

信頼性係数のベイズ推定

岡本安晴

日本女子大学人間社会学部

1. はじめに

テスト得点 X を真値 T と残差 E の和

$$X = T + E$$

で表したとき (Crocker & Algina, 1986, p. 107) 信頼性係数 ρ は

$$\rho = \sigma_T^2 / \sigma_X^2$$

で与えられる (ibid. p. 116) ρ の推定値としてクロンバックの係数があり、

$$\alpha = \left(1 - \sum_i s_i^2 / s_X^2 \right) M / (M - 1)$$

で算出される。ここで、テスト X は N 個の下位テスト X_1, \dots, X_N に分けられ $X = \sum X_i$ であり、 s_i^2 と s_X^2 は X_i と X の分散である。係数は ρ の下限を与える推定値 (ibid. p. 121) であるが、 ρ と

の差がどれくらいであるのかはテストの信頼性を評価する上で重要である。 ρ の推定を、項目 i の被験者 j における得点 X_{ij} の確率モデルに基づいてベイズ的に直接行い、その可能性を探る。まず、得点が連続量の場合を扱い、次に得点が $0 - 1$ の 2 値の場合について検討する。

2. モデル - 1 : 連続量の場合

得点 X_{ij} を

$$X_{ij} = T_j + E_{ij}; \quad i = 1, \dots, N, j = 1, \dots, M \quad (1)$$

とおく。 T_j は被験者 j の真値であり、 E_{ij} は残差である。

$$T_j \sim N(\mu_T, \sigma_T^2), \quad E_{ij} \sim N(0, \sigma_E^2) \quad (2)$$

とおくとき、 ρ は

$$\rho = \sigma_T^2 / (\sigma_T^2 + \sigma_E^2 / N)$$

で与えられる。 X_{ij} の確率を

$$P(X_{ij} | T_j, \mu_T, \sigma_T, \sigma_E) = P(X_{ij} - T_j | T_j, \mu_T, \sigma_T, \sigma_E) = P(X_{ij} - T_j | \sigma_E) \sim N(0, \sigma_E)$$

とおき、事前分布を

$$P_0(T_j, \mu_T, \sigma_T, \sigma_E) = P(T_j | \mu_T, \sigma_T) P(\mu_T) P(\sigma_T) P(\sigma_E)$$

$$\mu_T \sim U_{(-BM, +BM)}, \quad \sigma_T \sim U_{(0, +BT)}, \quad \sigma_E \sim U_{(0, +BE)}$$

とおく。このとき、事後分布は

$$\begin{aligned} P(T_j, \mu_T, \sigma_T, \sigma_E | X_{11}, \dots, X_{NM}) &\propto P(X_{11}, \dots, X_{NM} | T_j, \mu_T, \sigma_T, \sigma_E) P_0(T_j, \mu_T, \sigma_T, \sigma_E) \\ &= P_0(T_j, \mu_T, \sigma_T, \sigma_E) \prod_{j=1}^M \left\{ P(T_j | \mu_T, \sigma_T) \prod_{i=1}^N P(X_{ij} - T_j | \sigma_E) \right\} \end{aligned}$$

で与えられる。事後分布の推定を次の MCMC により行う。

3 . アルゴリズム - 1

以下の Metropolis - within Gibbs アルゴリズムで MCMC を行う。

A. MCMC の準備

$\sigma_T^{(0)}$ 、 $\mu_T^{(0)}$ 、 $\sigma_E^{(0)}$ 、 $T_j^{(0)}$ に初期値を設定し、 $\rho^{(0)} = (\sigma_T^{(0)})^2 / \{(\sigma_T^{(0)})^2 + (\sigma_E^{(0)})^2 / N\}$ とおく。Proposal distribution のパラメタ $\sigma_{\sigma T}$ 、 $\sigma_{\mu T}$ 、 $\sigma_{\sigma E}$ 、 σ_{Tj} の値を設定して、 $s \leftarrow 0$ とおく。

$$B. \quad y \sim N(\sigma_T^{(s)}, \sigma_{\sigma T}^2), \quad \sigma_T^{(s+1)} = \begin{cases} y & \text{with probability } \alpha_{\sigma T s} \\ \sigma_T^{(s)} & \text{otherwise} \end{cases}$$

$$C. \quad y \sim N(\mu_T^{(s)}, \sigma_{\mu T}^2), \quad \mu_T^{(s+1)} = \begin{cases} y & \text{with probability } \alpha_{\mu T s} \\ \mu_T^{(s)} & \text{otherwise} \end{cases}$$

$$D. \quad y \sim N(\sigma_E^{(s)}, \sigma_{\sigma E}^2), \quad \sigma_E^{(s+1)} = \begin{cases} y & \text{with probability } \alpha_{\sigma E s} \\ \sigma_E^{(s)} & \text{otherwise} \end{cases}$$

