



Modeling phenomena based on Epidemic diseases

Sharif University of technology

Winter/spring 2020

Homework2

Yasaman Asgari

Network based SIR Model

Introduction

In the first and second homework, we introduced an easy subtle model named “SIR” but in this homework we will look more precisely to have a better estimation in our society. In that model we assumed that all the people are completely connected to each other, but it’s not an exact assumption because if we look human or even any alive wight (creature), we realize that we are just connected to a small group of others which we call it **Neighbors**. In some studies, they’ve shown that in human societies the Neighbor’s population is about 10000 smaller than the overall population .So it really affect on the result of spreading a disease comparing to the previous Model .

Steps to model formulating

- **Definitions and assumptions**

The interconnection between components of the society is given by an **Adjacency Matrix** (named A in this report). In graph theory we use this term as a N^2 matrix which N is the total population. In this matrix we can assume that all the elements on the main diagonal are 1 and also the society graph is not directed-symmetric- ($\forall i, j$ which $1 \leq i, j \leq N : A(i, j) = A(j, i)$).Actually we name the components **Nodes** .The process is actually similar to model in HW1.We define three states for the components: **Susceptible** (if previously unexposed to the disease), **Infected** (if currently colonized by the disease), and **Recovered** (if they have successfully cleared the infection) .However, there are some differences .First, there is no actual rate in this model, every infected component can infect its neighbors with a **Probability** or will be recovered with another probability.

For Example, if we have three nodes –A (1), B (2), C (3)-and connections are (A, B), (A, C).The Adjacency Matrix and adjacency list will be as following.

$$\begin{pmatrix} 0 & 1 & 1 \\ 1 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix}$$

$$\{A : \{B, C\}, B : \{A\}, C : \{A\}\}$$

In this model, time passes quantically and we have **time steps** which tell us that it's a discrete model during time. In every step, we iterate on the components and make the required changes-S to I, I to R-. But there is a problem. The process using the N^2 matrix will be really slow. We should use another form of Adjacency Matrix (look at the example) which is a **Linked List**- a helpful data structure- contains all the infected population and their neighbors.

- **Assumptions**

In this homework our network is a **Grid Network** with $N = 256$ which every node is connected to the right, left, up and down nodes and it also has **periodical boundary condition**-nodes in the rightmost and leftmost are connected to each other and upmost and down most nodes as well. Average degree and also all nodes' degrees are 4. Initial infection population is 1 ($I_0 = 1$) which has been randomly selected among the society. P is the **infection probability** and r is the **recovery rate**. Disease strength is shown by $R_0 = \frac{p}{r}$.

