THE ESTIMATION OF DISPERSION FROM DIFFERENCES[1]

By Anthony P. Morse[2] and Frank E. Grubbs

Ballistic Research Laboratory, Aberdeen Proving Ground, Maryland

Summary. The estimation of variance by use of successive differences of higher order is discussed in this paper. Heretofore, attention has been focused, in published works, on estimates of variance obtained by employing the sum of squares of deviations from the mean and also by using mean square successive differences of the first order [1], [2], [3], [9]. A concise description of the method employing differences of any order with appropriate formulae for the precision of estimates so obtained and also a practical example on the use of the technique are given in section 11. Fundamental contributions to the estimation of variance from higher order differences, a study of the efficiency of the technique and proper orientation of the subject matter in the field of mathematical statistics are given in sections 2–10 of the paper.

1. Introduction. It frequently happens that successive observations, made at regular intervals of time, are subject to the same standard error while the means of the populations from which they are drawn display some kind of trend. The type of trend we speak of is brought about because of the manner in which we have to take measurements or because of variations in the measuring technique itself; or, again, the trend may be characteristic of the thing we are measuring. In any event, we may desire to eliminate the trend in order to study residual effects. As an example, it is desirable in the field of ballistics to evaluate the dispersion of machine guns firing from a moving airplane.

It may also happen that it is either inexpedient or impossible to estimate the standard error of the observations by the method of least squares, for in a large number of cases the type of trend is unknown. In this event a method employing differences of an appropriate order may prove valuable. The method consists merely of arranging the data in a vertical column in the order in which the observations were taken and then forming difference columns in the usual way of order 1, 2, up to say 5 or some other number depending on the peculiarities of the problem at hand and the number of the original observations. Next, sum the squares of the numbers in each column and divide the sum of squares of the pth order differences by $(n - p)\binom{2p}{p}$. When $n \geq 2$ and $p \geq 1$, the numbers thus arrived at are all unbiased estimates of the population variance σ^2 for the case where all the observations have the same expected value. In section 11 at the

[1] This paper is based substantially on a Ballistic Research Laboratory Report [10] of the same subject by Morse and has been prepared for publication by Grubbs at the suggestion of R. H. Kent. The authors are grateful to J. V. Lewis and H. L. Meyer for their many and varied comments, criticisms and suggestions.

[2] Now at the University of California, Berkeley, California.

end of the paper will be found a summary of this method, formulas by which the precision of the estimate of the variance σ^2 may be determined, and an example displaying the stability of this estimate with respect to p.

If a strong trend is present then the method of first differences will obviously yield an estimate of variance which is fictitiously large and the temptation to pass to higher order differences may quite reasonably be yielded to. As a matter of fact, unbiased estimates may be hoped for from pth order differences whenever there is good reason to suppose that the pth derivative of the trend function is small most of the time. However, even in the case of a sinusoidal trend where all derivatives have the same magnitude one may obtain good results from higher differences provided there are at least seven observations in each interval of length one period (see section 5 and Table II below). In connection with trends such as the sinusoidal type, the hopelessness of getting, say, even a fifth degree polynomial to fit over an interval of, say 20 periods is rather evident. It is for the above reasons that estimation of variance from higher order differences deserves consideration.

2. Historical comment. A brief historical development of the interest in successive differences as a means for estimating dispersion is given in [3]. This paper discusses the statistic

$$s^1 = \sqrt{\frac{\sum_{i=1}^{n/2} (x_{1i} - x_{2i-1})^2}{n}}$$

suggested by "Student" [W. S. Gossett] and E. S. Pearson and points out the relevant work of Jordan, Helmert, Vallier, Cranz, and Becker. It seems that Jordan devised methods based on sums of powers of the differences, whereas Helmert gave more careful consideration to the case of the first power, i.e. the sum of absolute differences. Reference [3] points out, however, that in these two cases all the $n(n - 1)/2$ differences that can be established from a sample of n observations were included in the estimates of dispersion recommended by Jordan and Helmert, so that the estimate was of no value in reducing the effect of a trend. Continuing the remarks of [3], we learn that in ballistics Vallier appears to have been the first to estimate dispersion from successive differences and that Cranz and Becker commended the mean successive difference

$$E_d = \frac{\sum_{i=1}^{n-1} | x_{i+1} - x_i |}{n - 1}$$

in estimating dispersion in range of guns since they were aware of variable external effects (such as tail winds) on a projectile. In this country, Bennett [1] appears to have suggested the use of successive differences independently of European ballisticians. In this connection, Bennett suggested that the probable

TABLE I

The Efficiency, $W(n, p)$, of $\delta^2_{n,p}$ As An Estimate of σ^2

n \ p	1	2	3	4	5	6	7	8	9	10
2	1.00000									
3	.80000	.50000								
4	.75000	.46154	.33333							
5	.72727	.46552	.32000	.25000						
6	.71429	.47213	.33149	.24427	.20000					
7	.70588	.47771	.34453	.25510	.19672	.16667				
8	.70000	.48214	.35537	.26871	.20633	.16471	.14286			
9	.69565	.48568	.36408	.28071	.21888	.17274	.14159	.12500		
10	.69231	.48855	.37113	.29071	.23058	.18385	.14830	.12414	.11111	
11	.68966	.49091	.37691	.29904	.24070	.19476	.15802	.12978	.11050	.10000
12	.68750	.49288	.38173	.30602	.24934	.20450	.16798	.13827	.11529	.09955
13	.68571	.49455	.38580	.31194	.25672	.21300	.17714	.14729	.12271	.10366
14	.68421	.49598	.38928	.31701	.26308	.22039	.18530	.15581	.13086	.11018
15	.68293	.49722	.39228	.32139	.26859	.22684	.19250	.16353	.13874	.11754
16	.68182	.49831	.39490	.32522	.27342	.23251	.19887	.17045	.14601	.12481
17	.68085	.49926	.39721	.32859	.27767	.23752	.20452	.17664	.15260	.13162
18	.68000	.50011	.39925	.33158	.28145	.24197	.20956	.18218	.15855	.13787
19	.67925	.50087	.40107	.33424	.28482	.24595	.21407	.18715	.16393	.14356
20	.67857	.50155	.40271	.33663	.28784	.24953	.21813	.19164	.16879	.14875
21	.67797	.50216	.40419	.33880	.29058	.25276	.22181	.19571	.17321	.15347
22	.67742	.50272	.40553	.34075	.29306	.25569	.22515	.19941	.17723	.15778
23	.67692	.50323	.40675	.34254	.29532	.25837	.22819	.20279	.18091	.16173
24	.67647	.50370	.40787	.34417	.29739	.26082	.23098	.20588	.18428	.16535
25	.67606	.50413	.40889	.34567	.29929	.26307	.23354	.20873	.18738	.16869
26	.67568	.50452	.40984	.34706	.30104	.26514	.23590	.21135	.19024	.17177
27	.67533	.50489	.41071	.34833	.30266	.26705	.23809	.21378	.19289	.17463
28	.67500	.50523	.41152	.34951	.30416	.26884	.24012	.21603	.19535	.17728
29	.67470	.50555	.41228	.35062	.30555	.27049	.24200	.21812	.19764	.17975
30	.67442	.50585	.41298	.35165	.30686	.27203	.24375	.22007	.19978	.18205
31	.67416	.50612	.41363	.35260	.30807	.27347	.24539	.22190	.20177	.18420
32	.67391	.50638	.41425	.35350	.30921	.27482	.24693	.22361	.20364	.18622
33	.67368	.50662	.41482	.35434	.31027	.27608	.24837	.22521	.20539	.18811
34	.67347	.50685	.41536	.35513	.31128	.27727	.24973	.22672	.20704	.18989
35	.67327	.50707	.41587	.35588	.31222	.27839	.25101	.22814	.20859	.19157
36	.67308	.50727	.41635	.35658	.31312	.27945	.25221	.22949	.21006	.19315
37	.67290	.50746	.41671	.35724	.31396	.28045	.25335	.23075	.21145	.19465
38	.67273	.50764	.41724	.35787	.31476	.28140	.25443	.23195	.21276	.19606
39	.67257	.50781	.41765	.35847	.31551	.28229	.25545	.23309	.21401	.19741
40	.67241	.50797	.41804	.35904	.31623	.28314	.25642	.23417	.21519	.19868

