ORSA Journal on Computing
Vol. 4, No. 3, Summer 1992

0899-1499/92/0403-0226 $01.25

Jackknifing under a Budget Constraint

PETER W. GLYNN / *Department of Operations Research, Stanford University, Stanford, CA 94305-4022, e-mail: glynn@leland.stanford.edu*

PHILIP HEIDELBERGER / *IBM T. J. Watson Research Center, Yorktown Heights, NY 10598, e-mail: berger@watson.ibm.com*

(Received: December 1989; final revision received: June 1991; accepted: July 1991)

In this paper, we consider the problem of estimating a parameter α that can be expressed as a nonlinear function of sample means. We develop a jackknife estimator for α that is appropriate to computational settings in which the total computer budget to be used is constrained. Despite the fact that the jackknifed observations are not i.i.d., we are able to show that our jackknife estimator reduces bias without increasing asymptotic variance. This makes the estimator particularly suitable for small sample applications. Because a special case of this estimator problem is that of estimating a ratio of two means, the results in this paper are pertinent to regenerative steady-state simulations.

Consider the problem of estimating, via simulation, the parameter $\alpha = g(\mu)$ where $g:\mathbb{R}^d \to \mathbb{R}$ is a (possibly) nonlinear function and μ is expressible as the mean of an $\mathbb{R}^d$-valued random vector X. This estimation problem arises in certain terminating simulation settings, as well as in steady-state regenerative simulations (see Section 2 for examples of specific applications). This paper is concerned with the problem of estimating α when a budget constraint t, representing the maximum amount of computer time to be used, is imposed on the simulation. Two major statistical difficulties arise in this estimation setting. First, the number of replications $N(t)$ of the r.v. X generated within the budget constraint t is random. Consequently, (deterministic) fixed sample size theory does not apply to this class of estimation problems. The second difficulty is that the nonlinearity in the function g produces bias in the estimation of α. A standard statistical technique used to deal with bias that emanates from nonlinearities of this kind is to jackknife the estimator (see, for example, Miller[19] and Iglehart[13]). However, the jackknife literature typically requires that the observations being jackknifed be i.i.d. Unfortunately, when a budget constraint is imposed, the observations are clearly no longer independent. The major contribution of this paper is to show that the dependence induced by the presence of the budget constraint does not destroy the bias-reducing properties of the jackknife. As indicated earlier, the jackknifed estimator that we obtain here has important implications for steady-state regenerative simulation. In particular, the estimator obtained here has the same bias-reducing properties as the regenerative low-bias estimator introduced by Meketon and Heidelberger.[17]

This paper is organized as follows. In Section 2, we precisely describe the estimation problem and give examples of various applications settings. We then proceed to describe the major results of this paper. Section 3 discusses the empirical behavior of our budget-constrained jackknife. We defer all proofs to Section 4.

1. Description of Main Results

Suppose that X is an $\mathbb{R}^d$-valued random vector with finite mean μ. As stated in the Introduction, our goal in this paper is to estimate $\alpha = g(\mu)$, subject to a budget constraint on the total amount t of computer time to be used. Throughout this paper, we assume that g: $\mathbb{R}^d \to \mathbb{R}$ has continuous second partial derivatives in some open neighborhood of μ. This type of estimation problem arises in several different problem settings.

Example 1. Given a stochastic system Y, we may be interested in the performance of the system over some (finite) time horizon T (T may be deterministic or it may be a stopping time with respect to Y; see p. 322 of Chung[2] for a definition of stopping time). Let Z be a real-valued performance measure that depends on Y over the above time horizon, so that Z can be represented as $Z = f(Y(s)$: $0 \leqslant s \leqslant T)$. Estimation of both the mean and variance of the performance measure Z are special cases of the estimation problem considered in this paper. To incorporate the estimation of $\alpha = EZ$ into our framework, we set $X = Z$ and $g(x) = x$. For the variance, we let $X = (Z^2, Z)$ and $g(x_1, x_2) = x_1 - x_2^2$. Note that in the case of the variance, g is a nonlinear mapping.

Example 2. Suppose that Y is a real-valued non-delayed regenerative process and that Z is the β-discounted cost associated with Y, namely

$$Z = \int_0^\infty \exp(-\beta s)Y(s)\,ds.$$

It is shown in Fox and Glynn[7] that $\alpha = EZ$ can be re-expressed as $\alpha = g(EX)$, where

$$X = \left(\int_0^\eta \exp(-\beta s)Y(s)\,ds,\quad \exp(-\beta\eta)\right),$$

Subject classifications: Simulation: statistical analysis; bias reduction probability: renewal processes.
Other key words: Bias; jackknife; regenerative simulation.

η is the first positive regeneration time of Y, and $g(x_1, x_2) = x_1(1 - x_2)^{-1}$. The advantage of this alternative representation is that the random vector X can be generated in finite time, whereas Z typically cannot. It is further shown in Fox and Glynn[7] that the expected β-discounted cost associated with a delayed regenerative process can also be expressed in the form $\alpha = g(EX)$, although the precise nature of X and g is then more complicated.

Example 3. Let Y be a non-delayed real-valued regenerative process and suppose that η is the first positive regeneration time of Y. Regenerative process theory (see, for example, Smith[21]) shows that the steady-state mean α of the process Y can be represented as the ratio

$$\alpha = E\int_0^{\eta} Y(s)\,ds/E\eta.$$

The above **ratio estimation problem** can be incorporated into our set-up by letting

$$X = \left(\int_0^{\eta} Y(s)\,ds, \quad \eta\right)$$

and putting $g(x_1, x_2) = x_1/x_2$. Thus, we conclude that the problem of estimating the steady-state mean of a regenerative stochastic process is a special case of the estimation problem considered in this paper. The observation that steady-state estimation for regenerative processes is a special case of ratio estimation lies at the heart of the regenerative method for steady-state simulation output analysis (see Crane and Lemoine[3] for a more complete description of the method).

