ACTUARIAL RESEARCH CLEARING HOUSE
1980 VOL. 1

Comments on the treatment of "improving the fit" in
Elements of Graduation, by M. D. Miller

Walter B. Lowrie, F.S.A.

On page 9 of the text, two expressions are given:

$$\sum(\theta''_x - \theta_x) \tag{1}$$

$$\sum x(\theta''_x - \theta_x) \tag{2}$$

We are told that these expressions should be close to zero to assure a good fit. (In the text, θ''_x denotes ungraduated deaths, θ_x the first graduation and θ'_x the second graduation.

Then on page 10, the transformation $q'_x = aq_x + b$ is given. The constants a and b are determined so that the deviations and accumulated deviations are zero, that is:

$$\sum\theta'_x = a\sum\theta_x + b\sum E_x = \sum\theta''_x \tag{3}$$

$$\sum{}^2\theta'_x = a\sum{}^2\theta_x + b\sum{}^2 E_x = \sum{}^2\theta''_x \tag{4}$$

This paper has three objectives:

(A) If (3) and (4) are satisfied, then

$$\sum_{x=\alpha}^{\omega}(\theta''_x - \theta'_x) = 0 \tag{1'}$$

and $$\sum_{x=\alpha}^{\omega}x(\theta''_x - \theta'_x) = 0 \tag{2'}$$

This result is independent of whether the accumulated deviations are "backward" or "forward".

(B) On page 10, there is a statement that "The sum of the deviations and the sum of the column of sub-totals should be close to zero. This is mathematically equivalent to the requirement that the sum and the first moment of the deviations be close to zero." (emphasis added) This statement is true but it is a bit general. It leads to objective (A) but gives the impression that $\sum^2(\theta''_x - \theta_x) = \sum_{x=\alpha}^{\omega} x(\theta''_x - \theta_x)$. This is not necessarily true. The proofs further on in the paper show that equality obtains only for "backward" sums when $\alpha = 1$ (see equation (5)).

(C) Some remarks are made about measures of fit.

To analyze the objectives we need the following:

1. Backward Sums. These are denoted by $\uparrow\sum^2\theta_x$:

$$\uparrow\sum{}^2\theta_x = \sum_{x=\alpha}^{\omega}\sum_{t=x}^{\omega}\theta_t =$$

$$\begin{array}{rrrrr}
\theta_\alpha + \theta_{\alpha+1} & + & \cdots & + \theta_{w-1} & + \theta_w \\
+ \theta_{\alpha+1} & + & \cdots & + \theta_{w-1} & + \theta_w \\
\ddots & & & \vdots & \vdots \\
 & & & + \theta_{w-1} & + \theta_w \\
 & & & & + \theta_w \\
\hline
\theta_\alpha + 2\theta_{\alpha+1} & + & \cdots & + (w-\alpha)\theta_{w-1} & + (w-\alpha+1)\theta_w
\end{array}$$

$$= (\alpha-(\alpha-1))\theta_\alpha + ((\alpha+1) - (\alpha-1))\theta_{\alpha+1} + \cdots$$

$$+((w-1) - (\alpha-1))\theta_{w-1} + ((w) - (\alpha-1))\theta_w$$

So

$$\uparrow\sum{}^2\theta_x = \sum_{x=\alpha}^{w} x\theta_x - (\alpha-1)\sum_{x=\alpha}^{w}\theta_x \tag{5}$$

If $\alpha = 1$, then $\uparrow\sum^2\theta_x = \sum_{x=\alpha}^{w} x\theta_x$ proving objective (B).

2. <u>Forward Sums</u>. These are denoted by $\downarrow\sum^2\theta_x$

$$\downarrow\sum^2\theta_x = \sum_{x=\alpha}^{w}\sum_{t=\alpha}^{x}\theta_t =$$

$$\begin{array}{l}
\theta_\alpha \\
+\ \theta_\alpha + \theta_{\alpha+1} \\
\quad\vdots \qquad \vdots \qquad \ddots \\
+\ \theta_\alpha + \theta_{\alpha+1} \ \cdots + \theta_{w-1} \\
+\ \theta_\alpha + \theta_{\alpha+1} \ \cdots + \theta_{w-1} + \theta_w \\
\hline
\end{array}$$

$$(w-\alpha+1)\theta \ + (w-\alpha)\theta_{+1} + \ldots + 2\theta_{w-1} + \theta_w$$

$$= ((w+1)-\alpha)\theta \ + ((w+1) - (\alpha+1))\theta_{+1} + \ldots$$

$$+ ((w+1) - (w-1))\theta_{w-1} + ((w+1) - w)\theta_w$$

So

$$\downarrow\sum^2\theta_x = (w+1)\sum_{x=\alpha}^{w}\theta_x - \sum_{x=\alpha}^{w}x\theta_x \tag{6}$$

It is clear that $\downarrow\sum^2\theta_x \neq \sum_{x=\alpha}^{w}x\theta_x$.

Going back to the original transformation:

$$q'_x = aq_x + b$$

$$\frac{\theta'_x}{E_x} = a \cdot \frac{\theta_x}{E_x} + b$$

$$\theta'_x = a\theta_x + bE_x \tag{7}$$

Summing both sides of equation (7) we get:

$$\sum_{x=\alpha}^{w}\theta'_x = a\sum_{x=\alpha}^{w}\theta_x + b\sum_{x=\alpha}^{w}E_x = \sum_{x=\alpha}^{w}\theta''_x \tag{8}$$

which is equation (3).

