

推定可能な母数のノンパラメトリック・ベイズ推定

慶大・理 大和 元 (Hajime Yamato)

1. 序

推定可能な母数のベイズ推定を, 2乗誤差に基づき, 事前分布として Dirichlet invariant process (Dala [1979]) を用いて行なう。

ノンパラメトリックなベイズ推定の為の事前分布として, Dirichlet process (Ferguson [1973]) との違いは, Dirichlet invariant process ではパラメータの値が分布に反映する点である。例えば, R^1 上でパラメータを原点対称にとれば, Dirichlet invariant process に従う分布は確率1で原点対称になる。 R^2 上で直線 $y=x$ に関して対称なパラメータをとれば, Dirichlet invariant process に従う分布は確率1で R^2 上で直線 $y=x$ に関して対称になる。

Dala [1979a, b, 1980] は Dirichlet invariant

process を用いて, R^1 上の対称な分布関数および center of symmetry の推定を行っている。

ここでは, 推定可能な母数の推定への応用を考へる。

2. Dirichlet invariant process

X は d -次元 Euclidean space を表わし, A は d -次元 Borel class を表わすものとする。 $G = \{g_1, g_2, \dots, g_k\}$ は $X \rightarrow X$ の measurable transformation からなる finite group である。 α は (X, A) 上の有限な measure で G の下で invariant, i.e., $\alpha(B) = \alpha(g_i B)$ for $\forall B \in A, i=1, \dots, k$ を満たすものとする。 簡単のため, $M = \alpha(X)$, $Q(\cdot) = \alpha(\cdot) / M$ で表すことにする。

X の measurable partition B_1, B_2, \dots, B_m が G -不変 (invariant) であるとは,

$$B_1 \cup \dots \cup B_m = X, \quad B_i \cap B_j = \phi \quad (i \neq j)$$

$B_j = g_i B_j$ for $j=1, \dots, m, i=1, \dots, k$ が成り立つことである。

定義 1 (Dala [1979 b])。 分布 P が次の (1), (2) を満たすとき, (X, A) 上でパラメータ α の Dirichlet G -invariant process に従うという。

(1) P は確率 1 で, G の下で不変である。

(2) \mathcal{G} の任意の \mathcal{G} -不変な measurable partition B_1, B_2, \dots, B_m に対して, $(P(B_1), \dots, P(B_m))$ は Dirichlet 分布 $(\alpha(B_1), \dots, \alpha(B_m))$ に従う。

分布 P が $(\mathcal{X}, \mathcal{A})$ 上でパラメータ α の Dirichlet \mathcal{G} -invariant process に従うことを, 簡単に $P \in DI(\alpha)$ と書くことにする。

P^* が $(\mathcal{X}, \mathcal{A})$ 上でパラメータ α の Dirichlet process に従うとき

$$P(A) \equiv \frac{1}{k} \sum_{j=1}^k P^*(g_j A) \quad \text{for } A \in \mathcal{A} \quad (1)$$

によって定義される P は $P \in DI(\alpha)$ 。

定義 2 (Dala [1979 b]). 任意 $m (= 1, 2, \dots)$, 及び $A_1, \dots, A_m, C_1, \dots, C_m (\in \mathcal{A})$ に対して

$$\begin{aligned} \Pr \{ X_1 \in C_1, \dots, X_m \in C_m \mid P(A_1), \dots, P(A_m), P(C_1), \dots, P(C_m) \} \\ = \prod_{i=1}^m P(C_i) \quad \text{a.s.} \end{aligned}$$

が成り立つとき, $X_1, \dots, X_m \in P$ が size m の標本という。

命題 1 (Daha [1979b]). $P \in DI(\alpha)$,

X は P から n の size 1 の標本とすると, 任意の $g \in G$ に対して, X と gX は分布 Q に従う。

命題 2 (Daha [1979b]). $P \in DI(\alpha)$,

X_1, \dots, X_n は P から n の size 1 の標本とすると, X_1, \dots, X_n が与えられたとき $P \in DI(\alpha + \sum_{i=1}^n \delta_{X_i}^g)$, ここで $\delta_{X_i}^g = \sum_{j=1}^k \delta_{g_j X_i} / k$ 。

上の命題 2 により, 標本 X_1, \dots, X_n を次の様に sequential に得られたものと見なすことが出来る:

- (1) $P \in DI(\alpha)$ から size 1 の標本 X_1 をとる。
- (2) $P \in DI(\alpha + \delta_{X_1}^g)$ から size 1 の標本 X_2 をとる。

...

