Journal of the Egyptian Mathematical Society 25 (2017) 191–196

Contents lists available at ScienceDirect

Journal of the Egyptian Mathematical Society

journal homepage: www.elsevier.com/locate/joems

Stochastic approximation with series of delayed observations

M.A. Mahmoud, A.A. Rasha, S.W. Waseem*

Department of Mathematics, Faculty of Science, Zagazig University, Zagazig, Egypt

ARTICLE INFO

Article history:
Received 8 July 2016
Revised 2 January 2017
Accepted 3 January 2017
Available online 26 January 2017

MSC:
62L20
62M05
62M10
62L10
62L12

Keywords:
Stochastic approximation
Series of observations
Loss systems
Multiservice observations
Modified procedure
Efficiency

ABSTRACT

The stochastic approximation procedure with series of delayed observations is investigated. The procedure is formed by modifying the Robbins–Monro stochastic approximation procedure to be applicable in the presence of series of delayed observations. The modified procedure depends on a new base concerning the relation between service time of the series and service times of its components. Two loss systems are introduced for application to the proposed procedure. This new situation can be applied to increase the production of items in many fields such as biological, medical, life time experiments, and some industrial projects, where items are realized after random time delays. The efficiency of the procedure is computed. Our proposal is general and we expect that it can be applied to any other stochastic approximation procedure.

1. Introduction

Stochastic approximation and related topics, as recursive parameter estimation, or up and down experimentation, belong to the so-called on-line methods, where the solution of the problem takes place in real time, in the course of the process of observation. Stochastic approximation is a procedure for finding the root of an equation, or the solution of a system of equations, where the values of the respective functions can only be observed (measured) with experimental errors, at recursively determined points. The procedure is nonparametric with respect to the type of the function as well as with respect to the distribution of the experimental error (if an information about this distribution is available, it can be made use of).

The procedure for finding the root is called the Robbins–Monro stochastic approximation procedure [1]. There is an extensive literature and a lot of papers on this topics (cf. [2–7]), we shall use the review papers [8–13] as references.

Typically, in the investigated stochastic approximation procedure, observations follow each other after fixed time-intervals; where the point of the next observation is corrected according to the result of the preceding one. However, in some situations, as in biological or lifetime experiments, it may happen that the result of an observation becomes known only after a random time delay.

Recently, the stochastic approximation has been used in clinical applications to find optimal dose as in [14].

It is meaningful to ask, whether and how stochastic approximation can then be applied. We answer the question for the modified stochastic approximation procedure with delayed series of delayed multiservice observations (or customers) by investigating two loss systems to be applicable in the presence of the modified Robbins–Monro stochastic approximation procedure. A series of delayed observations arrives to the system each time unit where the service time is an integer-valued random variable. Servers are parallel, and there is no waiting places if all servers are busy. According to this new approach, a server of one of the two loss systems cans serve a series of delayed multiservice observations. The number of served observations will be increased and the number of lost observations will be minimized. This approach is not applied in the papers [8,10–13]. In these papers the problem was discussed by applying special loss systems where servers cannot receive (serve) any observation during the time between any two consecutive arrivals.

* Corresponding author.
E-mail addresses: Mahm112000@yahoo.com (M.A. Mahmoud), Rasha_atwa75@hotmail.com (A.A. Rasha), dr.waseem85@hotmail.com (S.W. Waseem).

http://dx.doi.org/10.1016/j.joems.2017.01.004

However, in the proposed procedure, we introduced a new situation by investigating two loss systems, where the server cans receive an observation or more during the time between any two consecutive series of observations. In fact, the investigation of the mentioned two loss systems was inspired exactly by this application, where it will minimize the number of lost observations. It is proved that the service time of a series equals the sum of service times of its components (observations) where our procedure depends exactly on this new base. The service time of a series is independent of the number of its components and this can be used to increase the number of served observations of the series.

