

## 自己情報量に基づくレビュー点数の客観的な変換

照屋 良直 † 吉永 悠斗 † 山岸 祐己 † 和泉 舞 ‡ 高林 貴仁 ‡

† 静岡理科大学 情報学部 ‡ 株式会社良品計画 IT サービス部

## 1 はじめに

オンラインレビューサイトにおける諸問題としては、レビュー平均点の信頼性があげられる。確かに、平均点は単純かつ直感的に理解しやすい値ではあるが、各ユーザの主観的な評価に依るところが大きく、評価が甘いユーザと辛いユーザとでは、同一の点数における意味合いが異なることが推察される。よって、各ユーザが投稿した点数分布を考慮し、その点数が持つポジティブな意味合いとネガティブな意味合いを指標化することは重要であると言える。本手法は、回顧的 (retrospective) な枠組みによる時系列データ分析 [1] に類似した考え方で、各ユーザの点数確率分布における自己情報量 [2] を利用し、レビュー点数を客観的なポジティブスコアとネガティブスコアに変換することを目指す。

## 2 提案手法

レビューデータセットを

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

とする。  $c_n$  と  $u_n$  と  $s_n$  と  $t_n$  は、  $n$  番目に投稿されたレビューの、アイテムもしくはカテゴリ  $c_n \in \{1, \dots, i, \dots, I\}$  と、ユーザ  $u_n \in \{1, \dots, k, \dots, K\}$  と、レビュー点数  $s_{k,n} \in \{1, \dots, j, \dots, J\} = \mathcal{J}$  と、観測時刻  $t_1 \leq \dots \leq t_n \leq \dots \leq t_N$  をそれぞれ表す。  $n$  はタイムステップとし、  $\mathcal{N} = \{1, 2, \dots, N\}$  をタイムステップ集合とする。便宜上、  $s_{k,n}$  は  $J$ -次元ベクトルダミー変数として

$$s_{k,n,j} = \begin{cases} 1 & \text{if } s_{k,n} = j; \\ 0 & \text{otherwise.} \end{cases} \quad (2)$$

のように変換する。このとき、ユーザ  $k$  によってレビュー点数  $j$  が投稿される確率  $p_{k,j}$  が多項分布に従っていると仮定すると、  $p_{k,j}$  の最尤推定量は

$$\hat{p}_{k,j} = \frac{\sum_{n \in \mathcal{N}} s_{k,n,j}}{\sum_{n \in \mathcal{N}} \sum_{a \in \mathcal{J}} s_{k,n,a}} \quad (3)$$

のように与えられる。さらに、最尤推定量を用いたユーザ  $k$  の投稿点数確率分布を

$$\hat{\theta}_k = \{\hat{p}_{k,1}, \dots, \hat{p}_{k,j}, \dots, \hat{p}_{k,J}\} \quad (4)$$

とすれば、その累積分布関数

$$F(v; \hat{\theta}_k) = \sum_{\{j \in \mathcal{J} \mid j \leq v\}} \hat{p}_{k,j} \quad (5)$$

を考慮することができ、あるアイテムもしくはあるカテゴリ  $i$  に対して投稿されたレビュー点数  $s_{k,n}$  を、ユーザ  $k$  の自己情報量  $-\log F(s_{k,n}; \hat{\theta}_k)$  として変換することができる。レビュー点数  $1$  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  $-\log F(s_{k,n}; \hat{\theta}_k)$  は、ユーザ  $k$  にとって、点数  $s_{k,n}$  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 (5) を  $F_{neg}(s_{k,n}; \hat{\theta}_k)$  と表す。それに対し、点数  $j$  について逆順の累積分布関数

$$F_{pos}(v; \hat{\theta}_k) = \sum_{\{j \in \mathcal{J} \mid v \leq j\}} \hat{p}_{k,j} \quad (6)$$

を考えれば、同様に点数  $s_{k,n}$  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。また、提案情報量は、その加法性により、時系列データとして累積し続けることも可能であると考えられる。