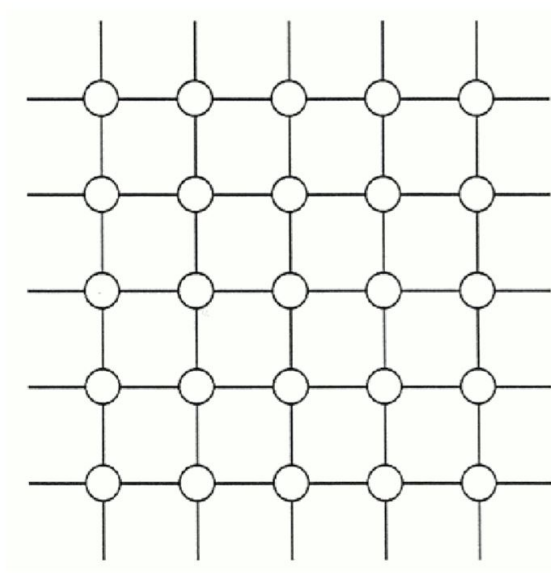


Image 1: 2-dimensional grid network

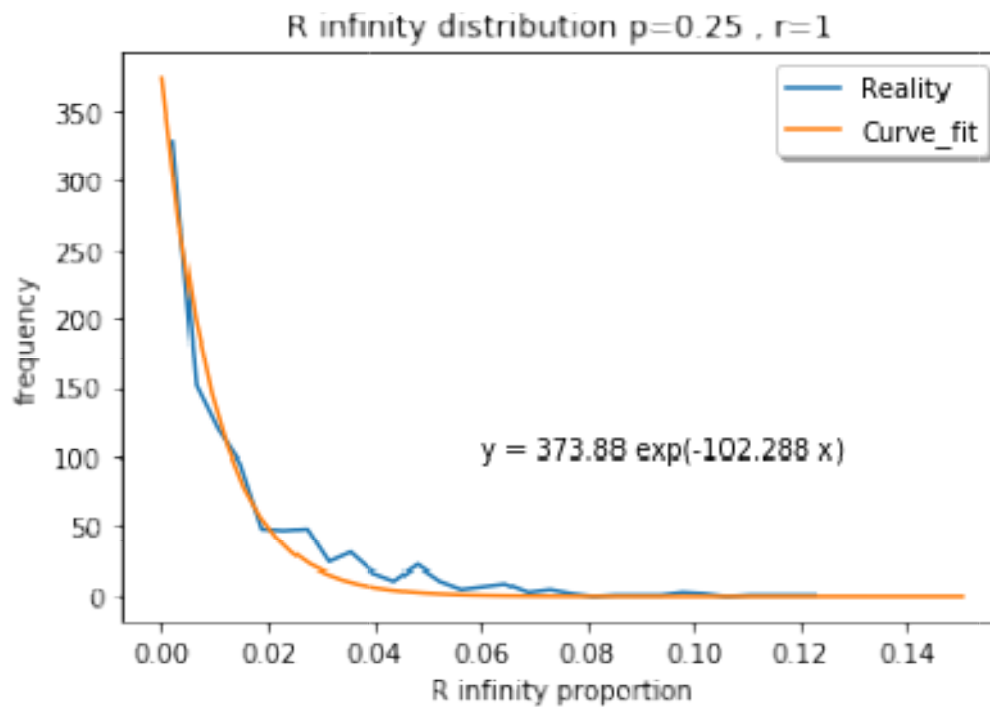
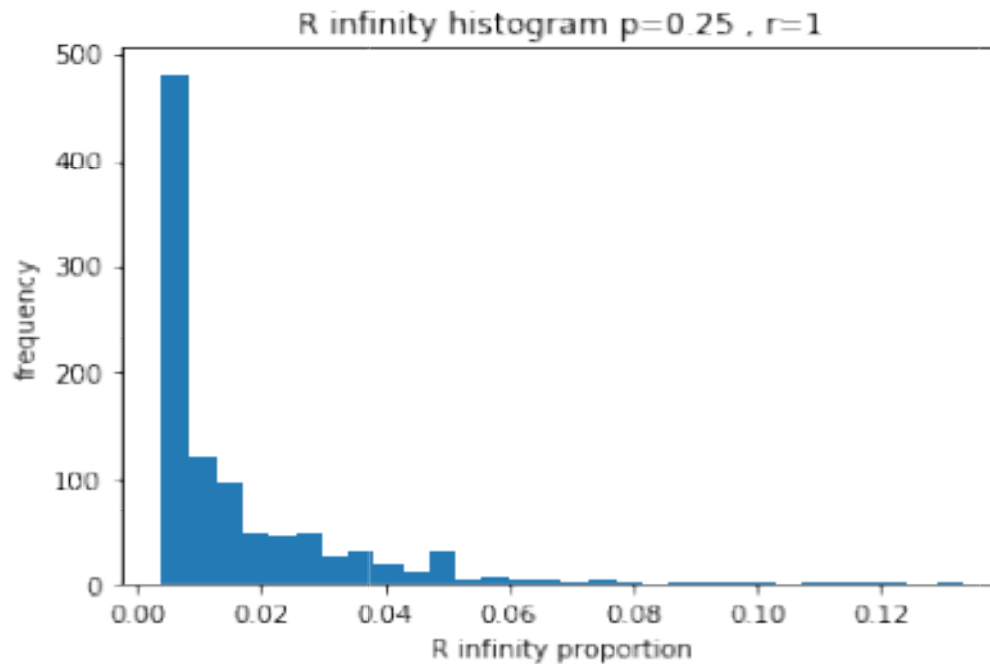
- **Future Condition problem**

For each node we have two attributes. First is the Initial condition before the time step, Second is the future condition. But why we use two attributes. Isn't the "condition" sufficient for the experiment? The answer is NO. Because we change condition with the next condition it could affect on system badly. For instance, assume that the I node will be recovered in this step, if we put R in the condition it can't infect other S neighbors and it isn't correct.

Results

- $P=0.25$, $r=1$, $I_0 = 1$

The second plot is derived from the first one. I've chosen the middle point of the bins and have counted the relevant data. Then I've used **Scipy** library the function **Curve fit** and as the image shows there is a good exponentiation fit for this histogram.



When we see this histogram we understand that with these initial arguments give us the R_∞ distribution like the exponentiation curves. First of all, I've found the error of fitting such curve and it's an array with element of (yfit-ydata) .Then I've squared them to make all of them positive. Then I found the mean and then square root .It's a good parameter to show how good our fit is. Then I used the RSquared formula to better describe the model .The formula is