Example 4. Suppose that rather than estimating the steady-state mean of a regenerative stochastic process, we now wish to estimate the steady-state variance. To be more precise, suppose that Y is a real-valued non-delayed regenerative process. Then, under suitable regularity conditions on Y, there exists a r.v. Y^* such that $Y(t) \Rightarrow Y^*$ as $t \to \infty$; Y^* has the steady-state distribution associated with the regenerative process Y. The problem of estimating $\alpha = EY^*$ was discussed in Example 3. It turns out that the regenerative structure of Y can also be fruitfully exploited to estimate $\alpha = \operatorname{var} Y^*$. More specifically, $\alpha = g(EX)$, where

$$X = \left(\int_0^{\eta} Y^2(s)\,ds, \quad \int_0^{\eta} Y(s)\,ds, \quad \eta\right),$$

η is the first positive regenerative time of Y, and $g(x_1, x_2, x_3) = (x_1/x_3) - (x_2/x_3)^2$. The problem of estimating $\alpha = \operatorname{var} Y^*$, when no budget constraint is present, is discussed in Glynn and Iglehart.[8]

Returning to the estimation problem at hand, our estimators are obtained by generating i.i.d. copies $X_1, X_2, \ldots$ of the r.v. X. Suppose that τ_i is the computer time required to generate X_i. We assume throughout this paper that the sequence of pairs $\{(X_n, \tau_n) : n \geq 1\}$ are i.i.d. However, we permit X_n and τ_n to be dependent r.v.'s. Indeed, in most applications, X_n and τ_n will be strongly correlated.

Given a budget constraint t, let $N(t) = \max\{n \geq 0: \sum_{i=1}^{n} \tau_i \leq t\}$ be the number of observations completed by time t. The classical estimator for α is then defined by

$$\alpha(t) = \begin{cases} g(\overline{X}(t)); & N(t) \geq 1 \\ 0; & N(t) = 0, \end{cases}$$

where $\overline{X}(t) = N(t)^{-1}\sum_{i=1}^{N(t)} X_i$. We adopt the convention that $\overline{X}(t) = 0$ if $N(t) = 0$. This estimator, while enjoying reasonable large-sample behavior, suffers from small-sample bias that can substantially degrade the performance of the estimator. To precisely describe the bias characteristics of $\alpha(t)$ requires some control on the growth of g.

Definition. Let $\|\cdot\|$ be the Euclidian norm on $\mathbb{R}^d$. Suppose X is an $\mathbb{R}^d$-valued r.v. such that $\mu = EX$ exists and is finite-valued. We say that g is polynomially dominated to degree r ($r \geq 0$) if there exist constants A and B such that

$$|g(x)| \leq A + B\|x - \mu\|^r$$

for all $x \in \mathscr{C}$, where $\mathscr{C}$ is the convex hull of the support of the distribution of X (see p. 31 of Chung[2] for a definition of the support of an $\mathbb{R}^d$-valued r.v.).

To obtain some feel for this condition, let us note that if g is bounded, then g is polynomially dominated to degree 0. This also occurs when the X_i's lie almost surely in some set on which g is bounded. This situation is illustrated by Example 3 when $|Y(s)| \leq M$, in which case $|\alpha(t)| \leq M$. Finally, we note that if all the partial derivatives of g of order r are globally bounded (i.e.,

$$\sup\left\{\left|\frac{\partial^r}{\partial_{x_1}^{i_1} \ldots \partial x_d^{i_d}} g(x)\right| : x \in \mathbb{R}^d\right\} < \infty$$

for all collections $(i_1, \ldots, i_d)$ such that $i_1 + \ldots + i_d = r$), then g is polynomially dominated to degree r. For $1 \leq i, j \leq d$, let

$$G_{ij} = \frac{\partial^2}{\partial x_i \partial x_j} g(\mu)$$

$$C_{ij} = E[(X_k(i) - \mu(i))(X_k(j) - \mu(j))],$$

where $X_k = (X_k(1), \ldots, X_k(d))$, $\mu = (\mu(1), \ldots, \mu(d))$. Set $\lambda = 1/E\tau_k$.

We are now ready to state our small-sample bias expansion for $\alpha(t)$. It is a generalization (to unbounded g) of Theorem 4.3 of Glynn and Heidelberger.[10]

Theorem 1. *Suppose that g is polynomially dominated to degree $r (r \geq 0)$ and $0 < \lambda < \infty$. Let $p = \max(2, r)$. If there exists $\delta > 0$ such that $E\|X\|^{2p+1+\delta} < \infty$ and $E\tau^{2p+\delta} < \infty$, then*

$$E\alpha(t) = \alpha + \frac{b}{t} + o\left(\frac{1}{t}\right)$$

as $t \to \infty$, where

$$b = \frac{\lambda^{-1}}{2} \sum_{i=1}^{d} \sum_{j=1}^{d} G_{ij} C_{ij}.$$

In view of the importance of steady-state simulation, we provide the following corollary.

Corollary 1. *Consider Example 3. If either:*

(*i*) $\sup\{|Y(s,\omega)|: \ s \geq 0, \ \omega \in \Omega\} < \infty$, $E\eta^{5+\delta} < \infty$, *and* $E\tau^{4+\delta} < \infty$ *or*
(*ii*) *there exists* $\epsilon > 0$ *such that* $P\{\eta \geq \epsilon\} = 1$, $E\|X\|^{5+\delta} < \infty$, *and* $E\tau^{4+\delta} < \infty$ *then*

$$E\alpha(t) = \alpha + b/t + o(1/t),$$

where

$$b = -\lambda^{-1} \cdot (E\eta)^{-2} \cdot E\left[\eta \int_0^\eta (Y(s) - \alpha)\right].$$

This generalizes a bias expression due to Meketon and Heidelberger[17]. Their discussion assumes that $\eta = \tau$ or, equivalently, that the amount of computer time expended to generate t units of simulated time is precisely equal to t. By contrast, we are not assuming here that computer time and simulated time are identical.

In addition to the bias, the quality of an estimator is also largely determined by its asymptotic variability. The next result characterizes the central limit behavior of $\alpha(t)$; its proof appears in Glynn and Heidelberger[10] as part of Theorem 4.1.

Theorem 2. *Suppose that* $0 < \lambda < \infty$ *and that* $E\|X\|^2 < \infty$. *Then,*

$$t^{1/2}(\alpha(t) - \alpha) \Rightarrow \sigma N(0,1)$$

as $t \to \infty$, *where*

$$\sigma^2 = \lambda^{-1} \sum_{i=1}^{d} \sum_{j=1}^{d} \frac{\partial g}{\partial x_i}(\mu) \frac{\partial g}{\partial x_j}(\mu) C_{ij}.$$

Assuming that $\alpha(t)$ is appropriately uniformly integrable, Theorem 2 implies that $\operatorname{var} \alpha(t) \sim \sigma^2/t$ as $t \to \infty$ (when $\sigma^2 > 0$).

