.00195	.00404
10	.00428	.000977	.00337

Table V

Comparison of exact and approximate solutions for distribution of number served in an M/D/1 busy period. ($\lambda = 2$ and $\mu = 8$)

k	$q_n(k)$ (exact)	$q_n(k)$ (1 moment approx.)	$q_n(k)$ (4 moment approx.)
1	.779	.750	.767
2	.151	.187	.169
3	.0443	.0469	.0433
4	.0153	.0117	.0127
5	.00583	.00293	.00426
6	.00235	.000732	.00169
7	.000990	.000183	.000682
8	.000430	.0000458	.000321
9	.000191	.0000114	.000167
10	.0000863	.00000286	.0000949

Table VI

Comparison of exact and approximate solutions for distribution of number served in an M/D/1 busy period. ($\lambda = 1$ and $\mu = 2$)

k	$q_n(k)$ (exact)	$q_n(k)$ (1 moment approx.)	$q_n(k)$ (4 moment approx.)
1	.606	.500	.589
2	.184	.250	.208
3	.0837	.125	.0868
4	.0451	.0625	.0420
5	.0267	.0312	.0230
6	.0168	.0156	.0140
7	.0110	.00781	.00927
8	.00744	.00391	.00657
9	.00515	.00195	.00490
10	.00363	.000977	.00378

Table VII

Comparison of exact and four-moment approximation for probability density of M/M/1 busy period length ($\lambda = 1$ and $\mu = 10$)

time	$q_b(t)$ (exact)	$q_b(t)$ (4 moment approx.)
.01	8.96	9.05
.03	7.22	7.27
.05	5.84	5.87
.07	4.74	4.75
.09	3.87	3.87
.11	3.17	3.16
.13	2.60	2.59
.15	2.14	2.14
.25	0.861	0.854
.35	0.373	0.369
.45	0.171	0.171

