2000年度日本オペレーションズ・リサーチ学会
春季研究発表会

危険度診断における総期待損失最小化モデルの一般化

01000170　青山学院大学　阿部 俊一*　ABE Shun-ichi
青山学院大学　劉 城剛**　LIU Cheng-gang

1．はじめに

構造部材等の非破壊検査データによる破壊強度レベルの格付け，防災診断データによる災害危険度レベルの判別，企業診断データによる経営健全度の格付け，医学検査データによる病名診断などは，いずれも判別分析モデルによって取扱うことができるが，本研究では文献[1]の総期待損失最小化モデルを更に一般化し，下記のモデルと記号を導入する．

2．確率モデルと記号

(1) $x=(x_1, x_2,\cdots, x_p)$：診断データベクトル；

(2) $f_j(x)$：母集団 $j\in M\equiv\{1,2,\cdots,m\}$ に属する個体（システム）に対するデータ x の確率密度関数；

(3) P_j：母集団 $j\in M$ に属する個体の出現確率；

(4) $\delta_k(x)=1$（データ x を持つ個体をその真の母集団 $j\in M$ の如何にかかわらず,母集団 k に属するものと判別するとき),
$\delta_k(x)=0$（その他のとき）；

(5) $C_{jk}(x)$：母集団 $j\in M$ に属し,データ x を持つ個体を母集団 $k\neq j$ に属するものと誤判別する時発生する期待損失で，x の区分的に連続な正の関数；特に，$C_{jj}(x)=0\ (j\in M)$ とおく；

(6) M に属する l 個の異なった要素を $\{k_1, k_2,\cdots, k_l\}$ $(2\leq l\leq m)$ とし，$\delta_{\{k_1,k_2,\cdots,k_l\}}(x)=1$（データ x を持つ個体の母集団をその真の母集団 $j\in M$ の如何にかかわらず，$k_1, k_2,\cdots, k_l$ のいずれかと判別するとき)，$\delta_{\{k_1,k_2,\cdots,k_l\}}(x)=0$（その他のとき）と定義し，こうした‘あいまいな判別’を許容するため，この種の判別に伴う期待費用を

(7) $C_{j\{k_1,k_2,\cdots,k_l\}}(x)\leq C_{jk_h}(x)\ \ (j\neq k_h, h=1,2,\cdots,l)$，
とする．ただし,簡単のために(6)，(7)における‘あいまいな判別’のための集合 $\{k_1, k_2,\cdots, k_l\}$ 全部(いま $n-m$ 個とする)を数え上げ，これらを改めて番号 $k=m+1, m+2,\cdots, n$ で表すことにし，$C_{j\{k_1,k_2,\cdots,k_{l_k}\}}$，$\delta_{j\{k_1,k_2,\cdots,k_{l_k}\}}$ をそれぞれ単に C_{jk}，$\delta_k (k=m+1, m+2, \cdots, n)$ とかくことにする；

(8) $D_k=\left\{x\,\middle|\,\delta_k(x)=1, x\in\Omega_k\right\}$
とする $(k\in N\equiv\{1, 2,\cdots, n\})$．ただし，

(9) $\Omega_k=\{x\mid f_k(x)>0\}$，$\Omega=\bigcup_{k=1}^{m}\Omega_k$．

3．提案モデルの数式化と最適化

前節の確率モデルと記号のもとで，判別対象の1個体(システム)当たりの総期待損失費用は

$$(10)\ C=\sum_{j=1}^{m}\sum_{k=1}^{n}P_j\int_{x\in\Omega}f_j(x)C_{jk}(x)\delta_k(x)dx$$
$$=\sum_{j=1}^{m}\sum_{k=1}^{n}P_j\int_{x\in D_k}f_j(x)C_{jk}(x)dx$$

