数理解析研究所講究録
第 150 巻 1972 年 1-22

統計空間(Statistics Space)の構造について

大阪市立大、理　工藤弘吉

§ 1. 統計空間の命名 —— Introduction.

本講究録の題目「統計的構造」は集合 X, その部分集合の σ 代数 $\mathbb{B}$, および $\mathbb{B}$ 上で与えられる確率測度 P_θ, $\theta \in \Theta$, の族 $\mathcal{P}$ の triplet $(X, \mathbb{B}, \mathcal{P})$ のことである. この場合 P_θ は母数空間と呼ばれる集合 Θ の元 θ から $\mathcal{P}$ の元への 1 意的な写像 $\theta \to P_\theta$ を表はしているとも考えられる. triplet $(X, \mathbb{B}, \mathcal{P})$ はかつて Blackwell [1], Kudō [2], LeCam [3], Sacksteder [4] 等によつて「統計的実験」(Statistical experiment) と呼ばれたものである. このように同じ triplet $(X, \mathbb{B}, \mathcal{P})$ が統計的実験と呼ばれたり統計的構造 (Statistical structure, Basu - Ghosh [5], Morimoto [6]) と呼ばれたりすることは用語の混乱であり, 数学研究の障害であるから, この混乱をときほごすことを考えなければならない. (この Symposium が行われた後で入手した LeCam [7] では, 写像 $\theta \to P_\theta$ を統計的実験と呼んでいる.) 例えば位相空間論では, 集合 X

と開集合族$\mathcal{G}$の pair $(X, \mathcal{G})$を位相空間，$\mathcal{G}$を位相構造と呼んでいるし，測度論では，集合X，その部分集合のσ代数$\mathbb{B}$，その上の測度μの triplet $(X, \mathbb{B}, \mu)$を測度空間，$\mathbb{B}$をボレル構造と呼んでいる．これらの先例にならつて

$(X, \mathbb{B}, \mathcal{P})$を統計空間（statistics space）

$\mathbb{B}$を統計的構造（statistical structure）

と呼んだら如何なものであろうか．Xは従来の通り標本空間（sample space）という．

次に位相空間論における連続関数，測度論における可測写像に対応するものとして，統計空間の研究では乱写像（random mapping）が定義される．この乱写像は，位相空間論で連続写像が位相の強弱を定義するように，統計空間の情報の多寡を定義する．この意味で，位相空間論の homeomorphism と同様に，統計空間の sufficiency を定義することが出来る．

統計量は乱写像の一種である．筆者は統計量は英語の statistic の訳であるし，また日常生活で統計と呼んでいるものは我々の云う統計量のとる値であるから，「統計量」という術語は「統計」に改めるべきであると思っている．（この意見をシンポジウムで表明したところ，これに反対する見解も多かつた．これは「統計量は確率変数であり，統計と云われているものは数値であるから，統計量と統計は全く異るものである」

ということに基づいているように思われる．しかし単なる数値を何故統計と呼ぶのであろうか．単なる数値を統計と呼ぶ理由はその背後に統計量と呼ばれる確率変数の概念がひそんでいるからではないであろうか．）統計空間の十分性は十分統計量の概念に結びつく．十分統計量は本講究録の主要関心事の一つである．

統計空間論としての漸近理論は未発展の分野の一つである．わづかに LeCam [7]，Kudō [8] がそれぞれこの方面の試みをしている．

統計空間論における先験分布の問題は mixture の問題として解体される．事実先験確率は —— 主観主義の立場に立たない限り —— それによる mixture を標本分布とする予測問題を引き起すに過ぎない．しかし母数空間にもボレル構造を想定するなら，それは統計的構造との duality の問題が起り，これも今後の問題となるであろう．

以上述べた事柄を順を追って概説しよう．

§2. 統計空間の性質

集合 X，X の部分集合の σ 代数 $\mathbb{B}$，$\mathbb{B}$ 上の確率測度の族 $\mathcal{P}$ からなる triplet $(X, \mathbb{B}, \mathcal{P})$ を<u>統計空間</u>と云い，X を<u>標本空間</u>，$\mathbb{B}$ を<u>統計的構造</u>と名づける．統計的構造 $\mathbb{B}$ に関する性質

のうちで最も重要なものはσ-generatedということである。

<u>定義2・1</u>　$\mathbb{B}$の可算個の元からなる部分代数$\mathbb{B}'$が存在して，$\mathbb{B}$が$\mathbb{B}'$を含む最小のσ代数となる　（すなわち$\mathbb{B}'$が$\mathbb{B}$をgenerateする）とき，$\mathbb{B}$は<u>σ-generated</u>であるといわれる。

Euclid空間のBorel fieldやその可算直積などはこの性質をもつが，本講究録の森本氏の論文にあらわれる有限母集団におけるすべての集合の作るσ代数，連続時間をもつ時系列の空間などはこの性質をもたない。

<u>定義2・2</u>　$\mathbb{B}$上の確率測度の族$\mathcal{P}$が$\mathbb{B}$上のあるσ-有限測度μに対して絶対連続な測度Pだけからなるとき，$\mathcal{P}$は（μによつて）<u>dominateされる</u>と云われる。

