

投票モデルの統計物理

SP-07120 神田 朋彦ヘンリー（非線形物理学講座）

1 序論

個人が意思決定をする際に他人の意見を参考にしているという事がある。しかし、その意見にどれほど影響を受けているかは人それぞれである。本研究では、投票実験を行い、得られた結果から人が他人の意見にどれほどの影響を受けるものなのかを観測してモデル化し、その統計的な性質を理解する事を目的とした。

2 実験内容

今回、我々は二択の投票問題の実験を実施した。詳しい内容は石澤の要旨を確認のこと。主な違いは参照人数のコントロールである。まず、過去の回答を参照しない上で回答を聞き、参照人数が1人の時は過去1人の答えを、参照人数が2人の時は過去2人の答えを二択の票数で教えていく。参照人数を1人、2人、3人、5人、7人、9人と「過去全て」と変えていきながら、再度回答を求めた。その際、解答を変えても良いものとした。より正答率を上げてもらうため、最終的に全ての条件のトータルでの正答率が高い被験者には報酬を払うことを周知した。以下参照人数を k と書くとする。

3 投票モデル

参照人数が0人の正答率は約60%だったので、約80%のクイズの解答を知らない人達と約20%の正解を知っている人達が存在すると推定できる。この正解を知らない人達が他の回答者の回答にどう影響されるのを見るために参照人数 k 人での正答率と被験者の正答率の関係性を調べた。

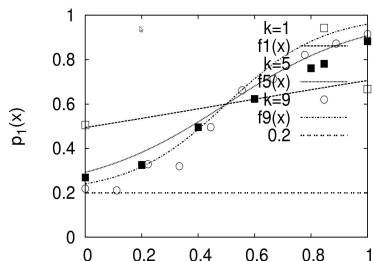


図1: 参照人数と正解者の割合による正答率の関係

$$f_k(x) = p_k(kx) \text{ とした。}$$

縦軸が正答率、横軸が参照人数の中の正解者の割合である。次のモデルがこの関係をよく記述することがわかった。

$$p_k(n) = (1-r) \cdot p + r \cdot \frac{1}{2} \cdot (\tanh(\lambda \cdot (\frac{n-k}{k+z} - \frac{1}{2})) + 1) \quad (1)$$

$(1-r)$ は情報を持ち、自らの情報で回答する人達の割合、 p はその正答率、 r は情報を持たず他の人の回答を見て自らの解答を決める人達の割合、 n は k 人のうち正解した人数を表している。 $(-\sim)$ と $z (0 \sim)$ は過去の回答を信用度を表す値である。 λ の値が高いほど、 z の値が低いほど過去の回答を信用する。 $p=1$ 、 $r=0.8$ とし、 λ 、 z は最小二乗法により、フィットさせて求めると、 $\lambda=6.5$ 、 $z=11$ となった。その場合も図1に示してある。

次にこの $p_k(n)$ に従って次々と投票するモデルを考える。すると、時間 t において正解の方に n 票はいつている確率は次のマスター方程式に従うことがわかる。

$$P(n, t) = (1-p_{t-1}(n)) \cdot P(n, t-1) + p_{t-1}(n-1) \cdot P(n-1, t-1)$$

第一項は時間 $t-1$ における n 票ある状態で t 番目の回答者が間違えることからくる寄与、第二項は時刻 t において $(n-1)$ 人正解の時 t

番目の回答者が正解することからくる寄与である。この時の定常状態を調べるため、平均場近似を用いる。定常状態の正答率を x とすると、 x は次の自己無撞着方程式を満たす。

$$x = (1-r) \cdot p + r \cdot \frac{1}{2} \cdot (\tanh(\lambda \cdot (x - \frac{1}{2})) + 1) \quad (2)$$

式(2)が持つ解の数から以下の相図を求めた。解が1個の時、無知な人たちは正解に投票し、ほぼ $(1-r)p+r$ に収束する。解が3個の時、無知な人たちの正答率は真ん中の不安定な解から上にある場合は正解に投票し、ほぼ $(1-r)p+r$ に収束する。下にある場合は不正解のほうに投票し、ほぼ $(1-r)p$ に収束する。

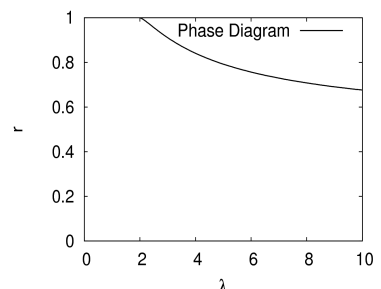


図2: 投票モデルの相図

図2の実線は解が一つの領域と3つの領域の境界を示している。境界の下領域では正答率の分布は1つのピークを持つ。相境界を越えてグラフの上側の領域に出たとき、正答率の分布に2つのピークができる。すなわち相転移が起こっているということになる。2を越えると相転移が起き、 x が上がるにつれて r の割合が低くても相転移が起こることがわかる。ここで、実験から得られた x と z の値を代入して、多数決が間違える確率 r を計算すると以下になった。(図3参照)

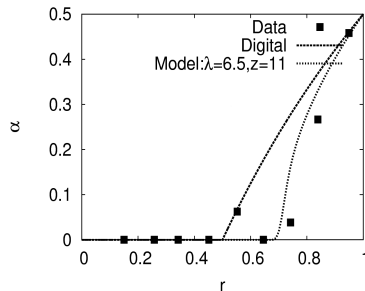


図3: 2つのモデルの r の値

r の値が0.5の時から急激に上に伸びているのがデジタルモデル(辻の要旨を参照)、0.7あたりから伸び始めているのが今回のモデルで、四角が実験結果である。グラフから、人がどのように投票するのはデジタルモデルよりも、式(1)の確率で投票する今回のモデルの方が一致していると言える。

4 今後の課題

今回の実験では被験者は31人の2グループであり、人数があまり多くない。そのため揺らぎにより r の推定がアバウトになってしまっている。今後は被験者の数を増やして、尚且つ r の値をインタビューにより確定した実験を行う必要がある。