

分配の正義への社会的価値関数アプローチ

清水裕士^a

要約

人々が不平等な社会状態を回避する傾向があることは、政治学、心理学、そして経済学などの多くの分野の研究で見出されている。一方で、自己が当事者でない場合においても、平等な分配が不平等な分配よりも選好されることが社会心理学において指摘されているが、そのメカニズムについては未だ明らかになっていない。本研究では、自己が当事者でない場合の報酬分配についての選好が、社会状態についての価値（以降、社会的価値）の比較によって生じること、そして社会的価値が、不平等の程度に応じて割り引かれるという観点から、社会選好とリスク選好を用いてモデルを構築した。Web を用いた実験から、社会選好とリスク選好の両方を仮定したモデルが最もデータにフィットしたこと、しかし現実的なパラメータの範囲では人々のマキシミン選好を説明できないことがあきらかとなった。

JEL 分類番号： D63, C51

キーワード： 分配の正義, マキシミン原理, 社会選好, ベイズ統計モデリング

^a 関西学院大学社会学部 hsimizu@kwansei.ac.jp

1. 問題

1.1. 望ましい社会状態とそれに対する個人の選好

どのような分配が望ましいのかについては、社会学、心理学、経済学、政治学など、多くの社会科学の研究で研究が行われている。厚生経済学では、人々の所得をどのように配分すべきかについて、公理的なアプローチによる規範的な分析がなされている。Samuelson(1947)は、ある社会状態への個人の効用から、社会的厚生を定義する、社会厚生関数アプローチをとった。社会的厚生関数を定める上で、どのような原理に基づいて個人の効用を写像するかが議論の対象とされる。具体的には、パレート原理、功利主義、マキシミン原理(Rawls, 1971)などが提案されている。

一方で、個人がどのような分配を選好するのかを扱った研究も多くある。最後通牒ゲームを用いた経済実験では、人々は他者に平等に分配することを好み、また他者からの不平等な分配を拒否する傾向があることが繰り返し示されている（たとえば Güth, Schmittberger, and Schwarze, 1982）。Fehr & Schmid (1999)は最後通牒ゲームの実験結果を受けて、自己が得た利益 $x \in \mathbb{R}$ の効用だけではなく、他者との不平等を避ける傾向も効用関数に含まれていることをモデルで示した。それは、

$$U(x_i | x_1 \dots x_n) = x_i - \alpha \frac{1}{n-1} \sum_{i \neq j} \max(x_j - x_i, 0) - \beta \frac{1}{n-1} \sum_{i \neq j} \max(x_i - x_j, 0)$$

である。個人 $i \in \{1, 2, \dots, n\}$ の効用 $U(x_i)$ は、自分を含めた n 人の個人 $j \in \{1, 2, \dots, n\}$ の得た利益の関数として表現される。 $\alpha \in (-\infty, \infty)$ は妬みパラメータ、 $\beta \in (-\infty, \infty)$ は哀れみパラメータと呼ばれ、これまでの実験結果から、 $\alpha > \beta > 0$ であることが想定されている。このモデルを以降 FS 効用関数と呼ぶ。

1.2. 社会心理学における公正判断についての研究

一方で社会心理学の分配正義研究では、人々がどのような社会状態を望ましいと判断するのか、つまり分配の正義についての個人の評価を対象とした研究も行われている。たとえば田村・亀田(2004)では、自己の利得が一切関与しない社会状態への公正判断を求められた場合でも、パレート効率的な分配よりも平等な分配が選ばれる場合があることを明らかにした。また Messick (1993)は、分配の正義についての意思決定のアンカーとして平等原理があり得ることを指摘している。

Kameda et al. (2016)においても中立な立場から社会状態を評価させ、効率的な社会状態よりもマキシミン原理に基づく公正判断が行われることを明らかにしている。また、このKameda et al. (2016)は、中立な立場における分配の正義のマキシミン選好と、くじについてのリスク選好に相同性があることを明らかにしている。その相同性は、視点取得に関わる

脳部位の活動と関連があることを見出している。このことは、マキシミン原理を採用する個人は、自身のリスク回避傾向の高さから、最も不遇な状態を自身に当てはめ、それを改善することに動機づけられる可能性を示している。

公正判断についての社会心理学的研究では、分配の正義の認知的アンカーとして、1. 平等原理、2. マキシミン原理、の2つが挙げられている。平等とマキシミンは混同されやすい原理 (Engelmann & Strobel, 2004) だが、後者が効率性を内包しているという意味では区別される。しかし、これまでの研究では公正判断についての意思決定モデルがないため、両者を理論的に明確に区別する、あるいは統合できる枠組みが提供されてこなかった。そこで本研究は社会状態についての公正評価についてのモデルを作成し、実験によって得られたデータによってその妥当性を検討することを目的とする。

1.3. モデル

本研究では、Kameda et al. (2016)の分配の正義における自己の他者への投影という知見を受け、以下の想定を置く。すなわち、人々は中立的な立場で社会状態を評価する場合でも、自分がもしその社会状態に置かれたときに得られうる期待効用 (以下、社会的価値) を最大化する社会状態を選好するだろうと仮定する。以下にモデルの詳細を述べる。