ノンパラメトリックな問題や母集団分布が離散的で連続な位置パラメータをもつ場合などはこの性質を持たない。またBasu-Ghosh [5]，Morimoto [6] でformulateされた有限母集団においてもdominateされない$\mathcal{P}$を考えている。

$\mathcal{P}$は様々な方法で位相を入れて位相空間とみなされる。これらの位相によるseparabilityは上記dominatednessと強い関係があることは広く知られている（[9] を見よ）。

$\mathcal{P}$の完備性（completeness）も1つの重要な性質であろう。しかしさらに重要な性質は次の正則性であろう。

確率空間（X，$\mathbb{B}$，P）において，$\mathbb{B}$の部分σ代数$\mathbb{A}$を考え，

任意のP可積分関数fに対して，$\mathcal{A}$可測関数$E_P[f|\mathcal{A}]$が存在して，$\int_A f\,dP = \int_A E_P[f\,|\mathcal{A}]\,dP$ for all $A \in \mathcal{A}$ が成立する．このE$_P[f|\mathcal{A}]$を，$\mathcal{A}$が与えられたfの条件付平均という．特にfが$B\,(\in \mathcal{B})$の特徴関数であるとき，$E_P[f\,|\mathcal{A}]$を$P(B|\mathcal{A})$とかき，$\mathcal{A}$が与えられた$B\,(\in \mathcal{B})$の条件付確率という．たとえP測度0の$A\,(\in \mathcal{A})$を除いても，必らずしも$P(\cdot:\mathcal{A})$は$\mathcal{B}$上の確率測度ではない．

定義2·3　統計空間$(X, \mathcal{B}, \mathcal{P})$は，$\mathcal{B}$のいかなる部分$\sigma$代数$\mathcal{A}$に対しても，$\mathcal{P}$のいかなる要素$P$に対しても，$\mathcal{B}$上の確率測度である条件付確率$P(\cdot:\mathcal{A})$をもつとき，この統計空間は正則(*regular*)であるといわれる．

Le Cam [3; p.1442] によると，Xが完備で可分な距離空間，局所コンパクト集合の直積空間で，それぞれBorel集合族，Baire集合族の直積族を$\mathcal{B}$ととり，さらに$\mathcal{P}$がdominatedなら，その統計空間は正則である．[cf. [3; p.1447, Corollary]].

§3　乱写像(Random Mapping)

ある集合$\Theta\,(\ni \theta)$から1つの統計空間$(X, \mathcal{B}, \mathcal{P})$の$\mathcal{P}$の上への写像$\theta \to P_\theta\,(\in \mathcal{P})$を$\Theta$の上の実験(Experiment)といい，$\Theta$を母

数空間という．Θの上の実験を$(X, \mathbb{B}, P., \Theta)$であらわす．

定義3・1　1つの母数空間Θ上の2つの実験$\mathcal{S}=(X, \mathbb{B}, P., \Theta)$, $\mathcal{R}=(Y, \mathbb{C}, Q., \Theta)$に対して，$\mathbb{C}\times X \to [0,1]$なる写像$T(C, x)$が

i)　$\forall C \in \mathbb{C}$: $T(C : x)$は$\mathbb{B}$可測関数

ii)　$\forall x \in X$: $T(C : x)$は$\mathbb{C}$上の確率測度

iii)　$\forall \theta \in \Theta$:　$Q_\theta(C) = \int_X T(C, x)\, dP_\theta$

をみたすとき，$T(C, x)$を$\mathcal{S}$から$\mathcal{R}$への乱写像という．

乱写像はしばしば *stochastic mapping* (*or transformation*), *transition measure*, *Markov kernel*, *randomized mapping* などという名前で呼ばれることもある．統計空間論における乱写像は位相空間論における連続写像のような役割をはたすものであろう．

$(X, \mathbb{B})$から$(Y, \mathbb{C})$への可測写像$t(x)$は乱写像の1例である．すなわち$T(C : x)=1$ $(t(x)\in C)$; $=0$ $(t(x)\notin C)$とおけば，Tは乱写像となる．上記の可測写像をしばしば統計(量)という．

乱写像Tによって，定義3・1のiii)の関係にあるとき，$Q.=TP.$とかく．実験$(X, \mathbb{B}, P., \Theta)$と$(Y, \mathbb{C}, Q., \Theta)$とに対して，$Q.=TP.$となる乱写像$T$が存在するとき，実験$(X,$

$\mathbb{B}$, $P_\cdot$, Θ) は実験 (Y, $\mathbb{C}$, $Q_\cdot$, Θ) に対して十分 (sufficient) という.

定義 3·3　実験 $\mathcal{S}$ が実験 $\mathcal{R}$ に対して十分であり, 逆に実験 $\mathcal{R}$ が $\mathcal{S}$ に対して十分であるとき, $\mathcal{S}$ は $\mathcal{R}$ と等価であるという.

実験の等価は位相空間論の同相ということに相応する概念であるだろう.

普通に用いられる十分統計量の概念はここでは次のように定義される.

