NOTES ON WHITTAKER-HENDERSON FORMULA A

NELS M. VALERIUS

NOTE 1

When Mr. Robert Henderson, 1873-1942, the distinguished life actuary and a Fellow of our Society, put forward "A New Method of Graduation" in 1924, in *Transactions* XXV of the Actuarial Society of America, the problem of starting values was resolved by calculating upward an auxiliary series which he labelled the u_x''' column, alongside the u_x'' column to be graduated. The u_x''' column terminated, in an extension of its upper end, in terms deemed acceptable for starting values. These were then copied to the same positions of the Henderson intermediate u_x' column to start that column off downward.

The formula for calculating the successive terms of the u_x''' column was the same basic formula that was thereupon used to work out the graduation itself, beginning from the starting values thus established; Henderson's intermediate u_x' column being next filled in downward by means of it, then the final u_x column upward.*

These starting values were not highly accurate, and the resulting $\{u_x\}$ was therefore not accurate at the upper end of the column. Mr. Henderson was not greatly concerned, witness his pun regarding "such unprofitable ends as the ends of a graduated mortality table" in his discussion, *T.A.S.A.* XXXIX, page 50, of C. A. Spoerl's comprehensive paper, "Whittaker-Henderson Formula A," *T.A.S.A.* XXXVIII, pp. 403-462.

In casualty and property actuarial work, unlike mortality tables, the ends and extensions of the ends of, for instance, graduated time series may be the more crucial parts. The potential of accuracy at the ends under Formula A, when it is used in this work, is a most attractive feature, not to overlook the preservation of moments inherent in an accurate A graduation.

The determination of accurate starting values has remained troublesome. Thus, the textbook for life actuarial students, sponsored by the Society of Actuaries, the monograph, *Elements of Graduation*, says (of the fourth-order difference equation graduation), "At the zero end of the series, the two needed values of u_x' cannot be determined accurately at the outset except by involved methods. If such methods are not to be resorted to, the

* A version of the formula is given in the Appendix herewith. Algebraic symbols used in the Notes may be identified there also.

first two graduated values must be estimated from the general run of the ungraduated values at the zero end. The graduation is completed using these estimated values and then corrected, if necessary."

Unexpectedly, a return to Henderson's method yields a solution. Henderson's up and down approach to the final column suggests that more of the same might provide any desired accuracy and this proves to be right. Actually, iteration of Henderson's u_x' series in alternately opposite directions will provide as perfected a Henderson's intermediate series $\{u_x'\}$ as one wishes, enabling a graduation with accuracy to as many places as one may have elected of any ungraduated series, short or long, provided it falls within the province of Formula A, that is, it has equal intervals, e.g., data of successive years, and graduation without formal recognition of weights is acceptable. Many series are of this kind.

No additional techniques are involved, obviously, to those used to work out the graduation in the usual way, from starting terms however established.

The graduator runs an auxiliary column like u_x''' upwards or downwards, starting opposite what segment of z consecutive terms of the u_x'' column he deems expedient; then at the extended end of such auxiliary column, he turns about and produces in reverse direction a tentative u_x' column, then another in reverse direction to it, and so on, until the required accuracy results. The process amounts to successive approximations to the final $\{u_x'\}$.

A point is reached where z successive terms in a column duplicate the corresponding terms in the last previous column of like direction and no more improvement is possible. The column between is an accurate u_x' column, as would also be the last partial column if filled out by copying from the previous column of like direction, so that accurate u_x' columns of both directions are available. The final u_x column must, of course, be worked out in the direction opposite to that of the u_x' column selected.

Having appropriated from Henderson's up u_x''' and down u_x' operations, the up and down (or down and up) sequence, his u_x''' column may, if one wishes, be replaced by any other plan for starting terms, including guesses. With approximate starting terms, one goes to work at once on a first tentative u_x' column, following it with others at will. Depending on the accuracy desired, it is not necessary to follow through to the point of maximum improvement. Just to adopt the bottom end of the first tentative u_x' column as the chosen start may be acceptable.

