I－2

ベイズ決定方式について

信州大学工学部　正　員　長　尚
信州大学工学部　正　員○小山　健

1．まえがき

期待損失最小化原則による決定方式を通常ベイズ決定方式と呼んでいる。これはこの解法がベイズの定理による事後確率の概念の助けを借りて誘導され，実際の計算過程でも事後確率を計算して利用しているからである。しかしながら本文では事後確率の概念の助けを借りなくても解法が誘導でき，したがって計算過程でも事後確率を利用しなくてよいことを示すと共に，その方が一般に計算の効率も良いことを指摘する。

2．解法原理及び手順

まず用語及び記号を次のように定義する。Θ（自然の状態）＝ $[\theta_1 \cdot \theta_i \cdot \theta_I]^T$ ・・（1），A（行動）＝ $[a_1 \cdot a_k \cdot a_K]$ ・・（2），Z（観測値）＝ $[z_1 \cdot z_l \cdot z_L]$ ・・（3），

$$\Lambda = \begin{bmatrix} L(\theta_1, a_1) & \cdot & L(\theta_1, a_K) \\ \cdot & L(\theta_i, a_k) & \cdot \\ L(\theta_I, a_1) & \cdot & L(\theta_I, a_K) \end{bmatrix} \quad (4) \quad P = \begin{bmatrix} p(\theta_1, z_1) & \cdot & p(\theta_1, z_L) \\ \cdot & p(\theta_i, z_l) & \cdot \\ p(\theta_I, z_1) & \cdot & p(\theta_I, z_L) \end{bmatrix} \quad (5),$$

Λ（損失関数），P（観測値の確率），W（自然の状態の事前確率）＝ $[w_1 \cdot w_i \cdot w_I]^T$ ・・・（6），D（決定方式）＝ $[d(z_1) \cdot d(z_l) \cdot d(z_L)]$ ・・（7）

各 z_l が観測されたという条件から決定されるD方式を採用した場合，i 状態の被る損失の期待値が危険度と呼ばれ，次の式で表される。$R(\theta_i, D) = \sum_{l=1}^{L} p(\theta_i, z_l)\, L(\theta_i, d(Z_l))$ ・・（8）
この危険度の期待値を，期待損失と言い，それは次のようになる。$E(R) = \sum_{i=1}^{I} \{w_i \sum_{l=1}^{L} p(\theta_i, z_l)\, L(\theta_i, d(Z_l))\} = \sum_{l=1}^{L} \{\sum_{i=1}^{I} w_i\, p(\theta_i, z_l)\, L(\theta_i, d(Z_l))\}$ ・・（9）　この式（9）が最小になるように決定方式Dを決めるのが，期待損失最小化原則による解法である。ところで式（9）の右辺の｛｝の中はその決めるべき方式の一つのエレメント $d(Z_l)$ のみが関係し，他のエレメントは関係していない。したがって式（9）の右辺全体を最小にするには，各 l について｛｝の中を別個に最小にする行動 a_k を選択すれば良いことになる。すなわち，$\min_{D} E(R) = \sum_{l=1}^{L} [\min_{a_k} \{\sum_{i=1}^{I} w_i\, p(\theta_i, z_l)\, L(\theta_i, a_k)\}]$ ・・（10）　となる。つまり式（9）で定義される期待損失を最小化するには，各観測毎の期待損失を別々に最小化すればよいことになる。そのためこの式（10）による具体的な解法手順は次のように非常に簡単になる。「まず z_l（l＝1・・・L）を固定し，すべての a_k（k＝1・・・K）について次の計算を行う。$x(k) = \sum_{i=1}^{I} w_i\, p(\theta_i, z_l)\, L(\theta_i, a_k)$ ・・・（11）　x(k) が最小になる a_k が $d(z_l)$ である。この操作をすべての z_l について行うと，決定方式の解Dが求められる。」

3．事後確率を用いる方法との比較

ここで，$p(z_l) = \sum_{i=1}^{I} w_i\, p(\theta_i, z_l)$ ・・（12）　とおき，これを用いて式（10）を次のように変形する。$\min_{D} E(R) = \sum_{l=1}^{L} [\min_{a_k} \{P(z_l) \sum_{i=1}^{I} ((w_i\, p(\theta_i, z_l)) / P(z_l))\, L(\theta_i, a_k)$

