

標準スコア投票方式におけるZスコア化後の処理について A proposal of processing methods after Z-scoring in the standard scoring method

城 真範[†], 浅野 健一郎^{‡,†}

Masanori Shiro, Kenichiro Asano

[†] (国) 産業技術総合研究所, [‡] (一社) 社会的健康戦略研究所

AIST, Social Health Strategy Research Institute

shiro@aist.go.jp

概要

複数の選択肢に対して回答者が順序を伴った評価を与えるアンケート・投票等について、適切な選択を行うために、その評価の正確な理解と分析のための方法を、回答者による入力と結果の分析の両面で提案する。本提案手法では、回答者はネガティブな評価も柔軟に入力することが可能であり、標準スコア投票方式の改良として、各回答をZスコアに直したあと、外れ値の影響を低減するための中央値の採用や、偏りを示すエントロピーやカルバック・ライブラー情報量を利用する。

キーワード：アンケート, 標準スコア投票方式, エントロピー, カルバック・ライブラー情報量

1. はじめに

1.1 問題意識

多数の人の判断・評価を一定程度合理的に集計し、何らかの選択を行う手段として、選択肢を提示してスコアを付けることが行われる。投票、能力評価、何らかの施策に対する満足度の評価、意思決定のための採決など、その利用場面は多岐である。本報告では複数の選択肢に対して回答者が順序を伴った評価を与えるアンケート・投票等について、適切な選択を行うために、その評価の正確な理解と分析のための提案をしたい。

ここで我々は回答方法と統計処理において、二つの方向を目指す。一つは選択肢に対する入力を可能な限り幅広く、かつ簡便にできるようデザインすること。もう一つは、結果を多面的に考察できるような枠組みを示すことである。

前者において問題と考える現状は得にネガティブな意思表示の反映がしにくい点である。ここでネガティブという意味は、ある選択肢を好意的なものとして評価することの逆で、ある選択肢を非好意的なものとし

て能動的に評価することを意味する。例えば投票においては通常、最も良いと思う単一の候補者に票をいれるが、有権者の中には、ある候補者以外は誰でも構わないと考える人もいるだろう。そうした意思是、単一の候補者しか選択できない現状の投票システムが正確に汲み取れない。また、一般的に心理学的なアンケートではポジティブな選択を求める場合が多い。例えば「次の中で好きなものを好きな順に選択せよ」といった設問は多いが、逆に「次の中で嫌いなものを嫌いな順に選択せよ」という設問は通常見かけない。しかし実際のところ、人の行動がポジティブな動機よりネガティブな動機に基づくという報告もある [1]。従って心理学的なアンケートにおいてもポジティブな評価のみならず、ネガティブな評価も簡便に示せる仕組みがあるべきだろう。

この問題は順位評価を入力できないという問題に一般化できる。社会政策のアンケートや政治的投票において死票を減らし、民意をより正確に反映するならば、選択肢（候補者）の全てに順位付けをし、順位の何らかの合理的な集計によって当選者を選ぶべきだろう。アンケートについても同様の論が可能である。問題はその合理的な集計とはどのような方法かという点に曖昧さが残ることである。

1.2 標準スコア投票方式とその問題点

民意を正確に反映できる投票方式は民主主義制度の根幹をなすため、その方法は古くから議論されてきた。選挙における実現はされていないものの、既にいくつかの有力な投票方式が提案されている。その中でも古くから広く知られているものは標準スコア投票方式 [2] と呼ばれる。これはまず有権者に負数を含む任意の実数値ですべての候補者に評点をつけてもらう。ネガティブな選択は負数を入力することで実現できる。ポジティブやネガティブな評価をしたくない場

合も、無回答を許容し、その値をゼロと仮定することで回答者の負荷を減らすことができ、すべての選択肢についての入力を強要しなくて良い。またどのような大きさの数でも自由に入力可能とし、回答値のスケール感において回答者自身の規格化処理を必要としない。そして、集計の際には入力されたスコアを Z スコア [3] に変換して単純加算することで集計する。投票において「この候補者だけには当選してほしい」といったネガティブな意思表示を可能にする。標準スコア投票方式は人が選択的な意思表示をする場合において一般に利用可能で、心理学のアンケートでも利用されることがある [4]。投票においては候補者間、有権者間で完全に平等・公平であることが求められるので、Z スコアに変換してから単純加算することで優劣を決することは適切である。

