Proceedings of the 1987 Winter Simulation Conference
A. Thesen, H. Grant, W. David Kelton (eds.)

CORRELATION AMONG ESTIMATORS OF THE VARIANCE OF THE SAMPLE MEAN

Bruce Schmeiser
Wheyming Tina Song
School of Industrial Engineering
Purdue University
West Lafayette, IN 47907, U.S.A.

ABSTRACT

Various types of estimators have been proposed for estimating the variance of the sample mean, a fundamental quantity in simulation output analysis. When used with low degrees of freedom, several of these estimators have little bias. But the low degrees of freedom correspond to high variance. One approach to creating estimators with smaller variance while maintaining the negligible bias is to use linear combinations of known estimators. Whether linear combinations provide improved estimators — and, if so, the choice of estimators to be included in the linear combination — depends upon the correlations among the various estimators. Linear combinations of estimators having high positive correlation would provide little improvement while combinations of independent estimators would provide substantial gain. We investigate the correlation among four well-known estimators as a function of the type of stochastic process generating the data, the sample size, the estimator type, and estimator parameters.

1. INTRODUCTION

Consider estimating the mean of a covariance-stationary process X using the sample mean $\bar{X} = n^{-1}\sum_{i=1}^{n} X_i$. Let ρ_h denote the lag-h autocorrelation $\mathrm{corr}(X_i, X_{i+h})$. The variance of the sample mean is $\mathrm{V}(\bar{X}) = n^{-1}V(X_i)[1 + 2\sum_{h=1}^{n-1}(1 - h/n)\rho_h]$. Estimating $\mathrm{V}(\bar{X})$ is central to confidence-interval, tolerance-interval, and ranking-and-selection procedures.

A variety of ideas have been used to estimate $\mathrm{V}(\bar{X})$ from the observations $X_1, X_2, \cdots, X_n$: direct estimator (Moran 1975), regression-based spectral-analysis estimator (Heidelberger and Welch 1981), ARMA time-series estimator (Fishman 1971, Schriber and Andrews 1984), nonoverlapping batch-means estimator (Schmeiser 1982), overlapping batch-means estimator (Meketon 1980, Meketon and Schmeiser 1984, Welch 1987), standardized time-series area and maximum estimators (Schruben 1983, Goldsman 1984), and the regenerative estimator (Crane and Iglehart 1975, Crane and Lemoine 1977). Goldsman and Schruben (1984), Goldsman, Kang and Sargent (1986), Goldsman and Meketon (1985), Law and Kelton (1982, 1984), and Song and Schmeiser (1987b) discuss relationships among various estimators.

Each type of estimator has its own strengths and weaknesses in terms of applicability and statistical properties. Each works well when underlying assumptions hold, but the statistical properties deteriorate as a function of the degree to which the assumptions are violated. Other than the regenerative estimator, each type of estimator has one or more parameters. These parameters determine the estimator's properties, in particular the tradeoff made between bias and variance. Usually the estimators are approximately asymptotically chi-square random variables and we speak of the number of degrees of freedom, which is inversely proportional to the estimator variance. Often the number of degrees of freedom is a function of a number of batches.

Choosing the value of the parameter for a particular estimator and application is difficult. Low degrees of freedom result in small bias and large variance; in terms of confidence intervals this implies good coverage probability and interval widths that are both large and highly variable. High degrees of freedom can result in large bias and small variance; in terms of confidence intervals this implies poor coverage probability and widths that are short and stable. The problem is to find the number of degrees of freedom that provides a good tradeoff between bias and variance.

When the number of degrees of freedom is determined by number of batches, the typical approach has involved repeatedly testing hypotheses of normality and independence on larger and larger batch sizes (smaller and smaller numbers of batches) in the hope of finding an appropriate batch size.

Our research program has taken a different approach. We hope to find estimators that use very large batches — so large that the analyst is usually confident without any hypothesis testing. Such batches might be one-quarter to one-half as long as the entire simulation run. A reasonable number of degrees of freedom is then obtained by combining different types of estimators with the same or different batch sizes. We refer to the estimators used to form the linear combination as *component* estimators.

One way to combine different component estimators is to form linear combinations with weights summing to one. If each component estimator has low degrees of freedom, the bias of the linear combination will be low. If the component estimators are not deterministically related and if weights close to the optimal (in the sense of smallest variance) values are used, the variance of the linear combination will be less than the variance of any individual estimator. If the correlations among the component estimators are not too close to one, then the variance of the linear combination will be considerably less than that of any one estimator.

