行政院國家科學委員會專題研究計畫成果報告

以回歸方法論求解起始偏誤問題

計畫類別：☑ 個別型計畫　　□ 整合型計畫

計畫編號： 96-2221-E-004-005-

執行期間： 中華民國96年8月1日至97年7月31日

計畫主持人：謝明華

執行單位：國立政治大學資訊管理學系

中華民國97年10月31日

以回歸方法論求解起始偏誤問題

計畫編號： 96-2221-E-004-005-
執行期間： 中華民國96年8月1日至97年7月31日
主持人： 謝明華
E-mail: mhsieh@nccu.edu.tw
執行單位： 國立政治大學資訊管理學系

一、中文摘要

我們提出數種以迴歸 (regression) 爲基礎的演算法。這些演算法可針對離散事件 (discrete-event) 模擬或馬可夫鏈蒙地卡羅 (Markov Chain Monte Carlo) 模擬中的均態 (steady-state) 參數提供信賴區間估計。在本計畫中, 我們將探討這些演算法的理論特性, 並將利用數個文獻中典型的實例來測試它們的實際效能 (empirical performance)。這些演算法可以在單一或是多個 CPU 的電腦上運算。

關鍵詞：迴歸, 離散事件模擬, 均態模擬, 馬可夫鏈蒙地卡羅

Abstract

We propose several regression-based procedures for building confidence interval estimators of steady-state parameters in discrete event simulations and Markov Chain Monte Carlo (MCMC) simulation. In this project, we will explore the theoretical properties of these procedures and test their empirical performance on various classical problems in discrete event and MCMC simulations. These procedures can be implemented on parallel processors or single processor architecture.

Keywords: Regression, discrete event simulation, steady-state simulation, Markov Chain Monte Carlo

二、研究目的與文獻探討

Let $Y = (Y(t) : t \geq 0)$ be a real-valued stochastic process representing the output of a simulation. Suppose that there exists a (deterministic) constant α for which

$$\bar{Y}(t) \equiv \frac{1}{t}\int_0^t Y(s)ds \Rightarrow \alpha \tag{1}$$

as $t \to \infty$, where $\Rightarrow$ denotes weak convergence. The steady-state simulation problem is concerned with computing the steady-state quantity α. (If the process of interest is a discrete-time sequence $(Y_n : n \geq 0)$, we can embed the process in continuous time by setting $Y(t) = Y_{\lfloor t \rfloor}$ for $t \geq 0$, where $\lfloor t \rfloor$ is the so-called "floor" of t.)

One of the principal complications associated with the steady-state simulation problem is that of the "initial transient". In particular, the process Y is typically initialized with an initial condition that is atypical of steady-state behavior, there by inducing an initial transient period during which the observations collected are biased relative to the steady-state mean α.

In great generality, it is known the bias of the time average estimator $\bar{Y}(t)$ can be expressed as

$$E\bar{Y}(t) = \alpha + \frac{b_1}{t} + \frac{b_2}{t^2} + \cdots + \frac{b_k}{t^k} + o(t^{-k}) \tag{2}$$

as $t \to \infty$, where $0 = b_2 = b_3 = \cdots$ and $o(f(t))$ represents a function having the property that $o(f(t))/f(t) \to 0$ as $t \to \infty$, see of Section 2. This supports the use of the approximation

$$E\bar{Y}(t) \approx \alpha + \frac{b}{t} \tag{3}$$

for large t, where $b = b_1$. Given the approximation (3), this suggests the possibility of developing an estimator for the steady-state quantity α that improves upon $\bar{Y}(t)$, based on regression methods in which α and b are treated as unknown regression parameters. This paper develops the theory and reports empirical findings based on this regression approach to reducing initial transient bias.

三、研究方法

My coauthor is Professor Peter Glynn. Professor Glynn is an expert in steady-state simulation. He had published a great number of influential papers about steady-state simulations. Therefore, I am looking forward to the cooperation opportunity. Below are examples of Professor Glynn's previous works in steady-state simulations.

