A SPECTRAL BASED TECHNIQUE FOR GENERATING CONFIDENCE INTERVALS FROM SIMULATION OUTPUTS

T. L. Moeller and P. D. Welch

IBM Thomas J. Watson Research Center

Yorktown Heights, New York 10549

ABSTRACT

A technique for generating confidence intervals on the common expectation of a sequence of correlated random variables is developed. The sequence is modelled as a covariance stationary process. In this situation the variance of the sample mean is proportional to the variance spectrum at zero frequency. This value of the spectrum is estimated by fitting a low order polynomial to the sample spectrum (periodogram) in the lower frequency region. The technique is applicable to both individual observations and batched data. Experimental results comparing it with the method of batch means are given for the steady state waiting time of the M/M/1 queue. The proposed technique gives valid confidence intervals of approximately the same average width as the method of batch means when the batch size is large enough for that method to be valid. It continues to give valid confidence intervals when the batch sizes are such that the method of batch means breaks down.

1. INTRODUCTION

We are interested in the output analysis of simulations of generalized queueing networks or job shop like systems. In particular we are interested in the problem of placing valid confidence intervals on the steady state parameters of such systems. In these situations the primary outputs are sequences of random variables. In developing procedures for generating valid confidence intervals one has to deal with two complicating phenomena:

1) there is a transient phase wherein the parameters of interest do not approximate their values in the steady state;

2) within a sequence, the primary random variables are correlated.

We will be proposing a method for dealing with the second of these complications. We will not discuss the first.

A number of procedures have been proposed for dealing with the correlated nature of the primary random variables. These procedures fall into two general categories. In the first, by some organization of the experimental protocol, independent or approximately independent random variables are generated and standard statistical techniques applicable to independent observations are applied. The methods of independent trials, batch means and regeneration points fall in this category. For discussions of these methods see [1] through [4]. The methods of this category are those in most common use. In the second category the methods attempt to estimate the effect of the correlation by some direct means. The most developed of these methods fits an autoregressive model to the sample sequence and generates confidence limits assuming the correlation structure of the estimated model. For a discussion of this see [1] and [5]. The application of spectral estimation techniques has been discussed, see [1] and [2], but does not seem to have been studied in any intensive way. The exploitation of spectral techniques is the subject of this paper.

We will view the simulation as yielding a sequence of random variables $X_1, \ldots, X_n$ with a common expectation

$$E(X_i) = \mu \tag{1}$$

which is unknown. We wish to generate a confidence interval for μ. We will assume that it is reasonable to model $X_i : i = 1, \ldots, N$ as a sample from a covariance stationary sequence with covariance function

$$\gamma(k) = E\{(X_i - \mu)(X_{i+k} - \mu)\}. \tag{2}$$

We let σ^2 be the common variance of the X_i's. Then

$$\sigma^2 = \mathrm{Var}\{X_i\} = \gamma(0). \tag{3}$$

Corresponding to $\gamma(k)$ we assume there is a spectral density $p(f)$. The sequence $\gamma(k)$, $k=0, \pm 1, \pm 2$ and the function $p(f)$ are a Fourier pair in the sense that

$$p(f) = \sum_{k=-\infty}^{\infty} \gamma(k) e^{-2\pi i f k} \tag{4}$$

and

$$\gamma(k) = \int_{-1/2}^{1/2} p(f)\, e^{2\pi i f k}\, df \tag{5}$$

In practice the sequence $X(j)$ could either be some primary sequence of random variables or a sequence derived from some primary sequence. As examples consider the following.

Example 1: Suppose we have a queueing network and we are interested in the steady state waiting time at the jth queue. Let

$W_{j,n}$ = waiting time of the nth customer departing from the jth queue.

