Centre for Statistical and Survey Methodology

The University of Wollongong

Working Paper

11-10

The Probability of an Out of Control Signal from Nelson's Supplementary Zig-Zag Test

David Griffiths, Martin Blunder, Chandra Gulati, Takeo Onizawa

Centre for Statistical & Survey Methodology, University of Wollongong, Wollongong NSW 2522. Phone +61 2 4221 5435, Fax +61 2 4221 4845. Email: anica@uow.edu.au

The Probability of an Out of Control Signal from Nelson's Supplementary Zig-Zag Test

David Griffiths, *Centre for Statistical and Survey Methodology, School of Mathematics and Applied Statistics, University of Wollongong, Australia. Email: david_griffiths@uow.edu.au*

Martin Bunder, *School of Mathematics and Applied Statistics, University of Wollongong, Australia. Email: martin_bunder@uow.edu.au*

Chandra Gulati, *Centre for Statistical and Survey Methodology, School of Mathematics and Applied Statistics, University of Wollongong, Australia. Email: cmg@uow.edu.au*

Takeo Onizawa, *Ex-student, School of Mathematics and Applied Statistics, University of Wollongong, Australia.*

Abstract

Nelson's 'supplementary runs' tests are widely used to augment the standard 'out of control' test for an $\bar{x}$ control chart, or a chart with individual values, to determine if any special causes exist. The fourth of Nelson's tests gives an out-of-control signal when fourteen points in a row follow a zig-zag pattern (alternating up and down); it is thus a signal that the process has negative autocorrelation. Using a recursive formula, the exact probability of a zig-zag sequence of length 14 is calculated for an in control process. This value does not appear in the SQC literature, but can be simply determined from results of André (1879, 1881, 1883), rediscovered by Entringer, (1966), which long precede the development of SQC. Two curious properties, relating the probabilities of zig-zag sequences of successive lengths, are also demonstrated.
AMS Subject Classification: 60-01 and 62P30
Keywords: $\bar{x}$ chart, Supplementary runs tests, Nelson's 'zig-zag' test, alternating permutations.

1 Introduction

Nelson (1984) suggested a set of supplementary tests for detecting out of control conditions in control charts. All of these tests are directed at finding unusual runs of points which signal that the process is not behaving as expected if in control. Nelson's eight tests are briefly described below. In all but two of the tests, the supplementary rule is based on all or some in a sequence of points being in a particular zone (or union of zones) in the control chart. These zones, and their associated probabilities, are shown in Table 1.

Table 1. . Control chart zones and associated probabilities

	Zone D_1	$P\{D_1\} = 0.00135$	
UCL			$\mu + 3\sigma$
	Zone A_1	$P\{A_1\} = 0.02140$	
			$\mu + 2\sigma$
	Zone B_1	$P\{B_1\} = 0.13591$	
			$\mu + \sigma$
	Zone C_1	$P\{C_1\} = 0.34134$	
CL			μ
	Zone C_2	$P\{C_2\} = 0.34134$	
			$\mu - \sigma$
	Zone B_2	$P\{B_2\} = 0.13591$	
			$\mu - 2\sigma$
	Zone A_2	$P\{A_2\} = 0.02140$	
LCL			$\mu - 3\sigma$
	Zone D_2	$P\{D_2\} = 0.00135$	

This table applies to single variable charts, in which the quality characteristic being plotted is assumed to follow a normal distribution with mean μ and standard deviation σ , provided the process is in control. With simple modification, the table can be applied to an $\bar{x}$ chart with mean μ and standard deviation. $\sigma/\sqrt{n}$. In such a control chart, the location and trend of observed values of the sample statistic is used to assess whether the process is in control.

2 Nelson's Eight Tests

Nelson described eight tests, Test 1 being the standard test of a single point beyond the control limits. The other seven tests supplement the standard test. His eight tests are defined by the following signals:
Test 1. A single point beyond zones A (that is, in zone D_1 or D_2);
Test 2. Nine points in a row on one side of centre line;
Test 3. Six points in a row, all increasing, or all decreasing;
Test 4. Fourteen points in a row, alternating up and down;
Test 5. (At least) two out of three points in a row both in zone A_1 or beyond, or both in A_2, or beyond;
Test 6. (At least) four out of five points in a row, on the same side of the centre line and also either in zone B_1 or beyond, or in B_2 or beyond;

Test 7. Fifteen points in a row, each in zone C_1 or in zone C_2;
Test 8. Eight points in a row, all beyond zones C, ie each in B_1 or beyond or B_2 or beyond.