1. Glynn and Iglehart (1988)

2. Glynn and Whitt (1991)
3. Glynn and Whitt (1994)
4. Glynn (1995)
5. Glynn and Heidelberger (1990)
6. Glynn and Heidelberger (1991a)
7. Glynn and Heidelberger (1991b)
8. Glynn and Heidelberger (1992a)
9. Glynn and Heidelberger (1992c)
10. Glynn and Heidelberger (1992b)

1 Related works

The steady-state simulation problem is a very important topic in simulation. It has been studied for a long time and is continue to be a hot research topic; see Law and Kelton (2000), Bratley et al. (1983), Banks et al. (2000), Meketon and Heidelberger (1982) Cash et al. (1992), Glynn (1995), Goldsman et al. (1994), Schruben (1982), Schruben et al. (1983), K. Preston White (1997), and Jr. et al. (2000), Hsieh et al. (2004), and Steiger et al. (2005).

We start with a justification of (2). It is typically the case that when (1) holds, there exists a random variable (rv) $Y(\infty)$ for which

$$Y(t) \Rightarrow Y(\infty) \tag{4}$$

as $t \to \infty$, where $\alpha = EY(\infty)$. (For example, (4) holds when Y is a positive recurrent regenerative process having a spread-out cycle-length distribution; see **?**, p. 351). In the presence of uniform integrability,

$$a(t) \stackrel{\triangle}{=} EY(t) - \alpha \to 0$$

as $t \to \infty$. The representation (2) then hinges on the question of how fast $a(t)$ converges to zero as $t \to \infty$.

Proposition 1.1 *If $a(t) = o(t^{-p})$ as $t \to \infty$ for $p > 1$ and $a(\cdot)$ is a bounded function, then*

$$E\bar{Y}(t) = \alpha + \frac{b_1}{t} + \frac{b_2}{t^2} + \cdots + \frac{b_k}{t^k} + o(t^{-k}) \tag{5}$$

as $t \to \infty$, where $k = \lfloor p \rfloor$ and $0 = b_2 = b_3 = \cdots$. If $a(t) = o(\exp(-rt))$ as $t \to \infty$ for $r > 0$ and $a(\cdot)$ is a bounded function, then $E\bar{Y}(t) = \alpha + o(\exp(-rt))$ as $t \to \infty$.

One sufficient condition guaranteeing the first set of hypotheses of Proposition 1.1 is that Y be a bounded regenerative process having a spread-out cycle length distribution with the $(p+1)$'th cycle length moment being finite. The second set of hypotheses follows for some $r > 0$ whenever Y is a bounded regenerative process having a spread-out cycle length distribution with some finite exponential cycle length moment. Such exponential convergence rates also hold whenever Y is a suitable real-valued functional of a geometrically ergodic Markov chain.

Several of our regression-based estimators will require knowledge of the covariance structure of the process Y. For $\varepsilon > 0$, set

$$\bar{Z}_\varepsilon(t) = \varepsilon \int_0^{t/\varepsilon} (Y(u) - \alpha) du.$$

Under modest mixing conditions on the process Y,

$$\varepsilon^{-1/2} \bar{Z}_\varepsilon \xrightarrow{fdd} \sigma B \tag{6}$$

as $\varepsilon \downarrow 0$, where $B = (B(t) : t \geq 0)$ is a standard Brownian motion (BM), and $\xrightarrow{fdd}$ denotes "weak convergence of the finite-dimensional distributions" (i.e. for each $n \geq 1$ and selection of time points $t_1 < \ldots < t_m$, $\varepsilon^{-1/2}(\bar{Z}_\varepsilon(t_1), \ldots, \bar{Z}_\varepsilon(t_m)) \Rightarrow (\sigma B(t_1), \ldots, \sigma B(t_m))$ as $\varepsilon \downarrow 0$). The variance parameter σ^2 appearing in (6) is often called the time-average variance constant (TAVC) of the process Y. Such convergence of the finite-dimensional distributions is implied by weak convergence in the function space $C[0, \infty)$, for which many sufficient conditions are known (ϕ-mixing, strong mixing, positive association, and strong invariance principle.

