資　　料

補整法覚え書*

藤　田　輝　昭**

最近では曲線の平滑化といえば，Spline関数によるそれを意味するかのようである．たとえば，穂坂教授[1]の (51) 式も，非再生型――与えられた点 (x_i, y_i) を必ずしも通らない――の natural spline 曲線[2]にほかならない．

しかし，spline 曲線に関する Schoenberg の最初の論文[3] (1946) が，接触補間公式の精密化と Greville の予想[4] (1944) の解決を目的にしたものであり，その背景には生保アクチュアリー (actuary) による補間および補整公式に関する半世紀以上の研究の積み重ねがあったことは意外に知られていない．その一因は，日本では補整法の教科書が比較的狭い範囲の中でしか circulate していないこと，その上 1940 年代以降の目ぼしい業績がまとまったテキストの形に整理されていない（英語でも日本語でも）ことにあるだろう．

一方，アクチュアリーの側は，Greville の有名な二論文[5),6)] (1949, 1954) によって，補整法の重要な公式が完全に整理されてしまって以来，新しい展開への熱意をすっかり失ったように見える．したがって，Schoenberg の前掲論文が，古典的な補整法から spline 曲線による平滑化理論への橋渡しの役をしているとも考えられる．

本稿では，次第に影が薄くなりつつあるこの“橋”の再構成を試みる．その目的は古典的補整法の遺産の中から，有用な部分を今日のアナリストやプログラマーが取り出せるように，その理論をこれらの人々に，accessible な形で提示することにある．

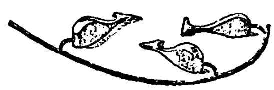
1. 補整と補間

死亡率の測定と生命表の作成は，古くからアクチュアリーの重要な職務のひとつであったし，最近では企業年金の設計にからんで，退職などによる脱退率の計算も必要になっている．観測データから得られるナマの死亡率（粗死亡率）や，粗脱退率が不規則な変動を含んでいて，そのままでは使えないことは一見して明らかなので，これを平滑化することが，まず必要になる．これを補整 (graduation) という．

すべての年令に共通に成り立つ“死亡法則”が見つかれば，問題はそのパラメータを決定することに還元される．そのような法則を見い出すことは，18 世紀から 19 世紀にかけて，いろいろと試みられた．中でも，今日まで生き残っているのが，Gompertz 曲線 (1825) であるが，全年令に単一の曲線をあてはめるのが無理なことはすぐにわかるので，今日では，データ数の少ない高年令に用いられるに過ぎない．

そうした高年令を除いた部分では，何らかの機械的な計算，またはグラフによる平滑化が必要になる．ここでグラフを使う方法が，Sprague (1886) によって強く支持されたことは，浅谷[7]も紹介しているとおりだが，その後も意外に根強い人気を保ってきた．後述の Whittaker-Henderson の補整法も，グラフ法への一近似と考えられ，“mechanico-graphic method”などと呼ばれた．1937 年には TASA*, **38**, p. 523 に，グラフ法のための道具として，**第 1 図**のような器具が紹介されている．――これが Spline にほかならない (Greville[2]のカバー表紙参照)．

actuaries.
There is here given a cut from the catalogue of Keuffel & Esser Co., New York City, showing a spline and spline weights (3¾ths & 8 lbs.).



Splines are made of xylonite, and also of wood, with a groove along one edge in which the finger of the weight fits. The weights

第 1 図

計算による補整法については，1940 年代になると，

* Notes on Graduation, by TERUAKI FUJITA (Mitsui Mutual Life)

** 三井生命保険相互会社

* 雑誌名の略称については文献表参照．

ありとあらゆる補整公式が出揃ってしまった．それらを分類すると，次のように大別できる．

（1）　連続的補間公式によるもの．

（2）　離散的補間公式によるもの．

（3）　定差方程式の解として得られるもの．

（1）に属するのが Sprague (1880) に始まる接触補間公式，（2）に属するのが Woolhouse (1870) 以来の，いわゆる linear compounding formula で，それぞれ Greville の前掲論文[5),6)]によって完全に整理された．（3）は Whittaker-Henderson の補整法にほかならない．