定義 3·4　実験 $\mathcal{S}$ から $\mathcal{R}$ への乱写像 T が十分であるとは, $\mathcal{S}$ と $\mathcal{R}$ とが等価であることを意味する.

T が特に (X, $\mathbb{B}$) から (Y, $\mathbb{C}$) への可測写像であるとき, 定義 3·4 の十分性は普通の十分性と一致する. 何故ならば, 統計(量) $T(x)$ の十分性は, 乱写像 $R_1(B|y)$ が存在して,

$$(*)\qquad \forall\theta\in\Theta,\ \forall B\in\mathbb{B},\ \forall C\in\mathbb{C}:\ P_\theta(B\cap T^{-1}C)=\int_C R_1(B|y)P_\theta T^{-1}(dy).$$

一方定義 3·4 の十分性は, 乱写像 $R_2(B|y)$ が存在して

$$(**)\qquad \forall\theta\in\Theta,\ \forall B\in\mathbb{B}:\ P_\theta(B)=\int_Y R_2(B|y)\,P_\theta T^{-1}(dy).$$

したがつて (*) の $C=Y$ とすれば (**) が成立することは直ちにわかる. 逆: (**) ⇨ (*) の証明は Le Cam [3; Prop. 9] にあるが, シンポジュムで竹内啓君が, $\mathcal{P}$ が dominated であ

ることを仮定して，この逆の別証明を与えたので，ここではそれを紹介する．この別証明は次の２つの良く知られた定理に基づく．

<u>定理3・1</u>　統計(量)Tに対して(*)をみたすR_1が存在するためには，Θの任意の２つの要素θ_1，θ_2の組を母数空間としたとき(*)をみたす$R_{\theta_1,\theta_2}=R_1(B|y)$が存在することが必要十分である，ただし$\mathcal{P}$はdominatedとする．

<u>定理3・2</u>　任意の統計(量)Tに対して，Tの値域Yのσ代数を$\mathbb{C}$とするとき，kを正数とするなら

$$\min_{B\in\mathbb{B}}\left(P_1(B)-kP_0(B)\right)\leq\min_{C\in\mathbb{C}}\left(P_1T^{-1}(C)-kP_0T^{-1}(C)\right).$$

特にすべての正数kに対して上式の等号がなりたつとき，またその時に限り，Tは(*)をみたす．ただしP_0，P_1は$\mathbb{B}$上の確率測度であり，$\Theta=\{0,1\}$とする．

<u>定理3・3</u>　$\mathcal{P}$がdominatedなら，(**)をみたすR_2が存在するとき(*)をみたすR_1が存在する．（注意：R_1は必ずしもR_2と一致しない．）

<u>竹内啓氏による定理3・3の証明</u>．$\mathcal{P}$がdominatedであるから，定理3・1によつて，$\Theta=\{0,1\}$と仮定する（一般性は失われない）．$C_k=\{y:\ g_1(y)\leq k\,g_0(y)\}$とおく，ここで$g_i(y)=dP_iT^{-1}/d(P_0T^{-1}+P_1T^{-1})$，$i=0,1$．その

とき, $\min_{C\in\mathbb{C}}[P_1T^{-1}(C)-kP_0T^{-1}(C)]=P_1T^{-1}(C_k)-kP_0T^{-1}(C_k)\leqq\int_Y R_2(B|y)\{P_1T^{-1}(dy)-kP_0T^{-1}(dy)\}=P_1(B)-kP_0(B)$, $B\in\mathbb{B}$. したがって $\min_{C\in\mathbb{C}}[P_1T^{-1}(C)-kP_0T^{-1}(C)]\leqq\min_{B\in\mathbb{B}}[P_1(B)-kP_0(B)]$. 故に定理3.2より, (*)をみたす$R_1$が存在する.

以下においても$\mathcal{P}$はdominated, λをdominate測度と仮定する. そのとき$\mathcal{P}=\{F_\theta(x)\lambda(dx):\theta\in\Theta\}$とすると, $(X,\mathbb{B},\mathcal{P})$から$(Y,\mathbb{C},\mathcal{Q})$への乱写像$T(C|x)$は

$$G_\theta(y)\mu(dy)=\int_X T(dy|x)F_\theta(x)\lambda(dx)$$

をみたす, ここで$\mathcal{Q}=\{G_\theta(y)\mu(dy):\theta\in\Theta\}$. 乱写像に関する(*)に対応する式は

$$(\dagger)\qquad R_1(dx|y)G_\theta(y)\mu(dy)=T(dy|x)F_\theta(y)\lambda(dx)$$

であり, (**)に対応する式は

$$(\dagger\dagger)\qquad F_\theta(x)\lambda(dx)=\int_Y R_2(dx|y)G_\theta(y)\mu(dy)$$

である. 乱写像に関するHalmos-Savageの定理(定理3.1)の拡張は次のようになる.

<u>定理3.4</u> 乱写像Tに対して$(\dagger)$をみたすR_1が存在するためには, Θの任意の2つの要素θ_1, θ_2の組を母数空間として

それに関して $R_{\theta_1,\theta_2}=R_1(B|y)$ が存在することが必要十分である.

