VARIANCES BASED ON DATA WITH DEAD TIME BETWEEN THE MEASUREMENTS *

James A. Barnes
Austron Incorporated
Boulder, Colorado 80301

David W. Allan
Time and Frequency Division
National Bureau of Standards
Boulder, Colorado 80303

ABSTRACT

The accepted definition of frequency stability in the time domain is the two-sample variance (or Allan variance). It is based on the measurement of average frequencies over adjacent time intervals, with no "dead time" between the intervals. The primary advantages of the Allan variance are that (1) it is convergent for many actual noise types for which the conventional variance is divergent; (2) it can distinguish between many important and different spectral noise types; (3) the two-sample approach relates to many practical implementations, for example, the rms change of an oscillator's frequency from one period to the next; and (4) Allan variances can be easily estimated at integer multiples of the sample interval.

In 1974 a table of bias functions which related variance estimates with various configurations of number of samples and dead time to the two-sample (or Allan) variance was published[1]. The tables were based on noises with pure power-law power spectral densities.

Often situations recur which unavoidably have distributed dead time between measurements, but still the conventional variances are not convergent. Some of these applications are outside of the time and frequency field. Also, the dead times are often distributed throughout a given average, and this distributed dead time is not treated in the 1974 tables.

This paper reviews the bias functions $B_1(N, r, \mu)$, and $B_2(r, \mu)$ and introduces a new bias function, $B_3(2, r, \mu)$, to handle the commonly occurring cases of the effect of distributed dead time on the computed variances. Some convenient and easy to interpret asymptotic limits are reported. The actual tables will be published elsewhere in book form.

INTRODUCTION

The most common statistics used are the sample mean and variance. These statistics indicate the approximate magnitude of a quantity and its uncertainty. For many situations a continuous function of time is sampled, or measured, at fairly regular instants. The sampling process (or measurement) is not normally instantaneous but takes a finite time interval to give an "average reading". If the underlying process (or noise) is random and uncorrelated in time, then the fluctuations are said to be "white" noise. In this situation, the sample mean and variance calculated by the conventional formulas,

$$m = \frac{1}{N}\sum_{n=1}^{N}\overline{y}_n$$

$$s^2 = \frac{1}{(n-1)}\sum_{n=1}^{N}(\overline{y}_n - m)^2 \tag{1}$$

provide the needed information. The "bar" over the y in (1) denotes the average over that finite time interval. As in science generally, the physical model determines the appropriate mathematical model. For the white noise model the sample mean and variance are the mainstays of most analyses.

Although white noise is a common model for many physical processes, other noises are becoming increasingly common. In the field of precise time and frequency measurement, for example, there are two quantities of great interest: instantaneous frequency and phase. These two quantities by definition are exactly related by a differential. (We are NOT considering Fourier frequencies at this point.) That is, the instantaneous frequency is the time rate of change of phase. Thus, if we were employing a model of white frequency noise (white FM) this is logically equivalent to phase noise which is the integral of the white noise— commonly called a Brownian motion or a random walk. Depending on whether we are currently interested in phase or frequency the sample mean and variance may or may not be appropriate.

By definition, white noise has a power spectral density (PSD) which does not vary with Fourier frequency. Since random walk noise is the integral of a white noise, it can be shown that the power spectral density of a random walk varies as $1/f^2$ (where f is a Fourier frequency). We encounter noise models whose power spectral densities are various power-laws of their Fourier frequencies. Flicker noise is a very common noise and is defined as a noise whose power spectral density varies as $1/f$ over a significant spectral range. If an oscillator's instantaneous frequency is well modeled by a flicker noise, then its phase would be the integral of the flicker noise. It would have a PSD which varied as $1/f^3$.

Noise models whose PSD's are power laws of the Fourier frequency but not integer exponents are possible but not common. This paper considers power-law PSD's of a quantity $y(t)$; $y(t)$ is a continuous sample function which can be measured at regular intervals. For noises whose PSD's vary as f^α with $\alpha \leq -1$ at low frequencies, the conventional sample mean and variance given in (1) diverge as N gets large[2,3]. This divergence renders the sample mean and variance nearly useless in some situations.

Although the sample mean and variance have limited usefulness, other time-domain statistics can be convergent and quite useful. The quantities which we consider in this paper depend greatly on the details of the sampling. Indeed, each sampling scheme has its own bias, and this is the origin of the bias functions discussed in this paper.

