

卒論テーマ候補：ゆびすま

池上 慧

2017 年 10 月 4 日

1 「ゆびすま」とは

「ゆびすま」とは 2 人以上で行われるゲームである。ここでは 2 人で行われるケースを想定する。プレイヤーは毎回「攻め」と「守り」の役目を交互に行う。プレイヤーは毎回好きな本数の親指を上げる。「攻め」のプレイヤーは今回上がる親指の本数を予想し、掛け声とともにその予想した数をコールしながら、自分でも好きな本数だけ親指を上げる。「守り」のプレイヤーは掛け声の後に親指を好きな本数だけあげる。「攻め」がコールした数と実際にあげられた親指の総数が等しかったなら「攻め」の勝ちであり、そうでなければ「引き分け」である。引き分けたら役割を交代してどちらかが勝つまで続けるものとする。(本来であれば勝てば腕を一本減らすことができ、先に二回勝利した方の勝ちというルールであるが、ここでは最初の一回の勝利が起こるまでの、2 本 vs 2 本の状況のみを想定する。)

以下ではコールできる指の本数を 2 までに限定した「限定ゆびすま」をまずは考える。このルールのもとでは、守り手の行動を $\{0, 1, 2\}$ だと正しく認識した場合に、混合戦略ナッシュ均衡で各アクションに割り振られる確率に大小が生じる。一方で後に提示する「通常ゆびすま」ではコールできる指の本数を 4 本まで許容していて、このルールのもとでは正しい認識を持った時、勝つ可能性のある手に振られる確率は全て等しくなる。この二つのルールによる差を何かに使えないか？

2 研究の意図するところ

ゆびすまで攻め手の戦略を決める肝は、「自分のコールと自分で揚げる親指の数の差」であることは明らかである。ここで「 a とコールしながら b 本の親指を上げる」という行動を (a, b) と表記することになると、 $\{(0, 0), (1, 1), (2, 2)\}$ は自分のコールと自分のゆびの数が等しい組 (set1) で、 $\{(1, 0), (2, 1)\}$ を自分のコールが自分のゆびの数よりも 1 つ多い組 (set2)、 $\{(2, 0)\}$ を自分のコールが自分のゆびの数よりも 2 つ多い組 (set3) の 3 つに戦略を分類することができる。

これらはそれぞれ同じ組の中では勝利確率は等しい。以下で計算するように、相手の戦略を $\{0, 1, 2\}$ 、すなわちゆびを上げる本数だと捉えたと、この全てをサポートに持つ混合戦略ナッシュ均衡がこのゲームの唯一のナッシュ均衡として存在し、その時以上のセットには全て等確率、すなわち $\frac{1}{3}$ ずつの確率が振られることになる。この時、各セットに含まれる戦略の数は異なっていることに注意すると、個別の戦略に振られる確率は set3 の中身、set2 の中身、set1 の中身の順番で大きくなるはずである。

しかし、この結果は直感に反するものである。 $\{(2, 0)\}$ を選ぶということは自分は 1 本も上げないで、相手が 2 本とも上げることを想定しているということであるが、これが実現する可能性は低そうに思えてしまう。実際、何も説明せず友人とこのゲームを繰り返し行った時、いきなり $\{(2, 0)\}$ が選ばれる回数は少なかった。

むしろ選ばれがちなのは set2、すなわち $\{(1, 0), (2, 1)\}$ の 2 つであった。本研究ではゆびすまにおいて set2 が戦略として選ばれやすいことの原因として、「攻め手が知らず知らずのうちに守り手の左右の手を別々の意思決定主体が操作するものだ」と想定してゲームをプレイしている」という仮説を検証する。

以下で計算するように、守り手の腕を別々の主体と想定した場合の混合戦略ナッシュ均衡は現実的なものが 3 つ存在し、そのどれもで set2 に高い確率が付与されている。逆に set1 や set3 には確率が付与されない均衡も存在し、それらの戦略を極端な手として嫌ってしまう傾向と合致していると言える。