Our goal is now to produce an estimator that has the same asymptotic variability as $\alpha(t)$, but has improved small-sample bias behavior. It is well known that one effective technique for dealing with bias associated with nonlinear functions of sample means of i.i.d. r.v.'s is to jackknife (see Miller[19] and Gray and Schucany[12] for further details). Unfortunately, the observations $X_1, X_2, \ldots, X_{N(t)}$ generated within the budget constraint t are not independent (and are not identically distributed as a function of t). In particular, the sum of the $N(t)$ τ_i's is constrained to be less than or equal to t, and hence the τ_i's (and consequently the X_i's) are dependent. Nevertheless, we will show that the dependency induced by the budget constraint is mild enough that the jackknife estimator continues to effectively reduce bias.

Let

$$\overline{X}_i(t) = \begin{cases} \dfrac{1}{N(t)-1} \displaystyle\sum_{\substack{j=1 \\ j \neq i}}^{N(t)} X_j; & N(t) \geq 2 \\ 0; & N(t) \leq 1, \end{cases}$$

$$\alpha_i(t) = \begin{cases} g(\overline{X}_i(t)); & N(t) \geq 2 \\ 0; & N(t) \leq 1 \end{cases}$$

$$\tilde{\alpha}_i(t) = N(t)\alpha(t) - (N(t) - 1)\alpha_i(t)$$

$$\alpha_J(t) = \begin{cases} \dfrac{1}{N(t)} \displaystyle\sum_{i=1}^{N(t)} \tilde{\alpha}_i(t); & N(t) \geq 1 \\ 0; & N(t) = 0. \end{cases}$$

Our next theorem asserts that the jackknife estimator $\alpha_J(t)$ does indeed effectively reduce the small-sample bias.

Theorem 3. *Suppose that* g *is polynomially dominated to degree* $r (r \geq 0)$ *and* $0 < \lambda < \infty$. *If there exists* $\delta > 0$ *such that* $E\exp(\delta\|X_1\|) < \infty$ *and* $E\exp(\delta\tau_1) < \infty$, *then*

$$E\alpha(t) = \alpha + o(1/t)$$

as $t \to \infty$.

In particular, Theorem 3 shows that the jackknifed estimator $\alpha_J(t)$ is an effective technique for reducing bias in regenerative steady-state simulations (see Example 3). Of course, the asymptotic nature of the result states only that the bias will be reduced for all sufficiently large computer budgets. For any given example, the bias may (in fact) be larger if the budget used is not big enough. This qualification is, however, typical of most of the low bias estimators that have been studied over the years. For example, this estimator provides a bias reduction that is qualitatively identical to that obtained when using the low-bias steady-state estimator proposed by Meketon and Heidelberger.[17] We note, however, that in contrast to the estimator proposed there, the budget constraint used here is specified in terms of computer time, whereas their estimator's budget constraint is determined by simulated time. Of course, a simulation time constraint is just a special case of a computer time constant (just set $\eta = \tau$ in Example 3). Thus, the estimator $\alpha_J(t)$ is a more generally applicable approach to obtaining bias reduction in steady-state regenerative simulations.

Our next goal is to show that the bias reduction obtained by using $\alpha_J(t)$ comes at no cost, in terms of additional asymptotic variability. In order to obtain a central limit theorem for $\alpha_J(t)$, we shall need to apply a "random time change" argument (to pass from the discrete time scale of observations, expressed in terms of n, to the continuous parameter computer time scale, expressed in terms of t). This basically requires proving a functional limit theorem for the jackknife estimator. Unfortunately, we have been unable to find such a limit theorem in the literature (although ordinary central limit theorems do appear; see, for example, Miller[18]). Thus, part (i) of our next theorem states a "strong approximation" result for the jackknife estimator $\alpha_J^*(t)$ given by

$$\alpha_J^*(n) = \frac{1}{n} \sum_{i=1}^{n} [ng(\overline{X}_n) - (n-1)g(\overline{X}_{in})],$$

where

$$\overline{X}_n = n^{-1} \sum_{j=1}^{n} X_j \quad \text{and} \quad \overline{X}_{in} = \sum_{\substack{j=1 \\ j \neq i}}^{n} X_j/(n-1).$$

A functional limit theorem for $\alpha_J^*(n)$ follows easily from the strong approximation result (see Csörgö and Révész[4]) and is stated as part (ii) of Theorem 4. The existence of this functional theorem also guarantees that sequential stopping rules, based on jackknife estimators, are asymptotically valid (see Glynn and Whitt[11]). Finally, part (iii) of Theorem 4 is the desired central limit theorem for $\alpha_J(t)$ and follows directly from part (ii) as a consequence of the "random time change" argument previously mentioned.

Theorem 4. *Suppose that $E\|X_1\|^p < \infty$ for some $p > 3$ and that $0 < \lambda < \infty$. Then,*

(1) *There exists a probability space $(\Omega, \mathcal{F}, P)$ supporting a sequence $\{\alpha_J^*(n)' : n \geq 1\}$ and an $\mathbb{R}^d$-valued Brownian motion B (with covariance matrix $C = (C_{ij}: 1 \leq i, j \leq d)$) such that:*

   (*a*) *$\{\alpha_J^*(n)' : n \geq 1\} \stackrel{\mathcal{D}}{=} \{\alpha_J^*(n): n \geq 1\}$, where $\stackrel{\mathcal{D}}{=}$ denotes equality in distribution*

   (*b*) *$\alpha_J^*(n)' = \alpha + \nabla g(\mu)B(n)/n + o(n^{1/p-1})$ a.s.*

(*ii*) *Let $\chi_n(t) = n^{1/2}(\alpha_J^*(nt) - \alpha)$ and $\chi(t) = \nabla g(\mu)B(t)/t$. Then, for $\epsilon > 0$,*

$$\chi_n \Rightarrow \chi$$

in the Skorohod space $D[\epsilon, \infty)$, where $D[\epsilon, \infty)$ is the space of right-continuous functions with left limits (see Ethier and Kurtz [6] for details on this space).

