数理解析研究所講究録
第150巻 1972年 23-36

最適領域の推定に関する統計的構造

統計数理研究所　　野田　一雄
多賀　保志

§1. 最適領域の決定問題

$(\mathscr{X}, \mathfrak{A})$を標本空間とし、その上の確率測度のある族$\mathscr{P}_1$が与えられているとする。

$\mathscr{X}$の上の$\mathfrak{A}$－可測な実関数のベクトル$\phi = (\phi_1, \phi_2, \cdots, \phi_\ell)$;

$$(1.1)\qquad \sum_{j=1}^{\ell} \phi_j(x) = 1, \qquad \phi_j(x) \geq 0 \qquad (j = 1, 2, \cdots, \ell)$$

を$\mathscr{X}$のℓ－分割とよぶ。　これらの全体をΦで表わすことにしよう。　そのうち制約条件$\mathbb{C}$を満たすものの全体を$\Phi(\mathbb{C}) \subset \Phi$で表わして、これが空でないことを仮定しておく。　また制約条件$\mathbb{C}$を置かないときは、Φを直接の対象と考えればよい。

いま、$\forall p \in \mathscr{P}_1$ について可積分であるような$\mathfrak{A}$－可測な実関数$g_i(x)$をとり、pに関して(g_i)とϕの内積の関数となるようなある目的関数$v(\phi, p)$を考える。　$\Phi(\mathbb{C})$の中に

この目的関数 $v(\phi, p)$ の最大値もしくは最小値を実現するものが存在するとき、これを $\mathfrak{X}$ の最適分割とよび、その一つを ϕ_p で表わすことにする。ここでは ϕ_p は $v(\phi, p)$ の最大値を実現するものとしよう。

このとき、ある $\phi \in \Phi(\mathbb{C})$ を選んだときの損失を

$$(1.2)\qquad L(\phi, p) = v(\phi_p ; p) - v(\phi, p) \geq 0$$

とする。

一般に $p \in \mathcal{P}_1$ は未知であるから，ϕ を決定する際に事前情報 s をとることを考える。s が値をとる標本空間を $(\mathcal{S}, \mathcal{B})$ とし，その上の確率測度の族 $\mathcal{Q}_1 = \{Q_p ; p \in \mathcal{P}_1\}$ が与えられているとする。つまり，X が $p \in \mathcal{P}_1$ に従うとき，s はこれとは独立に $Q_p \in \mathcal{Q}_1$ に従うと考える。

かくして，写像 $\dot{\phi} : \mathcal{S} \longrightarrow \Phi(\mathbb{C})$ を決定関数と考えるのであるが，これにはある種の可測性を仮定することによって，$\mathfrak{X} \times \mathcal{S}$ 上の $\mathcal{Q} \times \mathcal{B}$ －可測な実関数のベクトル $\hat{\phi} = (\hat{\phi}_1, \hat{\phi}_2, \cdots, \hat{\phi}_l)$;

$$(1.3)\qquad \sum_{j=1}^{l} \hat{\phi}_j(x, s) = 1, \qquad \hat{\phi}_j(x, s) \geq 0 \quad (j = 1, 2, \cdots, l)$$

をもって表わすことが出来る（[3]）。もちろん $\Phi(\mathbb{C})$ は，この決定の行動空間とみなされるわけである。

そこで，この場合，危険関数を

$$(1.4)\qquad r(\hat{\phi}, p) = \mathbb{E}_{Q_p}[L(\hat{\phi}, p)]$$

$$\hat{\phi} \in \hat{\Phi}(\mathbb{C}), \quad p \in \mathcal{P}_1$$

と考える。 (1.3) を満足する $\hat{\phi}$ の全体を $\hat{\Phi}$, そのうち制約条件 $\mathbb{C}$ を満足する部分空間を $\hat{\Phi}(\mathbb{C})$ で表わす。 直接には, $\hat{\Phi}(\mathbb{C})$ が決定関数の空間となるわけである。

