数理解析研究所講究録
第 223 巻 1974 年 62-72

変換群と分布の特徴づけ(I)

大阪市大　理学部　工藤　弘吉

筆者の得た二つの結果を紹介する．いづれも既知であるかどうかは今の所明らかではない．教えて頂ければ幸いである．

(A) この項は筆者の論文

On minimax invariant estimates of the transformation parameter, Nat. Sci. Rep., Ochanomizu University, 6 (1955), pp. 37—73.

の執筆中副産物として得られたものなので記号その他種々の考え方に上記論文との関連が多い．

n次元ユークリッド空間R^n，そのボレル集合の全体$\mathcal{B}$，$\mathcal{B}$上の1つの確率測度Pは固定されているものとし，Pはルベーグ測度について絶対連続とする．Pの密度関数を$\bar{f}(x_1, \cdots, x_n)$とする．$R^n$から$R^n$への変換

$$\sigma(x_1, x_2, \cdots, x_n) \longrightarrow$$

$$(a_\sigma x_1 + b_\sigma,\ a_\sigma x_2 + b_\sigma,\ \cdots,\ a_\sigma x_n + b_\sigma),$$

ここで　$0 < a_\sigma < \infty$,　$-\infty < b_\sigma < \infty$.

このような変換の全体Gは群をなし，σ，τ $(\in G)$ の積 $\sigma\tau$ については

$$a_{\sigma\tau} = a_\sigma a_\tau , \qquad b_{\sigma\tau} = a_\sigma b_\tau + b_\sigma$$

なる関係がある．

R^nの中の直線X_0：$x_1 = x_2 = \cdots = x_n$，点 $\boldsymbol{x}_0 = (1, 1, \cdots, 1)$ を定めると，

$$\boldsymbol{x} \notin X_0 \text{ に対して } \boldsymbol{x} = s\boldsymbol{z} + \bar{x}\,\boldsymbol{x}_0$$

と一意的に表わされる，ただしここで

$$\bar{x} = \frac{\sum_{i=1}^{n} x_i}{n} , \quad s = s(\boldsymbol{x}) = \sqrt{\frac{\sum_{i=1}^{n} (x_i - \bar{x})^2}{n}} ,$$

$$\boldsymbol{z} = \boldsymbol{z}(\boldsymbol{x}) = \left(\frac{x_1 - \bar{x}}{s} , \frac{x_2 - \bar{x}}{s} , \cdots , \frac{x_n - \bar{x}}{s} \right).$$

$\boldsymbol{z}$は超平面 $z_1 + z_2 + \cdots + z_n = 0$ の中の単位球D上の点である．このような表現によって，

(A 1) $$f(x_1, x_2, \cdots, x_n) = p(\bar{x}, s(\boldsymbol{x}), \boldsymbol{z}(\boldsymbol{x})),$$

$$dx_1\, dx_2 \cdots dx_n = s^{n-2}\, d\bar{x}\, ds\, d\boldsymbol{z}$$

となる，ここで $d\boldsymbol{z}$ は上記の球面Dの面素をあらわす．

X_0を軸とした回転の群を$\mathcal{O}$とし，その元をAであらわす．そのとき (i) すべての $\boldsymbol{x} \in R^n$，$A \in \mathcal{O}$ に対して $\overline{A\boldsymbol{x}} = \bar{x}$，$s(A\boldsymbol{x}) = s(\boldsymbol{x})$，(ii) $\boldsymbol{z}(\boldsymbol{x}) = \boldsymbol{z}(\boldsymbol{y})$ なら $\boldsymbol{z}(A\boldsymbol{x}) = \boldsymbol{z}(A\boldsymbol{y})$．(ii) から$\boldsymbol{z}$の空間$D$上に変換の群$\tilde{\mathcal{O}}$を，$A \to \tilde{A} \in \tilde{\mathcal{O}}$: $\boldsymbol{z}(A\boldsymbol{x})$

$=\tilde{A}\mathbf{z}(\mathbf{x})$ がなりたつように定義される。任意の $\mathbf{z}_1, \mathbf{z}_2$ に対して，$\tilde{A}\in\tilde{\mathcal{O}}$ を $\mathbf{z}_2=\tilde{A}\mathbf{z}_1$ がなりたつようにとれるから，Dは $\tilde{\mathcal{O}}$ の商空間である．その上，Dの面素 $d\mathbf{z}$ は $\tilde{\mathcal{O}}$ について不変測度である．

