死亡率の補整(graduation)

鷲見, 慎一
明治生命

出版情報：統計数理研究. 1 (2), pp.114-124, 1942-03-15. Research Association of Statistical Sciences

協同研究報告

死亡率の補整 (graduation)

會員 鷲 見 愼 一 (明治生命)

(昭和十六年九月十五日受理)

緒言 Walfordの保險大辭典[1]によれば，Graduationは思想的には死亡率に關する De Moivre の 1725年の假定に遡るべきことが擧げられており，1765年の Herr Lambert の Graphic method を以て死亡率の補整が實行された最初の方法としてゐる．1825年 B. Gompertz の彼の名を冠せられる法則の發表せられてより以後今日に到るまで，保險數理の方面に於いて不斷の努力が Graduation について爲されて來てゐる．勿論 Graduation は死亡率に限られたものでもなく，統計の一般的問題であり，特に時系列の方面に於て死亡率の場合と共通的な取扱を受け，思想的技術的に從來とも交互に接觸のあつたものである．

Graduation は Ajustment とか Smoothing of Data と同義に用ひられて居り，嚴密な定義は與へられてゐないが，統計的に試みにその定義を與へてみれば，補整可能な統計系列の數値から，高次の統計的推理を爲さんが爲に，系列の數學的期望値の實驗的近似を求めんとする統計的推理の一過程である とでも言ふことができよう．過去に爲されたる集計・總和・定差・乘積・比率等の統計的推理並に將來爲さんとする統計的推理によつて，今行ふべき補整手段も種々制肘を受けるものであり，補整の方法・結果の良否等の規準にも絶對的な要素を與へることとは出來ない．

補整可能な系列の意味は後でのべよう．F. Vinci 等も qualita graduabili, 或は又 graduatoria なる言葉を使つてゐる．[2] graduation なる字を Encyclopedia Britanica で引くと目盛を刻むこととなつており，物指・寒暖計等の度量衡器の目盛りの刻み，或は齒車の齒を刻むことを意味する．graduation formula はそれに必要なる公式を意味することになる．統計的に使用される樣になつたのは何時頃からかは詳でないが，その意味は上と同樣目盛の決定にあるとなしうるのではあるまいか．このことを肯定すれば同義語として用ひられる smoothing of Data よりは思想的に進步しており，ajustment, Ausgleichung, perequazione ほど意味のはつきりした言葉でもないが，又それほど意味の狹いものでもなく，思想的に整頓されてゐて良い言葉である．

1. 統計系列 一群の觀察點 $x_1, x_2, x_3, \cdots, x_n$ に對し，統計的に求められた觀察値 $u_1, u_2, u_3, \cdots, u_n$ を一般に統計系列と稱することにする．時點に對して與へられた特定の商品の價格等，所謂時系列或は保險會社の保險契約の保險金額別件數等は何れも統計系列と言ひ得る．又上に擧げた樣な集計蒐集によつて得られた生の數字のみならず，統計的演算によつて得られたる累積・乘積・比率等も，亦觀察値なる言葉を廣義に解釋すれば凡て統計系列と言ひうる．所謂系列統計 (serial statistics) と看做しうる統計は觀察點に對する觀察値なる關係より統計系列を與へると言ふことが出來る．

我々の補整の對象はかゝる統計系列であり，その內特に死亡率を取出してこゝに論ずるのである．

2. 補整の目的　統計系列の補整値につき平均・標準偏差・總計を計算するとき，之等を觀察された未補整の統計系列の對應する moments よりの數値と正確に一致せしめる樣にする場合がある．skewness, kurtosis 等の parameters も fitting が正しければ理論的のものと一致すべき樣にすることもある．

此等の parameters は系列を特徴づけるものであるから，觀察者は何故に parameters が保存されてゐるにも拘らず補整する必要があるかとの疑問を生ずることになる．

之に對して三つの理由が擧げられる．第一のものは，最も重要性が少いものであるが，凹凸のある標本値の代りに滑かな曲線を使ひたいと言ふことであり，第二のものは，最も重要なものであるが，統計の數學的發展の上には如何なる假定が必要であるかを知らねばならぬことである．此等二つは現象を fit する爲の（系列）曲線・型に關係してゐる．第三の理由は，他の二つの中間の重要性があるが，種々の方面で，a priori な理論を試驗する爲に此等理論の結果たる統計系列の efficacy を檢査することが必要となる爲である．[3]

3. 補整の理論　統計系列に對して我々は次の假定を設ける．即ち統計系列 $\{u_i\}$ は理想的な統計より得られると推定される眞の値 $\{E(u_i)\}$ と實際の統計に於て生ずるこの眞の値からの偏り $\{W_i\}$ とより成るものと假定するのである．これは $\{E(u_i)\}$ は universe における parameter を表はし，$\{u_i\}$ はその sample value であると云ふ思想に他ならない．こゝで

