

# Coefficient of unalikeability を使用した

## Unordered categorical 実験データの解析と示唆

1190491 竹下 知希

高知工科大学 経済・マネジメント学群

### 1. 概要

統計学では Ordered categorical variable data の分析に重きが置かれ、標準偏差に代表される様にそのバラつきを測る手法も確立されている。一方で、大小によって順序を決められない Unordered categorical variable data のバラつきを測る分析手法は、未だ十分に開発・確立されていない。本研究は Unordered categorical variable data のバラツキの測度に関する文献調査をし、幾つかの測度の中でも特に有用性が高いと言われている Coefficient of unalikeability を見出し、その測度を用いて大小では順序付けられない人々の選好や選択を分析する。農山漁村域と都市域の、価値観の多様性の比較から、都市域の人々の選好や選択のバラつきが高く、多様性が高くなると仮定した。その仮説を実証するため、Coefficient of unalikeability を用いて、異世代間ジレンマゲームにおける都市域と農山漁村域の意見のバラつき度合いを数値化した。数値化したデータを分析し、農山漁村域よりも都市域のほうが、意見や選択のバラツキ・多様性が高いと云う事が示唆された。

### 2. 序論

統計学では、標準偏差のように、大小がつけられるデータの分析手法が多く、ばらつきを計ることが出来る。しかし、Unordered categorical variable、という、牛丼を選ぶか、カレーを選ぶか、といった意見のような順番や大小のつけることが出来ないデータの変数の場合ではそういった分析手法が多くはみられなかった。

世の中には大小のつけられない物事での判断をすることが多く有る。筆者の所属する研究室の多くの実験では Unordered categorical variable data が出てくることが多く

あった。しかし、ばらつきについて調査することはほとんど無かった。そこで、Unordered categorical variable data の場合は、統計学における標準偏差のように、バラつきやバリエーションを計る指標がないのだろうかと考えた。本研究では、ネパールの都市域と農山漁村域における ISDG ゲーム（世代間持続可能性ジレンマゲーム）の実験データを Coefficient of unalikeability を用いてバラつきの度合いを数値化し、分析結果を考察し、都市域と農山漁村域における意見のバラつき度合いを数値化したとき、実際に都市域のほうが、多様性が高く示されるのかを実証していく。さらに、実験内で行われた話し合いの前と後での違いに対しても数値化を行い、考察する。

### 3. 実験手法

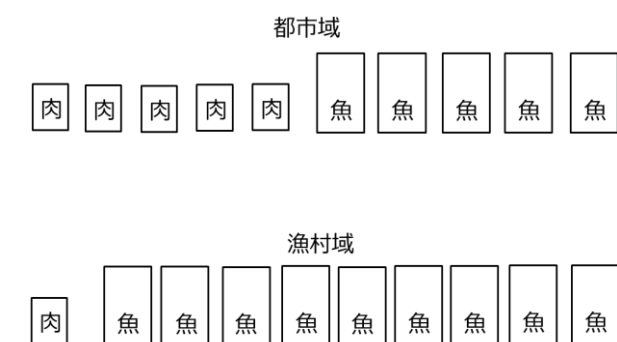
(図 1.) はネパールで、都市域 177 人と農山漁村域 186 人の計 363 人に ISDG ゲームを行い、個人的に A か B の、どの意見をチョイスするのかを調査した結果をまとめたものである。このデータを元に、意見のばらつきを調べるため Coefficient of unalikeability を使用した。ISDG ゲームでは、初めに A,B,N の意見をチョイスした後、3 人のグループで話し合いを行って貰い、もう一度意見をチョイスして貰った。

被験者	都市域	農山漁村域
363 人	177 人	186 人
意見		
A	51 人	21 人
B	119 人	160 人
N	7 人	5 人

(図 1.)

