Méthode du noyau dans l'évaluation de performances des systèmes d'attente

Smail Adjabi, Karima Lagha, Madjid Hassani

Méthode du noyau dans l'évaluation de performance des systèmes d'attente

Smail Adjabi, Karima Lagha et Madjid Hassani
Laboratoire LAMOS, Université de Béjaia,
06000 Béjaia, Algérie
E-mail : adjabi@hotmail.com

Résumé

L'évaluation de performances d'un système de files d'attente lorsque les lois des inter arrivées et de service sont générales (quelconques) est complexe car on a peu d'information sur ces lois. Pour cela, on estime ces lois par la méthode non paramétrique du noyau de Parzen-Rosenblatt adapté au cas où la taille de l'échantillon est aléatoire, ensuite, on évalue les performances de ce système. Les résultats sur les caractéristiques de performances obtenus par simulation montrent l'intérêt de l'estimateur de la densité par la méthode du noyau adapté.

Mots-clés : Évaluation de performances ; Système d'attente ; estimateur à noyau ; simulation.

Abstract

The queuing system performance evaluation, when the inter-arrival and service distribution are general (unspecified), is complex because the lack of information about these distributions. We use the kernel estimator of Parzen-Rosenblatt adapted to the random case to estimate density probability of distribution and we evaluate performance of various queuing systems by simulation. The results obtained and related to the characteristics of these systems show the interest of this approach.

Keywords : Performance evaluation ; queuing system ; kernel estimator ; simulation.

1 Introduction

L'évaluation de performances des systèmes de files d'attente est un problème très étudié dans la littérature. La difficulté de cette étude réside dans l'analyse théorique du système d'attente. En effet, il est facile d'évaluer les performances d'un système de files d'attente markovien ($M/M/1$), par contre il est difficile d'analyser le système non markovien G/G/1. Nous considérons dans ce travail, l'estimation des densités des lois des inter-arrivées et de service par l'estimateur à noyau de Parzen-Rosenblatt adapté au cas où la taille de l'échantillon est aléatoire. Pour corriger les effets de bord des fonctions à support $[0, \infty[$, on utilise dans l'estimation de la densité le noyau asymétrique Gamma. L'évaluation de performances des systèmes de files d'attente sera alors faite pas simulation.

2 Estimation séquentielle de la densité de probabilité

En général, pour estimer une densité de probabilité avec la méthode du noyau, on a besoin d'un échantillon de taille déterministe. Dans la pratique, il existe des situations ou la taille de l'échantillon peut être aléatoire. Par exemple, le nombre de véhicules qui passent par un endroit donné sur une période de temps $[0, t]$. Pour cela, une autre classe d'estimateurs existe pour estimer la densité d'une probabilité basée sur un échantillon de taille aléatoire.
Soit un échantillon $X_1, ..., X_{N_t}$ issu d'une v.a. X de fonction de densité de probabilité f. N_t est

une variable aléatoire discrète à valeurs entières, elle représente le nombre d'observations dans l'intervalle $(0,t]$. L'estimateur de $f(x)$ de Parzen-Rosenblatt adapté au cas où la taille de l'échantillon est aléatoire est de la forme :

$$f_{N_t}(x) = \frac{1}{N_t h(N_t)} \sum_{j=1}^{N_t} K\left(\frac{x - X_j}{h(N_t)}\right), \tag{1}$$

où h est un paramètre appelé paramètre de lissage qui satisfait : $h(n) \to 0$ quand $n \to \infty$.
La fonction $K(y)$ appelée noyau vérifie :

$$K(y) \geq 0, \int_{-\infty}^{+\infty} K(y)dy = 1, \sup_{y\in\mathbb{R}} K(y) < \infty, \ K(y) = K(-y) \text{ et } \lim_{y\to\infty} |yK(y)| = 0.$$

2.1 Propriétés de l'estimateur

Soit X une variable aléatoire de densité de probabilité f, et $X_1, ..., X_{N_t}$ un échantillon issu de X. Soit f_{N_t} l'estimateur de Parzen-Rosenblatt adapté défini dans (1).

- Si on suppose que la taille de l'échantillon N_t est indépendante des observations $X_1, X_2, \ldots, X_{N_t}$ et que $N_t \to \infty$ en Probabilité quand $t \to \infty$, on a alors les propriétés suivantes :

  Propriétés[11]
  L'estimateur $f_{N_t}(x)$ de $f(x)$ est asymptotiquement sans biais, consistent et uniformément consistent. De plus, cet estimateur est asymptotiquement normal, dans le sens :

$$\sqrt{N_t h(N_t)}\left(\hat{f}_{N_t}(x) - f(x)\right) \to \mathcal{N}\left(0, \ f(x)\int_{-\infty}^{\infty} K^2(y)dy\right).$$

- Si on suppose que la taille de l'échantillon N_t dépend des observations $X_1, X_2, \ldots, X_{N_t}$ et que $\frac{1}{t}N_t \to \pi$ en Probabilité quand $t \to \infty$, on a les propriétés suivantes :

  Propriétés [11]
  L'estimateur $f_{N_t}(x)$ de $f(x)$ est faiblement consistant, uniformément consistant et asymptotiquement normal.