(*iii*) *$t^{1/2}(\alpha_J(t) - \alpha) \Rightarrow \sigma N(0, 1)$as $t \to \infty$, where σ^2 is defined as in Theorem 2.*

In the presence of appropriate uniform integrability conditions, we find that part (iii) of Theorem 4 asserts that $\text{var}\, \alpha_J(t) \sim \sigma^2/t$ as $t \to \infty$ (provided $\sigma^2 > 0$). Contrasting Theorem 4 with Theorem 2, we therefore conclude that $\alpha_J(t)$ possesses the same asymptotic variability as does $\alpha(t)$. Since $\alpha_J(t)$ has superior small-sample bias characteristics, as compared to $\alpha(t)$, this suggests that $\alpha_J(t)$ will often be a superior estimator to $\alpha(t)$ in small-sample settings. In the standard i.i.d. ratio estimation context, this type of behavior was observed empirically by Iglehart.[13]

2. Empirical Results

In this section, we discuss the results of simulation experiments that compare, empirically, the properties of various estimators including the jackknife. We will restrict our attention to the problem of steady-state estimation in two queueing systems: The M/M/1 queue and a simple queueing network model of a computer system. For the M/M/1 queue we consider estimation of the mean steady-state waiting time $E[W]$. We let λ denote the arrival rate, μ the service rate and define $\rho = \lambda/\mu$ to be the traffic intensity. In our experiments we set $\rho = 0.9$ which we obtain with $\lambda = 0.9$ and $\mu = 1.0$. With these parameters, $E[W] = (1/\mu)\rho/(1-\rho) = 9.0$. We select the empty and idle state to be our regeneration state: the expected number of customers served per regenerative cycle is $1/(1-\rho) = 10.0$.

The queueing network model is shown in Figure 1. There are five queues, one representing a CPU and four representing I/O disk drives. The service discipline at all queues is FCFS and all service times are exponentially distributed. Jobs arrive to the system at the CPU according to a Poisson process with rate λ. Upon leaving the CPU, the job goes to I/O device i with probability p_i for $i = 1, \ldots, 4$. A job is then routed back to the CPU with probability p, or exits the system with probability $(1-p)$. We let S_0 denote the mean service time of a job at the CPU, and let S_i denote the mean service time of a job at the I/O device i. Under these assumptions, the model is a Jackson network having a product form solution (see Jackson[14] or, e.g., Chapter 3 of Lavenberg[16]) so that steady-state performance measures are easy to compute. We will be interested in estimating the expected system response time $E[R]$ which is the total expected time that a job spends in the system. For our experiments we set $\lambda = 1$, $p_i = 0.25$, $p = 0.75$, $S_0 = 0.2$, and $S_i = 0.5$ for $i = 1, \ldots, 4$. With these parameter settings, $E[R] = 8.0$ and the steady-state utilization of the CPU is 0.8, while the steady-state utilization of each I/O device is 0.5. The probability of an empty system (our regeneration state) is $0.2 \times (0.5)^4 = 0.0125$, which since arrivals are Poisson is also the fraction of arrivals to an empty system. Therefore, the expected number of arrivals per regenerative cycle is $80(=0.0125)$. We use the IBM Research Queueing Package (RESQ) to simulate this network (see Sauer, Macnair and Kurose[20]).

Let t denote the length of the simulation (in some time unit). We will consider a variety of estimators, some having "high" bias (i.e., bias of order $1/t$), and others having "low" bias (i.e., bias of order $o(1/t)$). Let $T_n = \tau_1 + \ldots + \tau_n$ be the time of the nth regeneration and let $k(t)$ denote the number of waiting (response) times completed by time

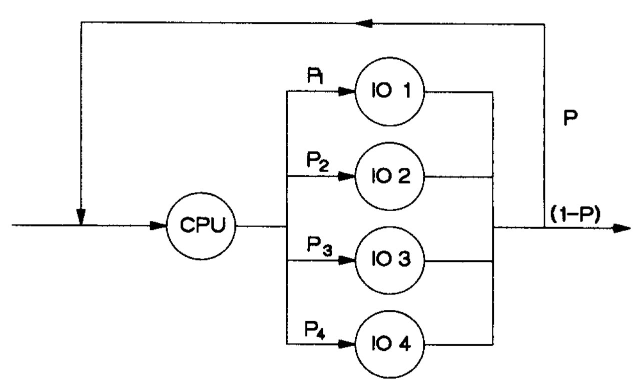

Figure 1. Queueing network model of a simple computer system.

t in the M/M/1 queue (queueing network, respectively). Let $\hat{\alpha}_{k(t)}$ denote the sample average of all waiting (response) times completed by time t. For example, in the M/M/1 queue

$$\hat{\alpha}_{k(t)} = \frac{\sum_{k=1}^{k(t)} W_k}{k(t)}.$$

Since we want to emphasize the dependence of the point estimates on the number of cycles completed, we will let α_n be the point estimate based on n regenerative cycles. With this notation, $\alpha_{N(t)} = \alpha(t)$, and $\alpha_{N(t)+1} = \alpha(T_{N(t)+1})$ is the point estimate based on $N(t) + 1$ cycles as suggested in Meketon and Heidelberger.[17] This estimator has bias of order $o(1/t)$ provided the steady-state performance measure can be expressed as $\alpha = E[X_k]/E[\eta_k]$ where $\eta_k = c \times \tau_k$ for some constant c, i.e., the length of the kth cycle in simulated time (η_k) is proportional to the length of the kth cycle in "CPU" time (τ_k). By "CPU" time we mean generally the time unit that determines the length of the simulation. Otherwise it has bias of order $1/t$ (see Glynn and Heidelberger[10] for a more complete explanation of this phenomenon). Letting $\tilde{N}(t) = \max(1, N(t))$, then $\alpha_{\tilde{N}(t)}$ is a point estimator which is described in Glynn and Heidelberger[10]: $\alpha_{\tilde{N}(t)}$ has the same leading term (of order $1/t$) in its bias expansion as does $\alpha_{N(t)}$. However, using $\tilde{N}(t)$ observations results in an unbiased estimate of a simple mean value, whereas using either $N(t)$ or $N(t) + 1$ observations generally results in a biased estimate of a simple mean. Glynn[9] has proposed an adjustment to $\hat{\alpha}_{k(t)}$ which can sometimes be used to reduce bias: we let $\tilde{\alpha}_k(t)$ denote this bias adjusted estimator. The estimator is based on a renewal-theoretic asymptotic expansion for the bias of the sample mean of a regenerative process. When it can be applied, the bias of $\tilde{\alpha}_{k(t)}$ is $o(1/t)$ as opposed to the order $1/t$ bias of $\hat{\alpha}_k(t)$.