TABLE I—*Continued*

n \ p	1	2	3	4	5	6	7	8	9	10
42	.67213	.50828	.41875	.36009	.31756	.28472	.25822	.23617	.21738	.20105
44	.67188	.50855	.41941	.36104	.31877	.28615	.25986	.23799	.21937	.20320
46	.67164	.50880	.42000	.36191	.31987	.28745	.26135	.23965	.22118	.20516
48	.67143	.50903	.42055	.36271	.32088	.28865	.26271	.24117	.22284	.20695
50	.67123	.50925	.42105	.36343	.32180	.28975	.26397	.24256	.22437	.20860
52	.67105	.50944	.42151	.36411	.32266	.29076	.26512	.24385	.22578	.21012
54	.67089	.50962	.42193	.36473	.32345	.29170	.26619	.24504	.22708	.21153
56	.67073	.50979	.42233	.36531	.32418	.29257	.26718	.24614	.22829	.21284
58	.67059	.50995	.42270	.36585	.32487	.29338	.26811	.24717	.22941	.21405
62	.67033	.51022	.42337	.36682	.32609	.29484	.26977	.24903	.23144	.21624
66	.67010	.51048	.42395	.36767	.32718	.29612	.27123	.25065	.23322	.21817
70	.66990	.51069	.42447	.36843	.32813	.29725	.27252	.25209	.23479	.21987
74	.66972	.51089	.42492	.36910	.32898	.29826	.27368	.25237	.23619	.22138
78	.66957	.51107	.42534	.36970	.32975	.29917	.27471	.25452	.23745	.22274
82	.66942	.51122	.42571	.37024	.33043	.29998	.27564	.25556	.23859	.22397
90	.66917	.51150	.42636	.37118	.33162	.30139	.27725	.25735	.24055	.22609
98	.66897	.51172	.42689	.37197	.33262	.30257	.27860	.25885	.24219	.22786
106	.66879	.51192	.42735	.37263	.33346	.30357	.27974	.26012	.24358	.22936
114	.66864	.51208	.42774	.37321	.33418	.30443	.28071	.26121	.24477	.23065
122	.66851	.51223	.42808	.37370	.33482	.30518	.28156	.26216	.24581	.23177
138	.66829	.51247	.42864	.37452	.33585	.30641	.28297	.26372	.24752	.23362
154	.66812	.51266	.42909	.37517	.33667	.30738	.28408	.26496	.24887	.23508
170	.66798	.51281	.42944	.37570	.33734	.30817	.28498	.26596	.24997	.23627
202	.66777	.51304	.43000	.37649	.33836	.30937	.28635	.26749	.25164	.23808
234	.66762	.51322	.43040	.37708	.33909	.31025	.28735	.26860	.25285	.23939
266	.66751	.51335	.43070	.37752	.33965	.31091	.28810	.26944	.25377	.24038
330	.66734	.51353	.43112	.37814	.34044	.31185	.28917	.27063	.25508	.24179
394	.66723	.51365	.43141	.37856	.34097	.31248	.28990	.27143	.25596	.24274
522	.66709	.51381	.43178	.37910	.34164	.31327	.29081	.27244	.25707	.24394
778	.66695	.51396	.43215	.37963	.34233	.31408	.29173	.27347	.25819	.24516
1290	.66684	.51409	.43245	.38007	.34288	.31474	.29248	.27430	.25910	.24613
2314	.66676	.51418	.43264	.38036	.34325	.31518	.29298	.27486	.25971	.24680
∞	.66667	.51429	.43290	.38073	.34372	.31573	.29361	.27556	.26048	.24763

error should be estimated from the root mean square successive differences as follows:

$$P.E. = .6745 \sqrt{\frac{\sum_{i=1}^{n-1} (x_{i+1} - x_i)^2}{2(n - 1)}}.$$

In 1940, J. von Neumann and R. H. Kent in [2] investigated further the estimation of probable error from mean square successive differences (sums of squares of first differences). J. von Neumann, R. H. Kent, H. R. Bellinson, and B. I. Hart [3] considered the distribution of

$$\delta^2 = \frac{1}{n - 1} \sum_{i=1}^{n-1} (x_{i+1} - x_i)^2$$

in a paper which appeared in June 1941. J. D. Williams [4] obtained the moments of $\eta = \frac{\delta^2}{s^2}$, where

$$s^2 = \frac{1}{n} \sum_{i=1}^{n} (x_i - \bar{x})^2,$$

and indicated that the rth moment of η is equal to the rth moment of δ^2 divided by the rth moment of s^2. The distribution of the ratio of the mean square successive difference to the variance has been published by J. von Neumann [5], [6] and B. I. Hart tabulated the probability integral and obtained percentage points for this statistic ([7], [8]). Indeed, it should be remarked that the statistical theory of successive differences is allied with the problem of serial correlation [9]. Finally, the use of squared differences of higher order than the first for estimating variance appears to have been suggested by A. A. Bennett. Quite independently, a treatment of the subject was given by Morse [10] in connection with problems on exterior ballistics. Various results on successive-difference estimation including significance tests have been given by Tintner [13]. One of Tintner's tests involves the use of selected sets of differences.