2.2 Choix du noyau K

Lorsque le support de la densité est $[0, \infty]$, on utilise un noyau flexible qui n'assigne pas de poids en dehors du support de la densité. En effet, le biais au bord de l'estimateur est du à l'allocation de poids par le noyau symétrique en dehors du support lorsque l'estimation est faite près du bord. pour éviter ce problème, mous utilisons un noyau asymétrique [3] comme le noyau Gamma de paramètres $(\frac{x}{h}+1, h)$ défini par :

$$K\left(\frac{x}{h}+1, h\right)(t) = \frac{t^{x/h} e^{-t/h}}{h^{\frac{x}{h}+1}\Gamma(\frac{x}{h}+1)},$$

où $\Gamma(.)$ est la fonction gamma et $h = h(n)$ est le paramètre de lissage

2.3 Choix du paramètre de lissage h

La méthode utilisée pour estimer le paramètre de lissage est la méthode validation croisée non biaisée (unbiased cross validation) proposée par Rudemo et Bowman [2]. Le critère consiste à choisir le paramètre de lissage qui minimise un estimateur de

$$ISE(h(n)) = \int_{\mathbb{R}} [f_n(x) - f(x)]^2 \, dx.$$

Ceci revient à minimiser la quantité :

$$UCV(h) = \int_{\mathbb{R}} f_n^2(x)dx - \frac{1}{n}\sum_{i=1}^{n} f_{n,i}X_i \ ,$$

où

$$f_{n,i}(x) = \frac{1}{(n-1)h}\sum_{\substack{1\leq j\leq n\\ i\neq j}} K\left(\frac{x-X_j}{h}\right).$$

3 Simulation

Nous simulons différents systèmes de files d'attente du type $M/M/1$, $M/G/1$, $G/M/1$ et $G/G/1$. On évalue les performances de ces systèmes sur une période de temps $[0,T]$. les caractéristiques de performances mesurées sont :
- $\overline{N}$: Nombre moyen de clients présents dans le système ;
- $\overline{T}$: Temps moyen de séjour d'un client dans le système ;
- $\overline{Q}$: Nombre moyen de clients en attente ;
- $\overline{W}$: Temps moyen d'attente d'un client dans la file.

Avec

$$\overline{T} = \frac{\overline{N}}{\lambda}\ , \quad \overline{W} = \frac{\overline{Q}}{\lambda}\ , \quad \overline{T} = \overline{W} + \frac{1}{\mu}\ ,$$

où λ et μ sont respectivement le taux d'arrivée des clients dans le système et le taux de service. Soient f et g les densités des temps des inter-arrivées et de service respectivement. Dans le but d'estimer f et g, on observe le système durant un intervalle de temps $[0,t]$. Soit N_t et M_t respectivement le nombre d'arrivées et le nombre des clients servis durant l'intervalle $(0,t]$. $X_1,...,X_{N_t}$ et $Y_1,...,Y_{M_t}$ sont deux échantillons des séquences des inter-arrivées et de service respectivement. Les estimateurs de f et g sont alors :

$$f_{N_t}(x) = \frac{1}{n_t}\sum_{j=1}^{N_t}\frac{1}{h(N_t)}K\left(\frac{x-X_j}{h(N_t)}\right), \tag{2}$$

$$g_{M_t}(x) = \frac{1}{M_t}\sum_{j=1}^{M_t}\frac{1}{h(M_t)}K\left(\frac{x-Y_j}{h(M_t)}\right). \tag{3}$$

Afin de simuler les différents systèmes et d'évaluer leur performances durant un intervalle de temps $(0,T]$, on procède comme suit :

1. Donne l'horizon de simulation T.
2. Génére les séquences d'inter-arrivées suivant une loi donnée ;
3. Génére les durées de service suivant une loi donnée ;
4. Calcule le nombre N_t de clients entrants et le nombre M_t de clients servis dans le système durant la période $(0,T]$
5. Estime $f(x)$ et $g(x)$ par $f_{N_t}(x)$ et $g_{M_t}(x)$;
6. Calcule la durée moyenne entre deux arrivées consécutives : $E(X) = \int x f_{N_t}(x)dx$;
7. Calcule le taux d'arrivé : $\lambda = \frac{1}{E(X)}$;
8. Calcule la durée moyenne de service : $E(Y) = \int y g_{M_t}(y)dy$;
9. Calcule le taux de service : $\mu = \frac{1}{E(Y)}$;
10. Calcule le coefficient d'utilisation du système : $\rho = \frac{\lambda}{\mu}$.
11. Si $\rho < 1$, à l'aide des formules analytiques, on évalue les performances du système étudié.

