27
8/30/79
240 UT IS

UCID- 18120

Lawrence Livermore Laboratory

A STATISTICAL EVALUATION OF THE PERFORMANCE OF THE TIME-DOMAIN PARET ALGORITHM

D. Lager

June 1, 1979

MASTER

EXTENT OF THIS DOCUMENT IS LIMITED

CONTENTS

Abstract	v
Introduction	1
The Structure Model	2
Performance Evaluation Methodology	7
Summary and Comments	23
Acknowledgment	25
Appendix--Bias Test	27
References	29

NOTICE

This report was prepared as an account of work sponsored by the United States Government. Neither the United States nor the United States Department of Energy, nor any of their employees, nor any of their contractors, subcontractors, or their employees, makes any warranty, express or implied, or assumes any legal liability or responsibility for the accuracy, completeness, or usefulness of any information appearing in this report. This report is not to be distributed outside the government.

Ref

LIST OF ILLUSTRATIONS

1. Eight-story structure being modeled	3
2. Eight degree-of-freedom model for eight-story structure	3
3. Node 1 displacement vs time	4
4. Node 1 magnitude spectrum	5
5. Coefficient of variation vs signal-to-noise ratio (SNR)	13
6. t-test for bias vs signal-to-noise ratio (SNR)	15
7. Root mean square error in damping vs NBEGIN	17
8. Coefficient of variation vs NPOLES	19
9. t-test for bias vs NPOLES	21

LIST OF TABLES

1. Modal parameters calculated by TDP using zero noise SAP IV data	5
2. Modal parameters calculated by SAP IV	6

ABSTRACT

A methodology was developed for evaluating the performance of the time-domain PARET (TDP) algorithm when processing waveforms contaminated with additive noise. The methodology is general; it may be applied to evaluating the performance of other algorithms and provides a good basis for comparison. The performance is evaluated by computing the bias, variability, and root mean square error in the parameters estimated by the algorithm when processing ensembles of waveforms. The waveforms used in this study were the displacements as a function of time from a computer simulation (by the SAP IV Program) of an eight-floor structure.

INTRODUCTION

The Nuclear Regulatory Commission (NRC) has funded Lawrence Livermore Laboratory to apply the PARET signal-processing technique^{1,2} to characterizing nuclear power plant structures so that the ability to predict how these structures respond to earthquakes might be improved. Since this problem raises issues of public safety, etc., it is important that the accuracy of the algorithm characterizing the structure be determined.

This report describes the first step toward that aim, the quantification of the accuracy (hereafter termed "performance") of the time-domain PARET algorithm (TDP) when processing the "measured" responses from a computer model of a structure with widely spaced (in frequency) modes of oscillation. Since, in real life, measurements are always contaminated by random noise, the performance must be evaluated statistically by running the algorithm for ensembles of waveforms with different sequences of random noise added. We express the performance in terms of the variability in the estimates of the parameters extracted from the waveforms and the root mean square (rms) error in the estimates. We also set up statistical tests to determine if the estimates are unbiased.

THE STRUCTURE MODEL

The TDP algorithm is a general parameter estimation method, which has been applied to many different problems.^{3,4,5} It is assumed the reader is familiar with its most recent implementation.¹ It is important that its accuracy be quantified for the range of parameters expected from the application of interest, since the algorithm may perform differently at estimating the parameters for, say, a radar target identification problem than for characterizing a mechanical structure. Both problems can be formulated in terms of an exponential series but the coefficients of the terms in the series will have different relative values. For this report we chose to apply the algorithm to the responses from a structure model known to have widely spaced modes of oscillation. A future report will discuss the performance for a more complicated structure having closely spaced modes.

The "structure" being modeled is the eight-story building shown in Fig. 1. A SAP IV⁶ computer model of the structure was constructed⁷ as shown in Fig. 2 after assuming we could ignore:

- out of plane displacements (assuming symmetry);
- longitudinal displacements;
- rotation (assuming rigid floor diaphragms).

The physical properties of the structure were:

- $E = 1.0 = \text{Young's modulus (k/ft}^2\text{)} \text{ (k is kips, a thousand-pound force)}$
- $I = 100 = \text{area moment of inertia (ft}^4\text{)}$
- $L = 1.0 = \text{distance between nodes (ft)}$
- $m = 1.0 = \text{lumped mass (ks}^2\text{/ft)}$.

