mwcg2.wpd 96/11/4

A Simple Algorithm for Approximating Confidence on the Modified Allan Variance and the Time Variance [1]

Marc A. Weiss
Time & Frequency Div., 847.5
National Institute of Standards and Technology
325 Broadway
Boulder, CO 80303
303-497-3261

Charles A. Greenhall [2]
Jet Propulsion Laboratory
California Institute of Technology
4800 Oak Grove Dr., 298-100
Pasadena, CA 91109
818-393-6944

ABSTRACT
An approximating algorithm for computing equivalent degrees of freedom of the Modified Allan Variance and its square root, the Modified Allan Deviation (MVAR and MDEV), and the Time Variance and Time Deviation (TVAR and TDEV) is presented, along with an algorithm for approximating the inverse chi-square distribution. These two algorithms allow relatively simple computations of confidence intervals on MDEV and TDEV, the latter currently used as a standard in the telecommunications industry. These algorithms enable users to present variance results with confidence intervals corresponding to any useful probability for most data lengths and noise types.

[2] The work of this author was performed at the Jet Propulsion Laboratory, California Institute of Technology, under a contract with the National Aeronautics and Space Administration

Introduction

We present here a simplified algorithm for calculating approximate confidence intervals on the modified Allan deviation, MDEV, and therefore also on the related time variance, TDEV. The algorithm has two parts: the first gives approximate equivalent degrees of freedom, edf, for the fully overlapped estimate of MVAR; the second gives approximate values of the inverse chi-squared distribution. An algorithm for estimating edf for the other measure commonly used in time and frequency metrology, the original Allan deviation, was published previously [1].

Confidence intervals are defined in terms of edf and the chi-squared distribution as follows. If s^2 denotes the usual sample variance of n independent and identically distributed Gaussian measurements (i.e., white noise) with actual variance σ^2, then it is well-known that

$$U = \frac{s^2}{\sigma^2}\cdot\nu. \tag{1}$$

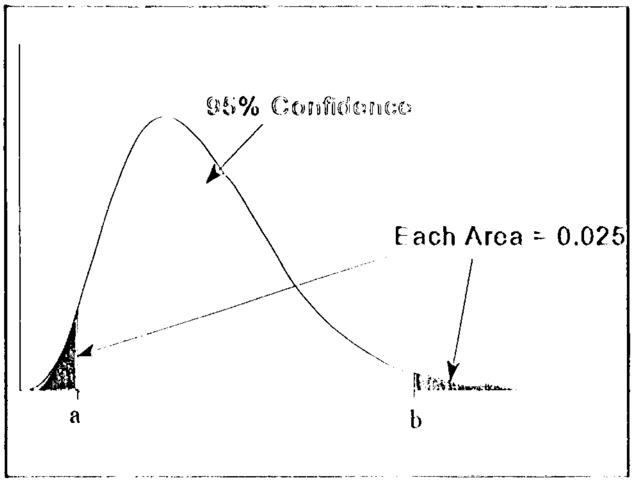

Figure 1 Finding the 95% confidence limits under the chi-squared distribution with 10 degrees of freedom.

has a chi-square distribution with $\nu = n-1$ degrees of freedom [2]. In the classical situation, the degrees of freedom associated with σ^2 are integer values depending only on the number of measurements, and exact confidence limits on the measurement variance are easily calculated using percentiles of the appropriate chi-square distribution. For example, figure 3 shows the chi-square distribution with 10 degrees of freedom, and also depicts the percentiles a and b that are needed to calculate a uncertainty bounds on σ^2 at the $p = 0.95$ confidence level from a particular s^2 based on 11 Gaussian measurements.

A 95% confidence interval is obtained as follows. First we find values a and b such that the probability is 0.95 that U of equation (1) lies between a and b. This condition is equivalent to the following:

$$\begin{gathered} a < \frac{s^2}{\sigma^2}\cdot\nu < b \\ \textit{if and only if} \\ \frac{\nu}{b} < \frac{\sigma^2}{s^2} < \frac{\nu}{a} \\ \textit{if and only if} \\ \frac{\nu}{b}\cdot s^2 < \sigma^2 < \frac{\nu}{a}\cdot s^2. \end{gathered} \tag{2}$$

The lower and upper bounds in the final inequality are confidence limits on the unknown variance σ^2. Note that the confidence factors ν/b and ν/a needed in the calculations are independent of the actual data. They give the magnitude of the confidence interval as a function of the number of points used to compute the variance. Hence we can compute these confidence factors for each noise type and various data lengths. The factors 1-ν/b and ν/a-1 give the multipliers for the magnitude of the lower and upper confidence intervals on the variances, respectively. For deviations such as TDEV, the corresponding multipliers are $1 - \sqrt{\nu/b}$ and $\sqrt{\nu/a} - 1$.

