一般統計推論から (その III)*

——統計的決定函数——

松 下 嘉 米 男

(1959 年 5 月受付)

The Mathematical Concept of Statistical Decision Function

Kameo MATUSITA

ここでは今迄われわれが理論的中心問題として述べてきた一般統計推論のうちから統計的決定函数の理論をとりあげ，その数理的構造の本質を明らかにしたいと思う．このことはまた今後の研究上重要なものと思われる．そのため必要な一連の基本的な仮定の設定から始め，この仮定の一連の系に基ずいて統計的決定函数の理論の体系を組みあげてみることとする．

1. 基 本 仮 定 I.

R は標本空間を示すものとする．

R に関する仮定 I

　R の部分集合から成る Borel 集合体 $\mathfrak{B}$ 及び $\mathfrak{B}$ の上の測度 m が定義されている．

Ω に関する仮定 I

　Ω の各分布整数 F は m に関して絶対連続， 即ち各 F は $F(E)=\int_E p(x)dm$ ($p(x)$ は密度函数) と表わされる．

Ω に次のような距離 ρ_Ω を導入する．即ち分布函数の系列 $\{F_n\}$ に対して，$\rho_\Omega(F_n, F)\to 0$ ($n\to\infty$) ならば，E に関して一様に $F_n(E)\to F(E)$ ——E は $\mathfrak{B}$ に属する任意の部分集合—— となるように入れる．この条件は次のようにも表わされる．

$$\int_R |p_1(x)-p_2(x)|dm \longrightarrow 0$$

ここに $p_1(x)$，$p_2(x)$ はそれぞれその密度函数を示すものとする． 故に $\rho_\Omega(F_1, F_2)=\int_R |p_1(x)-p_2(x)|dm$ とおくことができる．また $\sqrt{p(x)}$ は m に関して L^2–class に属するから，

$$\rho_\Omega(F_1, F_2)=\left\{\int(\sqrt{p_1(x)}-\sqrt{p_2(x)})^2dm\right\}^{\frac{1}{2}}$$

とおくこともできる．これによつて定義された距離が前述の性質をもつことは容易に示される．

Ω に関する仮定 II

　Ω は距離 ρ_Ω に関するコンパクトな距離空間である．

この仮定から次の系が得られる．

系．Ω は第二可算公理を充たす．

次に D は凡ての決定 (decision) の集合とし，$L(F, d)$ は F が真の分布函数であるとき，D の中の決定 d をとることによつて受ける損失を表わす損失函数であるとする．

L に関する仮定 I

---

* (III) は 1958 年，I. C. A. R. Statistical Wing に於ける講義のプリントから決定函数に関する部分を藤本熙と石井恵一が抄訳編集したものである．

$L(F, d)$ は積空間 $\Omega\times D$ の上で定義された負の値をとらない函数である．

L に関する仮定 II

D の凡ゆる d に対して，$L(F, d)$ は F に関する連続函数である．

L に関する仮定 III

$L(F, d)$ は有界な函数である．

$$0\leqq L(F, d)\leqq 1 \quad (\text{あらゆる } F \text{ 及び } d \text{ に対して})$$

と仮定する．次に D の要素 d_1, d_2 の間の距離を

$$\rho_D(d_1, d_2)=\max_{F\in\Omega}|L(F, d_1)-L(F, d_2)|$$

によつて与える．このようにして D は距離空間となる．

D に関する仮定 I

D は ρ_D に関するコンパクトな距離空間である．

この仮定から，

系．D は第二可算公理を充たす．

さて R から D の上への写像を統計的決定函数という．またこの統計的決定函数 $d=\phi(x)$ は D の任意の開集合 e に対して集合 $\{x\,;\,\phi(x)\in e\}$ が $\mathfrak{B}$ に属すならば，$\mathfrak{B}$–可測という．

系．$\mathfrak{B}$–可測な決定函数の系列 $\{\phi_n(x)\}$ が R の凡ゆる点 x において決定函数 $\phi(x)$ へ収束するときは，$\phi(x)$ もまた $\mathfrak{B}$–可測である．

