Nonparametric approximation of the characteristics of the $D/G/1$ queue with finite capacity

Faïrouz Afroun*

Laboratory of Pure and Applied Mathematics,
University of Tizi-Ouzou,
15000, Algeria
Email: afrounfairouz@gmail.com
*Corresponding author

Djamil Aïssani

Research Unit LaMOS (Modeling and Optimization of Systems),
University of Bejaia,
06000, Algeria
Fax: +213-34-21-51-88
Email: lamos_bejaia@hotmail.com

Djamel Hamadouche

Laboratory of Pure and Applied Mathematics,
University of Tizi-Ouzou,
15000, Algeria
Email: djhamad@yahoo.fr

Abstract: In this work, we consider the finite capacity of the $D/G/1$ queue. First, the modelling of the system in question by an embedded Discrete-Time Markov chain is considered. Secondly, the aim is to illustrate the use of the discrete kernel method for the estimation of the stationary characteristics of this chain, when the general distribution that governs it is an unknown function. To support and illustrate our proposals, two extensive simulation studies are carried out.

Keywords: deterministic queues; Markov chains; smoothing parameter; discrete kernels; errors; simulation.

Reference to this paper should be made as follows: Afroun, F., Aïssani, D. and Hamadouche, D. (2022) 'Nonparametric approximation of the characteristics of the $D/G/1$ queue with finite capacity', *Int. J. Computing Science and Mathematics*, Vol. 16, No. 2, pp.170–180.

Biographical notes: Faïrouz Afroun is a doctoral student at the Department of Mathematics at the University of Tizi-Ouzou, Algeria. She received a Master's degree in Applied Mathematics from the University of Bejaia, Algeria. She is a permanent researcher at the Laboratory of Pure and Applied Mathematics (LPAM)

at the University of Tizi-Ouzou. Her areas of research include Markov chains and their estimation using the parametric or nonparametric methods, queueing theory, stochastic modelling and statistics.

Djamil Aïssani is a Full Professor of Mathematics at the Department of Operations Research at the University of Bejaia, Algeria. He started his career at the University of Constantine in 1978. He received his PhD in 1983 from Kiev State University (Soviet Union). He is at the University of Bejaia since its opening in 1983/1984. He is the Director of Research, Head of the Faculty of Science and Engineering (1999 to 2000), Director of the Research Unit LaMOS, and Scientific Head of the Computer Science Doctorate School ReSyD (2003 to 2011), and he has taught in many universities (USTHB Algiers, Annaba, Rouen, Dijon, ENITA, EHESS Paris, CNAM Paris, etc.). He has published many papers on Markov chains, queueing systems, reliability theory, performance evaluation and their applications in electrical, telecommunication networks and computer systems.

Djamel Hamadouche is a Professor of Mathematics at the Department of Mathematics of the University of Tizi-Ouzou, Algeria. He started his career at the University of Algiers in 1987. He received his PhD in 1997 at University of Lille, France. He is at the University of Tizi-Ouzou since 1999. He is Director of Research, Member of the scientific council of the Faculty of Sciences since 2001, Director of the Research Laboratory of Mathematics (2007–2016). His research interests include Probability theory, Statistics, Markov chains and their applications, queueing theory and stochastic modelling.

1 Introduction

In practice, the deterministic notion in queuing models is not always used in the absolute sense, but it can also refer to the process where their random fluctuations (the variance) are low, i.e., refer to the process nearly deterministic. That is why, for instance, in the performance evaluation of modern telecommunication systems, the deterministic notion is often used. Indeed, to express the low fluctuations of the inter-arrival times, deterministic inter-arrival times queueing models are frequently applied for systems modelling in this field (Alfa, 2010; Roberts et al., 1996; Ott, 1987).

Theoretically, the evaluating of the performances of a queueing system is based on the different starting parameters describing it. However, in practice, the values of these parameters are known only in the form of a sample of data. In this sense, to evaluate the performances of a considered system, the use of statistical estimation techniques, that aim to provide an approximation for the (unknown) parameters values by exploiting the information provided by the sample, is inevitable.

In this paper, we propose to consider the $D/G/1/C$ queue modelled by an embedded Discrete-Time Markov Chain (DTMC) under the assumption that the distribution of a number of customers served between two consecutive arrivals is an unknown mass function. In addition, our goal is to estimate the transition matrix P associated to the $D/G/1/C$ queue using the kernel method. This choice is motivated by the fact that the transition

matrix allows us to deduct all performance measures (transient and stationary) of the model (Graham, 2014; Norris, 1997; Privault, 2018).

