On statistics which are almost sufficient from the viewpoint of the Fisher metrics

山口夏穂里 (Kaori YAMAGUCHI)
立命館大学大学院理工学研究科博士課程
前期課程基礎理工学専攻数理科学コース

概要

統計学の基本的な概念である十分統計量は，Ay-Jost-Lê-Schwachhöfer らによって Fisher 計量を用いて情報幾何学の分野で再定義された．本稿では，Ay-Jost-Lê-Schwachhöfer らが再定義した十分統計量を定量的に一般化し，その特徴づけを行う．これは立命館大学，野澤啓氏との共同研究である．

1 導入

1.1 先行研究

統計学者の Fisher は，パラメータづけされた確率分布の族が与えられたとき，その族に対する統計量の中でも，推定を行うための情報を漏らさず含んでいる統計量に着目した [F22]．その後 Neyman がこのような統計量を定式化し十分統計量と名付けた [N35]．その後，Peter らによって測度論の言葉のみで線形十分統計量と再定義され，Decell がその定義を一般化し，概線形十分計量が考えられた [DJ80]．Peter，Decell らは統計量が線形の場合で議論しているが，近年 Jost らは線形非線形に関わらずかつ無限次元の場合にも考えられる統計モデルを情報幾何学の分野で議論している．

2015 年 Jost らは，情報幾何の分野で，統計モデルを多様体の点でパラメータ付けられた確率密度関数の族を，無限次元の場合でも微分幾何的に扱えるように再定義した [AJLS15]．この統計モデルを（統計）多様体とみなすことで，Rao が導入した Fisher 計量という Riemann 計量を統計モデル上で扱える．また，Jost らは統計量によって誘導されたモデルの Fisher 計量が元のモデルの Fisher 計量と完全に一致する統計量として十分統計量を再定義した [AJLS17]．本稿では Jost らが情報幾何学の言葉で再定義した十分統計量の条件を定量的に緩めることによって一般化し，δ 概十分統計量と名付けその特徴づけを行う．

1.2　準備

まず，Jost らが再定義したパラメータ付き測度モデルの定義を紹介する．

定義 1.1 (パラメータ付き測度モデル)**.** パラメータ付き測度モデルとは，次を満たす三つ組 $(M, \Omega, \mathbf{p})$ のことである．

- M：有限次元または無限次元バナッハ多様体,
- Ω：可測空間,
- $\mathbf{p} : M \to \mathcal{M}(\Omega)$: C^1　ただし，$\mathcal{M}(\Omega) := \{\mu : \mu \text{は}\Omega\text{上有限測度}\}$.

パラメータ付き測度モデルとは，一言で述べると，バナッハ多様体でパラメトライズされた測度の族である．このパラメータ付き測度モデルの例として，以下のような n 回のコイントスのモデルを考える．

例 1.2. 三つ組 $(M, \Omega, \mathbf{p})$ として，以下を考える．$n \in \mathbb{N}$ に対して，

- $M = (0, 1)$,
- $\Omega = \{0, 1\}^n$,
- $\mathbf{p}(\xi) = \xi^{a(\omega)}(1-\xi)^{n-a(\omega)}\mu_0$，$\omega \in \Omega$　(ただし，μ_0 は数え上げ測度，$a(\omega)$ は ω に現れる 1 の数）．

するとこのとき，この三つ組はパラメータ付き測度モデルである．

この例では，コインの表が出る確率 $\xi \in M$ がパラメータである．
次に，統計量の定義を紹介する．

定義 1.3 (統計量)**.**　可測空間 Ω と Ω' に対して, 可測写像 $a : \Omega \to \Omega'$ を統計量という．

例 1.2. で与えた写像 $a(\omega)$ は統計量になっている．ここで，$\Omega' = \{0, 1, ..., n\}$ とすると（Ω' は可測空間)，例 1.2 の a は，Ω から Ω' への写像であるが，この統計量 a は例 1.2 のパラメータ付き測度モデルに対して，$\mathbf{p}(\xi) = a_*\mathbf{p}(\xi)$ とすると，次のような別のパラメータ付き測度モデルを誘導する．

例 1.4. 三つ組 $(M, \Omega, \mathbf{p})$ として，以下を考える．

- $M = (0, 1)$,
- $\Omega' = \{0, 1, ..., n\}$,
- $\mathbf{p}'(\xi) = {}_nC_k \xi^k (1-\xi)^{n-k}\mu_0$ $(\forall k \in \Omega')$　ただし，μ_0 は数え上げ測度.

