Journal of Mathematics and Statistics 1 (1): 35-39, 2005
ISSN 1549-3644

Kernel Density Estimation for Interdeparture Time of GI/G/1 Queues

Hsing Luh and Berlin Wu
National Chengchi University, Taipei 116, Taiwan ROC

Abstract: The departure process of a single queue has been studied since the 1960s. Due to its inherent complexity, closed form solutions for the distribution of the departure process are nearly intractable. In this study, kernel type estimators of the density of interdeparture time in a GI/G/1 queue are studied. Uniform strong consistency of the estimators in a GI/G/1 queue and their rates of convergence are obtained. The stochastic processes are shown to satisfy the strong mixing condition with random instants of sampling. With the analysis presented, we provide a novel analytic tool for studying the departure process in a general queueing model.

Key words: Strong Mixing, Density Estimators, Kernel, Bandwidth, GI/G/1, Departure Processes

INTRODUCTION

Queueing theory has been extensively used in today's applications in communication systems and flexible manufacturing networks. In order to obtain a good performance estimate of the system, instead of solving the optimal problem at the whole network, it is often preferable to study the waiting distribution at each station (isolated node). Describing the characteristic of the departure process at a node is thus one important issue in studying queueing networks, because the departure processes at one node may be considered as the arrival process at subsequent nodes.

Many literatures have been studied in this field. For examples, Daley [1] investigated departure processes from a GI/M/1 queue and studied the correlation structure. Bertsimas [2] pointed out the difficulty of analysis of a G/G/s queueing system and derived an algorithm of a relatively low order of complexity for the system-size, prearrival and post-departure probability distributions. Chang [3] showed that the Poisson process is the only stationary and ergodic process that induces identical distributions on the interdeparture times when the service times are exponentially distributed. Luh [4] provided an analytic tool for studying the departure process in a GI/G/1 queueing system.

Since network traffic is composed of complex random processes which may not conform to any known Markovian model as commonly adopted in queueing analysis, the sequence of interdeparture times may be nonstationary, and even have a long history. In the application of most real cases, at least certain kinds of (weakly) dependent should be considered in the process. Instead of conventional queueing approaches, many researchers have paid exceptional attention on the covariance structure. For example, Melamed *et. al.* [5] captured the autocorrelated traffic by TES (Transform-Expand-Sample). Hwang and Li [6] developed a statistical-match queueing (SMAQ) tool to study measurement-based traffic management problem. In advance queueing analysis, recent real-life traffic measurement indicates the significance of traffic macrodynamics to network performance. The macrodynamics, versus microdynamics, is defined for characterizing the traffic behavior on the coarse, versus refined, time scales at which the process is observed. Consequently, the Autoregressive/Moving Average (ARMA) process has been used to model the macrodynamic behavior of the arrival process in a queueing system. Kulkarni and Li [7] show the second-order statistics of the microdynamics are well captured by white noise in the arrival process with power spectrum which can have a significant impact on the queueing performance.

In the present research, we study kernel type estimators of the density of interdeparture time in a GI/G/1 queue. Uniform strong consistency of the estimators in a GI/G/1 queue and their rates of convergence are obtained. The stochastic processes are shown to satisfy the strong mixing condition with random instants of sampling.

Kernel Estimate of the Interdeparture Time: Statistically, the description of the departure process is usually written in terms of the interdeparture intervals $\{D_n\}$, where D_n is the time between the nth and (n+1)th departure epochs, n = 1,2,…. We shall confine our discussion to a more general GI/G/1 queueing system that implies a stationary and weak dependent sequence of positive random variables $\{D_n\}$ with finite mean $E(D_n)$. Let $P(\cdot)$ be the probability measure defined in Definition 1. In view of stationarity, define the distribution function of interdeparture times as:

$$f(t) = P(D_1 \le t).$$

This study is concerned with the estimation of the probability density function $f(t)$ for an interdeparture times process $\{D_n, n = 1, 2, ...\}$ on the basis of the discrete time samples $\{D(t_k)\}$, $1 \le k \le n$, where the sampling instants { t_k } are random. As an estimator of $f(t)$ we shall consider the kernel estimate defined by:

$$f_n(t) = (nb_n)^{-1} \sum_{j=1}^{n} K\left(\frac{t - D(t_j)}{b_n}\right), \qquad (1)$$

where K is a kernel function and { b_n } is a sequence of bandwidths tending to zero as n tends to infinity. Here f_n takes values in $R_+ = [0, \infty)$.