## 3 評価実験とまとめ

無印良品のネットストア\*におけるレビューデータセット  $N = 159458$  を評価実験の対象とし、アイテムランキングを生成した。今回は、アイテム  $i$  に投稿されたレビュー  $N_i = \{n \in \mathcal{N} \mid c_n = i\}$  におけるポジティブな意味での提案情報量の累積値

$$CP(i) = \sum_{m \in N_i} -\log F_{pos}(s_{k,m}; \hat{\theta}_k) \quad (7)$$

と、ネガティブな意味での提案情報量の累積値

$$CN(i) = \sum_{m \in N_i} -\log F_{neg}(s_{k,m}; \hat{\theta}_k) \quad (8)$$

を算出し、それらの差分  $CP(i) - CN(i)$  を提案評価値とした。また、比較として、レビュー平均点を投稿数で標準化した z-score [3] もアイテム  $i$  ごとに算出した。図 1 より、それぞれの評価値は、とくに上位と下位の部分で差異があるため、今回は各評価値における上位 10 アイテムと下位 10 アイテムを比較する。

表 1, 2 より、1 位はともに「シリコンジャムスプーン」となったが、z-score で 6 位の「シリコン調理スプーン」が提案評価値において 2 位となった。「シリコン調理スプーン」の平均点 (mean) は 4.299 と、それぞ

\*<https://www.muji.com/jp/ja/store>

Objective Conversion of Online Review Score by Self-Information  
 †Yoshinuo TERUYA †Yuto YOSHINAGA †Yuki YAMAGISHI  
 ‡Mai IZUMI ‡Takahito TAKABAYASHI  
 †Shizuoka Institute of Science and Technology  
 ‡Ryohin Keikaku Co., Ltd.

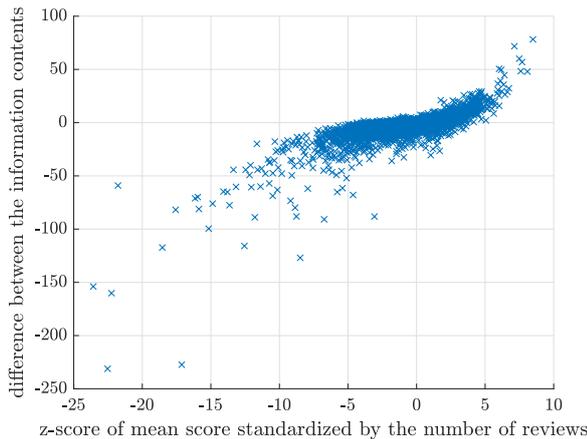


図 1: 標準化 z-score と提案評価値の比較

れの評価値における上位 10 アイテムの中では最も低くなっているが、各ユーザの満足度は高いことがうかがえる。また「洗顔用泡立てネット」、「大袋 塩チョコスティックパイ」、「ぶどうのクッキー」は、提案評価値においてそれぞれ 5 位、6 位、10 位となったが、z-score の上位 10 アイテムには出現していないため、これらのアイテムは、平均点や投稿数といった見かけ上の評価には表れないポジティブな要因を持っていることが示唆される。

表 1: 標準化 z-score 上位 10 アイテム

rank	item	reviews	mean	z-score
1	シリコンジャムスプーン	215	4.712	8.487
2	豆乳ビスケット	139	4.856	8.086
3	ステンレス マッシャー	137	4.818	7.693
4	ホホバオイル	292	4.531	7.601
5	素材を生かしたスナック ごぼう	200	4.645	7.487
6	シリコン調理スプーン	683	4.299	7.129
7	てんさい糖ビスケット	154	4.662	6.729
8	マイルドジェルクレンジング (大容量)	264	4.481	6.628
9	植林木ペーパー裏うつりにくいノート 5 冊組	120	4.725	6.449
10	携帯用メガネ拭き	138	4.667	6.408

