[INFO-F409] Learning Dynamics
First assignment

BUI QUANG PHUONG Quang Linh
Université libre de Bruxelles - ULB ID : 000427796
MA1 Computer Sciences

November 2018

Contents

1 The Hawk-Dove game

Conventions and notations

First of all, the Hawk-Dove game is modeled by the matrix presented in the Table 1.

	Hawk	Dove
Hawk	$\frac{V-D}{2}$ $\frac{V-D}{2}$	0 V
Dove	V 0	$\frac{V}{2}-T$ $\frac{V}{2}-T$

Table 1 – Payoff matrix of the Hawk-Dove game

The different actions of a player $i \in \{1, 2\}$ are denoted by the set $\mathcal{A} = \{H, D\}$ where H is the hawk action and D the dove action. Moreover, to denote the different actions payoff, we need an utility function of those actions. This utility function is then written :

$$u_i(a_i, a_{-i})$$

such that a_i is an action of player i and a_{-i} is the action of the other player where player i choses a_i. For instance, $u_1(Hawk, Dove) = V$ and $u_2(Hawk, Dove) = 0$).

In the case of mixed strategies, the notion of **expected value** of a payoff function is used and is written in the general case :

$$U_i(p_1, ..., p_k) = p_k u(a_k)$$

where i is a player, p_k is the probability that the other player chose the action a_k. In the case of the Hawk-Dove game, the expected value formula would be written such that $k = 2$ because it exists only 2 actions, i.e. :

$$U_i(p_1, p_2) = p_1 u(a_1) + p_2 u(a_2)$$

Furthermore, to find Nash equilibria, best responses have to be found. Those one will be highlighted in red for the line player considered as player one and green for the column player considered as player two.

1.1 Question 1 - Nash equilibrium

Statement

Find all the (mixed strategy) Nash equilibria of this game. How do the results change when the order of the parameters V, D and T is changed (V>D, D>T, etc.)?

Sign of V, D and T

Before starting the analysis, we should remind that V and D are representing respectively the fitness value of winning resources in fight and costs of injury thereby it would not make sense having negative values. V and D are then always positive values. Same for T which is the cost of wasting time. The time wasted is obviously always a positive value. To summarize,

$$V, D, T \geq 0$$

1.1.1 First case : $V > D$

In the first case, we consider that $V > D$. Therefore, we know that the value of $u_i(Hawk, Hawk) = \frac{V-D}{2}$ will be strictly positive. In this case, the choice of both players are quiet easy. As illustrated in Table 2, the Nash equilibrium is $(Hawk, Hawk) \in \mathcal{A}$. Indeed, both players are chosing *Hawk* because if one of them is switching to *Dove* then $u_i(Hawk, Hawk)$ which was a positive value becomes $u_1(Dove, Hawk)$ for player 1 or $u_2(Hawk, Dove)$ for player 2 which equals zero. Thus, they obviously prefer to pick $(Hawk, Hawk)$.

$$\boxed{\text{Nash equilibrium when } (V > D) = \{(Hawk, Hawk)\}}$$

	Hawk	Dove
Hawk	$\frac{V-D}{2} > 0$; $\frac{V-D}{2} > 0$	V ; 0
Dove	0 ; V	$\frac{V}{2} - T$; $\frac{V}{2} - T$

Table 2 – Nash equilibrium/Best responses in case 1

1.1.2 Second case : $V = D$

The second case considers the same value for V and D which means that $\frac{V-D}{2} = 0$. The Table 3 shows the Nash equilibria for that case. We can see that there doesn't exist only one Nash equilibrium, but three. $(Hawk, Hawk)$ stays a Nash equilibrium in this case, but $(Dove, Hawk)$ and $(Hawk, Dove)$ are now also Nash equilibria. In other words, if player 2 is chosing *Hawk*, the best response to it is whether $(Hawk, Hawk)$ or $(Dove, Hawk)$ because $u_1(Hawk, Hawk) = u_1(Dove, Hawk) = 0$. On the other hand, if player 2 is chosing *Dove*, the best response is only $(Hawk, Dove)$ since $u_1(Hawk, Dove) > u_1(Dove, Dove) \equiv V > \frac{V}{2} - T$. By symmetry, it is also valid for the opposite case where player 1's choice is known, so that the best responses are (symmetrically) equivalent which means $(Hawk, Hawk)$,$(Hawk, Dove)$ and $(Dove, Hawk)$ are the best responses for player 2.

