

真推定における admissibility について

早大. 理工. 草間 時武

§0. 真推定論における admissibility の研究にはいくつかの論文があるが

本講演ではその中で最も重要と思われる二つの論文

J. Sacks *Generalized Bayes solutions in estimation problems*

A.M.S. 34 (1963)

S. Karlin *Admissibility for estimation with quadratic loss.* A.M.S. 29 (1958)

と中に述べたい。

§1. 決定関数論において Bayes 解の全体 \mathcal{B} の regular topology := \mathcal{B}

の closure \mathcal{B}^c が完全類であることは良く知られている。(Bayes 解、

regular topology 等については 工藤教授の講演において説明されるはずで

ある) しかし \mathcal{B}^c に属する決定関数がどのようにして求められるのは難か

い問題である。J. Sacks は密度関数 $p(x, \omega)$ ($\omega \in \Omega$) と損失関数 $W(\omega, t)$

に制限を与えて \mathcal{B}^c に属する決定関数を求める方法を示した。 $F \in \text{a priori}$

measure (通常確率測度とする) とするとき F に関する Bayes 解は標本 x

に対して
$$\int_{\Omega} W(\omega, t) p(x, \omega) F(d\omega) / \int_{\Omega} p(x, \omega) F(d\omega) \dots \dots \dots (1)$$

を最小にする t を決定を行うものである。

ここで F が確率測度と仮定せず Ω の有界集合 S には $F(S) < \infty$ となる measure (したがって Ω が有界でなければ $F(\Omega) = \infty$ であってもよい) であるときほとんどすべての x に対して (1) を最小にする決定 t を行う決定関数 δ を Sacks は一般化された Bayes 解と名づけた。

B^x から $\delta(\pm\infty | x) \neq 0$ とする x の measure が 0 でないような \mathcal{S} のなかに集合 B_0 とする。後述する仮定 1 より B_0 は完全類である。これから述べるいくつかの仮定のもとで $\delta \in B_0$ は一般化された Bayes 解であることを示すことが主要な目的でありその証明の大ざっぱな outline を述べる。

μ を実数上の σ -finite measure とする。 $\Omega = \{\omega \mid \int e^{x\omega} \mu(dx) < \infty\}$ とし $\omega \in \Omega$ に対して密度関数 $p(x, \omega)$ を $p(x, \omega) = p(\omega) e^{x\omega}$ とする。証明の本質にかわりないので $\Omega = (-\infty, \infty)$ とし決定空間も $[-\infty, \infty]$ とする。

regular topology の意味で δ_n が δ に収束すれば;

すべての $x \in E$ ($\mu(E) > 0$) とすべての有界集合 C に対して

$$\lim_{n \rightarrow \infty} \delta_n(C | x) = 0$$

ならばほとんどすべての $x \in E$ とすべての有界集合 C に対して

$$\delta(C | x) = 0 \quad \dots \dots \dots (2)$$

また C がコンパクトのときすべての $x \in E$ で

$$\lim_{n \rightarrow \infty} \delta_n(C | x) = 1$$

ならばほとんどすべての $x \in E$ で

$$\delta(C | x) = 1 \quad \dots \dots \dots (3)$$

に注意しておく。(3) はこの議論で重要である。

仮定は損失関数 $W(\omega, t)$ に関するものである。

仮定 1 $0 \leq W(\omega, t) < \infty$, $W(\omega, t)$ は ω と t に別して連続, ω のコンパクト集合上で一様に $W(\omega, t) \rightarrow \infty$ ($t \rightarrow +\infty, -\infty$)