Historically, Roussas (1969) was the first to consider the kernel method in the estimation of Markov chains. Thereafter, several other authors have completed his results, but these results are restricted in a theoretical framework rather than a practical one. The use of the kernel method in the estimation of a transition matrix in a practical framework was considered by Gontijo et al. (2011). Recently, Cherfaoui et al. (2015a) addressed the problem of choosing the smoothing parameter in the context of the kernel estimation of the transition matrix. In this last work, in order to take into account the interaction of the different components of the $GI/M/1/N$ system, the authors proposed procedures for selecting the smoothing parameter, which are based on matrix norms. Moreover, they showed that the estimator of the chosen smoothing parameter, by minimising the matrix norm $\|.\|_2$, gives better results than the classical methods.

It should be noted that the works (Gontijo et al., 2011; Cherfaoui et al., 2015a) were carried out via continuous asymmetric kernel estimators (to estimate the distribution of the data defined on $\mathbb{R}^+$). However, in practice, several situations are modelled by Markov chains governing according to unknown discrete distributions. In this case, it is natural to estimate these unknown distributions using discrete kernels.

To estimate a discrete density function using the kernel method, the Dirac kernel estimator is often used by practitioners because of its simplicity and good asymptotic properties. However, this estimator is not suitable for small or medium sized samples. Besides the Dirac estimator, Aitchison and Aitken (1976) proposed another discrete kernel. The problem of this last kernel is that is only suitable for categorical data and finite discrete distributions. Recently, other discrete kernels have been proposed in Kokonendji et al. (2007); Kokonendji and Kiessé (2011) to estimate density functions with discrete support. The authors introduced the notion and the definition of a discrete kernel designed from a discrete probability distribution. This latest allowed a considerable development of the kernel method in the discrete case (see for instance the works (Djerroud et al., 2020; Belaid et al., 2018; Wansouwé et al., 2016; Somé and Kokonendji, 2016; Zougab et al., 2014)).

The objective of our work is to verify the validity of the conclusions reached in Cherfaoui et al. (2015a) on the Continuous-Time Markov Chain when we consider a DTMC. More precisely, we analyse the problem of the smoothing parameter choice via the minimisation of the matrix norms when we consider the discrete kernel estimator of the transition matrix describing the $D/G/1/C$ queue. To do this, we have developed explicit forms of the expressions, outcome from the three matrix norms $\|.\|_1$, $\|.\|_2$ and $\|.\|_\infty$, to minimise in order to select the smoothing parameter when estimating the matrix P. Besides, to support and illustrate our proposals two extensive simulations studies are carried out.

The remainder of this paper is organised as follows: In Section 2, we briefly introduce the stochastic model of $D/G/1/C$ queue. The problem of the kernel and the smoothing parameter choice in the estimation of the matrix P is detailed in Section 3. Before concluding, in Section 4, the simulation studies are carried out to show the main results of this paper.

2 Model description

Let us consider a $D/G/1/C$ queue where C $(C > 1)$ is the capacity of the system including the one who is in service. We assume that the service time distribution $S(t)$ is a

general distribution with rate μ and the inter-arrival time distribution $A(t)$ is a deterministic distribution function with mean λ^{-1} ($A(t) = \mathbf{1}_{\{t=\lambda^{-1}\}}$, with $\mathbf{1}_{\{.\}}$ is the indicator function). We assume also that the inter-arrival process and the service process are independent.

Consider the process $\{L(t),\ t \geq 0\}$, where $L(t)$ represents the number of customers in the system at time t. Obviously, $\{L(t),\ t \geq 0\}$ is not a Markov process. For this, we consider the embedded Markov chain defined at the new arrival epochs. Let $t_n = n\lambda^{-1}$ be a sequence representing the instants of customer arrivals and $N_n = N(t_n)$ the number of customers in the system at time t_n, then the state of the system at time t_{n+1} is given by:

$$N_{n+1} = min\{N_n + 1, C\} - X_n, \tag{1}$$

where $X_n \equiv X_n(\lambda^{-1})$ represents the number of customers served between two consecutive arrivals. Note that the random variable N_{n+1} depends only on N_n and X_n. In added, the random variables $\{X_n,\ n \in \mathbb{N}\}$, which are $i.i.d$ of common law:

$$a_x = f(x) = Pr(X_n = x),\ \ x = 0, 1, 2, 3, ... \tag{2}$$

with mean $\rho^{-1} = \mu\lambda^{-1}$, are independent of n and the state of the system before t_n. Therefore, the process $\{N_n,\ n \in \mathbb{N}\}$ is a homogeneous DTMC with the state space $E = \{0, 1, ..., C\}$ and the transition matrix, P, whose forme is as follow:

$$P = \begin{pmatrix} a_0 & a_1 & & \cdots & & a_{C-2} & a_{C-1} & b_C \\ a_0 & a_1 & & \cdots & & a_{C-2} & a_{C-1} & b_C \\ 0 & a_0 & & \cdots & & a_{C-3} & a_{C-2} & b_{C-1} \\ \vdots & \ddots & \ddots & & & & & \vdots \\ 0 & \cdots & 0 & a_0 & \cdots & a_{i-1} & a_i & b_{i+1} \\ \vdots & & & \ddots & \ddots & & & \vdots \\ 0 & & \cdots & & 0 & a_0 & a_1 & b_2 \\ 0 & & \cdots & & 0 & 0 & a_0 & b_1 \end{pmatrix};$$

where a_x is defined by (2) and $b_y = \sum_{y \geq x} a_x$.

