**TRANSACTIONS OF SOCIETY OF ACTUARIES
1982 VOL. 34**

AN EXTENSION OF THE WHITTAKER-HENDERSON METHOD OF GRADUATION

WALTER B. LOWRIE

ABSTRACT

The purpose of this paper is to introduce a practical method for interpolation and smoothing and give a number of numerical examples using the method. The method is a variation of the Whittaker-Henderson method. The criterion of smoothness used is a polynomial plus an exponential function. If interpolation is desired, then zero weights are used where the method is to supply values, and positive weights are used at the pivotal points. Programs in the APL language are given in Appendix II.

I. DEFINITIONS OF SMOOTHING AND INTERPOLATION

The numerical values resulting from an experiment can be too rough for practical use. The objective is to generate smoother values. It may also be necessary to fill in missing values. That is, we want a set of values, $u_1, u_2, \ldots, u_n$, defined on a grid of points (the positive integers), 1, 2, . . . , n, to be generated from experimental values $u''_1, \ldots, u''_n$. The experimental values may be defined on the entire grid or on a proper subset of the grid. When experimental values are given for the entire grid we have smoothing, and when they are defined on a proper subset of the grid we have interpolation. Whenever u''_i is not defined, w_i is taken to be zero in the expression below. This idea is mentioned by Kimeldorf and Jones ([3], p. 91). The interpolation process may, or may not, smooth the given data in addition to filling in missing values. When values are defined on a proper subset of the grid, these values are called pivotal values.

Discussion

The method we use results from minimizing

$$(1 - l)F + lF' + mS$$
$$= (1 - l) \sum_{i=1}^{n} w_i(u_i - u''_i)^2 + l \sum_{i=1}^{n} w'_i(u_i - s_i)^2$$
$$+ m \sum_{i=1}^{n-z} (\Delta^z u_i - r\Delta^{z-1} u_i)^2 ,$$

where, for $i = 1, 2, \ldots, n$,

u_i is a graduated value;
u''_i is a given data value;
s_i is a standard value (e.g., a prior mean value; see Kimeldorf and Jones [3]);
w_i is a weight applied to an ungraduated value; $w_i \geq 0$;
w'_i is a weight applied to a standard value; $w'_i > 0$;
$z = 1, 2, 3, \ldots$ is the degree of differencing chosen;
l ($0 \leq l \leq 1$) is the relative emphasis placed on the "standard" values;
m ($m > 0$) is the emphasis on smoothness (usually larger than the mean value of the weights, $\Sigma_{i=1}^{n} w_i/n$); and
r ($r > -1$) is discussed in the paper.

The function $(1 - l)F + lF' + mS$ is used to measure the adequacy of the graduation. Smaller values of this function correspond to better graduations, by our criteria. The weights (w_i) and (w'_i) allow varying emphasis on the data values. Using them, we can avoid undue emphasis where there is low exposure.

The term F is a measure of closeness between the graduated values and the ungraduated values in the least-squares sense.

The term F' is a measure of closeness of the graduated values to a set of standard values (s_i). These standard values are introduced to allow the graduator to impose his judgment on the graduated values. The standard values could come from published data relevant to the graduation, for example, a mortality table. They could also be constructed in light of the graduator's past experience.

The constant l allows the graduator to determine the relative emphasis between the closeness to the observed values (u''_i) and closeness to the standard values (s_i).

The term S is a measure of smoothness in the sense that, if it is small, the graduated values (u_i) are close to an exponential function plus a polynomial of degree $z - 2$ or less. For example, look at one term of S ($z = 3$):

$$\Delta^2[ai + b + (1 + r)^i] = r^2(1 + r)^i ,$$

$$\Delta^3[ai + b + (1 + r)^i] = r^3(1 + r)^i = r\Delta^2[ai + b + (1 + r)^i] ,$$

so

$$\Delta^3[ai + b + (1 + r)^i] - r\Delta^2[ai + b + (1 + r)^i] = 0 .$$

In general, if $u_i = P_{z-2}(i) + k(1 + r)^i$, where $P_{z-2}(i)$ is a polynomial of degree $z - 2$, or less, then S is zero. If the u_i'' are of this form, they are left invariant if $l = 0$. If $r = 0$, this process is the Whittaker-Henderson process and uses a polynomial of degree $z - 1$, or less, as a criterion of smoothness. If $z = 1$, the process uses an exponential function, $k(1 + r)^i$, as a criterion of smoothness.

The constant m allows the graduator to determine the relative emphasis on smoothness. If m is large, graduated values will be very close to an exponential function plus a polynomial.

Intuitively, the Whittaker-Henderson process generates values $u_1, u_2, \ldots, u_n$ that are closer to the function that serves as the criterion of fit.

Using matrix notation, we show in Appendix I that

$$(1 - l)F + lF' + mS = U^TAU - 2U^TB + C ,$$

where

$$W = \begin{bmatrix} w_1 & & & 0 \\ & w_2 & & \\ & & \ddots & \\ 0 & & & w_n \end{bmatrix}, \quad W' = \begin{bmatrix} w_1' & & & 0 \\ & w_2' & & \\ & & \ddots & \\ 0 & & & w_n' \end{bmatrix};$$

$$U = (u_1, u_2, \ldots, u_n)^T , \quad U'' = (u_1'', u_2'', \ldots, u_n'')^T ;$$

$$S = (s_1, s_2, \ldots, s_n)^T ;$$

$$KU = \begin{bmatrix} \Delta^z u_1 - r\Delta^{z-1}u_1 \\ \Delta^z u_2 - r\Delta^{z-1}u_2 \\ \cdot \\ \cdot \\ \cdot \\ \Delta^z u_{n-z} - r\Delta^{z-1}u_{n-z} \end{bmatrix} \quad \text{(a column vector of length } n - z) ;$$

$$A = (1 - l)W + lW' + mK^TK ;$$

$$B = (1 - l)WU'' + lW'S ;$$

C does not depend on U.

A theorem from linear algebra tells us that if the matrix A is positive definite, then

$$U^TAU - 2U^TB$$

takes its minimum value for the vector U that is defined by

$$AU = B .$$

A reference for this theorem is Noble ([5], p. 400).

Interpolation

Whittaker-Henderson methods have already been used to smooth data. It turns out that we can use these methods to interpolate as well. This is done by using positive weights at the pivotal points and zero weights where the method is to supply values.

In Appendix I, we prove that a set of generated values $u_1, u_2, \ldots, u_n$ exists minimizing $(1 - l)F + lF' + mS$, if the number of pivotal points is equal to or greater than z. The set of values, $u_1, u_2, \ldots, u_n$, also exists if $l > 0$ and all $w'_i > 0$. The values may also exist under other circumstances, but this matter is not dealt with in this paper.

II. NUMERICAL EXAMPLES

Three different groups of examples are dealt with in this paper. The first group uses female standard issue ordinary mortality data, nonmedical issues of 1963–77, from *TSA, 1979 Reports* ([4], pp. 45–49). The issue ages are 20–24, where there is relatively heavy exposure. The *data values* are determined by taking 1,000 times the actual deaths divided by the exposed to risk ($1{,}000q_x$). The *standard values* are determined by taking 1,000 times the expected deaths divided by the exposed to risk. The expected deaths are based on the 1965–70 Female Select Basic Table. Both the *weights* and the *standard weights* are the exposed-to-risk values.

The first set of tables (1A, 1B, and 1C) uses data from policy years 1–14 by policy year. This tests the smoothing ability of the Whittaker-Henderson formula without doing any interpolation.

The second set of tables (2A, 2B, and 2C) combines exposed to risk and combines actual and expected deaths in the following pattern:

Cell	Central Policy Year	Policy Years Combined
1	1	1
2	3	2–4
3	7	5–9
4	12	10–14

The data values and standard values for each cell are calculated similarly to the way they were calculated in the first set. For example,

$$1{,}000q_2 = \frac{d_1 + d_2 + d_3}{E_1 + E_2 + E_3} \times 1{,}000 \,,$$

where the d's are deaths and the E's are exposures. The second set is constructed to test the interpolation and smoothing capabilities of the Whittaker-Henderson formula on unequally spaced data.

TABLE 1A

FEMALE MORTALITY, ISSUE AGES 20–24 [4]: $z = 3$, $r = 0$, $l = 0.0$, $m = 1{,}000$

i	Standard Weights	Standard Values	Weights	Data Values	Graduated Values	Differences		
						First	Second	Third
1	2,115,646	.50008	2,115,646	.25288	.25295	.05608	−.07953	.16097
2	1,457,640	.54540	1,457,640	.30940	.30903	−.02345	.08145	−.06047
3	1,073,275	.52456	1,073,275	.28511	.28558	.05800	.02098	−.15427
4	882,728	.53471	882,728	.34325	.34359	.07898	−.13330	.17194
5	719,869	.55010	719,869	.42369	.42256	−.05432	.03865	.10761
6	570,331	.58036	570,331	.36821	.36824	−.01567	.14626	−.32691
7	472,780	.60916	472,780	.34900	.35257	.13059	−.18065	.55751
8	402,142	.63908	402,142	.49236	.48316	−.05006	.37686	−.93967
9	322,480	.66361	322,480	.41553	.43310	.32680	−.56281	.84721
10	276,229	.68059	276,229	.78196	.75990	−.23601	.28440	−.17601
11	232,221	.71484	232,221	.50814	.52389	.04839	.10840	−.35172
12	186,412	.77785	186,412	.57400	.57228	.15679	−.24333	....
13	162,016	.85177	162,016	.73450	.72907	−.08654	....	....
14	142,175	.94250	142,175	.64006	.64253	....	....	....

TABLE 1B

FEMALE MORTALITY, ISSUE AGES 20–24 [4]: $z = 3$, $r = 0$, $l = 0.0$, $m = 1{,}000{,}000$

i	Standard Weights	Standard Values	Weights	Data Values	Graduated Values	Differences		
						First	Second	Third
1	2,115,646	.50008	2,115,646	.25288	.25643	.03408	−.00703	.00752
2	1,457,640	.54540	1,457,640	.30940	.29052	.02705	.00049	−.00496
3	1,073,275	.52456	1,073,275	.28511	.31757	.02754	−.00447	−.00262
4	882,728	.53471	882,728	.34325	.34511	.02307	−.00709	.01618
5	719,869	.55010	719,869	.42369	.36818	.01598	.00909	.01148
6	570,331	.58036	570,331	.36821	.38415	.02507	.02057	−.00763
7	472,780	.60916	472,780	.34900	.40922	.04563	.01294	−.01268
8	402,142	.63908	402,142	.49236	.45485	.05857	.00026	−.01874
9	322,480	.66361	322,480	.41553	.51342	.05883	−.01848	.00574
10	276,229	.68059	276,229	.78196	.57225	.04035	−.01274	.00284
11	232,221	.71484	232,221	.50814	.61260	.02761	−.00991	−.00319
12	186,412	.77785	186,412	.57400	.64020	.01770	−.01310	....
13	162,016	.85177	162,016	.73450	.65790	.00460	....	....
14	142,175	.94250	142,175	.64006	.66250	....	....	....

The second group is the same as the first group, except that issue age 5–9 data were used. The exposure was much lower, so the data values ($1{,}000q_x$) showed much more statistical variation than in the first group. This is used as a test of how well the Whittaker-Henderson method handles rough data. These interpolations appear in Tables 3A and 3B (distinct data) and Tables 4A and 4B (combined data).

TABLE 1C

FEMALE MORTALITY, ISSUE AGES 20–24 [4]: $z = 3$, $r = 0.05$, $l = 0.2$, $m = 1{,}000{,}000$

i	Standard Weights	Standard Values	Weights	Data Values	Graduated Values	Differences		
						First	Second	Third
1	2,115,646	.50008	2,115,646	.25288	.30640	.03123	−.00787	.00783
2	1,457,640	.54540	1,457,640	.30940	.33762	.02335	−.00005	−.00208
3	1,073,275	.52456	1,073,275	.28511	.36097	.02330	−.00213	−.00153
4	882,728	.53471	882,728	.34325	.38428	.02117	−.00367	.01178
5	719,869	.55010	719,869	.42369	.40545	.01750	.00811	.00804
6	570,331	.58036	570,331	.36821	.42295	.02561	.01615	−.00607
7	472,780	.60916	472,780	.34900	.44856	.04176	.01008	−.00918
8	402,142	.63908	402,142	.49236	.49032	.05184	.00090	−.01332
9	322,480	.66361	322,480	.41553	.54217	.05275	−.01242	.00515
10	276,229	.68059	276,229	.78196	.59491	.04033	−.00728	.00240
11	232,221	.71484	232,221	.50814	.63524	.03305	−.00488	−.00260
12	186,412	.77785	186,412	.57400	.66829	.02817	−.00748	....
13	162,016	.85177	162,016	.73450	.69646	.02069	....	....
14	142,175	.94250	142,175	.64006	.71715	....	....	....