Unlike the other six tests, Test 3 and 4 do not relate to the zones described in Table 1.

Nelson's tests have gained widespread acceptance, and they have been incorporated in some commonly used statistical packages, including SAS, MINITAB and STATISTICA. Such tests have evolved over time. For example, the Western Electric Handbook (1956) suggests similar tests. They are now commonly incorporated in text books, e.g. Montgomery (1996), and in the quality manuals and practices of large manufacturing organisations, such as Ford.

When the process is in control, the probability that a sequence of the appropriate length will signal an out-of-control process has been previously determined for some of the tests by Walker, Philpot and Clement (1991). Champ and Woodall (1987) also report these or related probabilities for Tests 1, 5 and 6. The probabilities are given below. That for Test 4 has not previously appeared in the SQC literature. With the exception of Test 4, the brief detail of the calculation of each probability also explains its derivation.

$$
\begin{aligned}
P\{T_1\} &= P(\text{a point falls beyond zone } A_1 \text{ or beyond zone } A_2), \\
&= 2 \times P(D_1) = 0.00270. \\
P\{T_2\} &= 2P(C_1 \text{ or } B_1 \text{ or } A_1 \text{ or} D_1)^9 = 2 \times .5^9 = 0.00391. \\
P\{T_3\} &= (1+1)/6! = 0.00278,
\end{aligned}
$$

since there are 6!=720 possible orderings of six points, in only one of which there is a pattern of successive increases and in only one of which there is a pattern of successive decreases.

$$
\begin{aligned}
P\{T_4\} &= 398721962/14! = 0.00457, \text{ as derived later in this paper.} \\
P\{T_5\} &= 2 \times \begin{pmatrix} 3 \\ 2 \end{pmatrix} \times [\ P(A_1 \text{ or } D_1)]^2 \times [1 - P(A_1 \text{ or } D_1)] + 2[P(A_1 \text{ or } D_1)]^3 \\
&= 2 \times 3 \times 0.02275^2 \times 0.97725 + 2 \times 0.02275^3 = 0.00306. \\
P\{T_6\} &= 2 \times \begin{pmatrix} 5 \\ 4 \end{pmatrix} \times [\ P(D_1 \text{ or } A_1 \text{ or } B_1\ )]^4 \times [1 - P(D_1 \text{ or } A_1 \text{ or } B_1)] \\
&\quad +2[P(D_1 \text{ or } A_1 \text{ or } B_1)]^5 \\
&= 2 \times 5 \times 0.15865^4 \times 0.84135 + 2 \times 0.0.15865^5 = 0.00553. \\
P\{T_7\} &= [P(C_1 \text{ or } C_2)]^{15} = 0.68268^{15} = 0.00326. \\
P\{T_8\} &= [P(D_1 \text{ or } A_1 \text{ or } B_1 \text{ or } B_2 \text{ or } A_2 \text{ or} D_2)]^8 = 0.31732^8 = 0.00010
\end{aligned}
$$

3 Nelson's Test 4: Zig-Zag Runs

Although not spelled out by Nelson, the motivation of Test 4 was presumably not to provide another test for a shift in the process mean (as in Tests 2, 5 and 6) nor in the process spread (as in Tests 7 and 8), nor a trend upwards or downwards in the mean (as in Test 3). Rather, Test 4 appears to be targeted at detecting a process with negative autocorrelation. The purpose of this note is not to explore the properties of this test in meeting that purpose, but simply to calculate the probability of an out-of-control signal for a set of 14 points from an in-control process of independent observations.

4 The probability of an out-of-control signal for Nelson's Test 4

The number of ways that n points in a row will alternate up and down was denoted by $2A_n$ in André (1879, 1881, 1883), who noted that these numbers are generated by the series for $\tan(x+y)$ as follows:

$tan(\pi/4 + x/2) = \sum_{n=0}^{\infty} A_n x^n/(n!)$
This may alternatively be presented with the terms for n odd and n even presented as separate series:

$secx = \sum_{n=0}^{\infty} A_{2n} x^{2n}/((2n)!)$
$tanx = \sum_{n=0}^{\infty} A_{2n+1} x^{2n+1}/((2n+1)!))$

The sequence A_n is also related to the Euler numbers, E_n, and Bernoulli numbers B_n as follows:

$A_{2n} = |E_{2n}|,$
$A_{2n-1} = (-1)^{n-1} B_{2n} 2^{2n}(2^n - 1)/2^n$

Of course, both E_n and B_n can also be represented in series form. See, for example Abramowitz and Stegun (1970). Much of André's work was rediscovered by Entringer (1966). Here we use Entringer's simple algorithm to demonstrate the evaluation of A_n for small n, including the important value n = 14.
The method proceeds as follows. Let $T_n = 2A_n$ be the number of sequences with the zig-zag property among the n! sequences of length n involving the integers 1, 2, ..., n. Among these T_n sequences, let U_n be the number that start with an up

step and let D_n be the number that start with a down step. Then, from symmetry considerations, $D_n = U_n = A_n$, and also $T_n = D_n + U_n = 2A_n$. The required probability is $p_n = T_n/n!$.