するとこのとき, この三つ組 $(M, \Omega', \mathbf{p}')$ はパラメータ付き測度モデルであり, 例 1.2 の $(M, \Omega, \mathbf{p})$ から，統計量 a によって誘導されたモデルという．一般に元のモデル $(M, \Omega, \mathbf{p})$ から統計量

$a:\Omega\to\Omega'$ によって誘導されるモデル $(M,\Omega',\mathbf{p}')$ は多様体 M は変わらず，統計量によって可測空間が取り替えられ，測度の族は $\mathbf{p}'=a_*\mathbf{p}$ によって定義される．

次に，Jost らが十分統計量を情報幾何学の分野で再定義するのに用いた Fisher 計量を紹介するが，それには，測度の根，k 次可積分性という概念が必要なので，測度の根を説明した後に k 次可積分性の定義を紹介する．

Ω 上の有限測度の集合 $\mathcal{M}(\Omega)$ を順序関係 $\le$ を持つ順序集合として考える．ただし，ω_1, $\omega_2\in\mathcal{M}(\Omega)$ に対して，$\omega_1\le\omega_2$ とは ω_1 が ω_2 に dominate されているときをいう．$r>0$ に対して，$\mathcal{M}(\Omega)$ 上の帰納極限

$$\mathcal{S}^r(\Omega):=\lim_{\to}L^{1/r}(\Omega,\mu)$$

を考える．ただし，$L^{1/r}(\Omega,\mu)$ とは，固定された $r\in(0,1]$ に対して，$\int_\Omega|\phi|^{1/r}d\mu<\infty$ を満たす Ω 上可測関数全体のなす空間．ここで $\mu_1\le\mu_2$ のとき, 包含写像 $L^{1/r}(\Omega,\mu_1)\to L^{1/r}(\Omega,\mu_2)$ はバナッハ空間の等長写像になっている．したがって，$\mathcal{S}^r(\Omega)$ はバナッハ空間の構造を持つ．$\phi\in L^{1/r}(\Omega,\mu)$ に対して, 写像 $L^{1/r}(\Omega,\mu)\to\mathcal{S}^r(\Omega)$ の像を $\phi\mu^r$ と表し，このようにして μ の r-根 μ^r は定義される．また，$r>0$ に対して,

$$\mathcal{M}^r(\Omega):=\{\,\phi\mu^r\in\mathcal{S}^r(\Omega)\mid\mu\in\mathcal{M}(\Omega),\phi\ge0\,\}$$

とおく．

定義 1.5. $k\in\mathbb{N}$ に対して, パラメータ付き測度モデル $(M,\Omega,\mathbf{p})$ が k 次可積分であるとは，写像

$$\mathbf{p}^{1/k}:M\to\mathcal{M}^{1/k}(\Omega),\qquad\xi\mapsto\mathbf{p}(\xi)^{1/k}$$

が C^1 級であるときにいう．

定義 1.6 (Fisher 計量)**.** $(M,\Omega,\mathbf{p})$ を 2 次可積分パラメータ付き測度モデルとする．また，$\mathbf{p}(\xi)$ が $\mu_0\in\mathcal{M}(\Omega)$ によって dominate されているとする．すなわち，Ω 上の有限測度 μ_0 と密度関数 p に対して，$\mathbf{p}$ が $\mathbf{p}(\xi)=p(\omega;\xi)\mu_0$ と書けるとする．このとき M 上の Fisher 計量 $\mathfrak{g}$ は M 上の任意の接ベクトル v,w に対して，以下の式で定義される．

$$\mathfrak{g}(v,w)=\int_\Omega(\partial_v\log p(\cdot;\xi))(\partial_w\log p(\cdot;\xi))d\mathbf{p}(\xi).\tag{1}$$

この計量は Fisher 情報量の変わり具合の平均を取ることで定義されている．Fisher 計量と統計量について，次の単調性定理が成り立つ．

定理 1.7 ([AN00])**.** $(M,\Omega,\mathbf{p})$ を 2 次可積分パラメータ付き測度モデルとする．また，$\mathbf{p}(\xi)$ が測度 μ_0 によって dominate されているとする．このとき，以下の不等式が M の任意の接ベクトル v に対して成り立つ；

$$\mathfrak{g}'(v,v)\le\mathfrak{g}(v,v).\tag{2}$$

ただし，$\mathfrak{g}'$ は統計量 κ によって誘導されたモデルの Fisher 計量.