假定2 $w > t > s$ なら $W(w, t) < W(w, s)$ かつ

$$\inf_{w \geq t} [W(w, s) - W(w, t)] > 0$$

$w < t < s$ ならば $W(w, t) < W(w, s)$ かつ

$$\inf_{w \leq t} [W(w, s) - W(w, t)] > 0$$

假定3 すべて t と $\varepsilon > 0$ について

$$\sup_{w \leq 0} W(w, t) e^{\varepsilon w} < \infty, \quad \sup_{w \geq 0} W(w, t) e^{-\varepsilon w} < \infty,$$

$$\lim_{A \rightarrow \infty} \sup_{w \geq A} W(w, t) e^{-\varepsilon w} = 0, \quad \lim_{A \rightarrow \infty} \sup_{w \leq -A} W(w, t) e^{\varepsilon w} = 0$$

假定1~3 をみたす損失関数 W はたとえは

$$W(w, t) = |w - t|^\alpha \quad (\alpha > 0)$$

がその例である。

定理 $\delta_n (n=1, 2, \dots)$ を Bayes 解の列として regular topology の意味で $\delta_n \rightarrow \delta$ であり $\delta \in \mathcal{B}_0$ ならば "ほとんどすべての $x \in (z_0, z_1)$ について $\delta(B_x | x) = 1$ がなりたつ。

ここに B_x は ある measure F (有界集合 S には $F(S) < \infty$) に対して

$$B_x = \left\{ t \mid \int_{-\infty}^{\infty} W(w, t) p(x, w) F(dw) \text{ を最小にする} \right\}$$

と表わされる集合であり, z_0, z_1 は measure μ の support の下限と上限とする。

証明 先ず Lemma 5-2 を証明なしで述べ置く。

Lemma 1. S_n を a priori measure ξ_n に関する Bayes 解とする。 $\{\xi_n\}$ の部分列 $\{\xi_{n_k}\}$ と a での性質をみたすものがある。

$$\forall b > 0 \text{ の } b < \infty \text{ に対し } \limsup_{k \rightarrow \infty} \xi_{n_k}(-b, b) / \xi_{n_k}(-a, a) < \infty \quad \text{----- (4)}$$

Lemma 2. ほとんどの $x \in (z_0, z_1)$ に対し

$$\lim_{n \rightarrow \infty} \int_{-\infty}^{\infty} p(x, \omega) F_n(d\omega) = \int_{-\infty}^{\infty} p(x, \omega) F(d\omega) < \infty$$

また t のコンパクト集合 E 上 t 一様に

$$\lim_{n \rightarrow \infty} \int_{-\infty}^{\infty} W(\omega, t) p(x, \omega) F_n(d\omega) = \int_{-\infty}^{\infty} W(\omega, t) p(x, \omega) F(d\omega) < \infty.$$

$\int_{-\infty}^{\infty} W(\omega, t) p(x, \omega) F(d\omega)$ は t の連続関数である。

定理の証明の概略 Lemma 1 により (4) をみたす $\{\xi_{n_k}\}$ と a が存在する。 $F_{n_k} = \xi_{n_k} / \xi_{n_k}(-a, a)$ とおくと (4) より $\{F_{n_k}\}$ の部分列 $\{F_{n_{k_j}}\}$ と measure F が存在して $F_{n_{k_j}} \rightarrow F$ (弱収束の意味で)。 (4) より F は有界集合 S で $F(S) < \infty$ であり $F(-a, a) = 1$ なるから $F(\Omega) \neq 0$ 。 $\{\xi_{n_{k_j}}\}$ は最初の一列 $\{\xi_n\}$ と仮定しても一般性を失わないので $F_n \rightarrow F$ (弱収束) とする。 S_n は F_n に関する Bayes 解である。

$$H(t, x) = \int_{-\infty}^{\infty} W(\omega, t) p(x, \omega) F(d\omega), \quad H(t, x, n) = \int_{-\infty}^{\infty} W(\omega, t) p(x, \omega) F_n(d\omega)$$

$$m(x) = \inf_t H(t, x), \quad m(x, n) = \inf_t H(t, x, n)$$

$$B_{x, r} = \{t \mid H(t, x) \leq m(x) + r\}, \quad B_{x, r}^n = \{t \mid H(t, x, n) \leq m(x, n) + r\}$$

