MODELLING DECISION TIMES – ESTHER RODRIGO

July 30, 2018

The drift difussion model (DDM) is a type of sequential sampling model (SSM) which, until now, has been uniquely applied in impulsive / non-deliberated two-choice decisions.

These SSM models, and thus the DDM, assume that decisions are made by accumulating information until a threshold is reached. In order to model this accumulation of knowledge, DDM uses a random walk in two dimensions, starting at point z and finishing when it reaches a decision threshold (a or 0). Here, the first two parameters to model this behaviour have been named: a and z. The other two significant parameters will be defined as t (non-decision time) and v (drift rate).

The starting point z can be understood as an initial bias due to previous experiences or intuitive notions. On the other hand, the boundaries (a) represent the minimum accumulated knowledge needed to make a decision, so harder decisions will have wider boundaries than simple choices.

The speed by which the information is accumulated is defined as v (drift rate). In past papers, the drift rate v has been defined positive when the accumulation of information led to the right decision (and negative in the opposite case) but, as decisions in cooperation-defection experiments are not right or wrong, v will be defined as positive when the accumulation of knowledge leads to a cooperative choice (and negative on the contrary).

Apart from the drift rate, there is a source of stochastic noise in the decision process. Using the same parameters v and z, this noise will lead to different final decisions and different final times. The variability in decision times will produce the characteristic distribution of RT (reaction time) and the two possible final decisions will bring errors and the corresponding error RT distribution.

One last parameter t_0 is intended to measure the non-decision time.

The equation which represents the probability distribution:

$$P(t|v,a,z) = \frac{\pi}{a^2} \exp\left(-vza - \frac{v^2 t}{2}\right) \times \sum_{k=1}^{\infty} k \exp\left(-\frac{k^2 \pi^2 t}{2a^2}\right) \sin\left(k\pi z\right)$$

https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5112760/
https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4928591/
http://star.psy.ohio-state.edu/wp/?page_id=169

1 DATA VISUALIZATION

1.1 Weak PD with fixed partners

After importing and processing the data, we are left with the following dataframe:

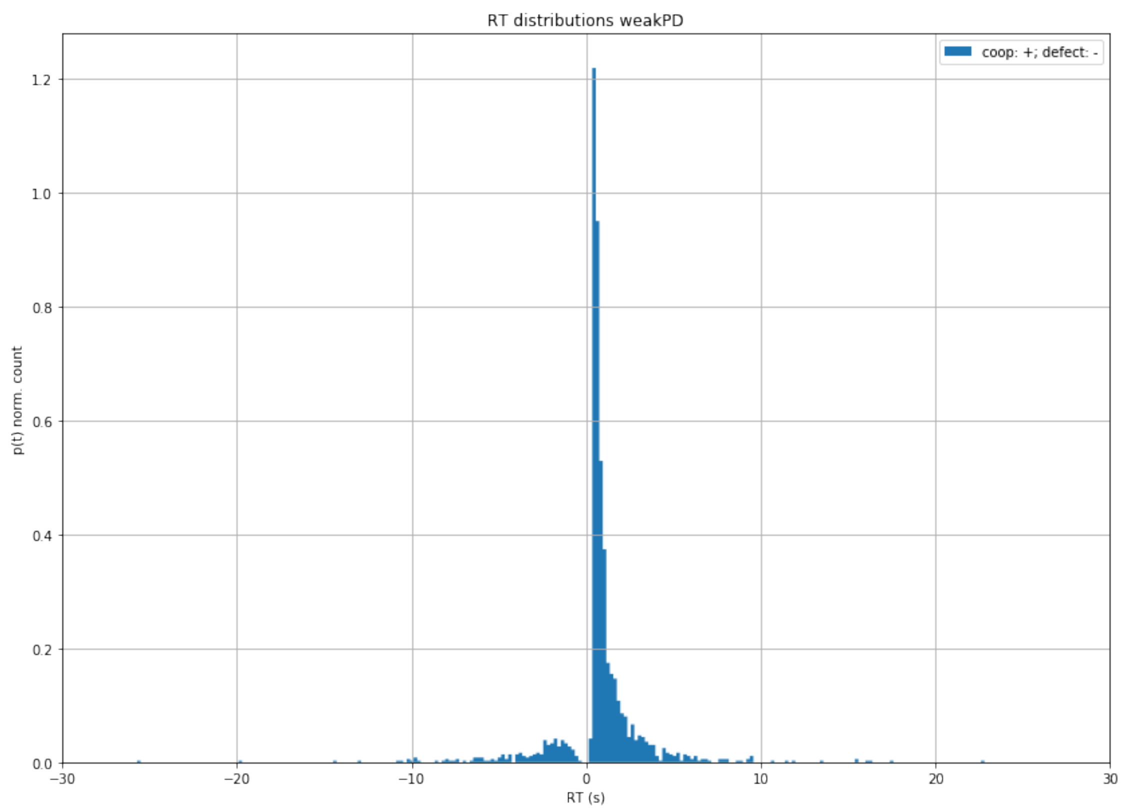
```
N of decisions are 1800
```

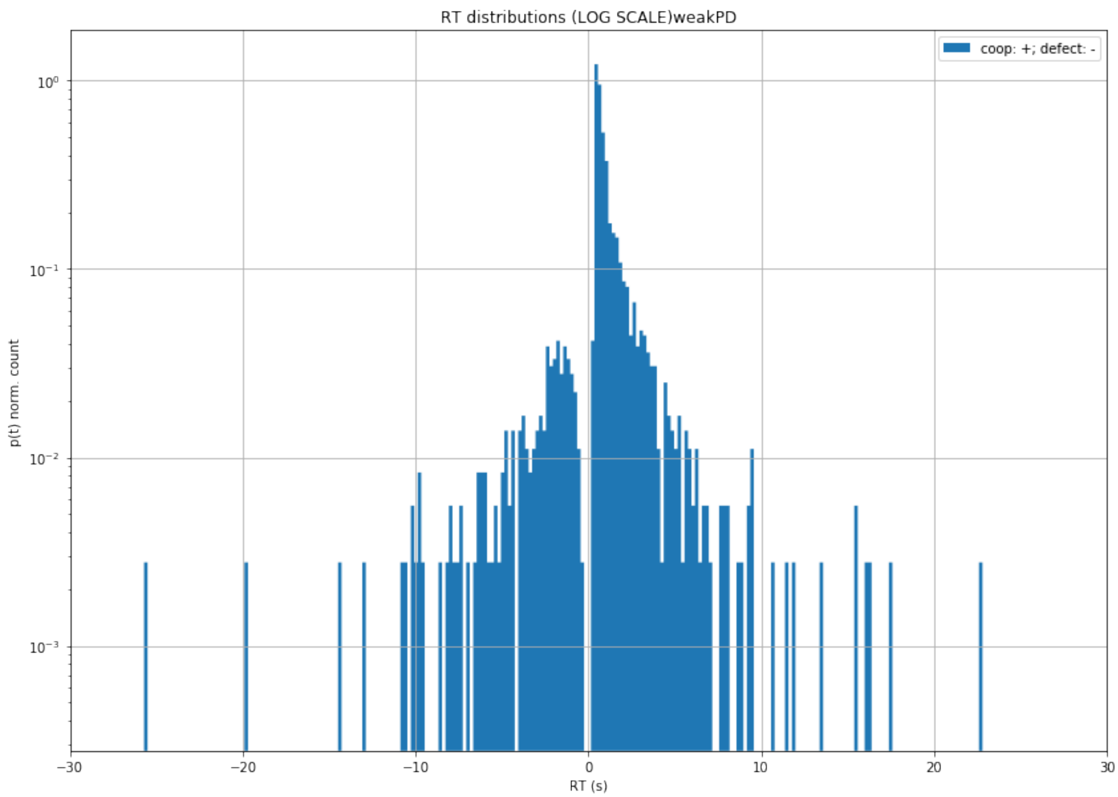
Out[68]:

	round	userid	response	payoff	rt
0	1	s3m7_usuario9	0	4	-4.452
1	1	s3m7_usuario18	1	0	9.443
2	1	s3m7_usuario7	1	3	2.213
3	1	s3m7_usuario10	1	3	8.101
4	1	s3m7_usuario13	1	3	5.043

1.1.1 RT distributions for cooperators and defectors + log scale

Collaborative game: much more decisions in the positive RT.





