JKAU: Eng. Sci., **vol. 10 no.1**, pp. 57-65 (1418 A.H. / 1998 A.D.)

A Maximum Entropy Analysis for $M^X/G/1$ Queueing System at Equilibrium

MATHU JAIN
Department of Mathematics, Indian Institute of Technology, N. Delhi, India

ABSTRACT. The maximum entropy principle (MEP) which is frequently used in information theory can be applied to construct most uncertain probability distribution subject to some constraints expressed by mean values. In this investigation, we have suggested probability distribution of possible states for $M^X/G/1$ queueing system in a maximum entropy condition subject to the expected number of customers. The expression for probability distribution involves mean arrival rate, first two moments of group size and service time distributions. We have established connections with classical queueing theory and operational analysis.

1. Introduction

An overview of the investigations on waiting lines shows that there is considerable amount of literature on queueing systems in steady state condition. Explicit results for probability distribution of the number of customers for the system in transient state are extremely difficult to obtain. Exact results for probability distribution in the possible states is not known for $M/G/1$ and even $M/E_k/1$ system. However, these results are known only for simple $M/M/1$ and $M/D/1$ models.

Maximum entropy models have been applied with varying degree of success to various fields including statistical mechanics, pattern recognition, operations research, economics, thermodynamics, biological, ecological and medical modelling, ... etc.[1]. Some efforts have also been made to estimate probability distribution for various models of queueing theory by using principle of maximum entropy. The prior information used in these investigations are in the form of mean values, e.g., mean arrival rate, mean service time and mean queue length.

The principle of maximum entropy (PME), as a measure of the amount of uncertainty – introduced earlier in information theory by Shannon[2] – was extended by Jaynes[3]. Shore[4] obtained axiomatic derivation of the principle of maximum entropy and the maximum cross entropy in system modelling. The study of $M/G/1$ and $G/M/1$ queueing system at equilibrium via PME was done by El-Affendi and Kouvatsos[5]. The $G/G/1$ queue was analysed by Kouvatsos[6-8] in detail. Walstra[9] employed the same principle to discuss general $G/G/1$ queueing network. The maximum entropy analysis for multiserver queueing system was established by Wu and Chang[10]. Kouvatsos and Aouel[11] studied $G/G/1$ queue with priority. They[12] also considered priority classes for $G/G/c$ queue and gave approximate results based on PME for the mean response times in steady state.

The maximum entropy analysis for queueing network was presented by Cantor[13] and Wu[14,15]. A constrained entropy optimization problem was described in Ref. [16]. Guiasu[17] presented a probabilistic model for an $M/G/1$ queueing system in a maximum entropy condition. He obtained mean queue size distribution by using the expected number of customers given by Pollaczek Khinchine formula[18].

Although many researchers have studied bulk queues, yet there is no analytical expression for the probability distribution of the number of customers in the system for even $M^X/M/1$ model. In this investigation an effort has been made to present the probability distribution of the possible states for an $M^X/G/1$ queue using a nonlinear programme related to Jaynes' principle[3]. We shall obtain queue size distribution in steady state by using MEP when the constraints involve only the first two moments of the inter-arrival and service time distribution.

2. The Principle of Maximum Entropy

We consider the finite discrete case of Shannon's entropy.

Let

$$H(p) = H_n(p_1, p_2, \dots, p_n) = -\sum_{k=1}^{n} p_k \log p_k \tag{1}$$

Here $H(p)$ measures the amount of uncertainty contained by the probability distribution

$$p = (p_1, p_2, \dots, p_n)$$

and is called system's entropy function. Shannon's entropy has the following property

$$H_n(p_1, p_2, \dots, p_n) \le H_n\left(\frac{1}{n}, \frac{1}{n}, \dots, \frac{1}{n}\right)$$

with equality iff

$$p_k = \frac{1}{n}, \quad (K = 1, 2, \dots, n)$$

The above relation shows that the uniform distribution is the most uncertain one when no constraint is imposed on the probability distribution. This result is equivalent to Laplace's principle of unsufficient reason, which states that the most reasonable strategy consists in attaching the same probability to different outcomes when we have no additional information about them.