($r \geq 0$) とおく。 $r = 0$ ならば $B_{x, 0} = B_x$ であり、 $S_n(B_{x, 0}^n \mid x) = 1$ ($n=1, 2, \dots$) である。 $S(B_{x, 0} \mid x) = 1$ (a.e. μ) であることと証明すればよい。

$$\lim_{r \rightarrow 0} S(B_{x,r} | x) = S(B_{x_0} | x) \text{ だから } S(B_{x_0} | x) = 1 \text{ (a.e. } \mu)$$

が成り立たないとするとき、 $\forall \varepsilon > 0, \exists x \in E_0$ ($E_0 \subset (z_0, z_1), \mu(E_0) > 0$)

$$t \quad S(B_{x,r} | x) \leq 1 - \varepsilon_0, \quad B_{x,r} \subset [-\sigma_0, \sigma_0] \quad \dots \dots \dots (5)$$

となる正数 $\varepsilon_0, \sigma_0, \varepsilon_0$ と E_0 が存在する。 仮定1 と Fatou の Lemma

より $H(t, x) (H(t, x, n)) \rightarrow \infty (t \rightarrow +\infty, -\infty)$ であるから $B_{x,r}$

$(B_{x,r}^n)$ は有界, $H(t, x) (H(t, x, n))$ は t の連続関数だから

$B_{x,r} (B_{x,r}^n)$ は closed (したがってコンパクトである。 したがって

σ_0 は存在する。

Lemma 2 により t のコンパクト集合上で $H(t, x, n)$ は $H(t, x)$ に
一様収束するから $m(x, n) \rightarrow m(x)$ 。 一様収束性から

$$B_{x,r}^n \subset [-\gamma_1, \gamma_1] \quad (n=1, 2, \dots) \quad x \in E_1, \quad r \leq r_0$$

となる $\gamma_1 > 0$ と $E_1 (\subset E_0, \mu(E_1) > 0)$ が存在する。 とわかる。

$r = r_0/4$ とする。 と十分大なる $n_0(x, r)$ が存在して

$$n \geq n_0(x, r) \longrightarrow B_{x, r/2}^n \subset B_{x,r}$$

$E_{kr} = \{x \mid B_{x, r/2}^n \subset B_{x,r}, n \geq k, x \in E_1\}$ とする。 $k \in \mathbb{N}$ と十分大

とすれば $\mu(E_{kr}) > 0$ 。 $x \in E_{kr}$ ならば

$$1 = S_n(B_{x, r/2}^n | x) \leq S_n(B_{x,r} | x) \quad \dots \dots \dots (6)$$

二つの場合に分けて考える。

(i) $x^* \in E_{kr}, \mu(\{x^*\}) > 0$ となる x^* が存在する場合。

(6)より

$$1 = \lim_{n \rightarrow \infty} S_n(B_{x^*, r/2}^n | x^*) \leq \lim_{n \rightarrow \infty} S_n(B_{x^*, r} | x^*)$$

$$\therefore \lim_{n \rightarrow \infty} S_n(B_{x^*, r} | x^*) = 1 \quad \dots \dots \dots (7)$$

$B_{x^*, r}$ がコンパクトであるから (7) より $S(B_{x^*, r} | x^*) = 1$ が成り

たに (5) に矛盾する。

(ii) $E_{k,r}$ が atom を持たぬ場合

$x^* \in E_{k,r}$ 且 $E_{k,r}$ に固有な密度が 1 の点 x^* が存在する。かかる x^* に
対しては どんな $\eta > 0$ に対しても $\mu\{[x^* - \eta, x^* + \eta] \cap E_{k,r}\} > 0$
がなりたつ。 Lemma 2 から $m(x)$ は連続関数 (x_0) である
から η を十分小さくとれば

$$y \in E^* = [x^* - \eta, x^* + \eta] \cap E_{k,r} \rightarrow B_{y,r} \subset B_{x^*, 2r} \subset B_{y, 4r}.$$

