

The Amount of Partial Information and Sufficiency

筑波大・数学 大谷内奈穂 (Nao Ohyauchi)
筑波大・数学 赤平 昌文 (Masafumi Akahira)

1 はじめに

Fisher([F25])は母数に対して統計量の補助性の概念を導入し、補助統計量の値が与えられたときに条件付原理を用いた。また、局外母数が存在するときに、Godambe([G80])は十分性と補助性について論じた。そしてそのようなときに、Bhappkar([Bh89], [Bh91])は補助統計量による条件付操作について論じ、さらに関心のある母数に対する統計量の部分的十分性(partial sufficiency)のいくつかの定義について考察した。なお部分的十分性については、Basu([Ba78])の論説がある。

本論では、局所母数が存在するとき、部分的情報量と十分性について論じ、部分的(情報)十分性の定義を与え、その特徴付けを行う。

2 部分的情報量と十分性

確率ベクトル $\mathbf{X} := (X_1, \dots, X_n)$ の (σ -有限測度 μ に関する) 同時確率密度関数 (joint probability density function 略して j.p.d.f.) を $f_{\theta, \eta}(\mathbf{x})$ とする。ただし、 $\mathbf{x} := (x_1, \dots, x_n)$, θ は関心のある実母数とし、 η は実ベクトル値局外母数とする。このとき、適当な正則条件の下で、Fisher 情報行列を

$$I^{\mathbf{X}} := \begin{pmatrix} I_{\theta\theta}^{\mathbf{X}} & I_{\theta\eta}^{\mathbf{X}} \\ I_{\eta\theta}^{\mathbf{X}} & I_{\eta\eta}^{\mathbf{X}} \end{pmatrix} \\ := \begin{pmatrix} E_{\theta} \left[\left\{ \frac{\partial}{\partial \theta} \log f_{\theta, \eta}(\mathbf{X}) \right\}^2 \right] & E_{\theta} \left[\left\{ \frac{\partial}{\partial \theta} \log f_{\theta, \eta}(\mathbf{X}) \right\} \left\{ \frac{\partial}{\partial \eta} \log f_{\theta, \eta}(\mathbf{X}) \right\} \right] \\ E_{\theta} \left[\left\{ \frac{\partial}{\partial \theta} \log f_{\theta, \eta}(\mathbf{X}) \right\} \left\{ \frac{\partial}{\partial \eta} \log f_{\theta, \eta}(\mathbf{X}) \right\} \right] & E_{\theta} \left[\left\{ \frac{\partial}{\partial \eta} \log f_{\theta, \eta}(\mathbf{X}) \right\} \left\{ \frac{\partial}{\partial \eta} \log f_{\theta, \eta}(\mathbf{X}) \right\} \right] \end{pmatrix}$$

によって定義する。ただし、 $\frac{\partial}{\partial \eta} \log f_{\theta, \eta}(\mathbf{X})$ は $\frac{\partial}{\partial \eta} \log f_{\theta, \eta}(\mathbf{X})$ の転置ベクトルとする。また、 θ に対する $\mathbf{X}$ の部分的情報量を

$$I_{\theta|\eta}^{\mathbf{X}} := I_{\theta\theta}^{\mathbf{X}} - I_{\theta\eta}^{\mathbf{X}} (I_{\eta\eta}^{\mathbf{X}})^{-1} I_{\eta\theta}^{\mathbf{X}} \quad (2.1)$$

で定義する。そして同様に、 θ に対する統計量 $T = t(\mathbf{X})$ の Fisher 情報行列を

$$I^T := \begin{pmatrix} I_{\theta\theta}^T & I_{\theta\eta}^T \\ I_{\eta\theta}^T & I_{\eta\eta}^T \end{pmatrix}$$

によって定義し、 θ に対する T の部分的情報量を

$$I_{\theta|\eta}^T := I_{\theta\theta}^T - I_{\theta\eta}^T (I_{\eta\eta}^T)^{-1} I_{\eta\theta}^T \quad (2.2)$$

で定義する。さらに、すべての θ とすべての η に対して

$$I_{\theta|\eta}^T = I_{\theta|\eta}^X$$

が成り立つとき、 T は部分的に(情報)十分であるという。

3 部分的(情報)十分性の特徴付け

この節では、 T が部分的に(情報)十分であるための必要十分条件を求める。まず、母数 θ と η の直交性の条件 $I_{\theta\eta}^X \equiv 0$ 、すなわち、すべての θ とすべての η に対して $I_{\theta\eta}^X = 0$ が成り立つと仮定する。一般には、 $I_{\theta\eta}^X \neq 0$ であるが、適当な関数 ϕ を用いて局外母数を変換 $\tilde{\eta} := \phi(\theta, \eta)$ によって再定義すれば、 $I_{\theta\tilde{\eta}}^X \equiv 0$ になることも多い。このとき次の定理が成り立つ。

