

論文審査の要旨

| | | | |
|---|--------------------|----|-------|
| 博士の専攻分野の名称 | 博 士 (理 学) | 氏名 | 稲 津 佑 |
| 学位授与の要件 | 学位規則第 4 条第 ①・2 項該当 | | |
| 論文題目 | | | |
| An unbiased C_p type criterion for ANOVA model with a tree order restriction (Tree order restriction が課せられた ANOVA モデルにおける不偏な C_p タイプの規準量) | | | |
| 論文審査担当者 | | | |
| 主 査 教 授 若木 宏文 審査委員 教 授 井上 昭彦 審査委員 教 授 田丸 博士 | | | |
| 〔論文審査の要旨〕 | | | |
| <p>分散分析 (Analysis of Variance, ANOVA) とは、多群の母平均の差異を分析する統計手法の一つである。例えば、ある農作物に肥料 A, B, C を与えて育て、収穫された農作物の平均重量間の与えた肥料による違いを分析するときに用いられる。肥料 A が従来使われてきた肥料で、肥料 B, C が開発中の肥料であるとき、比較対象である肥料 A を与えた作物を対照群と呼び、肥料 B, あるいは肥料 C を与えた作物を処理群などと呼ぶ。</p> <p>分散分析において、よく用いられる統計モデルは次のようなものである。対照群と処理群を合わせて k^* 個の群があり、第 i 群には N_i 個の観測変量 (作物の重量など) Y_{ij} ($j = 1, \dots, N_i$) があるとする。このとき、Y_{ij} ($i = 1, \dots, k^*; j = 1, \dots, N_i$) は独立に正規分布 $N(\mu_i, \sigma^2)$ に従う確率変数であると仮定する。</p> <p>第 1 群を対照群とし、第 i 群 ($i = 2, \dots, k^*$) に割り当てられた処理を、単に「処理 i」と呼ぶことにする。 Y_{ij} ($i = 1, \dots, k^*; j = 1, \dots, N_i$) の実際の観測値から、例えば、$k^* = 6$ として、</p> $\mu_1 < \mu_2 = \mu_3 < \mu_4 = \mu_5 = \mu_6$ <p>であると結論づけることができれば、「どの処理も効果があるが、処理 2 と処理 3 は同等、処理 4, 5, 6 は同等であり、処理 2,3 よりも 処理 4,5,6 の方が効果がある」といったことが言える。 $1 \leq k \leq k^*$ に対し、Q_1, \dots, Q_k を $J_{k^*} := \{1, \dots, k^*\}$ の分割、すなわち、Q_1, \dots, Q_k は空でない互いに素な自然数の集合で、$J_{k^*} = Q_1 \cup \dots \cup Q_k$ を満たすものとし、</p> <p style="text-align: center;">任意の $i, i' \in \{1, \dots, k^*\}$ に対して、i, i' が同じ Q_j に含まれるなら $\mu_i = \mu_{i'}$</p> <p>というモデルを考える。 J_{k^*} の異なる分割のそれぞれに、異なる統計モデルを考えることができ、可測値から特定のモデルを選択することによって、どの処理と処理が同等の効果を与えるかという結論を下すことができる。</p> <p>ところで、処理群の母平均 μ_i ($i \geq 2$) が、対照群の母平均 μ_1 以上であるという仮定を tree order restriction と呼ぶ。 tree order restriction が真である場合には、この制約</p> | | | |

条件下での母平均の最尤推定量は、通常の標本平均よりも精度が良いことが知られている。本論文の著者は、 J_{k^*} の分割によって定まるモデルに、tree order restriction を付加したモデルの選択問題を考え、 C_p 基準と呼ばれるモデル選択規準の予測平均 2 乗誤差に基づくリスクの推定量としての偏差を導出し、その偏差を修正した C_p 型のモデル選択規準 TOC_p を提案し、以下の定理を得た。

定理 以下で与えられる TOC_p はリスクに対する一様最小分散不偏推定量である：

$$\text{TOC}_p := (N - k^* - 2) \frac{\hat{\sigma}^2}{\sigma^2} + 2(\hat{m} + 1).$$

ただし、

$$\bar{Y}_i = \frac{1}{N_i} \sum_{j=1}^{N_i} Y_{ij}, \quad \hat{\sigma}^2 = \frac{1}{N} \sum_{i=1}^{k^*} \sum_{j=1}^{N_i} (Y_{ij} - \bar{Y}_i)^2, \quad \hat{m} = 1 + \sum_{a=2}^k 1_{\{\hat{\theta}_1 < \hat{\theta}_a\}}$$

であり、 $\hat{\sigma}^2$ は、モデルの下での σ^2 の最尤推定量、 $\hat{\theta}_j$ ($j = 1, \dots, k$) は Q_j に含まれる群に共通の母平均 $\theta_j = \mu_v$ ($\exists v \in Q_j$) の最尤推定量である。

tree order restriction のように母平均間に不等式制約がある場合、最尤推定量は標本平均 $\bar{Y}_1, \dots, \bar{Y}_{k^*}$ 間の $k^*!$ 通りの大小関係が成り立つそれぞれの領域で異なる表現形を持ち、リスクの推定量としての偏差の評価も、領域ごとに評価して足し合わせるが必要になる。例えば、

$$k^* = 4, \quad k = 3, \quad Q_1 = \{1\}, Q_2 = \{2\}, Q_3 = \{3, 4\}, \\ \bar{Y}_2 < \bar{Y}_1 < \bar{Y}_4 < \bar{Y}_3$$

の場合、 μ_1, \dots, μ_4 の最尤推定量は

$$\hat{\mu}_1 = \hat{\mu}_2 = \frac{N_1 \bar{Y}_1 + N_2 \bar{Y}_2}{N_1 + N_2}, \quad \hat{\mu}_3 = \hat{\mu}_4 = \frac{N_3 \bar{Y}_3 + N_4 \bar{Y}_4}{N_3 + N_4}$$

となる。この結果は、標本平均の順序のみならず、 J_{k^*} の分割の与え方によっても変わる。本論文では、一見して複雑な積分式で与えられる偏差が、 $P(\hat{\theta}_1 < \hat{\theta}_a)$ ($a = 2, \dots, k$) の和を用いて表わされるという予想を立て、その証明に成功した。また、提案された選択規準 TOC_p は、単純な形で使いやすく、またリスクの不偏推定量として理論的に最適なものであることも示されている。

以上、審査の結果、本論文の著者は博士（理学）の学位を授与される十分な資格があるものと認める。

公表論文

(1) An unbiased C_p type criterion for ANOVA model with a tree order restriction.

Y. Inatsu, *Hiroshima Mathematical Journal* (2017), to appear.