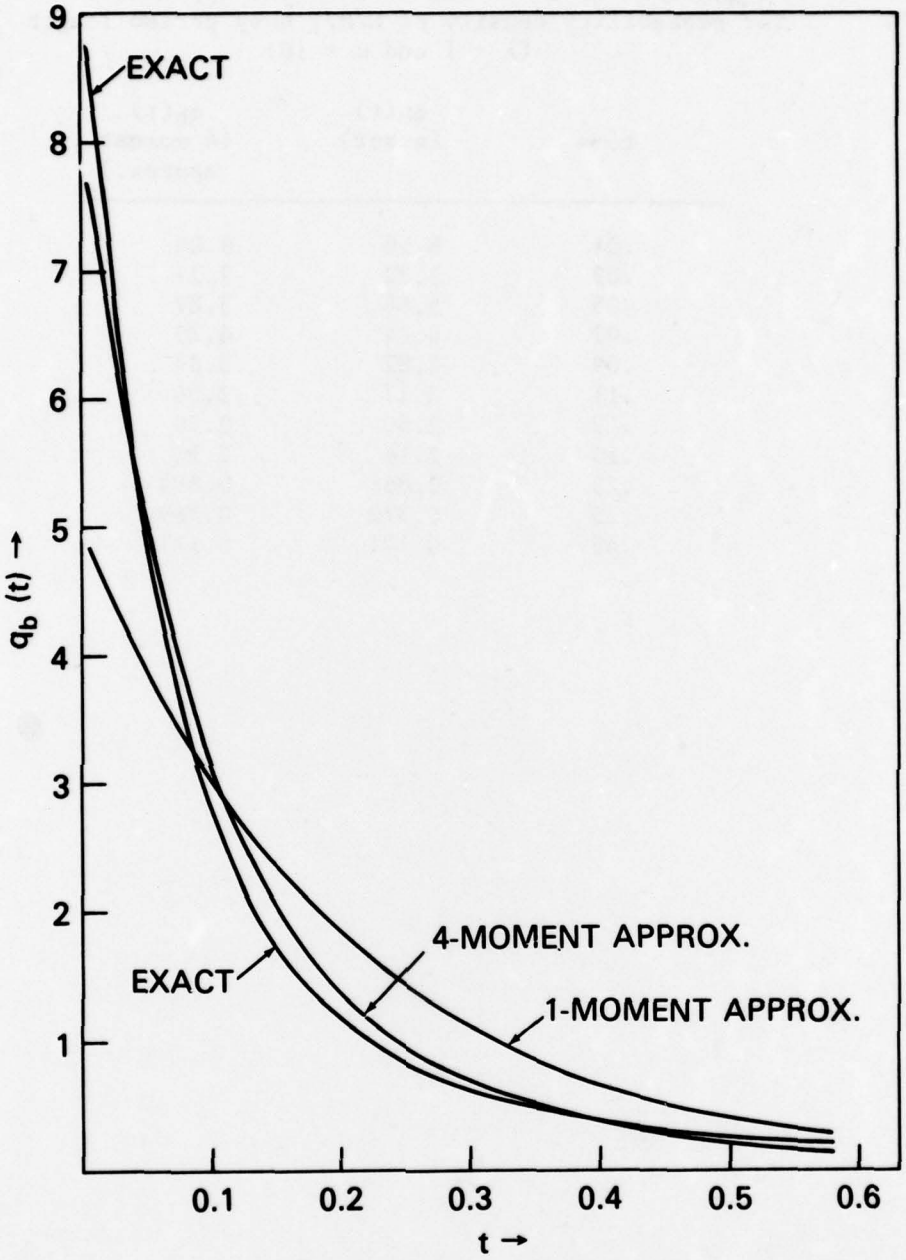


Fig. 1 - Exact and approximate M/M/1 busy period probability densities ($\lambda = 5$, $\mu = 10$)