The optimal weights are a simple function of the covariances between the component estimators. Let $\hat{\underline{V}} = (\hat{V}_1, \hat{V}_2, \cdots, \hat{V}_p)$ denote the vector of p component estimators and let $\Sigma_{\hat{\underline{V}}}$ denote the covariance matrix of $\hat{\underline{V}}$. Then $\underline{\alpha}$, the vector of weights that minimizes the variance of the linear combination $\hat{V} = \underline{\alpha}^t \hat{\underline{V}}$ subject to $\underline{1}^t \underline{\alpha} = 1$ is

$$\underline{\alpha} = (\Sigma_{\hat{\underline{V}}}^{-1} \underline{1}) \,/\, (\underline{1}^t \Sigma_{\hat{\underline{V}}}^{-1} \underline{1})$$

and the minimal variance is $(\underline{1}^t \Sigma_{\hat{\underline{V}}}^{-1} \underline{1})^{-1}$, where $\underline{1}$ is a vector of ones and assuming $\Sigma_{\hat{\underline{V}}}$ has full rank (i.e., no component estimators are deterministically linearly related).

The idea of linear combinations of component estimators is not new. For batch size n/k, where n is the run length and k is the number of batches, Schruben (1983) takes a linear combination of the standardized-time-series area estimator (with k degrees of freedom) and the nonoverlapping-batch-means estimator (with $k-1$ degrees of freedom). For large batch sizes these two estimators are asymptotically independent, so the combination has $2k-1$ degrees of freedom. Therefore the variance of the linear combination is about half that of either component estimator. Another example is the overlapping-batch-means variance estimator, which can be viewed as a linear combination of nonoverlapping-batch-means estimators (Meketon and Schmeiser 1984). The degrees of freedom can not be summed since the estimators are not independent, but complete overlapping (that is, using every possible batch of a specified batch size) results in fifty percent more degrees of freedom (asymptotically). Partial overlapping (taking fewer nonoverlapping batches in the linear combination) results in somewhat fewer degrees of freedom (Welch 1987).

In this paper, we report the results of some exploratory Monte Carlo experiments designed to investigate the correlations among some well-known estimators. The experimental factors are described in Section 2. Two particular experiments are described in Sections 3 and 4. Section 5 is a discussion.

2. EXPERIMENTAL FACTORS

Four types of estimators are considered: nonoverlapping batch means (NBM), overlapping batch means (OBM), standardized-time-series area (STS-A), and Heidelberger-Welch regression-based spectral estimator (HW). Each receives as input the sample size n and a vector of data $x_1, x_2, \cdots, x_n$. Except for HW, which has only six different degrees of freedom, either the number of batches or the degrees of freedom is an input, with the other calculated. (The appendices contain the NBM, OBM, and HW subroutines, which can be obtained from the first author. Note the denominator of the OBM estimator is different from the original used in Meketon and Schmeiser (1984). The OBM relationship between batch size and degrees of freedom was determined empirically. The HW routine was coded by James R. Wilson.)

Two types of data are considered: independent, identically distributed normal observations and steady-state first-order autoregressive normal observations with $\phi_1 = .8182$. This particular value of ϕ corresponds to $c = 1 + 2\sum_{h=1}^{\infty}\rho_h = 10$, a moderate value.

Four sample sizes are considered: n=100, 500, 2000, and 10000. In all cases $\mathrm{V}(\overline{X}) = 1$, so the variance of the observations, $\mathrm{V}(X)$, is a function of n.

For each combination of estimator type, data type, and sample size, correlations are estimated using 100 samples of size n to obtain each observed correlation. Replicating 100 times (that is, 10000 samples of size n) we obtain estimates of the correlation with standard error of less than .01, so the entries in the table have an error shown of one or two units in the last digit. The estimates corresponding to n=10000 are based on fewer observations and have standard errors of about .03. All observations are independent; that is, common random numbers are not used.

The first experiment, discussed in Section 3, considers each of the four estimator types parameterized for seven degrees of freedom. The second experiment, discussed in Section 4, considers each of the four estimator types parameterized for fixed number of batches (batch size) rather than fixed degrees of freedom.

3. EXPERIMENT 1: 7 DEGREES OF FREEDOM

Section 3.1 contains the results of a Monte-Carlo experiment in which each of the four estimator types has nominally seven degrees of freedom. The implications for linear combinations are discussed in Section 3.2.

3.1. Estimated Covariances and Correlations

Table 1 contains the estimated correlations.

Table 1: Estimated Correlations for Experiment 1. Nominal Seven Degrees of Freedom.