$$\boxed{\text{Nash equilibria when } (V = D) = \{(Hawk, Hawk), (Hawk, Dove), (Dove, Hawk)\}}$$

	Hawk	Dove
Hawk	$\frac{V-D}{2} = 0$; $\frac{V-D}{2} = 0$	V ; 0
Dove	0 ; V	$\frac{V}{2} - T$; $\frac{V}{2} - T$

Table 3 – Nash equilibria/Best responses in case 2

1.1.3 Third case : $V < D$

Finally, we are now considering that $V < D$. Therefore, $u_i(Hawk, Hawk) = \frac{V-D}{2}$ is now strictly a negative value. The Table 4 is showing the Nash equilibria for this case. Compared to the previous case, $(Hawk, Hawk)$ is not a Nash equilibrium anymore but $(Hawk, Dove)$ and $(Dove, Hawk)$ keep there. Indeed, when player 2 is playing *Hawk*, player 1 would play *Dove* since $u_1(Hawk, Hawk) < u_1(Dove, Hawk) \equiv \frac{V-D}{2} < 0$. In the case where player 2 is playing *Dove*, it doesn't change, it means that $u_1(Hawk, Dove) = V$ is still greater than $u_1(Dove, Dove) = \frac{V}{2} - T$. As reasoned in the previous case, this is also valid when player 2 is depending of player 1's choice thanks to the matrix's symmetry.

$$\boxed{\text{Nash equilibria when } (V < D) = \{(Hawk, Dove), (Dove, Hawk)\}}$$

	Hawk	Dove
Hawk	$\frac{V-D}{2} < 0$; $\frac{V-D}{2} < 0$	V ; 0
Dove	0 ; V	$\frac{V}{2} - T$; $\frac{V}{2} - T$

Table 4 – Nash equilibria/Best responses in case 3

1.1.4 What about the value of T ?

As you can see, the different cases are not depending of the value of T. In each case, $(Dove, Dove)$ will never be a Nash equilibrium since that $\frac{V}{2} - T$ will always be smaller than V (as reminder, V and $T > 0$), then $(Dove, Dove)$ will never be the best response.

1.1.5 Mixed strategy

To find the mixed strategy Nash equilibrium, we have to take into account the probabilities p and q which are respectively the probabilities of playing a certain action for player 1 and player 2. The general probabilities of a 2x2 game's matrix is shown in the Figure 1.

	L (q)	R (1-q)
T (p)	*pq*	*p(1-q)*
B (1-p)	*(1-p)q*	*(1-p)(1-q)*

Figure 1 – Probabilities of a 2x2 game's matrix

Thus, we have to compute the expected value of each case which means the player 1's expected payoff for the pure strategy $Hawk(p)$ and $Dove(1-p)$, same for player 2 i.e $Hawk(q)$ and $Dove(1-q)$.
Therefore, for player 1 we obtain those equations [1] :

$$U_1^H = q \cdot (\frac{V-D}{2}) + (1-q) \cdot V \tag{1}$$

$$U_1^D = (1-q) \cdot (\frac{V}{2} - T) \tag{2}$$

To find the value of probability q and to know which strategy is the best in which case, we make an inequality between the two equations to isolate q (here, we deal with (1) < (2) which means that

