Documento de trabajo
9407

INFERENCE IN ERGODIC QUEUES USING OCCUPATION CYCLES

Antonio Rodrigo

FACULTAD DE CIENCIAS ECONOMICAS Y EMPRESARIALES
UNIVERSIDAD COMPLUTENSE
Campus de Somosaguas 28223 MADRID

INFERENCE IN ERGODIC QUEUES USING OCCUPATION CYCLES

Antonio Rodrigo
Department of Economic Analysis.
Facultad de C.C. E.E. y Empresariales, U.C.M.
Campus de Somosaguas. 28223 Madrid.
e-mail: ececo03@sis.ucm.es

Abstract

We deal with inference problems on traffic intensity and on the mean number of customers in the system in steady state (q). We define the stopping time in terms of occupation cycles and apply the resulting estimators to a large class of stationary queues comparing it with alternative methods and with the integral estimator obtained directly from the sample in the Markovian case. Our method is quite efficient to estimate q. Also the asymptotic problems posed by Schruben and Kulkarni [7] do not appear. We observe that reductions in the sample information do not necessarily give worse estimations. We also give some numerical examples by simulating known models.

1. INTRODUCTION

Suppose a $G/G/1$ queue, (i.e. general interarrival times, general service times and an only channel) observed over a time interval $(0,T]$ where T is a random variable. Assume that the first customer arrives to an idle service area at time t=0. The sample data consist of the vector

$$V_T=\{n(T),\ m(T),\ u_i,\ s_j;\ 1\le i\le n(T),\ 1\le j\le m(T)\}$$

where $n(T)$ is the number of arrivals, $m(T)$ the number of departures and $\{u_i;1\le i\le n(T)\}$ and $\{s_j;1\le j\le m(T)\}$ are respectively the interarrival and service times during $(0,T]$. Suppose that the above vectors are independent sequences of independent and identically distributed random variables with densities $f(u;\theta)$ and $g(s;\phi)$, where θ and ϕ are unknown parameters . The likelihood function based on the data V_T is given by

$$L_T=\prod_{i=1}^{n(T)} f(u_i;\theta)\prod_{j=1}^{m(T)} g(s_j;\phi) \times\left[1-F_\theta\left(T-\sum_{i=1}^{n(T)}u_i\right)\right]\left[1-G_\phi\left(T-\gamma(T)-\sum_{j=1}^{m(T)}s_j\right)\right] \quad (1)$$

where F and G are the distribution functions corresponding to the densities f and g respectively and $\gamma(T)$ is the total idle period during $(0,T]$. Basawa and Prabhu [2] show that, under some general conditions which hold for a large class of queues and stopping rules, the likelihood function (1) is equivalent to

$$L_T=\prod_{i=1}^{n(T)} f(u_i;\theta)\prod_{j=1}^{m(T)} g(s_j;\phi) \quad (2)$$

With this function Basawa and Prabhu obtain expressions for the maximum likelihood (ML) estimators of θ and ϕ when the densities f and g belong to the exponential family.

In $M/M/1$ ergodic queues (i.e. Poisson arrivals and exponential service time) it is possible to consider some dependence between variables if the process $\{X(t);t\ge0\}$, number of customers in the system at time t, is an ergodic Markovian process. Using the results of Billingsley [3], asymptotically Gaussian consistent estimators have been obtained. In more general queues some authors reduce the problem to the Markovian case using the imbedded Markov chain.

In this paper we are interested on the estimation of the traffic intensity, ρ, and the mean number of customers in the system in stady state, q, for an ergodic queue. The traffic intensity is defined as $E[s]/E[u]$ and a $G/G/1$ queue is ergodic iff $\rho<1$ (see Asmussen [1]). In some cases we have expressions for q, for example in $M/M/1$ queues, $q=\rho/(1-\rho)$. Using the likelihoods (1) or (2) we can obtain estimators for $E[s]$ and $E[u]$, which can be used to give estimators for ρ and q. However, with positive probability, the estimator of ρ, $\hat{\rho}$, is greater than one and therefore it cannot be used to estimate q. In next section we give some estimators for the parameter ρ using occupation cycles in different situations which can be used to estimate q.

The estimation of q is not simple, because the random variables $n(T)$ and $m(T)$ are not independent. Schruben and Kulkarni [7] find some incongruities in the estimator of q when the above dependence between the variables is obviated. In section-3 we give some estimators of q for which is not present the problem of Schruben and Kulkarni. We show that a lower information does not necessarily give worse estimations. We also compare the estimator of q with the direct estimator (see Reynolds [6]). For this aim an expression for the variance of the estimator is needed. In the expression given by Reynolds in the $M/M/1$ case we found a mistake. The variance expression used here has been obtained as a particular case of the variance for an $M/G/1$ queue which can be obtained using the method given by Falin, Rodrigo and Vázquez [4]. We also give some numerical examples in these two sections. In the last section we consider a general case and give some results and indications for future research.