Moreover, the fact that the stochastic matrix P is ergodic then, the stationary distribution $\pi = (\pi_0,\ \pi_1, ..., \pi_C)$ exists and it is the solution of the system $\pi P = \pi$ with $\sum_{i=0}^{C} \pi_i = 1$.

3 Kernel estimation of transition matrix associated with the model

Let consider a $D/G/1/C$ queue described by a DTMC (see Section 2) and assume that the distribution of the number of the departures during two consecutive arrivals is an unknown function. That is to say, we only have a piece of partial information, on the distribution of the number of customers served between two consecutive arrivals, which is given in form of a sample of observations $x_1,\ x_2, ...,\ x_n$. In addition, our interest is the estimation of its transition matrix P using the kernel density estimator defined by:

$$\hat{f}(x) = \frac{1}{n} \sum_{i=1}^{n} K_{x,h}(X_i), \quad x \in \mathbb{N};$$

where $h \equiv h(n)$ is the smoothing parameter and $K_{x,h}$ is the discrete kernel of target x and the smoothing parameter h on the support $\aleph_{x,h} = \aleph_x$ (does not depend on h).

It is clear that the implementation of this technique requires to fixing beforehand the couple (K, h). For the choice of K, the problem is a priori easy. Indeed, the fact that $x \in \mathbb{N}$ and K has only a little influence on the quality of the estimator then it sufficient to use one of the following discrete kernels: Poisson kernel, Binomial kernel, Negative Binomial kernel, Triangular kernel, and Dirac kernel (Kokonendji et al., 2007; Kokonendji and Kiessé, 2011). For the choice of the smoothing parameter h, we propose to use two approaches, namely: the classical techniques and the matrix norms.

3.1 Choice of smoothing parameter via classical techniques

Let $X_1, \ldots, X_n$ be an n-sample $(i.i.d.)$ from the unknown distribution f. The idea, in this case, is to estimate the elements $f(x)$ from the sample without taking into account their repetition in the matrix P, that is to say it is enough to substitute $f(x)$ by its kernel estimator

$$\hat{f}(x) = \hat{P}r(X = x) = \frac{1}{n}\sum_{i=1}^{n} K_{x,h^*}(X_i), \qquad x \in \mathbb{N},$$

where h^* is the optimal smoothing parameter selected by the classical procedures and subsequently replace the elements $f(x)$ by their estimate $\hat{f}(x)$ in the matrix P to get $\hat{P}$. Below are the most classical methods used for choosing the smoothing parameter in the estimation of discrete density, in the case of a single sample.

1. Minimisation of the ISE: An appropriate choice for the smoothing parameter can be that minimising the integrated squared error (ISE), which is given by:

$$h^*_{ise} = \arg\min_h \left[\sum_{x\in\mathbb{N}} (\hat{f}(x) - f(x))^2\right] = \arg\min_h ISE(X_1, ..., X_n, n, h, K, f). \quad (3)$$

This technique has been detailed in Kokonendji et al. (2007); Kokonendji and Kiessé (2011).

2. Cross-validation: Another type of smoothing parameter selection techniques is those based on the cross-validation method, detailed initially by Scott and Terrell (1987), such as: Cross-validation by the least squares (UCV and BCV (Wansouwé et al., 2016)) and maximum likelihood cross validation (LCV). Recall that, the principle of these techniques is to estimate the density f at point X_i, using the Leave-one-out cross-validation.

3.2 The choice of smoothing parameter by matrix norms

In this case, the idea is to take into account the number of repetitions of $\hat{f}(x)$ in the matrix $\hat{P}$. To do this, we propose to use the matrix norms that allow us to include the repetitions of the quantities $\hat{f}(x)$ in $\hat{P}$. Hence, the optimal smoothing parameter can be calculated according to one of the following expressions that outcome, respectively, from the matrix norms $\|.\|_1$, $\|.\|_2$ and $\|.\|_\infty$:

$$h^*_1 = \arg\min_h \left[\max_{1\le i\le C}\left(\sum_{x=0}^{C-i}\left|\hat{f}(x) - f(x)\right| + \left|\sum_{x>C-i}\left(\hat{f}(x) - f(x)\right)\right|\right)\right]. \quad (4)$$