Damping was introduced as modal damping at 5% of critical.

The structure was excited by a pulse at $t = 0$ of one time-step duration (0.02 s) consisting of a unit amplitude-ground acceleration (simulating an earthquake). The response measured was the displacement of the first floor (node No. 1 in Fig. 2) as a function of time. The response and its frequency spectrum are

FIG. 1. Eight-story structure being modeled.

FIG. 2. Eight degree-of-freedom model for eight-story structure.

shown in Figs. 3 and 4, respectively. The spectrum shows that the strongest three modes occur at nominal frequencies of 1, 3, and 5 Hz, monotonically decreasing in amplitude with frequency. The 1-Hz mode is about an order of magnitude stronger than the 3-Hz mode.

The TDP algorithm was applied to the calculated response of the first floor and yielded the values for modal frequency, damping, and strength (denoted f , β , and S) shown in Table 1. The f and β for each mode as calculated by SAP IV are shown in Table 2. For the six strongest modes TDP agreed with SAP IV to at least three digits of accuracy. The discrepancies arise from the loss of significance by the method we used to transfer the data between SAP IV and TDP programs. However, the errors for the strongest modes were sufficiently small that we concluded TDP was an unbiased estimator for the case of zero noise. Consequently, for all the subsequent noise studies discussed in this report, the zero noise f , β , and S calculated by TDP are used as the true values.

FIG. 3. Node 1 displacement vs time.

FIG. 4. Node 1 magnitude spectrum.

TABLE 1. Modal parameters calculated by TDP using zero noise SAP IV data.

Mode No.	Frequency (Hz)	Damping (s^{-1})	Strength ($\times 10^4$)
1	1.0161	-0.31962	5.05445
2	3.0138	-0.94799	1.27570
3	4.9088	-1.5441	0.630835
4	6.6367	-2.0875	0.353047
5	8.1385	-2.5599	0.201670
6	9.3630	-2.9439	0.106650

TABLE 2. Modal parameters calculated by SAP IV.

Mode No.	Frequency (Hz)	Damping (s^{-1})
1	1.0174	-0.32003
2	3.0176	-0.94919
3	4.9150	-1.54603
4	6.6450	-2.09020
5	8.1487	-2.56320
6	9.3750	-2.94893
7	10.282	-3.23423
8	10.839	-3.40944

PERFORMANCE EVALUATION METHODOLOGY

Our main goal was to evaluate the performance of the TDP algorithm for estimating modal frequencies, damping, and mode shape, given data contaminated with "measurement" noise. The performance was found both as a function of the level of the noise and as a function of the control parameters for the algorithm.

The performance is quantified in the fashion of Smith and Lager⁸ by computing the statistical bias, variability (standard deviation), and rms error in the parameter estimates produced by the algorithm. The parameters being estimated are the frequency, damping, and strength (residue magnitude) for each of the natural modes. The residue magnitude corresponds to the value of the mode shape at a given point on the structure.

The performance is computed by applying the algorithm to an ensemble of waveforms that are identical with the exception that each is contaminated by a different sequence of random noise. The standard deviation of the noise for the ensemble is kept constant. This was accomplished by choosing an ensemble size of 10 and making 10 copies of the response waveform from the first node of the SAP IV model of the structure shown in Figs. 1 and 2. Each of the copies was then contaminated by adding white Gaussian noise with a fixed standard deviation. The TDP algorithm was then used to estimate the parameters for each of the waveforms.

The variability in the parameter estimates is expressed in terms of the unbiased standard deviation, σ , according to:

$$\sigma^2 = \frac{1}{N-1} \sum_{i=1}^N (x_i - \bar{x})^2 \quad (1)$$

where

$$\bar{x} = \text{mean value of the estimates over the ensemble} = \frac{1}{N} \sum_{i=1}^N x_i$$

x_i = parameter value estimated from i th waveform

$N = 10$ = number waveforms in ensemble.

For each of the parameters the bias b in the estimates was computed using:

$$b = x_t - \bar{x} \quad , \quad (2)$$

where

x_t = true value for parameter.