TABLE 2A

FEMALE MORTALITY, ISSUE AGES 20–24 [4], DATA COMBINED: $z = 3$, $r = 0$, $l = 0.0$, $m = 1{,}000$

i	Standard Weights	Standard Values	Weights	Data Values	Graduated Values	Differences		
						First	Second	Third
1	2,115,646	.50008	2,115,646	.25288	.25288	.03069	−.00374	.00058
2	1,457,640	.54540	0	....	.28357	.02695	−.00317	.00173
3	1,073,275	.52456	3,413,643	.31052	.31052	.02378	−.00144	.00240
4	882,728	.53471	0	....	.33430	.02234	.00097	.00260
5	719,869	.55010	0	....	.35664	.02331	.00356	.00231
6	570,331	.58036	0	....	.37995	.02687	.00537	.00154
7	472,780	.60916	2,487,602	.40682	.40682	.03274	.00740	.00092
8	402,142	.63908	0	....	.43955	.04014	.00833	.00046
9	322,480	.66361	0	....	.47969	.04847	.00879	.00015
10	276,229	.68059	0	....	.52816	.05726	.00894	−.00000
11	232,221	.71484	0	....	.58542	.06620	.00894	−.00000
12	186,412	.77785	999,053	.65162	.65162	.07514	.00894	....
13	162,016	.85177	0	....	.72676	.08408	....	....
14	142,175	.94250	0	....	.81084	....	....	....

The third group uses the 1969–71 United States male values of l_x. Monthly values are calculated by interpolation. Standard values are obtained by inserting 11 geometric means between each pair of annual values of l_x. These interpolations appear in Tables 5A, 5B, and 5C.

The graduations of the first set of data in the first group are shown in Tables 1A, 1B, and 1C. The graduation in Table 1A is not very good, since

TABLE 2B

FEMALE MORTALITY, ISSUE AGES 20–24 [4], DATA COMBINED:
$z = 3, r = 0, l = 0.0, m = 1{,}000{,}000$

i	Standard Weights	Standard Values	Weights	Data Values	Graduated Values	Differences		
						First	Second	Third
1	2,115,646	.50008	2,115,646	.25288	.25314	.03031	−.00355	.00056
2	1,457,640	.54540	0	....	.28345	.02676	−.00299	.00168
3	1,073,275	.52456	3,413,643	.31052	.31022	.02378	−.00130	.00234
4	882,728	.53471	0	....	.33400	.02248	.00104	.00253
5	719,869	.55010	0	....	.35647	.02351	.00356	.00225
6	570,331	.58036	0	....	.37999	.02708	.00581	.00150
7	472,780	.60916	2,487,602	.40682	.40707	.03289	.00731	.00090
8	402,142	.63908	0	....	.43995	.04020	.00821	.00045
9	322,480	.66361	0	....	.48015	.04840	.00865	.00015
10	276,229	.68059	0	....	.52855	.05706	.00880	.00000
11	232,221	.71484	0	....	.58561	.06586	.00880	.00000
12	186,412	.77785	999,053	.65162	.65147	.07466	.00880	....
13	162,016	.85177	0	....	.72613	.08347	....	....
14	142,175	.94250	0	....	.80960	....	....	....

TABLE 2C

FEMALE MORTALITY, ISSUE AGES 20–24 [4], DATA COMBINED:
$z = 3, r = 0.05, l = 0.2, m = 1{,}000{,}000$

i	Standard Weights	Standard Values	Weights	Data Values	Graduated Values	Differences		
						First	Second	Third
1	2,115,646	.50008	2,115,646	.25288	.30997	.02486	−.00590	.01513
2	1,457,640	.54540	....	....	.33483	.01896	.00922	−.01248
3	1,073,275	.52456	3,413,643	.31052	.35379	.02818	−.00325	−.00551
4	882,728	.53471	0	....	.38198	.02493	−.00876	.01037
5	719,869	.55010	0	....	.40690	.01617	.00161	.01555
6	570,331	.58036	0	....	.42307	.01778	.01716	−.00703
7	472,780	.60916	2,487,602	.40682	.44085	.03494	.01014	−.00808
8	402,142	.63908	0	....	.47579	.04508	.00206	−.00011
9	322,480	.66361	0	....	.52087	.04714	.00195	.00812
10	276,229	.68059	0	....	.56801	.04909	.01007	.01073
11	232,221	.71484	0	....	.61710	.05916	.02080	.00338
12	186,412	.77785	999,053	.65162	.67626	.07995	.02417	....
13	162,016	.85177	0	....	.75621	.10413	....	....
14	142,175	.94250	0	....	.86034	....	....	....

m is much lower than the mean value of the weights, but the fit is excellent. The graduated values oscillate. In Table 1B, m is raised to 1,000,000, giving a good graduation. Table 1C uses $l = 0.2$ and $r = 0.05$. This forces the graduated values to be closer to the standard values and closer to an exponential function plus a straight line, $P_1(i) + k(1.05)^i$. The fit is not as good as in Tables 1A and 1B, but the graduated values are close to values defined by a reasonable set of preconceived notions. Of course, other reasonable sets of preconceived notions exist.

TABLE 3A

FEMALE MORTALITY, ISSUE AGES 5–9 [4]: $z = 3$, $r = 0$, $l = 0.1$, $m = 1{,}000{,}000$

i	Standard Weights	Standard Values	Weights	Data Values	Graduated Values	Differences: First	Second	Third
1	341,105	.31955	341,105	.08209	.13173	.04099	−.03095	.00883
2	258,501	.28627	258,501	.21277	.17273	.01004	−.02212	.01425
3	189,815	.26341	189,815	.15805	.18277	−.01207	−.00786	.01895
4	183,797	.25572	183,797	.23395	.17069	−.01994	.01109	.01089
5	162,903	.26396	162,903	.06139	.15076	−.00885	.02198	.00135
6	130,979	.27485	130,979	.07635	.14191	.01313	.02333	−.00371
7	79,981	.30007	79,981	.10002	.15504	.03646	.01962	−.00146
8	62,220	.33751	62,220	.35358	.19150	.05608	.01816	−.00191
9	53,328	.37504	53,328	.26253	.24758	.07424	.01625	−.00644
10	50,526	.41563	50,526	.15833	.32183	.09050	.00981	−.00810
11	46,359	.47456	46,359	.40984	.41232	.10030	.00171	−.00708
12	42,641	.53939	42,641	.42213	.51262	.10201	−.00537	....
13	41,126	.58357	41,126	.97262	.61463	.09664	....	....
14	35,676	.58863	35,676	.50454	.71127	....	....	....

TABLE 3B

FEMALE MORTALITY, ISSUE AGES 5–9 [4]: $z = 3$, $r = 0$, $l = 0.5$, $m = 100{,}000{,}000$

i	Standard Weights	Standard Values	Weights	Data Values	Graduated Values	Differences: First	Second	Third
1	341,105	.31955	341,105	.08209	.23039	−.01982	.00837	.00010
2	258,501	.28627	258,501	.21277	.21057	−.01146	.00847	.00020
3	189,815	.26341	189,815	.15805	.19911	−.00299	.00867	.00028
4	183,797	.25572	183,797	.23395	.19612	.00568	.00895	.00025
5	162,903	.26396	162,903	.06139	.20180	.01463	.00920	.00017
6	130,979	.27485	130,979	.07635	.21644	.02383	.00937	.00009
7	79,781	.30007	79,981	.10002	.24027	.03320	.00946	.00006
8	62,220	.33751	62,220	.35358	.27347	.04266	.00952	.00002
9	53,328	.37504	53,328	.26253	.31613	.05218	.00954	−.00004
10	50,526	.41563	50,526	.15833	.36831	.06172	.00950	−.00005
11	46,359	.47456	46,359	.40984	.43003	.07123	.00945	−.00004
12	42,641	.53939	42,641	.42213	.50126	.08067	.00940	....
13	41,126	.58357	41,126	.97262	.58194	.09008	....	....
14	35,676	.58863	35,676	.50454	.67201	....	....	....

Tables 2A, 2B, and 2C correspond to Tables 1A, 1B, and 1C, respectively, since each pair uses the same parameters. The difference is that the data in Tables 2A, 2B, and 2C are combined into various pivotal points. The graduated values in Table 2A are very smooth. Assuming that the process of combining data has reduced offsetting random error, then the graduation in Table 2A is probably quite good. The graduation in Table

TABLE 4A

FEMALE MORTALITY, ISSUE AGES 5–9 [4], DATA COMBINED:
$z = 3, r = 0, l = 0.1, m = 1{,}000{,}000$

i	Standard Weights	Standard Values	Weights	Data Values	Graduated Values	Differences		
						First	Second	Third
1	341,105	.31955	341,105	.08209	.11785	.05605	−.03011	.00410
2	258,501	.28627	0	....	.17390	.02593	−.02601	.00940
3	189,815	.26341	632,113	.20250	.19983	−.00008	−.01661	.01317
4	183,797	.25572	0	....	.19976	−.01669	−.00344	.01439
5	162,903	.26396	0	....	.18307	−.02013	.01095	.01174
6	130,979	.27485	0	....	.16294	−.00918	.02269	.00375
7	79,981	.30007	489,411	.13077	.15375	.01351	.02644	−.00062
8	62,220	.33751	0	....	.16726	.03995	.02583	−.00243
9	53,328	.37504	0	....	.20721	.06577	.02340	−.00258
10 ...	50,526	.41563	0	....	.27298	.08917	.02082	−.00178
11 ...	46,359	.47456	0	....	.36215	.10999	.01903	−.00057
12 ...	42,641	.53939	216,328	.47613	.47214	.12902	.01846	....
13 ...	41,126	.58357	0	....	.60116	.14748	....	....
14 ...	35,676	.58863	0	....	.74864	....	....	....

TABLE 4B

FEMALE MORTALITY, ISSUE AGES 5–9 [4], DATA COMBINED:
$z = 3, r = 0, l = 0.5, m = 100{,}000{,}000$

i	Standard Weights	Standard Values	Weights	Data Values	Graduated Values	Differences		
						First	Second	Third
1	341,105	.31955	341,105	.08209	.23559	−.01884	.00735	.00012
2	258,501	.28627	0	....	.21676	−.01148	.00747	.00027
3	189,815	.26341	632,113	.20250	.20527	−.00402	.00774	.00040
4	183,797	.25572	0	....	.20125	.00372	.00813	.00046
5	162,903	.26396	0	....	.20497	.01185	.00859	.00040
6	130,979	.27485	0	....	.21682	.02044	.00899	.00020
7	79,981	.30007	489,411	.13077	.23726	.02943	.00919	.00007
8	62,220	.33751	0	....	.26670	.03862	.00926	.00000
9	53,328	.37504	0	....	.30532	.04788	.00927	−.00002
10 ...	50,526	.41563	0	....	.35320	.05715	.00925	−.00002
11 ...	46,359	.47456	0	....	.41034	.06639	.00923	−.00001
12 ...	42,641	.53939	216,328	.47613	.47674	.07562	.00922	....
13 ...	41,126	.58357	0	....	.55235	.08483	....	....
14 ...	35,676	.58863	0	....	.63719	....	....	....

TABLE 5A

MONTHLY VALUES OF l_x OBTAINED BY INTERPOLATION USING 1969–71 UNITED STATES MALE VALUES: $z = 3$, $r = -0.0029$, $l = 0.0$, $m = 10$

i	Standard Weights	Standard Values	Weights	Data Values	Graduated Values	Differences: First	Second	Third
1	1	100,000	1	100,000	99,999.83	−303.42	21.89	− .08
2	1	99,811	0	....	99,696.41	−281.53	21.81	− .11
3	1	99,622	0	....	99,414.88	−259.72	21.70	− .17
4	1	99,434	0	....	99,155.16	−238.02	21.53	− .23
5	1	99,246	0	....	98,917.14	−216.48	21.30	− .32
6	1	99,058	0	....	98,700.66	−195.19	20.98	− .42
7	1	98,871	0	....	98,505.47	−174.21	20.56	− .54
8	1	98,684	0	....	98,331.26	−153.65	20.01	− .68
9	1	98,498	0	....	98,177.61	−133.64	19.34	− .83
10	1	98,311	0	....	98,043.98	−114.30	18.50	−1.00
11	1	98,126	0	....	97,929.68	− 95.80	17.50	−1.19
12	1	97,940	0	....	97,833.88	− 78.30	16.31	−1.40
13	1	97,755	1	97,755	97,755.58	− 61.99	14.91	−1.56
14	1	97,744	0	....	97,693.59	− 47.08	13.35	−1.68
15	1	97,733	0	....	97,646.51	− 33.73	11.66	−1.77
16	1	97,722	0	....	97,612.78	− 22.07	9.90	−1.81
17	1	97,712	0	....	97,590.71	− 12.17	8.09	−1.80
18	1	97,701	0	....	97,578.54	− 4.08	6.29	−1.76
19	1	97,690	0	....	97,574.46	2.21	4.53	−1.68
20	1	97,679	0	....	97,576.68	6.74	2.85	−1.55
21	1	97,668	0	....	97,583.42	9.59	1.30	−1.38
22	1	97,657	0	....	97,593.01	10.90	− .08	−1.18
23	1	97,647	0	....	97,603.91	10.81	− 1.26	− .93
24	1	97,636	0	....	97,614.72	9.56	− 2.18	− .63
25	1	97,625	1	97,625	97,624.28	7.37	− 2.82	− .37
26	1	97,617	0	....	97,631.65	4.56	− 3.19	− .14
27	1	97,610	0	....	97,636.21	1.37	− 3.34	.05
28	1	97,602	0	....	97,637.57	− 1.97	− 3.28	.22
29	1	97,594	0	....	97,635.60	− 5.25	− 3.06	.36
30	1	97,587	0	....	97,630.35	− 8.32	− 2.71	.46
31	1	97,579	0	....	97,622.03	− 11.02	− 2.25	.54
32	1	97,571	0	....	97,611.01	− 13.27	− 1.71	.58
33	1	97,564	0	....	97,597.74	− 14.98	− 1.13	.59
34	1	97,556	0	....	97,582.76	− 16.11	− .54	.57
35	1	97,548	0	....	97,566.65	− 16.65	.03	.52
36	1	97,541	0	....	97,550.00	− 16.62	.55	.44
37	1	97,533	1	97,533	97,533.38	− 16.07	.99	.37
38	1	97,527	0	....	97,517.30	− 15.09	1.35	.30
39	1	97,520	0	....	97,502.22	− 13.73	1.65	.24
40	1	97,514	0	....	97,488.49	− 12.08	1.89	.18
41	1	97,508	0	....	97,476.41	− 10.19	2.07	.14
42	1	97,502	0	....	97,466.21	− 8.12	2.21	.09
43	1	97,495	0	....	97,458.09	− 5.92	2.30	.06
44	1	97,489	0	....	97,452.17	− 3.62	2.36	.03
45	1	97,483	0	....	97,448.56	− 1.26	2.39	.01
46	1	97,477	0	....	97,447.30	1.14	2.41	− .00
47	1	97,470	0	....	97,448.44	3.54	2.41	....
48	1	97,464	0	....	97,451.98	5.95	....	....
49	1	97,458	1	97,458	97,457.93	....	....	....