And suppose our simulation had generated $W_{j,n}$ $n=1, \ldots, M$. Then we could let

$$X_i = W_{j,i}$$

and $N = M$

or we could batch the observations in batches of size B and, assuming $M=KB$, let

$$X_i = \frac{1}{B} \sum_{n=1}^{B} W_{j,\,n+(i-1)B}$$

$$N = K.$$

Example 2: Consider the same queueing network and the same simulation output sequence, $W_{j,n}$ $n=1, \ldots, N$, discussed in Example 1 and suppose we are interested in the steady state probability that the waiting time at the jth queue is greater than C. Then we consider the sequence of random variables

$$\xi_{j,n} = 1 \quad \text{if } W_{j,n} > C$$
$$= 0 \quad \text{otherwise.}$$

Notice that

$$E\{\xi_{j,n}\} = \mathrm{Prob}\,\{W_{j,n} > C\}.$$

Again we could let

$$X_i = \xi_{j,i}$$

and $N = M$

or we could batch the observations and, assuming $M=KB$, let

$$X_i = \frac{1}{B} \sum_{n=1}^{B} \xi_{j,n+(i-1)B}$$

and $N = K$.

Example 3: Again consider the same queueing network of the earlier examples and suppose we are interested in the utilization of the jth queue. Then we consider the continuous process

$\nu_j(t) = 1$ if the jth server is busy at time t

$\nu_j(t) = 0$ otherwise.

Notice that

$$E\{\nu_j(t)\} = \text{Prob}\{\text{ jth server is busy at time t }\}$$

$$\approx \text{steady state utilization of the jth queue.}$$

In this case if T were the overall simulated time in approximate steady state we would choose a number of time intervals, K, where $KT_B=T$ and let

$$X_i = \int_{(i-1)T_B}^{i\,T_B} \nu_j(t)\,dt$$

and $N = K$.

2. THE METHOD

As discussed in Section 1 we assume we have a sample X_i : $i=1,\ldots,N$ from a covariance stationary process with covariance function $\gamma(k)$ and spectral density $p(f)$.

We let

$$\mu = E(X_i) \tag{6}$$

be the unknown, common expectation of the X_i. We wish to generate a confidence interval for μ from the $X_i : i = 1,\ldots,N$.

Now in this situation a reasonable estimator to use is the sample mean,

$$\hat{\mu} = \frac{1}{N}\sum_{i=1}^{N} X_i . \tag{7}$$

It is an unbiased estimator of μ whose variance is given by

$$\sigma^2_{\hat{\mu}} = \frac{1}{N}\sum_{k=-(N-1)}^{N-1} \frac{(N-k)}{N}\gamma(k) \tag{8}$$

$$= \frac{1}{N}\int_{-1/2}^{1/2} \frac{\sin^2 N\pi f}{N\sin^2 \pi f}\,p(f)\,df \tag{9}$$

The results (8) and (9) follow from standard theory on the variance and the spectral density of linear operators such as $\hat{\mu}$ applied to covariance stationary sequences (see e.g. [6]). For large N, (8) and (9) are given approximately by

$$\sigma^2_{\hat{\mu}} \approx \frac{1}{N}\sum_{k=-\infty}^{\infty}\gamma(k) \tag{10}$$

$$\approx \frac{p(0)}{N}. \tag{11}$$

We will use these approximations from here on.

To place a confidence interval on μ we generate an estimate $\hat{\sigma}^2_{\hat{\mu}}$ of the variance of $\hat{\mu}$ and assume that $(\mu-\hat{\mu})/\hat{\sigma}_{\hat{\mu}}$ is distributed as a normal random variable with zero mean and unit variance. A confidence interval at confidence level α is given as

$$(\hat{\mu} - \Phi(1-\alpha/2)\,\hat{\sigma}_{\hat{\mu}}\,,\ \hat{\mu} + \Phi(1-\alpha/2)\,\hat{\sigma}_{\hat{\mu}}) \tag{12}$$

where $\Phi(x)$ is the cumulative distribution function of the normal distribution with zero mean and unit variance.