Since the common time and frequency stability measures (AVAR, MVAR, TVAR) are calculated from data arising from non-white noise processes, the confidence limit procedure outlined above is an approximate method [3] that is based on approximating the distribution of U in (1) with the chi-square distribution with degrees of freedom

$$\nu = \frac{2(\sigma^2)^2}{Var\ (s^2)}. \qquad (3)$$

where σ^2 represents the appropriate stability measure (e.g., TVAR), s^2 represents its corresponding estimator, and $Var(s^2)$ is the variance of the s^2 estimator. The quantity ν is called the equivalent degrees of freedom, edf, since it need not be integer-valued.

In this contribution we have combined an algorithm previously published by Greenhall [4] for approximating equivalent degrees of freedom (edf) with an algorithm for approximating the inverse of the chi-squared distribution function. This latter algorithm was derived by Greenhall based on work of Barnes used in deriving tables in [5], but not published, and formulas from Abramowitz and Stegun (A&S) [6]. Previously, tables for confidence of TDEV and MDEV were published in [7]. These are exact computations for edf and the associated confidence intervals for various cases in computing TDEV and MDEV. We compare values approximating the exact edf and confidence factors in tables in [7], finding a worst case disagreement of -9.7% for the edf and +10.8% for the confidence intervals. Most cases are much better than that. The confidence intervals are pessimistic if they are too large and optimistic if they are too small. In many cases here, pessimism is better than optimism, since the true value of the variance is more certain to lie in a larger range than a smaller. For the comparison with the published tables the confidence intervals are no smaller than -3.3%.

Approximation for Equivalent Degrees of Freedom

This version of the formula is restricted to the case of the usual fully overlapped estimator of MVAR or TVAR ([8], Eq. (12); [4], Eq. (6), $m_1 = 1$).

Let:

N = number of time residuals,

m = averaging time / sample period.

M = N-3m+1, the number of terms summed in the estimate,

q = M/m.

Restrictions:

$N \geq 16$,

$m \leq N/5$.

The approximate edf is given by

$$\text{cdf} = \frac{a_0 q}{1 - \frac{a_1}{q}}, \quad (4)$$

where a_0 and a_1 are given in Table 1 as functions of m and the noise type.

Table 1. Coefficients for Approximate cdf Calculation

	$m = 1$		$m = 2$		$m > 2$	
Noise Type	a_0	a_1	a_0	a_1	a_0	a_1
White PM	0.514	0	0.935	0	1.225	0.589
Flicker PM	0.576	0	0.973	0	1.003	0.602
White FM	0.667	0	1.010	0	0.968	0.571
Flicker FM	0.811	0	1.027	0	0.947	0.416
Random-Walk FM	1.000	0	0.866	0	0.768	0.411

Under the assumptions given above, a maximum error of 11.1% in this approximation has been observed. Usually, it is much less.

Approximation for Inverse of Chi-Square Distribution

Let U be a chi-square random variable with df degrees of freedom (df can be nonintegral). Let $0 < p < 1$. Define x = x(p,df) as the 100·p percentile of the distribution of U; thus p is the probability that U < x. The algorithm below computes an approximation to x.

Restrictions:

$$df > 1,$$

$$0.005 \le p \le 0.995 .$$

Maximum observed error with these restrictions: 3%

```
if p <= .5 and df <= 10 then

    !Method: truncate power series in A&S [6] 26.4.6, invert by iteration
    a = df/2
        !Calculate G = Gamma(1+a) (A&S 6.1.35)
        constants
            c1 = -.5748646
            c2 =  .9512363
            c3 = -.6998588
            c4 =  .4245549
            c5 = -.1010678
        n = integer part of a
        y = a - n
        G = 1 + y*(c1 + y*(c2 + y*(c3 + y*(c4 + y*c5))))
        for k = 1 to n  !do nothing if n = 0
```

```
            G = G*(y + k)
        next k
    A = p*G
    u = 0
    for i = 1 to 7
                      u              u              u
        g = 1 + ----- * (1 + ----- * (1 + ----- ))
                    a + 1          a + 2          a + 3
        u = (A*exp(u)/g)^(1/a)
    next i
    x = 2*u

else

    !Method: A&S 26.4.17
    p1 = min(p, 1 - p)
        !Calculate X = inverse of normal distribution at 1-p1 (A&S 26.2.22)
        constants
            a0 = 2.30753
            a1 =  .27601
            b1 =  .99229
            b2 =  .04481
        t = sqrt(-2*log(p1))
                        a0 + a1*t
        X = t - ------------------
                  1 + t*(b1 + t*b2)
    s = sign(p - .5)     !sign(u) = 1 if u > 0, -1 if u < 0, 0 if u = 0
    b = 2/(9*df)
    x = df*(1 - b + s*X*sqrt(b))^3
```