A test can be performed (see Appendix) to determine if the algorithm has statistically significant bias by computing the statistic

$$t = \frac{b\sqrt{N}}{\sigma} \quad . \quad (3)$$

For an unbiased estimator, t will exceed the threshold value of 2.262 only 5 percent of the time.

The third performance measure, rms error, is calculated using:

$$\text{rms} = \sqrt{\frac{1}{N} \sum_{i=1}^N (x_t - x_i)^2} \quad . \quad (4)$$

The rms error can also be expressed in terms of the bias and standard deviation by:

$$\text{rms} = \sqrt{b^2 + \frac{(N-1)}{N} \sigma^2} \quad . \quad (5)$$

The true values for the parameters were obtained by using the time-domain PARET algorithm on the SAP IV waveform with zero noise added. The values agreed to several decimal places with those given by SAP IV, demonstrating the algorithm to be unbiased with zero noise present.

The noise added to each waveform in the ensemble was from a Gaussian distribution with zero mean and standard deviation σ_n . It was chosen both for being representative of the noise from measuring instruments and for its computational convenience. The values were computed by making use of the Central Limit Theorem and summing the numbers from a uniform distribution according to:

$$g(\sigma_n) = \sigma_n \sum_{i=1}^{12} (\mu_i - 0.5) , \quad (6)$$

where

$g(\sigma_n)$ = value from Gaussian distribution with standard deviation, σ_n

μ_i = value from a distribution uniform over interval 0 to 1.

The standard deviation of the noise σ_n was chosen as a fraction of the peak of the waveform

$$\sigma_n = F y_{\max} , \quad (7)$$

where

F = fraction of peak

$y_{\max}$ = maximum value of waveform $y(t)$, i.e., highest positive peak.

For the remainder of this report we define the signal-to-noise ratio as

$$\text{SNR} = 1/F , \quad (8)$$

the ratio of the waveform peak to the standard deviation of the noise.

This definition of SNR was chosen over the more usual definition⁹ of

$$\text{SNR} = \frac{P(y)}{P(n)} \quad , \quad (9)$$

where

$P(y)$ is power in waveform $y(t)$

$P(n)$ is power in the noise $n(t)$,

since the power ratio is not independent of the time duration over which it is computed. For times after the waveform has decayed to zero (but the noise has not) the SNR can be made to decrease arbitrarily by extending the time for computing the power. We chose the simple fraction of peak definition above to circumvent this difficulty.

Also, measurement instruments tend to produce a noise level that is more a function of the peak value (since the range setting is chosen to have the peak give a near full scale reading) than the actual power within the waveform.

The parameters being estimated were the frequency, damping, and strength (f , B , and S) for each mode in the waveform described by the exponential model,

$$y(t) = \sum_{i=1}^M A_i e^{\alpha_i t} \quad , \quad (10)$$

where

α_i = i th mode or natural frequency = $\beta_i + j2\pi f_i$

j = $\sqrt{-1}$

B = damping

f = frequency

A = complex residue

S = $|A|$

M = number of modes.

The modal damping value β is expressed in units of inverse time (s^{-1}) rather than the usual

$$F_c = \frac{\beta}{\sqrt{\beta^2 + (2\pi f)^2}} \quad (11)$$

where F_c is the fraction of critical damping. The parameters β and f are estimated as separated quantities by the algorithms, so we decided to evaluate their performance as separate quantities.

The performance of the algorithm was evaluated for three types of preprocessing of the waveforms--none, filtered, and autocorrelated. For the case of filter preprocessing each of the waveforms was truncation-filtered with a passband from 0 to 6.06 Hz, which was accomplished by computing the Fast Fourier Transform (FFT) of each of the waveforms, setting the frequency components outside the passband to zero, and performing an inverse FFT to produce a new time waveform. For the case of autocorrelation preprocessing, the normalized autocorrelation function,

$$R(\tau) = \frac{\int_{-\infty}^{\infty} x(t)x(t + \tau)dt}{R(0)} \quad (12)$$

where

$R(\tau)$ = autocorrelation value at time τ

$x(t)$ = value of waveform at time t

was computed for each of the waveforms in terms of the discrete sample autocorrelation function according to:

$$R_i = \frac{\sum_{j=1}^{N-i+1} x_j x_{i+j-1}}{\sum_{j=1}^N x_j^2} \quad i = 1, 2, \dots, N \quad (13)$$

where N = number of points in waveform.