As we remarked in the introduction we will proceed by a quite direct application of the theory of spectral estimation. Our estimate of $\hat{\sigma}^2_{\hat{\mu}}$ will be obtained as

$$\hat{\sigma}^2_{\hat{\mu}} = \frac{\hat{p}(0)}{N} \tag{13}$$

where $\hat{p}(0)$ is generated by applying polynomial regression to the logarithim of the periodogram. The main assumption behind the method is that $\log(p(f))$ be a smooth function in the sense that it can be approximated by a low order polynomial over a substantial frequency range which includes zero.

The periodogram of the sequence $X(j)$, $j=1,\ldots,N$ is defined as

$$I(n/N) = \frac{\left|\sum_{j=1}^{N} X(j)\,e^{2\pi i(j-1)n/N}\right|^2}{N} \tag{14}$$

$$n = 0, 1, \ldots, [N/2]$$

where

$$[N/2] = \text{largest integer} \leq N/2$$

$$\text{and } i = (-1)^{1/2} .$$

For a wide class of covariance stationary processes (see e.g. [7]) which include all Gaussian processes we have asymptotically (as N gets large) that for n, m=1 ,..., [N/2]

$$E\{I(n/N)\} = p(n/N) , \tag{15}$$

$$\text{Var}\{I(n/N)\} = p^2(n/N) , \tag{16}$$

$$\text{and Cov}\{I(n/N), I(m/N)\} = 0 , \quad n \neq m . \tag{17}$$

Further, again asymptotically, the I(n/N) are distributed as a constant times a χ^2 random variable with 2 degrees of freedom. If we assume that p(f) is smooth in the lower frequency region then Equations (15) and (17) suggest that one could apply polynomial regression to estimate p(f) and take $\hat{p}(0)$ as the y axis intercept of the estimated $\hat{p}(f)$. However as Equation (16) indicates, if p(f) is varying, the variance is not constant. Further the χ_2^2 distribution is quite skewed. Because of this it is more common to consider log(I(n/N)) : n=1 ,..., [N/2].

If we let

$$C = E\{\log(\chi_2^2/2)\} = .577 \tag{18}$$

then

$$E\{\log(I(n/N)) - C\} = \log(p(n/N)) , \tag{19}$$

$$\text{Var}\{\log(I(n/N)) - C\} = \text{Var}\{\chi_2^2\} , \tag{20}$$

$$n = 1, \ldots, [N/2] ,$$

and we see that log(I(n/N))−C has log(p(n/N)) as expectation and a constant variance. Furthermore the distribution of log (χ_2^2) is more nearly normal than the distribution of χ_2^2. Hence log(I(n/N))−C is a natural candidate for the application of regression techniques.

With this background the proposed procedure is as follows:

1) calculate I(n/N) for n=1, 2 ,..., K where K = minimum ([N/2] , 100);
2) calculate f(n) = log{I(n/N)−C} for n=1, 2 ,..., K ;
3) fit a second degree polynomial $g(n) = a_o+a_1n+a_2n^2$ to f(n) using standard regression (least squares) techniques;
4) let the fitted polynomial be $\hat{g}(n) = \hat{a}_o + \hat{a}_1n+\hat{a}_2n^2$; we use as our estimate, $\hat{p}(0) = e^{\hat{g}(0)} = e^{\hat{a}_0}$;
5) finally $\hat{\sigma}_{\hat{\mu}} = (\hat{p}(0)/N)^{1/2}$ and the confidence interval is given by (12).

The major assumption of this approach is that log(p(f)) is a smooth function which can be approximated by a second degree polynomial over the range 0 to K/N. The value I(0) is not used because it only reflects the unknown mean. Fast Fourier algorithms can be used to calculate I(n/N). For a discussion of the relationship between the periodogram and the FFT algorithm see [8].

