NASA CONTRACTOR REPORT

NASA CR-226

MEAN AND MEAN SQUARE MEASUREMENTS OF NONSTATIONARY RANDOM PROCESSES

by George P. Thrall and Julius S. Bendat

Prepared under Contract No. NAS 5-4590 *by*
MEASUREMENT ANALYSIS CORPORATION
Los Angeles, Calif.
for

NATIONAL AERONAUTICS AND SPACE ADMINISTRATION • WASHINGTON, D. C. • MAY 1965

NASA CR-226

MEAN AND MEAN SQUARE MEASUREMENTS OF NONSTATIONARY RANDOM PROCESSES

By George P. Thrall and Julius S. Bendat

Prepared under Contract No. NAS 5-4590 by
MEASUREMENT ANALYSIS CORPORATION
Los Angeles, Calif.

for

NATIONAL AERONAUTICS AND SPACE ADMINISTRATION

For sale by the Clearinghouse for Federal Scientific and Technical Information
Springfield, Virginia 22151 – Price $2.00

MEAN AND MEAN SQUARE MEASUREMENTS
OF NONSTATIONARY RANDOM PROCESSES

CONTENTS

MEAN AND MEAN SQUARE MEASUREMENTS
OF NONSTATIONARY RANDOM PROCESSES

1. INTRODUCTION

A random process $\{x_i(t)\}$, $i = 1, 2, 3, \ldots,$ is an ensemble of functions of a single variable t which can be characterized through its statistical properties. A typical random process is pictured in Figure 1. The variable t is time or any other parameter of interest.

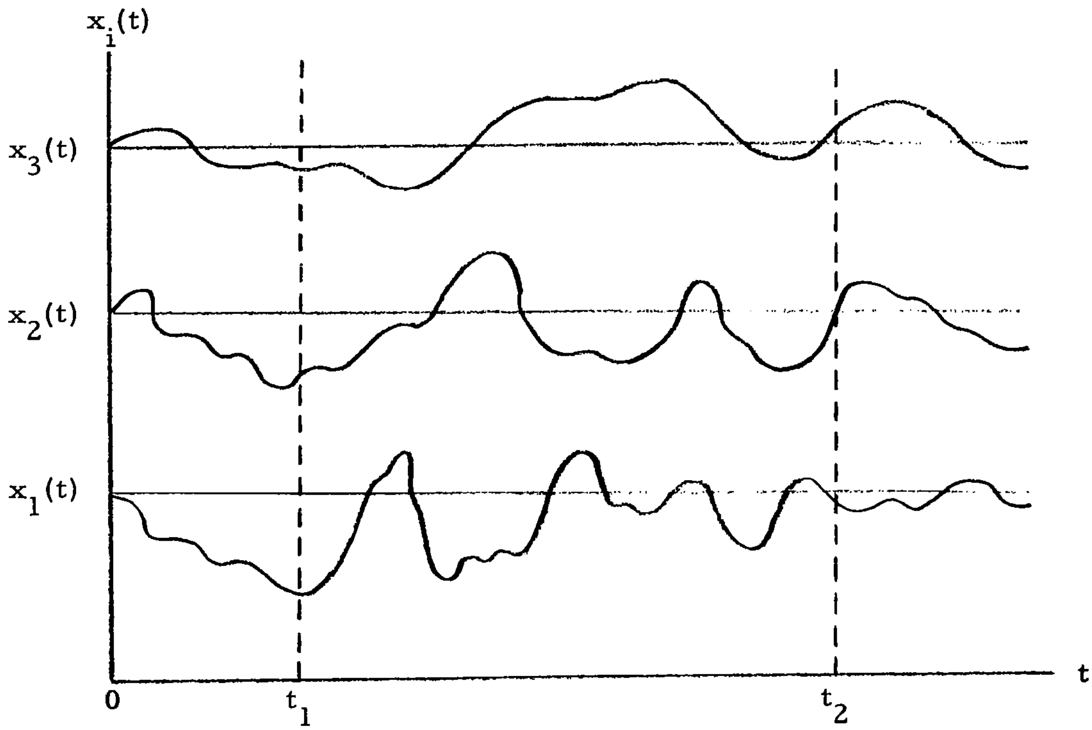


Figure 1. Random Process

The set of amplitude values at a particular time t_1 , denoted by $\{x_i(t_1)\}$, can be combined together in appropriate ways to determine their mean value, mean square value, and higher moments, and thus a complete probability distribution at t_1 , where the probability distribution yields the probability that the amplitude values at t_1 will lie in any specified amplitude range.

For example, the mean value at t_1 is defined by the ensemble average, denoted by the expected value $E[\ ]$, namely,

$$\mu_x(t) = E\left[x(t_1)\right] = \lim_{N\to\infty} \frac{1}{N} \sum_{i=1}^{N} x_i(t_1) \tag{1}$$

A different time $t = t_2$ can be selected, and similar statistical calculations may be carried out for the set of amplitude values $\left\{x_i(t_2)\right\}$, $i = 1, 2, \ldots$ In general, significantly different results would be obtained for the two sets $\left\{x_i(t_1)\right\}$ and $\left\{x_i(t_2)\right\}$. That is to say, the statistical results would not be invariant with respect to translations in time. Random processes of this general category are known as nonstationary random processes. The processes are said to be stationary when statistical properties do not change with time. Much past analytical work assumed a stationary hypothesis because it simplified the further derivations. This report is concerned with methods for analyzing nonstationary data.

Nonstationary data are of common occurrence in different physical situations. Such data are obtained when an environment changes suddenly, as in transient operating conditions, or when properties of a system are altered to such a degree that its response is changed, as under fatigue effects or through adaptive mechanisms.

In Section 2, techniques for estimating nonstationary mean values are presented using methods of ensemble averaging and orthogonal function approximations. The assumptions are made that the time t in each record $x_i(t)$ is measured from a well-defined origin and that the set of records $\left\{x_i(t)\right\}$ are statistically independent. Expressions are developed for the expected value of each estimate as well as the corresponding mean square estimation error. Further, it is shown that the orthogonal function approximation technique will produce smaller errors than ensemble averaging if a bias error can be made small.

Section 3 considers the corresponding estimation problem for nonstationary mean square values. Three techniques are analyzed: ensemble averaging, orthogonal approximation, and short time averaging. Two examples are given of the application of the orthogonal approximation technique to the estimation of nonstationary mean square values; and a comparison is made between these results and ensemble averaging. For short time averaging techniques, a special type of nonstationary random process is considered in some detail consisting of a time varying amplification of a stationary process. Mathematical formulas are obtained which yield the bias error and an upper bound on the mean square measurement error. Because of the complexity of the formulas, numerical results are difficult to obtain without computers.

In Section 4, recommendations are given for a computer based simulation of the various estimation procedures for typical nonstationary random processes of interest. It is indicated that the results of the simulation program would lead to a rational basis for optimally choosing an estimation procedure in many applications.

2. MEAN VALUE MEASUREMENTS

2.1 ENSEMBLE AVERAGING

For nonstationary data, a basic statistical problem is to determine how the average (mean) value changes with time. Mean values can be estimated by using an average response computer that performs the following operation to calculate a sample mean value from a sample of size N, namely for N records $\{x_i(t)\ ;\ 0 \le t \le T\ ;\ i = 1, 2, \ldots, N\}$ from a nonstationary process $x(t)$, compute

$$m(t) = \frac{1}{N} \sum_{i=1}^{N} x_i(t) \tag{2}$$

The quantity $m(t)$ will differ over different choices of the N samples $\{x_i(t)\}$. Consequently one must investigate how closely an arbitrary measurement, $m(t)$, approximates the true mean value $\mu(t)$ which is given by the expected value

$$\mu(t) = E\left[m(t)\right] = \frac{1}{N} \sum_{i=1}^{N} E\left[x_i(t)\right] = \mu_x(t) \tag{3}$$

Note that $m(t)$ is an unbiased estimate of the true mean value since $\mu(t) = \mu_x(t)$.

A measure of the error involved in estimating $\mu(t)$ by $m(t)$ is the variance of $m(t)$ given by

$$\sigma_m^2(t) = E\left[m(t) - \mu(t)\right]^2 \tag{4}$$

The square root of the variance, $\sigma_m(t)$, called the "standard deviation," provides the indicator that determines how closely a set of measurements of $m(t)$ clusters about its mean value $\mu(t)$.

In most practical applications, the N samples used to compute m(t) are statistically independent and this will be assumed here. Upon expanding Eq. (4), it is seen that

$$\sigma_m^2(t) = E\left[\frac{1}{N}\sum_{i=1}^{N}\left[x_i(t) - \mu(t)\right]^2\right]$$

$$= \frac{1}{N^2}\sum_{i,j=1}^{N} E\left[\left[(x_i(t) - \mu(t)\right]\left[x_j(t) - \mu(t)\right]\right]$$

$$= \frac{1}{N^2}\sum_{i=1}^{N}\sigma_x^2(t)$$

$$= \frac{\sigma_x^2(t)}{N} \tag{5}$$

where $\sigma_x^2(t)$ is the variance associated with the nonstationary process x(t). It should be noted that the independent sample assumption causes all the cross-product terms $(i \neq j)$ in the double sum of Eq. (5) to be zero.