3. Definitions and notations. Suppose the observations $x_1, x_2, x_3, \cdots x_n$ are made at times $a = t_1 < t_2 < t_2 < \cdots < t_n = b$ and the t_i are uniformly spaced without error. Let $f(t_i)$ be the true trend so that $\eta_i = f(t_i)$ is the mean of the population from which x_i is drawn and $\epsilon_i = x_i - \eta_i$ is a random error. Further, let p be a non-negative integer less than n and denote to the ith backward difference of order p of x by $\Delta^p x_i$, i.e.

$$\Delta^p x_i = \Delta^{p-1} x_i - \Delta^{p-1} x_{i-1} = \sum_{r=0}^{p} (-1)^r \binom{p}{r} x_{i-r},$$

$$\text{where } \binom{m}{n} = \frac{m!}{n!(m - n)!}; \qquad \text{and } i \geq p + 1.$$

We define the following:

(1) $$\delta^2_{n,p} = \frac{1}{\binom{2p}{p}(n-p)} \sum_{i=p+1}^{n} (\Delta^p \epsilon_i)^2;$$

(2) $$d^2_{n,p} = \frac{1}{\binom{2p}{p}(n-p)} \sum_{i=p+1}^{n} (\Delta^p x_i)^2;$$

(3) $$\nu^2_{n,p} = \frac{1}{\binom{2p}{p}(n-p)} \sum_{i=p+1}^{n} (\Delta^p \eta_i)^2;$$

(4) $$k_{n,p} = \frac{2}{\binom{2p}{p}(n-p)} \sum_{i=p+1}^{n} (\Delta^p \eta_i)(\Delta^p \epsilon_i).$$

By $E(u)$ we will mean the expected value of u, whereas the variance of u will be denoted by

$$\mathrm{Var}\,(u) = E\{u - E(u)\}^2.$$

Basically, we shall assume that the ϵ_i are sufficiently Gaussian and independent that

$$E(\epsilon_i) = E(\epsilon_i^3) = 0, \qquad E(\epsilon_i^2) = \sigma^2,$$
$$\mu_4 = E(\epsilon_i^4) = 3\sigma^4,$$
$$E(\epsilon_i^\alpha \epsilon_j^\beta) = E(\epsilon_i^\alpha)E(\epsilon_j^\beta),$$

whenever i, j, α and β are positive integers for which

$$i \neq j, \qquad 1 \leq i \leq n, \qquad 1 \leq j \leq n.$$

4. Expected values. We will now determine the mean or expected values of $\delta^2_{n,p}$ and $d^2_{n,p}$.

$$E(\delta^2_{n,p}) = \frac{1}{\binom{2p}{p}(n-p)} \sum_{i=p+1}^{n} E\left\{\sum_{r=0}^{p} (-1)^r \binom{p}{r} \epsilon_{i-r}\right\}^2,$$

$$E(\delta^2_{n,p}) = \frac{1}{\binom{2p}{p}} \sum_{r=0}^{p} \binom{p}{r}^2 \sigma^2.$$

or

(5) $$E(\delta^2_{n,p}) = \sigma^2.$$

(see Lemma 1.3 of section 6 below),

Continuing, we have

$$E(d_{n,p}^2) = \frac{1}{\binom{2p}{p}(n-p)} E\left\{\sum_{i=p+1}^{n} (\Delta^p \epsilon_i + \Delta^p \eta_i)^2\right\},$$

$$E(d_{n,p}^2) = \frac{1}{\binom{2p}{p}(n-p)} \left\{(n-p)\binom{2p}{p}\sigma^2 + \sum_{i=p+1}^{n} (\Delta^p \eta_i)^2\right\},$$

or

$$E(d_{n,p}^2) = \sigma^2 + \nu_{n,p}^2 . \tag{6}$$

Consequently, we observe, $d_{n,p}^2$ is on the average larger than σ^2 by the quantity $\nu_{n,p}^2$. In a particular problem, therefore, we are faced with the situation of choosing that combination of n and p which (i) regulates the size of $\nu_{n,p}^2$ and (ii) gives the desired precision of our estimate of variance.

5. The magnitude of $\nu_{n,p}^2$. In order to study the size of $\nu_{n,p}^2$, we will derive for this quantity an upper bound which will indicate the applicability of the method of differences to non-polynomial as well as polynomial trends.

Now,

$$\Delta^p \eta_i = \Delta^p f(t_i) = \int_{t_{i-1}}^{t_i} \int_0^h \cdots \int_0^h f^{(p)}(y_1 - y_2 - \cdots - y_p)\, dy_p\, dy_{p-1} \cdots dy_1 ,$$

where $t_r - t_{r-1} = h$, by straightforward integration. It will be convenient tô change the order of integration; thus

$$\Delta^p f(t_i) = \int_0^h \cdots \int_0^h \int_{t_{i-1}}^{t_i} f^{(p)}(y_1 - y_2 - \cdots - y_p)\, dy_1\, dy_p \cdots dy_2 .$$

Since, from Schwarz's inequality it is clear that

$$\left\{\int_\alpha^\beta g(s)\, ds\right\}^2 \le (\beta - \alpha) \int_\alpha^\beta \{g(s)\}^2\, ds$$

whenever α and β are real numbers and g is integrable, we have

$$\{\Delta^p \eta_i\}^2 \le h^p \int_0^h \cdots \int_0^h \int_{t_{i-1}}^{t_i} \{f^{(p)}(y_1 - y_2 - \cdots - y_p)\}^2\, dy_1\, dy_p \cdots dy_2 .$$

Also,

$$\sum_{i=p+1}^{n} \{\Delta^p \eta_i\}^2 \le h^p \int_0^h \cdots \int_0^h \int_{t_p}^{t_n} \{f^{(p)}(y_1 - y_2 - \cdots - y_p)\}^2\, dy_1\, dy_p \cdots dy_2 .$$

But for $0 \le r \le (p-1)h = t_p - a$ we have

$$\int_{t_p}^{t_n} \{f^{(p)}(y_1 - r)\}^2\, dy_1 = \int_{t_p - r}^{t_n - r} \{f^{(p)}(s)\}^2\, ds \le \int_a^b \{f^{(p)}(s)\}^2\, ds.$$

Consequently

$$\sum_{i=p+1}^{n} (\Delta^p \eta_i)^2 \leq h^p \int_0^h \cdots \int_0^h \int_a^b \{f^{(p)}(s)\}^2 \, ds \, dy_p \cdots dy_2 = h^{2p-1} \int_a^b \{f^{(p)}(s)\}^2 \, ds.$$

Since $h = \dfrac{b-a}{n-1}$, we have finally

$$\nu_{n,p}^2 \leq \frac{1}{\binom{2p}{p}} \left(\frac{b-a}{n-p}\right) \left(\frac{b-a}{n-1}\right)^{2p-1} \int_a^b \frac{\{f^{(p)}(s)\}^2 \, ds}{b-a}, \tag{7}$$

which is an upper bound for $\nu_{n,p}^2$ in terms of the average value of the square of the pth derivative of the trend function f.

If the trend function f is of the polynomial form,

$$f(t) = \sum_{r=0}^{p} a_r t^r$$

then the effect of the trend can be eliminated from our observations by estimating dispersion from $(p + 1)$st differences. However, if it is *known* that the trend is of polynomial form, then an estimate of dispersion based on least squares would, of course, be better. In fact, it will be shown later that the precision of $\delta_{n,p}^2$ decreases markedly as p increases. The use of $d_{n,p}^2$ as an estimate of σ^2 is primarily of value when the type of trend is unknown; however, even when the type of trend is known the computational simplicity of $d_{n,p}^2$ may offset to some extent its lack of optimum precision.

Let us reflect on the magnitude of $\nu_{n,p}^2$ over a single period of a sinusoidal trend, say $f(t) = \sin t$. In (7) we set $a = 0$, $b = 2\pi$ and secure

$$\nu_{n,p}^2 \leq \frac{\pi}{\binom{2p}{p}(n-p)} \left(\frac{2\pi}{n-1}\right)^{2p-1}.$$

Taking n to be the number of observations for a complete period, a tabulation of the upper bound for $\nu_{n,p}^2$ for this case is given in Table II. Thus, when there are about seven or more observations in each interval of length one period, estimation of dispersion from higher order differences may prove of considerable value even for this rather extreme type of trend.