En se basant sur cet algorithme, nous allons simuler différents systèmes de files d'attente du type $(M/M/1, M/G/1, G/M/1, G/G/1)$. Dans le but de voir le comportement aléatoire des séquences des inter-arrivées et des durées de service, le système sera simulé plusieurs fois suivant des horizons de plus en plus grand (100, 500, 1000 et 10000).
Les résultats obtenus par la simulation sont donnés sous forme de tableaux(Tab. 1, Tab.2, Tab. 3 et Tab. 4) où

1. Horizon : Horizon de la simulation ;
2. N_t **:** Nombre de clients arrivés dans le système ;
3. M_t **:** Nombre de clients servis dans le système ;
4. EX_{th} **:** Durée moyenne théorique entre deux arrivées consécutives ;
5. EY_{th} **:** Durée moyenne théorique des services ;
6. $\overline{N}_{th}$ **:** Nombre moyen théorique de clients présent dans le système ;
7. $\overline{T}_{th}$: Temps moyen théorique d'un client dans le système ;
8. $\overline{Q}_{th}$ **:** Nombre moyen théorique de clients en attente ;
9. $\overline{W}_{th}$ **:** Temps moyen théorique d'attente d'un client dans la file ;
10. EX_s **:** Durée moyenne entre deux arrivées consécutives donnée par la simulation ;
11. EY_s **:** Durée moyenne des services donnée par la simulation ;
12. $\overline{N}_s$ **:** Nombre moyen de clients présent dans le système, donné par la simulation ;
13. $\overline{T}_s$: Temps moyen de séjour d'un client dans le système, donné par la simulation ;
14. $\overline{Q}_s$ **:** Nombre moyen de clients en attente, donné par la simulation ;
15. $\overline{W}_s$ **:** Temps moyen d'attente d'un client dans la file donné par la simulation ;
16. $\overline{Mi}$ **:** Moyenne de la période d'inoccupation ;
17. $\overline{Vari}$ **:** Variance de la période d'inoccupation.

Système 1 (M/M/1)														
Loi des inter-arrivées : Exponentielle de paramètre ($\lambda = 0.3$)														
Loi des durées de service : Exponentielle de paramètre ($\mu = 0.4$)														
			Résultats de simulation						Résultats théoriques					
Horizon	N_t	M_t	EX_s	EY_s	$\overline{N}_s$	$\overline{Q}_s$	$\overline{T}_s$	$\overline{W}_s$	EX_{th}	EY_{th}	$\overline{N}_{th}$	$\overline{Q}_{th}$	$\overline{T}_{th}$	$\overline{W}_{th}$
[0,100]	30	27	3.38	2.60	3.35	2.58	11.35	8.74	3.33	2.5	3	2.25	10	7.5
[0,500]	152	151	3.29	2.56	3.47	2.70	11.48	8.91						
[0,1000]	293	291	3.39	2.59	3.22	2.46	10.94	8.35						
[0,10000]	2991	2990	3.34	2.51	3.01	2.25	10.05	7.55						

Tab. 1 Résultats du système M/M/1

Système 2 (G/M/1)														
Loi des inter-arrivées : Weibull de paramètres ($a = 5, b = 3$)														
Loi des durées de service : Exponentielle de paramètre ($\mu = 0.4$)														
			Résultats de simulation						Résultats théoriques					
Horizon	N_t	M_t	EX_s	EY_s	$\overline{N}_s$	$\overline{Q}_s$	$\overline{T}_s$	$\overline{W}_s$	EX_{th}	EY_{th}	$\overline{N}_{th}$	$\overline{Q}_{th}$	$\overline{T}_{th}$	$\overline{W}_{th}$
[0,100]	37	36	2.75	2.61	2.29	1.34	6.31	3.70	2.75	2.50	2.46	1.55	6.77	4.27
[0,500]	181	170	2.76	2.51	2.65	1.73	7.31	4.79						
[0,1000]	365	359	2.74	2.52	2.52	1.60	6.93	4.40						
[0,10000]	3632	3631	2.75	2.50	2.49	1.58	6.86	4.35						

