Proceedings of the 1985 Winter Simulation Conference
D. Gantz, G. Blais, S. Solomon (eds.)

THE CORRELATION BETWEEN MEAN AND VARIANCE ESTIMATORS

Keebom Kang
Department of I.E.
University of Miami
Coral Gables, FL 33124

David Goldsman
School of ISyE
Georgia Institute of Technology
Atlanta, GA 30332

ABSTRACT

One of the pitfalls encountered when using confidence interval estimators for the mean of a stationary stochastic process is that the mean and variance estimators ($\bar{X}$ and V) may be correlated. We derive Corr($\bar{X}$,V) for various variance estimators and stochastic processes, and we examine the effects of this correlation upon confidence interval estimator performance. Among the variance estimators under consideration are those arising from the methods of batched means and standardized time series. Both small sample and asymptotic results are reported.

1. INTRODUCTION

An active area of simulation output analysis research involves estimation of confidence intervals for the mean of a stationary stochastic process, $X_1, X_2, \ldots, X_n$. Approximate $100(1-\alpha)\%$ confidence interval estimators (c.i.e.'s) for the underlying process mean μ are typically given by:

$$\mu \in \bar{\bar{X}}_n \pm t_{1-\frac{1}{2}\alpha}\sqrt{V}\ , \qquad (1\text{-}1)$$

where $\bar{\bar{X}}_n \equiv \Sigma X_i/n$, V is an estimator for $\mathrm{Var}(\bar{\bar{X}}_n)$, and $t_{1-\frac{1}{2}\alpha}$ is the appropriate quantile of a t-distribution.

The purpose of this paper is to investigate properties of Corr($\bar{\bar{X}}_n$,V) for various variance estimators and stochastic processes of interest (Section 3). We also examine the effects of this correlation upon c.i.e. performance (Section 4).

2. BACKGROUND

2.1 Confidence Intervals

In order for the c.i.e.'s given by (1-1) to be (asymptotically) valid, three requirements must hold:

(1) $\bar{\bar{X}}_n \approx \mathrm{Nor}(\mu, \sigma^2/n)$

(2) $\bar{\bar{X}}_n$ and V are approximately independent

(3) $V \approx \sigma^2\chi^2(d)/d$, where $\sigma^2 \equiv \lim_{n\to\infty} n\mathrm{Var}(\bar{\bar{X}}_n)$ and d is the appropriate degrees of freedom.

These requirements guarantee that $(\bar{\bar{X}}_n-\mu)/\sqrt{V}$ has a t-distribution. Thus, any violation of the requirements can affect the validity of the c.i.e. Eq.(1-1).

Example 1: Violation of the normality assumption in terms of skewness results in the skewness of $(\bar{\bar{X}}_n-\mu)/\sqrt{V}$ [Johnson (1978)]; so this statistic no longer has the t-distribution.

Example 2: Nonzero correlation between $\bar{\bar{X}}_n$ and V results in asymmetric confidence interval coverage (which will be discussed in more detail later). For example, suppose we wish to estimate the expected customer waiting time for an M/M/1 queueing model with ρ = 0.8. (The theoretical expected waiting time is 3.2). Assume that we obtain two estimates, 3.0 and 3.4. If Corr($\bar{\bar{X}}_n$,V) > 0, then a confidence interval based on the 3.4 point estimate will have greater chance of covering the true parameter, 3.2, than will a confidence interval based on the 3.0 estimate.

Example 3: If the variance estimator V is not distributed as a chi-squared random variable (times the appropriate constant), then $(\bar{\bar{X}}_n-\mu)/\sqrt{V}$ no longer has the t-distribution.

There are a number of problems associated with the estimation of the mean of a stationary simulation process. The most serious pitfall is related to the fact that, in many simulations, the X_i's are serially correlated [Law (1977)]. This serial correlation can result in violations of requirements (2) and (3). Requirement

(1) does not pose a major problem in confidence interval estimation since, by a central limit theorem, the sample mean of the X_i's becomes approximately normal as n becomes large.

Over the last two decades, a number of confidence interval estimation methodologies have been proposed and studied: batched means, independent replications, ARMA time series modeling, spectral representation, regeneration, standardized time series [Schruben (1983)], and overlapping batched means [Meketon and Schmeiser (1984)]. Details concerning the first five methodologies can be found in, e.g., Bratley, Fox, and Schrage (1983). In this paper, we concentrate primarily on variance estimators arising from the methods of batched means, independent replications, and standardized time series.