In any case, if $(\bar{Z}_\varepsilon(t)^2 : \varepsilon \geq 1)$ is uniformly integrable, then (6) implies that

$$\begin{aligned} \mathrm{Cov}\left(\frac{\bar{Z}_\varepsilon(s)}{s}, \frac{\bar{Z}_\varepsilon(t)}{t}\right) &= \mathrm{Cov}(\bar{Y}(s/\varepsilon), \bar{Y}(t/\varepsilon)) \\ &\sim \varepsilon \mathrm{Cov}\left(\frac{\sigma B(s)}{s}, \frac{\sigma B(t)}{t}\right) \\ &= \varepsilon \sigma^2 \min(1/s, 1/t) \end{aligned} \tag{7}$$

as $\varepsilon \downarrow 0$, where $\sim$ means that the ratio of the left-hand side to the right-hand side converges to 1 as $\varepsilon \downarrow 0$. The above description of the asymptotic covariance structure of $\bar{Y}(\cdot)$ will play a key role in our theoretical analyses.

2 Regression via Ordinary Least Squares

We start with the most obvious regression-based means of exploiting (3). In particular, given a simulation of Y up to time t, consider estimating α

and b by minimizing the sum of squares

$$\sum_{i=1}^{m} (\bar{Y}(t_i) - \alpha - \frac{b}{t_i})^2, \tag{8}$$

where $0 < t_1 < \cdots < t_m = t$. The minimizer $\hat{\alpha}_1 = \hat{\alpha}_1(t)$ of (8) is given by

$$\hat{\alpha}_1(t) = \frac{(\sum_{i=1}^{m} \frac{1}{t_i^2})(\sum_{j=1}^{m} \bar{Y}(t_j)) - (\sum_{i=1}^{m} \frac{\bar{Y}(t_i)}{t_i})(\sum_{j=1}^{m} \frac{1}{t_j})}{m \sum_{i=1}^{m} \frac{1}{t_i^2} - (\sum_{j=1}^{m} \frac{1}{t_j})^2}$$

Put $s_i = t_i/t$ and $s_0 = 0$. For $m = 2$, $\hat{\alpha}_1(t)$ simplifies to

$$\hat{\alpha}_1(t) = \frac{\bar{Y}(t) - s_1 \bar{Y}(s_1 t)}{1 - s_1} = \frac{1}{t - t_1} \int_{t_1}^{t} Y(u) du,$$

so that $\hat{\alpha}_1(t)$ corresponds then to the time-average obtained when the first interval $[0, t_1]$ of observations is deleted from the sample. Furthermore, in the presence of (7),

$$\frac{E(\hat{\alpha}_1(t) - \alpha)^2}{E(\bar{Y}(t) - \alpha)^2} \to \frac{1}{1 - s_1}$$

as $t \to \infty$, establishing that $\hat{\alpha}_1(t)$ is inferior (in terms of mean square error (MSE)) to the time-average $\bar{Y}(t)$ when t is large.

For $m = 3$ and $s_1 = 1/3$, $s_2 = 1/3$, and $s_3 = 1$,

$$\begin{aligned} \hat{\alpha}_1(t) &= \bar{Y}(t/3) \left(\frac{-17}{26} \right) + \bar{Y}(2t/3) \left(\frac{8}{13} \right) + \bar{Y}(t) \left(\frac{27}{26} \right) \\ &= \frac{\int_{t/3}^{2t/3} Y(s) ds}{(t/3)} \left(\frac{17}{26} \right) + \frac{\int_{2t/3}^{t} Y(s) ds}{(t/3)} \left(\frac{9}{26} \right). \end{aligned}$$

Here, $\hat{\alpha}_1(t)$ deletes the initial interval $[0, t/3]$ but weights observations from two subsequent intervals $[t/3, 2t/3]$ and $[2t/3, t]$ unequally. Also

$$\frac{E(\hat{\alpha}_1(t) - \alpha)^2}{E(\bar{Y}(t) - \alpha)^2} \to \frac{555}{313},$$

establishing again that the original time-average estimator $\bar{Y}(t)$ is preferable when t is large.