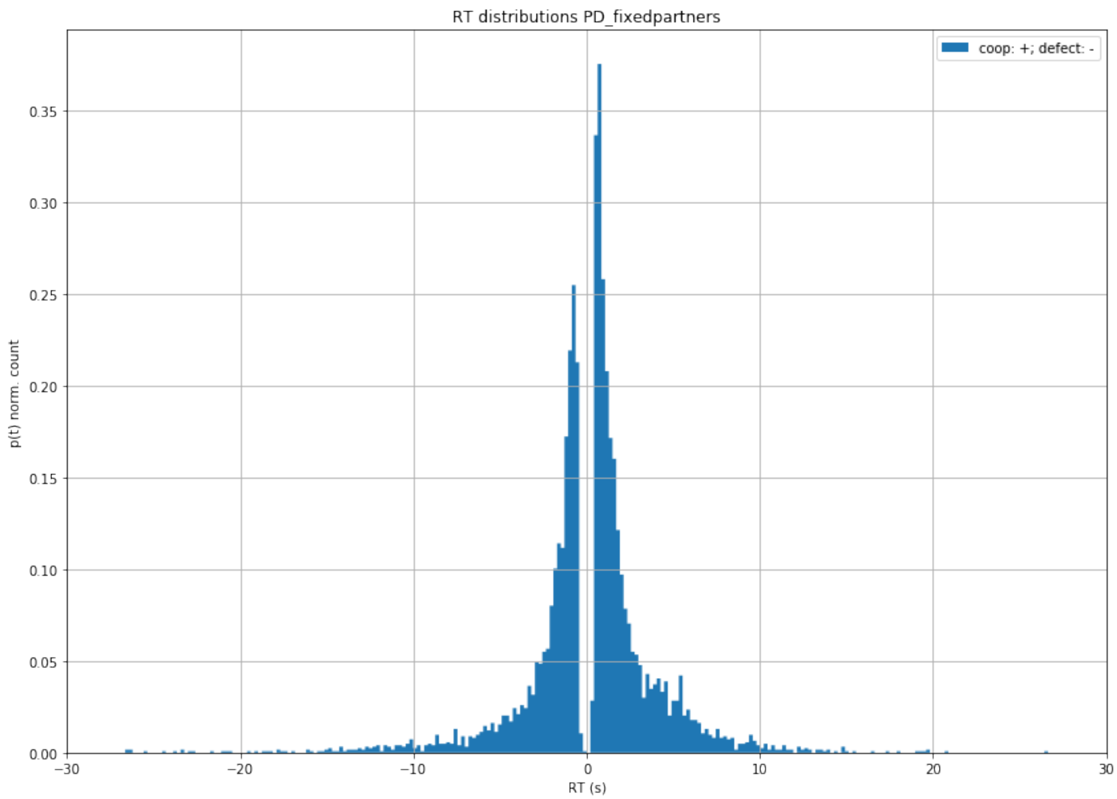
1.2 Fixed partners

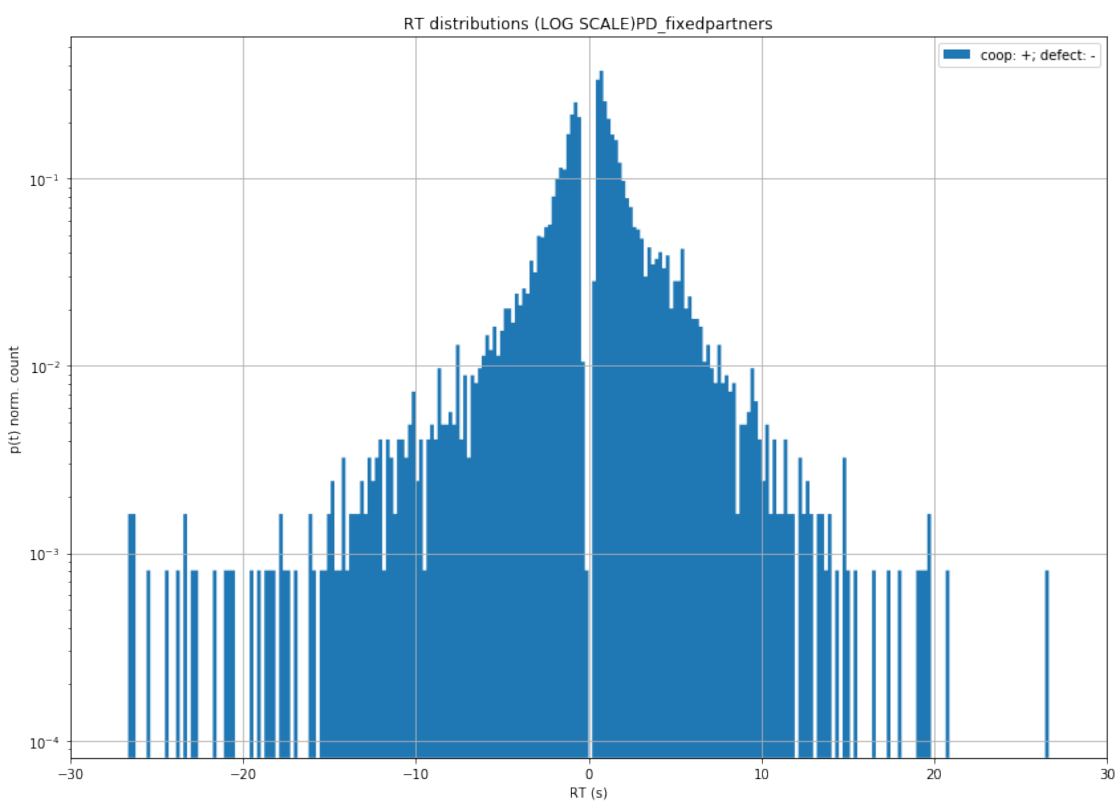
```
N of decisions are  5800
```

	round	userid	response	payoff	rt
0	1	s4m8_usuario2	0	4	-4.600
1	1	s4m8_usuario7	1	0	4.055
2	1	s4m8_usuario14	1	3	5.394
3	1	s4m8_usuario13	1	3	7.806
4	1	s4m8_usuario9	0	4	-15.067

1.2.1 RT distributions for cooperators and defectors + log scale

And the distributions of time are still pro-collaboration but less obvious than in the previous game.


RT distributions PD_fixedpartners
coop: +; defect: -
p(t) norm. count
RT (s)
0.00
0.05
0.10
0.15
0.20
0.25
0.30
0.35
−30
−20
−10
0
10
20
30


RT distributions (LOG SCALE)PD_fixedpartners
coop: +; defect: -
p(t) norm. count
RT (s)
10^{-1}
10^{-2}
10^{-3}
10^{-4}
−30
−20
−10
0
10
20
30

1.3 Changing partners

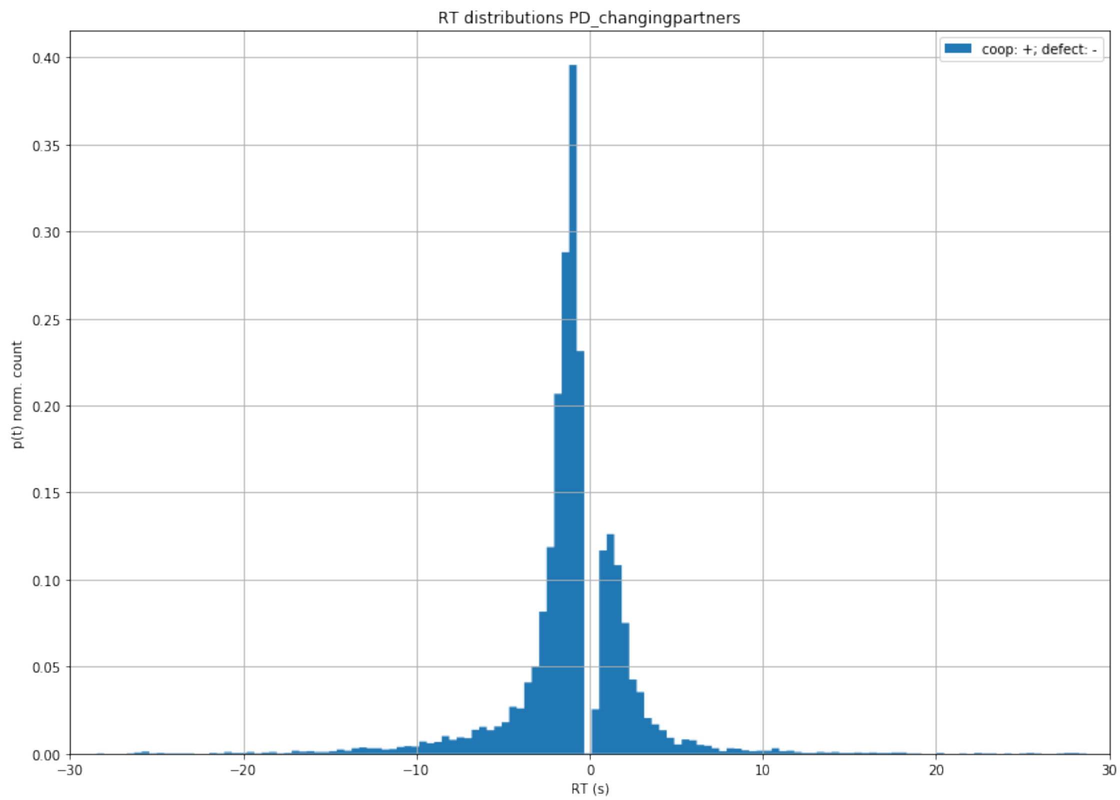
```
N of decisions are 9600
```

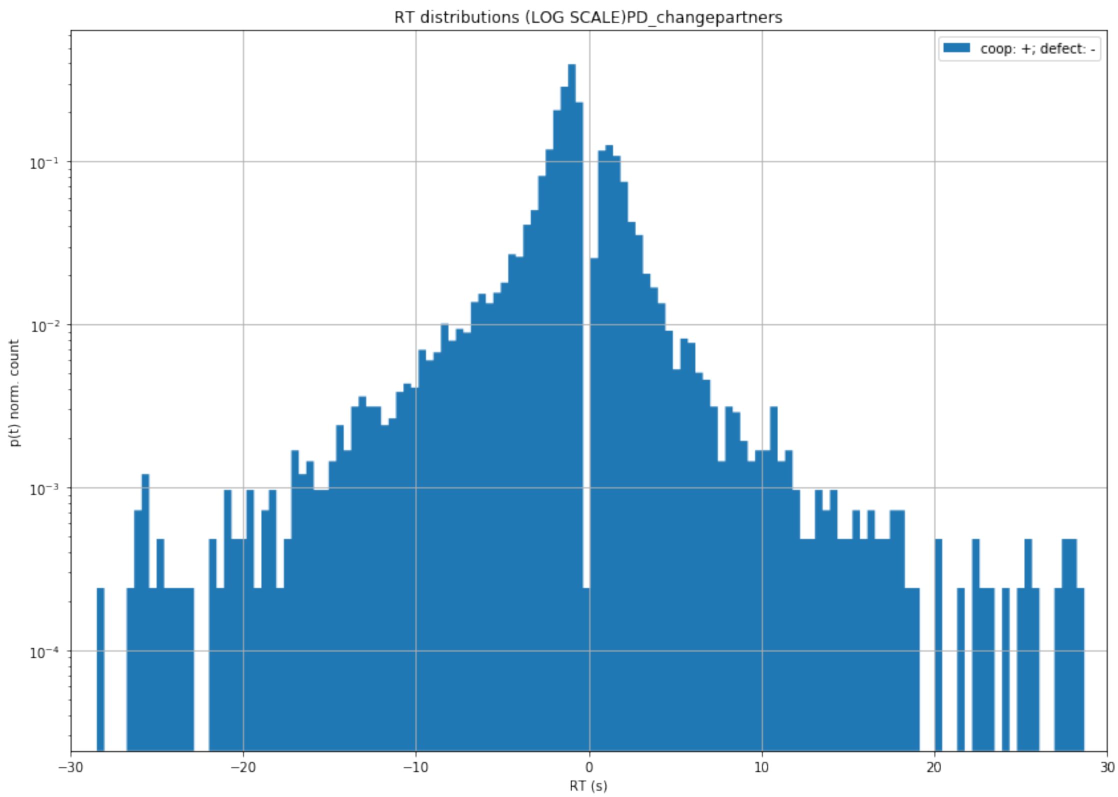
Out[4]:

	round	userid	response	rt
0	1	s5m20_usuario1	0	-6.678
1	1	s5m20_usuario2	1	4.601
2	1	s5m20_usuario3	0	-12.777
3	1	s5m20_usuario4	1	2.399
4	1	s5m20_usuario5	0	-17.088

1.3.1 RT distributions for cooperators and defectors + log scale

Obvious bias towards defection.