しかし心理学的アンケートや社会政策のアンケートでは、完全な公平性の実現よりも、結果の特徴に応じた機動的な選択をしたい場合がある。例えば単純なスコアの大きさよりも、特定の支持者ばかりが支持している選択肢、つまり得られたスコアが大きく偏っている選択肢の除外を優先したいことがある。また逆にそうした偏ったスコアを得た選択肢だけを優先的に取り出したいこともある。得点そのものより、選択肢の得点順位に応じて資源配分を行う場合もある。そこで標準スコア投票方式において最後に Z スコアを加算するという部分を拡張する必要が生じる。なお、我々は最終的に何らかの選択を行う手段としての投票・アンケートを想定しているので、そもそも得点の分布を何らかの特徴的数値で代表させない分析は除外するものとする。

標準スコア投票方式における Z スコアの単純な加算は、統計的にはスコアの平均値 (Mean) を求めることに等しい。つまりある候補者についての Z スコアの統計分布を、その平均値で代表させていると解釈できる。ところが平均値による分布の代表には「外れ値」による影響が古くから指摘される [5]。このための対応も様々な提案されているが、簡単な方法は最頻値 (Median) による代表である (順位に直してから平均値を算出する方法も有効であるとされる)。

一方でばらつきの尺度は一般的には標準偏差によって与えられる。標準偏差は 0 以上の値をとるが、Z スコアを利用した場合、標準偏差はスコアの次元を持ち、選択肢の数に依存して上限値が変わる。このため例えば質問セットの中で回答に応じて枝間が変わる場合などは分析を煩雑にする。Z スコア化によって回答者一人あたりの回答を規格化し、選択肢の構造によら

ない分析の方法が必要である。

2. 提案方法

2.1 平均と分散の代替

本報告では、標準スコア投票方式の改良として、Z スコア化のあとに平均値を取るのではなく最頻値を取る方法を提案する。また、並行して標準偏差以外で分布の不確実性 (ランダム性) の大きさを定量化する方法として、エントロピー (Entropy) とカルバック・ライブラー情報量 [6] (Kullback-Leibler divergence。以下 KL) の利用を提案する。エントロピーの値が大きな場合、ある選択肢 (候補者) に対する評価が割れていることを意味する。また、KL は何らかの基準分布からの擬似的な距離を示すもので、本報告では一様分布からの KL (KL based uniform) と、回答結果の全 Z スコアの分布を基準とした KL (KL based owndata) の二つを考えたい。以下にエントロピーと KL の定義を示す。

- 分布 P に対するエントロピー

$$H(P) = - \sum_{i=1}^n p_i \log p_i$$

- 2 つの確率分布 $P = \{p_1, p_2, \dots, p_n\}$ および $Q = \{q_1, q_2, \dots, q_n\}$ に対する KL :

$$D_{KL}(P \parallel Q) = \sum_{i=1}^n p_i \log \frac{p_i}{q_i}$$

この定義からわかることは、エントロピーはデルタ関数 [7] で表現される分布からの KL に対応することである。

2.2 架空例による実験

本報告の確認として人工的な例を示す。スコアを平均値で代表させにくい例として二峰性をもつ分布を作成する。具体的には選択肢の数を K とした場合、各回答を 2 つの正弦関数と 2 つの余弦関数の和として、各周期関数のパラメータを次のように乱数を使って作成した。

- 周期の範囲: $[5, K/2]$ の一様乱数
- 位相の範囲: $[0, 2\pi]$ の一様乱数
- 振幅の範囲: $[0.5, 2.0]$ の一様乱数

回答者数は 10,000 とし、選択肢 (Choice) K を 7 とし、回答者毎に Z スコア化された 10,000 個の回答セットを生成した。

3. 結果

結果を示す。まず図 1 は 7 つの選択肢 (Choice) が得た Z スコアの分布である。選択肢には便宜的に 1 から 7 の番号を振ってある。それぞれの選択肢において二峰性のサンプルが作成できていることを確認した。

