【創立 70 周年記念大会】
第 I セッション
レジュメ：　董　普

Order restricted Model の社会保障制度への応用

― 正確な死亡率算出にむけた **Isotonic regression** の活用

董　普（東北財経大学金融学院）

中国の経済発展に伴い、保険が国民日常生活に占める位置はますます重要となってきた。近代的中国保険制度の発足は遅いものの、近年には高い成長を示している。中国における生命保険市場は大きく変化し、人口の高齢化に伴う死亡保険市場の成熟化と企業年金など生存保障市場の成長である。

他方、保険の供給者が増加し、競争が一段と厳しくなった。こうした環境変化を背景に、保険会社は競争力を強化するために、より厳密な保険料算出方式の導入が必要となってきている。

死亡保障保険の保険料に影響する大きな要因として、（予定）死亡率と（予定）利率の 2 つがあげられる。無配当の死亡保障保険において、実際の死亡率が予定死亡率より高い場合は、保険者が不利となり、逆の場合は保険契約者が不利となる。

今後、将来の死亡率予測は保険料の計算において更に重要となる。本論では、将来の死亡率をより正確に予測する一つの方法を提案する。

1．問題提起

（1）問題 1：

生命保険を購入するとき、まず保険会社が考慮に入れるのは被保険者の健康状況である。健康状況がよいと思われる人の死亡率は、健康状況がそれほどよくない人の死亡率とは異なる。ここでは、保険会社の健康状況への考察を選択にみなす。

x 歳かつ選択された人のその以降年度の死亡率は $q_{[x]+j}, j = 0,1,\Lambda$ とする。

その中、x歳は選択年齢であり、$q_{[x]+j}$はx歳かつ選択された人の現在年齢が$x+j$歳で、$x+j+1$前での死亡確率である(eg.　$q_{[38]+2}$は38歳かつ選択された人の現在年齢が40歳で、41歳前での死亡確率を表す)。

当然の考え方で、同じ年齢の死亡率は選択時間の推移に伴って増大しつつある。つまり、

$$q_{[x-k+1]+k-1} \geq \Lambda \geq q_{[x-1]+1} \geq q_{[x]}$$

(eg.　$q_{[35]+5} \geq \mathrm{L} \geq q_{[39]+1} \geq q_{[40]}$)

x-i+1歳かつ選択のx歳に生きているn_i人に対して、記号を簡素化するため、$\mu_i = q_{[x-i+1]+i-1}, i = 1,2,\Lambda\ k.$とする。

(1.1)式からわかるように、

$$\mu_1 \leq \mu_2 \leq \Lambda \leq \mu_k \qquad (1.2)$$

その意味は、同じx歳の人に対して、年齢がx-s歳の人を選択することがx歳の人を選択することより長い時間をかかり、死亡率がより高いのである。ここの$\mu = (\mu_1, \mu_2, \Lambda\ \mu_k)$は総体背景を表すパラメータである。

x_{ij}はx-i+1歳かつ選択のx歳に生きているn_i人において、j人目のx+1歳前の生死状況を表す。j=1,2,…n_i。その中

$$x_{ij} = \begin{cases} 1 & \text{死亡} \\ 0 & \text{生存} \end{cases}$$

かくして、x_{ij}は密度関数が$\mu_i^x (1-\mu_i)^x = \exp\{x \ln\frac{\mu_i}{1-\mu_i} + \ln(1-\mu_i)\}$であるサンプルデータである。

通常、サンプルデータを用いて、μ_iを推定する(この推定は制限条件なしの最大尤度推定)

$$\bar{x}_i = \frac{1}{n_i}\sum_{j=1}^{n_i} x_{ij}$$

多くの推定は$\mu_1 \leq \mu_2 \leq \Lambda \leq \mu_k$を満たさず、それは実情とは異なる。

$n_i\bar{x}_i$ が二項分布に従うことは証明されうる。したがって、尤度関数は以下のようになる。

$$\prod_{i=1}^{k}\mu_i^{n_i\bar{x}_i}(1-\mu_i)^{n_i(i-\bar{x}_i)} \qquad (1.3)$$

(1.2)に制限される(1.3)の最大値は、(1.2)に制限される $\mu=(\mu_1,\mu_2,\Lambda\ \mu_k)$ の最大尤度推定である。つまり、以下の数学モデルである。

$$\begin{aligned}&\max\prod_{i=1}^{k}\mu_i^{n_i\bar{x}_i}(1-\mu_i)^{n_i(i-\bar{x}_i)}\\&st\,\mu_1\le \mathrm{L}\ \le\mu_k\end{aligned} \qquad \text{(A)}$$