2 Bin size

For any histogram, the bin size follows eq (https://www.fmrib.ox.ac.uk/datasets/techrep/tr00mj2/tr00mj2/nod

$$W = 3.49\sigma N^{-1/3}$$

here σ equals 2,94 and N equals 17200.

$$W = 0.39749341805822797$$

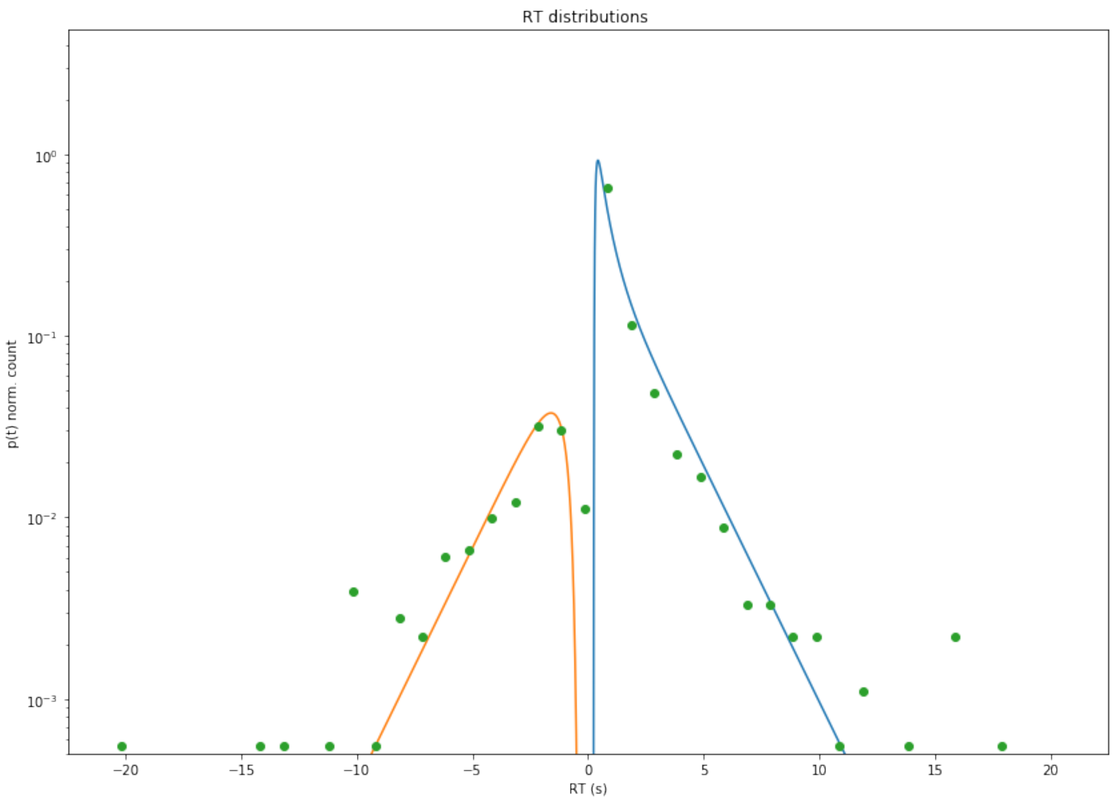
so the number of bins has to be

$$60/W = 150$$

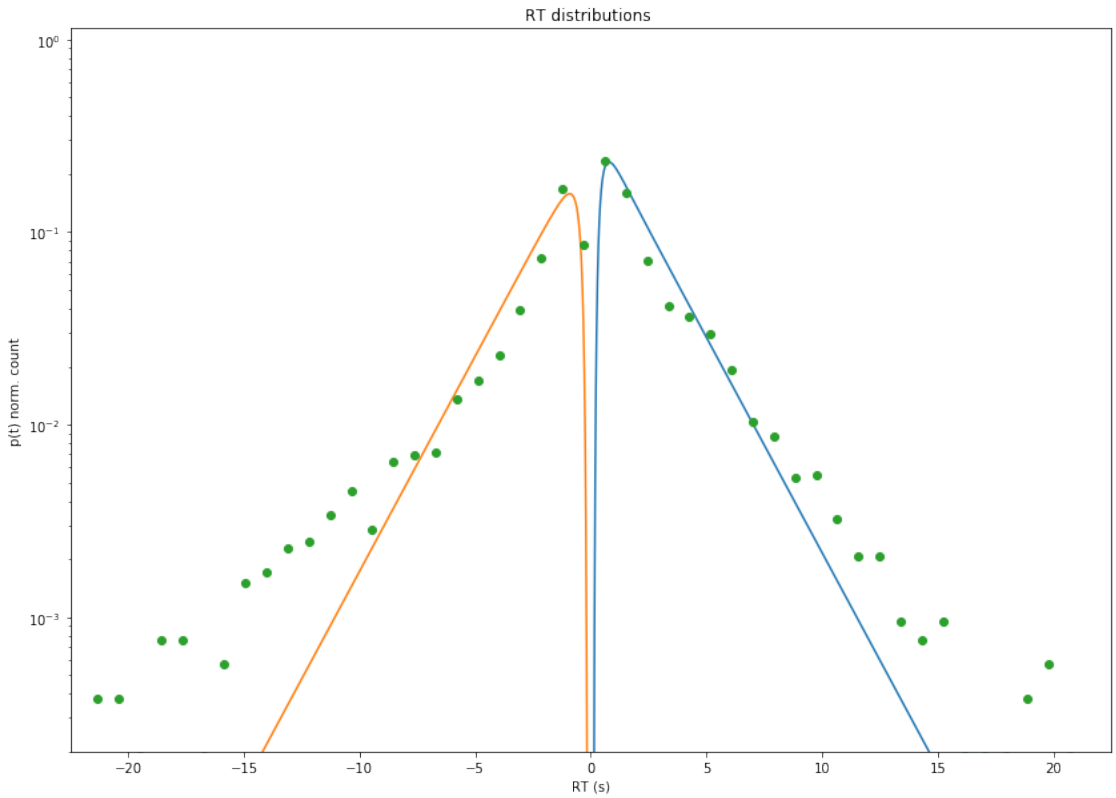
as the data is split in 2 sides (positive for cooperation times and negative for defection times) we use 300 bins.

3 MODEL FITTING

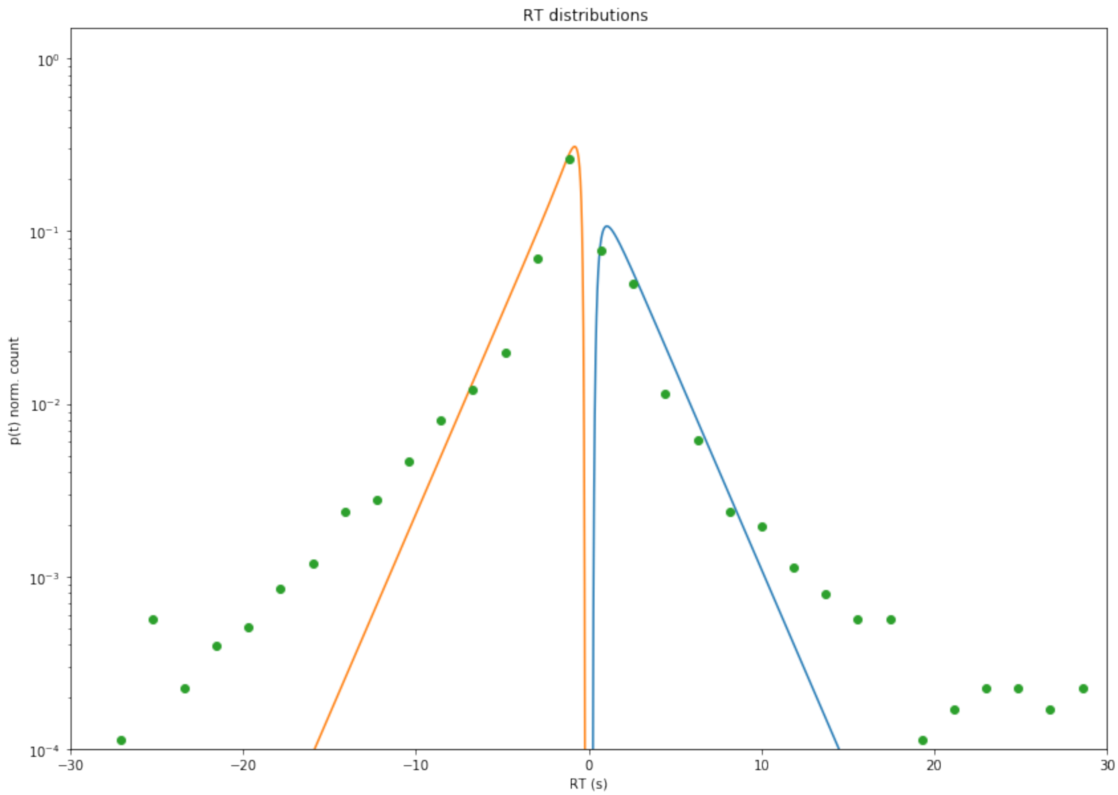
3.1 Weak PD with fixed partners



3.2 PD with fixed partners



3.3 PD with changing partners



4 FITTING QUALITY

In order to assess the quality of the data, a statistical test is carried out over the following parameters: drift rate v, initial bias z, non-decision time t and decision boundary a. A first statistical analysis can be done with the autocorrelation, the trace and the histogram for each parameter.

Autocorrelation coefficient at lag k is obtained from the following eq.

$$r_k = \frac{\sum_{i=1}^{N-k}(x_i - \bar{x})(x_{i+k} - \bar{x})}{\sum_{i=1}^{N}(x_i - \bar{x})^2}$$

These autocorrelation plots show how random our data is. As expected, there are no significant autocorrelations in any of the plots (except the first parameters which are always 1 by def).

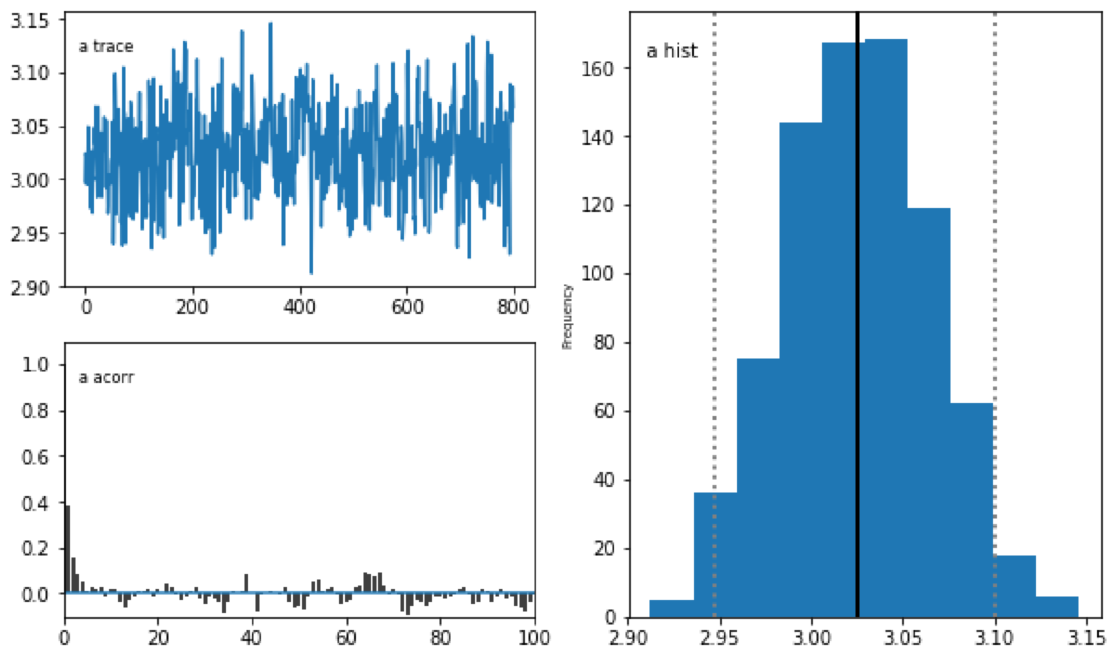
More explanation of autocorrelations in
https://www.itl.nist.gov/div898/handbook/eda/section3/autocop1.htm

On the other hand, the trace shows how stationary the obtained parameters are. As some data values are discarded on the first rounds, we need to make sure that the parameter has become steady in order to assure its veracity.