2.2 Batched Means

Suppose that we divide the stationary stochastic process $X_1,\ldots,X_n$ into $b > 1$ contiguous, nonoverlapping *batches*. For ease of exposition, assume that each batch is of length m (so that $n = bm$). Denote

$$\bar{X}_{i,m} \equiv \frac{1}{m}\sum_{j=1}^{m} X_{(i-1)m+j}$$

as the i-th *batched mean*, $i=1,\ldots,b$.

$$\underbrace{X_1,\ldots,X_m}_{\text{batch 1}},\underbrace{X_{m+1},\ldots,X_{2m}}_{\text{batch 2}},\ldots,\underbrace{X_{(b-1)m+1},\ldots,X_n}_{\text{batch m}}$$

$$\bar{X}_{1,m} \qquad \bar{X}_{2,m} \qquad \cdots \qquad \bar{X}_{b,m}$$

Assuming that the batched means are approximately i.i.d. normal random variables with unknown mean μ and variance σ^2/m, then a $100(1-\alpha)\%$ c.i.e. for μ is given by

$$\mu \in \bar{\bar{X}}_n \pm t_{b-1,1-\frac{1}{2}\alpha}(V_{0,b}/n)^{\frac{1}{2}}, \qquad (2\text{-}1)$$

where $t_{d,\gamma}$ is the upper γ quantile of a t-distribution with d degrees of freedom and $V_{0,b} \equiv m\sum_{j=1}^{b}[\,\bar{X}_{j,m} - \bar{\bar{X}}_n\,]^2/(b-1)$ is the *classical batched means* estimator for σ^2.

In the method of *independent replications*, we conduct b independent runs of the simulation, each of length m. The replicate means are computed, and Eq.(2-1) is used to calculate confidence intervals for μ.

replication 1: $X_1,X_2,\ldots,X_m \to \bar{X}_{1,m}$

replication 2: $X_{m+1},X_{m+2},\ldots,X_{2m} \to \bar{X}_{2,m}$

$\vdots$

replication b: $X_{(b-1)m+1},\ldots,X_{bm} \to \bar{X}_{b,m}$

As $m \to \infty$, the methods of batched means and independent replications effectively are equivalent.

2.3 Standardized Time Series

The method of *standardized time series* (STS) uses a process central limit theorem to (asymptotically) transform the stochastic process of interest into a so-called *Brownian bridge*. Properties of Brownian bridges are then used to estimate confidence intervals for the mean.

Consider a stationary process $X_1,\ldots,X_n$ [satisfying other mild assumptions from Schruben (1983)], which is divided into b batches of size m. Denote the j-th cumulative mean from batch i as:

$$\bar{X}_{i,j} \equiv \frac{1}{j}\sum_{k=1}^{j} X_{(i-1)m+k}$$

($\bar{X}_{i,m}$ is the i-th batched mean). For all i and j, let

$$S_{i,j} \equiv \bar{X}_{i,m} - \bar{X}_{i,j} \quad \text{and}$$

$$T_{i,m}(t) \equiv \frac{\|mt\| S_{i,\|mt\|}}{\sigma\sqrt{m}},$$

where $\|.\|$ is the greatest integer function and $T_{i,m}(t)$ is the *standardized time series* from the i-th batch. Schruben shows that as $m \to \infty$, $T_{i,m}(t) \xrightarrow{\mathcal{D}} B_t$, $t \in [0,1]$, where B_t is a standard Brownian bridge (i.e., Brownian motion which starts and stops at zero).

Finally, define for all batches,

$$\hat{A}_i \equiv \sum_{j=1}^{m} jS_{i,j},$$

$$\hat{K}_i \equiv \underset{j}{\operatorname{argmax}}\,\{jS_{i,j}\}, \quad \text{and}$$

$$\hat{S}_i \equiv \hat{K}_i S_{i,\hat{K}_i}.$$

Then the following are estimators for σ^2:

(0) Classical batched means estimator: $V_{0,b}$ from Eq.(2-1).

(1) Area estimator:

$$V_{1,b} \equiv \frac{12}{(m^3-m)b} \sum_{i=1}^{b} \hat{A}_i^2 \approx \sigma^2 \chi^2(b)/b, \quad b \geq 1.$$

(2) Maximum estimator:

$$V_{3,b} \equiv \frac{m}{3b} \sum_{i=1}^{b} \frac{\hat{S}_i^2}{\hat{K}_i(m-\hat{K}_i)} \approx \frac{\sigma^2 \chi^2(3b)}{3b}, \quad b \geq 1.$$

To construct confidence intervals based on the area and maximum estimators, we simply use Eq.(2-1) with the appropriate variance estimator and degrees of freedom. Goldsman and Schruben (1984) show that the STS c.i.e.'s (area and maximum) possess certain advantages over the classical batched means estimator.

