CHAPTER VII

Criteria for the Stability of the Variances of the Series of Finite Differences

We have shown in the last chapter how it is possible to find the standard error of the difference between the variances of two consecutive series of finite differences. From the point of view of probability, it appears that these variances become stable in long series if consecutive variances do not differ from each other by more than about three times the standard error of the difference (see above, pages 33 f.). Criteria have been given for this by Professor O. Anderson[1] of Sofia, Bulgaria, and they have been recently improved by Dr. R. Zaycoff[2] of Sofia. We are going to use in this connection Dr. Zaycoff's first criterion, which he himself considers the better one.

The comparison is made in the following way. We form the criterion (standard-error ratio):

$$R_k = \frac{(V_k - V_{k+1})}{V_k} Q_k .$$

This is the difference between two consecutive variances divided by its standard error. The value of the difference between two consecutive variances of the series of finite differences (for instance the difference between the variances of the series of first finite differences and the series of second finite differences, $V_1 - V_2$) is divided by the variance of the difference of lower order (in our example by the variance of the series of the first differences, V_1). The result is multiplied by the quantity Q_k with the index which corresponds to the lower of the series of finite differences (in our case by Q_1). We shall use the quantities Q_k rather than the approximations H_{kN} if we wish to be quite accurate and if our distributions show considerable kurtosis. We do this in our case for all the variances of the series of finite differences of wheat-flour prices up to the tenth. The last criterion R_9, for instance, is the value of the difference between the variances of the ninth and tenth finite differences, $V_9 - V_{10}$, divided by the vari-

[1] O. Anderson, *Die Korrelationsrechnung in der Konjunkturforschung*, Bonn, 1929, pp. 56 ff., 112 ff. See also below, Appendix II, pp. 139 ff., and Appendix I, Sections C and D, for a summary of computations.

[2] R. Zaycoff, "Ueber die Ausschaltung der Zufälligen Komponente nach der 'Variate-Difference' Methode," *Publications of the Statistical Institute for Economic Research*, State University of Sofia, 1937, No. 1, pp. 78 ff.

ance of the ninth finite difference, V_9, and multiplied by the quantity Q_9 as described in the previous chapter (Table 22).

We have again in our example of annual wheat-flour prices, from Table 17: $V_9 = 0.3334$ and $V_{10} = 0.3243$. We get from Table 22: $Q_9 = 21.64$. Hence the criterion R_9 for the difference between the variances of the ninth and tenth differences is: $R_9 = (V_9 - V_{10}) Q_9 / V_9 = (0.3334 - 0.3243)(21.64)/0.3334 = 0.5906$. This value and all other values for the exact criterion R_k for the annual wheat-flour prices are tabulated in Table 22.

TABLE 22

DIFFERENCE ANALYSIS
ANNUAL AMERICAN WHEAT-FLOUR PRICES, 1890-1937

Order of Difference	Variance	Kurtosis		Standard Error Ratio	Approximate Standard Error Ratio
k	V_k	D_k	Q_k	R_k	R^0_k
0	4.7969	33.4473	6.73	5.7459	5.7878
1	0.7020	4.4384	12.30	4.5863	5.0514
2	0.4402	1.1000	16.34	1.7488	1.9368
3	0.3931	1.1920	18.30	0.7637	0.8952
4	0.3767	1.1974	19.70	0.5492	0.6668
5	0.3662	1.2609	20.49	0.6376	0.8012
6	0.3548	1.3041	20.87	0.2764	0.3582
7	0.3501	1.2733	21.37	0.4579	0.5987
8	0.3426	1.2460	21.60	0.5800	0.7668
9	0.3334	1.2141	21.64	0.5906	0.7885
10	0.3243	....	....	....	....