この場合, ϕ_p のミニマックス推定量 $\tilde{\phi}$ は,

$$(1.5)\qquad \inf_{\hat{\phi}\in\hat{\Phi}(\mathbb{C})}\ \sup_{p\in\mathcal{P}_1} r(\hat{\phi}, p)$$

を実現するもので, これが当面の目標となるであろう。

このミニマックス $\tilde{\phi}$ を求めるために, Bayes 推定量との関係を論じるには, 次のような先験測度 ξ のある集合 Ξ_1 を考えていく。 すなわち, $\mathcal{P}_1$ の部分集合から作られる σ-代数 $\mathcal{C}_1$ を適当にとって, $(\mathcal{P}_1, \mathcal{C}_1)$ 上の確率測度 ξ のある族 Ξ_1 を選ぶ。

この場合, Bayes 危険は,

$$(1.6)\qquad r(\hat{\phi}, \xi) = \mathbb{E}_\xi[r(\hat{\phi}, p)]$$
$$\hat{\phi}\in\hat{\Phi}(\mathbb{C}),\ \xi\in\Xi_1$$

となり, その inf を実現する $\hat{\phi}\in\hat{\Phi}(\mathbb{C})$ を ϕ_p の Bayes 推定量とよび $\hat{\phi}_\xi$ で表わすことにしよう。

一般には, 先験確率測度 ξ は, $(\mathcal{P}_1\times\mathcal{Q}_1, \mathcal{C}_1\times\mathcal{D}_1)$ の上の測度と考えるべきであるが, この場合 p と Q_p との独立性を仮定しているので, 上記のように表現してよい。

さて, 以上の設定の下で, 次のようなことが問題となって

くる。

[1°] 一般的な条件の下で，最適分割 ϕ_p の存在およびその explicit form を求めること。

[2°] 一般的な条件の下で，Bayes 推定量 $\hat{\phi}_\xi$ の存在およびその explicit form を求めること。

[3°] 一般的な条件の下で，ミニマックス推定量の存在およびその explicit form を求めること。

特に適当に $\{\xi_i : i=1,2,\cdots\} \subset \Xi_1$ を選出することによって

$$(1.7) \qquad \sup_{p \in \mathcal{P}_1} r(\tilde{\phi}, p) = \lim_{i \to \infty} r(\hat{\phi}_{\xi_i}, \xi_i)$$

なる関係を求めること。

[4°] 有限個ないしは可算個の $\mathcal{P}_0 = \{p_i\} \subset \mathcal{P}_1$ を適当にとることによって，族 $\mathcal{P}_0$ についての Bayes 推定量ならびにミニマックス推定量が $\hat{\phi}_\xi$，$\tilde{\phi}$ の近似になっているようにすること。

特に $\mathcal{P}_1$ に適当な距離を導入することにより，その ε－被覆を上述の問題に適用すること。

[5°] $\tilde{\phi}$ は (1.7) の関係から許容性をもつことになるが，更に $\hat{\phi}_\xi$，$\tilde{\phi}$ の consistency，すなわち，S のサンプル・サイズ n を増大させるとき，危険関数が 0 に収束せしめるような性質をもつことを示すこと。

以上の問題が[2],[3]において、次のような定式化の下で議論された。

νを$(\mathscr{X}, \mathfrak{A})$上のある与えられた$\sigma$-有限な測度とする。$f_i(x)$を$\mathfrak{A}$-可測で、$\nu$および$\forall p \in \mathscr{P}_1$について可積分な実関数, $g_{i'}(x)$を$\forall p \in \mathscr{P}_1$について可積分な実関数とする。

このとき，制約条件を，c_{ij}をある定数として，

$$(1.8) \qquad \tau_{ij}(\phi) = \int_{\mathscr{X}} f_i(x)\phi_j(x)\nu(dx) \leq c_{ij} \quad (\text{or} = c_{ij})$$
$$(i=1,2,\cdots,m\,;\ j=1,2,\cdots,l)$$

によって定める。

次に目的関数vは、まず

$$(1.9) \qquad \psi_{ij}(\phi, p) = \int_{\mathscr{X}} g_{i'}(x)\phi_j(x)p(dx)$$
$$(i'=1,2,\cdots,k\,;\ j=1,2,\cdots,l)$$