- (n) $P \in DI(\alpha + \delta_{X_1}^g + \dots + \delta_{X_{n-1}}^g)$ から size 1 の標本 X_n をとる。

従って, 命題 1 から, size n の標本 X_1, \dots, X_n の分布を次の様に考えることが出来る:

- (1) $X_1 \sim Q$
- (2) $X_2 \sim \frac{1}{M+1} (MQ + \frac{1}{k} \sum_{j=1}^k \delta_{g_j X_1})$, given X_1

...

$$(m) \quad X_m \sim \frac{1}{M+n-1} \left(MQ + \frac{1}{k} \sum_{j=1}^k (\int_{g_j} X_1 + \dots + \int_{g_j} X_{m-1}) \right) \\ , \text{ given } X_1, \dots, X_{m-1}.$$

3. ベイズ推定

$h(x)$ を $(\mathcal{X}, \mathcal{A})$ 上の measurable real valued f とし $\int_{\mathcal{X}} h(x) dQ(x) < \infty$ を満たすものとする。このとき、 $P \in DI(\mathcal{A})$ が \mathcal{X} の size 1 の標本を X とすると

$$\int_{\mathcal{X}} h(x) dP(x) = E[h(X) | P] \text{ であり, 命題1により}$$

$$E \int_{\mathcal{X}} h(x) dP(x) = E h(X) = \int_{\mathcal{X}} h(x) dQ(x) \quad (2)$$

が成り立つ。従って, degree 1 の推定可能な母数

$$\theta_1 = \int_{\mathcal{X}} h(x) dP(x)$$

に対して, $P \in DI(\mathcal{A})$ が \mathcal{X} の size n の標本 X_1, \dots, X_n に基づく θ_1 のベイズ推定量は, (2) と命題2より

$$\hat{\theta}_1 = E[\theta_1 | X_1, \dots, X_n]$$

$$= \int_{\mathcal{X}} h(x) d \frac{\alpha(x) + \sum_{i=1}^n \int_{X_i}^g(x)}{M+n}$$

$$= \frac{M}{M+n} \int_{\mathcal{X}} h(x) dQ(x) + \frac{1}{(M+n)k} \sum_{i=1}^n \sum_{j=1}^k h(g_j X_i).$$

例. $(\mathcal{X}, \mathcal{A}) = (R', B')$ とし, $G = \{e, g\}$, $e(x) = x$, $g(x) = 2c - x$ for $\forall x \in R'$ とする。但し, c は定数。

任意に固定した t に対して $h(x) = 1 (x \leq t), = 0 (x > t)$ とおくと, $\theta_1 = F(t)$, 但し F は分布 D に対応する分布関数である。 $F(t)$ のベイズ推定量は

$$\hat{F}(t) = \frac{M}{M+n} F_0(t) + \frac{1}{(M+n)k} \sum_{i=1}^n \left\{ \int_{X_i}^{(-\infty, t]} + \int_{2c-X_i}^{(-\infty, t]} \right\}$$

で与えられる。ここで, $\gamma_n = M/(M+n)$, $F_0(t) = Q((-\infty, t])$

Dalal [1979] は $F(t)$ のベイズ推定量として, 積分2乗誤差を用いて, 上と同じ $\hat{F}(t)$ を得ている。Dalal [1979] の方法は $\hat{F}(t) = E[F(t) | X_1, \dots, X_n]$ 1, Prop. 2 と (2) の特別な場合にわたる

$$E P(A) = Q(A) \text{ for } A \in \mathcal{A}$$

を用いている。ただし, $P \in \mathcal{D}_I(\mathcal{A})$ 。

次に, degree α の推定可能な母数のベイズ推定について考之まう (Yamato [1985])。

補題 1. $P \in \mathcal{D}_I(\mathcal{A})$,

$h(x, y)$: measurable real-valued ft. on $(\mathcal{X}^2, \mathcal{A}^2)$

& symmetric in x, y

$$\int_{\mathcal{X}^2} h(x, y) dQ(x) dQ(y), \int_{\mathcal{X}} h(x, gx) dQ(x) < \infty \quad (\forall g \in G)$$