The service time distribution of the loss system, whose arrivals are series of observations, depends on the service time distribution of the loss system whose arrivals are multiservice observations. The probability service time of a series equals the sum of products of probabilities service times of observations without delay terminated by with delay. If the service times of observations are terminated with delay, then all next observations are lost, but if they are terminated without delay, then there is no loss of any observation. The exact bounded service time distribution of a series of delayed observations is obtained by approximating the unbounded geometric random time delay distribution based on the procedure [8], by the bounded service time distribution based on the modified procedure. The approximation depends on the number of multiservice observations of the series as well as it depends on the efficiency of the modified procedure where the number of observations can be increased. This new approach minimizes the number of lost observations and increases the number of served observations, therefore it can be applied to increase the production of items in many fields where items are realized after random time delays. The efficiency of the proposed procedure and its approximation by the efficiency of the procedure [8] are calculated where the results obtained show that the approximation seems to be acceptable. This proves that our proposal can serve as a model of the Robbins–Monro stochastic approximation procedure.

The proposed procedure is more general than the procedure given by the paper [13]. The results in paper [13] can be obtained in the special case that each series contains one observation only.

The investigated procedure is new and we expect that it can be applied to any stochastic approximation or recursive estimation procedure.

2. The modified stochastic approximation procedure with delayed series of delayed multiservice observations for independent random time delay distributions

The stochastic approximation procedure with delayed observations is modified to be applicable in the presence of delayed series of delayed multiservice observations. This application will increase the number of served observations in each series. To obtain the exact service time distribution of the series of delayed observation, the unbounded geometric random time delay distribution based on the procedure [8], is approximated by the bounded service time distribution based on the modified procedure. Two loss systems are investigated where arrivals are series of multiservice observations and a series of delayed observations arrives each time unit, service time is an integer-valued random variable, servers are parallel stations, and there is no waiting places if all servers are busy.

2.1. Stochastic approximation procedure with delayed observations

The Robbins–Monro stochastic approximation procedure with delayed observations has been investigated previously for a geometrical delay distribution [8]. To eliminate or diminish time losses due to delays of observations, it has been proposed that experiments (or observations) are allocated into K parallel series in the following way.

The experiment is based on three essential elements, that is, deterministic arrivals, K parallel series, and no queue. The K series are either open or closed at points $x_{n_k}^{(k)}$, where $1 \le k \le K$, $n_k - 1$ is the number of observations realized in the kth series up to time $n-0$ (i.e. immediately before time n). At the beginning, i. e., before time $n = 1$, all series are open, all n_k are equal to 1, and all $x_1^{(k)}$ are equal to the same constant. At time n, an experiment is made at point $x_{n_i}^{(i)}$, where i is the open series with the smallest n_i and smallest i among them. The ith series is then closed at the same $x_{n_i}^{(i)}$ till time $(n+t(n)+1)-0$ when it opens at the point

$$x_{n_i+1}^{(i)} = x_{n_i}^{(i)} - a_{n_i}\left(r\left(x_{n_i}^{(i)}\right) + e\left(n+t(n)+1, x_{n_i}^{(i)}\right)\right).$$

Here $r(x)$ is a function whose zero point θ is to be found; $e(\upsilon+1, x)$ is the observational error (noise) corresponding to an observation of r made at point x, which becomes known during the interval $[\upsilon,\ \upsilon+1]$; $x_{n_i}^{(i)}$ is the current approximation to θ in the ith series at time $n-0$; $a_n, n \in N$, is a zero sequence of positive constants, typically $a_n = \frac{a}{n+n_0}$, where n_0 is non-negative; and $t(n)$ is [the integer part of] the delay of the result of an experiment made at time n.

If there is no open series at time $n-0$, no experiment is made at time n and a time-loss is thus incurred. If l is the steady state probability of such a time loss, its complement, $e = 1-l$, is called the efficiency of the procedure.