この定理は直ちに証明出来る.

<u>定理3・5</u>　(†) をみたす乱写像 R_1 が存在するためには (††) をみたす R_2 が存在することが必要十分である.

<u>証明</u>　可測空間 $(X\times Y,\ \mathbb{B}\times\mathbb{C})$ 上の確率測度 $\nu_\theta(dx,dy)=T(dy|x)F_\theta(x)\lambda(dx)$, $\theta\in\Theta$, の全体は統計空間を作り, $f(x,y)=y$ は $(X\times Y,\ \mathbb{B}\times\mathbb{C},\ \nu_\theta,\ \theta\in\Theta)$ から $(\ \ ,\ \mathbb{C},\ G_\theta(y)\mu(dy),\ \theta\in\Theta)$ への統計(量)である. しかも (††) を用いると, $\overline{R}(A|y)=\int_X T(A_x|x)R_2(dx|y)$, $A\subset X\times Y$, $A\in\mathbb{B}\times\mathbb{C}$, $A_x=\{y:(x,y)\in A\}$, とおけば

$$\int_A \nu_\theta(dx,dy)=\int_Y \overline{R}(A|y)G_\theta(y)\mu(dy)$$

が成立する. これは f が定義3・4 の意味で十分であることを示めしている ($\overline{R}$ は $(**)$ の R_2 の役割をしている). したがつて定理3・3 が使えて

$$\nu_\theta(A\cap f^{-1}(C))=\int_C \overline{\overline{R}}(A|y)G_\theta(y)\mu(dy),\ C\in\mathbb{C},$$

をみたす $(Y,\mathbb{C},q)$ から $(X\times Y,\ \mathbb{B}\times\mathbb{C},\ \nu_\cdot,\Theta)$ への乱写像 $\overline{\overline{R}}$ が存在する. この $\overline{\overline{R}}(A|y)$ は $y\notin R_x$ に対しては 0 で,

$y \in A_x$ に対しては A_x の測度となるから, $R_2(B:y) = \overline{R}(B \times Y : y)$ とおくと, R_2は乱写像で(††)をみたす. 逆は明らかである.

次の定理は乱写像に関するNeymanの因子分解定理の拡張である.

<u>定理3·6</u> $T(C|x)$が$(X, \mathbb{B}, F_\theta \lambda(dx), \theta \in \Theta)$からある統計空間$(Y, \mathbb{C}, G_\theta(y)\mu(dy), \theta \in \Theta)$への十分乱写像であるためには, λ可積分関数$f(x)$と, 各θに対し, $\mathbb{C}$可測関数$K_\theta(y) \geqq 0$が存在して, 直~~積~~測空間$(X \times Y, \mathbb{B} \times \mathbb{C})$上で確率測度$T(dy|x)F_\theta(x)\lambda(dx)$に対して殆んど常に

$$F_\theta(x) = K_\theta(y) f(x)$$

がなりたつことが必要十分である.

<u>証明</u> Tが十分乱写像であれば, 定理3·5から(†)が成立する$R_1(B|y)$が存在する. しかるにHalmos-Savage [10]によれば, $\{F_\theta(x)\lambda(dx)\}$はその可算要素の凸結合$\lambda_0(dx) = \sum_{i=1}^\infty c_i F_{\theta_i}(x)\lambda(dx)$で同等(互いに絶対連続の意味で)なものが存在する. このとき$f(x) = \sum_{i=1}^\infty c_i F_{\theta_i}(x)$, $g(y) = \sum_{i=1}^\infty c_i G_{\theta_i}(y)$とおくと

$$T(dy|x) f(x) \lambda(dx) = R_1(dx|y) g(y) \mu(dy).$$

したがつて $(+)$ から，測度 $T(dy|x)f(x)\lambda(dx)$ で 0 の集合を除けば $F_\theta(x)/f(x)=G_\theta(y)/g(y)$ が成立し，我々の条件が必要であることが証明された．次にわれわれの条件がみたされると仮定して T が十分乱写像であることを証明する．$k=\int_X f(x)\lambda(dx)$，$\overline{f}(x)=f(x)/k$，$\overline{K_\theta}(y)=kK_\theta(y)$ とおき，改めて $\overline{f}$，$\overline{K_\theta}$ を f，K_θ と書くことにする．このとき $\mu(dy)=\int_X T(dy|x)f(x)\lambda(dx)$ とおけば，測度 $\int_X T(dy|x)F_\theta(x)\lambda(dx)$ は μ に関して絶対連続となるから，その Radon-Nikodym 微分 $G_\theta(y)$ が定義される．次に任意の $B\in\mathbb{B}$ に対して $\int_B T(dy|x)f(x)\lambda(dx)$ は μ に関して絶対連続であるから，$R(dx|y)\mu(dy)=T(dy|x)f(x)\lambda(dx)$ が成立する条件付確率 $R(dx|y)$ が存在する．これらのことを用いて計算すると

$$\int_Y R(dx|y)G_\theta(y)d\mu=\int_Y R(dx|y)\int_X T(dy|\overline{x})F_\theta(\overline{x})\lambda(d\overline{x})$$

