Sequential Estimation of Variance in Steady-State Simulation

Adriaan Schmidt[1], Krzysztof Pawlikowski[1], and Don McNickle[2]

Departments of
[1]Computer Science and Software Engineering
[2]Management
University of Canterbury
Christchurch, New Zealand

Technical Report
TR-COSC 03/08
University of Canterbury
December 2008

Contents

List of Tables

List of Figures

Chapter 1

Introduction

Today, many studies of communication networks rely on simulation conducted to assess their performance. Steady-state simulation is used to draw conclusions about the long-run behaviour of stable systems.

Current methodology of analysis of output data from steady-state simulation focuses almost exclusively on the offline estimation of the *steady-state means* of the parameters under investigation. Thus, the literature on "variance estimation" mostly deals with the estimation of the *variance of the mean*, which is needed to construct a confidence interval of the estimated mean values. So far, little work has been done on the estimation of the steady-state variance of simulated processes.

In the performance analysis of communication networks, we find applications where the *packet delay variation* or *jitter* is of interest. In audio or video streaming applications, networking packets should take approximately the same time to arrive at their destination; the delay itself is less important (see e.g. Tanenbaum, 2003). To find the jitter of a communication link, the variance of the packet delay times needs to be estimated.

The theoretical background of this research includes sequential steady-state simulation, stochastic processes, basic results on the estimation of the steady-state mean, and stochastic properties of the variance. These are briefly summarised in Chapter 2.

The aim of this research is the sequential (online) estimation of the *steady-state variance*, along with the *variance of the variance* which is used to construct a confidence interval of the estimate. To this end, we propose and evaluate several variance estimators in Chapter 4.

As an additional focus, we investigate the initial transient period of simulation output, and try to find methods of automated, sequential detection of the end of this period in Chapter 3.

The research that led to this report was based on *Akaroa2*, an automated, parallel simulation controller developed at the *University of Canterbury* in Christchurch, New Zealand. We implement the proposed variance estimators with the help of the Akaroa2 framework, and assess their performance experimentally. The results of these experiments are presented in Chapter 5.

Chapter 2

Background

This chapter introduces the theoretical background used in the remainder of this report. It outlines the type of simulation experiments considered and the properties of the stochastic processes we expect as output from these experiments. It then gives a short overview of important properties of estimators and states some existing results for the estimation of mean values and variances. Finally, it briefly introduces the concept of *Multiple Replications in Parallel*.

2.1 Discrete-Event Simulation

When investigating a system using a simulation study, the first step is to create a model that abstractly describes the system. From the model, we create a simulation program which, when run, takes random numbers as input, and produces as output observations of the process of interest. These observations are then analysed using statistical methods, and from this we try to draw conclusions about the performance of the original (real-world) system. In the case of *sequential output analysis*, the analysis procedure also determines the run length of the simulation (see Section 2.3). Figure 2.1 illustrates this general setup.

Two types of discrete-event simulation can be distinguished: *terminating simulation* and *steady-state simulation*. In a terminating simulation, the model under investigation has a natural stopping point, at which the simulation ends. This can be a specific time or a certain event. The initial conditions of the model, reflected in the initial state of the simulation program, have to be carefully chosen as they can have great influence on the outcome of the simulation.

In *steady-state simulation*, the focus lies on analysing the long-run behaviour

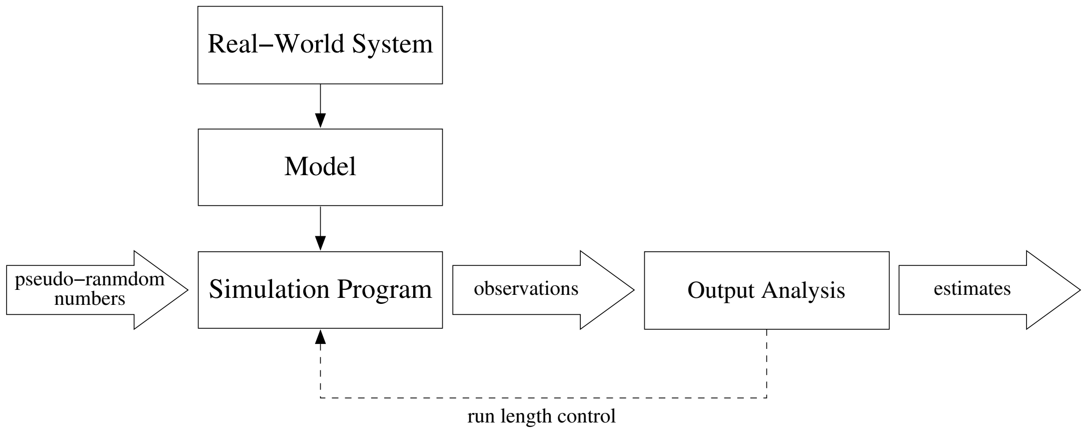


Figure 2.1: General Setup of a Simulation Study Using Online Sequential Output Data Analysis

of a stable system. Usually, the initial state of the model is not representative of this behaviour, and there is a warm-up period before the model reaches its steady state. The observations during this *initial transient period* are believed to often bias the results, and are discarded.

This report deals with the automated analysis of the output of a steady-state simulation.

2.1.1 Stochastic Processes

The output of a discrete-event simulation is a sequence of *observations*[1]

$$\{x\} = x_1, x_2, \ldots .$$

Since a simulation program takes random numbers as input, the observations it produces are also random, and can be regarded as the realisation of a discrete-time stochastic process

$$\{X\} = X_1, X_2, \ldots .$$

We observe two phases in the output sequence of a simulation: an *initial transient* phase, in which the observations are very much dependent on the initial conditions of the simulation program, and a *stationary* or *steady-state* phase, in

[1]Throughout this report, we denote random variables by uppercase letters and observations, which are realisations of random variables, by lowercase letters.

which the influence of the initial conditions is no longer visible. We denote the last observation of the initial transient phase by x_{n_0}:

$$\{x\} = \underbrace{x_1, x_2, \ldots, x_{n_0}}_{\text{initial transient}}, \underbrace{x_{n_0+1}, x_{n_0+2}, \ldots}_{\text{steady state}}.$$

n_0 is a random variable, and is also referred to as the *truncation point* for the initial transient phase of simulation.

The length of the initial transient phase depends on the influence of the initial conditions on the system's behaviour, and varies from replication to replication. In the mean, we can see convergence to a steady state, but for a single replication, no universally accurate length of the initial transient phase can be determined.

Once a simulation has entered its steady state, we assume that its output data can be represented by a wide-sense stationary and ergodic process. This means that for $i, j > n_0$, the mean

$$\nu = \mathrm{E}[X_i]$$

is the same for all i, and the covariance

$$\mathrm{Cov}[X_i, X_j] = \mathrm{E}[(X_i - \nu)(X_j - \nu)]$$

depends only on the distance $|i - j|$, and not on the absolute values of i and j. Furthermore, we can deduce the statistical properties of the process from one long run of the simulation experiment.

Throughout this report, we use the following symbols to describe parameters of a process:

$$\begin{aligned} \nu &= \mathrm{E}[X_\infty] && \text{is the steady-state mean} \\ \sigma^2 &= \mathrm{Var}[X_\infty] && \text{is the steady-state variance} \\ \rho_j &= \frac{\mathrm{Cov}[X_i, X_{i+j}]}{\sigma^2} && \text{is the autocorrelation of lag } j \end{aligned}$$

It is important to note that simulation output data are usually correlated, so when analysing the output of a simulation, we have to assume that $\rho_j \neq 0$.

2.2 Estimation

In general, the estimate of a performance measure consists of a *point estimate* and an *interval estimate*, defining a region in which the actual parameter lies with a given probability.

The estimator of a parameter θ is $\hat{\Theta}$, and we denote the point estimate by $\hat{\theta}$, and the lower and upper bounds of the confidence interval by $\hat{\theta}_l$ and $\hat{\theta}_h$, respectively. We can then write

$$\Pr\left[\hat{\theta}_l \leq \theta \leq \hat{\theta}_h\right] \geq 1 - \alpha.$$

The probability $1 - \alpha$ is called the *confidence level.* If the confidence interval is symmetric around the point estimate $\hat{\theta}$, it is often described by a single parameter Δ, defining the half width of the interval:

$$\Pr\left[\hat{\theta} - \Delta \leq \theta \leq \hat{\theta} + \Delta\right] \geq 1 - \alpha.$$

Given a point estimate and a confidence interval, we can give two measures of precision: the *absolute precision* ϵ_a is defined as the half-width of the confidence interval, Δ, and the *relative precision* is defined as

$$\epsilon_r = \frac{\Delta}{\hat{\theta}}.$$

In the investigation of estimators, several properties are of special interest:

The *bias* describes a systematic deviation of the estimator from the real value of the parameter:

$$\text{Bias}[\hat{\Theta}] = \text{E}[\hat{\Theta} - \theta].$$

An estimator is *unbiased*, if

$$\text{E}[\hat{\Theta}] = \theta.$$

The *variance* of the estimator measures the (squared) deviation of the estimator from its mean:

$$\text{Var}[\hat{\Theta}] = \text{E}[(\hat{\Theta} - \text{E}[\hat{\Theta}])^2].$$

The *mean square error* (MSE) describes the (squared) deviation of the estimator from the real value of the parameter:

$$\text{MSE}[\hat{\Theta}] = \text{E}[(\hat{\Theta} - \theta)^2].$$

With the above definitions, the MSE can also be expressed as

$$\mathrm{MSE}[\hat{\Theta}] = \left(\mathrm{Bias}[\hat{\Theta}]\right)^2 + \mathrm{Var}[\hat{\Theta}].$$

Strictly speaking, when we talk about an *estimator* $\hat{\Theta}$, we usually mean a *series of estimators* $\hat{\Theta}(n)$, $n = 1, 2, \ldots$ that estimate the parameter θ from the first n observations of the process. Such a series of estimators is *consistent*, if for all ε

$$\lim_{n\to\infty} \Pr[|\hat{\Theta} - \theta| \geq \varepsilon] = 0.$$

Lehmann (1983, p. 332, Theorem 1.1) shows that to prove consistency, it is sufficient to show that the MSE of the estimator $\rightarrow 0$ as the sample size $n \rightarrow \infty$. For an unbiased estimator $\hat{\Theta}$, this means that the estimator is consistent if $\mathrm{Var}[\hat{\Theta}] \rightarrow 0$, as $n \rightarrow \infty$.

2.2.1 Mean Value Estimation

The well-known consistent estimator of the mean ν is the sample average

$$\overline{X}(n) = \frac{1}{n}\sum_{j=1}^{n} X_j. \tag{2.1}$$

The variance of $\overline{X}(n)$ is (see Law and Kelton, 1991, p. 285)

$$\mathrm{Var}[\overline{X}(n)] = \frac{\sigma^2}{n}\left(1 + 2\sum_{j=1}^{n-1}(1 - j/n)\rho_j\right). \tag{2.2}$$

It is easy to see that $\mathrm{Var}[\overline{X}(n)] = \frac{\sigma^2}{n}$ for uncorrelated observations. Since simulation output data are usually correlated, using $\frac{\sigma^2}{n}$ as the variance of the mean value leads to wrong confidence intervals. In the case of queueing systems, the output data usually have positive correlation coefficients $\rho_j > 0$, so $\frac{\sigma^2}{n}$ underestimates the true variance of the mean, and the resulting confidence intervals will be optimistically narrow.

A number of methods are known to overcome this problem, among which are several methods based on batched means, approaches using standardized time series, and estimators using spectral analysis. A survey of techniques has been published by Pawlikowski (1990).

Given the variance of the mean, $\mathrm{Var}[\overline{X}(n)]$, we define the half-width of the confidence interval at a confidence level of $1-\alpha$ as

$$\Delta_{\overline{X}(n)} = t_{d,1-\alpha/2}\sqrt{\mathrm{Var}[\overline{X}(n)]},$$

where $t_{d,1-\alpha/2}$ is the $1-\frac{\alpha}{2}$ quantile of the t-distribution with d degrees of freedom. The exact variance is usually not known, and an estimate has to be used instead. Depending on the method of estimation, different ways of obtaining $\mathrm{Var}[\overline{X}(n)]$ and the degrees of freedom of the t-distribution can be used.

2.2.2 Variance Estimation

For *independent* and *identically distributed* random variables, the well-known consistent point estimate of the variance is

$$S^2(n) = \frac{1}{n-1}\sum_{j=1}^{n}(X_j - \overline{X}(n))^2. \tag{2.3}$$

In this case, the variance of $S^2(n)$ is (see Wilks, 1962, p. 200)

$$\mathrm{Var}[S^2(n)] = \frac{1}{n}\left(\mu_4 - \frac{n-3}{n-1}\,\sigma^4\right), \tag{2.4}$$

where μ_4 is the fourth central moment of the steady-state distribution, and σ^2 is the steady-state variance.

When dealing with *correlated* observations, however, $S^2(n)$ is no longer unbiased. Based on a result by Anderson (1971, p. 448), Law and Kelton (1991, p. 284) show that

$$\mathrm{E}[S^2(n)] = \sigma^2\left(1 - 2\,\frac{\sum_{j=1}^{n-1}(1-j/n)\rho_j}{n-1}\right). \tag{2.5}$$

Note that in this case, $S^2(n)$ is asymptotically unbiased for $n \to \infty$, but it is not easy to say at which point n becomes large enough to consider the bias negligible. Also, in the case of correlated observations, estimating $\mathrm{Var}[S^2(n)]$ becomes highly impractical.

It is difficult to make a general statement about the distribution of $S^2(n)$. If the X_i are independent and have a normal distribution, we know that $S^2(n)$, in this case the sum of squares of normal random variables, follows a χ^2-distribution with $n-1$ degrees of freedom. Trivedi (2002, p. 663) notes that for mean values, the violation of the assumption of normally distributed random variables does not severely affect the generated confidence intervals; however, the confidence intervals of the variance derived using the χ^2-distribution can be poor when the distribution of the X_i is significantly different from the normal distribution.

For the case of independent and identically distributed random variables, Cramér (1946, Section 28.1) shows that the sums of any of their moments follows the central limit theorem, and the distribution is asymptotically normal. Since we can only use $S^2(n)$ if we are dealing with uncorrelated observations, we will in this case assume that $S^2(n)$ is normally distributed for large n.[2]

2.2.3 Assessing the Quality of Estimators

Apart from the bias, variance and mean square error, which make valuable statements about the quality of an estimator, another useful and important measure is the *coverage*. It describes the frequency with which the true value of a parameter is covered by the confidence interval. At a confidence level of $1-\alpha$, the theoretical (expected) coverage is also $1-\alpha$. However, due to assumptions made about the process, which might not be met in practice, actual coverage can deviate from its expected value. Coverage is assessed experimentally, and is used as the main instrument of analysing the estimators presented in Chapter 4 of this report.

2.3 Sequential Analysis

When running a simulation to evaluate the performance of a system, we are concerned with the correctness of the results. Therefore, we are not only interested in finding an estimate for a performance parameter, but also a confidence interval, inside which the parameter resides with a given probability (see Pawlikowski, Jeong, and Lee, 2002).

Three parameters occur in the result of the output data analysis: the simulation

[2] If a process is ϕ-mixing (see Billingsley, 1968), then the central limit theorem holds even in the case of correlated observations. Because the ϕ-mixing property is not easy to work with, we do not go into more detail on this. Note, however, that most processes encountered in practice are, in fact, ϕ-mixing.

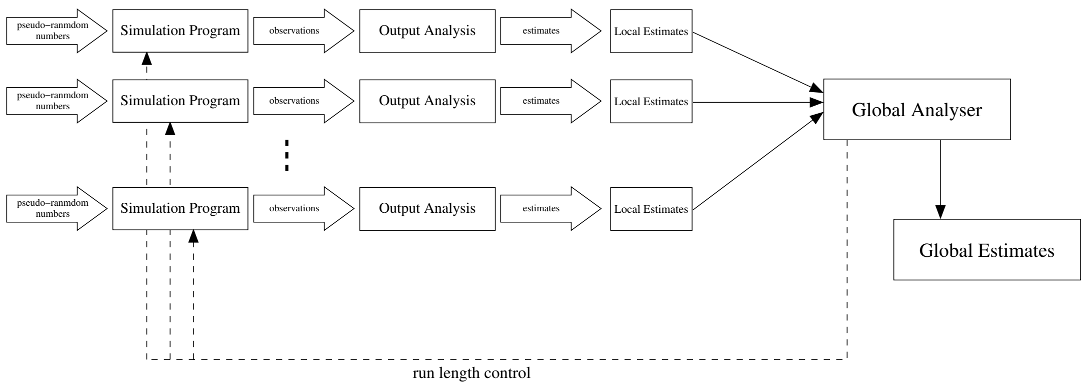


Figure 2.2: General Setup of a Simulation Study Using Sequential Online Output Data Analysis and Multiple Replications in Parallel

run length, the confidence level $1 - \alpha$, and the width of the confidence interval.

When running a simulation experiment, we can chose values for any two of these, and will, in this way, influence the third. If, for example, the confidence level and the simulation run length are fixed, there is no way of knowing in advance the width of the confidence interval.

In sequential simulation, the confidence level and the required precision are given, and the simulation is run until those requirements are met. This makes it necessary to periodically check the current width of the confidence interval at a given confidence level, and then either stop the simulation, if the required precision is reached, or otherwise continue it to collect more observations.

2.3.1 Multiple Replications in Parallel

Time needed to complete a simulation can be significantly reduced if the simulation is executed in parallel. For mean values, Pawlikowski, Yau, and McNickle (1994) describe how the method of *Multiple Replications in Parallel* can be used to combine local estimates of several computers executing the same simulation program with different streams of pseudo-random numbers into one global estimate of the mean value.

Figure 2.2 shows the setup of a simulation experiment using Multiple Replications in Parallel. Several instances of the same simulation program are run simultaneously, each using different pseudo-random numbers and thus producing different estimates. Within the ith instance of simulation, the local estimate consists of a point estimate $\hat{\nu}_i$, the variance of the estimate $\mathrm{Var}[\hat{\nu}_i]$, the degrees of freedom d_i, and the number of observations, n_i, used to obtain the estimate. It is assumed that the estimates are t-distributed.

A global analyser then combines these local estimates into one global estimate of the mean

$$\hat{\nu} = \frac{\sum_i n_i \hat{\nu}_i}{\sum_i n_i},$$

with a variance of

$$\mathrm{Var}[\hat{\nu}] = \frac{\sum_i n_i^2 \, \mathrm{Var}[\hat{\nu}_i]}{\left(\sum_i n_i\right)^2}.$$

The resulting estimate is assumed to have a t-distribution with $\sum_i d_i$ degrees of freedom. When the global estimate has reached the desired precision, all instances of the simulation program are stopped.

Pawlikowski and McNickle (2001) show that the speedup achieved using this method follows a truncated version of Amdahl's law. For practical applications and a reasonable degree of parallelisation, the speedup is almost linear in the number of computers used.