To find θ, the average of the current approximations over all series, $\theta_n = \frac{1}{K}\sum_{k=1}^{K} x_{n_k}^{(k)}$, is chosen as a global approximation of θ at time $n-0$. Under the usual assumptions on function r and error $e(n, x)$, not repeated here, and under the independence of delays $t(n)$, the normed approximation $n^{\frac{1}{2}}(\theta_n - \theta)$ is asymptotically normally distributed with parameters 0 and σ/e, where σ^2 is the asymptotic variance of the same normed approximation in a procedure without delays. Hence, e is also the relative asymptotic efficiency of θ_n as a statistical estimator.

2.2. Description of the first loss system with delayed series of observations

Consider the service system $GI/GI/K/0$ where both the inter-arrival and the service times are non-negative integer values, but their distributions are unspecified, otherwise. K is the total number of servers (or of parallel service stations) in the system. 0 means that there are no waiting places and series are lost if all servers are busy. We will confine ourselves to the case of purely deterministic inter-arrival times, where one series of observations comes each time unit $n = 1, 2, \ldots$ sharp. Such a service system is called a loss system with delayed series of observations and is denoted by $D/GI/K/0$. The service time t is assumed to be an integer valued random variable that however could have originated from a continuous one by off-rounding. The service time t will be rounded down to 0, if the service of a series of observations, who came at time n, is finished by or immediately before the time $n+1$ (i.e., at time $n+1-0$), rounded down to 1, if the service is finished by the time $n+2$ but not before time $n+1$, etc.. Apparently, rounding up (to the next larger integer) would be more natural, but we are inspired by the application treated later on, where a service time not exceeding one time is considered as standard, and where t plays the role of a delay, i.e. of an excess over one time unit.

Denote by $P_0; P_{1,}; P_2; \ldots$ the distribution of the rounded down service time of a series of observations or by $P_0; P_{1,}; P_2, \ldots; P_T$, if the (rounded down) service time, of a series of observations, cannot exceed T time units.

Describe the state of the system at time $n-0$ by a K- duple of integers

$$00\dots0y_{j+1}y_{j+2},\dots,y_K, \tag{1}$$

with

$$0<y_{j+1}\le y_{j+2}\le\dots\le y_K,$$

where j denotes the number of servers free immediately before time n (i.e., at time $n-0$), while $y_{j+1}y_{j+2},\dots,y_K$ denote the remaining service times of the occupied servers in a nondecreasing order. That is, y_i equals 1, if the corresponding server being occupied by a series of observations at time $n-0$, will be freed by time $n+1-0$; y_i equals 2, if he is freed by time $n+2-0$ but not before time $n+1$, etc..

A standard argument of the queuing theory shows that the system with states (1) represents a time-homogeneous, finite or countable Markov chain.

2.3. Description of the second loss system with s observations

Consider the service system $GI_1/GI_1/1/0$ with s observations, where both the interarrival time and the service time may take on only nonnegative values $1/s$ and t respectively, but their distributions are unspecified, otherwise. The total number of servers in the system is 1 only, where all the s observations will be served by this server. In this case, the system has multiservice for the s observations. 0 means that there are no waiting places. Such a service system is a loss system with s multiservice observations and will be denoted by $D_1/GI_1/1/0$, where an observation arrives each $1/s$ time unit. The service time t will be rounded down to 0, if the service of an observation, who came at time n, is finished by the time $n+\frac{1}{s}$; rounded down to 1, if the service is finished by the time $n+\frac{1}{s}+1$ but not before $n+1$, etc.. In general, the service time t will be rounded down to T, if the service of an observation (who came at time n) is finished by the time $n+\frac{1}{s}+T$, where the service time of an observation cannot exceed T time units.

2.4. Application of the two loss systems to the stochastic approximation procedure with delayed observations

Allocation of experiments into K series is not the only possible approach to the problem of stochastic approximation with delays. An alternative approach not discussed in the papers [8,10–13], will be sketched here.

The new approach is investigated by introducing two loss systems with series of delayed multiservice observations and independent random service time. To do this, we modify the Robbins–Monro stochastic approximation procedure to be applicable in the presence of the introduced two loss systems.