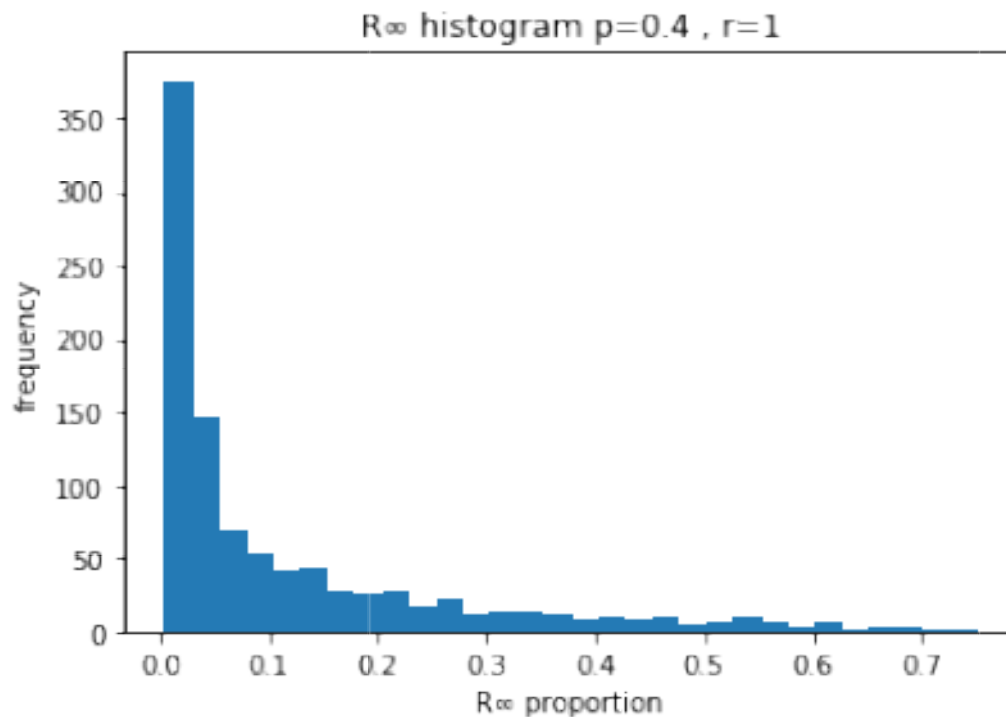
$$R^2 = 1 - \frac{\text{var}(\text{model})}{\text{var}(\text{ydata})}$$

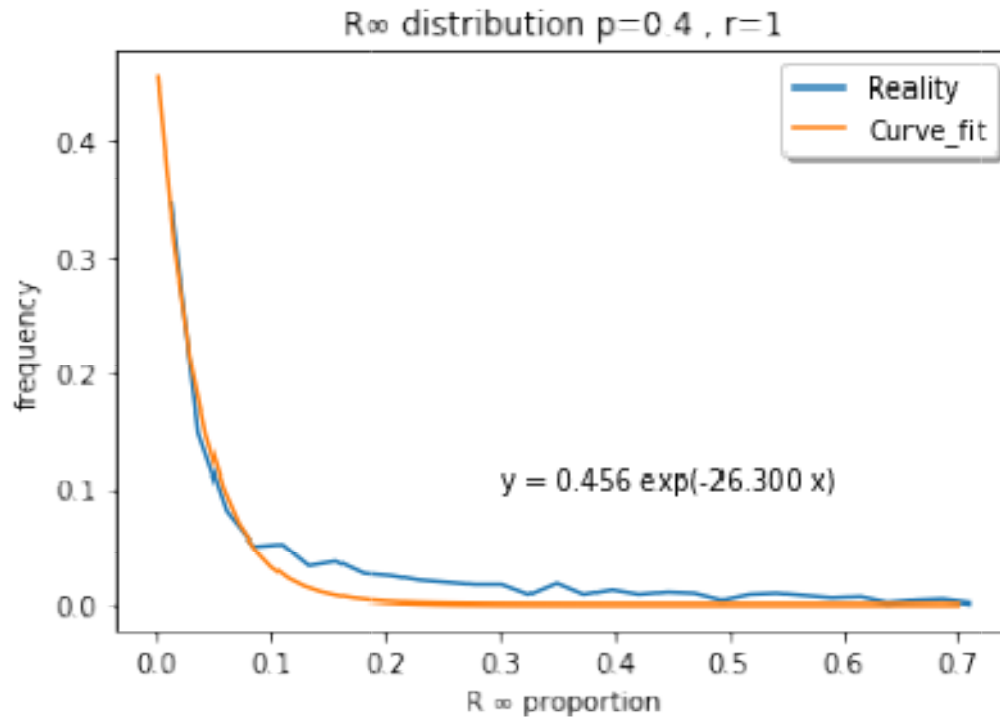
Results are as below:

- RSME(Root Mean Squared Error) = 0.00961570550152927 (really small)
- RSquared =0.9910057250057928 (really high which means it's a good model)

The histogram tells us the with this given parameters system collapses more probable because the R_∞ between 0 , 0.02 happen more frequently it means that disease spreads out rarely and most of the time the infected population recovered on early stages of the prevalence.