$$h_2^* = \arg\min_h \left[\sum_{x=0}^{C-1} (C-x+1) \left(\hat{f}(x) - f(x) \right)^2 + \sum_{i=1}^{C} \left(\sum_{x \geq i} \left(\hat{f}(x) - f(x) \right) \right)^2 + \left(\sum_{x \geq C} \left(\hat{f}(x) - f(x) \right) \right)^2 \right]. \tag{5}$$

$$h_3^* = \arg\min_h \left[\max_{0 \leq j \leq C} \left(\begin{array}{l} \left\{ \sum\limits_{x=0}^{j} \left| \hat{f}(x) - f(x) \right| + \left| \hat{f}(j) - f(j) \right| \right\} \mathbf{1}_{\{j \neq 0\}} \\ + \left\{ \sum\limits_{i=1}^{C} \left| \sum\limits_{x \geq i} \left(\hat{f}(x) - f(x) \right) \right| + \left| \sum\limits_{x \geq C} \left(\hat{f}(x) - f(x) \right) \right| \right\} \mathbf{1}_{\{j=0\}} \end{array} \right) \right] \tag{6}$$

with $\mathbf{1}_{\{.\}}$ is the indicator function.

4 Numerical application

The purpose of this section is to analyse numerically the impact of the smoothing parameter, chosen by the minimisation of matrix norms, on the performances of the kernel estimator of the transition matrix associated with the embedded Markov chain N in two situations, namely: the variation of the sample size and the variation of the traffic intensity ρ. To do this, we designed a simulator, under the Matlab environment, whose main steps are as follows:

- *Step 1.* Estimate h_{opt} using expressions (3)–(6) for a n-sample generated from f.
- *Step 2.* Calculate $\hat{P}$, $\hat{\pi}_C$ and $\hat{L}_s$ for h_{opt} obtained in Step 1.
- *Step 3.* Calculate h^*, $\overline{\pi}_C$ and $\overline{L}_s$ the average of h_{opt}, $\hat{\pi}_C$ and $\hat{L}_s$, and their variance.

The quantities $\hat{\pi}_C$ and $\hat{L}_s$ are the estimators of the loss-probability and the queue length respectively, which are obtained respectively by solving the system of equations $\pi \hat{P} = \pi$ (under the condition $\sum_{i=0}^{C} \hat{\pi}_i = 1$) and the expression $L_s = \sum_{i=0}^{C} i \hat{\pi}_i$.

4.1 Effect of the sample size on the estimator's performances

In this application, we focused on the effect of the choice of the smoothing parameter, on the stationary characteristics estimators of our model, according to the sample size. For this, we fix: the capacity of the system $C = 10$, the number of simulations $m = 1000$, the sample sizes $n \in \{50,\ 100;\ 250;\ 500;\ 1000\}$, the service rate $\mu = 1$, the traffic intensity (the arrival rate) $\rho = \lambda/\mu = \lambda = 0.75$ and the distribution $f \in \{\text{Poisson}(\rho^{-1}), \text{Geometric } (1/(1+\rho^{-1})), \text{Binomial } (C, \rho^{-1}/C)\}$, and the kernel $K \in \{Poisson;\ Negative\ Binomial;\ Triangular\}$ noted respectively K_{P_o}, K_{NB} and K_T. For all $(x, y) \in \mathbb{N}^2$ and $h > 0$, these kernels are defined as follow:

- $K_{P_o(x+h)}(y) = e^{-(x+h)}\frac{(x+h)^y}{y!}$, with $\aleph_{x,h} = \mathbb{N}$.
- $K_{BN\left(x+1,\frac{x+1}{2x+1+h}\right)}(y) = \frac{(x+y)!}{y!x!}\left(\frac{x+h}{2x+1+h}\right)^y\left(\frac{x+1}{2x+1+h}\right)^{x+1}$, with $\aleph_{x,h} = \mathbb{N}$.
- $K_{T_{(b,h,x)}}(y) = \frac{(b+1)^h - |y-x|^h}{(2b+1)(b+1)^h - 2\sum\limits_{j=0}^{b} j^h}\mathbf{1}_{\{|y-x|<b\}}$, with $\aleph_{x,h} = \{x,\ x \pm 1, \ldots,\ x \pm b\}$, $b \in \mathbb{N}$ and $\mathbf{1}_{\{.\}}$ is the indicator function. For the numerical application, we chose $b = 2$.

The exact characteristics of the considered model, for the above parameters, are given in Table 1 while a sample of the simulation results is ranked in Tables 2–5.

Table 1 Exact value of L_s and π_C according to the distribution f

Distribution	*Poisson*		*Geometric*		*Binomial*	
characteristic	L_s	π_C	L_s	π_C	L_s	π_C
Value	1.1813	6.38×10^{-4}	2.5149	146.99×10^{-4}	1.0140	3.05×10^{-4}

Table 2 Estimation of smoothing parameters h^*: case of the Binomial distribution.