	OBM		STS-A		HW		
	iid normal	AR(1) $c=10$	iid normal	AR(1) $c=10$	iid normal	AR(1) $c=10$	sample size n
NBM	.72	.90	.23	.33	.53	.60	100
	.71	.77	.26	.32	.52	.56	500
	.72	.73	.26	.27	.53	.50	2000
	.68	.69	.24	.23	.50	.53	10000
OBM			.17	.28	.55	.52	100
			.22	.23	.57	.57	500
			.19	.21	.58	.55	2000
			.24	.19		.59	10000
STS-A					.38	.45	100
					.39	.44	500
					.39	.37	2000
						.41	10000

The experimental results lead to several observations.

Observation 1. The asymptotic correlations are reached quickly. In each case of iid data, a sample size of n=100 yields about the same correlation as larger sample sizes. In each case of dependent data, the asymptotic values have been reached by sample size n=2000.

Observation 2. The largest deviations from the asymptotic correlations occur in the most difficult case, n=100 and c=10. Although not shown, this case has the most severe biases: −40%, −30%, −70%, and −20%, respectively, for NBM, OBM, STS-A, and HW.

Observation 3. The asymptotic correlations are the same for both types of data. Using the view that c dependent observations play the role of a single independent observation, we would expect more or less the same behavior for n=100 independent observations as for n=1000 dependent observations, which is not inconsistent with the tabled values.

Observation 4. The NBM and STS-A estimators are not asymptotically independent in this experiment. This does not contradict Schruben (1983), who uses asymptotic independence, because he assumes equal batch sizes and we are using equal degrees of freedom.

3.2. Optimal Linear Combinations

Let us suppose, based on the third observation, that asymptotic correlations do not depend on the type of data process (for some large set of processes satisfying some reasonable mixing and moment restrictions). Then we can calculate the optimal linear combinations of estimators and the variance of the optimal linear combination. Assume that for NBM, OBM, STS-A, and HW estimators respectively we have the component covariance matrix

$$\Sigma_{\underline{\hat{V}}} = \begin{bmatrix} .28 & .20 & .07 & .14 \\ .20 & .28 & .06 & .15 \\ .07 & .06 & .28 & .10 \\ .14 & .15 & .10 & .25 \end{bmatrix},$$

based on the correlations shown in Table 1 and estimated variances. (Since $V(\bar{X}) = 1$, an estimator with seven degrees of freedom should have a variance of $2/7 \approx .286$, so HW is providing a variance more consistent with 8 degrees of freedom.) Given Σ, the results that follow are deterministic.

The resulting optimal linear combination has weights $\underline{\alpha} = (.19, .24, .39, .17)$. The minimal variance is .15, a 40% decrease in the variance compared to HW. These values are shown in Table 2 along with corresponding optimal values taking combinations of three estimators and of two estimators.

Negative correlation between component estimators is unreasonable to expect, since they are functions of the same data. Instead, independence among component estimators provides a reasonable benchmark. Then for pairs of estimators, the best variance that could be hoped for is about .13, which is not obtained by any pair since all are positively correlated. The best pair is (OBM, STS-A) at

.17, with (NBM, STS-A) and (STS-A, HW) close behind.

Table 2. Optimal Weights and Associated Variances for Linear Combinations of Pairs, Triples, and All Component Estimators. Nominal Seven Degrees of Freedom.

NBM	OBM	STS-A	HW	Variance
.5	.5			.24
	.43		.57	.21
.44			.56	.20
		.45	.55	.18
.5		.5		.18
	.5	.5		.17
.29	.23		.48	.20
.32		.38	.30	.16
	.33	.39	.27	.16
.25	.30	.45		.16
.19	.24	.37	.23	.15

Of the triples, the immediate observation is that STS-A should always be included, since all three of its triples have a variance of about .16 while the triple without STS-A has a variance of .20.

And of course using all four estimators in the linear combination is the best with a variance of .15, which corresponds to about 13 degrees of freedom.

These empirical results using identical degrees of freedom for each component estimator make one appreciate the asymptotic independence between STS-A and NBM obtained when the batch sizes are identical. With seven degrees of freedom for STS-A, there are seven batches. With these seven batches, NBM has six degrees of freedom. Therefore the (STS-A, NBM) covariance matrix is diagonal with STS-A variance .28 and NBM variance $(7/6)\times.28=.33$. The optimal weights are .54 for STS-A and .46 for NBM, with variance .15. That is, the optimal linear combination of NBM and STS-A has about the same variance as with the optimal linear combination of all four estimators restricted to having seven degrees of freedom. In addition, in the (NBM, STS-A) estimator the NBM batches are longer (one-seventh of the sample rather than one-eighth), which should improve bias.