The proposal of this Note will be made clearer by reference to Exhibit I, where the Specimen Graduation by Formula A on page 39 of *Elements of Graduation* is reworked under the method of this Note.

In Exhibit I, to start the auxiliary column u_x''' off upward, z terms, two in this case, u_{16}''' and u_{17}''', were devised by a preliminary smoothing of the corresponding segment of $\{u_x''\}$. The work was carried forward with two places of decimals and so displayed; as one or two extra places beyond those retained are advisable, however (even more if a is quite large), the result is not accurate in the last place.

This judgment may be confirmed by reference to the Society of Actuaries' Study Notes for Part V, wherein the graduation by the "second Henderson method" is given to five places. That result, rounded back to two places, differs, but by no more than .01 or .02, as to most of the terms.

Pursuing the proposal's potential for accuracy, the writer has reproduced the Study Notes' five-place result by the method of Exhibit I, retaining six places of decimals after the initial $u_{16}''' = 105$ and $u_{17}''' = 117$; u_{15}''' becoming 94.166667 in place of 94.17, and so on. It was necessary to go to $\{_3u_x'\}$, for a satisfactory $\{u_x'\}$. The work is not reproduced here, being mentioned only as an example of extreme accuracy achieved by this method.

Exhibit II is included for comparison of the usual method (unless use of matrices and computers should today be called the usual methods). It is the monograph's numerical example, pages 38-39, with slight improvements in the arithmetic and with Spoerl's corrections carried out, for comparison to Exhibit I. Faulty arithmetic on page 39 marred Spoerl's corrections, in the writer's opinion, and the recalculation is provided for an equitable comparison of results. The corrected u_0 and u_1 are given on page 38 of the monograph as 27.34 and 29.76, respectively, instead of the values on Exhibit II of 27.39 and 29.80, which accord with the Study Notes' five-place results of 27.39625 and 29.80043.

In the monograph, and in Exhibit II, an auxiliary column u_x^* is read from a graph drawn among the upper end values of $\{u_x''\}$ and extended upward. The extension is a straight line since $z = 2$, rather than an arc as it would be were $z = 3$ (see Table on page 5). Note that the graph approximates the true u's extended, and, using the fact that true $u_x =$ true u_x' in the outer extension where $x < -a$, thereby approximates the true starting u''s.

It might not be amiss to remind that a large a emphasizes smoothness, a small a, fidelity or closeness of fit. The value of z (1, 2, or 3) fixes the relative simplicity or complexity of the specific Formula A: $z = 1$ being so simple as hardly to be useful at all, "the trivial case," as Spoerl calls it, page 413; $z = 2$ being widely useful; and $z = 3$ being rather elegant. One, two, or three starting terms are required, respectively, and the graduated results tend to be a succession of segments whose first, second, or third differences are zero, respectively. Their respective extensions are exactly such segments, save the inner *a* terms of the extensions at the starting end.

TABLE

z	$2z$	*Starting Terms*	*Extensions and Tendency*
1	2	1	$\Delta u_x = 0$
2	4	2	$\Delta^2 u_x = 0$
3	6	3	$\Delta^3 u_x = 0$

In this Note, the intermedite variable is u'_{x-a}, like Henderson's u'_{x-n}. In the monograph, the variable is termed u_x', shedding what Mr. Kingsland Camp, F.S.A., has called the "displacement" in his manual. *The Whittaker-Henderson Graduation Processes.* The newer usage is neater and advantageous in several respects, particularly in accommodating fractional *a*'s. Nevertheless, with the alternating sequences proposed, the symmetry of the chosen format seems preferable. It also adheres to Mr. Spoerl's cited paper, an advantage still, as this remains the prime reference for Formula A. Iterating $\{u_x'\}$ is, of course, equally valid either way. Exhibit IIA is included to illustrate the work with u_x' as the intermediate variable.