The described two loss systems can serve as a model of the stochastic approximation procedure [8] with delayed observations (in case of geometric delay distribution) and with allocation of experiments into K series. The series of delayed observations in the two loss systems are observations in the approximation procedure; the service time of a series is the delay of an observation in the approximation procedure; the K servers of series of delayed observations are the K series of experiments; and the non-existence of a waiting room is the impossibility to accept a series of delayed observations if all the K servers are occupied. According to this approach, a new situation can be obtained by approximating the unbounded geometric delay distribution of the stochastic approximation procedure [8], by the service time distribution of the modified stochastic approximation procedure with series of delayed observations; using the method of moments.

3. Investigation of the compound service time distribution of the introduced loss systems

To find the compound service time distribution of the two loss systems, we prove the following theorem.

Theorem. the compound service times t_1 of s observations equals the service time t of the series of s observations.

Proof. the compound service times t_1 of s observations will be rounded down to 0, if the service of an observation who came at time $n+\frac{x-1}{s}$ is finished by the time $n+\frac{x}{s}$ (i.e., before time $n+\frac{x}{s}$ or at time $+\frac{x}{s}-0$), where the xth observation is served without delay; for all $x=1,2,\dots,s$. In this case, the service of a series of s observations who started at time n is finished by the time $n+1$, and the service time of the series will be rounded down to 0 where it equals the compound service times t_1 of s observations. The compound service times t_1 of s observations will be rounded down to 1, if the xth observation is served without delay by the time $n+\frac{x}{s}$, for all $x=1,2,\dots,s-i$; $i=1,2,\dots,s$, and the $(s-i+1)$th observation is served with time delay (equals one time unit) and finishes its service by the time $n+\frac{x+1}{s}+1$. The observations coming during the service of the $(s-i+1)$th delayed observation will be lost for there are no waiting places. Consequently, the service of a series of s observations is finished by the time $n+2$, and the service time of the series will be rounded down to 1 where it equals the compound service times t_1 of at most s observations. Here at most s observations will be served for, the service with delay of the $(s-i+1)$th observation makes the next coming observations are lost, where the number of lost observations may be $1,2,\dots,s-1$. In general, the compound service times t_1 of s observations will be rounded down to $t=1,2,\dots,T$, if the xth observation is served without delay by the time $n+\frac{x}{s}$; for all $x=1,2,\dots,s-i$; $i=1,2,\dots,s$, and the $(s-i+1)$th observation is served with time delay (equals t time units) and finishes its service by the time $n+\frac{x+1}{s}+t$. Therefore, the service of a series of s observations is finished by the time $n+t+1$, and the service time of the series will be rounded down t where it equals the compound of the service times t_1 of at most s observations. This completes the proof of the theorem.

The compound service time distribution of the introduced two loss systems is investigated in the following way.

3.1. Service time distribution of the second loss system with s observations

Denote by $p_0;p_1;p_2;\dots$, the distribution of the rounded down service time of observations or by $p_0;p_1;p_2;\dots;p_T$, if the (rounded down) service time of an observation cannot exceed T time units.

Let $A_i(t)$ be the event that the xth observation is served without delay with probability p_0, for all $x=1,2,\dots,s-i$; $i=1,2,\dots,s$, and the $(s-i+1)$th observation is served with time delay $t=1,2,\dots T$ time units, with probability p_t. Since the service times of the s observations are independent random variables; it can be seen that

$$\begin{aligned} P(A_i(t)) &= p_0^{s-i}p_t, \quad i=1,2,\dots,s; \qquad t=1,2,\dots,T \\ &= p_0^s, \quad i=1; \qquad t=0. \end{aligned} \tag{2}$$

3.2. Service time distribution of the first loss system with delayed series of s observations

Since the events $A_i(t)$; $i=1,2,\dots,s$; $t=1,2,\dots,T$, are the random service times of a series of s observations, it can be seen that the events are mutually exclusive. The service time distribution of the series is given by the following equation.

$$P\left(\overset{s}{\underset{i=1}{\cup}} A_i(t) = \sum_{i=1}^{s} p(A_i(t)\right) = p_t \sum_{i=1}^{s} p_0^{s-i} = \frac{p_t\left(1-p_0^s\right)}{(1-p_0)} = P_t; \quad t = 1, 2, \ldots, T, \tag{3}$$

where, P_t is the service time distribution of delayed series of s observations for $t = 1, \ldots, T$.