Density estimation has been studied extensively since the works of Rosenblatt [8] and Parzen [9]. Under dependent situations, kernel type density estimators have been investigated by Masry [10, 11], Robinson [12], Roussas [13] and Tran [14, 15] for various weakly dependent processes. Gyorfi *et al.* [6] studied the uniform convergence and the L_1 convergence under different mixing conditions.

The purpose of this study is to establish weak conditions under which f_n converges uniformly on R_+ to f a.s. We also obtain sharp rates of convergence of f_n to f.

Assumptions and Preliminaries: Consider the GI/G/1 queue, in which the arrivals form a renewal process and the service time are independently and identically distributed. Without loss generality, service times are assumed independent of the arrival process, and the arrival rates is strictly less than the service rate. Let A_n, S_n, T_n and W_n be the interarrival time, service time, flow time, waiting time of the nth customer. Since the flow time equals service time plus waiting time, we have:

$$T_n = S_n + W_n$$

Moreover, for each n, we have $D_n = S_{n+1} + (A_n - T_n)^+$, where $x^+ \overset{\Delta}{=} \max(x,0)$ which implies

$$D_{n+1} = W_{n+1} - W_n + S_{n+1} - S_n + A_n \qquad (2)$$

Let A, S, T W and D be a generic interarrival time, service time, steady-state flow time, steady-state waiting time, and interdeparture time. Therefore in steady-state we have

$$T \overset{d}{=} S + W \overset{d}{=} S + (T - A)^+$$

and $$D \overset{d}{=} S + (A - T)^+$$

where: $\overset{d}{=}$ means equal in distribution.

Because S and A are bounded, the density of D is bounded as well. Recall the strong mixing condition which is defined by Tran [15].

Definition 1. Let X_k, k=...,-1,0,1,..., be a strictly stationary sequence of random variables defined on a probability space (Ω, F, P) and taking values in R. Let $F_{-\infty}^{0}$ and F_m^{∞} denote, respectively, the σ-fields generated by X_k, $k \le 0$ and by X_k, $k \ge m$. Then X_k is strong mixing if

$$\alpha(m) = \sup\{|P(A \cap B) - P(A)P(B)| : A \in F_{-\infty}^{0}, B \in F_m^{\infty}\} \downarrow 0, \text{ as } m \to \infty. \qquad (3)$$

The strong mixing condition is well known to be weaker than many dependence conditions, for example, the absolutely regular condition or the ϕ-mixing condition. For more information on strong mixing processes, see Rosenblatt [8], or Roussas [13].

We first write a weaker condition of (3) according to the evolutionary development of $\{D_n\}$. From (2), the covariance of D gives:

$$\begin{aligned}
&\mathrm{E}(D_k, D_{k+h}) - E(D_k)E(D_{k+h}) \qquad (4)\\
&= \mathrm{E}(S_{k+h-1}, W_{k-1}) - \mathrm{E}(S_{k+h}, W_{k-1}) - \mathrm{E}(A_{k+h-1}, W_{k-1}) - \mathrm{E}(S_{k+h-1}, W_k)\\
&+ \mathrm{E}(S_{k+h}, W_k) + \mathrm{E}(A_{k+h-1}, W_k) + \mathrm{E}(S_{k-1}, W_{k+h-1}) - \mathrm{E}(S_k, W_{k+h-1})\\
&+ \mathrm{E}(W_{k-1}, W_{k+h-1}) - \mathrm{E}(S_{k-1}, W_{k+h}) + \mathrm{E}(S_k, W_{k+h})\\
&- E(S_{k+h-1})E(W_{k-1}) + E(S_{k+h})E(W_{k-1}) + E(A_{k+h-1})E(W_{k-1}) + E(S_{k+h-1})E(W_k)\\
&- E(S_{k+h})E(W_k) - \mathrm{E}(A_{k+h-1})E(W_k) - \mathrm{E}(S_{k-1})E(W_{k+h-1}) + \mathrm{E}(S_k)E(W_{k+h-1})\\
&- \mathrm{E}(W_{k-1})E(W_{k+h-1}) + \mathrm{E}(S_{k-1})E(W_{k+h}) - \mathrm{E}(S_k)E(W_{k+h})
\end{aligned}$$

