INITIAL TRANSIENT PERIOD DETECTION FOR STEADY-STATE QUANTILE ESTIMATION

Jong-Suk R. Lee, Don McNickle[†] and Krzysztof Pawlikowski
Department of Computer Science and [†]Department of Management
University of Canterbury
Christchurch, New Zealand
E-mail: {ruth, krys}@cosc.canterbury.ac.nz

KEYWORDS

Steady-State Simulation, Quantile Estimation, Initial Transient Period Detection

ABSTRACT

Performance of stochastic dynamic systems such as computer and telecommunication networks is often assessed by quantiles of, for example, system's response times or experienced delays. Thus, estimation of quantiles during stochastic simulation of computer and telecommunication networks is an important issue. However, unlike estimation of mean values, estimation of quantiles requires that simulation output data (observations) have to be stored for processing in several steps. Thus, QE (quantile estimation) requires large amounts of computer storage as well as longer computation time than for example, estimation of mean values. Several approaches for estimating quantiles in the method of RCs (regenerative cycles) and other methods have been proposed to reduce these problems. For sequential steady-state QE using the methods, like batch means and spectral analysis, one must first eliminate the effect of the initial transient period. However, heuristic rules or statistical tests specifically for detecting the initial transient period when estimating the steady-state quantiles have not been developed yet. In this paper, we study properties of statistical tests for detecting the initial transient period (originally developed for estimating mean values), when applied to estimating steady-state quantiles in a sequential simulation. We have also investigated statistical tests in terms of appropriateness for implementing them in a fully automated simulation package.

1 INTRODUCTION

Simulators are often more concerned with the extreme performance of simulated systems characterised by quantiles, than with its average behaviour, when simulating a dynamic stochastic system, such as a computer or a telecommunication network. Quantiles are convenient measures of entire ranges of values of simulation output data, and are particularly useful for planning necessary capacities for various resources, comparing the overall performance of alternative designs or establishing minimum standards of performance. Therefore, from a practical point of view, the problem of reliable estimators of quantiles during stochastic simulation is quite important.

The pth quantile of a random variable X equals Q_p if

$$Pr[X \leq Q_p] = p.$$

Let $x_1 \leq x_2 \leq \ldots \leq x_n$, $x_i \geq 0$, be the ordered sequence of n observations of a random variable X, collected during the simulation. The usual point estimator of the pth quantile is given by

$$\hat{Q}_p(n) = x_{\lfloor n*p+1 \rfloor}, \qquad 0 < p < 1 \tag{1}$$

where $\lfloor z \rfloor$ denotes the integral part of z.

The problem with using this formula is that, especially in the case of correlated sequences of observations, the length of sample sequence required for achieving an adequate statistical error of the pth quantile $\hat{Q}_p(n)$ can be very large and impossible to predict in advance. Direct application of Equation (1) in quantile estimation (QE) requires large amounts of computer storage for storing the entire sequence of observations since sorting through the entire sequence of output data is required whenever a new observation is recorded (Heidelberger and Lewis 1984; Jain and Chlamtac 1985).

Clearly, repeating storing and sorting of the entire sequence for a QE is impractical in the long runs required for estimating steady-state parameters. Several approaches for avoiding these storing and sorting difficulties when estimating quantiles and their statistical error in the method of RCs and other methods have been proposed for non-sequential procedures in Heidelberger and Lewis 1984; Iglehart 1976; Jain and Chlamtac 1985; Seila 1982.

Another problem commonly encountered in simulating systems, especially using method other than RCs, is that the performance measures observed during the initial transient period of a simulation run are not typical of the system's true behaviour. It normally takes some time for the effect of the starting conditions to become insignificant and for the simulation model to reach steady-state. If the problem is to obtain performance measures for the model under steady-state conditions, then the analyst must eliminate the initialisation bias of observations collected during an initial transient period of a simulation run.