さて確率測度Pは $\sigma\in G$，$A\in\mathcal{O}$ によって

$$\bar{\sigma}P(E)=P(\sigma^{-1}E)=\frac{1}{a_\sigma^n}\int_{E'}p\left(\frac{\bar{\mathbf{x}}-b_\sigma}{a_\sigma},\frac{s}{a_\sigma},\mathbf{z}\right)\times s^{n-2}\,d\bar{\mathbf{x}}\,ds\,d\mathbf{z}$$

$$\bar{A}\bar{\sigma}P(E)=P(\sigma^{-1}A^{-1}E)=\frac{1}{a_\sigma^n}\int_{E'}p\left(\frac{\bar{\mathbf{x}}-b_\sigma}{a_\sigma},\frac{s}{a_\sigma},\tilde{A}^{-1}\mathbf{z}\right)\times s^{n-2}\,d\bar{\mathbf{x}}\,ds\,d\mathbf{z}$$

という新しい確率分布 $\bar{\sigma}P$，$\bar{A}\bar{\sigma}P$ に転換される，ただし E' は写像 $\mathbf{x}=(x_1,x_2,\cdots,x_n)\to(\bar{x},s(\mathbf{x}),\mathbf{z}(\mathbf{x}))$ による $E(\in\mathcal{B})$ の像である．

<u>定理</u>　<u>$x_1,x_2,\cdots,x_n$ は独立同一分布 $f(x)$ に従い，$f(x)$ が正で連続微分可能であるとする．そのとき σP が変換群 $\mathcal{O}$ で不変であるためには，$f(x)=(2\pi a^2)^{-\frac{1}{2}}\exp\left(-\frac{1}{2a^2}(x-b)^2\right)$ となることが必要十分である．</u>

<u>証明</u>　十分であることは明らか．必要であることを示そう．$f(x)$ が連続であるから，(A1)から $p(\bar{x},s,\mathbf{z})$ も連続であり，$\bar{A}\bar{\sigma}P=\bar{\sigma}P$ より $p(\bar{x},s,\mathbf{z})=p(\bar{x},s,\tilde{A}^{-1}\mathbf{z})$ がなり

たつ。$\bar{A}\bar{\sigma}P=\bar{\sigma}P$ がすべての $A\in\mathcal{O}$ でなりたって，D が $\mathcal{O}$ で等質であるから，$p(\bar{x},\ s,\ \mathbf{z})$ は $\mathbf{z}$ を含まない。この関数の対数を $g(n\bar{x},\ n(s^2+\bar{x}^2))$ であらわし，一方 $f(x)$ の対数を $l(x)$ であらわすなら，(A1)から

$$\sum_{i=1}^{n} l(x_i)=g\Big(\sum_{i=1}^{n} x_i\ ,\ \sum_{i=1}^{n} x_i^2\Big).$$

この式の両辺を x_i で微分すると

$$\frac{d\,l(x_i)}{dx_i}=h\Big(\sum_{i=1}^{n} x_i,\sum_{i=1}^{n} x_i^2\Big)+2x_i\,k\Big(\sum_{i=1}^{n} x_i,\sum_{i=1}^{n} x_i^2\Big),$$
$$i=1,2,\cdots,n,$$

ここで $h(\alpha,\beta)=\frac{\partial}{\partial\alpha}g(\alpha,\beta)$，$k(\alpha,\beta)=\frac{\partial}{\partial\beta}g(\alpha,\beta)$。この式から $h(\alpha,\beta)$，$k(\alpha,\beta)$ を消去すると，

$$\begin{vmatrix} l'(\alpha) & \alpha & 1 \\ l'(\beta) & \beta & 1 \\ l'(\gamma) & \gamma & 1 \end{vmatrix}=0.$$

したがって $l(\alpha)$ は次の形となる：

$$l(\alpha)=a\alpha^2+b\alpha+c.$$

すなわち，$f(x)=\exp(ax^2+bx+c)$ となり，$f(x)$ は正規分布である。

R^n における変換群 $\mathcal{O}$ の跡 (trajectory) は $(\bar{\mathbf{x}},\ s(\mathbf{x}))$ であり，したがって $\mathcal{O}$ に関して不変な測度 σP の全体 $\{\sigma P:\ \sigma\in G\}$ に対して $(\bar{\mathbf{x}},\ s(\mathbf{x}))$ は十分統計量である（レーマン（渋谷・竹内訳）統計的検定論，岩波書店： 277ページ，19;

または本項の最初に引用した拙著の Th. 1.7. 参照)。また逆に$\{\bar{A}\bar{\sigma}P : A \in \mathcal{O}, \sigma \in G\}$における$(\bar{X}, s(X))$の十分性から$\sigma P$の$\mathcal{O}$不変性がでるから，定理から直ちに

系　定理の条件の下では，$\{\bar{A}\bar{\sigma}P : A \in \mathcal{O}, \sigma \in G\}$において$(\bar{X}, s(X))$が十分統計量であるためには，母集団分布$f(x)$が正規分布であることが必要十分である.