と表される。この値を最小化する最適な $D_k=D_k^*$ を求める手順は以下の通りである．

$$(11)\ g_k(x)=\sum_{j=1}^{m}P_jf_j(x)C_{jk}(x)\ \ (k\in N\equiv\{1,2,\cdots,n\})$$

$$(12)\ g_{k^*}(x)=\min_{k\in N}g_k(x)$$

となる整数 k^* が存在するが，同一の $x\in\Omega$ に対して k^* が2個以上存在すれば

(13) 任意の1個を選び，$k^*=k^*(x)$ とおき，

(14) $D_k^*=\{x\mid k^*(x)=k, x\in\Omega_k\}$　$(k\in N)$
とおけばよい．

4．$M=\{1,2\}$, $N=\{1, 2, 3\}$ のモデル例など

$M=\{1,2\}$ の場合，与えられた母集団は $j=1, 2$ で表される2個だけであるが，データ $x=(x_1, x_2,\cdots, x_p)$ を持つ個体（システム）を，それぞれ，母集団1，2と判別する x の最適領域 D_1^*, D_2^* のほかに，その個体を母集団1または2に属するものと‘あいまいに判別する’あるいは‘明確な判別を保留する’第3の領域として最適な D_3^* を求めようとするのである．誤判別及びあいまいな判別（判別保留）の総期待費用(10)は，記号(1)～(9)を用いて，この例では

*青山学院大学大学院理工学研究科経営工学専攻　教授　**同学生　〒157−8572 東京都世田谷区千歳台6−16−1　TEL 03-5384-1111(代)

(15) $C=\int_{x\in D_2} P_1 f_1(x) C_{12}(x)dx+\int_{x\in D_1} P_2 f_2(x) C_{21}(x)dx$
$+\int_{x\in D_3}[P_1 f_1(x) C_{13}(x)+P_2 f_2(x) C_{23}(x)]dx$

となる．ただし，ここに D_3 は上記のあいまいな判別(判別保留)に対応する領域とする．また

(16) $\rho_*(x)=\dfrac{P_1 f_1(x)}{P_2 f_2(x)},\ \ \rho_0(x)=\dfrac{C_{21}(x)}{C_{12}(x)},$

$\rho_1(x)=\dfrac{\alpha_2(x)}{C_{13}(x)},\quad \rho_2(x)=\dfrac{C_{23}(x)}{\alpha_1(x)},$