The key to a simple approach to finding p_n is to further subdivide the T_n sequences according to the number with which they start. Thus, let $T(n, k)$ be the number of sequences of length n which have the zig-zag property and which start with the integer k. Similarly, let $U(n, k)$ and $D(n, k)$ be the corresponding number of such sequences beginning respectively with an up-step or down-step. Then, $T(n, k) = U(n, k) + D(n, k)$.

Note that $T_n = \sum_{k=0}^{n} T(n, k)$, $U_n = \sum_{k=0}^{n} U(n, k)$, and $D_n = \sum_{k=0}^{n} D(n, k)$. Symmetry conditions require that $U(n, k) = D(n, n + 1 - k)$.

Consideration of particular values of $U(n, k)$ and $D(n.k)$ is useful at this stage. Note firstly that zig-zag sequences only exist for $n \geq 3$. Secondly, $D(n, 1) = U(n, n) = 0$, since it is not possible to step down (up) from the lowest (highest) number. Enumeration of $D(n, k)$ and $U(n, k)$ for all k and $n \leq 6$ is given in Table 2, as can be established for each n by an examination of all $n!$ sequences. For example, $U(4, 2) = 2$, since the only two zig-zag sequences starting with a step up from 2 are (2, 3, 1, 4) and (2, 4, 1, 3), and $D(5, 3) = 4$, the relevant sequences being (3, 1, 5, 2, 4), (3, 1, 4, 2, 5), (3, 2, 5, 1, 4) and (3, 2, 4, 1, 5).

Table 2. The numbers, $D(n, k)$ and $U(n, k)$ of zig zag sequences of different types

	$D(n.k)$							$U(n.k)$						
k	1	2	3	4	5	6	7	1	2	3	4	5	6	7
n														
3	0	1	1					1	1	0				
4	0	1	2	2				2	2	1	0			
5	0	2	4	5	5			5	5	4	2	0		
6	0	5	10	14	16	16		16	16	14	10	5	0	
7	0	16	32	46	56	61	61	61	61	56	46	32	16	0

An analytic approach is useful to generalise the patterns which appear in Table 2. Now, every zigzag sequence of length n starting with a down step can be generated by adding a down step to the front of a sequence of length $n - 1$ starting with an up step. Consideration of all possible sequences beginning with any j and of length $n - 1$ leads to the following result for $k = 2, 3, \ldots, n$

$D(n, k) = \sum_{j=n-k+1}^{n-1} U(n - 1, j - 1)$.

Recall that for k=1, $D(n, 1) = 0$. From this relationship between the D's and the

U's, it follows that for $k \geq 2$.

$D(n,k) = \sum_{j=n-k+1}^{n-1} D(n-1,j)$.

This formula provides a means of writing down the values in a row of a table of values of $D(n,k)$ as partial sums of elements in the row above, and hence generating the values of T_3, T_4,.... In row n of Table 2, the sum of the elements ($D's$ and $U's$) is T_n, which can also be found directly from the table, since $2D(n+1,n) = 2D_n = 2A_n = T_n$. The values of T_n, along with the probabilities, p_n, of a zig-zag sequence of length n, for $n = 3(1)14$ are given in Table 3. The required result is $p_{14} = T_{14}/14! = 0.0046$.

Table 3. Probabilities of zig-zag sequences of different lengths, and associated quantities

n	T_n	p_n	p_{n-1}/p_n	$r_n = p_{n-1}/p_n - \pi/2$	$-r_{n-1}/r_n$
3	4	0.666666667	1.5		
4	10	0.416666667	1.6	0.0292	
5	32	0.266666667	1.5625	-0.0083	3.520
6	122	0.169444444	1.57377	0.00297	2.789
7	544	0.107936508	1.56985	-0.00094	3.153
8	2770	0.068700397	1.57112	0.000323	2.922
9	15872	0.043738977	1.57069	-0.000106	3.051
10	101042	0.027844466	1.57083	0.0000356	2.972
11	707584	0.017726471	1.57078	-0.0000118	3.018
12	5405530	0.011284994	1.5708003	0.00000395	2.990
13	44736512	0.007184256	1.5707950	-0.00000131	3.006
14	398721962	0.004573638	1.5707968	0.000000438	2.996

5 Convergence of ratios

As an intriguing aside, an interesting limiting relationship among the p_n's was rediscovered while investigating this matter. Also shown in Table 3 are the ratios p_n/p_{n-1} ; these values demonstrate that $\lim_{n\to\infty} \frac{p_n}{p_{n-1}} = \frac{\pi}{2}$, and that convergence is quite rapid. A formal proof of this is given by André (1883).