To do this, ideally, observations should be collected only after the system has reached steady-state. A problem with this approach is that how one can recognise that steady-state has been achieved. A number of ways to estimate the length of the initial transient period of steady-state simulations have been proposed in (Cash et al. 1992; Goldsman et al. 1994; Schruben 1982; Schruben et al. 1983). Basic problems related with the existence of initial transient periods can be found, for example, in Pawlikowski 1990; Randhawa and Baxter 1992; Wilson and Pritsker 1978. The length of the initial transient period has traditionally been determined using different heuristic rules.

More precise measures of the length of the initial transient could be obtained by using various statistical tests invented to test the stationarity of data sequences. These tests operate in a hypothesis testing framework, formally testing the null hypothesis that *there is no initialisation bias in the output mean* against the alternate hypothesis that initialisation bias in the output exists.

Our motivation is to find a robust initial transient period detector especially for steady-state QE which could be used in practical applications of sequential steady-state simulation and could also be implemented in a fully automated simulation package like Akaroa-2 (Ewing et al. 1999). In Section 2, we have summarised two statistical tests for detecting initialisation bias. In Section 3, we discuss the automatic detection of the initial transient period for steady-state QE. We also present the numerical results of two sequential detection methods of initial transient period, especially applying them in a fully automated simulation package Akaroa-2 (Ewing et al. 1999) in Section 4. Finally, our findings are summarised in Section 5.

2 STATISTICAL TESTS

In this section, we summarise two statistical tests for detecting the initial transient period in steady-state mean estimation.

2.1 Schruben Test

Stationarity tests based on variance estimation methods using standardised time series were first proposed by Schruben (Schruben 1982) using the maximum estimator of the variance, and improved on by Schruben et al. (Schruben et al. 1983) especially using a weighted area estimator.

Rejection or acceptance of the hypothesis that the given subsequence of observations is stationary, or equivalently, that the initial transient period is not included in the observations, depends on the statistic calculated from the considered sequence. Despite the sophisticated theory hidden behind these tests they are quite simple numerically, and can be applied to a wide class of simulated processes.

A practical problem when implementing one of these tests is that they require a priori knowledge of the steady-state variance σ^2 of the simulated process which is not normally available when the test is applied because the system is still in its initial transient period. These tests solve this problem by estimating the steady-state variance over the latter portion of the collected data. This is done on the assumption that this latter portion of data is more representative of the steady-state behaviour of the system, thus giving a better estimate of the steady-state variance. The effectiveness of this test is strongly dependent on how accurate the variance estimator is.

2.2 Goldsman, Schruben and Swain Test

The Goldsman, Schruben and Swain (GSS) test which is a natural generalisations of previously proposed tests in (Schruben 1982) and (Schruben et al. 1983), was proposed and tested in (Goldsman et al. 1994). In this test, observations $x_1, x_2, \cdots, x_n$ are divided into b batches of length m (assume $n = bm$). Variance estimators based on the first b' batches are compared to the corresponding estimators from the remaining $b - b'$ batches. If these estimators are deemed to be significantly different, then we say that an initial transient is present, and would likely bias point estimators of the steady-state mean.

The empirical results of the GSS test using different variance estimators of the batch means, standardised time series, and linear combinations of these estimators give a good agreement with the more extensive studies in Cash et al. 1992; Schruben 1982; Schruben et al. 1983. Furthermore, they stated that using the maximum estimator, and the combination of batch means and maximum estimator are the most powerful tests for detecting the presence of initial transient bias. For the GSS test, the compromise choice of $b \approx 8$ with the large enough batch size m and $b'/b \approx 0.75$ were recommended in Cash et al. 1992.

3 AUTOMATIC DETECTION OF INITIAL TRANSIENT PERIOD FOR STEADY-STATE QE

A common way of dealing with such an initialisation bias is to take a large (possibly wasteful) number of observations in an attempt to overwhelm the initialisation effects. Another method is simply to delete (truncate) a portion of the output from the beginning of the simulation run. This presumably allows the simulation to warm up before data are retained for analysis. The simulation experimenter would then hope that the biasing observations had been eliminated or their effects ameliorated. Of course, if the output is truncated too early, then significant bias might still be present. If it is truncated too late, then good observations are lost.