$\mathfrak{D}$ に関する仮定

$\mathfrak{B}$–可測なあらゆる決定函数の集合 $\mathfrak{D}$ が，実験する者の選択しうる決定函数の全体である．

2. 補　　題

前述の仮定にしたがつて，次の一連の補題が証明できる．

補題 1.　$L(F, d)$ はコンパクトな距離空間 $\Omega\times D$ の上の連続函数である．

補題 2.　Ω の任意の F に対して，$L(F,\phi(x))$ は x の可測函数である．

補題 3.　$F_n\to F$ (距離 ρ_Ω の意味で)，$\phi_n(x)\to\phi(x)$ $(n\to\infty)$ のとき，$r(F_n, \phi_n)\to r(F, \phi)$, $0\leqq r(F,\phi)\leqq 1$. ここに $r(F, \phi)$ は

$$r(F, \phi)=\int_R L(F, \phi(x))dF(x)$$

つまり危険度函数である．

補題 4.　$r(F, \phi)$ は ϕ に関して一様に，Ω における F の函数として連続である．

次に決定函数の空間 $\mathfrak{D}$ の2要素 ϕ_1, ϕ_2 の間の距離を

$$\rho_{\mathfrak{D}}(\phi_1, \phi_2)=\max_{F\in\Omega}|r(F, \phi_1)-r(F, \phi_2)|$$

によつて定義する．

$\rho_{\mathfrak{D}}(\phi_1, \phi_2)\leqq 1$ は明らかである．次に $\mathfrak{D}^*$ は距離 $\rho_{\mathfrak{D}}$ が 0 であるような要素を同一視して $\mathfrak{D}$ から得られた距離空間であるとする．

補題 5.　$\mathfrak{D}^*$ はコンディショナリイ・コンパクト空間である (任意の列が基本列を部分列として含む).

補題 6.　$\phi_n(x)\to\phi(x)$ $(n\to\infty)$ のとき，F に関して一様に $r(F, \phi_n)\to r(F, \phi)$ である．

$\gamma(\mathfrak{D}^*)$ は $\mathfrak{D}^*$ における凡ての基本列の作る距離空間であるとする．二つの基本列 $\phi_1^*=\{\phi_{1n}\}$, $\phi_2^*=\{\phi_{2n}\}$ の間の距離は $\rho_{\mathfrak{D}}(\phi_1^*, \phi_2^*)=\lim_{n\to\infty}\rho_{\mathfrak{D}}(\phi_{1n}, \phi_{2n})$ で与えられる．このとき $\gamma(\mathfrak{D}^*)$ は完備かつ有界であり，コンパクトとなる．$\mathfrak{D}^*$ は明らかに $\gamma(\mathfrak{D}^*)$ に埋込まれ，$\mathfrak{D}^*$ の閉苞が $\gamma(\mathfrak{D}^*)$ である．また $\Omega\times\mathfrak{D}^*$ の上で定義された $r(F, \phi)$ は連続的に $\Omega\times\gamma(\mathfrak{D}^*)$ の上に拡張される．それ故決定函数の概念を全空間 $\gamma(\mathfrak{D}^*)$ へ拡げることができる．$\mathfrak{D}^*$ に属さぬ $\gamma(\mathfrak{D}^*)$ の要素 ϕ^* は出

意の正数 ϵ に対し，またすべての F に対して，

$$|r(F,\ \phi^*)-r(F,\ \phi)|<\epsilon$$

となるような $\mathfrak{D}^*$ の中の決定函数 ϕ が存在することを意味する．

標本空間 R が可算集合である場合のように，$\mathfrak{D}^*$ がはじめからコンパクトであるときには，勿論 $\mathfrak{D}^*=\gamma(\mathfrak{D}^*)$ である．以下 $\gamma(\mathfrak{D}^*)$ を単に $\mathfrak{D}^*$ とかくことにする．従つて，$\mathfrak{D}^*$ はコンパクトな距離空間を表わす．そして，

$\mathfrak{D}^*$ は第二可算公理を充たす．

3. 基本仮定 II.

上に仮定した損失函数の有界性を除くためには他の仮定を付け加えなければならない．このような仮定の系の一つとして次のようなものが考えられる．

R に関する仮定 I

Ω に関する仮定 I

Ω に関する仮定 II

L に関する仮定 I

D に関する仮定 II

D は二要素の間の距離が

$$\rho_D(d_1,\ d_2)=\max_{F\in\Omega}|L(F,\ d_1)-L(F,\ d_2)|$$

によつて与えられる局所コンパクトな距離空間である．（そうすると補題 1 及び 2 が得られる．）

$\mathfrak{D}^*$ に関する仮定

実験する者が選択出来る決定函数の族 $\mathfrak{D}^*$ は，二点（決定函数）ϕ_1, ϕ_2 の間の距離が

$$\rho_{\mathfrak{D}}(\phi_1,\ \phi_2)=\max_{F\in\Omega}|r(F,\ \phi_1)-r(F,\ \phi_2)|$$