Akaroa2

The automated simulation controller Akaroa2 implements the procedure of Multiple Replications in Parallel to speed up simulations (see Ewing, Pawlikowski, and McNickle, 2003). The three main functions of Akaroa2 are the estimation of mean values, the stopping of simulations once a required precision of estimates is reached, and the parallel execution of the simulation program on multiple computers in a local area network. To calculate confidence intervals for its estimates, Akaroa2 implements both the method of nonoverlapping batch means (see e.g. Pawlikowski, 1990) and the method of spectral analysis, as proposed by Heidelberger and Welch (1981). Akaroa2 also detects the end of the initial transient period by a combination of a heuristic and a statistical test.

The variance estimators developed in this report were integrated into the Akaroa2 framework, which made it possible to use several subroutines already present in Akaroa2. Among the existing functions used are the mean value estimation procedure based on spectral analysis, a routine to test if a sequence of batch means is uncorrelated, and the detection of the initial transient.

Chapter 3

Detection of the Initial Transient

In steady-state simulation, the focus lies on estimating the long run behaviour of a stable system. When running a simulation, initial values have to be chosen for the model variables, and these are generally not typical of this long run behaviour. Thus, the first observations collected are usually not representative of the system in steady state, and can introduce a bias to the estimates.

Reliable automated detection of the end of the initial transient remains an open problem. There are some heuristic “rules of thumb” to decide when a process has reached *stationarity of the mean value*. In addition, a few statistical tests are known which try to test the stationarity of a sequence of observations.

There is an ongoing discussion, whether initial observations should be discarded or not. Findings by McNickle, Ewing, and Pawlikowski (2008) indicate that, with the long simulation run lengths made possible by recent advances in computing technology, the problem of bias due to an initial transient period may become insignificant. However, McNickle et al. (2008) also show that discarding initial observations may have other benefits in sequential simulation, especially improving the accuracy of the simulation run length.

When estimating the steady-state variance, we want the process to have reached *variance stationarity* before we begin collecting observations for estimation. No previous work has been done on this problem, and here we do not try to develop new methods for detecting the truncation point of a process. Instead, we try to adapt existing methods that are used in the estimation of mean values. As a starting point, we take the methods implemented in the simulation package Akaroa2 (Ewing et al., 2003), where detection of the initial transient is attempted using a combination of a heuristic and a statistical test.

3.1 Initial transient period

Two measures are known that can describe the initial transient period of a queueing simulation theoretically. The *relaxation time* is a measure of the rate at which the mean waiting time tends to its steady-state value, which has been analysed for a number of different processes. Based on a result by Odoni and Roth (1983), Jackway and de Silva (1992) formulate the length of the initial transient period of the waiting times in an M/M/1 queue as

$$\tau_n = \frac{8}{2.8(1-\sqrt{\rho})^2},$$

where ρ is the system load.

The other, more comprehensible measure is derived from the expected waiting time of the customers in the system. Kelton and Law (1985) show how to calculate the expected waiting time of the nth customer. We then say that the initial transient period ends with the first customer whose expected waiting time lies within a small distance of the steady-state value. This method can easily be modified to find the expected variance of the waiting time of the nth customer.

Figure 3.1 shows the theoretical measures of the length of the initial transient period for the waiting time in the M/M/1 queue. For the expected waiting time and variance of the nth customer, it is assumed that stationarity is reached when the value reaches 99% of the steady-state value. We see that the measure based on the expected waiting time closely matches the relaxation time estimate. Furthermore, the expected variance of the waiting time reaches its steady-state value approximately 1.5 times slower than the expected waiting time. This would, at least in this case, make the theoretical length of the initial transient period with regard to the process variance about 50% longer than that with regard to the process mean.

3.2 Heuristic Rules

A number of heuristic rules have been proposed to decide when a process has reached stationarity of its mean value. Pawlikowski (1990) summarises some of those rules; we try to adapt one of them to find the end of the initial transient with regard to the variance.

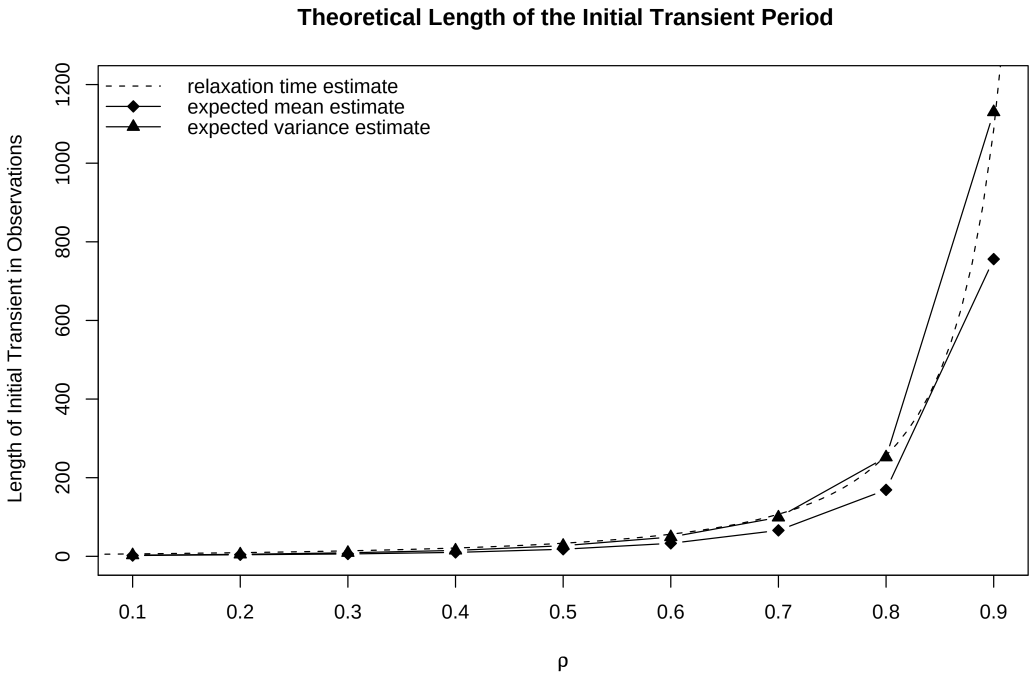


Figure 3.1: Theoretical Length of the Initial Transient Period of the Waiting Time in the M/M/1 Queue

3.2.1 Rule of "25 Crossings"

We use rule R5 of Pawlikowski (1990), which is also implemented in Akaroa2, as a starting point in our investigation:

> *The initial transient period is over after n_0 observations if the time series $x_1, x_2, \ldots, x_{n_0}$ crosses the mean $\overline{X}(n_0)$ k times.* (Pawlikowski, 1990)

The heuristic was evaluated by Gafarian, Ancker, and Morisaku (1978), and they recommend a value of $k = 25$.

To adapt this method to find the end of the initial transient period of a sequence $\{x\}$ with regard to the second moment of the process, we apply it to the modified sequence $\{y\}$, defined by

$$y_i = x_i^2.$$

3.3 Statistical Tests

A number of statistical tests have been proposed to determine if a series of observations contains initial (transient) observations that can cause a bias in the estimation of the mean value.

We investigate in detail a test by Schruben, Singh, and Tierney (1983) that is based on the concept of *standardised time series*. Other tests proposed include a more recent test by Goldsman, Schruben, and Swain (1994), and a randomisation test by Yücesan (1993).

3.3.1 Schruben Test

The test proposed by Schruben et al. (1983) is designed to decide if a given sequence of observations is stationary, or if it has an initialisation phase that might cause a bias. It is based on the test statistic

$$T = \sqrt{\frac{45}{n^3 \hat{\sigma}_0^2}} \sum_{k=1}^{n} \left(1 - \frac{k}{n}\right) k \left(\overline{X}(n) - \overline{X}(k)\right), \tag{3.1}$$

with n being the length of the tested sequence, $\overline{X}(i) = \frac{1}{i} \sum_{k=1}^{i} X_k$ the average of the first i observations of the process, and $\hat{\sigma}_0^2$ an estimate of the *steady state variance constant* $\sigma_0^2 = \lim_{n\to\infty} n \operatorname{Var}[\overline{X}(n)]$. If $|T| < t_{d,1-\alpha/2}$, the hypothesis of the absence of initialisation bias is accepted, otherwise the sample is rejected as having an initialisation bias. The degrees of freedom d of the t-distribution are calculated during the estimation of the variance constant σ_0^2.

The constant σ_0^2 is characteristic of the process, and has to be estimated from the observations. Because we assume that later observations are more representative of the process in steady state than earlier ones, we only use the second half of the tested sequence to estimate σ_0^2.

In their paper, Schruben et al. (1983) report that the test detects initialisation bias in the waiting times of an M/M/1 queueing system, if it is applied to the resulting process when averaging over ten replications. When using unmodified data from a single replication, however, the test seems to fail. In addition, the test statistic seems to be very sensitive to the size of the window of observations it is applied to. For different system loads, appropriate window sizes can be experimentally determined (as in the case of averaging over ten replications

Window Size	Proportion of Rejected Samples			
	Biased Sample		Unbiased Sample	
	$\rho = 0.9$	$\rho = 0.95$	$\rho = 0.9$	$\rho = 0.95$
200	0.40	0.47	0.23	0.29
400	0.36	0.46	0.26	0.29
600	0.37	0.46	0.25	0.30
800	0.27	0.41	0.23	0.28
1000	0.29	0.40	0.20	0.25
1200	0.25	0.29	0.17	0.22
1400	0.22	0.38	0.18	0.23
1600	0.19	0.33	0.19	0.22
1800	0.26	0.32	0.15	0.27
2000	0.16	0.27	0.18	0.26

Table 3.1: Evaluation of the Test Proposed by Schruben et al. (1983), Applied to a Single Replication of the Waiting Times in an M/M/1 Queue

reported in Schruben et al. (1983)), but we were unable to find a window size that yields good results for a large range of system loads.

Table 3.1 shows the results of the experiments with the Schruben test, applied to a single replication of the waiting times in an M/M/1 queue with loads of $\rho = 0.9$ and $\rho = 0.95$. For different window sizes, the test was applied to a biased and an unbiased sample. The biased sample was obtained from the beginning of the simulation output, and the unbiased sample was taken after the simulation had already produced $10,000$ observations. The results were obtained using sequential simulation with an absolute error of 0.05 at the 0.9 confidence level. Although in general, a slightly larger proportion of the biased than the unbiased samples is rejected, the test can not be called reliable in this case.

Because our goal is to find the end of the initial transient, given only a *single replication*, we conclude that we will note use this test in the context of a single server queueing system, as it is doubtful that it will perform better for more complex models.

3.4 Summary

The aim of this chapter was to find a detection method of the initial transient of the process variance. It is reasonable to suspect that applying a working method of detection of the truncation point for mean values to the squares of the original observations may yield a useful detection method. This could, however, not be verified, because no reliable method of detecting the truncation point for mean values could be found.

There are heuristics that help determine the length of the initial transient period of the mean value. One of those, the rule of "25 Crossings", was modified to find the truncation point with regard to the variance. Because heuristic rules tend to underestimate the length of the initial transient of the mean in highly loaded systems, (see Gafarian et al., 1978) it can not be expected that the modified rule will work reliably if used in simulation models with a high load.

Experience during the simulation experiments presented in Chapter 5 has shown that the overall results of the variance estimation are acceptable, although no reliable methods for detecting variance stationarity were used. Instead, the procedure used in Chapter 5 is the one that had already been implemented in Akaroa2: a combination of the unmodified "25 Crossings" heuristic and the Schruben test of Section 3.3.1, applied to a single replication of the process.

Chapter 4

Estimation of the Steady-State Variance

This chapter introduces several approaches for the estimation of the steady-state variance of a simulated process.

In the first section, an approach is presented, which uses existing estimators of the process mean to obtain point and interval estimates of the variance. The remaining sections of this chapter introduce different methods for a direct estimation of the process variance, based on the sample variance S^2. We have seen in Chapter 2 that, if the observations are correlated and the sample size finite, S^2 is a biased estimator of σ^2.

The proposed estimators deal with this problem in different ways: The estimator $\hat{\sigma}^2_{IR}$ uses independent replications of the process. Estimators $\hat{\sigma}^2_{SP}$ and $\hat{\sigma}^2_{VB}$ use uncorrelated observations from one replication to estimate the variance, whereas $\hat{\sigma}^2_{BM}$ compensates for the bias of S^2.

In this chapter, we assume that any initial transient phase has already been removed from the process, and that $x_1, x_2, \ldots$ are the first observations in steady state.

4.1 Variance as a Mean Value

This first approach to variance estimation results from the definition of the variance as

$$\begin{aligned}\mathrm{Var}[X] &= E[(X - E[X])^2] &(4.1)\\ &= E[X^2] - (E[X])^2. &(4.2)\end{aligned}$$

From this, we derive two estimators: $\hat{\sigma}^2_{M1}$ from (4.1) and $\hat{\sigma}^2_{M2}$ from (4.2). At first glance, these two methods could be expected to be equal, but we will see that they calculate the confidence interval in different ways and produce different results in terms of coverage. Both estimators make use of existing procedures for the estimation of means to give point and interval estimates of the steady-state variance. This makes the implementation of these two estimators easy.

To obtain estimates of the mean, a number of different techniques could be employed; here, we use a method based on spectral analysis, as proposed by Heidelberger and Welch (1981). The method has been shown to produce good results of coverage of the mean value by Pawlikowski et al. (1998).

The resulting expressions for the lower and upper bounds of the confidence interval are

$$\overline{X}_l(n) = \overline{X}(n) - t_{d,1-\alpha/2}\sqrt{\mathrm{Var}[\overline{X}(n)]}$$
$$\overline{X}_h(n) = \overline{X}(n) + t_{d,1-\alpha/2}\sqrt{\mathrm{Var}[\overline{X}(n)]}$$

where the details of obtaining $\mathrm{Var}[\overline{X}(n)]$ and the degrees of freedom d of the t-distribution can be found in the paper by Heidelberger and Welch. Other methods of obtaining point and interval estimates of the mean may be used, in which case different expressions for $\overline{X}_l(n)$ and $\overline{X}_h(n)$ result.

4.1.1 The Estimator $\hat{\sigma}^2_{M1}$

For the first estimator, based on (4.1), we transform the observations:

$$y_i = (x_i - \bar{x}(i))^2,$$

where x_i denotes the ith original observation, and $\bar{x}(i) = \frac{1}{i}\sum_{j=1}^{i} x_j$ is the average of the first i observations. We then estimate the variance of the original sequence $\{x\}$ as the mean value of the new sequence $\{y\}$:

$$\hat{\sigma}^2_{M1}(n) = \frac{1}{n}\sum_{i=1}^{n} y_i.$$

The confidence interval is then symmetric:

$$\bar{y}_l(n) < \sigma^2 < \bar{y}_h(n),$$

where $\bar{y}_l(n)$ and $\bar{y}_h(n)$ are calculated by the chosen mean value estimator.

Algorithm 1 Estimator $\hat{\sigma}^2_{M1}$

Require: Instance of the Mean Value Estimator

$\Sigma \leftarrow 0$, $n \leftarrow 0$

repeat

$x \leftarrow$ GetObservation

$\Sigma \leftarrow \Sigma + x$

$n \leftarrow n + 1$

$y \leftarrow (x - \Sigma/n)^2$

SubmitToEstimatorOfMean(y)

$\hat{\sigma}^2_{M1}$, $\Delta \leftarrow$ CurrentEstimateOfMean

until RequiredPrecision($\hat{\sigma}^2_{M1}$, Δ)

return $\hat{\sigma}^2_{M1}$, $[\hat{\sigma}^2_{M1} - \Delta, \hat{\sigma}^2_{M1} + \Delta]$

4.1.2 The Estimator $\hat{\sigma}^2_{M2}$

This method makes use of (4.2), separately estimating $\mathrm{E}[X]$, using

$$\bar{x}(n) = \frac{1}{n}\sum_{i=1}^{n} x_i$$

and obtaining a symmetric confidence interval

$$\bar{x}_l(n) < \mathrm{E}[X] < \bar{x}_h(n),$$

and $\mathrm{E}[X^2]$, using

$$\overline{x^2}(n) = \frac{1}{n}\sum_{i=1}^{n} x_i^2$$

with a symmetric confidence interval

$$\overline{x^2}_l(n) < \mathrm{E}[X^2] < \overline{x^2}_h(n).$$

The two estimates are then combined into

$$\hat{\sigma}^2_{M2}(n) = \overline{x^2}(n) - (\bar{x}(n))^2\,,$$

with an asymmetric confidence interval constructed according to Law (1983):

$$\overline{x^2}_l(n) - (\bar{x}_l(n))^2 < \sigma^2 < \overline{x^2}_h(n) - (\bar{x}_h(n))^2\,.$$

Again, $\bar{x}_l(n)$, $\bar{x}_h(n)$, $\overline{x_{2l}}(n)$, and $\overline{x^2}_h(n)$ are calculated by the chosen estimator of the mean.

Algorithm 2 Estimator $\hat{\sigma}^2_{M2}$

Require: Two Instances of the Mean Value Estimator
  repeat
    $x \leftarrow$ GetObservation
    SubmitToEstimatorOfMean1(x)
    SubmitToEstimatorOfMean2(x^2)
    $\hat{\nu}_1$, $\Delta_1 \leftarrow$ CurrentEstimateOfMean1
    $\hat{\nu}_2$, $\Delta_2 \leftarrow$ CurrentEstimateOfMean2
    $\hat{\sigma}^2_{M2} \leftarrow \hat{\nu}_2 - \hat{\nu}_1^2$
    $\Delta_l \leftarrow (\hat{\nu}_2 - \Delta_2) - (\hat{\nu}_1 - \Delta_1)^2$
    $\Delta_h \leftarrow (\hat{\nu}_2 + \Delta_2) - (\hat{\nu}_1 + \Delta_1)^2$
  until RequiredPrecision($\hat{\sigma}^2_{M2}$, $[\Delta_l, \Delta_h]$)
  return $\hat{\sigma}^2_{M2}, [\Delta_l, \Delta_h]$

4.2 Uncorrelated Observations

Once the simulated process has entered its steady state, the variance of a single random variable is

$$\mathrm{Var}[X_i] = \sigma^2, \quad \forall i > n_0,$$

where n_0 denotes the end of the inital transient period.

It is possible to use (2.3), the sample variance $S^2(n)$, as a point estimator, if the observations used are (almost) uncorrelated. Only in this case is $S^2(n)$ unbiased, and its variance, as expressed in (2.4), can be used to construct a confidence interval.

One solution to ensuring that observations are uncorrelated is to run independent replications of the simulation model, and to only use one single observation from the steady state of each replication. Another possibility is to use observations from one long simulation run that are far apart, and assume those to be uncorrelated.

Both of these methods have the obvious disadvantage that they discard most of the generated observations, either as part of the initial transient period or because they are correlated. However, they avoid statistical problems typical for correlated data.

4.2.1 The estimator $\hat{\sigma}^2_{IR}$

This estimator runs independent replications of the simulation model, and uses only one single observation from the steady state of each replication.

The resulting observations are

$$\begin{array}{lll} x_{1,1}, \quad x_{1,2}, \quad \ldots & & \text{first replication} \\ x_{2,1}, \quad x_{2,2}, \quad \ldots & & \text{second replication} \\ \quad\quad\vdots & & \\ x_{i,1}, \quad x_{i,2}, \quad \ldots & & i\text{th replication} \\ \quad\quad\vdots & & \end{array}$$

and the secondary data used to estimate the variance are the observations

$$y_i = x_{i,n_{0,i}},$$

where $n_{0,i}$ is the truncation point of the ith replication.