[1] U_i^j means the player i's expected payoff for the pure strategy j

we will find the condition where playing $Hawk$ is less optimal than playing $Dove$):

$$q \cdot (\frac{V-D}{2}) + (1-q) \cdot V < (1-q) \cdot (\frac{V}{2} - T) \tag{3}$$

$$\frac{qV}{2} - \frac{qD}{2} + V - qV < \frac{V}{2} - T - \frac{qV}{2} + qT$$

$$-\frac{qD}{2} < -\frac{V}{2} - T + qT$$

$$\frac{qD}{2} + qT > \frac{V}{2} + T$$

$$q \cdot (\frac{D}{2} + T) > \frac{V}{2} + T$$

$$q > \frac{\frac{V}{2} + T}{\frac{D}{2} + T}$$

$$\boxed{q > \frac{V+2T}{D+2T}} \tag{4}$$

Therefore, when $q > \frac{V+2T}{D+2T}$, playing $Dove$ is the best choice because $U_1^H < U_1^D$. On the contrary, when $q < \frac{V+2T}{D+2T}$ playing $Hawk$ is then the best choice because $U_1^H > U_1^D$.

By symmetry, we can deduct the equations for player 2 :

$$U_2^H = p \cdot (\frac{V-D}{2}) + (1-p) \cdot V \tag{5}$$

$$U_2^D = (1-p) \cdot (\frac{V}{2} - T) \tag{6}$$

therefore, the value of p will be the same as q too :

$$p = \frac{V+2T}{D+2T} \tag{7}$$

which means that :

$$p = q = \frac{V+2T}{D+2T} \in [0,1] \tag{8}$$

To summarize, for player one (two), when :

- $q(p) > \frac{V+2T}{D+2T}$, the best response set is $Dove$ or $p(q) = 0$
- $q(p) = \frac{V+2T}{D+2T}$, the best response set is the set of all $p(q)$ values in [0,1]
- $q(p) < \frac{V+2T}{D+2T}$, the best response set is $Hawk$ or $p(q) = 1$

1.2 Question 2 - Mixed strategy drawing

Statement

Under which conditions does displaying become more beneficial than escalating? Draw the set of all mixed strategies.

Graph of all mixed strategies

Based on the calculations of probabilities p and q in the subsubsection 1.1.5, the graph of all mixed strategies for the Hawk-Dove game is presented in the Figure 2. The green lines are the set of player 1's best mixed strategies while the red ones are the set of player 2's best mixed strategies. The blue circles are the mixed strategies possible Nash equilibria of the game. Thus, we have three possible Nash equilibria that are defined by the sets $\{(1,0);(1,0)\},\{(0,1),(0,1)\}$ and $\{(\frac{V+2T}{D+2T},\frac{V+2T}{D+2T}),(\frac{V+2T}{D+2T},\frac{V+2T}{D+2T})\}$. However, the only one is $\{(\frac{V+2T}{D+2T},\frac{V+2T}{D+2T}),(\frac{V+2T}{D+2T},\frac{V+2T}{D+2T})\}$.

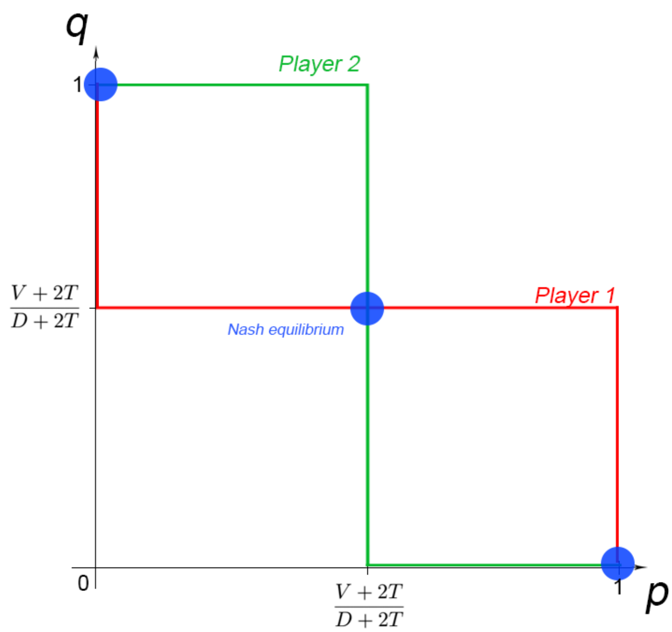

Figure 2 – Mixed strategies graph of the Hawk-Dove game

Remark Note that $\frac{V+2T}{D+2T}$ has to be a value between 0 and 1 otherwise it wouldn't be a valid probability. Therefore, to respect this condition, necessarily $V \leq D$.