Jaynes[3] extended this principle by introducing the principle of maximum entropy, which maximizes the entropy (1) subject to constraints

$$E(f) = \sum_{k=1}^{n} f_k \; p_k \tag{2}$$

where f_k, k = 1, 2, ... , n, are suitable functions reflecting the weights of p_k. In general the number of these functions are less than the number of possible states.

Also

$$\sum_{k=1}^{n} p_k = 1 \tag{3}$$

According to PME, we choose the probability distribution containing the largest amount of uncertainty subject to constraints by the given information.

The solution of non-linear programme (1) subject to constraints (2) and (3) is given by[3]

$$p_k = \frac{1}{\phi(\beta_o)} \exp\left(-\beta_k f_k\right) \quad , \quad k = 1, 2, \ldots, n$$

where $\phi(\beta_o) = \sum_{k=1}^{n} \exp\left(-\beta_k f_k\right)$

and β_k is the unique solution of the equation

$$\frac{d}{d\beta_k} \log \varnothing\left(\beta_k\right) = - E(f) \tag{4}$$

In general, there is no analytical expression for the solution of the Equation (4). But, in case of queueing theory, we can get simple expressions for probability distribution satisfying PME when prior information are available in terms of mean values.

Now, we give the statement of a lemma which can be proved by using Taylor's theorem[17].

Lemma: For any $t > 0$, there is τ depending on t, between 1 and t, such that

$$g(t) = t \log t = (t-1) + \frac{1}{2\tau}(t-1)^2 \tag{5}$$

This lemma enables us to prove the following theorem :

Theorem 1: If the arrival rate is λ, then the PME implies that the inter-arrival time follows an exponential distribution.

Proof : See Appendix for the proof of theorem.

This theorem shows that, if the only information available about the input of a queueing system is the arrival rate, the most uncertain distribution for the inter-arrival time is the exponential distribution. Similarly, we can prove that the most uncertain distribution for the service time is exponential if the only information available is the service rate.

Theorem 2 : Let L be the expected number of customers in the system. Then by using the MEP, the probability distribution of state N of the system is

$$p_n = \text{Prob.}\,(N = n) = \frac{L^n}{(1+L)^{n+1}} \qquad (6)$$

$$(n = 0, 1, 2, \ldots)$$

3. Maximum Entropy in Single Server Bulk Queueing Models

3.1 The $M^X/M/1$ Model

Consider $M^X/M/1$ queue with mean arrival rate λ and mean service time $1/\mu$. The customers are assumed to arrive in group of random size with mean $\bar{a}$ and finite variance σ_a^2. The average number of customers in steady state is given by[18]

$$L = \frac{\lambda}{2\mu}\left[\frac{(\sigma_a^2 + \bar{a}^2 + \bar{a})}{\bar{a}(1-\rho)}\right]$$

$$= \frac{\lambda}{\mu - \bar{a}}\left[\frac{(\sigma_a^2 + \bar{a}^2 + \bar{a})}{2\,\bar{a}}\right]$$

Applying the MEP, the probability distribution of the system states in steady state is

$$p_n = \frac{L^n}{(1+L)^n}$$

$$= \frac{2\bar{a}(1-\theta\bar{a})(\sigma_a^2 + \bar{a}^2 + \bar{a})\theta^n}{2\bar{a} + (\sigma_a^2\bar{a}^2 + \bar{a})^{n+1}} \qquad (7)$$

where $\theta = \dfrac{\lambda}{\mu}$

In case of geometrically distributed batches, we have

$$\bar{a} = \frac{1}{1-\alpha}\ , \ \sigma_a^2 = \frac{1}{(1-\alpha)^2}$$

where α is the parameter of the geometric distribution so that (7) reduces to

$$p_n = 1 - \frac{\lambda}{(1-\alpha)}\left(\frac{\theta}{1-\alpha}\right)^n \tag{8}$$

The probability of having no customer in the system can be obtained by substituting $n = 0$ in (8) and is given by

$$p_0 = \frac{2\bar{a}\,(1-\theta\bar{a})}{2\bar{a}+\theta(\sigma_a^2-\bar{a}^2+\bar{a})} \tag{9}$$

