

## 対戦実験による「繰り返し囚人のジレンマ」の戦略分析

( 入会予定 ) 関西大学 \*西山 武 NISHIYAMA Takeshi  
 関西大学 山田奈津子 YAMADA Natuko  
 ( 01401144 ) 関西大学 中井 暉久 NAKAI Teruhisa

## 1. はじめに

社会生活において様々な主体は、それぞれの目的を達成しようと「意思決定」をおこなう。そして多くの場合、目的の達成度は自分自身の意思決定ばかりでなく、他の主体の意思決定にも依存する。そのような社会構造の中で「社会的ジレンマ」という状況が生まれることがある。

「社会的ジレンマ」の基本定義として、ロビン・ドウズ (Robyn M. Dawes) が以下のような3点を示している [1]。

1. 意思決定可能な個体間で、互いに「協力」・「非協力」が選択できるような状況がある。
2. 各個体にとっては「協力」を選択するよりも「非協力」を選択する方が望ましい結果が得られる。
3. 全個体が自分にとって有利な「非協力」を選択した場合の結果は、全個体が「協力」を選択した結果よりも悪いものになってしまう。

そこで本研究では、このような意思決定の相互依存関係を数理的に分析する「ゲーム理論」を用い、実際の社会におけるジレンマの状況に対する人間の行動を、ゲーム理論の視点から検証する。

## 2. 囚人のジレンマ・ゲーム [2]

Table 1. 「囚人のジレンマ」型の利得行列

利得行列		Player II	
		協調	裏切り
Player I	協調	(10, 10)	(0, 15)
	裏切り	(15, 0)	(1, 1)

囚人のジレンマでは、相手が協調もしくは裏切りのどちらを選んでも、自分は協調よりも裏切りを選ぶ方が得をする(裏切りの魅力)。しかし、この結果「自分だけが裏切りひば、最も得をすることができるが、二人ともが裏切りを選んでしまうと、お互い1の利得しか得ることができない。それではお互いが協調を選ぶ方が、高い利得を得ることができるのではないか…」というジレンマ状況が生まれる。さらにこのジレンマ状況が繰り返されることで、将来的な総利得を高めることを考慮する必要が生じ、「裏切り」よりも「協調」を選択する動機が強まる。これが囚人のジレンマの繰り返しによって発生する、新たなジレンマ状況である。

## 【戦略特性変数の定義】

1980年、ロバート・アクセルロッド (Robert Axelrod) による「コンピュータ選手権」が行われた [3]。これは、プログラムによって構成された戦略同士を、コンピュータ上で総当たり対戦させるものであり、この「コンピュータ選手権」で最高の得点(総利得)を獲得した戦略が「しっぺ返し (TIT FOR TAT: 以下 TFT)」というものである。TFT とは、最初の回は協調し、その後は前の

回の対戦で相手をとった手をまねるという単純な戦略であり、以下のような特徴・戦略特性がある。

1. 協調を基本とし、自分からは決して裏切りを出さない。
2. 相手が裏切りを出した時、すぐに反応して次回裏切りを出す。
3. 1回裏切りを返したら、次回からは再び協調の姿勢にもどる。

そしてアクセルロッドは、この選手権の優勝戦略である TFT が人間にとっても最良の戦略であるという結論に至り、その特徴を文章によって示したのである。そこで本研究では新たに、これらの特徴を「戦略特性変数」というパラメータとして定義し、「人間対コンピュータ」の対戦実験によって、プレーヤー戦略の特徴を数値的に表現・解析する。

## 戦略特性変数

繰り返し数  $n$  回の囚人のジレンマ・ゲームにおいて、各プレーヤーの出した戦略「協調 (C)」「裏切り (D)」をもとに、以下の5つの数値をカウントする。

- $\alpha$  : 相手がCを出した回数。
- $\beta$  : 相手のCの直後に、自分がCを出した回数。
- $\gamma$  : 相手のDの直後に、自分がDを出した回数。
- $\delta$  : 相手が続けてD→Cと出した回数。
- $\epsilon$  : 相手のD→Cの後に、自分がCを出した回数。

そしてこれら5つの変数をもとに、次の3つの戦略特性変数…「善良度  $\theta$ 」「敏感度  $\nu$ 」「寛容度  $\xi$ 」を求めらる。

善良度  $\theta$  : 自分から先に裏切りなかつた割合。

$$\theta = \beta / \alpha \quad \dots (1)$$

敏感度  $\nu$  : 相手の裏切りに対して、すぐに裏切りを返した割合。

$$\nu = \gamma / (n - \alpha) \quad \dots (2)$$

寛容度  $\xi$  : 相手の改悛 (D→C) に対して、根に持たず協調を返した割合。

$$\xi = \epsilon / \delta \quad \dots (3)$$

## 3. 対戦実験

## 【実験方法】

ジレンマ状況での人間行動を「戦略特性変数」を用いて分析するため、人間とプログラムによる対戦実験をおこなう。

1. ゲームのルール、利得行列等の説明。
2. 被験者には、対戦相手9名と Table 1 の利得行列を用いた繰り返し数50回の「繰り返し囚人のジレンマ・ゲーム」で対戦をおこなってもらう。(対戦1)
3. 被験者の獲得利得・戦略特性変数等、解析に必要なデータをとる。(データ1)
4. データ1を被験者に提示。対戦1における結果(獲得利得と各々の順位)を説明し、さらに高い利得を目指して、再度、9人の対戦相手と対戦をおこなってもらう。(対戦2)