2B looks good in relation to that in Table 1B. The graduations obtained by grouping data seem to be quite good, since the pivotal values are smooth. Imposing preconceived notions on the grouped data makes the graduated values in Table 2C depart more from the pivotal values than was the case in Tables 2A and 2B. Further, the graduation in Table 2C does not look as good as that in Tables 2A and 2B.

TABLE 5B

MONTHLY VALUES OF l_x OBTAINED BY INTERPOLATION USING 1969–71 UNITED STATES MALE VALUES: $z = 3$, $r = -0.0029$, $l = 0.0$, $m = 10$

i	Standard Weights	Standard Values	Weights	Data Values	Graduated Values	Differences		
						First	Second	Third
1	1	97,755	1	97,755	97,755.00	−12.60	.33	−.00
2	1	97,744	0	....	97,742.39	−12.27	.33	−.00
3	1	97,733	0	....	97,730.12	−11.95	.33	−.00
4	1	97,722	0	....	97,718.17	−11.62	.33	−.00
5	1	97,712	0	....	97,706.56	−11.29	.32	−.00
6	1	97,701	0	....	97,695.26	−10.97	.32	−.00
7	1	97,690	0	....	97,684.29	−10.65	.32	−.00
8	1	97,679	0	....	97,673.64	−10.34	.31	−.01
9	1	97,668	0	....	97,663.30	−10.02	.30	−.01
10	1	97,657	0	....	97,653.28	− 9.72	.30	−.01
11	1	97,647	0	....	97,643.56	− 9.42	.29	−.01
12	1	97,636	0	....	97,634.14	− 9.13	.28	−.01
13	1	97,625	1	97,625	97,625.00	− 8.86	.27	−.01
14	1	97,617	0	....	97,616.15	− 8.59	.25	−.01
15	1	97,610	0	....	97,607.56	− 8.34	.24	−.02
16	1	97,602	0	....	97,599.22	− 8.10	.22	−.02
17	1	97,594	0	....	97,591.12	− 7.88	.21	−.02
18	1	97,587	0	....	97,583.25	− 7.67	.19	−.02
19	1	97,579	0	....	97,575.58	− 7.48	.17	−.02
20	1	97,571	0	....	97,568.10	− 7.31	.16	−.02
21	1	97,564	0	....	97,560.79	− 7.15	.14	−.01
22	1	97,556	0	....	97,553.64	− 7.01	.13	−.01
23	1	97,548	0	....	97,546.64	− 6.88	.11	−.01
24	1	97,541	0	....	97,539.76	− 6.76	.10	−.01
25	1	97,533	1	97,533	97,533.00	− 6.66	.09	−.01
26	1	97,527	0	....	97,526.34	− 6.57	.08	−.01
27	1	97,520	0	....	97,519.77	− 6.48	.08	−.01
28	1	97,514	0	....	97,513.29	− 6.40	.07	−.00
29	1	97,508	0	....	97,506.88	− 6.33	.07	−.00
30	1	97,502	0	....	97,500.55	− 6.26	.06	−.00
31	1	97,495	0	....	97,494.29	− 6.20	.06	−.00
32	1	97,489	0	....	97,488.09	− 6.14	.06	−.00
33	1	97,483	0	....	97,481.95	− 6.08	.06	−.00
34	1	97,477	0	....	97,475.87	− 6.02	.06	−.00
35	1	97,470	0	....	97,469.86	− 5.96	.06	....
36	1	97,464	0	....	97,463.90	− 5.90	....	....
37	1	97,458	1	97,458	97,458.00	....	....	....

TABLE 5C

MONTHLY VALUES OF l_x OBTAINED BY INTERPOLATION USING 1969–71 UNITED STATES MALE VALUES: $z = 3$, $r = 0$, $l = 0.9$, $m = 10$

i	Standard Weights	Standard Values	Weights	Data Values	Graduated Values	Differences: First	Differences: Second	Differences: Third
1	1	100,000	1	100,000	99,997.66	−186.51	.20	− .23
2	1	99,811	0	....	99,811.15	−186.31	− .04	− .68
3	1	99,622	0	....	99,624.84	−186.35	− .72	− 1.12
4	1	99,434	0	....	99,438.49	−187.07	− 1.85	− 1.14
5	1	99,246	0	....	99,251.42	−188.92	− 2.99	− .25
6	1	99,058	0	....	99,062.50	−191.90	− 3.23	1.92
7	1	98,871	0	....	98,870.60	−195.14	− 1.31	5.32
8	1	98,684	0	....	98,675.46	−196.45	4.01	9.17
9	1	98,498	0	....	98,479.01	−192.44	13.18	11.78
10	1	98,311	0	....	98,286.57	−179.26	24.96	10.91
11	1	98,126	0	....	98,107.31	−154.30	35.86	4.91
12	1	97,940	0	....	97,953.01	−118.44	40.77	− 5.06
13	1	97,755	1	97,755	97,834.57	− 77.67	35.71	−11.04
14	1	97,744	0	....	97,756.90	− 41.96	24.68	−11.88
15	1	97,733	0	....	97,714.94	− 17.28	12.79	− 9.25
16	1	97,722	0	....	97,697.65	− 4.49	3.54	− 5.37
17	1	97,712	0	....	97,693.16	− .95	− 1.83	− 1.91
18	1	97,701	0	....	97,692.21	− 2.78	− 3.74	.35
19	1	97,690	0	....	97,689.44	− 6.52	− 3.39	1.37
20	1	97,679	0	....	97,682.92	− 9.91	− 2.02	1.49
21	1	97,668	0	....	97,673.01	− 11.93	− .53	1.13
22	1	97,657	0	....	97,661.09	− 12.45	.61	.61
23	1	97,647	0	....	97,648.64	− 11.85	1.22	.12
24	1	97,636	0	....	97,636.79	− 10.63	1.34	− .26
25	1	97,625	1	97,625	97,626.16	− 9.28	1.08	− .42
26	1	97,617	0	....	97,616.88	− 8.21	.66	− .40
27	1	97,610	0	....	97,608.67	− 7.55	.26	− .27
28	1	97,602	0	....	97,601.12	− 7.29	− .02	− .13
29	1	97,594	0	....	97,593.83	− 7.31	− .15	− .02
30	1	97,587	0	....	97,586.52	− 7.46	− .17	.05
31	1	97,579	0	....	97,579.06	− 7.63	− .12	.10
32	1	97,571	0	....	97,571.43	− 7.75	− .02	.12
33	1	97,564	0	....	97,563.68	− 7.77	.10	.12
34	1	97,556	0	....	97,555.91	− 7.68	.22	.10
35	1	97,548	0	....	97,548.23	− 7.46	.31	.04
36	1	97,541	0	....	97,540.77	− 7.15	.35	− .05
37	1	97,533	1	97,533	97,533.62	− 6.80	.30	− .10
38	1	97,527	0	....	97,526.82	− 6.50	.20	− .10
39	1	97,520	0	....	97,520.33	− 6.29	.10	− .08
40	1	97,514	0	....	97,514.03	− 6.19	.03	− .04
41	1	97,508	0	....	97,507.84	− 6.16	− .02	− .01
42	1	97,502	0	....	97,501.67	− 6.18	− .03	.00
43	1	97,495	0	....	97,495.49	− 6.21	− .03	.01
44	1	97,489	0	....	97,489.28	− 6.24	− .02	.01
45	1	97,483	0	....	97,483.04	− 6.26	− .01	.01
46	1	97,477	0	....	97,476.78	− 6.27	− .00	.00
47	1	97,470	0	....	97,470.52	− 6.27	− .00	....
48	1	97,464	0	....	97,464.25	− 6.27	....	....
49	1	97,458	1	97,458	97,457.98	....	....	....

The second group of graduations (Tables 3 and 4) uses rougher data, so higher values of m must be used. Tables 3A and 4A show poor graduations for $m = 1{,}000{,}000$. The graduated values oscillate. The standard values (expected mortality) decrease, then increase, with duration, so the graduated values are allowed to decrease, then increase. This is accomplished by making $l = 0.5$ and $m = 100{,}000{,}000$. The value $l = 0.5$ puts considerable emphasis on the standard values. The resulting graduations are quite smooth and look reasonable. These graduations use $r = 0$, because the graduated values should decrease, then increase, whereas an exponential function cannot do both. Comparing Tables 3B and 4B, we see that combining the data did not make much difference in this group of data, whereas grouping produced a significant difference between Tables 1B and 2B.

The third group (Tables 5A and 5B) shows the results of attempting to obtain monthly values of l_x by interpolation. If l_0 is included, the interpolation oscillates (in Table 5A), since the change from l_0 to l_1 is large in relation to the change from l_1 to l_2. This same behavior also occurs if osculatory interpolation is used. Table 5B uses the same parameters as Table 5A, but l_0 is omitted, so the resulting graduation is much better.

Table 5C puts very heavy weight on the standard values ($l = 0.9$). The resulting interpolation is reasonably good. The standard values (geometric means) comprise a good interpolation by themselves if the interpolation must pass through the pivotal points exactly.

III. COMMENTS ON THE COMPUTER ASPECTS

The interpolations in Tables 5A and 5C utilize 49×49 matrices and are relatively cheap (on the order of $1.50). Working with larger matrices is disproportionately more expensive. This indicates that smoothing and interpolation should be done in chunks and blended together. The size of the chunks depends on the programs and computer system used. Numerical instability is not observed in solving the matrix equation $AU = B$ for 49×49 matrices. The Choleski factorization method is used because it is far cheaper than the built-in matrix inverse function in the APL language for larger systems.

IV. FUTURE DIRECTIONS

There are a number of computer packages that minimize a quadratic form ($U^TAU - 2U^TB$) subject to one or more sets of inequality constraints.

Possible constraints are the following:

1. $u_i \geqslant 0$ for $i = 1, 2, \ldots, n$.
2. $u_i \geqslant c$ for $i = 1, 2, \ldots, n$.
3. $u_{i+1} \geqslant u_i$ for $i = 1, 2, \ldots, n - 1$.
4. $u_{i+1} \geqslant ku_i$ for $i = 1, 2, \ldots, n - 1; k > 1$.
5. $u_{i+1} \leqslant ku_i$ for $i = 1, 2, \ldots, n - 1; 0 < k \leqslant 1$.

By using combinations of these constraints, desirable features can be built into a graduation.

V. ACKNOWLEDGMENTS

I would like to thank the members of the Committee on Papers, who read this paper, for their valuable suggestions.

APPENDIX I

MATHEMATICAL PROOFS

We are given values (data) $u''_1, u''_2, \ldots, u''_n$. The data may be given for the entire grid of positive integers $1, 2, \ldots, n$ or for a proper subset of the grid.

Choose

n weights $w_1, w_2, \ldots, w_{n-1}, w_n$; $w_i \geqslant 0, i = 1, 2, \ldots, n$;
n weights $w'_1, w'_2, \ldots, w'_{n-1}, w'_n$; $w'_i > 0, i = 1, 2, \ldots, n$;
n standard values $s_1, s_2, \ldots, s_n$;
$l\ (0 \leqslant l \leqslant 1)$; $z = 1, 2, 3, \ldots$;
$m\ (m > 0)$; $r > -1$.