E. For ($j = 1$ to M)

$$y \sim N(T_j^{(s)}, \sigma_{Tj}^2), \quad T_j^{(s+1)} = \begin{cases} y & \text{with probability } \alpha_{Tj s} \\ T_j^{(s)} & \text{otherwise} \end{cases}$$

$$F. \quad \rho^{(s+1)} \leftarrow (\sigma_T^{(s+1)})^2 / \{(\sigma_T^{(s+1)})^2 + (\sigma_E^{(s+1)})^2 / N\}$$

G. もし $s+1 < N\text{Simu}$ ならば、 $s \leftarrow s+1$ と設定して、ステップ B に戻る。

4 . 分析例 1

アルゴリズム - 1 による ρ の推定の可能性を、真の信頼性係数の値がわかっている人工データによって確認する。式(1)において、 $N = 20$ 、 $M = 2000$ とおき、式(2)において $\mu_T = 50$ 、 $\sigma_T = 10$ 、 $\sigma_E = 20$ とおいたモデルでデータを生成した。このとき

$$\rho = 10^2 / (10^2 + 20^2 / 20) \approx 0.833$$

である。生成したデータに対する 係数は

$$\alpha \approx 0.833$$

であり、 係数により信頼性係数 ρ の値の推定が行えることが確認できた。アルゴリズム - 1 を $N\text{Simu} = 10000$ で実行した。先頭の 1000 試行を burn-in として捨て、 $\rho^{(s)}$ の平均値 $\hat{\rho}$ と 95%HPD 区間を算出し、

$$\hat{\rho} \approx 0.833, \quad 95\%HPD \text{ 区間} = (0.8234, 0.8435)$$

を得た。これは、アルゴリズム - 1 によるベイズ的分析によって ρ の点推定値および 95%HPD 区間を得ることができることを示す結果である。

ここでのデータ X_{ij} は連続量のものであるが、次にデータが 2 値の場合について検討する。

5 . モデル - 2 : 2 値データの場合

得点 Y_{ij} が 0 - 1 の 2 値の場合のモデルを、式(1)に対応して式(3)で与える。

$$Y_{ij} = \begin{cases} 1 & \text{if } T_j + E_{ij} > c \\ 0 & \text{otherwise} \end{cases} \quad (3)$$

ここで、 c は Y_{ij} が値 1 をとるための基準値である。このとき、

$$\begin{aligned} P(Y_{ij} = 1 \mid c, T_j, \sigma_E) &= P(T_j + E_{ij} > c \mid c, T_j, \sigma_E) = P(E_{ij} > c - T_j \mid c, T_j, \sigma_E) \\ &= \Phi(-(c - T_j)/\sigma_E) \end{aligned} \quad (4)$$

$$P(Y_{ij} = 0 \mid c, T_j, \sigma_E) = \Phi((c - T_j)/\sigma_E) \quad (5)$$

となる。ここで、 $\Phi(z)$ は累積標準正規分布関数である。式 (4) と式 (5) をまとめると

$$P(Y_{ij} \mid c, T_j, \sigma_E) = \Phi(-(c - T_j)/\sigma_E)^{Y_{ij}} \Phi((c - T_j)/\sigma_E)^{1-Y_{ij}}$$

となる。式 (4) および式 (5) には原点と単位の任意性があるので、 $\mu_T = 0$, $\sigma_T = 1$ として、

$$T_j \sim N(0, 1), \quad E_{ij} \sim N(0, \sigma_E^2)$$

とおく。事前分布を

$$P_0(c, T_j, \sigma_E) = P(c)P(T_j)P(\sigma_E); \quad c \sim U_{(-Bc, +Bc)}, \quad T_j \sim N(0, 1), \quad \sigma_E \sim U_{(0, +BE)}$$

とおくと、事後分布は

$$\begin{aligned} P(c, T_j, \sigma_E \mid Y_{11}, \dots, Y_{NM}) &\propto P(Y_{11}, \dots, Y_{NM} \mid c, T_j, \sigma_E) P_0(c, T_j, \sigma_E) \\ &= P_0(c, T_j, \sigma_E) \prod_{j=1}^M \prod_{i=1}^N P(Y_{ij} \mid c, T_j, \sigma_E) \\ &= P_0(c, T_j, \sigma_E) \prod_{j=1}^M \prod_{i=1}^N \Phi(-(c - T_j)/\sigma_E)^{Y_{ij}} \Phi((c - T_j)/\sigma_E)^{1-Y_{ij}} \end{aligned}$$