As was shown in Ref. [1], a knowledge of the mean value and variance for the random variable m(t) at any time t enables one to answer questions concerning the range of the results at any time t without knowing the exact probability distribution function for m(t). From the Tchebycheff inequality, which applies to arbitrary general situations, one may state with 89% confidence, for example, that an observed measurement for m(t) lies inside the range $\left[\mu(t) - 3\sigma_m(t),\ \mu(t) + 3\sigma_m(t)\right]$. In equation form, for any constant k, the Tchebycheff inequality is

$$\mathrm{Prob}\left[\left|m(t) - \mu(t)\right| \geqq k\sigma_m(t)\right] < \frac{1}{k^2} \tag{6}$$

Thus, for $k = 3$, this probability is at most $(1:9)$, giving the above 89% confidence limits. See Figure 2.

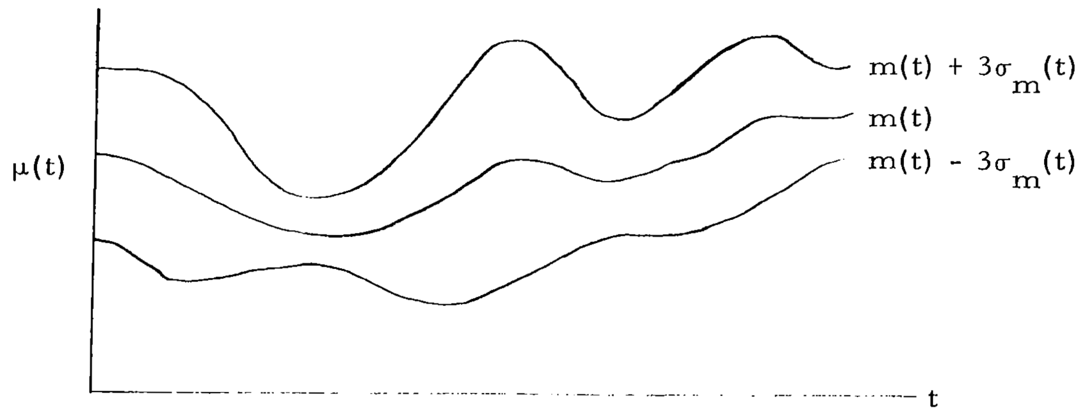


Figure 2. 89% Confidence Limits for Arbitrary Distribution based upon Tchebycheff Inequality

A stronger statement can be made if one can justify an assumption that $m(t)$ follows a normal (Gaussian) distribution at any value of t. For this special case, a 95% confidence band is given by the range $\left[\mu(t) - 2\sigma_m(t),\ \mu(t) + 2\sigma_m(t)\right]$. Thus an observed measurement for $m(t)$ in the Gaussian case yields a greater confidence of being close to the theoretical mean value than in the case where the underlying probability distribution is unknown.

Consider the range for $\mu(t)$ as given by Eq. (6), namely

$$|m(t) - \mu(t)| \leqq k\sigma_m(t) \tag{7}$$

where k is a constant.

Solving for $\mu(t)$ yields the two extreme range values:

$$\mu(t) = m(t) \pm k\sigma_m(t) \tag{8}$$

Thus, one would measure $m(t)$ and then apply the above formula to estimate $\mu(t)$ for various conditions. Also, note from Eq. (5) that by increasing N, one may guarantee that $m(t)$ will fall close to $\mu(t)$ regardless of the magnitude of $\sigma_x(t)$ and the underlying distribution. This effect is illustrated by the curves shown in Figure 3 which are plots of the ratio $\mu(t)/m(t)$ as a function of $\left(\sqrt{N}\ \mu(t)/\sigma_x(t)\right)$. Two cases are considered. Case 1 applied to arbitrary probability distributions and sets $k = 3$, corresponding to an 89% confidence band as given by the Tchebycheff inequality. Case 2 applies to a Gaussian probability distribution and sets $k = 2$, corresponding to a 95% confidence band. The lower and upper limits used for Figure 3 are shown in Table 1.

$\frac{\sqrt{N}\ \mu(t)}{\sigma_x(t)}$	Case 1		Case 2	
	Lower Limit	Upper Limit	Lower Limit	Upper Limit
2	--	--	0.50	∞
3	0.50	∞	0.60	3.00
4	0.57	4.00	0.67	2.00
5	0.63	2.50	0.71	1.67
6	0.67	2.00	0.75	1.50
8	0.73	1.60	0.80	1.33
10	0.77	1.43	0.83	1.25
12	0.80	1.33	0.86	1.20
15	0.83	1.25	0.88	1.15
18	0.86	1.20	0.90	1.12
20	0.87	1.18	0.91	1.11
25	0.89	1.14	0.93	1.09

Table 1. Data for Figure 3

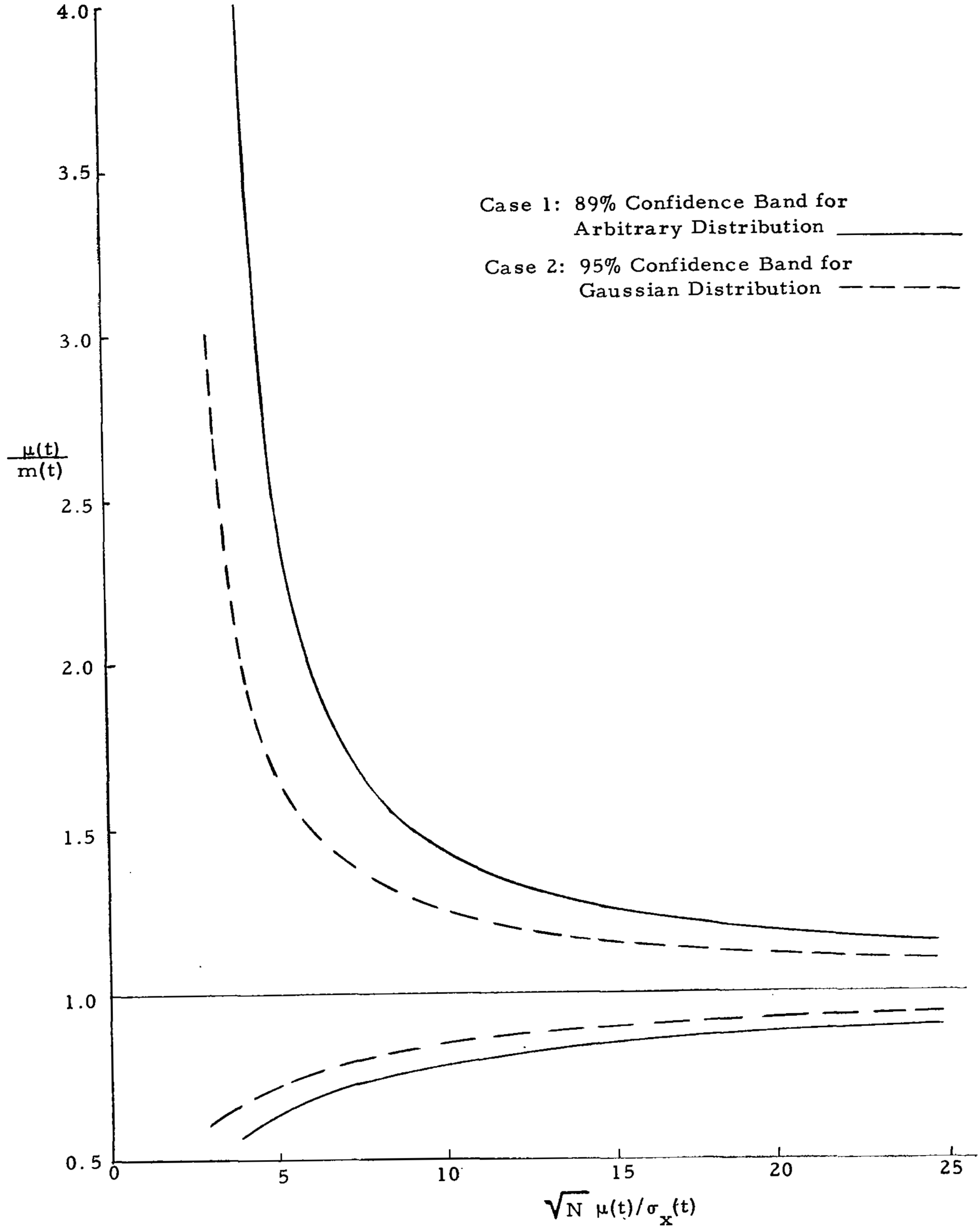


Figure 3. Confidence Bands for $\mu(t)/m(t)$ as function of $\sqrt{N}\ \mu(t)/\sigma_x(t)$

2.2 ORTHOGONAL FUNCTION APPROXIMATION

The second approach is based upon fitting an Kth order orthogonal function to $m(t)$ and using the resulting expansion to estimate $\mu(t)$. This method has the advantage that a double smoothing effect is achieved since $m(t)$ is an average of the N measurements, and the coefficients of the expansion are obtained by a time average.