Find $u_1, u_2, \ldots, u_n$ such that

$$(1 - l)F + lF' + mS \text{ is minimized },$$

where

$$F = \sum_{i=1}^{n} w_i(u_i - u''_i)^2 ,$$

$$F' = \sum_{i=1}^{n} w'_i(u_i - s_i)^2 ,$$

$$S = \sum_{i=1}^{n-z} (\Delta^z u_i - r\Delta^{z-1} u_i)^2 .$$

Special Cases

Case 1.—If $z \geqslant 2$, then $S = 0$ if $u_i = P_{z-2}(i) + k(1 + r)^i$. If the u_i'''s are of this form, then they are left invariant.

Case 2.—If, in addition, $r = 0$, then $S = 0$ for $u_i = P_{z-1}(i)$, where P_{z-1} is a polynomial of degree $z - 1$ (or less).

Case 3.—If $z = 1$, then $S = 0$ for $u_i = k(1 + r)^i$. [Note: $\Delta^1 u_i - r\Delta^0 u_i = u_{i+1} - (1 + r)u_i$. This criterion was suggested by Camp [1] and others. If the u_i'''s are of this form, they are left invariant for any k.]

We now prove that

$$(1 - l)F + lF' + mS = U^T AU - 2U^T B + C ,$$

where

$$W = \begin{bmatrix} w_1 & & & 0 \\ & w_2 & & \\ & & \ddots & \\ 0 & & & w_n \end{bmatrix}, \quad W' = \begin{bmatrix} w_1' & & & 0 \\ & w_2' & & \\ & & \ddots & \\ 0 & & & w_n' \end{bmatrix};$$

This type of matrix can be referred to as $W = \text{diag}(w_1, w_2, \ldots, w_n)$.

$$U = (u_1, u_2, \ldots, u_n)^T , \quad U'' = (u_1'', u_2'', \ldots, u_n'')^T ;$$

$$S = (s_1, s_2, \ldots, s_n)^T ;$$

$$KU = \begin{bmatrix} \Delta^z u_1 - r\Delta^{z-1}u_1 \\ \Delta^z u_2 - r\Delta^{z-1}u_2 \\ \cdot \\ \cdot \\ \cdot \\ \Delta^z u_{n-z} - r\Delta^{z-1}u_{n-z} \end{bmatrix} \quad \text{(a column vector of length } n - z\text{)} ;$$

$$A = (1 - l)W + lW' + mK^T K ;$$

$$B = (1 - l)WU'' + lW'S ;$$

$$C = (1 - l)U''^T WU'' + lS^T W'S ;$$

C does not depend on U.

Note that if $X = (x_1, x_2, \ldots, x_n)^T$, a column vector, then

$$X^T X = \sum_{i=1}^{n} x_i^2 ,$$

and if W is as defined above, then

$$X^TWX = \sum_{i=1}^{n} w_i x_i^2 \; ;$$

letting $X = U - U''$, we have

$$F = \sum_{i=1}^{n} w_i(u_i - u_i'')^2 = (U - U'')^T W(U - U'') \, .$$

If $X = U - S$, we have

$$F' = \sum_{i=1}^{n} w_i'(u_i - s_i)^2 = (U - S)^T W'(U - S) \, .$$

If $X = KU$,

$$S = \sum_{i=1}^{n-z} (\Delta^z u_i - r\Delta^{z-1} u_i)^2 = (KU)^T(KU) = U^TK^TKU \, ,$$

so

$$\begin{aligned} (1 - l)F + lF' + mS = {} & (1 - l)(U - U'')^T W(U - U'') \\ & + l(U - S)^T W'(U - S) + mU^TK^TKU \, . \end{aligned} \tag{1}$$

Using the fact that $(U - U'')^T = U^T - U''^T$ and $W(U - U'') = WU - WU''$, we have

$$\begin{aligned} F = (U - U'')^T W(U - U'') &= (U - U'')^T WU - (U - U'')^T WU'' \\ &= (U^T - U''^T)WU - (U^T - U''^T)WU'' \\ &= U^TWU - U''^TWU - U^TWU'' + U''^TWU'' \\ &= U^TWU - 2U^TWU'' + U''^TWU'' \, . \end{aligned}$$

So

$$\begin{aligned} (1 - l)F + lF' + mS = {} & (1 - l)(U^TWU - 2U^TWU'' + U''^TWU'') \\ & + l(U^TW'U - 2U^TW'S + S^TW'S) \\ & + mU^TK^TKU \end{aligned}$$

$$\begin{aligned}
&= (1 - l)U^TWU + lU^TW'U + mU^TK^TKU \\
&\quad - 2(1 - l)U^TWU'' - 2lU^TW'S \\
&\quad + (1 - l)U''^TWU'' + lS^TW'S \\
&= U^T[(1 - l)W + lW' + mK^TK]U \\
&\quad - 2U^T[(1 - l)WU'' + lW'S] \\
&\quad + (1 - l)U''^TWU'' + lS^TW'S \\
&= U^TAU - 2U^TB + C\ .
\end{aligned}$$

A theorem from linear algebra tells us that if the matrix A is positive definite, then

$$U^TAU - 2U^T[(1 - l)WU'' + lW'S] = U^TAU - 2U^TB$$

takes its minimum value for the vector U that is defined by

$$AU = (1 - l)WU'' + lW'S$$

or

$$AU = B\ .$$

A reference for this theorem is Noble ([5], p. 400).

POSITIVE DEFINITENESS

In this section we will show that A is positive definite if any of the following is true:

1. The degree of differencing z is less than or equal to the number of pivotal points.
2. All the weights w_i ($i = 1, 2, \ldots, n$) are strictly positive ($l < 1$).
3. All the weights w'_i ($i = 1, 2, \ldots, n$) are strictly positive and $l > 0$.

Definition of Positive Definiteness

The symmetric matrix A is positive definite if $X^TAX > 0$ for all vectors $X \neq (0, 0, \ldots, 0)^T$. Note that $X^TAX = 0$ if $X = (0, 0, \ldots, 0)^T$.

Conditions under Which the Matrix A Is Positive Definite

Let X be any vector with n real elements $X = (x_1, x_2, \ldots, x_n)^T$. Then $X^TA = (1 - l)X^TW + lX^TW' + mX^TK^TK$, and

$$X^TAX = (1 - l)X^TWX + lX^TW'X + mX^TK^TKX , \quad m > 0$$

$$= (1 - l)\sum_{i=1}^{n} w_i x_i^2 + l \sum_{i=1}^{n} w_i' x_i^2 \tag{2}$$

$$+ m \sum_{i=1}^{n-z} (\Delta^z x_i - r\Delta^{z-1} x_i)^2 .$$

If $l > 0$ and $w_i' > 0$ for $i = 1, 2, \ldots, n - 1, n$, then $X^TAX > 0$ unless $X = 0$. This is clear from equation (2). These conditions can always be stipulated. Thus A is positive definite if these conditions hold.

If $l < 1$ and $w_i > 0$ for $i = 1, 2, \ldots, n - 1, n$, then $X^TAX > 0$ unless $X = 0$. This situation occurs in smoothing. Then A is positive definite.

However, in interpolation, some of the w_i's are 0. Suppose $l = 0$; then

$$X^TAX = \sum_{i=1}^{n} w_i x_i^2 + m \sum_{i=1}^{n-z} (\Delta^z x_i - r\Delta^{z-1} x_i)^2 . \tag{3}$$

Since both terms on the right-hand side of equation (3) are nonnegative, $X^TAX = 0$ if and only if both terms are zero. This can happen only if the u_i's have a functional form where $S = 0$ *and* this function is zero at each of the pivotal points. That is, the function must have as many distinct zeros as there are pivotal points. There are three different functions to look at.

Case 1

$$x_i = P_{z-2}(i) + k(1 + r)^i , \quad z = 2, 3, \ldots ; \quad r \neq 0 .$$

Case 2

$$x_i = P_{z-1}(i) , \quad z = 2, 3, \ldots ; \quad r = 0 .$$

Case 3

$$x_i = k(1 + r)^i , \quad z = 1 .$$

If case 2 holds, then $x_i = P_{z-1}(i)$ can have, at most, $z - 1$ positive distinct roots, since it is a polynomial of degree $z - 1$.

If case 3 holds, then $x_i = k(1 + r)^i$ has no real roots, since $1 + r > 0$.

We will show that, if case 1 holds, $x_i = P_{z-2}(i) + k(1 + r)^i$ has at most $z - 1$ distinct positive roots. Use induction on z.

Let $z = 1$ (case 3); then $x_i = k(1 + r)^i$ has no real roots.

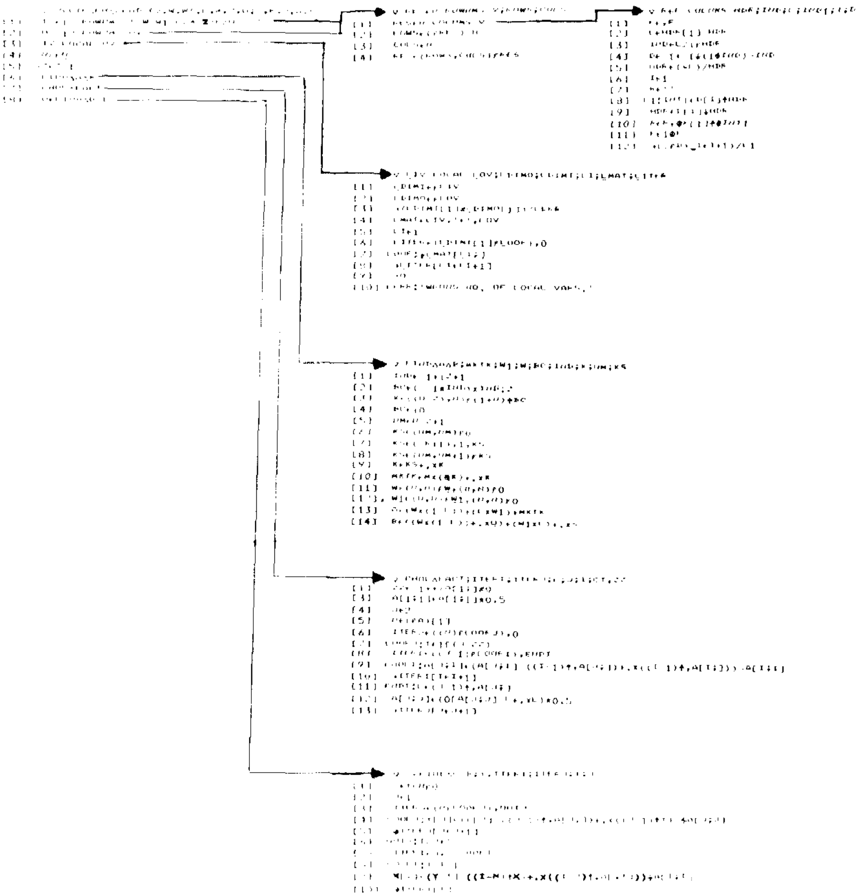
Let $z = 2$ (case 1); then $x_i = P_0(i) + k(1 + r)^i$ has at most one positive root, since $P_0(i)$ is a constant. This is clear from geometry, since $k(1 + r)^i$ is either always increasing or always decreasing ($r < 0$), so it can attain the constant value, $-P_0(i)$, only once.

Assume that $P_{z-2}(i) + k(1 + r)^i$ has at most $z - 1$ positive distinct roots. We must show that $P_{z-1}(i) + k(1 + r)^i$ has at most z positive distinct roots. Assume the contrary, that is, that $P_{z-1}(i) + k(1 + r)^i$ has $z + 1$ positive distinct roots. The derivative of $P_{z-1}(i) + k(1 + r)^i$ is $P_{z-2}(i) + k \ln(1 + r)(1 + r)^i$. Using Rolle's theorem z times, it can be shown that $P_{z-2}(i) + k \ln(1 + r)(1 + r)^i$ has z positive distinct roots, which contradicts the inductive assumption.

So $X^T AX > 0$ ($X \neq 0$) as long as the number of pivotal points (positive weights) is greater than the maximum possible number of positive distinct roots ($z - 1$). Thus A is positive definite.

APPENDIX II

APL PROGRAMS



PROGRAMMING NOTES

DIFΔEQΔGRAD

The first three lines set up the following correspondences:

$$\underline{W} \leftarrow O\underline{W}$$
$$\underline{W}1 \leftarrow O\underline{W}1$$
$$L \leftarrow OL$$
$$M \leftarrow OM$$

$$Z \leftarrow OZ$$
$$R \leftarrow OR$$
$$\underline{U} \leftarrow O\underline{U}$$
$$S \leftarrow OS$$

The variables on the left are local variables, and those on the right are global. This is a device for allowing an APL function to have more than two arguments. The sixth line calls the function *FINDΔAΔB* that calculates the matrix A and the vector B. The seventh line calls the function *CHOLΔFACT* that factors the matrix A into LL^T and stores L in the lower triangle of A. The function *CHOLΔFACT* assumes that A is a band matrix and calculates the band width of A. The eighth line calls the function *FINDΔU* that calculates the graduated values, U, such that $AU = B$.