ここまでの論述からわかるように、x 歳かつ選択された人のその以降年度の死亡率を予測することは、数学モデルの解を求める作業となる。

（2）問題 2：

μ_x を x 歳人の一年以内の死亡率とする。かくして、$\mu_x\le\mu_{x+1}\le \mathrm{L}\ \le\mu_{x+k-1}$. である。

たとえば、$\mu_{70}\le\mu_{71}\le \mathrm{L}\ \le\mu_{84}$ は 70 歳の人の一年以内の死亡率が 71 歳人の同死亡率より小さいことを意味する。その他はこれによって類推しうる。

x_{xj} は x 歳の n_x 人において、j 人目の一年以内の生死状況を表す。j=1,2,… n_x。その中、

$$x_{xj}=\begin{cases}1 & \text{死亡}\\0 & \text{生存}\end{cases}$$

かくして、x_{ij} は密度関数が $\mu_x{}^y(1-\mu_x)^y=\exp\{y\ln\dfrac{\mu_x}{1-\mu_x}+\ln(1-\mu_x)\}$ であるサンプルデータである。

問題 1 の検討と同じように、死亡率 $\mu_x,\mu_{x+1},\mathrm{L}\ ,\mu_{x+k-1}$ への推定は数学モデルの解を求める作業となる。

$$\max \prod_{i=x}^{x+k-1} \mu_i^{n_i \bar{x}_i} (1-\mu_i)^{n_i (i-\bar{x}_i)}$$
$$st\mu_x \leq \text{L} \leq \mu_{x+k-1}$$

ここでは、$\bar{x}_y = \dfrac{1}{n_y} \sum_{j=1}^{n_y} x_{yj}$

2. 問題解決

問題の解決の向け、次の準備を行う。

（１）Isotonic regression の定義

$X = (x_1, x_2, \Lambda, x_k)$ は既定関数とする。もし $X^* = (x_1^*, x_2^*, \Lambda, x_k^*) \in G$ 、かつ

$$\sum_{i=1}^{k} (x_i - x_i^*)^2 w_i = \min_{Y \in G} \sum_{i=1}^{k} (x_i - y_i)^2 w_i$$

であれば、X^* は simple semi-order の下での$(X,\ W)$における Isotonic regression である。ここで、$G = \{y_1, \Lambda, y_k) \mid y_1 \leq \Lambda \leq y_k\}$ 、$W = (w_1, w_2, \Lambda, w_k)$ 、$w_i > 0$ は既定の weighted functions である。

（２）Isotonic regression の解を求める

Isotonic regression の解を求める最適な方法は Ayer et al(1955)に提出された PAVA 計算方法である。

B を $K = (1, 2, \cdots, k)$の部分集合とする。$A_V(B) = \sum_{i \in B} x_i w_i / \sum_{i \in B} w_i$ 。

（３）PAVA 計算方法

① もし $x \in G$, であれば、$x^* = x$。

② もし $x_j > x_{j+1}$, を成立させる j が存在すれば B={j,j+1}, $x_B = A_V(B), \omega_B = \omega_j + \omega_{j+1}$. として、

$\tilde{x} = (x_1, \cdots, x_{j-1}, x_B, x_{j+1}, \cdots, x_k), \tilde{\omega} = (\omega_1, \cdots, \omega_{j-1}, \omega_B, \omega_{j+1}, \cdots, \omega_k)$, とする。

③ ②の繰り返しを通して、K を l 個の $B_1, \cdots B_l$, に分解し、$A_V(B_1) < \cdots < A_V(B_l)$. を満たさせる。かくして、

$x_i^* = A_V(B_t), i \in B_t, t = 1, \cdots, l$。

上述した方法から得られるのは Isotonic regression であり、それは文献で詳しく証明されている。

ある σ 有限測度は以下の密度関数の指数分布族を有する。

$$f(y;\theta,\tau) = \exp\{f_1(\theta)f_2(\tau)H(y,\tau) + s(y,\tau) + q(\theta,\tau)\} \qquad (2.1)$$

