OPEN ACCESS Freely available online

PLOS | ONE

Maximum Entropy Principle Based Estimation of Performance Distribution in Queueing Theory

Dayi He, Ran Li, Qi Huang, Ping Lei*

School of Humanities & Economic Management, Lab of Resources & Environment Management, China University of Geosciences (Beijing), Beijing, P. R. China

Abstract

In related research on queuing systems, in order to determine the system state, there is a widespread practice to assume that the system is stable and that distributions of the customer arrival ratio and service ratio are known information. In this study, the queuing system is looked at as a black box without any assumptions on the distribution of the arrival and service ratios and only keeping the assumption on the stability of the queuing system. By applying the principle of maximum entropy, the performance distribution of queuing systems is derived from some easily accessible indexes, such as the capacity of the system, the mean number of customers in the system, and the mean utilization of the servers. Some special cases are modeled and their performance distributions are derived. Using the chi-square goodness of fit test, the accuracy and generality for practical purposes of the principle of maximum entropy approach is demonstrated.

Citation: He D, Li R, Huang Q, Lei P (2014) Maximum Entropy Principle Based Estimation of Performance Distribution in Queueing Theory. PLoS ONE 9(9): e106965. doi:10.1371/journal.pone.0106965

Editor: Christof Markus Aegerter, University of Zurich, Switzerland

Received June 24, 2014; **Accepted** August 8, 2014; **Published** September 10, 2014

Data Availability: The authors confirm that all data underlying the findings are fully available without restriction. All relevant data are within the paper.

Funding: This research is supported by the Key Laboratory of Carrying Capacity Assessment for Resource and Environment, MLR (CCA2012.05) and supported by the Fundamental Research Funds for the Central Universities (2-9-2012-86). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing Interests: The authors have declared that no competing interests exist.

* Email: leipingleiping@hotmail.com

Introduction

Queuing theory is mainly regarded as a branch of applied probability theory. Its applications are in different fields, such as communication networks, computer systems, machine plants, and services. Fig. 1 is a typical queuing system with a single server.

Queuing theory tries to answer questions like the mean waiting time in the queue, the mean system response time (waiting time in the queue plus service time), the mean utilization of the service facility, the distribution of the number of customers in the queue, and the distribution of the number of customers in the system. These questions are mainly investigated in a stochastic scenario, where, for example, the inter-arrival times of the customers or the serving times are assumed to be random typically Poisson arrivals and to have exponent distribution serving times.

However, usually we are mainly interested in steady state solutions (see Figure 2); that is, where the system after a long running time tends to reach a stable state in which, for example, the distribution of customers in the system does not change (limiting distribution).

In a canonical way, the steady state of system performance can be derived from assumptions on the distribution of inter-arrival times and service times. Hence, these assumptions are the basic requirement for analyzing queuing systems. However, in practical situations the pre-assumed distributions are difficult to satisfy or to acquire. To some degree, this fact limits the practical applications of queuing theory.

The maximum entropy principle is applicable to queuing theory because very often only partial information is available about the probability distributions. With respect to queuing theory, the principle of maximum entropy has been applied to solving numerous systems including, but not limited to, $M/G/1$ and $M/G/1$ queues ([1,2]), finite and infinite capacity $G/G1$ queues ([3,4]), multi-server queues ([5,6]), multiple class queues with priorities ([7]), and queues with vacation ([8–10]) and queuing networks ([11–13]). In fact, since the early 1970's many attempts have been made to apply the method of maximum entropy in the field of queuing theory. Ferdinand [14] used the method to derive the equilibrium solution of the $M/M/1/N$ system by analogy with statistical mechanics. Shore [15] built an abstract model from which he determined the maximum entropy solution of the $M/M/\infty$ and $M/M/\infty//N$ systems. Bard [16] applied entropy maximization to a class of problems in the performance evaluation of computer systems. El-Affendi and Kouvatsos [2] used the maximum entropy principle to analyze the $M/G/I-$ and $G/M/1$-queuing systems at equilibrium. Alfa and Chen [17] developed a discrete time approach model for obtaining the expected queue length of the $M(t)/G/1$ queue. Arizono, Cui, and Ohta [18] analyzed $M/M/S$ using the maximum entropy principle. Falin, Martin, and Artalejo [19] presented information on theoretic approximations for the $M/G/1$ queue with retrials. Kouvatsos and Tabet-Aouel [20] applied entropy maximization to characterize the distributional form of the steady-state probabilities of a $G/G/c/PR$ queue with $c(\geq 2)$ parallel servers and $R(\geq 2)$ priority classes under a pre-emptive resume (PR) rule. Tadj and Hamdi [21] dealt with a quorum queuing system with a threshold level $r(\geq 1)$ that regulates the beginning and ending of idle and busy periods as follows: an idle period starts when the queue size drops below level r and a busy period starts as soon as the queue accumulates the same number r. The single server