(1)　$$u_i = E(u_i) + W_i \quad i = 1, 2, \cdots\cdots$$

であるが，この關係を基として $\{u_i\}$ より $\{E(u_i)\}$ を求めんとするのが我々の問題としようとする操作即ち補整である．言換へれば如何にすれば u_i を分離して $E(u_i)$ と W_i とに爲しうるかであるが，今與へられてゐるものは u_i のみであるから，$E(u_i)$ 及 W_i に對して更に假定を設けなければならない．この假定の性質が種々の補整方法を生ぜしめる．次に簡單に若干の補整方法に就て之を考へてみよう．

4. 補整の方法　a <u>解析的方法</u>　死亡率に關する知識より豫想せられる系列の性質を利用して補整する方法に解析的方法がある．これには有名なる Makeham の法則を利用してもよいし，或は Pearson の曲線を利用することもあらう．何れに於ても死亡率の相近接せる年齢間に於ける zig-zag を完全に否定することを共通點としてゐる．

かゝる見地から問題を扱ふことに反對すべき理由は全然見當らない．事實統計が理想統計に近づけば近づく程，死亡率は滑かとなるのであるから，その理想型として前述の如き解析式を假定することは意味のあることであり，數學的に死亡率を取扱ふ上からは最も便利であるべきではあるが，必然的に上の樣な解析式の一つであるべき强力な理由は之も又存在しないのである．かく理想統計の取るべき解析式に根本的假定を設けることは，實際現實に與へられた統計が簡單な上の如き解析式では充分に表現し得ざることが常識となつてゐる今日に於ては，既往の解析的方法は必ずしも最良とは言へないのである．最近此の方面に於ては Makeham の法則を改良せんとした試みが多數爲されてゐるが，式の單純性を保持せしめつつ，應用を主眼として發展せしめられつつある．

b 機械的方法　滑らかさを解析的方法程極度に求めず，zig-zag を單に機械的に減少せしめる方法に所謂加合法がある．これ等は凡て求めんとする $E\{u_i\}$ を，u_i を中心に兩側に n 個の u_{i-n}, $u_{i-n+1}, \cdots, u_{i-1}, u_i, u_{i+1}, \cdots, u_{i+n-1}, u_{i+n}$ なる 2_{n+1} 個の數値の或る加重平均と考へてゐるのである．夫々の加合式に於ける重みは一定してゐるから，この方法によつて得られた個々の結果を比較して種々の加合式の中より個々の統計に對して最良と認められるものを苦心して探す必要を生じ，且其の良否の判定にも多分の主觀が介在する餘地を殘してゐる．

5. 補整値の良否の規準　補整値の良否の判定には完全な規準が確立せられてゐるものではなく，將來これが完成するものとも考へられない．この規準は ignoramus であるばかりでなく，ignorabimus でさへあると考へられる．現在行はれてゐる最も普遍的な規準は smoothness と closeness 即ち滑かさと近密さとである．滑かさに就ては既に一寸觸れてあり，概念は得られたものとして近密さにつき若干考へてみよう．

近密さと言へば，統計から與へられた數字に近いと言ふ意味であるから，死亡率に於ては統計から得られた値，即ち粗成死亡率に近いと言ふ意味にも解釋出來るが，更に遡つて考へることも出來る．今統計から得られた系列が α 歳から ω 歳まで，若年より高年齢へと年齢順に並べられたとき解析的方法に於ては，滑らかさは完全であるから，近密さを最も良くする爲に，考へてゐる解析式を最小自乘法により單純に fit する．この解析式で表される曲線の內，最小自乘の意味で最も近密な曲線がかくして得られる．この場合 $\sum\{u_i-E(u_i)\}^2$ を最小ならしむる如き $E(u_i)$ を與へる解析式を求めるのであれば，個々の u_i の精確度 (precision) は等しいものと假定してゐるのである．然し死亡率算定の材料は，延人員と死亡人員の集計整理されたものより單純な比をとることによつて得られるものであるから，延人員を考慮して精確度に重みを付けることが考へられる．精確度が觀察數の平方根に比例して增加するとの假定の下に $\sum E_x\left(\frac{\theta_x}{E_x}-q_x\right)^2$ を最小ならしめる樣に最小自乘法を行つた解析式を求める方が良いのである．こゝに E_x は x 歳の延人員，θ_x は x 歳の死亡數，q_x は求める死亡率である．死亡率に於ては，延人員の變動は高年齢に於ては特に急激であり，死亡率に凹凸が多いのも此の部分であるから，延人員を考慮せずに解析式を fit せんとするときは，精確度を均等にする爲に粗成死亡率を豫め適宜處理してから最小自乘法を用ふべきである．