NOTE 2

Graduations by Formula A share the property that least squares fittings of lines and parabolas exhibit (these may be looked upon as the cases of ultimate smoothness of fourth-order and sixth-order Whittaker-Henderson applications), viz., additive elements of a series, such as an underwriting history of annual losses, expenses, and underwriting gain or loss, and their sum or premium, may be separately graduated with the same choice of constants, with results that are still additive, so that the graduated annual elements add up to the graduated annual sums. Another example might be a history of losses analyzed by kinds. In other words, these multiplicative processes follow the distributive law for multiplication.

NOTE 3

Extensions of the graduated series may be useful. In casualty actuarial work, smoothing or graduation of a time series is quite likely to be used for prediction of the future. Interest concentrates on extension of the series, and the smoothing process is assumed to reveal, more or less, an underlying law.

A straight line or parabola, fitted by least squares, is sometimes calculated and extended. As said in connection with Note 2, these are the extreme cases of Whittaker-Henderson A graduations, resulting when the constant called a herein approaches infinity. If such extensions are valid, extensions derived with lesser a's should have validity also. The difference would be that the crude or observed terms nearest the end influence the extension more than the others, whereas in the usual least squares fitting, all observed terms influence the result alike.

The calculation of valid extensions of any length involves the retention of more places than required for a graduation over the observed interval only, that is, if the extensions are to be considered meaningful and not merely auxiliaries of the graduation process. The reason is that the extended terms are in a difference series of $(z-1)$th order and the rounding error in the difference used cumulates. To illustrate, extensions from Exhibit II and extensions from the "second Henderson" method are given below. In the case of sixth-order graduations, the rounding error would cause a greater divergence.

x	*Exhibit II*	*"Second Henderson"*
−4	17.75	17.77953
−3	20.16	20.18371
−2	22.57	22.58789
−1	24.98	24.99207
0_α	27.39	27.39625
1	29.80	29.80043
17	117.36	117.36378
18_ω	126.74	126.74849
19	136.12	136.13320
20	145.50	145.51791
21	154.88	154.90262
22	164.26	164.28733

APPENDIX

The formula by which the individual terms of the u''', u', and u series are calculated, corresponding to a given u'' series, may be applied in the following forms:

$$
\begin{aligned}
Au_x''' &= Bu_{x+1}''' - Cu_{x+2}''' + Du_{x+3}''' + Eu_{x-a}'' \\
Au_x' &= Bu_{x-1}' - Cu_{x-2}' + Du_{x-3}' + Eu_{x+a}'' \\
Au_x &= Bu_{x+1} - Cu_{x+2} + Du_{x+3} + Eu_{x-a}'
\end{aligned}
$$

The value z may be 1, 2, or 3 in a Whittaker-Henderson A graduation and the value a may be any positive number.

When $z = 1$, $A = (a + 1)$; $B = a$; $C = 0$; $D = 0$; $E = 1$

When $z = 2$, $A = (a + 1)(a + 2)$; $B = 2a(a + 2)$; $C = a(a + 1)$; $D = 0$; $E = 2$

When $z = 3$, $A = (a + 1)(a + 2)^2(a + 3)$

$B = a(a + 2)(a + 3)(3a + 5)$

$C = a(a + 1)(a + 3)(3a + 4)$

$D = a(a + 1)^2(a + 2)$

$E = 4(2a + 3)$

These are adapted from C. A. Spoerl's paper, page 408. As in Note 1, the intermediate variable in this Appendix is u_{x-a}', like Henderson's u_{x-n}'. If the variable is termed u_x', all a's would drop out of the subscripts.