このとき,

$$E \int_{\mathcal{X}^2} h(x, y) dP(x) dP(y)$$

$$= \frac{1}{M+1} \left[M \int_{\mathcal{X}^2} h(x, y) dQ(x) dQ(y) + \frac{1}{k} \sum_{j=1}^k \int_{\mathcal{X}} h(x, g_j x) dQ(x) \right].$$

(略証) $X_1, X_2 \in P \in DI(\alpha)$ からの size 2 の標本と
 すると

$$\begin{aligned} E \int_{\mathcal{X}^2} h(x, y) dP(x) dP(y) &= E E[h(X_1, X_2) | P] \\ &= E h(X_1, X_2) \\ &= E E[h(X_1, X_2) | X_1] \end{aligned}$$

4頁の下段のよりに考えよことにより,

$$= E \frac{1}{M+1} \left[M \int_{\mathcal{X}} h(X_1, x_2) dQ(x_2) + \frac{1}{k} \sum_{j=1}^k h(X_1, g_j X_1) \right].$$

さらに, X_1 について (分布 Q をもつ) 期待値をとる。

定理1. $P \in DI(\alpha)$,

X_1, \dots, X_n : P からの size n の標本

$$\theta_2 = \int_{\mathcal{X}^2} h(x, y) dP(x) dP(y)$$

$h(x, y)$ は補題1の条件を満たすものとする。

このとき, θ_2 のベイズ推定量 $\hat{\theta}_2$ は次のようになり。

$$\begin{aligned}
\hat{\theta}_2 &= \frac{M+n}{M+n+1} \left[\int_{\mathcal{X}^2} g_m^2 h(x, y) dQ(x) dQ(y) \right. \\
&\quad + \frac{2g_m(1-g_m)}{nk} \sum_{i=1}^n \sum_{j=1}^k \int_{\mathcal{X}} h(x, g_j X_i) dQ(x) \\
&\quad \left. + \frac{(1-g_m)^2}{n^2 k^2} \sum_{i_1, i_2} \sum_{j_1, j_2} h(g_{j_1} X_{i_1}, g_{j_2} X_{i_2}) \right] \\
&\quad + \frac{1}{(M+n+1)k} \left[g_m \sum_{j=1}^k \int_{\mathcal{X}} h(x, g_j x) dQ(x) \right. \\
&\quad \left. + \frac{1-g_m}{nk} \sum_i \sum_{j_1, j_2} h(g_{j_1} X_i, g_{j_2} X_i) \right].
\end{aligned}$$

(略証) X_1, \dots, X_n の条件の下で $P \in DI(\alpha + \sum_{i=1}^n \beta_{X_i}^g)$ となることに注意して, $\hat{\theta}_2 = E[\theta_2 | X_1, \dots, X_n] \wedge$ 補題1を用いる。

一般には、次の様になる。

補題2. $P \in DI(\alpha)$,

$h(x_1, \dots, x_s)$: measurable real-valued ft. on $(\mathcal{X}^s, \mathcal{A}^s)$ & symmetric in x_1, \dots, x_s .

このとき,

$$E \int_{\mathcal{X}^s} h(x_1, \dots, x_s) \prod_{i=1}^s dP(x_i)$$

$$= \sum_C \frac{s! M^{\sum m(i)}}{k^{s - \sum m(i)} M^{(s)} \prod_{i=1}^s [m(i)! i^{m(i)}]}$$

$$\times \int_{\mathcal{X}^{\sum m(i)}} \sum_f h(x_{11}, \dots, x_{1m(1)}, x_{21}, f_1 x_{21}, x_{22}, f_2 x_{22}, \dots, \\ x_{s1}, \dots, f_{s - \sum m(i)} x_{s1}) \prod_{i=1}^s \prod_{j=1}^{m(i)} dQ(x_{ij}),$$