In the M/M/1 queue simulations, we let t be the number of customers served. In this case $k(t) = t$ and $\eta_k = \tau_k$ so that $\alpha_{N(t)+1}$ is a low bias estimator. Also, $\tilde{\alpha}_k(t)$ is easy to implement in this case. Thus we have three low bias estimators ($\tilde{\alpha}_k(t)$, $\alpha_{N(t)+1}$, and $\alpha_J(t)$) and three high bias estimators ($\alpha_{N(t)}$, $\alpha_{\tilde{N}(t)}$ and $\hat{\alpha}_{k(t)}$) in this case. For the queueing network model, we let t be simulated time. In this case, η_k is the number of jobs processed by the network in the kth cycle and τ_k is the length of the kth cycle in simulated time. Since we are no longer assured that $\eta_k = c \times \tau_k$, $\alpha_{N(t)+1}$ has bias of order $1/t$. In addition, it is not clear that $\tilde{\alpha}_k(t)$ can be efficiently implemented. Thus in this case, we have only one low bias estimator ($\alpha_J(t)$) and four high bias estimators ($\alpha_{N(t)}$, $\alpha_{\tilde{N}(t)}$, $\alpha_{N(t)+1}$ and $\hat{\alpha}_{k(t)}$).

Tables I and II report the results of simulation experiments for the M/M/1 queue and the queueing network, respectively. For each t, R i.i.d. replications were performed. For each estimator, the sample average over the R replications was computed, along with an estimate of its standard deviation. These can be used to form confidence intervals for the expected value of an estimator. For example, from Table I, an approximate 90% confidence interval for $E[\alpha_{N(t)}]$ for $t = 500$ is $5.618 \pm 1.645 \times 0.015$. Since $E[W] = 9.00$, we conclude that $\alpha_{N(t)}$ is biased for this value of t. In Table I, the results are as expected: the low (theoretical) bias estimators are generally closer to the steady-state value than are the high bias estimators. The only exception is that $\alpha_J(t)$ has more bias then $\hat{\alpha}_{k(t)}$ for the shortest run ($t = 500$). This relates to the comment that follows Theorem 3, in which it was pointed out that we have established only that the jackknife estimator must reduce bias when the run-length chosen is sufficiently large. The jackknife also has higher variance than the other estimators for short runs. For intermediate to long runs, the low bias estimators actually do reduce bias without a significant increase in variance.

Table I. Means and Standard Deviations in Simulations of the M / M / 1 Queue with $\rho = 0.9$, $E[W] = 9.0$

	$t = 500$ $R = 40{,}000$	$t = 1{,}000$ $R = 20{,}000$	$t = 2{,}500$ $R = 20{,}000$	$t = 5{,}000$ $R = 20{,}000$	$t = 10{,}000$ $R = 10{,}000$
$\hat{\alpha}_{k(t)}$	7.319 0.022	8.049 0.030	8.614 0.023	8.796 0.017	8.891 0.018
$\alpha_{N(t)}$	5.618 0.015	7.061 0.024	8.238 0.022	8.615 0.017	8.800 0.017
$\alpha_{\tilde{N}(t)}$	5.646 0.016	7.063 0.024	8.238 0.022	8.615 0.017	8.800 0.017
$\alpha_{N(t)+1}$	8.113 0.025	8.584 0.032	8.893 0.024	8.956 0.018	8/971 0.018
$\alpha_J(t)$	6.796 0.025	8.123 0.037	8.848 0.029	8.943 0.020	8.970 0.019
$\tilde{\alpha}_k(t)$	8.211 0.032	8.611 0.038	8.883 0.026	8.945 0.018	8.971 0.018

Table II. Means and Standard Deviations in Simulations of the Queueing Network Model of a Computer System $E[R] = 8.0$

	$t = 500$ $R = 4{,}000$	$t = 1{,}000$ $R = 2{,}000$	$t = 2{,}000$ $R = 1{,}000$	$t = 4{,}000$ $R = 500$
$\hat{\alpha}_{k(t)}$	7.620	7.819	7.943	7.943
	0.025	0.028	0.028	0.029
$\alpha_{N(t)}$	6.926	7.628	7.885	7.917
	0.031	0.030	0.029	0.030
$\alpha_{\tilde{N}(t)}$	7.090	7.639	7.885	7.917
	0.028	0.029	0.029	0.030
$\alpha_{N(t)+1}$	7.870	7.969	8.004	7.977
	0.024	0.027	0.028	0.029
$\alpha_J(t)$	7.860	8.054	8.049	7.991
	0.046	0.041	0.033	0.031

Similar observations can be obtained from Table II for the queueing network model. Although only the jackknife has provably low bias in this example, $\alpha_{N(t)+1}$ also performs well.

3. Proofs

Proof of Theorem 1. Fix $\epsilon > 0$ such that g has bounded second partial derivatives over the set $\{x: \|x - \mu\| \leqslant \epsilon\}$. Let $I(t) = I(\|\bar{X}(t) - \mu\| > \epsilon)$ and put $I_c(t) = 1 - I(t)$. Since $0 < \lambda < \infty$, it follows that $N(t) \to \infty$ a.s. Furthermore, $n^{-1}\sum_{i=1}^n X_i \to \mu$ a.s. by the strong law of large numbers. Hence, it is evident that $\bar{X}(t) \to \mu$ a.s. as $t \to \infty$, so that $I(t) = 0$ for t sufficiently large a.s. We will now show that $E\alpha(t)I(t) = o(1/t)$.

We first note that since g is polynomially dominated to degree r and $\bar{X}(t) \in \mathscr{C}$, we have that

$$E|\alpha(t)|I(t) \leqslant AEI(t) + BE\|\bar{X}(t) - \mu\|^r I(t). \quad (3.1)$$

Now, on $\{I(t) = 1\}$, $\|\bar{X}(t) - \mu\|/\epsilon \geqslant 1$ so

$$EI(t) > \frac{1}{t^{p/2}\epsilon^p} E[t^{p/2}\|\bar{X}(t) - \mu\|^p I(t)]. \quad (3.2)$$

Also, a similar argument shows that

$$E\|\bar{X}(t) - \mu\|^r I(t) \leqslant \frac{1}{t^{p/2}\epsilon^{p-r}} E[t^{p/2}\|\bar{X}(t) - \mu\|^p I(t)]. \quad (3.3)$$

By Theorem 4.7 of Glynn and Heidelberger,[10] $\{t^{p/2}\|\bar{X}(t) - \mu\|^p: t > t_0\}$ is uniformly integrable for some finite t_0 (the argument given there easily extends to noninteger values of p). Since $I(t) \to 0$ a.s., it follows that

$$E[t^{p/2}\|\bar{X}(t) - \mu\|^p I(t)] = o(1) \quad (3.4)$$

as $t \to \infty$. Combining (3.1)–(3.4), we conclude that $E|\alpha(t)|I(t) = o(t^{-p/2})$ as $t \to \infty$.