によつて与えられるようなコンパクトな距離空間である．

そのすべての点は条件：

すべての F に対して $\int_R L^2(F,\ \phi)p(x)dm \leqq G$ ここに G はある定数である

を充たす．

4. 補　　題

主な結果の証明のために，次の補題が必要になる．

補題 7. (Helly の定理)

$\{\mu_n\}$ は第二可算公理を充たすコンパクトな距離空間の凡ての開集合によつて生成されるボレル (Borel) 集合体の上の分布函数の系列で，通常の意味で μ へ収束するものであるとする．（通常の意味でというのは——$\mu(\overline{E}-E^\circ)=0$ のような任意の集合 E に対して，$\mu(E)$ へ収束することを意味する．ここで $\overline{E}$ 及び E° は E の閉苞及び開核を示す．）そのとき任意の連続函数 $f(x)$ 及び $\mu(\overline{E}-E^\circ)=0$ なる任意の集合 E に対して

$$\int_E f(x)d\mu_n \to \int_E f(x)d\mu\ \ (n\to\infty).$$

補題 8. 第二可算公理を充たすコンパクトな距離空間の凡ての開集合によつて生成されるボレル集合体の上の凡ての分布の集合は通常の収束によつて与えられる位相に関してコンパクトな空間である．

空間 R のボレル集合体 B に関して可測な集合 E が測度 μ のアトムであるというのは，

$\mu(E)\neq 0$ で，かつ任意の可測な $E'\subseteq E$ に対して $\mu(E')=0$ かまたは $\mu(E')=\mu(E)$ となることであると定義する．

補題 9. (J. Ville の定理)

m, n 型の，実数を要素とする行列 (a_{ij}) に対して次のことが成立つとする．

任意の負でない数の一組 $\xi_1, \xi_2, \cdots, \xi_m$ に対して，$\sum_{i=1}^{m}\xi_i a_{ij}\ (j=1,2,\cdots,n)$ の中で，少くとも一つ負でない値をとるものがある．

そうすると，負でない一組 $\eta, \cdots, \eta_n$ で，その和が1となり，しかもすべての $\sum_{j=1}^{n}a_{ij}\eta_j\ (n=1, \cdots, m)$ が負にならないようなものが存在する．

補題 10. $\mu_1, \mu_2, \cdots, \mu_p$ を有限かつアトムをもたない測度，$\eta_1, \eta_2, \cdots, \eta_n$ を $\sum_{i=1}^{n}\eta_i=1$ をみたす負でない数の組とする．そうすると，n 個の互いに素な部分集合 $S_1, S_2, \cdots, S_n$ による R の分割で

$$\mu_k(S_j)=\eta_i\mu_k(R), \quad (j=1, \cdots, n; k=1, \cdots, p)$$

となるものが存在する．

5. 主な結果

この節では以上述べてきた仮定から統計的決定函数の問題についての最適解の存在することを示すことができ，またその主な性質を導くことのできることについて述べる．

$\mathfrak{S}$, $\mathfrak{F}$ はそれぞれ Ω, $\mathfrak{D}^*$ の凡ての開集合によつて生成されるボレル集合体であるとする．そこで集合体 $\mathfrak{S}, \mathfrak{D}^*$ の上の分布 μ, δ を考え，それらを Ω, $\mathfrak{D}^*$ の上の分布とよぶ．μ またはこのア・プリオリの分布，δ は確率化された（または混合型）決定函数ともよぶ．$\Omega\times\mathfrak{D}^*$ は計量化可測であり，$r(F, \phi)$ は $\Omega\times\mathfrak{D}^*$ の上で連続であるから，$r(F, \phi)$ は $\mathfrak{S}\times\mathfrak{F}$ 可測であり，

$$r(\mu, \delta)=\int_{\Omega}\int_{\mathfrak{D}^*}r(F, \phi)d\mu d\delta$$

を考えることができる．

M 及び $\varDelta$ は Ω 及び $\mathfrak{D}^*$ の上の凡ての分布の集合を表わすとすると，M, $\varDelta$ は通常の収束によつて与えられる位相に関してコンパクトである．

補題 7 から

定理 I. $r(\mu, \delta)$ はコンパクト空間 $M\times\varDelta$ の上の連続函数である．

またこの結果として

定理 II. $\max_{\mu\in M} r(\mu, \delta)$ 及び $\min_{\delta\in\Delta} r(\mu, \delta)$ はそれぞれ $\varDelta$ 及び M の上の連続函数である．