Most of the proposed heuristic rules and statistical tests for detecting the initial transient period are developed for the case of estimating the steady-state mean of the system. Specific heuristic rules or statistical tests of detecting the initial transient period especially for estimating steady-state quantiles have not been developed yet.

Two methods for automatic detection of the length of the initial transient period have been proposed: one proposed by Pawlikowski (Pawlikowski 1990) based on Schruben's test (Schruben 1982) with the method of spectral analysis used for the variance estimator (Heidelberger and Welch 1981), and another one proposed by Jackway and deSilva (Jackway and deSilva 1992) based on the same Schruben's test (Schruben 1982) with an autoregressive method used for the variance estimator. The former initial transient period detection method is implemented in a simulation package Akaroa-2 (Ewing et al. 1999). For estimating the steady-state mean, the initial transient period is well automatically and sequentially detected in Akaroa-2.

A sequential procedure for estimating steady-state quantiles based on the method of spectral analysis for the variance of the quantile estimates has been proposed in (Raatikainen 1990). The procedure has not been fully automated since the length of the initial transient period for the steady-state QE has to be decided beforehand. Another method for estimating steady-state quantiles in sequential simulation has been proposed in (Lee et al. 1999). Unlike the method proposed by Raatikainen (Raatikainen 1990), this approach uses a fully automated detection procedure for the initial transient period. Therefore, we have an immediate question if the currently available detection methods of the initial transient period proposed for estimating the steady-state mean are directly applicable for detecting the initial transient period especially when estimating the steady-state quantiles.

To check whether the initial transient period detection method implemented in Akaroa-2 is suitable for sequential QE or not, we can validate the initial transient period detection method of Akaroa-2 in a simple way. One suggestion for the validation is that just after the initial transient period the value of (say) the 0.9 quantile of the waiting time in the queue for an $M/M/1/\infty$ queueing system with the traffic intensity $\rho = 0.8$ should be very close to its theoretical steady-state 0.9 quantile, which can be calculated by

$$\max\left(0, \frac{E[w]}{\rho} \ln[10 * \rho]\right), \qquad (2)$$

where $E[w]$ is the steady-state mean waiting time in the queue for the $M/M/1/\infty$ queueing system. If not, some of the results on quantiles may be suspect since it means that the initial transient period detection method in Akaroa-2 applied when estimating steady-state quantiles does not work properly, i.e., initialisation bias still exists even after deleting all observations collected in the initial transient period.

4 NUMERICAL RESULTS

First, to see the suitability of the initial transient period detection method implemented in Akaroa-2 when estimating the steady-state quantiles, we did the above simple suggestion. The results are depicted in Figure 1. The empirical values have been obtained from 100,000 independent waiting times of the first recorded customer over 100,000 independent simulation runs for the $M/M/1/\infty$ queueing system, after passing the initial transient period which has been suggested by the initial transient period detection method originally implemented in Akaroa-2.

The empirical 90th quantile and the empirical mean waiting time in the queue for the first recorded cus-

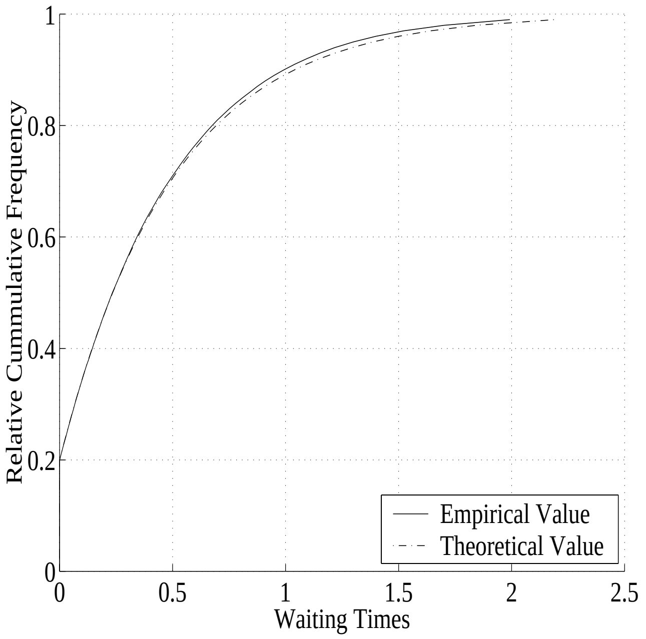

Figure 1: Comparisons of theoretical and empirical values when using Schruben's test for detecting an initial transient period ($M/M/1/\infty$ at traffic intensity $\rho = 0.8$)