Let $\left\{ P_k(t) \,,\ k = 0, 1, \ldots \right\}$ be a complete set of orthonormal functions defined on the interval $(0, T)$. Thus

$$\int_0^T P_i(t)\, P_j(t)\, dt = \begin{cases} 0 \,, & i \neq j \\ 1 \,, & i = j \end{cases} \tag{9}$$

Let $L_K(t)$ be a linear combination of the first K members of $P_k(t)$:

$$L_K(t) = \sum_{k=0}^{K} a_k P_k(t) \tag{10}$$

where the coefficients a_k are to be determined as follows. Consider the quantity

$$\Delta = \int_0^T \left[m(t) - L_K(t) \right]^2 dt \tag{11}$$

which is a measure of the total distance between $m(t)$ and $L_K(t)$. The coefficients a_k are to be chosen to minimize Δ. Without going into the detailed calculations, it may be shown that the minimum value of Δ occurs when

$$\frac{\partial \Delta}{\partial a_k} = 0 \quad , \qquad k = 0, \ldots, K \tag{12}$$

from which it follows that

$$a_k = \int_0^T m(t)\, P_k(t)\, dt \tag{13}$$

At this point it is convenient to compute the first two moments of a_k. From Eq. (13),

$$E\left[a_k\right] = \int_0^T E\left[m(t)\right] P_k(t)\, dt = \int_0^T \mu(t)\, P_k(t)\, dt = b_k \tag{14}$$

where b_k is defined to be the kth coefficient in the orthogonal expansion of $\mu(t)$. The second moment of a_k is

$$E\left[a_k^2\right] = \int_0^T \int_0^T E\left[m(u)\, m(v)\right] P_k(u)\, P_k(v)\, du\, dv$$

$$= \int_0^T \int_0^T \left[\mu(u)\, \mu(v) + \frac{R_x(u, v)}{N}\right] P_k(u)\, P_k(v)\, du\, dv \tag{15}$$

where $R_x(u, v)$ is the covariance function of the nonstationary random process $x(t)$. Using Eq. (14), Eq. (15) may be expressed as

$$E\left[a_k^2\right] = b_k^2 + \frac{1}{N} \int_0^T \int_0^T R_x(u, v)\, P_k(u)\, P_k(v)\, du\, dv \tag{16}$$

The covariance function $R_x(u, v)$ is defined by

$$R_x(u, v) = E\left\{\left[x(u) - \mu_x(u)\right]\left[x(v) - \mu_x(v)\right]\right\}$$

$$= E\left[x(u)\, x(v)\right] - \mu_x(u)\, \mu_x(v)$$

and should not be confused with the correlation function which is defined by $E\left[x(u)\,x(v)\right]$ alone. When either of the mean values $\mu_x(u)$ or $\mu_x(v)$ is zero, then the covariance function $R_x(u, v)$ is identical to the correlation function $E\left[x(u)\,x(v)\right]$.

The expected value of $L_K(t)$ may now be found by application of Eqs. (10) and (14). Thus,

$$E\left[L_K(t)\right] = \sum_{k=0}^{K} E\left[a_k\right] P_k(t) = \sum_{k=0}^{K} b_k P_k(t) \tag{17}$$

Equation (17) indicates that, unless $\mu(t)$ can be represented exactly by a Kth order expansion of the $P_k(t)$, $L_K(t)$ is a biased estimate of $\mu(t)$. The effect of bias will be shown more clearly after the mean square error has been determined.

Rather than use the error criterion at a particular value of t discussed in the previous section, the one to be employed here will be the integrated mean square difference between $L_K(t)$ and $\mu(t)$ for all t in the range $0 \leq t \leq T$, which is defined by

$$\epsilon^2 = \int_0^T E\left[L_K(t) - \mu(t)\right]^2 dt \tag{18}$$

There are several reasons for this choice. First, it is a measure of the total "distance" between $L_K(t)$ and $\mu(t)$ for all t in the range $0 \leq t \leq T$, and secondly, it leads to a separation of the bias and noise errors.

Expanding Eq. (18), it is seen that

$$\epsilon^2 = \int_0^T E\left[L_K^2(t)\right] dt - 2\int_0^T E\left[L_K(t)\right] \mu(t)\, dt + \int_0^T \mu^2(t)\, dt \tag{19}$$

The first two integrals of Eq. (19) may be evaluated as follows:

$$\int_0^T E\left[L_K(t)\right] \mu(t)\, dt = \int_0^T \sum_{k=0}^{K} b_k P_k(t)\, \mu(t)\, dt$$

$$= \sum_{k=0}^{K} b_k \int_0^T \mu(t)\, P_k(t)\, dt$$

$$= \sum_{k=0}^{K} b_k^2 \tag{20}$$

and

$$\int_0^T E\left[L_K^2(t)\right] dt = \int_0^T \sum_{i,\, j=0}^{K} E\left[a_i a_j\right] P_i(t)\, P_j(t)\, dt$$

$$= \sum_{k=0}^{K} E\left[a_k^2\right]$$

$$= \sum_{k=0}^{K} \left[b_k^2 + \frac{1}{N} \int_0^T \int_0^T R_x(u,v)\, P_k(u)\, P_k(v)\, du\, dv\right] \tag{21}$$

where the last equality of Eq. (21) follows from Eq. (16). Substitution of Eq. (20) and (21) into Eq. (19) shows that

$$\epsilon^2 = \int_0^T \mu^2(t)\, dt - \sum_{k=0}^{K} b_k^2 + \frac{1}{N} \sum_{k=0}^{K} \int_0^T \int_0^T R_x(u,\, v)\, P_k(u)\, P_k(v)\, du\, dv \tag{22}$$

Equation (22) has several interesting properties which resulted from the use of the orthogonal expansion. Let ϵ_K^2 represent the first two terms of ϵ^2, called the truncation error. It is clear that these two terms involve only $\mu(t)$ and the coefficients in the orthogonal expansion of $\mu(t)$, and thus are independent of the noise process $x(t)$. The truncation error ϵ_K^2 in using a finite orthogonal expansion is thus completely isolated from the error caused by noise. This property is of great practical importance since it permits independent investigations of the "signal" and "noise" to be carried out.

It should be noted that ϵ_K^2 is always positive or zero. This follows from the fact that, for any orthonormal system, Bessel's inequality, Ref. [2, p. 51], applies so that

$$\sum_{k=0}^{K} b_k^2 \leq \int_0^T \mu^2(t)\, dt \tag{23}$$

where the equality sign holds if $\mu(t)$ can be represented exactly by a Kth degree expansion.

The final term in the expression for ϵ^2 is the contribution of the noise to the mean square error and will be denoted by ϵ_o^2. Unless $\mu(t)$ and $R(u, v)$ are known, it is not possible to evaluate the double integrals; however, an upper bound may be obtained by using the fact that $R(u, v) \leq R(u, u)$. Thus

$$\epsilon_o^2 = \frac{1}{N} \sum_{k=0}^{N} \int_0^T \int_0^T R_x(u, v)\, P_k(u)\, P_k(v)\, du\, dv$$

$$\leq \frac{1}{N} \sum_{k=0}^{N} \int_0^T R_x(u, u)\, P_k(u) \int_0^T P_k(v)\, dv\, du \tag{24}$$

Since $P_0(v) = \frac{1}{\sqrt{T}}$, the orthogonality property of the set $P_k(v)$ shows that

$$\sqrt{T} \int_0^T P_k(v) P_0(v)\, dv = \int_0^T P_k(v)\, dv = 0 \tag{25}$$

Thus all terms in the second line of Eq. (24) are zero except the k = 0 term, and

$$\epsilon_o^2 \leq \frac{1}{N} \int_0^T R_x(u, u)\, du = \frac{1}{N} \int_0^T \sigma_x^2(u)\, du \tag{26}$$

If the integrated mean square error criterion is applied to the ensemble averaging technique, it is easily shown that the resulting mean square error is exactly equal to the right side of Eq. (26). This means that by choosing a suitable value of K to minimize ϵ_K , an improved estimate of $\mu(t)$ will be obtained in almost all cases through the use of the orthogonal expansion.

The orthogonal expansion technique will be illustrated in a subsequent section on the estimation of mean square values.