For general values of m, the following result provides an exact expression for the limiting ratio of the MSE of $\hat{\alpha}_1(t)$ to that of $\bar{Y}(t)$.

Proposition 2.1 *Assume (5) and (8). Then,*

$$E\hat{\alpha}_1(t) = \alpha + o(t^{-k}) \tag{9}$$

and

$$\frac{E(\hat{\alpha}_1(t)-\alpha)^2}{E(\bar{Y}(t)-\alpha)^2} \to \sum_{i=1}^{m}(s_i - s_{i-1})\beta_i^2 \tag{10}$$

as $t \to \infty$, where

$$\beta_i = \frac{(\sum_{j=i}^{m}\frac{1}{s_j})(\sum_{k=1}^{m}\frac{1}{s_k^2}) - (\sum_{j=i}^{m}\frac{1}{s_j^2})(\sum_{k=1}^{m}\frac{1}{s_k})}{m(\sum_{k=1}^{m}\frac{1}{s_k^2}) - (\sum_{k=1}^{m}\frac{1}{s_k})^2}$$

Our main result in this section establishes that the ordinary least squares regression estimator always exhibits a larger asymptotic MSE than that of $\bar{Y}(t)$ (except in the trivial case in which $m = 1$ and $\hat{\alpha}_1(t) = \bar{Y}(t)$, in which case the MSE coincides).

Theorem 2.2 *If $m > 1$,*

$$\sum_{i=1}^{m}(s_i - s_{i-1})\beta_i^2 > 1.$$

3 Regression via Generalized Least Squares

In the presence of (6), the random vector $(\bar{Y}(t_1), \ldots, \bar{Y}(t_m))$ has a distribution that is approximately Gaussian when the t_i's are large. This suggests that α and b should be estimated by maximizing the (approximate) Gaussian likelihood of $(\bar{Y}(t_1), \ldots, \bar{Y}(t_m))$ over α and b. This leads to the generalized least squares problem

$$\min_{\alpha,b} \sum_{i=1}^{m}\sum_{j=1}^{m}(\bar{Y}(t_i) - \alpha - \frac{b}{t_i})\Gamma_{ij}(\bar{Y}(t_j) - \alpha - \frac{b}{t_j}) \tag{11}$$

where $\Gamma = (\Gamma_{ij} : 1 \le i,j \le m)$ is the inverse of the covariance matrix of the random vector $(\bar{Y}(t_1), \ldots, \bar{Y}(t_m))$, with corresponding minimizers $\hat{\alpha}_2(t)$ and $\hat{b}_2(t)$. Of course, in general, the (exact) covariance matrix is unknown. This suggests two possibilities. One is to use the asymptotic description of the covariance matrix given by (7); the alternative is to use a covariance matrix that is estimated from the simulated data.

We consider first the case in which the covariance is obtained from (7).

Proposition 3.1 *Let $0 < v_1 < \cdots < v_m$. The inverse of the $m \times m$ matrix $\tilde{C}$ with (i,j)'th entry given by $\min(v_i, v_j)$ is the matrix $\tilde{\Gamma} = (\tilde{\Gamma}_{ij} : 1 \le i,j \le m)$ with*

$$\begin{cases} \tilde{\Gamma}_{i,i} = \frac{v_{i+1}-v_{i-1}}{(v_{i+1}-v_i)(v_i-v_{i-1})}, & 1 \le i \le m-1, \\ \tilde{\Gamma}_{m,m} = \frac{1}{v_m - v_{m-1}}, & \\ \tilde{\Gamma}_{i,i+1} = -\frac{1}{v_{i+1}-v_i}, & 1 \le i \le m-1, \\ \tilde{\Gamma}_{i,i-1} = -\frac{1}{v_i - v_{i-1}}, & 2 \le i \le m, \end{cases}$$

and $\tilde{\Gamma}_{i,j} = 0$ *otherwise, where* $v_0 \triangleq 0$.

4 Numerical examples

In this section, we present computational results which illustrate the theoretical properties of the regression based estimators mentioned above. We start with a description of the (three) stochastic processes and associated paramaters.