In other words

$$\sum_{x=\alpha}^{w}(\theta''_x - \theta'_x) = 0 \qquad (1')$$

This is the first part of objective (A).

Now multiply equation (7) by x and sum:

$$\sum_{x=\alpha}^{w} x\theta'_x = a\sum_{x=\alpha}^{w} x\theta_x + b\sum_{x=\alpha}^{w} xE_x = \sum_{x=\alpha}^{w} x\theta''_x \qquad (9)$$

If equation (9) is satisfied, then $\sum_{x=\alpha}^{w} x(\theta''_x - \theta'_x) = 0$ (2')

The final step is to show that if equation (8) holds, then equation (9) is the same as equation (4). Since equation (9) implies equation (2') then equation (4) will imply equation (2') and the second part of objective (A) will be proven. This is true for forward sums and backward sums.

<u>Proof for Backward Sums</u>

Substitute equation (5) in equation (9):

$$a\left[\uparrow\sum{}^2\theta_x + (\alpha-1)\sum_{x=\alpha}^{w}\theta_x\right] + b\left[\uparrow\sum{}^2E_x + (\alpha-1)\sum_{x=\alpha}^{w}E_x\right]$$
$$= \uparrow\sum{}^2\theta''_x + (\alpha-1)\sum_{x=\alpha}^{w}\theta''_x$$

$$a\left[\uparrow\sum{}^2\theta_x\right] + b\left[\uparrow\sum{}^2E_x\right] + (\alpha-1)\left\{a\sum_{x=\alpha}^{w}\theta_x + b\sum_{x=\alpha}^{w}E_x\right\}$$
$$= \uparrow\sum{}^2\theta''_x + (\alpha-1)\sum_{x=\alpha}^{w}\theta''_x$$

The terms multiplied by $\alpha-1$ drop out since equation (8) holds, so equation (9) becomes:

$$a\left[\uparrow\sum{}^2\theta_x\right] + b\left[\uparrow\sum{}^2E_x\right] = \uparrow\sum{}^2\theta''_x$$

which is equation (4).

The proof for Forward Sums is similar. The same constants ("a" and "b") result from using Forward Sums.

(C) Comments about measures of fit

Actually $\sum(\theta''_x - \theta_x)$ being small is not necessarily a good measure of fit since deviations of opposite sign may cancel.

The condition $\sum x(\theta''_x - \theta_x) = 0$ assures that the average age at death of the ungraduated values is the same as the average age at death of the graduated values. This is true if we define average age at death as:

$$\frac{\sum x\theta_x}{\sum \theta_x} \qquad \text{for graduated deaths}$$

$$\frac{\sum x\theta''_x}{\sum \theta''_x} \qquad \text{for ungraduated deaths.}$$

So if $\sum(\theta''_x - \theta_x) = 0$ and if $\sum x(\theta''_x - \theta_x) = 0$, then the average ages at death are equal. However, as a condition of fit, $\sum x(\theta''_x - \theta_x)$ seems poor since the higher ages are emphasized. This is a particular problem where the exposure is low. Also, deviations of opposite sign can cancel.

The conditions:

$$\sum \left|\theta''_x - \theta_x\right| \tag{10}$$

$$\text{and} \quad \sum(\theta''_x - \theta_x)^2 \tag{11}$$

are probably better measures of fit than $\sum(\theta''_x - \theta_x)$ and $\sum x(\theta''_x - \theta_x)$ since they don't allow deviations of opposite sign to cancel. Condition (11) may over-emphasize large deviations. This usually occurs when the exposure is small. Unfortunately, the mathematics involved to utilize condition (10) is more complicated. Recently, linear programming methods have been used to deal with conditions using absolute values. See "A Linear Programming Approach to Graduation" by Donald R. Schutte, TSA XXX p.407.

To go even further, the conditions:

$$\sum E_x \left| \theta''_x - \theta_x \right|$$

$$\text{and} \sum E_x \left(\theta''_x - \theta_x \right)^2$$

seem to be the best measures of fit since they de-emphasize large deviations where the exposure is small.

Similar conditions are used (in part), to generate the Whittaker-Henderson type a and b methods:

$$\text{Type a: minimize } \sum (q_y - q''_y)^2 + h \sum (\Delta^2 q_y)^2 \quad , \quad h > 0$$

$$\text{Type b: minimize } \sum E_y (q_y - q''_y)^2 + k \sum (\Delta^z q_y)^2 , \quad k > 0$$

As a by-product of the Type b formula, it can be shown that $\sum (\theta''_x - \theta_x) = 0$ and $\sum x(\theta''_x - \theta_x) = 0$.

The χ^2 test can also be used to test the adequacy of fit, after a graduation is performed. Let the observed frequencies be $\theta''_\alpha, \ldots \theta''_w$ and the (estimated) expected frequencies be $\theta_\alpha, \ldots \theta_w$. Then

$$\chi^2 = \sum_{x=\alpha}^{w} \frac{(\theta_x - \theta''_x)^2}{\theta_x}$$

with $w - \alpha$ degrees of freedom.

It is hard to generalize these comments. Some work should be done to predict the effects of using a given measure of fit.

The students in my graduation class at the University of Nebraska - Lincoln, and my colleague, Warren Luckner made valuable comments on this paper.