3. EXPERIMENTAL RESULTS

As a test of the procedure we simulated the M/M/1 queue and estimated the steady state waiting time. The X_i were averages over batches as discussed in Example 1 of Section 1. Let

$$W_n = \text{waiting time of the n'th departing customer} . \tag{22}$$

We are interested in

$$\mu = \lim_{n\to\infty} E\{W_n\}. \qquad (23)$$

We suppose we have W_n, n=1 ,..., M = NB and that

$$X_i = \frac{1}{B} \sum_{n=1}^{B} W_{n+(i-1)B} . \qquad (24)$$

Hence the random sequence considered; X_i : i=1 ,..., N; is a sequence of N batch means with batch size B and

$$E(X_i) \approx \mu . \qquad (25)$$

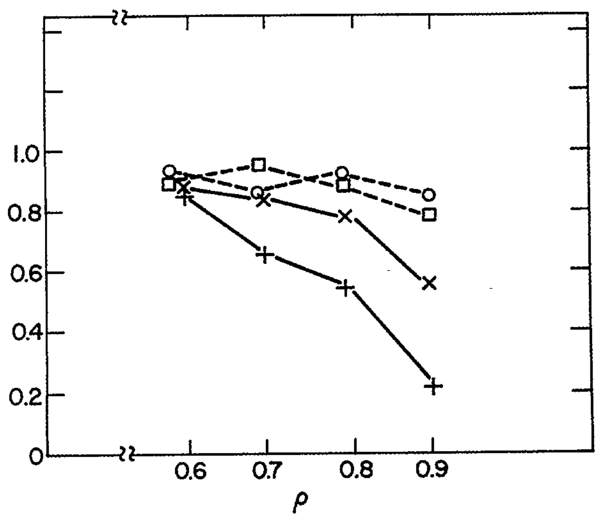
We compare the proposed method with the "method of batch means". In the method of batch means one assumes the batch size B to be large enough to make the means approximately uncorrelated. We first consider a sequence of experiments where the method of batch means breaks down because of correlation between the batches while the proposed method continues to provide valid confidence intervals.

Fifty independent simulations of the M/M/1 queue were run for the eight different conditions given in Table 1. The quantity ρ is the utilization. The number of service completions, M, was adjusted so that, for each ρ, it would be equal to 1000 x E{number of customers served in a busy period}. For each of the 50 independent simulations 90% confidence intervals were calculated using both the spectral method and the method of batch means. In Figure 1 the observed coverages are plotted. It can be seen that the method of batch means breaks down as ρ increases. As would be expected, it breaks down more drastically for the batch size of 20 than for the batch size of 100. The spectral method provides approximately 90% coverage in every instance.

ρ	B	N	M
.6	20	375	7500
.6	100	75	7500
.7	20	500	10000
.7	100	100	10000
.8	20	750	15000
.8	100	150	15000
.9	20	1500	30000
.9	100	300	30000

Table 1

Experimental Values for the Results of Figure 1



+ batch means: 20 events/batch
× batch means: 100 events/batch
○ spectral method: 20 events/batch
□ spectral method: 100 events/batch

Figure 1: Coverage for Spectral Method and "Batch Means" method for Batch Sizes of 20 and 100 as a function of ρ.

As an illustration of the basic function underlying the spectral method we have in Figure 2 plotted a particular instance of I(n/N), log(I(n/N)) and the quadratic fit. This is for the case ρ=.9, B=20, N=1500, and M=30,000. Notice the variance stabilizing effect of the logarithmic transformation and the uncorrelated appearance of the periodogram fluctuations.

As a second comparison of the two methods a sequence of conditions were considered with batch size large enough so that the method of batch means provided valid confidence intervals over the full range of ρ. We wish to show that the spectral method also provides proper coverage under these conditions and compare the average widths of the confidence intervals generated by the two methods. The conditions are tabulated in Table 2. The batch sizes are 50 and 100 times the expected number of service completions in a busy period. In Figure 3 we compare the coverage of 90% confidence intervals generated by the two methods. This coverage is based on 50 independent simulations under each set of conditions. Notice that both methods provide proper coverage. In Figure 4 we plot R(ρ) defined as

$$R(\rho) = \frac{\text{average confidence interval width (spectral method)}}{\text{average confidence interval width (batch means)}} .$$

Notice that R(ρ) fluctuates in the region 1.0 to 1.4. Hence there is a slight increase in the average width of the confidence intervals generated by the spectral method compared to those generated by the method of batch means when the method of batch means works properly.