In the classical $M^X/M/1$ model, we have

$$\overline{p_0} = 1 - \frac{\lambda\bar{a}}{\mu} \tag{10}$$

which is independent of the choice of batch size distribution. When σ_a^2 is small, then

$$\bar{p}_0 \cong p_0$$

3.2 The $M^X/G/1$ Model

The mean queue length L of $M^X/G/1$ queue is given by Pollaczek-Khinchine formula[10]

$$L = \frac{\lambda\bar{a}}{2\mu(1-\rho)}\left[\bar{a}(C_a^2+1)+\rho(C_b^2-1)+1\right] \tag{11}$$

where C_a^2 and C_b^2 are the square coefficients of variation of group size and service time distributions, respectively.

Here $\rho = \frac{\lambda\bar{a}}{\mu}$

The expression (11) can be put in the form

$$L = \frac{\rho}{1-\rho}(F+\rho G)$$

where $F = \frac{\bar{a}(C_a^2+1)+1}{2}$, $G = \frac{C_b^2-1}{2}$

Using theorem 2, the probability of n customers in the system is

$$p_n = \frac{2^{n+1}\lambda^{n+1}\mu\,(\mu-\lambda)\,(\mu F+\lambda G)^n}{[2\mu^2+\{\lambda\mu\bar{a}\,(C_b^2+1)-1\}+\lambda^2\,(C_b^2-1)]^{n+1}} \tag{12}$$

4. Some Particular Cases

4.1 $M/G/1$ Model

In this case $C_a^2=0, \bar{a}=1$. Substituting these values and $C_b^2 = a_b^2\,\mu^2$, where σ_b^2 is the variance of service time distribution, in Equation (12), we have

$$p_n = \frac{2\mu(\mu-\lambda)\lambda^n(2\mu+\sigma^2\mu^2\lambda-\lambda)^2}{(2\mu^2-\lambda^2+\lambda^2\mu^2\sigma_b^2)^{n+1}} \tag{13}$$

which is the same result obtained by Guiasu[17] using the MEP.

4.2 $M^X/E_k/1$ Model

Here $C_b^2 = \frac{1}{k}$, so that

$$p_n = \frac{2\mu(\mu-\lambda)\lambda^n\{2\mu F+\lambda(1-k)/k\}^n}{[2\mu^2+\lambda\mu\{\bar{a}(1/k+1)-1\}+\lambda^2(1/k-1)]^{n+1}} \tag{14}$$

The expression for p_n in explicit form is not known by using classical birth death approach.

4.3 $M^B/G/1$ Model

In this model, customers are assumed to arrive in constant batch size B and the result can be obtained by substituting $C_a^2 = 0$ and $\bar{a} = B$ in(12)

4.4 $M^X/D/1$ Model

If service time is constant, then $C_b^2 = 0$ and (12) reduces to

$$p_n = \frac{2\mu(\mu-\lambda)\,\lambda^n\,(2\mu F-\lambda)^n}{[2\mu^2+\lambda\mu\{\bar{a}(C_a^2+1)-1\}-\lambda^2]^{n+1}} \tag{15}$$

For single arriving customers, *i.e.*, for $M/D/1$ model, by putting $C_a^2 = 0$, our results becomes

$$p_n = \frac{2\mu(\mu-\lambda)\,\lambda^n\,(2\mu-\lambda)^n}{(2\mu^2-\lambda^2)^{n+1}} \tag{16}$$

which is the same result given by Guiasu[17] for $M/D/1$ model.

5. Discussion

In this investigation, we have constructed the most uncertain probability distribution for $M^X/G/1$ model subject to constraints expressed in terms of mean values, and variances of batch size and service time distributions. The extension of the MEP to $G^X/G/1$ is currently the subject of further study. The MEP can also be used to obtain the probability distribution of the number of customers and servers in double ended queueing systems.

Since exact results for the probability distribution of the number of customers in the system are not available for even the simple M^X/M/1 bulk model, therefore a comparison can be made by taking approximate results based on diffusion approximation technique.