と置く。$k \times l$-実数値行列空間のある凸部分集合$\mathfrak{Z}$をとることにより、$\psi_{ij}(\phi, p)$の値がすべての$\phi \in \mathfrak{F}(\mathfrak{C})$, $p \in \mathscr{P}_1$について$\mathfrak{Z}$の中に実現されることを仮定しておく。このとき$\mathfrak{Z}$上の実関数$h(z)$をとり，

$$(1.10) \qquad v(\phi, p) = h_0[\psi_{ij}(\phi, p)]$$

と定める。

かくして，$\mathfrak{A}$, $\mathfrak{B}$, および$\mathscr{P}_1$, $\mathscr{Q}_1$およびんに一般的な条件を置くことによって，[1°]～[5°]の議論が展開される。

さて、ここでは、[5°]に関して、[3]でとられた方法に

依らず，ミニマックス$\hat{\phi}$の収束の速さを求めることを考えよう。

§2. ミニマックス推定量$\tilde{\phi}_n$の収束の速さ

事前情報S_nのサンプル・サイズをnとする。すなわち，$(S_n, \mathcal{B}, \mathcal{Q}_1) = (\mathcal{X}^n, \mathcal{A}^n, \mathcal{P}_1^n)$ なる積空間において考える。

また，簡便のため，ここでは，制約条件を

$$(2.1) \quad \int_{\mathcal{X}} f(x)\phi(x)\nu(dx) \leq c \quad (\alpha = c)$$

とし，目的関数を

$$(2.2) \quad v(\phi, p) = \int_{\mathcal{X}} g(x)\phi(x)p(dx)$$

と置く。ここでは，Φは，$0 \leq \phi(x) \leq 1$ なる$\mathcal{A}$－可測関数の全体とする。

また損失は，

$$(2.3) \quad L(\phi, p) = \int_{\mathcal{X}} g(\phi_p - \phi)\,p(dx)$$

となり，危険関数は，

$$(2.4) \quad r(\hat{\phi}, p) = \int_{\mathcal{X}^n}\int_{\mathcal{X}} g(\phi_p(x) - \hat{\phi}(x, s))\,p(dx)\,p^n(ds)$$

となる。

さて，S_nをもとにした経験分布を$\hat{P}_n(\cdot \mid s)$で表わすことにする。

まず幾つかの lemma を用意しよう。

$\boxed{\text{Lemma 2.1}}$ (Dudley [1] の Proposition の拡張)

$g(x)$を$\forall p \in \mathcal{P}_1$について自乗可積分であるとする。 そして

(2.5) $$G_p(A ; g) = \int_A g(x)p(dx), \quad A \in \mathfrak{A}, \quad p \in \mathcal{P}_1.$$

と記すことにしよう。

このとき，$A \in \mathfrak{A}$の任意の有限個の分割

$$\{A_j ; j=1,2,\cdots,m\}$$

$$A = \bigcup_{j=1}^{m} A_j, \quad A_i \cap A_j = \phi \quad \text{for } i \neq j$$

について，

(2.6) $$E_{p^n} \sum_{j=1}^{m} \left| \int_{A_j} g(x)\{\hat{p}_n(dx|s) - p(dx)\} \right| \leq \left[\frac{m}{n} G_p(A ; g^2)\right]^{\frac{1}{2}}$$

さて，$(\mathfrak{X}, \mathfrak{A})$上の確率測度で，$f(x)$を可積分，$g(x)$を自乗可積分とするようなものの全体を$\mathcal{P}_2 (\supset \mathcal{P}_1)$としよう。

このとき，$\mathcal{P}_2$に次のような距離ρを導入する。

(2.7). $$\rho(p_1, p_2) = \sup_{\phi \in \Phi} \left| \int_{\mathfrak{X}} g(x)\phi(x)(p_1(dx) - p_2(dx)) \right|, \quad p_1, p_2 \in \mathcal{P}_2.$$