Tab. 2 Résultats du système Weibull/M/1

Système 3 (M/G/1)														
Loi des inter-arrivées : Exponentielle de paramètre ($\lambda = 0.3$)														
Loi des durées de service : Cox2 de paramètres ($\mu_1 = 3, \mu_2 = 1, a = 0.2$)														
			Résultats de simulation						Résultats théoriques					
Horizon	N_t	M_t	EX_s	EY_s	$\overline{N}_s$	$\overline{Q}_s$	$\overline{T}_s$	$\overline{W}_s$	EX_{th}	EY_{th}	$\overline{N}_{th}$	$\overline{Q}_{th}$	$\overline{T}_{th}$	$\overline{W}_{th}$
[0,100]	28	27	3.67	0.63	0.19	0.026	0.73	0.097	3.33	0.53	0.18	0.024	0.61	0.082
[0,500]	151	150	3.33	0.52	0.18	0.022	0.60	0.074						
[0,1000]	300	299	3.34	0.53	0.18	0.024	0.61	0.080						
[0,10000]	3010	3009	3.32	0.52	0.18	0.023	0.60	0.079						

Tab. 3 Résultats du système M/Cox2/1

Système 4 ($G/G/1$)												
Loi des inter-arrivées : Erlang de paramètres ($\mu = 2, K = 2$)												
Loi des durées de service : Hyperexponentielle de paramètres ($\mu_1 = 2, \mu_2 = 1.5, a = 0.6$)												
Horizon	N_t	M_t	EX_{th}	EX_s	EY_{th}	EY_s	Mi	$Vari$	N	Q	T	W
[0,100]	101	100		0.99		0.67	0.89	0.73	1.65	0.97	1.64	0.97
[0,500]	504	503	1	0.99	0.56	0.61	0.78	0.52	1.32	0.70	1.31	0.70
[0,1000]	1001	999		0.99		0.59	0.84	0.46	1.12	0.53	1.12	0.53
[0,10000]	9964	9963		1.00		0.59	0.78	0.43	1.12	0.53	1.13	0.53

Tab. 4 Résultats du système $E_2/H_2/1$

Interprétation des résultats

D'après les résultats des tableaux précédents, on remarque que les caractéristiques des systèmes, obtenus par simulation en utilisant l'estimateur à noyau Parzen-Rosenblatt adapté au cas où la taille de l'échantillon est aléatoire pour les densités des temps des inter-arrivées et de service sont proches des caractéristiques théoriques. Les résultats sont meilleurs lorsque l'horizon T est grand.

Conclusion

Dans ce travail nous avons étudié la performance des systèmes de files d'attente en estimant les densités des temps des temps des inter-arrivées et de service par l'estimateur à noyau de Parzen-Rosenblatt adapté au cas où la taille de l'échantillon est aléatoire. Nous avons simulé différents systèmes d'attente et évalué leur performances. Les résultats obtenus montrent que l'utilisation de l'estimateur à noyau lorsque la taille de l'échantillon est aléatoire est possible et donne de bons résultats. Les résultats sont encore meilleurs lorsque l'horizon est grand.

Références

[1] Anscombe, F.J. (1952). Large sample theorie of sequentiel estimation. *Proc. Comb. Philos. Soc*, *48*, 600–607.

[2] Bowman A.W. (1984). An alternative method of cross-validation for the smoothing of density estimates, Biometrika, 71, 353-360.

[3] Chen, S.X. (2000). Probability density function estimation using gamma kernels, Ann. Inst. Statist. Math. 52(3), 471-480.

[4] Deheuvels, P. (1974). Conditions nécessaires et suffisantes de convergence ponctuelle presque sûre et uniforme presque sûre des estimateurs de la densité, C.R. Acad. Sci., Paris 278. sér. A :1217-1220.

[5] Deheuvels, P. (1990). Laws of iterated logarithm for density estimators. In : Nonparametric Functional Estimation and Related Topics, Roussas, G. (Ed), Kluwer, Dordrecht., 19-20.

[6] Hall, P. (1981). Laws of the Iterated Logarithm for Nonparametric Density Estimators, Z. Wahrscheinlichkeitstheorie verw. Gebiete, 56 :47-61.

[7] Parzen, E. (1962). On estimation of a probability density and mode, Ann. Math. Statist., 33 :1065-1076.

[8] Prakasa Rao, B.L.S. (1983). Non parametric Functional Estimation, Probability and Mathematical Statistics. A serie of Monographs and Textbooks. 72 PRA 83a.

[9] Renyi, A. (1973). On the asymptotic distribution of the sum of a random number of independent random variable. *Acta. Math. cad. Sci. Hung* , *8*, 193–199.

[10] Schuster, E. F. (1985). Incorporating support constraints into non parametric estimation of densities, Communication in Statistics. Theory and Methods, 14, 1123-1136.

[11] Srivastava, R.C. (1973). Estimation of Probability Density Function based on Random Number of Observations with Applications, Int. Stat. Rev., 41 :77-86.