6. Some combinatorial relations. Although we will ultimately establish expressions for the variances of $\delta_{n,p}^2$ and $d_{n,p}^2$, it appears desirable to give first a number of combinatorial relations which present themselves in the computation of moments. The relations are easily checked and most of them are possibly well known. Nevertheless, it will be convenient to record them for reference and in some instances to give proofs. In what follows it will be understood that $\binom{p}{q} = 0$ whenever p and q are not such integers that $0 \leq q \leq p$.

TABLE II

p \ n	5	6	7	8	9	10
1	.617	.395	.274	.201	.154	.110
2	.676	.260	.120	.063	.036	.016
3	.751	.164	.049	.018	.008	.002
4	.106	.111	.021	.005	.002	.0003
5	—	.098	.009	.002	.0004	.0000

LEMMA 1.1. $q\binom{p}{q} = p\binom{p-1}{q-1}$.

LEMMA 1.2. $\binom{p}{r} = \binom{p}{p-r}$.

LEMMA 1.3. $\sum_r \binom{p}{r}\binom{p}{r+s} = \binom{2p}{p+s}$.

PROOF:

$$\sum_s \binom{2p}{s} x^s = (1+x)^{2p} = \{(1+x)^p\}^2 = \left\{\sum_s \binom{p}{s} x^s\right\}^2$$
$$= \sum_s \sum_r \binom{p}{r}\binom{p}{s-r} x^s.$$

Hence

$$\binom{2p}{s} = \sum_r \binom{p}{r}\binom{p}{s-r},$$

and

$$\binom{2p}{p+s} = \sum_r \binom{p}{r}\binom{p}{p+s-r} = \sum_r \binom{p}{r}\binom{p}{r-s} = \sum_r \binom{p}{r+s}\binom{p}{r}.$$

LEMMA 1.4. *If* $p^2 + r^2 > 0$ *then* $\binom{p}{r} = \binom{p-1}{r} + \binom{p-1}{r-1}$.

LEMMA 1.5. $(p-2r)\binom{p}{r} = p\left\{\binom{p-1}{r} - \binom{p-1}{r-1}\right\}$.

LEMMA 1.6. $(p-2r)\binom{p}{r}^2 = p\left\{\binom{p-1}{r}^2 - \binom{p-1}{r-1}^2\right\}$.

PROOF: Multiply, using 1.4 and 1.5.

LEMMA 1.7.[3] $r\binom{2p}{p+r}^2 = p\left\{\binom{2p-1}{p-r}^2 - \binom{2p-1}{p-r-1}^2\right\}$.

[3] Major A. A. Bennett communicated this Lemma.

PROOF: $(s-2t)\binom{s}{t}^2 = s\left\{\binom{s-1}{t}^2 - \binom{s-1}{t-1}^2\right\}$ from 1.6.

Put $s = 2p$, $t = p - r$, then

$$2r\binom{2p}{p+r}^2 = 2p\left\{\binom{2p-1}{p-r}^2 - \binom{2p-1}{p-r-1}^2\right\}.$$

LEMMA 1.8. *If f is a function, i, n, p are integers and $p + 1 \leq i \leq n$, then*

$$\sum_{r=1}^{n}\binom{p}{i-r}f(i-r) = \sum_{r=0}^{p}\binom{p}{r}f(r).$$

PROOF:

$$\sum_{r=1}^{n}\binom{p}{i-r}f(i-r) = \sum_{s=i-n}^{i-1}\binom{p}{s}f(s) = \sum_{r=0}^{p}\binom{p}{r}f(r).$$

LEMMA 1.9. *If $-\infty < A(r, s) = A(s, r) < \infty$ for each integer r and s, then*

$$E\left(\left\{\sum_{r=1}^{n}\sum_{s=1}^{n}A(r,s)\epsilon_r\epsilon_s\right\}^2\right) = (\mu_4 - 3\sigma^4)\sum_{r=1}^{n}A(r,r)^2 + \sigma^4\left\{\sum_{r=1}^{n}A(r,r)\right\}^2 + 2\sigma^4\sum_{r=1}^{n}\sum_{s=1}^{n}A(r,s)^2.$$

PROOF: Let $N(r, s) = 1$ when $r < s$ and let $N(r, s) = 0$ otherwise. Clearly

$$\sum_{r=1}^{n}\sum_{s=1}^{n}A(r,s)\epsilon_r\epsilon_s = \sum_{r=1}^{n}A(r,r)\epsilon_r^2 + 2\sum_{r=1}^{n}\sum_{s=1}^{n}N(r,s)A(r,s)\epsilon_r\epsilon_s,$$

and

$$E\left(\left\{\sum_{r=1}^{n}\sum_{s=1}^{n}A(r,s)\epsilon_r\epsilon_s\right\}^2\right) = E\left(\left\{\sum_{r=1}^{n}A(r,r)\epsilon_r^2\right\}^2\right) + 4E\left(\left\{\sum_{r=1}^{n}\sum_{s=1}^{n}N(r,s)A(r,s)\epsilon_r\epsilon_s\right\}^2\right).$$

Now

$$E\left(\left\{\sum_{r=1}^{n}A(r,r)\epsilon_r^2\right\}^2\right) = (\mu_4 - \sigma^4)\sum_{r=1}^{n}A(r,r)^2 + \sigma^4\left\{\sum_{r=1}^{n}A(r,r)\right\}^2,$$

and

$$\begin{aligned} 4E\left(\left\{\sum_{r=1}^{n}\sum_{s=1}^{n}N(r,s)A(r,s)\epsilon_r\epsilon_s\right\}^2\right) &= 4\sigma^4\sum_{r=1}^{n}\sum_{s=1}^{n}N(r,s)A(r,s)^2 \\ &= 2\sigma^4\sum_{r=1}^{n}\sum_{s=1}^{n}A(r,s)^2 - 2\sigma^4\sum_{r=1}^{n}A(r,r)^2 \end{aligned}$$

The last three relations combine to yield the desired result.

LEMMA 1.10. $\binom{2p}{p}^2 (n - p)^2 E(\delta_{n,p}^4)$

$$= (\mu_4 - 3\sigma^4) \sum_{r=1}^{n} \left\{ \sum_{i=p+1}^{n} \binom{p}{i-r}^2 \right\}^2 + \sigma^4 \left\{ \sum_{r=1}^{n} \sum_{i=p+1}^{n} \binom{p}{i-r}^2 \right\}^2 + 2\sigma^4 \sum_{r=1}^{n} \sum_{s=1}^{n} \left\{ \sum_{i=p+1}^{n} \binom{p}{i-r}\binom{p}{i-s} \right\}^2.$$

PROOF: Helped by 1.8, check that

$$(\Delta_i^p \epsilon)^2 = \left\{ \sum_{r=0}^{p} (-1)^r \binom{p}{r} \epsilon_{i-r} \right\}^2 = \left\{ \sum_{r=1}^{n} (-1)^{i-r} \binom{p}{i-r} \epsilon_r \right\}^2$$
$$= \sum_{r=1}^{n} \sum_{s=1}^{n} (-1)^{r+s} \binom{p}{i-r}\binom{p}{i-s} \epsilon_r \epsilon_s .$$

Therefore

$$\binom{2p}{p} (n - p)\delta_{n,p}^2 = \sum_{r=1}^{n} \sum_{s=1}^{n} \left\{ (-1)^{r+s} \sum_{i=p+1}^{n} \binom{p}{i-r}\binom{p}{i-s} \right\} \epsilon_r \epsilon_s .$$

Let

$$A(r, s) = (-1)^{r+s} \sum_{i=p+1}^{n} \binom{p}{i-r}\binom{p}{i-s},$$

and apply 1.9 to complete the proof.