Sample Input

```
OS
/OW/OW1/OL/OM/OZ/OR/OU/OS/

OW
1 0 0 0 0 0 0 0 0 0 0 0 1 0 0 0 0 0 0 0 0 0 0 0 1 0 0 0 0 0 0 0 0 0 0 0 1
OW1
1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
OL
0
OM
10
OZ
4
OR
¯0.0029
OU
97755 0 0 0 0 0 0 0 0 0 0 0 97625 0 0 0 0 0 0 0 0 0 0 0 97533 0 0 0 0 0 0 0 0 0 0 0 97458
OS
97755 97744.16006 97733.32132 97722.48378 97711.64744 97700.81231 97689.97838 97679.14564 97668.31411 97657.48378
      97646.65465 97635.82673 97625 97617.33002 97609.66064 97601.99197 97594.3237 97586.65613 97578.98916 97571.32279
      97563.65703 97555.99187 97548.32731 97540.66335 97533 97526.7478 97520.49599 97514.24459 97507.99359 97501.74299
      97495.49279 97489.24299 97482.99359 97476.74459 97470.49599 97464.2478 97458
```

NOTES ON DATA

The following tabulation shows correspondences between APL global variables used for inputting data and mathematical variables.

Mathematical Variable	APL Variable		
	Name	Shape	Size
Weights (W)	$O\underline{W}$	Vector	n
Standard weights (W')	$O\underline{W}1$	Vector	n
Emphasis on standard (l)	OL	Scalar	
Emphasis on smoothness (m)	OM	Scalar	
Degree (z)	OZ	Scalar	
Exponential constant (r)	OR	Scalar	
Data (U'')	$O\underline{U}$	Vector	n
Standard (S)	OS	Vector	n

REFERENCES

1. Camp, K. "New Possibilities in Graduation," *TSA*, VII (1955), 6.
2. Greville, T. N. E. Part 5 Study Notes—Graduation. Society of Actuaries, 1974.
3. Kimeldorf, G. S., and Jones, D. A. "Bayesian Graduation," *TSA*, XIX (1967), 66.
4. "Mortality under Standard Ordinary Insurance Issues between 1977 and 1978 Anniversaries," *TSA, 1979 Reports*. 1980.
5. Noble, B. *Applied Linear Algebra*, Englewood Cliffs, N.J.: Prentice-Hall, 1969.

DISCUSSION OF PRECEDING PAPER

FRANK E. KNORR:

I would like to congratulate Mr. Lowrie on the fine work he did demonstrating the versatility of the Whittaker-Henderson method of graduation. When I was studying for Part 5, this was never mentioned as an interpolation method. At that time it was easy to see that when its smoothness factor is increased indefinitely, the graduated values resemble the points of a polynomial. That is, the normal Whittaker-Henderson method causes the graduated values to approach the smooth points of a polynomial. Mr. Lowrie has shown how this method can be extended in three ways:

1. By having the graduated values approach the smooth points of a standard table.
2. By having the graduated values approach the smooth points of an exponential curve (as well as a curve which can be expressed as the sum of an exponential function and a polynomial).
3. By using this method to interpolate between pivotal points, that is, using zero for weights at nonpivotal points.

My discussion addresses each of these three extensions.

As a student of APL, I would also like to express my appreciation for the programs included in Appendix II. However, I would have taken the easy (yet expensive) way out by condensing the last two programs into one line:

$$U \leftarrow B \boxdiv A \ .$$

Using Standard Values

It should be pointed out that when we are dealing with standardized values, one nice property of the Whittaker-Henderson graduation is lost. That property can be expressed in equation form as

$$\Sigma W_i u_i = \Sigma W_i u_i'' \ ; \quad \Sigma i W_i u_i = \Sigma i W_i u_i'' \ .$$

This preserves the total number of deaths and the average age at death if crude mortality rates are graduated and the exposures are used as the weights ([2], p. 54).

For example, the raw data used in Tables 1A, 1B, and 1C of the paper had a total of 3,258 deaths and an average duration at death of 1.98 years. When the graduated mortality rates from Table 1A are applied to the exposures, the total number of deaths is still 3,258 and the average duration at death also remains 1.98. The same holds true for the graduated rates from Table 1B, since neither of these involve the standard values (i.e., l is set equal to zero). However, once the standard values are involved, this property is lost, as in Table 1C where the graduated rates applied to the exposures yield a total of 3,635 deaths and an average duration at death of 2.15 years. If one thinks of the standard mortality rates as a smooth goal that can be approached by increasing l and decreasing the smoothness factor m, then one would expect that the total deaths and the average duration also approach those of the standard data, which are 5,145 and 2.81, respectively.

In Tables 2A, 2B, and 2C of the paper the data were grouped to central durations. (This grouping can be thought of as another form of graduation of the same data as in Tables 1A, 1B, and 1C.) The data were grouped in a way that preserved the number of deaths but increased the average duration to 2.06 years. When the graduated rates are applied to the weights used in the three graduations the resulting number of deaths and average durations are the following:

	Table 2A	Table 2B	Table 2C
Total number of deaths	3,258	3,258	3,636
Average duration at death (years)..............	2.06	2.06	2.23

Using Exponential Smoothness

During my work with the Committee to Recommend New Disability Tables for Valuation, we were faced with the task of graduating data whose underlying pattern was believed to be more in the form of an exponential curve than of a polynomial. Our approach was slightly different from Mr. Lowrie's and can be expressed as a three-step process:

1. Let $v''_i = \log_{10} u''_i$, for all i.
2. Graduate the v'''s yielding v's.
3. Let the graduated values $u_i = 10^{v_i}$, for all i.

Under this method, increasing the smoothness factor would result in graduated values that would approach points on a smooth curve that can be expressed as

$$10^{P_{z-1}(i)} ,$$

where $P_{z-1}(i)$ is a polynomial, while the method demonstrated in this paper would result in values that would approach

$$P_{z-2}(i) + k(1 + r)^i .$$

The 1969–71 United States Male l_x values were graduated using our three-step method, and the results are presented here as Table 5D. A comparison with the graduated values of Table 5A is also shown.

It should also be pointed out here that the relationships

$$\Sigma W_i u_i = \Sigma W_i u_i'' \quad \text{and} \quad \Sigma i W_i u_i = \Sigma i W_i u_i''$$

still hold for Mr. Lowrie's method, but not for our three-step method.

However, the relationships

$$\Sigma W_i \log u_i = \Sigma W_i \log u_i'' \quad \text{and} \quad \Sigma i W_i \log u_i = \Sigma i W_i \log u_i''$$

hold instead. From a practical point of view this property is probably useful only if survivorship rates are being graduated with weights all equal to one:

$$\begin{aligned}
\log {}_t p_x'' &= \log p_x'' p_{x+1}'' p_{x+2}'' \cdots p_{x+t-1}'' \\
&= \Sigma \log p_{x+i-1}'' \\
&= \Sigma W_i \log u_i'' \\
&= \Sigma W_i \log u_i \\
&= \Sigma \log p_{x+i-1} \\
&= \log p_x p_{x+1} p_{x+2} \cdots p_{x+t-1} \\
&= \log {}_t p_x \,.
\end{aligned}$$

Therefore,

$${}_t p_x = {}_t p_x'' \,;$$

the probability of survivorship from the beginning to the end of the table would be preserved.

Using Whittaker-Henderson to Interpolate

The work of the Committee to Recommend New Disability Tables for Valuation also involved using the Whittaker-Henderson graduation method to interpolate termination rates between durations. However, I am not aware of any mention of this property before Mr. Lowrie's paper. His

TABLE 5D

MONTHLY VALUES OF l_x, OBTAINED BY INTERPOLATION USING 1969–71 UNITED STATES MALE VALUES: THREE-STEP PROCESS, $z = 3, r = 0, l = 0, m = 10$

t	Graduated Values from Table 5A	Graduated Values Using 3-Step Process	Difference between Graduated Values	Differences: First	Second	Third
1	99,999.83	99,999.82	−.01	−304.92	22.48	−.20
2	99,696.41	99,694.90	−1.51	−282.44	22.27	−.22
3	99,414.88	99,412.45	−2.42	−260.17	22.05	−.26
4	99,155.16	99,152.29	−2.88	−238.12	21.79	−.32
5	98,917.14	98,914.17	−2.98	−216.33	21.47	−.39
6	98,700.66	98,697.84	−2.82	−194.86	21.08	−.48
7	98,505.47	98,502.98	−2.50	−173.78	20.60	−.59
8	98,331.26	98,329.20	−2.07	−153.18	20.01	−.71
9	98,177.61	98,176.02	−1.59	−133.16	19.30	−.86
10	98,043.98	98,042.86	−1.12	−113.86	18.44	−1.02
11	97,929.68	97,929.00	−.68	−95.42	17.43	−1.20
12	97,833.88	97,833.58	−.30	−77.99	16.23	−1.40
13	97,755.58	97,755.59	.01	−61.76	14.83	−1.55
14	97,693.59	97,693.83	.24	−46.93	13.28	−1.67
15	97,646.51	97,646.90	.39	−33.65	11.60	−1.75
16	97,612.78	97,613.26	.48	−22.04	9.85	−1.79
17	97,590.71	97,591.21	.50	−12.19	8.06	−1.79
18	97,578.54	97,579.02	.48	−4.13	6.27	−1.75
19	97,574.46	97,574.88	.42	2.14	4.52	−1.67
20	97,576.68	97,577.02	.35	6.66	2.85	−1.54
21	97,583.42	97,583.68	.26	9.51	1.31	−1.38
22	97,593.01	97,593.19	.18	10.82	−.07	−1.17
23	97,603.91	97,604.01	.11	10.75	−1.24	−.93
24	97,614.72	97,614.77	.04	9.51	−2.17	−.64
25	97,624.28	97,624.27	.00	7.34	−2.80	−.38
26	97,631.65	97,631.61	−.04	4.54	−3.18	−.15
27	97,636.21	97,636.15	−.06	1.36	−3.33	.05
28	97,637.57	97,637.51	−.07	−1.97	−3.28	.22
29	97,635.60	97,635.54	−.07	−5.25	−3.06	.35
30	97,630.35	97,630.29	−.06	−8.30	−2.70	.46
31	97,622.03	97,621.99	−.05	−11.01	−2.25	.53
32	97,611.01	97,610.98	−.03	−13.25	−1.71	.58
33	97,597.74	97,597.72	−.02	−14.97	−1.14	.59
34	97,582.76	97,582.76	−.01	−16.10	−.55	.57
35	97,566.65	97,566.65	.00	−16.65	.02	.52
36	97,550.00	97,550.00	.00	−16.63	.54	.44
37	97,533.38	97,533.37	.00	−16.09	.98	.37
38	97,517.30	97,517.29	−.01	−15.10	1.35	.30
39	97,502.22	97,502.18	−.03	−13.76	1.65	.24
40	97,488.49	97,488.43	−.06	−12.11	1.89	.19
41	97,476.41	97,476.32	−.09	−10.22	2.07	.14
42	97,466.21	97,466.09	−.12	−8.15	2.21	.10
43	97,458.09	97,457.94	−.15	−5.94	2.31	.07
44	97,452.17	97,452.00	−.18	−3.63	2.38	.04
45	97,448.56	97,448.37	−.19	−1.25	2.42	.02
46	97,447.30	97,447.12	−.18	1.17	2.44	.01
47	97,448.44	97,448.28	−.15	3.60	2.44	.00
48	97,451.98	97,451.89	−.10	6.05	.00	.00
49	97,457.93	97,457.93	.00	.00	.00	.00

work in demonstrating this property is commendable in itself. When combined with the other extensions presented here, it is truly a noteworthy effort.

MARK EVANS:

Professor Lowrie's paper appears to advance graduation technology. Putting the Lowrie modification of the Whittaker-Henderson method into practice, I encountered two problems which others are likely to face.

Often when graduating data, one desires that the sum of the products of the weights and the graduated values be very close to the sum of the products of the weights and the data values. Under the Lowrie modification this may not be the case when l is not zero. Fortunately, one may overcome this problem by simply multiplying the standard values by a constant so that when applied to the weights they reproduce the sum of the weighted data values. This is expressed by the following formula:

$$S_i' = S_i x \frac{\sum_{i=1}^{n} U_i'' W_i}{\sum_{i=1}^{n} S_i W_i} .$$

Note that the standard weights are not used in this calculation. Appendix I contains programs that have been altered slightly. Line 5 in both *DIFΔEQΔGRAD* and *GRADSS* will perform the above calculation upon removal of the comment symbol. The effect of this adjustment can be seen in comparing Reports 1 and 2 in Appendix II. The sum of the products of the graduated values is 158 percent of the corresponding sum for the data values in Report 1, which duplicates Professor Lowrie's original Table 1C. The corresponding number for Report 2 is 100 percent. Report 2 is identical with Report 1 except that the standard values have been adjusted.

The practitioner may feel ill equipped to make a proper choice for r. One possible approach is to develop an iterative technique which finds the r which produces the smallest value for mS. Program *GRADSS* in Appendix I finds such an r using a successive bisection approach. The results of this approach are in Report 3 of Appendix II. A printout of the iterative process is also contained in Appendix II. Note that this approach values the number of third differences in excess of 0.005 from 7 to 1. Refinements in the search for r may be an addition to the list of future directions.