ここで，$E_{p^n}\rho(\hat{p}_n, p)$を評価したいのであるが，そのために次のような準備をする。

$\mathfrak{X}$を separable metric space とし，$\mathfrak{A}$をその Borel-field としておく。

このとき，$\forall \varepsilon > 0$に対して，直径2ε以下なる被覆$\{U_j\}$

をとり，これが高々可算集合であるようにすることが出来る。
しかも，$\eta(\varepsilon)>0$ に対し，

$$(2.8)\qquad G_p(W(\varepsilon,p):(g)) < \eta(\varepsilon)$$

なる $W(\varepsilon,p)\in\mathfrak{O}$ を除けば，$\mathfrak{X}\sim W(\varepsilon,p)$ は有限個の $\{U_g\}$ でおおわれているわけである。 そこでその minimal number を $N(p,\varepsilon,\eta(\varepsilon))$ で表わすことにしよう。

$\boxed{\text{Lemma 2.2}}$

$k>2$ なる実数に対し，$0<K(p)<\infty$ があって，任意の $0<\varepsilon\leq 1$ について

$$(2.9)\qquad N(p,\varepsilon,\varepsilon^{k/(k-2)}) \leq K(p)\varepsilon^{-k}$$

が満たされているものとする。

このとき $0<M(k,K(p),G_p(\mathfrak{X}:g^2))<\infty$ があって

$$(2.10)\qquad \mathbb{E}_{p^n}\rho(\hat{P}_n(\cdot|s),p) \leq M n^{-\frac{1}{k}},\quad \text{for } p\in\mathcal{P}_2$$

が成立する。

次に，$s\in\mathcal{S}_n=\mathfrak{X}^n$ を固定するとき，$p_0\in\mathcal{P}_1$ を適当に選んで

$$(2.11)\qquad \mathcal{P}_0(s) \Leftarrow \{p : p\in\mathcal{P}_1,\ \rho(\hat{P}_n(\cdot|s),p)\leq\rho(\hat{P}_n(\cdot|s),p_0)\}$$

を定める。 このとき，$\mathcal{P}_0(s)$ は $\mathcal{P}_1$ の有界集合となるが，これが全有界であるように p_0 がとれることを仮定しよう。

この仮定のもとで，$\mathcal{P}_0(s)$ について，$\forall\varepsilon>0$ に対し，直径 2ε

以下の有限個の被覆 $\{\mathcal{P}_{\varepsilon,i} : i=1,2,\cdots,m(\varepsilon,s)\}$ をとる。各 $\mathcal{P}_{\varepsilon,i}$ から代表元 $p_{\varepsilon,i}\in\mathcal{P}_{\varepsilon,i}$ を選んで固定する。このうち、$\rho(p_{\varepsilon,i},\hat{p}_n(\cdot|s))$ の最小値を実現するものを一つ固定して $\tilde{p}_n(\cdot|s)$ と記すことにしよう。この $\tilde{p}_n(\cdot|s)$ については $\mathcal{B}$-可測性をいっておく必要があるが、ここでは触れないでおく。

Lemma 2.3

Lemma 2.2 の仮定を仮定し、$\mathcal{P}_0(s)$ が全有界であるように $p_0\in\mathcal{P}_1$ がとれることを仮定すれば

(2.12) $$E_{p^n}\rho(\tilde{p}_n(\cdot|s),p)\le 4M(k,p)\,n^{-\frac{1}{k}} \quad \text{for } p\in\mathcal{P}_1.$$

が成立する。ただし、k, $M(k,p)$ は Lemma 2.2 で与えられるものである。

次に、制約条件 (2.1) のもとで、$s\in\mathcal{S}_n$ を固定するとき

(2.13) $$v(\hat{\phi},p)=\int_{\mathcal{X}} g(x)\hat{\phi}(x,s)\tilde{p}_n(dx|s)$$

の sup を実現する $\hat{\phi}\in\hat{\Phi}(\mathbb{C})$ を $\hat{\phi}_{\tilde{p}_n}(x,s)$ と記すことにしよう。

Lemma 2.4

$s\in\mathcal{S}_n$ を固定して考える。$\nu\ll p_\nu$ とするとき、$(\mathcal{X},\mathfrak{A})$ 上の確率測度 $\tilde{p}_n^*$ を

