

順序尺度の確率分布における情報量を活用した GPA に代わる評価指標の提案

山崎 綾一郎† 吉永 悠人§ 山岸 祐己†‡¶ 谷川 邦† 足立 智子†
高林 貴仁¶

† 静岡理科大学 § 遠州鉄道株式会社 ‡ 浜松医科大学 ¶ 株式会社良品計画

1 はじめに

GPA は教員による段階評価 (S, A, B...) をもとに算出されるが, これらの評価は教員の主観に依存するため, 客観性に欠けるという課題がある. また, 段階評価は本来順序尺度であるにもかかわらず, GPA 算出時には間隔尺度として扱われており, 統計的整合性にも問題がある. また, GPA は 1 年次と卒業時の成績に強い相関があることが示されており, 初期の成績順位構造を長期にわたり固定化させる指標であると指摘されている [1]. 主観の評価を客観的评价へと変換するとして, 項目反応理論 (IRT) [2] が知られているが, 最低 300 程度のサンプルサイズやハイパーパラメータの設定が必要であり, 実用面での導入は容易ではない.

そこで本研究では, 順序尺度に基づく確率分布の情報量に着目し, ハイパーパラメータを必要としない新たな成績評価モデルを提案する. 提案手法は, 各講義の成績を累積自己情報量 [3] に変換することで, 評価分布に依存せず共通の尺度で扱うことができ, 基本的に余剰単位は有利に働くよう設計されている.

2 提案手法

複数の講義における成績データセットを

$$\mathcal{D} = \{(c_1, s_{u_1,1}, t_1), \dots, (c_N, s_{u_N,N}, t_N)\} \quad (1)$$

とする. c_n と u_n と $s_{k,n}$ と t_n は, n 番目に確定した成績の, 対象学生 $c_n \in \{1, \dots, i, \dots, I\}$ と, 対象講義 $u_n \in \{1, \dots, k, \dots, K\}$ と, 成績 $s_{k,n} \in \{1, \dots, j, \dots, J\} = \mathcal{J}$ と, 観測時刻 $t_1 \leq \dots \leq t_n \leq \dots \leq t_N$ をそれぞれ表す. ここで, 成績は例えば, 不可・可・良・優・秀や, D・C・B・A・S のような段階評価を考え, それらを順に $\{1, \dots, j, \dots, J\}$ のように置き換えるものとする. このとき, 段階評価の仕方が複数存在する場合は, 講義ごとの \mathcal{J}_k を設定することも可能である. n はタイムステップとし, $N = \{1, 2, \dots, N\}$ をタイムステップ集合とする. 便宜上, $s_{k,n}$ が j をとる場合に 1, そ

れ以外の場合には 0 となるような J 次元のダミー変数 $s_{k,n,j}$ を定義する. このとき, 評価を順序カテゴリとして扱う場合でも, 多項分布として扱う場合でも, ある講義 k が m 番目に評価 j を与える確率 $p_{k,m,j}$ に関する対数尤度関数は以下ようになる.

$$\begin{aligned} \log L(\theta_{k,m}) &= \sum_{\{n \in N | n \leq m\}} \sum_{j \in \mathcal{J}} s_{k,n,j} \log p_{k,m,j} \quad (2) \\ &= \sum_{j=1}^J \left(\sum_{\{n \in N | n \leq m\}} s_{k,n,j} \right) \log p_{k,m,j} \quad (3) \end{aligned}$$

講義 k における m 番目の成績 j の評価が付与される確率 $p_{k,m,j}$ が多項分布に従っていると仮定すると, $p_{k,m,j}$ の最尤推定量は

$$\hat{p}_{k,m,j} = \frac{\sum_{\{n \in N | n \leq m\}} s_{k,n,j}}{\sum_{\{n \in N | n \leq m\}} \sum_{a \in \mathcal{J}} s_{k,n,a}} \quad (4)$$