$4r = r_0$ 故に (5) より

$$S(B_{x^*, 2r} | y) \leq S(B_{y, 4r} | y) \leq 1 - \varepsilon_0 \quad (y \in E^*) \dots (8)$$

$E^* \subset E_{k,r}$ 故に $m \geq \varepsilon_0$ ならば $y \in E^*$ に対して

$$1 = S_n(B_{y, r_0}^n | y) \leq S_n(B_{y, r} | y) \leq S_n(B_{x^*, 2r} | y) \quad \text{したがって}$$

$$(3) \text{より } S(B_{x^*, 2r} | y) = 1 \quad \text{これは (8) と矛盾する。}$$

したがって (5) は正しいにたない。 故に $S(B_{x_0} | x) = 1$ (a.e. μ)

系 1 \mathcal{F} は有限集合 S ならば $F(S) < \infty$ とする measure F の,

その z の t と x に対して $H(t, x) < \infty$ とする F 全体とする。 $F \in \mathcal{F}$

に対して \mathcal{B}_F と $S(B_z | x) = 1$ (a.e. μ) とする S の全体とする。

$\mu(\{z_0\} \cup \{z, \frac{1}{2}\}) = 0$ とする仮定 1 ~ 3 のもとに

$$\mathcal{B}_0 \subset \mathcal{B}^* = \bigcup \{ \mathcal{B}_F \mid F \in \mathcal{F} \} \quad \text{したがって } \mathcal{B}^* \text{ は完全}$$

類である。

系 2 系 1 につけておいて $W(w, t)$ が $w \in \text{fix}$ と t の強の意味
の凸関数となるとき $\mathcal{B}_0 = \mathcal{B}^*$ となる。

系 1 は定理からすぐわかる ($\mu(\{z_0\} \cup \{z, \frac{1}{2}\}) = 0$ の仮定が必要
なことは講演の時の説明に中する)

系2は W が凸関数であるから B_2 は唯一からなる集合であり, (したがって) B_F は唯一の決定関数 δ_F からなる. δ_F は明らかに non-randomized である. $\delta_F(\pm\infty|x) = 0$ ($a.e. \mu$) だから $\delta_F \in B^*$ を示せば $\delta_F \in B$ であることがわかる. $F_n \in F$ の $[-n, n]$ への restriction とし $\psi_n = F_n/F[-n, n]$ とする. δ_n を確率測度 ψ_n に対する Bayes 解とすると $\delta_n \rightarrow \delta_F$ ($n \rightarrow \infty$) がなりたつ. 故に $\delta_F \in B^*$.

この問題において本質的な役割を演ずるのは Lemma 2 のように思われる (したがって密度関数がある一般的な場合にもこの定理はなりたつように思われる).

§2. Karlin の仕事は 臆推定論における admissibility の問題に大きな寄与をしたと思われる. 密度関数 $p(x, w)$ に対して $h(w)$ を推定する問題を考える. 損失関数は quadratic loss とする. そのときは non-randomized な決定関数のみを考慮せよ. $a(x)$ が admissible な推定量であることを証明するのに Karlin は次のような方法を用いた.

$$r(w, a) = \int_X [a(x) - h(w)]^2 p(x, w) d\mu(x). \quad \text{この危険関数と}$$

なるが $r(w, b) \leq r(w, a)$ ($w \in \Omega$) な推定量 $b(x)$ が存在

すれば $b(x) = a(x)$ ($a.e. \mu$) であることを証明すればよい.

$$r(w, b) \leq r(w, a) \quad (w \in \Omega) \text{ より}$$

$$\int_X [b(x) - a(x)]^2 p(x, w) d\mu(x) \leq 2 \int_X [a(x) - b(x)] [a(x) - h(w)] p(x, w) d\mu(x) \quad \dots \dots \dots (9)$$