To conclude, as calculated in the subsubsection 1.1.5 and represented in the Figure 2, if a player i plays $Hawk$ with a probability $< \frac{V+2T}{D+2T}$ then the opponent player would play $Hawk$ because the expected value for playing $Hawk$ is higher than the expected value of playing $Dove$. Similarly, playing $Dove$ is then an optimal choice when the player i chooses $Hawk$ with a probability $> \frac{V+2T}{D+2T}$.

2 Which social dilemma ?

Statement

Player A knows he's confronted with one of three social dilemma's; a prisoner's dilemma, a snowdrift game or stag-hunt game (see above). In each game he needs to decide whether to cooperate (C) or defect (D), yet he is not sure in which he actually is. He's sure that each game is equally likely. The other player, player B, knows in which game he's playing. Determine the pure Nash equilibria using the Bayesian game analysis discussed in the course.

Data (payoff matrix) of each game

The payoff matrix of each game is given in the statement and is taken again in the Figure 3.

Prisonners dilemma

	C	D
C	2,2	0,5
D	5,0	1,1

Stag-Hunt game

	C	D
C	5,5	0,2
D	2,0	1,1

Snowdrift game

	C	D
C	2,2	1,5
D	5,1	0,0

Figure 3 – Payoff matrices of the prisonner's dilemma, stag-hunt and snowdrift game

Resolution

The problem presented here is a *Bayesian* game problem where a player A doesn't know which game will he play while the second player B knows it. For all cases, the player A has the choice between cooperating (C) or defecting (D). To find the Nash equilibria in such a problem, the first thing to do is to enumerate all the possibles combinations of player A's choice in each game. Let's define the set of all player i's action's $\mathcal{A}_i$. In that case, $\mathcal{A}_A$ will then take those values : $\mathcal{A}_A = \{(C,C,C),(C,C,D),(C,D,C),(D,C,C),(C,D,D),(D,C,D),(D,D,C),(D,D,D)\}$. Concerning $\mathcal{A}_B$, player B has only two choices because he knows in which game he's playing, so it's simple as $\mathcal{A}_B = \{C,D\}$.

We know have all the possible actions for both player, the next step is calculating the payoff values for each combinations between an action of player A and player B. To do that, we represent $\mathcal{A}_A$ in the columns of the matrix and $\mathcal{A}_B$ in the lines of the matrix. Moreover, we know that each game is equally likely which means that every game has a probability of $\frac{1}{3}$ to be played by player A. Table 5 is presenting the payoff matrix of player A for the current approached Bayesian problem. Note that the payoff values are multiplied by 3 to avoid fractions to make the matrix more clear.

	CCC	CCD	CDC	DCC	CDD	DCD	DDC	DDD
C	9	8	4	7	3	6	2	1
D	12	7	11	8	6	3	7	2

Table 5 – Payoff matrix of player A for the Bayesian game problem

Let's now evaluate the best responses for player A (thanks to the matrix just shown above) and for player B. To find a Nash equilibrium, the best responses of both players should match. Best responses are illustrated in Figure 4 as well as the matched ones which mean the pure Nash equilibria of the problem.

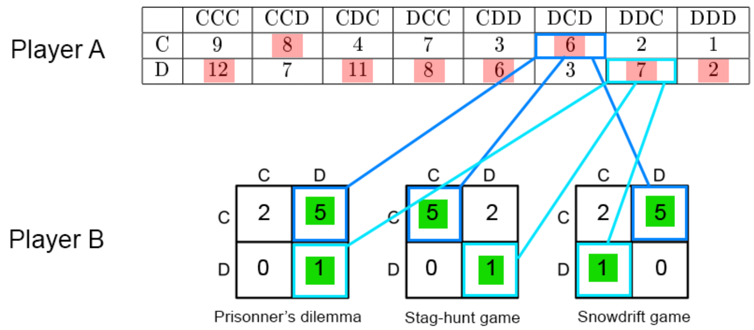

Figure 4 – Nash equilibria and best responses of the problem "Which social dilemma?" – the best responses for player A are highlighted in red, the player B's ones are highlighted in green. The Nash equilibria are brought out in blue and cyan.