3. MEAN SQUARE VALUE MEASUREMENTS

3.1 ENSEMBLE AVERAGING

A similar analysis to the one given in Section 2.1 will now be carried out to determine how the nonstationary mean square values change with time. This can be estimated by using a mean square value device that performs the following operation on the N available sample records $\{x_i(t)\}$, $i = 1, 2, \ldots, N$, namely to define $g_i(t) = x_i^2(t)$ and to compute

$$g(t) = \frac{1}{N} \sum_{i=1}^{N} g_i(t) = \frac{1}{N} \sum_{i=1}^{N} x_i^2(t) \tag{27}$$

The quantity $g(t)$ is an estimate of the true mean square value of the random process $\{x_i(t)\}$ at any time t since the mean value of $g(t)$ is

$$\begin{aligned}
\mu_g(t) = E\left[g(t)\right] &= \frac{1}{N} \sum_{i=1}^{N} E\left[x_i^2(t)\right] \\
&= \frac{1}{N} \sum_{i=1}^{N} \left[\mu_x^2(t) + \sigma_x^2(t)\right] \\
&= \mu_x^2(t) + \sigma_x^2(t)
\end{aligned} \tag{28}$$

The variance associated with a set of estimates $g(t)$ will now be calculated. By definition, the variance

$$\sigma_g^2(t) = E\left[g^2(t)\right] - \mu_g^2(t) \tag{29}$$

where $\mu_g(t)$ is given by Eq. (28) and where

$$E\left[g^2(t)\right] = \frac{1}{N^2} \sum_{i,j=1}^{N} E\left[x_i^2(t)\, x_j^2(t)\right] \tag{30}$$

$$= \frac{1}{N^2}\left[\sum_{i=1}^{N} E\left[x_i^4(t)\right] + \sum_{\substack{i,j=1\\ i\neq j}}^{N} E\left[x_i^2(t)\, x_j^2(t)\right]\right] \tag{31}$$

Thus the problem reduces to evaluation of the ensemble averages appearing in Eq. (31).

In order to obtain reasonable closed form answers, it will be assumed now that the set of values $\left\{x_i(t)\right\}$ at any time t follows a Gaussian distribution with mean value $\mu_x(t)$ and variance $\sigma_x^2(t)$. One can then derive for the ensemble averages

$$E\left[x_i^4(t)\right] = 3\left[\sigma_x^2(t) + \mu_x^2(t)\right]^2 - 2\mu_x^4(t) \tag{32}$$

$$E\left[x_i^2(t)\, x_j^2(t)\right] = \left[\sigma_x^2(t) + \mu_x^2(t)\right]^2 \quad \text{for } i \neq j \tag{33}$$

Substitution into Eqs. (31) and (29) yields the result

$$\sigma_g^2(t) = \frac{2}{N}\left[\sigma_x^4(t) + 2\mu_x^2(t)\,\sigma_x^2(t)\right] \tag{34}$$

Thus $\sigma_g^2(t)$ approaches zero as N approaches infinity so that $g(t)$ is a consistent estimate of the mean value $\mu_g(t)$.

For the case in which the mean value $\mu_x(t)$ is equal to zero, estimation of mean square values and variances are equivalent since variances represent mean square values about the mean. Equation (28) now becomes

$$\mu_g(t) = \sigma_x^2(t) \tag{35}$$

Therefore, $g(t)$ is an unbiased estimate of the variance if and only if the mean value is zero. The corresponding variance of $g(t)$ becomes

$$\sigma_g^2(t) = \frac{2}{N}\,\sigma_x^4(t) = \frac{2}{N}\,\mu_g^2(t) \tag{36}$$

Confidence curves similar to those obtained in Section 2.1 may be obtained from the Tchebycheff inequality. For any positive constant, k,

$$\text{Prob}\left[\,|\,g(t) - \mu_g(t)\,| \ge k\,\sigma_g(t)\right] \le \frac{1}{k^2} \tag{37}$$

where $\mu_g(t)$ and $\sigma_g(t)$ represent the true mean value and standard deviation associated with the set of measurements $\left\{g(t)\right\}$. A value of $k = 3$ corresponds to an 89% confidence band. The end limits are found from

$$\begin{aligned} g(t) &= \mu_g(t) \pm k\,\sigma_g(t) \\ &= \left(1 \pm \sqrt{\frac{2}{N}}\;k\right)\mu_g(t) \end{aligned} \tag{38}$$

Solving for $\mu_g(t)$ yields

$$\mu_g(t) = \frac{g(t)}{1 \pm \sqrt{\frac{2}{N}}\;k} \tag{39}$$

assuming that $\sqrt{\frac{N}{2}} > k$. This equation indicates how $\mu_g(t)$ is related to an estimate $g(t)$.

For example, if $k = 3$ and $N = 50$, then there is an 89% confidence that $\mu_g(t)$ lies in the range bounded by

$$\mu_g(t) = \frac{g(t)}{1 \pm 0.60} = \left[0.625\,g(t) \text{ to } 2.5\,g(t)\right] \tag{40}$$

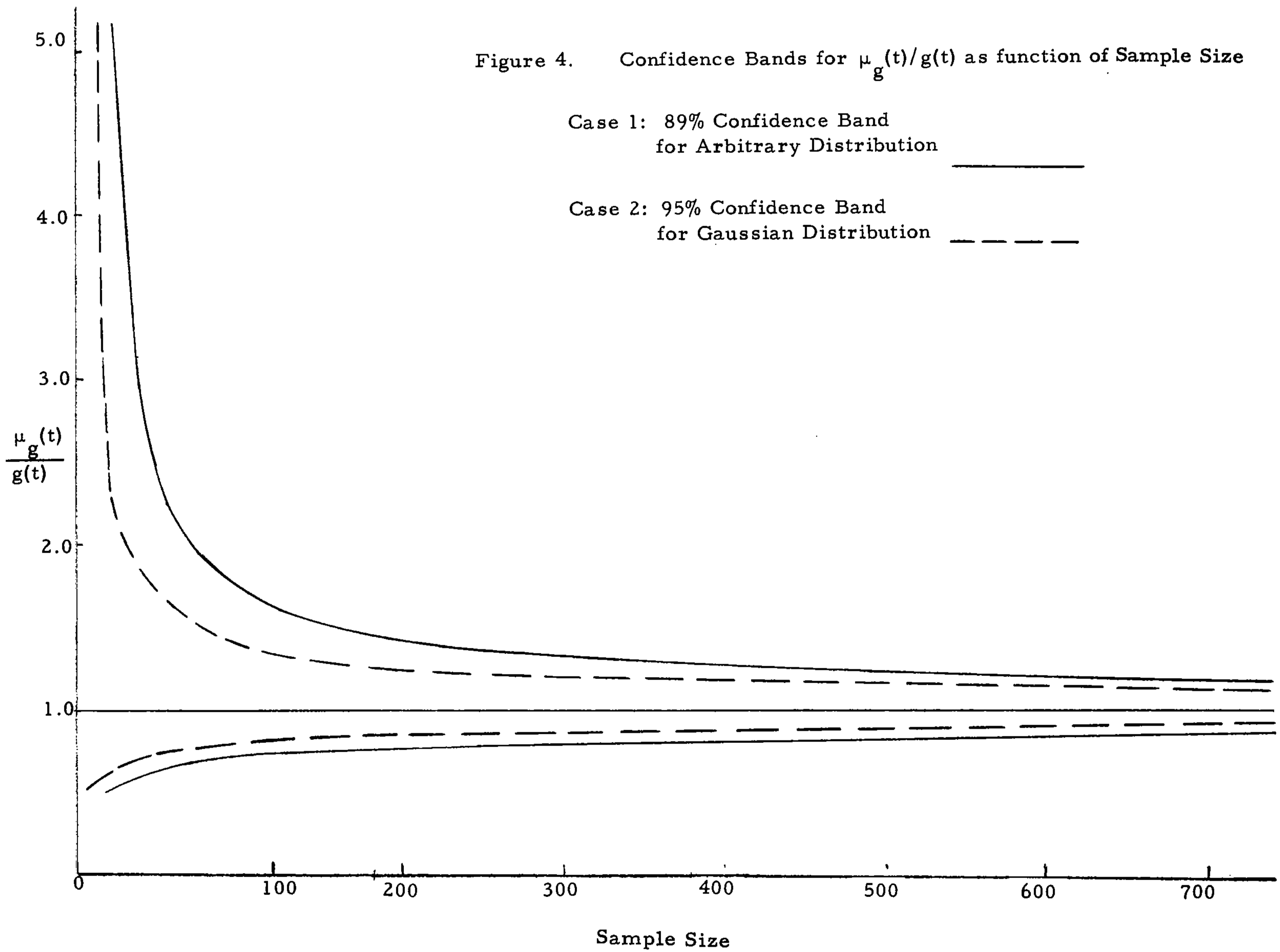


Figure 4. Confidence Bands for $\mu_g(t)/g(t)$ as function of Sample Size

Special tables and confidence curves can be generated similar to Table 1 and Figure 3.

Figure 4 based on Eq. (39) is a plot of the ratio $\mu_g(t)/g(t)$ as a function of the number of samples N. Two cases are considered. Case 1 applies to arbitrary probability distributions for $g(t)$ and sets $k = 3$ corresponding to an 89% confidence band as given by the Tchebycheff inequality. Case 2 applies to a Gaussian probability distribution for $g(t)$ and sets $k = 2$ corresponding to a 95% Gaussian confidence band.