Those criteria R_0, R_1, R_2, etc., are arranged in a series. We are going to consider as reasonably stable the variance beginning from which the R_k becomes numerically smaller than 3 and stays more or less so. The R's for our annual wheat-flour prices are given in Table 22, where also some of the calculations are exhibited. Whereas R_0 and R_1 are greater than 4, R_2 and all the following parameters are certainly smaller than 2 and become very small indeed for higher differences. Hence we can conclude that we have eliminated the mathematical expectation or the nonrandom element of our wheat-flour prices with reasonable accuracy in the second or third difference in the particular case under consideration. We can say that in all probability beginning with the second or third difference there are left only some remainders of the nonrandom element which we can neglect for our purposes (see above, pages 33 f.). An alternative procedure is given in Chapter VIII.

We shall now try to find the best approximation to the true random variance of the original series. In order to do this we take the variance of the series of finite differences beginning from which we can be reasonably sure that we have eliminated the nonrandom element to the desired degree. In our case, for example, we should take the variance of the second finite difference, which is equal to 0.4402 (Table 22). We could say that this probably represents a good approximation to the true variance of the pure random element in our time series, as defined above (Chapter I, Section C).

If we desire to avoid a great deal of calculation and if we have reason to think that the kurtosis of our series is not very great, we may use the approximations H_{kN} instead of the better estimates Q_k. We shall then form a criterion (approximate standard-error ratio):

$$R_k^0 = \frac{(V_k - V_{k+1})}{V_k} H_{kN} .$$

The procedure is also shown in Table 22. It is the same as the former, except that we use the approximate instead of the more accurate estimates of the standard errors.

In order to show the calculations, let us use, for instance, our previous example and calculate the approximate criterion, R_4^0, for the difference between the variances of the fourth and fifth difference series of annual wheat-flour prices. We have from Table 22: $V_4 =$ 0.3767 and $V_5 = 0.3662$ as estimates for the two variances. Table 20 gives for $k = 4$ and $N = 48$, by interpolation, $H_{4N} = 23.928$. Hence the approximate criterion $R_4^0 = (0.3767 - 0.3662)(23.928)/0.3767 =$ 0.6668. The values of this approximate criterion R_k^0 are also given in Table 22. They yield the same result as before. It appears that we probably eliminate our mathematical expectation or nonrandom element in the second or third difference, since R_2^0 and the following values are all smaller than 2.

We have accomplished our goal, for many purposes, if we reach this stage of our analysis. If we want, for instance, to make a comparison of our annual wheat-flour prices with some other statistical data, we must know the limits of errors and inaccuracies involved in this statistical comparison. We have seen that it is very improbable that we shall get a deviation from the true value that is greater than three times the standard deviation of the random element, which is approximately 0.66. We are going to be wrong in only about three cases out of a thousand if we indicate the probable limits of a deviation due to the random variation from the mathematical expectation by $\pm 3\sqrt{V_2} = \pm 1.99$. (See above, pages 33 f.)

This is extremely important if we wish to compare two statistical series and if we desire to know whether differences can be explained by the random variation. The same can easily be done for certain derived statistical series which we calculate from our original series. We reproduce, in Appendix IV, the standard errors of some statistical series, some of which have been given previously by the author.[3] They should facilitate the comparison of certain parameters calculated from the original series by giving their standard errors. It is, for instance, possible to indicate the standard error resulting from the random variation that is retained after calculating a moving average, a seasonal index, a trend, etc. All those statistical estimates are most important because they give the desired statistical measurements for definite economic purposes. They should always be treated from the point of view of probability. The variate difference method can give us a reasonably good estimate of the random variation in those parameters which is due to the erratic element as defined above. It enables us also to make statistical tests of hypotheses.

The same type of analysis is shown in Tables 23 to 25 for the other prices. We show the difference analysis for annual wool prices in Table 23. Both the accurate criterion R_k and the approximate criterion R_k^0 indicate that we have already eliminated the nonrandom ele-