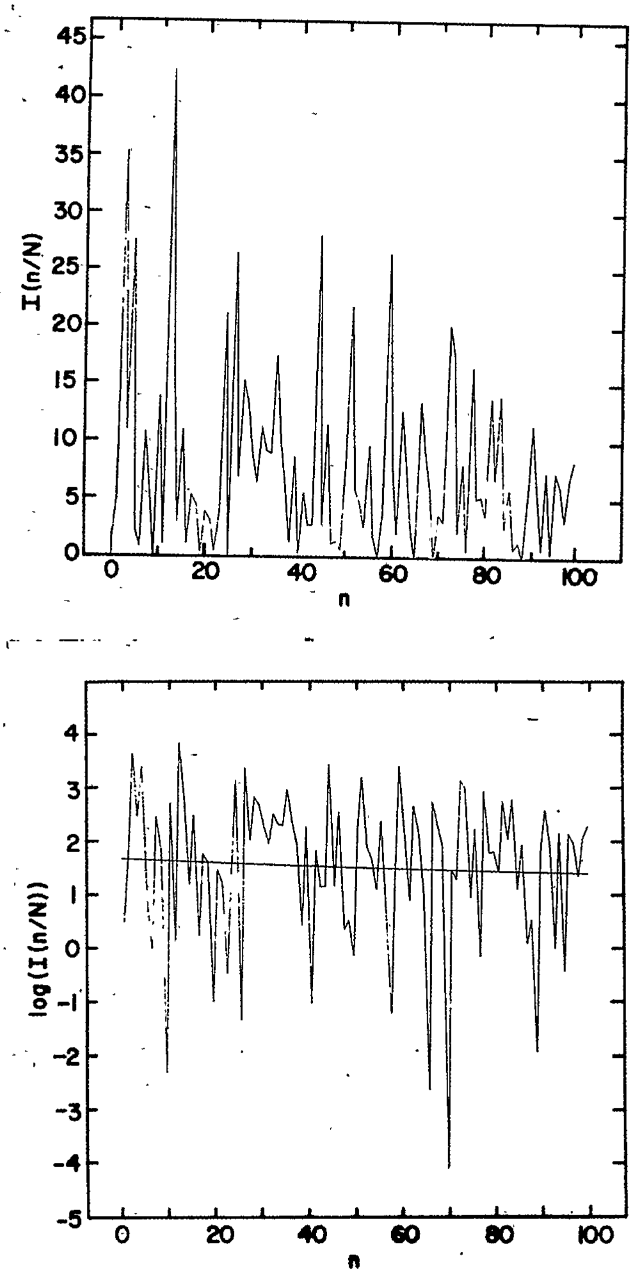
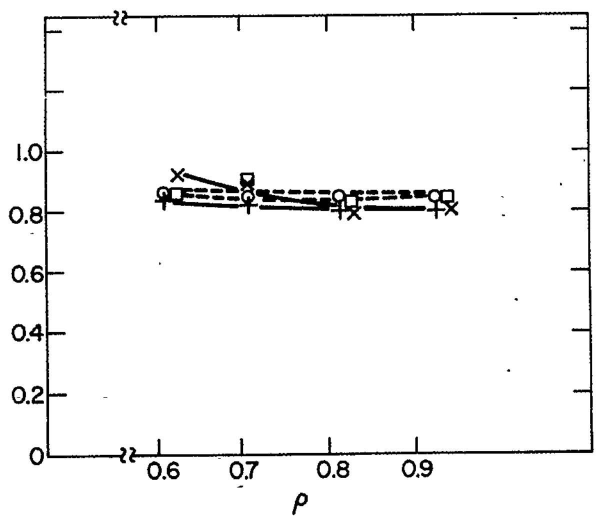


Figure 2: An Example of I(n/N), log(I (n/N)) and Quadratic Fit for ρ = .9, B = 20, N = 1500.