LEMMA 1.11.

$$\sum_{r=1}^{n} \sum_{s=1}^{n} \left\{ \sum_{i=p+1}^{n} \binom{p}{i-r}\binom{p}{i-s} \right\}^2$$
$$= (n - p) \sum_{r=p-n}^{n-p} \binom{2p}{p+r}^2 - 2p \binom{2p-1}{p}^2 + 2p \binom{2p-1}{n}^2 .$$

PROOF.

$$\sum_{r=1}^{n} \sum_{s=1}^{n} \left\{ \sum_{i=p+1}^{n} \binom{p}{i-r}\binom{p}{i-s} \right\}^2$$
$$= \sum_{i=p+1}^{n} \sum_{j=p+1}^{n} \sum_{r=1}^{n} \sum_{s=1}^{n} \binom{p}{i-s}\binom{p}{j-s}\binom{p}{i-r}\binom{p}{j-r}$$
$$= \sum_{i=p+1}^{n} \sum_{j=p+1}^{n} \sum_{r=1}^{n} \sum_{s=0}^{p} \binom{p}{s}\binom{p}{s+j-i}\binom{p}{i-r}\binom{p}{j-r}, \text{ using 1.8;}$$
$$= \sum_{i=p+1}^{n} \sum_{j=p+1}^{n} \sum_{r=0}^{p} \sum_{s=0}^{p} \binom{p}{s}\binom{p}{s+j-i}\binom{p}{r}\binom{p}{r+j-i}, \text{ using 1.8 again;}$$

$$= \sum_{i=p+1}^{n} \sum_{j=p+1}^{n} \sum_{r} \binom{p}{r} \binom{p}{r+j-i} \sum_{s} \binom{p}{s} \binom{p}{s+j-i}$$

$$= \sum_{i=p+1}^{n} \sum_{j=p+1}^{n} \left\{ \sum_{r} \binom{p}{r} \binom{p}{r+j-i} \right\}^2 = \sum_{i=p+1}^{n} \sum_{j=p+1}^{n} \binom{2p}{p+j-i}^2, \text{ from 1.3;}$$

$$= \sum_{r=p+1-n}^{n-p-1} (n - p - |r|) \binom{2p}{p+r}^2$$

$$= \sum_{r=p-n}^{n-p} (n - p - |r|) \binom{2p}{p+r}^2$$

$$= (n-p) \sum_{r=p-n}^{n-p} \binom{2p}{p+r}^2 - 2 \sum_{r=0}^{n-p} r \binom{2p}{p+r}^2$$

$$= (n-p) \sum_{r=p-n}^{n-p} \binom{2p}{p+r}^2 - 2 \sum_{r=0}^{n-p} p \left\{ \binom{2p-1}{p-r}^2 - \binom{2p-1}{p-r-1}^2 \right\}, \text{ using 1.7;}$$

$$= (n-p) \sum_{r=p-n}^{n-p} \binom{2p}{p+r}^2 - 2p \left\{ \binom{2p-1}{p}^2 - \binom{2p-1}{2p-n-1}^2 \right\}$$

$$= (n-p) \sum_{r=p-n}^{n-p} \binom{2p}{p+r}^2 - 2p \binom{2p-1}{p}^2 + 2p \binom{2p-1}{n}^2.$$

LEMMA 1.12.

$$\sum_{r=1}^{n} \sum_{i=p+1}^{n} \binom{p}{i-r}^2 = (n-p) \binom{2p}{p}.$$

PROOF.

$$\sum_{r=1}^{n} \sum_{i=p+1}^{n} \binom{p}{i-r}^2 = \sum_{i=p+1}^{n} \sum_{r=1}^{n} \binom{p}{i-r}^2 = \sum_{i=p+1}^{n} \sum_{r=0}^{p} \binom{p}{r}^2, \text{ from 1.8;}$$

$$= (n-p) \sum_{r} \binom{p}{r}^2 = (n-p) \binom{2p}{p}.$$

7. The variances of $\delta^2_{n,p}$ and $d^2_{n,p}$. In order to get some idea as to the efficiency of the statistics $\delta^2_{n,p}$ and $d^2_{n,p}$, we will examine their variances. We have

$$\binom{2p}{p}^2 (n-p)^2 \operatorname{Var}(\delta^2_{n,p}) = \binom{2p}{p}^2 (n-p)^2 \{E(\delta^4_{n,p}) - [E(\delta^2_{n,p})]^2\}$$

$$= \binom{2p}{p}^2 (n-p)^2 \sigma^4 + 2(n-p)\sigma^4 \sum_{r=p-n}^{n-p} \binom{2p}{p+r}^2 - 4p\sigma^4 \binom{2p-1}{p}^2$$

$$+ 4p\sigma^4 \binom{2p-1}{n}^2$$

with the aid of Lemmas 1.10, 1.11, 1.12 and using the relation $\mu_4 - 3\sigma^4 = 0$. Thus,

$$\binom{2p}{p}^2 (n-p)^2 \operatorname{Var}(\delta^2_{n,p}) = 2(n-p)\sigma^4 \sum_{r=p-n}^{n-p} \binom{2p}{p+r}^2 - 4p\sigma^4\binom{2p-1}{p}^2 + 4p\sigma^4\binom{2p-1}{n}^2. \tag{8}$$

If $2p \leq n$, then

$$\sum_{r=p-n}^{n-p}\binom{2p}{p+r}^2 \sum_r \binom{2p}{p+r}^2 = \sum_r \binom{2p}{r} = \binom{4p}{2p}.$$

Moreover, $\binom{2p-1}{n} = 0.$

Therefore,

$$\binom{2p}{p}^2 (n-p)^2 \operatorname{Var}(\delta^2_{n,p}) = 2(n-p)\binom{4p}{2p}\sigma^4 - 4p\binom{2p-1}{p}^2\sigma^4 \tag{9}$$

when $2p \leq n$.
As for the variance of $d^2_{n,p}$, we have

$$\operatorname{Var}(d^2_{n,p}) = E\{d^2_{n,p} - \nu^2_{n,p} - \sigma^2\}^2 = E\{\delta^2_{n,p} + k_{n,p} + \nu^2_{n,p} - \nu^2_{n,p} - \sigma^2\}^2$$
$$= E\{\delta^2_{n,p} - \sigma^2) + k_{n,p}\}^2,$$

or

$$\operatorname{Var}(d^2_{n,p}) = \operatorname{Var}(\delta_{n,p})^2 + E(k^2_{n,p}), \tag{10}$$

since $E[(\delta^2_{n,p} - \sigma^2)k_{n,p}] = 0$.
However, from Schwarz's inequality, it is guaranteed that

$$E(k^2_{n,p}) \leq 4\nu^2_{n,p}\sigma^2.$$

Thus

$$\operatorname{Var} d^2_{n,p} \leq \operatorname{Var}(\delta^2_{n,p}) + 4\nu^2_{n,p}\sigma^2. \tag{11}$$

An upper bound has already been given for $\nu^2_{n,p}$ in section 5 above.