表 2: 提案評価値上位 10 アイテム

rank	item	reviews	mean	proposed
1	シリコンジャムスプーン	215	4.712	78.21
2	シリコン調理スプーン	683	4.299	71.80
3	素材を生かしたスナック ごぼう	200	4.645	60.10
4	ステンレス マッシャー	137	4.818	56.88
5	洗顔用泡立てネット	191	4.518	50.53
6	大袋 塩チョコスティックパイ	155	4.600	49.66
7	ホホバオイル	292	4.531	48.35
8	豆乳ビスケット	139	4.856	47.98
9	携帯用メガネ拭き	138	4.667	44.70
10	ぶどうのクッキー	101	4.752	39.32

表 3, 4 より、「掃除用品システム・フローリングモップ」、「アクリル冷水筒」、「フタが外せる P E T 詰替ボトル」などは似たような順位になったが、z-score で 7 位の「綿であったかUネック八分袖 T シャツ」が提案評価値において 2 位となった。「綿であったかUネック八分袖 T シャツ」の平均点 (mean) は 2.915 と、z-score

の下位 10 アイテムの中では最も高くなっているが、各ユーザの満足度は低いことがうかがえる。また、提案評価値の下位 10 アイテムで出現している「繰り返し使える 2 枚組・三層マスク」と「素材を生かしたカレー バターチキン」の平均点は 3.364 と 3.397 であり、それぞれの評価値における下位 10 アイテムの中ではかなり高くなっている。すなわち、これら 2 アイテムは、平均点や投稿数といった見かけ上の評価には表れないネガティブな要因を持っていることが示唆される。

表 3: 標準化 z-score 下位 10 アイテム

rank	item	reviews	mean	z-score
1	掃除用品システム・フローリングモップ	196	1.658	-23.57
2	アクリル冷水筒	469	2.527	-22.53
3	フタが外せる P E T 詰替ボトル	269	2.100	-22.24
4	掃除用品システム・バス用替えスポンジ	161	1.615	-21.77
5	かかと滑り止め付き 綿混 5 本指フットカバー	169	2.006	-18.54
6	スプレーヘッド	163	2.074	-17.57
7	綿であったかUネック八分袖 T シャツ	518	2.915	-17.13
8	エイジングケア化粧水 高保湿タイプ (大容量)	90	1.633	-16.15
9	掃除用品システム アルミ伸縮式ボール	117	1.940	-15.95
10	掃除用品システム・バス用スポンジ	121	1.983	-15.87

表 4: 提案評価値下位 10 アイテム

rank	item	reviews	mean	proposed
1	アクリル冷水筒	469	2.527	-231.19
2	綿であったかUネック八分袖 T シャツ	518	2.915	-227.36
3	フタが外せる P E T 詰替ボトル	269	2.100	-160.16
4	掃除用品システム・フローリングモップ	196	1.658	-153.88
5	繰り返し使える 2 枚組・三層マスク	409	3.364	-126.96
6	かかと滑り止め付き 綿混 5 本指フットカバー	169	2.006	-117.33
7	アルミ洗濯用ハンガー	213	2.770	-115.85
8	掃除用品システム・カーペットクリーナー用替えテープ	201	2.488	-99.57
9	素材を生かしたカレー バターチキン	290	3.397	-90.85
10	シルク入りカップ入りタンクトップ	200	2.805	-88.914

### 参考文献

- [1] Rebecca Killick, Paul Fearnhead, and I.A. Eckley. Optimal detection of changepoints with a linear computational cost. *Journal of the American Statistical Association*, Vol. 107, pp. 1590–1598, 12 2012.
- [2] C. E. Shannon. A mathematical theory of communication. *The Bell System Technical Journal*, Vol. 27, No. 3, pp. 379–423, 1948.
- [3] Douglas G Altman and J Martin Bland. Standard deviations and standard errors. *BMJ*, Vol. 331, No. 7521, p. 903, 2005.