NOTE ON GRADUATION BY ADJUSTED AVERAGE.

BY

ROBERT HENDERSON.

The expression "graduation by adjusted average" is used to include those methods by which the graduated value of a function is determined by adding together a number of adjacent terms each multiplied by a numerical factor. Summation formulas are thus included as particular cases but other formulas are intended to be included which cannot be expressed in the form of successive summations. We will first investigate the formula which will arise when the graduated value is determined by fitting an algebraic function of the third degree to the ungraduated values. It is known that in the case of an algebraic function the method of least squares and the method of moments give identical results.

Let U_0 denote the ungraduated value of the function corresponding to the graduated value to be determined and suppose that only terms from U_{-n} to U_n are to be taken into account and that W_x is the weight to be assigned to the term U_x for all values of x.

Then if $a + bx + cx^2 + dx^3$ be the algebraic function determined and consequently a the graduated value of U_0 we have the following four equations in a, b, c and d:

$$\Sigma_{-n}^{+n}(a + bx + cx^2 + dx^3) W_x = \Sigma_{-n}^{+n} W_x U_x,$$

$$\Sigma_{-n}^{+n}(ax + bx^2 + cx^3 + dx^4) W_x = \Sigma_{-n}^{+n} x W_x U_x,$$

$$\Sigma_{-n}^{+n}(ax^2 + bx^3 + cx^4 + dx^5) W_x = \Sigma_{-n}^{+n} x^2 W_x U_x,$$

$$\Sigma_{-n}^{+n}(ax^3 + bx^4 + cx^5 + dx^6) W_x = \Sigma_{-n}^{+n} x^3 W_x U_x.$$

Or if we designate $\Sigma_{-n}^{+n} x^r W_x$ by s_r,

$$s_0 a + s_1 b + s_2 c + s_3 d = \Sigma_{-n}^{+n} W_x U_x,$$

$$s_1 a + s_2 b + s_3 c + s_4 d = \Sigma_{-n}^{+n} x W_x U_x,$$

$$s_2 a + s_3 b + s_4 c + s_5 d = \Sigma_{-n}^{+n} x^2 W_x U_x,$$

$$s_3 a + s_4 b + s_5 c + s_6 d = \Sigma_{-n}^{+n} x^3 W_x U_x.$$

These are linear equations in a, b, c and d, therefore we have:

$$a = h\Sigma_{-n}^{+n} W_x U_x + j\Sigma_{-n}^{+n} x W_x U_x + k\Sigma_{-n}^{+n} x^2 W_x U_x + l\Sigma_{-n}^{+n} x^3 W_x U_x$$
$$= \Sigma_{-n}^{+n}(h + jx + kx^2 + lx^3) W_x U_x,$$

provided h, j, k and l are determined so that

$$hs_0 + js_1 + ks_2 + ls_3 \text{ or } \Sigma_{-n}^{+n}(h + jx + kx^2 + lx^3) W_x = 1,$$
$$hs_1 + js_2 + ks_3 + ls_4 \text{ or } \Sigma_{-n}^{+n} x(h + jx + kx^2 + lx^3) W_x = 0,$$
$$hs_2 + js_3 + ks_4 + ls_5 \text{ or } \Sigma_{-n}^{+n} x^2(h + jx + kx^2 + lx^3) W_x = 0,$$
$$hs_3 + js_4 + ks_5 + ls_6 \text{ or } \Sigma_{-n}^{+n} x^3(h + jx + kx^2 + lx^3) W_x = 0.$$

If then we write V_x for $(h + jx + kx^2 + lx^3) W_x$ and put U'_0 for the graduated value of U_0, we have

$$U'_0 = \Sigma_{-n}^{+n} V_x U_x,$$

where

$$\Sigma_{-n}^{+n} V_x = 1,$$
$$\Sigma_{-n}^{+n} x V_x = 0,$$
$$\Sigma_{-n}^{+n} x^2 V_x = 0,$$
$$\Sigma_{-n}^{+n} x^3 V_x = 0.$$

Thus far no limitations have been placed on the values of W_x so that none are placed on those of V_x except the four conditions expressed above. If, however, we suppose that the values of W_x are all necessarily positive then the values of V_x will change sign along with $(h + jx + kx^2 + lx^3)$. Conversely any series of values of V_x, satisfying these four conditions and changing sign not more than three times, may be considered as derived in this manner from a series of values of W_x of the form $V_x/(h + jx + kx^2 + lx^3)$, where the function $h + jx + kx^2 + lx^3$ changes sign in the same intervals as V_x.