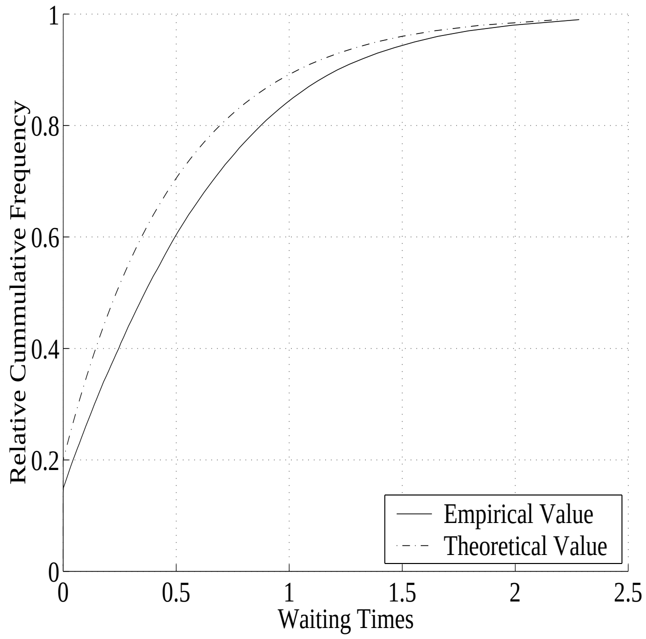

Figure 2: Comparisons of theoretical and empirical values when using the GSS test for detecting an initial transient period ($M/M/1/\infty$ at traffic intensity $\rho = 0.8$)

tomer after passing the initial transient period were calculated from 100,000 independent waiting times recorded after passing the initial transient period from each of 100,000 simulation runs, and compared with the theoretical steady-state values, which can be calculated by using Equation (2), of the mean waiting time in the queue for the $M/M/1/\infty$ queueing system.

The simple experiment is showing that the 90th quantile of the $M/M/1/\infty$ queueing system when estimating the mean waiting time in the queue is indistinguishable from its steady-state quantile - they are extremely close; see Figure 1. In such a situation we have the first evidence that the length of initial transients for quantiles and means should be close. This experiment is very modest, but at least gives some justification for using the initial transient period detection method, originally developed and applied for detecting the initial transient period in the case of estimating the steady-state means, implemented in Akaroa-2 especially when estimating quantiles. Therefore, an initial transient period detector developed for applying when estimating the steady-state means does seem to work adequately for quantiles as well.

Secondly, to find the better statistical test for detecting an initial transient period for steady-state quantiles, we have also investigated the performance of the GSS test based on the maximum estimator of the standardised time series which was reported as giving the best performance in (Cash et al. 1992; Goldsman et al. 1994). The GSS test is also implemented in the simulation package Akaroa-2 (Ewing et al. 1999). The empirical and the theoretical results of the GSS test and Schruben's test are depicted in Figure 2.

When comparing Figure 1 and Figure 2, the GSS test, based on the maximum estimator of the standardised time series, clearly shows much worse performance than Schruben's test, based on the method of spectral analysis, since the empirical values are much far from the theoretical values.

Another comparison of Schruben's test and the GSS test is presented in Table 1. These empirical statistics are obtained from 100,000 independent simulation replications (the same data of Figure 1 and Figure 2) for $M/M/1/\infty$ queueing system at traffic intensity $\rho = 0.8$ with the required relative statistical error of 10% or less, at the 0.95 confidence level.