それ故，

$$\min_{\delta\in\Delta}\max_{\mu\in M} r(\mu, \delta)=r(\mu_0, \delta_0)$$

なる μ_0, δ_0 が存在する．

このような δ_0 はミニマックス解と呼ばれる．ミニマックス解はある意味で統計的決定の問題への最良な解であることが判る．

任意の $\mu\in M$ に対して，$r(\mu, \delta)$ を最小にする一つの $\delta_\mu\in\varDelta$ が存在する．このことは定理 I の直接の結果である．このような δ_μ は μ に関するベイズ解といわれる．

定理 III. 任意のア・プリオリの分布 μ に対して，μ に関するベイズ解 δ_μ が存在する．

定理 IV. μ に関する任意のベイズ解 δ_μ に対して

$$\delta_\mu\{\phi_0; r(\mu, \phi_0)=\min_{\phi} r(\mu, \phi)\}=1.$$

ここに $\quad r(\mu, \phi)=\int_{\Omega}r(F, \phi)d\mu.$

定理 V. $\{\delta_n\}$ はそれぞれがア・プリオリの分布 μ_n に関するベイズ解であるような，ア・プリオリの分布 $\{\mu_n\}$ の系列に対応するベイズ解の系列であるとする．また，$\{\delta_n\}$ は分布 δ_0 へ収束するとする．このとき δ_0 もまたあるア・プリオリの分布に関するベイズ解である．即ちベイズ解の系列の極限分布もまたベイズ解である．

$\max_{\mu}\min_{\delta} r(\mu, \delta)=r(\lambda, \delta_\lambda)$ のとき λ を最も都合の悪い分布 (least favourable distribution) とよぶ．

定理 VI. 最も都合の悪い分布が存在する．

定理 VII. $r_0=\max_{\mu}\ \min_{\delta}\ r(\mu,\delta)$ とするとき，

$$\Omega \text{ の凡ての } F \text{ に対して}\quad r_0 \geqq r(F, \delta_\lambda)=\int_{\mathfrak{D}^*} r(F, \phi)d\delta_\lambda$$

を充たす任意の最も都合の悪い分布に関するベイズ解が存在する．

定理 VIII. $\min_{\delta}\max_{\mu} r(\mu, \delta)=\max_{\mu}\min_{\delta} r(\mu, \delta)$

定理 IX. λ を任意の最も都合の悪い分布，δ_0 を任意のミニマックス解とする．このとき (λ, δ_0) は $r(\mu, \delta)$ の鞍状点，即ち $\max_{\mu} r(\mu, \delta_0)=r(\lambda, \delta_0)$，$\min_{\delta} r(\lambda, \delta_0)=r(\lambda, \delta_0)$ である．

定理 X. 任意のミニマックス解はあらゆる最も都合の悪い分布に関するベイズ解である．

定理 XI. 任意の最も都合の悪い分布 λ 及び任意のミニマックス解 δ_0 に対して

$$\delta_0\{\phi_0\,;\, r(\lambda, \phi_0)=\min_{\phi} r(\lambda, \phi)\}=1.$$

定理 XII. 任意のミニマックス解 δ_0 は，Ω の凡ての F に対して $r(F, \delta_0)\leqq r_0$ を満足する．逆に若し δ_0 が凡ての F に対して $r(F, \delta_0)\leqq r_0$ であるような $\mathfrak{D}^*$ の上の分布であれば，δ_0 は一つのミニマックス解である．

定理 XIII. 決定函数 δ_0 がミニマックス解であるための必要かつ充分な条件は，任意の決定函数 δ 及び凡ての F に対して

$$r(F, \delta_0)\leqq \max_{F} r(F, \delta)$$

であることである．

定理 XIV. 任意の正数 ε に対し，$\mathfrak{D}_1, \cdots, \mathfrak{D}_n$ を，$\mathfrak{D}^*$ の部分集合で次のような条件をみたすものとする．

$$\mathfrak{D}^*=\mathfrak{D}_1+\cdots+\mathfrak{D}_n,\quad \mathfrak{D}_i\cap\mathfrak{D}_j=0\ (i\neq j),$$

$$\phi, \phi'\in\mathfrak{D}_i \text{ に対し,}\quad |r(F, \phi)-r(F, \phi')|<\varepsilon.$$

また $\phi_1, \cdots, \phi_n$ を，それぞれ $\mathfrak{D}_1, \cdots, \mathfrak{D}_n$ からとつた任意の要素とする．そうすると，$\delta(\phi_1, \phi_2, \cdots, \phi_n)=1$ なる混合型決定函数 δ で，F に関して一様に $|r(F, \delta)-r(F, \delta_0)|<\varepsilon$ となるものが存在する．ここに，δ_0 はミニマックス解とする．