3.2 ORTHOGONAL FUNCTION APPROXIMATION

In Section 2.2 the analytical basis for orthogonal function approximation is established and therefore will not be repeated here. The corresponding results for mean square value estimation by an orthogonal expansion $L_K(t)$ may be obtained by appropriate modification of the equations of Section 2.2.

As in previous sections, it is assumed that N independent sample records $\{x_i(t)\}$ are available. For simplicity it is also assumed that the underlying random process has zero mean. The quantity $g(t)$ is again defined by Eq. (27), namely

$$g(t) = \frac{1}{N} \sum_{i=1}^{N} x_i^2(t) \tag{41}$$

Thus, if $m(t)$ is replaced by $g(t)$ in Eq. (13) of Section 2.2, the coefficients of the orthogonal expansion of $g(t)$ are given by

$$a_k = \int_0^T g(t)\, P_k(t)\, dt \tag{42}$$

Upon taking the expected value of Eq. (42), it is seen from Eq. (35) that

$$E\left[a_k\right] = \int_0^T \sigma_x^2(t)\, P_k(t)\, dt = b_k \tag{43}$$

where b_k is now the kth coefficient in the orthogonal expansion of $\sigma_x^2(t)$. In order to calculate the second moment of a_k, it is necessary to have the fourth moment of $x(t)$ which in general will not be known.

For the important special case of a Gaussian distribution, however, the fourth moment is expressible in terms of the second moment. Therefore, in the remainder of this section, the Gaussian assumption will be made. Then, it may be shown that

$$E\left[a_k^2\right] = \int_0^T \int_0^T \sigma_x^2(u)\, \sigma_x^2(v) \left[1 + \frac{2}{N}\, \rho_x^2(u, v)\right] P_k(u)\, P_k(v)\, du\, dv \tag{44}$$

where $\rho_x(u, v)$ is the normalized covariance function of the nonstationary random process $x(t)$.

The normalized covariance function $\rho_x(u, v)$ is defined by

$$\rho_x(u, v) = \frac{R_x(u, v)}{\left[R_x(u, u)\, R_x(v, v)\right]^{\frac{1}{2}}}$$

where $R_x(u, v)$ is the unnormalized covariance function. For all u and v, it may be shown that

$$-1 \leqq \rho_x(u, v) \leqq 1$$

Using the above results, the first and second moments of the mean square estimate, $L_K(t)$, are found to be

$$E\left[L_K(t)\right] = \sum_{k=1}^{K} b_k P_k(t) \tag{45}$$

This is the same expression as was obtained in Eq. (17) of Section 2.2. Thus, unless $\sigma_x^2(t)$ can be represented exactly by a Kth order expansion of the $P_k(t)$, the quantity $L_K(t)$ is a biased estimate of $\sigma_x^2(t)$.

The integrated mean square error will again be used to measure how close $L_K(t)$ is to $\sigma_x^2(t)$. From the definition presented previously, the integrated mean square error ϵ^2 is given by

$$\epsilon^2 = \int_0^T \left[L_K(t) - \sigma_x^2(t)\right]^2 dt \tag{46}$$

An expansion of Eq. (46) similar to the one carried out in Section 2.2 which led to Eq. (22) shows that ϵ^2 may be written as

$$\epsilon^2 = \int_0^T \sigma_x^4(t)\, dt - \sum_{k=1}^{K} b_k^2 + \frac{2}{N} \sum_{k=1}^{K} \int_0^T \int_0^T \sigma_x^2(u)\, \sigma_x^2(v)\, \rho_x^2(u, v)\, P_k(u)\, P_k(v)\, du\, dv \tag{47}$$

Here again, it is seen that the errors caused by truncation and noise are completely separated. Thus, as before ϵ^2 can be written as

$$\epsilon^2 = \epsilon_K^2 + \epsilon_o^2 \tag{48}$$

and it may be shown that

$$\epsilon_o^2 < \frac{2}{N} \int_0^T \sigma_x^4(t)\, dt \tag{49}$$

indicating that if the truncation error, ϵ_K^2, is small, the orthogonal function approximation technique will produce lower errors than ensemble averaging.

The orthogonal function approximation technique will now be illustrated by two examples to show how this approach may be applied to the estimation of nonstationary mean square values. For simplicity, the time interval of interest is chosen to be of unit length and the normalized covariance function to be stationary (i.e., $\rho(u, v) = \rho(u - v)$.

Example 1:

Suppose it is known that a random process with zero mean has a mean square value which can be represented by a second order polynomial, i.e.,

$$\sigma^2(t) = c_0 + c_1 t + c_2 t^2 \quad , \qquad 0 \leq t \leq 1 \tag{50}$$

where the coefficients are unknown but constant. This could have been determined by computer simulation of the physical process, for example. Since $\sigma^2(t)$ is a polynomial on the interval (0, 1), the best set of orthonormal functions to use is the set of orthonormal polynomials on the interval (0, 1) since the smallest number of terms are then required to estimate $\sigma^2(t)$. On the interval (0, 1), the first three orthonormal polynomials are

$$\begin{aligned} P_0(t) &= 1 \\ P_1(t) &= \sqrt{3}\ (2t - 1) \\ P_2(t) &= \sqrt{5}\ (6t^2 - 6t + 1) \end{aligned} \tag{51}$$

After $g(t)$, as defined by Eq. (41), has been measured experimentally, the coefficients of the estimate of $\sigma^2(t)$ would be determined from

$$a_m = \int_0^1 g(t)\, P_m(t)\, dt \quad , \qquad i = 0, 1, 2 \tag{52}$$

so that the estimate of $\sigma^2(t)$ becomes

$$L_M(t) = \sum_{m=0}^{2} a_m P_m(t) \tag{53}$$

It is easily seen that $L_M(t)$ is an unbiased estimate of $\sigma^2(t)$ since

$$E\left[L_M(t)\right] = \sum_{m=0}^{2} E\left[a_m\right] P_m(t)$$

$$= \sum_{m=0}^{2} b_m P_m(t) \tag{54}$$

where

$$b_0 = \int_0^1 A^2(t)\,P_0(t)\,dt = c_0 + \frac{c_1}{2} + \frac{c_2}{3}$$

$$b_1 = \int_0^1 A^2(t)\,P_1(t)\,dt = \frac{c_1 + c_2}{2\sqrt{3}}$$

$$b_2 = \int_0^1 A^2(t)\,P_2(t)\,dt = \frac{c_2}{6\sqrt{5}}$$

Because $\sigma^2(t)$ is a second order polynomial, it is completely specified by a second order orthogonal polynomial expansion. Thus the truncation error ϵ_K^2 is zero and the mean square error expression of Eq. (47) reduces to

$$\epsilon_1^2 = \frac{2}{N} \sum_{m=0}^{2} \int_0^1 \int_0^1 \sigma^2(u)\,\sigma^2(v)\,\rho^2(u - v)\,P_m(u)\,P_m(v)\,du\,dv \tag{55}$$

where $\rho(\tau)$ is the normalized stationary covariance function. Unfortunately, Eq. (55) is quite tedious to evaluate even for a simple covariance function such as $\exp(-K|\tau|)$ and the resulting expression would not be subject to easy interpretation. In evaluating the above mean square error for actual experimental data, the computations would best be carried out on a digital computer.

Example 2:

To illustrate the smoothing effect of an orthogonal expansion in the estimation of $\sigma^2(t)$, the following example has been developed which allows the calculation of ϵ_o^2 without undue effort.

Let $\sigma^2(t)$ have the form shown in Figure 5. Since $\sigma^2(t)$ consists of a set of step functions, it is convenient to choose an orthonormal set which is also made up of step functions. One set of such functions is comprised of Walsh functions, Ref. [3, p. 20-21] ; the first four of which are shown in Figure 6. The values of the coefficients of the expansion of $\sigma^2(t)$ in terms of the Walsh functions are

$$
\begin{aligned}
b_0 &= \int_0^1 \sigma^2(t)\, P_0(t)\, dt = \frac{3}{4} \\
b_1 &= \int_0^1 \sigma^2(t)\, P_1(t)\, dt = -\frac{1}{4} \qquad (56) \\
b_j &= 0 \quad , \qquad j = 2, 3, \ldots
\end{aligned}
$$

Direct calculation shows that

$$\int_0^1 \sigma^4(t)\, dt = \frac{1}{8} + \frac{1}{2} = \frac{5}{8}$$

and

$$b_0^2 + b_1^2 = \frac{9}{16} + \frac{1}{16} = \frac{5}{8}$$

so that the truncation error is zero if the first two terms in the expansion are used to represent $\sigma^2(t)$. The resulting expression for the integrated mean square error is

$$\epsilon_1^2 = \frac{2}{N} \int_0^1 \int_0^1 \left[f_0(u,v) + f_1(u, v) \right] \rho^2(u - v)\, du\, dv \tag{57}$$

where

$$f_i(u, v) = \sigma^2(u)\,\sigma^2(v)\, P_i(u)\, P_i(v) \quad ; \quad i = 1, 2$$

Because of the jumps in both the Walsh functions and $\sigma^2(t)$, the region over which the integration is to be carried out in Eq. (57) must be divided into four parts as shown in Figure 7. In each region the values of f_0 and f_1 are constants which are tabulated below.