The mean waiting times in steady-state when applying Schruben's test and the GSS test; (see the second column of the third and the fourth rows of Table 1,) are calculated by averaging 100,000 independent waiting times collected just after the initial transient period from 100,000 independent simulation replications. The

Table 1: Statistics obtained after applying Schruben's test and the GSS test for detecting the initial transient period

	Waiting Time	Total Observations	Transient Observations	Relative Length of Transient
Theory	0.4	47,443	326	0.69%
Schruben's Test	0.384	42,104	455	1.08%
GSS Test	0.507	41,434	390	0.94%

numbers of total observations and transient observations in the third and the fourth column of the third and the fourth rows of Table 1 are obtained by averaging the final total observations and the final transient observations recorded from 100,000 independent sequential steady-state simulation executed in Akaroa-2 with Schruben's test and the GSS test for detecting the initial transient period, respectively. The last column of the third and the fourth rows of Table 1 is the proportion of transient observations (the fourth column) over total observations (the third column).

The theoretical convergence of the mean waiting time in the $M/M/1/\infty$ queueing system at traffic intensity ρ = 0.8 to steady-state can be depicted in Figure 3 (McNickle 1991). (Note, we have assumed 6 initial customers at time zero since it gives interesting non monotone convergence to steady-state.) If we make the assumption that we are in steady-state when the mean waiting time is very close (within 0.0001 and 0.0005) of the steady-state value, then we can achieve steady-state within 0.0001 and 0.0005 after 432 and 326 customers, respectively.The theoretical mean waiting time and the theoretically required total and transient observations, with the required relative statistical error of 10% or less at the 0.95 confidence level, in the $M/M/1/\infty$ queueing system at traffic intensity $\rho = 0.8$ are also presented along with the empirically obtained results in Table 1.

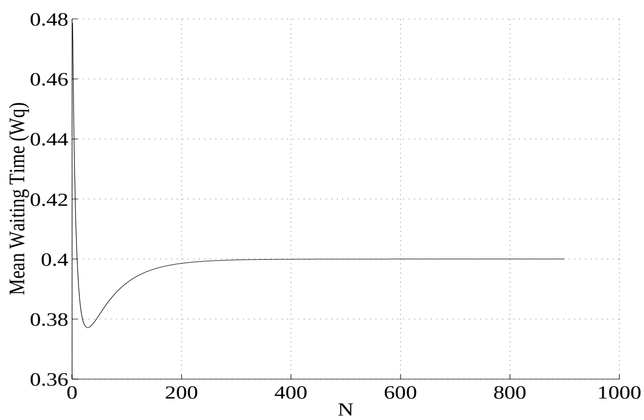

Figure 3: Theoretical convergence of waiting time in $M/M/1/\infty$ queueing system at load 0.8 (Initial customer number at time zero = 6)

As we can see from Table 1, the mean waiting time obtained from Schruben's test is much closer to the theoretical value than the mean waiting time obtained from the GSS test. This suggests that Schruben's test for detecting the initial transient period may be better than the GSS test since the mean waiting time calculated from 100,000 first customers recorded after passing the initial transient period should be very close to the theoretical value. To eliminate or reduce the effect of the initial transient period for estimating steady-state parameters, in general, deleting a little bit longer length of the initial transient period detected using Schruben's test seems desirable.

5 CONCLUSIONS

In this paper, we have studied properties of statistical tests of detecting the initial transient period, especially in terms of the propriety of the current state of the art in the initial transient period detection methods when applying them for estimating steady-state quantiles in a sequential steady-state simulation. The experiment is done in a very modest way, but at least gives some justification for using Schruben's test for detecting the initial transient period implemented in Akaroa-2 when estimating quantiles. Therefore, an initial transient period detector originally developed and applied for estimating steady-state means does seem to work adequately for estimating steady-state quantiles as well. Schruben's test for detecting the initial transient period seems better than the GSS test since Schruben's test suggests a longer end point of the initial transient period.

REFERENCES

Cash, C. R.; D. G. Dippold; J. M. Long; and W. P. Pollard. 1992. "Evaluation of Tests for Initial-Condition Bias." *Proceedings of the 1992 Winter Simulation Conference*, 577-585.