To conclude, there exists 2 pure Nash equilibria in this problem. The first one is (C, DCD) (i.e player A cooperating in the second game only while player B is cooperating) and the second one is (D, DDC) (i.e player B cooperating in the last game only while player B is defecting).

3 Games in finite population

Reminder of notations

First, let's recap all the notations which will be used further. Denote by

- Z, the size of the population
- m, the number of rounds of the prisoner's dilemma
- p_{ij}, the *fermi* function which i is an individual and j a population
- β, the intensity of selection
- $\{C, D, TFT, RANDOM\}$, respectively "always plays C", "always plays D", "play C first and then chose whichever choice the second player chose last time", "play random whether C or D each time" the 4 possibles strategies

The payoff matrix of a 2-person prisoner's dilemma is illustrated below in Figure 5

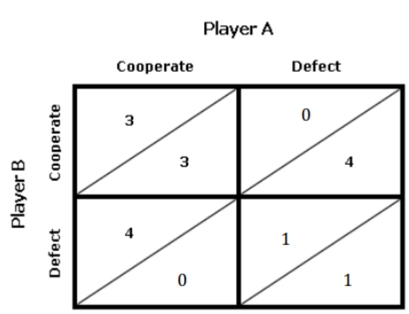

Figure 5 – Payoff matrix of a 2-person prisoner's dilemma

3.1 Question 1 - Expected payoff of strategies

Statement

Complete the notebook with the payoff table and copy in your answer document the expected payoff of each strategy when facing against each other strategy. Use the provided python notebook ("2IPD.ipnb") to calculate the fixation probabilities of each strategy for $\beta = 10$ and paste the stationary distribution in your answer document.

3.1.1 Expected payoff of each strategy

Here are the calculations of each strategy facing against each other strategy with $m = 10$ for a player i. As reminder, $U_i(a_i, a_{-i})$ denotes the expected payoff value of the strategy a_i against the strategy a_{-i} for m rounds such that payoffs for

- Reward for mutual cooperation (C vs C) $\equiv R = 3$
- Punishment for mutual defection (D vs D) $\equiv P = 1$
- Suckers payoff for unilateral cooperation (C vs D) $\equiv S = 0$
- Temptation to defect (D vs C) $\equiv T = 4$

$$U_i(C,C) = R \cdot m = 30 \tag{9}$$
$$U_i(C,D) = S \cdot m = 0 \tag{10}$$
$$U_i(C,TFT) = R \cdot m = 30 \tag{11}$$
$$U_i(C,RANDOM) = \left[\frac{1}{2} \cdot R + \frac{1}{2} \cdot S\right] \cdot m = (\frac{1}{2} \cdot 3 + \frac{1}{2} \cdot 0) \cdot 10 = 15 \tag{12}$$

$$U_i(D,C) = T \cdot m = 40 \tag{13}$$
$$U_i(D,D) = P \cdot m = 10 \tag{14}$$
$$U_i(D,TFT) = T + P \cdot (m-1) = 4 + 9 = 13 \tag{15}$$
$$U_i(D,RANDOM) = \left[\frac{1}{2} \cdot T + \frac{1}{2} \cdot P\right] = (\frac{1}{2} \cdot 4 + \frac{1}{2} \cdot 1) \cdot 10 = 25 \tag{16}$$

$$U_i(TFT,C) = R \cdot m = 30 \tag{17}$$
$$U_i(TFT,D) = S + P \cdot (m-1) = 0 + 9 = 9 \tag{18}$$
$$U_i(TFT,TFT) = R \cdot m = 30 \tag{19}$$
$$\begin{aligned} U_i(TFT,RANDOM) &= \left[\frac{1}{2} \cdot R + \frac{1}{2} \cdot S\right] + \left[\frac{1}{4} \cdot R + \frac{1}{4} \cdot S + \frac{1}{4} \cdot T + \frac{1}{4} + \cdot P\right] \cdot (m-1) \\ &= \frac{5}{2} + \left[0 + \frac{3}{4} + 1 + \frac{1}{4}\right] \cdot 9 = 19.5 \end{aligned} \tag{20}$$