It is obvious that $\mathrm{Var}[Y_i] = \mathrm{Var}[X_\infty] = \sigma^2$, and because the y_i are obtained from independent replications of the process, and are thus independent and identically distributed, the sample variance $S^2(n)$ can be used as an unbiased, consistent estimate of σ^2, and a confidence interval can be calculated according to (2.4).

To be practically used in a sequential procedure, (2.3) and (2.4), $S^2(n)$ and its variance, are transformed, and the true values of the process mean and variance, ν and σ^2, are replaced by their estimates $\bar{y}$ and $\hat{\sigma}^2_{IR}$, respectively.

This leads to the point estimate

$$\hat{\sigma}^2_{IR} = \frac{1}{n-1}\sum_{j=1}^{n} y_j^2 - \frac{n}{n-1}\,\bar{y}^2,$$

where

$$\bar{y} = \frac{1}{n}\sum_{j=1}^{n} y_j$$

is the sample mean of the sequence $\{y\}$.

The variance of $\hat{\sigma}^2_{IR}$ is estimated using (2.4):

$$\begin{aligned}\mathrm{Var}[\hat{\sigma}^2_{IR}] &= \frac{1}{n}\left(\sum_{j=1}^{n}(y_j-\bar{y})^4 - \frac{n-3}{n-1}\,\hat{\sigma}^4_{IR}\right)\\ &= \frac{1}{n^2}\sum_{j=1}^{n} y_j^4 - \frac{4}{n^2}\bar{y}\sum_{j=1}^{n} y_j^3 + \frac{6}{n^2}\bar{y}^2\sum_{j=1}^{n} y_j^2 - \frac{3}{n}\bar{y}^4 - \frac{n-3}{n(n-1)}\,\hat{\sigma}^4_{IR},\end{aligned}$$

and the resulting confidence interval at the $1-\alpha$ confidence level is

$$\hat{\sigma}^2_{IR} - \Delta_{IR} < \sigma^2 < \hat{\sigma}^2_{IR} + \Delta_{IR},$$

with

$$\Delta_{IR} = z_{1-\alpha/2}\sqrt{\mathrm{Var}[\hat{\sigma}^2_{IR}]}.$$

Because no general results exist on the distribution of $S^2(n)$, and we consider cases in which n is large, we assume a normal distribution, and $z_{1-\alpha/2}$ is the $1-\frac{\alpha}{2}$ quantile of the standard normal distribution.

We can see that an implementation of this estimator only needs to save the sums over y_j, y_j^2, y_j^3, and y_j^4 to calculate the point and interval estimates.

The main problem of this estimator, apart from the many observations it discards as part of the initial transient periods, is that it relies on the accurate detection of the truncation point n_0. As seen in Chapter 3, no truly reliable method of automated detection of the truncation point was found. Therefore, it can be expected that the estimator $\hat{\sigma}^2_{IR}$ is not useful for practical purposes.

Algorithm 3 Estimator $\hat{\sigma}^2_{IR}$

$n \leftarrow 0$
$\Sigma y \leftarrow 0,\ \Sigma y^2 \leftarrow 0,\ \Sigma y^3 \leftarrow 0,\ \Sigma y^4 \leftarrow 0$
repeat
  RestartSimulationModel
  repeat
    $x \leftarrow$ GetObservation
  until InitialTransientPeriodOver
  $y \leftarrow$ GetObservation
  $n \leftarrow n + 1$
  $\Sigma y \leftarrow \Sigma y + y$
  $\Sigma y^2 \leftarrow \Sigma y^2 + y^2$
  $\Sigma y^3 \leftarrow \Sigma y^3 + y^3$
  $\Sigma y^4 \leftarrow \Sigma y^4 + y^4$
  $\bar{y} \leftarrow \Sigma y / n$
  $\hat{\sigma}^2_{IR} \leftarrow \Sigma y^2/(n-1) - n\bar{y}^2/(n-1)$
  $v \leftarrow \Sigma y^4/n^2 - 4\bar{y}\Sigma y^3/n^2 + 6\bar{y}\Sigma y^2/n^2 - 3\bar{y}^4/n - (\hat{\sigma}^2)^2(n-3)/n/(n-1)$
  $\Delta \leftarrow z_{1-\alpha/2}\sqrt{v}$
until RequiredPrecision($\hat{\sigma}^2_{IR}$, Δ)
return $\hat{\sigma}^2_{IR}$, $[\hat{\sigma}^2_{IR} - \Delta, \hat{\sigma}^2_{IR} + \Delta]$

4.2.2 The estimator $\hat{\sigma}^2_{SP}$

If the absolute values of the autocorrelation coefficients of the process $\{X\}$, $|\rho_j|$, decrease monotonically, then it is possible to obtain an approximately uncorrelated sub-sequence of the original observations by considering observations that are far apart.

Assuming that we know the lag k_0, at which the correlation between observations becomes negligible, we analyse the observations

$$y_i = x_{k_0 i}, \quad i = 1, 2, \dots. \tag{4.3}$$

Because we assume the process $\{X\}$ to be stationary, with $\mathrm{Var}[X_i] = \sigma^2$ for all i, it is obvious that $\{Y\}$ is also a stationary process with $\mathrm{Var}[Y_i] = \sigma^2$.

To find an appropriate value for the spacing k_0, we successively test different values by extracting the respective subsequences $\{y\}$ and analysing their autocorrelation. Once a value of k_0 is found that yields a sequence $\{y\}$ with non-significant autocorrelation, this value is accepted, and estimation of the variance is started.

As a criterion for the non-significance of the autocorrelation it is required that the autocorrelation coefficients up to a lag of L are sufficiently small. The procedure to test for autocorrelation in a sequence $\{y\}$ is the same procedure used to test batch means for autocorrelation in the method of batch means, as it is implemented in Akaroa2. It is described in detail by Pawlikowski (1990). For the experiments presented in this report, we use a value of $L = 20$.

Before a certain value of k_0 is tested, enough observations have to be collected to make a reliable estimate of the autocorrelation of $\{y\}$ possible. The test is performed as soon as $50L$ observations of the sequence $\{y\}$, or $50Lk_0$ observations of the original sequence $\{x\}$, have been collected.

Estimation of the variance of $\{y\}$ is performed using the method already described in Section 4.2.1, leading to an unbiased, consistent estimate of σ^2.

Just like the estimator $\hat{\sigma}^2_{IR}$, this method of estimation is inefficient, because all observations x_j, $j \neq k_0 i$ are discarded. In addition, it has to rely on the correct estimation of autocorrelations to find k_0. If the resulting value of k_0 is too small, the sequence $\{y\}$ will still be correlated, and $S^2(n)$ can no longer be used as an estimator of σ^2.

Algorithm 4 Estimator $\hat{\sigma}^2_{SP}$

```
n ← 0
Σy ← 0, Σy² ← 0, Σy³ ← 0, Σy⁴ ← 0
k₀ ← DetermineK0
repeat
   for i = 1 to k₀ − 1 do
      x ← GetObservation
   end for
   y ← GetObservation
   n ← n + 1
   Σy ← Σy + y
   Σy² ← Σy² + y²
   Σy³ ← Σy³ + y³
   Σy⁴ ← Σy⁴ + y⁴
   ȳ ← Σy/n
   σ̂²_SP ← Σy²/(n − 1) − nȳ²/(n − 1)
   v ← Σy⁴/n² − 4ȳΣy³/n² + 6ȳΣy²/n² − 3ȳ⁴/n − (σ̂²)²(n − 3)/n/(n − 1)
   Δ ← z_{1−α/2} √v
until RequiredPrecision(σ̂²_SP, Δ)
return σ̂²_SP, [σ̂²_SP − Δ, σ̂²_SP + Δ]
```

4.3 "Vertical Batches"

In this section, we extend the estimator $\hat{\sigma}^2_{SP}$ to make it more efficient by using all observations from the steady-state phase of the process.

The new estimator batches observations, but unlike conventional batching procedures, which collect consecutive observations into batches and continually increase the number of batches, this method uses a fixed number of batches and places consecutive observations into different batches.

$$\begin{array}{ccccc}
X_1, & X_2, & X_3, & \dots, & X_k, \\
X_{k+1}, & X_{k+2}, & X_{k+3}, & \dots, & X_{2k}, \\
\vdots & \vdots & \vdots & & \vdots \\
X_{ki+1}, & X_{ki+2}, & X_{ki+3}, & \dots, & X_{ki+k}, \\
\vdots & \vdots & \vdots & & \vdots \\
\underbrace{}_{\text{batch 1}} & \underbrace{}_{\text{batch 2}} & \underbrace{}_{\text{batch 3}} & & \underbrace{}_{\text{batch k}}
\end{array}$$

We use $k \geq k_0$ batches, where k_0 is the lag at which the correlation between observations is no longer significant, as seen in the previous section. Each batch contains observations that are far apart and therefore approximately uncorrelated, so we can use (2.4), the sample variance S^2, as an unbiased estimate of the variance of each subsequence, S^2_j. It is easy to see that the sample mean of those variances is an unbiased estimate of the steady-state variance.

To find the variance of this estimator, it is important to notice that the S^2_j are correlated, and thus $\mathrm{Var}[\hat{\sigma}^2] \neq \frac{\mathrm{Var}[S^2_j]}{k}$.

Instead, we express the variance according to (2.2):

$$\mathrm{Var}[\hat{\sigma}^2] = \frac{\mathrm{Var}[S^2_j]}{k} \left(1 + 2 \sum_{i=1}^{k_0-1} (1 - i/k) R_i \right), \tag{4.4}$$

where R_i is the lag i autocorrelation of the sequence S^2_j. Because k_0 was selected such that the autocorrelation coefficients ρ_i of the original sequence $\{X\}$ are negligible at lags $i \geq k_0$, it is reasonable to assume that also the coefficients R_i are not significant at lags $i \geq k_0$.

As the number of observations n grows, all the S^2_j converge consistently to σ^2, and their variance $\mathrm{Var}[S^2_j] \to 0$ as $n \to \infty$. Thus, also $\mathrm{Var}[\hat{\sigma}^2] \to 0$ as $n \to \infty$, which means that the proposed estimate is unbiased and consistent.

4.3.1 The estimator $\hat{\sigma}^2_{VB}$

We consider $k = \kappa k_0$ batches $\{y_j\}$:

$$y_{j,i} = x_{j+ki}, \quad j = 0, \ldots, k-1$$

An appropriate value for the parameter κ will be determined experimentally.

After collecting n observations, we have k batches of length $m = \lfloor n/k \rfloor$ to form an estimate of the steady-state variance:

$$\hat{\sigma}^2_{VB}(n) = \frac{1}{k} \sum_{j=0}^{k-1} s^2_j(m),$$

where

$$s^2_j(m) = \left(\frac{1}{m-1} \sum_{i=1}^{m} y^2_{j,i} \right) - \frac{m}{m-1}\, \bar{y}^2_j$$

is the sample variance of the batch $\{y_j\}$, and

$$\bar{y}_j = \frac{1}{m} \sum_{i=1}^{m} y_{j,i}$$

is its sample average.

To estimate the variance of $\hat{\sigma}^2_{VB}$, we use (4.4), replacing $\mathrm{Var}[S^2_j]$ with the sample variance of s^2_j, and R_i with their estimates $\hat{R}_i$:

$$\mathrm{Var}[\hat{\sigma}^2_{VB}] = \frac{1}{k(k-1)} \sum_{j=1}^{k} \left(s^2_j - \overline{s^2} \right)^2 \left(1 + 2 \sum_{i=1}^{k_0 - 1} (1 - i/k) \hat{R}_i \right),$$

where $\overline{s^2} = \frac{1}{k} \sum_{j=1}^{k} s^2_j$ is the sample average of the batch variances s^2_j.

The autocorrelation coefficients are estimated as $\hat{R}_i = \hat{C}_i / \hat{C}_0$, where $\hat{C}_i$ is the lag i sample autocovariance of the sequence s^2_j:

$$\hat{C}_i = \frac{1}{k-i} \sum_{j=1}^{k-i} \left(s^2_j - \overline{s^2} \right) \left(s^2_{j+i} - \overline{s^2} \right)$$

We use the mean of the secondary data points s^2_j as an estimate of the variance, which justifies treating $\hat{\sigma}^2_{VB}$ as t-distributed with $k-1$ degrees of freedom. The resulting confidence interval is

$$\hat{\sigma}^2_{VB} - \Delta_{VB} < \sigma^2 < \hat{\sigma}^2_{VB} + \Delta_{VB},$$

with

$$\Delta_{VB} = t_{k-1,1-\alpha/2}\sqrt{\mathrm{Var}[\hat{\sigma}^2_{VB}]}.$$

To find a good value for the parameter κ, we note that estimating the coefficients R_i up to a lag of $k_0 - 1$ requires a sufficiently large sample of batch variances s_j^2, which means that the value of κ has to be sufficiently large. On the other hand, a large value of κ leads to more, and thus smaller batches, which makes the estimation of the batch variances less reliable.

Algorithm 5 Estimator $\hat{\sigma}^2_{VB}$

```
m ← 0
k₀ ← DetermineK0
k ← κk₀
Σy_1 ... Σy_k ← 0 ... 0
Σy_1^2 ... Σy_k^2 ← 0 ... 0
Σy_1^3 ... Σy_k^3 ← 0 ... 0
Σy_1^4 ... Σy_k^4 ← 0 ... 0
repeat
   m ← m + 1
   for i = 1 to k do
      y ← GetObservation
      Σy_i ← Σy_i + y
      Σy_i^2 ← Σy_i^2 + y^2
      Σy_i^3 ← Σy_i^3 + y^3
      Σy_i^4 ← Σy_i^4 + y^4
      ȳ_i ← Σy_i/m
      σ̂_i^2 ← Σy_i^2/(m − 1) − m ȳ_i^2/(m − 1)
   end for
   σ̂_VB^2 ← Sum(σ̂_i^2)/k
   v ← Autocovariance(σ̂_1^2 ... σ̂_k^2, 0)
   for j = 1 to k₀ − 1 do
      v ← v + 2 · (1 − j/k)·Autocovariance(σ̂_1^2 ... σ̂_k^2, j)
   end for
   Δ ← t_{k−1,1−α/2} √(v/k)
until RequiredPrecision(σ̂_VB^2, Δ)
return σ̂_VB^2, [σ̂_VB^2 − Δ, σ̂_VB^2 + Δ]
```

Algorithm 6 AUTOCOVARIANCE($x_1, \ldots, x_n, j$)

{Calculate the lag j autocovariance of the sequence $\{x\}$}
$\bar{x} \leftarrow$ SUM$(x_1, \ldots, x_n)/n$
$\Sigma \leftarrow 0$
for $i = 1$ to $n - j$ **do**
  $\Sigma \leftarrow \Sigma + (x_i - \bar{x})(x_{i+j} - \bar{x})$
end for
return $\Sigma/(n - j)$

4.4 A Batch Means Approach

Feldman, Deuermeyer, and Yang (1996), in an unpublished report, describe a procedure of estimating the steady-state variance, based on the method of batch means. Their main contribution is the possibility of compensating for the bias of the variance of a sample $S^2(m)$ by adding the variance of the mean $\overline{X}(m)$.

One can think of this as splitting the variance into two components, a *local variance* describing the short term variations of the process, and a *global variance* representing the long term variations. The local variance is calculated as the mean variance inside equal-sized batches, and the global variance is the variance of the means of the same batches.

To apply the method, the observations are grouped into b equal-sized batches of length m. A remarkable fact about the proposed point estimator is that is seems to be independent of the batch size m. Experiments confirm that it even works with batches as small as two observations.

For each batch j, consisting of $X_{j,1}, \ldots, X_{j,m}$, we calculate the sample mean and sample variance as

$$\overline{X}_j = \frac{1}{n} \sum_{k=1}^{n} X_{j,k} \quad \text{and} \tag{4.5}$$

$$S_j^2 = \frac{1}{n} \sum_{k=1}^{n} (X_{j,k} - \overline{X}_j)^2. \tag{4.6}$$

From these we form the following statistics:

$$\overline{\overline{X}} = \frac{1}{b} \sum_{j=1}^{b} \overline{X}_j, \tag{4.7}$$

$$S_X^2 = \frac{1}{b-1} \sum_{j=1}^{b} (\overline{X}_j - \overline{\overline{X}})^2, \tag{4.8}$$

$$\overline{V} = \frac{1}{b} \sum_{j=1}^{b} S_j^2. \tag{4.9}$$

$\overline{\overline{X}}$ and S_X^2 are the sample mean and sample variance of the batch means, and $\overline{V}$ is the mean of the batch variances.

We now define the point estimator as

$$\hat{\sigma}^2 = \overline{V} + S_X^2. \tag{4.10}$$

We know that $\overline{V}$ is an unbiased estimator of $\mathrm{E}[S_j^2] = \mathrm{E}[\frac{m-1}{m}S^2(m)]$, and that S_X^2 is an unbiased estimator of $\mathrm{Var}[\overline{X}]$. So we can see that

$$\begin{aligned} \mathrm{E}[\hat{\sigma}^2] &= \mathrm{E}[\overline{V}] + \mathrm{E}[S_X^2] \\ &= \mathrm{E}[\frac{m-1}{m}\, S^2(m)] + \mathrm{Var}[\overline{X}] \\ &= \frac{m-1}{m}\,\sigma^2 - \frac{2\sigma^2}{m}\,\xi_m + \frac{1}{m}\,\sigma^2 + \frac{2\sigma^2}{m}\,\xi_m \qquad \text{from (2.2) and (2.5)} \\ &= \sigma^2, \end{aligned}$$

where $\xi_m = \sum_{j=1}^{m-1}(1 - j/m)\rho_j$ are the autocorrelation terms found in (2.2) and (2.5). Thus, $\hat{\sigma}^2$ is an unbiased estimator of the steady-state variance.

To show that $\hat{\sigma}^2$ is a consistent estimate, Feldman et al. (1996) consider its variance

$$\begin{aligned} \mathrm{Var}[\hat{\sigma}^2] &= \mathrm{Var}[\overline{V}] + \mathrm{Var}[S_X^2] + 2\,\mathrm{Cov}[\overline{V}, S_X^2] \\ &\leq \mathrm{Var}[\overline{V}] + \mathrm{Var}[S_X^2] + 2|\,\mathrm{Cov}[\overline{V}, S_X^2]| \\ &\leq \mathrm{Var}[\overline{V}] + \mathrm{Var}[S_X^2] + 2\sqrt{\mathrm{Var}[\overline{V}]\,\mathrm{Var}[S_X^2]}. \end{aligned}$$

We know that $\overline{V}$ and S_X^2 are consistent because they are the sample mean and the sample variance of random samples. This means that

$$\begin{aligned} \mathrm{Var}[\overline{V}] &\to 0 \text{ as } n \to \infty, \text{ and} \\ \mathrm{Var}[S_X^2] &\to 0 \text{ as } n \to \infty. \end{aligned}$$

Therefore, $\mathrm{Var}[\hat{\sigma}^2] \to 0$, as $n \to \infty$, and thus $\hat{\sigma}^2$ is an unbiased, constistent estimate of the steady-state variance.