- 対戦2について、被験者の獲得利得・戦略特性変数等のデータをとる。(データ2)
  - データ1とデータ2を比較・考察する。
- 今回の実験では被験者の対戦相手として、コンピュータ・プログラムによる戦略9つを用意した。
- A: 全面協調戦略: 常に「協調」を出し続ける戦略  
 B: TFT  
 C: ためらめ戦略: ランダムに「協調」「裏切り」を選ぶ戦略  
 D: 一度目は「裏切り」で、それ以降は相手の手をまねる戦略  
 E: 25回まではためらめ、それ以降は「協調」する戦略  
 F: 2回連続の「裏切り」にのみ反応するTFT型の戦略  
 G: 25回目までの合計利得によって、その後の手を「全面協調」か「全面裏切り」に変化させる戦略  
 H: 相手の「協調」「裏切り」の各回数をカウントし、その回数に応じて各回の自分の戦略を変える戦略  
 I: 10回目までの相手の出方によって、以降の手を「協調」「裏切り」のどちらかに決定してしまう戦略

**【実験の前提条件】**

- 対戦相手(プログラム)の出現順序はランダムに決定される。
- 繰り返し数(50回で終了)は被験者には知らされていない。
- 対戦相手の特徴・性質は被験者には知らされていない。
- 会話等で対戦相手とコミュニケーションをとることは不可能。

**【実験結果】**

各戦略特性変数・総利得について、データ1とデータ2を用いた統計的検定をおこなった結果、被験者の戦略特性に関して次のような変化が観測された。(Table 2)

- 「**敏感度**」の減少…自分の利得を高めるために、相手の出方・戦略の構成をじっくりと伺い、即応することを避けている。
- 「**寛容度**」の増加…自分が相手に対して寛容に対処することで、相手の協調を引き出そうと考えている。

Table 2. 1回目と2回目のデータに関する検定

	検定内容	1回目	2回目	有意差
1	被験者全員: 獲得利得の差	378.3968	441.7698	**
2	被験者全員: 善良度 $\theta$ の差	0.4056	0.4532	
3	被験者全員: 敏感度 $\nu$ の差	0.7925	0.6815	*
4	被験者全員: 寛容度 $\xi$ の差	0.3139	0.4958	**
5	利得上昇者: 善良度 $\theta$ の差	0.3552	0.4443	
6	利得上昇者: 敏感度 $\nu$ の差	0.8353	0.6870	*
7	利得上昇者: 寛容度 $\xi$ の差	0.2774	0.4964	**
8	利得下降者: 善良度 $\theta$ の差	0.6994	0.4857	
9	利得下降者: 敏感度 $\nu$ の差	0.6355	0.6612	
10	利得下降者: 寛容度 $\xi$ の差	0.4480	0.4938	

\*\* :  $\alpha = 0.01$ , \* :  $\alpha = 0.05$

注) 網掛け部分は数値の高い方を示す

また最終結果を見ると、TFTを上回るような高い利得を得た被験者も何名が存在した。そこで、このTFTを上回る高い利得を得た被験者の上位6名について、その戦略の特徴がどのようにTFTと異なるのかを次のような比較によって検証する。

- 利得上位者6名の獲得利得(対戦1、対戦2)を、対戦相手A~Iごとに分割する。
- TFTが獲得した利得を対戦相手A~Iごとに分割する。

2. 各対戦相手A~IごとのTFTの獲得利得と被験者平均利得について、検定・グラフによる比較を行う。

この比較の結果、次のような特徴が見られた。(Table 3)

- 特定の対戦相手に対しては「TFT」の獲得利得を大きく上回る…コンピュータが臨機応変に対応できないのに対し、人間は相手の特性を考慮して、対戦ごとに柔軟に対応することができる。



Fig.1 TFTとの比較 (1回目の利得)



Fig.2 TFTとの比較 (2回目の利得)

Table 3. TFTとの比較

	対戦 1回目			対戦 2回目		
	有意差	TFT	被験者	有意差	TFT	被験者
A	**	500	7142	**	500	6792
B	*	500	2725	*	500	3828
C	*	287	3490	**	287	3605
D	**	375	1042	**	375	3523
E	**	362	5313	**	362	5143
F		287	3522	*	287	5430
G	*	433	2720	**	433	4908
H	**	495	1337	**	495	4010
I	*	414	5465	**	414	5707

\*\* :  $\alpha = 0.01$ , \* :  $\alpha = 0.05$

注) 網掛け部分は対戦1と対戦2で特に変化があった部分

**4. まとめ**

アクセルロッドの結論によると、前述の「善良性」・「感受性」・「寛容性」といった3つの特徴を持つ「TFT」こそが、最適の理想に近いプレーヤー戦略であると考え、これら3つの特徴の重要性を述べている。今回の実験の結果から、人間が作り上げる戦略においては「基本的には協調の姿勢で相手の協調を引き出されることに期待しながら、その反面、相手によってはその協調に付け込んで自分は裏切り、自分の利得を高めることも忘れない」といったような「柔軟」な戦略特性が重視されている点が示された。そして同時にこの結果は、プレーヤーの選択する戦略が、対戦相手のとる戦略の構成・特徴に起因して様々に変化することを示している。

**《参考文献》**

- [1] Robyn M. Dawes, "Social dilemmas," Annual Review of Psychology, 31, 169-193 (1980)
- [2] Davis, Morton D. "Game Theory: A Nontechnical Introduction" New York, Basic Books (1970)
- [3] Axelrod, R. and William D. Hamilton. "The Evolution of Cooperation." Science 211, 1390-96 (1981)