Region	α	β	γ	δ
$f_0(u, v)$	1/4	1/2	1/2	1
$f_1(u, v)$	1/4	-1/2	-1/2	1

Table 3

From the table it is seen that the integrals over β and γ cancel and Eq. (57) reduces to

$$\epsilon_1^2 = \frac{2}{N}\left[\frac{1}{2} \int_\alpha \rho^2(u - v)\, du\, dv + 2 \int_\delta \rho^2(u - v)\, du\, dv \right] \tag{58}$$

Upon making the change of variable

$$\begin{aligned} r &= u - v \\ s &= u + v \end{aligned} \quad , \qquad du\, dv = \frac{dr\, ds}{2} \tag{59}$$

the regions α and δ are mapped onto the $r - s$ plane as shown in Figure 8, and the expression for ϵ_1^2 becomes

$$\begin{aligned} \epsilon_1^2 &= \frac{1}{N}\left[\frac{1}{2} \int_\alpha \rho^2(r)\, dr\, ds + 2 \int_\delta \rho^2(r)\, dr\, ds \right] \\ &= \frac{1}{N}\left[\left\{ \frac{1}{2} \int_{\alpha_1} + \frac{1}{2} \int_{\alpha_1} + 2 \int_{\delta_1} + 2 \int_{\delta_2} \right\} \rho^2(r)\, dr\, ds \right] \end{aligned} \tag{60}$$

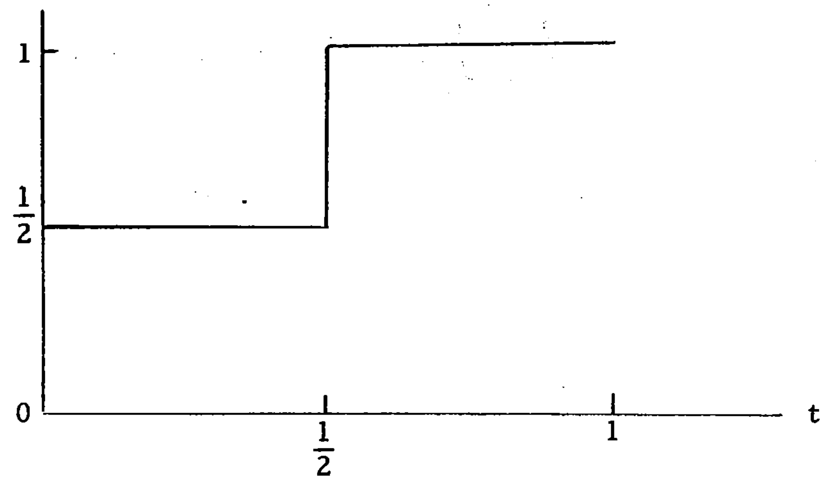


Figure 5. Mean Square Value of Random Process

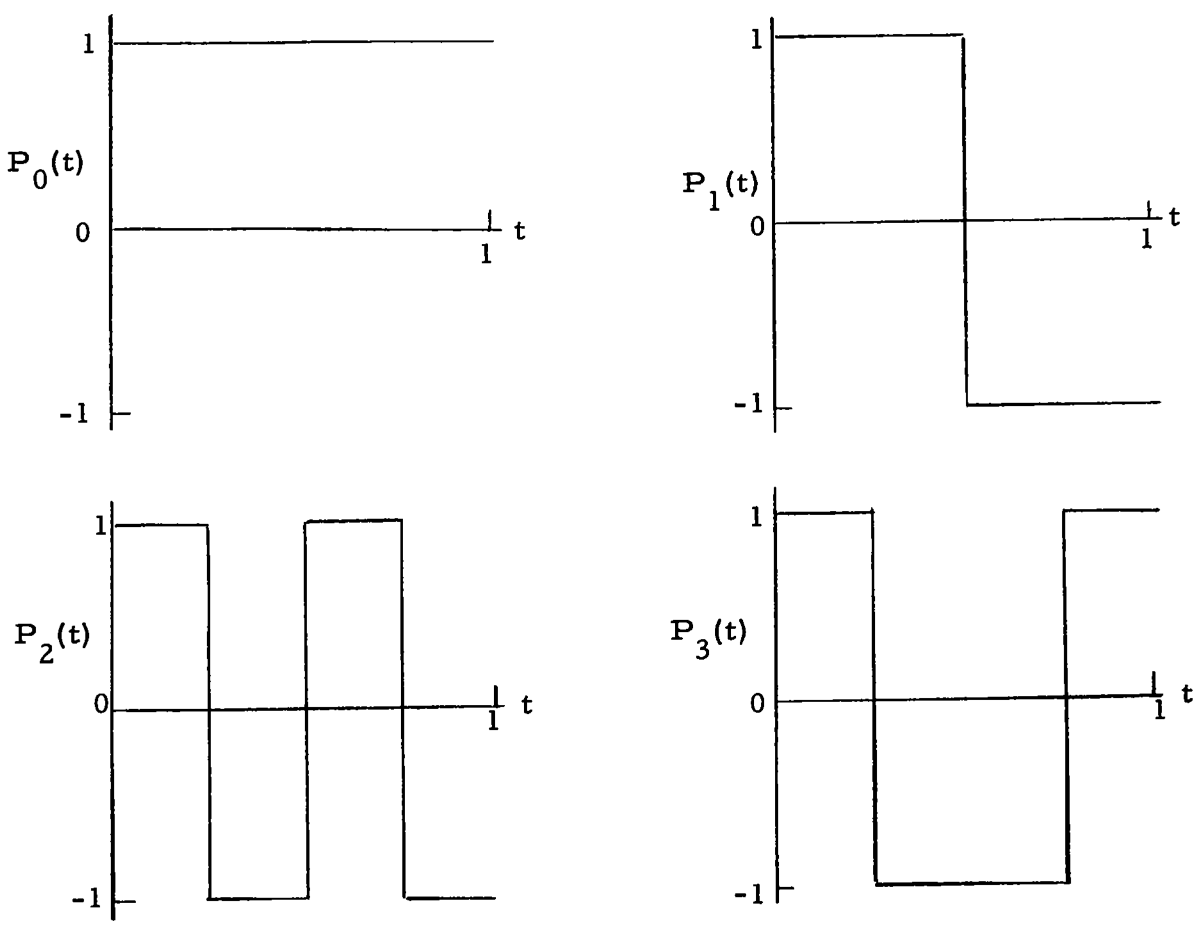


Figure 6. Walsh Functions

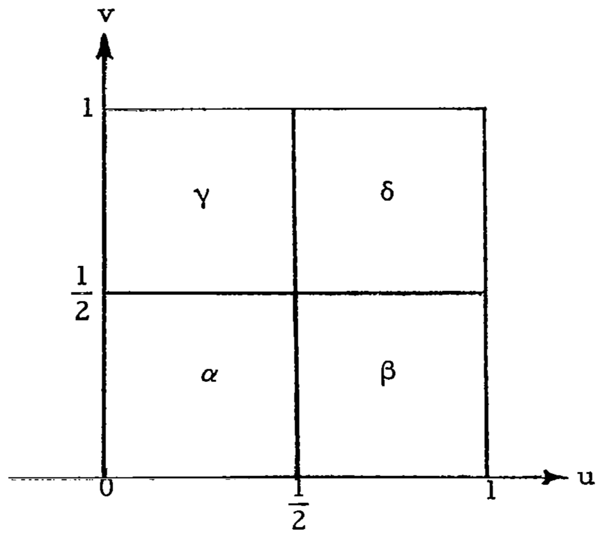


Figure 7. Region of Integration

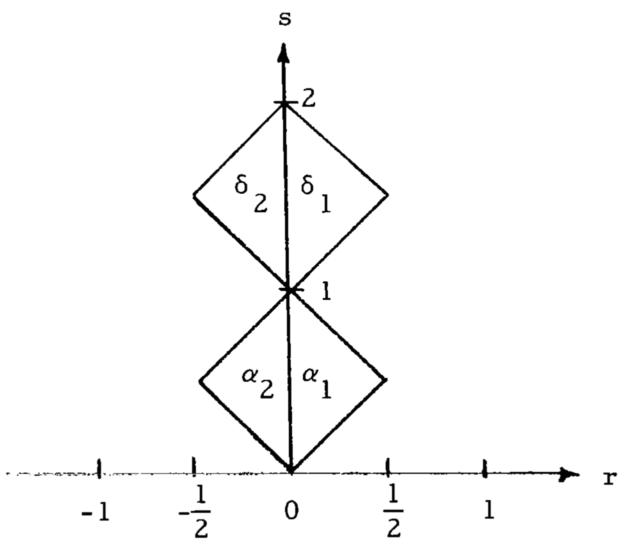


Figure 8. Transformed Region of Integration

Without going into the details, it may be shown that the four integrals in Eq. (60) are numerically equal. Thus, by adding the coefficients,

$$\epsilon_1^2 = \frac{5}{N} \int_{\alpha_1} \rho^2(r)\, dr\, ds = \frac{5}{N} \int_0^{\frac{1}{2}} \int_r^{1-r} \rho^2(r)\, ds\, dr \tag{61}$$