To obtain a confidence interval, we consider the statistic

$$Y_j = S_j^2 + \frac{b}{b-1}\left(\overline{X}_j - \overline{\overline{X}}\right)^2, \tag{4.11}$$

whose sample mean is equal to the estimator $\hat{\sigma}^2$. In the next section, we present the estimator $\hat{\sigma}^2_{BM}$ as the sample mean of Y_j, and obtain a confidence interval using the sample variance of Y_j:

$$S_Y^2 = \frac{1}{b-1}\sum_{j=1}^{b}\left(Y_j - \hat{\sigma}^2\right)^2. \tag{4.12}$$

While the point estimator performs well independently of the batch size m, the interval estimator does not. Equation (4.12) only accurately describes the variance of the estimator, if the Y_j form an uncorrelated series. Experiments indicate that the Y_j are uncorrelated if the batch means $\overline{X}_j$ are uncorrelated. This is the same condition as needed to estimate a confidence interval of the mean in the "ordinary" batch means method.

4.4.1 The estimator $\hat{\sigma}^2_{BM}$

To be used in a sequential procedure, (4.11) and (4.12) are transformed. The resulting point estimator, constructed as the sample mean of y_j, is

$$\begin{aligned}\hat{\sigma}^2_{BM} &= \frac{1}{b}\sum_{j=1}^{b} y_j \\ &= \frac{1}{b}\sum_{j=1}^{b} s_j^2 + \frac{1}{b-1}\sum_{j=1}^{b} (\bar{x}_j - \bar{\bar{x}})^2 \\ &= \frac{1}{b}\sum_{j=1}^{b} s_j^2 + \frac{1}{b-1}\sum_{j=1}^{b} \bar{x}_j^2 - \frac{b}{b-1}\,\bar{\bar{x}}^2.\end{aligned}$$

To construct a confidence interval, we use the sample variance of the y_j:

$$\begin{aligned}s_Y^2 &= \frac{1}{b-1}\sum_{j=1}^{b} \left(y_j - \hat{\sigma}^2_{BM}\right)^2 \\ &= \frac{1}{b-1}\sum_{j=1}^{b} \left(y_j^2 - 2y_j\hat{\sigma}^2_{BM} + \hat{\sigma}^4_{BM}\right) \\ &= \frac{1}{b-1}\sum_{j=1}^{b} y_j^2 - \frac{b}{b-1}\,\hat{\sigma}^4_{BM}\end{aligned}$$

with

$$\sum_{j=1}^{b} y_j^2 = \sum_{j=1}^{b} \left(s_j^2 + \frac{b}{b-1} \left(\bar{x}_j - \bar{\bar{x}} \right)^2 \right)^2$$
$$= \sum_{j=1}^{b} s_j^4 + \frac{2b}{b-1} \left(\sum_{j=1}^{b} s_j^2 \bar{x}_j^2 - 2\bar{\bar{x}} \sum_{j=1}^{b} s_j^2 \bar{x}_j + \bar{\bar{X}}^2 \sum_{j=1}^{b} s_j^2 \right)$$
$$+ \left(\frac{b}{b-1} \right)^2 \sum_{j=1}^{b} \left(\bar{x}_j - \bar{\bar{x}} \right)^4 ,$$

where

$$\sum_{j=1}^{b} \left(\bar{x}_j - \bar{\bar{x}} \right)^4 = \sum_{j=1}^{b} \bar{x}_j^4 - 4\bar{\bar{x}} \sum_{j=1}^{b} \bar{x}_j^3 + 6\bar{\bar{x}}^2 \sum_{j=1}^{b} \bar{x}_j^2 - 3b\bar{\bar{x}}^4.$$

Now we see that an implementation of this estimator only needs to store the sums over s_j^2, s_j^4, $\bar{x}_j$, $\bar{x}_j^2$, $\bar{x}_j^3$, $\bar{x}_j^4$, $s_j^2\bar{x}_j$, and $s_j^2\bar{x}_j^2$ to calculate the estimates.

The resulting confidence interval is

$$\hat{\sigma}_{BM}^2 - \Delta_{BM} < \sigma^2 < \hat{\sigma}_{BM}^2 + \Delta_{BM},$$

with

$$\Delta_{BM} = t_{b-1,1-\alpha/2} \sqrt{s_Y^2 / b}$$

where $t_{b-1,1-\alpha/2}$ is the $1 - \frac{\alpha}{2}$ quantile of the t-distribution with $b-1$ degrees of freedom. The use of the t-distribution is justified by the fact that we use the sample mean of the secondary data points y_j as an estimate of σ^2.

Algorithm 7 Estimator $\hat{\sigma}^2_{BM}$

$m \leftarrow$ DETERMINEBATCHSIZE {detailed procedure in Pawlikowski (1990)}
$\Sigma\bar{x} \leftarrow 0, \Sigma\bar{x}^2 \leftarrow 0, \Sigma\bar{x}^3 \leftarrow 0, \Sigma\bar{x}^4 \leftarrow 0$
$\Sigma s^2 \leftarrow 0, \Sigma s^4 \leftarrow 0, \Sigma s^2\bar{x} \leftarrow 0, \Sigma s^2\bar{x}^2 \leftarrow 0$
$b \leftarrow 0$ {Batch Count}
repeat
  $\Sigma x \leftarrow 0$, $\Sigma x^2 \leftarrow 0$
  for $i = 1$ to m **do** {collect batch of observations}
    $x \leftarrow$ GETOBSERVATION
    $\Sigma x \leftarrow \Sigma x + x$
    $\Sigma x^2 \leftarrow \Sigma x^2 + x^2$
  end for
  $b \leftarrow b + 1$
  $\bar{x} \leftarrow \Sigma x / m$
  $s^2 \leftarrow \Sigma x^2 - (\Sigma x)^2$
  $\Sigma\bar{x} \leftarrow \Sigma\bar{x} + \bar{x}$
  $\Sigma\bar{x}^2 \leftarrow \Sigma\bar{x}^2 + \bar{x}^2$
  $\Sigma\bar{x}^3 \leftarrow \Sigma\bar{x}^3 + \bar{x}^3$
  $\Sigma\bar{x}^4 \leftarrow \Sigma\bar{x}^4 + \bar{x}^4$
  $\Sigma s^2 \leftarrow \Sigma s^2 + s^2$
  $\Sigma s^4 \leftarrow \Sigma s^4 + (s^2)^2$
  $\Sigma s^2\bar{x} \leftarrow \Sigma s^2\bar{x} + s^2 \cdot \bar{x}$
  $\Sigma s^2\bar{x}^2 \leftarrow \Sigma s^2\bar{x}^2 + s^2 \cdot \bar{x}^2$
  $\nu \leftarrow \Sigma\bar{x}/b$
  $\hat{\sigma}^2_{BM} \leftarrow \Sigma s^2/b + \Sigma\bar{x}^2/(b-1) - b\nu^2/(b-1)$
  $\Sigma y^2 \leftarrow \Sigma s^4 + \frac{2b}{b-1}(\Sigma s^2\bar{x}^2 - 2\nu\Sigma s^2\bar{x} + \nu^2\Sigma s^2)$
    $+(\frac{c}{c-1})^2(\Sigma\bar{x}^4 - 4\nu\Sigma\bar{x}^3 + 6\nu^2\Sigma\bar{x}^2 - 3b\nu^4)$
  $S^2_Y \leftarrow \frac{1}{c-1}\Sigma y^2 - \frac{c}{c-1}(\hat{\sigma}^2)^2$
  $\Delta \leftarrow t_{b-1,1-\alpha/2}\sqrt{S^2_Y/b}$
until REQUIREDPRECISION($\hat{\sigma}^2_{BM}$, Δ)
return $\hat{\sigma}^2_{BM}$, $[\hat{\sigma}^2_{BM} - \Delta, \hat{\sigma}^2_{BM} + \Delta]$

4.5 Summary

In this chapter we proposed several estimators of the steady-state variance. Some of their characteristics are summarised in Table 4.1.

Because they discard many observations, and can therefore be expected to require long simulation runs, the estimators $\hat{\sigma}^2_{IR}$ and $\hat{\sigma}^2_{SP}$ can not be recommended for practical use. The estimators $\hat{\sigma}^2_{M1}$ and $\hat{\sigma}^2_{M2}$ are easiest to implement, and because of that, they deserve special interest. The remaining two estimators both rely on the estimation of autocorrelations to determine their parameters, $\hat{\sigma}^2_{VB}$ to determine the number of batches and $\hat{\sigma}^2_{BM}$ to find an appropriate batch size. Experimental assessment of the estimators will show if this influences their performance. In addition, in the estimator $\hat{\sigma}^2_{VB}$, the estimation of the autocorrelations of the batch variances directly influences the generation of the confidence interval, which could prove problematic.

Estimator	**Remarks**
$\hat{\sigma}^2_{M1}$	• easy implementation • uses existing mean value estimator • does not discard observations
$\hat{\sigma}^2_{M2}$	• easy implementation • uses existing mean value estimator • does not discard observations
$\hat{\sigma}^2_{IR}$	• inefficient, discards many observations • relies on the quality of initial transient detection
$\hat{\sigma}^2_{SP}$	• inefficient, discards many observations • relies on initial detection of the autocorrelation
$\hat{\sigma}^2_{VB}$	• based on $\hat{\sigma}^2_{SP}$ • does not discard observations • relies on initial detection of autocorrelation
$\hat{\sigma}^2_{BM}$	• sound mathematical background • easy implementation • does not discard observations • has to rely on finding an appropriate batch size

Table 4.1: Summary of Variance Estimators

Chapter 5

Comparative Study of Estimators

In this chapter, performance of the proposed methods is assessed using several different queueing models. The most important performance measure of the estimators is their coverage of confidence intervals. In addition, we assess the sample sizes required by the various estimators to finish the sequential simulation.

Several of the estimators have parameters for which no universally optimal values can be found. Depending on the process, appropriate values are determined automatically at the beginning of a simulation run. We examine the values of the parameters that were determined by the estimators, and if possible compare them to theoretical results.

5.1 Reference Models

We use implementations of the single server queueing models *M/M/1*, *M/E_2/1*, and *M/H_2/1*. Although only three models are considered, they cover a broad range of possible processes. The service times of the models have a coefficient of variation of $C_{st} = 1$ (M/M/1 queue), $C_{st} < 1$ (M/E_2/1 queue), or $C_{st} > 1$ (M/H_2/1 queue). The performance measure studied is the *waiting time* (time a customer spends waiting in queue) in steady state.

Especially the M/M/1 queue has been studied intensively, and many results are available from queueing theory. For the other two models, theoretical results that go beyond the expected means and variances of the waiting time are more difficult to obtain, and where other theoretical results (e.g. the autocorrelation structure) are needed, we only consider the M/M/1 queue.

To obtain theoretical values of the mean and variance of the waiting time, the Pollaczek-Khintchine transform formula $W_q(s)$ for the waiting time distribution is used (see e.g. Gross and Harris, 1985, p. 274).

As with any Laplace-Stieltjes transform, differentiating $W_q(s)$ n times, and evaluating it at $s = 0$ yields the nth moment of the waiting time distribution. Thus, the mean is

$$\nu = \lim_{s \to 0} -\frac{dW_q(s)}{ds}, \tag{5.1}$$

and the variance

$$\sigma^2 = \lim_{s \to 0} \frac{d^2 W_q(s)}{ds^2} - \nu^2. \tag{5.2}$$

The symbolic algebra package Maple (Waterloo Maple Inc., 2008) was used to evaluate these formulas.

5.1.1 The M/M/1 Queue

The arrivals of this single-server queue are modeled as a Poisson process with rate λ, the service time has an exponential distribution with rate μ, the utilisation of the queue is $\rho = \lambda/\mu$.

The Pollaczek-Khintchine transform formula for the waiting time distribution is

$$W_q(s) = \frac{\left(1 - \frac{\lambda}{\mu}\right) s}{s - \lambda \left(1 - \frac{\mu}{\mu + s}\right)},$$

which yields the expected values of its mean and variance:

$$\nu = \frac{\lambda}{\mu(\mu - \lambda)} \tag{5.3}$$

$$\sigma^2 = \frac{\lambda(2\mu - \lambda)}{\mu^2(\mu - \lambda)^2} \tag{5.4}$$

Equation (5.3) states a well-known result (see Gross and Harris, 1985, p. 77). Daley (1968) gives formulae for the autocorrelation coefficients of the waiting times in the M/M/1 queue, and Law (1977) extends those results to find the autocorrelation of batch means of the waiting times. We use these results in the investigation of the performance of the estimators $\hat{\sigma}^2_{SP}$, $\hat{\sigma}^2_{VB}$, and $\hat{\sigma}^2_{BM}$.

Table 5.1 specifies the working conditions of the M/M/1 queueing system in the simulation experiments reported here.

System Load ρ	Arrival Rate λ	Service Rate μ	Mean Waiting Time ν	Variance of Waiting Time σ^2
0.1	0.1	1	0.111	0.235
0.2	0.2	1	0.25	0.563
0.3	0.3	1	0.429	1.041
0.4	0.4	1	0.667	1.778
0.5	0.5	1	1	3
0.6	0.6	1	1.5	5.25
0.7	0.7	1	2.333	10.111
0.8	0.8	1	4	24
0.9	0.9	1	9	99

Table 5.1: Working Conditions of the Simulated M/M/1 Queueing System

5.1.2 The M/E_2/1 Queue

The service time has an *Erlang* distribution with two equal stages of service. Arrivals occur according to a Poisson proces with rate λ, and the service time of each stage of service is exponentially distributed with rate μ. The resulting system load is $\rho = 2\lambda/\mu$.

For the M/E_2/1 queue, the Pollaczek-Khintchine transform formula for the waiting time distribution is

$$W_q(s) = \frac{\left(1 - \frac{2\lambda}{\mu}\right) s}{s - \lambda \left(1 - \frac{\mu^2}{(\mu+s)^2}\right)},$$

from which we find the expected values of its mean and variance:

$$\nu = \frac{3\lambda}{\mu(\mu - 2\lambda)} \tag{5.5}$$

$$\sigma^2 = \frac{\lambda(8\mu - 7\lambda)}{\mu^2(\mu - 2\lambda)^2} \tag{5.6}$$

The values chosen for the experiments are summarised in Table 5.2.

System Load ρ	Arrival Rate λ	Service Rate μ	Mean Waiting Time ν	Variance of Waiting Time σ^2
0.1	0.1	2	0.083	0.118
0.2	0.2	2	0.188	0.285
0.3	0.3	2	0.321	1.032
0.4	0.4	2	0.5	0.917
0.5	0.5	2	0.75	1.563
0.6	0.6	2	1.125	2.766
0.7	0.7	2	1.75	5.396
0.8	0.8	2	3	13
0.9	0.9	2	7.75	54.563

Table 5.2: Working Conditions of the Simulated $M/E_2/1$ Queueing System

5.1.3 The $M/H_2/1$ Queue

Again, arrivals are modeled as a Poisson process, but now the service time has a hyperexponential distribution. A customer is served with exponential service time, and the rate is μ_1 with probability p, and μ_2 with probability $1-p$. Values of the parameters are chosen so that the service time has a mean of 1 and a variance of 5:

$$\begin{aligned} p &= \frac{1}{2} + \sqrt{1/3} \approx 0.908 \\ \mu_1 &= 1 + \sqrt{2/3} \approx 1.816 \\ \mu_2 &= 1 - \sqrt{2/3} \approx 0.184 \end{aligned}$$

For the $M/H_2/1$ queue, the Pollaczek-Khintchine transform formula for the waiting time distribution is

$$W_q(s) = \frac{\left(1 - \lambda\left(\frac{p}{\mu_1} + \frac{1-p}{\mu_2}\right)\right) s}{s - \lambda\left(1 - \frac{p\mu_1}{\mu_1+s} - \frac{(1-p)\mu_2}{\mu_2+s}\right)}.$$

System Load ρ	Arrival Rate λ	Service Parameters p	μ_1	μ_2	Mean Waiting Time ν	Variance of Waiting Time σ^2
0.1	0.1	0.908	1.816	0.184	0.333	3.444
0.2	0.2	0.908	1.816	0.184	0.75	8.063
0.3	0.3	0.908	1.816	0.184	1.286	14.51
0.4	0.4	0.908	1.816	0.184	2	24
0.5	0.5	0.908	1.816	0.184	3	39
0.6	0.6	0.908	1.816	0.184	4.5	65.25
0.7	0.7	0.908	1.816	0.184	7	119
0.8	0.8	0.908	1.816	0.184	12	264
0.9	0.9	0.908	1.816	0.184	27	999

Table 5.3: Working Conditions of the Simulated M/H_2/1 Queueing System

This leads to expected values of its mean and variance:

$$\nu = \frac{\lambda(\mu_1^2 - \mu_1^2 p + \mu_2^2 p)}{\mu_1\mu_2(\mu_1\mu_2 - \lambda\mu_2 p - \lambda\mu_1 + \lambda\mu_1 p)}$$

$$\sigma^2 = \frac{\lambda}{\mu_1^2\mu_2^2(\mu_1\mu_2 - \lambda\mu_2 p - \lambda\mu_1 + \lambda\mu_1 p)}\Big(- 2\mu_1^4 - 2\mu_1\mu_2^3 p^2\lambda - 2\mu_1^2\mu_2^2 p\lambda + 2\mu_1\mu_2^3 p\lambda + 2\mu_1^3\mu_2\lambda p - 2\mu_1^3\mu_2 p^2\lambda - 2\mu_2\mu_2^4 p + 2\mu_1^4\mu_2 p + 2\mu_1^2\mu_2^2 p^2\lambda + \lambda\mu_1^4 - 2\lambda\mu_1^4 p + \lambda\mu_1^4 p^2 + \lambda p^2\mu_2^4\Big)$$

The values chosen for the experiments are summarised in Table 5.3.

5.2 Detection of the Initial Transient

5.2.1 Rule of "25 Crossings"

We have seen in Section 3.2 that, in case of the M/M/1 queueing system, we can expect the initial transient period of the variance to be about 50% longer than that of the mean value.

Figure 5.1 shows the results of applying the rule of "25 Crossings" to the original process and to the squared process. As reference, the plot includes the relaxation

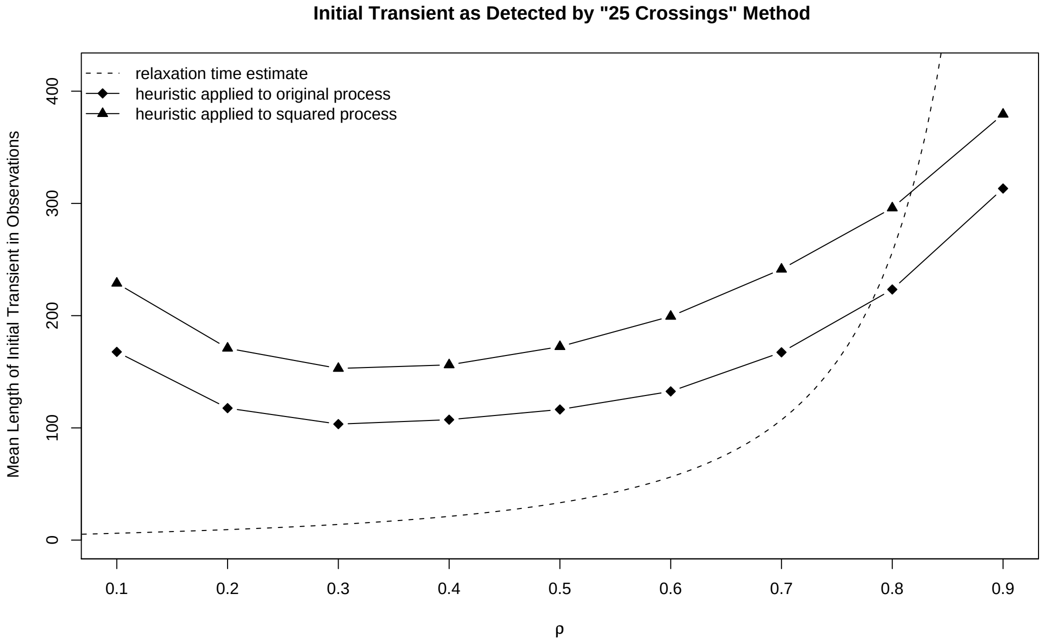


Figure 5.1: Length of the Initial Transient Period of the Waiting Time in an M/M/1 Queue as Detected by the "25 Crossings" Heuristic

time estimate of the length of the initial transient period of the mean. We see that the squared process takes about 1.3 times longer to reach the 25th crossing than the original process.