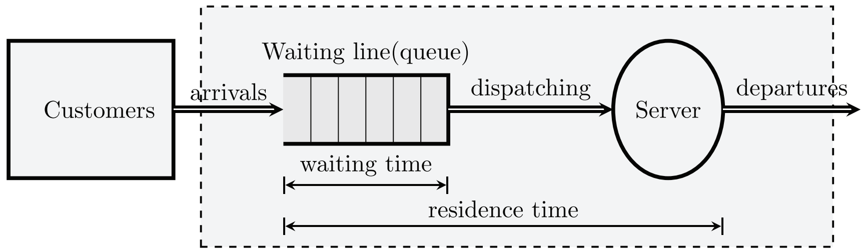

Figure 1. Queuing System Structure with Single-Server.
doi:10.1371/journal.pone.0106965.g001

processes r customers in one batch. They denoted this system as $M/G^r/1$ by Kendall's notation. They presented a maximum entropy model to determine the distribution of related random variables with known server utilization and mean queue length.

The work mentioned above applied the principle of maximum entropy to certain kinds of queuing systems based on assumptions concerning inter-arrivals and server times, which limited the practical applications of queuing theory. In fact, if a queue system is stable, this is not necessary when the maximum entropy method is applied to a certain queuing system. In this study, we consider the queuing system as a black box and derive a performance index for the queuing system by the principle of maximum entropy only on the assumption that the queue is stable instead of making assumptions on the distribution of inter-arrival times and service times. Meanwhile, from the viewpoint of expanding the practical application of queuing systems, we use some easily accessible indexes of queuing systems, such as the capacity of the system, the mean of customers in the queue, and the mean utilization of the system. Based on these indexes and the principle of maximum entropy, optimization models are then developed to derive the performance of queuing systems.

This paper is organized as follows. Section 2 is a simple review of the maximum entropy principle. Sections 3 and 4 develop different maximum entropy models with known mean numbers of customers and the average value of busy periods under unlimited and limited server capacity, respectively. Section 5 compares our results with general models with known assumptions on the distributions of inter-arrival times and server times by the χ^2 goodness of fit test. The last section concludes.

Methods: The Principle of Maximum Entropy

The principle of maximum entropy provides a solution to the old problem of the assignment of a probability distribution to a random variable that avoids bias while satisfying given or known information about the random variable. Jaynes is credited with having formalized the principle of maximum entropy in [22].

Mathematically the principle can be presented as follows: consider a system Θ that has a finite or countable infinite set θ of possible states $\theta_0, \theta_1, \cdots, \theta_n, \cdots$. Let $p(\theta_i) = p_i$ be the probability that the system Θ is in state θ_i. Suppose all that is known about these probabilities are $(m+1)$ constraints of the form

$$\sum_{i=0}^{n} p_i = 1, \quad p_i \geq 0 \tag{1}$$

$$\sum_{i=0}^{n} f_k(\theta) p_i = M_k, \qquad 1 \leq k \leq m < \infty \tag{2}$$

where $\{M_k\}$ are expectations defined on a set of suitable functions $\{f_k(\theta)\}$, which can be looked at as the known information. Since, in general, the number of constraints is less than the number of possible states, one is faced with an infinite number of distributions $\{p_i\}$ that satisfy these constraints. The problem is which one to choose.