Note that T_n, A_n and S_{n+1} are independent of each other; therefore, W_{k-1}, A_{k-1} and S_{k+h-1} are independent of each other as well. Thus, (4) becomes

$$\begin{aligned}
&\mathrm{E}(S_{k-1}, W_{k+h-1} - W_{k+h}) + \mathrm{E}(S_k, W_{k+h} - W_{k+h-1}) - \mathrm{E}(S_{k-1})E(W_{k+h-1} - W_{w+h})\\
&- \mathrm{E}(S_k)E(W_{k+h} - W_{k+h-1}) + \mathrm{E}(W_{k-1}, W_{k+h-1}) - \mathrm{E}(W_{k-1})E(W_{k+h-1}) \qquad (5)
\end{aligned}$$

Since Cov $(S_{k-1}, W_{k+h-1} - W_{k+h}) \to 0$, Cov $(W_{k-1}, W_{k+h-1}) \to 0$ as $h \to \infty$, it has (5) approach to 0 and so does (4) as $h \to \infty$.

Thus, in order to have a general result for the density function f, we should give the following assumptions for the kernel K and the process $f(t)$. Let the letter C to denote a generic constant. All limits are taken as $n \to \infty$ unless indicated otherwise. To prove the main theorem, we need the following assumptions.

Assumption 1: The kernel K is a probability density function satisfies $|K(x) - K(y)| < C|x - y|$.

We assume that the sampling instants $\{t_k\}$ are random, constituting a renewal process on $[0, \infty)$. Let $\{\tau_k\}$, $1 \le k < \infty$, be a sequence of i.i.d. random variables with a common distribution $G(x)$ on $[0, \infty)$ with $G(0) = 0$ and a finite mean $\int_0^\infty x dG(x) = 1/\beta < \infty$. The sampling instants are defined $t_k = \sum_{i=1}^k \tau_i$ by, $k = 1, 2, \dots$.

Let $G_k(x)$ be the cumulative distribution function of t_k. If $G(x)$ is absolutely continuous with density $g(x)$ then G_k has a derivative, say, $g_k(x)$, which is the probability density function of t_k. Define

$m(t) = 2\sum_{k=1}^\infty k g_k(t), \qquad t > 0$

The quantity m is often referred to in the renewal theory literature as second-order factorial density.

Assumption 2: The renewal-type sampling instants $\{t_k\}$ have an intensity density $g(x)$ on $[0, \infty)$ and the second-order differential $m(x)$ satisfies $m(x) \le C(1 + x)$ on R_+.

Lemma 1: Suppose A(x) and S(x) are bounded and satisfied with Lipschitz condition. Then, the density $f(x)$ is bounded and satisfied with Lipschitz condition.

Proof : By (2), we have $D = S + (A - T)^+ \le$ S + A $\le$ C since S and A are bounded. $| f(x) - f(y) |$ = $| S(x) - S(y) + (A(x) - T(x))^+ - (A(y) - T(y))^+ | \le | S(x) - S(y) | + | A(x) - A(y) | \le$ C $| x - y |$

Assumption 3: Suppose the joint probability density $f(x, y; \tau)$ of (D_0, D_τ) exists. There exists some constants C such that it satisfies

$$\int_0^\infty f(x, y; \tau + s) g(\tau) d\tau \le \mathrm{C} < \infty$$

for all x,y and $s \ge 0$.