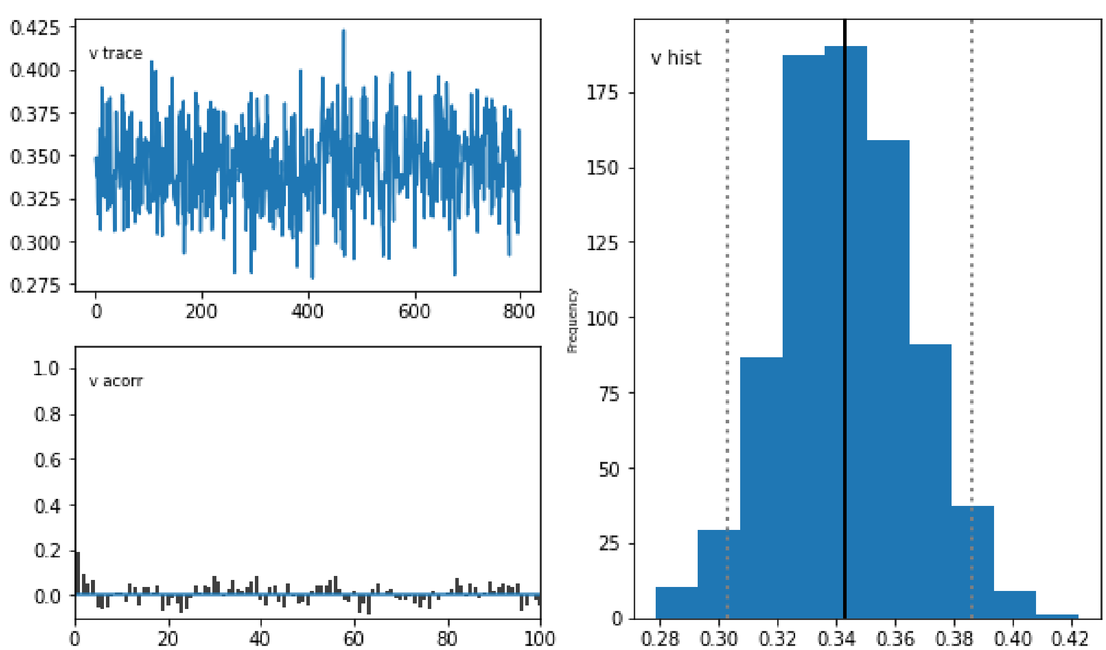
The parameter histogram shows, for different simulations, what was the final value. A normal Gaussian distribution implies an appropriate final value.