さらに限定合理性の文脈で提案されている各種解概念との比較を行うことで、限定合理性の一形態として相手主体の誤認があり得ることを確認する。

3 限定ゆびすまのナッシュ均衡

計算は別紙。

(q, r, s) で case1 において守り手が $\{0, 1, 2\}$ をそれぞれ取る確率を、 p で case2 において守り手の左右の腕が取る混合戦略、すなわち親指を上げる確率を、 (x, y, z) で攻め手が取る混合戦略、すなわち set1, set2, set3 をとる確率を表すとする。この時、それぞれのケースでいかが混合戦略ナッシュ均衡となる。

case1 : 相手が統一された意思決定主体だと想定

$$\begin{cases} (q, r, s) = (\frac{1}{3}, \frac{1}{3}, \frac{1}{3}) \\ (x, y, z) = (\frac{1}{3}, \frac{1}{3}, \frac{1}{3}) \end{cases} \quad (1)$$

case2 : 相手の左右の手を別々の意思決定主体だと想定

$$\begin{cases} p = \frac{1}{3} \\ (x, y, z) = (\frac{1}{3}, \frac{2}{3}, 0) \end{cases} \quad (2)$$

$$\begin{cases} p = \frac{1}{2} \\ (x, y, z) = (0, 1, 0) \end{cases} \quad (3)$$

$$\begin{cases} p = \frac{2}{3} \\ (x, y, z) = (0, \frac{2}{3}, \frac{1}{3}) \end{cases} \quad (4)$$

4 データへのフィットの比べ方

ゆびすまを実際に対戦させてデータを集めることを想定する。そのデータに対して先の二つのモデルの説明度合いを比較してどちらのモデルが採択されるかを調べたい。単純に考えられるのは以下のように尤度を比較すること。

case1 に対しては通常の意味で対数尤度が計算できる。case2 に対しては以下のように EM アルゴリズムを用いてその対数尤度を計算する。 $i = 1, \dots, n$ を n 回のプレイの index とする。 Y_i を攻め手が i 回目のプレイで行った戦略を示す確率変数であるとする。すなわちそのサポートは $\{(0, 0), (1, 1), (2, 2), (1, 0), (2, 1), (2, 0)\}$ の 6 つである。 $j = 1, 2, 3$ で均衡の index として、それぞれ先の均衡 (2), (3), (4) を示すとする。この時、 $f_j(y_i)$ は均衡 j がプレイ i で実現している時に、攻め手が出す戦略の確率分布関数である。先に定義された戦略の set 内では均等に確率が分配されるとすると、それぞれ以下ようになる。

$$f_1(y_i) = \begin{cases} \frac{1}{9} & y_i \in \text{set1} \\ \frac{1}{3} & y_i \in \text{set2} \\ 0 & y_i \in \text{set3} \end{cases} \quad f_2(y_i) = \begin{cases} 0 & y_i \in \text{set1} \\ \frac{1}{2} & y_i \in \text{set2} \\ 0 & y_i \in \text{set3} \end{cases} \quad f_3(y_i) = \begin{cases} 0 & y_i \in \text{set1} \\ \frac{1}{3} & y_i \in \text{set2} \\ \frac{1}{3} & y_i \in \text{set3} \end{cases} \quad (5)$$

ここで均衡 j が選ばれる確率を ψ_j で書くとすると、 Y_i は混合分布である $\sum_j \psi_j f_j(y_i)$ に従うとできる。missing variable として各プレイがどの均衡から生じたものなのを示した確率変数 u_i を考えると、完全なデータが与えられた時の尤度は、 $I(\cdot)$ を指示関数とすると以下のように書ける。

$$L(Y, u|\psi) = \prod_i \prod_j (\psi_j \cdot f_j(y_i))^{I(u_i=j)} \quad (6)$$

Q 関数の定義は以下である。

$$Q(\psi|\psi^{(0)}) = E_{\psi^{(0)}} [\log L(Y, u|\psi) | Y]$$

Q 関数はパラメータについて最大化されるので、パラメータに関連する部分だけ残せば最大化のための一階条件を問題なく得られる。従って以下では ψ に関係のない項は省略してイコールで記す。ただし、 ψ^0 は EM アルゴリズムない

で一回前の iteration で得られたパラメータの値であり、期待値は missing variable について取られているとする。

$$\begin{aligned} E_{\psi^{(0)}} [\log L(Y, u|\psi) | Y] &= \sum_i \sum_j E_{\psi^{(0)}} [I(u_i = j) \log \psi_j | Y] \\ &= \sum_j \log \psi_j \sum_i E_{\psi^{(0)}} [I(u_i = j) | Y] \end{aligned} \quad (7)$$