Assume that the normalized covariance is given by

$$\rho(r) = \exp\,(-\gamma\,|r|\,) \tag{62}$$

Then, substitution of Eq. (62) into Eq. (61) leads to the result that

$$\epsilon_1^2 = \frac{5}{N} \int_0^{\frac{1}{2}} (1 - 2r)\, \exp\,(-2\gamma\, r)\, dr = \frac{5}{N} \left[\frac{e^{-\gamma} + \gamma - 1}{2\gamma^2} \right] \tag{63}$$

Equation (63) is shown in Figure 9 where the integrated mean square error is plotted as a function of the parameter γ. Since γ is directly proportional to the effective width of the stationary noise spectrum, it is clear that the orthogonal function expansion is much more effective in smoothing if the noise is wideband. It is likely that this result is true for most cases of practical interest; however, a general proof is not given here.

As a numerical illustration of these results, suppose $\gamma = 5$ and $N = 100$. From Figure 9, the orthogonal approximation approach produces an integrated mean square error of 0.004 while the error for ensemble averaging is 0.0125.

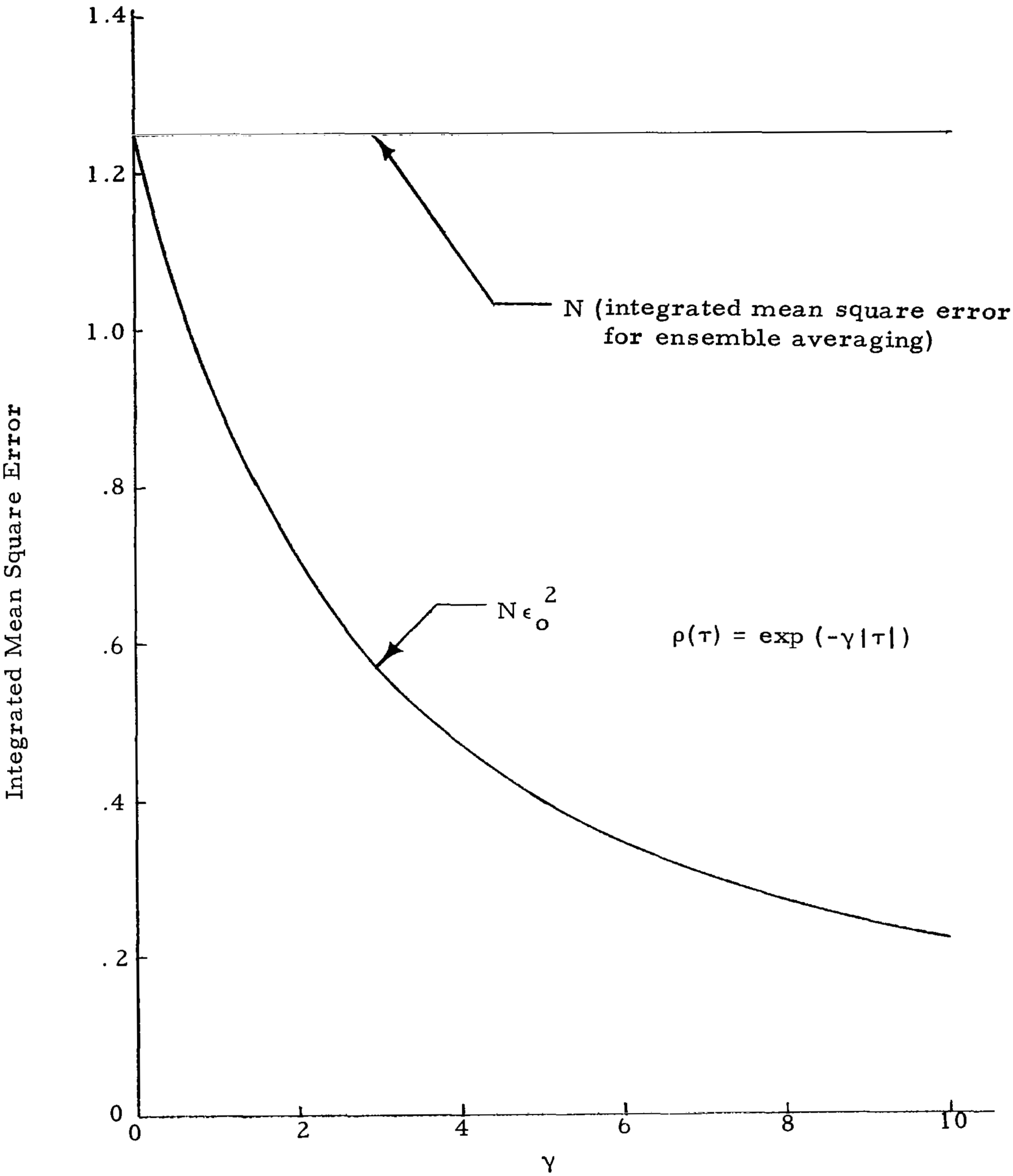


Figure 9.. Integrated Mean Square Error

3.3 SHORT TIME AVERAGING

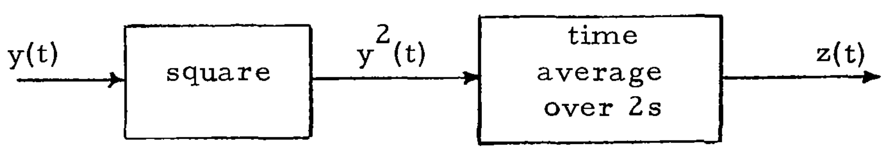
One technique which is currently employed to estimate nonstationary mean square values when only one or a few samples are available consists of obtaining a continuous short time average. The effect of a time average is to smooth the random fluctuations over the averaging interval and thus reduce the uncertainty in the estimate. In the analysis presented below, explicit expressions for the bias error and mean square error will be developed for processes with nonstationary mean square values.

It will be assumed that the nonstationary process has zero mean value, and is of the form

$$y(t) = A(t)\,x(t) \tag{64}$$

where $x(t)$ is a zero mean Gaussian process and $A(t)$ is any integrable function. The covariance function of $x(t)$ is $R_x(t_1, t_2)$ and it is assumed that $R_x(t, t) \equiv R_x(t) \equiv 1$. Many physical situations may be represented by the nonstationary process of Eq. (64).

Suppose that $y(t)$ is first squared and then averaged over a time 2s, as shown below.



The resulting random process $z(t)$ is thus defined by

$$z(t) = \frac{1}{2s} \int_{t-s}^{t+s} y^2(u)\, du \tag{65}$$

The expected value of $z(t)$ may be found as follows.

$$E\left[z(t)\right] = \frac{1}{2s} \int_{t-s}^{t+s} A^2(u)\, E\left[x^2(u)\right] du$$

$$= \frac{1}{2s} \int_{t-s}^{t+s} A^2(u)\, du \tag{66}$$

since $E\left[x^2(u)\right] = R(u) = 1$.

In order to determine the bias error, assume that $A^2(t)$ can be represented in the interval $(t-s,\ t+s)$ by a power series of order N, namely,

$$A^2(t) = \sum_{n=0}^{N} a_n t^n \tag{67}$$

Substitution of Eq. (67) into Eq. (66) gives

$$E\left[z(t)\right] = \frac{1}{2s} \sum_{n=0}^{N} a_n \int_{t-s}^{t+s} u^n\, du$$

$$= \sum_{n=0}^{N} \frac{a_n}{2s(n+1)} \left[(t+s)^{n+1} - (t-s)^{n+1} \right]$$

$$= \sum_{n=0}^{N} a_n \sum_{m=0}^{n}{}^{e} \frac{\binom{n+1}{m+1}}{n+1}\, t^{n-m}\, s^m \tag{68}$$

where the third equality is obtained after some algebraic manipulation and the "e" within the second summation means that only the even values of m are used.