6. 補整の方針　今滑かさの規準として $E(u_x)$ の k 次定差，即ち $\Delta^k E(u_x)$ の自乘の和 $\sum_{i=\alpha}^{\omega}(\Delta^k E(u_i))^2$ を採用する．即ち單純な $n-1$ 次多項式を規準にとるのである．このことは $n-1$ 次多項式を fit したときの滑かさを最良のものとすることを意味してゐる．

然らば多項式を規準にとることが許されたものとして，一體何次の多項式を採用すべきかと言ふことが次の問題となる．又今この問題が解決されたとすれば，第二の問題はこの多項式を規準にして如何に合理的に補整を爲すべきかと言ふことである．確率論の見地から此等にその問題に對して如何なる方針が與へられるか，以下このことに就て論じたい．

7. 多項式の次數の決定[4]　統計系列 u_i が與へられてゐるとき，u_i の一次定差は

$$\Delta u_i = u_{i+1} - u_i .$$

二次定差は一次定差の定差で

$$\Delta^2 u_i = \Delta u_{i+1} - \Delta u_i = u_{i+2} - 2u_{i+1} + u_i .$$

一般に k 次定差は

(2) $$\Delta_K u_i = {}_K C_0 u_{i+K} - {}_K C_1 u_{i+K-1} + {}_K C_2 u_{i+K-2} - \cdots\cdots + (-1)^K_K C_K u_i .$$

こゝに ${}_K C_\alpha$ は K 個の内 α を一時にとつた組合せで二項係數である.

$$ {}_K C_\alpha = \frac{K(K-1)(K-2)\cdots\cdots(K-\alpha+1)}{1\cdot 2\cdot 3\cdots\cdots \alpha} .$$

今 u_i が smooth part $m_i = E(u_i)$ と random element $x_i = W_i$ の和で表されるのであるから

(3) $$u_i = m_i + x_i .$$

右邊の二つの要素の間には相關々係がないものとする.

(4) $$E(m_i x_j) = 0 \quad i, j = 1, 2, \cdots, N .$$

random part の element は互に相關々係がない.

(5) $$E(x_i x_j) = 0 \quad i, j = 1, 2, \cdots, N, i \neq j .$$

x の數學的期望値は零で, その population variance は σ である.

(6) $$E(x_i) = 0 \qquad i = 1, 2, \cdots, N ,$$

(7) $$E(x_i^2) = \sigma^2 = \mu_2^{(0)} \quad i = 1, 2, \cdots, N .$$

x の K 次定差は (6) 及 x_i の線形なること (2) から

(8) $$E(\Delta^K x_i) = 0$$

定差の平方の數學的期望値は

(9) $$\begin{aligned} E(\Delta^K x_i)^2 = \mu_2^{(K)} &= E({}_K C_0 \cdot x_{i+K} - {}_K C_1 \cdot x_{i+K-1} + \cdots\cdots)^2 \\ &= E({}_K C_0^2 \cdot x_{i+K}^2 - {}_K C_1^2 \cdot x_{i+K-1} + \cdots - 2{}_K C_0 {}_K C_1 \cdot x_{i+K} \cdot x_{i+K-1} + \cdots) \\ &= {}_{2K} C_K E(x_i^2) = {}_{2K} C_K \mu_2^{(0)} = {}_{2K} C_K \cdot \sigma^2 \end{aligned}$$

である. 何となれば (5) 及

(10) $${}_K C_K^2 + {}_K C_{K-1}^2 + \cdots\cdots + {}_K C_1^2 + {}_K C_0^2 = {}_{2K} C_K$$

であるからである. Random element の K 次定差の population variance を $\sigma_K^2 = \mu_2^{(K)} / {}_{2K} C_K$ で表すと

(11) $$\sigma_0^2 = \sigma_1^2 = \cdots\cdots = \sigma^2$$

が得られる.

若し u の K_0 次定差で non-random element, 又は數學的期望値が消去出來たとすれば近似的に

(12) $$V_{K_0} = V_{K_0+1} = V_{K_0+2} .$$

こゝに V_K は σ_K^2 の實驗的近似値である.

即ち K_0 次以上の次數の u_i の定差に於ては random element のみが殘り從つて (12) が成立する. このことを我々が統計的に試驗しようと言ふのである.