From Eq. (3),

$$P_0 = 1 - \sum_{t=1}^{T} P_t = 1 - \frac{\left(1-p_0^s\right)}{(1-p_0)} \sum_{t=1}^{T} p_t = p_0^s. \tag{4}$$

Therefore, and by Eqs. (3) and (4), the compound service time distribution of the introduced two loss systems is equal to P_t for all $t = 0, 1, \ldots, T$, where

$$\begin{aligned} P_t &= p_0^s; & t &= 0, \\ &= \frac{p_t\left(1-p_0^s\right)}{(1-p_0)}; & t &= 1, 2, \ldots, T\ , \end{aligned} \tag{5}$$

where t is the compound service time of the s observations.

3.3. A special compound service time distribution of the introduced two loss systems

A special service time distribution of series of s observations is obtained by considering the compound service times of the s observations in the case that the service time of the first observation equals t for $t = 1, 2, \ldots, T$ time units. Since there is no waiting places exist in the introduced two loss systems and the server is busy by the service of the first observation for $t = 1, 2, \ldots, T$ time units, then the next coming xth observations for $x = 2, 3, \ldots, s$ will be lost. In this case, no observations are served without delay, and

$$p_0 = 0. \tag{6}$$

Substitute (6) into (5) to get the compound service time distribution P_t of the delayed series of s observations under the assumption that the first observation is the only served one with time delay equals t for $t = 1, 2, \ldots, T$ time units, and the rest $(s-1)$ observations are lost, where

$$\begin{aligned} P_t &= p_t,\ \ t = 1, 2, \ldots, T \\ &= 0,\ \ t = 0, \end{aligned} \tag{7}$$

The $(s-1)$ lost observations are not allowed to be served by other service station of the same loss system even they were free, for this will affect both the number of served arrivals of the two loss systems and the efficiency of the modified stochastic approximation procedure with delayed series of s delayed observations.

As a result of this case, each series of s observations will become a series of one delayed observation and by the last theorem, the service time of each series equals the service time of its delayed observation. Therefore, the service time distribution P_t of a series equals the service time distribution p_t of its delayed observation. We conclude that the number of served observations of a series of s observations can be maximized if all or at least $(s-1)$ of them are served without delay, where this will minimizes the number of lost observations of the series to become zero.

4. Methodology

The first topic of the work was studying mathematical models of two special service systems.

As the considered systems can be viewed as compound Markov chains, it was first of all the theory of finite or countable (discrete-time) compound Markov chains that were made use of. The compound states of the chain reachable from the initial zero compound state, and the equations necessary for the stationary distribution were found. Some matrix algebra and methods of solving systems of linear equations were utilized in connection with the stationary distribution and with asymptotic efficiency.

The second topic of the work was investigating the modified stochastic approximation procedure with delayed series of delayed observations and with allocation of the series into parallel series. Here, the results obtained for the above mentioned service systems became main tool of the investigation. Moreover, results on almost sure convergence and on asymptotic normality, known for the non-delayed stochastic approximation, were made use of.

4.1. The stationary distribution of the compound Markov chain

The set S of compound states that can be reached from the initial 00…0 under the assumption $P_i > 0$, for all $0 \le i \le T$, together with the corresponding matrix of transition probabilities P will be called the basic compound Markov chain. We can easily see, still assuming$P_i > 0$, for all $0 \le i \le T$, that the basic compound Markov chain is irreducible, ergodic. Hence, there is a unique stationary distribution π, determined by the system of equations

$$P^{\mathrm{T}}\pi = \pi, \tag{8}$$

where T denotes the transpose of the matrix P.