$$=\int_Y R(dx|y)\int_X T(dy|\overline{x})K_\theta(y)f(\overline{x})\lambda(d\overline{x})$$

$$=\int_Y R(dx|y)\int_X K_\theta(y)R(d\overline{x}|y)\mu(dy)$$

$$=\int_Y R(dx|y)K_\theta(y)\mu(dy)$$

$$= \int_Y K_\theta(y) f(x) T(dy|x) \lambda(dx)$$

$$= \int_Y F_\theta(x) T(dy|x) \lambda(dx) = F_\theta(x) \lambda(dx)$$

となつて，Rは(††)のR_2の役目をはたし，Tは十分乱写像であることがわかる．

§4 統計空間の表現

本節では$\mathcal{P}$はσ有限な測度λでdominateされると仮定し，$f_\theta(x)$をその一般密度関数$dP_\theta/d\lambda$とする．そのとき

$$\tilde{f}_x(\theta) = f_\theta(x)$$

とおくと，$\tilde{f}.$はXからR^Θの中への写像$x \longrightarrow \tilde{f}_x(\cdot)$とみなされ，$\tilde{f}.$による$X$の像$Y=\{\tilde{f}_x : x \in X\}$ $(\subset R^\Theta)$は統計(量)$\tilde{f}.$の値域である．R^Θに各点収束位相を導入して，YをR^Θの部分位相空間とみなすことができるから，そのボレル集合族またはベール集合族を$\mathbb{C}$とすると，$\tilde{f}.$は$(X, \mathbb{B})$から$(Y, \mathbb{C})$への可測写像となる．

$u \in R^\Theta$に対して$u(\theta) \in R$を対応させる写像をφ_θとすると，

$$\varphi_\theta(\tilde{f}_x) = f_\theta(x)$$

かつ

$$P_\theta(B) = \int_B \varphi_\theta(\tilde{f}_x)\,\lambda(dx)$$

がなりたつ．したがつてNeymanの因子分解定理により，$\tilde{f}.$は十分統計(量)となる．統計(量)$\tilde{f}.$によつて，$\mathcal{P}$を$(Y, \mathbb{C})$上に移した分布空間を$\mathcal{Q}$とすると，上のことは実験$(X, \mathbb{B}, \mathcal{P})$と$(Y, \mathbb{C}, \mathcal{Q})$とが等価であることを示めしている．$(Y, \mathbb{C}, \mathcal{Q})$を$(X, \mathbb{B}, \mathcal{P})$の<u>表現</u>と云うこともできる．

なお$\tilde{f}.$は最小十分統計量であることも証明される(Dynkin [11])．

次にΘを実数空間とし（n次元ユークリッド空間としてもよいが，簡単のため$n=1$とする）$f_\theta(x) > 0$，θについて連続微分可能とする．そのときYからR^{Θ}の中への対応

$$i:\ \tilde{f}. \longrightarrow \frac{d \log \tilde{f}.}{d\theta} \quad (= Z.\ とおく)$$

は1：1であるから，iも十分統計(量)であつて，$Z.$の統計空間も1つの表現である．この表現統計空間は$Z.(\theta) \to E_\theta[Z.(\theta)^2]$によつて，Fisherの情報量に結びつく．

なお本講究録で草間氏はYに別種の位相を入れて論じている．

§5 新しい統計空間の構成

(A) <u>Truncated experiment</u>. 与えられた統計空間 $(X, \mathbb{B}, \mathcal{P})$ および X の与えられた部分集合 Y に対して, $\mathbb{C} = \{C \cap Y : C \in \mathbb{B}\}$, $\mathcal{Q} = \{P_\theta(dx \cap Y)/P_\theta(Y) : P_\theta \in \mathcal{P}\}$ を作り, 新しく統計空間 $(Y, \mathbb{C}, \mathcal{Q})$ を考えるとき, この空間を truncated experiment という, ただし Y は $\mathbb{B}$ 可測とする.

(B) <u>部分統計空間</u>. 与えられた統計空間 $(X, \mathbb{B}, \mathcal{P})$ に対して, $\mathbb{B}$ の部分 σ 代数 $\mathbb{C}$, $\mathcal{P}$ の要素 P_θ の $\mathbb{C}$ への縮小 Q_θ の全体 $\mathcal{Q}$ を用いて作られる新しい統計空間 $(X, \mathbb{B}, \mathcal{Q})$ のこと.

(C) <u>合成統計空間</u>, <u>周辺統計空間</u>, <u>条件付統計空間</u> 2つの可測空間 $(X, \mathbb{B})$, $(Y, \mathbb{C})$ の直積空間 $(X \times Y, \mathbb{B} \times \mathbb{C})$ 上の統計空間 $(X \times Y, \mathbb{B} \times \mathbb{C}, \mathcal{P})$ が与えられたとする. このとき, $\mathcal{P}^X = \{P(dx \times Y) : P \in \mathcal{P}\}$ とおき, 新しい統計空間 $(X, \mathbb{B}, \mathcal{P}^X)$ を考える. このとき $(X, \mathbb{B}, \mathcal{P}^X)$ のことを X に関する周辺統計空間という. さらに x を定めたとき $C (\in \mathbb{C})$ の確率測度であり C を定めたとき x の $\mathbb{B}$ 可測関数である $Q_\theta(C : x)$ が存在して

$$P_\theta(B \times C) = \int_B Q_\theta(C : x)\, P_\theta(dx \times Y)$$

がなりたつとき, $\mathcal{Q}_x^Y = \{Q_\theta^Y(dy : x) : \theta \in \Theta\}$ とおく. $(Y, \mathbb{C}, \mathcal{Q}_x^Y)$ を与えられた x に対する条件付実験という.