大標本に於ては標準偏差を正確に計算出來るときは, 二つの重要なる試驗方法がある. 一つは凡

ての分布に對して成立する Tchebycheff の不等式である．即ち若し t を正の常數とし x が random variable でその variance が σ^2 なるものとすれば，$|x-E(x)|\leqq t\sigma$ なることの確率 P は

$$P \geqq 1-\frac{1}{t^2}$$

である．若し x が正常分布の random variable ならば，その確率法則を

$$p(x)\,dx = \frac{1}{\sigma\sqrt{\pi}}\exp\left(-\frac{[x-E(x)]^2}{2\sigma^2}\right)dx$$

と考へるとき，$|t|=|[x-E(x)]/\sigma|<t_0$ の確率は次の積分で與へられる．

$$P=\int_{-t_0}^{t}\frac{1}{\sqrt{2\pi}}\exp\left(\frac{-t^2}{2}\right)dt\,.$$

この積分の數値は表に出來上つてゐることは衆知のことである（この試驗の方法は大標本に於てのみ正確に計算出來る眞の population variance を知ることが必要なることに注意せねばならぬ．

相次ぐ定差系列の間の差に關して試驗を行ふには，V_K-V_{K+1} なる差の標準偏差を見出さねばならぬ．その平方は

$$e_K^2 = E(V_K-V_{K+1})^2 \tag{13}$$

で，こゝに V_K は次の標本定差で眞の値 σ_K^2 に對する實驗的近似である．

元系列 u の平均値は

$$\bar{u}=\frac{\sum_1^N u_i}{N}, \tag{14}$$

又 K 次差の α 乘の和は

$$S_\alpha^{(K)}=\sum_{i=1}^{N-K}(\varDelta^K u_i)^\alpha \qquad K=0,1,2,\cdots \tag{15}$$

である．こゝに元系列は零次差である．

Tchuprow, Anderson, Fisher 及び Zaycoff の式から，元系列の population variance σ^2 の unbiased estimate は

$$V_0=\frac{\sum_1^N (u_i-\bar{u})^2}{N-1}=\frac{S_2^{(0)}-N\bar{u}^2}{N-1} \tag{16}$$

である．

高次差の variances σ_K^2 の unbiased estimate は (9) から，x の眞の平均又は數學的期望値が (6) 式から零なることを考へると

$$V_K=\frac{\sum_{i=1}^{N-K}(\varDelta^K u_i)^2}{(N-K)_{2K}C_K}S_2^{(K)}A_{KN},\quad A_{KN}=\frac{1}{(N-K)_{2K}C_K} \tag{17}$$

である．

同樣にして元系列の kurtosis に對する estimate は $\mu_4'-3(\sigma_0^2)^2$ の近似評價で

$$\begin{aligned} D_0 &= \frac{1}{1-\frac{4}{N}+\frac{6}{N^2}-\frac{3}{N^3}}\left\{\frac{1}{N}\sum_{i=1}^{N}(u_i-\bar{u})^4-3\left[\frac{(N-1)}{N}\right]^2V_0^2\right\} \\ &= (m_4-E_N V_0^2)F_N; \\ m_4 &= (S_4^{(0)}/N)-(4S_3^{(0)}\bar{u}/N)+(6S_2^{(0)}\bar{u}^2/N)-3\bar{u}^4; \end{aligned} \tag{18}$$

$$E_N = 3\left(\frac{N-1}{N}\right)^2; \quad F_N = \frac{1}{1-\frac{4}{N}+\frac{6}{N^2}-\frac{3}{N^3}}$$

であり，定差系列に對しては $\mu_4^{(K)}-3(\sigma_K^2)$ は次式で評價されるであらう．

$$\text{(19)} \qquad D_K = \frac{1}{\sum_{i=0}^{K}({}_KC_i)^4}\left\{\frac{1}{(N-K)}\sum_{i=1}^{N-K}(\Delta^K u_i)^4-3({}_{2K}C_K)^2 V_K^2\right\}$$

$$= \left[\frac{S_4^{(K)}}{N-K}-B_K V_K^2\right]C_K; \quad B_K = 3({}_{2K}C_K)^2; \quad C_K = \frac{1}{\sum_{i=0}^{K}({}_KC_i)^4}$$

である．常數 $A_{KN}, E_N, F_N, B_K, C_K$ は Tintner: Variate Difference Method, 1940 の第 10, 18, 19, 14 及 16 表で與へられてゐる．

Anderson は相次ぐ定差系列の二つの variance の差の標準偏差の自乘を次の形で與へてゐる．

$$\text{(20)} \qquad e_K^2 = E(V_{K+1}-\sigma^2)^2+E(V_K-\sigma^2)^2-2E(V_K-\sigma^2)(V_{K+1}-\sigma^2);$$

$$E(V_K-\sigma^2)^2 = \frac{\mu_4-3(\sigma^2)^2}{N-K}\left[1-\frac{2S_K}{({}_{2K}C_K)^2(N-K)}\right] + \frac{2(\sigma^2)^2}{N-K}\left[\frac{{}_{4K}C_{2K}}{({}_{2K}C_K)^2}-\frac{K}{2(N-K)}\right];$$

$$S_{(K)} = \sum_{i=0}^{K-1}({}_KC_i\cdot{}_KC_{i+1})^2+2\sum_{i=0}^{K-2}(({}_KC_i\cdot{}_KC_{i+2})^2+\cdots\cdots+{}_KC_KC_0\cdot{}_KC_K))^2$$