8. The efficiency of $\delta^2_{n,p}$. It is appropriate to consider the efficiency (as defined by Fisher [11]) of the statistic $\delta^2_{n,p}$. In this sense, the efficiency of $\delta^2_{n,p}$ is given by

$$W(n,p) = \frac{\operatorname{Var} s^2_n}{\operatorname{Var} \delta^2_{n,p}}. \qquad \text{where } s^2_n = \frac{\sum (x_i - \bar{x})^2}{n-1}.$$

Accordingly,

$$W(n, p) = \frac{2\sigma^4}{(n-1)\,\mathrm{Var}\,(\delta^2_{n,p})}$$

or

$$W(n, p)$$

$$= \frac{(n-p)^2 \binom{2p}{p}^2}{(n-1)\left\{(n-p)\sum_{r=p-n}^{n-p}\binom{2p}{p+r}^2 - 2p\binom{2p-1}{p}^2 + 2p\binom{2p-1}{n}^2\right\}}. \tag{12}$$

If $2p \leq n$

$$W(n. p) = \frac{(n-p)^2\binom{2p}{p}^2}{(n-1)\left\{(n-p)\binom{4p}{2p} - 2p\binom{2p-1}{p}^2\right\}}, \qquad \text{from (9);} \tag{13}$$

or

$$W(n, p) = \frac{\binom{2p}{p}^2}{\binom{4p}{2p}\left\{1 - \frac{p-1}{n-p}\right\}\left\{1 - \frac{2p\binom{2p-1}{p}^2}{(n-p)\binom{4p}{2p}}\right\}}, \qquad \text{if } 2p \leq n. \tag{14}$$

Formulas (12) and (13) were used in preparing Table I given at the end of the paper. For convenience in using formulas (1) and (2) the binomial coefficients $\binom{2p}{p}$ for $0 \leq p \leq 10$ are given in Table III.

If $n \geq 2$, then

$$W(n, 1) = \frac{2}{3}\cdot\frac{1}{1 - \frac{1}{3n-3}} = \frac{2(n-1)}{3n-4}, \tag{15}$$

as was pointed out by von Neumann, Kent, Bellinson, and Hart in [3].

If $n \geq 4$, then

$$W(n, 2) = \frac{18}{35}\frac{1}{\left\{1 + \frac{1}{n-2}\right\}\left\{1 - \frac{18}{35(n-2)}\right\}} = \frac{18(n-2)^2}{(n-1)(35n-88)}. \tag{16}$$

As a limiting value for n, we have

$$W(\infty, p) = \lim_{n \to \infty} W(n, p) = \frac{\binom{2p}{p}^2}{\binom{4p}{2p}}. \tag{17}$$

Using Stirling's formula for the approximation to the factorial, we have

$$\lim_{p \to \infty} \sqrt{p}\, W(\infty, p) = \sqrt{\frac{2}{\pi}}.$$

Thus, as $p \to \infty$, $W(\infty, p)$ tends to zero and is asymptotically equal to $\sqrt{\frac{2}{\pi p}}$

TABLE III

The Binomial Coefficient $\binom{2p}{p}$

p	$\binom{2p}{p}$
0	1
1	2
2	6
3	20
4	70
5	252
6	924
7	3432
8	12870
9	48620
10	184756

For the case $n \geq 2$, $p \geq 1$ and f *constant*, then $s_n^2 = \frac{\Sigma (x_i - \bar{x})^2}{n - 1}$ and $\delta_{n,p}^2$ and $d_{n,p}^2$ are all unbiased estimates of the population variance σ^2. Moreover, for this case

$$W(n, p) = \frac{\text{Var}\,(s_n^2)}{\text{Var}\,(\delta_{n,p}^2)} = \frac{\text{Var}\,(s_n^2)}{\text{Var}\,(d_{n,p}^2)}.$$

Using s_m^2 based on $m - 1$ degrees of freedom and keeping the trend, f, constant, then m and n may be chosen so that approximately

$$\text{Var}\,(s_m^2) = \text{Var}\,(d_{n,p}^2)$$

and for a normal population this means that

$$m = 1 + (n - 1) W(n, p).$$

Using Table I, it may be seen that for constant trend, f, the worth of $d^2_{50,10}$ as an estimate of σ^2 for a normal population is about the same as that of s^2_{11}, whereas that of $d^2_{50,1}$ is about equivalent to s^2_{40}. However, if the trend f is not constant then the worth of s^2_n as an estimate of σ^2 is diminished while that of $d^2_{n,p}$ is increased.

Similarly, if the trend is cubic over 20 observations then least squares gives an unbiased estimate of σ^2 based on 16 degrees of freedom, whereas $d^2_{20,4}$ gives an estimate equivalent in precision to about 6.4 degrees of freedom. However, if only eight observations follow a cubic trend, then least squares furnish an unbiased estimate of σ^2 based on four degrees of freedom whereas $d^2_{8,4}$ furnishes an estimate equivalent to about 1.9 degrees of freedom. Thus, in the case of 20 observations, cubic least squares is, so to speak, 2.5 times as valuable as $d^2_{20,4}$; in the case of eight observations, cubic least squares is 2.1 times as valuable as $d^2_{8,4}$.

It might be mentioned that the method of differences is of value in estimating goodness of fit. If the fit is good, then our estimate of σ^2 derived from least squares should on the average be equal to the estimate derived from a suitable $d^2_{n,p}$. If the fit is poor then $d^2_{n,p}$ will be smaller on the average than the former.

9. The approximate probable error in estimating σ from differences. The approximate standard error of $\delta_{n,p}$ is given by the relation

$$\text{S.E. } (\delta_{n,p}) \sim \frac{1}{2}\frac{\text{S.E. } (\delta^2_{n,p})}{\sigma} = \frac{\sigma}{\sqrt{2(n-1)W(n,p)}}.$$

If p has been so chosen that $\nu^2_{n,p}$ is suitably small then [see equation (11)] some confidence may be put in the approximate formulas:

$$\text{S.E } (d_{n,p}) = \frac{\sigma}{\sqrt{2(n-1)W(n,p)}} \tag{18}$$

$$\text{P.E. } (d_{n,p}) = \frac{.6745\sigma}{\sqrt{2(n-1)W(n,p)}}. \tag{19}$$

Formula (19) was used in preparing Table IV which gives the approximate probable error to be feared in using $d_{n,p}$ as an estimate of σ. This table should yield interesting information whenever p has been chosen so that $d^2_{n,p}$ is a suitably unbiased estimate of σ^2.