4.1 Weak PD with fixed partners

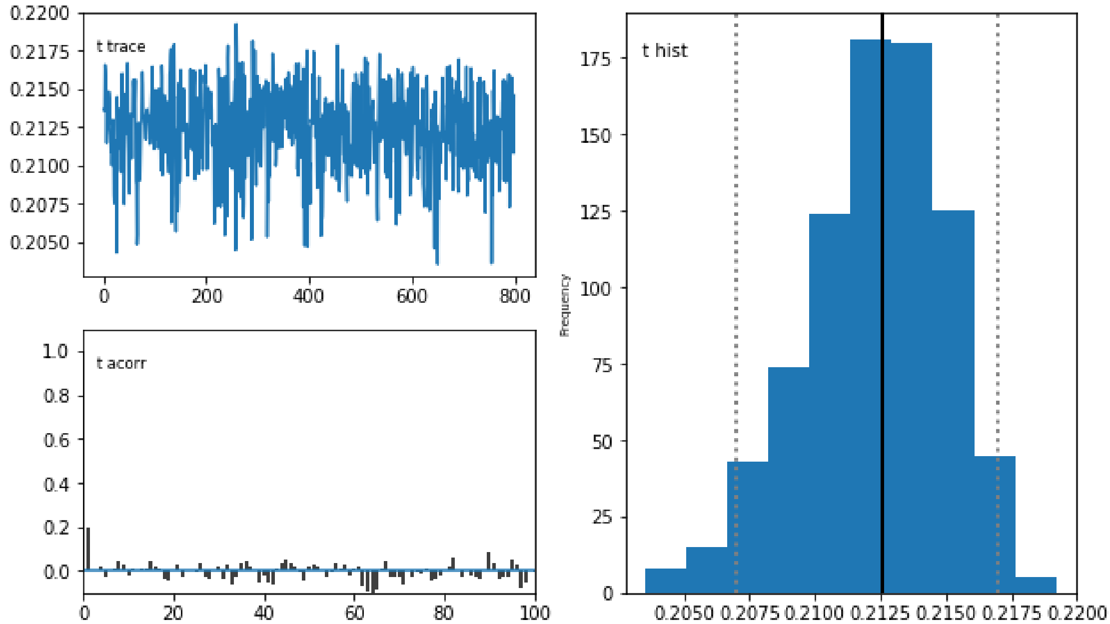
4.1.1 a analysis



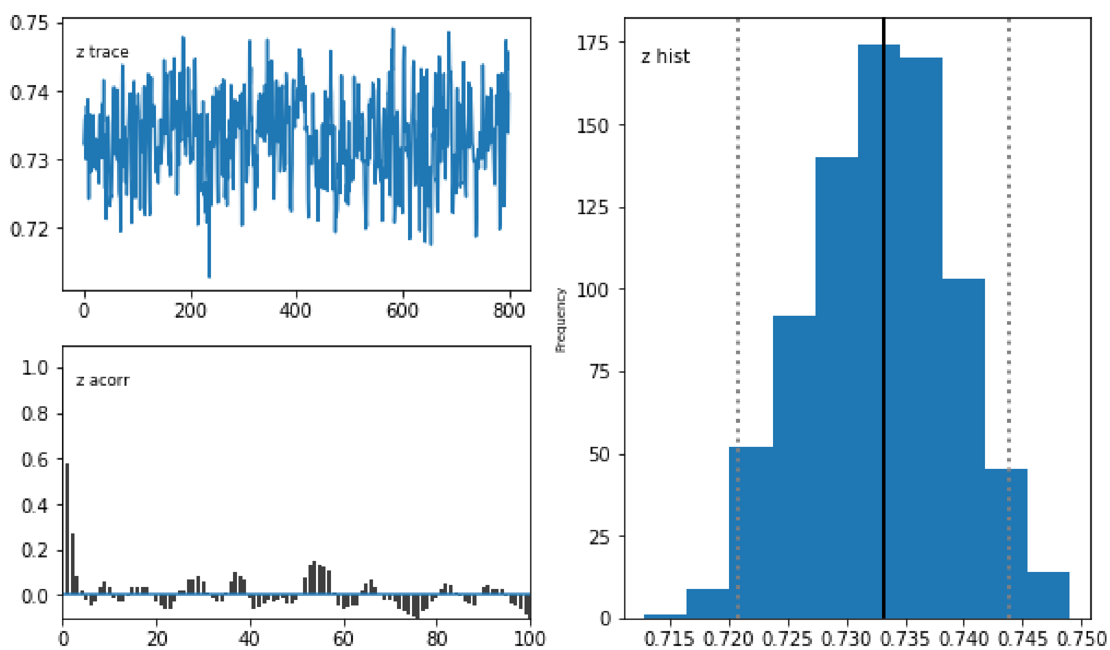
4.1.2 v analysis



4.1.3 t analysis

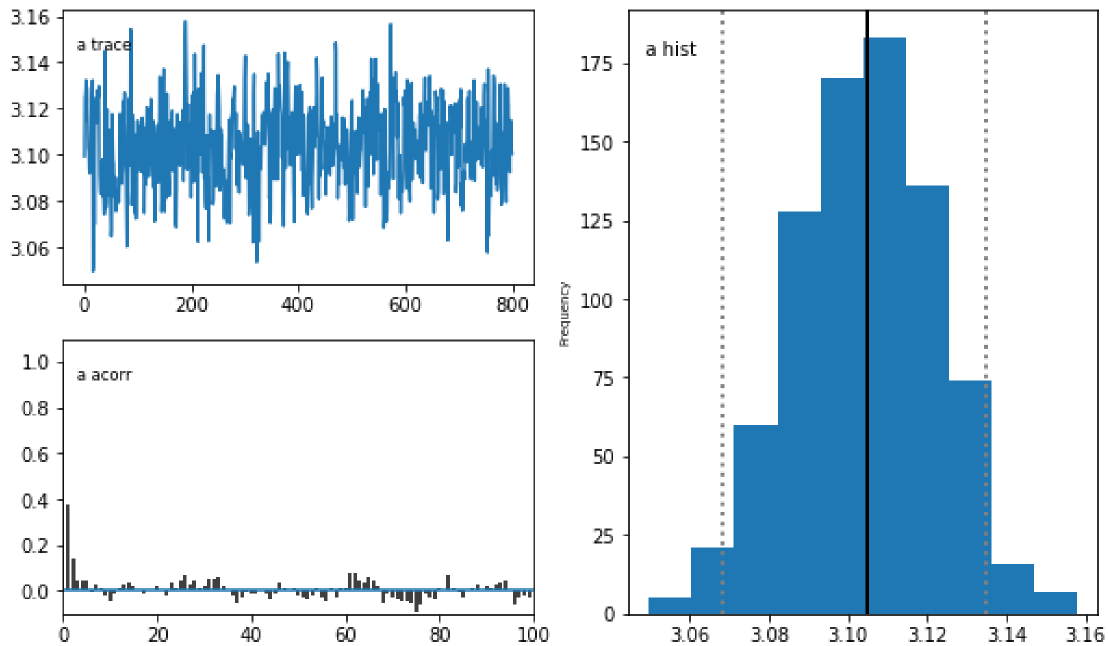


4.1.4 z analysis

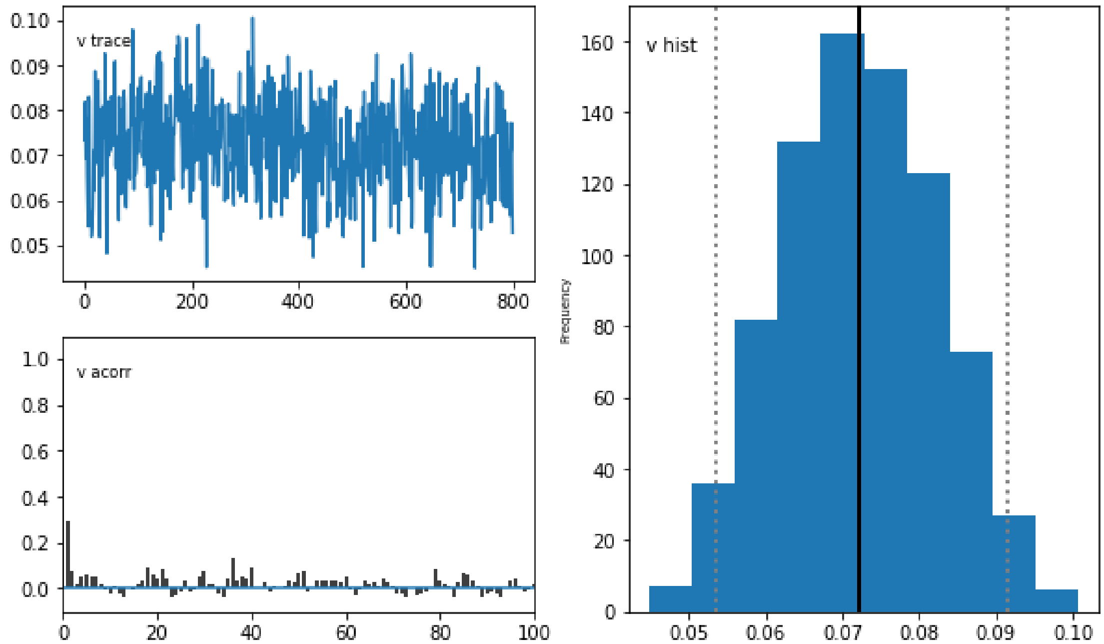


4.2 PD with fixed partners

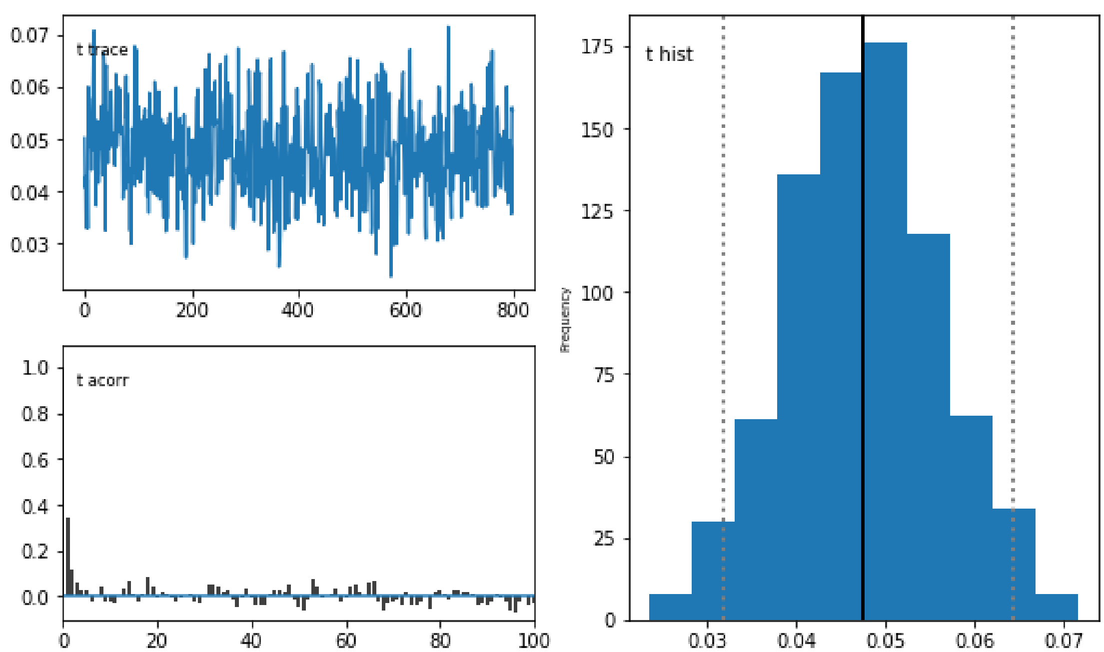
4.2.1 a analysis



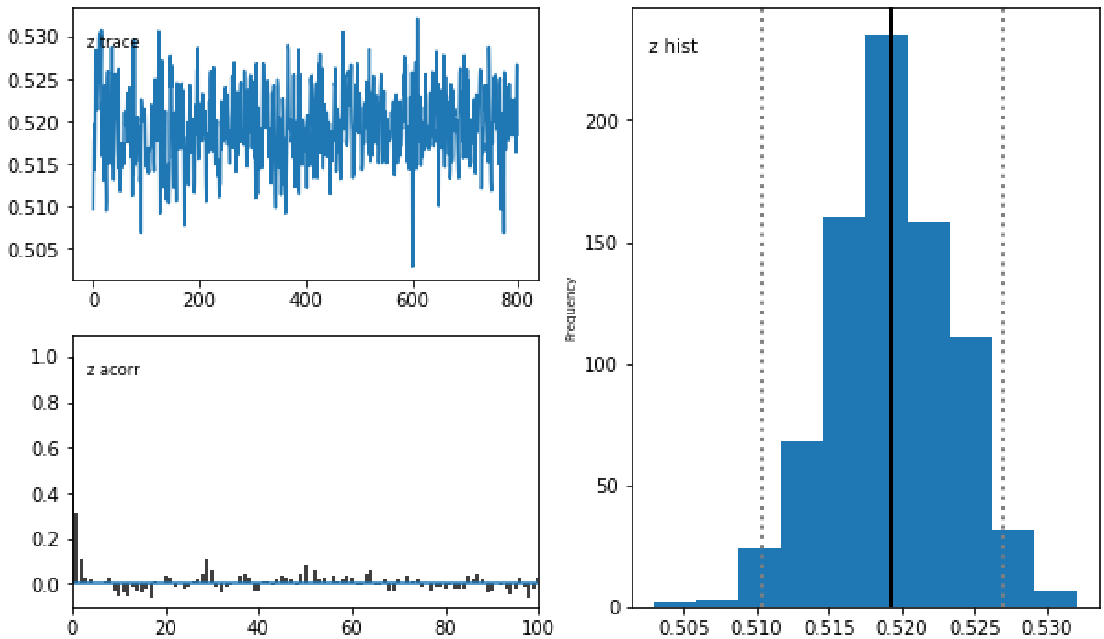
4.2.2 v analysis



4.2.3 t analysis

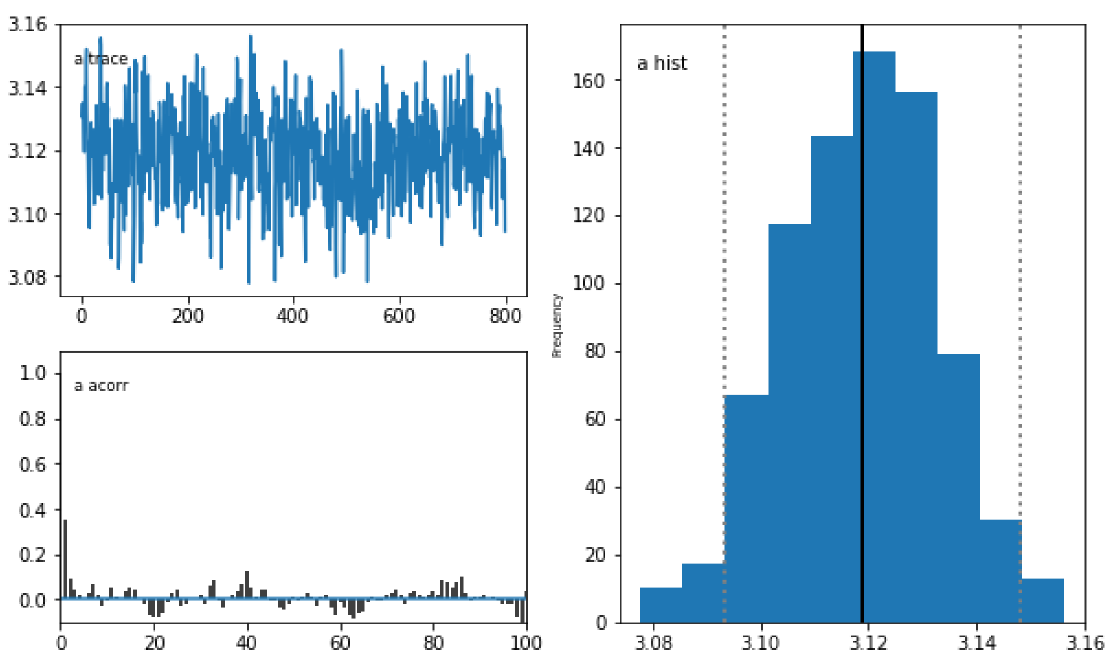


4.2.4 z analysis

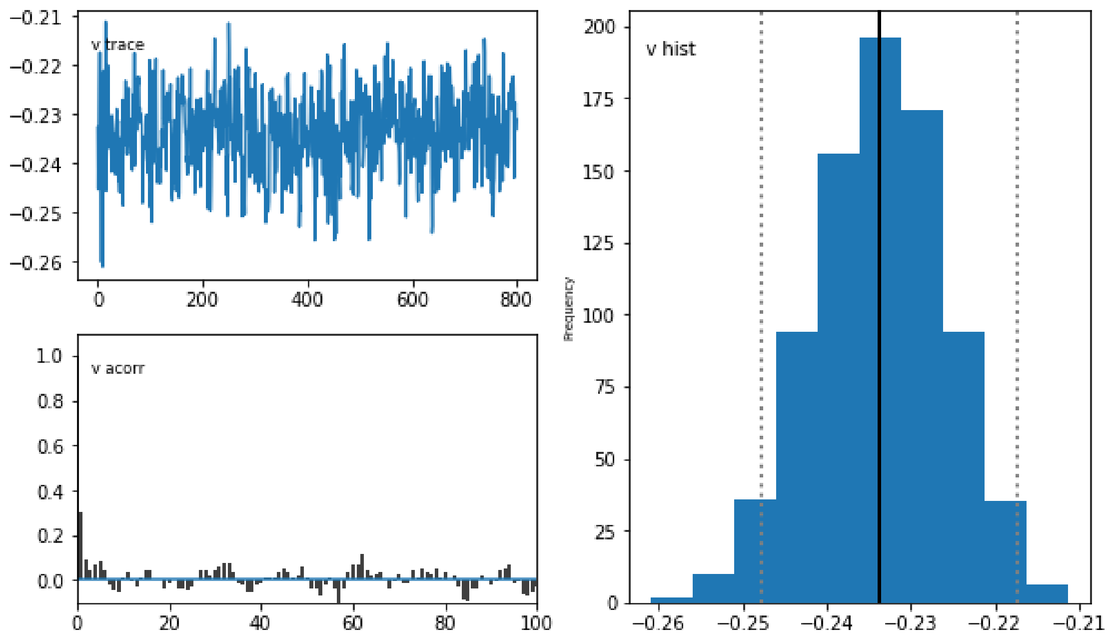