ただし, 下式の全ての積分の存在を仮定する。また, \sum_C は $\sum_{i=1}^s i m(i) = s$ を満たす全ての s 々の非負整数の組 $m(1), \dots, m(s)$ についての和を表す。 \sum_f は \mathcal{G} での $f_1, \dots, f_{s - \sum m(i)}$ ($\in \mathcal{G}$) についての和を表す。

(略証) 補題1の方法ではなく, Dirichlet process P^* が

$$P^*(\cdot) = \sum_{j=1}^{\infty} P_j \int V_j(\cdot),$$

と表わされることと, 3章(1)とを用いる。ただし, V_1, V_2, \dots は \mathcal{X} の値をとる i.i.d. な r.v. で分布 Q をもつ。 P_1, P_2, \dots は $M = \alpha(\mathcal{X})$ を通してのみ α に依存する r.v. で V_1, V_2, \dots に独立, 且つ $\sum_{j=1}^{\infty} P_j = 1, P_j \geq 0$ ($j=1, 2, \dots$) を満たす (Ferguson [1973])。

積分 $\int_{\mathcal{X}^s} h(x_1, \dots, x_s) \prod_{i=1}^s dP(x_i)$ を V_1, V_2, \dots と P_1, P_2, \dots を用いて展開した上で, 任意の整数 $u, r(1), \dots, r(u)$ に対して, $t = r(1) + \dots + r(u)$ とおくと

$$\mathbb{E} \sum_{j^{(1)}, \dots, j^{(u)}} P_{j^{(1)}}^{r(1)} \dots P_{j^{(u)}}^{r(u)} = (r(1)-1)! \dots (r(u)-1)! M^u / M^{(t)},$$

$$(M^{(t)} = M(M+1) \cdots (M+t-1))$$

が成り立つ (Yamato [1984]) によると, Q が G -invariant であることを用いる。

degree s の推定可能な母数 $\theta_s = \int_{x^s} h(x_1, \dots, x_s) \prod_{i=1}^s dP(x_i)$ に対して (h は x_1, \dots, x_s について対称),

$$h_g(x_1, \dots, x_s) = \frac{1}{k^s} \sum_{f \in G} h(f_1 x_1, \dots, f_s x_s)$$

よると, G -invariant な P に対して

$$\theta_s = \int_{x^s} h_g(x_1, \dots, x_s) \prod_{i=1}^s dP(x_i)$$

と表わされる。命題 2 に注意して, 上式に補題 2 を用いると, h_g が G -invariant であることから,

定理 2. $P \in DI(\alpha)$,

X_1, \dots, X_n : P からの size n の標本

h について, 補題 2 の右辺の積分の存在を仮定する。

このとき, 推定可能な母数 θ_s のバイズ推定量 $\hat{\theta}_s$ は次のようになる。

$$\hat{\theta}_s = \sum_c \frac{s! (M+n)^{\sum m(i)}}{(M+n)^{(s)} \prod_{i=1}^s [m(i)! i^{m(i)}]} \times$$

$$\int_{x \in \Sigma^{m(i)}} h_g(x_{11}, \dots, x_{1, m(1)}, x_{21}, x_{21}, \dots, x_{2, m(2)}, x_{2, m(2)}, \dots, x_{s1}, \dots, x_{s1}) \prod_{i=1}^s \prod_{j=1}^{m(i)} d\hat{P}_n(x_{ij}),$$

ただし, $\hat{P}_n = \delta_n Q + (1 - \delta_n) P_n^*$, $P_n^* = \frac{1}{n} \sum_{i=1}^n \delta_{x_i} g = \frac{1}{n} \sum_{i=1}^n \sum_{j=1}^s \delta_{g_j x_i}$.