2.4 Some Time Series Processes of Interest

We call a process $\{X_t: t \in Z_+\}$ an *autoregressive moving average process* of orders p and q [ARMA(p,q)] if it is implicitly defined by:

$$\sum_{i=0}^{p} \varphi_i X_{t-i} = \sum_{j=0}^{q} \theta_j \varepsilon_{t-j},$$

where φ_0, φ_p, θ_0, and θ_q are nonzero, and the ε_t's are uncorrelated random variables with mean 0 and variance σ_ε^2. Usually, the ε_t's are assumed to be i.i.d. normal. ARMA models have been found to adequately approximate many processes encountered in practice [cf. Box and Jenkins (1976)]. A *moving average* process of order q [MA(q)] is an ARMA(0,q) process, and an *autoregressive* process of order p [AR(p)] is an ARMA(p,0) process.

It is useful to specifically define the following ARMA-type processes:

(1) MA(1): $X_i = \varepsilon_i + \alpha\varepsilon_{i-1}$, where $\varepsilon_i \sim$ i.i.d. Nor(0,1).

(2) AR(1): $X_i = \alpha X_{i-1} + \varepsilon_i$, where $\varepsilon_i \sim$ i.i.d. $\mathrm{Nor}(0,\tau^2(1-\alpha^2))$, $\tau^2 \equiv \mathrm{Var}(X_i)$.

(3) EAR(1): $X_i = \begin{cases} \alpha X_{i-1} & \text{w.p. } \alpha \\ \alpha X_{i-1} + \varepsilon_i & \text{w.p. } 1-\alpha \end{cases},$

where $\varepsilon_i \sim$ i.i.d. Exp(λ) and $0 \leq \alpha < 1$. The correlation structure of the EAR(1) process is the same as that of the AR(1). If $\alpha=0$, then $X_i \sim$ i.i.d. Exp(λ).

3. SOME RESULTS CONCERNING $\mathrm{Corr}(\bar{\bar{X}}_n, V)$

We study the correlation between the sample mean and various estimators of the process variance σ^2. Analytical and empirical results for several simple processses are presented.

Theorem 3-1: Consider any i.i.d. process $X_1,\ldots,X_n$ with $E[X_i] = 0$ and $E[X_i^3] < \infty$. Define $S^2 \equiv \Sigma_i(X_i - \bar{\bar{X}}_n)^2/(n-1)$. Then $\mathrm{Cov}(\bar{\bar{X}}_n, S^2) = E[X_i^3]/n$. (Hence, the covariance is dependent on the skewness of the X_i's.)

Proof: See Kang (1984), p. 98.

Remark: Identifying the supposedly i.i.d. batched (or replicated) means with the X_i's in the above theorem, we see that the covariance between $\bar{\bar{X}}_n$ and $V_{0,b}$ is dependent on the skewness of the batched means.

Theorem 3-2: For any mean 0 covariance stationary ARMA(p,q) process with i.i.d. symmetric about 0 noise, $\mathrm{Corr}(\bar{\bar{X}}_n, V_{0,b}) = 0$.

Proof: See Kang (1984), p. 97.

Theorem 3-3: Under the same conditions as in Theorem 3-2, $\mathrm{Corr}(\bar{\bar{X}}_n, V_{1,b}) = 0$.

Proof: See Goldsman (1984), p. 106.

Remarks: (i) In particular, $\mathrm{Corr}(\bar{\bar{X}}_n, V_{0,b}) = \mathrm{Corr}(\bar{\bar{X}}_n, V_{1,b}) = 0$ for the MA(1) and AR(1) processes.

(ii) Although the AR(1) and EAR(1) processes have the same correlation structure, the above results do not apply for the EAR(1) process; analytical results for this process are somewhat tedious and are deferred to Kang and Goldsman (1985).

Empirical Results:

(1) $\mathrm{Corr}(\bar{\bar{X}}_n, V_{i,b})$ (i=0,1,3) is *nonzero* for the EAR(1) process. For example, consider $V_{1,1}$, the area estimator based on one batch. Table 3.1 gives estimated values of $\mathrm{Corr}(\bar{\bar{X}}_n, V_{1,1})$ for various n and α. We see that the correlation is significantly greater than zero for small sample size n. As sample size increases, the correlation decreases. It also appears that the correlation increases as α approaches 1.