THE ALLAN VARIANCE

As a result of the recognition that the conventional sample variance fails to converge as the number of samples, N, grows, Allan suggested that one can set $N = 2$ and average many of these two-sample variances to get a convergent and stable measure of the spread of the quantity in question[3]. This is what has come to be called the Allan variance.

More specifically let's consider a sample function of time as indicated in Fig. 1. A measurement consists of averaging $y(t)$ over the interval, τ. The next measurement begins a time T after the beginning of the previous measurement interval. There is no logical reason why T must be as large as τ or larger — if $T < \tau$ then the second measurement begins before the first is completed, which is unusual but possible. When $T = \tau$ there is no "dead time" between measurements.

The accepted definition of the Allan variance is the expected value of a two-sample variance with NO dead time between successive measurements. In symbols, the Allan variance is given by:

$$\sigma_y^2(\tau) = 1/2E[(\bar{y}_1 - \bar{y}_2)^2] \tag{2}$$

where there is no dead time between the two sample averages for the Allan variance and the $E[\cdot]$ is the expectation value.

THE BIAS FUNCTION $B_1(N, r, \mu)$

Define N to be the number of sample averages of $y(t)$ used in (1) to estimate a sample variance ($N = 2$ for an Allan variance). Also define r to be the ratio of T to τ ($r = 1$ for no dead time between measurements). The parameter μ is related to the exponent of the power-law of the PSD of the process $y(t)$. It has been shown that if α is the exponent in the power-law spectrum for $y(t)$, then the Allan variance varies as τ raised to the μ power, where α and μ are related as shown in Fig. 2[2–4]. One can use estimates of μ to infer α, the spectral type. The ambiguity in α for $\mu = -2$ has been resolved by using a modified $\sigma_y^2(\tau)$ [5–7].

Often data cannot be taken without dead time between sample averages, and it might be useful to consider other than two-sample variances. We will define the bias function $B_1(N, r, \mu)$ by the ratio:

$$B_1(N, r, \mu) = \frac{\sigma^2(N, T, \tau)}{\sigma^2(2, T, \tau)} \tag{3}$$

where $\sigma^2(N, T, \tau)$ is the expected sample variance given in (1) above, based on N measurements at intervals T and averaged over a time τ and $r = T/\tau$. In words, $B_1(N, T, \mu)$ is the ratio of the expected variance for N measurements to the expected variance for two samples (everything else held constant). Implicitly, the variances on the right in (3) do depend on the noise type (the value of μ or α) even though not shown as an independent variable. N, r andμ, are shown as a independent variables for all of the bias functions in this paper because the values of the ratio of these variances explicitly depends on μ as will be derived later in the paper. Allan showed that if N and r are held constant, then the α, μ relationship shown in Fig. 2 is the same; that is, one can still infer the spectral type from the τ dependence using the equation $\alpha = -\mu - 1, \quad -2 \le \mu < 2$[3].

THE BIAS FUNCTION $B_2(r, \mu)$

The bias function $B_2(r, \mu)$ is defined by the relation

$$B_2(r, \mu) = \frac{\sigma^2(2, T, \tau)}{\sigma^2(2, \tau, \tau)} = \frac{\sigma^2(2, T, \tau)}{\sigma_y^2(\tau)} \tag{4}$$

In words, $B_2(r, \mu)$ is the ratio of the expected variance with dead time to that without dead time (with $N = 2$ and τ the same for both variances). A plot of the $B_2(r, \mu)$ function is shown in Fig. 3. The bias functions B_1 and B_2 represent biases relative to $N = 2$ rather than infinity; that is, the ratio of the N sample variance (with or without dead time) to the Allan variance and the ratio of the two-sample dead-time variance to the Allan variance.

THE BIAS FUNCTION $B_3(N, M, r, \mu)$

Consider the case where measurements are available with dead time between each pair of measurements. The measurements were averaged over the time interval τ_o, and it may not be convenient to retake the data. We might want to estimate the Allan variance at, say, multiples of the averaging time τ_o. If we average groups of the measurements of $y(t)$, then the dead times between each initial measurement are distributed throughout the new average measurement rather than accumulating at the end (see Fig. 4). Define:

$$\bar{\bar{y}}_i = (1/M) \sum_{n=i}^{M+i-1} \bar{y}_n, \tag{5}$$

where $\bar{y}_n$ are the raw or original measurements based on dead time $T_o - \tau_o$.

Also define the variance with distributed dead time as:

$$\sigma^2(2, MT_o, M\tau_o) = 1/2E[(\bar{\bar{y}}_i - \bar{\bar{y}}_{i+M})^2], \tag{6}$$

where $\tau = M\tau_o$ *and* $T = MT_o$.