The first stochastic process is the waiting time process in the $M/M/1$ queue with server utilization $\rho = 0.9$, service rate $\mu = 1$, and an empty-and-idle initial condition.

We next turn to review the specific estimators for α that we shall consider in this paper. Given a realization of Y up to time t, our first steady-state estimator is the time average estimator

$$\bar{Y}(t) \equiv \frac{1}{t}\int_0^t Y(s)ds$$

Given selection of time points $t_1 < \cdots < t_m = t$, our regression based estimators are all linear combinations of $\bar{Y}(t_i)$'s.

Let T denote matrix transpostion and set

$$\bar{Y} = (\bar{Y}(t_1), \bar{Y}(t_2), \ldots, \bar{Y}(t_m))^T, \quad A = \begin{pmatrix} 1 & 1/t_1 \\ 1 & 1/t_2 \\ \vdots & \vdots \\ 1 & 1/t_m \end{pmatrix}, \quad \text{and } \theta = \begin{pmatrix} \alpha \\ b \end{pmatrix}.$$

Then the ordinary least square problem (8) can be rewritten as

$$\min_{\theta} \quad (\bar{Y} - A\theta)^T(\bar{Y} - A\theta) \tag{12}$$

It is well known that the solution θ^* to (12) must satisfy the normal equation (Golub and Van Loan, 1996, p. 238):

$$(A^T A)\theta^* = A^T\bar{Y} \tag{13}$$

It is clear that A has full rank, thus $\theta^* = (A^T A)^{-1}A^T\bar{Y}$. Our first regression-based estimator $\alpha_1(t)$ is therefore just the first element of θ^*, i.e. $\theta^*(1)$.

Let Σ be the covariance matrix of $\bar{Y}$. We have shown that, when Σ is assumed to be the asymptotic matrix obtained from (7), the estimators based on the ideas of generalized least squares and best unbiased linear estimator lead to the same standard truncated time-average estimator $\hat{\alpha}_2(t)$ (equation (**??**)). However, we are also interested in investigating the performance of these estimators when Σ is estimated by p independent replications of $\bar{Y}$. We denote this sample covariance matrix $\hat{C}$ (equation (**??**). Here, we shall show that, when Σ is estimated by $\hat{C}$, the estimators based on the ideas of generalized least squares and best unbiased linear estimator also lead to the same estimator. Actually, as long as Σ is positive definite, both approaches will produce identical estimator.

We rewrite (11) as

$$\min_{\theta} \quad (\bar{Y} - A\theta)^T \Sigma^{-1} (\bar{Y} - A\theta) \tag{14}$$

Let R be Σ's Choleskey triangle, i.e., $\Sigma = RR^T$. Set $\tilde{Y} = R^{-1}\bar{Y}$ and $\tilde{A} = R^{-1}A$. Then, program (14) can be written as

$$\min_{\theta} \quad (\tilde{Y} - \tilde{A}\theta)^T (\tilde{Y} - \tilde{A}\theta)$$

Again, it is easy to see that $\tilde{A}$ has full rank. Therefore, above program has an unique solution

$$\begin{aligned} \hat{\theta} &= (\tilde{A}^T \tilde{A})^{-1} \tilde{A}^T \tilde{Y} \\ &= (A^T R^{-T} R^{-1} A)^{-1} A^T R^{-T} R^{-1} \bar{Y} \\ &= (A^T \Sigma^{-1} A)^{-1} A^T \Sigma^{-1} \bar{Y} \end{aligned}$$

Use $\hat{C}$ in place of Σ, we obtain our next estimator $\hat{\alpha}_3(t)$. In particuilar,

$$\hat{\alpha}_3(t) = e^T (A^T \hat{C}^{-1} A)^{-1} A^T \hat{C}^{-1} \bar{Y}, \tag{15}$$

where $e^T = (1, 0)$.