Similar results can be obtained for $V_{0,b}$ and $V_{3,b}$ [cf. Kang and Goldsman (1985)].

n	α = 0.0	0.1	0.5	0.9
10	.338	.353	.452	.425
50	.213	.241	.285	.441
200	.123	.100	.122	.300
500	.075	.083	.085	.170
1000				.139
2500				.097

Table 3.1

Estimated $\mathrm{Corr}(\bar{\bar{X}}_n, V_{1,1})$ for the EAR(1) process for various n and α. Each entry is based on 1000 independent simulation runs. (All standard errors were less than 0.035.)

(2) Analytical expressions for $\mathrm{Corr}(\bar{\bar{X}}_n, V_{3,b})$ are not generally tractable. Nevertheless, as can be seen in Table 3.2, $\mathrm{Corr}(\bar{\bar{X}}_n, V_{3,1})$ is very close to zero for the AR(1) case. [The same is true for the MA(1) process.]

n	α = -0.9	-0.5	0.0	-0.5	-0.9
10	-.037	-.010	-.036	-.035	-.049
50	.009	-.045	-.025	.023	.040
200	-.052	-.008	-.060	-.028	.042

Table 3.2

Estimated $\mathrm{Corr}(\bar{\bar{X}}_n, V_{3,1})$ for the AR(1) process for various n and α. Each entry is based on 1000 independent simulation runs. (All standard errors were less than 0.037.)

A reasonable conjecture is that asymmetry in the underlying distribution of the stochastic process causes $\mathrm{Corr}(\bar{\bar{X}}_n, V_{i,b})$ to be nonzero. Further, the correlation approaches zero as the sample size increases (this being obvious for the classical batched means case by a central limit theorem).

4. IMPLICATIONS OF $\mathrm{Corr}(\bar{\bar{X}}_n, V)$ ON CONFIDENCE INTERVAL PERFORMANCE

Suppose we are working with one of the aforementioned c.i.e.'s. Denote the endpoints of the c.i.e. by $[L_\alpha, U_\alpha]$, where $1-\alpha$ is the nominal probability of *coverage*; i.e., we desire $\Pr\{L_\alpha \le \mu \le U_\alpha\} = 1-\alpha$. Note that L_α and U_α are random variables. Also, define:

$$E(\alpha) \equiv |\Pr\{L_\alpha \le \mu \le U_\alpha\} - (1-\alpha)|,$$

$$E^l(\alpha) \equiv \Pr\{L_\alpha > \mu\} - \alpha/2, \text{ and}$$

$$E^u(\alpha) \equiv \Pr\{U_\alpha < \mu\} - \alpha/2.$$

$E(\alpha)$ is called the *coverage error function*; it is the difference between the actual coverage and the nominal coverage of the parameter μ. A c.i.e. is said to be *symmetric* if $E^l(\alpha) = E^u(\alpha)$. Although coverage is of primary importance as a c.i.e. performance criterion, Glynn (1982) comments that symmetry of coverage is also of some import. He also argues that since $E(\alpha)$ converges to zero faster than $E^u(\alpha)$ or $E^l(\alpha)$, asymmetric coverage can occur even if the actual coverage is close to the nominal value $1-\alpha$ [see the results for process (2) in Table 3.3 below].

Several authors [e.g., Fishman (1978)] remark that $\mathrm{Corr}(\bar{\bar{X}}_n, V)$ can play an important role in the performance of confidence interval procedures. In the present section, we give examples which show that this correlation may or may not have a significant effect on confidence interval coverage. An example is also given showing that this correlation can have a significant effect on the symmetry of the coverage.

The following three time series were examined in order to study the relationships between $\mathrm{Corr}(\bar{\bar{X}}_n, V_{0,b})$ and c.i.e. coverage, and between $\mathrm{Corr}(\bar{\bar{X}}_n, V_{0,b})$ and symmetry of the coverage.

(1) AR(1): $X_t = 0.95 * X_{t-1} + \varepsilon_t$, where $\varepsilon_t \sim$ i.i.d. Nor(0,1).

(2) $X_t \sim$ i.i.d. Exp(λ).

(3) Customer waiting times in queue for an M/M/1 model with $\rho = 0.8$.

[These three processes were chosen so as to compare c.i.e. performance for processes with different levels of dependency and

$\mathrm{Corr}(\bar{\bar{X}}_n, V_{0,b})$: In (1), the X_t's are highly dependent but $\mathrm{Corr}(\bar{\bar{X}}_n, V_{0,b}) = 0$. In (2), the X_t's are independent but the correlation is positive. The X_t's are dependent and the correlation is positive in case (3).]