Coefficient of unalikeability とは、観測値が互いにどの程度異なっているかに焦点をあて分析する手法である。例えば、都市域 10 人と漁村域 10 人に、夕飯に肉を食べるか、魚を食べるか、というアンケートを実施したとしよう。その結果、都市域では肉を食べると答えた人が 5 人、魚を食べると答えた人が 5 人。漁村域では肉を食べると答えた人が 1 人、魚を食べると答えた人が 9 人であった。(図 2.)

肉か魚どちらを食べるか



(図 2)

このデータを縦と横に並べる。縦と横で交差する場所で、同じ選択肢ならば 0、違う選択肢ならば 1 と記載する。(図 3.1・3.2)

都市域

	肉	肉	肉	肉	肉	魚	魚	魚	魚	魚
肉	0	0	0	0	0	1	1	1	1	1
肉	0	0	0	0	0	1	1	1	1	1
肉	0	0	0	0	0	1	1	1	1	1
肉	0	0	0	0	0	1	1	1	1	1
肉	0	0	0	0	0	1	1	1	1	1
魚	1	1	1	1	1	0	0	0	0	0
魚	1	1	1	1	1	0	0	0	0	0
魚	1	1	1	1	1	0	0	0	0	0
魚	1	1	1	1	1	0	0	0	0	0
魚	1	1	1	1	1	0	0	0	0	0

(図 3.1)

漁村域

	肉	魚	魚	魚	魚	魚	魚	魚	魚	魚
肉	0	1	1	1	1	1	1	1	1	1
魚	1	0	0	0	0	0	0	0	0	0
魚	1	0	0	0	0	0	0	0	0	0
魚	1	0	0	0	0	0	0	0	0	0
魚	1	0	0	0	0	0	0	0	0	0
魚	1	0	0	0	0	0	0	0	0	0
魚	1	0	0	0	0	0	0	0	0	0
魚	1	0	0	0	0	0	0	0	0	0
魚	1	0	0	0	0	0	0	0	0	0
魚	1	0	0	0	0	0	0	0	0	0

(図 3.2)

交差した場所に出来た数字の数を分母に取り、その中の 1 の数を分子に取り、その数値がより 1 に近ければ近いほど、そのデータの中で異なった意見が存在していると言える。逆に、0 に近ければ近いほど同じ意見が多いといえる。これを利用し、意見のバラつき度合いを数値化し、意見がどれだけ異なっているのかを導き出すことが出来る。それによりどちらの多様性が高いかを判断することが出来る。上記の例だと、都市域のデータは分母が  $10 \times 10$  の 100。分子が  $5 \times 5$  の部分が二つあるので、 $5 \times 5 \times 2$  の 50 となる。漁村域のデータは分母が  $10 \times 10$  の 100。分子が  $1 \times 9$  の部分が二つあるので  $1 \times 9 \times 2$  で 18 となる。都市域では 0.5 漁村域では 0.18 という数値から、都市域の方が 1 に近く、多様性が高いといえる。(図 4)

都市域	50/100	0.5
農山漁村域	18/100	0.18

(図 4)

上記の理論をネパールでの ISDG ゲームでの実験結果に対して使用し、実際に都市域のほうが農山漁村域に対して意見の多様性が高く示されるのかどうかを数値化して対比していく。そして、話し合いを行う前と、話し合いを行った後での意見のバラつき度合いも調べ、考察していく。

#### 4. 結果

ネパールで行った ISDG ゲーム実験の被験者は合計 363 人、内都市域の被験者は 186 人、農山漁村域の被験者は 177 人である。都市域の被験者のデータを縦と横に並べ、交差した部分が同じ意見ならば 0、その中で異なった回答の部分を 1 として座標を作った。そのデータから意見のばらつきを数値化する。そのうち A の意見、B の意見とともに N (no answer) も含め三つの選択肢で集計した。図 5 は都市域と農山漁村域の集計結果をまとめ計算した内訳である。都市域の分母となる全体の数は  $177 \times 177 = 31329$  同じ意見を選んでる部分は A と A の  $51 \times 51$ 、B と B の  $119 \times 119$ 、N と N の  $7 \times 7$  を足した 16811。違う意見を選んでる部分は A と B が 2 箇所  $51 \times 119 \times 2$ 、A と N が 2 箇所  $51 \times 7 \times 2$ 、B と N が 2 箇所  $119 \times 7 \times 2$  を足した 14518 という