その中、y はある集合 A から取得し、θ はある区間 $(\underline{\theta}, \bar{\theta})$ に属する。$\tau \in T$ は嫌なパラメータである。θ の推計問題を考慮するため、以下の条件が成立すると仮定する。

1、f_1 と $q(\cdot, \tau)$ は、$(\underline{\theta}, \bar{\theta})$ での 2 次連続導関数を有する。

2、$f_1'(\theta) > 0, \forall \theta \in (\underline{\theta}, \bar{\theta})$, $f_2(\tau) > 0, \forall \tau \in T$

3、$q'(\theta,\tau) = -\theta f_1'(\theta) f_2(\tau), \forall \theta \in (\underline{\theta}, \bar{\theta}), \forall \tau \in T$

上記の条件から、

$E(H(Y,\tau)) = \theta$、$\mathrm{Var}(H(Y,\tau)) = (f_1'(\theta) f_2(\tau))^{-1}$　　が得られる。

今、仮に k グループがある。y_{ij}、$j = 1, \cdots, n_i$ を密度関数が $f(y_i;\theta_i,\tau_i)$ であるサンプルデータとする。$i = 1, \cdots, k$。条件 1、2、3 から、θ_i の最大尤度推定は以下のようになる。

$$\hat{\theta}_i = \frac{1}{n_i} \sum_{j=1}^{n_i} H(y_{ij}, \tau_i) \qquad (2.2)$$

$\theta = (\theta_1, \theta_2, \cdots, \theta_k)'$ とする。もし θ の分量が semi-order に制限されれば、θ の最大尤度推定は以下の結果に示される。

定理：semi-order 制限がある場合、θ の最大尤度推定は $\hat{\theta}$ の Isotonic regression である。その中、$\hat{\theta} = (\hat{\theta}_1, \hat{\theta}_2, \cdots, \hat{\theta}_k)'$、$\theta_i$ は(2.2)からみえる。weighted functions は $\omega = (n_1 f_2(\tau_1), \cdots, n_k f_2(\tau_k))'$ である。

定理の証明は Robertson et al.(1988)を参照する。

定理：モデル

$$\max \prod_{i=1}^{k} \mu_i^{n_i \bar{x}_i} (1-\mu_i)^{n_i (i-\bar{x}_i)}$$
$$st \mu_1 \le \cdots \le \mu_k$$

の解を求める問題は、$\mu_1 \le \mu_2 \le \cdots \le \mu_k$ 制限の下で、$\sum_{i=1}^{k} n_i (\bar{x} - \mu_i)^2$ の最小値を求める問題と同じである。それは $X = (\bar{x}_1, \cdots \bar{x}_k)$ の Isotonic regression とは同じであり、その中の weighted functions は $N = (n_1, n_2, \ldots n_k)$である。

証明：$X = \begin{cases} 1 & \text{死亡} \\ 0 & \text{生存} \end{cases}$ とする。その中、死亡の確率は μ である。密度関数は

$\mu^x (1-\mu)^x = \exp\{x \ln \frac{\mu}{1-\mu} + \ln(1-\mu)\}$ となる。

(2．1)式に対応して、ここの μ をθにみなし、x を y にみなしている。$s(x, \tau) = 0$、$q(\mu, \tau) = \ln(1-\mu)$、$H(x, \tau) = x$、$f_2(\tau) = 1$、$f_1(\mu) = \ln(\mu/(1-\mu))$。

1、$f_1(\mu) = \ln(\mu/(1-\mu))$ と $q(\mu, \tau) = \ln(1-\mu)$は、$(\underline{0,\ 1})$ での 2 次連続導関数を有する。

2、$f_1'(\mu) = \frac{1}{\mu(1-\mu)} > 0, \forall \mu \in (\underline{0,\ 1})$, $f_2(\tau) = 1 > 0, \forall \tau \in T$

3、$q'(\mu, \tau) = -\frac{1}{(1-\mu)}$ であるため、$q'(\mu, \tau) = -\mu f_1'(\mu) f_2(\tau), \forall \mu \in (\underline{0,\ 1})\ \forall \tau \in T$ である。

公準の条件を満たしている。(2.2)式に対応して、$\hat{\theta}_i = \frac{1}{n_i} \sum_{j=1}^{n_i} x_{ij} = \bar{x}_i,\quad i = 1, \cdots, k.$ である。

公準からわかるように、

$$\max \prod_{i=1}^{k} \mu_i^{n_i \bar{x}_i} (1-\mu_i)^{n_i (i-\bar{x}_i)}$$
$$st\mu_1 \le \mathrm{L} \le \mu_k$$