We conducted 1000 runs of the three time series. From each of the three groups of 1000 runs, we calculated 1000 confidence intervals for the appropriate process mean. We were then able to estimate actual coverage of the mean, $E^u(0.1)$, $E^l(0.1)$, and $\mathrm{Corr}(\bar{\bar{X}}_n, V_{0,b})$ for the three time series. These results are summarized in Table 3.3.

Experiment	(1)	(2)	(3)
Coverage of the mean	0.162	0.892	0.837
$E^u(\alpha)$	0.361	-0.033	-0.032
$E^l(\alpha)$	0.377	0.045	0.095
Corr	-0.003	0.729	0.595

Table 3.3

Coverage, $E^u(\alpha)$, $E^l(\alpha)$, and $\mathrm{Corr}(\bar{\bar{X}}_n, V_{0,b})$ for the three experiments described in Section 4 ($\alpha = 0.10$).

The entries in Table 3.3 show that poor coverage can be obtained when $\mathrm{Corr}(\bar{\bar{X}}_n, V_{0,b})$ is nearly zero [process (1)], and good coverage (i.e, actual coverage close to the nominal coverage) can be obtained when this correlation is quite positive [process (2)]. Thus, $\mathrm{Corr}(\bar{\bar{X}}_n, V)$ may or may not have a significant effect on the coverage when the point estimator of μ is unbiased. However, Glynn (1982) points out that nonzero $\mathrm{Corr}(\bar{\bar{X}}_n, V)$ can be very detrimental (in terms of coverage) when the point estimator of μ is biased (as in the case of the regenerative confidence interval method).

As is illustrated in Table 3.3 by the results for processes (2) and (3) [and as is argued in Example 2 of Section 1], nonzero $\mathrm{Corr}(\bar{\bar{X}}_n, V)$ appears to cause asymmetric coverage.

5. SUMMARY

In this paper, we analytically and empirically studied $\mathrm{Corr}(\bar{\bar{X}}_n, V)$ for various variance estimators and stochastic processes. We examined the effects of this correlation upon confidence interval estimator coverage and symmetry of coverage. From limited Monte Carlo work, small sample and "asymptotic" results were reported.

REFERENCES

Box, G.E.P. and G.M. Jenkins (1976). *Time Series Analysis: Forecasting and Control.* Holden—Day, San Francisco.

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

Fishman, G.S. (1978). "Grouping Observations in Digital Simulation". *Management Science*, 24, 510—521.

Glynn, P.W. (1982). "Coverage Error for Confidence Intervals Arising in Simulation Output Analysis". *Proc. Winter Simul. Conf.*, II, 369—375.

Goldsman, D. (1984). *On Using Standardized Time Series to Analyze Stochastic Processes.* Ph.D. Dissertation, School of O.R. & I.E., Cornell University, Ithaca, NY.

Goldsman, D. and L. Schruben (1984). "Asymptotic Properties of Some Confidence Interval Estimators for Simulation Output". *Management Science*, 30, 1217—1225.

Johnson, N.J. (1978). "Modified t-test and Confidence Intervals for Asymmetrical Populations". *JASA*, 73, 536—544.

Kang, K. (1984). *Confidence Interval Estimation via Batch Means and Time Series Modeling.* Ph.D. Dissertation, School of Industrial Engineering, Purdue University, West Lafayette, IN.

Kang, K. and D. Goldsman (1985). "The Correlation between Mean and Variance Estimators". School of ISyE Tech. Report, Georgia Tech, Atlanta, GA.

Law, A.M. (1977). "Confidence Intervals in Discrete Event Simulation: A Comparison of Replication and Batch Means". *Naval Res. Log. Quart.*, 24, 667—678.

Meketon, M.S. and B.W. Schmeiser (1984). "Overlapping Batch Means: Something for Nothing?" *Proc. Winter Simul. Conf.*, 227–230.

Schruben, L. (1983). "Confidence Interval Estimation Using Standardized Time Series". *Operations Research*, 31, 1090–1108.

KEEBOM KANG

Keebom Kang is an Assistant Professor of Industrial Engineering at the University of Miami. He received his Ph.D. from the School of Industrial Engineering, Purdue University. His research interests include simulation output analysis and stochastic modeling.

Department of Industrial Engineering
University of Miami
Coral Gables, FL 33124
(305) 284-2370

DAVE GOLDSMAN

Dave Goldsman is an Assistant Professor of Industrial and Systems Engineering at Georgia Tech. He holds a Ph.D. from the School of O.R. & I.E., Cornell University. His research interests include simulation output analysis, and ranking and selection.

School of ISyE
Georgia Tech
Atlanta, GA 30332
(404) 894-2365