4.3 PD with changing partners

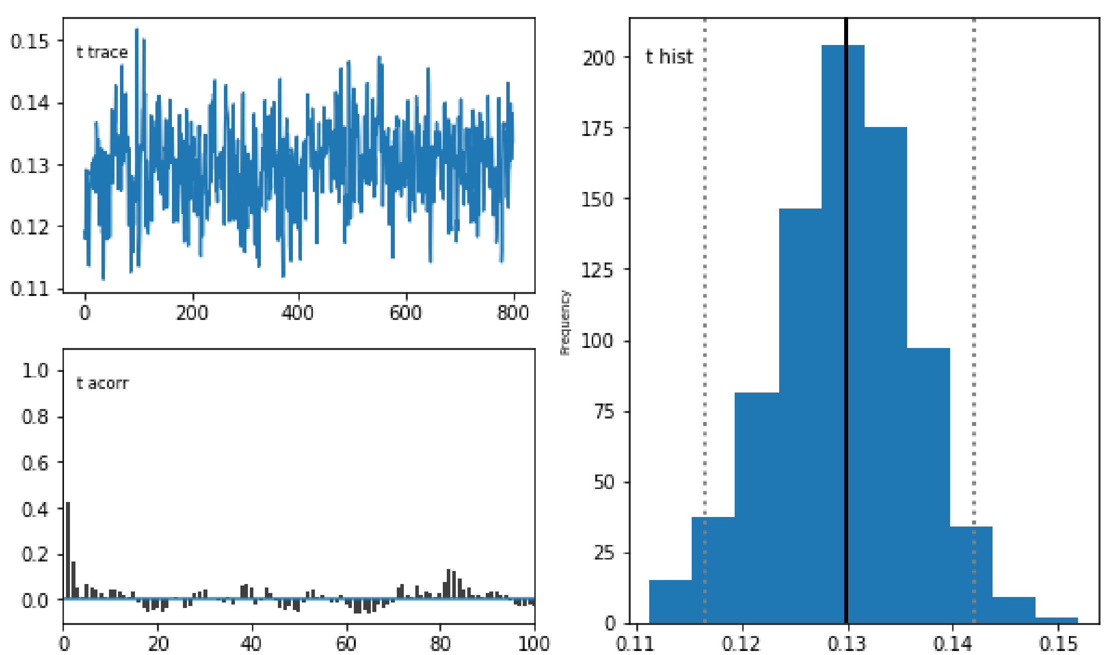
4.3.1 a analysis



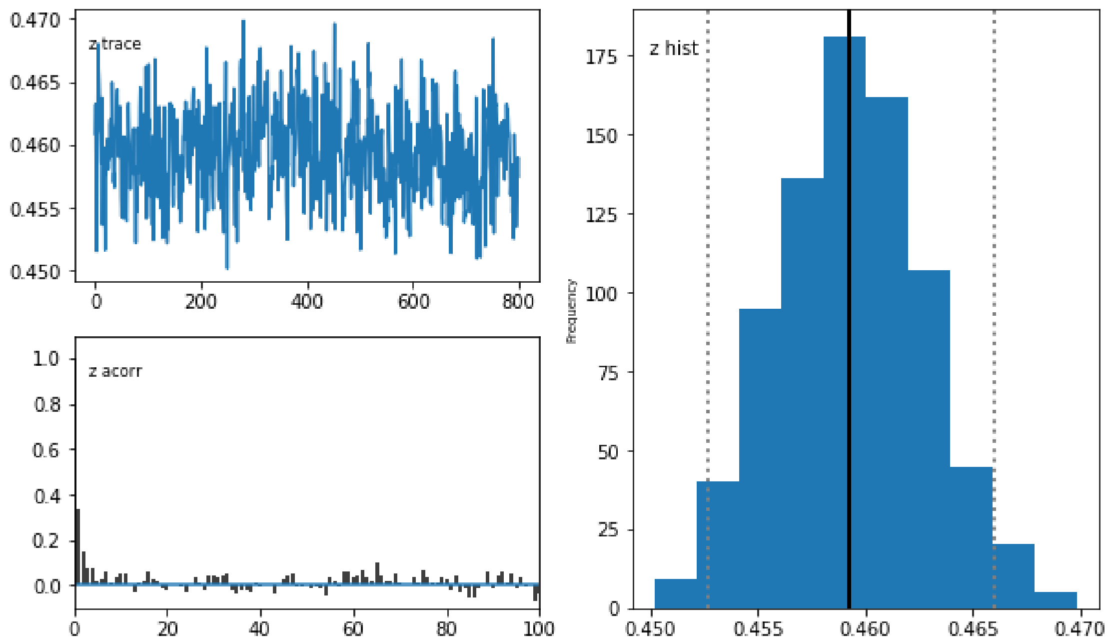
4.3.2 v analysis



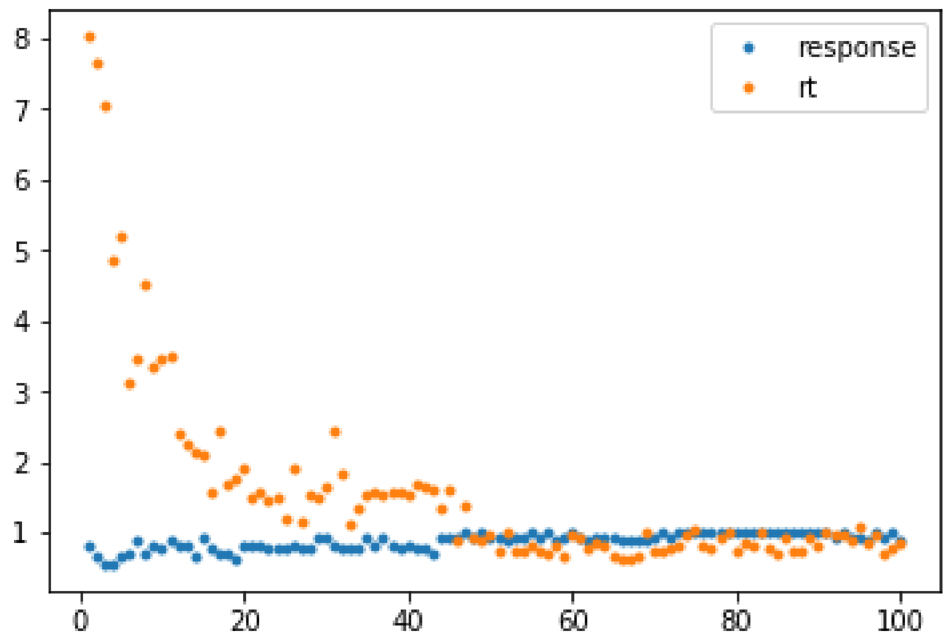
4.3.3 t analysis



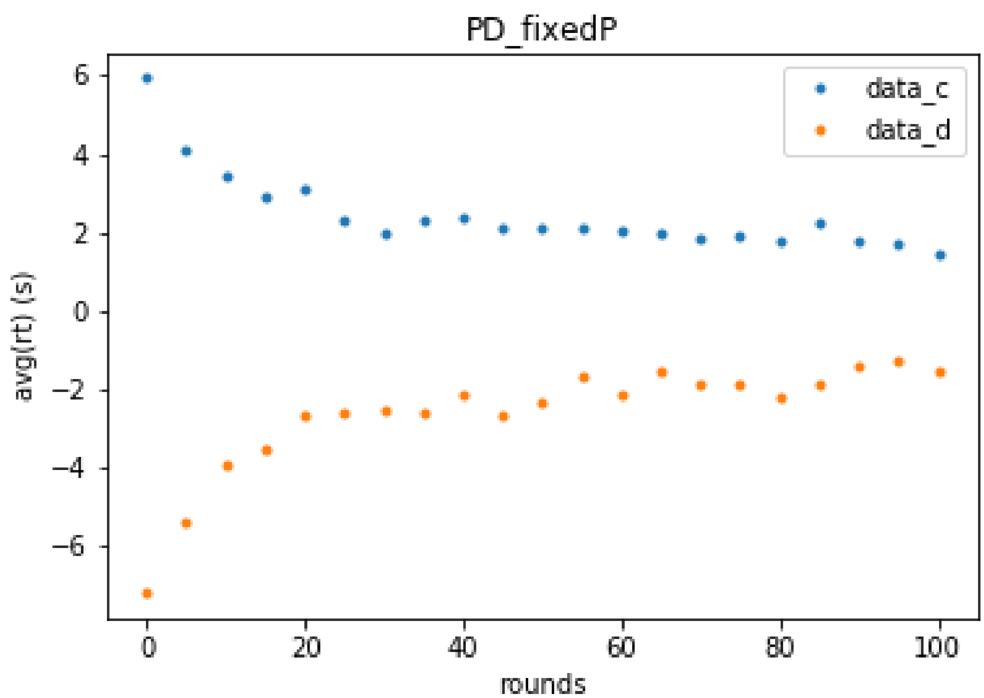
4.3.4 z analysis

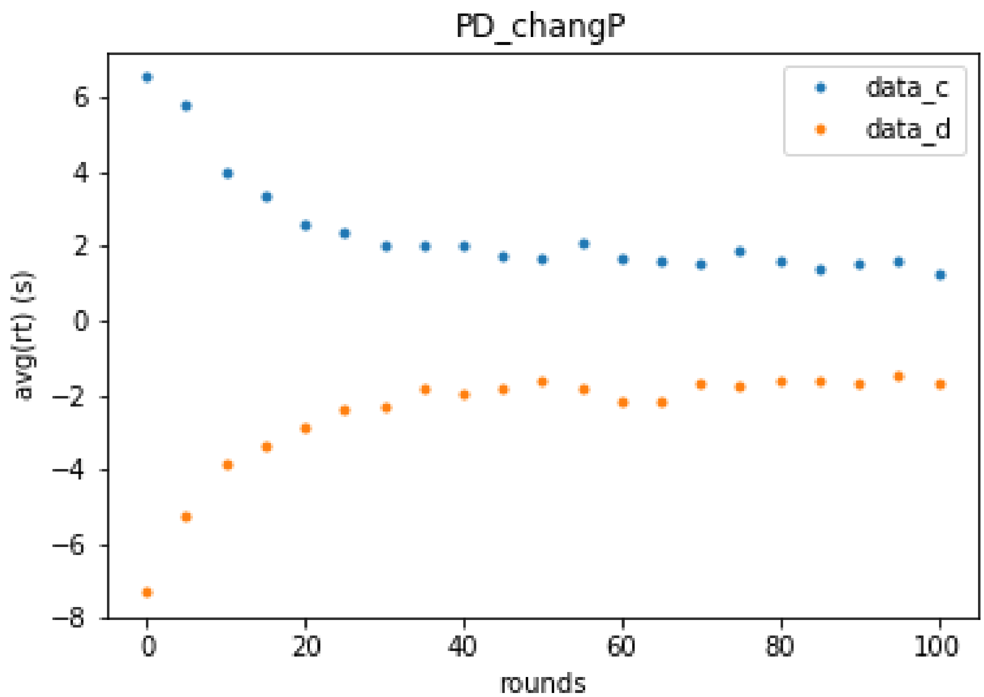


5 GROUP BY ROUND: RT (orange) by round



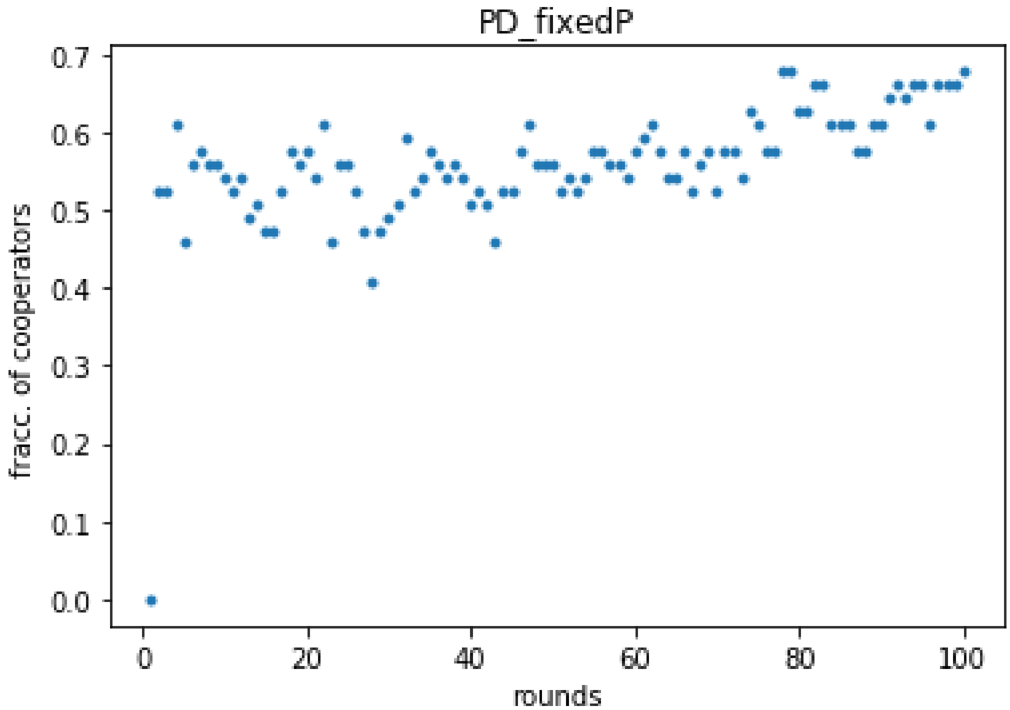
In the first rounds, (for both datasets) takes less time to decide coop than defection. Then it stabilizes around 2 s for later rounds.