の解を求める問題は、$\mu_1 \le \mu_2 \le \Lambda \le \mu_k$ 制限の下で、$\sum_{i=1}^{k} n_i (\bar{x} - \mu_i)^2$ の最小値を求める問題と同じである。それは $X = (\bar{x}_1, \Lambda\ \bar{x}_k)$ の Isotonic regression とは同じであり、その中の weighted functions は $N = (n_1, n_2, \ldots n_k)$ である。

○Miller の事例

年齢 x	観測データ数 n_x	観測死亡数 h_x	観測死亡率（最初の推計） u_x
70	135	6	0.044
71	143	12	0.084
72	140	10	0.071
73	144	11	0.076
74	149	6	0.040
75	154	16	0.104
76	150	24	0.160
77	139	8	0.058
78	145	16	0.110
79	140	13	0.093
80	137	19	0.139
81	136	21	0.154
82	126	23	0.183
83	126	26	0.206
84	109	26	0.239

先験結果によれば、死亡率の序列は以下の要素を満たすのである。

1. 年齢が重ねるにつれて、増加しつつある。
2. 高齢者区間において、曲線がより著しく上昇している。

PAVA 計算方法により、以下の結果が得られる。

年齢 x	観測データ数 n_x	観測死亡数 h_x	観測死亡率（最初の推計） u_x
70	135	6	0.044
71~72	283	22	0.078
73	144	11	0.076
74	149	6	0.040
75	154	16	0.104
76	150	24	0.160
77	139	8	0.058
78	145	16	0.110
79	140	13	0.093
80	137	19	0.139
81	136	21	0.154
82	126	23	0.183
83	126	26	0.206
84	109	26	0.239

年齢 x	観測データ数 n_x	観測死亡数 h_x	観測死亡率（最初の推計） u_x
70	135	6	0.044
71~73	427	33	0.077
74	149	6	0.040
75	154	16	0.104
76	150	24	0.160
77	139	8	0.058
78	145	16	0.110
79	140	13	0.093
80	137	19	0.139
81	136	21	0.154
82	126	23	0.183
83	126	26	0.206
84	109	26	0.239

年齢 x	観測データ数 n_x	観測死亡数 h_x	観測死亡率（最初の推計） u_x
70	135	6	0.044
71~74	576	39	0.068
75	154	16	0.104
76	150	24	0.160
77	139	8	0.058
78	145	16	0.110
79	140	13	0.093
80	137	19	0.139
81	136	21	0.154
82	126	23	0.183
83	126	26	0.206
84	109	26	0.239

年齢 x	観測データ数 n_x	観測死亡数 h_x	観測死亡率（最初の推計） u_x
70	135	6	0.044
71~74	576	39	0.068
75	154	16	0.104
76~77	289	32	0.111
78	145	16	0.110
79	140	13	0.093
80	137	19	0.139
81	136	21	0.154
82	126	23	0.183
83	126	26	0.206
84	109	26	0.239

年齢 x	観測データ数 n_x	観測死亡数 h_x	観測死亡率（最初の推計） u_x
70	135	6	0.044
71~74	576	39	0.068
75	154	16	0.104
76~78	434	48	0.111
79	140	13	0.093

80	137	19	0.139
81	136	21	0.154
82	126	23	0.183
83	126	26	0.206
84	109	26	0.239

年齢 x	観測データ数 n_x	観測死亡数 h_x	観測死亡率（最初の推計） u_x
70	135	6	0.044
71~74	576	39	0.068
75	154	16	0.104
76~79	574	61	0.106
80	137	19	0.139
81	136	21	0.154
82	126	23	0.183
83	126	26	0.206
84	109	26	0.239

３．新たに提案する計算方法

まず、先行研究における修正（Smoothing）に対する定義を見てみよう。

（1）Andrews と Nesbitt の定義：自身の規律がある自然現象の若干の観測データから、規則的な修正を通して、その現象を代表するものを求める。

（2）Miller の定義：修正（Smoothing）はこのような信頼できる方法である。その方法を用いて、ある不規則な序列を有する連続変数から、規則的な修正序列が得られる。その得られる序列は観測データと一致である。

大部分の定義は以下のような基本思想がある。つまり、観測データは常に不規則な序列である。これらのデータが修正されるべきと思われるが、修正後の値は観測データに大きく離れていないのである。それは Fit という。