これをパラメータ ψ_1, ψ_2, ψ_3 について最大化することを考える。一階条件より

$$\psi_j = \frac{1}{n} \sum_i E_{\psi^{(0)}} [I(u_i = j) | Y] \quad (8)$$

である。

ここで、ベイズの公式より

$$E_{\psi^{(0)}} [I(u_i = j) | Y] = P_{\psi^{(0)}}(u_i = j | Y) = \frac{p_{\psi^{(0)}}(y_i | u_i = j) \cdot p_{\psi^{(0)}}(u_i = j)}{p_{\psi^{(0)}}(y_i)} = \frac{f_j(y_i) \cdot \psi_j^{(0)}}{\sum_j \psi_j^{(0)} f_j(y_i)}$$

を得るので、先の結果と合わせることでパラメータの更新式として以下を得る。

$$\forall j \quad \psi_j^{(1)} = \frac{1}{n} \sum_i \frac{f_j(y_i) \cdot \psi_j^{(0)}}{\sum_j \psi_j^{(0)} f_j(y_i)} \quad (9)$$

これによって収束させたパラメータについて先の Q 関数の値を計算し、それを持って case2 の対数尤度とする。case1 の対数尤度と case2 の Q 関数の値を比較してモデルのフィットを比べる。

5 通常ゆびすまのナッシュ均衡

先の限定ゆびすまの時と導かれる混合戦略ナッシュ均衡における各行動集合に割り振られる確率は同じである。守り手についても同様の確率を付与する混合戦略がナッシュ均衡となっている。しかし、通常ゆびすまにおいては先に置いた set がそれぞれ set1 が $\{(0, 0), (1, 1), (2, 2)\}$ 、set2 が $\{(1, 0), (2, 1), (3, 2)\}$ 、set3 が $\{(2, 0), (3, 1), (4, 2)\}$ となり、集合ないの要素の数が同じであるため、各行動に割り振られる確率自体は限定ゆびすまと異なる。データへのフィットの計算も同様に行えるが、確率関数は限定ゆびすまの時と異なるものとなるので注意が必用。

6 守り手のデータを活用する

今までの話は全て攻め手の戦略について注目してきた。データのフィットも攻め手のとった行動に関してのデータのみで行えるものである。しかし守り手の行動を観察することで、「相手の意思決定主体を誤って認識してしまう」という現象が無意識に起こっているか自覚の下で起こっているのかを判断することができるかもしれない。

攻め手がわのデータを用いて、攻め手が守り手の指を別々のものとしてみているというモデルが支持された状況を考える。この時、守り手のとった行動のデータから case1 のモデル、すなわち攻め手が意思決定主体を正しく認識している時のモデルが支持されたとすると、意思決定主体を誤ってしまうのは自分がせめてになった時のみであり、守り手になった時には相手が自分の意思決定主体が統一されていることを理解した上で行動していることを想定していることになる。これはつまり自分が守り手である時は、自分が攻め手に成った時には守り手の意思決定主体を正しく認識して行動するであろうと思っているということを意味する。この場合には意思決定主体の誤認は攻め手に成った時に無意識化でなされている現象だということができるであろう。

守り手がわのデータのフィットを見るのは攻め手がわのデータを用いたものと同様にして行う。(ゆえに同様に妥当性を考えないといけない。)

7 alternative models

ナッシュ均衡以外の解を考える。2 人完備情報ゲームに対して適用できる解概念として、Osborn and Rubinshtein (1998) で提案された S(k) Equilibrium や McKelvey and Palfrey (1995) で提案された quantal response equilibrium

などがある。しかし、先のように攻め手のアクションとして、勝利できる相手の本数別に行動の組をくくった時点で、そのゲームは2人対称ゼロサムゲームとなる。具体的にそのゲームの利得表は以下のである。このような利得構造を

		守り		
		0	1	2
攻め	set1	(1,-1)	(0,0)	(0,0)
	set2	(0,0)	(1,-1)	(0,0)
	set3	(0,0)	(0,0)	(1,-1)

図1 利得表

持つゲームにおいては、既存の限定合理性に基づく均衡概念では、想定されるような set2 に偏った分布は道ミクことはできない。

従って、対抗するモデルとして考える限定合理性は、自分の行動のうちでどれとどれが無差別なものであるかを認識するそのあり方に置かれるべきであることわかる。ここである二つの行動が攻め手にとって無差別であるとは、その二つの行動を取った時の利得が全く同じであることを意味している。以上より本研究で対立するモデルは以下の4つとできる。