2. ESTIMATORS FOR THE TRAFFIC INTENSITY

Suppose observed an $M/M/1$ ergodic queue on the interval $(0,T]$ and let $\theta\equiv\lambda$ and $\phi\equiv\mu$ be the rates of arrival and service respectively. The ML estimators for λ and μ are given by $n(T)/T$ and $m(T)/S(T)$ respectively, where $S(T)=T-\gamma(T)$ is the total occupation time during $(0,T]$. Thus, $\hat{\rho}_1=n(T)S(T)/(Tm(T))$ can be considered as an estimator of ρ. However, as we discussed above, it is not adequate for our purposes. Now, if we consider T as the moment where the process $\{X(t);t\geq 0\}$ leaves the state zero by N-th time, i.e., if we observe the queue during N occupation cycles, we have the following theorem

Theorem-1. Suppose N occupation cycles of an $M/M/1$ ergodic queue have been observed, then

$$\hat{\lambda}_2=\frac{m(T)}{T},\quad \hat{\mu}_2=\frac{m(T)}{S(T)}\quad\text{and}\quad \hat{\rho}_2=\frac{S(T)}{T} \tag{3}$$

are strongly consistent estimators of the true values of the parameters λ, μ and ρ, with $\hat{\rho}_2<1$ a.s. Moreover, if $T\uparrow\infty$ $(N\uparrow\infty)$ the random variable $N^{1/2}(\hat{\lambda}-\lambda, \hat{\mu}-\mu)$ is a bivariate Normal distribution and $\hat{\lambda}$ and $\hat{\mu}$ are asymptotically independent.

Proof. Note that $n(T)=m(T)$ and the proof follows from Basawa and Prabhu [2]. ■

Remark-1. The above theorem can be proved using the classical theory of maximum likelihood, because the vector V_T can be written as N independent and identically distributed random vectors

$$y_1=\{n_1,u_{1j},s_{1j},\ j=1,\dots,n_1\}$$
$$\vdots$$
$$y_N=\{n_N,u_{Nj},s_{Nj},\ j=1,\dots,n_N\}$$

where n_i is the number of arrivals (equivalently departures) and u_{ij} and s_{ij} the interarrival and service times of each customer in each cycle respectively. Thus the likelihood (1) can be written as

$$L_2=\prod_{i=1}^{N} h(y_i;\eta),\quad \eta=(\theta,\phi) \tag{4}$$

where $h(y_i;\eta)$ is the density function of y_i.

Remark-2. If we observe the system in a general interval $(0,T]$ and provided that T is large enough, it will be possible to approach the complete sample by the sequence $y_1,\dots,y_{N(T)}$ where $N(T)$ is the number of complete cycles observed during the interval $(0,T]$. This approach is possible in a considerable class of queues when some regularity conditions are satisfied (see [2]). In the following we write $M\equiv M(N)\equiv n(T)=m(T)$, $S\equiv S(T)$ and $T\equiv T(N)$.

Remark-3. It can be proved that (S,T) is a minimal sufficient statistic for the parameter ρ, provided that $h(y_i;\eta)$ is the likelihood function of the observation y_i. However, only when T is a fixed time and in the ergodic Markovian case, it has been proved that the function (2) is a true likelihood function (see [3]).

The above estimator of ρ does not depend on the number of arrival. If we consider that all customers wait in queue a positive time until they begin to be served except for the first one of each cycle, we have that the function (4) can be written as

$$L_3=\prod_{i=1}^{N}(\lambda\mu)^{n_i}\exp(-\lambda u_{i.}-\mu s_{i.})\rho^{n_i-1}(1+\rho)^{1-2n_i} \tag{5}$$

where $u_{i.}$ and $s_{i.}$ are the duration of i-th cycle and occupation period respectively. Using the expression (5) as likelihood function we get

Theorem-2. Let $\hat{\rho}_3$ be the unique solution of the equation $2MT\rho^2+NT\rho-(2M-N)S=0$ in the interval $(0,1)$. Then $\hat{\rho}_3$, $\hat{\mu}_3=2M/U$ and $\hat{\lambda}_3=\hat{\mu}_3\hat{\rho}_3$, where $U=\hat{\rho}_3T+S$, are strongly consistent estimators for the parameters. Moreover, the vector $N^{1/2}(\hat{\lambda}_3-\lambda,\hat{\mu}_3-\mu)$ is asymtotically Gaussian but $\hat{\lambda}_3$ and $\hat{\mu}_3$ are not asymtotically independent. The covariance matrix is given by

$$\frac{1}{1+3\rho}\begin{pmatrix}(1+\rho)^{-1}(2\lambda^2+\mu^2+\lambda\mu)\rho^2 & \lambda^2 \\ \lambda^2 & (2\lambda+\mu)\lambda\rho\end{pmatrix}$$

Proof. It follows from theorem-1 and the expression for the estimator $\hat{\rho}_3$.∎

Theorem-3. If only the number of arrivals in each occupation cycle is observed, then $\hat{\rho}_4=(M-N)/M$ is the ML estimator for ρ which is strongly consistent and asymptotically Gaussian with variance $\rho(1-\rho^2)$.