Acknowledgement

This work is supported by UGC Career Award Project Grant No. F. 10-29/93 (SA-III). The author is grateful to the referees for their helpful suggestions and constructive comments for the improvement of the paper.

References

[1] **Kapoor, J.N.,** *Maximum Entropy Models in Science and Engineering*, Wiley Eastern, New Delhi (1989).

[2] **Shannon, C.E.,** A mathematical theory of communication, *Bell. Syst. Tech. J.*, **27:** 379-423, 623-656 (1948).

[3] **Jaynes, E.T.,** Prior probabilities, *IEEE Trans Systems Sci. Cybern.*, **SSC-4:** 227-241 (1968).

[4] **Shore, J.E.,** Information theoretic approximations for $M/G/1$ and $G/G/1$ queueing systems, *Acta Info.*, **17:** 43-61 (1982).

[5] **El-Affendi, M.A.** and **Kouvatsos, D.D.,** A maximum entropy analysis of the M/G/1 and G/M/1 queueing systems at equilibrium, *Acta Info.*, **10:** 339-355 (1983).

[6] **Kouvatsos, D.D.,** Maximum entropy methods for general queueing networks, *In:* **D. Potier** (Ed.), *Modelling Techniques and Tools for Performance Analysis*, North Holland, Amsterdam, pp. 589-608 (1985).

[7] **Kouvatsos, D.D.,** A universal maximum entropy algorithm for the analysis of general closed networks. *In:* **T. Hasegawa** *et al.* (Eds.), *Computer Networking and Performance Evaluation*, North Holland, Amsterdam, pp. 113-124 (1986).

[8] **Kouvatsos, D.D.,** A maximum entropy analysis of the $G/G/1$ queue at equilibrium, *J. Opl. Res. Soc.*, **39**(2): 183-200 (1988).

[9] **Walstra, R.,** *Iterative Analysis of Networking of Queues*, Ph.D., Thesis, Technical Report CSEI-166, University of Toronto, Canada (1984).

[10] **Wu, J.S.** and **Chang, W.C.,** Maximum entropy analysis of multi-server queueing systems, *J. Oper. Res. Soc.*, **40**(9): 815-826 (1989).

[11] **Kouvatsos, D.D.** and **Tabet-Aouel, N.,** A maximum entropy priority approximation for a stable $G/G/1$ queue, *Acta Info.*, **26:** 241-267 (1989).

[12] **Tabel-Aouel, N.** and **Kouvatsos, D.D.,** On an approximation to the mean response times of priority classes in a stable G/G/C/PR queue, *J. Opl. Res. Soc.*, **43**(3): 227-239 (1992).

[13] **Cantor, J.L.,** *Information Theoretic Analysis for a Multi-Server Queueing System of Equilibrium with Applications to Queue in Tandem*, M.Sc. Thesis, University of Maryland, U.S.A. (1984).

[14] **Wu, J.S.,** *Maximum Entropy Analysis of Queues and Queueing Networks*. Ph.D. Thesis, University of Calgary, Canada (1988).

[15] **Wu, J.S.,** Maximum entropy analysis of open queueing network with group arrivals, *J. Opl. Res.*, **43**(11): 1063-1078 (1992).

[16] **Tsao, H.S.J., Fang, S.C.** and **Lee, D.N.,** *On the Optimal Entropy Analysis*. E.J. Oper. Res., **59:** 324-329 (1992).

[17] **Guiasu, S.,** Maximum entropy condition in queueing theory, *J. Opl. Res. Soc.*, **37:** 293-301 (1986).

[18] **Chawdhry, M.L.** and **Templeton, J.G.C.,** *A first course in bulk queues*, John Wiley & Sons, New York (1988).