		ISE		$\|.\|_1$		$\|.\|_2$		$\|.\|_\infty$	
K	n	h^*	$Var(h^*)$	h^*	$Var(h^*)$	h^*	$Var(h^*)$	h^*	$Var(h^*)$
K_{P_o}	100	0.3450	0.0482	0.2454	0.0477	0.4731	0.0570	1.5245	0.1484
	500	0.3030	0.0112	0.1597	0.0080	0.4429	0.0126	1.4841	0.0323
	1000	0.3070	0.0058	0.1636	0.0045	0.4485	0.0064	1.4924	0.0106
K_{NB}	100	2.5008	0.4686	3.4811	0.6580	3.2666	0.3926	7.9421	0.4766
	500	2.5595	0.0733	3.5231	0.1047	3.3025	0.0728	7.9830	0.0835
	1000	2.5739	0.0424	3.5363	0.0645	3.3185	0.0421	8.0075	0.0458
K_T	100	0.1224	0.0147	0.1211	0.0195	0.1133	0.0136	0.1093	0.0123
	500	0.0247	0.0011	0.0313	0.0015	0.0233	0.0011	0.0209	0.0009
	1000	0.0196	0.0006	0.0216	0.0007	0.0186	0.0005	0.0173	0.0005

Table 3 Estimation of the mean number of customer in the system: case of the Poisson distribution

		ISE			$\|.\|_1$			$\|.\|_2$			$\|.\|_\infty$		
K	n	$\overline{L}_s$	Var	MSE	$\overline{L}_s$	Var	MSE	$\overline{L}_s$	Var	MSE	$\overline{L}_s$	Var	MSE
K_{P_o}	100	0.7887	0.0051	0.1592	0.7871	0.1258	0.2812	0.8698	0.0035	**0.1006**	1.7185	0.0540	0.3425
	500	0.7934	0.0533	0.2038	0.7417	0.1747	0.3680	0.8845	0.0490	**0.1371**	1.7351	0.0511	0.3577
	1000	0.7675	0.0003	0.1716	0.6823	0.0001	0.2491	0.8598	0.0003	**0.1036**	1.7276	0.0272	0.3256
K_{NB}	100	0.7966	0.0002	0.1482	0.8828	0.0024	0.0915	0.9127	0.0001	**0.0723**	2.5172	0.0073	1.7917
	500	0.7944	0.0000	0.1498	0.8779	0.0006	0.0927	0.9114	0.0000	**0.0729**	2.5238	0.0014	1.8036
	1000	0.7942	0.0000	0.1499	0.8770	0.0002	0.0928	0.9115	0.0000	**0.0728**	2.5219	0.0006	1.7978
K_T	100	0.9109	0.0429	0.1160	0.9007	0.0481	0.1268	0.9276	0.0424	**0.1067**	0.9258	0.0452	0.1104
	500	1.0634	0.0222	0.0361	1.0471	0.0262	0.0443	1.0704	0.0201	**0.0324**	1.0713	0.0285	0.0406
	1000	1.1163	0.0092	0.0135	1.1122	0.0125	0.0173	1.1219	0.0083	**0.0118**	1.1229	0.0110	0.0144

Table 4 Estimation of the mean number of customer in the system: case of the Geometric distribution

		ISE			$\|.\|_1$			$\|.\|_2$			$\|.\|_\infty$		
K	n	$\overline{L}_s$	Var	MSE	$\overline{L}_s$	Var	MSE	$\overline{L}_s$	Var	MSE	$\overline{L}_s$	Var	MSE
K_{P_o}	100	2.5978	0.1570	0.1638	2.6105	0.1489	0.1581	2.5614	0.1391	0.1413	2.3643	0.0664	**0.0891**
	500	2.2907	0.3730	0.4233	2.2990	0.3732	0.4198	2.2697	0.3483	0.4084	2.2340	0.1870	**0.2659**
	1000	2.1723	0.4257	0.5431	2.1935	0.4168	0.5201	2.1702	0.3955	0.5144	2.1149	0.2720	**0.4320**
K_{NB}	100	1.9208	0.0082	0.3612	1.9147	0.0068	0.3671	2.0419	0.0050	**0.2288**	3.0021	0.0044	0.2417
	500	1.9042	0.0021	0.3751	1.8980	0.0018	0.3825	2.0301	0.0012	**0.2363**	3.0115	0.0011	0.2477
	1000	1.9061	0.0011	0.3718	1.9006	0.0010	0.3784	2.0311	0.0007	**0.2347**	3.0077	0.0006	0.2434
K_T	100	2.1619	0.2636	0.3882	2.1787	0.2793	0.3924	2.1887	0.2466	0.3530	2.2043	0.2060	**0.3024**
	500	2.3696	0.0612	0.0824	2.3967	0.0651	0.0790	2.3771	0.0574	0.0764	2.3810	0.0488	**0.0667**
	1000	2.4297	0.0416	0.0489	2.4395	0.0406	0.0463	2.4324	0.0382	0.0450	2.4309	0.0289	**0.0359**