(B) 次に述べる事柄は23年前に和文で発表したことなので，新しいものでないが，まだ一般には知られていないようであるから，紹介の意味で結果だけを記しておく，証明その他くわしいことは下記の論文を参照されたい:

α-estimate の分布関数，統計数理研究所講究録第7巻(1951), pp. 335–354.

$(X, \mathcal{B})$を可測空間，その上の確率測度の1径数族$\mathcal{P} = \{P_\theta : -\infty < \theta < \infty\}$が与えられている.

仮定1　$\mathcal{P}$の要素P_θは互に絶対連続である.

この仮定のもとでは

$$P_\theta(E) = \int_E f(x, \theta)\, P_0(dx)$$

なる$f(x, \theta) = \dfrac{dP_\theta}{dP_0}$がある. この関数は任意の$\theta, \theta'$に対して次の性質をもつ.

$$P_{\theta'}(E) = \int_E (f(x, \theta')/f(x, \theta))\, P_\theta(dx).$$

<u>仮定2</u>　<u>どんなθについてもP_θは nonatomic である.</u>

この仮定のもとでは，仮説P_θ・対立仮説$P_{\theta'}$に対する最強力検定をどんな水準ででも作れる．$\theta>0$を任意にとって，これを固定する．仮説P_0・対立仮説P_θに対する最強力水準α検定の棄却域をR_αとする．αを動かしてこのようなR_αの全体（棄却域系）$\mathbb{R}=\{R_\alpha : 0\leqq\alpha\leqq 1\}$を次の条件をみたすように作れることは明らかである：

$0\leqq\alpha\leqq 1$なるαに対し1つ，ただ1つ，の$R_\alpha(\in\mathbb{R})$が対応する．

$R_0=\phi$（空集合），　$R_1=X$，

$P_0(R_\alpha)=\alpha$，

$\alpha_1<\alpha_2$なら$R_{\alpha_1}\subset R_{\alpha_2}$．

このような$\mathbb{R}$が与えられると，その要素を棄却域にする検定の検定力(power)は$P_\theta(R_\alpha)$である．これは後に$\gamma(\alpha\,|\,\theta)$であらわされることになるが，その前に最も重要な1つの仮定をのべる．

<u>仮定3</u>　<u>任意のθ'，θ''（$\theta'<\theta''$）および任意の$k(>0)$に対して，α（$0\leqq\alpha\leqq 1$）が定まり，</u>

$$P_0(R_\alpha \,\triangle\, Q_k(\theta',\theta''))=0,$$

<u>ここで</u> $Q_k(\theta',\theta'')=\{x\in X : (f(x,\theta'')/f(x,\theta'))\geqq k\}$

<u>であり，$\triangle$は集合の対称差を示す．</u>

仮定3は，$\mathbb{R}$はθのえらび方に無関係に作りうることを示している（ただし$\theta>0$）．例えば，$P_{\theta'}$，$P_{\theta''}$を仮説，対立仮説（ただし$\theta'<\theta''$）とする検定問題では，$\alpha=P_{\theta'}(R_{\alpha'})$なる$\alpha'$（$\in\mathbb{R}$）をとると，これは最強力水準$\alpha$の棄却域となる．もし$\theta''<\theta'$なら，$P_{\theta'}(R_{\alpha''})=1-\alpha$となる$R_{\alpha''}$（$\in\mathbb{R}$）をとるとき，$X-R_{\alpha''}$は最強力水準$\alpha$の棄却域となり，その検定力は$1-P_{\theta''}(R_{\alpha''})$となる．このように$P_\theta(R_\alpha)$は，仮定3のもとでは，極めて重要な働きをするので，

$$-\infty<\theta<\infty \quad \text{で} \quad \gamma(\alpha|\theta)=P_\theta(R_\alpha)$$