もし与えられたxに対する条件付実験がYに関する周辺実験に一致するとき，Xに関する実験とYに関する実験は独立であるといい，そうでないときそれらは従属であるという．

（D）　混合実験．合成実験$(X\times Y,\ \mathbb{B}\times\mathbb{C},\ \mathcal{P})$において$Y$に関する周辺実験の分布$P_\theta(X\times dy)$が$\theta$に関係ないとき，$X$に関する周辺実験を与えられた$y$に対する条件付実験の混合実験という．

（E）　極限実験．極限実験を考える際標本空間を共通にとつて考えることができる．それは例えば§4で述べた表現統計空間を考えるなどの方法でその場合にreduceできるからである．統計的実験列$\{(X,\ \mathbb{B}_n,\ P_{\theta n},\ \theta\in\Theta)\}$および同じく実験$(X,\ \mathbb{B}_0,\ P_\theta,\ \theta\in\Theta)$が与えられたものとする．これらの実験に対して

（1）すべての自然数nに対してσ代数$\mathbb{A}_n$があり，

$\mathbb{A}_n\supset\mathbb{B}_n$，　$P_{\theta n}$が$\mathbb{A}_n$まで拡張され，

$\mathbb{A}_1\subset\mathbb{A}_2\subset\cdots\subset\mathbb{A}_n\subset\mathbb{A}_{n+1}\subset\cdots$ ，　$\mathbb{A}_0=[\bigcup_{n=1}^{\infty}\mathbb{A}_n]^\sigma$

$\mathbb{B}_0\subset\mathbb{A}_0$，　P_θは$\mathbb{A}_0$まで拡張される．

ただし$[\ \ ]^\sigma$は$[\ ]$内の集合族を含む最小のσ代数．

（2）　$\lim_{n\to\infty}\|P_{\theta n}-P_\theta\|_{\mathbb{B}_n}=0$

ただし$\|\mu\|_{\mathbb{B}}$は signed measure μ の$\mathbb{B}$への縮小の全変分．

（3）　Θの点列$\{\theta_i\}$，　正数列$\{c_i\}$，$\sum_{i=1}^{\infty}c_i=1$，に対して

$$\lambda_n = \sum_{i=1}^{\infty} c_i P_{\theta_i n} \quad , \quad \lambda = \sum_{i=1}^{\infty} c_i P_{\theta_i}$$

とおくとき，$\lambda_1, \lambda_2, \cdots, \lambda$は対応するという．$\lambda_1, \lambda_2, \cdots,$ λが対応するとき，いかなる$f \in \bigcup_{n=1}^{\infty} M_n$に対しても

$$(\S) \qquad \lim_{n \to \infty} \int_X |E_{\lambda_n}[f \mid \mathbb{B}_n]| \lambda_n(dx) = \int_X |E_\lambda[f \mid \mathbb{B}_0]| \lambda(dx),$$

ただしM_nはA_n可測な有界関数の全体で；$E_\lambda[f \mid \mathbb{B}]$は確率測度$\lambda$による与えられた$\mathbb{B}$に対する条件付平均をあらわす．を満すとき，実験$(X, \mathbb{B}_0, P_\theta, \theta \in \Theta)$を実験列$\{(X, \mathbb{B}_n, P_{\theta n}, \theta \in \Theta)\}$の極限という．

$f \in M_n$に対して

$$\left| \int_X |E_{\lambda_n}[f \mid \mathbb{B}_n]| \lambda_n(dx) - \int_X |E_\lambda[f \mid \mathbb{B}_n]| \lambda(dx) \right|$$

$$= \left| \ \left\| \int f \lambda_n(dx) \right\|_{\mathbb{B}_n} - \left\| \int f \ \lambda(dx) \right\|_{\mathbb{B}_n} \right|$$

$$\leqq \left\| \int f \ (\lambda_n - \lambda)(dx) \right\|_{\mathbb{B}_n} = \left(\max_{x \in X} |f(x)| \right) \| \lambda_n - \lambda \|_{\mathbb{B}_n}$$

$$\leqq \left(\max_{x \in X} |f(x)| \right) \sum_{i=1}^{\infty} c_i \| P_{\theta_i n} - P_{\theta_i} \|_{\mathbb{B}_n} \to 0 .$$

ここで$\int f \ \lambda(dx)$等はfの不定積分を表わす．したがって上の$(\S)$は

$$\lim_{n \to \infty} \int_X |E_\lambda[f \mid \mathbb{B}_n] \lambda(dx) = \int_X |E_\lambda[f \mid \mathbb{B}_0]| \lambda(dx)$$