Suppose now that the values of W_x are symmetrical about W_0, so that $W_x = W_{-x}$. Then, where r is odd, we have $s_r = 0$ since positive and negative terms will cancel one another, so that the four equations in a, b, c and d reduce to

$$s_0 a + s_2 c = \Sigma_{-n}^{+n} W_x U_x,$$
$$s_2 b + s_4 d = \Sigma_{-n}^{+n} x W_x U_x,$$
$$s_2 a + s_4 c = \Sigma_{-n}^{+n} x^2 W_x U_x,$$
$$s_4 b + s_6 d = \Sigma_{-n}^{+n} x^3 W_x U_x.$$

Since a appears only in the first and third of these equations we have

$$a = h\Sigma_{-n}^{+n} W_x U_x + k\Sigma_{-n}^{+n} x^2 W_x U_x$$
$$= \Sigma_{-n}^{+n} (h + kx^2) W_x U_x,$$

where

$$hs_0 + ks_2 \text{ or } \Sigma_{-n}^{+n} (h + kx^2) W_x = 1,$$
$$hs_2 + ks_4 \text{ or } \Sigma_{-n}^{+n} x^2 (h + kx^2) W_x = 0.$$

Here

$$V_x = (h + kx^2) W_x.$$

Hence the values of V_x are also symmetrical. Conversely any symmetrical graduation formula such that $V_x = V_{-x}$ with two changes of sign may be considered as derived from a symmetrical set of weights W_x of the form $V_x/(h + kx^2)$ where $h + kx^2$ changes sign in the same intervals as V_x.

It is interesting to determine the relative weights which would produce certain well known summation formulas.

In Woolhouse's formula V_0 to V_4 inclusive are positive, V_0 being equal to $\frac{1}{5}$. V_5 is zero and V_6 and V_7 are negative. This suggests $(25 - x^2)/125$ as the transforming factor. This gives the following series of values for W_0 to W_7 inclusive, W_{-x} being, it will be remembered, equal to W_x for all values of x.

$$1;\ 1;\ 1;\ \tfrac{7}{16};\ \tfrac{6}{18};\ ?;\ \tfrac{4}{22};\ \tfrac{3}{24}.$$

The value of W_5 is indeterminate but the analogy of the neighboring terms suggests the value $\frac{5}{20}$ or $\frac{1}{4}$.

Similarly in Higham's formula we have for the values of W_0 to W_8 inclusive,

$$1;\ 1;\ \tfrac{6}{7};\ \tfrac{5}{8};\ \tfrac{1}{3};\ \tfrac{1}{4};\ \tfrac{2}{11};\ \tfrac{1}{12};\ \tfrac{1}{39}.$$

And in Karup's formula we have for the values of W_0 to W_9 inclusive,

$$1;\ \tfrac{19}{20};\ \tfrac{29}{35};\ \tfrac{53}{80};\ \tfrac{7}{15};\ \tfrac{1}{4};\ \tfrac{8}{55};\ \tfrac{3}{40};\ \tfrac{2}{65};\ \tfrac{1}{140}.$$

In all of these cases the vanishing of V_5 gives a suggestion of the transforming factor, but in the case of Spencer's 21 term formula none of the terms vanish. The change of sign, however, takes place between $x = 5$ and $x = 6$ and investigation shows that in order that W_x should always decrease as x increases numerically the change of sign in the transforming factor must occur between $x^2 = 29.5$ and $x^2 = 33.25$.

Taking $(30 - x^2)/175$ as the factor the series of weights is

$$1;\ 57/58;\ 47/52;\ 11/14;\ 9/14;\ 3/5;\ 1/6;\ 5/38;\ 5/68;\ 1/34;\ 1/140,$$

and taking $2(33 - x^2)/385$ as the factor it is

$$1;\ 627/640;\ 517/580;\ 121/160;\ 99/170;\ 33/80;\ 11/30;\ 11/64;\ 11/124;$$
$$11/320;\ 11/1340.$$

This result illustrates the fact that the same series of values of V_x may be derived from many different series of values of W_x.

Let us now revert to the case where the values of V_x are not necessarily symmetrical and suppose that graduated values are determined for a number of successive terms, the same values of V_x being used for each. Then generally we have

$$U'_y = \Sigma_{x=-\infty}^{x=+\infty} V_x U_{y+x} = \Sigma_{z=-\infty}^{z=+\infty} V_{z-y} U_z, \text{ where } z = y + x.$$

While the above summation is indicated as extending infinitely in both directions the terms under the summation will vanish where $V_x = 0$.

Let us now examine the conditions for the smoothest possible graduated series. In doing so we shall tentatively adopt the usual criterion of smoothness, namely the smallness of the mean square of the error in the third difference. Then

$$\Delta^3 U'_y = \Sigma_{y=-\infty}^{y=+\infty} U_z \Delta^3 V_{z-y},$$

where the variable on both sides is y. Then the mean square of the error in $\Delta^3 U'_y$ is

$$\Sigma_{-\infty}^{+\infty} e_z^2 (\Delta^3 V_{z-y})^2,$$

where e_z^2 is the mean square of the error in U_z. Summing this for all values of y and assuming that, except for a finite range of values of z, e_z^2 vanishes we have for the sum

$$\Sigma_{y=-\infty}^{y=+\infty} \Sigma_{z=-\infty}^{z=+\infty} e_z^2 (\Delta^3 V_{z-y})^2 = \Sigma_{z=-\infty}^{z=+\infty} e_z^2 \Sigma_{y=-\infty}^{y=+\infty} (\Delta^3 V_{z-y})^2$$
$$= \Sigma_{z=-\infty}^{z=+\infty} e_z^2 \Sigma_{x=-\infty}^{x=+\infty} (\Delta^3 V_{x-3})^2 \text{ since } \Delta^3 V_{z-y} (y \text{ variable}) = -\Delta^3 V_{x-3}.$$

But $\Sigma_{x=-\infty}^{x=+\infty} (\Delta^3 V_{x-3})^2$ is independent of z so that the expression reduces to the product of two factors, one of which, $\Sigma_{z=-\infty}^{z=+\infty} e_z^2$, is independent of the graduation formula. It is only necessary therefore to consider the conditions for a minimum value of $\Sigma_{x=-\infty}^{x=+\infty}(\Delta^3 V_{x-3})^2$.