そのほか、アクチュアリーたちは死亡率を修正するとき、常に先験情報を使う。また、それは真実変化率のある時点で形成される円滑な序列と信じる。

以上の考慮をベースに：

Fit：

$$\frac{(\bar{X}_0-p_0)^2}{\frac{p_0(1-p_0)}{n_0}}+\frac{(\bar{X}_1-p_1)^2}{\frac{p_1(1-p_1)}{n_1}}+\frac{(\bar{X}_2-p_2)^2}{\frac{p_2(1-p_2)}{n_2}}+\cdots\frac{(\bar{X}_{104}-p_{104})^2}{\frac{p_{104}(1-p_{104})}{n_{104}}} \qquad (1.1)$$

先験情報

$p_0 \geq p_1 \geq p_2 \geq \cdots \geq p_{11} \leq p_{12} \leq \cdots \leq p_{104} < 1$ （1.2）

円滑性　　$S=\sum_i (\Delta^4 p_i)^2$

（1.3）

つまり、(1.2)の制限条件の下で、(1.1)の最小値を求めるとともに、(1.3)が比較的に小さいことを検証する。詳しくは、以下の数学モデルである。

$$\min\{\frac{(\bar{X}_0-p_0)^2}{\frac{p_0(1-p_0)}{n_0}}+\frac{(\bar{X}_1-p_1)^2}{\frac{p_1(1-p_1)}{n_1}}+\frac{(\bar{X}_2-p_2)^2}{\frac{p_2(1-p_2)}{n_2}}+\cdots\frac{(\bar{X}_{104}-p_{104})^2}{\frac{p_{104}(1-p_{104})}{n_{104}}}\} \qquad \text{モデルA}$$

St: 0< $p_0 \geq p_1 \geq p_2 \geq \cdots \geq p_{11} \leq p_{12} \leq \cdots \leq p_{104} < 1$

ここの $\bar{X}_i$ は i 母体目の粗死亡率をあらわし、p_i は i 母体目の死亡率である。

計算を簡素化するために、以下の計算方法が提出される。

$$\min\{\frac{(\bar{X}_1-p_1)^2}{\frac{p_1(1-p_1)}{n_1}}+\frac{(\bar{X}_2-p_2)^2}{\frac{p_2(1-p_2)}{n_2}}+\cdots\frac{(\bar{X}_k-p_k)^2}{\frac{p_k(1-p_k)}{n_k}}\} \qquad (1.4)$$

St:0< $p_1 \leq p_2 \leq \cdots \leq p_k < 1$

(1.4)式は

$$\frac{n_1(\bar{X}_1-p_1)^2}{p_1(1-p_1)}+\frac{n_2(\bar{X}_2-p_2)^2}{p_2(1-p_2)}+\cdots\frac{n_k(\bar{X}_k-p_k)^2}{p_k(1-p_k)}$$ となる。 （1.5）

【創立70周年記念大会】
第Ⅰセッション
レジュメ：　董　普

公準1：$0 \le x \le 1$ のとき、方程式 $(1-2x)u^2 + 2x^2u - x^2 = 0$ は二つの根がある。

$u_1 = x, u_2 = \dfrac{x}{2x-1}$。$g(u) = (1-2x)u^2 + 2x^2u - x^2$ とする。かくして、$0 < u < x$ の時、$g(u) < 0$。$x < u < 1$ の時、$g(u) > 0.$

証明：二次関数を用いて、簡単に証明できる。

公準2：関数 $f(u) = \dfrac{n(x-u)^2}{u(1-u)}$ を証明する。ここで、$0<x<1$、$f(u)$ は $(0,x)$ において単調減少し、$x < u < 1$ の時に単調増加する。そして $u=x$ のときには最小値となる。

証明：$f'(u) = \dfrac{n}{u^2(1-u)^2}\{(u-x)[(1-2x)u+x]\}$

もし $0 \le x \le 1/2$ であれば、$(1-2x)u + x > 0$。
$0 < u < x$ の場合、$f'(u) < 0$、$x < u < 1$ の場合、$f'(u) > 0$。したがって、$f(u)$ は $(0,x)$ において単調減少し、$x < u < 1$ の時に単調増加する。そして $u=x$ のときには最小値となる。