2.　接触補間公式

与えられた等間隔データを $u_1, u_2, \cdots\cdots, u_n$ とする．接触補間公式による補整値 v_x は

i）区間 $(-\infty, \infty)$ において，x の高々 $m(\geqq 1)$ 次の多項式を接続したものであり

ii）v_x は区間 $(-\infty, \infty)$ で $k(\geqq 1)$ 次の導関数まで連続

という条件を満たさなければならない．

多項式の接続の場所は $x=1, 2, \cdots\cdots, n$，すなわち，区間 $(i, i+1)$ の端点である場合と，$x=1^1/_2, 2^1/_2, \cdots\cdots$ という中間の点である場合が考えられる．前者の場合を endpoint 式，後者を midpoint 式という．いずれの場合も k が接続（接触）の次数と呼ばれる．

"補間" 公式であるための条件は，それらの多項式の係数が，u_i の有限個の定差の1次結合としてあらわされ，その形が各区間で一様であることである．u_i の定差は原系列 $u_1, u_2, \cdots\cdots, u_n$ の1次式としてあらわすことができるから，これを整理すると，結局，v_x は次の形になる：

$$v_x=\sum_{i=-\infty}^{\infty} L(x-i)u_i \tag{1}$$

ただし……，u_{-1}, u_0 および u_{n+1}, u_{n+2}, ……は必要に応じて補充するものとし，$L(t)$ は t の多項式である（t の全区間で同一形式とは限らない）．通常 $L(t)$ は $t=0$ を軸として対称：

$$L(t)=L(-t) \tag{2}$$

であり，かつ $|t|$ のある値以上でゼロである．加えて，u_i が一定値なら v_x も同じ値でなければならないから

$$\sum_{i=-\infty}^{\infty} L(x-i)=1 \tag{3}$$

（1）により，v_x の値を決めるのに参加する u の項数，すなわち，$L(x-i)\neq 0$ であるような u_i の数の最大値が決まる．この項数を，この補間公式の"項数"という．

$L(t)$ が t の0以外の整数値について0であり，したがって，（3）から $L(0)=1$ のときには $v_i=u_i$ $(i=1, 2, \cdots, n)$ となり，原系列が完全に再生される．これを再生型の補間公式，そうでない場合を非再生型という．

さらに，u_x が x の高々 r 次の多項式によって与えられたときに $v_x=u_x$ となる場合，公式は"r 次の多項式を再生する"，"r 階の定差まで正しい"，または単に"r 階まで正しい"という．r の最大値が degree of reproduction である．

実用上は（1）のほかに Everett 型の補間公式の形で用いることが多い．この場合[4)] endpoint 式は

$$\left.\begin{array}{l} v_{i+x}=F(1-x)u_i+F(x)u_{i+1} \quad 0\leqq x\leqq 1 \\ F(x)=f_0(x)+f_2(x)\delta^2+f_4(x)\delta^4+\cdots\cdots \end{array}\right\} \tag{4}$$

の形になる．ここで f_0, f_2, ……は x の多項式であり，δ は Sheppard の中央定差の記号とする．

また，midpoint 式は

$$\left.\begin{array}{l} v_{i+x}=u_{i+1/2}+G(x)u_{i+1}-G(1-x)u_i \quad 0\leqq x\leqq 1 \\ G(x)=f_1(x)\delta+f_3(x)\delta^3+\cdots\cdots \end{array}\right\} \tag{5}$$

の形になる．どちらの場合にも，F または G の中に含まれる x の次数が，補間多項式の次数である．

こうして，接触補間公式は次の各要素によって特徴づけられる：

（1）　接触の次数．

（2）　endpoint 式か midpoint 式か．

（3）　再生型か非再生型か．

（4）　補間多項式の次数．

（5）　その項数．

（6）　degree of reproduction.

これらの基準による補間公式の分類と特徴づけの問題は，Greville[6)] によって完全に解決された．これを「Greville の規則」と名づけよう．

補間公式（1）における関数 $L(t)$ を，この公式の basic function と呼ぶ．これに対応して，次のような作用素を考える：

$$G=\int_{-\infty}^{\infty} L(t)E^{-t}dt \tag{6}$$

これを，補間公式（1）の特性作用素 (characteristic operator) という（E はいうまでもなく，$Ef(x)=f(x+1)$ によって定義される前進作用素）．