We now turn to $E\alpha(t)I_c(t)$. However, for that term, we can apply exactly the same argument as that used to prove Theorem 4.7 of Glynn and Heidelberger.[10] It shows that $E\alpha(t)I_c(t) = \alpha + b/t + o(1/t)$ as $t \to \infty$. This clearly completes the proof of our result.

Proof of Theorem 3. For purposes of simplifying our notation, we confine our presentation to the case where $d = 1$. The general case can be attacked in precisely the same way as the scalar case.

We start by showing that $\bar{X}_{in}$ is uniformly close to $\bar{X}_n$, $1 \leqslant i \leqslant n$ (this is a standard technique in the jackknifing literature; however, the standard references only give uniformity in probability, rather than with probability one). We first note that for $n \geqslant 1$,

$$\begin{aligned}(n-1)(\bar{X}_n - \bar{X}_{in}) &= n\bar{X}_n - (n-1)\bar{X}_{in} - \bar{X}_n \\ &= \sum_{j=1}^n X_j - \sum_{\substack{j=1 \\ j\neq i}}^n X_j - \bar{X}_n \\ &= X_i - \bar{X}_n. \end{aligned} \quad (3.5)$$

Hence,

$$\begin{aligned}\max_{1\leqslant i\leqslant n} |\bar{X}_n - \bar{X}_{in}| &\leqslant \frac{1}{n-1}\max_{1\leqslant i\leqslant n}|X_i - \bar{X}_n| \\ &\leqslant \frac{1}{n-1}\max_{1\leqslant i\leqslant n}|X_i| + \frac{1}{n-1}|\bar{X}_n|.\end{aligned}$$

If $E|X|^p < \infty$, it follows that $n^{-1}\sum_{i=1}^n |X_i|^p \to E|X|^p < \infty$ a.s. as $n \to \infty$. Consequently $|X_n|^p/n \to 0$ a.s., from which it is evident that $\max_{1\leqslant i\leqslant n}|X_i|^p/n \to 0$ a.s. i.e., $\max_{1\leqslant i\leqslant n}|X_i| = o(n^{1/p})$ a.s. Thus, we conlude that if $E|X|^p < \infty$,

$$\max_{1\leqslant i\leqslant n}|\bar{X}_n - \bar{X}_{in}| = o(n^{1/p-1}) \quad \text{a.s.} \quad (3.7)$$

We now proceed to prove our theorem. Fix $\epsilon > 0$ so that g is twice continuously differentiable on $\{x : |x - \mu| \leqslant 3\epsilon\}$. Since $N(t) \to \infty$ a.s., (3.7) implies that $\tilde{I}_c(t) \to 1$ a.s., where $\tilde{I}_c(t) = I(|\bar{X}(t) - \mu| \leqslant \epsilon,\ |\bar{X}(t) - \bar{X}_i(t)| \leqslant \epsilon,\ 1 \leqslant i \leqslant N(t),\ N(t) \geqslant 2)$. We now write $E\alpha_J(t)$ as $E\alpha_J(t) = E\alpha_J(t)\tilde{I}_c(t) + E\alpha_J(t)\tilde{I}(t)$, where $\tilde{I}(t) = 1 - \tilde{I}_c(t)$, and analyze each piece separately.

If $|\bar{X}_n - \mu| \leqslant \epsilon$, $|\bar{X}_{in} - \bar{X}_n| \leqslant \epsilon$, we can expand $g(\bar{X}_{in})$ in a Taylor series about $\bar{X}_n$, yielding

$$\begin{aligned}&ng(\bar{X}_n) - (n-1)g(\bar{X}_{in}) \\ &\quad = g(\bar{X}_n) - (n-1)[g(\bar{X}_{in}) - g(\bar{X}_n)] \\ &\quad = g(\bar{X}_n) - (n-1)\Big[g'(\bar{X}_n)(\bar{X}_{in} - \bar{X}_n) \\ &\qquad + \frac{g''(\xi_{in})}{2}(\bar{X}_{in} - \bar{X}_n)^2\Big],\end{aligned}$$

where ξ_{in} lies between $\bar{X}_n$ and $\bar{X}_{in}$. We further note that the above expansions are uniform in i (i.e., $\max\{|\xi_{in} - \bar{X}_n| : 1 \leqslant i \leqslant n\} \to 0$ a.s. as $n \to \infty$) by (3.7). Hence, for n sufficiently large, (3.5) implies that we have

$$\alpha_J^*(n) = h(\bar{X}_n) + g'(\bar{X}_n)\cdot\frac{1}{n}\sum_{i=1}^{n}(X_i - \bar{X}_n) - \frac{1}{n(n-1)}\sum_{i=1}^{n} g''(\xi_{in})(X_i - \bar{X}_n)^2/2 = g(\bar{X}_n) - \frac{1}{n(n-1)}\sum_{i=1}^{n} g''(\xi_{in})(X_i - \bar{X}_n)^2/2. \quad (3.8)$$

Since $\alpha_J(t) = \alpha_J^*(N(t))$, we obtain

$$E\alpha_J(t)\tilde{I}_c(t) = E\alpha(t) - Eg(\bar{X}(t))\tilde{I}(t) - E\left[\frac{1}{N(t)(N(t)-1)}\cdot\sum_{i=1}^{N(t)} g''(\xi_{iN(t)})(X_i - \bar{X}(t))^2\tilde{I}_c(t)/2\right] \quad (3.9)$$

Theorem 1 implies that $E\alpha(t) = \alpha + b/t + o(1/t)$. We will show later that $EN(t)|g(\bar{X}(t))|\tilde{I}(t) = o(1/t)$. Since $g(\bar{X}(t)) = 0$ on $\{N(t) = 0\}$, it is clear that $|g(\bar{X}(t))| \leqslant N(t)|g(\bar{X}(t))|$. Thus, the estimate on $EN(t)|g(\bar{X}(t))|\tilde{I}(t)$ will show that $Eg(\bar{X}(t))\tilde{I}(t) = o(1/t)$. As for the third term on the right-hand side of (3.9), note that

$$\left(\frac{t}{N(t)}\right)\frac{1}{N(t)-1}\sum_{i=1}^{N(t)} g''(\xi_{iN(t)})(X_i - \bar{X}(t))^2\tilde{I}_c(t)/2 \to \lambda^{-1}g''(\mu)E(X_1-\mu)^2/2 = b \quad \text{a.s.} \quad (3.10)$$

as $t \to \infty$. Thus, we will be able to conclude that $E\alpha_J(t)\tilde{I}_c(t) = \alpha + o(1/t)$, once we show that the third turn, when multiplied by t, is uniformly integrable. Let $M = \max\{|g''(x)| : |x-\mu| \leqslant 3\epsilon\}$. Then, the left-hand side of (3.10) is dominated by

$$M\left(\frac{t}{N(t)}\right)\frac{1}{N(t)}\sum_{i=1}^{N(t)}(X_i - \bar{X}(t))^2 I(N(t) \geqslant 1). \quad (3.11)$$