4. EXPERIMENT 2: EQUAL BATCH SIZES

Since equal batch sizes looked good in the last subsection, we now investigate using seven equal batches and then three equal batches. Since HW doesn't use batches in this sense, for HW we use seven and three degrees of freedom. Since $n=100$ for iid normal data provided asymptotic values in Experment 1, we consider only such small samples here. If indeed the estimated values are close to the asymptotic correlations, then the results apply more generally.

4.1. Estimated Covariances and Correlations

For seven batches the estimated covariance and correlation matrices for NBM, OBM, STS-A, and HW are

$$\begin{bmatrix} .33 & .20 & .00 & .15 \\ & .22 & .06 & .14 \\ & & .29 & .10 \\ & & & .25 \end{bmatrix} \qquad \begin{bmatrix} 1. & .77 & .00 & .51 \\ & 1. & .24 & .61 \\ & & 1. & .38 \\ & & & 1. \end{bmatrix}.$$

For three batches the estimated covariance and correlation matrices are

$$\begin{bmatrix} .97 & .50 & .00 & .28 \\ & .54 & .19 & .26 \\ & & .66 & .22 \\ & & & .51 \end{bmatrix} \qquad \begin{bmatrix} 1. & .68 & .00 & .38 \\ & 1. & .06 & .50 \\ & & 1. & .38 \\ & & & 1. \end{bmatrix}.$$

The only practically significant change from the first experiment is the expected drop to zero of the correlation between NBM and STS-A. The variances are no longer similar since the nominal degrees of freedom differ.

4.2. Optimal Linear Combinations

First a note about what is not important in this section. The important comparison is not between the variances obtained here with fixed batch size and the variances obtained in the last section using seven degrees of freedom. The issue is whether the idea of linear combinations provides an improvement compared to the set of component estimators used. Therefore, we compare the variances of the linear combination to the variances of the component estimators: .33, .22, .29, .25.

Table 3, which is analogous to Table 2, shows the performance of the various optimal linear combinations calculated from the estimated covariance matrix.

Table 3. Optimal Weights and Associated Variances for Linear Combinations of Pairs, Triples, and All Component Estimators. Batch Size is One-Seventh the Run Length.

NBM	OBM	STS-A	HW	Variance
.13	.87			.22
	.58		.42	.19
.36			.64	.21
		.44	.56	.18
.47		.53		.15
	.59	.41		.15
.05	.54		.41	.19
.36		.44	.20	.15
	.45	.35	.20	.15
.25	.31	.44		.15
.21	.26	.39	.14	.14

Table 4. Optimal Weights and Associated Variances for Linear Combinations of Pairs, Triples, and All Component Estimators. Batch Size is One-Third the Run Length.

NBM	OBM	STS-A	HW	Variance
.08	.92			.54
	.47		.53	.39
.25			.75	.45
		.40	.60	.40
.40		.60		.39
	.57	.43		.39
.02	.46		.53	.39
.25		.39	.36	.34
	.35	.29	.36	.34
.22	.32	.47		.37
.14	.21	.33	.32	.33

Once again we see that, although the linear combination of all four component estimators has the smallest variance, pairs can do quite well. In particular, (NBM, STS-A) is very good due to their independence and (OBM, STS-A) is equally good due to their small correlation and the small OBM variance. And again any triple containing STS-A works very well. The variance of .14 obtained by combining four component estimators is about 64 percent of the OBM variance, so the improvement is not as great as obtained with equal degrees of freedom. The variance of .14 corresponds to 14 degrees of freedom.

Now we investigate batches one-third the length of the run. Table 4, which is analogous to Tables 2 and 3, shows the performance of the various optimal linear combinations calculated from the estimated covariance matrix.

Here for the first time no pair of component estimators is close to the variance possible using three or four component estimators. The triples (NBM, STS-A, HW) and (OBM, STS-A, HW) have variance .34, which is almost as low as the .33 obtained using all four. The .33 is 65 percent of the variance of HW and .85 percent of the variance obtained from Schruben's (NBM, STS-A) estimator. The variance of .33 corresponds to 6 degrees of freedom.

5. DISCUSSION

We have empirically explored the use of linear combinations of known estimators of the variance of the sample mean. Three observations are obvious: (a) The asymptotic correlations between estimators of the variance of the sample mean do not appear to depend upon the underlying process. (b) The asymptotic correlations appear to hold more or less as soon as the number of degrees of freedom is low enough to provide asymptotic unbiasedness. (c) In each of the three cases, the best linear combinations produced about two degrees of freedom per batch or per degree of freedom in the component estimators.