及び

$$\text{(21)} \qquad E(V_K-\sigma^2)(V_{K+1}-\sigma^2) = \frac{\mu_4-3(\sigma^2)^2}{N-K}\left[1-\frac{2S'_{(K)}}{{}_{2K}C_K\cdot{}_{2K+2}C_{K+1}(N-K-1)}\right]$$

$$+\frac{2(\sigma^2)^2}{N-K}\cdot\left[\frac{{}_{4K+1}C_{2K}}{{}_{2K}C_K\cdot{}_{2K+2}C_{K+1}}\cdot\frac{2N-2K-1}{N-K-1}-\frac{K+1}{2(N-K-1)}\right];$$

$$S'_{(K)} = \sum_{i=0}^{K-1}({}_KC_i\cdot{}_{K+1}C_{i+2})^2+2\sum_{i=0}^{K-2}(({}_KC_i\cdot{}_{K+1}C_{i+3})^2+\cdots+{}_KC_KC_0\cdot{}_{K+1}C_{K+1})^2.$$

Anderson 及び Zaycoff は $K=0,1,2,\cdots,10$ に對する此等の面倒な式を計算した．我々は眞の population variance σ^2 の近似値として V_K を用ひ，又眞の population の kurtosis $\mu_4-3(\sigma^2)^2$ の best empirical な近似として D_K を用ふる．又標準偏差の實驗的な近似としては

$$\text{(22)} \qquad e_K = \frac{V_K}{Q_K}$$

を採用し，こゝに

$$Q_K = \frac{H_{KN}}{\sqrt{1+J_{KN}G_K}}$$

$$H_{KN} = \sqrt{\frac{(N-K)(N-K-1)}{b'_K N+b''_K-\frac{N}{(N-K)(N-K-1)}}},$$

$$G_K = D_K/V_K^2$$

$$J_{KN} = \frac{b_K-\left(C_K+\frac{C'_K}{N}\right)\frac{N}{(N-K)(N-K-1)}}{b'_K N+b''_K-\frac{N}{(N-K)(N-K-1)}}.$$

この b, b', b'', C 及び C' は他の量の計算の爲にも必要であるが，Zaycoff が計算したものがあり Tintner の第 39 表に轉載してある．

若し分布が正常なるか，或は正常に近い場合には kurtosis は標準偏差と比べれば無視しうる故 G_K は小さくなる．それで我々は次の近似を使ふことが出來る．

$$e_K^0 = \frac{V_K}{H_{KN}}. \tag{23}$$

Anderson は大なる N に對し $K \geqq 6$ なるとき次の近似式を與へてゐる．

$$(e_K^0)^2 = \frac{3(K+1)V_K^2\sqrt{2\pi K}}{2(2K+1)^3(N-K-1)}. \tag{24}$$

正確な式 (22) 又は近似式 (23) に基づき二つの規準 R_K と R_K^0 を作る．これは，相次ぐ定差の variance の差と，その標準誤差の比を表すものである．

$$R_K = \frac{V_K - V_{K+1}}{e_K} = \frac{(V_K - V_{K+1})Q_K}{V_K} \tag{25}$$

$$R_{K0} = \frac{V_K - V_{K+1}}{e_K^0} = \frac{(V_K - V_{K+1})H_{KN}}{V_K}. \tag{26}$$

此等の値は，高次定差の variance が多少とも等しく，從つて non-random 分子が即ち數學的期望値が消去されることの確率に關する判定に用ひられる．この問題は fiducial limit の見地から決定さるべきものである．一般に $|R_K|$ 又は $|R_K^0|$ が 3 より小ならば充分である．

8. 定差方程式の解による補整[5] 以上に於て補整の規準とすべき多項式の次數が決定したのであるから，次に第二の問題に移るのであるが，こゝでは Whittaker の思想を基礎として問題を進めてゆきたい．

今 σ を小さい常數値とするとき，次の假定を考へる：

"u_α は $E(u_\alpha)$ と $E(u_\alpha)+\sigma$ の間にあり，$u_{\alpha+1}$ は $E(u_{\alpha+1})$ と $E(u_{\alpha+1})+\sigma$ の間にあり，一般に $u_{\alpha+i}$ は $E(u_{\alpha+i})$ と $E(u_{\alpha+i})+\sigma$ の間にあり，最後に u_ω が $E(u_\omega)$ と $E(u_\omega)+\sigma$ の間にある．"