The maximum entropy principle states that, of all the distributions satisfying the constraints supplied by the given information, the minimally prejudiced distribution that should be chosen is the one that maximizes the system entropy,

$$H(p) = -\sum_{i=0}^{n} p_i \ln p_i. \tag{3}$$

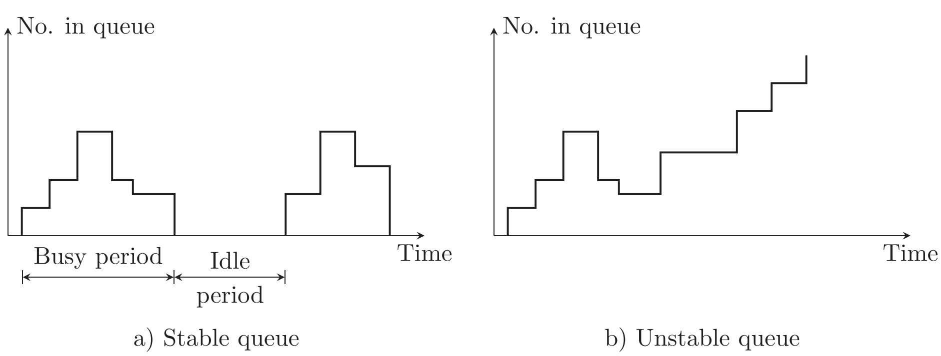

Figure 2. Stable and Unstable Queueing System.
doi:10.1371/journal.pone.0106965.g002

To sum up, the maximum entropy principle can be described as the following model for discrete variants.

$$\max\ H(p) = -\sum_{i=0}^{n} p_i \ln p_i$$
$$\text{s.t.} \begin{cases} \sum_{i=0}^{n} p_i = 1 \\ \sum_{i=0}^{n} f_k(\theta) p_i = M_k, & 1 \le k \le m < \infty \\ p_i \ge 0, & i = 0,1,2,\cdots,n \end{cases} \tag{4}$$

This is a natural extension of Laplace's famous principle of insufficient reason, which postulates that the uniform distribution is the most satisfactory representation of our knowledge when we know nothing about the random variable except that each probability is non-negative and that the sum of the probabilities is unity.

The maximization of (3) subject to constraints (1) and (2) can be solved by using the Lagrangian method of undetermined multipliers leading to the solution:

$$p_i = \exp(-\lambda_0 - \lambda_1 f_1(\theta) - \cdots - \lambda_m f_m(\theta)),\ i = 0,1,2,\cdots,n \tag{5}$$

where $\{\lambda_k\}$ are the Lagrangian multipliers corresponding to the set of constraints (1) and (2). More details on the maximum entropy principle and its generalizations can be found in [23,24].

Especially, if only the expected value E is known, the second constraint converts to

$$\sum_{i=0}^{n} i p_i = E$$

and according to (5) the estimated distribution is

$$p_i = \exp(-\lambda_0 - \lambda_1 i),\ i = 0,1,2,\cdots,n$$

To be simple, let

$$\alpha = \exp(-\lambda_0), \quad \beta = \exp(-\lambda_1)$$

then

$$p_i = \alpha\beta^i, \quad i = 1,2,\cdots,n$$

Substituting the above equation into the constraints, we can get

$$\begin{cases} \sum_{i=0}^{n} \alpha\beta^i = 1 \\ \sum_{i=0}^{n} i\alpha\beta^i = E \end{cases}$$

Hence,

$$\sum_{i=0}^{n} i\alpha\beta^i = E\sum_{i=0}^{n} \alpha\beta^i \Rightarrow \sum_{i=0}^{n} (i-E)\beta^i = 0$$

where the coefficients of β in the above equation are increasing, and $(0-E)<0$ and there exists an n that makes $(n-E)>0$, so that there is only one change in sign in the coefficients of β. According to Descartes's rule of signs, there is only one positive real root to the above equation. This indicates the solution uniqueness of the maximum entropy estimation problem even if there are only expected value and unit and non-negative requirements, which provides the basis for the later study in this article.

The maximum entropy approach to queuing systems is based on finding a maximum-entropy performance distribution based on the knowledge of some moments of the distribution concerned. To simplify, we will only discuss a queuing system with a single server and infinite customers, and where the dispatching rule is FIFO (First In First Out). The queuing systems are classified by server capacity into two types: queuing systems with either unlimited or limited server capacity. In each type, we will estimate the distribution of the system state by the maximum entropy principle from the mean number of customers in the system and the average value of a busy period.

Discussion

1. Queuing system with unlimited server capacity

If it has unlimited server capacity, a queuing system can serve as many customers as possible. Generally speaking, it is hard to discover the distributions of customer arrivals and departures. Hence, for a queuing system it is easy and practical to acquire knowledge of the mean number of customers and the busy periods etc. if the queue is stable.