This indicates that the modified "25 Crossings" rule performs almost as well for the process variance as the original rule does for the process mean. For system loads above $\rho = 0.8$, however, the results suggest that the rules underestimate the length of the initial transient period.

5.3 Estimation of the Steady-State Variance

To assess the performance of the various estimators presented in Chapter 4, the sequential form of coverage analysis, as described by Pawlikowski, Ewing, and McNickle (1998), is used. The procedure implements three rules:

1. Analysis of coverage is done sequentially.
2. A minumum number of "bad" confidence intervals must be recorded.
3. Too short simulation runs are discarded.

Unless stated otherwise, the studies presented in this section use a minumum number of bad confidence intervals of 200, and discard simulation runs that are shorter than the mean run length minus one standard deviation. This leads to about 10–15% of the simulation runs being discarded (please refer to the appendix for exact numbers). The coverage analysis is stopped upon reaching a relative precision of $\epsilon_r = 0.05$ at the 0.95 confidence level.

5.3.1 Simulation Run Length

The different estimators converge to the desired precision with different rates. Figure 5.2 shows the mean number of observations needed for the estimator to reach a relative precision of $\epsilon_r = 0.05$ at the 0.95 confidence level. The model used is the M/M/1 queue.

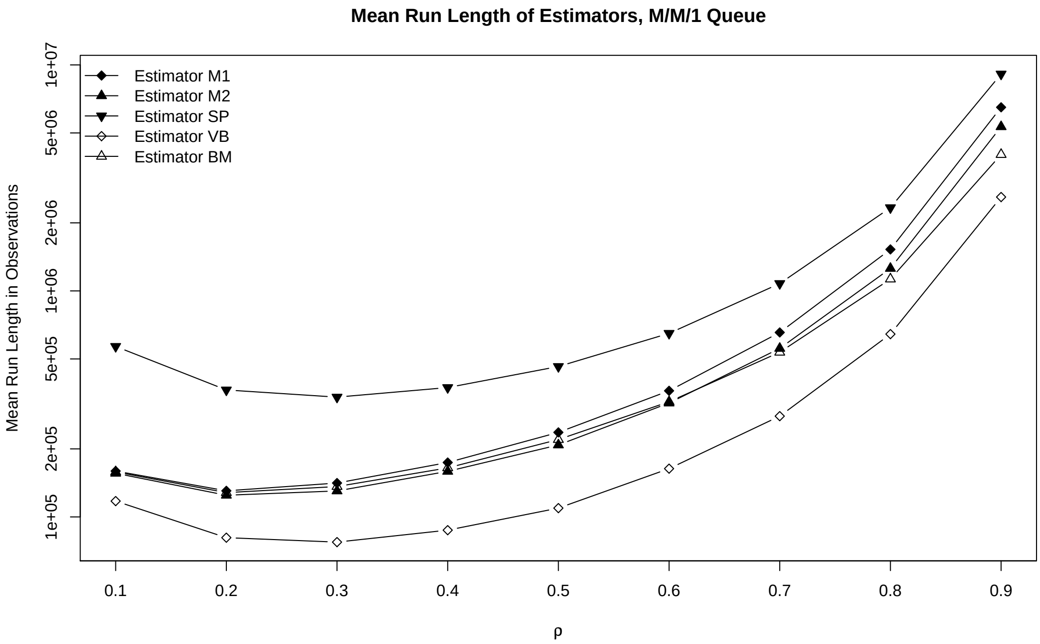


Figure 5.2: Comparison of Simulation Run Length (Logarithmic Scale)

The estimators $\hat{\sigma}^2_{M1}$, $\hat{\sigma}^2_{M2}$ and $\hat{\sigma}^2_{BM}$ show a similar behaviour, with $\hat{\sigma}^2_{BM}$ having slightly shorter run lengths than the other two.

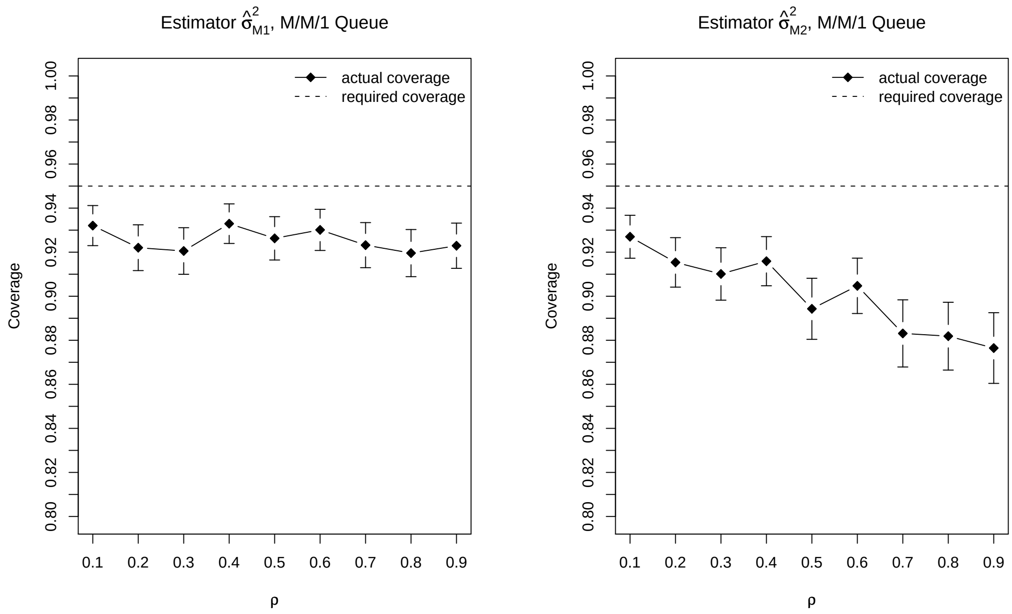
As was to be expected, the estimator $\hat{\sigma}^2_{SP}$ has the longest run length. The reason for this is the large number of discarded observations. The Estimator

$\hat{\sigma}^2_{VB}$, that was designed to improve on this by using all collected observations, has significantly shorter run lengths.

5.3.2 Experimental Coverage Analysis

The results of the coverage analysis is displayed over a range of system loads. Unless noted otherwise, confidence intervals are calculated at the 0.95 confidence level. When more than one curve in displayed in the same diagram, the plot does not contain the confidence intervals, but the points have still been calculated with the same precision. Details about the statistical properties of the data used to draw the figures can be found in the appendix.

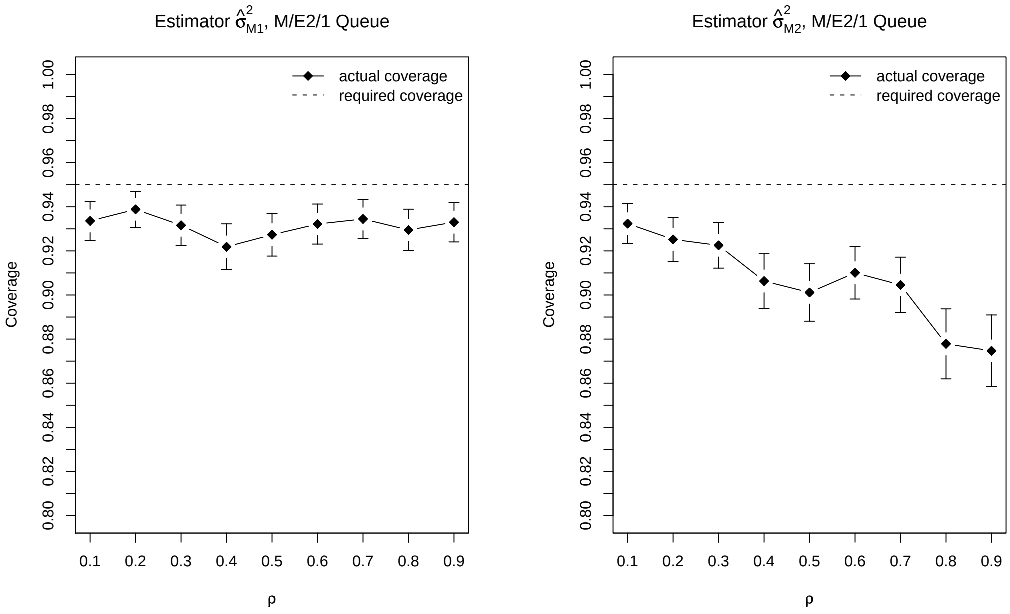
Estimators $\hat{\sigma}^2_{M1}$ and $\hat{\sigma}^2_{M2}$



(a) Estimator $\hat{\sigma}^2_{M1}$ (b) Estimator $\hat{\sigma}^2_{M2}$

Figure 5.3: Coverage of the Estimators $\hat{\sigma}^2_{M1}$ and $\hat{\sigma}^2_{M2}$, M/M/1 Queue

The results of the coverage analysis when treating the variance as a mean value, as described in Section 4.1, can be found in Figures 5.3 (M/M/1 queue), 5.4 ($M/E_2/1$ queue), and 5.5 ($M/H_2/1$ queue).

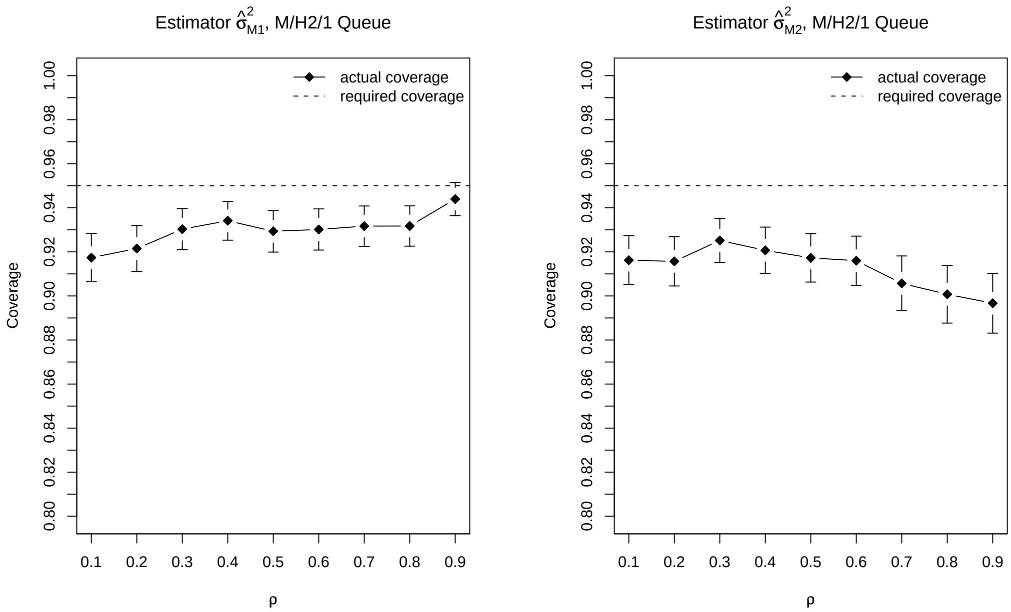


(a) Estimator $\hat{\sigma}^2_{M1}$ (b) Estimator $\hat{\sigma}^2_{M2}$

Figure 5.4: Coverage of the Estimators $\hat{\sigma}^2_{M1}$ and $\hat{\sigma}^2_{M2}$, M/E_2/1 Queue

We can see that the estimator $\hat{\sigma}^2_{M1}$ is superior to $\hat{\sigma}^2_{M2}$. While its coverage remains at an almost constant level even for a broad range of system loads, that of $\hat{\sigma}^2_{M2}$ drops significantly for highly loaded systems.

Especially considering that these estimators represent a practical approach to variance estimation, and that they are easy to implement, the performance of $\hat{\sigma}^2_{M1}$ is remarkable, even if the coverage is slightly below the required level.



(a) Estimator $\hat{\sigma}^2_{M1}$ (b) Estimator $\hat{\sigma}^2_{M2}$

Figure 5.5: Coverage of the Estimators $\hat{\sigma}^2_{M1}$ and $\hat{\sigma}^2_{M2}$, M/H_2/1 Queue

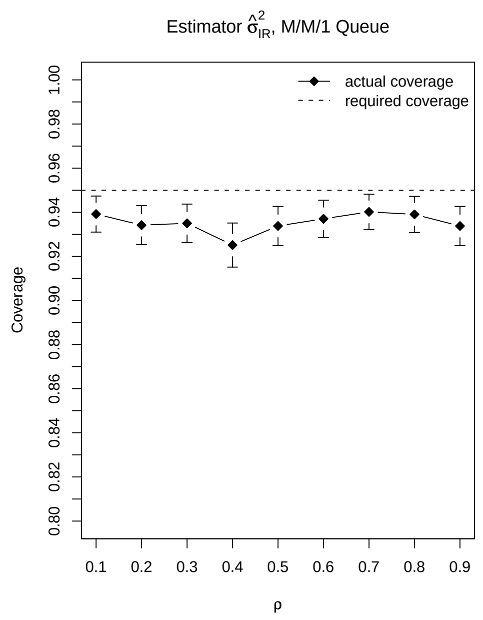


Figure 5.6: Coverage of Estimator $\hat{\sigma}^2_{IR}$, M/M/1 Queue, Fixed Length of Initial Transient Period

Estimator $\hat{\sigma}^2_{IR}$

We already saw in Section 4.2.1 that the estimator $\hat{\sigma}^2_{IR}$ is not useful for pratical purposes because it discards many observations. In addition, it relies on a functioning method of detection of the initial transient phase. Chapter 3 showed that such a method is currently not available; therefore, the estimator was not assessed in its proposed form. Instead, a fixed length was assumed for the initial transient period, obtained from the expected variance of the nth customer (see Figure 3.1 on page 15).

From the results in Figure 5.6 we conclude that the estimator $\hat{\sigma}^2_{IR}$ would produce valid results in terms of coverage, if a reliable method of detection of the initial transient phase can be found.

Estimator $\hat{\sigma}^2_{SP}$

In this estimator, the parameter k_0 describes the lag at which the autocorrelation of the observations becomes negligible. A value for k_0 is determined automatically, and the procedure used directly influences the performance of the estimator. To test whether the estimator finds reasonable values for k_0, we need theoretical values of this parameter.

For the M/M/1 queue, it is possible to calculate the theoretical values of the autocorrelation coefficients of the waiting time (see Daley, 1968, esp. equation 32, and Table 1, p. 697). We then say that the theoretical value of k_0 is the lag at which the value of the autocorrelation first drops below 0.01: $k_{0,\text{theory}} = \min\{i \mid \rho_i < 0.01\}$.

A comparison of these theoretical values of k_0 to the ones determined by the estimator $\hat{\sigma}^2_{SP}$ is shown in Figure 5.7. We see that the mean values found for the parameter k_0 closely match the ones we expect from theoretical calculations.

The results of the coverage analysis of the estimator $\hat{\sigma}^2_{SP}$ can be found in Figures 5.8(a), 5.9(a) and 5.10(a). The estimator performs well in terms of coverage for all three models tested. Nevertheless, because of the many observations it

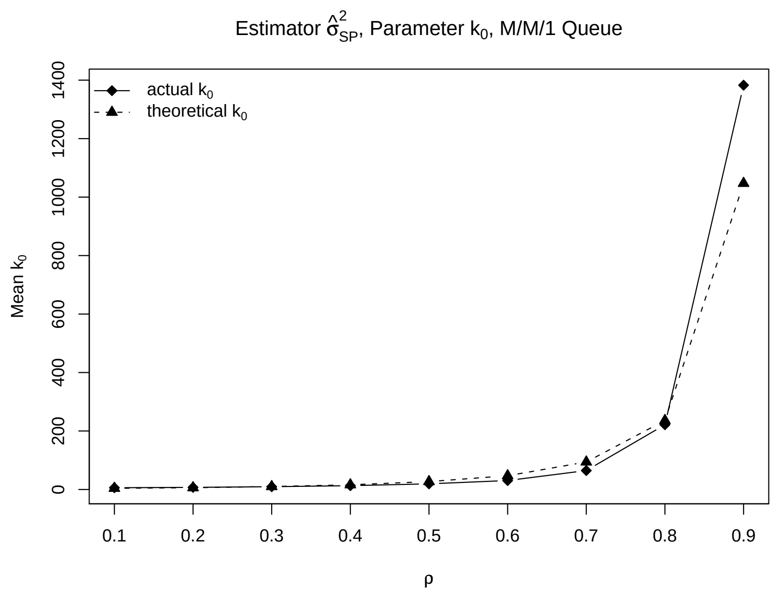
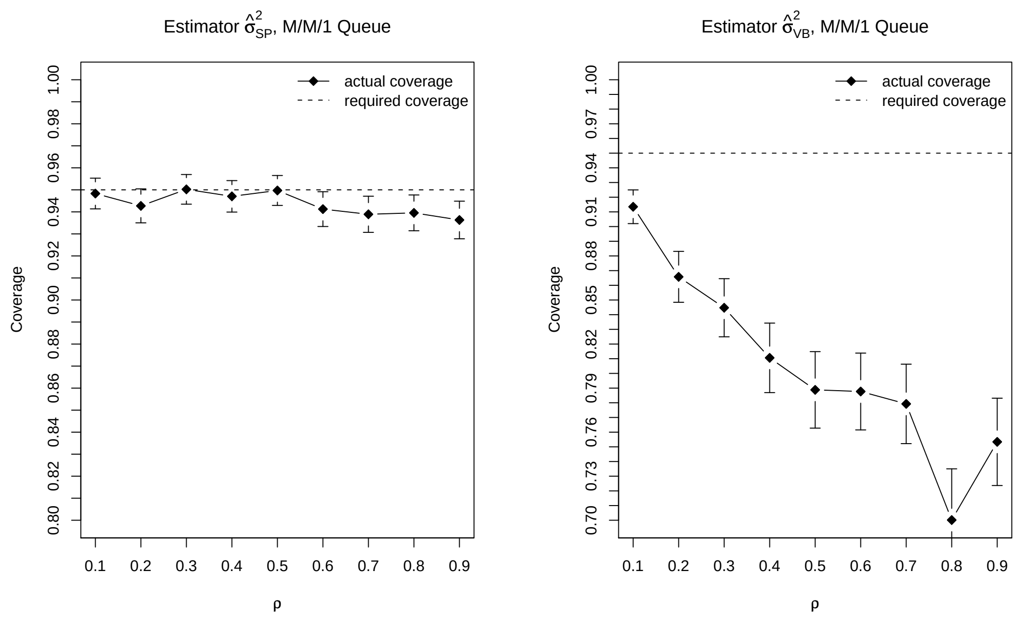