もし $1/2 < x \le 1$ であれば、$\{\dfrac{x}{2x-1}\}' = -\dfrac{1}{(2x-1)^2} < 0$、$\dfrac{x}{2x-1}$ は単調減少する。

したがって、$\dfrac{x}{2x-1} > 1 > x$。

$0 < u < x$ のとき、公準1からわかるように、$(u-x) < 0, [(1-2x)u + x] > 0$。$f(u)$ は $(0,x)$ において単調減少である。$x < u < 1$ のとき、公準1からわかるように、$(u-x) > 0, [(1-2x)u + x] > 0$。$f(u)$ は $x < u < 1$ の時、単調増加である。$u=x$ のときには最小値となる。

定理 1：仮に関数 $f(p_1, p_2, \mathrm{L}\ , p_k) = \dfrac{(\bar{X}_1 - p_1)^2}{\dfrac{p_1(1-p_1)}{n_1}} + \dfrac{(\bar{X}_2 - p_2)^2}{\dfrac{p_2(1-p_2)}{n_2}} + \mathrm{L}\ \dfrac{(\bar{X}_k - p_k)^2}{\dfrac{p_k(1-p_k)}{n_k}}$ は $0< p_1$

$\leq p_2 \leq \cdots \leq p_k < 1$ の制限の下での最小値点 $p_1{*} \leq p{*}_2 \leq \Lambda\ \leq p{*}_k$ を満たす。もし $\bar{X}_j > \bar{X}_{j+1}$ であれば、 $p{*}_j = p{*}_{j+1}$ である。

証明：仮に $p{*}_j < p{*}_{j+1}$ である。そうすると、以下の三種類の可能性がある。

1° $p{*}_j < p{*}_{j+1} \leq \bar{X}_j$

2° $p{*}_j \leq \bar{X}_{j+1} < \bar{X}_j < p{*}_{j+1}$

3° $\bar{X}_{j+1} < p{*}_j < p{*}_{j+1}$

1° $p{*}_j < p{*}_{j+1} \leq \bar{X}_j$ の場合、まずは

$y{*}_1 = p_1{*},\ y{*}_2 = p{*}_2, \Lambda\ , y{*}_{j-1} = p{*}_{j-1},\ y{*}_{j+1} = p{*}_{j+1}, \Lambda\ , y{*}_k = p{*}_k$ とする。

$y{*}_j$ は $p{*}_j < y{*}_j < p{*}_{j+1}$ を満たす任意値であり、公準 1 から証明できるように、

$$f(y{*}_1, y{*}_2, \Lambda\ , y{*}_k) - f(p{*}_1, p{*}_2, \Lambda\ , p{*}_k) < 0$$

と $p_1{*} \leq p{*}_2 \leq \Lambda\ \leq p{*}_k$ とは関数である。

$f(p_1, p_2, \Lambda\ , p_k) = \dfrac{(\bar{X}_1 - p_1)^2}{\dfrac{p_1(1-p_1)}{n_1}} + \dfrac{(\bar{X}_2 - p_2)^2}{\dfrac{p_2(1-p_2)}{n_2}} + \Lambda\ \dfrac{(\bar{X}_k - p_k)^2}{\dfrac{p_k(1-p_k)}{n_k}}$ は $0< p_1 \leq p_2 \leq \cdots \leq p_k < 1$

の制限の下での最小値点を満たされず、仮定と矛盾する。

その他は同じように証明されうる。

公準 3： $0 < \mathrm{a_i} < 1, \mathrm{i} = 1, \mathrm{L}\ , \mathrm{m}$ のとき、方程式

$((\sum_{\mathrm{i}=1}^{\mathrm{m}} \mathrm{n_i} - 2\sum_{\mathrm{i}=1}^{\mathrm{m}} \mathrm{n_i a_i}) u^2 + 2\sum_{\mathrm{i}=1}^{\mathrm{m}} \mathrm{n_i a_i}^2 u - \sum_{\mathrm{i}=1}^{\mathrm{m}} \mathrm{n_i a_i} = 0$ は二つの根がある。

$$u_1=\frac{-\sum_{i=1}^{m}n_i a_i^2-\sqrt{(\sum_{i=1}^{m}n_i a_i^2)\sum_{i=1}^{m}n_i(a_i-1)^2}}{\sum_{i=1}^{m}n_i-2\sum_{i=1}^{m}n_i a_i},$$

$$u_2=\frac{-\sum_{i=1}^{m}n_i a_i^2+\sqrt{(\sum_{i=1}^{m}n_i a_i^2)\sum_{i=1}^{m}n_i(a_i-1)^2}}{\sum_{i=1}^{m}n_i-2\sum_{i=1}^{m}n_i a_i}$$ 。