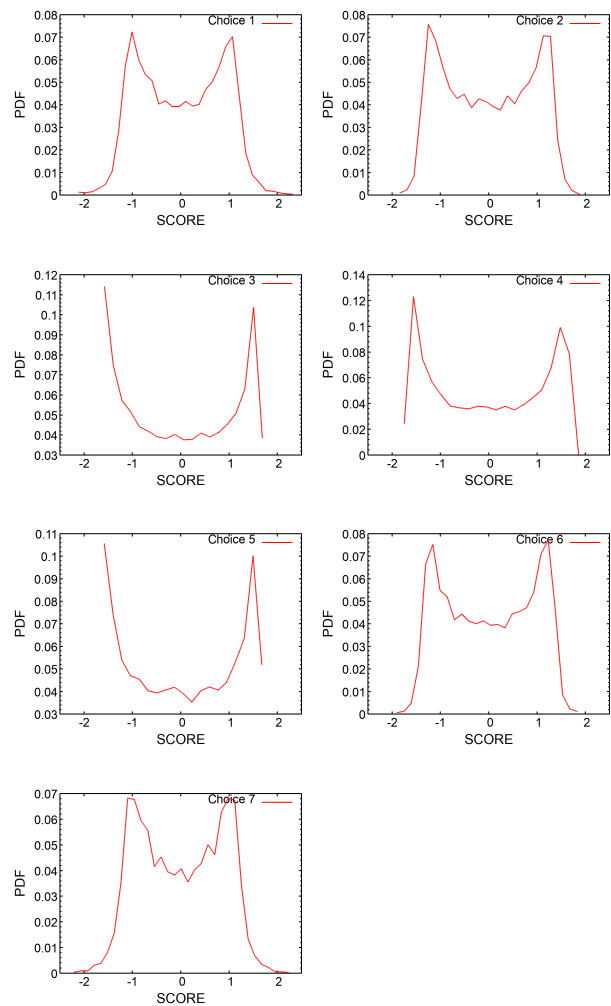


図 1：各候補が得たスコア分布（候補番号順）

次に、これらの分布をもとにした平均と最頻値、および、エントロピー、2 つの KL を示す。図 2 が平均と最頻値、図 3 がエントロピーと 2 つの KL である。どちらも横軸は選択肢 (Choice) である。

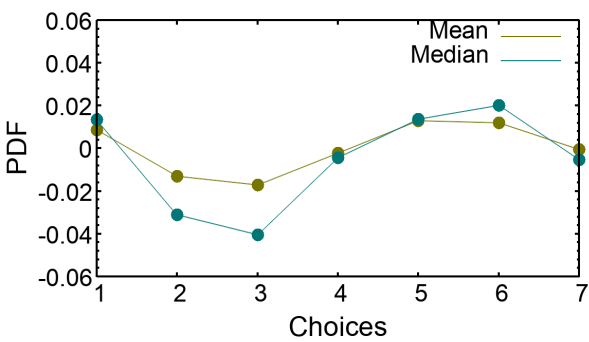


図 2：平均と最頻値

図 2 では選択肢 5 と 6 に置いて平均と最頻値の違いで順位の入れ替えが起こっている。このように単峰性でない分布においては分布の振る舞いを平均で代表させることが必ずしも適切でない場合が存在する。

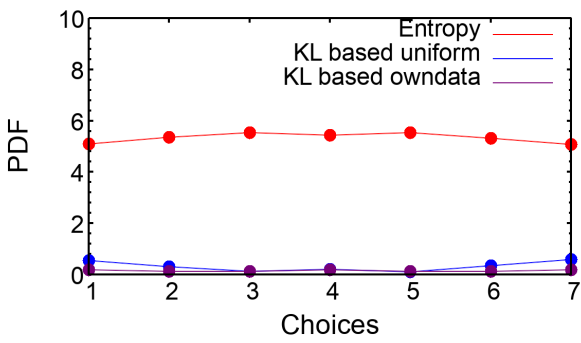


図 3：エントロピーと 2 つの KL

図 3 ではエントロピーの値が大きく、KL の値が小さいが、これはサンプルの性質によるものである。

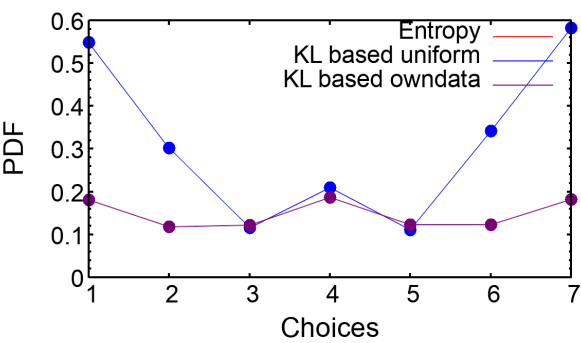


図 4：図 3 の拡大 (2 つの KL)

図 4 は図 3 の下側を拡大したものである。KL based owndata は 10000×7 個のサンプル全体の分布を基準にした KL である。一様分布やデルタ関数分布を基準にした KL に比べれば変化が少なくなっている。回収された全回答内に対して、ある選択肢がもつ分布の相対