Figure 5.7: Estimator $\hat{\sigma}^2_{SP}$, Parameter k_0



(a) Estimator $\hat{\sigma}^2_{SP}$

(b) Estimator $\hat{\sigma}^2_{VB}$

Figure 5.8: Coverage of Estimators $\hat{\sigma}^2_{SP}$ and $\hat{\sigma}^2_{VB}$, M/M/1 Queue

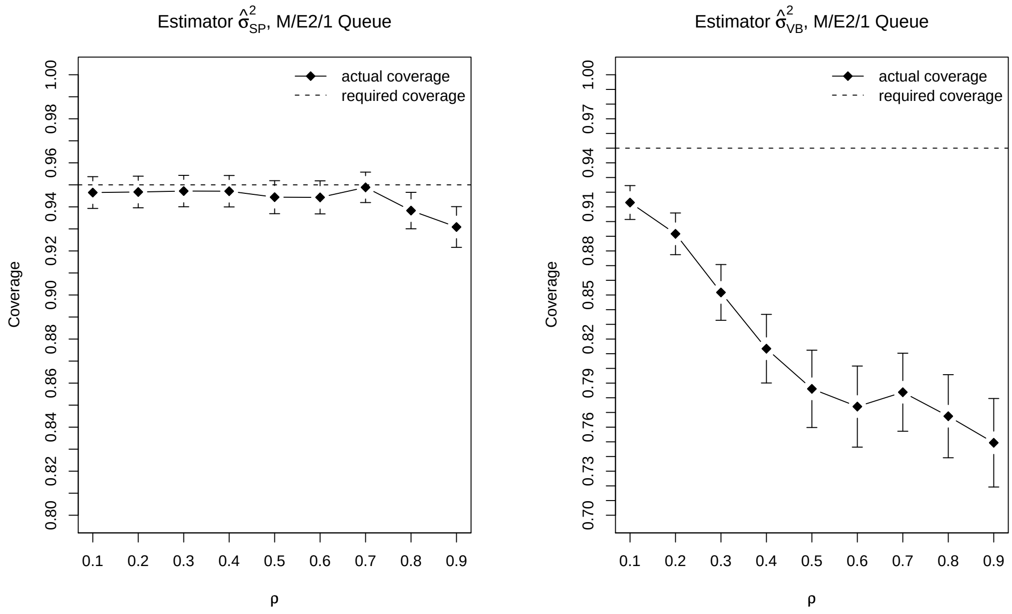
discards, it can not be recommended for practical use in its current form.

Estimator $\hat{\sigma}^2_{VB}$

In the estimator $\hat{\sigma}^2_{VB}$, the parameter κ plays an important role. It determines the number of batches used, and thus influences the sample size for the calculation of the autocorrelation of the batch variances S^2_j. It is expected that large values of κ lead to better estimates of the autocorrelation and thus to more accurate confidence intervals. However, as κ increases, the number of batches grows, and the batch size is reduced. This leads to inaccurate estimates of the batch variances S^2_j.

The coverage for different values of κ is depicted in Figure 5.11. We see that coverage is best for values $\kappa \approx 50$. This is the value used in the coverage experiments shown in Figures 5.8(b), 5.9(b) and 5.10(b).

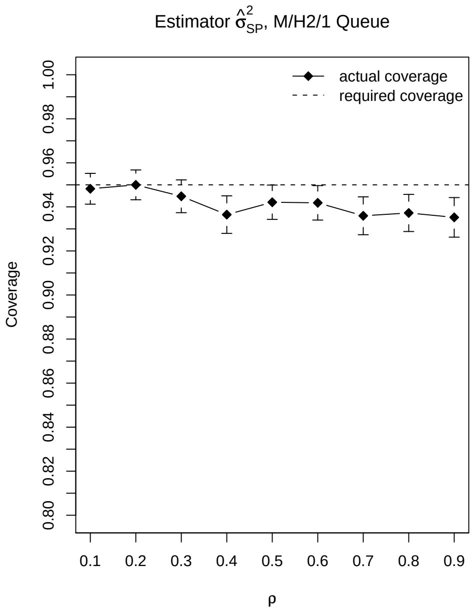
The estimator $\hat{\sigma}^2_{VB}$ was designed to improve the estimator $\hat{\sigma}^2_{SP}$ by not discarding observations. While this improvement is visible in simulation run length (see Section 5.3.1), the estimator does not perform well in terms of coverage.



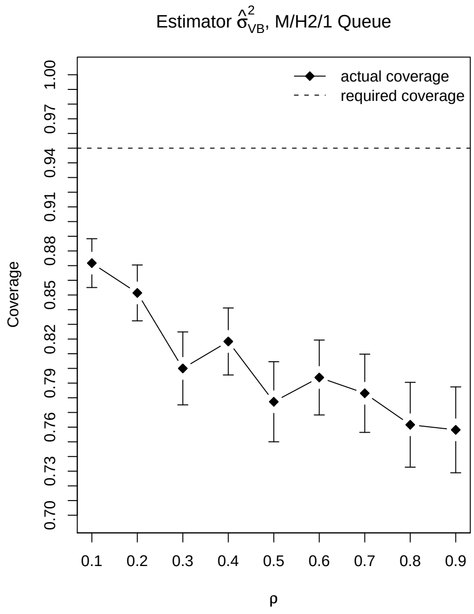
(a) Estimator $\hat{\sigma}^2_{SP}$ (b) Estimator $\hat{\sigma}^2_{VB}$

Figure 5.9: Coverage of Estimators $\hat{\sigma}^2_{SP}$ and $\hat{\sigma}^2_{VB}$, M/E_2/1 Queue

The analysis of the estimator $\hat{\sigma}^2_{SP}$ has shown that the general idea of considering "far apart" observations is valid. It is possible that further investigation will lead to a better method of making the estimator more efficient. By exploiting the correlation between the batch variances S_j^2, it might be possible to obtain a reduction in variance, similar to that observed in the methods of overlapping batch means (see Meketon and Schmeiser, 1984).



(a) Estimator $\hat{\sigma}^2_{SP}$



(b) Estimator $\hat{\sigma}^2_{VB}$

Figure 5.10: Coverage of Estimators $\hat{\sigma}^2_{SP}$ and $\hat{\sigma}^2_{VB}$, M/H_2/1 Queue

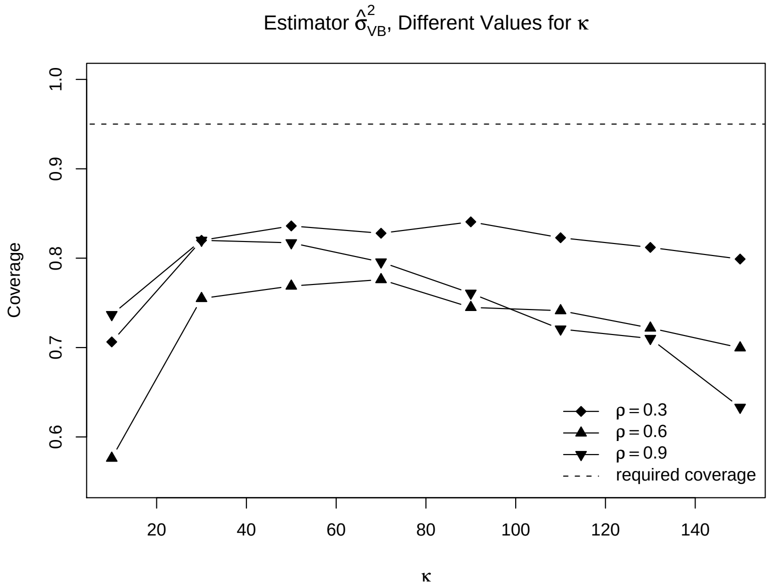


Figure 5.11: Coverage of Estimator $\hat{\sigma}^2_{VB}$, Different Values for Parameter κ

Estimator $\hat{\sigma}^2_{BM}$

The only parameter of the estimator $\hat{\sigma}^2_{BM}$ is the batch size. Batches need to be large enough, so that their means are approximately uncorrelated.

Figure 5.12 shows that as the system load increases, larger batches are needed to produce good coverage. It can be seen that almost optimal coverage can be obtained even for high system loads.

When the batch size is detected automatically by testing the batch means for autocorrelation, the determined size is often smaller than is needed to produce optimal coverage. Figures 5.13 and 5.14(a) show the results of the coverage analysis using automated detection of the batch size. A drop in coverage can be seen for high system loads.

Theoretical results on the autocorrelation of batch means are provided by Law (1977), who extends results by Daley (1968), and gives formulae to calculate the correlation coefficients of batch means as a function of batch size. We assume that batch means are uncorrelated when their lag 1 autocorrelation drops below a certain level ϵ, and analyse the minimum batch size needed to produce

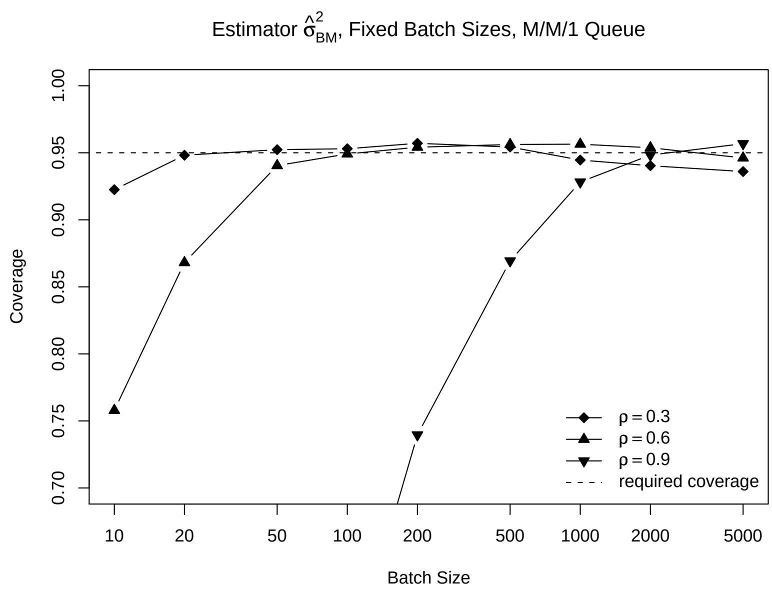
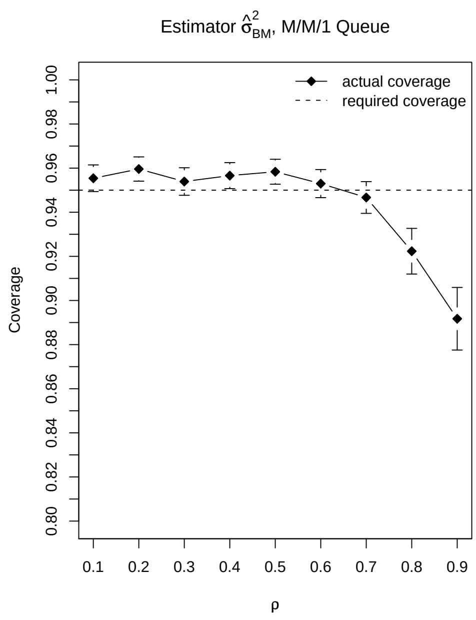
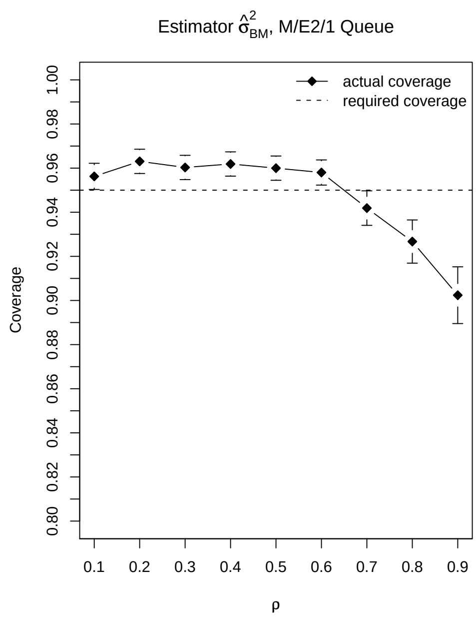


Figure 5.12: Coverage of the Estimator $\hat{\sigma}^2_{BM}$ for Different Fixed Batch Sizes (Logarithmic Scale), M/M/1 Queue.



(a) M/M/1 Queue



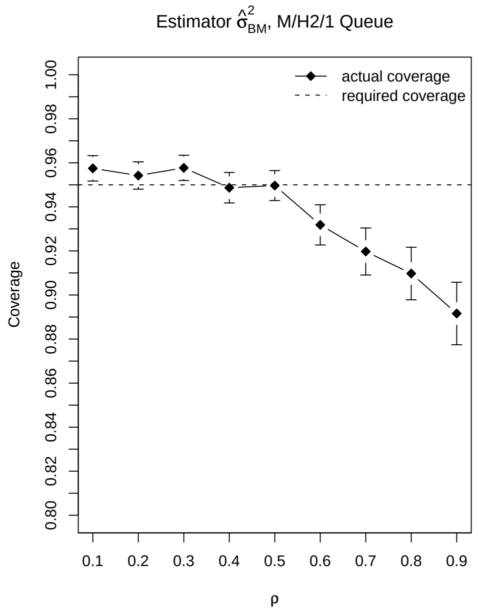
(b) $M/E_2/1$ Queue

Figure 5.13: Coverage of Estimator $\hat{\sigma}^2_{BM}$, Automated Detection of Batch Size

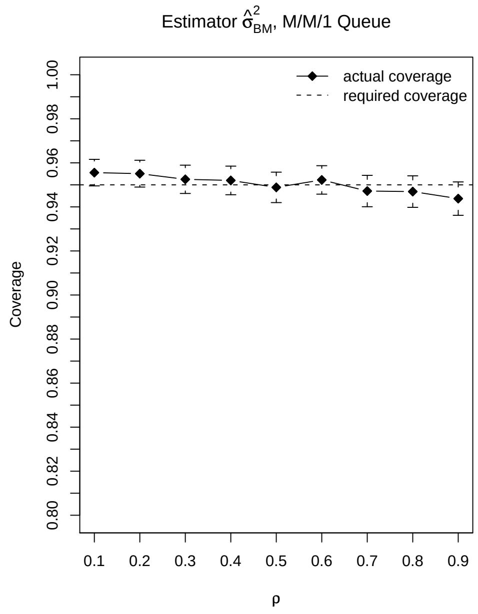
uncorrelated batch means. A comparison of the mean batch sizes determined by the estimator to the theoretical batch sizes for a value of $\epsilon = 0.05$ can be found in Figure 5.15. We see that for high system loads, the batch sizes determined by the estimator are too small to produce uncorrelated batch means. This explains the drop in coverage for these high system loads.

For the estimation of mean values using the method of batch means, McNickle, Ewing, and Pawlikowski (2004) observe a very similar drop in coverage and find that it is almost entirely due to the correlation in the batch means. Figure 5.16 is taken from their paper and shows the coverage of the ordinary batch means estimator with automated detection of batch size. In the case of variance estimation using the estimator $\hat{\sigma}^2_{BM}$, there might be additional sources of error that lead to the reduction of coverage, but the parallels to the results of McNickle et al. (2004) suggest that the correlation of the batch means is the dominating reason for the drop in coverage.

To confirm this, the batch size used by estimator $\hat{\sigma}^2_{BM}$ is changed. After automated detection of the batch size, the result is multiplied by three before it is used. The resulting coverage for the M/M/1 queueing model is shown in Figure



(a) M/H_2/1 Queue, Automated Detection of Batch Size



(b) M/M/1 Queue, Triple Batch Size

Figure 5.14: Coverage of Estimator $\hat{\sigma}^2_{BM}$

5.14(b). As expected, coverage has improved for high system loads.

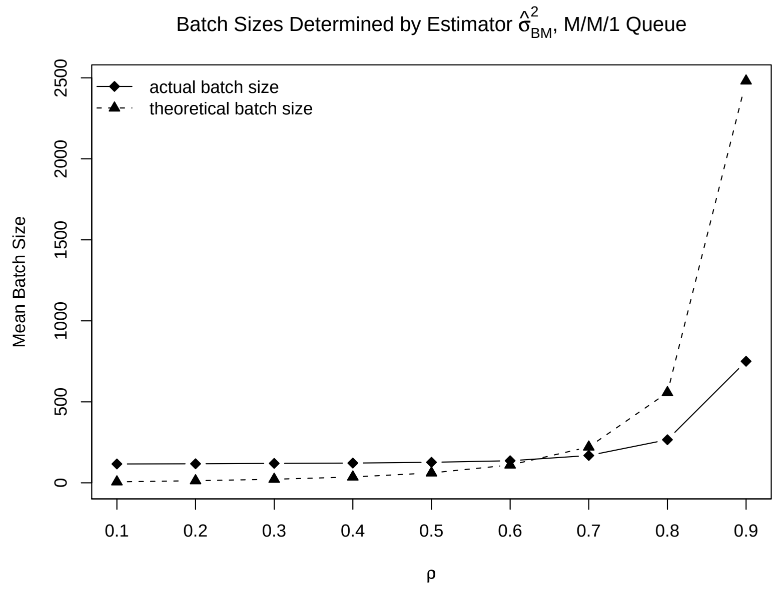


Figure 5.15: Batch Sizes Automatically Determined by Estimator $\hat{\sigma}^2_{BM}$, M/M/1 Queue

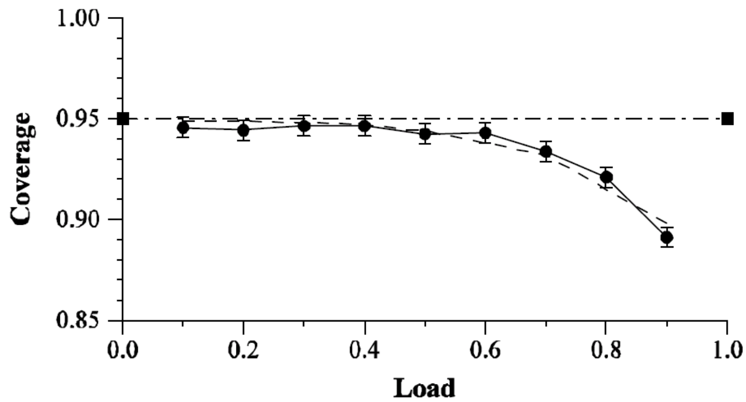
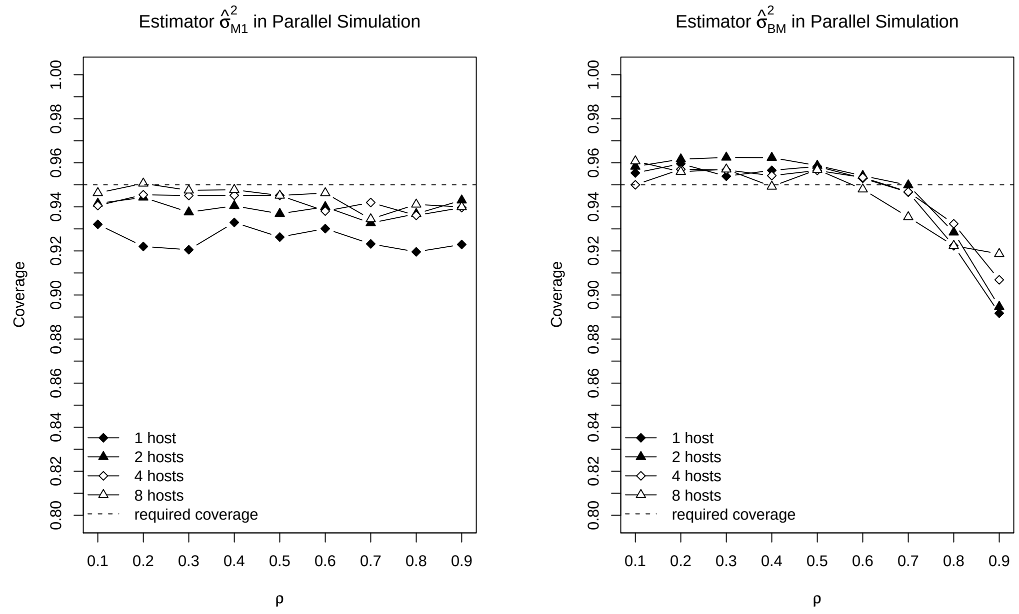


Figure 5.16: Coverage of Batch Means from McNickle et al. (2004). The dashed line represents the expected coverage calculated from Law (1977).