定理 1 $I_{\theta\eta}^X \equiv 0$ と仮定する。このとき $I_{\theta|\eta}^X \equiv I_{\theta|\eta}^T$ であるための必要十分条件は、 $I_{\theta\theta}^T \equiv I_{\theta\theta}^X$ である。

証明の概略 まず、行列 $I^X - I^T$ は非負定符号になる。そこで、 $I_{\theta|\eta}^X \equiv I_{\theta|\eta}^T$ とすれば、 $I_{\theta\eta}^X \equiv 0$ という条件を用いて、(2.1)と(2.2)から

$$I_{\theta\theta}^X \equiv I_{\theta\theta}^T - I_{\theta\eta}^T (I_{\eta\eta}^T)^{-1} I_{\eta\theta}^T$$

になる。また、 $I_{\theta\theta}^X \geq I_{\theta\theta}^T$ で $(I_{\eta\eta}^T)^{-1}$ が正定値であるから、 $I_{\theta\theta}^X \equiv I_{\theta\theta}^T$ になり、 $I_{\theta\eta}^T \equiv 0$ となる。

逆については、まず、 $I_{\theta\eta}^X \equiv 0$ であるから、(2.1)と(2.2)から行列

$$I^X - I^T \equiv \begin{pmatrix} I_{\theta\theta}^X - I_{\theta\theta}^T & -I_{\theta\eta}^T \\ -I_{\eta\theta}^T & I_{\eta\eta}^X - I_{\eta\eta}^T \end{pmatrix}$$

は非負定符号となる。そして、 $I_{\theta\theta}^X \equiv I_{\theta\theta}^T$ とすれば、 $I_{\theta\eta}^T \equiv 0$ となる。□

次に、 T が与えられたときの X の条件付確率密度関数(conditional p.d.f.略してc.p.d.f)を $h_{\theta,\eta}(\mathbf{x}|T)$ 、 T のp.d.f.を $g_{\theta,\eta}(T)$ とすると

$$f_{\theta,\eta}(\mathbf{x}) = h_{\theta,\eta}(\mathbf{x}|t) g_{\theta,\eta}(t) \quad a.e. \quad (3.1)$$

となる。また、

$$I_{\theta\theta}^X - I_{\theta\theta}^T = E \left[E \left[\left\{ \frac{\partial}{\partial \theta} \log h_{\theta,\eta}(\mathbf{X}|T) \right\}^2 \mid T \right] \right]$$

になる(Fisher[F25], Rao[R61]). よって, $I_{\theta\theta}^{\mathbf{X}} \equiv I_{\theta\theta}^T$ ならば

$$\frac{\partial}{\partial \theta} \log h_{\theta, \eta}(\mathbf{x}|t) = 0 \quad a.e.$$

となるから, $h_{\theta, \eta}(\mathbf{x}|t)$ は θ に無関係になる. 従って, (3.1)から

$$f_{\theta, \eta}(\mathbf{x}) = h_{\eta}(\mathbf{x}|t) g_{\theta, \eta}(t) \quad a.e. \quad (3.2)$$

になる. 逆に, (3.2)が成り立つとすれば, $I_{\theta\theta}^T \equiv I_{\theta\theta}^{\mathbf{X}}$ になる. よって, 次の定理を得る.

定理 2 $I_{\theta\eta}^{\mathbf{X}} \equiv 0$ と仮定する. このとき, 統計量 $T = t(\mathbf{X})$ が θ に対して部分的(情報)十分であるための必要十分条件は, $\mathbf{X}$ のj.p.d.f.が

$$f_{\theta, \eta}(\mathbf{x}) = h_{\eta}(\mathbf{x}|t) g_{\theta, \eta}(t) \quad a.e.$$

となることである.

上の定理 2 は, Neyman型因子分解定理であり, 非常に有用である.

例 1 $X_1, \dots, X_n$ をたがいに独立に, いずれも正規分布 $N(\mu, \sigma^2)$ に従う確率変数とする. このとき, $\mathbf{X} := (X_1, \dots, X_n)$ のj.p.d.f.は

$$f_{\theta, \eta}(\mathbf{x}) = \frac{1}{(\sqrt{2\pi}\eta)^n} \exp \left\{ -\frac{1}{2\eta} \sum_{i=1}^n (x_i - \bar{x})^2 \right\} \exp \left\{ -\frac{n}{2\eta^2} (\bar{x} - \theta)^2 \right\}$$

になる. ただし, $\mathbf{x} := (x_1, \dots, x_n)$, $\bar{x} = \sum_{i=1}^n x_i/n$ とする. よって, 定理 2 より, $\bar{X}$ は θ に対して部分的(情報)十分である.