結果を示した。

農山漁村域の分母となる全体の数は  $186 \times 186 = 34596$ 。同じ意見を選んでいる部分は A と A の  $21 \times 21$ 、B と B の  $160 \times 160$ 、N と N の  $5 \times 5$  を足した 26066。違う意見を選んでいる部分は A と B が 2 箇所  $21 \times 160 \times 2$ 、A と N が 2 箇所  $21 \times 5 \times 2$ 、B と N が 2 箇所  $160 \times 5 \times 2$  を足した 8530 という結果を示した。(図 5)

これに Coefficient of unalikeability を用いると、都市域の分母になる数は全体の 31329 分子になる数は 1 になる座標の数 14518 である。農山漁村域の分母になる数は全体の 34596 分子になる数は同じく 1 になる座標の数 8530 である。よってこれらを計算し数値化すると都市域は 0.46 農山漁村域は 0.25 という結果になった。(図 6)

都市域	農山漁村域
全体	全体
$177 \times 177$ = 31329	$186 \times 186$ = 34596
0 になる	0 になる
$(51 \times 51) +$ $(119 \times 119) +$ $(7 \times 7)$ = 16811	$(21 \times 21) +$ $(160 \times 160) +$ $(5 \times 5)$ = 26066
1 になる	1 になる
$(51 \times 119 \times 2) +$ $(51 \times 7 \times 2) +$ $(119 \times 7 \times 2)$ = 14518	$(21 \times 160 \times 2) +$ $(21 \times 5 \times 2) +$ $(160 \times 5 \times 2)$ = 8530

(図 5)

都市域	14518/31329	0.46
農山漁村域	8530/34596	0.25

(図 6\* 小数点第 3 位以下切り捨て)

この結果より都市域の数値が農山漁村域の数値よりも 1 に近いことから、都市域のほうが意見のチョイスのほうがバラツ

きの度合いが高いといえる結果を示した。

図 7 と図 8 は ISDG ゲームでのデータにおける、話し合いの前と後に行った回答の内訳である。図 7 は話し合いの前に A と答えた人、B と答えた人、N と答えた人を、話し合いの後に意見が変わった人別に分類してある。図 8 では、話し合いの後に意見が変わった人も変わらなかった人も含め、最終的にどの意見を選択しているのかを分類している。この分類された内訳に対し、Coefficient of unalikeability を使用し、話し合いを行う前と、後で、意見のバラつき度合いにどういった変化が起きたのかを調べた。(図 9)

前	都市域	農山漁村域
A	30 (16.95%)	17 (9.14%)
A	12 (6.78%)	2 (1.08%)
A	9 (5.08%)	2 (1.08%)
B	99 (55.93%)	146 (78.49%)
B	11 (6.21%)	4 (2.15%)
B	9 (5.08%)	10 (5.38%)
N	2 (1.13%)	0 (0.00%)
N	3 (1.69%)	1 (0.54%)
N	2 (1.13%)	4 (2.15%)

(図 7)

後	都市域	農山漁村域
A→A	30 (16.95%)	17 (9.14%)
B→A	11 (6.21%)	4 (2.15%)
N→A	3 (1.69%)	1 (0.54%)
B→B	99 (55.93%)	146 (78.49%)
A→B	12 (6.78%)	2 (1.08%)
N→B	2 (1.13%)	4 (2.15%)
N→N	2 (1.13%)	0 (0.00%)
A→N	9 (5.08%)	2 (1.08%)
B→N	9 (5.08%)	10 (5.38%)

(図 8)