APPENDIX I

The APL programs in this appendix are merely Professor Lowrie's with minor modifications and enhancements. *DIFΔEQΔGRAD* has been altered to take advantage of input algorithms for which some systems are particularly well suited. The program *REPORT* is not shown, since it utilizes capabilities that are not common to most APL systems. It merely formats the output and performs no calculations.

```
      141                                                                    INPUT
       ∇ U←DIFΔEQΔGRAD CONTROL;W̲;W̲1;L;M;Z;U̲;S;R;N;A;B;I;ITERI
[1]    I←0
[2]   LOOPI:⍎CONTROL[I←I+1;]
[3]    →(I<8)/LOOPI
[4]    N←⍴U̲
[5]   ⍝S←S×(+/U̲×W̲)÷+/S×W̲
[6]    Z←Z-1
[7]    FINDΔAΔB
[8]    CHOLΔFACT
[9]    U←FINDΔU B
[10]   PRT←REPORT⍉(8,N)⍴W̲1,S,W̲,U̲,U,(N↑(1 FINDΔPOLY 1)+.×U),(N↑(2 FINDΔPOLY 2)
       +.×U),N↑(3 FINDΔPOLY 3)+.×U
[11]   PRT←PRT,[1] 133↑'0 Z=',(⍕Z+1),'  R=',(⍕R),'  L=',(⍕L),'  M=',⍕M
       ∇
```

```
      157                                                                    INPUT
       ∇ FINDΔAΔB;MKTK;W1;W;BC;IND;K;NM;KS
[1]    K←Z FINDΔPOLY Z
[2]    BC←⍳0
[3]    NM←N-Z+1
[4]    KS←(NM,NM)⍴0
[5]    KS←(-R+1),1,KS
[6]    KS←(NM,NM+1)⍴KS
[7]    K←KS+.×K
[8]    MKTK←M×(⍉K)+.×K
[9]    W←(N,N)⍴W̲,(N,N)⍴0
[10]   W1←(N,N)⍴W̲1,(N,N)⍴0
[11]   A←(W×1-L)+(L×W1)+MKTK
[12]   B←((W×1-L)+.×U̲)+(W1×L)+.×S
       ∇
```

```
      172                                                                    INPUT
       ∇ CHOLΔFACT;ITERI;ITERJ;L;J;I;CT;ZZ
[1]    ZZ←¯1++/A[1;]≠0
[2]    A[1;1]←A[1;1]*÷2
[3]    J←2
[4]    N←(⍴A)[1]
[5]    ITERJ←(N⍴LOOPJ),0
[6]   LOOPJ:I←1⌈J-ZZ
[7]    ITERI←(1↓J⍴LOOPI),ENDI
[8]   LOOPI:A[J;I]←(A[J;I]-((I-1)↑,A[J;])+.×(I-1)↑,A[I;])÷A[I;I]
[9]    →ITERI[I←I+1]
[10]  ENDI:L←(J-1)↑,A[J;]
[11]   A[J;J]←(0⌈A[J;J]-L+.×L)*÷2
[12]   →ITERJ[J←J+1]
       ∇
```

```
      188                                                                    INPUT
       ∇ X←FIND∆U B;Y;ITERI;ITERJ;I;J
[1]    X←Y←Nρ0
[2]    J←1
[3]    ITERJ←(NρLOOPJ),OUTJ
[4]   LOOPJ:Y[J]←(B[J]-((J-1)↑,A[J;])+.×(J-1)↑Y)÷A[J;J]
[5]    →ITERJ[J←J+1]
[6]   OUTJ:I←N+1
[7]    ITERI←0,NρLOOPI
[8]   LOOPI:I←I-1
[9]    X[I]←(Y[I]-((I-N)↑X)+.×(I-N)↑,A[;I])÷A[I;I]
[10]   →ITERI[1]
       ∇

      201                                                                    INPUT
       ∇ K←Z FIND∆POLY Y;IND
[1]    K←((N-Z),N)ρ(1+N)↑((1-2×2|Y)×¯1*IND)×(IND←0,⍳Y)!Y
       ∇

LINE 205        COLUMN 1                                                     INPUT
       ∇ U←GRADSS CONTROL;W;W1;L;M;Z;U;S;R;N;A;B;I;ITERI;SS;O;O3
[1]    I←0
[2]   LOOPI:⍎CONTROL[I←I+1;]
[3]    →(I<8)/LOOPI
[4]    N←ρU
[5]   ⍝S←S×(+/U×W)÷+/S×W
[6]    Z←Z-1
[7]    R←1.024×⌽2-⍳3
[8]    SS←(FIND∆SS R[1]),(FIND∆SS R[2]),FIND∆SS R[3]
[9]    PRTR← 2 70 ↑(2 7 ρ'R:      1000SS:'), 0 3 ⍕ 2 3 ρR,SS×1000
[10] LOOPR:⎕← ¯2 70 ↑PRTR
[11]   →(2=1↑O←⍋SS)/MID
[12]   R←O3⌽(R[1+O3]+-/O3⌽R),R←2↑(O3←3=1↑O)↓R
[13]   SS←O3⌽(FIND∆SS R[O[1]]),SS[1 2 +O3]
[14]   →RET
[15] MID:R←(R,0.5×+/R←2↑(O3←3=O[2])↓R)[1 3 2]
[16]   SS←((2↑O3↓SS),FIND∆SS R[2])[1 3 2]
[17] RET:PRTR←PRTR,[1] 2 70 ↑(2 7 ρ'R:      1000SS:'), 0 3 ⍕ 2 3 ρR,SS×1000
[18]   →(0.001<-/R[2 1])/LOOPR
[19]   R←R[(⍋SS)[1]]
[20]   FIND∆A∆B
[21]   CHOL∆FACT

LINE 227        COLUMN 1                                                     INPUT
[22]   U←FIND∆U B
[23]   PRT←REPORT⍉(8,N)ρW1,S,W,U,U,(N↑(1 FIND∆POLY 1)+.×U),(N↑(2 FIND∆POLY 2)
       +.×U),N↑(3 FIND∆POLY 3)+.×U
[24]   PRT←PRT,[1] 133↑'0 Z=',(⍕Z+1),'  R=',(⍕R),'  L=',(⍕L),'  M=',⍕M
       ∇

      254                                                                    INPUT
       ∇ SS←FIND∆SS R
[1]    FIND∆A∆B
[2]    CHOL∆FACT
[3]    U←FIND∆U B
[4]    SS←+/((((Z+1) FIND∆POLY Z+1)+.×U)-R×((Z+1) FIND∆POLY Z)+.×U)*2
       ∇

      260                                                                    INPUT
       CON1A
W←T1AW
W1←T1AW
L←0
M←1000
Z←3
R←0
U←T1AU
S←T1AS
       CON1C
W←T1AW
W1←T1AW
L←.2
M←1000000
Z←3
R←.05
U←T1AU
S←T1AS
```

APPENDIX II

This appendix contains computer output from the APL programs in Appendix I. Reports 1, 2, and 3 are described in the main part of the discussion. Report 4 shows the combined effect of using adjusted standard values and a recalculated r. Following Report 4 is a printout of the iterative process.

REPORT 1

LOWRIE MODIFICATION OF WHITTAKER-HENDERSON
ORIGINAL TABLE 1C

i	Standard Weights	Standard Values	Weights	Data Values	Graduated Values	First Difference	Second Difference	Third Difference
1	2,115,646	.50008	2,115,646	.25288	.30640	.03122	−.00787	.00782
2	1,457,640	.54540	1,457,640	.30940	.33762	.02335	−.00005	−.00209
3	1,073,275	.52456	1,073,275	.28511	.36097	.02330	−.00214	−.00154
4	882,728	.53471	882,728	.34325	.38428	.02117	−.00367	.01178
5	719,869	.55010	719,869	.42369	.40545	.01750	.00810	.00804
6	570,331	.58036	570,331	.36821	.42294	.02560	.01615	−.00606
7	472,780	.60916	472,780	.34900	.44854	.04174	.01009	−.00917
8	402,142	.63908	402,142	.49236	.49029	.05183	.00092	−.01333
9	322,480	.66361	322,480	.41533	.54212	.05275	−.01241	.00514
10	276,229	.68059	276,229	.78196	.59487	.04034	−.00727	.00239
11	232,221	.71484	232,221	.50814	.63521	.03307	−.00488	−.00261
12	186,412	.77785	186,412	.57400	.66828	.02819	−.00749	.00000
13	162,016	.85177	162,016	.73450	.69647	.02070	.00000	.00000
14	142,175	.94250	142,175	.64006	.71716	.00000	.00000	.00000

NOTE.—$z = 3$, $r = 0.05$, $l = 0.2$, $m = 1{,}000{,}000$.

REPORT 2

LOWRIE MODIFICATION OF WHITTAKER-HENDERSON
TABLE 1C WITH ADJUSTED STANDARD VALUES

i	Standard Weights	Standard Values	Weights	Data Values	Graduated Values	First Difference	Second Difference	Third Difference
1	2,115,646	.31666	2,115,646	.25288	.26929	.02971	−.00696	.00701
2	1,457,640	.34536	1,457,640	.30940	.29900	.02275	.00005	−.00270
3	1,073,275	.33216	1,073,275	.28511	.32175	.02281	−.00265	−.00165
4	882,728	.33859	882,728	.34325	.34456	.02016	−.00429	.01206
5	719,869	.34834	719,869	.42369	.36472	.01587	.00777	.00835
6	570,331	.36750	570,331	.36821	.38059	.02364	.01612	−.00599
7	472,780	.38574	472,780	.34900	.40422	.03975	.01013	−.00942
8	402,142	.40468	402,142	.49236	.44397	.04988	.00071	−.01378
9	322,480	.42021	322,480	.41533	.49385	.05059	−.01307	.00474
10	276,229	.43097	276,229	.78196	.54444	.03752	−.00833	.00217
11	232,221	.45265	232,221	.50814	.58196	.02919	−.00616	−.00271
12	186,412	.49255	186,412	.57400	.61115	.02303	−.00887	.00000
13	162,016	.53936	162,016	.73450	.63418	.01416	.00000	.00000
14	142,175	.59681	142,175	.64006	.64834	.00000	.00000	.00000

NOTE.—$z = 3$, $r = 0.05$, $l = 0.2$, $m = 1{,}000{,}000$.

REPORT 3

Lowrie Modification of Whittaker-Henderson Table 1C with Calculated r

i	Standard Weights	Standard Values	Weights	Data Values	Graduated Values	First Difference	Second Difference	Third Difference
1......	2,115,646	.50008	2,115,646	.25288	.30863	.02544	−.00188	.00213
2......	1,457,640	.54540	1,457,640	.30940	.33406	.02355	.00025	−.00001
3......	1,073,275	.52456	1,073,275	.28511	.35762	.02380	.00024	.00052
4......	882,728	.53471	882,728	.34325	.38142	.02404	.00076	.00416
5......	719,869	.55010	719,869	.42369	.40546	.02480	.00492	.00244
6......	570,331	.58036	570,331	.36821	.43025	.02971	.00736	−.00202
7......	472,780	.60916	472,780	.34900	.45997	.03707	.00534	−.00287
8......	402,142	.62908	402,142	.49236	.49704	.04241	.00247	−.00476
9......	322,480	.66361	322,480	.41533	.53946	.04488	−.00229	−.00031
10.....	276,229	.68059	276,229	.78196	.58434	.04259	−.00260	−.00400
11.....	232,221	.71484	232,221	.50814	.62693	.03999	−.00659	−.01252
12.....	186,412	.77785	186,412	.57400	.66692	.03340	−.01912	.00000
13.....	162,016	.85177	162,016	.73450	.70032	.01428	.00000	.00000
14.....	142,175	.94250	142,175	.64006	.71460	.00000	.00000	.00000

NOTE.—$z = 3$, $r = 1.596$, $l = 0.2$, $m = 1,000,000$.

REPORT 4

Lowrie Modification of Whittaker-Henderson Table 1C with Adjustment and Calculated r

i	Standard Weights	Standard Values	Weights	Data Values	Graduated Values	First Difference	Second Difference	Third Difference
1......	2,115,646	.31666	2,115,646	.25288	.27139	.02398	−.00124	.00165
2......	1,457,640	.34536	1,457,640	.30940	.29537	.02274	.00041	−.00004
3......	1,073,275	.33216	1,073,275	.28511	.31811	.02315	.00037	.00043
4......	882,728	.33859	882,728	.34325	.34126	.02351	.00080	.00335
5......	719,869	.34834	719,869	.42369	.36477	.02431	.00415	.00194
6......	570,331	.36750	570,331	.36821	.38908	.02846	.00609	−.00162
7......	472,780	.38574	472,780	.34900	.41754	.03455	.00447	−.00227
8......	402,142	.40468	402,142	.49236	.45209	.03903	.00221	−.00394
9......	322,480	.42021	322,480	.41533	.49112	.04123	−.00173	−.00069
10.....	276,229	.43097	276,229	.78196	.53236	.03950	−.00242	−.00494
11.....	232,221	.45265	232,221	.50814	.57186	.03708	−.00736	−.01640
12.....	186,412	.49255	186,412	.57400	.60895	.02973	−.02376	.00000
13.....	162,016	.53936	162,016	.73450	.63868	.00597	.00000	.00000
14.....	142,175	.59681	142,175	.64006	.64465	.00000	.00000	.00000

NOTE.—$z = 3$, $r = 1.974$, $l = 0.2$, $m = 1,000,000$.