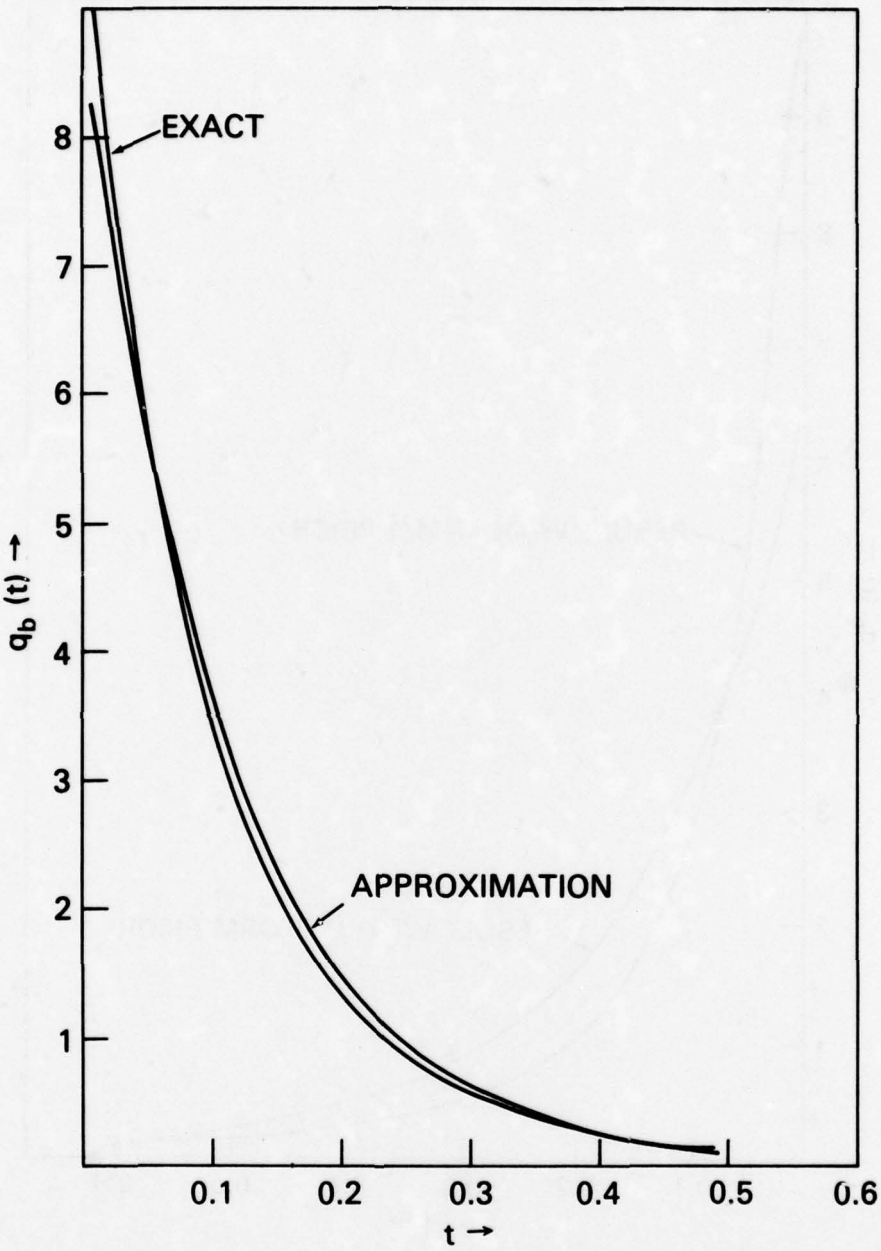


Fig. 2 - Exact and 1-moment approximation for M/M/1 busy period probability density ($\lambda = 1$, $\mu = 10$)