The differences between the ratios p_n/p_{n-1} and the limiting value of $\pi/2$ are also presented in Table 3; a further curiosity emerges, namely that the sequence of ratios $-r_{n-1}/r_n$ appears to converge to 3. Proof of this is not readily obtained using the methods and results of André (1879, 1881, 1883) or Entringer (1966), nor, seemingly, by other means.

6 Discussion

A common modification to some of Nelson's tests has been to vary the length of a sequence that defines a signal. For example, in the SAS procedure PROC SHEWHART, Test 2 is modified by the use of eight points instead of nine. The choice of the number of points in a sequence is arbitrary; the motivation for particular choices for the different tests might be to make the signal probabilities similar for in control processes. It is a simple matter to alter three of the tests to give a set of probabilities all as close as possible to 0.003. The tests are changed as follows. In Test 4, the length of the zig-zag sequence is increased to 15, and the resulting probability is 0.00291. If the requirement for Test 6 is changed from 4 out of 5 successive points to 5 out of 7 points, then the corresponding probability is 0.00318. Finally, for a pattern involving not 8, but 5 points in a row, the probability for Test 8 becomes 0.00322.

In practice, a subset of Nelson's seven supplementary tests (or various modifications of them) is often used in process monitoring. The tests are not independent. When considered as part of a longer sequence, a point may simultaneously signal an out-of-control process according to two or more tests. The simultaneous application of these tests has been addressed by Champ and Woodall (1987) and Walker et al. (1991). Both papers explore the effect of a shift in the process mean on some of the probabilities of an out-of control signal from Nelson's tests. Champ and Woodall (1987) have also investigated the overall probability of an out-of-control signal for some limited combinations of tests.

As with the other tests proposed by Nelson, there would be some interest in finding such quantities as average run length and false alarm probabilities for sequences of various lengths. These are not strongly pertinent to this paper and are not pursued here.

Acknowledgements

We thank Sat Gupta for giving us an opportunity to contribute to this volume in honor of Professor H.C. Gupta. Chandra Gulati was privileged to have taken two courses, in Probability and Stochastic Processes, given by H.C. Gupta during his study for a Masters Degree in Statistics at the University of Delhi. David Griffiths and H.C. Gupta shared the same distinguished doctoral advisor, Professor M.S. Bartlett. Discussion with Professor Alf van der Poorten led us to the previous derivations of the key result in this paper. Takeo Onizawa completed an Honours Master degree in Total Quality Management, and he is currently living in Japan.

References:

Abramowitz, M. and Stegun, I. 1970. *Handbook of Mathematical Functions*, Dover, New York

André, D. 1879. Développements de séc et de tang x, *Comptes Rendus* 88, 965-967.

André, D. 1881. Sur les permuataions alternes, *Journal Mathematique* 7, 167-184.

André, D. 1883. Probabilité pour qu'une permutation donnée de n letters soit une permutation alternée, *Comptes Rendus* 97, 983-984.

Champ, C.W. and Woodall, W. 1987. Exact results for Shewhart control charts with supplementary runs rules, *Technometrics* 29, 393-399.

Comtet, L. 1974. *Advanced Combinatorics*, D. Reidel, Dordrecht.

Entringer, R.C. 1966. A combinatorial interpretation of the Euler and Bernoulli numbers, *Nieuw Archief voor Wiskunde* 14, 241-246.

Montgomery, D.C. 1996. *Introduction to Statistical Quality Control.* 3rd edition, Wiley, New York.

Nelson, L.S. 1984. The Shewhart control chart-tests for special causes, *Journal of Quality Technology* 16, 237-239.

Walker, E., Philpot, J.W. and Clement, J. 1991. False signal rates for the Shewhart control chart with supplementary runs tests, *Journal of Quality Technology* 23, 247-252.

Western Electric 1956. *Statistical Quality Control Handbook*, Western Electric Corporation, Indianapolis, Ind.