If indeed the correlations among estimators don't vary much with sample size or process type, then the optimal weights can be computed once for any set of component estimators, thereby avoiding computational problems for users and the statistical issues related to using estimated weights.

Schruben's linear combination of NBM and STS-A deserves special mention. Except when the batch size of the component estimators was low (three), this combination did as almost as well as combinations involving three or all four estimators. The reason is their asymptotic independence, a property shared with no other pair of estimators studied here.

So what do we know about the future of using linear combinations of several component estimators, each with low degrees of freedom and therefore low bias? Can we somehow build a composite estimator with degrees of freedom more than twice what the component estimators offer? The empirical results here are pessimistic when we restrict the component estimators to have similar batch sizes or degrees of freedom.

For two reasons, however, it may make sense to use quite different batch sizes. One reason is that the component estimators could be of the same type but differ in batch size or degrees of freedom, as discussed in Song and Schmeiser (1987) for OBM. The second reason is that the batch size should reflect what is known about the bias properties of the component estimators. For example, STS-A requires larger batches than NBM or OBM because each batch needs to resemble Brownian motion, a stronger requirement than placed on a batch mean.

REFERENCES

Crane, M.A. and Iglehart, D.L. (1975). Simulating stable stochastic systems, III: Regenerative processes and discrete-event simulations. *Operations Research* **23**, 33-45.

Crane, M.A. and Lemoine, A.J. (1977). *An Introduction to the Regenerative Method for Simulation Analysis.* Springer-Verlag, Berlin.

Fishman, G.S. (1971). Estimating sample size in computer simulation experiments. *Management Science* **18**, 21-38.

Goldsman, D. (1984). On using standardized time series to analyze stochastic processes. Unpublished Ph.D. Dissertation, School of Operations Research and Industrial Engineering, Cornell University, Ithaca, New York.

Goldsman, D. and Schruben, L.W. (1984). Asymptotic properties of some confidence interval estimators for simulation output. *Management Science* **30**, 1217-1225.

Goldsman, D., Kang, K. and Sargent, R.G. (1986). Large and small sample comparisons of various variance estimators. In: *Proceedings of the 1986 Winter Simulation Conference* (J. Wilson, J. Henriksen and S. Roberts, eds.), 278-284.

Goldsman, D. and Meketon, M.S. (1985). A comparison of several variance estimators. Report J-85-12, School of Industrial and Systems Engineering, Georgia Institute of Technology, Atlanta, Georgia.

Heidelberger, P. and Welch, P.D. (1981). A spectral method for confidence interval generation and run length control in simulation. *Communications of the ACM* **24**, 233-245.

Law, A.M. and Kelton, W.D. (1982). Confidence intervals for steady-state simulations, II: A survey of sequential procedures, *Management Science* **28**, 550-562.

Law, A.M. and Kelton, W.D. (1984). Confidence intervals for steady-state simulations, I: A survey of fixed sample size procedures. *Operations Research* **32**, 1221-1239.

Meketon, M.S. (1980). The variance time curve: Theory, estimation and application. Unpublished Ph.D. Dissertation, School of Operations Research and Industrial Engineering, Cornell University, Ithaca, New York.

Meketon, M.S. and Schmeiser, B. (1984). Overlapping batch means: Something for nothing? In: *Proceedings of the 1984 Winter Simulation Conference,* (S. Sheppard, U. Pooch, and D. Pegden, eds.), 227-230.

Moran, P.A.P. (1975). The estimation of standard errors in Monte Carlo simulation experiments. *Biometrika* **62**, 1-4.

Schmeiser, B.W. (1982). Batch-size effects in the analysis of simulation output. *Operations Research* **30**, 556-568.

Schriber, T.J. and Andrews, R.W. (1984). ARMA-based confidence interval procedures for simulation output analysis. *Americal Journal of Mathematical and Management Science* **4**, 345-373.

Schruben, L.W. (1983). Confidence interval estimation using standardized time series. *Operations Research* **31**, 1090-1108.

Song, W.-M. T. and Schmeiser, B. (1987a). Dispersion matrices and optimal linear combinations of variance estimators of the sample mean. Research Memorandum 87-1, School of Industrial Engineering, Purdue University, West Lafayette, Indiana 47907.

Song, W.-M.T. and Schmeiser, B. (1987b). On the dispersion matrix of variance estimators of the sample mean in the analysis of simulation output.

Welch, P.D. (1987). On the relationship between batch means, overlapping batch means and spectral estimation. In: *Proceedings of the 1987 Winter Simulation Conference* (A. Thesen, H. Grant, and D. Kelton, eds.), to appear.