The bias error in estimating $A^2(t)$ is defined by

$$b(t, s) = E\left[z(t)\right] - A^2(t)$$

$$= \sum_{n=0}^{N} a_n \sum_{\substack{e \\ m=2}}^{n} \frac{\binom{n+1}{m+1}}{n+1} t^{n-m} s^m \qquad (69)$$

If $b(t, s)$ is written in the form

$$b(t, s) = \sum_{n=0}^{N} b_n(t, s) \qquad (70)$$

then the first few terms of the series of Eq. (70) are given by

$$b_0 = 0$$

$$b_1 = 0$$

$$b_2 = \frac{a_2 s^2}{3}$$

$$b_3 = a_3 t s^2$$

$$b_4 = a_4 (2t^2 s^2 + \frac{s^4}{5})$$

$$b_5 = a_5 \left(\frac{10t^3 s^2}{3} + t s^4 \right)$$

$$b_6 = a_6 \left(\frac{t^4 s^2}{5} + 3t^2 s^4 + \frac{s^6}{7} \right)$$

$$b_7 = a_7 \left(\frac{t^5 s^2}{7} + 7t^3 s^4 + t s^6 \right)$$

It is clear from the above development that if $A^2(t)$ is a polynomial of moderate order, a large bias error can be introduced into the estimation procedure. In most practical applications, however, the variations in $A^2(t)$ are not too complicated within an interval of length $2s$ since s is usually small. This means that, within each interval of length $2s$, $A^2(t)$ can be considered to be a polynomial of degree two or less and the resulting bias error will be negligible or zero.

The mean square error in estimating $A^2(t)$ by a short time average may be found as follows: Let $\epsilon^2(t)$ denote the expected value of the square of the difference between $z(t)$ and $A^2(t)$, then

$$\begin{aligned}\epsilon^2(t) &= E\left[z(t) - A^2(t)\right]^2 \\ &= E\left[z^2(t)\right] - 2A^2(t)\, E\left[z(t)\right] + A^4(t) \\ &= E\left[z^2(t)\right] - 2A^2(t)\left[A^2(t) + b(t, s)\right] + A^4(t) \\ &= E\left[z^2(t)\right] - A^4(t) - 2A^2(t)\, b(t, s) \qquad (71)\end{aligned}$$

From the definition of $z(t)$,

$$\begin{aligned}E\left[z^2(t)\right] &= \frac{1}{4s^2} \iint_{t-s}^{t+s} E\left[y^2(u)\, y^2(v)\right] du\, dv \\ &= \frac{1}{4s^2} \iint_{t-s}^{t+s} A^2(u)\, A^2(v)\, E\left[x^2(u)\, x^2(v)\right] du\, dv \qquad (72)\end{aligned}$$

When $x(t)$ is assumed to be Gaussian, as is the case here, a well known result for Gaussian processes gives

$$E\left[x^2(u)\, x^2(v)\right] = R^2(u)\, R^2(v) + 2R^2(u, v) = 1 + 2R^2(u, v) \qquad (73)$$

since $R(u) = R(v) = 1$.

Therefore, Eq. (72) becomes

$$E\left[z^2(t)\right] = \frac{1}{4s^2} \iint_{t-s}^{t+s} A^2(u)\,A^2(v)\left[1 + 2R^2(u, v)\right] du\,dv$$

$$= \left[\frac{1}{2s}\int_{t-s}^{t+s} A^2(u)\,du\right]^2 + \frac{1}{2s^2}\iint_{t-s}^{t+s} A^2(u)\,A^2(v)\,R^2(u, v)\,du\,dv$$

$$= \left[A^2(t) + b(t, s)\right]^2 + \frac{1}{2s^2}\iint_{t-s}^{t+s} A^2(u)\,A^2(v)\,R^2(u, v)\,du\,dv \tag{74}$$

Substitution of this result into Eq. (71) shows that

$$\epsilon^2(t) = b^2(t, s) + \frac{1}{2s^2}\iint_{t-s}^{t+s} A^2(u)\,A^2(v)\,R^2(u, v)\,du\,dv \tag{75}$$

The double integral occurring in Eq. (75) is quite difficult to evaluate even for simple cases. However, it is possible to establish an upper bound. The Schwarz inequality, Ref. [2, p. 49], states that for any two integrable functions f and g

$$\left(\int fg\right)^2 \leq \left(\int f^2\right)\left(\int g^2\right) \tag{76}$$

In terms of the double integral of Eq. (75), the Schwarz inequality shows that

$$\frac{1}{2s^2}\iint_{t-s}^{t+s} A^2(u)\,A^2(v)\,R^2(u,v)\,du\,dv < \frac{1}{2s^2}\left[\iint_{t-s}^{t+s} A^4(u)A^4(v)\,du\,dv \iint_{t-s}^{t+s} R^4(u,v)\,du\,dv\right]^{\frac{1}{2}}$$

$$= \frac{1}{2s^2}\int_{t-s}^{t+s} A^4(u)\,du\left[\iint_{t-s}^{t+s} R^4(u,v)\,du\,dv\right]^{\frac{1}{2}} \tag{77}$$

Using the above results, it is now possible to examine certain limiting cases for the mean square error and bias. As a first step, consider the limiting case as s approaches zero. From Eq. (69),

$$\lim_{s \to 0} b(t, s) = 0 \tag{78}$$

and from Eq. (75),

$$\lim_{s \to 0} \epsilon = \sqrt{2}\, A^2(t) \tag{79}$$

Note that the rms error is $\sqrt{2}$ times the quantity $A^2(t)$ being measured for small s.

Further examination of Eq. (75) indicates that the mean square error will become quite large for large values of s because of the increasing bias effect. Thus it is of interest to determine if an optimum value of s exists such that $\epsilon^2(t)$ is a minimum.

For simplicity, let the double integral in Eq. (75) be denoted by $\frac{1}{s^2} K(t, s)$, then

$$\epsilon^2(t, s) = b^2(t, s) + \frac{K(t, s)}{s^2} \tag{80}$$

The derivative of ϵ^2 with respect to s is given by

$$\frac{d\epsilon^2}{ds} = 2b\,b' + \frac{s^2 K' - 2sK}{s^4} \tag{81}$$

For small s, the bias term (2b b') will in most cases be zero or negligible. Hence, any value s_1, such that

$$\left.\frac{d\epsilon^2}{ds}\right|_{s=s_1} = 0 \tag{82}$$

will be close to a corresponding value s_2, where

$$s_2^2 K'(t, s_2) - 2s_2 K(t, s_2) = 0 \tag{83}$$

Now if a value s_2 can be found, it will be a first approximation to the optimum averaging time s, and it is likely that $\epsilon^2(t, s_2)$ will be close to $\epsilon^2(t, s_1)$.

From Eq. (83) it follows that

$$\frac{K'(t, s_2)}{K(t, s_2)} = \frac{2}{s_2}$$

or

$$K(t, s_2) = Cs_2^2 \tag{84}$$

where C is a constant of integration. The value of C may be found from Eq. (79) to be $(2A^4(t))$, and the value of s_2 is given by the solution of the equation

$$K(t, s_2) = 2A^4(t)\, s_2^2 \tag{85a}$$

In terms of the original functions

$$\iint_{t-s_2}^{t+s_2} A^2(u)\, A^2(v)\, R^2(u, v)\, du\, dv = 2A^4(t)\, s_2^2 \tag{85b}$$

The considerations presented above clearly illustrate the analytic difficulties associated with an error analysis of short time averaging techniques. In order to further investigate this, and other estimation procedures as well, it is advisable to simulate typical cases of interest on a computer and observe the effects of averaging time on the mean square error. This approach also has the advantage of permitting easy variation of the input parameters to test the sensitivity of the mean square error to these variations.

4. RECOMMENDATIONS FOR SUCCEEDING RESEARCH

In deciding which estimation procedure to implement for an experimental program, consideration must be given to the over-all costs involved. While certain of the techniques described in this report will theoretically give better results than others, they may be much more expensive to set up because of the complexity of operations. Thus, trade-offs exist between the number of measurements made, the cost of each measurement, and the cost of implementing the estimation procedure. Because of the high cost of many experimental programs, it is very desirable to select the estimation procedure which will minimize the total cost.

To gain a better understanding of the relative performance of the various estimation techniques, the following program is suggested. First, since mathematical difficulties preclude an exact error analysis in most cases, a computer simulation study should be made. Nonstationary processes of the type expected in actual physical situations would be used and a comparative evaluation of each estimation technique made for each type of process. One of the principal results of such a simulation would be to give the experimenter a better understanding of which estimation technique is best for his particular application. Quantitatively, the simulation would result in establishing required sample sizes to produce the same value of the mean square error. Also, it should be possible to determine optimum values for the parameters of the estimation procedures.

REFERENCES

1. Bendat, J. S., Enochson, L. D., Klein, G. H., and A. G. Piersol. "Advanced Concepts of Stochastic Processes and Statistics for Flight Vehicle Vibration Estimation and Measurement," ASD-TDR-62-973, Aeronautical Systems Division, AFSC, USAF, Wright-Patterson AFB, Ohio. December 1962. ASTIA AD 297 031.

2. Courant, R., and D. Hilbert, Methods of Mathematical Physics, vol. I, Interscience Publishers, New York. 1953.

3. Wiener, N., Nonlinear Problems in Random Theory, John Wiley and Sons, Inc. New York. 1958.

NASA-Langley, 1965 CR-226