Proof. It follows from Kleinock [5] and the Maximum Likelihood Theory.∎

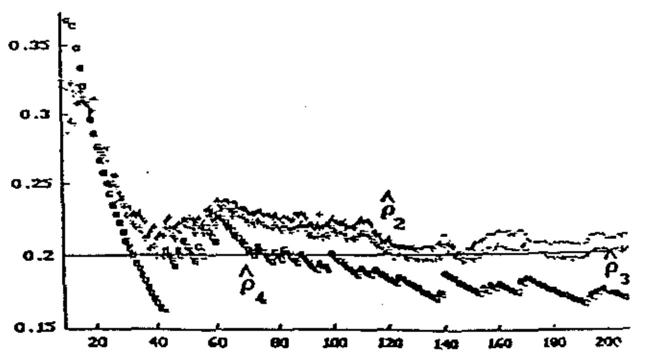
Theorem-4. If $M\neq N$, then we have one of the following relations: $\hat{\rho}_2<\hat{\rho}_3<\hat{\rho}_4$ or $\hat{\rho}_4<\hat{\rho}_3<\hat{\rho}_2$ or $\hat{\rho}_2=\hat{\rho}_3=\hat{\rho}_4$.

Proof. It follows from the expression of $\hat{\rho}_3$.∎

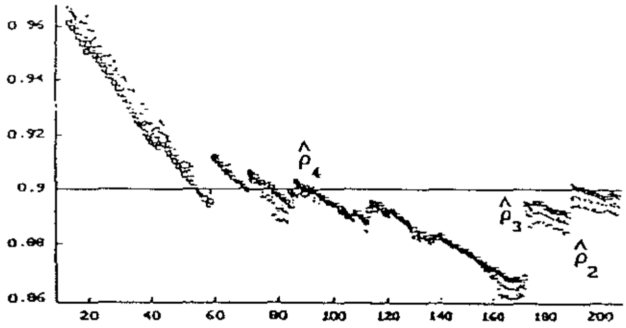
Remark-4. Obviosly the estimator $\hat{\rho}_4$ has the greatest variance, but it gives 'good' estimations, above all when ρ is near one (see graphic-II). Moreover, in the estimation of q, the corresponding estimator of $\hat{\rho}_4$ has better properties in some cases. Another advantage is the possibility of obtaining unbiased estimators for both parameters λ and μ.

Numerical examples

Traffic estimations for two $M/M/1$ ergodic queues, simulated during 200 occupation cycles, are showed when $\rho=.2$ (low traffic) and $\rho=.9$ (heavy traffic).



GRAPHIC-I: low traffic, $\rho=.2$.



GRAPHIC-II: heavy traffic, $\rho=.9$.

It is observed that $\hat{\rho}_3$ is always between $\hat{\rho}_2$ and $\hat{\rho}_4$ (see theorem-4) and, if ρ is away zero, the estimator $\hat{\rho}_4$ is as 'good' as the others, even better in some cases. Note also that there are not quite differences between $\hat{\rho}_2$ and $\hat{\rho}_3$, but usually $\hat{\rho}_3$ is preferable.

3. ESTIMATORS FOR THE MEAN NUMBER OF CUSTOMERS

If $\hat{\rho}$ is an estimator for ρ, to estimate q it is necessary that $\hat{\rho}<1$. Then, if we use as estimators $\hat{\rho}_2$, $\hat{\rho}_3$ or $\hat{\rho}_4$ we will get $\hat{q}_2$, $\hat{q}_3$ or $\hat{q}_4$ as estimators of q using the relation $q=\rho/(1-\rho)$. To obtain expressions for the errors we will make some weak modifications to get unbiased estimators that will denote with the same notation.

Theorem-5. $\hat{q}_2=(N-1)S/(N(T-S))$ and $\hat{q}_4=(M-N)/N$ are unbiased estimators of q whose variances are given by

$$Var[\hat{q}_2]=\frac{(2N-1-\rho)\rho^2}{N(N-2)(1-\rho)^3},\ N\geq 3,\quad Var[\hat{q}_4]=\frac{1}{N}\,\frac{\rho(1+\rho)}{(1-\rho)^3}$$

Proof. Using the generating functions of S, M and $T-S$ and the independence between S and $T-S$.∎

Corollary-6. $Var[\hat{q}_4]\leq Var[\hat{q}_2]$ iff $N\geq(1+\rho)(2-\rho)/(1-\rho)$.