(a) Estimator $\hat{\sigma}^2_{M1}$ (b) Estimator $\hat{\sigma}^2_{BM}$

Figure 5.17: Coverage of the Estimators $\hat{\sigma}^2_{M1}$ and $\hat{\sigma}^2_{BM}$ Using Multiple Replications in Parallel, M/M/1 Queue

5.3.3 Multiple Replications in Parallel

The previous section showed that the two estimators $\hat{\sigma}^2_{M1}$ and $\hat{\sigma}^2_{BM}$ produce good results in terms of the coverage of confidence intervals. In addition, both estimators were tested in a parallel simulation scenario with Multiple Replications in Parallel. The model was run on 2, 4, and 8 hosts, and estimates from the replications were combined into a global estimate, as discussed in Section 2.3.1.

The original version of Multiple Replications in Parallel, as published by Pawlikowski et al. (1994), was intended for the estimation of mean values. But since $\hat{\sigma}^2_{M1}$ estimates the variance as a mean value, and $\hat{\sigma}^2_{BM}$ estimates it as the sample mean of secondary data points, we can justify using the same method of generating a global estimate from several local estimates.

Figure 5.17 shows the results of the coverage analysis of the M/M/1 queueing model for the estimators $\hat{\sigma}^2_{M1}$ and $\hat{\sigma}^2_{BM}$. The estimator $\hat{\sigma}^2_{BM}$ was also tested in its "triple batch size" version; the result is shown in Figure 5.18.

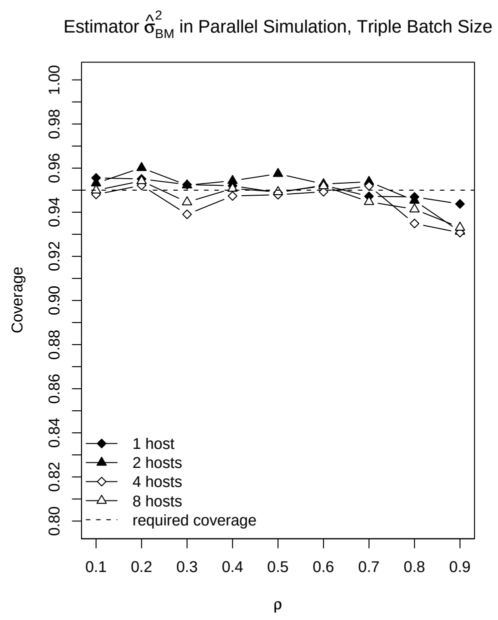


Figure 5.18: Coverage of the Estimator $\hat{\sigma}^2_{BM}$ Using Triple Batch Sizes and Multiple Replications in Parallel, M/M/1 Queue

We see that executing the simulation in parallel produces valid results, and that in fact the use of more than one host generally improves the coverage of confidence intervals.

Chapter 6

Conclusion

Several solutions were found to solve the task of estimating the steady-state variance of a simulated process. We proposed six different estimators and compared their performance using simulation studies. Table 6.1 summarises the main results of these studies.

Some of the estimators were successfully used in parallel simulation, following the concept of Multiple Replications in Parallel.

The detection of the end of the initial transient phase with regard to the process variance was investigated, and a heuristic was proposed. However, no generally reliable method of detection of the initial transient period was found.

6.1 Future Work

An attempt to improve the efficiency of the estimator $\hat{\sigma}^2_{SP}$ resulted in a new estimator $\hat{\sigma}^2_{VB}$ that did not perform well in terms of coverage. Further investigation is needed to find better methods of calculating a confidence interval from the "Vertical Batches" method of variance estimation.

A weakness of the estimator $\hat{\sigma}^2_{BM}$ is the automated detection of the batch size. Simple multiplication of the detected size by three produced good coverage for the M/M/1 queueing model at the tested system loads; however, it would be good to find a more reliable procedure for the detection of appropriate batch sizes.

So far, the proposed estimators have only been tested on three reference models. Further testing is needed to verify that they are usable for a broad range of

Estimator	**Remarks**
$\hat{\sigma}^2_{M1}$	• good coverage
$\hat{\sigma}^2_{M2}$	• coverage not as good as $\hat{\sigma}^2_{M1}$
$\hat{\sigma}^2_{IR}$	• not usable in practice
$\hat{\sigma}^2_{SP}$	• good coverage • needs large sample size
$\hat{\sigma}^2_{VB}$	• attempt to improve $\hat{\sigma}^2_{SP}$ • coverage is not acceptable
$\hat{\sigma}^2_{BM}$	• good coverage for not too high system loads • increasing the batch size leads to better coverage for highly loaded systems

Table 6.1: Summary of the Results of the Simulation Studies

applications.

Further study is also needed in the area of the detection of the initial transient phase of simulation with regard to the process variance.

Appendix A

Results of the Simulation Studies

The experimental results presented in this report were obatined using a combination of C++ programs, the Akaroa2 package (Ewing et al., 2003), and scripts written in R (R Development Core Team, 2007). The figures were created using R.

The following tables contain the data used to plot the figures found in Chapter 5, and give additional information on the results of the respective simulation experiments.

We use the following symbols:

ρ	system load
N_r	total number of replicated simulation runs
$\bar{l}_r$	mean run length of simulations (number of observations)
N_a	number of runs of acceptable length
p_d	proportion of discarded runs of too short length
$\bar{c}$	mean coverage
Δ_c	half width of confidence interval of coverage at 0.95 confidence level
l_{it}	(fixed) length of the initial transient period (estimator $\hat{\sigma}^2_{IR}$ only)
$\bar{k}_0$	mean value of parameter k_0 (estimators $\hat{\sigma}^2_{SP}$ and $\hat{\sigma}^2_{VB}$ only)
$\bar{m}$	mean batch size (estimator $\hat{\sigma}^2_{BM}$ only)

ρ	N_r	$\bar{l}_r$	N_a	p_d	$\bar{c}$	Δ_c
0.1	3422	159565.5	2944	0.140	0.932	0.0091
0.2	2979	130453.2	2565	0.139	0.922	0.0104
0.3	2860	141227.4	2517	0.12	0.92	0.0106
0.4	3441	174081.3	2982	0.133	0.933	0.00898
0.5	3111	236612.9	2713	0.128	0.926	0.00984
0.6	3301	361426.0	2862	0.133	0.93	0.00935
0.7	2950	654935.6	2604	0.117	0.923	0.0102
0.8	2860	1526187.4	2487	0.130	0.92	0.0107
0.9	2943	6492721.0	2595	0.118	0.923	0.0103

Table A.1: Simulation Results: Estimator $\hat{\sigma}^2_{M1}$, M/M/1 Queue

ρ	N_r	$\bar{l}_r$	N_a	p_d	$\bar{c}$	Δ_c
0.1	3464	115896.0	3012	0.130	0.934	0.0089
0.2	3792	94043.49	3270	0.138	0.939	0.00822
0.3	3333	101289.9	2926	0.122	0.932	0.00915
0.4	2938	126003.7	2560	0.129	0.922	0.0104
0.5	3154	173895.1	2766	0.123	0.927	0.00968
0.6	3366	271413.5	2949	0.124	0.932	0.00908
0.7	3490	489542.7	3053	0.125	0.934	0.00878
0.8	3246	1212045.5	2837	0.126	0.93	0.00943
0.9	3371	5700725.7	2987	0.114	0.933	0.00897

Table A.2: Simulation Results: Estimator $\hat{\sigma}^2_{M1}$, M/E_2/1 Queue

ρ	N_r	$\bar{l}_r$	N_a	p_d	$\bar{c}$	Δ_c
0.1	2794	576458.4	2421	0.134	0.917	0.0110
0.2	2929	559461.6	2548	0.13	0.922	0.0104
0.3	3299	631533.2	2870	0.13	0.93	0.00932
0.4	3482	762826.8	3037	0.128	0.934	0.00883
0.5	3241	1020592.3	2831	0.127	0.93	0.00944
0.6	3312	1475524.9	2864	0.135	0.93	0.00934
0.7	3365	2532903.9	2928	0.130	0.932	0.00914
0.8	3353	5442252.5	2930	0.126	0.932	0.00914
0.9	4068	23836873.2	3570	0.122	0.944	0.00755

Table A.3: Simulation Results: Estimator $\hat{\sigma}^2_{M1}$, M/H_2/1 Queue

ρ	N_r	$\bar{l}_r$	N_a	p_d	$\bar{c}$	Δ_c
0.1	3995	156463.8	3422	0.143	0.942	0.00786
0.2	4187	127709.9	3591	0.142	0.944	0.0075
0.3	3712	136782.6	3204	0.137	0.938	0.00838
0.4	3869	168604.3	3359	0.132	0.94	0.008
0.5	3649	231215.3	3169	0.132	0.937	0.00847
0.6	3849	350404.4	3333	0.134	0.94	0.00807
0.7	3391	620616.3	2970	0.124	0.933	0.00902
0.8	3622	1473178.3	3167	0.126	0.937	0.00848
0.9	4021	6321199.9	3508	0.128	0.943	0.00768

Table A.4: Simulation Results: Estimator $\hat{\sigma}^2_{M1}$, M/M/1 Queue, 2 Hosts

ρ	N_r	$\bar{l}_r$	N_a	p_d	$\bar{c}$	Δ_c
0.1	3922	156641.9	3366	0.142	0.94	0.00799
0.2	4276	127924.8	3670	0.142	0.946	0.00735
0.3	4261	137738.1	3646	0.144	0.945	0.0074
0.4	4241	169973.6	3652	0.139	0.945	0.00738
0.5	4185	230796.8	3649	0.128	0.945	0.00739
0.6	3655	355419.6	3231	0.116	0.938	0.00831
0.7	3854	622529.1	3447	0.106	0.942	0.00781
0.8	3498	1472881.1	3131	0.105	0.936	0.00857
0.9	3747	6292133.9	3314	0.116	0.94	0.00811

Table A.5: Simulation Results: Estimator $\hat{\sigma}^2_{M1}$, M/M/1 Queue, 4 Hosts

ρ	N_r	$\bar{l}_r$	N_a	p_d	$\bar{c}$	Δ_c
0.1	4383	159730.3	3745	0.146	0.946	0.00722
0.2	4682	131819.0	4053	0.134	0.95	0.00667
0.3	4425	141388.6	3811	0.139	0.948	0.00708
0.4	4408	174821.8	3829	0.131	0.948	0.00705
0.5	4129	236358.3	3648	0.116	0.945	0.00739
0.6	4147	354997.8	3723	0.102	0.946	0.00725
0.7	3405	614614.9	3051	0.104	0.934	0.00879
0.8	3761	1460034.9	3397	0.0968	0.941	0.00792
0.9	3659	6340499.9	3331	0.0896	0.94	0.00807

Table A.6: Simulation Results: Estimator $\hat{\sigma}^2_{M1}$, M/M/1 Queue, 8 Hosts

ρ	N_r	$\bar{l}_r$	N_a	p_d	$\bar{c}$	Δ_c
0.1	3179	155970.4	2740	0.138	0.927	0.00975
0.2	2736	124700.6	2363	0.136	0.915	0.0112
0.3	2531	130294.6	2225	0.121	0.91	0.0119
0.4	2717	159075.4	2378	0.125	0.916	0.0112
0.5	2157	208136.3	1892	0.123	0.894	0.0139
0.6	2385	319059.6	2099	0.12	0.905	0.0126
0.7	1925	556740.3	1711	0.111	0.883	0.0152
0.8	1940	1257572.7	1693	0.127	0.882	0.0154
0.9	1813	5326784.1	1619	0.107	0.876	0.0160

Table A.7: Simulation Results: Estimator $\hat{\sigma}^2_{M2}$, M/M/1 Queue

ρ	N_r	$\bar{l}_r$	N_a	p_d	$\bar{c}$	Δ_c
0.1	3409	112390.0	2958	0.132	0.932	0.00905
0.2	3104	89223.25	2675	0.138	0.925	0.00997
0.3	2932	93660.58	2581	0.120	0.923	0.0103
0.4	2429	114197.1	2135	0.121	0.906	0.0124
0.5	2316	151818.9	2023	0.127	0.901	0.0130
0.6	2523	231521.6	2224	0.119	0.91	0.0119
0.7	2408	412048.7	2096	0.130	0.905	0.0126
0.8	1868	922983.3	1637	0.124	0.878	0.0159
0.9	1804	3974583.6	1596	0.115	0.875	0.0163

Table A.8: Simulation Results: Estimator $\hat{\sigma}^2_{M2}$, M/E_2/1 Queue

ρ	N_r	$\bar{l}_r$	N_a	p_d	$\bar{c}$	Δ_c
0.1	2754	567853.7	2387	0.133	0.916	0.0111
0.2	2749	546208.5	2384	0.133	0.916	0.0112
0.3	3063	606151.8	2673	0.127	0.925	0.00998
0.4	2897	723722.1	2522	0.129	0.92	0.0106
0.5	2754	943514.2	2418	0.122	0.917	0.0110
0.6	2736	1345144.7	2381	0.130	0.916	0.0111
0.7	2422	2274286.5	2121	0.124	0.906	0.0124
0.8	2317	4769161.2	2015	0.130	0.9	0.0131
0.9	2208	17324261.7	1936	0.123	0.897	0.0136

Table A.9: Simulation Results: Estimator $\hat{\sigma}^2_{M2}$, M/H_2/1 Queue

ρ	l_{it}	N_r	$\bar{l}_r$	N_a	p_d	$\bar{c}$	Δ_c
0.1	4	3893	90061.26	3288	0.155	0.94	0.00817
0.2	6	3581	44307.95	3037	0.152	0.934	0.00883
0.3	10	3635	29364.43	3076	0.154	0.935	0.00872
0.4	16	3158	22023.02	2671	0.154	0.925	0.00999
0.5	27	3544	17914.63	3020	0.148	0.934	0.00887
0.6	49	3721	15287.90	3176	0.146	0.937	0.00845
0.7	99	3976	13574.05	3342	0.159	0.94	0.00805
0.8	252	3852	12581.04	3280	0.148	0.939	0.0082
0.9	1130	3555	11995.23	3019	0.151	0.934	0.00888

Table A.10: Simulation Results: Estimator $\hat{\sigma}^2_{IR}$, M/M/1 Queue, Fixed Length of Initial Transient Period

ρ	N_r	$\bar{l}_r$	N_a	p_d	$\bar{k}_0$	$\bar{c}$	Δ_c
0.1	4236	568029.7	3871	0.0862	6.6	0.948	0.00698
0.2	3995	364515.0	3492	0.126	8.6	0.943	0.00771
0.3	4586	338841.7	4019	0.124	11.9	0.95	0.00673
0.4	4348	373522.2	3778	0.131	17.4	0.947	0.00714
0.5	4532	462866.5	3978	0.122	26.4	0.95	0.0068
0.6	3926	648946.7	3405	0.133	43.3	0.941	0.0079
0.7	3763	1078117.9	3274	0.13	80.6	0.939	0.0082
0.8	3804	2334350.9	3309	0.13	187.7	0.94	0.00812
0.9	3639	9110571.6	3141	0.137	771.2	0.936	0.00854

Table A.11: Simulation Results: Estimator $\hat{\sigma}^2_{SP}$, M/M/1 Queue

ρ	N_r	$\bar{l}_r$	N_a	p_d	$\bar{k}_0$	$\bar{c}$	Δ_c
0.1	4323	409197.7	3739	0.135	6.2	0.947	0.00722
0.2	4674	266053.5	3757	0.196	7.9	0.947	0.00718
0.3	4513	249315.2	3787	0.161	10.5	0.947	0.00713
0.4	4367	273554.2	3782	0.134	14.5	0.947	0.00714
0.5	4137	346372.2	3598	0.130	21.8	0.944	0.00749
0.6	4162	497603.8	3592	0.137	35.3	0.944	0.0075
0.7	4506	829112.2	3912	0.132	64.2	0.949	0.0069
0.8	3724	1773436.9	3243	0.129	146.1	0.938	0.00828
0.9	3246	6954090.9	2893	0.109	588.0	0.93	0.00925

Table A.12: Simulation Results: Estimator $\hat{\sigma}^2_{SP}$, M/E_2/1 Queue

ρ	N_r	$\bar{l}_r$	N_a	p_d	$\bar{k}_0$	$\bar{c}$	Δ_c
0.1	4210	1747547.5	3863	0.0824	10.8	0.948	0.00699
0.2	4609	1303795.9	4000	0.132	17.4	0.95	0.00676
0.3	4192	1291397.7	3624	0.135	27.5	0.945	0.00744
0.4	3608	1431362.0	3149	0.127	43.2	0.936	0.00852
0.5	3945	1741375.2	3455	0.124	70.0	0.942	0.00779
0.6	3932	2345324.5	3439	0.125	119.6	0.942	0.00783
0.7	3597	3636324.1	3123	0.132	227.4	0.936	0.00859
0.8	3608	7313727.9	3187	0.117	536.5	0.937	0.00842
0.9	3282	26805842.2	2888	0.12	2203.9	0.935	0.00898

Table A.13: Simulation Results: Estimator $\hat{\sigma}^2_{SP}$, M/H_2/1 Queue

ρ	N_r	$\bar{l}_r$	N_a	p_d	$\bar{k}_0$	$\bar{c}$	Δ_c
0.1	2710	117501.0	2312	0.147	6.4	0.913	0.0115
0.2	1730	80937.92	1490	0.139	8.3	0.866	0.0173
0.3	1489	77353.19	1288	0.135	11.6	0.845	0.0198
0.4	1219	87371.7	1056	0.134	16.6	0.81	0.0237
0.5	1107	109376.6	947	0.145	25.6	0.789	0.0260
0.6	1095	163527.4	942	0.140	41.6	0.788	0.0262
0.7	1068	279161.6	906	0.152	78.4	0.78	0.0271
0.8	765	644022.6	667	0.128	183.1	0.7	0.0349
0.9	917	2602474.9	811	0.116	741.9	0.753	0.0297