のように与えられる. さらに, 最尤推定量を用いた講義 k の成績確率分布を

$$\hat{\theta}_{k,m} = \{\hat{p}_{k,m,1}, \dots, \hat{p}_{k,m,j}, \dots, \hat{p}_{k,m,J}\} \quad (5)$$

とすれば, その累積分布関数

$$F(v; \hat{\theta}_{k,m}) = \sum_{\{j \in \mathcal{J} | j \leq v\}} \hat{p}_{k,m,j} \quad (6)$$

を考察することができ, ある学生 i に対して付与された成績 $s_{k,n}$ を, そのときの講義 k の累積自己情報量 $-\log F(s_{k,n}; \hat{\theta}_{k,m})$ として変換することができる. 成績 1 が最低評価を意味し, J が最高評価を意味している場合, 提案情報量 $-\log F(s_{k,n}; \hat{\theta}_{k,m})$ は, m 番目の成績が確定したときの講義 k にとって, 成績 $s_{k,n}$ がいかに珍しく評価が低い (ネガティブ) かを示していることになるため, 以降では式 (6) を $F_{neg}(s_{k,n}; \hat{\theta}_{k,m})$ と表す. それに対し, 成績 j について逆順の累積分布関数

$$F_{pos}(v; \hat{\theta}_{k,m}) = \sum_{\{j \in \mathcal{J} | v \leq j\}} \hat{p}_{k,m,j} \quad (7)$$

を考えれば, 同様に成績 $s_{k,n}$ がいかに珍しく評価が高い (ポジティブ) かを示していることになる. ここで, m 番目に確定した成績までに 1 人の学生にしか成績が付与されていない, もしくは複数人の学生に対し同

Proposal of an Alternative Evaluation Index to GPA Utilizing the Information Content in the Probability Distribution of Ordinal Scales

†Ryoichiro YAMAZAKI §Yuto YOSHINAGA ‡‡¶Yuki YAMAGISHI †Hiroshi TANIGAWA †Tomoko ADACHI ¶Takahito TAKABAYASHI

†Shizuoka Institute of Science and Technology

§Enshu Railway Co., Ltd

‡Hamamatsu University School of Medicine

¶Ryohin Keikaku Co., Ltd.

じ成績 j しか付与していない講義 k の提案情報量は 0 になる。このとき、提案モデルは観測数 N に対して $O(N)$ で全ての学生の時系列データを生成可能であり、データの追加や更新があった場合でも、 $O(N)$ でそれらを再計算することが可能である。

3 評価実験とまとめ

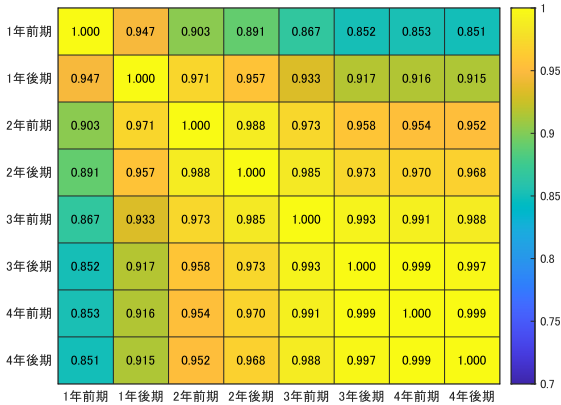
匿名化された学生 $I = 165$ が履修した講義 $K = 186$ における成績 4 年分 $N = 10835$ を評価実験の対象とした。今回は、 w 番目までに学生 i に付与された成績 $N_{i,w} = \{n \in N \mid c_n = i, n \leq w\}$ におけるポジティブな意味での提案情報量の累積値

$$CP(i, w) = \sum_{m \in N_{i,w}} -\log F_{pos}(s_{k,m}; \hat{\theta}_{k,m}) \quad (8)$$