と同等である．なおM_nはL^1空間$L^1(X, A_n, \lambda)$で稠密であ

るから，［8］の Theorem 2 から次の定理が得られる．

定理5・1　実験$(X, \mathbb{B}_0, P_\theta, \theta\in\Theta)$が実験列$\{(X, \mathbb{B}_n, P_{\theta n}, \theta\in\Theta)\}$の極限であり，$\mathcal{P}=\{P_\theta : \theta\in\Theta\}$が$\mathbb{A}_0$上で dominated であるとする．そのとき各$(X, \mathbb{B}_n, P_{\theta n}, \theta\in\Theta)$が実験$(X, \mathbb{A}_n, P_{\theta n}, \theta\in\Theta)$に対して十分であれば，$(X, \mathbb{B}_0, P_\theta, \theta\in\Theta)$は$(X, \mathbb{A}_0, P_\theta, \theta\in\Theta)$に対して十分である．

ここで定義された極限実験が LeCam［7］の定義した極限実験とどのような関係にあるかということはまだ一切明らかではない．

§6　標本空間上の変換群により不変な統計空間．

省略

§7　標本空間と母数空間との間の duality．

母数空間上の先験確率は統計学において重要な概念の一つである．これを考えるために，Θにはその部分集合のσ代数$\mathcal{a}$を固定して考え，標本分布P_θはすべての$A\in\mathbb{A}$に対して$P_\theta(A)$がθの関数として$\mathcal{a}$可測なものであるとする，ただし標本空間はX，統計的構造は$\mathbb{A}$，標本分布P_θは$\mathbb{A}$上の確率測度であるとしている．そこで$\mathcal{a}$上の確率測度ξ（これを先験分布という）に関して，$(X\times\Theta, \mathbb{A}\times\mathcal{a})$上の確率分布

$$\tilde{\xi}(A\times\alpha)=\int_{\alpha}P_{\theta}(A)\,\xi(d\theta),\quad \alpha\in\mathcal{A},\ A\in\mathbb{A},$$

を定義する.

<u>補題</u>. 確率空間$(Z,\mathcal{Z},\nu)$の$\mathcal{Z}$の部分σ代数$\mathcal{F}$, $\mathcal{G}$に対して,

すべての$G\in\mathcal{G}$に対して $\nu(G|\mathcal{F})=\nu(G|\mathcal{F}\cap\mathcal{G})$
ν-a.e.

がなりたつときまたそのときに限り

すべての$F\in\mathcal{F}$に対して $\nu(F|\mathcal{G})=\nu(F|\mathcal{F}\cap\mathcal{G})$
ν-a.e.

証明は次の2つの式から明かである.

$$\int_F\nu(G|\mathcal{F})\,d\nu=\nu(F\cap G)=\int_G\nu(F|\mathcal{G})\,d\nu$$

$$\int_F\nu(G|\mathcal{F}\cap\mathcal{G})\,d\nu=\int_Z\nu(G|\mathcal{F}\cap\mathcal{G})\nu(F|\mathcal{F}\cap\mathcal{G})\,d\nu$$

$$=\int_G\nu(F|\mathcal{F}\cap\mathcal{G})\,d\nu.$$

この補題を確率空間$(X\times\Theta,\mathbb{A}\times\mathcal{A},\tilde{\xi})$, $\mathbb{A}\times\mathcal{A}$の部分σ代数$\mathbb{C}\times\mathcal{B}$, $\mathbb{B}\times\mathcal{C}$に応用すると

<u>定理7.1</u> すべての$B\in\mathbb{B}$, $\gamma\in\mathcal{C}$に対して

$$\tilde{\xi}(B\times\gamma|\mathbb{C}\times\mathcal{B})=\tilde{\xi}(B\times\gamma|(\mathbb{B}\cap\mathbb{C})\times(\mathcal{B}\cap\mathcal{C}))$$

$$\tilde{\xi}\text{-a.e.}$$

がなりたつときまたそのときに限り

すべての $C \in \mathbb{C}$, $\beta \in \mathcal{B}$ に対して

$$\tilde{\xi}(C\times\beta \mid \mathbb{B}\times\mathcal{C}) = \tilde{\xi}(C\times\beta \mid (\mathbb{B}\cap\mathbb{C})\times(\mathcal{B}\cap\mathcal{C}))$$
$$\tilde{\xi}\text{-a.e.}$$

この定理7·1の2つの同値な命題のいづれかがなりたつとき，次のような記号であらわす：

$$\mathbb{B}:\mathbb{C} \cong \mathcal{B}:\mathcal{C} \quad (\xi)$$

またすべての ξ について上式がなりたつときは (ξ) を除く.