Let the mean number of customers in the system under steady state be L_s (≥ 0), and p_i is the probability of the fact that there are $i(i=0,1,2,3,\cdots,)$ customers in the queuing system. According to the maximum entropy principle, the following model can be established if there is no more information.

$$\max\ S = -\sum_{i=0}^{\infty} p_i \ln p_i$$
$$\text{s.t.} \begin{cases} \sum_{i=0}^{\infty} p_i = 1, & \text{(a)} \\ \sum_{i=0}^{\infty} i p_i = L_s, & \text{(b)} \\ p_i \ge 0, & i = 0,1,2,\cdots, \end{cases} \tag{6}$$

Based on the method in Section 2, the distribution of the system state is

$$p_i = \mathrm{e}^{-(1+\lambda_0)-\lambda_1 i} = \mathrm{e}^{-(1+\lambda_0)}(\mathrm{e}^{-\lambda_1})^i := ab^i \tag{7}$$

with (6.a) and (6.b),

$$\begin{cases} \sum_{i=0}^{\infty} ab^i = 1 \\ \sum_{i=0}^{\infty} iab^i = L_s \end{cases} \tag{8}$$

As we know, the above equations have only one positive real root. As for $\sum_{i=0}^{\infty} b^i = \frac{1}{1-b}$ $(b<1)$, we can get

$$\sum_{i=0}^{\infty} ib^i = \frac{b}{(1-b)^2}$$

so with (8)

$$\frac{L_s}{1-b} = \frac{b}{(1-b)^2}$$

then,

$$\begin{cases} b = \frac{L_s}{1+L_s} \\ a = \frac{1}{1+L_s} \end{cases} \tag{9}$$

so

$$p_i = ab^i = \frac{1}{1+L_s}\left(\frac{L_s}{1+L_s}\right)^i \tag{10}$$

Let

$$\rho = \frac{L_s}{1+L_s} \tag{11}$$

then

$$p_i = (1-\rho)\rho^i \tag{12}$$

The results in (12) reflect the probability distribution function of a single server queuing system with a known mean number of customers and without limitation on system capacity, from which we can obtain the performance indexes of this kind of queuing system. It should be noted that the results presented here are coherent with a $M/M/1$ queuing system. The intrinsic reason is that the maximum entropy distribution with a known non-negative mean value is a Poisson distribution [25].

Another situation is that the queuing system has unlimited server capacity and a known mean server utilization of of $1-p_0$. Under this situation, the maximum entropy model is changed to be:

$$\max \ S = -\sum_{i=0}^{\infty} p_i \ln p_i$$
$$\text{s.t.} \begin{cases} \sum_{i=1}^{\infty} p_i = 1 - p_0 \\ \sum_{i=0}^{\infty} ip_i = L_s \\ p_i \geq 0, \ i = 0,1,2,\cdots, \end{cases} \tag{13}$$

With a similar approach, we can get

$$p_i = ab^i \tag{14}$$

where

$$\frac{ab}{1-b} = 1 - p_0, \qquad \frac{ab}{(1-b)^2} = L_s \tag{15}$$

that makes

$$a = \frac{(1-p_0)^2}{L_s - 1 + p_0}, \qquad 1 - b = \frac{1-p_0}{L_s} \tag{16}$$

then

$$p_i = \frac{(1-p_0)^2}{L_s - 1 + p_0}\left(\frac{L_s - 1 + p_0}{L_s}\right)^i \quad i = 1,2,\cdots, \tag{17}$$

Then, with a known L_s and p_0, the system performance can be achieved.

2. Queuing system with limited server capacity

In this section we will study the situation where the queuing system has a limited capacity; that is, there are N customers at most in the system, $N+1$ customers will leave, and the other assumptions are the same as before.