一般に（6）の形の作用素を“連続作用素”という．すなわち，その一般形は

$$K=\int_{-\infty}^{\infty} f(t)E^{-t}dt \qquad (7)$$

ここで，f はある有界な区間〔a, b〕の外部で 0 である．関数 f を K の basic function と呼ぶ．

（1）の補間作用素は一般に

$$J=\sum_j c_j E^{-\lambda_j} \qquad (8)$$

の形にあらわすことができる．この形の作用素を“離散作用素”と呼ぶ．

離散作用素（8）の“span”は

$$\mathrm{span}(J)=\max_j \lambda_j-\min_j \lambda_j \qquad (9)$$

によって定義される．連続作用素の場合には，$f(t)$ が non-zero であるような最小の区間を〔a, b〕とするとき，$b-a$ がその span である．

連続作用素 (7) の“trace” $t(K)$ は，次の式によって定義される離散作用素である：

$$t(K)=\sum_{i=-\infty}^{\infty} f(i)E^{-i} \qquad (10)$$

特に，Gが（6）のような特性作用素であれば

$$t(G)u_i=v_i \qquad (11)$$

以上の準備のもとで，Greville の規則を次のように述べることができる．

〔1〕 xについて連続な補間公式が，r 階まで正しいための必要十分条件は，次の（a），（b）が同時に成り立つことである：

（a） 特性作用素 G が $M^{r+1}H$ の形であること．ここで M は次のような連続作用素：

$$M=\int_{-1/2}^{1/2} E^t dt$$

Hは次の形の作用素である：

$$H=J_{-q+r}D^{-q+r}+J_{-q+r+1}D^{-q+r+1}+\cdots+J_{r-p-1}D^{r-p-1} \qquad (12)$$

ただし，Jは離散作用素，Dは微分作用素．

（b） 上の H を D のベキ級数として展開したとき，D^r までの各項が M^{-r-1} を同様に展開したときの各項と一致する．

〔2〕 補間公式の項数を s，作用素 H の span を hとすると

$$s=h+r+1 \qquad (13)$$

〔3〕（a） q次の補間多項式がr階まで正しく，少なくともp次の接触をするためには，Hが(12)の形であり，かつ $J_{-q+r}\neq 0$ であることが必要十分．

（b） 上の場合

$$q\geq r,\quad q\geq p+1 \qquad (14)$$

（c） (12)の右辺をDのベキに展開したとき，負のベキに対応する項は互いに相殺して，ゼロにならなければならない．

〔4〕（a） rが偶数の場合，この補間公式が endpoint 式であるための必要十分条件は，(12)の各Jが次の形であること：

$$a_0\mu+a_1\delta+a_2\mu\delta^2+a_3\delta^3+\cdots\cdots \qquad (15)$$

ただし，μ は通常の“平均値”作用素：

$$\mu=\frac{1}{2}(E^{1/2}+E^{-1/2})$$

(15) の形の離散作用素を Bessel 型という．

また，midpoint 式であるための必要十分条件は，各Jが次の形であること：

$$a_0+a_1\mu\delta+a_2\delta^2+a_3\mu\delta^3+\cdots\cdots \qquad (16)$$

この形の離散作用素を Stirling 型という．

（b） rが奇数の場合には，（a）における Bessel 型と Stirling 型が逆になる．

〔5〕（a） 特性作用素の trace $t(G)$ は Stirling 型で，$a_0=1$，かつ δ の1次からr次までの項がない．対称な補間公式の場合にはさらに，δ の偶数次の項しかあらわれない．

（b） 対称な補間公式でsが奇数なら，$t(G)$ の span は $s-1$ をこえない．

（c） 対称な補間公式でsが偶数，$p\geq 0$ なら，$t(G)$ の span は $s-2$ をこえない．

（d） 再生型であるための必要十分条件は $t(G)=1$．

特性作用素と補間公式(1)，さらに（4），（5）との間の変換は容易にできる．後者が与えられたときには，特性作用素によって，その性格を分析することができるし，逆に望ましい特徴を持つ公式を簡単に作ることもできる．その実例については Greville[6] 参照．

3. Linear Compounding

この種の公式も（1）の形に書くことができる．ただし，この場合には $L(x-j)$ が x の離散的な値，通常は区間〔$i, i+1$〕を m 等分した点に対してのみ定義されていることである．したがって，Lは数式ではなく係数表の形で与えられる．

