2-step 単調欠測データの下での T^2 型検定統計量の漸近分布について

東京理科大・理・院　川崎　玉恵
東京理科大・理　　瀬尾　隆

1 標本問題における 2-step 単調欠測データの下での平均ベクトルの仮説検定統計量の漸近分布について議論する．2-step 単調欠測データとは

$$\begin{pmatrix} x_{11}^{(1)} & x_{12}^{(1)} & \cdots & x_{1p_1}^{(1)} & x_{1p_1+1}^{(1)} & \cdots & x_{1p}^{(1)} \\ \vdots & \vdots & & \vdots & \vdots & & \vdots \\ x_{n_11}^{(1)} & x_{n_12}^{(1)} & \cdots & x_{n_1p_1}^{(1)} & x_{n_1p_1+1}^{(1)} & \cdots & x_{n_1p}^{(1)} \\ x_{11}^{(2)} & x_{12}^{(2)} & \cdots & x_{1p_1}^{(2)} & * & \cdots & * \\ \vdots & \vdots & & \vdots & \vdots & & \vdots \\ x_{n_21}^{(2)} & x_{n_22}^{(2)} & \cdots & x_{n_2p_1}^{(2)} & * & \cdots & * \end{pmatrix} = \begin{pmatrix} \boldsymbol{x}_{11}^{(1)\prime} & \boldsymbol{x}_{21}^{(1)\prime} \\ \vdots & \vdots \\ \boldsymbol{x}_{1n_1}^{(1)\prime} & \boldsymbol{x}_{2n_1}^{(1)\prime} \\ \boldsymbol{x}_{1}^{(2)\prime} & \boldsymbol{*} \\ \vdots & \vdots \\ \boldsymbol{x}_{n_2}^{(2)\prime} & \boldsymbol{*} \end{pmatrix}$$

のような形をしており，n_1 個の標本ベクトル $\boldsymbol{x}_1^{(1)}, \boldsymbol{x}_2^{(1)}, \ldots, \boldsymbol{x}_{n_1}^{(1)}$ がそれぞれ独立に $N_p(\boldsymbol{\mu}, \Sigma)$ に従い，n_2 個の標本ベクトル $\boldsymbol{x}_1^{(2)}, \boldsymbol{x}_2^{(2)}, \ldots, \boldsymbol{x}_{n_2}^{(1)}$ がそれぞれ独立に $N_{p_1}(\boldsymbol{\mu}_1, \Sigma_{11})$ に従うとする．ただし，“$*$” は欠測値を表し，$\boldsymbol{x}_j^{(1)} = (\boldsymbol{x}_{1j}^{(1)\prime}, \boldsymbol{x}_{2j}^{(1)\prime})'$ である．また，$\boldsymbol{\mu}_1$ は $\boldsymbol{\mu}$ の p_1 次元部分ベクトル，Σ_{11} は Σ の $p_1 \times p_1$ ブロック行列とし，n_1 個の標本ベクトルと n_2 個の標本ベクトルはそれぞれ独立な分布に従っているとする．このときの平均ベクトルの仮説検定問題 $H_0 : \boldsymbol{\mu} = \boldsymbol{\mu}_0$ vs. $H_1 : \boldsymbol{\mu} \neq \boldsymbol{\mu}_0$ に対する検定統計量として，次のような Hotelling の T^2 型統計量を用いる．ただし，$\boldsymbol{\mu}_0$ は既知ベクトルで，$n = n_1 + n_2$ である．

$$T^2 = (\widehat{\boldsymbol{\mu}} - \boldsymbol{\mu}_0)' \begin{pmatrix} \widehat{\Sigma}_{11}/n & \widehat{\Sigma}_{12}/n \\ \widehat{\Sigma}_{21}/n & \widehat{\mathrm{Cov}}(\widehat{\boldsymbol{\mu}}_2) \end{pmatrix}^{-1} (\widehat{\boldsymbol{\mu}} - \boldsymbol{\mu}_0)$$

このときの最尤推定量 $\widehat{\boldsymbol{\mu}}$ と $\widehat{\Sigma}$，$\widehat{\mathrm{Cov}}(\widehat{\boldsymbol{\mu}}_2)$ については Kanda and Fujikoshi (1998) によって与えられている．

本報告では，Kawasaki and Seo (2015) で議論されている T^2 型統計量に対する統計量の展開を利用し，T^2 型統計量の漸近分布についての議論を行う．具体的には，確率展開を $N_1, N_2 \to \infty$，$N_i/N \to$positive constants のもとで

$$\boldsymbol{z}^{(1)} = \sqrt{N_1} \begin{pmatrix} \overline{\boldsymbol{x}}_1^{(1)} - \boldsymbol{\mu}_1 \\ \overline{\boldsymbol{x}}_2^{(1)} - \boldsymbol{\mu}_2 \end{pmatrix},\ \boldsymbol{z}^{(2)} = \sqrt{N_2}(\overline{\boldsymbol{x}}^{(2)} - \boldsymbol{\mu}_1),\ V^{(1)} = \sqrt{N_1}(S^{(1)} - I_p),\ V^{(2)} = \sqrt{N_2}(S^{(2)} - I_{p_1})$$

とおき，T^2 型統計量の展開を行っている．ただし，$N_i = n_i - 1$，$N = N_1 + N_2$ とする．さらに，T^2 型統計量に対するバイアス修正統計量と Fujikoshi (2000) による修正 Bartlett 変換統計量についても議論する．結果は川崎，瀬尾 (2014) で提案したものとは異なるもので，カイ 2 乗近似への収束をより改良したものになっている．モンテカルロ・シミュレーションによる数値的評価の結果や考察，今後の課題については当日報告する．

参考文献

[1] Fujikoshi, Y. (2000). Transformations with improved chi-square approximations. *J. Multivariate Anal.*, **72**, 249–263.

[2] Kanda, T. and Fujikoshi, Y. (1998). Some basic properties of the MLE's for a multivariate normal distribution with monotone missing data. *Amer. J. Math. Manage. Sci.*, **18**, 161–190.

[3] Kawasaki, T. and Seo, T. (2015). Bias correction for T^2 type statistic with two-step monotone missing data. *to appear in Statistics.*

[4] 川崎玉恵，瀬尾隆 (2014)．2-step 単調欠測データのもとでの平均ベクトルに対する検定統計量のバイアス修正．2014 年度 統計関連学会連合大会講演報告集，64.