Firstly, we will consider the queuing system with only a known mean customer L_s. Then, the maximum entropy model is:

$$\max \ S = -\sum_{i=0}^{N} p_i \ln p_i$$
$$\text{s.t.} \begin{cases} \sum_{i=0}^{N} p_i = 1, & \text{(a)} \\ \sum_{i=0}^{N} ip_i = L_s, & \text{(b)} \\ p_i \geq 0, & i = 0,1,2,\cdots,N \end{cases} \tag{18}$$

By a similar method, we come to

$$p_i = ab^i, \quad i = 0,1,2,\cdots,N \tag{19}$$

where a and b are the roots of the following equations

$$\begin{cases} a\sum_{i=0}^{N} b^i = 1, \\ a\sum_{i=0}^{N} ib^i = L_s \end{cases} \tag{20}$$

As we know, the above equations have only one positive root. And we define

$$f(b) = \sum_{i=0}^{N} (i - L_s)b^i = 0 \tag{21}$$

And because

$$f(0) = 0, f(1) = N\left(\frac{N+1}{2} - L_s\right), f(\infty) > 0 \tag{22}$$

we can come to the following results:

i) if $L_s < \frac{N}{2}$, then $b < 1$,

ii) if $L_s = \frac{N}{2}$, then $b = 1$ and the distribution of steady state is an average distribution,

iii) if $L_s > \frac{N}{2}$, then $b > 1$.

Table 1 and Table 2 demonstrate the above results. In Table 1 the known mean number of customers is assumed to be 6 and in Table 2 the number is assumed to be 8. In Table 1, the capacity of the system is set to be $N = 6$, when L_s is set to be different values, we can get a' and b'value by solving (20). Then, by using (19) the value of p_i can be calculated. Then, the maximum entropy can be calculated too. By similar approach, Table 2 can be achieved.

If the system with limited server capacity has information of mean customer number L_s and busy period $1 - p_0$, the maximum entropy model will be

$$\max\ S = -\sum_{i=0}^{N} p_i \ln p_i$$
$$\text{s.t.} \begin{cases} \sum_{i=1}^{N} p_i = 1 - p_0 \\ \sum_{i=0}^{N} ip_i = L_s \\ p_i \geq 0,\ i = 0,1,2,\cdots,N \end{cases} \tag{23}$$

So we get

$$p_i = cd^i \qquad i = 1,2,\cdots,N \tag{24}$$

where

Table 1. Maximum Entropy Distribution ($N = 6$).

L_s	b	a	p_0	p_1	p_2	p_3	p_4	p_5	p_6	S_{max}
0.5	0.335	0.666	0.666	0.223	0.075	0.025	0.008	0.003	0.001	0.954
1.0	0.517	0.488	0.488	0.252	0.130	0.067	0.035	0.018	0.009	1.378
1.5	0.651	0.367	0.367	0.239	0.156	0.101	0.066	0.043	0.028	1.646
2.0	0.768	0.276	0.276	0.212	0.162	0.125	0.096	0.074	0.057	1.818
2.5	0.881	0.202	0.202	0.178	0.157	0.138	0.122	0.107	0.095	1.915
3.0	**1.000**	**0.143**	**0.143**	**0.143**	**0.143**	**0.143**	**0.143**	**0.143**	**0.143**	**1.946**
3.5	1.135	0.095	0.095	0.107	0.122	0.138	0.157	0.178	0.202	1.915
4.0	1.302	0.057	0.057	0.074	0.096	0.125	0.162	0.212	0.276	1.818
4.5	1.536	0.028	0.028	0.043	0.066	0.101	0.156	0.239	0.367	1.646
5.0	1.935	0.009	0.009	0.018	0.035	0.067	0.130	0.252	0.488	1.378
5.5	2.987	0.001	0.001	0.003	0.008	0.025	0.075	0.223	0.666	0.954

doi:10.1371/journal.pone.0106965.t001

Table 2. Maximum Entropy Distribution ($N=8$).