例 2 $X_1, \dots, X_n$ を確率変数とし, $z_{ji} (j = 1, \dots, k; i = 1, \dots, n)$ を定数の値をとる独立変数とする. そして, 線形回帰モデル

$$X_i = \beta_0 + \beta_1 z_{1i} + \dots + \beta_k z_{ki} + \epsilon_i \quad (i = 1, \dots, n)$$

を考える. ただし, $\epsilon_1, \dots, \epsilon_n$ はたがいに独立に, いずれも $N(0, \sigma^2)$ に従う確率変数とし, σ^2 は既知とする. このとき, $\mathbf{X} := (X_1, \dots, X_n)$ のj.p.d.f.は

$$f(\mathbf{x}) = \frac{1}{(\sqrt{2\pi}\sigma)^n} \exp \left\{ -\frac{1}{2\sigma^2} \sum_{i=1}^n (x_i - \beta_0 - \beta_1 z_{1i} - \dots - \beta_k z_{ki})^2 \right\}$$

になる. ただし, $\mathbf{x} := (x_1, \dots, x_n)$ とする. ここで, β_1 を関心のある母数とし, 他のすべての母数を局外母数とする. 通常, β_1 と $(\beta_0, \beta_2, \dots, \beta_p)$ は(情報)直交しない. そこで

$$\sum_{i=1}^n \omega_i = 0, \quad \sum_{i=1}^n \omega_i z_{ji} = 0 \quad (j = 2, \dots, k)$$

となるように

$$\omega_i := z_{1i} - c_1 - c_2 z_{2i} - \cdots - c_k z_{ki} \quad (i = 1, \dots, n)$$

をとり,

$$\gamma_0 := \beta_0 + c_1 \beta_1, \quad \gamma_2 := \beta_2 + c_2 \beta_1, \quad \dots, \quad \gamma_k := \beta_k + c_k \beta_1$$

と変換する. このとき

$$\begin{aligned} f(\mathbf{x}) &= \frac{1}{(\sqrt{2\pi}\sigma)^n} \exp \left\{ -\frac{1}{2\sigma^2} \sum_{i=1}^n (x_i - \gamma_0 - \beta_1 \omega_i - \gamma_2 z_{2i} - \cdots - \gamma_k z_{ki})^2 \right\} \\ &= \frac{1}{(\sqrt{2\pi}\sigma)^n} \exp \left\{ -\frac{1}{2\sigma^2} \left(\sum_{i=1}^n x_i^2 - 2\beta_1 \sum_{i=1}^n \omega_i x_i + \beta_1^2 \sum_{i=1}^n \omega_i^2 \right) \right\} \\ &\quad \cdot \exp \left\{ -\frac{1}{2\sigma^2} \sum_{i=1}^n (\gamma_0 + \gamma_2 z_{2i} + \cdots + \gamma_k z_{ki})^2 \right\} \end{aligned}$$

になる. よって, β_1 は他のすべての母数 $\gamma_0, \gamma_2, \dots, \gamma_k$ に直交していることが分かり, また, 定理 2 より, $T := \sum_{i=1}^n \omega_i X_i$ またはそれと同等な最小 2 乗推定量

$$\hat{\beta}_1 := \sum_{i=1}^n \omega_i X_i \bigg/ \sum_{i=1}^n \omega_i^2$$

が β_1 に対する部分的(情報)十分になる.

参考文献

- [Ba78] Basu, D.(1978). On partial sufficiency : A review. *J. Statist. Plann. Inference* **2**, 1-13.
- [Bh89] Bhapkar, V.P.(1989). Conditioning on ancillary statistics and loss of information in the presence of nuisance parameters. *J. Statist. Plann. Inference* **21**, 139-160.
- [Bh91] Bhapkar, V.P.(1991). Loss of information in the presence of nuisance parameters and partial sufficiency. *J. Statist. Plann. Inference* **28**, 185-203.
- [F25] Fisher, R.A.(1925). Theory of statistical estimation. *Proc. Camb. Phil. Soc.*, **22**, 700-725.
- [G80] Godambe, V.P.(1980). On sufficiency and ancillarity in the presence of a nuisance parameter. *Biometrika* **67**, 155-162.