We can now define B_3 as the ratio of the N-sample variance with distributed dead time to the N-sample variance with dead time accumulated at the end as in Fig. 1:

$$B_3(N, M, r, \mu) = \frac{\sigma^2(N, MT_o, M\tau_o)}{\sigma^2(N, T, \tau)} \tag{7}$$

Although $B_3(N, M, r, \mu)$ is defined for general N, the tables here confine treatment to the case where $N = 2$. There is little value in extending the tables to include general N, and the work becomes excessive. Though the variance on the right in (7) depend explicitly on N, T snf τ, the ratio $B_3(N, M, R, \mu)$ can be shown to depend on the ratio $r = T/\tau$, and explicitly on μ as developed later in this paper.

In words, $B_3(2, M, r, \mu)$ is the ratio of the expected two-sample variance with distributed dead time, as shown in Fig. 4, to the expected two-sample variance with all the dead time grouped together as shown in Fig. 1. Both the numerator and the denominator have the same total averaging time and dead time, but they are distributed differently. The product $B_2(r, \mu) \cdot B_3(2, M, r, \mu)$ is the distributed dead-time variance over the Allan variance for a particular T, τ, M and μ.

Some useful asymptotic forms of B_3 can be found. In the case of large M and $M > r$, we may write that:

$$B_3 \simeq \frac{1+\mu}{3}, \quad 1 \leq \mu \leq 2,$$

$$B_3 \simeq \frac{4\ln(2)}{2\ln(r) + 3}, \quad \mu = 0. \tag{8}$$

One simple and important conclusion from these two equations is that for the cases of flicker noise FM and random walk noise FM the τ^μ dependence is the same whether or not there is distributed dead time for the asymptotic limit. The values of the variances only differ by a constant. This conclusion is also true for white noise FM, and in this case the constant is 1.

In the case $r >> 1$, and $-2 \leq \mu \leq -1$, we may write for the asymptotic behavior of B_3:

$$B_3 \simeq M^\alpha, \quad \alpha = -\mu - 1. \tag{9}$$

In this region of power-law spectrum the B_3 function has an M^α dependence for an f^α spectrum.

THE BIAS FUNCTIONS

The bias functions can be written fairly simply by first defining the function:

$$F(A) = 2A^{\mu+2} - (A+1)^{\mu+2} - |(A-1)|^{\mu+2} \tag{10}$$

The bias functions become:

$$B_1(N,r,\mu) = \frac{1 + \sum_{n=1}^{N-1} \frac{N-n}{N(N-1)} F(nr)}{1 + (1/2)F(r)}, \tag{11}$$

$$B_2(r,\mu) = \frac{1 + (1/2)F(r)}{2(1-2^\mu)}, \tag{12}$$

$$B_3(2,M,r<\mu) = \frac{2M^{\mu+2} + F(MR)\cdot M - \sum_{n=1}^{M-1}(M-n)[2F(nr) - (F(M+n)r) - F(M-n)r)]}{(M^{\mu+2})[F(r)+2]}. \tag{13}$$

For $f = 0$, Equations (11), (12), and (13) are indeterminate of form 0/0 and must be evaluated by L'Hospital's rule. Special attention must also be given when expressions of the form 0^0 arise. We verified a random sampling of the table entries using noise simulation and Monte Carlo techniques. No errors were detected. The results in this paper differ somewhat from those in Ref. 3.

CONCLUSION

For some important power-law spectral density models often used in characterizing precision oscillators ($Sy(f) \sim f^\alpha$, $\alpha = -2,-1,0,+1,+2$), we have studied the effects on variances when dead time exists between the frequency samples and the frequency samples are averaged to increase the integration time. Since dead time between measurements is a common problem throughout metrology, the analysis here has broader applicability than just time and frequency regarding traditional variance analysis.

Heretofore the two-sample zero dead-time Allan variance has been shown to have some convenient theoretical properties in relation to power-law spectra as the integration or sample time is varied (if $\sigma_y^2(\tau) \sim \tau^\mu$, then $\alpha = -\mu - 1$, $\mu > -2$). Since $\sigma_y(\tau)$, by definition, is estimated from data with no dead-time, the sample or integration time can be unambiguously changed to investigate the τ dependence. In this paper we have shown that in the asymptotic limit of several samples being averaged with dead time present in the data, the τ dependence is the same and the $\alpha = -\mu - 1$ relationship still remains valid for white noise frequency modulation (FM) ($\mu = -1$, $\alpha = 0$), flicker noise FM ($\mu = 0$, $\alpha = -1$), and for random walk FM ($\mu = +1$, $\alpha = -2$). The asymptotic limit is approached as the number of samples averaged times the initial data sample time τ_o becomes larger than the dead time ($M > r$). The variances so obtained only differ by a constant, which can be calculated as given in the paper.