このため，m等分した間隔を1と読み直した方が処理が簡単になる．事実，死亡表の補整の場合には，5

才間隔で与えられた原データから，１才間隔の補整値を求めるのが普通である．特に，$m=1$ の場合には，原データの平滑化のみが行なわれるので，これを狭義の補整公式ということもある．いずれにせよ，(１)は次のように書き替えられる：

$$v_i=\sum_{k=-\infty}^{\infty} L_{i-km}u_{km} \tag{17}$$

もちろん，$|k|$ の十分大きな値については $L_{i-km}=0$ とし，また，定数系列の再生を当然の条件と考えれば

$$\sum_{k=-\infty}^{\infty} L_{i-km}=1 \tag{18}$$

u_{km} は m 項間隔で与えられた値だが，これとは別に系列 $U_x(x=\cdots,0,1,2,\cdots\cdots)$ が与えられたとき

$$V_i=\frac{1}{m}\sum_{t=-\infty}^{\infty} L_{i-t}U_t \tag{19}$$

は，(18)により狭義の補整公式になっている．これを“補間公式 (17) に対応する補整公式”と呼ぶ．

作用素Gを形式的に

$$G=\sum_{t=-\infty}^{\infty} L_tE^{-t} \tag{20}$$

によって定義すると，(19)は

$$V_i=\frac{1}{m}GU_i \tag{21}$$

と書ける．このような形の G を linear compound operator という．

連続補間公式の場合と同様に，u_x が高々 r 次の多項式によって与えられたとき，x の与えられた値について，$v_x=u_x$ が得られるならば，補間公式 (17) は r 階まで正しいという．

これよりゆるい条件として，$u_x=P(x)$，P は高々 r 次の多項式，のときに

$$v_x=P(x)+(\text{高々 } r-1 \text{ 次の多項式})$$

が得られるならば，補間公式 (17) は“次数rを保存する”という (Schoenberg[3] p. 61)．

次の二つの事実が，本質的には Vaughan[8] によって述べられたので，Greville[5]はこれを「Vaughan の原理」と名づけた．

〔１〕 離散補間公式 (17) が次数rを保存するための必要十分条件は，対応する補整作用素 (20)が，ある linear compound operator H について

$$G=[m]^{r+1}H \tag{22}$$

の形にあらわされること．いうまでもなく $[m]$ は summation operator：

$$[m]=E^{-\frac{m-1}{2}}+E^{-\frac{m-3}{2}}+\cdots\cdots+E^{\frac{m-1}{2}} \tag{23}$$

である．

〔２〕 補間公式 (17) がr階まで正しいための必要十分条件は

（a） それが次数rを保存し，かつ

（b） 対応する補整公式 (19) がr階まで正しいこと．

この最後の条件は，$V_i-U_i=1/m(G-m)U_i$ が U_i の $r+1$ 階以上の定差だけを含むこと，形式的には $G-m$ の中央定差 δ による展開が δ^{r+1} を因子として含むことである．〔１〕の条件を合わせて考え，Bessel の補間公式から得られる

$$[m]=\frac{E^{m/2}-E^{-m/2}}{\delta}=m+\frac{m(m^2-1)}{24}\delta^2+\frac{m(m^2-1)(m^2-9)}{1920}\delta^4+\cdots\cdots \tag{24}$$

を使うと，rによってHの最初の方の項が決まってしまう．

こうして

（１） 再生型か非再生型か．

（２） degree of reproduction.

（３） 項数．

に加えて

（４） k を与えて $\sum[\Delta^kv_i]^2$ を最小にする．

という平滑化条件を設定すると，Hの形が決まる．

条件（４）は古くから平滑度の目安として用いられてきた．もし，v_x が k 次の多項式であれば，$(\Delta^kv_i)^2$ は一般に non-zero だが，$k-1$ 次以下ならゼロになる．したがって，degree of reproduction をrとすると，$k=r+1$ を取るのが適当だろう (Wolfenden, TASA, **16**, 109-10; Greville, RAIA **34**, 39)．

$k=0$ の場合は最小二乗近似を求める問題にほかならず，Lidstone[9]はチェビシェフの選別直交多項式 (q 関数) によって係数を求める方法を与えている．