で与えられる。また、信頼性係数 ρ は

$$\rho = \sigma_T^2 / (\sigma_T^2 + \sigma_E^2 / N) = 1 / (1 + \sigma_E^2 / N)$$

で与えられる。

事後分布の推定を次の MCMC により行う。

6 . アルゴリズム - 2

2 値データ用にアルゴリズム - 1 を変更した以下のアルゴリズムで MCMC を行う。

A. MCMC の準備

$c^{(0)}$ 、 $T_j^{(0)}$ 、 $\sigma_E^{(0)}$ に初期値を設定し、 $\rho^{(0)} = 1 / \{1 + (\sigma_E^{(0)})^2 / N\}$ とおく。Proposal

distribution のパラメタ σ_c 、 σ_{Tj} 、 σ_{σ_E} の値を設定して、 $s \leftarrow 0$ とおく。

$$B. \quad y \sim N(\sigma_E^{(s)}, \sigma_{\sigma_E}^2), \quad \sigma_E^{(s+1)} = \begin{cases} y & \text{with probability } \alpha_{\sigma_E s} \\ \sigma_E^{(s)} & \text{otherwise} \end{cases}$$

$$C. \quad y \sim N(c^{(s)}, \sigma_c^2), \quad c^{(s+1)} = \begin{cases} y & \text{with probability } \alpha_{cs} \\ c^{(s)} & \text{otherwise} \end{cases}$$

D. For ($j = 1$ to M)

$$y \sim N(T_j^{(s)}, \sigma_{Tj}^2), \quad T_j^{(s+1)} = \begin{cases} y & \text{with probability } \alpha_{Tj s} \\ T_j^{(s)} & \text{otherwise} \end{cases}$$

$$E. \quad \rho^{(s+1)} = 1 / \left(1 + \left(\sigma_E^{(s+1)} \right)^2 / N \right)$$

F. もし $s + 1 < N\text{Simu}$ ならば、 $s \leftarrow s + 1$ と設定して、ステップ B に戻る。

7 . 分析例 2

2 値データ Y_{ij} に対するアルゴリズム 2 による信頼性係数の推定の可能性を、真の信頼性係数の値がわかっている人工データによって確認する。まず、分析例 1 におけるデータ X_{ij} を

$$Y_{ij} = \begin{cases} 1 & \text{if } X_{ij} > 50 \\ 0 & \text{otherwise} \end{cases}$$

と 2 値化した。この Y_{ij} に対して 係数を算出すると

$$\alpha \approx 0.752$$

となった。アルゴリズム 2 によって信頼性係数 (事後分布の平均値) と 95%PHD 区間を求めると、

$$\hat{\rho} = 0.836, \quad 95\%HPD \text{ 区間} = (0.8251, 0.8488)$$

を得た。真値 $\rho \approx 0.833$ に対してよい推定であるといえる。

上の分析で用いたデータセットとは別に、同じパラメタでデータ X_{ij} を作成した。このデータセットに対する 係数は

$$\alpha \approx 0.831$$

であった。2 値化を上の場合とは異なる基準値による次式で行った。

$$Y_{ij} = \begin{cases} 1 & \text{if } X_{ij} > 33.56 \\ 0 & \text{otherwise} \end{cases}$$

このとき、 係数は

$$\alpha \approx 0.6952$$

となった。基準値を 50 とした場合の 2 値データより 係数はより小さい値となっている。アルゴリズム 2 によって信頼性係数 (事後分布の平均値) と 95%PHD 区間を求めると、

$$\hat{\rho} = 0.821, \quad 95\%HPD \text{ 区間} = (0.8070, 0.8350)$$

を得た。真値 $\rho \approx 0.833$ に対してよい推定であるといえる。

8 . まとめ

分析例 1 および分析例 2 の成功は、データ生成に適合する確率モデルによりベイズ的方法で信頼性係数の推定を行うことの可能性を示唆するものである。ベイズ的方法は事後分布から 95%HPD 区間などを算出することができるので、点推定値のみを算出する 係数より豊かな情報を提供する推定法であり、今後さらにこの方法を開発していくことの必要性が確認されたといえる。

引用文献

Crocker, L. and Algina, J. (1986). Introduction to classical and modern test theory. Wadsworth.