が導かれる 単調増加関数 $F(w)$ で

$$\int_{\Omega} h(w) p(x, w) dF(w) = a(x) \int_{\Omega} p(x, w) dF(w) \quad \dots (10)$$

とみなすものが存在するとする。(9)の両辺を dF で積分し、積分順序の交換 $\mu(x)$ を用いれば

$$\int_X [b(x) - a(x)]^2 \left[\int_{\Omega} p(x, w) dF(w) \right] d\mu(x) \leq 0 \quad (w \in \Omega)$$

がなりたつ。故に $b(x) = a(x)$ (a.e. μ)。

この応用にはいくつかの応用がある。

先ず X が密度関数 $\beta(w)e^{wX}$ を持つ場合と考える。 Ω は

$$\int_{-\infty}^{\infty} e^{wX} d\mu(x) < \infty \text{ となる } w \text{ の全体で区間である。 } \theta(w) = E_w(X)$$

$= -\beta'(w)/\beta(w)$ と大小1の標本 X から推定する問題を考えれば十分である。

Exponential family における十分統計量はまた exponential family に属するからである。

定理2 $\underline{w}, \bar{w} \in \Omega$ の下限, 上限とする (\underline{w}, \bar{w}) の内点 c とする

$$b \text{ にとり } \int_c^b \beta^{-\lambda}(w) dw \rightarrow +\infty \quad (b \rightarrow \bar{w})$$

$$\int_a^c \beta^{-\lambda}(w) dw \rightarrow +\infty \quad (a \rightarrow \underline{w})$$

$a, b \in \Omega$

がなりたてば $(\frac{1}{\lambda+1})X$ は $\theta(w)$ の admissible な推定量である

この定理から Ω と μ の様子によって色々な場合に結果が得られる。

(A) $\Omega = (-\infty, \infty)$ と $\mu((0, \infty)) > 0, \mu((-\infty, 0)) > 0$ ならば

$0 < \delta \leq 1$ なる δ に対して δX は admissible な推定量である

(B) $\Omega = (-\infty, \infty)$ と $\mu(\{0\}) > 0$ ならば $0 < \delta \leq 1$ なる δ に対して

γ は admissible な推定量である。

0) $\Omega = (-\infty, \infty)$ だけしか与えられていないときは γ の admissibility がわかる。

一般に $\Omega = (a, \infty)$, 又は $(-\infty, b)$ (a, b は実数) のときは γ は admissible な推定量にならない場合が多い。

$$\gamma = p(x, w) = \begin{cases} f(w)r(x) & 0 \leq x \leq w \\ 0, & \text{otherwise} \end{cases}$$

の場合を考える。 $\Rightarrow r(x) > 0$, $f^{-1}(w) = \int_0^w r(x) dx < \infty$ ($w \in (0, \infty)$), $f(w) \rightarrow 0$ ($w \rightarrow \infty$) とする。

$[1/f(w)]^\alpha$ ($\alpha > 0$) を推定することにし $[1/f(x)]^\alpha$ なる形の推定量のみを考える。

定理 $r[1/f(x)]^\alpha$ なる形の推定量 γ が $[1/f(w)]^\alpha$ の admissible な推定量 となるのは $\alpha = (2\alpha + 1)/(\alpha + 1)$ の時のみである。

次の Translation parameter の推定問題を考える。 λ は密度関数 $p(x, w) = p(x + w)$ を持つとする ($p(\xi)$ は既知)。大きさ 1 の標本 x をもとにして $-w$ を推定するのに推定量 $g(x) = x$ の admissible であることが次の条件によって示される。

$$\int_{-\infty}^{\infty} \xi^2 p(\xi) d\xi < \infty, \int_{-\infty}^{\infty} \xi^2 p^2(\xi) d\xi < \infty, \int_{-\infty}^{\infty} \xi p(\xi) d\xi = 0. \dots\dots$$

このとき $g(x) = x$ は $|g(x) - x| \leq M < \infty$ をみたす推定量 $g(x)$ の全体のなす admissible である。