Table A.14: Simulation Results: Estimator $\hat{\sigma}^2_{VB}$, M/M/1 Queue

ρ	N_r	$\bar{l}_r$	N_a	p_d	$\bar{k}_0$	$\bar{c}$	Δ_c
0.1	2677	87216.77	2298	0.142	5.9	0.913	0.0115
0.2	2157	60196.1	1846	0.144	7.4	0.892	0.0142
0.3	1550	57124.65	1349	0.130	10.0	0.852	0.019
0.4	1235	64099.27	1072	0.132	14.1	0.813	0.0234
0.5	1068	81644.85	935	0.125	21.2	0.786	0.0263
0.6	1015	121542.8	885	0.128	34.2	0.774	0.0276
0.7	1093	213085.1	925	0.154	63.4	0.784	0.0266
0.8	1000	477600.7	860	0.14	142.3	0.767	0.0283
0.9	890	2028923.7	798	0.103	584.3	0.75	0.0301

Table A.15: Simulation Results: Estimator $\hat{\sigma}^2_{VB}$, M/E_2/1 Queue

ρ	N_r	$\bar{l}_r$	N_a	p_d	$\bar{k}_0$	$\bar{c}$	Δ_c
0.1	1844	342775.4	1559	0.155	10.4	0.872	0.0166
0.2	1566	287442.9	1346	0.140	16.8	0.851	0.0190
0.3	1160	289583.5	1000	0.138	26.7	0.8	0.0248
0.4	1286	338086.8	1101	0.144	41.6	0.818	0.0228
0.5	1031	427970.7	898	0.129	68.3	0.777	0.0273
0.6	1132	602230.7	970	0.143	117.6	0.794	0.0255
0.7	1068	979000.1	922	0.137	223.1	0.783	0.0267
0.8	958	2064435.9	839	0.124	526.2	0.762	0.0289
0.9	951	7923091.1	827	0.130	2170.0	0.758	0.0292

Table A.16: Simulation Results: Estimator $\hat{\sigma}^2_{VB}$, M/H_2/1 Queue

ρ	$\bar{m}$	N_r	$\bar{l}_r$	N_a	p_d	$\bar{c}$	Δ_c
0.3	10	2968	108529.7	2580	0.131	0.922	0.0103
0.3	20	4407	123528.3	3860	0.124	0.948	0.007
0.3	50	4774	133202.1	4194	0.121	0.952	0.00645
0.3	100	4850	136369.0	4256	0.122	0.953	0.00636
0.3	200	5363	138076.2	4659	0.131	0.957	0.00582
0.3	500	4911	138596.7	4376	0.109	0.954	0.00619
0.3	1000	3963	137182.6	3609	0.0893	0.945	0.00747
0.3	2000	3747	135548.2	3353	0.105	0.94	0.00802
0.3	5000	3657	135387.4	3125	0.145	0.936	0.00859
0.6	10	954	118438.7	826	0.134	0.758	0.0293
0.6	20	1776	188031.8	1516	0.146	0.868	0.0171
0.6	50	3854	276421.4	3358	0.129	0.94	0.00801
0.6	100	4474	317103.8	3931	0.121	0.95	0.00687
0.6	200	4990	338497.0	4362	0.126	0.954	0.00621
0.6	500	5219	351383.7	4565	0.125	0.956	0.00594
0.6	1000	5166	354291.5	4589	0.112	0.956	0.00591
0.6	2000	4810	353492.7	4326	0.101	0.954	0.00626
0.6	5000	4168	349153.1	3713	0.109	0.946	0.00726
0.9	10	6433	186095.3	6433	0	0.193	0.00964
0.9	20	3983	287227.6	3983	0	0.279	0.0139
0.9	50	2125	594817.6	2047	0.0367	0.429	0.0215
0.9	100	1308	1078532.6	1210	0.0749	0.56	0.028
0.9	200	864	1893916.5	768	0.111	0.74	0.0311
0.9	500	1727	3495030.5	1531	0.113	0.87	0.0169
0.9	1000	3200	4724645.1	2786	0.129	0.928	0.0096
0.9	2000	4423	5525862.2	3868	0.125	0.948	0.00698
0.9	5000	5246	6050275.8	4629	0.118	0.957	0.00586

Table A.17: Simulation Results: Estimator $\hat{\sigma}^2_{BM}$, M/M/1 Queue, Fixed Batch Sizes

ρ	N_r	$\bar{l}_r$	N_a	p_d	$\bar{m}$	$\bar{c}$	Δ_c
0.1	5180	158131.4	4485	0.134	116.4771	0.955	0.00604
0.2	5712	128073.4	4949	0.134	117.8218	0.96	0.00549
0.3	4949	136259.9	4340	0.123	120.6797	0.954	0.00624
0.4	5258	164799.7	4608	0.124	122.4935	0.957	0.00589
0.5	5518	219976.4	4802	0.130	128.0092	0.958	0.00565
0.6	4841	323267.7	4253	0.121	138.4082	0.953	0.00636
0.7	4284	535344.6	3751	0.124	173.5137	0.947	0.0072
0.8	3002	1128131.3	2575	0.142	278.4466	0.922	0.0103
0.9	2133	4012287.1	1847	0.134	796.7244	0.892	0.0142

Table A.18: Simulation Results: Estimator $\hat{\sigma}^2_{BM}$, M/M/1 Queue

ρ	N_r	$\bar{l}_r$	N_a	p_d	$\bar{m}$	$\bar{c}$	Δ_c
0.1	5180	158131.4	4485	0.134	116.5	0.955	0.00604
0.2	5712	128073.4	4949	0.134	117.8	0.96	0.00549
0.3	4949	136259.9	4340	0.123	120.7	0.954	0.00624
0.4	5258	164799.7	4608	0.124	122.5	0.957	0.00589
0.5	5518	219976.4	4802	0.130	128.0	0.958	0.00565
0.6	4841	323267.7	4253	0.121	138.4	0.953	0.00636
0.7	4284	535344.6	3751	0.124	173.5	0.947	0.0072
0.8	3002	1128131.3	2575	0.142	278.4	0.922	0.0103
0.9	2133	4012287.1	1847	0.134	796.7	0.892	0.0142

Table A.19: Simulation Results: Estimator $\hat{\sigma}^2_{BM}$, M/E_2/1 Queue

ρ	N_r	$\bar{l}_r$	N_a	p_d	$\bar{m}$	$\bar{c}$	Δ_c
0.1	5403	573397.0	4706	0.129	128.8	0.958	0.00577
0.2	4941	542739.8	4371	0.115	130.3	0.954	0.0062
0.3	5361	591407.0	4730	0.118	135.4	0.958	0.00574
0.4	4470	686771.6	3899	0.128	144.0	0.949	0.00693
0.5	4548	845355.6	3975	0.126	168.3	0.95	0.0068
0.6	3353	1136562.7	2934	0.125	214.7	0.932	0.00912
0.7	2884	1724031.9	2492	0.136	312.6	0.92	0.0107
0.8	2558	3314494.3	2216	0.134	594.0	0.91	0.0119
0.9	2123	11477138.1	1845	0.131	2010.7	0.892	0.0142

Table A.20: Simulation Results: Estimator $\hat{\sigma}^2_{BM}$, M/H_2/1 Queue

ρ	N_r	$\bar{l}_r$	N_a	p_d	$\bar{m}$	$\bar{c}$	Δ_c
0.1	5501	158511.8	4799	0.128	117.0	0.958	0.00566
0.2	6038	128601.4	5216	0.136	117.4	0.962	0.00521
0.3	6135	136408.3	5326	0.132	119.8	0.962	0.00511
0.4	6064	164978.9	5312	0.124	122.4	0.962	0.00512
0.5	5535	221365.1	4849	0.124	127.0	0.959	0.0056
0.6	5003	325385.4	4360	0.129	137.6	0.954	0.00621
0.7	4543	540155.4	3988	0.122	173.0	0.95	0.00678
0.8	3255	1127863.0	2795	0.141	275.1	0.928	0.00956
0.9	2163	4075855.4	1899	0.122	784.8	0.895	0.0138

Table A.21: Simulation Results: Estimator $\hat{\sigma}^2_{BM}$, M/M/1 Queue, 2 Hosts

ρ	N_r	$\bar{l}_r$	N_a	p_d	$\bar{m}$	$\bar{c}$	Δ_c
0.1	4619	161407.6	4001	0.134	114.9	0.95	0.00676
0.2	5397	132447.9	4658	0.137	114.7	0.957	0.00582
0.3	5390	140625.6	4627	0.142	116.1	0.957	0.00586
0.4	5034	168769.1	4365	0.133	119.2	0.954	0.0062
0.5	5294	222845.0	4602	0.131	125.2	0.957	0.00589
0.6	4881	329039.2	4287	0.122	136.2	0.953	0.00632
0.7	4284	544497.9	3761	0.122	170.5	0.947	0.00717
0.8	3388	1157388.9	2955	0.128	273.6	0.932	0.00906
0.9	2462	4165891.9	2148	0.128	775.1	0.907	0.0123

Table A.22: Simulation Results: Estimator $\hat{\sigma}^2_{BM}$, M/M/1 Queue, 4 Hosts

ρ	N_r	$\bar{l}_r$	N_a	p_d	$\bar{m}$	$\bar{c}$	Δ_c
0.1	5904	165116.5	5096	0.137	111.3	0.96	0.00533
0.2	5265	135854.6	4540	0.138	111.1	0.956	0.00597
0.3	5357	143665.3	4657	0.131	111.7	0.957	0.00582
0.4	4561	171560.0	3942	0.136	113.8	0.95	0.00685
0.5	5352	224558.5	4643	0.132	117.9	0.957	0.00584
0.6	4404	327938.1	3846	0.127	128.0	0.948	0.00702
0.7	3557	548429.3	3095	0.13	162.4	0.935	0.00867
0.8	2981	1178855.9	2576	0.136	267.4	0.922	0.0103
0.9	2852	4301986.0	2459	0.138	759.0	0.919	0.0108

Table A.23: Simulation Results: Estimator $\hat{\sigma}^2_{BM}$, M/M/1 Queue, 8 Hosts

ρ	N_r	$\bar{l}_r$	N_a	p_d	$\bar{m}$	$\bar{c}$	Δ_c
0.1	5065	347225.2	4500	0.112	358.5	0.956	0.00602
0.2	4950	312371.1	4453	0.100	357.4	0.955	0.00609
0.3	4709	323868.5	4212	0.106	365.9	0.953	0.00643
0.4	4634	358401.0	4168	0.101	369.9	0.952	0.00649
0.5	4267	424745.4	3911	0.0834	386.9	0.949	0.0069
0.6	4693	551522.1	4188	0.108	417.4	0.952	0.00646
0.7	4268	832195.9	3788	0.112	521.3	0.947	0.00712
0.8	4249	1674829.2	3770	0.113	831.9	0.947	0.00716
0.9	3956	6192263.0	3556	0.101	2397.8	0.944	0.00758

Table A.24: Simulation Results: Estimator $\hat{\sigma}^2_{BM}$, M/M/1 Queue, Triple Batch Size

ρ	N_r	$\bar{l}_r$	N_a	p_d	$\bar{m}$	$\bar{c}$	Δ_c
0.1	4797	503122.3	4269	0.11	356.6	0.953	0.00634
0.2	5685	464273.8	5016	0.118	357.3	0.96	0.00542
0.3	4629	476861.9	4189	0.095	363.2	0.952	0.00646
0.4	4879	518039.7	4373	0.104	369.8	0.954	0.0062
0.5	5177	592980.4	4700	0.0921	385.9	0.957	0.00577
0.6	4741	733230.7	4235	0.107	417.4	0.953	0.00639
0.7	4785	1047690.5	4331	0.0949	519.5	0.954	0.00625
0.8	4073	1967212.5	3657	0.102	826.3	0.945	0.00737
0.9	3189	6854264.1	2909	0.0878	2346.6	0.931	0.0092

Table A.25: Simulation Results: Estimator $\hat{\sigma}^2_{BM}$, M/M/1 Queue, Triple Batch Size, 2 Hosts

ρ	N_r	$\bar{l}_r$	N_a	p_d	$\bar{m}$	$\bar{c}$	Δ_c
0.1	4176	778289.9	3850	0.078	349.1	0.948	0.00701
0.2	4936	721140.1	4182	0.153	347.7	0.952	0.00647
0.3	3773	740068.5	3281	0.130	353.7	0.939	0.00819
0.4	4167	796766.6	3801	0.0878	361.7	0.947	0.0071
0.5	4227	893061.6	3841	0.0913	380.7	0.948	0.00703
0.6	4456	1060849.7	3963	0.111	416.0	0.95	0.00683
0.7	4609	1420966.9	4160	0.0974	516.8	0.952	0.0065
0.8	3415	2477834.4	3072	0.100	824.5	0.935	0.00873
0.9	3196	8002809.0	2887	0.0967	2330.2	0.93	0.00927

Table A.26: Simulation Results: Estimator $\hat{\sigma}^2_{BM}$, M/M/1 Queue, Triple Batch Size, 4 Hosts

ρ	N_r	$\bar{l}_r$	N_a	p_d	$\bar{m}$	$\bar{c}$	Δ_c
0.1	4702	1060006.6	3996	0.15	333.6	0.95	0.00676
0.2	5192	895364.9	4365	0.159	333.0	0.954	0.0062
0.3	4277	924194.1	3608	0.156	334.9	0.945	0.00747
0.4	4784	1077111.3	4060	0.151	340.7	0.95	0.00666
0.5	4709	1306670.9	3937	0.164	360.5	0.95	0.00686
0.6	4845	1599987.7	4157	0.142	402.2	0.952	0.00651
0.7	4080	2059158.3	3618	0.113	506.5	0.945	0.00745
0.8	3859	3332560.1	3411	0.116	814.4	0.941	0.00789
0.9	3317	9855626.5	2988	0.0992	2303.6	0.933	0.00897

Table A.27: Simulation Results: Estimator $\hat{\sigma}^2_{BM}$, M/M/1 Queue, Triple Batch Size, 8 Hosts

Bibliography

Anderson, T. W. *The Statistical Analysis of Time Series*. John Wiley & Sons, 1971.

Billingsley, P. *Convergence of Probability Measures*. John Wiley & Sons, 1968.

Cramér, Harald. *Mathematical Methods of Statistics*. Princeton University Press, 1946.

Daley, D. J. The serial correlation coefficients of waiting times in a stationary single server queue. *Journal of the Australian Mathematical Society*, 8:683–699, 1968.

Ewing, Greg, Pawlikowski, Krzysztof, and McNickle, Don. *Akaroa 2.8, User's Manual*, April 2003. `http://akaroa.canterbury.ac.nz`.

Feldman, Richard M., Deuermeyer, Bryan L., and Yang, Yoonjung. Estimation of process variance by simulation. unpublished paper, 1996.

Gafarian, A. V., Ancker, C. J., and Morisaku, T. Evaluation of commonly used rules for detecting "steady state" in computer simulation. *Naval Research Logistics Quarterly*, 25:511–529, 1978.

Goldsman, D., Schruben, L. W., and Swain, J. J. Tests for transient means in simulated time series. *Naval Research Logistics*, 41:171–187, 1994.

Gross, Donald and Harris, Carl M. *Fundamentals of Queueing Theory*. John Wiley & Sons, 2nd edition, 1985.

Heidelberger, Philip and Welch, Peter D. A spectral method for confidence interval generation and run length control in simulations. *Communications of the ACM*, 25:233–245, 1981.

Jackway, Paul and de Silva, Basil. A methodology for initialisation bias reduction in computer simulation output. *Asia-Pacific Journal of Operational Research*, 9:87–100, 1992.

Kelton, W. David and Law, Averill M. The transient behavior of the m/m/s queue, with implications for steady-state simulation. *Operations Research*, 33:378–396, 1985.

Law, Averill M. Confidence intervals in discrete event simulation: A comparison of replication and batch means. *Naval Research Logistics Quarterly*, 24:667–678, 1977.

Law, Averill M. Statistical analysis of simulation output data. *Operations Research*, 31(6):983–1029, 1983.

Law, Averill M. and Kelton, W. David. *Simulation Modeling and Analysis*. McGraw-Hill, 2nd edition, 1991.

Lehmann, Erich L. *Theory of Point Estimation*. John Wiley & Sons, 1983.

McNickle, Don, Ewing, Gregory, and Pawlikowski, Krzysztof. Refining spectral analysis for confidence interval estimation in sequential simulation. In *Proceedings of the European Simulation Symposium, ESS 2004, Budapest, Hungary*, pages 99–103, Oct. 2004.

McNickle, Don, Ewing, Greg, and Pawlikowski, Krzysztof. Some effects of transient deletion on sequential steady-state simulation. paper in preparation, 2008.

Meketon, Marc S. and Schmeiser, Bruce. Overlapping batch means: Something for nothing? In *Proceedings of the 1984 Winter Simulation Conference*, pages 227–230. IEEE, 1984.

Odoni, A. R. and Roth, E. An empirical investigation of the transient behaviour of stationary queueing systems. *Operations Research*, 31:432–455, 1983.

Pawlikowski, Krzysztof. Steady-state simulation of queueing processes: A survey of problems and solutions. *ACM Computing Surveys*, 22(2):123–170, 1990.

Pawlikowski, Krzysztof and McNickle, Don. Speeding up stoachstic discrete-event simulation. In *Proceedings of the European Simulation Symposium, ESS 2001, Marseille, France*, pages 132–138, Oct. 2001.

Pawlikowski, Krzysztof, Yau, Victor, and McNickle, Don. Distributed stochastic discrete-event simulation in parallel time streams. In *Proceedings of the 1994 Winter Simulation Conference*, pages 723–730, 1994.

Pawlikowski, Krzysztof, Ewing, Greg, and McNickle, Don. Coverage of confidence intervals in sequential steady-state simulation. *Journal of Simulation Practice and Theory*, 6(3):255–267, 1998.

Pawlikowski, Krzysztof, Jeong, H.-D. Joshua, and Lee, J-S. Ruth. On credibility of simulation studies of telecommunication networks. *IEEE Communications Magazine*, pages 132–139, January 2002.

R Development Core Team. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria, 2007. ISBN 3-900051-07-0. `http://www.R-project.org`.

Schruben, L., Singh, H., and Tierney, L. Optimal tests for initialization bias in simulation output. *Operations Reserarch*, 31(6):1167–1178, November 1983.

Tanenbaum, Andrew S. *Computer Networks*. Prentice Hall, 4th edition, 2003.

Trivedi, Kishor S. *Probability and Statistics with Reliability, Queuing, and Computer Science Applications*. John Wiley & Sons, 2nd edition, 2002.

Waterloo Maple Inc. *Maple, Version 12*, 2008.

Wilks, Samuel S. *Mathematical Statistics*. John Wiley & Sons, 2nd edition, 1962.

Yücesan, Enver. Randomization tests for initialization bias in simulation output. *Naval Research Logistics*, 40:643–663, 1993.