RMSEA に対する BCa 法の適用

紺谷幸弘，狩野裕
大阪大学 大学院基礎工学研究科

本発表では，構造方程式モデリング (Structural equation modeling; SEM) の適合度指標の一つである RMSEA に対する，BCa 法 (Efron, 1987) の有用性を数値実験により検討した．

適合度指標は，モデルのデータに対する当てはまりの良さの指標であり，これまでに様々なものが考案されている．適合度指標の代表的な例としては，CFI や GFI，AGFI がある．RMSEA(Root Mean Square Error of Approximation) もまた，現在用いられている代表的な適合度指標の一つである．Steiger and Fouladi (1997) によれば，RMSEA は区間推定されるべき指標であるが，従来の区間推定法では高い精度が期待できない．

BCa 法は，非常に広い族に対し高い精度を持つ区間推定法であり，RMSEA の区間推定においても高い精度が期待されるが，数値実験からは従来法の優位性が示唆された．

1 RMSEA の従来の区間推定法と BCa 法

RMSEA は，母 RMSEA と呼ばれる量 ε (Steiger and Lind, 1980) の推定量の一つである．母 RMSEA は，未知の母数ベクトル $\boldsymbol{\theta} \in \mathbb{R}^q$ で規定される，共分散行列 $\boldsymbol{\Sigma}(\boldsymbol{\theta})$ が，真の母共分散行列 $\boldsymbol{\Sigma}_0$ にどれだけ近いかという観点からモデルの適合を評価する適合度指標である．

RMSEA における区間推定の従来法は，漸近的な結果に基づいた議論によって導出されている．しかし，標本のサイズが小さい場合には，漸近論と実際の間の乖離が大きくなるので，漸近論に基づいた従来の区間推定では高い精度が期待できない．

多くの場合において，最も自然な区間推定の方法は，興味のある未知母数の推定量（の関数で与えられる量）の分布を利用することである．しかしながら，一般にこの分布は未知であるので，その代替としてサンプルサイズが大きいときの分布（極限分布）が用いられることが多い（極限分布による近似）．極限分布による近似を用いた方法の短所は，極限分布と実際の分布の乖離が大きい場合，構成される信頼区間の，名義上の信頼率と実際の被覆確率の間のズレ（被覆誤差）が大きくなることである．被覆誤差は，統計量の真の分布とそれを近似する分布の間のズレ（近似誤差）の大きさに比例する．したがって，近似誤差を小さくすることによって，区間推定法の改良を実現することができる．

BCa 法 (Efron, 1987) は，bootstrap 法を用いた区間推定法の一つであり，percentile 法 (Efron, 1979) を改良したものである．percentile 法は，極限分布に基づく信頼区間と同程度の被覆誤差を持つが，極限分布さえ不明な統計量に対しても適用できることが長所の一つである．percentile 法の導出も極限分布による近似を用いた議論によって行われており，この近似の精度が良くないため，被覆誤差が大きくなっている．BCa 法は，統計量のバイアスや，統計量の標準偏差と未知母数の関数関係を考慮することによって，近似の誤差を小さくし，percentile 法より小さな被覆誤差を達成している．

したがって，RMSEA の信頼区間構成に BCa 法を用いることによって，従来法に比べて被覆誤差が小さい信頼区間が得られると考えられる．本発表では，数値実験により，従来法と BCa 法の被覆誤差を比較・検討した．

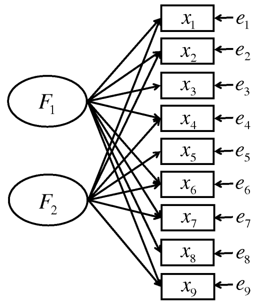

図 1: 数値実験においてデータに当てはめる直交モデル．F_1 は x_5，F_2 は x_3 と x_8 を除く全ての観測変数 x_i にパスがひかれている．(母 RMSEA $\varepsilon = 0.076$)

2 数値実験

2.1 設定

母平均ベクトルとして $\boldsymbol{\mu} = \mathbf{0}$，母共分散行列として $\boldsymbol{\Omega} = \boldsymbol{\Lambda}\boldsymbol{\Lambda}' + \boldsymbol{\Psi}$ を持つ 9 変量正規標本 $N_9(\mathbf{0}, \boldsymbol{\Omega})$ について検討する．ここに，

$$\boldsymbol{\Lambda} = \begin{pmatrix} .668 & .692 & .500 & .839 & .700 & .800 & .670 & .442 & .775 \\ .304 & .236 & .287 & -.321 & -.319 & -.372 & .385 & .245 & .424 \end{pmatrix}'$$

$$\boldsymbol{\psi} = (\psi_i) = (.462 \quad .465 \quad .667 \quad .192 \quad .408 \quad .221 \quad .403 \quad .745 \quad .219)'$$

で，これは Emmett (1949) の相関行列から，因子分析法における最尤法によって得られた，因子負荷行列 $\boldsymbol{\Lambda} = (\lambda_{ik})$，および独自分散 $\boldsymbol{\Psi} = (\psi_i)$ である．