A knowledge of the appropriate power-law spectral model (which can, in principle, be estimated from the data) is required to translate a distributed dead-time variance to the corresponding value of the Allan variance.

REFERENCES

1. J.A. Barnes, "Tables of Bias Functions, B1 and B2, for Variances Based on Finite Samples of Processes with Power Law Spectral Densities," NBS Tech Note 375 (1969).

2. J.A. Barnes, "Atomic Timekeeping and the Statistics of Precision Signal Generation," IEEE Proc. 54, No. 2, pp. 207-220, Feb. 1966.

3. D.W. Allan, "Statistics of Atomic Frequency Standards," IEEE Proc. 54, No. 2, pp. 221-230, Feb. 1966.

4. R.F.C. Vessot, et.al., "The Specification of Oscillator Characteristics from Measurements Made in the Frequency Domain," IEEE Proc. 54, No. 2, pp. 199-207, Feb. 1966.

5. D.W. Allan and J.A. Barnes, "A Modified "Allan Variance" With Increased Oscillator Characterization Ability," Proc. 35th Annual Frequency Control Symposium, USAERADCOM, Ft. Monmouth, NJ, May 1981, pp. 470-475.

6. D.W. Allan, "Time and Frequency (Time-Domain) Characterization, Estimation, and Prediction of Precision Clocks and Oscillators," IEEE Transactions on UFFC, Vol.UFFC-34, No. 6, November 1987, pp. 647–654.

7. P. Lesage and T. Ayi, "Characterization of Frequency Stability: Analysis of the Modified Allan Variance and Properties of its Estimate," IEEE Trans. Instrum. Meas., IM-33, no. 4, pp. 332-336, Dec. 1984.

8. N.D. Faulkner and E.V.I. Mestre, "Time-Domain Analysis of Frequency Stability Using Nonzero Dead-Time Counter Techniques." IEEE Trans. on Instrum. & Meas., IM-34, 144-151 (1985).

Table of Some Bias Function Identities

$$
\begin{aligned}
B_1(2,4,\mu) &= 1.000 \\
B_1(N,r,2) &= (N(N+1))/6 \\
B_1(N,1,1) &= N/2 \\
B_1(N,1,\mu) &= (N(1-N^{\mu}))/[(2(N-1)(1-2^{\mu})] \quad \text{for } \mu \neq 0 \\
&= N\ln(N)/[2(N-1)\ln(2)] \quad \text{for } \mu = 0 \\
B_1(N,1,\mu) &= 1.000 \quad \text{for } \mu < 0 \\
&= [2/N(N-1)] \sum_{n-1}^{N-1} (N-n)\cdot n^{\mu} \quad \text{for } \mu > 0 \\
B_1(N,r,-1) &= 1.000 \quad \text{if } r > 1 \\
B_1(N,r,-2) &= 1.000 \quad \text{if } r = 1 \text{ or } 0 \\
B_2(0,\mu) &= 0 \\
B_2(1,\mu) &= 1.000 \\
B_2(r,2) &= r^2 \\
B_2(r,1) &= (3r-1)/2 \text{ if } r > 1 \\
B_2(r,-1) &= r \text{ if } 0 < r < 1 \\
&= 1.000 \text{ if } r > 1 \\
B_2(r,-2) &= 0 \quad \text{if } r = 0 \\
&= 1 \text{ if } r = 1 \\
&= 2/3 \quad \text{otherwise} \\
B_3(2,M,1,\mu) &= 1.000 \\
B_3(2,M,r,-2) &= M \\
B_3(2,r,\mu) &= 1.000 \\
B_3(2,M,r,2) &= 1.000 \\
B_3(2,M,r,-1) &= 1.000 \quad \text{for } r > 1
\end{aligned}
$$

QUESTIONS AND ANSWERS

James Barnes, Austron: I have a brief comment on the history of this. The bias functions, B_1 and B_2, were published some 22 years ago. We considered B_3 at that time, but decided that there was no use for it. We didn't want to encourage people to do that. However, there are people outside the time and frequency community who are acquiring data in this lousy form. They need to be able analyze this data.

Mr. Allan: That's a very good point. The biggest need for it is outside the time and frequency field.