The differential coefficient of this sum with respect to V_x is

$$2\{\Delta^3 V_{x-3} - 3\Delta^3 V_{x-2} + 3\Delta^3 V_{x-1} - \Delta^3 V_x\} = -2\Delta^6 V_{x-3}$$

and therefore the complete differential is

$$\Sigma_{x=-\infty}^{x=+\infty} - 2\Delta^6 V_{x-3}\delta V_x.$$

But owing to the four conditions limiting V_x we have, where all values of V_x except those from V_{-n} to V_n inclusive vanish:

$$\Sigma_{x=-n}^{x=n}\delta V_x = 0,$$

$$\Sigma_{x=-n}^{x=n} x\delta V_x = 0,$$

$$\Sigma_{x=-n}^{x=n} x^2\delta V_x = 0,$$

$$\Sigma_{x=-n}^{x=n} x^3\delta V_x = 0.$$

For the smoothest possible series with a graduation formula extending from $-n$ to n inclusive we must have

$$\Sigma_{x=-n}^{x=+n}\Delta^6 V_{x-3}\delta V_x = 0$$

for all values of the differentials satisfying these four equations, or in other words we must have a relation of the form

$$\Delta^6 V_{x-3} = \kappa + \lambda x + \mu x^2 + \nu x^3$$

for values of x from $-n$ to n inclusive.

But $\Delta^6 V_{x-3}$ includes terms from V_{x-3} to V_{x+3} inclusive. Therefore V_x must be an algebraic function of not more than the ninth degree for values of x from $-(n+3)$ to $n+3$ inclusive.

But we know that V_x vanishes for six values of x, viz., $-(n+3)$; $-(n+2)$; $-(n+1)$; $n+1$; $n+2$ and $n+3$; therefore $\{(n+1)^2 - x^2\}\{(n+2)^2 - x^2\}\{(n+3)^2 - x^2\}$ must be a factor of V_x which must therefore take the form

$$V_x = \{(n+1)^2 - x^2\}\{n+2)^2 - x^2\}\{n+3)^2 - x^2\}(h + jx + kx^2 + lx^3),$$

where h, j, k and l are to be determined from the four conditions limiting V_x.

This is seen to be exactly the formula which would be arrived at by assigning weights $W_x = \{(n+1)^2 - x^2\}\{(n+2)^2 - x^2\}$

$\{(n+3)^2-x^2\}$ and as the values of W_x are symmetrical we see that j and l vanish and the expression reduces to

$$V_x=(h+kx^2)W_x.$$

Suppose now, to obtain another view of the matter we approach it from the standpoint of a smooth series of weights and determine the condition for a minimum value of $\Sigma_{-\infty}^{+\infty}(\Delta^3 W_x)^2$ subject to the condition that all values of W_x except those from W_{-n} to W_n inclusive should vanish and that $\Sigma_{x=-n}^{x=n} W_x=K$. In this case we have by reasoning similar to the above a relation $\Delta^6 W_{x-3}=k$ for values of x from $-n$ to n inclusive. Hence W_x is of the sixth degree from $-(n+3)$ to $n+3$ inclusive. But W_x vanishes for the six values of x, $-(n+3)$; $-(n+2)$; $-(n+1)$; $(n+1)$; $(n+2)$ and $(n+3)$; hence W_x takes the form

$$W_x=W\{(n+1)^2-x^2\}\{(n+2)^2-x^2\}\{(n+3)^2-x^2\}.$$

The same formula is thus obtained as before, since only relative weights affect the final result. We thus see that the smoothest possible graduated series from a formula of given range is obtained by assigning the smoothest possible series of weights to the successive terms.

In the preceding investigation it has been assumed that the weight assigned to a term depends only on its distance from the term for which the graduated value is being determined. In view, however, of the fact that the intrinsic weights of the terms of the ungraduated series are sometimes known to vary from term to term it may be desirable to take these variations into account and to assign to each term a weight calculated as the product of two factors one $w_z=w_{y+x}$ peculiar to the term itself and the other $W_x=W_{z-y}$ dependent only on the distance. We must then substitute $W_x w_{y+x}$ for W_x and consider the series as probably not symmetrical. Analogy would suggest that the smoothest series of graduated values would be obtained by giving to w_z values inversely proportional to e_z^2 and to W_x the same values as before. The problem of testing this suggestion is left to the readers.