$\alpha_1(x)=C_{12}(x)-C_{13}(x), \alpha_2(x)=C_{21}(x)-C_{23}(x),$

$\pi(x)=\dfrac{\alpha_1(x)\alpha_2(x)}{C_{13}(x)C_{23}(x)}$

と定義すると，文献[1]と同様にして

(17) $\rho_1(x)<\rho_0(x)<\rho_2(x) \iff \pi(x)<1,$

(18) $\rho_1(x)>\rho_0(x)>\rho_2(x) \iff \pi(x)>1,$

(19) $\rho_0(x)=\rho_1(x)=\rho_2(x) \iff \pi(x)=1$

となることを証明することができる．そこで，

(20) $D_1^-=\{x\ |\rho_*(x)>\rho_0(x), \pi(x)\le 1, x\in\Omega_1\},$

$D_2^-=\{x\ |\rho_*(x)<\rho_0(x), \pi(x)\le 1, x\in\Omega_2\},$

$D_3^-=\{x\ |\rho_*(x)=\rho_0(x), \pi(x)\le 1, x\in\Omega\};$

(21) $D_1^+=\{x\ |\rho_*(x)>\rho_1(x), \pi(x)>1, x\in\Omega_1\},$

$D_2^+=\{x\ |\rho_*(x)<\rho_2(x), \pi(x)>1, x\in\Omega_2\},$

$D_3^+=\{x\ |\rho_1(x)\ge\rho_*(x)\ge\rho_2(x), \pi(x)>1, x\in\Omega\};$

とおくと，(15)を最小化する最適な領域は

(22) $D_1^*=D_1^-\cup D_1^+, D_2^*=D_2^-\cup D_2^+, D_3^*=D_3^-\cup D_3^+$

で与えられることを示すことができる．

例1　文献[1]のモデルは $m=p=2$，$f_1(x)$ と $f_2(x)$ がいずれも共通の分散･共分散行列を持つ２次元の正規確率密度関数で

(23) $\rho_0=1/3$，$\rho_1=1$，$\rho_2=1/5$，

$\pi(x)=5>1$，$D_1^-=D_2^-=D_3^-=\phi$

となる特別なモデルであり，

(24) $\theta_1=\ln\rho_1=0$，$\theta_2=\ln\rho_2=-1.6094$，

(25) $Z(x)=0.2325x_1+0.006441x_2-15.541$

に対して

(26) $D_1^*=\{x\,|\,Z(x)>\theta_1, x\in\Omega_1\},$

$D_2^*=\{x\,|\,Z(x)<\theta_2, x\in\Omega_2\},$

$D_3^*=\{x\,|\,\theta_1\le Z(x)\le\theta_2, x\in\Omega\}.$

例2　(15)において，特に

(27) $C_{13}(x)=C_{23}(x)=\lambda_0>0,\ \ 0<\lambda_j(x)<\infty,$

(28) $\alpha_j(x)=\dfrac{\lambda_j(x)P_j f_j(x)}{P_1 f_1(x)+P_2 f_2(x)}\ \ (j=1,2)$

とおくと，以下の各式が得られる：

(29) $\alpha_1(x)=\lambda_1(x)\,\rho_*(x)/(1+\rho_*(x)),$

$\alpha_2(x)=\lambda_2(x)/(1+\rho_*(x)),$

(30) $\pi(x)=\lambda_1(x)\,\lambda_2(x)\,\rho_*(x)/\{\lambda_0(1+\rho_*(x))\}^2.$

(31) $\pi(x)\le 1 \iff z^2+\{2-\lambda_0^{-2}\lambda_1(x)\lambda_2(x)\}z+1\ge 0,$

ただし，$z=\rho_*(x)$．これから

(ｉ)すべての $x\in\Omega$ に対して $\lambda_1(x)\lambda_2(x)\le 4\lambda_0^2$ が成立する場合，(15)を最小化する(22)の領域は

(32) $\begin{cases} D_1^*=D_1^-=\{x\,|\,\rho_*(x)>\rho_0(x), x\in\Omega_1\},\\ D_2^*=D_2^-=\{x\,|\,\rho_*(x)<\rho_0(x), x\in\Omega_2\},\\ D_3^*=D_3^-=\{x\,|\,\rho_*(x)=\rho_0(x), x\in\Omega\ \}.\end{cases}$

となり，D_1^* と D_2^* の判別境界(判別関数) $D_3{}^*$ は

(33) $\rho_*(x)=\rho_0(x)=\sqrt{\dfrac{\lambda_0+\lambda_2(x)}{\lambda_0+\lambda_1(x)}}$

で与えられる．

(ⅱ)任意の $x\in\Omega$ に対して $\lambda_1(x)\lambda_2(x)>4\lambda_0^2$ の場合，

$B-1=\{\lambda_1(x)\,\lambda_2(x)\ -4\lambda_0^2\}/(2\lambda_0^2)>0,$

$\theta_1, \theta_2=B\pm\sqrt{B^2-1}\quad(\theta_2>\theta_1>0)$

となり，(30),(31)から

(34) $\pi(x)\le 1 \iff \rho_*(x)\le\theta_1$ または $\rho_*(x)\ge\theta_2,$

(35) $\pi(x)>1 \iff \theta_1<\rho_*(x)<\theta_2$

となる．従って最適な判別領域は(20),(21),(34),(35)で特徴付けられる $D_j^-, D_j^+ (j=1,2,3)$ を用いて(22)で与えられる．

(ⅲ)任意の $x\in\Omega$ に対して上の(i)または(ii)の条件が成立する場合は,上記の(i)と(ii)の簡単な応用,

例3　$M=\{1,2,3\}, N=\{1,2,\cdots,5\}$ の例として,例2において確率密度 $f_3(x)$ を持つ母集団3を追加し，あいまいな判別領域, $4=\{1,2\}$, $5=\{2,3\}$,を許容するモデルを考える(詳細は発表の際に示す).

参考文献：
[1]阿部俊一："２群判別で判別保留を考慮した総期待費用最小化モデル，" 日本 OR 学会 1997 年度秋季研究発表会アブストラクト集，pp.152－153.