APPENDIX A: THE NBM ESTIMATOR

```
      subroutine nobm(x,n,m,df,vxbar)
c     Tina Song  may 22, 1987  purdue university
c     nonoverlapping batch mean estimator
c       of the variance of x bar
      real  x(n)
c......parameter definitions
c  ....input
c       n:     number of observations
c       x:     vector of observations
c  ....option
c       df:    approximate degrees of freedom
c                for vxbar (2.5 .lt. df)
c       m:     batch size
c  ....output
c       vxbar: estimated variance of the
c                sample mean
c......variable definitions
c       k:     number of batches (n/m of df+1)
c       i:     index for data points
c       j:     index for batches
c       sumx:  sum of data points 1 to n
c       xbar:  sample average
c       bsum:  sum of data points (j-1)m+1
c                to jm  (batch sum)
c       bsum2: square of the difference of the
c                batch mean and xbar
c       sum2:  sum of the square of the dif-
c                ference of the batch mean and
c                xbar (ie. sum of bsum2)
c.....calculate m or df or both
      if (m .gt. 0) then
         k=n/m
         df=float(k-1)
      elseif (df .gt. 0) then
         k=df+1.
         m=n/k
      else
         df=7.
         k=8
         m=n/k
      endif
c.....calculate the nobm estimator of
c       the variance of x bar
      sumx=0.
      do 10 i=1,n
 10      sumx=sumx+x(i)
         xbar=sumx/n
      sum2=0.
      do 40 j=1,k
         bsum=0.
         do 30 i=((j-1)*m)+1,j*m
 30         bsum=bsum+x(i)
            bsum2=((bsum/m)-xbar)**2
 40         sum2=sum2+bsum2
      vxbar=(sum2/(k-1))/(n/float(m))
      return
      end
```

APPENDIX B: THE OBM ESTIMATOR

```
      subroutine obm(x,n,m,df,vxbar)
c     bruce schmeiser  purdue university
c     overlapping batch-means estimator
c       of the variance of x bar
c     reference:  Meketon and Schmeiser,
c       1986 wsc proceedings
      real x(n)
c.....parameter definitions
c  ....input
c       n:     number of observations
c       x:     vector of observations
c  ....option
c       df:    approximate degrees of freedom
c                for vxbar (2.5 .lt. df)
c       m:     batch size
c  ....output
c       vxbar: estimated variance of
c                the sample mean
c.....variable definitions
c       i:     index for both data points
c                and batches
c       sumx:  sum of data points 1 to i
c       sumd:  sum of data points 1 to i-m
c       bsum:  sum of data points i-m+1 to i
c                (batch sum)
c       epsdf: accuracy for degrees of freedom
c       sum:   sum of the batch sums
c       sum2:  sum of the squared batch sums
c       xbar:  sample average
c       sumbm: sum of the batch means
c       sum2bm:sum of the squared batch means
c.....calculate m or df or both
      data epsdf/.1/
      if (m .gt. 0)  then
        if (m .eq. 1) then
          df = n-1
        else
          fk = float(n)/m
          df = 1.5 * (fk-1.)
     1           * (1. + (fk-1.)**(-.5-.6*fk))
        endif
      elseif (df .gt. 0.) then
        fk = df/1.5 + 1.
        do 5 j=1,100
          fk = (df/1.5 - (fk-1.)**(.5-.6*fk)) + 1.
          dftest = 1.5 * (fk-1.)
     1           * (1. + (fk-1.)**(-.5-.6*fk))
          if (abs(df-dftest).lt.epsdf/n)  go to 6
    5   continue
        stop 666
    6   m = n/fk + .5
      else
        df = 7.
        m = n/5.6666
      endif
c.......process the first m observations
c          -- the first complete batch
      sumx = 0.
      do 10 i=1,m
   10   sumx = sumx + x(i)
      sum = sumx
      sum2 = sum*sum
c.......process observations m+1 through n
      sumd = 0.
      do 20 i=m+1,n
        sumd = sumd + x(i-m)
        sumx = sumx+ x(i)
        bsum = sumx - sumd
        sum = sum + bsum
   20 sum2 = sum2 + bsum*bsum
c.......convert batch sums to batch means
      xbar = sumx / n
      sumbm = sum / m
      sum2bm = sum2 / (m*m)
c.......calculate the OLB estimator
      vxbar = ((sum2bm - xbar *
     1          (2*sumbm - (n-m+1)*xbar)))
     2        / (((n-m+1.)*(n-m))/m)
      return
      end
```