10. Remarks. We have presented a useful technique for estimating variance from higher order differences and have given the precision of our estimate. The method of estimating variance from higher order differences appears to be quite valuable in cases where the type of trend in our observations is unknown. A considerable field of work remains concerning a complete investigation of the distribution and other properties of the statistic $d^2_{n,p}$. In this connection, Baer [12] has already published a study on the stochastic limit of $\frac{n}{n-1}d^2_{n,1}$. It is hoped that others will contribute to the problem of estimating dispersion

TABLE IV

*The Probable Error In Estimating σ From Differences**

n \ p	0	1	2	3	4	5	6	7	8	9	10
1	.4769										
2	.3373	.4769									
3	.2753	.3771	.4769								
4	.2384	.3180	.4054	.4769							
5	.2133	.2796	.3495	.4215	.4769						
6	.1948	.2524	.3104	.3704	.4404	.4769					
7	.1803	.2317	.2817	.3318	.3855	.4390	.4769				
8	.1686	.2154	.2596	.3024	.3477	.3969	.4442	.4769			
9	.1589	.2022	.2420	.2794	.3183	.3604	.4057	.4481	.4769		
10	.1508	.1911	.2274	.2610	.2948	.3311	.3708	.4128	.4513	.4769	
11	.1438	.1816	.2153	.2457	.2758	.3074	.3417	.3794	.4186	.4537	.4769
12	.1376	.1734	.2048	.2328	.2599	.2880	.3180	.3508	.3867	.4234	.4558
13	.1323	.1663	.1958	.2217	.2465	.2717	.2983	.3272	.3587	.3930	.4276
14	.1274	.1599	.1878	.2120	.2350	.2579	.2818	.3073	.3351	.3656	.3984
15	.1231	.1542	.1808	.2035	.2248	.2459	.2677	.2905	.3152	.3423	.3718
16	.1192	.1491	.1744	.1960	.2159	.2355	.2554	.2761	.2983	.3223	.3485
17	.1156	.1445	.1687	.1892	.2080	.2262	.2447	.2637	.2837	.3052	.3286
18	.1124	.1403	.1636	.1831	.2009	.2180	.2352	.2527	.2710	.2905	.3116
19	.1094	.1364	.1589	.1775	.1945	.2106	.2267	.2430	.2599	.2777	.2967
20	.1066	.1328	.1545	.1724	.1886	.2040	.2191	.2343	.2500	.2663	.2837
21	.1040	.1295	.1505	.1677	.1832	.1978	.2121	.2264	.2411	.2562	.2722
22	.1016	.1265	.1468	.1634	.1783	.1922	.2058	.2193	.2331	.2472	.2620
23	.0994	.1236	.1433	.1594	.1738	.1871	.2000	.2129	.2258	.2391	.2529
24	.0973	.1209	.1401	.1557	.1695	.1824	.1948	.2069	.2191	.2316	.2446
25	.0954	.1184	.1371	.1522	.1656	.1779	.1898	.2015	.2131	.2249	.2370
26	.0935	.1160	.1343	.1490	.1619	.1739	.1853	.1964	.2075	.2187	.2301
27	.0918	.1138	.1316	.1459	.1585	.1700	.1810	.1917	.2023	.2130	.2238
28	.0902	.1117	.1291	.1431	.1553	.1664	.1770	.1873	.1975	.2077	.2180
29	.0885	.1097	.1268	.1404	.1522	.1631	.1733	.1832	.1930	.2028	.2126
30	.0871	.1078	.1245	.1378	.1493	.1599	.1698	.1794	.1888	.1981	.2076
31	.0857	.1060	.1224	.1354	.1466	.1569	.1665	.1758	.1848	.1938	.2029
32	.0843	.1043	.1204	.1331	.1441	.1540	.1634	.1724	.1811	.1898	.1985
33	.0831	.1027	.1184	.1309	.1416	.1514	.1605	.1692	.1776	.1860	.1944
34	.0818	.1012	.1166	.1288	.1393	.1488	.1577	.1661	.1744	.1825	.1905
35	.0807	.0999	.1149	.1268	.1371	.1464	.1550	.1632	.1713	.1791	.1869
36	.0795	.0983	.1132	.1249	.1350	.1441	.1525	.1605	.1683	.1759	.1834
37	.0784	.0969	.1116	.1231	.1330	.1418	.1501	.1579	.1655	.1729	.1802
38	.0774	.0956	.1101	.1214	.1311	.1397	.1478	.1555	.1628	.1700	.1771
39	.0764	.0943	.1086	.1197	.1292	.1377	.1456	.1531	.1603	.1673	.1741
40	.0754	.0931	.1072	.1181	.1274	.1358	.1435	.1508	.1578	.1646	.1713

TABLE IV—*Continued*

n \ p	0	1	2	3	4	5	6	7	8	9	10
42	.0736	.0909	.1045	.1151	.1241	.1322	.1396	.1466	.1533	.1597	.1661
44	.0719	.0887	.1020	.1123	.1211	.1288	.1360	.1427	.1491	.1553	.1613
46	.0703	.0868	.0997	.1097	.1182	.1257	.1326	.1391	.1453	.1512	.1570
48	.0689	.0849	.0975	.1073	.1155	.1228	.1295	.1357	.1417	.1474	.1529
50	.0675	.0832	.0955	.1050	.1130	.1201	.1266	.1326	.1383	.1438	.1492
52	.0661	.0815	.0936	.1029	.1107	.1176	.1238	.1297	.1352	.1405	.1457
54	.0649	.0800	.0918	.1009	.1085	.1152	.1213	.1270	.1323	.1375	.1425
56	.0637	.0785	.0901	.0990	.1064	.1129	.1189	.1244	.1296	.1346	.1394
58	.0626	.0771	.0885	.0972	.1045	.1108	.1166	.1220	.1271	.1319	.1366
62	.0606	.0746	.0855	.0939	.1008	.1069	.1125	.1176	.1224	.1270	.1313
66	.0587	.0723	.0828	.0909	.0975	.1034	.1087	.1136	.1182	.1225	.1266
70	.0570	.0702	.0804	.0881	.0946	.1002	.1053	.1100	.1144	.1185	.1224
74	.0554	.0682	.0781	.0856	.0919	.0973	.1022	.1067	.1109	.1149	.1186
78	.0540	.0664	.0760	.0833	.0894	.0947	.0994	.1037	.1077	.1115	.1152
82	.0527	.0648	.0741	.0812	.0871	.0922	.0968	.1009	.1048	.1085	.1120
90	.0503	.0618	.0707	.0774	.0830	.0878	.0921	.0960	.0997	.1031	.1063
98	.0482	.0592	.0677	.0741	.0794	.0840	.0880	.0917	.0952	.0984	.1014
106	.0463	.0569	.0650	.0712	.0762	.0806	.0845	.0880	.0913	.0943	.0972
114	.0447	.0549	.0627	.0686	.0734	.0776	.0813	.0847	.0878	.0907	.0934
122	.0432	.0530	.0606	.0663	.0709	.0749	.0785	.0817	.0847	.0875	.0900
138	.0406	.0498	.0569	.0622	.0666	.0703	.0736	.0766	.0794	.0819	.0843
154	.0384	.0472	.0538	.0589	.0630	.0664	.0695	.0723	.0749	.0773	.0795
170	.0366	.0449	.0512	.0560	.0599	.0632	.0661	.0687	.0711	.0734	.0755
202	.0336	.0412	.0470	.0513	.0548	.0578	.0605	.0629	.0650	.0671	.0689
234	.0312	.0382	.0436	.0476	.0509	.0537	.0561	.0583	.0603	.0621	.0639
266	.0292	.0359	.0409	.0446	.0477	.0503	.0525	.0546	.0565	.0582	.0598
330	.0262	.0322	.0367	.0400	.0428	.0451	.0471	.0489	.0505	.0521	.0535
394	.0240	.0295	.0336	.0366	.0391	.0412	.0430	.0447	.0462	.0475	.0488
522	.0209	.0256	.0292	.0318	.0339	.0357	.0373	.0387	.0400	.0412	.0423
778	.0171	.0210	.0239	.0260	.0278	.0292	.0305	.0317	.0327	.0337	.0346
1290	.0133	.0163	.0185	.0202	.0216	.0227	.0237	.0246	.0254	.0261	.0268
2314	.0099	.0121	.0138	.0151	.0161	.0169	.0177	.0183	.0189	.0195	.0200