L_s	b	a	p_0	p_1	p_2	p_3	p_4	p_5	p_6	p_7	p_8	S_{max}
0.5	0.334	0.667	0.667	0.222	0.074	0.025	0.008	0.003	0.001	0.000	0.000	0.955
1.0	0.505	0.496	0.496	0.251	0.126	0.064	0.032	0.016	0.008	0.004	0.002	1.384
1.5	0.618	0.387	0.387	0.239	0.148	0.092	0.057	0.035	0.022	0.013	0.008	1.671
2.0	0.707	0.306	0.306	0.217	0.153	0.108	0.077	0.054	0.038	0.027	0.019	1.876
2.5	0.785	0.243	0.243	0.190	0.149	0.117	0.092	0.072	0.057	0.044	0.035	2.022
3.0	0.857	0.191	0.191	0.163	0.140	0.120	0.103	0.088	0.075	0.065	0.055	2.121
3.5	0.927	0.148	0.148	0.137	0.127	0.118	0.109	0.101	0.094	0.087	0.081	2.178
4.0	**1.000**	**0.111**	**0.111**	**0.111**	**0.111**	**0.111**	**0.111**	**0.111**	**0.111**	**0.111**	**0.111**	**2.197**
4.5	1.079	0.081	0.081	0.087	0.094	0.101	0.109	0.118	0.127	0.137	0.148	2.178
5.0	1.168	0.055	0.055	0.065	0.075	0.088	0.103	0.120	0.140	0.163	0.191	2.121
5.5	1.275	0.035	0.035	0.044	0.057	0.072	0.092	0.117	0.149	0.190	0.243	2.022
6.0	1.414	0.019	0.019	0.027	0.038	0.054	0.077	0.108	0.153	0.217	0.306	1.876
6.5	1.617	0.008	0.008	0.013	0.022	0.035	0.057	0.092	0.148	0.239	0.387	1.671
7.0	1.981	0.002	0.002	0.004	0.008	0.016	0.032	0.064	0.126	0.251	0.496	1.384
7.5	2.998	0.000	0.000	0.000	0.001	0.003	0.008	0.025	0.074	0.222	0.667	0.955

doi:10.1371/journal.pone.0106965.t002

$$c\sum_{i=1}^{N} d^i = 1-p_0,\ c\sum_{i=1}^{N} id^i = L_s \tag{25}$$

that is

$$\frac{cd(1-d^N)}{1-d} = 1-p_0,\ c\frac{d-(N+1)d^{N+1}+Nd^{N+2}}{(1-d)^2} = L_s \tag{26}$$

The solution is the root of the following equation:

$$\frac{1-(N+1)d^N+Nd^{N+1}}{1-d-d^N+d^{N+1}} = \frac{L_s}{1-p_0} \tag{27}$$

By equation (27) with known N, L_s, and p_0, the value of d can be decided and then c can be decided too by equation (26): then we can get the steady state distribution $\{p_i\}$. Similar results can be derived:

i) if $\frac{L_s}{1-p_0} > \frac{N+1}{2}$, then $d>1$,

ii) if $\frac{L_s}{1-p_0} = \frac{N+1}{2}$, then $d=1$,

iii) if $\frac{L_s}{1-p_0} < \frac{N+1}{2}$, then $d<1$.

In Tables 3 and 4, we assume $N=6$ and p_0 equals 0.1 and 0.01, respectively; then we calculate the system performance distribution by changing the value of L_s. In Table 3, by using (26) the value of c and d can be got. Then substituting c and d into (24) the value of p_i can be calculated and the corresponding entropy can be calculated too. In similar procedure, we can get Table 4.

3. The chi-square goodness of fit test

Without any assumptions on the distribution of inter-arrival times and server times, we deduced the performance distributions of the queuing system by the maximum entropy principle above. Is this method effective and feasible? We will test our method by the χ^2 goodness of fit test to determine this.

Taking model (18) as an example, if we know the distribution of inter-arrival times and server times follow the Poisson process, then we get a $M/M/1/N$ queuing system. Its performance distribution is

$$\begin{cases} p_0 = \dfrac{1-\rho}{1-\rho^{n+1}}, & \rho = \dfrac{\lambda}{\mu} \neq 1 \\ p_n = p_0\rho^n & \end{cases} \tag{28}$$

and the mean number of customers in the system is

$$L_s = \sum_{n=0}^{N} np_n = \frac{\rho}{1-\rho} - \frac{(N+1)\rho^{N+1}}{1-\rho^{N+1}} \tag{29}$$

We can decide the value of N and L_s and get the value of ρ by solving (29), the theoretic distribution can be calculated by using (28). The results are shown in Table 5.

On the other side, if we only know N and L_s of the queuing system, according to (18) we can arrive at the maximum entropy distribution as shown in Table 1. Comparing Table 5 with

Table 3. Maximum Entropy Distribution ($N=6$, $p_0=0.1$).