(2.14) $$\tilde{p}_n^*(dx|s)=\frac{1}{2}\left[p_\nu(dx)+\tilde{p}_n(dx|s)\right].$$

でもって定義する。このとき，$\hat{P}_n$，ν の $\tilde{P}_n^*$ に関する一般化された密度を $\frac{d\hat{P}_n}{d\tilde{P}_n^*}(x,s)$，$\frac{d\nu}{d\tilde{P}_n^*}(x,s)$ として，これらが $\mathcal{A}\times\mathcal{B}$－可測であることを仮定しておく。

このとき，(2.13) の sup を実現する $\hat{\phi}_{\tilde{P}_n}(x,s)\in\hat{\Phi}$ が存在し，またこれは，$\tilde{P}_n^*$－測度 0 を除外して，

$$(2.15)\quad \hat{\phi}_{\tilde{P}_n}(x,s)=\begin{cases}1 & \text{for}\quad g(x)\frac{d\hat{P}_n}{d\tilde{P}_n^*}(x,s)-\hat{\lambda}_n(s)f(x)\\ & \qquad\times\frac{d\nu}{d\tilde{P}_n^*}(x,s)>0\\ 0 & \text{for}\qquad\qquad {}^{\prime\prime}\qquad\qquad <0\end{cases}$$

となることが必要十分である。ただし，$\hat{\lambda}_n(s)$ は，$\mathcal{B}$－可測関数で (2.15) を満たす $\hat{\phi}_{\tilde{P}_n}(x,s)$ が制約条件 (2.1) を満足するように定められる。

Lemma 2.5

Lemma 2.3，2.4 の仮定のもとで

$$(2.16)\qquad r(\hat{\phi}_{\tilde{P}_n},p)\le 4M(k,p)n^{-\frac{1}{k}},\qquad p\in\mathcal{P}_1$$

が成立する。

以上の準備のもとで，ミニマックス推定量 $\tilde{T}_n$ の収束の速さは次のように示される。

$\boxed{\text{Theorem 2.6}}$

次のような仮定を置く。

1) Lemma 2.2の仮定。特に $K(p)$. $G_p(\mathfrak{X}: g^2)$ が $p \in \mathcal{P}_1$ について有界であること。したがって，

(2.17) $$M(k, p) \leq \bar{M}(k) < \infty \quad \text{for all } p \in \mathcal{P}_1$$

なる $\bar{M}(k)$ の存在を仮定しておく。

2) 距離 ρ によって $\mathcal{P}_0(s)$ が全有界となるように $p_0 \in \mathcal{P}_1$ が選べる。

3) Lemma 2.4 の $\dfrac{d\hat{p}_n}{d\hat{p}_n^*}$，$\dfrac{d\nu}{d\hat{p}_n^*}$ の $\mathcal{A} \times \mathcal{B}$-可測性を仮定する。

以上の仮定のもとでミニマックス推定量 $\hat{\varphi}_n(x, s)$ について

(2.18) $$r(\hat{\varphi}_n, p) \leq 4\bar{M}(k) n^{-\frac{1}{2k}} \quad \text{for all } p \in \mathcal{P}_1.$$

が成立する。

(証). 以上のlemmaにより，

(2.19) $$\sup_{p \in \mathcal{P}_1} r(\hat{\varphi}_n, p) = \inf_{\hat{\varphi} \in \hat{\Phi}(\mathbb{C})} \sup_{p \in \mathcal{P}_1} r(\hat{\varphi}_n, p)$$

$$\leq \sup_{p \in \mathcal{P}_1} r(\hat{\varphi}_{\hat{p}_n}, p) \leq 4\bar{M}(k) n^{-\frac{1}{2k}}$$

したがって，(2.18)の成立を見る。

$\boxed{\text{Proposition 2.7}}$ (Dudleyのpropositionの拡張)