4.2. Efficiency of the two service systems

If $T < K$, then each state α contains at least one 0 and in this case no observation will be lost. For $T \ge K$, the unknowns π_α with α containing no 0·s can be easily eliminated from the system (8), as successive transitions from these states to states containing 0·s occur deterministically, with probability 1. One of the remaining equations can always be deleted as superfluous; another one is to be added, namely the requirement

$$\sum_{\alpha} \pi_\alpha = 1.$$

Solving the reduced system of equations and summing the coordinates of the solution, we get the loss probability l,

$$l = \sum_{\alpha} \pi_\alpha$$

α containing no 0, or complementarily, the efficiency e of the two service systems, where

$$e = \sum_{\alpha} \pi_\alpha; \qquad \alpha \text{ contains at least one } 0. \tag{9}$$

Note that,

$$l + e = 1.$$

4.3. Applications of the modified stochastic approximation procedure

Two applications (examples) of the modified procedure are given here, the first depends on delayed series of large number of observations; the second depends on an approximation based on geometric time delay distribution of observations.

4.3.1. Application of the modified procedure to series of large number of observations

Assume that $T = K = 2$; the number of observations $s = 30$; and the service time distribution of an observation equals

Table 1
Comparisons between the exact percentage efficiency e of the modified procedure with delayed series of 30 compound observations each has large probability p_0; and approximate efficiency e_1 of a procedure with geometrically distributed time delay and parameter p.

$s=1$									$K=2$		$K=3$		$K=4$		$K=5$		$K=6$	
p_0	p_1	p_2	p_3	p_4	p_5	p_6	$E(t)$	p	e	e_1	e	e_1	e	e_1	e	e_1	e	e_1
0.96	0.035	0.005	0	0	0	0	0.045	0.957										
0.96	0.020	0.015	0.005	0	0	0	0.065	0.939	99.9									
0.96	0.016	0.010	0.009	0.005	0	0	0.083	0.923	99.8									
0.96	0.015	0.009	0.007	0.005	0.004	0	0.086	0.921	99.7	99.9								
0.96	0.014	0.008	0.007	0.006	0.003	0.002	0.102	0.907	99.6	99.9								
$s=30$																		
0.29	0.62	0.09	0	0	0	0	0.80	0.556	94.6	88.5								
0.29	0.35	0.26	0.10	0	0	0	1.17	0.461	81.5	79.7	98.1	95.5						
0.29	0.28	0.18	0.16	0.09	0	0	1.48	0.403	73.1	72.3	92.9	91.3	99.4	98.4				
0.29	0.26	0.16	0.12	0.09	0.08	0	1.70	0.370	68.7	68.7	89.4	88.4			99.9	99.7		
0.29	0.25	0.14	0.12	0.11	0.05	0.04	1.82	0.355	66.8	66.5	86.9	86.1						

100.0 in all empty cells, and $(P_0, \ldots\ldots\ldots, P_6)$ are approximated to two digits.

$(p_0, p_1, p_2) = (0.7, 0.2, 0.1)$. The compound service time distribution of a series of 30 observations is obtained by substituting $s = 30$, and $(p_0, p_1, p_2) = (0.7, 0.2, 0.1)$ into Eq. (5) to get the service time distribution $(P_0, P_1, P_2) = (2.25^{-5}, 0.6667, 0.316)$ of the series.

Substitute $(P_0,\ P_1,\ P_2)$ in the reduced system of equations (8) and follow the steps stated in Section 4.2 to get the explicit solution for π_α, α contains at least one 0, as $\pi_{00} = 0.14, \pi_{01} = 0.47$, and $\pi_{02} = 0.19$. To obtain the efficiency e, of the modified stochastic approximation procedure, substitute the values of π_{00}, π_{01}, π_{02} in Eq. (9) to get $e = 0.8$.

As a result of this application, the modified procedure can serve a delayed series of large number of observations with high efficiency e. This application can be used to increase the production of items with high efficiency in many fields such as biological, medical, life time experiments, and some industrial projects, where items are realized after random time delays.