第三者に配分された報酬分配の望ましさを人々が選好する状況において、その選好が社会的価値関数によって算出された社会的価値 V の比較によって実現されると想定する。社会状態の評価者 $k \in \{1, 2, \dots, K\}$ が判断する社会的価値 V_k は、分配されるメンバー $i \in \{1, 2, \dots, n\}$ の分配された利益に基づく、評価者 k の効用関数 V を用いて、

$$V_k(x) = V\{U_k(x_1), \dots, U_k(x_i), \dots, U_k(x_n)\}$$

と表されると仮定する。このとき、効用関数 U_k はFS効用関数であると仮定する。すなわち、分配されるメンバーの利得を $x = \{x_1, x_2, \dots, x_n\}$ としたとき、評価者 k が推測するメンバー i の効用 $U_k(x_i)$ は、評価者 k 自身のFS効用関数のパラメータ α_k, β_k を用いて計算される。FS効用関数のパラメータ α と β は、Fehr & Schmidt(1999)では個人差が想定されていないが、本研究ではそれが評価者ごとに異なると仮定する。

期待効用の計算において、Kameda et al. (2016)ではリスク選好とマキシミン選好に相同性があることが指摘されていることから、累積プロスペクト理論(Tversky. and Kahneman, 1992)のリスク選好を考慮した期待効用、つまりシヨケ期待値を採用する。累積プロスペクト理論では、確率荷重関数を $w(p) = \frac{p^\gamma}{(p^\gamma + (1-p)^\gamma)^{1/\gamma}}$ とおき、その期待効用 $\sum \pi_i U_k(x_i)$ の重み π_i を $\pi_i = w(p_i + \sum_{j+1}^n p_j) - w(\sum_{j+1}^n p_j)$ とする。ただし、 $x_1 < x_2 \dots < x_n$ である。シヨケ期待値は、 $p_i = 1/n$ のとき、 γ が小さくなるにつれて単調減少し、 $\gamma \rightarrow 0$ のときもっとも不遇な x_1 に

一致する。つまりリスク選好の程度に応じてマキシミンの選好に変化することが示される。

ここで、社会選好を仮定する場合と、リスク選好を仮定する場合のそれぞれについて社会的価値関数を導出する。まず最も単純に、報酬の期待効用を社会的価値とする場合、つまり社会選好もリスク選好も考慮しない場合、社会的価値は $V_k(x) = \sum_{i=1}^n \frac{1}{n} x_i = \mu$ である。次に、社会選好を仮定した社会的価値関数は

$$V_k(x) = \sum_{i=1}^n \frac{1}{n} U_k(x_i)$$

と表される。効用関数 U_k が FS 効用関数であることから、これを整理すると

$$V_k(x) = \mu \{1 - (\alpha_k + \beta_k)G\}$$

と整理できる。なお、 G はジニ係数であり、 $G = \frac{1}{2\mu n(n-1)} \sum_{i=1}^n \sum_{j \neq i}^n |x_j - x_i|$ である。この社会的価値関数が意味することは、社会的価値が与えられた報酬の平均値にたいして、 $\{1 - (\alpha_k + \beta_k)G\}$ だけ割り引かれることである。これを社会的価値の社会選好・期待効用モデルと呼ぶ。続いて、リスク選好のみを考慮するモデルは、

$$V_k(x) = \sum_{i=1}^n \pi_i x_i = \mu' = \mu \left\{ \frac{\mu'}{\mu} \right\}$$

である。以降、シヨケ期待値を μ' と表記する。これをリスク選好効用モデルと呼ぶ。このモデルでは、報酬の期待値から μ'/μ だけ割り引かれている。最後に、社会選好とリスク選好の両方を想定する社会選好・リスク選好モデルは、

$$V_k^i = \mu \left\{ \frac{\mu'}{\mu} - (\alpha_k + \beta_k)G_1' - (\alpha_k - \beta_k)G_2' \right\}$$

であり、ただし、 $G_1' = \frac{1}{2\mu(n-1)} \sum_{i=1}^n \pi_i \sum_{j \neq i}^n |x_j - x_i|$ 、 $G_2' = \frac{1}{2\mu(n-1)} n(\mu - \mu')$ である。 G_2' は $\gamma = 1$ のとき、 $\mu' = \mu$ となるため 0 となる。

これらのモデルについて、実験のデータからパラメータの推定及びモデル比較を行った。

2. 方法

2.1. データ取得

実験は、クラウドワークスに登録している 627 名を対象に、Web による調査形式で行った。調査画面は Qualtrics によって作成された。実験報酬は一人あたり 60 円であった。

社会選好の測定 FS モデルのパラメータの推定のため、水野・清水 (2020) の手法を用いて、自己が関与する報酬分配についての選好課題を行った。

社会状態の公正判断 続いて、回答者は自己が関与しない、架空の 3 人の人物への報酬分配についての選好について回答を求められた。選択肢は、左は常に 3 人に対して平等な分配が、左側は常に不平等な分配額が提示された。平等な分配額は、5 万ずつの分配から 5