	都市域	農山漁村域
話し合い前	=0.46	=0.25
話し合い後	=0.52	=0.31

(図 9)

話し合い前の数値よりも、話し合い後の数値のほうが都市域、農山漁村域ともに数値が高く示された。都市域は 0.06 上昇し、農山漁村域でも 0.06 の上昇が見られた。両方とも数値が上昇したが、話し合い前後を比べても都市域のほうが、多様性が高いと示された。

## 5. 考察

結果より、都市域のほうが、多様性が高く示された。農山漁村域と都市域の文化レベルは大きく違う。経済学上、人というのは力そのものである。人が沢山集まることで多様性と競争が生まれ、総体的にレベルは過疎地域よりも当然上がると考えられる。豊富な貨幣や物資も集まりやすい。

一方、農山漁村域だからこそ資本主義に塗れない自由な発想が台頭することや、漁業や畜産での実力での上がれるというメリットもある。名を挙げたくて、わざわざ地方に活躍の場を求める人々もいる。

しかしながら、都市域では、日々文化の切磋琢磨が資本主義的な激しい競争が起きており、技術の革新的な発明や大きなビジネスが生まれている。地域における多様な文化芸術を振興するなど、文化による多様な価値観の形成と地域の包摂的環境の推進による文化芸術の社会的価値の醸成を図る政策も行われている。もちろん、農山漁村域に比べ都市域のほうが人口も多い。当然そこに住む人々も、人種、思想、文化と多種多様になる。そのため、人々の考え方や意見のチョイスにバラつきが大きく生じたとも考えられる。都市域と農山漁村域では約 2 倍もの差が生じた。本研究では、どの意見を選ぶか、ということには重きを置いておらず、チョイスする意見が、どれほどばらつくかということに注目しているため、都市域と農山漁村域の人々が、どういった意見を選びやすいのかということは調査することは出来なかった。データを見ると農山漁村域の人々は B の意見を多くの人が選ぶ傾向にあったが、今回の数値化による比較では、何故 B の意見を選んだのかは分からなかった。同じように都市域でも B の意

見のほうが選ぶ人が多かったが、何故かは分からなかった。

しかし、Coefficient of unalikeability を使用し実際に意見の選択のバラつき度合いが数値化し比較することが出来た。さらに、都市域と農山漁村域において 2 倍近い差が出たことによって都市域のほうが、多様化が進んでいることを実際に確認することが出来た。それにより、どういった環境の人々の意見がどれだけ多様化しているかを調べることが出来る。これにより、事前に意見調査を行い、その地域で行われる政策や、商品などのマーケティングにも活かすことが出来るのではないかと。多様性が高く示される地域では様々な意見に対応した政策や商品が必要になる。しかし、今回の農山漁村域のような、多様性が低く示された地域では、統一される傾向のあるニーズに対してアプローチすることが可能になるだろう。

さらに、話し合いの前のチョイスと話し合いの後でのチョイスを比較することも出来た。都市域と農山漁村域での数値を見ると、両方とも同じ値だけ数値が上昇していた。この結果より、話し合いを行うことによって、多様性に違いはあるが、都市域、農山漁村域ともに意見が同程度多様化することが分かった。

本研究では、ネパールにおける ISDG（世代間持続可能性ジレンマゲーム）ゲームのデータを使用し意見のバラつき度合いを調べたが、用意された回答に一般常識や偏見が加味されるような意見があるならば、また違った数値が出ていたかもしれない、どの意見が集約するのかを考慮したうえで、ある程度定義付けした尺度を用いた場合、より信憑性の高い多様性を調べる一つの指標となり、研究においてだけでなく企業のマーケティングや国民の意識調査などでの利用が可能になるのではないだろうか。

## 6. 参考文献

- ・ Kader, G. and Perry, M.(2007) Variability for Categorical Variables. Journal of statistics education15

文部科学省（2018）平成 29 年度 文部科学白書