The subscripts in the formula as shown assume the u''' and u columns are being worked out upward and the u' column is being worked out downward. Since the columns might just as well be worked out in the opposite directions, that is, u_x''' downward, u_x' upward, and u_x downward, and further since Note 1 proposes iteration of the u' series in reverse directions, the subscripts on the right hand may also be:

$$
\begin{gathered}
x - 1, x - 2, x - 3, x + a \\
x + 1, x + 2, x + 3, x - a \\
x - 1, x - 2, x - 3, x + a
\end{gathered}
$$

The coefficients resulting from certain choices of values for z and a are tabulated:

z	a	A	B	C	D	E
1	1	2	1	0	0	1
1	2	3	2	0	0	1
1	3	4	3	0	0	1
2	1	3	3	− 1	0	1
2	2	6	8	− 3	0	1
2	3	10	15	− 6	0	1
2	4	15	24	− 10	0	1
3	1	18	24	− 14	3	5
3	2	60	110	− 75	18	7
3	3	100	210	−156	40	6
3	4	315	714	−560	150	11

The case, $z = 1$, is hardly useful. The case, $z = 2$, being a good deal easier to work with than the case $z = 3$, is widely used. The case, $z = 3$, makes a fine graduation. It is to be used when the differences in some parts of the distribution are large relative to those in other parts, as in a bell-shaped distribution or one with large differences at one or both ends.

Note that some of the choices of constants are more convenient or useful than others. The combination $z = 2$, $a = 1$ makes for a very rapid calculation and would be a good choice for becoming familiar with the process, as well as for practical applications. Where $a = 3$, whether z is 2 or 3, A is a power of 10 so it is not necessary to divide through by A or use fractional multipliers. This is quite convenient, and these are among the most useful combinations.

The problem of starting terms is treated in Note 1. At the end away from the start, the calculation of any of these series but the final one runs into the problem that the u_x''s give out before the series can be completed. The rest of the column is filled in by means of a difference series continuing such a series established by the last z terms it was possible to calculate by formula.

The columns are carried $(x + a)$ terms beyond the end of $\{u_x''\}$ so that the outer z terms may serve as successively corrected starting terms. When one of the columns has been selected as the final u_x' column, back over which the u_x column will be developed, it is usual to "turn the corner" and start $\{u_x\}$ by copying the last z terms of $\{u_x'\}$, excluding any extension, into the u_x column, but no harm is done by turning farther out, as the last z terms of $\{u_x'\}$ proper and the terms of the extension are all of the same difference series.

The terms appended have z^{th} differences of zero and we easily caculated by means of the formulas as next given, prime marks omitted:

If $z = 1$, $u_x = u_{x-1}$,

If $z = 2$, $u_x = 2u_{x-1} - u_{x-2}$,

If $z = 3$, $u_x = 3u_{x-1} - 3u_{x-2} + u_{x-3}$,

provided the columns are downward. Again, when the columns are in reverse direction and developing upward, the subscripts on the right-hand side become $x + 1$, $x + 2$, and $x + 3$.

EXHIBIT I

Specimen Graduation by Formula A
Method of Note 1; Henderson Type Start

x	u'''_x	u''_x	$_1u'_x$	$_2u'_x$ (Final u'_x)	$_3u'_x$	u_x
-4	17.60		17.60	17.74	17.74	
-3	20.05		20.05	20.15	20.15	
-2	22.50		23.60	22.56	23.66	
-1	24.95		25.44	24.97	25.47	
0	27.40	34	27.29	27.38	27.30	27.38
1	29.85	24	30.33	29.79	30.33	29.79
2	32.30	31	31.80	32.20	31.79	32.57
3	34.75	40	35.40	34.61	35.39	35.76
4	39.41	30	39.30	39.22		39.32
5	43.59	49	42.70	43.36		43.44
6	47.75	48	48.45	47.52		47.79
7	53.48	48	52.92	53.33		52.36
8	57.11	67	57.50	57.18		57.15
9	61.67	58	62.71	62.16		61.92
10	66.23	67	67.53	67.41		66.98
11	69.27	75	71.35	71.43		72.42
12	74.60	76	78.37	77.99		78.32
13	79.72	76	85.48	84.44		84.91
14	85.73	102	91.62	91.52		92.29
15	94.17	100	98.59	100.18		100.07
16	105**	101	107.98	109.44		108.38
17	117 *	115	117.37	116.81		117.37
18		134	126.76	126.62		126.76
19			136.15	137.37		
20			145.54	146.75		
21			154.93	154.93		
22			164.32	164.32		
		1275				1274.98