APPENDIX C: THE HW ESTIMATOR

```
      subroutine hwspec(x,n,fft,f0,dof,ier)
c
c***********************************************************************
c*                                                                     *
c*    Subroutine hwspec calculates the Heidelberger-Welch estimator .  *
c*    f0 for the spectrum at zero frequency corresponding to an input  *
c*    time series [x(j) : j = 1, ...., n] of length n.                 *
c*                                                                     *
c*    A quadratic polynomial is fitted to the first 25 points of       *
c*    log-averaged periodogram using ordinary least squares; then      *
c*    the intercept estimator is exponentiated and adjusted for the    *
c*    bias introduced by the exponential transformation in order to    *
c*    to yield the final estimator of the spectrum at zero frequency.  *
c*                                                                     *
c*    Reference: Heidelberger, P. and Welch, P. D. (1981). A spectral  *
c*    method for confidence interval generation and run length con-    *
c*    trol in simulations. Communications of the ACM, 24, 233-245.     *
c*                                                                     *
c*    This program was written by:                                     *
c*                                                                     *
c*                  James R. Wilson                                    *
c*                  School of Industrial Engineering                   *
c*                  Purdue University                                  *
c*                  West Lafayette, IN  47907, U.S.A.                  *
c*                  (317) 494-5408                                     *
c*                                                                     *
c***********************************************************************
c
c
c**** Inputs:
c
c         x -- The array containing the input time series.  Must consist
c              of at least 100 observations.  Note that x(.) is real and
c              must be dimensioned to n (at least) in the calling
c              program.
c
c         n -- The length of the input time series.  Although the
c              internal work arrays of hwspec are currently
c              configured to accomodate up to 10000 observations,
c              hwspec should be able to process much longer series
c              (up to n = 2**20) without modification.
c              However, if there is inadequate space in the internal
c              internal work arrays wk(.) and iwk(.) to handle a
c              particular series, then an error message is
c              written to the standard output device telling the user
c              how to redimension these arrays by resetting the
c              value of the symbolic constant iwkmax in the first
c              parameter statement below.
c
c
c**** Outputs:
c
c       fft -- The Fast Fourier Transform of the time series
c               [x(1), ..., x(n)].  Note that fft(.) is complex
c               and must be dimensioned to n (at least) in the
c               calling program.
c
c        f0 -- The Heidelberger-Welch estimator of the limit
c
c                            -
c                   lim  n*Var[ X(n) ],
c                 n->inf
c
c                    -
c              where X(n) is the sample mean of the series.
c              This is the spectrum evaluated at zero frequency.
c
c       dof -- The "effective" degrees of freedom associated
c               with the estimator f0.
c
c       ier -- Error indicator:
c              ier = 1 if n is less than 100.
c              ier = 2 if n is greater than 10000 AND there
c                      is inadequate space in wk(.) and iwk(.)
c                      to compute the required FFT.  An error
c                      message reports the required dimension
c                      for wk(.) and iwk(.) as specified by the
c                      symbolic constant iwkmax.
c
c
c**** Notes: This program requires the IMSL routines vdcps, fftcc,
c            becovm, and rlmul to run.
c
c            The following parameter statement defines the constants
c            of the procedure:
c
c
c
c       kpts    -- The number of points in the log-averaged
c                  periodogram that are used to fit a polynomial in
c                  the (averaged) Fourier frequencies.  Although the
c                  default value is 25, this constant can be reset
c                  to 50 as described by Heidelberger and Welch (1981).
c
c       ideg    -- The degree of the polynomial fitted to the log-
c                  averaged periodogram.  Although the default value is
c                  2, this constant can be set to 1, 2, or 3 as
c                  described by Heidelberger and Welch (1981).
c                  If either kpts or ideg is reset by the user, hwspec
c                  automatically resets the appropriate constants
c                  for computing f0 and dof.
c
c       iwkmax -- The length of the work storage arrays wk(.) and
c                  iwk(.) used in computing the FFT.  If error code 2
c                  occurs in the execution of hwspec, then the user
c                  should reset iwkmax to the required value and
c                  recompile hwspec.
c
      parameter (kpts=25, ideg=2, iwkmax=60150)
c
c***********************************************************************
c
      parameter (idp1=ideg+1, issm=(idp1*(idp1+1))/2)
      parameter (ivmax=(ideg*(ideg+1))/2, jc=kpts/25)
      parameter (iumax=2*kpts, adj1=0.270)
      complex fft(n)
      dimension x(n),y(kpts),g(kpts,idp1),phat(iumax),c1(3,2),c2(3,2)
      dimension wk(iwkmax),iwk(iwkmax),nbr(6), temp(idp1), vmeans(idp1)
      dimension gssm(issm), beta(idp1,7), varb(ivmax), anova(14)
      dimension ipf(13), iexp(13), ipwr(13)
      equivalence (y(1),g(1,idp1)), (wk(1),iwk(1))
      data c1/0.948,0.882,0.784,0.974,0.941,0.895/
      data c2/18.,7.,3.,37.,16.,8./
c
c**** Check for acceptable time series length...
c
      if (n .lt. 100) then
         ier = 1
         f0 = -1.0
         return
      endif
c
c**** Compute the space requirements for the FFT...
c
      call vdcps(n,npf,ipf,iexp,ipwr)
      m = 0
      k = 1
      jk = 1
      do 100 i = 1, npf
         m = m + iexp(i)
         k = max0(k,ipf(i))
         jk = jk*( ipf(i)**(iexp(i)/2) )
100   continue
      kb = ( n/(jk**2) ) - 2
      imax = 3*m + 3 + max0(4*m+7+6*k, kb+1+2*jk)
      if (imax .gt. iwkmax) then
         ier = 2
         print *, ' Inadequate work storage to compute the FFT'
         print *, ' Reset iwkmax = ',imax, ' and recompile hwspec'
         f0 = -1.0
         return
      endif
c
c**** Compute the FFT of the time series...
c
      do 200 i = 1, n
         fft(i) = cmplx(x(i))
200   continue
      call fftcc(fft,n,iwk,wk)
c
c**** Compute the corresponding periodogram...
c
      do 300 iu = 1, iumax
         phat(iu) = (cabs(fft(iu+1))**2)/n
300   continue
c
c**** Compute the log-averaged periodogram and the corresponding
c**** average of adjacent Fourier frequencies...
c
      do 400 iu = 1, kpts
         iu2 = 2*iu
         iu2m1 = iu2 - 1
         avfreq = ( float(iu2m1 + iu2) )/(2.0*n)
         avper = ( phat(iu2m1) + phat(iu2) )/2.
         y(iu) = alog(avper) + adj1
         do 350 jv = 1, ideg
            g(iu,jv) = avfreq**jv
350      continue
400   continue
c
c**** Perform a quadratic regression of the log-averaged periodogram
c     on the correpsonding averaged frequency to obtain the estimated
c     intercept...
c
      nbr(1) = idp1
      nbr(2) = kpts
      nbr(3) = kpts
      nbr(4) = 1
      nbr(5) = 1
      nbr(6) = 1
      call becovm(g,kpts,nbr,temp,vmeans,gssm,ier1)
      alpha = 0.05
      call rlmul (gssm,vmeans,kpts,ideg,alpha,anova,beta,idp1,varb,
     $            ier)
c
c**** Exponentiate to untransform the intercept estimate, then
c     adjust for the bias introduced by the exponential
c     transformation...
c
      f0 = c1(ideg,jc)*exp(beta(idp1,1))
      dof = c2(ideg,jc)
      return
      end
```