Table 5 Estimation of the loss probability: case of the Poisson distribution ($\times 10^{-4}$)

		ISE		$\|.\|_1$		$\|.\|_2$		$\|.\|_\infty$	
K	n	$\overline{\pi}_C$	MSE	$\overline{\pi}_C$	MSE	$\overline{\pi}_C$	MSE	$\overline{\pi}_C$	MSE
K_{P_o}	100	1.4857	0.0025	6.6148	0.1435	2.3455	**0.0017**	42.9779	0.1992
	500	3.7503	0.0649	8.7334	0.3592	4.6890	**0.0637**	44.2053	0.2138
	1000	1.1858	0.0027	0.6199	0.0033	2.1327	**0.0018**	42.2591	0.1474
K_{NB}	100	1.5796	0.0023	2.7321	0.0014	3.1424	**0.0011**	154.8834	2.2385
	500	1.5535	0.0023	2.6066	0.0014	3.1178	**0.0011**	155.5688	2.2316
	1000	1.5511	0.0023	2.5836	0.0014	3.1184	**0.0011**	155.0893	2.2140
K_T	100	2.9368	0.0019	2.9682	0.0024	3.1272	**0.0017**	3.2056	0.0022
	500	4.7239	0.0010	4.5481	0.0013	4.7968	**0.0009**	5.0260	0.0016
	1000	5.3219	0.0005	5.3239	0.0007	5.4061	**0.0004**	5.4968	0.0007

From the obtained results, we notice that:

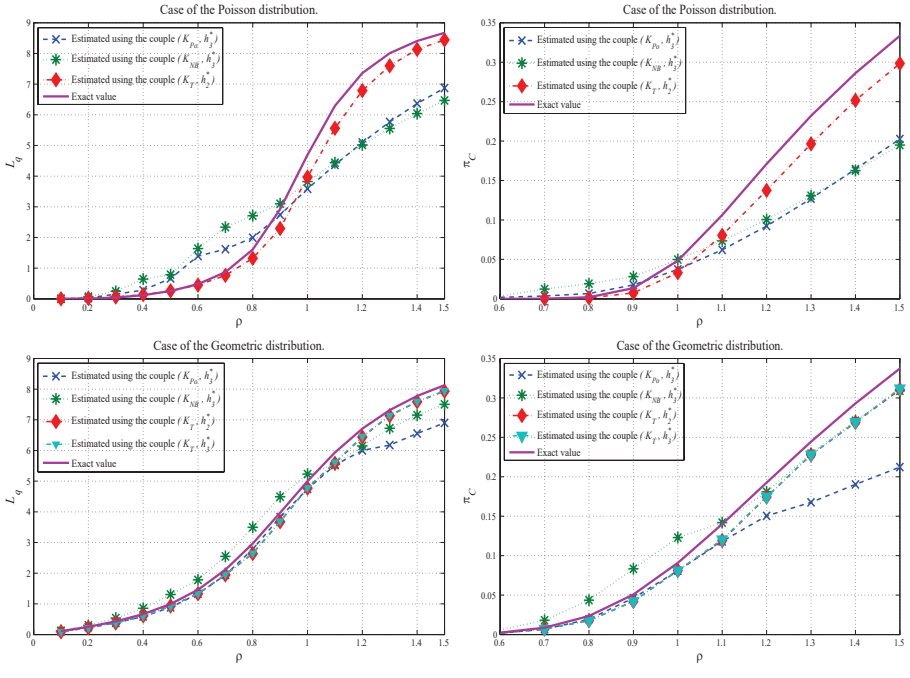
- Except for the case of use of K_T where the convergence property of the optimal smoothing parameter h^* according to n, is satisfied, when using K_{P_0} or K_{NB} this property is not satisfied and this regardless of the used selection procedure and the distribution f.
- In the case of the Poisson and the Geometric distributions, the behaviour of the variance of the optimal smoothing parameter, $Var(h^*)$, is not regular according n when we use K_{P_0} to construct the estimator $\hat{P}$. While, in the case of K_{NB} and K_T, the behaviour of $Var(h^*)$ is sensitive and converges to zero according to n, and this regardless of f and the selection procedure used. Besides, the smallest $Var(h^*)$ is noticed when using the K_T.
- In the case of the Poisson and the Binomial distribution, the h^* which provides us estimators, of L_q and π_C, with a smaller MSE is that defined in equation (5) when we use the K_T. Also, it is preferable to avoid the use of the kernels K_{P_0} and K_{NB}. If f is a Geometric distribution, to obtain a more efficient estimators, of L_q and π_C, it is preferable to construct $\hat{P}$, by using the couple $(K_T,\ h_3^*)$ if $n \geq 250$ and the couple $(K_{P_0},\ h_3^*)$ when $n < 250$.