* 1/3 (101 + 115 + 134) = 117

** 1/3 (100 + 101 + 115) = 105

$z = 2$; $a = 2$

$$u'''_x = 1/6\,(8u'''_{x+1} - 3u'''_{x+2} + u''_{x-2})$$

$$u'_x = 1/6\,(8u'_{x\mp1} - 3u'_{x\pm2} + u''_{x\pm2})$$

$$u_x = 1/6\,(8u_{x-1} - 3u_{x-2} + u'_{x+2})$$

EXHIBIT II Specimen Graduation by Formula A
Page 39 of "Elements of Graduation" Reworked
Spoerl's Correction Applied

x	u^*_x	u''_x	u'_x	u_x	u^*_x	u'_x	u_x
-4	14 →		14.00		17.75	17.75	
-3	17 →		17.00		20.16	20.16	
-2	20		21.33		22.57	23.67	
-1	23		23.94		24.98	25.48	
0	26	34	26.42	26.40	27.39	27.30	27.39
1	29	24	29.92	29.27	29.80	30.33	29.80
2		31	31.68	32.36		31.79	32.57
3		40	35.45	35.74		35.39	35.75
4		30	39.43	39.39		39.29	39.30
5		49	42.85	43.54		42.69	43.41
6		48	48.58	47.89		48.44	47.76
7		48	53.02	52.44		52.91	52.34
8		67	57.57	57.21		57.49	57.14
9		58	62.75	61.97		62.70	61.92
10		67	67.55	67.02		67.52	66.99
11		75	71.36	72.45		71.34	72.43
12		76	78.37	78.34		78.36	78.33
13		76	85.48	84.92		85.48	84.92
14		102	91.62	92.30		91.63	92.30
15		100	98.59	100.07		98.60	100.07
16		101	107.98	108.38		107.98	108.38
17		115	117.37 →	117.37		117.36	117.36
18		134	126.76 →	126.76		126.74	126.74
		1275		1273.82			1274.90

Corrections per formula 5.37, page 37, "Elements of Graduation"

Corrected u_0 = 26.40 + 9/2 (26.40 - 26) - 3 (29.27 - 29) = 27.39

Corrected u_1 = 29.27 + 7/3 (26.40 - 26) - 3/2 (29.27 - 29) = 29.80

$z = 2$; $a = 2$

EXHIBIT II A Specimen Graduation by Formula A
Method of Note 1; with Monograph's Start & Placement

x		u^*_x	u''_x	${}_1u'_x$	${}_2u'_x$	${}_3u'_x$	u_x
					17.74		
					20.15		
-4		14			22.56		
-3		17			24.97		
-2		20		14.00	27.38	17.74	
-1		23		17.00	29.79	20.15	
0	α	26	34	21.33	32.20	23.66	27.39
1		29	24	23.94	34.61	25.47	29.80
2			31	26.42	39.22	27.30	32.57
3			40	29.92	43.36	30.33	35.75
4			30	31.68	47.52	31.79	39.30
5			49	35.45	53.33	35.39	43.41
6			48	39.43	57.18	39.29	47.76
7			48	42.85	62.16	42.69	52.34
8			67	48.58	67.41	48.44	57.14
9			58	53.02	71.43	52.91	61.92
10			67	57.57	77.99	57.49	66.99
11			75	62.75	84.44	62.70	72.43
12			76	67.55	91.52	67.52	78.33
13			76	71.36	100.18	71.34	84.92
14			102	78.37	109.44	78.36	92.30
15			100	85.48	116.81	85.48	100.07
16			101	91.62	126.62	91.63	108.36
17			115	98.59	137.37	98.60	117.36
18	ω		134	107.98	146.75	107.98	126.74
19				117.37	154.93		
20				126.76	164.32		
21				136.15			
22				145.54			
				154.93			
				164.32			
Σ			1275				1274.90