とおく．この記号を用いると，仮説$P_{\theta'}$，対立仮説$P_{\theta''}$に対する最強力水準α検定の検定力は

$\theta'<\theta''$ のとき　$\gamma(\alpha'|\theta'')$，ただし$\alpha=\gamma(\alpha'|\theta')$

$\theta'>\theta''$ のとき　$1-\gamma(\alpha'|\theta'')$，ただし$\gamma(\alpha'|\theta')=1-\alpha$。

また$\gamma(\alpha|\theta)$，$\theta>0$，は次の条件(i)—(iii)をみたす：(i) $\gamma(0|\theta)=0$，$\gamma(1|\theta)=1$，(ii) $\gamma(\alpha|\theta)$はすべての$\theta>0$で，αの単調増加な凸関数，(iii) $\gamma(\alpha|\theta)>\alpha$ $(0<\alpha<1)$．なお$\gamma(\alpha|0)\equiv\alpha$であることに注意せよ．

パラメータは分布の固有名詞のようなものであるから，その意味からいえば，その名づけ方は任意である．次の仮定はそういうわけでパラメータの付け方を与えるものである．

仮定4　$\gamma(\gamma(\alpha|\theta)|\theta')=\gamma(\alpha|\theta+\theta')$.

この仮定があると，$\gamma(\gamma(\alpha|-\theta)|\theta)=\gamma(\alpha|0)=\alpha$ であるから，仮説 $P_{\theta'}$・対立仮説 $P_{\theta''}$ に対する最強力水準 α 検定の検定力は

$\theta'<\theta''$ のとき　$\gamma(\alpha|\theta''-\theta')$

$\theta'>\theta''$ のとき　$1-\gamma(1-\alpha|\theta''-\theta')$

となる．また $0<\alpha<1$ なら $\gamma(\alpha|\theta)$ は θ の単調増加関数で，$\theta_1<\theta_2$ なら $R_{\gamma(\alpha|-\theta_1)}\supset R_{\gamma(\alpha|-\theta_2)}$ がなりたつ．

いま α $(0<\alpha<1)$ に対して，x の関数

$$\hat{\theta}_\alpha(x)=\sup\{\theta:\ x\in R_{\gamma(\alpha|-\theta)}\}$$

を定義すると，この関数は可測であり，$\mathcal{P}$ に対する十分統計量で，標本分布 P_θ のもとでの $\hat{\theta}_\alpha(x)$ の累積分布関数は

$$D_\alpha(\hat{\theta}|\theta)=P_\theta\{\hat{\theta}_\alpha(x)\leqq\hat{\theta}\}=1-\gamma(\alpha|\theta-\hat{\theta})$$

であり，したがって

(B1)　　$D_\alpha(\hat{\theta}|\theta)=D_\alpha(\hat{\theta}-\theta|0)$.

仮定1－4のもとでは，$D_\alpha(\theta|0)$ が連続な導関数 $\varphi_\alpha(\theta)$ をもつことが証明される．この関数 $\varphi_\alpha(\theta)$ はもちろん $\hat{\theta}_\alpha(x)$ の分布密度関数である．また(B1)より $\frac{d}{d\hat{\theta}}D_\alpha(\hat{\theta}|\theta)=\varphi_\alpha(\hat{\theta}-\theta)$ であるから，θ は $\hat{\theta}_\alpha(x)$ の分布に関しては位置母数の役割をはたす．$\hat{\theta}_\alpha(x)$ を θ の推定として用いることができるわけで，これを<u>α 推定</u>と名づけることにする．

次に1つのα ($0<\alpha<1$) を固定して

(B2) $$\Phi(\theta)=\gamma(\alpha|\theta)$$

とおくと，$\Phi(\theta)$は$(-\infty,\infty)$から$(0,1)$の上への1対1対応であり，$\Phi(0)=\alpha$，$\Phi(-\infty)=0$，$\Phi(+\infty)=1$.
さらに$\Phi(\theta)$の逆関数は不定積分の形で

$$\Phi^{-1}(\alpha')=\int\frac{d\alpha}{\frac{\partial}{\partial\theta}\gamma(\alpha'|\theta)\Big|_{\theta=0}}$$

とあらわされる. ここで不定積分に含まれる任意定数は$\Phi(\theta)$を定義する場合に用いられたαに由来する. そして

(B2′) $$\gamma(\alpha'|\theta)=\Phi(\Phi^{-1}(\alpha')+\theta)$$

が成立する. また$\hat{\theta}_\alpha(x)$が$\mathcal{P}$に対する十分統計量であることより，$f(x,\theta)$は$\hat{\theta}_\alpha(x)$とθだけの関数であるから，