Appendix A

Proof of Theorem 1

Suppose f_a be the probability density function (p.d.f.) of the inter-arrival time T_a. We have to determine f_a by using the PME such that

$$\text{Max } H \doteq -\int_0^\infty f_a(t) \log f_a(t)\, dt \tag{A1}$$

subject to the following constraints

i) The normalizing constraints

$$f_a(t)\, dt = 1 \tag{A2}$$

ii) The mean arrival rate satisfies

$$\frac{1}{\lambda} = \int_0^\infty t f_a(t)\, dt \tag{A3}$$

By introducing Lagrange's multipliers $\alpha > 0$, $\beta > 0$ and using (5), we have

$$- H + \alpha.1 + \beta.\frac{1}{\lambda} = \int_0^{\infty} f_a(t)\,[\log f_a(t) + \alpha + \beta t]\,dt = \int_0^{\infty} \exp(-\alpha - \beta t)\,[f_a(t)\exp(\alpha + \beta t) - 1]\,dt$$

$$+ \int_0^{\infty} \exp(-\alpha - \beta t)\,\frac{1}{2\tau(t)}\,[f_a(t)\exp(\alpha + \beta t) - 1]^2\,dt \geq \int_0^{\infty} \exp(-\alpha - \beta t)\,[f_a(t)\exp(\alpha + \beta t) - 1]\,dt$$

$$= 1 - \int_0^{\infty} \exp(-\alpha - \beta t)\,dt \qquad \text{(A4)}$$

with equality iff

$$f_a(t) = \exp(-\alpha - \beta t)\ , \quad (t > 0) \qquad \text{(A5)}$$

Using (A2) and (A5), we get

$$\exp(\alpha) = \int_0^{\infty} \exp(-\beta t)\,dt = \frac{1}{\beta}$$

$$\text{or } \alpha = -\log \beta$$

So that (A5) becomes

$$f_a(t) = \beta \exp(-\beta t) \qquad \text{(A6)}$$

By (A3) and (A6), we have

$$\beta = \frac{1}{\lambda}$$

Therefore (A6) gives

$$f_a(t) = \frac{1}{\lambda}\exp(-t/\lambda)$$

A.2 Proof of Theorem 2

We have to maximize the discrete countable entropy

$$H = -\sum_{n=0}^{\infty} p_n \log p_n \qquad \text{(A7)}$$

$$\text{Subject to } \sum_{n=0}^{\infty} p_n = 1 \qquad \text{(A8)}$$

$$\text{and } L = \sum_{n=0}^{\infty} n p_n \qquad \text{(A9)}$$

Solving (A7) subject to (A8) and (A9). (Similar to the proof of theorem 1), we get

$$p_n = \exp(-\alpha - \beta n) \qquad \text{(A10)}$$

$$(n = 0, 1, \ldots),\ (\alpha > 0, \beta > 0)$$

$$\text{and } \exp(-\beta) = \frac{L}{1 + L} \qquad \text{(A11)}$$

Using (A8) and (A10), we get

$$\exp(-\alpha) = \left[\sum_{n=0}^{\infty} \exp(-\beta n)\right]^{-1} \qquad \text{(A12)}$$

$$= 1 - \exp(-\beta)$$

Now from (A10) – (A12), we get result (6).

تحليــل باستخــدام القيمــة العــظمى للانتــروبيــا لنـظــام الأرتــال (الطوابــير) $M^X/G/1$ عنـد الاتـزان

ماتيــو جــين
قسم الرياضيات ، المعهد التقني الهندي
نيــــــو دلهي - الهنـــــد

المستخلص . يمكن تطبيق مبدأ الانتروبيا العظمى المستخدم كثيرًا في نظرية المعلومات لإنشاء أكثر التوزيعات الاحتمالية ريبة ، شريطة الالتزام بقيود تحددها القيم المتوسطة . قمنا في هذا البحث باقتراح توزيع احتمالي للحالات الممكنة في نظام الأرتال (الطوابير) $M^X/G/1$ المشروط بقيمة عظمى للانتروبيا ، والمقيد بعدد متوقع للعملاء . ويتضمن التعبير الناتج للتوزيع الاحتمالي متوسط معدل الوصول والعزمين الأولين لتوزيعي حجم الزمرة وزمن الخدمة . وقد قمنا بربط دراستنا هذه بكل من نظرية الأرتال التقليدية والتحليل التشغيلي .