千円ずつ小さくなり、最終的に 5 千円ずつの分配となった。不平等な分配額は、(4 万, 5 万, 6 万), (3 万, 5 万, 7 万), (2 万, 5 万, 8 万), (1 万, 5 万, 9 万), (千, 5 万 9 千 9 百) の 5 パターンが用意された。

2.2. 実験手続き

本研究では分配タイプへの選好を測定するのではなく、時間選好などの測定に用いられる主観的等価点を推定する手法を用いた。完全に平等な分配額についての社会的価値は、その分配額の平均値と一致する。この前提のもとで、平等分配の選択肢と、不平等分配の選択肢、たとえば $x = (40000, 50000, 60000)$ の 2 つを提示し、どちらを選好するか回答者に尋ねる。このとき、平等分配の額は、最初は 5 万ずつだが、そのあと 4 万, 3 万と減少していった。また本研究では、ベイズ統計モデリングによってモデルパラメータの推定を行った。また、今回のデータではリスク選好について測定されていないため、リスク選好パラメータ γ についてはモデルから推定した。

3. 結果

社会状態の公正判断について、二肢課題における反応 $Z = 0, 1$ を選択する確率 $\theta \in [0, 1]$ を、以下のようにベルヌーイ分布を用いて、

$$Z_k \sim \text{Bernoulli}(\theta_k), \quad \theta_k = \frac{\exp(\lambda_k V_k^i)}{\exp(\lambda_k V_k^i) + \exp(\lambda_k V_k^e)}$$

と仮定した。上記の式中の $V_k(x)$ はそれぞれのモデルの式から計算される。推定は `rstan2.18` による自動微分変分推論を用いて行った。

モデル比較のために、各モデルで情報量基準として負の対数周辺尤度を `bridge sampling` 法によって推定した。結果、社会選好とリスク選好の両方を想定したモデルが最も情報量基準が小さかった。

表 1. モデル比較の結果

モデル	負の対数周辺尤度
社会選好期待効用モデル	9599.46
リスク選好効用モデル	13583.48
社会選好・リスク選好モデル	6133.79

モデルで推定された社会的価値は、データの主観的等価点と 0.7~0.9 程度の相関が得られ、高い予測力を持つことが示された。しかし、社会選好・リスク選好モデルにおいて推定されたリスク選好パラメータ γ は 0.16 と、Tversky and Kahneman (1992) で得られた 0.61 に比べては非常に小さかった。これは、リスク選好だけでは、人々のマキシミ的な

公正判断を十分説明できていないことを意味する。そこで $\gamma = 0.61$ に固定したときに、 p_i を $1/n$ ではなくフリー推定したところ、最も報酬の少ない人の確率が0.97となった。この結果から、人々は社会選好とリスク選好では説明できない、「最も不遇な人に自主的な関心を示す」マキシミンの認知バイアス(Kameda, et al. 2016)を持っていることが示唆された。

4.考察

本研究では、自己が関与しない中立的な立場における公正判断のメカニズムを社会選好とリスク選好を用いてモデル化した。ベイズ統計モデリングによるモデル比較の結果、社会選好もリスク選好も考慮するモデルが妥当であることが示された。しかし一方で、現実的なリスク選好のパラメータの範囲では、人々のマキシミンのバイアスは説明できないことも示された。今後は、リスク選好の測定を行い、その個人差も考慮した上でモデルが妥当であるのかを確かめる必要がある。

引用文献

- Fehr, E. and Schmidt, K.M. 1999. A theory of fairness, competition, and cooperation. *The Quarterly Journal of Economics*, 114, 817-868.
- Güth, W., Schmittberger, R. and Schwarze, B. 1982. An Experimental Analysis of Ultimatum Bargaining. *Journal of Economic Behavior and Organization*, 3, 367-388.
- Tversky, A. and Kahneman, D. 1992. Advances in prospect theory: Cumulative representation of uncertainty. *Journal of Risk and Uncertainty*. 5,297-323.
- Kameda, T., Inukai, K., Higuchi, S., Ogawa, A., Kim, H., Matsuda, T., & Sakagami, M. 2016. Rawlsian maximin rule operates as a common cognitive anchor in distributive justice and risky decisions. *Proceedings of the National Academy of Sciences*, 113, 11817-11822.
- Messick, D. M., and Schell, T. 1992. Evidence for an equality heuristic in social decision making. *Acta Psychologica*, 80, 311-323.
- Rawls, J. 1971. *A Theory of Justice*. Harvard Univ. Press, Cambridge, MA.
- Samuelson, P. A. 1947. *Foundations of economic analysis*. Harvard University Press. Cambridge, MA, US.
- 田村亮・亀田達也 2004. 「寡きを患えず，均しからずを患う」?: グループの意思決定におけるパレート原理の作用, *社会心理学研究*, 20, 26-34.