L_s	p_1	p_2	p_3	p_4	p_5	p_6	S_{max}
1.00	0.810	0.081	0.008	0.001	0.000	0.000	0.650
1.50	0.533	0.219	0.090	0.037	0.015	0.006	1.331
2.00	0.374	0.226	0.137	0.083	0.050	0.030	1.669
2.50	0.260	0.201	0.155	0.120	0.093	0.072	1.856
3.00	0.172	0.163	0.154	0.145	0.137	0.129	1.933
3.15	**0.150**	**0.150**	**0.150**	**0.150**	**0.150**	**0.150**	**1.938**
3.50	0.104	0.119	0.137	0.156	0.179	0.205	1.914
4.00	0.053	0.075	0.106	0.151	0.213	0.302	1.795
4.50	0.018	0.035	0.065	0.122	0.229	0.430	1.556
5.00	0.002	0.006	0.019	0.060	0.193	0.621	1.127

doi:10.1371/journal.pone.0106965.t003

Table 4. Maximum Entropy Distribution ($N=6$, $p_0=0.01$).

L_s	p_1	p_2	p_3	p_4	p_5	p_6	$S_{\max}$
1.000	0.980	0.010	0.000	0.000	0.000	0.000	0.112
1.500	0.650	0.224	0.077	0.027	0.009	0.003	1.016
2.000	0.467	0.252	0.136	0.073	0.040	0.021	1.422
2.500	0.338	0.236	0.165	0.115	0.080	0.056	1.663
3.000	0.239	0.203	0.172	0.146	0.124	0.105	1.792
3.500	0.160	0.162	0.164	0.166	0.168	0.170	1.830
3.465	**0.165**	**0.165**	**0.165**	**0.165**	**0.165**	**0.165**	**1.830**
4.000	0.098	0.118	0.143	0.172	0.208	0.252	1.780
4.500	0.050	0.074	0.110	0.162	0.240	0.354	1.636
5.000	0.018	0.035	0.067	0.130	0.252	0.489	1.377
5.500	0.002	0.006	0.020	0.066	0.212	0.683	0.938

doi:10.1371/journal.pone.0106965.t004

Table 5. Theoretic Distribution of $M/M/1/N$ ($N=6$).

L_s	ρ	p_0	p_1	p_2	p_3	p_4	p_5	p_6
0.5	0.335	0.665	0.223	0.075	0.025	0.008	0.003	0.001
1.0	0.517	0.488	0.252	0.130	0.067	0.035	0.018	0.009
1.5	0.651	0.367	0.239	0.156	0.101	0.066	0.043	0.028
2.0	0.768	0.276	0.212	0.162	0.125	0.096	0.074	0.057
2.5	0.881	0.202	0.178	0.157	0.138	0.122	0.107	0.095
3.0	**1.000**	**0.143**	**0.143**	**0.143**	**0.143**	**0.143**	**0.143**	**0.143**
3.5	1.135	0.095	0.107	0.122	0.138	0.157	0.178	0.202
4.0	1.302	0.057	0.074	0.096	0.125	0.162	0.212	0.276
4.5	1.536	0.028	0.043	0.066	0.101	0.156	0.239	0.367
5.0	1.934	0.009	0.018	0.035	0.067	0.130	0.252	0.488
5.5	2.984	0.001	0.003	0.008	0.025	0.075	0.223	0.665

doi:10.1371/journal.pone.0106965.t005

Table 1, it can be found that the maximum entropy distribution is very close to the theoretic distribution; that is, the χ^2 is almost equal to 0. Hence, the maximum entropy distribution is a good estimation.

Conclusions

Queuing system analysis is usually based on some assumptions about the distributions of inter-arrival times and server times. This study shows that there is no need to assume those distributions, and if a queuing system is looked at as a black box, the system performance can be estimated by the maximum entropy principle with some easily accessible macro-level indexes. In this paper, some common queuing system are studied including queuing system with unlimited server capacity and queuing system with limited server capacity. By utilizing the principle of maximum entropy, and with known information of some easily accessible macro-level indexes such as mean number of customers in the system L_s, system capacity N and mean server utilization of of $1-p_0$, we demonstrate that maximum entropy method is a feasible and effective approach to estimate the system performance distribution.

However, our study focused on single server queuing systems. For further research, multi-server queuing systems should be taken into consideration. For multi-server queuing system, more factors, for example queuing rules, server layout, system capacity et.al., should be considered. Hence, it will be more complicated. However, with assuming on those factors and observed indexes as presented in this paper, we can that our methods will be applicable in those circumstances also.