この假定を "假定 H" と呼ぶ．統計より $u_\alpha, u_{\alpha+1}, \cdots, u_\omega$ を得る以前に於ては，K 次差の平方の和

$$S = \sum_\alpha^\omega (\Delta^K E(u_i))^2$$

の小なること，即ち滑かさ以外に假定 H の確率に對する手掛りはないのである．

從つて，正常法則に類似に假定 H の a priori な確率が

$$Ce^{-\lambda^2 S}\sigma^{\omega-\alpha+1}$$

であると考へることが出來る．

次に假定 H が眞なりとの想定の下に，統計より得られた測度が $u_\alpha, u_{\alpha+1}, \cdots, u_\omega$ なることの a priori の確率を考へる．

統計より得られた量 u_α の眞の値はこの假定の下では $E(u_\alpha)$ であるから，u_α と $u_\alpha+\sigma$ の間の値が實際に統計より得られることの確率は（誤差の正常法則を假定すれば）

$$\frac{h_\alpha}{\sqrt{\pi}} e^{-h^2_\alpha(u_\alpha - E(u_\alpha))^2}\sigma.$$

こゝに h_α は統計量の精確度を表す常數である．同様にして，第二の量 $E(u_{\alpha+1})$ に對して統計より得られる値が $u_{\alpha+1}$ と $u_{\alpha+1}+\sigma$ の間にあることの確率は

(A)
$$\frac{h_{\alpha+1}}{\sqrt{\pi}} e^{-h^2_{\alpha+1}(u_{\alpha+1}-E(u_{\alpha+1}))^2} \sigma$$

で，こゝに h_2 はこの統計量の精確度の測定である．かくして假定 H が眞なりとの想定の上で，第一の統計量が u_α と $u_\alpha+\sigma$ の間に，第二の統計量が $u_{\alpha+1}$ と $u_{\alpha+1}+\sigma$ の間に，以下同様にして ω 歳に對する統計量が u_ω と $u_\omega+\sigma$ の間にあることの確率は，a priori には

(B)
$$\frac{h_\alpha h_{\alpha+1} \cdots h_\omega}{(\sqrt{\pi})^{\omega-\alpha+1}} e^{-F} \sigma^{\omega-\alpha+1}$$

で，こゝに F は

$$F = h_\alpha^2(u_\alpha - E(u_\alpha))^2 + h_{\alpha+1}^2(u_{\alpha+1} - E(u_{\alpha+1}))^2 + \cdots + h_\omega^2(u_\omega - E(u_\omega))^2$$

である．

S 及 F なる和によつて補整値の滑かさと，未補整値への補整値の忠實さを數字的に夫々表すことができる．

次に我々は Inductive Probability の定理を應用する．或る統計量は一定數の假定の內何れによつても說明しうるが，この假定の內眞なるものは唯一つしかないと考へ，更に S 番目の假定は統計がとられる前は p_S であることが，何等かの知識から敎へられており，且この假定が眞なるものであるとの假定の下には統計量の確率が P_S であると考へる．すると統計量を考慮に入れると，S 番目の假定の確率は

$$\frac{p_S P_S}{\sum p_S P_S}$$

で，こゝに $\sum$ は凡ての假定に關しての和を示す．このことから，統計がとられる前の最も確からしい假定は p_S が最大となるものであるが，統計がとられた後の最も確からしい假定は，積 $P_S p_S$ が最大となる如き假定である．この定理を今の問題に應用し且 (A) 及 (B) を組合せれば

$$\frac{c h_\alpha h_{\alpha+1} \cdots h_\omega}{(\sqrt{\pi})^{\omega-\alpha+1}} e^{-\lambda^2 S - F} \sigma^{2(\omega-\alpha+1)}$$

が最大となる如き假定が最も確からしい：即ち $E(u_i)$ なる量の最も確からしい組 $E(u_\alpha), E(u_{\alpha+1}), \cdots, E(u_\omega)$ は

(C)
$$\lambda^2 S + F$$

が最小となるものである．

(C) を $E(u_\alpha), E(u_{\alpha+1}), \cdots\cdots, E(u_\omega)$ で微分してこの最小値を求めるのである．微分を實行すれば

(27)
$$\begin{cases} 0 = h_\alpha^2 u_\alpha - h_\alpha^2 E(u_\alpha) - \lambda^2 \Delta^K E(u_\alpha), \\ 0 = h_{\alpha+1}^2 u_{\alpha+1} - h_{\alpha+1}^2 E(u_{\alpha+1}) + K\lambda^2 \Delta^K E(u_{\alpha+1}), \\ 0 = h_{\alpha+2}^2 u_{\alpha+2} - h_{\alpha+2}^2 E(u_{\alpha+2}) - {}_K C_{K-2} \lambda^2 \Delta^K E(u_\alpha) + {}_K C_{K-1} \lambda^2 \Delta^K E(u_{\alpha+1}) - {}_K C_K \lambda^2 \Delta^3 E(u_{\alpha+2}) \end{cases}$$