The TOP algorithm was used to process each of the waveforms. All the control parameters were held constant at the values previously determined as optimum for the unfiltered data as there was no significant improvement by using different values for processing the filtered and autocorrelated waveforms. The parameter settings were:

NPOLES = 9
 FMAX = 8 Hz
 NBEGIN = 15 (the first peak)
 NWINDOW = 10
 IWINDOW = 1.

The performance obtained by the different preprocessing methods is summarized in Fig. 5. It is expressed in terms of the coefficient of variation of the parameter estimates

$$C_x = \frac{\sigma_x}{x_{\text{true}}} , \quad (14)$$

where x is the parameter being estimated ($x = f$ for frequency, β for damping, S for strength) and x_{true} is the true value of the parameter.

The coefficient of variation is plotted in Fig. 5 as a function of the signal-to-noise ratio, SNR, defined by

$$\text{SNR} = \frac{y_{\text{max}}}{\sigma_n} , \quad (15)$$

where

y_{max} is maximum value of the waveform,

σ_n is the standard deviation of the Gaussian noise added to the waveform.

All the plots in Fig. 5 are to the same log-log scales--4 decades in C_x and 2 decades in SNR. The curves for the case of no preprocessing are labeled U (for unfiltered); those for filter preprocessing are labeled F; and those for autocorrelation preprocessing are labeled A. Figures 5(a), (b), and (c) show

The modal damping value β is expressed in units of inverse time (s^{-1}) rather than the usual

$$F_c = \frac{\beta}{\sqrt{\beta^2 + (2\pi f)^2}} \quad , \quad (11)$$

where F_c is the fraction of critical damping. The parameters β and f are estimated as separated quantities by the algorithms, so we decided to evaluate their performance as separate quantities.

The performance of the algorithm was evaluated for three types of preprocessing of the waveforms--none, filtered, and autocorrelated. For the case of filter preprocessing each of the waveforms was truncation-filtered with a passband from 0 to 6.06 Hz, which was accomplished by computing the Fast Fourier Transform (FFT) of each of the waveforms, setting the frequency components outside the passband to zero, and performing an inverse FFT to produce a new time waveform. For the case of autocorrelation preprocessing, the normalized autocorrelation function,

$$R(\tau) = \frac{\int_{-\infty}^{\infty} x(t)x(t+\tau)dt}{R(0)} \quad , \quad (12)$$

where

$R(\tau)$ = autocorrelation value at time τ

$x(t)$ = value of waveform at time t

was computed for each of the waveforms in terms of the discrete sample autocorrelation function according to:

$$R_i = \frac{\sum_{j=1}^{N-i+1} x_j x_{i+j-1}}{\sum_{j=1}^N x_j^2} \quad i = 1, 2, \dots, N \quad , \quad (13)$$

where N = number of points in waveform.

The TDP algorithm was used to process each of the waveforms. All the control parameters were held constant at the values previously determined as optimum for the unfiltered data as there was no significant improvement by using different values for processing the filtered and autocorrelated waveforms. The parameter settings were:

NPOLES = 9
 FMAX = 8 Hz
 NBEGIN = 15 (the first peak)
 NWINDOW = 10
 IWINDOW = 1.

The performance obtained by the different preprocessing methods is summarized in Fig. 5. It is expressed in terms of the coefficient of variation of the parameter estimates

$$C_x = \frac{\sigma_x}{x_{\text{true}}} , \quad (14)$$

where x is the parameter being estimated ($x = f$ for frequency, β for damping, S for strength) and x_{true} is the true value of the parameter.

The coefficient of variation is plotted in Fig. 5 as a function of the signal-to-noise ratio, SNR, defined by

$$\text{SNR} = \frac{y_{\text{max}}}{\sigma_n} , \quad (15)$$

where

y_{max} is maximum value of the waveform,

σ_n is the standard deviation of the Gaussian noise added to the waveform.