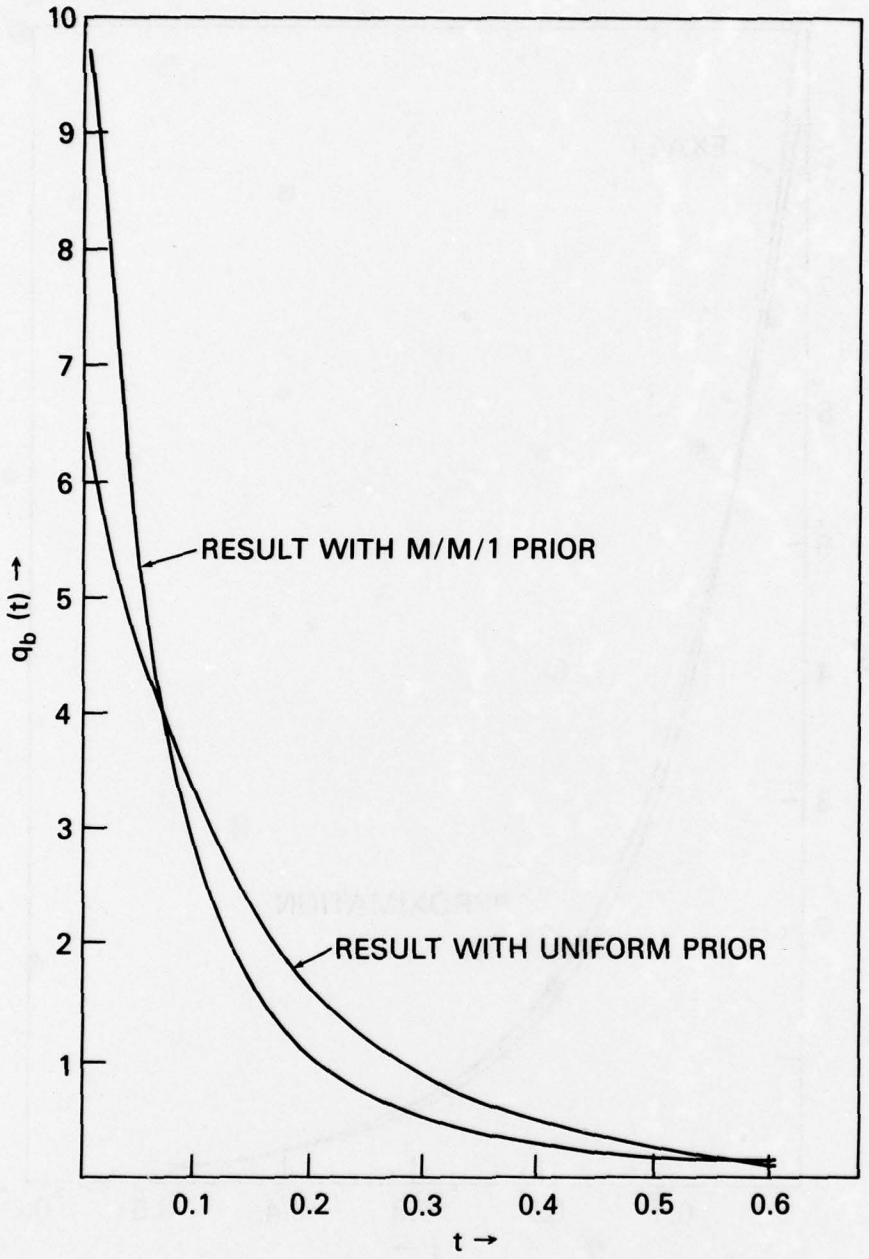


Fig. 3 - Two-moment approximations for M/G/1 busy period probability density using uniform and M/M/1 priors ($\lambda = 5$, $s_1 = .1$, $s_2 = .04$)

References

1. J. W. Cohen, The Single Server Queue, Amsterdam, North-Holland, 1969.

2. D. R. Cox and W. L. Smith, Queues, London, Chapman and Hall, 1961.

3. L. Kleinrock, Queueing Systems, Vol I: Theory, New York, John Wiley, 1975.

4. J. E. Shore and R. W. Johnson, "Axiomatic Derivation of the Principle of Maximum Entropy and the Principle of Minimum Cross-Entropy," NRL Memorandum Report 3898, Dec 1978, Naval Research Laboratory, Washington, D. C. 20375.

5. J. E. Shore and R. W. Johnson, "Axiomatic Derivation of the Principle of Maximum Entropy and the Principle of Minimum Cross-Entropy," IEEE Trans. Inf. Theory, to be published.

6. E. T. Jaynes, "Information Theory and Statistical Mechanics I," Phys. Rev. 106, 1957, pp. 620-630.

7. V. E. Benes, Mathematical Theory of Connecting Networks and Telephone Traffic, Academic Press, New York, 1965.

8. J. E. Shore, "Derivation of Equilibrium and Time-Dependent Solutions to M/M/ //N and M/M/ Queueing Systems Using Entropy Maximization, Proceedings 1978 National Computer Conference, AFIPS, 1978, pp. 483-487.

9. I. J. Good, "Maximum Entropy for Hypothesis Formulation, Especially for Multidimensional Contingency Tables," Annals Math. Stat. 34, 1963, pp. 911-934.

10. I. J. Good, Probability and the Weighing of Evidence, Charles Griffen, London, 1950.

11. S. Kullback, Information Theory and Statistics, Wiley, New York, 1959.

12. A. Hobson and B. Cheng, "A Comparison of the Shannon and Kullback Information Measures," J. Stat. Phys. 7, No. 4, 1973, pp. 301-310.

13. R. W. Johnson, "Axiomatic Characterization of the Directed Divergences and Their Linear Combinations," accepted by IEEE Trans. Inf. Theory.

14. E. T. Jaynes, "Prior Probabilities," IEEE Trans. on Systems Science and Cybernetics SSC-4, 1968, pp. 227-241.

15. C. E. Shannon, "A Mathematical Theory of Communication," Bell System Tech. Jour., 27, 1948, pp. 379-423.

16. A. Hobson, "A New Theorem of Information Theory," J. Stat. Phys. 1, No. 3, 1969, pp. 383-391.

17. R. W. Johnson, "APL Functions for Determining Probability Distributions by the Principles of Maximum Entropy and Minimum Cross-Entropy," Proceedings APL 79, May 1979.

18. A. E. Ephremides, private communication.