である．今 $\begin{Bmatrix} u_{\omega+1}, u_{\omega+2}, \cdots, u_{\omega+K} \\ u_{\alpha-1}, u_{\alpha-2}, \cdots, u_{\alpha-K} \end{Bmatrix}$ が K 次定差が零となる系列を作るとすれば，上の關係は凡て次の形で表される．こゝに $\epsilon_i = \frac{h_i^2}{\lambda^2}$

(28)
$$\epsilon_i E(u_x) - \Delta^{2K} E(u_{x-3}) = \epsilon_i E_x$$

こゝで統計量の精確度を凡て等しいと假定すれば，$\epsilon_j=\epsilon$ であつて，

$$\epsilon E(u_x)-\Delta^{2K}E(u_{x-3})=\epsilon E_x \tag{29}$$

を得る．この式によつて moments は保存される．

9. 補整の實際　我々は前二節で與へた方法により，規準と爲すべき多項式を決定し，次で (28) の方程式を解くことによつて補整値を求めることを，實際の例について試みたいと思ふ．その爲に商工省日本經驗件數男子三年截斷表の粗成死亡率を用ひた結果は次の通である．

多項式の次數の決定　この爲に使用すべき死亡率の範圍は，死亡率に於ては時系列における如き個々の觀察點における同質性 (Homogeneity) が考へられず，觀察値の精度が或る程度まで延人員數を規準にとつて表示しうることを考慮すれば，自ら限定せられざるを得ない．幸にして最も材料が安定せる部分は最も延人員數が一樣に集中されてゐるし，この部分の精確度が一樣に大で安定してゐると考へられ，强力なる補整がこの部分に向けらるべきとこは論を俟たないのである．多項式の次數を決定するのはこの部分であるものとしたい．第6節の方法を前記の表の 20 歳より 29 歳までとつて之に適用すれば第1表の如くなる．

第　1　表

$N=40$ （20歳より 59歳まで）

k	V_K	D_K	G_K	Q_K	R_K
1	77265	+1.00303	+1.68017	12.185	+ 10.701
2	09409	− 02474	−2.79548	17.259	+ 3.997
3	07230	− 00647	−1.23946	19.539	+ 711
4	06967	− 00553	−1.14021	21.587	+ 1.097
5	06613	− 00575	−1.31579	23.192	+ 1.343
6	06230	− 00704	−1.81443	21.445	+ 1.394
7	05825	− 00868	−2.56047	26.162	+ 1.856
8	05412	− 00869	−2.97603	27.059	+ 2.005
9	05011	− 00785	−3.12749	27.378	− 585
10	05118	− 01190	−4.54198	29.494	

こゝで定められた多項式とは，その多項式で全範圍を最小自乘法等で fit せんとする意味ではなく，その次數の定差が安定することの意味である．即ち定差の和を最小ならしめ且粗成死亡率とも近密性を保持せしめんとするものである．その爲には Tintner の如く[6] Sheppard の方法によることもよからうが，彼も言つてゐる樣に，定差の安定する次數と使用すべき式の次數との間に關係がなく，Anderson の意見のまゝに經驗的に之を決定してゐるのである．こゝで Whittaker-Henderson の思想と方法を用ふれば，この間の關聯が思想的につくのである．

Henderson の type B の補整[7]を行つた結果を第2表に擧げよう．こゝに Σ_1, Σ_2 は夫々今行つた補整値と公式表の補整値に關しての總和である．こゝに (28) における h^2 としては經過契約人員（延人員）を以て之に當てた．計算の詳細は省略するが，Henderson の論文及教科書[8]にある方法を一部不用と思はれる箇所を省き，其他はそのまゝ踏襲したのである．結果を公式の死亡率[9]と比較すれば，滑らかさに於てはこゝに擧げたものの方がはるかに勝り，近密度に於ても自乘の和と云ふ規準に於ては新しい方法の方が勝れてゐるが，實際死亡數との符號をつけた差の累積に於ては公式の表の方がはるかに勝つてゐる．最後のものを最高の規準にとらんとするときは，今迄のべて

來た思想のみでは不充分であるが，既に多くの假定の上に立つて $E(u_x)$ を求めて來たのであるから統一した理論を，かくの如き新しい假定を導入した上で立てることは相當の困難が伴ふのではあるまいか．