モデル1 相手主体を正しく認識し利得構造も正しく把握している

モデル2 相手主体を誤って認識しているが利得構造は正しく把握している

モデル3 相手主体は正しく認識しているが利得構造を誤認している

モデル4 相手主体を誤って認識し利得構造も誤認している

ここまで述べてきたのは利得構造は正しく把握しているモデル1とモデル2についてである。以下では利得構造の誤認という意味での限定合理性をどのようにモデルに取り入れるかを考えるが、誤差項を加えるようなランダムな誤認の仕方ではとる行動の分布に偏りが出る状況を記述するのは難しく、またゆびすまのような単純な利得構造において実際の報酬に対する誤認というのは考えにくい。ここで述べている「利得構造の誤認」はつまるところ「どの戦略が同程度に勝ちやすく、勝ちやすさにはどの程度違いがあるか」についての誤認である。そこで、ゆびすまの攻め手のとる行動が「宣言」と「あげる指の数」という二つの要素から構成されていることに注目し、別の方向性からある戦略の勝ちやすさについての誤認が生まれる状況を記述することを考える。

通常ゆびすまのケースを考える。この時攻め手の行動は以下のような表の上に置くことができる。これは列方向には

		宣言数による組み分け				
		p_0	p_1	p_2	p_1	p_0
指の数による 組み分け	r_0	[0,0]	[1,0]	[2,0]		
	r_1		[1,1]	[2,1]	[3,1]	
	r_0			[2,2]	[3,2]	[4,2]

図2 組み分け

宣言数で分割し、行方向にはあげるゆびの数で分割したものである。本来は宣言数とゆびの数のペア、このケースでは9ペア、が攻め手の行動として定義されているが、このような複数の手段の組み合わせを一つの行動として見なすのは人間の認知上困難であることが考えられる。その時とられていると認知的過程として、組み合わせを構成する各要素について別々に決定し、それぞれがうまく連携するように行動を決める過程をここでは想定する。すなわち、「勝ちやすい宣言数の選び方」と「勝ちやすいゆび数の選び方」がお互いに矛盾しないように決定され、その積として図2上に配置された本来の戦略が取られる確率が決定される、という状況を考えてみる。

まずモデル3のケース、すなわち相手主体を正しく認識し相手の取りうる行動は $\{0, 1, 2\}$ のいずれかであると考えているケースを扱う。宣言数について取りうる行動の組は $\{0, 1, 2, 3, 4\}$ の5つであるが、対称性より守り手は指の本数0, 2については無差別であるので、攻め手の宣言数については $\{0, 4\}$, $\{1, 3\}$, $\{2\}$ がそれぞれ無差別な行動の組となる。したがって、均衡下に置いてはそれぞれに置く確率が等しくなるはずなので、図2のように各宣言数に均衡確率 (p_0, p_1, p_2) を置く。ただし $2p_0 + 2p_1 + p_2 = 1$ であることに注意する。あげる指の数についても対称性を利用して同様に考えるこ

とができる。すなわちあげる指の数については $\{0, 2\}, \{1\}$ が無差別な行動の組であり、それぞれに図 2 のように確率 (r_0, r_1) を割り振る。ここで $2r_0 + r_1 = 1$ であることに注意する。

宣言数サイドとあげる指の数サイドで別々に混合戦略ナッシュ均衡を求めることを考える。ただし、内部でお互いの均衡確率を利用するスキームを用い、お互いが連携するような均衡をいっぺんに求めることを考える。宣言数サイドから具体的に見ていく。守り手の各行動に紐付いた期待利得は以下のように書ける。

$$\text{expected payoff when the saver chooses 0} = p_0 + \frac{r_1}{r_0 + r_1}p_1 + r_0p_2$$

$$\text{expected payoff when the saver chooses 1} = \frac{r_0}{r_0 + r_1}p_1 + r_1p_2 + \frac{r_0}{r_1 + r_0}p_1$$

$$\text{expected payoff when the saver chooses 2} = r_0p_2 + \frac{r_1}{r_1 + r_0}p_1 + p_0$$

8 考えたいこと

- 「相手の意思決定に関する勘違いが最適な選択からずれた行動を誘導する」事例としては面白いと思うが、ゆびすまという文脈が陳腐すぎる気がする。どのような書き方をすれば面白くなるか。
- 限定ゆびすまでは通常のゆびすまと異なり正しい認識の下でも最適な戦略の導く行動の分布に偏りがある。この違いを面白い形で利用できないか。
- 無差別な戦略の組をくくる際の限定合理性を考えた既存のモデルはあるか？