Let $w = (w_1, \ldots . w_m)$ be the weights of $\bar{Y}$. Best linear unbiased estimator for α can be obtained by solving the quadratic program

$$\begin{aligned} \min_{w} \quad & w^T \Sigma w \\ \text{subject to} \quad & A^T w = e \end{aligned} \tag{16}$$

From the theory of optimization (Gill, Murray, and Wright 1991), program (16) has an unique solution w^* if there exists $\lambda \in \Re^2$, such that

$$A\lambda = \Sigma w^*$$

Above equation basically says that the gradient of $w^T \Sigma w$ at w^* lies in the column space of A. With this optimal condition for (16) and knowinf Σ is positive definite, we have

$$w^* = \Sigma^{-1} A\lambda \tag{17}$$

表 1: Estimated MSE of Various Estimators for the $M/M/1$ Queue Waiting Time Process: $m = 3$, MSE is estimated by 1000 independent replications.

Batch size	MSE($\bar{Y}$)	MSE($\hat{\alpha}_1$)	MSE($\hat{\alpha}_2$)	MSE($\hat{\alpha}_1$)/MSE($\bar{Y}$)	MSE($\hat{\alpha}_2$)/MSE($\bar{Y}$)
100	27.6963	40.3183	41.9098	1.4557	1.5132
500	14.5060	23.6877	23.6909	1.6330	1.6332
2500	4.2540	7.7698	6.7426	1.8265	1.5850
12500	0.8659	1.4544	1.3076	1.6795	1.5101
62500	0.2044	0.3402	0.3077	1.6641	1.5050
312500	0.0393	0.0648	0.0592	1.6476	1.5058

Multiplying both sides by A^T, we find that

$$e = Aw^* = A^T\Sigma^{-1}A\lambda$$

Now using the fact A has full rank, we can invert $A^T\Sigma^{-1}A$ and write

$$\lambda = (A^T\Sigma^{-1}A)^{-1}e$$

Next using above equation and (17) to obatin

$$w^* = \Sigma^{-1}A(A^T\Sigma^{-1}A)^{-1}e$$

It follows that, when $\hat{C}$ is used in place of Σ, the best linear unbiased estimator for α is

$$(w^*)^T\bar{Y} = (\hat{C}^{-1}A(A^T\hat{C}^{-1}A)^{-1}e)^T\bar{Y} = e^T(A^T\hat{C}^{-1}A)^{-1}A^T\hat{C}^{-1}\bar{Y},$$

which is just $\alpha_3(t)$.

Our first set of numerical experiments produce $\bar{Y}(t), \hat{\alpha}_1(t)$, and $\hat{\alpha}_2(t)$ from single replication of Y up to time t. W e use a simple rule to select time points selection. In particular, we select batch size and the number of time points m and set $t_i = i\times$ batch size.

The results of this project is a joint work with Professor Glynn at Stanford University. We will submit the results of this project to a suitable journal for publication in the near future.

參考文獻

Banks, Jerry, Carson, John S., Nelson, Barry L., and Nicol, David M. (2000), *Discrete-Event System Simulation*, Upper Saddle River, New Jersey: Prentice-Hall, Inc., 3d edition.

表 2: Estimated MSE of Various Estimators for the $M/M/1$ Queue Waiting Time Process: $m = 15$, MSE is estimated by 1000 independent replications.

Batch size	MSE($\bar{Y}$)	MSE($\hat{\alpha}_1$)	MSE($\hat{\alpha}_2$)	MSE($\hat{\alpha}_1$)/MSE($\bar{Y}$)	MSE($\hat{\alpha}_2$)/MSE($\bar{Y}$)
100	17.7043	18.5690	19.9858	1.0488	1.1289
500	4.1671	5.7623	4.5698	1.3828	1.0966
2500	0.9192	1.3712	0.9926	1.4918	1.0799
12500	0.1928	0.2835	0.2047	1.4707	1.0619
62500	0.0383	0.0568	0.0414	1.4820	1.0792

表 3: Estimated MSE of Various Estimators for the $M/M/1$ Queue Waiting Time Process: $m = 75$, MSE is estimated by 1000 independent replications.