TABLE 23

DIFFERENCE ANALYSIS
ANNUAL WOOL PRICES, 1890-1937

Order of Difference	Variance	Kurtosis		Standard Error Ratio	Approximate Standard Error Ratio
k	V_k	D_k	Q_k	R_k	R_k^0
0	0.1069	0.020 93	6.72	4.9775	5.0234
1	0.02769	0.009 810	11.87	0.7671	0.8758
2	0.02590	0.006 434	15.39	–0.2020	–0.2377
3	0.02624	0.007 861	17.22	–0.0918	–0.1145
4	0.02638	0.008 374	18.47	0.0840	0.1088
5	0.02626	0.008 802	19.26	0.1981	0.2647
6	0.02600	0.008 990	19.76	0.1673	0.2289
7	0.02577	0.008 934	20.17	0.1879	0.2603
8	0.02553	0.008 739	20.48	0.2727	0.3802
9	0.02519	0.008 493	20.64	0.2622	0.3671
10	0.02487	....	....	....	....

[3] G. Tintner, *Prices in the Trade Cycle*, Vienna, 1935, pp. 81 ff.

ment in the first or second finite differences, since both become, and stay, smaller than 1 in absolute value, i.e., without regard to sign, for $k = 2, 3$, etc.

TABLE 24

DIFFERENCE ANALYSIS
MONTHLY WOOL PRICES, 1890-1937

Order of Difference	Variance	Kurtosis		Standard Error Ratio	Approximate Standard Error Ratio
k	V_k	D_k	Q_k	R_k	R^0_k
0	0.112 1	0.047 6626	23.87	23.4370	23.5141
1	0.002 054	0.000 9100	41.75	16.6066	20.1061
2	0.001 237	0.000 1925	58.56	6.6750	8.1811
3	0.001 096	0.000 1904	66.93	4.0912	5.5411
4	0.001 029	0.000 1559	76.57	3.7434	5.2731
5	0.000 9787	0.000 1444	83.41	2.8804	4.2771
6	0.000 9449	0.000 1341	89.69	2.1642	3.3485
7	0.000 9221	0.000 1269	95.30	1.7775	2.8489
8	0.000 9049	0.000 1213	100.38	1.3979	2.3108
9	0.000 8923	0.000 1170	105.01	0.9650	1.6396
10	0.000 8841	....	....	....	....

Table 24 gives the same type of analysis for monthly wool prices. The accurate criterion would indicate that the mathematical expectation has been eliminated in the fifth difference, whereas the series

TABLE 25

DIFFERENCE ANALYSIS
ANNUAL RAW-SILK PRICES, 1890-1937

Order of Difference	Variance	Kurtosis		Standard Error Ratio	Approximate Standard Error Ratio
k	V_k	D_k	Q_k	R_k	R^0_k
0	2.8914	5.020 9	6.76	5.8598	5.8775
1	0.3849	0.429 2	13.10	3.8639	3.9943
2	0.2714	0.163 4	17.35	2.5383	2.6489
3	0.2317	0.086 24	20.67	2.5330	2.6297
4	0.2033	0.047 31	23.19	2.3842	2.4599
5	0.1824	0.012 60	25.45	2.2599	2.2862
6	0.1662	–0.004 80	27.20	2.1113	2.0993
7	0.1533	–0.016 70	28.66	2.0750	2.0241
8	0.1422	–0.021 80	29.70	2.0464	1.9677
9	0.1324	–0.023 45	30.39	1.7213	1.6367
10	0.1249	....	....	....	....

of the R_k^0 becomes numerically smaller than 3 only for $k = 7$ and higher, indicating that we succeed only in the seventh difference in getting rid of the nonrandom element.

Table 25 gives the difference analysis for the annual raw-silk prices. Both the exact criterion R_k and the approximation R_k^0 point to the second difference as the difference in which only traces of the nonrandom or smooth element remain.

Table 24 gives as an estimate of the true random variance of the monthly wool prices: $V_6 = 0.000\,9449$. From Table 23 for the yearly wool prices: $V_1 = 0.02769$. The random variance of the annual prices is here almost 30 times as great as that of the monthly prices.

The explanation of this may be attempted in the following way: some economic causes, which may appear permanent from the point of view of the shorter run, i.e., in the monthly series, become nonpermanent and hence part of the random element from the point of view of a longer run, i.e., in the annual series. (See also above, Chapter I, Section C.)