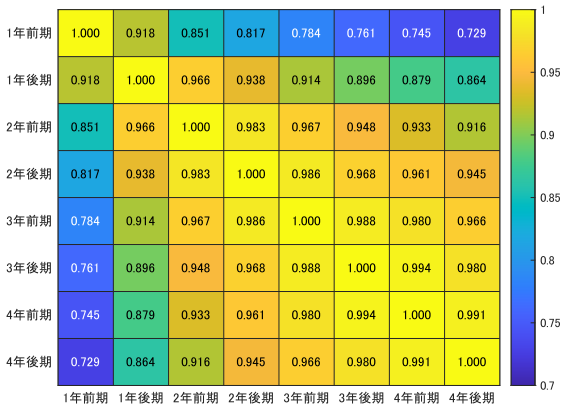
と、ネガティブな意味での提案情報量の累積値

$$CN(i, w) = \sum_{m \in N_{i,w}} -\log F_{neg}(s_{k,m}; \hat{\theta}_{k,m}) \quad (9)$$

を算出し、その最終値 $CP(i, N)$, $CN(i, N)$ を GPA と比較した。本実験での GPA は、1 年次からの通算成績を示す累積 GPA を GPA と表記することとする。



(a) GPA の半期別順位相関

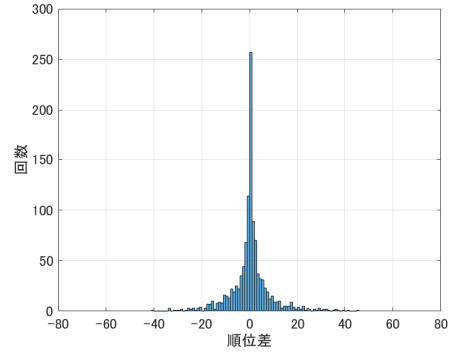


(b) $CP(i, N)$ の半期別順位相関

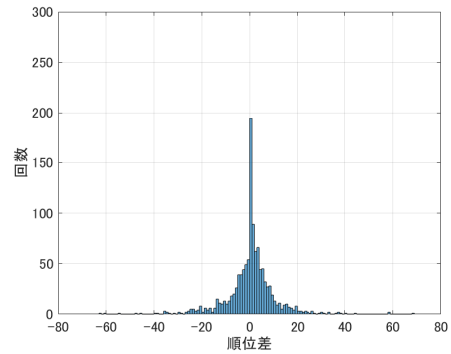
図 1: 半期別順位相関

図 1a から、1 年前期と 4 年後期の順位相関 (kendall) [4] が 0.729 であったのに対し、図 1b では 0.851 であるこ

とがわかる。順位相関自体が大きな差ではないものの、GPA は序盤で形成された成績順位が卒業時まで強く維持される傾向がある一方、 $CP(i, w)$ は順位変動が GPA と比較して許容されやすいことがうかがえる。



(a) GPA の半期毎順位差ヒストグラム



(b) $CP(i, N)$ の半期毎順位差ヒストグラム

図 2: 半期毎順位差ヒストグラム

図 2b から、 $CP(i, w)$ に基づく半期毎の順位変動は、図 2a (GPA) よりも多く含まれている。よって、提案手法は GPA と比較して初期成績に依存しにくく、成績が一時的に低下してもその後の巻き返しが評価に反映されやすい特性を有することが示唆された。

参考文献

- [1] Paul A Westrick, Jessica P Marini, Linda Young, Helen Ng, and Emily J Shaw. The consequences of a low first-year grade point average on later college outcomes. *College Board*, 2023.
- [2] Shenghai Dai, Thao Thu Vo, Olasunkanmi James Kehinde, Haixia He, Yu Xue, Cihan Demir, and Xiaolin Wang. Performance of polytomous irt models with rating scale data: An investigation over sample size, instrument length, and missing data. In *Frontiers in Education*, p. 721963. Frontiers Media SA, 2021.
- [3] C. E. Shannon. A mathematical theory of communication. *The Bell System Technical Journal*, Vol. 27, No. 3, pp. 379–423, 1948.
- [4] M. G. Kendall. *Rank Correlation Methods*. Charles Griffin, London, 1970.