4. 例

例1. $(X, A) = (R^1, B^1) \times L$, $G = \{e, g\}$, $e(x) = x$, $g(x) = -x$ for $x \in R^1$ とする。パラメータ α を G -不変、即ち原点に関して対称にとることにより, $P \in DI(\alpha)$ は確率1で原点に関して対称となる。degree 2 の推定可能な

数
$$\Delta = \int_{R^2} |x - y| dP(x) dP(y)$$

を考へよう, これは分布 P の coefficient of mean difference として知られてゐる。 $\int |x| dQ(x) < \infty$ とすると, Δ のバイズ推定量は

$$\hat{\Delta} = \frac{M+n}{M+n+1} \left[\delta_n^2 \int_{R^2} |x-y| dQ(x) dQ(y) + \frac{\delta_n(1-\delta_n)}{n} \sum_{i=1}^n \int_{R^1} (|x-X_{i1}| + |x+X_{i1}|) dQ(x) \right]$$

$$+ \frac{(1-\delta_n)^2}{2n^2} \sum_{i,j} (|X_i - X_j| + |X_i + X_j|) \\ + \frac{1}{M+n+1} \left[\delta_n \int_{\mathbb{R}^1} |x| dQ(x) + \frac{1-\delta_n}{n} \sum_i |X_i| \right].$$

Q を固定して, $M \rightarrow 0$ としたとき, 極限ベイズ推定量は

$$\Delta^* = \frac{1}{n(n+1)} \left[\sum_{i < j} (|X_i - X_j| + |X_i + X_j|) + 2 \sum_i |X_i| \right]$$

で与えられる。他方, Dirichlet process を用いたときの極限ベイズ推定量は

$$\Delta^{**} = \frac{2}{n(n+1)} \sum_{i < j} |X_i - X_j|$$

で与えられる (Yamato [1977])。

例2. $(\mathcal{X}, \mathcal{A}) = (\mathbb{R}^2, \mathcal{B}^2)$ とし, $G = \{e, g\}$, $e(x, y) = (x, y)$ & $g(x, y) = (y, x)$ for $(x, y) \in \mathbb{R}^2$ とする。パラメータ α を G -不変, 即ち直線 $y=x$ に関して対称にとることにより, $P \in \mathcal{D}I(\alpha)$ は確率1で直線 $y=x$ に関して対称になる。

degree 2 の推定可解な母数として, 命題 P の共分散

$$\Theta_2 = \int_{\mathbb{R}^2} x_1 x_2 dP(x_1, x_2) - \int_{\mathbb{R}} x_1 dP(x_1, x_2) \int_{\mathbb{R}} x_2 dP(x_1, x_2)$$

のベイズ推定を考えよう。

$\int_{R^2} x_1^2 dQ(x_1, x_2) < \infty$ とおくと, θ のベイズ推定量は

$$\begin{aligned} \hat{\theta} &= \frac{M+n}{M+n+1} \left\{ \delta_m^2 \text{Cov}(Q) \right. \\ &\quad + \frac{\delta_m(1-\delta_m)}{n} \sum_{i=1}^n \int_{R^2} (x_1 - X_i^{(1)})(x_2 - X_i^{(2)}) dQ(x_1, x_2) \\ &\quad + \frac{(1-\delta_m)^2}{4n^2} \sum_{i,j} [(X_i^{(1)} - X_j^{(1)})(X_i^{(2)} - X_j^{(2)}) + \\ &\qquad\qquad\qquad (X_i^{(1)} - X_j^{(2)})(X_i^{(2)} - X_j^{(1)})] \left. \right\} \\ &\quad - \frac{1}{2(M+n+1)} \left[\delta_m \text{Var}(Q) + \frac{1-\delta_m}{2n} \sum_i (X_i^{(1)} - X_i^{(2)})^2 \right], \end{aligned}$$

但し, $(X_1^{(1)}, X_1^{(2)}), \dots, (X_n^{(1)}, X_n^{(2)})$ は P から n の標本を表わす。

Q を固定して, $M \rightarrow 0$ としたとき, 極限ベイズ推定量は

$$\begin{aligned} \theta^* &= \frac{1}{4n(n+1)} \sum_{i,j} [(X_i^{(1)} - X_j^{(1)})(X_i^{(2)} - X_j^{(2)}) + (X_i^{(1)} - X_j^{(2)})(X_i^{(2)} - X_j^{(1)})] \\ &\quad - \frac{1}{4n(n+1)} \sum_i (X_i^{(1)} - X_i^{(2)})^2 \end{aligned}$$

で与えられる。他方, Dirichlet process を用いたときの極限ベイズ推定量は

$$\theta^{**} = \frac{1}{2n(n+1)} \sum_{i,j} (X_i^{(1)} - X_j^{(1)})(X_i^{(2)} - X_j^{(2)}) .$$