$g(u)=((\sum_{i=1}^{m}n_i-2\sum_{i=1}^{m}n_i a_i)u^2+2\sum_{i=1}^{m}n_i a_i^2 u-\sum_{i=1}^{m}n_i a_i$ とする。かくして、$0<u<u_2$ の時、$g(u)<0$。$u_2<u<1$ の時、$g(u)>0.$

証明：二次関数を用いて、簡単に証明できる。

定理 2：関数 $f(u)=\dfrac{\sum_{i=1}^{m}n_i(a_i-u)^2}{u(1-u)}$ を証明する。ここで、$0<a_i<1, i=1,\ldots,m$ とする。$f(u)$ は $0<u<u_2$ において単調減少し、$u_2<u<1$ において単調増加する。そして、$u=u_2$ のとき、最小値となる。

証明：公準 3 と公準 2 の証明方法と類似して、以下のような式が得られる。

$$\min\{\frac{(\bar{X}_1-p_1)^2}{\frac{p_1(1-p_1)}{n_1}}+\frac{(\bar{X}_2-p_2)^2}{\frac{p_2(1-p_2)}{n_2}}+\cdots\frac{(\bar{X}_k-p_k)^2}{\frac{p_k(1-p_k)}{n_k}}\}$$

St:0< $p_1\le p_2\le\cdots\le p_k<1$

この式は以下のように改める。

$$\frac{n_1(\bar{X}_1-p_1)^2}{p_1(1-p_1)}+\frac{n_2(\bar{X}_2-p_2)^2}{p_2(1-p_2)}+\cdots\frac{n_k(\bar{X}_k-p_k)^2}{p_k(1-p_k)}\qquad \text{モデル B}$$

St:0< $p_1\le p_2\le\cdots\le p_k<1$

モデルBの解を求める計算方法（提案）

1°もし $\bar{X}_1 \le \bar{X}_2 \le \cdots \le \bar{X}_k$ であれば、$(p_1^*, p_2^* \cdots, p_k^*) = (\bar{X}_1, \bar{X}_2, \cdots, \bar{X}_k)$ である。成立しなければ、次に転入する。

2° もし $\bar{X}_j > \bar{X}_{j+1}$ が存在すれば、かつ

$$\bar{X}_1 \le \cdots \le \bar{X}_{j-1} \le \frac{-\sum_{i=j}^{j+1} n_i\bar{x}_i^2 + \sqrt{(\sum_{i=j}^{j+1} n_i\bar{x}_i^2)\sum_{i=j}^{j+1} n_i(\bar{x}_i-1)^2}}{\sum_{i=j}^{j+1} n_i - 2\sum_{i=j}^{j+1} n_i\bar{x}_i} \le \bar{X}_{j+2} \le \cdots \le \bar{X}_k$$ であれば、

$$(p_1^*, p_2^* \cdots, p_{j-1}^*) = (\bar{X}_1, \bar{X}_2, \cdots, \bar{X}_{j-1}),\quad p_{j+1}^* = p_j^* = \frac{-\sum_{i=j}^{j+1} n_i\bar{x}_i^2 + \sqrt{(\sum_{i=j}^{j+1} n_i\bar{x}_i^2)\sum_{i=j}^{j+1} n_i(\bar{x}_i-1)^2}}{\sum_{i=j}^{j+1} n_i - 2\sum_{i=j}^{j+1} n_i\bar{x}_i}$$ である。

$(p_{j+2}^*, \cdots, p_k^*) = (\bar{X}_{j+2}, \cdots, \bar{X}_k)$ が成立しなければ、

$$B = \{j, j+1\},\ \bar{X}_B = \frac{-\sum_{i=j}^{j+1} n_i\bar{x}_i^2 + \sqrt{(\sum_{i=j}^{j+1} n_i\bar{x}_i^2)\sum_{i=j}^{j+1} n_i(\bar{x}_i-1)^2}}{\sum_{i=j}^{j+1} n_i - 2\sum_{i=j}^{j+1} n_i\bar{x}_i}$$ である。