* If $d_{n,p}^2$ is a sufficiently unbiased estimate of σ^2, then the approximate probable error to be feared in using $d_{n,p}$ as an estimate of σ may be obtained by multiplying the following tabular entries by σ.

when observed data display trends as it is believed that the method of differences deserves much attention. In particular, it is hoped that someone will have the time and ingenuity to calculate the distribution of the statistic

$$\frac{\delta^2_{n,p}}{\delta^2_{n,p+1}}.^{4}$$

Were this done, an admirable criterion would be at hand for gauging the significance of a change in the estimate of σ^2 as we pass from differences of order p to those of order $p + 1$. Of course, useful information in this connection could be obtained from a knowledge of the distributions of $\delta^2_{n,p}$ and $\delta^2_{n,p+1}$; in fact their variances as herein calculated give us a basis for somewhat reasonable conclusions. An expression for the standard error of the difference between the estimates of σ^2 from two consecutive series of finite differences is given in [13, Chapter VI].

In connection with testing goodness of fit, it would be valuable also to know the distribution of

$$\frac{S^2_{n,p}}{\delta^2_{n,p+1}},$$

where $S^2_{n,p}$ is the estimate of variance derived from the least squares fitting of a polynomial of degree p.

For convenience of reference, we conclude the paper with

11. A concise description of the method and its precision. It frequently happens that successive observations made at regular intervals are subject to the same standard error σ while the means of the populations from which they are drawn display a trend. We give here a method of estimating the variance σ^2 and of determining the precision of our estimate. This method is primarily of value when the trend is unknown; however even when the type of trend is known, its computational simplicity may make the method advantageous.

The method. Arrange the data in a vertical column and then in the usual way form difference columns of order 1, 2, $\cdots$, p. Sum the squares of the pth order differences and divide by the number $(n - p)\binom{2p}{p}$. Our estimate of σ^2 is the number $d^2_{n,p}$, where

$$d^2_{n,p} = \frac{1}{\binom{2p}{p}(n - p)} \sum_{i=p+1}^{n} (\Delta^p x_i)^2.$$

---

[4] Dixon [9] gives moments of the statistic $\dfrac{\sum_{i=1}^{n} (x_i - 2x_{i+1} + x_{i+2})^2}{\sum_{i=1}^{n} (x_i - x_{i+1})^2}$ where $x_{n+1} = x_1$ and $x_{n+2} = x_2$.

The precision. The precision of this estimate may be determined from the following information (which has been derived in the present paper):

$$E(d^2_{n,p}) = \sigma^2 + \nu^2_{n,p} ;$$

$$\nu^2_{n,p} \leq \frac{1}{\binom{2p}{p}} \left(\frac{b-a}{n-p}\right)\left(\frac{b-a}{n-1}\right)^{2p-1} \int_a^b \frac{[f^{(p)}(s)]^2\, ds}{b-a} ;$$

$$\text{Var}\,(d^2_{n,p}) \leq \text{Var}\,(\delta^2_{n,p}) + 4\nu^2_{n,p}\sigma^2 ;$$

$$\text{Var}\,(\delta^2_{n,p}) = \frac{2\sigma^4}{(n-1)W(n,p)},$$

where $W(n, p)$ is given in Table I.

TABLE V

p	σ_x	σ_y	σ_z
1	18.90	184.62	11.22
2	1.21	1.88	10.56
3	.88	1.85	10.30
4	.87	1.84	10.12
5	.86	1.83	10.01

In case $\nu^2_{n,p}$ is sufficiently small (this is determined by the requirements of the given problem), then Table IV may be used directly to determine the approximate probable error in using $d_{n,p}$ as an estimate of σ.

An example. As a practical example of the use of the method of differences when the trend is unknown and of the stability of the statistic $d^2_{n,p}$ with respect to p, we mention a recent problem at Aberdeen Proving Ground which had to do with estimating the accuracy with which certain photographic measurements locate a moving object. Ballistic Cameras were used to determine horizontal x and y, and vertical z coordinates (all in feet) of an airplane traveling about 160 mph at an elevation of about 35,000 feet. An automatic pilot was in use in the airplane as it flew over a three mile course. At one second intervals for a period of 70 seconds two Ballistic Cameras, 5000 feet apart, were used to locate the plane. Since the plane was traveling pretty much in the y direction one would expect: that first differences would yield a standard error in y far in excess of its true one; that second differences would furnish a much better estimate; and that perhaps third differences would yield a still more trustworthy one. No matter what order of difference is used we never expect such an estimate to be too small. In this problem, the standard errors in x, y, z as estimated from differences of certain orders, p, were as given in Table V.

REFERENCES

[1] A. A. Bennett, unpublished report to the Chief of Ordnance, U. S. Army, circa 1918.
[2] R. H. Kent and J. von Neumann, "The estimation of the probable error from successive differences", Ballistic Res. Lab. Report No. 175, Feb., 1940.
[3] John von Neumann, R. H. Kent, H. R. Bellinson, and B. I. Hart, "The mean square successive difference", *Annals of Math. Stat.*, Vol. 12 (1941), pp. 153–162.
[4] J. D. Williams, "Moments of the ratio of the mean square successive difference to the mean square difference in samples from a normal universe", *Annals of Math. Stat.*, Vol. 12 (1941), pp. 239–241.
[5] John von Neumann, "Distribution of the ratio of the mean square successive difference to the variance", *Annals of Math. Stat.*, Vol. 12 (1941), pp. 367–395.
[6] John von Neumann, "A further remark concerning the distribution of the ratio of the mean square successive difference to the variance", *Annals of Math. Stat.*, Vol. 13 (1942), pp. 86–88.
[7] B. I. Hart, "Tabulation of the probabilities for the ratio of the mean square successive difference to the variance", *Annals of Math. Stat.*, Vol. 13 (1942), pp. 207–214.
[8] B. I. Hart, "Significance levels for the ratio of the mean square successive difference to the variance", *Annals of Math. Stat.*, Vol. 13 (1942), pp. 445–446.
[9] Wilfrid J. Dixon, "Further contributions to the problem of serial correlation", *Annals of Math. Stat.*, Vol. 15 (1944), pp. 119–144.
[10] Anthony P. Morse, "The estimation of dispersion from differences", Ballistic Res. Lab. Report No. 557, July, 1945.
[11] R. A. Fisher, *Phil. Trans. A.*, Vol. 222 (1922), p. 316.
[12] Reinhold Baer, "Sampling from a changing population", *Annals of Math. Stat.*, Vol. 16 (1945), pp. 348–361.
[13] G. Tintner, *The variate Difference Method*, (Cowles Commission Monograph No. 5), Principia Press, Bloomington, Indiana, 1940.