例3. 例2と同じ仮定の下で, degree 2の推定可能な
母数 $\theta = \int_{\mathbb{R}^4} s(x_1 - y_1) s(x_2 - y_2) dP(x_1, x_2) dP(y_1, y_2)$

のベイズ推定を考えよう。ここで, $s(x) = 1 (x > 0), = 0$
($x = 0$), $= -1 (x < 0)$ 。

θ のベイズ推定量は,

$$\begin{aligned} \hat{\theta} &= \frac{M+n}{M+n+1} \left\{ \beta_n^2 \int_{\mathbb{R}^4} s(x_1 - y_1) s(x_2 - y_2) dQ(x_1, x_2) dQ(y_1, y_2) \right. \\ &\quad + \frac{2\beta_n(1-\beta_n)}{n} \sum_i \int_{\mathbb{R}^2} s(x_1 - X_i^{(1)}) s(x_2 - X_i^{(2)}) dQ(x_1, x_2) \\ &\quad \left. + \frac{(1-\beta_n)^2}{2n^2} \sum_{i,j} [s(X_i^{(1)} - X_j^{(1)}) s(X_i^{(2)} - X_j^{(2)}) + s(X_i^{(1)} - X_j^{(2)}) s(X_i^{(2)} - X_j^{(1)})] \right\} \\ &\quad - \frac{1}{2(M+n+1)} [\beta_n Q\{(x_1, x_2) \mid x_1 \neq x_2\} \\ &\quad \quad + \frac{1-\beta_n}{n} \times \text{no. of } \{i : X_i^{(1)} \neq X_i^{(2)}, i=1, \dots, n\}]. \end{aligned}$$

Q を固定して, $M \rightarrow \infty$ とし, 極限ベイズ推定量は

$$\begin{aligned} \theta^* &= \frac{1}{2n(n+1)} \sum_{i,j} [s(X_i^{(1)} - X_j^{(1)}) s(X_i^{(2)} - X_j^{(2)}) + s(X_i^{(1)} - X_j^{(2)}) s(X_i^{(2)} - X_j^{(1)})] \\ &\quad - \frac{1}{2n(n+1)} \times \text{no. of } \{i : X_i^{(1)} \neq X_i^{(2)}, i=1, \dots, n\}. \end{aligned}$$

と与えられる。他方, Dirichlet processを用いたときの極

限ベイズ推定量は

$$\theta^{**} = \frac{1}{n(n+1)} \sum_{i,j} s(X_i^{(1)} - X_j^{(1)}) s(X_i^{(2)} - X_j^{(2)})$$

で与えられる (Yamato [1977])。なお,

$$U = \frac{1}{n(n-1)} \sum_{i,j} s(X_i^{(1)} - X_j^{(1)}) s(X_i^{(2)} - X_j^{(2)})$$

は difference-sign covariance χ (7) である。

References

- [1] Ferguson, T.S. (1973). A Bayesian analysis of some non-parametric problems. Ann. Statist. 1, 209-230.
- [2] Dalal, S.R. (1979a). Nonparametric and robust Bayes estimation of location. Proc. International Symposium on Optimizing Methods in Statistics (J.S. Rustagi Ed.). New York: Academic Press.
- [3] Dalal, S.R. (1979b). Dirichlet invariant process and applications to nonparametric estimation of symmetric distribution functions. Stoch. Proc. and Appl., 9, 99-107.
- [4] Dalal, S.R. (1980). Non-parametric Bayes decision theory. Bayesian Statistics (J.M. Bernardo, M.H. Degroot, D.V. Lindley & A.F.M. Smith, Eds.). Valencia, Spain: University Press.

- [5] Yamato, H. (1977). Relations between limiting Bayes estimates and the U-statistics for estimable parameters of degrees 2 and 3. Commun. in Statist., A, 6, 55-66.
- [6] Yamato, H. (1984). Characteristic functions of means of distributiond chosen from a Dirichlet process. Ann. Probability, 12, 262-267.
- [7] Yamato, H. (1985). Bayes estimation of estimable parameter with a Dirichlet invariant process. (submitted)