Batch size	MSE($\bar{Y}$)	MSE($\hat{\alpha}_1$)	MSE($\hat{\alpha}_2$)	MSE($\hat{\alpha}_1$)/MSE($\bar{Y}$)	MSE($\hat{\alpha}_2$)/MSE($\bar{Y}$)
100	4.4359	6.9598	4.5307	1.5690	1.0214
500	0.8902	1.5170	0.9110	1.7040	1.0233
2500	0.1954	0.3147	0.1988	1.6101	1.0171
12500	0.0381	0.0608	0.0387	1.5968	1.0146

Bratley, P., Fox, B. L., and Schrage, L. (1983), *A Guide To Simulation*, New York: Springer-Verlag.

Cash, Charles R., Nelson, Barry L., Dippold, David G., Long, J. Mark, and Pollard, William P. (1992), "Evaluation of tests for initial-condition bias", in *WSC '92: Proceedings of the 24th conference on Winter simulation*, 577–585, New York, NY, USA: ACM Press.

Glynn, Peter W. (1995), "Some new results on the initial transient problem", in *Proceedings of the 1995 Winter Simulation Conference*, 165–170, Arlington, VA USA.

Glynn, Peter W. and Heidelberger, P. (1990), "Bias properties of budget constrained monte carlo simulations", *Operations Research*, 38, 801–814.

——— (1991a), "Analysis of initial transient deletion for replicated steady-state simulations", *Operations Research Letters*, 10, 437–443.

——— (1991b), "Analysis of parallel, replicated simulations under a completion time constraint", *ACM Trans. on Modeling and Computer Simulation*, 1, 3–23.

——— (1992a), “Analysis of initial transient deletion for parallel steady-state simulations”, *SIAM Journal Scientific Stat. Computing*, 13, 909–922.

——— (1992b), “Experiments with initial transient deletion for parallel, replicated steady-state simulations”, *Management Science*, 38, 400–418.

——— (1992c), “Jackknifing under a budget constraint”, *ORSA Journal on Computing*, 4, 226–234.

Glynn, Peter W. and Iglehart, Donald L. (1988), “Simulation methods for queues: an overview”, *Queueing Systems Theory Appl.*, 3(3), 221–255.

Glynn, Peter W. and Whitt, Ward (1991), “Estimating the asymptotic variance with batch means”, *Oper. Res. Lett.*, 10(8), 431–435.

——— (1994), “Logarithmic asymptotics for steady-state tail probabilities in a single-server queue”, *J. Appl. Probab.*, 31A, 131–156, studies in applied probability.

Goldsman, D., Schruben, L. W., and Swain, J.J. (1994), “Tests for transient means in simulated time series”, *Naval Res. Logist. Quart.*, 41, 171–187.

Hsieh, M.-H., Iglehart, D. L., and Glynn, P. W. (2004), “Empirical performance of bias-reducing estimators for regenerative steady-state simulations”, *ACM Trans. on Modeling and Computer Simulation*, 14, 325–343.

Jr., K. P. White, Cobb, M. J., and Spratt, Stephen C. (2000), “A comparison of five steady-state truncation heuristics for simulation”, in *WSC '00: Proceedings of the 32nd conference on Winter simulation*, 755–760, San Diego, CA, USA: Society for Computer Simulation International.

K. Preston White, Jr. (1997), “An effective truncation heuristic for bias reduction in simulation output”, *Simulation*, 69(6), 323–334.

Law, Averill M. and Kelton, W. David (2000), *Simulation Modeling & Analysis*, New York: McGraw-Hill, Inc, 3d edition.

Meketon, M.S. and Heidelberger, P. (1982), “A renewal theoretic approach to bias reduction in regenerative simulations”, *Manage. Sci.*, 26, 173–181.

Schruben, L. W. (1982), “Detecting initialization bias in simulation output”, *Oper. Res.*, 30(3), 151–153.

Schruben, L. W., Singh, H., and Tierney, L. (1983), “Optimal tests for initialization bais in simulation output”, *Oper. Res.*, 31(6), 1167–1178.

Steiger, Natalie M., Lada, Emily K., Wilson, James R., Joines, Jeffrey A., Alexopoulos, Christos, and Goldsman, David (2005), “ASAP3: a batch means procedure for steady-state simulation analysis”, *ACM Trans. Model. Comput. Simul.*, 15(1), 39–73.