一般に，任意の δ に対して，$\sum\eta_i=1$，$\eta_i\geqq 0$ $(i=1, 2, \cdots, n)$ で，F に関して一様に $|r(F, \delta)-\sum_i r(F, \phi_i)\eta_i|<\varepsilon$ となる $\eta=(\eta_1, \cdots, \eta_n)$ が存在する．故に，大きさ ε の差異の範囲内では，有限個の決定函数を考えればよい．

定理 XV. 任意の最も都合の悪い分布 λ と，任意のミニマックス解 δ_0 に対して，

$$(*)\qquad r(F, \delta_0)=r_0$$

が，Ω_λ に属するすべての F に対して成り立つ．ここに，Ω_λ は，F を含む任意の開集合が，正の λ-測度をもつような F の集合である．逆に，ある最も都合の悪い分布 λ に対して，(*) が成り立つならば，δ_0 はミニマックス解である．

Ω のすべての F に対して

$$r(F, \delta^*) \leqq r(F, \delta)$$

で，かつ少なくとも一つの F_0 に対して

$$r(F_0, \delta^*) < r(F_0, \delta)$$

であるとき，決定函数 δ^* は決定函数 δ よりも一様に良いという．決定函数 δ^* よりも一様に良い決定函数がないとき，決定函数 δ^* はアドミッシブル (admissible) であるという．決定函数の族 C が，C に属さない任意の $\delta(\in \mathfrak{D}^*)$ に対して，δ より一様に良い決定函数を要素としてもつとき，C は $\mathfrak{D}^*$ に関して完備 (complete) であるという．

定理 XVI. あらゆるベイズ解の族は $\mathfrak{D}^*$ に関して完備である．

今まで確率化された決定函数——それは，もちろん，確率化されていない決定函数の概念を含むものであるが——を考えてきた．しかし，次の定理によつて，ある条件の下では，確率化された決定函数は考えなくてもすますことができる．

定理 XVII. Ω の各 F にアトムがなく，かつ Ω の任意の有限個の分布 $F_1, \cdots, F_i$ に対し，$F_1, \cdots, F_i$ に関してなされる決定が有限個しかないならば，任意の確率化された決定函数 δ に対して次のような“純”決定函数 $\phi_{(x)}$ が存在する．

Ω のすべての F に対し $r(F, \phi) = r(F, \delta)$

6. 費用を考慮する場合

さて，決定函数を採るとき，実験の費用を考慮に入れてみよう．費用は，明らかに ϕ 及び得られた標本 x の函数である．これを $c(x, \phi)$ で表わすことにする．許される費用の限界は普通，実験の前から定められている．これを定めるのに二つの方法がある．一つは，各 x に対して $c(x, \phi)$ を制限することであり，他の一つは，$c(x, \phi)$ の平均値を制限することである．どちらの場合にも次の仮定をおく．

c に関する仮定 I. 費用函数 $c(x, \phi)$ は，x を固定したとき ϕ の連続函数である．

上の二つの場合のうち，第一の場合には，$c(x, \phi) \leqq K$ なる定数 K が定められる．そのとき，すべての x に対して $c(x, \phi) \leqq K$ となるような決定函数の全体 $\mathfrak{D}^{**}$ は $\mathfrak{D}^*$ の閉部分集合を作る．したがつて，$\mathfrak{D}^{**}$ は第二可算公理をみたすコンパクトな距離空間である．故に，$\mathfrak{D}^*$ の代りに $\mathfrak{D}^{**}$ を考えることにより，今までの理論が適用できる．

第二の場合には，更に次のような二つの仮定を置けばよい．

c に関する仮定 II. 費用函数 $c(x, \phi)$ は，ϕ の函数とみなしたとき，x に関して一様に，連続である．

c に関する仮定 III. 費用函数 $c(x, \phi)$ は x の可測函数である．

これらの仮定があれば，費用の F に関する平均値 $\int_R c(x, \phi)dF$ が定義できて，ϕ の連続函数になる．すべての F に対して $\int_R c(x, \phi)dF \leqq K$ をみたす決定函数の全体は，コンパクトな距離空間を作り，そこで再び前の理論が適用できる．

7. 確率化の二方法

決定函数の確率化に二つの方法がある．一つはここで今まで考えた方法であり，他の一つは確率化された決定函数として，標本空間から決定空間上の確率分布全体の作る空間の上への写像を考える方法である．ところで，この二つの確率化の方法は同等であることを示すことができる．しかし，ここではこの議論に立ち入ることはやめ，これでこの稿をとじることにする．

統 計 数 理 研 究 所