$\mathbb{O}$ を $\mathbb{A}$ の最小の部分 σ 代数 $\{\phi, X\}$ とし，$\mathcal{O}$ を $\mathcal{A}$ の最小の σ 代数 $\{\phi, \Theta\}$ とする. そのとき

$$\tilde{\xi}(A\times\Theta \mid \mathbb{O}\times\mathcal{A}) = P_\theta(A)$$
$$\tilde{\xi}(A\times\Theta \mid \mathbb{C}\times\mathcal{A}) = P_\theta(A \mid \mathbb{C})$$
$$\tilde{\xi}(X\times\beta \mid \mathbb{O}\times\mathcal{O}) = \xi(\beta)$$
$$\tilde{\xi}(X\times\beta \mid \mathbb{O}\times\mathcal{C}) = \xi(\beta \mid \mathcal{C})$$
$$\tilde{\xi}(A\times\Theta \mid \mathbb{O}\times\mathcal{O}) = \int_\Theta P_\theta(A)\,\xi(d\theta)$$
$$\tilde{\xi}(A\times\beta \mid \mathbb{O}\times\mathcal{O}) = \int_\beta P_\theta(A)\,\xi(d\theta)$$
$$\tilde{\xi}(A\times\Theta \mid \mathbb{A}\times\mathcal{O}) = \text{posterior distr. } \xi(A \mid x)$$

が $\tilde{\xi}$-a.e. で成立する. 故に $(\forall\xi : \xi(N)=0)$ が $N=\phi$ のときは

$$\mathbb{C} \text{ が十分}\sigma\text{代数} \Longleftrightarrow \mathbb{A}:\mathbb{C} \cong \mathcal{A}:\mathcal{O}.$$

$\mathbb{C}$が ancillary ⇔ $\mathbb{C}:\mathbb{O} \cong \mathcal{A}:\mathcal{O}$.

（$\mathbb{C}$のすべての要素Cに対して$P_\theta(C)$がθに無関係であるとき，$\mathbb{C}$は ancillary σ-algebra という．Basu [12]をみよ）.

$\mathbb{B}$と$\mathbb{C}$が独立 ⇔ $\mathbb{B}:\mathbb{C} \cong \mathcal{A}:\mathcal{A}$.

すべての$B \in \mathbb{B}$について$P_\theta(B)$は$\mathcal{B}$可測

$$\Leftrightarrow \mathbb{B}:\mathbb{O} \cong \mathcal{A}:\mathcal{B}.$$

$\mathcal{B}$が parametrically sufficient（$\forall A \in \mathbb{A}$: $P_\theta(A)$が$\mathcal{B}$可測なこと．Kudō [13], Barankin [14], Barankin-Kudō [15]をみよ）

$$\Leftrightarrow \mathbb{A}:\mathbb{O} \cong \mathcal{A}:\mathcal{B}$$

したがって$\mathcal{B}$が parametrically sufficient であるための必要十分条件は，母数空間と標本空間を逆に考えたとき，$\mathcal{B}$が十分σ代数となることである．

なお（$\forall \xi : \xi(N) = 0$）と$N = \phi$とが同等であるための1つの十分条件については Barankin-Kudō [16]をみよ．

REFERENCES

[1] Blackwell, D. and Girshick, M.A.: Theory of Games and Statistical Decisions, John Wiley & Sons, 1954.

[2] Kudō, H.: Dependent experiments and sufficient statistics, Nat. Sci. Rep. Ochanomizu Univ. 4(1953), pp.151-163.

[3] LeCam, L.: Sufficiency and approximate sufficiency, Ann. Math. Stat. 35(1964), pp.1419-1455.

[4] Sacksteder, R.: A note on statistical equivalence, Ann.

Math. Stat. 38(1967), pp.787-794.

[5] Basu, D. and Ghosh, J.K.: Sufficient statistics in samplin from a finite universe, 36th Session of I.S.I. 1967.

[6] Morimoto, H.: 本講究録掲載論文.

[7] LeCam, L.: Limits of experiments, Unpublished. 1971

[8] Kudō, H.: On an approximation to a sufficient statistic including a concept of asymptotic sufficiency, Journ. Fac. Sci. Univ. of Tokyo, Sec. I, 17(1970), 273-290.

[9] Berger, A.: Remark on separable spaces of probability measures, Ann. Math. Stat. 22(1951), pp.119-120.

[10] Halmos, P.R. and Savage, L.J.: Applications of the Radon-Nikodym theorem to the theory of sufficient statistics, Ann. Math. Stat. 20(1949), pp.225-241.

[11] Dynkin, E. B.: On sufficient and necessary statistics for families of probability distributions, Doklady Acad. Nauk SSSR 75(1950), pp.68-90 (英訳は Amer. Math. Soc., Selected Translation in Math. Stat. and Prob. 1(1961), pp.17-40.)

[12] Basu, D.: The family of ancillary statistics, Sankhya 21 (1959), pp. 247-256.

[13] Kudō, H.: On partial prior information and the property of parametricsufficiency, Proc. 5th Berkeley Symp. Math. Stat. and Prob. I, 1965.

[14] Barankin, E.W.: Sufficient parameters: solution of the minimal dimensionality problem, Ann. Inst. Stat. Math. 12(1960), pp.91-118.

[15] Barankin, E.W. and Kudō, H.: A general theorem on sufficiency, unpublished. 1965

[16] Barankin, E.W. and Kudō, H.: A remark on universl measurability with respect to a subalgebra induced by a statistic, unpublished. 1966