4.3.2. Application of the Robbins–Monro procedure to the loss system with series of observations

This application has been done by approximating the efficiency e_1 of the procedure [8], based on geometrical random time delay distribution of observations, by the efficiency e of the modified procedure with delayed series where each series contains 30 observations. The states S_1 of the procedure [8] are the number of servers free at time $n-0$ (or immediately before time n), where S_1 together with the corresponding matrix of transition probabilities P_1 is a Markov chain. Assume that the number of servers K equals 2, it can be seen that

$$S_1 = \{0, 1, 2\},$$
$$P_1 = \begin{pmatrix} q^2 & 2pq & p^2 \\ q^2 & 2pq & p^2 \\ 0 & q & p \end{pmatrix}, \qquad (10)$$

The asymptotic efficiency e_1 is approximated by the asymptotic efficiency e, where e is obtained by the application given in Section 4.3.1 and equals 0.8. The method of moments is used to find the approximation where the first moment μ of the geometric random time delay distribution of procedure [8] is set equals to the first moment $E(t)$ of the compound service time distribution of the investigated loss system to get

$$\mu = E(t),$$

or,

$$(q/p) = \sum_{t=0}^{2} tP_t,$$
$$p = 1\Big/\left(1 + \sum_{t=0}^{2} tP_t\right). \qquad (11)$$

Insert the service time distribution $(P_0, P_1, P_2) = (2.25^{-5}, 0.6667, 0.316)$, obtained by the application given in Section 4.3.1, into Eq. (11) to get the approximated parameter p where $p = 0.435$, and $q = 1 - p = 0.565$. Substitute $p,\ q$ in Eq. (10) to get the approximated transition matrix

$$P_1 = \begin{pmatrix} 0.319 & 0.492 & 0.189 \\ 0.319 & 0.492 & 0.189 \\ 0 & 0.565 & 0.435 \end{pmatrix}.$$

At last, substitute the approximated transition matrix P_1 into the reduced system of equations (8) and follow the steps stated in section (4.2) to solve for the unknowns π_0, π_1, π_2. It can be seen that the approximated efficiency $e_1 = 0.761$, where

$$e_1 = 1 - \pi_0$$

The results obtained show that the approximation of the efficiency e_1 by the efficiency $e = 0.8$ is acceptable and consequently, the investigated loss system with delayed series of s observations can serve as a model of the Robbins–Monro stochastic approximation procedure with delayed observations. This model modifies the stochastic procedure to be applicable in the presence of delayed series of s compound observations. Our procedure is called the modified stochastic approximation procedure with delayed series of s compound observations.

4.4. New results and discussion

Table 1 gives a comparison between the exact efficiency e of the modified procedure with delayed series of 30 delayed compound observations, and the approximate efficiency e_1 of the stochastic procedure [8] with delayed observations. In Table 2, the efficiency e in the case $s = 1$ is approximately the same as in the case $s = 30$ and this gives an important application by increasing the production of items with high efficiency in many fields such as biological, medical, life time experiments, and industrial projects where items are realized after random time delays.

The results obtained by Tables 1 and 2 seem to be satisfactory, with the difference between the two efficiencies e and e_1 decreasing. Consequently, this show in general that the modified stochastic approximation procedure with delayed series of s compound

Table 2
Comparisons between the exact percentage efficiency e of the modified procedure with delayed series of 30 compound observations each has small probability p_0; and approximate efficiency e_1, of a procedure with geometrically distributed time delay and parameter p.