Ewing, G.; K. Pawlikowski; and D. McNickle. 1999. "Akaroa-2: Exploiting Network Computing by Distributed Stochastic Simulation." *13th European Simulation Multiconference* (Warsaw, Poland), 175-181.

Goldsman, D.; L. W. Schruben; and J. J. Swain. 1994. "Tests for Transient Means in Simulated Time Series." *Naval research Logistics Quarter*, 41, 171-187.

Heidelberger, P. and P. A. W. Lewis. 1984. "Quantile Estimation in Dependent Sequences." *Operations Research*, 32, no. 1: 185-209.

Heidelberger, P. and P. D. Welch. 1981. "A Spectral Method for Confidence Interval Generation and Run Length Control in Simulations." *Communications of the ACM* 25: 233-245.

Iglehart, D. L. 1976. "Simulating Stable Stochastic Systems, VI: Quantile Estimation." *Journal of the Association for Computing Machinery*, 23, no. 2: 347-360.

Jackway, P. T. and B. M. deSilva. 1992. "A Methodology for Initialization Bias Reduction in Computer Simulation Output." *Asia-Pacific Journal of Operational Research*, 9, 87-100.

Jain, R. and I. Chlamtac. 1985. "The P^2 Algorithm for Dynamic Calculation of Quantiles and Histograms Without Storing Observations." *Communications of the ACM*, 28, no. 10: 1076-1085.

Lee, J. R.; D. C. McNickle; and K. Pawlikowski. 1999. "Quantile Estimation in Sequential Steady-State Simulation." *13th European Simulation Multiconference* (Warsaw, Poland), 168-174.

McNickle, D. C. 1991. "Estimating the Average Delay of the First N Customers in an $M/Erlang/1$ Queue." *Asia-Pacific Journal of Operational Research*, 8, No. 1: 44-53.

Pawlikowski, K. 1990. "Steady-State Simulation of Queueing Processes: A Survey of Problems and Solutions." *ACM Computing Surveys*, no. 2: 123-170.

Raatikainen, K. E. E. 1990. "Sequential Procedure for Simultaneous Estimation of Several Percentiles." *Transactions of The Society for Computer Simulation*, 7, no. 1: 21-44.

Randhawa, S. U. and L. K. Baxter. 1992. "A Study in the Application of Schriber's Truncation Rule to Simulation Output." *Transactions of The Society for Computer Simulation*, 9, no. 3: 175-192.

Schruben, L. W. 1982. "Detecting Initialization Bias in Simulation Output." *Operations Research*, 30, no. 3: 569-590.

Schruben, L. W.; H. Singh; and L. Tierney. 1983. "Optimal Tests for Initialization Bias in Simulation Output." *Operations Research*, 31, 1167-1178.

Seila, A. F. 1982. "A Batching Approach to Quantile Estimation in Regenerative Simulations." *Management Science*, 28, no. 5: 573-581.

Wilson, J. R. and A. A. B. Pritsker. 1978. "A Survey of Research on the Simulation Startup Problem." *Simulation*, 31, 55-58.

AUTHORS' BIOGRAPHIES

JONG-SUK RUTH LEE is currently a PhD student in computer science from the University of Canterbury, Christchurch, New Zealand and a student member of IEEE. Her research interests are sequential quantitative stochastic simulation and distributed simulation of telecommunication networks.

DON MCNICKLE is a Senior Lecturer in Management Science on the Department of Management, University of Canterbury, New Zealand. His research interests include queueing theory and statistical aspects of simulation. He received a PhD in Mathematics from the University of Auckland, New Zealand. Member of ORSA.

KRZYSZTOF PAWLIKOWSKI is an Associate Professor (Reader) in Computer Science at University of Canterbury, Christchurch, New Zealand. He received his PhD degree in Computer Engineering from the Technical University of Gdansk, Poland. The author of over 80 research papers and four books. His research interests include stochastic simulation, distributed processing, ATM, optical and wireless telecommunication networks, and teletraffic modelling. Senior member of IEEE.