Numerical Example

Before giving tabular results, we show by example how they are used and how they are calculated by the algorithms given above. Assume the situation of the last line of Table II: White PM noise, 1025 time residuals, averaging time = 128 sample periods. Suppose that an MDEV value s is computed by a fully overlapped estimate. The tabulated 95% lower and upper factors are 33.89% and 104.1%. Therefore, a 95% confidence interval for the true MDEV σ is 0.661s to 2.041s.

The tabulated edf and confidence factors are obtained as follows: N = 1025, m = 128, M = 1025 - 3*128 + 1 = 642 (the number of summands in the estimate), q = M/m = 5.0156, a_0 = 1.225, a_1 = 0.589 from Table I, edf = 6.9617 from (4). For 95% confidence we need to compute the 2.5% and 97.5% chi-square levels. The inverse chi-square algorithm, with df = 6.9617 and p = 0.025, gives x = 1.6720 as the 2.5% level, denoted by a in (2). Similarly, the 97.5% level is 15.928, denoted by b. The computed confidence factors are $1 - \sqrt{edf/b} = 0.3389$, $\sqrt{edf/a} - 1 = 1.0405$. (Note that the values in Table II were computed from values of a_0 and a_1 having more significant digits than the ones given in Table I.)

Results

The following data are the results for White PM with fully overlapped estimates. Table II gives the approximate edf and confidence factors. Table III gives the percentage errors from the exact values. The errors for White PM are the largest.

Table II
Approximate edf and Confidence Factors
Noise type: White PM

N	m	edf	lower 68%	upper 68%	lower 95%	upper 95%
17.00	1.000	7.714	17.74	39.14	32.79	94.61
17.00	2.000	5.610	19.67	50.41	36.24	128.7
33.00	1.000	15.94	13.65	23.38	25.52	52.42
33.00	2.000	13.09	14.71	26.67	27.40	60.97
33.00	4.000	7.543	17.87	39.82	33.03	96.58
65.00	1.000	32.40	10.29	14.96	19.46	31.99
65.00	2.000	28.05	10.91	16.33	20.60	35.19
65.00	4.000	17.29	13.23	22.17	24.78	49.37
65.00	8.000	7.241	18.12	41.10	33.47	100.3
129.0	1.000	65.31	7.622	9.916	14.58	20.63
129.0	2.000	57.97	8.030	10.62	15.33	22.17
129.0	4.000	36.86	9.746	13.84	18.48	29.42
129.0	8.000	16.98	13.33	22.43	24.94	50.03
129.0	16.00	7.091	18.24	41.78	33.69	102.3
257.0	1.000	131.1	5.579	6.715	10.76	13.74
257.0	2.000	117.8	5.857	7.123	11.28	14.60
257.0	4.000	76.04	7.128	9.095	13.66	18.84
257.0	8.000	36.55	9.780	13.91	18.54	29.58
257.0	16.00	16.83	13.37	22.56	25.02	50.36
257.0	32.00	7.016	18.31	42.13	33.80	103.3
513.0	1.000	262.8	4.045	4.610	7.856	9.332
513.0	2.000	237.5	4.241	4.867	8.229	9.864
513.0	4.000	154.4	5.177	6.141	10.00	12.53
513.0	8.000	75.73	7.141	9.116	13.68	18.89
513.0	16.00	36.40	9.798	13.95	18.57	29.66
513.0	32.00	16.75	13.40	22.63	25.07	50.53
513.0	64.00	6.978	18.34	42.30	33.86	103.8
1025.	1.000	526.1	2.913	3.195	5.685	6.421
1025.	2.000	476.9	3.052	3.363	5.954	6.766
1025.	4.000	311.1	3.737	4.214	7.267	8.513
1025.	8.000	154.1	5.182	6.148	10.01	12.54
1025.	16.00	75.58	7.148	9.126	13.69	18.91
1025.	32.00	36.32	9.806	13.97	18.59	29.70
1025.	64.00	16.71	13.41	22.67	25.09	50.62
1025.	128.0	6.959	18.36	42.39	33.89	104.1