$$U_i(RANDOM, C) = \left[\frac{1}{2}R + \frac{1}{2} \cdot T\right] \cdot m = \left[\frac{1}{2} \cdot 3 + \frac{1}{2} \cdot 4\right] \cdot 10 = 35 \tag{21}$$

$$U_i(RANDOM, D) = \left[\frac{1}{2} \cdot S + \frac{1}{2} \cdot P\right] \cdot m = \left[\frac{1}{2} \cdot 0 + \frac{1}{2} \cdot 1\right] \cdot 10 = 5 \tag{22}$$

$$\begin{aligned} U_i(RANDOM, TFT) &= \left[\frac{1}{2} \cdot R + \frac{1}{2} \cdot T\right] + \left[\frac{1}{4} \cdot R + \frac{1}{4} \cdot S + \frac{1}{4} \cdot T + \frac{1}{4} \cdot P\right] \cdot (m-1) \\ &= \frac{7}{2} + \left[0 + \frac{3}{4} + 1 + \frac{1}{4}\right] \cdot 9 = 21.5 \end{aligned} \tag{23}$$

$$\begin{aligned} U_i(RANDOM, RANDOM) &= \left[\frac{1}{4} \cdot R + \frac{1}{4} \cdot S + \frac{1}{4} \cdot T + \frac{1}{4} \cdot P\right] \cdot m \\ &= \left[0 + \frac{3}{4} + 1 + \frac{1}{4}\right] \cdot 10 = 20 \end{aligned} \tag{24}$$

To summerize and to make it more clear, Table 6 recaps all the expected payoff values for each strategy against every other strategy for $m = 10$.

Strat 1 \ Strat 2	C	D	TFT	RANDOM
C	30	0	30	15
D	40	10	13	25
TFT	30	9	30	19.5
RANDOM	35	5	21.5	20

Table 6 – Expected payoff values of 2-players prisoner's dilemma when $m = 10$

3.1.2 Fixation probabilities

The fixation probabilities are given in the matrix below. Respectively from left to right : C, D, TFT and $RANDOM$.

$$\begin{bmatrix} 2.65746147e-10 & 9.99999973e-01 & 2.68402668e-08 & 2.65744154e-10 \end{bmatrix}$$

3.1.3 Stationary distribution

The stationary distribution when $\beta = 10$ is presented in the Figure 6

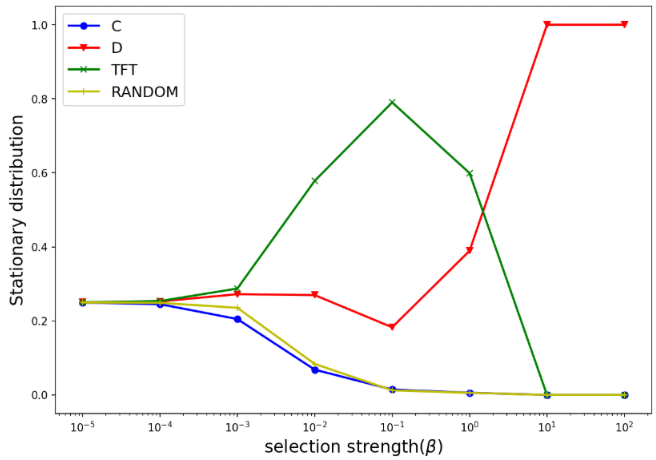

Figure 6 – stationary distribution when $\beta = 10$ and $m = 10$

3.2 Question 2 - Stationary distribution for different β

Statement

Draw the stationary distribution for these 4 strategies when the mutation rate $\mu \to 0$ for different values of β. Paste the graph generated by the provided python notebook in your answer document

3.2.1 Result

The mutation rate μ is denoted by the variable $drift$ in the Python notebook. When setting it at $\beta = 0.01$ (which means $\mu \to 0$), the result of the stationary distribution for different values of β are given below :