This process clearly converges a.s. to $M\lambda^{-1}E(X_1-\mu)^2$. We claim that it is appropriately uniformly integrable. Using the conditional exchangeability of $X_1,\ldots,X_{N(t)}$ (see Section 2 of Glynn and Heidelberger[10]), we find that the expectation of (3.11) is

$$MtE\left[N(t)^{-1}(X_1 - \bar{X}(t))^2 I(N(t)\geqslant 1)\right]. \quad (3.12)$$

To show the uniform integrability of (3.11), it suffices to show that (3.12) converges to $M\lambda^{-1}E(X_1-\mu)^2$. But

$$E\left[(t/N(t))^2(X_1 - \bar{X}(t))^4 I(N(t)\geqslant 1)\right] \leqslant E^{1/2}[(t/N(t))^4 I(N(t)\geqslant 1)] - E^{1/2}\left[(X_1 - \bar{X}(t))^4\right] \leqslant E^{1/2}[(t/N(t))^4] \cdot\left[2^8E(X_1-\mu)^8 + 2^8E(\bar{X}(t)-\mu)^8\right]^{1/2},$$

which is bounded in t by Corollary 11 and Theorem 6 of Glynn and Heidelberg.[10] Since the process $(t/N(t))(X_1 - \bar{X}(t))^2 I(N(t)\geqslant 1)$ appearing in (3.12) has uniformly bounded second moment, this implies that (3.12) does indeed converge to $M\lambda^{-1}E(X_1-\mu)^2$. Hence, (3.11) (and thus (3.10)) are also uniformly integrable. Consequently, $E\alpha_J(t)\tilde{I}_c(t) = \alpha + o(1/t)$.

To complete the proof, we need to show that $E\alpha_J(t)\tilde{I}(t) = o(1/t)$. Now, the conditional exchangeability of $X_1,\ldots,X_{N(t)}$ implies that each of the terms $\tilde{\alpha}_i(t)(1 \leqslant i \leqslant N(t))$ have identical distributions when conditioned on $N(t)$. Because $\tilde{I}(t)$ is symmetric in $X_1,\ldots,X_{N(t)}$, we conclude that

$$E\alpha_J(t)\tilde{I}(t) = E\tilde{\alpha}_1(t)\tilde{I}(t).$$

Since g is polynomially dominated to degree r, we find that

$$|\tilde{\alpha}_1(t)|\tilde{I}(t) \leqslant N(t)|g(\bar{X}(t))|\tilde{I}(t) + N(t)|g(\bar{X}_1(t))|\tilde{I}(t) \leqslant N(t)[2A\tilde{I}(t) + B(|\bar{X}(t)-\mu|^r + |\bar{X}_1(t)-\mu|^r]. \quad (3.13)$$

Also, we have that for $q > 0$,

$$\tilde{I}(t) \leqslant I(|\bar{X}(t)-\mu| > \epsilon) + \sum_{i=1}^{N(t)} I(|\bar{X}(t) - \bar{X}_i(t)| > \epsilon, N(t)\geqslant 2) + I(N(t)\leqslant 1) \leqslant |\bar{X}(t)-\mu|^q/\epsilon^q + \sum_{i=1}^{N(t)}|\bar{X}(t)-\bar{X}_i(t)|^q/\epsilon^q I(N(t)\geqslant 2) + I(\tau_1+\tau_2 > t) \leqslant |\bar{X}(t)-\mu|^q/\epsilon^q + 2^q\sum_{i=1}^{N(t)}|X_i - \bar{X}(t)|^q/N(t)^q\epsilon^q I(N(t)\geqslant 1) + I(\tau_1+\tau_2 > t), \quad (3.14)$$

where we used (3.5) and the fact that $(N(t)-1)^{-1} \leqslant 2N(t)^{-1}$ on $\{N(t)\geqslant 2\}$ to obtain the last inequality. Similarly, we find that

$$|\bar{X}_1(t)-\mu|^r \leqslant 2^r|\bar{X}_1(t)-\bar{X}(t)|^r + 2^r|\bar{X}(t)-\mu|^r \leqslant 2^{2r}N(t)^{-r}|X_1 - \bar{X}(t)|^r + 2^r|\bar{X}(t)-\mu|^r. \quad (3.15)$$

Now, the conditional exchangeability of $X_1,\ldots,X_{N(t)}$ guarantees (see Glynn and Heidelberger[10]) that

$$E\sum_{i=1}^{N(t)}|X_i - \bar{X}(t)|^q N(t)^{1-q} I(N(t)\geqslant 1) = E\left[N(t)^{2-q}|X_1 - \bar{X}(t)|^q;\ N(t)\geqslant 1\right]. \quad (3.16)$$

Note that if g is polynomially dominated to degree r, then it is also polynomially dominated to degree r' for any $r' \geqslant r$. Hence, we have the freedom to choose r and q as large as we wish in (3.13)–(3.16). Combining (3.13) to (3.16), we therefore find that in order to conclude that $E|\tilde{\alpha}_1(t)|\tilde{I}(t) = o(1/t)$, it suffices to show that

$$E|\overline{X}(t) - \mu|^p N(t) = o(1/t) \qquad (3.17)$$

$$E|X_1 - \overline{X}(t)|^p N(t)^{2-p} = o(1/t) \qquad (3.18)$$

$$P\{\tau_1 + \tau_2 > t\} = o(1/t) \qquad (3.19)$$

for p sufficiently large. To obtain (3.17), we apply the Cauchy-Schwarz inequality, Theorem 6 of Glynn and Heidelberger,[10] and Theorem 2.3 of Janson [15] (to conclude that $E^{1/2}N^2(t) = O(t)$). Relation (3.18) is handled similarly to (3.12) (choose $p \geqslant 4$). Finally, to obtain (3.19), note that Markov's inequality and the independence of τ_1 and τ_2 prove that

$$P\{\tau_1 + \tau_2 > t\} \leqslant e^{-\delta t} E \exp(\delta(\tau_1 + \tau_2)) = e^{-\delta t}(E \exp(\delta \tau_1))^2.$$

Hence, (3.17)–(3.19) are proved, showing that $E\tilde{\alpha}_1(t)\tilde{I}(t) = o(1/t)$ and completing the proof of the theorem.