第 2 表

x	$E(u_x)$	u_x	x	$E(u_x)$	u_x	x	$E(u_x)$	u_x
13	3176	4319	41	10278	10575	69	84161	80767
14	4463	3802	42	10817	10553	70	91790	93568
15	5858	6183	43	11524	11495	71	100051	102217
16	7303	6521	44	12342	12494	72	108737	109225
17	8693	8751	45	13157	13101	73	117793	116243
18	9838	10036	46	13993	14172	74	127262	127120
19	10571	10776	47	14923	14549	75	137206	140507
20	10839	10900	48	16070	16257	76	147707	142067
21	10727	10529	49	17313	17385	77	158875	156339
22	10397	10698	50	18615	18640	78	170751	175513
23	9996	10340	51	20038	19906	79	183333	185816
24	9677	9459	52	21662	21630	80	196646	198859
25	9422	9702	53	23484	23583	81	210760	217237
26	9103	9264	54	25455	25405	82	225770	183365
27	8714	8625	55	27560	27740	83	241782	299094
28	8373	8336	56	29808	29486	84	258862	258621
29	8164	7966	57	32264	32660	85	277107	210526
30	8074	8215	58	34944	34704	86	296617	311475
31	7979	8556	59	37940	38172	87	317466	242424
32	7854	7968	60	41289	40198	88	339711	250000
33	7804	7662	61	44968	46219	89	363383	333333
34	7981	7738	62	48801	48408	90	388478	571429
35	8342	8496	63	52785	53319	91	415000	1000000
36	8633	8948	64	56965	55686			
37	8774	8510	65	61449	61969			
38	8985	8931	66	66257	67857			
39	9332	9434	67	71490	70581			
40	9768	9633	68	77400	77509			

$\sum_1(\Delta^3)^2=792000$

$\sum_2(\Delta^3)^2=3174300$

$\sum_1(q-q')^2=153066$

$\sum_2(q-q')^2=210416$

尚以上の規準の他に，豫定と實際の死亡數の差の累差の符號の變る回數，特定五歳群に關しての豫定と實際が特によく fit してゐること等を規準にとることがある．此等は凡て全體としての近密さの爲に，部分的な近密さが著しく疎外されることを避ける爲であつて，此點あまりに部分的な近密さを强調すると，逆に全體としての近密さが破壞せられるばかりでなく，滑らかさの點に於て非常なる不合理を生ぜしめることになり易いから注意しなければならないと思ふ．此等の規準は結果的に導入せられたものであるから，その理論的な研究は今後に俟たるべきものであり，只今の所では初めに扱つた二つの規準が最も合理的なものである．

最後に附言しておきたいことは，緒言で論じた如く，補整は一聯の統計推理中の一過程であるから，次の過程によつて，行ふべき補整の方法を決定すべきことであり，今迄のべた方法は，出來る

だけ数學的期望値に近い近似値を求めんとした方法であり，Henderson を中心として米國に於て最近特に進歩した方法である.[10] 之に對して英國に於ては連續な導函數を有し，次の計算に於て解析的方法が取りうる樣に，解析式による補整に就て最近迄論ぜられてゐる現狀にある.[11] 尙此の他にも必要なだけの導函數の連續性を有せしめることも考へられてゐる.[12] 此等を上に言つた樣な見方から考へてみるのも面白いと思つて考へついたまゝを述べてみたのである.

註

(1) Walford: The Insurance Cyclopaedia 1871-78.

(2) F. Vinci: Manuale di statistica, Volume primo. pag. 175. Bologna. 1936.

(3) Kenny: Journal of the American Statistical Association, vol. XXVI, March, 1931. Supplement. p. 36.

(4) Tintner: Variate Difference Methods, 1940, Appendix II.

(5) Whittaker and Robinson: Calculus of Observation. p. 303.

(6) Tintner, op. cit. p. 100.

(7) Henderson: Further Remarks on Graduation, T. A. S. A. XXVI. p. 52. 1924.

(8) Henderson: Mathematical Theory of Graduation, 1934.

(9) 日本商工省經理生命表，明治四十五年－昭和二年.

(10) Henderson: A New Method of Graduation Transactions of Actuarial Society of America. XXV. p. 29. 1923.
Henderson: 前揭 (7).
Henderson: Actuarial Studies No. 4. 初版：Theory of Graduation, 1919；改訂增補版，前揭 (8).
Spoerl: Henderson's Graduation Formula B., T. A. S. A. XXXII. Part I. No. 85. pp. 60-72.
Spoerl: The Whittaker-Henderson Graduation Formula A., T. A. S. A. Vol. XX No. 98.

(11) G. J. Lidstone: Some Properties of Makeham's Second Law of Mortality and the Double and Triple Geometric Laws. Journal of Institute of Actuaries Vol. LXVIII. Part IV. 1937.
E. Olifiers: Graduation by Double Geometric Laws supplementary to Makeham's Basic Curves. With an Applicasion to the Graduation of the A. 1924-29 (Ultimate) Tables. J. I. A. Vol. LXVIII. Part IV. 1937.
H. O'Brien: Application of Uniform Quarity etc. J.I.A. Vol. LXVIII. Part IV. 1937. pp. 541-550.

(12) Hans Cl. Nybölle: On pseudo-analytical Graduation J.I.A. Vol. LXVI. Part. No. 314. pp. 63-87.