(B3) $$f(x,\theta)=\lambda(-\hat{\theta}_\alpha(x),\theta)$$

とおくと，

(B4) $$\lambda(r,s)=\frac{\Phi'(r+s)}{\Phi'(r)}$$

となる. こうして

<u>定理</u>　<u>$(\mathcal{X},\mathcal{B},\mathcal{P})$が仮定1－4をみたせば，$\mathcal{P}$の密度関数の比$f(x,\theta)$は(B2)－(B4)によって定まる形となる，ここで$\hat{\theta}_\alpha(x)$は$\mathcal{X}$から$(-\infty,\infty)$の上への写像である.</u>

<u>仮定5</u>　<u>$\lambda(r,s)$はsに関して偏微分可能で，</u>

$$l(r,s)=\frac{\partial}{\partial s}\log\lambda(r,s)$$

はrの連続関数である.

仮定6 $\int_{-\infty}^{\infty}\exp\left(\int l(r,0)\,dr\right)dr<\infty.$

この2つの仮定があれば,(B4)をみたす$\Phi(r)$は

$$\Phi(r)=\frac{\int_{-\infty}^{r}\exp\left(\int l(q,0)\,dq\right)dq}{\int_{-\infty}^{\infty}\exp\left(\int l(q,0)\,dq\right)dq}$$

としてもとめられる. さらに仮定1—6から

仮定7 $l(r,s)=l(r+s,0).$

仮定8 $\Phi(r)$の特性関数の0点の集合は内点をもたない.

がいえるが,仮定5—8は,仮定1—4をみたす統計問題の特徴でもある. すなわち

定理 仮定5—7をみたす$\lambda(r,s)$,および(B4)によって定まる$\Phi(r)$が仮定8をみたすものとする. このとき(B2′)で定まる$\gamma(\alpha|\theta)$に対して,仮定1—4をみたす統計問題$(X,\mathcal{B},\mathcal{P})$を構成することができる.

次の定理は仮定1—6をみたす統計問題は本質的に位置母数をもつ統計問題であることを示している.

定理 $(X,\mathcal{B},\mathcal{P})$が仮定1—6をみたすとする. もし$X$上の統計量Tの値域$\mathcal{T}$が$(-\infty,\infty)$と1対1対応$u(t)$をもち$u(T(x))$が$\alpha$推定$\hat{\theta}_\alpha(x)$と一致するなら,$\mathcal{T}$上に1径数変換群$G=\{\sigma^s:-\infty<s<\infty\}$ $(\sigma^s\sigma^r=\sigma^{s+r})$を導入

して，$T(x)$ の周辺分布 P_θ^T を

$$P_\theta^T(S) = P_0^T(\sigma^{-\theta}S),\quad S\subset\mathfrak{T},\ T^{-1}S\in\mathcal{B},$$

とすることができる．

この定理でいう$\mathfrak{T}$上の変換σ^θは

$$\sigma^\theta t = u^{-1}(u(t)-\theta)$$

として定義される．なお $\sigma^\theta T(R_\alpha) = T(R_{\gamma(\alpha|-\theta)})$ がなりたつことも注目に値する．

特に$X=R^n$ $(n\geq 2)$で，$x\,(\in X)$が独立同一分布に従う確率変数$x_1, x_2, \cdots, x_n$を成分とするベクトルである場合は，（x_iの分布に連続性等のゆるい条件をつければ）十分統計量$\widehat{\theta}_\alpha(x)$の存在から，母集団分布（$x_i$の分布）は指数型であることがわかる（Koopmanの定理）．一方$\widehat{\theta}_\alpha(x)$の分布に関しては$\theta$は位置母数である．故に拙著数理統計学（改訂版），現代数学講座23，共立出版，定理3·7より，母集団分布は正規型かガンマ型となってしまう．

注意　Borges - Pfanzagl 論文（AMS 36, 1965）は分布の型を指数型と限定し，標本空間上の変換群から誘導される母数群を考えた．それに対して本項で考えた問題では，分布の型を限定せずまた標本空間上の変換群からfreeな母数を考え，群論的な性質（仮定4）をその代りに設定した．筆者にはこの2つの結果が統合されうるように思えて仕方がない．（完）