4.2 Effect of the traffic intensity on the estimator's performances

The purpose of this second application is to analyse numerically the effect of ρ on the quality of the estimates h^*, $\overline{\pi}_C$ and $\overline{L}_s$. To do this, we set $n = 200$ and $\rho = \lambda \in [0.1\ ;\ 1.5]$ while the rest of the parameters are identical to those of the first application. A sample of obtained results is presented in Figure 1. From the results, we notice that:

- For a low traffic intensities ($\rho < 0.6$), in all the situations considered, we obtain practically the same characteristics (bias, variance and MSE) of the estimators $\overline{L}_s$ and $\overline{\pi}_C$.
- Contrary to the case of K_T, where h^* is inversely proportional to ρ, when using K_{P_0} or K_{NB} the parameter h^* is proportional to ρ. We also notice that the behaviour of $Var(h^*)$ according to ρ is not regular.
- In the case of the Poisson and the Binomial distributions, the h^* which provides us estimators of L_q and π_C with a small MSE according to the used kernel are: $(K_{P_0},\ h_3^*)$, $(K_{NB},\ h_3^*)$ and (K_T, h_2^*). But overall, whether for $\rho \geq 0.6$, it is preferable to use the couple $(K_T,\ h_2^*)$ and avoid the couple (K_{NB}, h^*). While, in the case f is a Geometric distribution the h^* which provide us estimators of L_q and π_C with a small MSE according to the used kernel are: $(K_{P_0},\ h_3^*)$, $(K_{NB},\ h_3^*)$, and $(K_T,\ h_3^*)$ or $(K_T,\ h_2^*)$. But overall, it is preferable to use K_T and select the smoothing parameter by formulas (5) or (6).

Figure 1 Variation of the model characteristics against the traffic intensity ρ (see online version for colours)



5 Concluding remarks

In this paper, we considered the choice of the smoothing parameter by procedures that are based on the minimisation of matrix norms in the kernel estimation of a transition matrix associated with an embedded DTMC, describing a $D/G/1/C$ queue. Our simulation results allow us to conclude, on the one hand, that the smoothing parameters selected by the minimisation of matrix norms (in particular the quadratic matrix norm) provide us, in general, more efficient estimators and, on the other hand, that the choice of the kernel is of great importance, where it is preferable to use K_T and avoid K_{NB} in the context addressed.

In order to explain the paradoxical behaviour of h^* according to n in certain situations, let us recall that the global bias and the global variance of the estimator of a probability mass function f, denoted respectively $IBias$ and $IVar$, are defined as follow:

$$\begin{aligned} IBias(\hat{f}) &= h\, B\left(h, K, f, f''\right) \\ &= h\, \left\{ \sum_{x\in\mathbb{N}} \left[f\left\{\mathbb{E}\left(\mathcal{K}_x, h\right)\right\} - f(x) + \frac{V(K,x,h)}{2h} f''(x) \right] \right\}. \end{aligned} \tag{7}$$

$$IVar(\hat{f}) = \frac{1}{n} E(\mathcal{K}_{x,h}) = \frac{1}{n} \sum_{x\in\mathbb{N}} f(x) Pr(\mathcal{K}_{x,h} = x). \tag{8}$$

where $\mathcal{K}_{x,h}$ is the random variable of law $K_{x,h}$ defined on $\aleph_{x,h}$ and

$$V(K,x,h) = \begin{cases} x+h, & \text{if } K = K_{P_0}; \\ \frac{x-(x-1)h}{x+1}, & \text{if } K = K_{NB}; \\ \frac{(2x+1)x+(3x+1)h}{x+1}, & \text{if } K = K_T. \end{cases} \tag{9}$$

From the expression (8), we note that if $n \to \infty$, then $IVar(\hat{f}) \to 0$ whatever the kernel used (Kokonendji and Kiessé, 2011). Consequently, the behaviour of the h^* according to n can only be justified by the behaviour of $IBias(\hat{f})$. From the expression (7), it is clear that $IBias(\hat{f}) \to 0$ if only $h \to 0$ when n tends to infinity and $B(h, K, f, f'')$ is a finite quantity. However, the expression (9), indicate that the quantity $B(h, K, f, f'')$ tends to zero only if the smoothing parameter h is sufficiently large and this fact that $V(x,h)/2h$ is a decreasing function according to h. At this stage, we conclude that h must be chosen so as to have a compromise between h and $B(h, K, f, f'')$ and which is not necessarily a zero at the limit. This conclusion perfectly reflects the behaviour of our results.