Greville[10]はこの方法を $k>1$ の場合に拡張して，最大平滑式を求めた．これらの公式はわが国でもよく使われている．

4. Whittaker-Henderson の補整法

いま述べた Greville の式にしても，linear compounding という枠内での最大平滑化に過ぎない．より自然なアプローチとして，Whittaker は 1919 年に次の量：

$$\epsilon\Sigma(v_i-u_i)^2+\Sigma(\Delta^3v_i)^2 \qquad (\epsilon>0)$$

を最小にするように補整値 v_i を決めるという問題を

考えた[11]. 一般に

$$\epsilon\Sigma(v_i-u_i)^2+\Sigma(\Delta^k v_i)^2 \quad (\epsilon>0) \qquad (25)$$

を最小にする問題は, 定差方程式

$$\epsilon(v_i-u_i)+(-1)^k\delta^{2k}v_i=0 \qquad (26)$$

を解く問題に帰着することがすぐにわかる. この定差方程式の解法は, まず Hendrson[12]によって取り上げられたので, (25)は一般に Whittaker-Henderson の公式Aと呼ばれる.

Henderson[13]はこれを拡張して, i 才人口 E_i を重みとして採用し, 補正項 $\Delta^2 v_1$ を加えた

$$\epsilon\Sigma E_i(v_i-u_i)^2+\Sigma(\Delta^3 v_i)^2+0.1(\Delta^2 v_1)^2$$

という式を考えた. そこで一般に

$$\Sigma h_i(v_i-u_i)^2+\sum_{r=1}^{k}\lambda_r\Sigma(\Delta^r v_i)^2 \quad (h_i, \lambda_r>0) \qquad (27)$$

を, Whittaker-Henderson の公式Bと呼ぶ.

これらの問題は Henderson, Aitken, Spoerl などによって攻撃されたが, いずれにせよ, 戦前の計算手段では決して簡単とはいえなかったに違いない. それでも, 発想の自然さという利点のために, いくつかの生命表で実用化されている.

この補整法が " mechanico-graphic " と形容されたことはすでに触れたが, この mechanico の意味について Wells[14] (1937) が, 次のような解釈を与えている. すなわち, 与えられた点 $(1, u_1)$, ……, (n, u_n) にバネの一端を固定し, 他端に張力を与えたピアノ線かゴム紐をひっかける. この系の平衡状態が近似的に Whittaker-Henderson の公式Bを最小にする.

5. Bayesian Graduation

補整法に対する比較的新しい接近として, ベイジアンの決定理論を応用する試みがある[15]. 観測値 $u_i (i=1, 2, \cdots, n)$ に対応する真の値を W_i とすると, ベイジアンの理論では W_i も確率変数と考えられる. 正規分布の仮定の下で, $W_1, \cdots, W_n$ の事前確率 (主観的確率) にもとづく平均値ベクトルをm, 共分散行列をAとしよう.

死亡率の測定のような形の実験では, W_i の実現値が w_i であったとき, 各観測値 u_i は w_i を平均値とする独立な二項分布に従う. これを正規分布で近似し, その共分散行列をBとする. したがって, Bは対角行列である.

簡単な計算によって, ベイズの定理による事後確率は

$$P_{W|U}(w|u)=K\exp\left\{-\frac{1}{2}[(u-w)'B^{-1}(u-w)+(w-m)'A^{-1}(w-m)]\right\} \qquad (28)$$

の形になる. ただし, Kは定数, uは観測値の列ベクトル, w は w_i の列ベクトルである. この値を最大にするには

$$S=(u-w)'B^{-1}(u-w)+(w-m)'A^{-1}(w-m) \qquad (29)$$

を最小にする w の値を求めればよい. B^{-1} の対角要素を e_i, $A^{-1}=[z_{ij}]$ とすると

$$S=\Sigma e_i(u_i-w_i)^2+\sum_{i,j}z_{ij}(w_i-m_i)(w_j-m_j) \qquad (30)$$

となる.

この式と Whittaker-Henderson の公式 B との類似は明らかだが, それだけでなく Schoenberg[16]による次の結果を思い出させる.