Denote

$$\psi(n,1) = (\log n)^{1/2} / (nb_n)^{1/2}$$

Assumption 4: For some $\ell > 0$, $(\psi(n,1))^{-1} \sup_{x \ge n^\ell} |f(x)| = O(1)$,

$\sum_{n=1}^{\infty} n(1-\int_{x\geq n^{\ell}} f(x)dx) < \infty$.

Assumption 5: $(\psi(n,1)b_n)^{-1} \sup_{|x|\geq n^{\ell}} K(x/b_n) = O(1)$.

Uniform Convergence of f_n: The following lemmas are needed in the proof of Theory. The proofs of these lemmas can be found in the Wu [17] based on assumptions.

Lemma 2: Suppose Assumptions 1-5 hold and b_n tends to zero slowly enough that $nb_n/\log n \to \infty$. We have

$$\sup_{x\leq 2n^{\ell}} \left|f_n(x) - Ef_n(x)\right| = O(\psi(n,1)) \quad \text{a.s. as } n\to\infty.$$

Lemma 3: Suppose the condition of Lemma 2 holds. Then

$$\sup_{x>2n^{\ell}} \left|f_n(x) - f(x)\right| = O(\psi(n,1)) \quad \text{a.s. as } n\to\infty.$$

Theorem: Suppose all assumptions hold. Further assume $\int |x|\,|K(x)|dx < \infty$ and $(\psi(n,1))^{-1} b_n = O(1)$. We have

$$\sup_{x\in R} \left|f_n(x) - f(x)\right| = O(\psi(n,1)) \quad \text{a.s.}$$

Proof: Since $\int K(x)dx = 1$ and $\int |x|\,|K(x)|dx < \infty$, following Roussas [15, p. 141], we have

$$\sup_{x\in R_+} \left|Ef_n(x) - f(x)\right| \leq C\, b_n,$$

which implies $\sup_{x\in R_+} \left|Ef_n(x) - f(x)\right| < \mathrm{C}\, O(\psi(n,1))$ by letting $b_n = O(\psi(n,1))$. From Lemmas 2 and 3, it produces

$$\begin{aligned}
&\sup_{x\in R_+} \left|f_n(x) - f(x)\right| \\
&\leq \sup_{x\in R_+} \left|f_n(x) - Ef_n(x)\right| + \sup_{x\in R_+} \left|Ef_n(x) - f(x)\right| \\
&\leq \sup_{x\leq 2n^{\ell}} \left|f_n(x) - Ef_n(x)\right| + \sup_{x>2n^{\ell}} \left|f_n(x) - Ef_n(x)\right| + \sup_{x\in R_+} \left|Ef_n(x) - f(x)\right| \quad \text{a.s.} \\
&\leq O(\psi(n,1)) + \sup_{x>2n^{\ell}} \left|f_n(x) - Ef_n(x)\right| \quad \text{a.s.} \\
&\leq O(\psi(n,1)) + \sup_{x>2n^{\ell}} \left|f_n(x) - f(x)\right| + \sup_{x>2n^{\ell}} \left|f(x) - Ef_n(x)\right| \quad \text{a.s.} \\
&\leq O(\psi(n,1)).
\end{aligned}$$

Remark: Consider an M/M/1 model where the case that { τ_k } constitutes an ordinary renewal process with { τ_k } having an exponential density function $\theta e^{-\theta x}$ (θ >0). In this case t_k has the Gamma density function, namely

$$\theta(\theta x)^{k-1} e^{-\theta x} / \Gamma(k), \qquad \Gamma(k) = \int_0^{\infty} x^{k-1} e^{-x}.$$

From Theorem 2, it follows that f_n can achieve the uniform rate of convergence on R_+ of order $(n^{-1}\log n)^{1/3}$.

CONCLUSION

In this study, we study the departure process of the GI/G/1 queue. We use kernel type estimators for the stochastic processes to prove the uniform strong consistency of the estimators and their rates of convergence. With the analysis presented, we provide a novel analytic tool for studying the departure process in a general queueing model. The stochastic processes are assumed to satisfy the strong mixing condition with the sampling instants which are random.