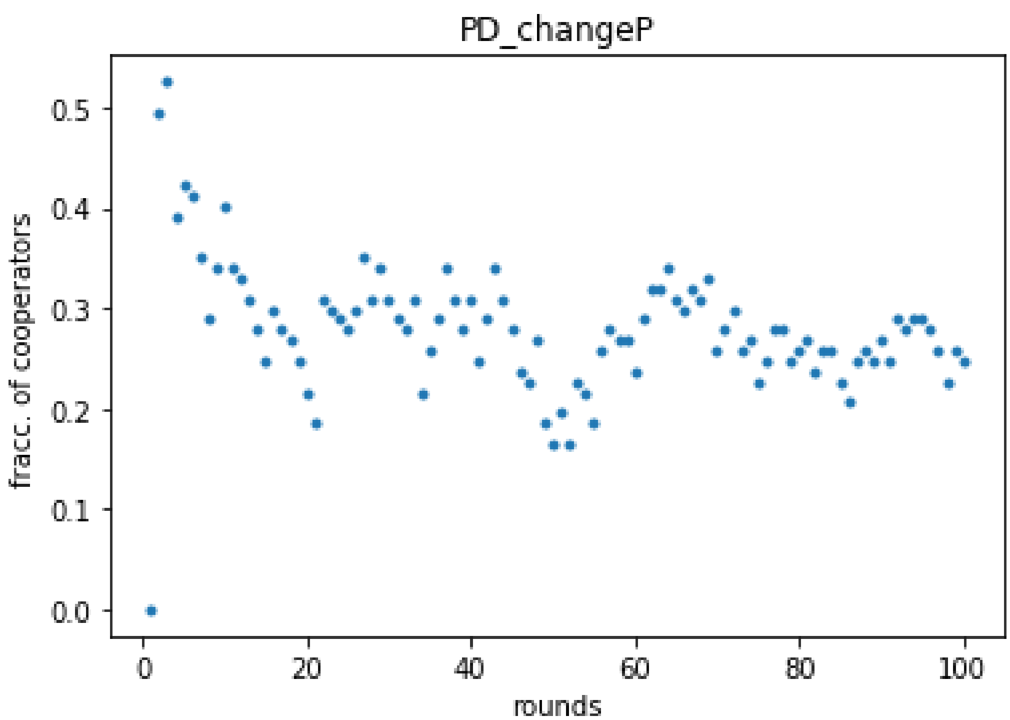




6 Fraction of cooperators

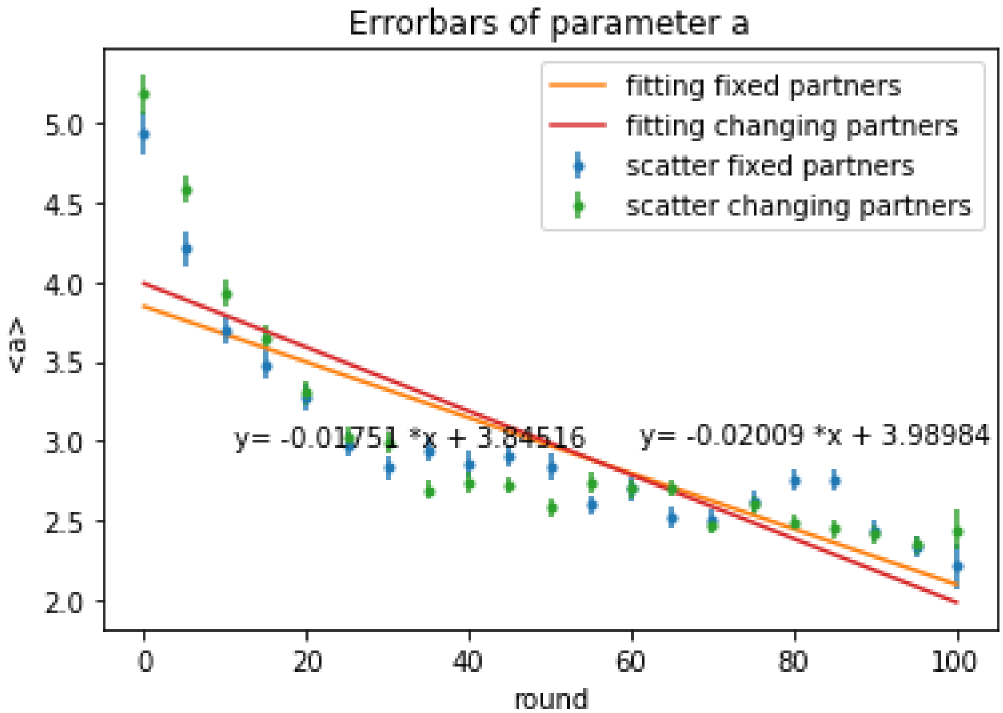
Grows for fixed partners but decreases for changing.


PD_fixedP
fracc. of cooperators
rounds
0.0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0
20
40
60
80
100

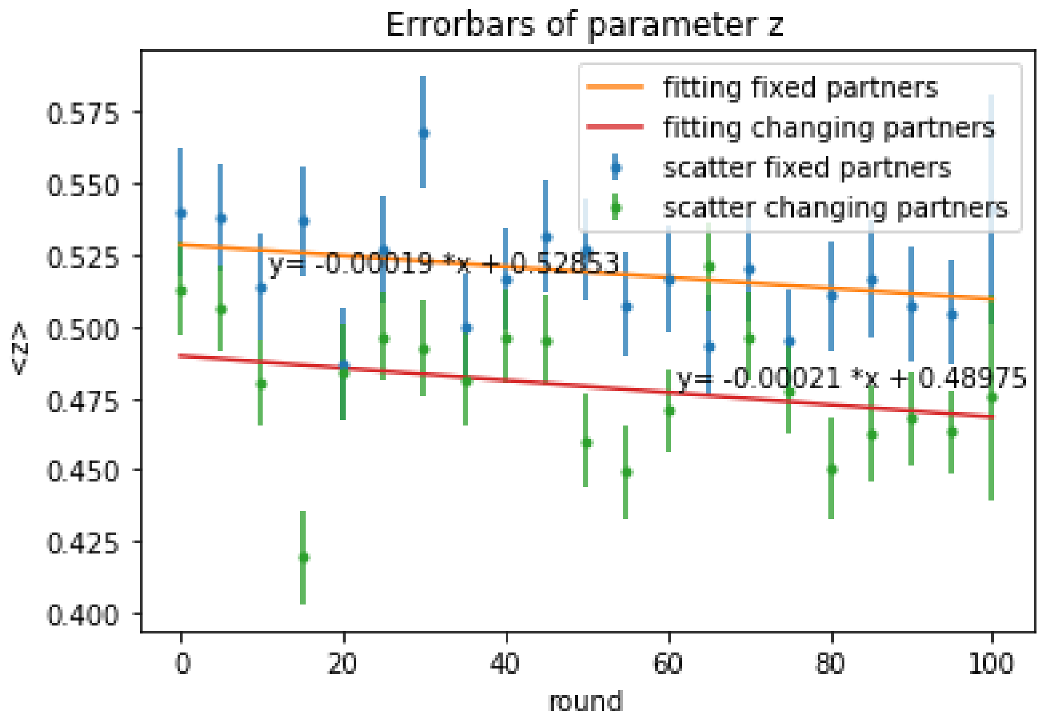

PD_changeP
fracc. of cooperators
rounds
0.0
0.1
0.2
0.3
0.4
0.5
0
20
40
60
80
100

7 Evolution of the parameters by round

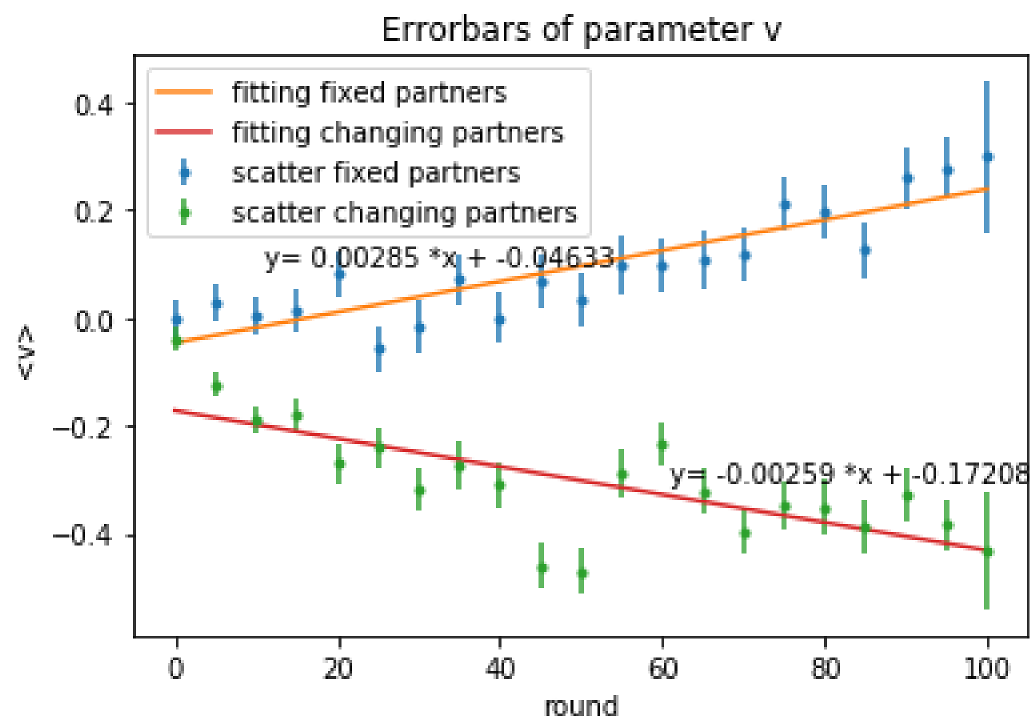
a defines how difficult the game is, decreases with time for both games (which makes sense, the more you play the easier it gets).