この定理をもう少し直観的に説明すると，統計量で誘導されたモデルの Fisher 計量は，元のモデルの Fisher 計量を超えないということである．差 $\mathfrak{g}(v,v)-\mathfrak{g}'(v,v)$ は 2 次情報損失と言われる．次のように Jost らはこの情報損失が 0 であるとき，その統計量を十分統計量と呼んだ．

定義 1.8 (十分統計量)**.** 2 次可積分パラメータ付き測度モデル $(M,\Omega,\mathbf{p})$ 上の統計量 κ が十分であるとは，M 上の任意の接ベクトル v に対して，

$$\mathfrak{g}'(v,v)=\mathfrak{g}(v,v)$$

が成り立つときにいう．

n 回のコイントスを考えたとき，例 1.2. で定義した a は十分統計量である．ノルムを計算するとそれが確かめられる．

今回得た主結果においては Jost らが定義した十分統計量を定量的に一般化した δ 概十分統計量を導入し，その特徴づけを得た．まず，δ 概十分統計量の定義を紹介する．

定義 1.9 (δ 概十分統計量)**.** $0<\delta\leq 1$ に対して， 統計量 κ が δ 概十分であるとは，M 上の任意の接ベクトル v に対して，

$$\delta^2\mathfrak{g}(v,v)\leq\mathfrak{g}'(v,v)\leq\mathfrak{g}(v,v)$$

が成り立つときにいう．

二つ目の不等式は上で紹介した単調性定理によりいつでも成り立っており，これに加えて $\mathfrak{g}$ と $\mathfrak{g}'$ が双リプシッツ同値になるような統計量を δ 概十分統計量と定義した．実際にこの定義を満たす統計量があることを以下の例で示す．

例 1.10. 以下のパラメータ付き測度モデルを考える．$M=(0,1)$, $\Omega=\{0,1\}^2$，$\mathbf{p}(\xi)=p(\cdot;\xi)\mu_0$, ただし μ_0 は Ω 上数え上げ測度．また，密度関数 $p(\cdot;\xi)$ は，$p(\cdot;\xi)=\xi^{a(\omega)}(1-\xi)^{2-a(\omega)}\quad(\omega\in\Omega)$, ただし，$a(\omega)$ は $\omega\in\Omega$ の 1 が出た回数と定める．

ここで，$\Omega'=\{0,1\}$ とし，統計量 $\kappa:\Omega\to\Omega'$ を $\kappa(\omega_1,\omega_2)=\omega_1,\quad\forall(\omega_1,\omega_2)\in\Omega$ と定めると，

$$\|\partial_\xi\log p(\xi)\|_2^2=\frac{2}{\xi(1-\xi)},$$
$$\|\partial_\xi\log\kappa^*p'(\xi)\|_2^2=\frac{1}{\xi(1-\xi)}.$$

となり，$\frac{\|\partial_\xi\log\kappa^*p'\|_2^2}{\|\partial_\xi\log p\|_2^2}=\frac{1}{2}$ であるので, κ は $\frac{1}{\sqrt{2}}$ 概十分統計量であり， $\forall\delta>\frac{1}{\sqrt{2}}$ に対して，δ 概十分統計量ではない．

2　主結果

1 章では δ 概十分統計量を定義し，2 章ではその特徴づけを紹介するが，この特徴付けは Jost らが示した十分統計量の特徴づけとパラレルに作ったので，まずは Jost らが得た十分統計量の特徴

づけを紹介する．

定理 2.1 ([AJLS17])**.** $(M,\Omega,\mathbf{p})$ を Ω 上の有限測度 μ_0 に対して $\mathbf{p}(\xi)=p(\omega;\xi)\mu_0$ の形で書ける 2 次可積分パラメータ付き測度モデルとする．また，密度関数 p は正とする．$\kappa\colon\Omega\to\Omega'$ を統計量とし，$p'(\omega;\xi)$ を $\kappa_*\mu_0$ に対する $\kappa_*\mathbf{p}(\xi)$ の密度関数とする．そして，M は連結とする．このとき以下は同値である．