$s=1$																		
									$K=2$		$K=3$		$K=4$		$K=5$		$K=6$	
p_0	p_1	p_2	p_3	p_4	p_5	p_6	$E(t)$	p	e	e_1	e	e_1	e	e_1	e	e_1	e	e_1
0.04	0.4	0.56	0	0	0	0	1.52	0.397	74.4	71.2								
0.04	0.32	0.32	0.32	0	0	0	1.92	0.342	64.8	63.3	88.6	84.2						
0.04	0.24	0.24	0.24	0.24	0.24	0	0	2.4	0.294	56.4	55.6	79.1	76.6	94.8	90.6			
0.04	0.192	0.192	0.192	0.192	0.192	0	2.88	0.258	49.9	49.4	71.2	69.7			97.7	94.3		
0.04	0.16	0.16	0.16	0.16	0.16	.16	3.36	0.229	44.7	44.4	64.5	63.6					99	97
$s=30$																		
0	0.417	0.583	0	0	0	0	1.58	0.387	73.1	69.8								
0	0.333	0.333	0.334	0	0	0	2.00	0.333	63.5	61.9	87.5	82.9						
0	0.250	0.250	0.250	0.250	0	0	2.50	0.286	55.0	54.2	77.7	75.1	94.0	89.5				
0	0.2	0.2	0.2	0.2	0.2	0	3.00	0.250	48.5	48.1	69.6	68.1			97.3	93.4		
0	0.167	0.167	0.167	0.167	0.166	0.166	3.50	0.229	43.4	43.1	62.9	61.9					99	96

100.0 in all empty cells, and $(P_0, \ldots\ldots\ldots, P_6)$ are approximated to three digits.

observations can serve as a model of the Robbins–Monro stochastic approximation procedure with delayed observations. Accordingly, the modified procedure improves the Robbins–Monro procedure by diminishing the time losses due to the delay of observations. The results in Tables 1 and 2 are calculated by using the Matlab program (version 7) and Maple program (version 16).

The results obtained by Tables 1 and 2 prove that the investigated procedure is more general than the procedure given by the paper [13] for all the results given by this paper can be obtained in the special case that each series contains one observation only.

In conclusion, our approach is new and we expect that it can be applied to other stochastic approximation or recursive estimation procedures.

References

[1] H. Robbins, S. Monro, A stochastic approximation method, Ann. Math. Stat. 22 (1951) 400–407.
[2] A.E. Albert, L.A. Gamder, Stochastic Approximation and Nonlinear Regression, MIT Press, Cambridge, 1967.
[3] Ya.I. Albert, C.E. Chidume, J. Li, Stochastic approximation method for fixed point problems, Appl. Math. 3 (2012) 2123–2132.
[4] R. Combes, An introduction to stochastic approximation, Literature (2013) 1–9.
[5] J. Harold, G. Kushner, G. Yin, Stochastic approximation and recursive algorithm and applications, Appl. Math. 35 (1997) xxi, 417.
[6] M.B. Nevel'son, R.Z. Khasmin'skii, Stochastic Approximation and Recursive Estimation, Nauka, Moscow, 1972 English trans. Amer. Soc.
[7] Z. Xu, Y. Dai, A stochastic approximation frame algorithm with adaptive directions, Numer. Math. Theory Methods Appl. 1 (2008) 460–474.
[8] V. Dupač, U. Herkenrath, Stochastic approximation with delayed observations, Biometrika 72 (1985) 683–685.
[9] A.M. Mahmoud, On Robbins–Monro stochastic approximation with delayed observations, in: Proceedings of the Fourth Prague Symposium on Asymptotic Statistics, 1988, pp. 399–403.
[10] M.A. Mahmoud, A.A. Rasha, Stochastic approximation with compound delayed observations, Math. Comput. Appl. 10 (2005) 283–289.
[11] M.A. Mahmoud, A.A. Rasha, Investigation of the process D / GI / K / 0 loss system for independent random time Delay distribution, in: Proceedings of the 9th International Conference on Ordered Statistical Data and their Applications, Zagazig University, 2010.
[12] M.A. Mahmoud, A.A. Rasha, Stochastic approximation and compound delayed observations with independent random time delay distribution, Arab J. Sci. Eng. 36 (2011) 1549–1558.
[13] M.A. Mahmoud, A.A. Rasha, S.W. Waseem, Stochastic approximation with delayed components of observations and exact decreasing random time delay distribution, Sylwan J. 159 (5) (2015) 138–146 (ISSN: 0039-7660).
[14] Y. Cheung, M.S.V. Elkind, Stochastic approximation with virtual observations for dose-finding on discrete levels, Biometrika 97 (2010) 109–121.