Table III

Percentage Error: (Approximate-Correct)/Correct

White PM

N	m	cdf	lower 68%	upper 68%	lower 95%	upper 95%
17.00	1.000	-3.426	0.2167	2.221	0.9884	3.178
17.00	2.000	-9.743	2.224	8.438	3.030	10.82
33.00	1.000	-1.647	-0.1272	0.2460	0.4867	1.416
33.00	2.000	-4.167	0.7902	2.026	1.409	3.668
33.00	4.000	3.427	-2.067	-2.940	-1.226	-3.074
65.00	1.000	-0.8256	-0.3389	-0.2002	0.2024	0.4626
65.00	2.000	-1.951	0.1303	0.5085	0.6388	1.291
65.00	4.000	3.878	-2.181	-3.262	-1.523	-2.614
65.00	8.000	-3.550	0.2623	2.416	1.012	3.383
129.0	1.000	-0.4043	-0.4403	-0.4065	0.03683	0.1016
129.0	2.000	-0.9504	-0.2007	-0.1216	0.2545	0.4675
129.0	4.000	4.033	-2.284	-2.982	-1.712	-2.577
129.0	8.000	-2.917	0.4245	1.134	0.9731	2.324
129.0	16.00	-5.303	0.8504	3.895	1.600	5.169
257.0	1.000	-0.1949	-0.5022	-0.4934	-0.08395	-0.07942
257.0	2.000	-0.4919	-0.3668	-0.3280	0.04159	0.09829
257.0	4.000	4.101	-2.351	-2.859	-1.874	-2.472
257.0	8.000	-2.707	0.5069	0.9700	0.9826	1.684
257.0	16.00	-4.608	1.079	2.286	1.639	3.645
257.0	32.00	-5.754	0.9825	4.297	1.761	5.651
513.0	1.000	-0.1132	-0.5156	-0.5110	-0.09529	-0.1023
513.0	2.000	-0.2498	-0.4671	-0.4547	-0.03606	-0.02684
513.0	4.000	4.113	-2.407	-2.751	-1.922	-2.424
513.0	8.000	-2.612	0.5655	0.8747	0.9799	1.437
513.0	16.00	-4.387	1.246	2.038	1.708	2.842
513.0	32.00	-5.041	1.257	2.590	1.809	3.994
513.0	64.00	-5.880	1.050	4.400	1.812	5.797
1025.	1.000	-0.05344	-0.5274	-0.5436	-0.1347	-0.1431
1025.	2.000	-0.1088	-0.5118	-0.5195	-0.09570	-0.1039
1025.	4.000	4.163	-2.462	-2.703	-2.013	-2.351
1025.	8.000	-2.596	0.6215	0.8352	1.038	1.315
1025.	16.00	-4.285	1.341	1.869	1.743	2.494
1025.	32.00	-4.788	1.419	2.241	1.850	3.089
1025.	64.00	-5.150	1.269	2.650	1.853	4.084
1025.	128.0	-5.905	1.028	4.441	1.807	5.829

References

[1] C. A. Greenhall, "Recipes for degrees of freedom of frequency stability estimators," *IEEE Trans. Instrum. Meas.*, vol. 40, pp. 994-999, 1991.

[2] P. G. Hoel, *Introduction to Mathematical Statistics*, New York: John Wiley & Sons, 1971.

[3] Eds. D. B. Sullivan, D. W. Allan, D. A. Howe, and F. L. Walls, *Characterization of Clocks and Oscillators,* National Institute of Standards and Technology Tech Note 1337, 1990.

[4] C. A. Greenhall, "Estimating the Modified Allan Variance," *Proc. 49th Frequency Control Symposium*, pp. 346-353, San Francisco, 1995.

[5] D.A. Howe, D.W. Allan, and J. Barnes "Properties of Signal Sources and Measurement Methods," *Proc. 35th Frequency Control Symposium*, 1981.

[6] M. Abramowitz and I. A. Stegun, *Handbook of Mathematical Functions*, Dover, New York, 1972. Republication, with corrections, of work originally published by the National Bureau of Standards, 1964.

[7] M. A. Weiss, F. L. Walls, C. A. Greenhall, T. Walter, "Confidence on the Modified Allan Variance and the Time Variance," *Proc. 9th European Frequency and Time Forum*, 1995.

[8] D. W. Allan and J. A. Barnes, "A modified Allan variance with increased oscillator characterization ability," *Proc. 35th Frequency Control Symposium*, pp. 470-474, 1981.