もしこれ以外の $g(x)$ については (9) が成り立たないことが証明できる。

ば $g(x) = x$ は δ の推定量の δ が δ admissible になる。
 これは例として $p(\xi)$ が $\xi < a, \xi > b, -\infty < a < b < \infty$ のとき $= 0$,
 その他のとき ≥ 0 のときである。また $p(x, w) = p(x+w)$ が単調
 尤度比を持つときは $|g(x) - x| \leq M < \infty$ とみたす $g(x)$ が δ を考
 慮する。一般に δ_0 と δ が違わない δ の方が δ は δ_0
 が admissible になると言うとき local admissible と言うことに
 すると, local admissibility の研究は一つの方向であろう。

今までは quadratic loss の場合のみと考へてきたが $W(w, a)$
 $= (a+w)^{2N}$ のときの translation parameter problem と考へる。

$$(1) p(\xi) \text{ が symmetric, } (2) \int_{-\infty}^{\infty} \xi^{2N} p(\xi) d\xi < \infty, \int_{-\infty}^{\infty} \xi^{2N} p^2(\xi) d\xi < \infty$$

のとき $|g(x) - x| \leq M < \infty$ ならば $g(x)$ の方が δ は admissible であ
 る。

最後に X の $p(x, w) = p(x+w)$ にしたときその n 個の観測結
 果 x_1, \dots, x_n から $-w$ を推定する δ と考へる Pitman の
 推定量は

$$\delta^*(x_1, x_2, \dots, x_n) = x_1 - T(y_2, \dots, y_n)$$

$$T(y_2, \dots, y_n) = \frac{\int_{-\infty}^{\infty} \xi p(\xi) \prod_{i=2}^n p(x_i + \xi) d\xi}{\int_{-\infty}^{\infty} p(\xi) \prod_{i=2}^n p(x_i + \xi) d\xi}$$

である Girshick Savage は quadratic loss で δ^* が minmax
 (unique) である δ と示し, Blackwell は $X \in w$ の整數値を
 とり $p(x+w)$ が有限個の δ で $\delta > 0$ である場合には
 δ^* の admissibility を証明した Karlin は δ^* の admissibility

次の三つの場合に証明(1)の条件は ξ のために成り立つ

ある。

$$\text{Case 1. } p(\xi) \begin{cases} \geq 0 & -\infty < a \leq \xi \leq b < \infty \\ = 0 & \xi < a, b < \xi \end{cases}$$

$$\Rightarrow \text{これは } |g(x_1, y_1, \dots, y_n) - g^*(x_1, y_1, \dots, y_n)| \leq M < \infty$$

なる g の方が g^* は admissible である。

$$\text{Case 2 } X = \{0, \pm 1, \pm 2, \dots\},$$

$$\Omega = \{0, \pm 1, \pm 2, \dots\}$$

$$P\{X=c | w\} = p(c+w), \quad \sum_{j=-\infty}^{\infty} p(j) = 1.$$

\Rightarrow これは

$$\sum_{j=-\infty}^{\infty} j^2 \sqrt{p(j)} < \infty \quad \text{なる假定のもとで } g^* \text{ の admissibility を}$$

証明する。

$$\text{Case 3. } 0 \leq p(\xi) \leq C, \quad \int_{-\infty}^{\infty} \xi^2 \sqrt{p(\xi)} d\xi < \infty \quad \text{かつ}$$

$$\frac{p(\xi) p(y_2 + \xi) \dots p(y_n + \xi)}{\int_{-\infty}^{\infty} p(\theta) p(y_2 + \theta) \dots p(y_n + \theta) d\theta} \leq C' \quad \text{かつ } \wedge 2 \text{ の}$$

ξ, y_2, \dots, y_n でありたつ場合

$$\Rightarrow \text{これは } |g(x_1, y_1, \dots, y_n) - g^*(x_1, y_1, \dots, y_n)| \leq M < \infty$$

となり g の方が g^* は admissible である。