ITERATIONS FOR RECALCULATED *R*

R:	-1.024	.000	1.024
1000SS:	1.213	.658	.491
R:	.000	1.024	2.048
1000SS:	.658	.491	.482
R:	1.024	2.048	3.072
1000SS:	.491	.482	.524
R:	1.024	1.536	2.048
1000SS:	.491	.475	.482
R:	1.536	1.792	2.048
1000SS:	.475	.476	.482
R:	1.280	1.536	1.792
1000SS:	.480	.475	.476
R:	1.536	1.664	1.792
1000SS:	.475	.475	.476
R:	1.408	1.536	1.664
1000SS:	.476	.475	.475
R:	1.536	1.600	1.664
1000SS:	.475	.475	.475
R:	1.536	1.568	1.600
1000SS:	.475	.475	.475
R:	1.568	1.600	1.632
1000SS:	.475	.475	.475
R:	1.568	1.584	1.600
1000SS:	.475	.475	.475
R:	1.584	1.600	1.616
1000SS:	.475	.475	.475
R:	1.584	1.592	1.600
1000SS:	.475	.475	.475
R:	1.592	1.600	1.608
1000SS:	.475	.475	.475
R:	1.592	1.596	1.600
1000SS:	.475	.475	.475
R:	1.596	1.598	1.600
1000SS:	.475	.475	.475
R:	1.594	1.596	1.598
1000SS:	.475	.475	.475
R:	1.596	1.597	1.598
1000SS:	.475	.475	.475

ITERATIONS FOR RECALCULATED *R* AND STANDARD VALUE ADJUSTMENT

R:	-1.024	.000	1.024
1000SS:	1.221	.672	.484
R:	.000	1.024	2.048
1000SS:	.672	.484	.449
R:	1.024	2.048	3.072
1000SS:	.484	.449	.468
R:	2.048	2.560	3.072
1000SS:	.449	.456	.468
R:	1.536	2.048	2.560
1000SS:	.455	.449	.456
R:	1.536	1.792	2.048
1000SS:	.455	.450	.449
R:	1.792	2.048	2.304
1000SS:	.450	.449	.452
R:	1.792	1.920	2.048
1000SS:	.450	.449	.449
R:	1.920	1.984	2.048
1000SS:	.449	.449	.449
R:	1.920	1.952	1.984
1000SS:	.449	.449	.449
R:	1.952	1.984	2.016
1000SS:	.449	.449	.449
R:	1.952	1.968	1.984
1000SS:	.449	.449	.449
R:	1.968	1.976	1.984
1000SS:	.449	.449	.449
R:	1.968	1.972	1.976
1000SS:	.449	.449	.449
R:	1.972	1.976	1.980
1000SS:	.449	.449	.449
R:	1.972	1.974	1.976
1000SS:	.449	.449	.449
R:	1.974	1.975	1.976
1000SS:	.449	.449	.449

CHARLES E. CHITTENDEN:

Mr. Lowrie's paper provides an opportunity to reflect on the role of graduation in practical actuarial work. I believe the "graduationists" have developed methods far more elaborate than most actuarial applications require. The work of graduation is to substitute appearances for facts. This alchemy in reverse should be effected only when the situation warrants.

Mr. Lowrie is right that experimental results can be too rough for practical use. The whole purpose of graduation, as I see it, is to make results usable. However, some graduationists have gone too far in justifying their craft. For example, Morton D. Miller's *Elements of Graduation* says (p. 6):

> We have reason to believe that most of the laws of nature do not exhibit irregular variations or sharp breaks, but may be expressed in terms of regular and continuous functions. This conclusion . . . is adequately supported by experiment. Nevertheless, any set of observations of a series of measurements corresponding to a physical law will exhibit irregularities; some will be positive, some negative.

How can this be? If experimental results always turn out jagged, then the empirical evidence clearly supports the view that reality itself is jagged. How could the conclusion that reality is smooth be supported by experiments the results of which are jagged? Graduation does not necessarily bring us closer to Truth; it merely gives us results we can live with. We do not want to explain why premium rates at age 53 are less than those at age 52, and graduation gets us off that hook. But if that is all graduation accomplishes, can't it be done with spline and weights rather than with 49×49 matrices and APL routines?

Do the same actuaries who take blind stabs at their long-range interest assumptions really need to have their mortality tables graduated by an extension of the Whittaker-Henderson formula which uses 49×49 matrices? If I used a 49×49 matrix and was 99.9 percent confident that each entry was correct, the probability of all my entries being correct would be less than 10 percent. Besides, I do not think I am 99.9 percent confident of anything.

Perhaps Mr. Lowrie's method is easier to use than I imagine. If so, it may be fine for what I would call "formal graduations," such as might be used in the construction of CSO tables. A fancy graduation may be justified if the data are ample enough to be credible and the results of using the table are visible enough to the public to require smoothness. But formal graduations are rarely appropriate in routine actuarial projects, and we should not delude ourselves that mathematical sophistication will improve accuracy. Papers on graduation can have the unfortunate effect of encouraging actuaries to overuse fancy techniques. My reaction to an elaborate graduation of the results of some small study would be definite but not positive.

In everyday work, data are often scanty and the need for smoothness less than compelling, so that graduation, if it is performed at all, should be simple, quick, and inexpensive. And since graduation is merely cosmetic, no justifications are required.

ELIAS S. W. SHIU:

I offer the following technical comments and references with respect to the Whittaker-Henderson graduation technique extended in Mr. Lowrie's paper.

1. *Fit*

Let $w + w' \neq 0$; then the expression

$$w(u - v)^2 + w'(u - s)^2 - (w + w')\left(u - \frac{wv + w's}{w + w'}\right)^2$$

is independent of u. Thus if we put

$$\begin{aligned} u_i^* &= 0 , && \text{if } (1 - l)w_i + lw_i' = 0 \\ &= \frac{(1 - l)w_i u_i'' + lw_i' s_i}{(1 - l)w_i + lw_i'} , && \text{otherwise} , \end{aligned}$$

then the minimization of

$$(1 - l)F + lF' + mS$$

is equivalent to the minimization of

$$F^* + mS ,$$

where

$$F^* = \Sigma_i \left[(1 - l)w_i + lw_i'\right](u_i - u_i^*)^2 .$$

2. *Smoothness*

The smoothness function S in the paper has been proposed by R. Henderson ([6], p. 43):

The same principle may be applied with other measures of departure from smoothness such as $\Sigma(\Delta^2 q_x)^2$ or $\Sigma(\Delta^3 q_x)^2 + l\Sigma(\Delta^2 q_x)^2$ or $\Sigma(r\Delta^2 q_x - r^{-1}\Delta^2 q_{x+1})^2$ the last form being suggested by the assumption that through most of the table q_x would be approximately of the form $A + Hx + Bc^x$ where $c = r^2$ as a table following that law would be exactly reproduced. These assumptions would give, instead of $-\delta^6 q_x$, respectively $\delta^4 q_x$, $-\delta^6 q_x + l\delta^4 q_x$ and $-\delta_6 q_x + (r^{-1} - r)^2 \delta^4 q_x$ with appropriate modifications at the ends of the series.

The smoothness function S has also been studied by K. Camp [2]. Since

$$\Delta^z u_i - r\Delta^{z-1} u_i = \Delta^{z-1} u_{i+1} - (1 + r)\Delta^{z-1} u_i ,$$

in Camp's terminology ([2], p. 19) we are *requiring that the* $(z - 1)$*th differences of the graduated series be constrained toward a geometric progression with constant ratio* $1 + r$.

Remarks:

1. The term *Camp form* has been used to describe the expression arising from such a smoothness criterion ([7], p. 25). However, the term *Henderson-Camp form* may be more appropriate.
2. In the Henderson-Camp method, the ratio of the geometric series toward which a given order of differences is to be constrained is determined by a preliminary graduation; thus S is really another criterion for fit.
3. The function S is a special case of

$$\Sigma_i \, [p(E)u_i]^2 ,$$

where $p(E)$ is a polynomial in E, the forward-shift operator. In [5] there is a FORTRAN program (written in 1961) for quartic polynomials p. For a method to determine the coefficients of p see [5], page 393.

4. As quoted above, Henderson also proposed minimizing

$$\alpha\Sigma_i \, (\Delta^z u_i)^2 + \beta\Sigma_i \, (\Delta^{z-1} u_i)^2 + \ldots .$$

This has been called by C. A. Spoerl [12] the *mixed difference case* and can also be treated by the FORTRAN program in [5].

5. It is interesting to note that A. C. Cragoe ([2], p. 413) programmed the Choleski square-root method for Whittaker-Henderson graduation on an IBM 650 computer nearly thirty years ago.

3. *Alternative Proof*

We now derive the results in the first half of Appendix I by multivariable calculus as in [11]. We shall follow the notation in the paper except for vectors which will be denoted by small letters.

Let

$$V = (1 - l)W + lW' ,$$

$$\boldsymbol{u} = (u_1, \ldots , u_n)^T ,$$

and

$$\boldsymbol{u}^* = (u_1^*, \ldots, u_n^*)^T .$$

Thus we wish to minimize the function

$$\begin{aligned} f(\boldsymbol{u}) &= F^* + mS \\ &= (\boldsymbol{u} - \boldsymbol{u}^*)^T V(\boldsymbol{u} - \boldsymbol{u}^*) + m\boldsymbol{u}^T K^T K\boldsymbol{u} . \end{aligned} \tag{3.1}$$

Since f is quadratic in $\boldsymbol{u}$, all derivatives of order three and higher are zero. Hence we have the three-term Taylor expansion

$$f(\boldsymbol{u} + \epsilon) = f(\boldsymbol{u}) + \epsilon^T \nabla f(\boldsymbol{u}) + \tfrac{1}{2}\epsilon^T H \epsilon ,$$

where the gradient vector

$$\nabla f(\boldsymbol{u}) = 2V(\boldsymbol{u} - \boldsymbol{u}^*) + 2mK^T K\boldsymbol{u} ,$$

and the Hessian matrix

$$\begin{aligned} H &= 2(V + mK^T K) \\ &= 2A . \end{aligned}$$

Since

$$\epsilon^T H \epsilon \geqslant \mathbf{0} \quad \text{for all } \epsilon \in \boldsymbol{R}^n ,$$

a vector $\boldsymbol{u} \in \boldsymbol{R}^n$ is a minimum if and only if

$$\nabla f(\boldsymbol{u}) = \mathbf{0} ,$$

or

$$(V + mK^T K)\boldsymbol{u} = V\boldsymbol{u}^* . \tag{3.2}$$

It is easy to check that equation (3.2) is the same as the equation

$$AU = B$$

in the paper.

4. *Positive Definiteness of A*

It is perhaps useful to rephrase the results in the second half of Appendix I in terms of nullspaces (kernels).

For a matrix M, let $N(M)$ denote its nullspace. We say a matrix M is nonnegative definite if M is symmetric and $\boldsymbol{x}^T M \boldsymbol{x} \geqslant 0$ for all $\boldsymbol{x}$ ([8], p. 94).

LEMMA 1. *Let M be a nonnegative definite matrix; then*

$$N(M) = \{\boldsymbol{x} \mid \boldsymbol{x}^T M \boldsymbol{x} = 0\} .$$

COROLLARY. *Let C and D be two nonnegative definite matrices; then*

$$N(C + D) = N(C) \cap N(D) .$$

The matrix

$$A = V + mK^T K$$

is positive definite if and only if $N(A)$ is just the zero vector.

$$\begin{aligned} N(A) &= N(V) \cap N(K^T K) \\ &= N(V) \cap N(K) \\ &= N((1 - l)W) \cap N(lW') \cap N(K) . \end{aligned}$$

Thus the results in the paper follow from the following fact ([10], p. 48, no. 75):

Let the polynomials $p_1(x), p_2(x), \ldots, p_k(x)$ be $\not\equiv 0$ and of degree $z_1 - 1, z_2 - 1, \ldots, z_k - 1$, respectively. Let the real constants $a_1, a_2, \ldots, a_k$ be distinct. Then the function

$$p_1(x)e^{a_1 x} + p_2(x)e^{a_2 x} + \ldots + p_k(x)e^{a_k x}$$

has at most $z_1 + z_2 + \ldots + z_k - 1$ real zeros.

5. *Existence of Minimum*

Equation (3.2),

$$A\boldsymbol{u} = V\boldsymbol{u}^* ,$$

always has a solution. Of course, if A is invertible, then

$$\boldsymbol{u} = A^{-1}V\boldsymbol{u}^* .$$

LEMMA 2. *If M is a symmetric matrix, then the range of M, $R(M)$, is the orthogonal complement of $N(M)$.*

Since

$$N(A) \subseteq N(V) ,$$

by Lemma 2

$$R(V) \subseteq R(A) .$$

Thus equation (3.2) always has a solution. Furthermore, the solution is unique if we require that it be a vector in $R(A)$.