この状況の下で，以下の手続きを行った．

(1) $\boldsymbol{X}_1, \ldots, \boldsymbol{X}_{150} \overset{\text{i.i.d.}}{\sim} N_9(\mathbf{0}, \boldsymbol{\Omega})$ から，不偏共分散行列 $\mathbf{S}$ を得る．

表 2.1: 数値実験の結果得られた各区間推定法の $\hat{P}(\alpha)$

α	従来法	percentile 法	BCa 法
0.05	0.04	0.54	0.07
0.95	0.99	0.99	0.54

(2) (1) で得られる $\mathbf{S}$ に対し，図 1 の 2 因子直交モデルを当てはめ，その母 RMSEA ϵ の両側等確率 90%区間 $\hat{I}(0.9)$ を (i) 従来法，および (ii) percentile 法，(iii) BCa 法により構成する．ここに，

$$\hat{I}(0.9) = \left[\hat{J}(0.05),\, \hat{J}(0.95)\right] \tag{2.1}$$

で，$\hat{J}(\alpha)$ は

$$\Pr\left(\hat{J}(\alpha) \geq \varepsilon\right) \approx \alpha \tag{2.2}$$

が期待される点である．

ただし，上の手続きは次の設定の下で行った．

- 母数ベクトル $\boldsymbol{\theta}$ の推定には最尤法を用いる
- bootstrap 法の resampling 手続きについて，9 変量正規分布 $N_9(\mathbf{0}, \boldsymbol{S})$ に従う $\boldsymbol{X}_1^*, \ldots, \boldsymbol{X}_{150}^*$ を bootstrap 標本として用いる．ここに，$\boldsymbol{S}$ は標本 $\boldsymbol{X}_1, \ldots, \boldsymbol{X}_{150}$ から得られた不偏共分散行列である．
- bootstrap resampling の回数は 500 回

(3) (1)〜(2) の手続きを 100 回繰り返す．

(4) (2) で挙げた各方法について，次の量を計算する．

$$\hat{P}(\alpha) = \frac{\#\left\{\hat{J}_k(\alpha) \geq \varepsilon;\, k = 1, \ldots, 100\right\}}{100} \tag{2.3}$$

ここに，$\hat{J}_k(\alpha)$ は (3) の手続きにおいて，k 番目に得られた $\hat{J}(\alpha)$ を表す．

(2.2) より，この $\hat{P}(\alpha)$ の値が α に近いほど，当該区間推定法が優れていることが示唆される．

上の数値実験は，統計解析ソフトウェア R（バージョン 2.7.2）を用いて行った．詳細は割愛するが，R の挙動が不安定になるのを避けるために，使用するデータ $\boldsymbol{X}_1, \ldots, \boldsymbol{X}_{150}$ および bootstrap 法で用いるリサンプリングデータ（bootstrap 標本）に対し，ある制約を設けた．

2.2 結果

表 2.1 は，数値実験の結果得られた，各区間推定法の $\hat{P}(\alpha)$ である．$\hat{P}(\alpha)$ の定義 (2.3) から，$\alpha \approx \hat{P}(\alpha)$ となることが望ましく，この差が小さいほど当該区間推定法の精度が高いことが示唆される．したがって，今回検討した三つの区間推定法では，従来法が精度の点で最も優れ，bootstrap 法を利用した二つの方法，percentile 法，BCa 法は従来法に比べ，かなり劣ることが示唆される．

2.3 考察

今回，RMSEA に対しては，bootstrap 法による区間推定法が全く機能しないという結果が示唆されたが，この原因を bootstrap 分布の構成が適切に行われていないことにあると考えるならば，bootstrap 分布が適切に構成できなかった原因はどこにあるのだろうか．考えられるものとして次が挙げられる．

(1) bootstrap 法における繰り返し回数が不足していること
（bootstrap 分布の構成には，bootstrap resampling の回数が 1000〜2000 程度必要であると言われているが，今回は 500 回しか行っていない）

(2) 検討回数が少ないこと
（今回の結果は，検討回数が 100 と少ないことによる，「偏り」のある結果かも知れない）

(3) RMSEA は BCa 法の直接の適用が不可能であること

(1),(2) は数値実験の設定に関する事項であり，次回これらの点に配慮して実験を行うことによって，疑問は解消される．(3) は根本的，かつ理論的な事項であり，非常に検証する価値のあるものである．

参考文献

[1] Efron, B. (1979) Bootstrap methods: another look at the jackknife, *Ann. Statist.* **7**, pp. 1-26

[2] Efron, B. (1987) Better confidence intervals, *Journal of the American Statistical Association*, **82**, No.397(Mar., 1987), pp. 171–185

[3] Emmett, W. G.(1949) Factor analysis by Lawley's method of maximum likelihood, *British J. Psych. Statist. Sec.* **2**, pp. 90–97

[4] Steiger, J. H. and Fouladi, R. T. (1997). Noncentrality interval estimation and the evaluation of statistical models. In *What If There Were No Significance Tests?*. (Harlow, L. L., Mulaik, S. A. and Steiger, J. A., Eds.), pp. 221–258. Mahwah, NJ: Erlbaum.

[5] Steiger, J. H. and Lind, J. C. (1980). *Statistically based tests for the number of common factors.* Paper presented at the annual meeting of the Psychometric Society, Iowa.