$(\bar{X}_1, \bar{X}_2, \cdots \bar{X}_{j-1}, \bar{X}_B, \bar{X}_{j+2}, \cdots, \bar{X}_k)$ に対して、次に転入する。

3° もし $\bar{X}_{B_s} > \bar{X}_{B_{s+1}}$ であれば、

$$B = B_S \cup B_{S+1},\ \bar{X}_B = \frac{-\sum_{i \in B_S \cup B_{S+1}} n_i\bar{x}_i^2 + \sqrt{(\sum_{i \in B_S \cup B_{S+1}} n_i\bar{x}_i^2)\sum_{i \in B_S \cup B_{S+1}} n_i(\bar{x}_i-1)^2}}{\sum_{i \in B_S \cup B_{S+1}} n_i - 2\sum_{i \in B_S \cup B_{S+1}} n_i\bar{x}_i}$$ である。

3の繰り返しを通して、集合をm個の B_1, $B_2, \cdots$, B_m に分解し、

$\bar{X}_{B_1} \le \bar{X}_{B_2} \le \cdots \le \bar{X}_{B_m}$ を満たさせる。かくして、$p_i^* = \bar{X}_{B_t}, i \in B_t, i = 1, \cdots, k$ となる。

証明：定理1と定理2から証明できる。

上記の計算方法からわかるように、k回で最終結果が得られる。

自身が関与した参考文献

1.. 盧玉貞・董普(2004)、「多維保序回帰及最大似然估計」、『大連海事大学学報』12 号、Vol.102

2. 董普・盧玉貞(2003)、「正態分布的均值被簡単樹半序約束方差為未知的最大似然估計」、『数学的実践和認識』11 号、Vol.33

3. 董普(2003)、「多維正態分布均値在序約束下的仮説検験」、『数学進展』ISSN 1000-0917、Vol.32、No.1

4. 董普・盧玉貞(2003)、「正態分布的均値和方差別分別被簡単樹半序和簡単半序約束下的保序最大似然估計」、『応用概率統計』ISSN 1001-4268、Vol.19、No.1

5. 盧玉貞・董普(2002)、「選択死亡率的一種估計方法」、『中国高教論衆』ISBNT-5437-3063-4/G.2788、Vol.22

6. 董普・盧玉貞(2001) 、「多維保序回帰和正態分布均値和方差的估計」、『中国高教論衆』ISBN7-5437-3003-O/G·2751、Vol.22、No.2

7. 第二作者(2000)、「K=2 時多維保序回帰」、『大連海事大学学報』1 号、96-99

8. 「対三個正態母体均値被簡単半序約束方差為未知的最大似然估計」、『CSIAM 2000 中国工業与応用数学学会第六次大会論文集』、2000 年、272-275

9. 第二作者(2000)、「均値被半序約束方差為未知的最大似然估計」、『瀋陽建築工程学院学報』2 号、150-151

10. 第二作者(1998)、「正態分布的均値被簡単半序約束方差為未知的最大似然估計」、『CSIAM 1998 中国工業与応用数学学会第五次大会論文集』、402-406

11. 「在簡単半序約束下、多維保序回帰」、『CSIAM 1998 中国工業与応用数学学会第五次大会論文集』、1998 年、415-419

12. 第二作者(1997)、「有限制条件下 Poisson 分布均値的估計」、『瀋陽建築工程学院学報』1 号 76-80

13. 第二作者(1996)、「半序約束下正態均値和方差的估計」、『瀋陽建築工程学院学報』3 号、367-372

筆者の提案する計算方法の応用

中国財政部：基本養老保険収入、支出に関する計算と統計分析システム

中国財政部：中国基本医療保険の計算モデルと応用研究

REFERENCES

[1] SHI, N.Z., (1993). Isotonic regression and maximum likelihood estimate, Chinese Journal of Applied Probability and statistics VOL. 9, NO. 2, MAY (1993) 203-215.

[2] SHI, N.Z., (1993). Maximum likelihood of isotonic normal means with unknown variances; Journal of Multivariate Analysisi (1988), 64, 183-195.

[3] Barlow, R.E, Bartholoremew, D.J., Bremner, J.M. and Brunk, H.D., Statistical Inference under Order Restrictions: The Theory and Application of Isotonic Regression. New York: Wiley, 1972.

[4] Ayer, M., Brunk, H.D., Ewing, G.M., Reid, W.T. and Silverman, E., Ann. Nath. Statist.,26 (1955), 641-647.