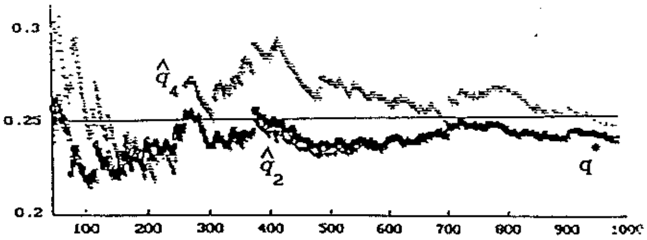
Proof. It follows from $Var[\hat{q}_4]$ and $Var[\hat{q}_2]$.∎

Theorem-7. Let q^* be the integral estimator of q (see [6]). Then, if $T\uparrow\infty$, $Var[q^*]\simeq 2\rho(1+\rho)/(\mu(1-\rho)^4T)$, $Var[\hat{q}_2]<Var[q^*]$ always and $Var[\hat{q}_4]\leq Var[q^*]$ iff $\rho\geq.5$.

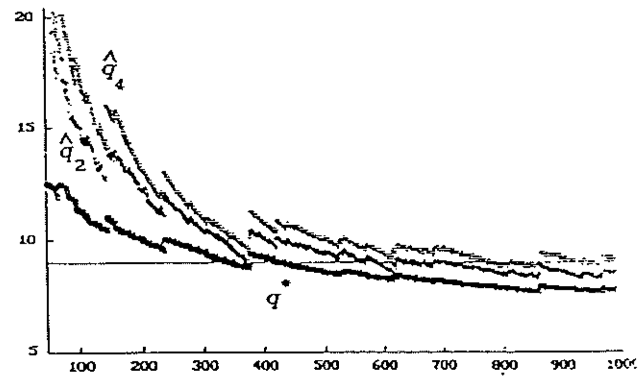
Proof. $Var[q^*]$ can be obtained as a particular case of a $M/G/1$ queue following the method given in [4]. If T (or N) is large, it is verified that $N\simeq\lambda(1-\rho)T$ and the above relations are satisfied.∎

Numerical examples.

Here, we simulate 1000 occupation cycles under low and heavy traffic as in the previous section to compare the estimators $\hat{q}_2$, $\hat{q}_4$ and q^*.



GRAPHIC-III: low traffic, ρ=.2, q=.25.



GRAPHIC-IV: heavy traffic, ρ=.9, q=9.

We observe that $\hat{q}_2$ and $\hat{q}_4$ give similar estimations when $\rho\simeq1$ and only with low traffic $\hat{q}_2$ looks better. More differences appear with q^* when ρ is close to one.

4. ESTIMATORS IN GENERAL QUEUES

In $G/G/1$ queue systems the expression (4) remains valid, even when k channels are considered in the service area, impatient customers, retrials or any other possible variation in the model. The only necessary hypothesis is the existence of an ergodic regenerative process (see [1]). Then, looking at the queue between consecutive regenerative instants, we have divided the observation vector in N independent and identically distributed random variables (in the above sections $\{X(t);t\geq0\}$ is the regenerative process and the moments when it leaves the state zero are the regeneration instants). In some cases the regenerative process is a vector and the distribution function of the variable y is too complex even unknown, but this problem arises independently of the stopping rule that we are used.

It would be interesting to consider some particular cases and to compare the estimations for ρ and q in the same way as in the previous sections, for example with distributions belonging to the exponential family and/or deterministic service time. Note that in these cases we have other relations for q which can be unknown in some cases (see [1]).

BIBLIOGRAPHY

[1] S. Asmussen. "Applied probability and queues", Wiley, New York, 1987.

[2] I. V. Basawa and N. U. Prabhu. "Large sample inference from simple server queues", Queueing Systems, 1, 289-304, 1988.

[3] P. Billingsley. "Statistical inference for Markov processes", The University of Chicago Press, Chicago, 1961.

[4] G. Falin, A. Rodrigo and M. Vazquez. "A new model for the $M/G/1$ retrial queue", Preprint, 1993.

[5] L. Kleinrock. "Queueing Systems", Vol. I, Wiley, New York, 1974.

[6] J. F. Reynolds. "Asymptotic properties of mean length estimators for finite Markov queue", Oper. Res. 20, 1, 52-57, 1972.

[7] L. Schruben and R. Kulkarni. "Some consecuences of estimating parameters for the $M/M/1$ queue", Oper. Res. Letters, 1, 2, 75-78, 1982.