AUTHOR'S BIOGRAPHIES

Bruce Schmeiser is a professor in the School of Industrial Engineering at Purdue University. He received his undergraduate degree in mathematical sciences and master's degree in industrial and management engineering at The University of Iowa. His Ph.D. is from the School of Industrial and Systems Engineering at the Georgia Institute of Technology. He is the *Operations Research* area editor in simulation and has served in editorial positions of *IIE Transactions, Communications in Statistics, B: Simulation and Computation, Journal of Quality Technology, American Journal of Mathematical and Management Sciences,* and the *Handbook of Industrial Engineering.* He is the past chairman of the TIMS College on Simulation and Gaming. He represents ORSA on the Winter Simulation Conference Board of Directors, currently serving as vice-chairman. His research interests are the probabilistic and statistical aspects of digital-computer stochastic simulation, including input modeling, random-variate generation, output analysis, and variance reduction.

Bruce Schmeiser
School of Industrial Engineering
Purdue University
West Lafayette, IN 47907, U.S.A.
(317) 494-5422
schmeise@gb.ecn.purdue.edu (arpanet)

Wheyming Tina Song is a Ph.D. candidate in the School of Industrial Engineering at Purdue University, majoring in simulation and applied probability. She received her undergraduate degree in statistics and master's degree in industrial management at Cheng-Kung University in Taiwan. She then received her master's degrees in applied mathematics and industrial engineering from the University of Pittsburgh.

Wheyming Tina Song
School of Industrial Engineering
Purdue University
West Lafayette, IN 47907, U.S.A.
(317) 494-5170
wheyming@gb.ecn.purdue.edu (arpanet)