$\beta = 10$ **:**

$$\begin{bmatrix} 2.65746147e-10 & 9.99999973e-01 & 2.68402668e-08 & 2.65744154e-10 \end{bmatrix}$$

$\beta = 5$ **:**

$$\begin{bmatrix} 7.10392023e-06 & 9.99268310e-01 & 7.17495943e-04 & 7.09002624e-06 \end{bmatrix}$$

$\beta = 1$ **:**

$$\begin{bmatrix} 0.00594699 & 0.38929365 & 0.5991636 & 0.00559576 \end{bmatrix}$$

What we can see from that results is that more the value of β is small, more the stationary distribution of the strategies are high. It grows quite exponentially.

3.2.2 Graph

Concerning the graph of the stationary distribution, whatever the value of β it doesn't change anything. Therefore, the graph for every β is given by the Figure 6.

3.3 Question 3 - Best strategy

Statement

Which strategy is more successful ?

3.3.1 Best strategy

The best strategy is "Always play D". In the graph presented in Figure 6, we can see that more the selection strength β is high, more the strategy D is increasing which shows that more important the payoff is in the selection process more D is benefit.

3.4 Question 4 - $m = 1$

Statement

What happens if $m = 1$?

3.4.1 New expected payoffs

Of course, modifying m impacts the values of the expected payoff. Here are presented these values when $m = 1$. Obviously, it will also impacts the fixation probabilities values and the stationary distribution. It will be shown in the subsubsection 3.4.2 and subsubsection 3.4.3.

Strat 1 \ Strat 2	C	D	TFT	RANDOM
C	3	0	3	1.5
D	4	1	4	2.5
TFT	3	0	3	1.5
RANDOM	2.5	0.5	2.5	2

Table 7 – Expected payoff values of 2-players prisoner's dilemma when $m = 1$

3.4.2 Fixation probabilities

The fixation probabilities where $m = 1$ are given in the matrix below. Respectively from left to right : C, D, TFT and $RANDOM$.

$$\begin{bmatrix} 3.04072135e-17 & 1.00000000e+00 & 1.04488881e-17 & 2.61829550e-27 \end{bmatrix}$$

3.4.3 Stationary distribution

The stationary distribution when $\beta = 10$ and $m = 1$ is presented in the Figure 7

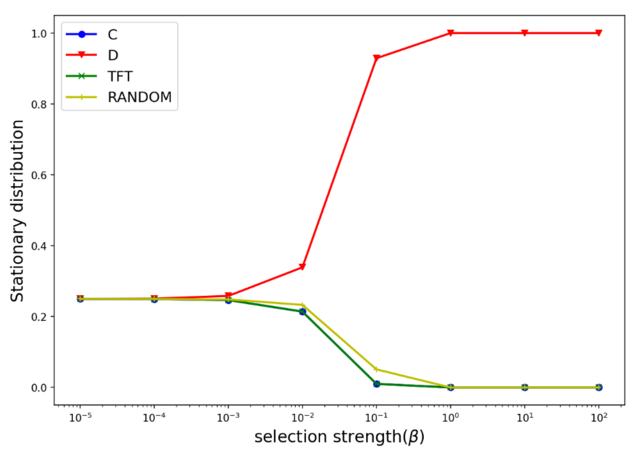

Figure 7 – stationary distribution when $\beta = 10$ and $m = 1$

3.4.4 Best strategy

Therefore, we can also find the best stragegy where $m = 1$ thanks to the previous results. However, the best strategy is still "Always play D" as observed in the graph. That's the only strategy where the strategy's stationary distribution increases when the intensity of selection increases.

3.5 Question 5 - Moran process

Statement

Implement the Moran process algorithm. Repeat the simulation 10 times for 10^3 *generations.*

3.5.1 Result

The stationary distribution given by the simulation is illustrated in the Figure 8. Of course, we have to keep in mind that it's not the only possible graph because it's a simulation with random values. However, the behaviour of all these graphs are the same. In this case, the behaviour of the graph is beneficial for the "TFT" strategy. Compared to the graph obtained in the question 2, we have then a different optimal strategy which was "Only defect" and becomes "TFT" here.