The time scales of traffic measurement on a network are directly related to the time and space complexity of the model used to describe the departure process. We have rigorously proved the "rate process" or the accumulated departures in a time interval may be taken to describe the departure process. Intuitively, this result can also be extrapolated to self-similar output processes since the strong mixing property represents an asymptotic behavior of the second-order statistics. The generation of strong mixing of traffic sequences and its associated queueing analysis require further study.

There are many studies about the problem of queuing system construction for various types of models. Some of them are only suitable for independent observations or special cases. Some of them have too strong assumptions that could not be easily reached. The weakness of the independent concept for distributed density function of service time clearly resides in the complexity of statistical computation. Unlike the traditional methods, estimation of a GI/G/1 queue system applies the concept of strong mixing to cooperate the realization structure change and dynamic heredity. This research liberates us from the independent-based process and thus fewer assumptions of the system will be made.

Finally, in spite of the realistic performance for the strong mixing property, there remain some problems for further studies. For example:

* The convergence of the proof for GI/G/1 model and the proposed assumptions have not been well used. This needs further investigation.
* To find an efficient procedure for the outliers as well as the intervention that make the structural change.

However, in order to give the popular questions, such as adaptive modeling, what if a hush point occurs, and combined forecasting, a satisfied answer, we believe Theorem suggested in this study will be a worthwhile approach and will stimulate more future empirical work in the GI/G/1 system.

REFERENCES

1. Daley, D.J., 1968. The correlation structure for the out process of single server queueing systems. Annals of Mathematical Statistics, 39: 1007-1019.
2. Bertsimas, D., 1990. An analytic approach to a general class of G/G/s queueing systems, Oper. Res. 38 : 139-155.
3. Chang, C., 1994. On the input-output map of a G/G/1 queue. J. Applied Probability, 31: 1128-1133.
4. Luh, H., 1999. On derivation of N-step departure processes. European J. Operations Res. 118: 194-212.
5. Melamed, B., Q. Ren and Sengupta, 1995. Modeling and analysis of a single server queue with autocorrected traffic. Proc. IEEE INFOCOM'95, 2: 634-642.
6. Hwang, C. and S. Li, 1995. On the convergence of traffic measurement and queueing analysis: A statistic-MAtch queueing (SMAQ) tool. Proc. of IEEE INFOCOM'95, 2: 602-623.
7. Kulkarni, L.A. and S. Li, 1998. Measurement-based traffic modeling: Capturing important statistics COMMUN. Statist.-Stocastic Models, 14: 1113-1150.
8. Rosenblatt, M., 1956. Remarks on some nonparametric estimators of a density function. Ann. Math. Statist., 27: 832-837.
9. Parzen, E., 1962. On estimation of a probability density function and mode. Ann Math. Statist., 33: 1065-1076.
10. Masry, E., 1983. Probability density estimation from sampled data. IEEE Trans. Inform. Theory, IT-29: 699-709.
11. Masry, E., 1986. Recursive probability density estimations for weakly dependent processes. IEEE Trans. Inform. Theory, IT-32: 254-267.
12. Robinson, P.M., 1983. Nonparametric estimates for time series. J. Time Series Anal., 4: 185-197.
13. Roussas, G. G., 1988. Nonparametric estimation in mixing sequences of random variables. J. Statist. Plann. Inference, 18: 135-149.
14. Tran, L. T., 1989b. The L_1 convergence of kernel density estimators under dependence. Can. J. Statist., 17: 197-208.
15. Tran, L.T., 1983. Nonparametric function estimation for time series by local average estimators. Ann. Statist., 40: 1040-1057.
16. Gyorfi, L., W. Hardle, P. Sarda and P. Vieu, 1989. Nonparametric Curve Estimation From time Series. Springer-Verlag, New York.
17. Wu, B., 1997. Kernel density estimation under weak dependence with sampled data. J. Statistical Planning and Inference, 61: 141-154.