Proof of Theorem 4. As in the proof of Theorem 3, we specialize to $d = 1$ in order to simplify the notation; the general case follows precisely the same form as the scalar proof.

Our starting point is (3.8). Noting that $n^{-1}\Sigma_{i=1}^n g''(\xi_{in})(X_i - \overline{X}_n)^2 \to g''(\mu)E(X_1 - \mu)^2$ a.s. as $n \to \infty$, we conclude that

$$\alpha_J^*(n) = g(\overline{X}_n) + O(n^{-1}) \quad \text{a.s.} \qquad (3.20)$$

as $n \to \infty$. We now expand $g(\overline{X}_n)$ in a Taylor series about μ:

$$g(\overline{X}_n) = \alpha + g'(\mu)(\overline{X}_n - \mu) + \frac{g''(\xi_n)}{2}(\overline{X}_n - \mu)^2, \qquad (3.21)$$

where ξ_n lies between $\overline{X}_n$ and μ. Now, by the law of the iterated logarithm, $\overline{X}_n - \mu = O(\sqrt{\log\log n/n})$ a.s. Combining (3.20) and (3.21), we then get

$$\alpha_J^*(n) = \alpha + g'(\mu)(\overline{X}_n - \mu) + o(n^{1/p-1}) \quad \text{a.s.} \quad (3.22)$$

as $n \to \infty$. We now appeal to Theorem 1 of Einmahl[5] to guarantee existence of a probability space $(\Omega, \mathcal{F}, P)$ supporting r.v.'s $\{X_n': n \geqslant 1\} \stackrel{\mathcal{D}}{=} \{X_n: n \geqslant 1\}$ such that $X_1' + \ldots + X_n' - n\mu = B(n) + o(n^{1/p})$ a.s. (In what follows, we adopt the usual tradition of assuming $(\Omega, \mathcal{F}, P)$ as our original probability space; this can be done without any essential loss of generality.) Hence

$$\overline{X}_n - \mu = \frac{B(n)}{n} + o(n^{1/p-1}) \quad \text{a.s.} \qquad (3.23)$$

Combining (3.22) and (3.23) yields part (i) of the theorem.

Obtaining part (ii) is now easy. Note that

$$\chi_n(t) = g'(\mu)B(nt)/n^{1/2}t + o(n^{1/p-1/2}t^{1/p-1}) \quad \text{a.s.}$$

but $B(n\cdot)/n^{1/2} \stackrel{\mathcal{D}}{=} B(\cdot)$ and the term $o(n^{1/p-1/2}t^{1/p-1})$ goes to zero uniformly in $t \geqslant \epsilon$, proving the result. Part (iii) is now an immediate consequence of a standard "random time change" argument (see Billingsley,[11] p. 44).

Acknowledgments

The first author's contribution to this paper was written while he was visiting the Department of Mathematics and Statistics at Carleton University, Ottawa, Canada. He wishes to thank the Department for its hospitality. Research was supported by the IBM Corporation under SUR-SST Contract 12480042, the U.S. Army Research Office under Contract DAAL-03-88-K-0063, and by a Natural Sciences and Engineering Research Council of Canada Grant.

References

1. P. Billingsley, 1968. *Convergence of Probability Measures*, John Wiley, New York.
2. K.L. Chung, 1974. *A Course in Probability Theory*, Academic Press, New York.
3. M.A. Crane and A.J. LeMoine, 1977. *An Introduction to the Regenerative Method for Simulation Analysis* (Lecture Notes in Control and Information Sciences), Springer-Verlag, New York.
4. M. Csörgö and P. Révéz, 1981. *Strong Approximations in Probability and Statistics*, Academic Press, New York.
5. U. Einmahl, 1989. Extensions of results of Komlós, Major, and Tusnáday to the multivariate case, *Journal of Multivariate Analysis 28*, 20–68.
6. S.N. Ethier and T.G. Kurtz, 1986. *Markov Processes: Characterization and Convergence*, John Wiley, New York.
7. B.L. Fox and P.W. Glynn, 1989. Simulating Discounted Costs, *Management Science 35*, 1297–1315.
8. P.W. Glynn and D.L. Iglehart, 1986. Estimation of Steady-State Central Moments by the Regenerative Method of Simulation, *Operations Research Letters 5*, 271–276.
9. P.W. Glynn, 1989. A Low Bias Steady-State Estimator for Equilibrium Processes, Technical Report, Department of Operations Research, Stanford University, Stanford, CA.
10. P.W. Glynn and P. Heidelberger, 1990. Bias Properties of Budget Constrained Monte Carlo Simulations, *Operations Research 38*, 801–815.
11. P.W. Glynn and W. Whitt, 1992. The Asymptotic Validity of Sequential Stopping Rules for Stochastic Simulations, *Annals of Applied Probability 2*, 180–198.
12. H.L. Gray and W.R. Schucany, 1972. *The Generalized Jackknife Statistic*, Marcel Dekker, New York.
13. D.L. Iglehart, 1975. Simulating Stable Stochastic Systems, V: Comparison of Ration Estimators, *Naval Research Logistics Quarterly 22*, 553–565.
14. J.R. Jackson, 1957. Networks of Waiting Lines, *Operations Research 5*, 518–521.

15. S. JANSON, 1983. Renewal Theory for m-Dependent Random Variables, *Annals of Probability 11*, 558–568.
16. S.S. LAVENBERG, (ed.), 1983. *Computer Performance Modeling Handbook*, Academic Press, New York.
17. M.S. MEKETON and P. HEIDELBERGER, 1982. A Renewal Theoretic Approach to Bias Reduction in Regenerative Simulation, *Management Science 26*, 173–181.
18. R.G. MILLER, 1964. A Trustworthy Jackknife, *Annals of Mathematic Statistics 35*, 1594–1605.
19. R.G. MILLER, 1974. The Jackknife—A Review, *Biometrika 61*, 1–15.
20. C.H. SAUER, E.A. MACNAIR and J.F. KUROSE, 1984. Queueing Network Simulations of Computer Communications, *IEEE Journal on Selected Areas in Communications SAC-2*, 203–219.
21. W.L. SMITH, 1955. Regenerative Stochastic Processes, *Proceedings of the Royal Society Series A 232*, 6–31.