Regarding the Triangular kernel, it should be noted that the estimator designed via this kernel is free from the problem of the convergence of its local and global bias. Indeed, when we consider the Triangular kernel the expression (**??**), which can be rewritten as follows:

$$IBias\left(\hat{f}\right) = \sum_{x\in\mathbb{N}} \left[f(x) \left\{ \frac{(b+1)^h}{P(b,h)} - 1 \right\} + \sum_{y\in\aleph_{x,h}\setminus\{x\}} f(y) Pr\left(\mathcal{K}_{T_{(b,h,x)}} = y\right) \right], \tag{10}$$

tend to zero when h tends to zero (for more details see (Kokonendji et al., 2007)).

The negative impact of the $\hat{f}$ bias on the performances of the kernel estimator of a Markov chain transition matrix has already been reported by Cherfaoui et al. (2015 b) in the case $x \in \mathbb{R}_+$. Consequently, in the light of our results and those of Cherfaoui et al. (2015 b), it is natural to consider the introduction of the bias-reduction techniques for a possible improvement in the performances of the kernel estimator of a transition matrix.

References

Aitchison, J. and Aitken, C. (1976) 'Multivariate binary discrimination by the kernel method', *Biometrika*, Vol. 63, No. 3, pp.413–420.

Alfa, A. (2010) *Queueing Theory for Telecommunications: Discrete Time Modelling of a Single Node System*, Springer, New York.

Belaid, N., Adjabi, S., Kokonendji, C. and Zougab, N. (2018) 'Bayesian adaptive bandwidth selector for multivariate discrete kernel estimator', *Communications in Statistics – Theory and Methods*, Vol. 47, No. 12, pp.2988–3001.

Cherfaoui, M., Boualem, M., Aïssani, D. and Adjabi, S. (2015a) 'Choix du paramètre de lissage dans l'estimation à noyau d'une matrice de transition d'un processus semi-markovien', *C. R. Acad. Sci. Paris, Ser. I*, Vol. 353, No. 3, pp.273–277.

Cherfaoui, M., Boualem, M., Aïssani, D. and Adjabi, S. (2015b) 'Quelques propriétés des estimateurs à noyaux gamma pour des échantillons de petites tailles', *Afrika Statistika*, Vol. 10, No. 1, pp.763–776.

Djerroud, L., Kiessé, T. and Adjabi, S. (2020) 'Semiparametric multiple kernel estimators and model diagnostics for count regression functions', *Communications in Statistics – Theory and Methods*, Vol. 49, No. 9, pp.2131–2157.

Gontijo, G., Atuncar, G., Cruz, F. and Kerbache, L. (2011) 'Performance evaluation and dimensioning of $GI^{[X]}/M/c/N$ systems through kernel estimation', *Mathematical Problems in Engineering*, Vol. 2011, 20 pages.

Graham, C. (2014) *Markov Chains: Analytic and Monte Carlo Computations*, Wiley Series in Probability and Statistics. Wiley.

Kokonendji, C. and Kiessé, T. (2011) 'Discrete associated kernels method and extensions', *Statistical Methodology*, Vol. 8, No. 6, pp.497–516.

Kokonendji, C., Kiessé, T. and Zocchi, S. (2007) 'Discrete triangular distributions and non-parametric estimation for probability mass function', *Journal of Nonparametric Statistics*, Vol. 19, No. 6-8, pp.241–254.

Norris, J.R. (1997) *Markov chains. Cambridge Series in Statistical and Probabilistic Mathematics*, Cambridge University Press, Cambridge.

Ott, T. (1987) 'Simple inequalities for the D/G/1 queue', *Operations Research*, Vol. 35, No. 4, pp.589–597.

Privault, N. (2018) *Understanding Markov Chains: Examples and Applications*, Springer Undergraduate Mathematics Series. Springer Singapore.

Roberts, J., Mocci, U. and Virtamo, J. (1996) *Broadband Network Teletraffic*, Final Report of Action COST 242, Springer, Berling.

Roussas, G. (1969) 'Nonparametric estimation in markov processes', *Ann. Inst. Statist. Math.*, Vol. 21, pp.73–87.

Scott, D.W. and Terrell, G.R. (1987) 'Biased and unbiased cross-validation in density estimation', *Journal of the American Statistical Association*, Vol. 82, No. 400, pp.1131–1146.

Somé, S. and Kokonendji, C. (2016) 'Effects of associated kernels in nonparametric multiple regressions', *Journal of Statistical Theory and Practice*, Vol. 10, No. 2, pp.456–471.

Wansouwé, W., Somé, S. and Kokonendji, C. (2016) 'An R package for discrete and continuous associated kernel estimations', *The R Journal*, Vol. 8, No. 2, pp.258–276.

Zougab, N., Adjabi, S. and Kokonendji, C. (2014) 'Bayesian approach in nonparametric count regression with binomial kernel', *Communications in Statistics - Simulation and Computation*, Vol. 43, No. 5, pp.1052–1063.