Author Contributions

Conceived and designed the experiments: DH. Performed the experiments: RL. Analyzed the data: QH. Contributed reagents/materials/analysis tools: PL. Contributed to the writing of the manuscript: DH PL.

References

1. Shore JE (1982) Information theoretic approximations for M/G/1 and G/G/1 queuing systems. Acta Informatica 17: 43–61.
2. El-Affendi MA, Kouvatsos DD (1983) A maximum entropy analysis of the M/G/1 and G/M/1 queueing systems at equilibrium. Acta Informatica 19: 339–355.
3. Kouvatsos DD (1986) Maximum entropy and the G/G/1/N queue. Acta Informatic 23: 545–565.
4. Kouvatsos DD (1988) A maximum entropy analysis of the G/G/1 queue at equilibrium. The Journal of the Operational Research Society 39: 183–200.
5. Kouvatsos DD, Almond J (1988) Maximum entropy two-station cyclic queues with multiple general servers. Acta Informatic 26: 241–267.
6. Wu JS, Chan WC (1989) Maximum entropy analysis of multiple-server queueing systems. The Journal of the Operational Research Society 40: 815–825.
7. Kouvatsos DD, Tabet-Aouel N (1989) A maximum entropy priority approximation for a stable G/G/1 queue. Acta Informatica 27: 247–286.
8. Skianis C, Kouvatsos DD (1998) Arbitrary open queueing networks with server vacation periods and blocking. Annals of Operations Research 79: 143–180.
9. Wang KH, Chuang SL, Pearn WL (2002) Maximum entropy analysis to the N policy M/G/1 queueing system with a removable server. Applied Mathematical Modelling 26: 1151–1162.
10. Ke JC, Lin CH (2008) Maximum entropy approach for batch-arrival queue under N policy with an un-reliable server and single vacation. Journal of Computational and Applied Mathematics 221: 1–15.
11. Walstra RJ (1985) Nonexponential networks of queues: a maximum entropy analysis. SIGMETRICS Performance Evaluation Review 13: 27–37.
12. Wu JS (1992) Maximum entropy analysis of open queueing networks with group arrivals. The Journal of the Operational Research Society 43: 1063–1078.
13. Kouvatsos DD, Awan I (2003) Entropy maximisation and open queueing networks with priorities and blocking. Performance Evaluation 51: 191–227.
14. Ferdinand AE (1970) A statistical mechanical approach to systems analysis. IBM Jnl. Res. Develop. 14: 539–547.
15. Shore JE (1978) Derivation of equilibrium and time-dependent solutions to $M/M/\infty//N$ and $M/M/\infty$ queueing systems using entropy maximisation. In Proc. Natl. Comput. Conference; Ghosh S. P., Liu L. Y., Eds; AFIPS Press: Anheim, CA, pp. 483–487.
16. Bard Y (1980) Estimation of state probabilities using the maximum entropy principle. IBM Jnl. Res. Develop. 24: 563–569.
17. Alfa AS, Chen M (1991) Approximating queue lengths in $M(t)/G/1$ queue using the maximum entropy principle. Acta Informatica 28: 801–815.
18. Arizono Y, Cui Y, Ohta H (1991) An analysis of $M/M/S$ queueing systems based on the maximum entropy principle. Journal of the Operational Research Society 41: 69–73.
19. Falin GI, Martin M, Artalejo JR (1994) Information theoretic approximations for the $M/G/1$ retrial queue. Acta Informatica 31: 559–571.
20. Kouvatsos DD, Tabet-Aouel N (1994) An ME-based approximation for multi-server queues with preemptive priority. European Journal of Operational Research 77: 496–515.
21. Tadj L, Hamdi A (2001) Maximum entropy solutions to a quorum queueing system. Mathematical and Computer Modelling 34: 19–27.
22. Jaynes ET (1957) Information theory and statistical mechanics. Physical Review 106: 620–630.
23. Kapur JN, Kesavan HK (1992) Entropy optimization principles with applications. San Diego: Academic Press Inc.
24. Golan A, Judge G, Miller D (1996) Maximum Entropy Econometrics: Robust Estimation with Limited Data. NY: John Wiley & Sons Ltd.
25. He DY, Qiu WH (2002) Solving continuous stochastic decision-making problem under uncertainty by the maximum entropy formalism. Theories and Practices of System Engineering 22: 97–100. (in Chinese)