$\mathfrak{X}$ を p 次元ユークリッド空間，$\mathcal{A}$ をそのBorel-fieldとする。

このとき，

(2.20) $\alpha = pk/(k-p)(k-2)$, $k>p$, $k>2$

なるαについて

(2.21) $G_p(\mathcal{X} : |x|^{\alpha}|g(x)|) < \infty$

が成立すれば Lemma 2.2 の仮定が満足される。

$\boxed{例}$

〈1〉 $(\mathcal{X}, \mathcal{O}\!\mathcal{L})$ をユークリッド空間として，その上の確率測度の族 $\mathcal{P}_1 = \{P_{(\mu,\sigma^2)} ; -k_1 \le \mu \le k_2, \ \delta \le \sigma \le k_3\}$ を $P_{(\mu,\sigma^2)}$ が正規分布 $N(\mu, \sigma^2)$ であるようにとっておく。ただし，k_1, k_2, k_3, δ は適当な正数である。

このとき，Theorem 2.6 の条件 2), 3) が満足されることは明白であろう。

条件 1) については，Proposition 2.7 の適用を考えればよい。$g(x)$ が簡単に x などであれば Proposition 2.7 の仮定を満足していることは見やすい。

$G_p(\mathcal{X} : g^2)$, $L(P)$ は，例えば $k=3$ としておくと，それぞれ $\mu^2+\sigma^2$, $\int_{\mathcal{X}} |x|^{5/2} P_{(\mu,\sigma^2)}(dx)$ となるが，いずれも $-k_1 \le \mu \le k_2$, $\delta \le \sigma \le k_3$ で有界（連続）である。したがって Theorem 2.6 が，この場合 成立する。

$g(x)$ が連続関数の場合は，適当な区間で多項式近似を考えればよい。

<2>

特にいま，cを正数として，制約条件を

$$(2.22)\quad \int_{\mathcal{X}} \phi(x)\,\nu(dx) \leq c$$

とし，目的関数を

$$(2.23)\quad \int_{\mathcal{X}} x\,\phi(x)\,p_{(\mu,1^2)}(dx),\quad -K_1 \leq \mu \leq K_2$$

としよう。　ただし，νはルベック測度である。

このとき，最適領域ϕ_pは，

$$(2.24)\quad \phi_p(x) = \begin{cases} 1 & \text{for } x\,p_{(\mu,1^2)}(x) - \lambda_\mu > 0, \\ 0 & \text{for } \quad \text{〃} \quad < 0 \end{cases}$$

で与えられる。　ただしλ_μは(2.22)を満足するような正の数である。　また$p_{(\mu,1^2)}$は$P_{(\mu,1^2)}$の密度

$$p_{(\mu,1^2)} = \frac{1}{\sqrt{2\pi}} \exp\left[-\frac{(x-\mu)^2}{2}\right]$$

である。

このミニマックス推定量$\tilde{\phi}_n$は，

$$(2.25)\quad \tilde{\phi}_n(x,s) = \begin{cases} 1 & \text{for } x\,p_n(x,s) - \tilde{\lambda}_n(s) > 0, \\ 0 & \text{for } \quad \text{〃} \quad < 0 \end{cases}$$

$$p_n(x,s) = \frac{\sqrt{n}}{\sqrt{2\pi}\sqrt{n+1}} \exp\left[-\frac{(x-\bar{x})^2}{2\cdot\frac{n+1}{n}}\right],$$

$$\bar{x} = \sum_{i=1}^{n} x_i / n, \quad s = (x_1, x_2, \cdots, x_n)$$

として与えられる。　ただし，$\tilde{\lambda}_n(s)$は正値可測関数で制約条件(2.22)を満足するように定められる。

文献

[1] Dudley, R.M: "The speed of mean Glivenko-Canteli convergence", A.M.S., Vol. 40, No. 1, 40~50, (1969)

[2] Noda, K & Taga, Y.; "Bayes and minimax estimation methods for the optimum decomposition of a sample space based on prior information", Review of I.S.I., Vol. 38; 3, 344-350 (1970)

[3] Noda, K & Taga, Y.; "Minimax estimation methods for the optimum decomposition of a sample space based on prior information"
(A.I.S.M. Vol. 23. No. 1 に発表予定 (1971)).