Schoenberg は, Whittaker-Henderson の公式 A において, $\Sigma(\Delta^k v_i)^2$ を積分によっておきかえた式:

$$\Sigma(v_i-u_i)^2+\frac{\epsilon}{(k!)^2}\int_{-\infty}^{\infty}[v_x^{(k)}]^2dx \qquad (31)$$

を最小にする問題を考え, この問題が次の式を最小にする問題に帰着することを示した:

$$\Sigma(v_i-u_i)^2+\epsilon\sum_{i,j}\lambda_{ij}\Delta^k v_i\Delta^k v_j \qquad (32)$$

ここで λ_{ij} はある正値行列 $[l_{ij}]$ の逆行列の元として定義される. $\Delta^k v_i$ は $v_1, \cdots, v_n$ の1次式になるから, (32) も (30) と同様に, v_i の2次形式の形に整理できる.

いうまでもなく (31) [したがって (32)] の解は, spline 曲線によって与えられる (穂坂[1]の (44) 式と同じものであることに注意). 一方, (30) の解は

$$(A^{-1}+B^{-1})v=B^{-1}u+A^{-1}m$$

によって与えられる(vは解の列ベクトル). これは, さらに

$$v-m=u-m-BA^{-1}(v-m)$$

すなわち

$$v_i=u_i-e_i^{-1}z_{ii}(v_i-m_i) \qquad (33)$$

と書くことができる. よく知られているように, 3次の非再生型 spline 補整では

$$v_i=u_i-\lambda_i\delta^2 v_i'' \\ =u_i-6\lambda_i(\delta^2 v_i-v_i'') \qquad (34)$$

だから (Greville[2] p.19, 穂坂[1] (50) 式), 結局ベイジアン補整も spline 補整の近似と考えてよい.

ただし, m, A および B の n^2+2n 個の値は, 補

整者が指定しなければならない．もし，mを既存の死亡表から取り，Aの係数をうまく決めれば，よい補整値が得られるだろう．しかし，Aの係数をどのように決めるかが実は大問題であろう．H.M. Markowitz のポートフォリオ・セレクションでも同様に，主観確率にもとづく共分散行列を与えなければならないところに難点があって，広範囲な実用化がさまたげられてきたことを思い出そう．主観確率を出発点とする決定理論の，これは致命的弱点なのであろうか．

参考文献

1) 穂坂 衛："曲線，曲面の合成および平滑化理論"，情報処理，**10** (1969)，121-131.
2) T. N. E. Greville (ed.): **Theory and Applications of Spline Functions,** Academic Press 1969.
3) I. J. Schoenberg: Contributions to the problem of approximation of equidistant data by analytic functions, Quart. Appl. Math. **4** (1946), Part A, 45-99, Part B, 112-141.
4) T. N. E. Greville: The general theory of osculatory interpolation, TASA, **45** (1944), 202-265.
5) T. N. E. Greville: On the derivation of discrete interpolation formulas, TSA **1** (1949), 343-357.
6) T. N. E. Greville and H. Vaughan: Polynomial interpolation in terms of symbolic operators, TSA **6** (1954), 413-476.
7) 浅谷輝雄：「生命保険の歴史」四季社 (1957), p. 298.
8) H. Vaughan: Some notes on interpolation, JIA, **72** (1946), 482-497.
9) G. J. Lidstone: Notes on orthogonal polynomials and their application to least-square methods of (1) Fitting polynomial curves to data, (2) Graduation by weighted means, JIA, **64** (1933), 128-159.
10) T. N. E. Greville: Actuarial Note: Adjusted average graduation formulas of maximum smoothness, RAIA, **36** (1947), 249-264.
11) E. T. Whittaker: On a new method of graduation, Proc. Edin. Math. Soc., **41** (1923), 63.
12) R. Henderson: A new method of graduation, TASA, **25** (1924), 29.
13) R. Henderson: Further remarks on graduation, TASA **26** (1925), 52-57.
14) E. H. Wells: The mechanical side of mechanico-graphic graduation, TASA, **38** (1937), 384-401.
15) G. S. Kimeldorf and D. A. Jones: Bayesian graduation, TSA, **19** (1967), 66-112.
16) I. J. Schoenberg: Spline functions and the problem of graduation, Proc. Nat. Acad. Sci., U. S. A., **52** (1964), 947-950.

雑誌名略称

JIA＝Journal of the Institute of Actuaries

RAIA＝The Record of the American Institute of Actuaries

TASA＝Transactions of the Actuarial Society of America

TSA＝Transactions, Society of Actuaries

(昭和 44 年8月 12 日受付)