All the plots in Fig. 5 are to the same log-log scales--4 decades in C_x and 2 decades in SNR. The curves for the case of no preprocessing are labeled U (for unfiltered); those for filter preprocessing are labeled F; and those for autocorrelation preprocessing are labeled A. Figures 5(a), (b), and (c) show

FIG. 5. Coefficient of variation vs signal-to-noise ratio (SNR).

the variability in the estimates of the frequency C_f , for the 1-, 3-, and 5-Hz modes, respectively. Figures 5 (d), (e), and (f) show the variability in damping C_β for the 1-, 3-, and 5-Hz modes, respectively. Figures 5 (g), (h), and (i) show the variability in the strength C_s for the 1-, 3-, and 5-Hz modes, respectively. No strengths were calculated for the autocorrelated waveforms.

Several conclusions can be drawn from Fig. 5:

- There is no significant difference in coefficient of variation between the unfiltered and filtered data--this will probably not be true for data with closely spaced modes.
- As expected, autocorrelation improves the estimates for the dominant low-frequency mode and degrades the estimates for the weaker high-frequency mode.
- The estimates of the frequencies are an order of magnitude better than those for damping and strength, which are about equal.
- The estimates for the parameters of the 3-Hz mode are an order of magnitude worse than for the 1-Hz mode, reflecting the fact that its strength is about an order of magnitude less. This means the SNR should be defined on a per-mode basis rather than on a per-waveform basis.

A second measure of performance of the algorithm is used to determine if there is a significant bias in the parameter estimates. This is done by performing a t-test against the hypothesis that the estimates are unbiased as described in the Appendix.

The values of t_x (the statistic t for the parameter x) are plotted in Fig. 6 as a function of SNR with the threshold ($t = 2.262$) shown as a dotted line. An unbiased estimator will exceed the threshold only 5 percent of the time. Figures 6 (a), (b), and (c) show t_f the bias test in the frequency estimates for the 1-, 3-, and 5-Hz modes, respectively; Fig. 6 (d), (e), and (f) show t_β the bias test for the damping estimates for the 1-, 3-, and 5-Hz modes, respectively; Fig. 6 (g), (h), and (i) show t_s the bias test in the strength estimates for the 1-, 3-, and 5-Hz modes, respectively. The U, F, and A labels for the curves have the same meaning as in Fig. 5.

FIG. 6. t-test for bias vs signal-to-noise ratio (SNR).

Two conclusions may be drawn from Fig. 6:

- The bias in the estimates is independent of the SNR (all the curves are horizontal lines).
- The TDP algorithm is definably unbiased for the 1- and 3-Hz modes but shows some bias for the 5-Hz mode. Whether this is a function of the relative strengths and frequencies of the modes or the total number of modes will have to be investigated during the closely spaced modes study.

The performance of the TDP algorithm was also evaluated as a function of NBEGIN, the point in the waveform to begin the processing. For this study, the unfiltered ensemble with a signal-to-noise ratio of 100 was used. The other parameter settings were:

```
NPOLES = 9
FMAX   = 8 Hz
NWINDOW = 1 .
```

Note that no sliding-window averaging was done.

The performance was quantified in terms of the normalized rms error in the estimate of the damping:

$$R_B = \frac{\text{rms}_B}{B_{\text{TRUE}}} \quad (16)$$

The variation of R_B as a function of the beginning point for processing is shown in Fig. 7 (a) for the damping estimates of the 1-, 3-, and 5-Hz modes. The values for R_B for the 1-Hz mode are indicated by the character 1 in the plot, while the values for the 3- and 5-Hz modes are indicated by the 3's and 5's, respectively. A sample waveform from the ensemble is plotted in Fig. 7 (b) according to the same time scale showing the variation of performance with the peaks and zero crossings of the waveform.

FIG. 7. Root mean square error in damping vs NBEGIN.

Several conclusions can be reached from observing Fig. 7:

- The ability to estimate the damping for the more highly damped modes (the 5-Hz mode damps out five times faster than the 1-Hz mode) degrades rapidly after the first peak in the waveform.
- There is a considerable variability in the estimates for the weaker modes as a function of NBEGIN. Averaging the results over several values of NBEGIN (i.e., using the sliding-window technique) should reduce the variability by a factor proportional to the square root of the number of "windows" being averaged.
- The performance degrades with an apparently exponential trend as a function of NBEGIN. This is expected since the signal-to-noise ratio is decreasing due to the exponential waveform decay.
- The performance is very poor for NBEGIN = 1 and 2 since they are the times that the excitation is present on the structure.
- There seems to be some evidence that the performance is better for values of NBEGIN that occur at zero crossings for the mode of interest. This can be more thoroughly investigated at a later date.