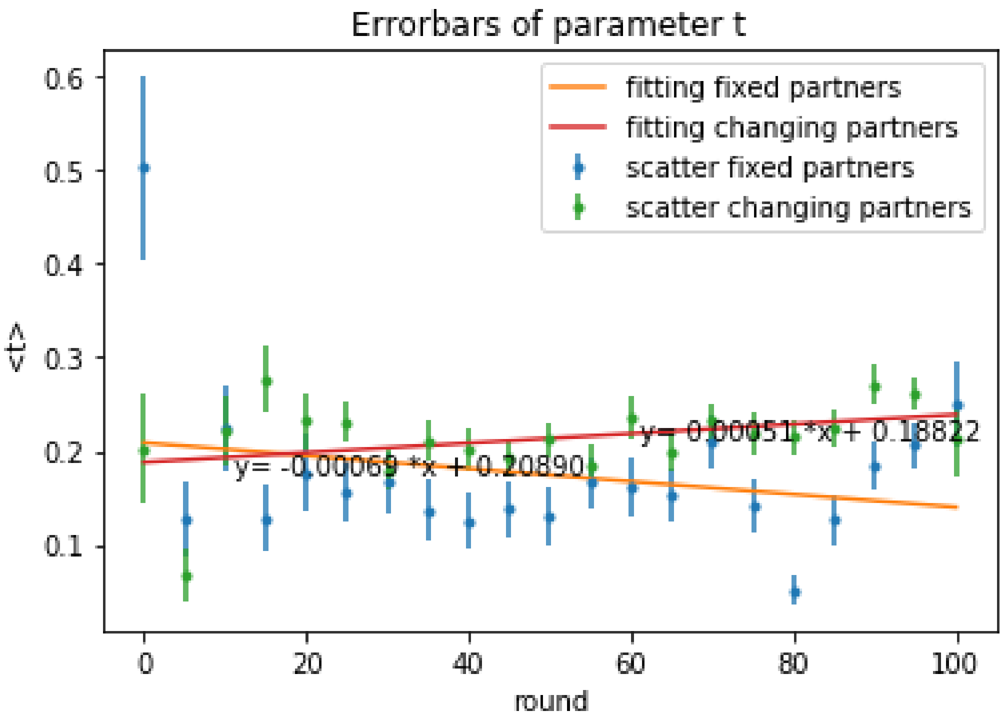
For fixed partners the initial decision is to cooperate (z>0.5) but after, it decreases. For both games people evolve towards defection.



v grows for fixed partners but decreases with time for changing partners. Higher coop with time, how does this fit with lower z per rounds?



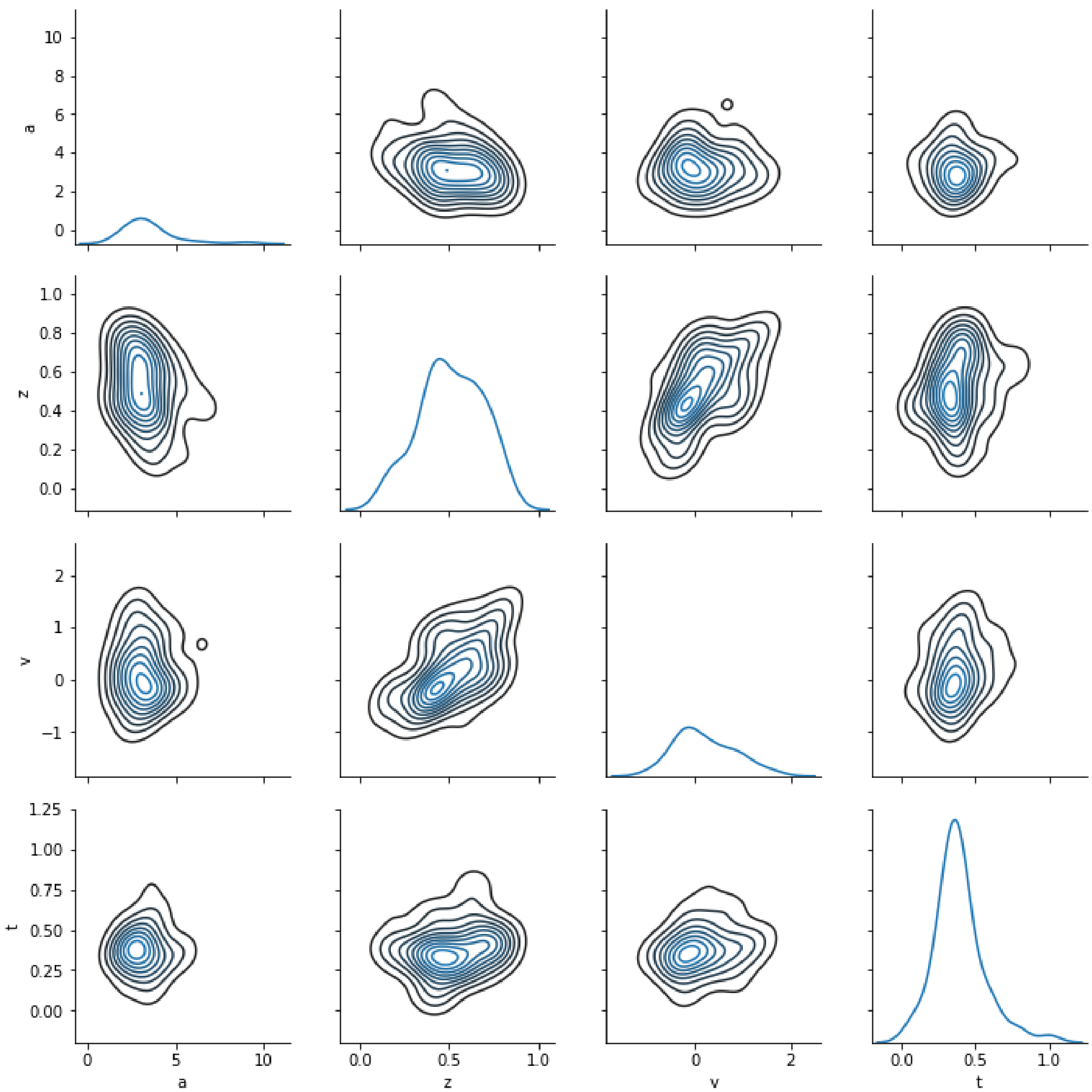
t incerases with time? loss of focus/interest on the game?


Errorbars of parameter t
fitting fixed partners
fitting changing partners
scatter fixed partners
scatter changing partners
y= 0.00051 *x + 0.18822
y= -0.00069 *x + 0.20890
<t>
0.6
0.5
0.4
0.3
0.2
0.1
0
20
40
60
80
100
round

8 GROUP BY PLAYERS

8.1 Correlations

8.1.1 Correlations fixed partners



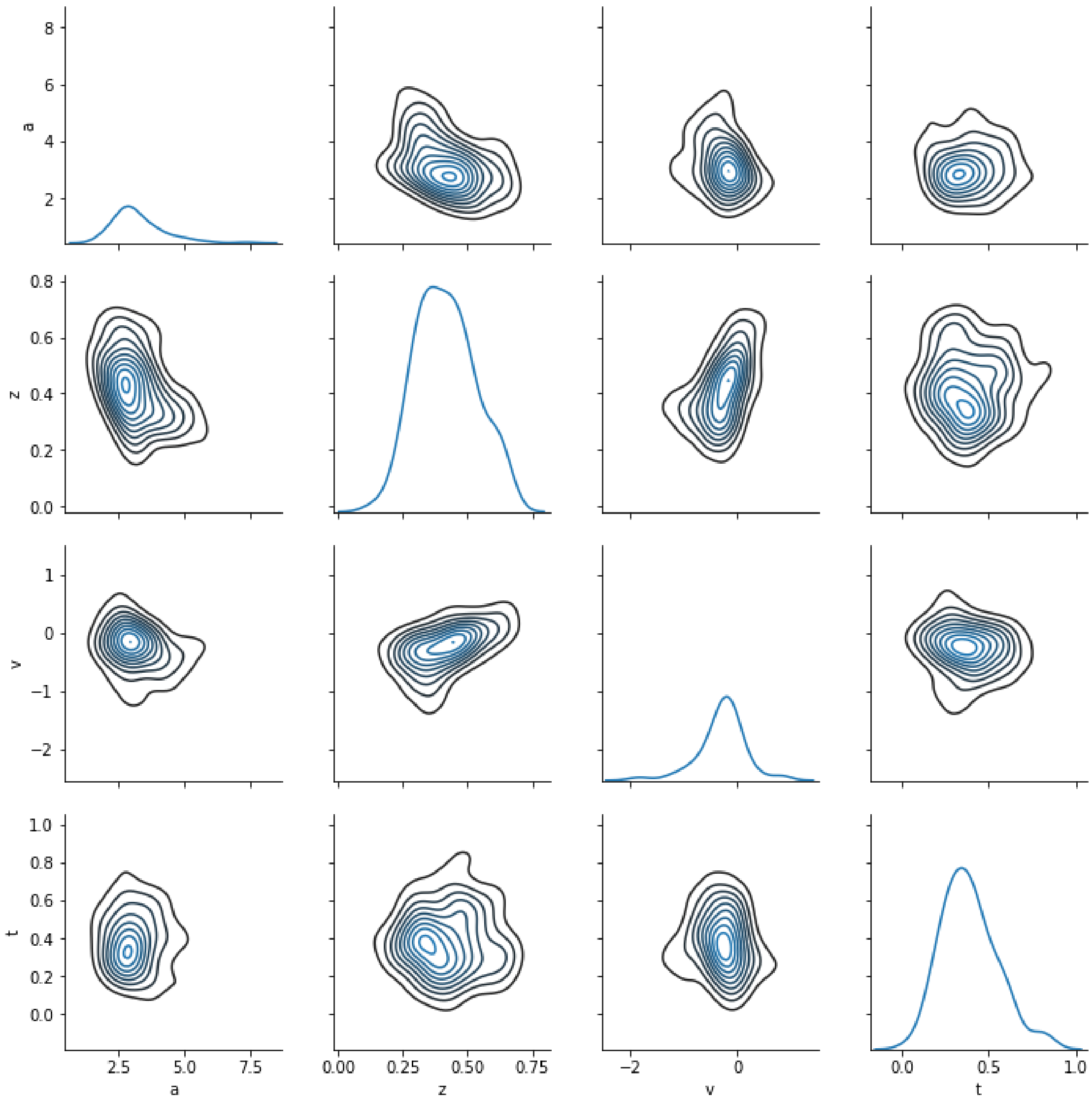
We observe a clear correlation between v and z for both fixed and changing partners. We obtain the following values for Pearson correlation coefficient and its pvalue:

$$0.43620745416795215, 0.0006195414012371265 7$$

significant!

8.1.2 Correlations changing partners

Pearson corr coef and p value: 0.37218761672582679, 0.00018846910708224403



8.2 Significant parameters: PCA

In [ ]:
```
print(X4)
print(pd.DataFrame(np.cov(X4)))


from sklearn import decomposition

pca = decomposition.PCA()
pca.fit(X4)
print(pca.explained_variance_)

# As we can see, all the components are independent and therefore significant
```

After this, we did clustering of people by their a,v,z and t parameters. We discard t as it is an game-unrelated parameter. Kmeans, MeanShift, agglomerative clustering, Gaussian Mixture and DBSCAN did not give too nice clusters and we are still working on this (and the significance of this clusters).