In theory, "only defect" is the optimal strategy as shown before because playing only D will never make the other player get more payoff than him but in practice, it seems that "TFT" is the best strategy to adopt. To explain this, "TFT" is a win-win strategy and applying this strategy will never make the player lose. "TFT" is a strategy that promotes the cooperating strategy and punish the defecting strategy. Thus, "TFT" will force the other player to cooperate if he wants higher payoff, i.e if player 2 plays only defect, he knows that the player 1 who plays "TFT" will make him win only 1 per round while he could win 3 if he cooperates. So naturally, the second player will cooperate, then "TFT" is the most benefit strategy and allows both player to have good payoffs. This phenomenon is also observable in the Figure 8, we see that when the distribution of D decreases, the "TFT" one is increasing and vice versa.

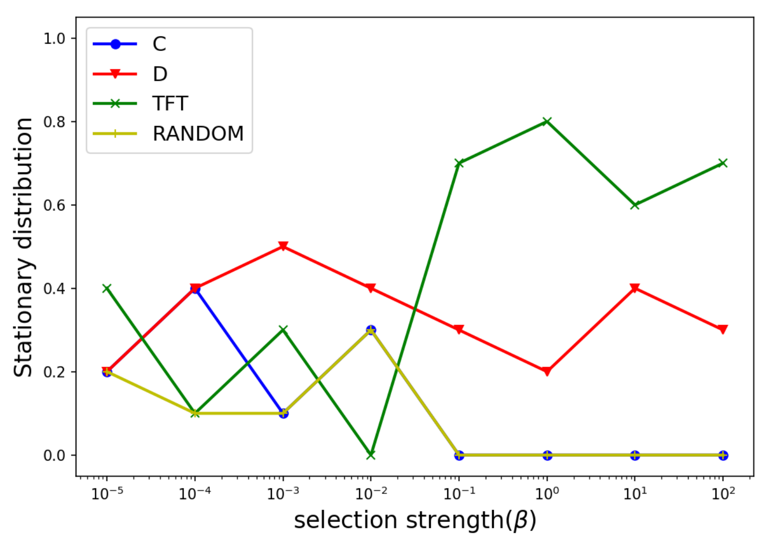

Figure 8 – stationary distribution with Moran process algorithm

3.5.2 Question 6 - Mutation rate

Statement

What happens if the mutation rate is bigger than zero?

3.5.3 Result

The value of μ put at 10^{-3} gives us the result shown in Figure 9. What we can see from the graph is that the same general appearance seems to appear even we have some randomness. Indeed, the "TFT" strategy is still the best one followed by "Only Defect" and the 2 others. Nevertheless, this randomness decreases a bit the gap between the different strategies. Thus, we can sometimes have a different result between those strategies, for instance when $\beta = 10$, the D strategy has a lower stationary distribution than C while it never happens in the previous section when the randomness was null. However, this randomness is not enough large to totally modify the general behaviour and the optimal strategy.

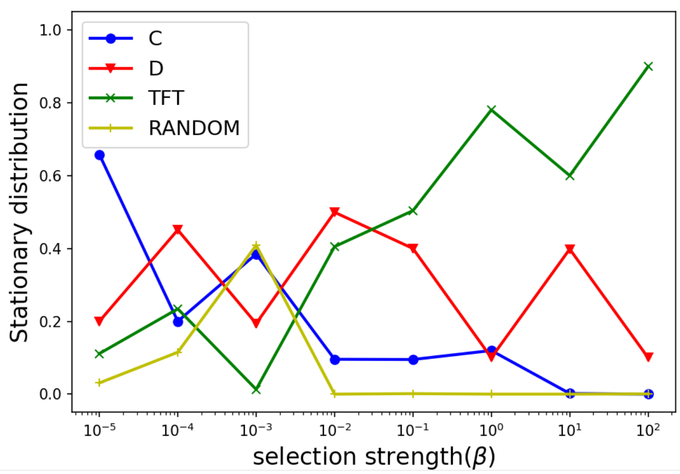

Figure 9 – stationary distribution with $\mu > 0$