A third evaluation of the performance of the TDP algorithm was done vs the number of poles (i.e., the number of terms in the exponential series) used to fit the data. The data used were the ensemble with a signal-to-noise ratio of 100. The parameters held constant for this study were:

NBEGIN = 15
FMAX = 8
NWINDOW = 10
IWINDOW = 1.

The performance was quantified in terms of the coefficient of variation for the 1-, 3-, and 5-Hz modes, as shown in Fig. 8. The coefficient of variation for the frequency estimates, C_f , are shown in Fig. 8 (a), (b), and (c) for the 1-, 3-, and 5-Hz modes, respectively. The coefficient of variation for damping C_β is shown in (d), (e), and (f) while that for strength, C_s , is shown in (g), (h), and (i). A vertical dashed line is plotted at NPOLES = 9,

Mode frequencies — Hz

FIG. 8. Coefficient of variation vs NPOLES.

the value chosen as giving optimum performance. Two curves are shown on each of the plots, one for no preprocessing (labeled U) and one for filtered preprocessing (labeled F).

For each of the plots the number of poles requested was varied from 6 to 18. In principle, if there were no noise in the data only 6 poles would be required to estimate accurately the frequency, damping, and strength for the three modes. An observation of Fig. 8 leads to the following conclusions:

- The performance is very poor (by orders of magnitude) for NPOLES less than 9 but reasonably constant for NPOLES greater than 9. Whether this means the optimum NPOLES is (3 plus) or (1.5 times) that in the data is a subject for further research.
- The filter preprocessing reduces the variability by orders of magnitude for NPOLES less than 9 but has little effect for NPOLES greater than 9.
- The filter serves the very useful function of reducing the variability in the estimates when the user is forced to request NPOLES near the true number of poles in the data, which occurs when data are processed from a structure with many modes.

The question of bias in the estimates is addressed with the aid of Fig. 9. The performance in this case is expressed in terms of a t-test for bias (see Appendix) in the frequency, damping, and strength for each mode. The $P = 0.05$ threshold value (2.262) is plotted as a horizontal dotted line.

The t values for the frequency estimates are shown in Fig. 9 (a), (b), and (c) for the 1-, 3-, and 5-Hz modes, respectively. The t values for damping are shown in (d), (e), and (f), and those for strength are shown in (g), (h), and (i).

A vertical dashed line is plotted at NPOLES = 9, the value chosen as giving optimum performance. Two curves are shown on each plot, one for no processing (labeled U) and one for filter preprocessing (labeled F).

Mode frequencies — Hz

FIG. 9. t-test for bias vs NPOLES.

A careful perusal of Fig. 9 leads to the following conclusions:

- For NPOLES less than 9, both filtered and unfiltered data produce biased estimates.
- For most of the parameters being estimated there is a definite minimum in the bias at NPOLES = 9.
- For the 5-Hz mode the bias in the parameter estimates from the filtered data decreases substantially as NPOLES is increased.

SUMMARY AND COMMENTS

The performance of the TDP algorithm is most strongly dependent on signal-to-noise ratio (SNR) in the waveform being processed when the SNR is calculated on a per-mode basis. The relationship for the variability in the estimates is a straight line when plotted on log-log scales, with an order of magnitude decrease in variability caused by an order of magnitude increase in SNR. The best results are obtained when the processing begins at the first peak in the waveform, most likely because the SNR is best at that point and degrades rapidly as the signal decays.

The TDP algorithm evaluation methodology is a good general method for comparing algorithms that process signals that are subject to noise contamination. By expressing performance in terms of variability in the estimates it is possible to derive confidence intervals for the estimates even for algorithms that do not themselves provide them. It is also possible to compare the actual variability against the algorithms' estimates of the variability, for those that do provide them. For example, a version of TDP computes an estimate of the variability when using the sliding-window technique. These estimates could be compared with the values obtained in this report.