ρ	B	N	M
.6	125	60	7500
.6	250	60	15000
.7	167	60	10020
.7	334	60	20040
.8	250	60	15000
.8	500	60	30000
.9	500	60	30000
.9	1000	60	60000

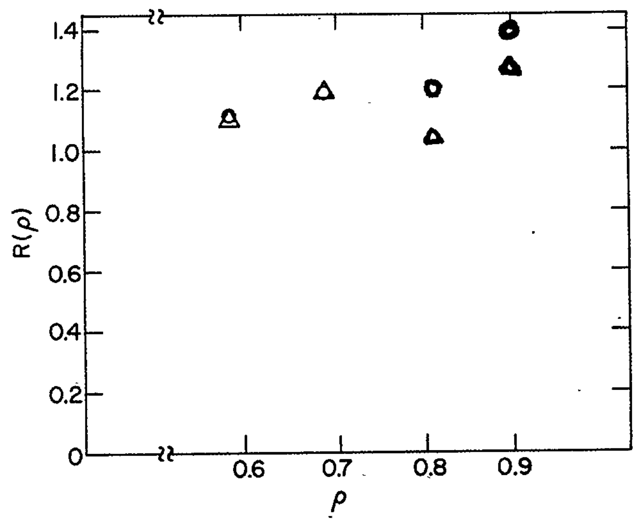
Table 2

Experimental Values for the Results of Figures 3 and 4



+ batch means: 50 E{busy p.}/batch
× batch means: 100 E{busy p.}/batch
o spectral method: 50 E{busy p.}/batch
□ spectral method: 100 E{busy p.}/batch

Figure 3: Coverage for Spectral Method and "Batch Means" method for Batch Sizes of 50 and 100 times E{number of service completions in a busy period} as a Function of ρ.



o 50 E{busy period}/batch
Δ 100 E{busy period}/batch

Figure 4: R(ρ) vs. ρ.

4. SUMMARY

We have outlined a method for placing confidence intervals on the common expectation of a sequence of observations which can be reasonably modelled as a sample sequence from a covariance stationary process. The main assumption underlying the method is that log(p(f)), where p(f) is the spectral density, be a smooth function which can be approximated by a low order polynomial. We have shown the method to work properly on simulations of the M/M/1 queue. Evidence so far indicates that it provides a method of generating confidence intervals from batched data whose validity does not depend upon the batch size. Experimentation is continuing on more complex systems and alternative methods of fitting smooth curves to the logarithm of the periodogram and to the periodogram itself.

5. REFERENCES

[1] Fishman, G. S., "Concepts and Methods in Discrete Event Digital Simulation", Wiley, N.Y., 1973.

[2] Mihram, G. A., "Simulation: Statistical Foundations and Methodology", Academic Press, N.Y., 1972.

[3] Law, A. M., "Confidence Intervals in Discrete Event Simulation: A State-of-the-Art Survey", Technical Report No. 77-14, Dept. of Ind. Eng., U. of Wisconsin, Madison.

[4] Iglehart, D. L., "The Regenerative Method for Simulation Analysis", Current Trends in Programming Methodology, Vol III, Software Modelling and Its Impact on Performance, Prentice-Hall, Englewood Cliffs, N.J., 1977.

[5] Fishman, G. S., "Estimating Sample Size in Computing Simulation Experiments", Management Science, Vol. 18, No. 1, Sept. 1971.

[6] Jenkins, G. M. and Watts, D. G., "Spectral Analysis and its Applications", Holden Day, San Francisco, 1968.

[7] Olshen, R. A., "Asymptotic Properties of the Periodogram of a Discrete Stationary Process", J. Appl. Prob., Vol. 4, 508, 1967.

[8] Cooley, J. W., Lewis, P. A. W. and Welch, P. D., "The Application of the Fast Fourier Transform Algorithm to the Estimation of Spectra and Cross Spectra", J. Sound Vib., Vol. 12, No. 3, 1970.