(i) κ が $(M,\Omega,\mathbf{p})$ に対して十分である．
(ii) M の任意の接ベクトル v に対して $\partial_v \log p=\partial_v \log\kappa^* p'$ が成り立つ．
(iii) 写像 $\xi\mapsto\frac{p(\cdot;\xi)}{p'(\kappa(\cdot);\xi)}$ が定値写像.
(iv) $\exists s:\Omega'\times M\to\mathbb{R}$, $\exists t\in L^1(\Omega,\mu_0)$ s.t.

$$p(\omega;\xi)=s(\kappa(\omega);\xi)t(\omega),\qquad \mu_0\text{-}a.e.\ \omega\in\Omega,\forall\xi\in M.$$

前定理に対して，δ 概十分統計量を定理 2.1 とパラレルに特徴付けられ，これが主結果である．

定理 2.2 ([NY23])**.** $(M,\Omega,\mathbf{p})$ を，密度関数 p が正になるような Ω 上の有限測度 μ_0 に対して $\mathbf{p}(\xi)=p(\omega;\xi)\mu_0$ の形で書ける 2 次可積分パラメータ付き測度モデルとする．また，M を Fisher 計量を備える C^2 級有限次元多様体とする． さらに，統計量 $\kappa:\Omega\to\Omega'$ および $0<\delta\leq 1$ に対して，以下は同値である．

(i) κ が $(M,\Omega,\mathbf{p})$ に対して δ 概十分である．
(ii) M の任意の接ベクトル v に対して，

$$\left\|\partial_v \log\frac{p(\cdot;\xi)}{\kappa^* p'(\cdot;\xi)}\right\|\leq\sqrt{1-\delta^2}\|\partial_v\log p(\cdot;\xi)\| \tag{3}$$

が成り立つ．ただし，$\|\cdot\|$ は L^2 ノルムである．
(iii) 写像 $M\to L^2(\Omega,\mu_0);\xi\mapsto\log\frac{p(\cdot;\xi)}{p'(\kappa(\cdot);\xi)}$ が Fisher 計量と L^2 計量によって定義される距離に対して $\sqrt{1-\delta^2}$ 局所リプシッツである．
(iv) $\exists s:\Omega'\times M\to\mathbb{R}$, $\exists t:\Omega\times M\to\mathbb{R}_{>0}$ s.t. $\log t(\cdot;\xi)\in L^2(\Omega,\mu_0)\ \forall\xi\in M$,
   - $p(\omega;\xi)=s(\kappa(\omega);\xi)t(\omega;\xi)$ μ_0-a.e. $\omega\in\Omega,\forall\xi\in M$,
   - the map $M\to L^2(\Omega,\mu_0);\xi\mapsto\log t(\cdot;\xi)$ が Fisher 計量と L^2 計量によって定義される距離に対して $\sqrt{1-\delta^2}$ 局所リプシッツである．

証明は [NY23] を参照されたい．これらの結果を用いて，コイントスのモデルの統計量について調べられる．

参考文献

[F22] Fisher, R.A., On the mathematical foundations of theoretical statistics *Philos. Trans. R. Soc. Lond. Ser. A Math. Phys. Eng. Sci.* containing Papers of a Mathematical or Physical Character 222, 309–368. `http://doi.org/10.1098/rsta.1922.0009`

[N35] Neyman, J., Sur un teorema concernente le considette statistiche sufficienti, *Giorn. Istit. Ital. Att.* **6** 320–334, (1935).

[DJ80] Decell, H.P., Jr., Sufficient, almost sufficient statistics and applications. *Proceedings of the First World Conference on Mathematics at the Service of Man (Barcelona, 1977)*, Vol. I, pp. 541–550, Univ. Politec., Barcelona, 1980.

[AJLS15] Ay, N., Jost, J., Lê, H.V., Schwachhöfer, L., *Parametrized measure models*, Bernoulli **24**(3) 1692–1725, (2015). `https://doi.org/10.3150/16-BEJ910`

[AJLS17] Ay, N., Jost, J., Lê, H.V., Schwachhöfer, L., *Information Geometry*, Ergeb. Math. Grenzgeb. (3) Folge. A Series of Modern Surveys in Mathematics, 64. Springer, Cham, 2017.

[AN00] Amari, S., Nagaoka, H., *Methods of Information Geometry*, Translated from the 1993 Japanese original by Daishi Harada. Transl. Math. Monogr. 191. American Mathematical Society, Providence, RI; Oxford University Press, Oxford, 2000.

[NY23] Yamaguchi, K, Nozawa, H On statistics which are almost sufficient from the viewpoint of the Fisher metrics, arXiv(2023)