ACKNOWLEDGMENT

I wish to acknowledge Dennis Fisher and Michael Posehn for allowing me access to their computational facilities. I also want to thank Edward Luebke for making the task of finding my coding errors a trivial one. I also wish to thank Warren Smith and James Candy for discussing with me the implications of the statistical tests used in this report--they helped my understanding immeasurably.

APPENDIX--BIAS TEST

Given n samples of x , a Gaussian random variable with mean $\bar{x}$ and standard deviation σ representing the estimates of the parameter μ by some algorithm, we wish to determine if that algorithm exhibits a bias error in the estimates. That is, we wish to determine if $\bar{x}$, the true mean of x , is equal to μ , the true value of the parameter. But we can only compute an estimate of $\bar{x}$, so we must test the hypothesis

$$H_0: \bar{x} = \mu$$

against

$$H_1: \bar{x} \neq \mu$$

with probability α of falsely rejecting H_0 . We need to set a threshold τ so that, if H_0 is true, then the deviation in estimate of the mean $\hat{x}$ will exceed the threshold with probability α [i.e., $P(|\hat{x} - \mu| > \tau) = \alpha$].

Since we assume H_0 is true, then x is Gaussian and distributed $N(\bar{x}, \sigma^2)$, and from Brownlee (see Ref. 10, p. 295), $\hat{x}$ is distributed $t(\hat{x}, \hat{\sigma}^2/n, n-1)$, where

$$\hat{\sigma}^2 = \frac{1}{n-1} \sum_{i=1}^n (x_i - \hat{x})^2 \quad . \quad (A-1)$$

Since the standard t-table¹⁰ is for $t(0, 1, n-1)$ the quantities in hypothesis test must be normalized to zero mean and unit variance, giving

$$P \left\{ \left| \frac{\hat{x} - \mu}{\hat{\sigma}/\sqrt{n}} \right| > \tau' \right\} = \alpha \quad , \quad (A-2)$$

where $\tau' = \tau/(\hat{\sigma}/\sqrt{n})$ is from the table. For the case of $n = 10$ and $\alpha = 0.05$, then $\tau' = 2.262$. This implies that the quantity

$$t = \frac{\hat{b}\sqrt{N}}{\hat{\sigma}}, \quad (\text{A-3})$$

where

$$\hat{b} = \hat{\bar{x}} - \mu$$

will exceed the threshold only 5 percent of the time for an unbiased estimator (i.e., for $\bar{x} = \mu$).

REFERENCES

1. D. L. Lager, Using PARET (Parameter Estimation Technique) for the Determination of Mode Shapes, Lawrence Livermore Laboratory, Livermore, Calif., UCID-17898 (1978).
2. A. J. Poggio, M. L. Van Blanicum, E. K. Miller, and R. Mittra, "Evaluation of a Processing Technique for Transient Data," IEEE Trans. on Antennas and Propagation AP-26, 1 (1978).
3. R. J. Lytle, A Remote Method for Determining the Thicknesses and Constitutive Properties of Planar Layered Media, Lawrence Livermore Laboratory, Livermore, Calif., UCRL-51789 (1975).
4. E. K. Miller, F. J. Deadrick, H. G. Hudson, A. J. Poggio, J. A. Landt, Radar Target Classification Using Temporal-Mode Analysis, Lawrence Livermore Laboratory, Livermore, Calif., UCRL-51825 (1975).
5. E. K. Miller, D. L. Lager, Prony's Method for Angle Domain, Lawrence Livermore Laboratory, Livermore, Calif., UCID-17502 (1977).
6. S. Sackett, SAP IV User's Manual, Lawrence Livermore Laboratory, Livermore, Calif., UCID-18226 (1979).
7. R. C. Murray, Lawrence Livermore Laboratory, Livermore, Calif., private communication (October 1978).
8. W. Smith, D. L. Lager, Parametric Characterization of Random Processes Using Prony's Method, Lawrence Livermore Laboratory, Livermore, Calif., UCRL-52673 (1979).
9. M. Schwartz, Information, Transmission, Modulation and Noise (McGraw-Hill, New York, 1959), p. 226.
10. K. A. Brownlee, Statistical Theory and Methodology in Science and Engineering (Wiley and Sons, New York, 1965).

JK:ew