6. *Nonnegativity of Solutions*

It is pointed out in the paper that if m is large, graduated values will be very close to an exponential function plus a polynomial. (See also [3], sec. 2, and [4], p. 91.) Hence, solutions of equation (3.2) may contain negative values. We should consider the following problem:

Find $\boldsymbol{u} \in \boldsymbol{R}^n_+ = \{\boldsymbol{x} \in \boldsymbol{R}^n \mid \boldsymbol{x} \geqslant \boldsymbol{0}\}$ such that for each $\boldsymbol{x} \in \boldsymbol{R}^n_+$

$$f(\boldsymbol{u}) \leqslant f(\boldsymbol{x}) ,$$

where the function f is defined by (3.1).

The following result is an immediate consequence of the Kuhn-Tucker theorem. An elementary analytical proof can be given (see [9]), but for $n = 2$ it is quite instructive to derive this theorem geometrically by sketching level curves (see [13] or [1], chap. 5).

THEOREM. *Let $g: \boldsymbol{R}^n \longrightarrow \boldsymbol{R}$ be a differentiable convex function. Let $\boldsymbol{u} \in \boldsymbol{R}^n_+$; then*

$$g(\boldsymbol{u}) \leqslant g(\boldsymbol{x}) \quad \text{for all } \boldsymbol{x} \in \boldsymbol{R}^n_+$$

if and only if

$$\nabla g(\boldsymbol{u}) \geqslant \boldsymbol{0}$$

and

$$\boldsymbol{u}^T \nabla g(\boldsymbol{u}) = 0 .$$

Therefore, in a Whittaker-Henderson graduation, if we restrict the graduated series to be nonnegative, we should seek only those nonnegative vectors $\boldsymbol{u}$ such that

$$A\boldsymbol{u} \geqslant V\boldsymbol{u}^*$$

and

$$u^T A u = u^T V u^* \ .$$

7. *Quadratic Programming*

As suggested in Section V of the paper, a Whittaker-Henderson graduation should be formulated as a quadratic-programming problem. Quite a few techniques are available to solve a quadratic-programming problem. Most of them transform it to a linear programming so that the simplex algorithm can be used; the Theil and Van de Panne method ([13]; [1], chap. 5) is different.

In the Theil and Van de Panne method, a systematic search is made for those solutions that lie on the boundaries of structural and nonnegativity constraints. The unconstrained problem is solved first. Then the constraints are added one at a time until a solution satisfying all constraints is found. Not all possible combinations of constraints need to be considered because of several ingenious rules devised by Theil and Van de Panne. This method seems to be most suitable for our problem since the solution probably lies on relatively few of the constraining hyperplanes. Indeed, if the solution lies in the interior (an unconstrained minimum), the method converges in one step—that is, we simply solve equation (3.2).

REFERENCES

1. BOOT, J. C. G. *Quadratic Programming: Algorithms—Anomalies—Applications*. Amsterdam: North-Holland, 1964.
2. CAMP, K. "New Possibilities in Graduation," *TSA*, VII (1955), 6–30; Discussion, 413–16.
3. CHAN, F. Y.; CHAN, L. K.; and MEAD, E. R. "Properties and Modifications of Whittaker-Henderson Graduation," *Scandinavian Actuarial Journal*, 1982, pp. 57–61.
4. GERBER, H. U. *An Introduction to Mathematical Risk Theory*. Homewood, Ill.: Richard D. Irwin, 1980.
5. GREVILLE, T. N. E. "A Fortran Program for Generalized Whittaker-Henderson Graduation," *ARCH*, 1982.1, 385–401.
6. HENDERSON, R. *Mathematical Theory of Graduation*. 2d ed. New York: Actuarial Society of America, 1938.
7. HICKMAN, J. C., and MILLER, R. B. "Notes on Bayesian Graduation," *TSA*, XXIX (1977), 7–21; Discussion, 23–49.
8. LANCASTER, P. *Theory of Matrices*. New York: Academic Press, 1969.
9. MANGASARIAN, O. L. "Pseudo-Convex Functions," *SIAM J. Control*, III (1965), 281–90.

10. PÓLYA, G., and SZEGÖ, G. *Aufgaben und Lehrsätze aus der Analysis*, vol. 2. Berlin: Springer, 1925.
11. SHIU, E. S. W. "Matrix Whittaker-Henderson Graduation Formula," *ARCH*, 1977.1.
12. SPOERL, C. A. "The Whittaker-Henderson Graduation Formula A: The Mixed Difference Case," *TASA*, XLII (1941), 292–313; Discussion, XLIII (1942), 68–80.
13. THEIL, H., and VAN DE PANNE, C. "Quadratic Programming as an Extension of Classical Quadratic Maximization," *Management Science*, VII (1960), 1–20.

F. Y. CHAN, L. K. CHAN,* AND E. R. MEAD:

It is interesting to see that standard values and exponential smoothing are incorporated into the traditional Whittaker-Henderson graduation method.

We will show that Mr. Lowrie's extension can be expressed in the form of the traditional method plus terms not involving the graduated values. Consequently, the following results for the traditional method can be obtained for the extension: (1) Greville's [3] solution procedure, (2) Chan, Chan, and Mead's [1] properties and modifications, and (3) Gerber's [2] Pareto-optimal property.

Notation used by Mr. Lowrie will be used throughout the discussion.

Mr. Lowrie proposes the following formula:

$$G(U) = F(U) + mS(U) ,$$

where

$$F(U) = (1 - l)(U - U'')^T W(U - U'') + l(U - S)^T W'(U - S) ,$$
$$S(U) = U^T K^T K U .$$

It can be expressed as

$$G(U) = \hat{F}(U) + mS(U) - \hat{U}''^T \hat{W} \hat{U}'' + (1 - l)U''^T W U'' + lS^T W' S ,$$

where

$$\hat{W} = (1 - l)W + lW' , \quad \hat{U}'' = \hat{W}^{-1}[1 - l)WU'' + lW'S] ,$$
$$\hat{F}(U) = (U - \hat{U}'')^T W(U - \hat{U}'') .$$

* Dr. L. K. Chan, not a member of the Society, is Professor and Head, Department of Statistics, University of Manitoba, Winnipeg, Manitoba, Canada.

The last three terms of $G(U)$ do not involve U. Hence, for a given m, minimizing $G(U)$ is equivalent to minimizing

$$\hat{G}(U) = \hat{F}(U) + mS(U) .$$

Since W and W' are diagonal matrices, $\hat{W}$ is diagonal. Assume that its diagonal elements are positive (this is the case when all $w_i > 0$ and $l < 1$, or all $w'_i > 0$ and $l > 0$). $S(U)$ is a sum of squares, so K^TK is positive semidefinite. Therefore, $\hat{G}(U)$ is te form of the traditional Whittaker-Henderson formula, and $\hat{W}$ and K^TK have the required properties.

1. Using Greville's solution procedure, one obtains

$$(\hat{W} + mK^TK)U = \hat{W}\hat{U}'' ,$$

which is the same as Lowrie's $AU = B$.

2. The following theorem follows from the arguments used in the proof of Theorem 1 by Chan, Chan, and Mead.

> THEOREM. *If for each real number $m > 0$, U_m ($\neq \hat{U}''$) is the value of U which minimizes $G(U)$, then $G(U_m)$, $F(U_m)$, and $S(U_m)$ are, respectively, nondecreasing, increasing, and decreasing functions of m.*

One can also obtain results similar to Chan, Chan, and Mead's Theorems 2 and 3 regarding the modified graduation problem: to minimize $F(U)$ subject to the constraint that $S(U)$ does not exceed a given tolerance level or vice versa.

3. Using Gerber's proof ([2], p. 91), one can show that U_m is Pareto-optimal in the sense that there does not exist a U such that $F(U) \leq F(U_m)$ and $S(U) \leq S(U_m)$ with at least one of the inequalities being strict.

REFERENCES

1. CHAN, F. Y.; CHAN, L. K.; and MEAD, E. R. "Properties and Modifications of Whittaker-Henderson Graduation," *Scandinavian Actuarial Journal*, 1982, pp. 57–61.
2. GERBER, H. U. *An Introduction to Mathematical Risk Theory.* Homewood, Ill.: Richard D. Irwin, 1979.
3. GREVILLE, T. N. E. Part 5 Study Notes—Graduation. *Society of Actuaries*, 1974.

(AUTHOR'S REVIEW OF DISCUSSION)

WALTER B. LOWRIE:

I wish to thank all those who submitted discussions to my paper as well as Mr. Richard London for pointing out certain errors in the galley proof.

Frank Knorr's discussion points out that my last two programs could be condensed to

$$U \leftarrow B \boxminus A \; .$$

I started this way, but knew from previous work that the Choleski factorization is much more efficient if A has more than twenty-three rows (columns). Of course, A must be positive definite. Mr. Knorr points out that the total number of deaths and the average age at death are not preserved in "my" method. Mr. Evans also mentions this problem and proposes a solution which works very well. This solution is proper if one assumes that the "standard" values have the proper shape but the level has to be adjusted to match the data. I will return to this subject after commenting on Mr. Shiu's discussion. Mr. Evans points out that the determination of r is difficult. His approach is good. I estimated r by forming ratios of u_{x+1}/u_x and averaging these ratios in the middle of the range. This is a crude method but would provide one of the starting points for a better method, such as Newton's method or golden section search (see Gill et al. [1], p. 90). Looking at Mr. Evans's results, I am surprised at the relative insensitivity of the objective function to the value of r.

I would like to correct Mr. Chittenden's idea that one must supply all the entries in a 49 × 49 matrix. One enters only the observed and standard values, the corresponding weights, and a few constants. (These values must be entered correctly—even if the method is graphical.) Then the computer calculates the 49 × 49 matrix with far higher accuracy than 99.9 percent. Mr. Chittenden is concerned with the philosophy of graduation. His questions are worth considering—essentially: do we need a high-powered method to do something that is simple? The motto of the Society tells us to seek the truth (facts for appearances). In this case, then, we are trying to estimate "true" underlying mortality rates. Graduation can be regarded as a procedure to eliminate random fluctuations in observed data. Sometimes we deliberately ignore the truth for practical reasons (e.g., the dip in mortality for males in their twenties) to make the results of a graduation more practical. On the other hand, our a priori knowledge and the observed data must also be reflected in the final results.

Graphical methods are excellent for performing a graduation or reviewing a graduation performed by other means. The effects of the observed data, a priori knowledge, and practicality can be evaluated from a graph. "My" method, among others, can perform a graduation almost instantly with a cost of one to three dollars. Modifications can be made and reruns done very quickly. The computer can be programmed to produce graphical results to review the graduation. The computer can calculate numbers and draw graphs more quickly and cheaply than a human can. The computer does not mind doing the same thing over and over again. I believe "my" method is practical once the program is installed in the computer. Also, the research underlying new methods can uncover new truths or expose new aspects of a problem.

Mr. Shiu and Messrs. Chan, Chan, and Mead show that "my" method of using standard values *and* observed values is equivalent to using composite weights and composite values. Mr. Shiu shows that the composite weights and values are (respectively):

$$w_i^* = (1 - l)w_i + lw_i' ,$$

$$u_i^* = 0 \qquad \text{if } (1 - l)w_i + lw_i' = 0$$

$$= \frac{(1 - l)w_i u_i'' + lw_i' s_i}{(1 - l)w_i + lw_i'} \qquad \text{otherwise} .$$

Then, as long as the matrix A is positive definite, the results in Mr. Greville's study note can be applied. In particular,

$$\Sigma_i w_i^* u_i^* = \Sigma_i w_i^* u_i ; \quad \Sigma_i iw_i^* u_i^* = \Sigma_i iw_i^* u_i . \tag{1}$$

If the composite values u_i^* are our best estimates of the true values (using a credibility argument, say) then using the composite weights w_i^* gives us the relationships 1). I have a feeling that Mr. Evans's approach is more appropriate, however, but I have no proof. There is much to be done, as is mentioned in the paper and in Mr. Shiu's discussion, in minimizing the fit and smoothness function subject to appropriate constraints.

Mr. Shiu and Messrs. Chan, Chan, and Mead's results lead to the following analysis to show what really happens with a modified Whittaker-Henderson method. Minimizing

$$F^* + mS ,$$

where

$$F^* = \Sigma_i w_i^*(u_i - u_i^*)^2 ,$$

$$S = \Sigma_i (\Delta^z u_i - r\Delta^{z-1} u_i)^2 ,$$

can be alternately viewed as a penalty optimization problem.

The solution to the penalty problem is a function fit by least squares to the points u_i^* with the weights w_i^*. This function satisfies the constraints in S and has the form

$$u_i + P_{z-2}(i) + k(1 + r)^i .$$

For a given value of m, the corresponding U_m is an approximation to the least-squares function. As m gets larger, if the U_m approach a limit, then the limit is the solution to the least-squares problem (see Luenberger [2], p. 280). In other words, increasing m gives a function that is closer to the least-squares function.

REFERENCES

1. Gill, P. E.; Murray, W.; Wright, M. H. *Practical Optimization.* New York: Academic Press, 1981.
2. Luenberger, D. G. *Introduction to Linear and Non-linear Programming.* Reading, Mass.: Addison-Wesley, 1973.