的な偏りを示したものである。

4. 考察

KL は統計分布間の擬距離を意味する。デルタ関数で表現される分布（エントロピーゼロ）からの KL がエントロピーであり、一様分布のエントロピーが最大である。選択肢が K 個ある離散一様分布のエントロピーは $\log_2 K$ である。したがって選択肢の数 K が増えても KL を $\log_2 K$ で規格化することで分布のばらつきを比較することができるようになる。KL は標準偏差と異なりスコアの次元を持たないので、回答者の選択によって枝間が変化する場合や、追加の設問が加わる場合にも柔軟な分析を可能にする。なおエントロピーと一様分布を基準にした KL を、みやすさのため Z スコアで規格化して示したのが図 5 である。規格化によって 2 つの分布が横軸を中心に対称になっていることがわかる。

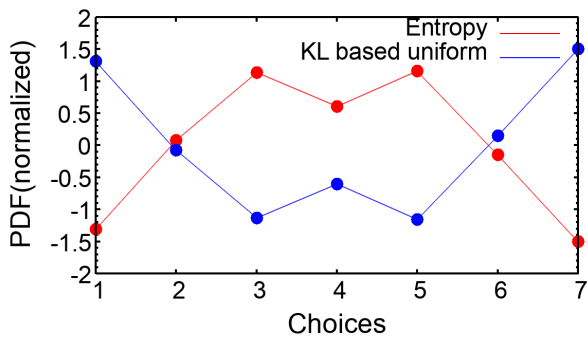


図 5：エントロピーと一様分布を基準にした KL

本報告では主に標準スコア投票方式の解釈に焦点を当て、極端な分布であれば外れ値が多ければ平均と最頻値に差が出ることに、無次元化されたばらつきによる分析を提案した。また実際のアンケートでは設問毎に選択肢数が異なり、条件づけされた複雑な枝間も存在する。こうした場合に、分布のばらつきを無次元化することで、アンケート・投票の結果を深く汲み取れる可能性があることを示した。それは例えば、マイノリティの意見を作為なく抽出したり、逆にマジョリティを優先する政策の根拠としても使えるだろう。

今後の課題として重要なものは無効回答についてである。標準スコア投票方式は回答者の入力に高い自由度を与える点で回答者の意思をよく汲み取ることができるが、回答者が全選択肢に完全に同じスコアを与えたり、全選択肢を白紙回答した場合は標準偏差の計算ができなくなり Z スコア化ができなくなる。こうした場合、その回答は無効とせざる得ない。無効回答の扱いを具体的にどうするかは今後の課題である。

5. 終わりに

複数の選択肢に対して回答者が順序を伴った評価を与えるアンケート・投票等について、適切な選択を行うために、その評価の正確な理解と分析のための方法を、回答者による入力と結果の分析の両面で提案した。本提案手法では、回答者はネガティブな評価も柔軟に入力することが可能であり、標準スコア投票方式の改良として、各回答を Z スコアに直したあと、外れ値の影響を低減するための中央値の採用や、偏りを示すエントロピーやカルバック・ライブラー情報量の利用を提案した。本研究においては産業技術総合研究所の平山淳一郎氏に有益なアドバイスを頂いた。

参考文献

- [1] Roy F Baumeister, Ellen Bratslavsky, Catrin Finkenauer, and Kathleen D Vohs. Bad is stronger than good. *Review of general psychology*, Vol. 5, No. 4, pp. 323–370, 2001.
- [2] Samuel Merrill III. *Making Multicandidate Elections More Democratic*. Princeton University Press, 1988.
- [3] Ronald Aylmer Fisher. Statistical methods for research workers. In *Breakthroughs in statistics: Methodology and distribution*, pp. 66–70. Springer, 1970.
- [4] Niels Mahrholz and Pascal Belin. Perceptual dimensions of voice judgments: Trustworthiness, attractiveness, and dominance. *Journal of Nonverbal Behavior*, Vol. 46, No. 4, pp. 501–522, 2022.
- [5] John Tukey. Bias and confidence in not quite large samples. *Ann. Math. Statist.*, Vol. 29, p. 614, 1958.
- [6] Solomon Kullback and Richard A Leibler. On information and sufficiency. *The annals of mathematical statistics*, Vol. 22, No. 1, pp. 79–86, 1951.
- [7] Paul Adrien Maurice Dirac. The physical interpretation of the quantum dynamics. *Proceedings of the Royal Society of London. Series A, Containing Papers of a Mathematical and Physical Character*, Vol. 113, No. 765, pp. 621–641, 1927.