$\}\ ] = \sum_{\ell=1}^{L} [P(z_\ell) \min_{a_k} \{ \sum_{i=1}^{I} w_i'(z_\ell) L(\theta_i, a_k)\}]$ ・・（13） ここに $w_i'(z_\ell)$はz_ℓ を観測した時の，ベイズの定理によるθ_i の事後確率である。この式（13）は，式（10）による“各観測毎の期待損失最小化”とは，z_ℓ の観測が条件となったθ_i の事後確率による“各観測毎の事後期待損失最小化”という意味を，結果として含んでいるということを示している。これは前述したようにDの決定にはz_ℓ の観測が条件になっていることと，自然の状態Θはどれか一つのθ_i に絞られる筈であるという状況下[1]である観測z_ℓ がなされたとき，ベイズの定理によって事前確率w_i は事後確率$w_i'(z_\ell)$に修正されることとからの，当然の帰結である。ただし｛｝中の各観測毎の期待損失の値は両者では異なる。何故なら式（13）から明らかなように，後者は観測毎に基準化された確率が用いられているからである。この式（13）による解法が「ベイズ決定方式の解法」として通常の教科書・参考書などに示されている。例えばある著書[2]によるとその解法手順は次のようになっている。「まずv_1，・・・，v_I を固定して（$v_i \geq 0$，$\sum_{i=1}^{I} v_i = 1$），$\min_{a_k} E\{R(V)\} = \min_{a_k} \sum_{i=1}^{I} v_i L(\theta_i, a_k)$ ・・（14） の関係から種々なVに対応するa_k を求めておく。次いで事後確率を計算し，この関係を利用して，決定方式Dの解を求める。」 式（14）においてVに事前確率Wを用いると観測値がない場合の決定方式ということになるので，これをノー・データ問題と呼んでいる。この場合は観測値z_ℓ に無関係な一定の行動a_k が解となる。Vに事後確率Wを用いると観測値がある場合の決定方式となり，これを求めるには種々な確率に対して予めノー・データ問題の解を用意しておくと便利なこともある。そこでこのような解法手順が考えられているようである。なお式（14）でVにWを用いたときのノー・データ問題の最小期待損失 min E｛R（W）｝は，式（10）のデータがある場合の期待損失 min E（R）においてz_ℓ に無関係に一定のa_k を選択したことに相当しているので（z_ℓ 毎に別なa_k を選択した方が期待損失は一般に小さい），これらの間には次のような関係がある。

$\min E\{R(W)\} \geq \min E(R)$ ・・（15） つまりデータを利用することによって，それがない場合の解とは異なる行動がその解の中に含まれる場合には，期待損失はデータがないときより小さくなり，それだけデータの効用があることになる。ところで，この事後確率を用いた解法も前述の式（10）による解法も，式の誘導過程から明らかなように，いずれも式（9）で定義された期待損失を最小にすることを原則に用いているから，本質的な差はなく，当然同じ解が得られる。しからば事後確率を用いた解法の方が計算が容易かと言うと，必ずしもそうではない。確かに，自然の状態の数Iが2～3の場合には，ノー・データ問題の解を予め一つの図等で用意しておけるメリットはあるが，計算の手間は式（10）によるものと大差はない。Iが4以上となると，ノー・データ問題の解を予め用意しておくことは不可能で，事後確率を別に計算する必要がある分だけ，事後確率を用いた解法の方が計算の効率が落ちる。したがって通常の教科書・参考書で，例えば前記書[2]の119ページにあるように，期待損失“最小化の問題を解析的に解かねばならないが，それには一般的な方法がない．そのために考えられた有力な方法が事後確率の方法である”として，いかにも事後確率を用いなければうまく解けないというようなような説明は，適切ではない。なお計算例による比較は講演時に発表する。

参考文献　1）長尚：ベイズの定理の適用について，土木学会論文集第 350号，1984.　2）松原望：意志決定の基礎，朝倉書店，1977.