- $P=0.4, r=1, I_0 = 1$





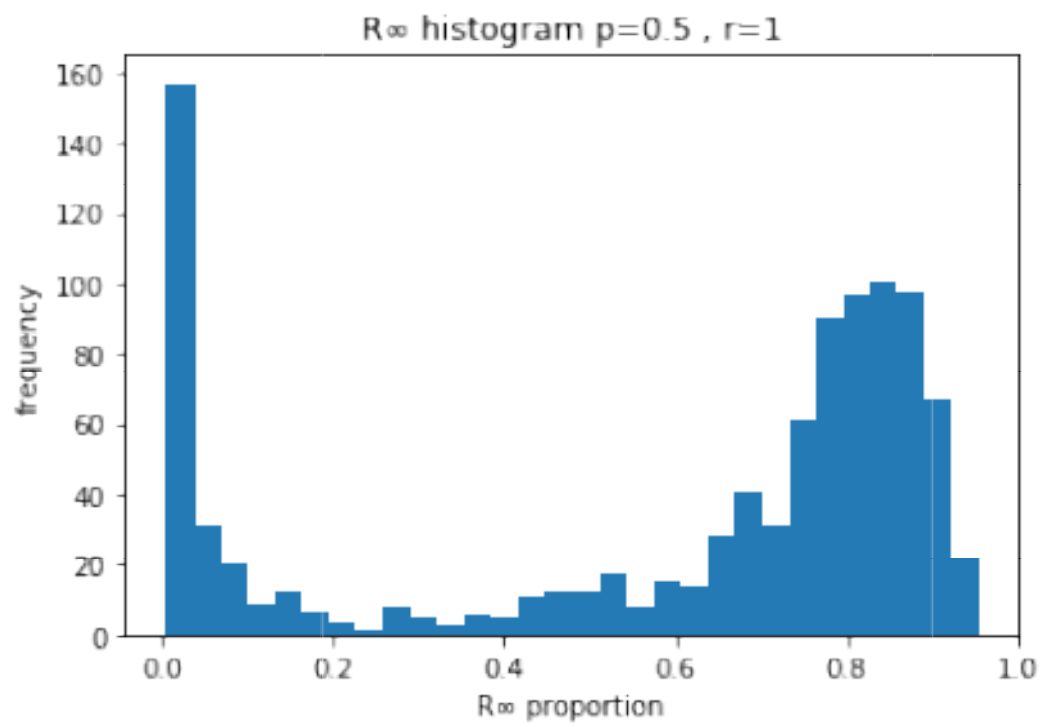
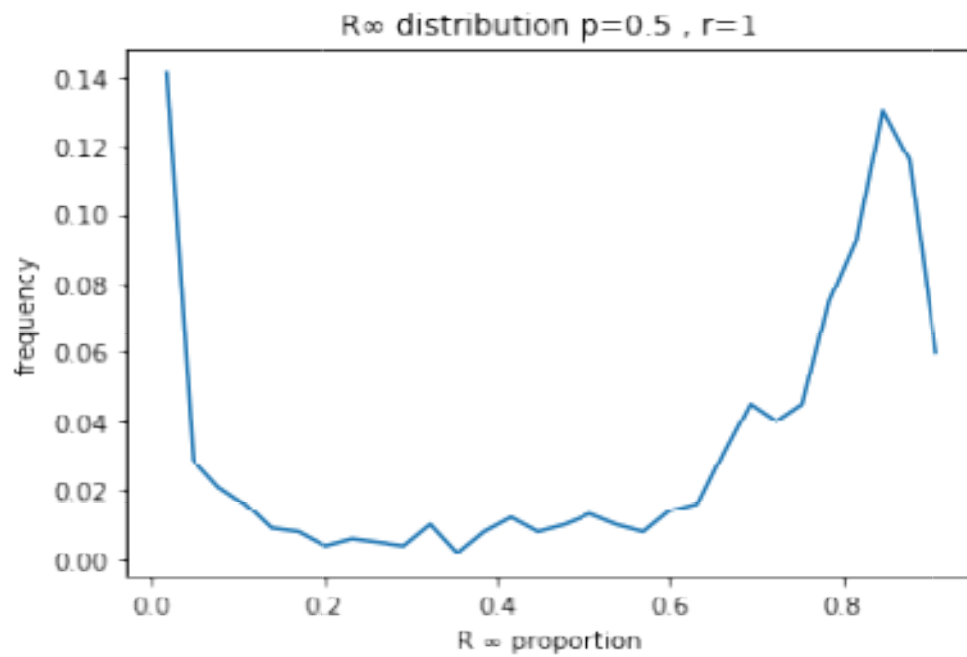
The results are as below: 0.014857312247783009

- RSME(Root Mean Squared Error) = 0.014857312247783009 (really small)
- RSquared = 0.9910057250057928 (really high which means it's a good model)

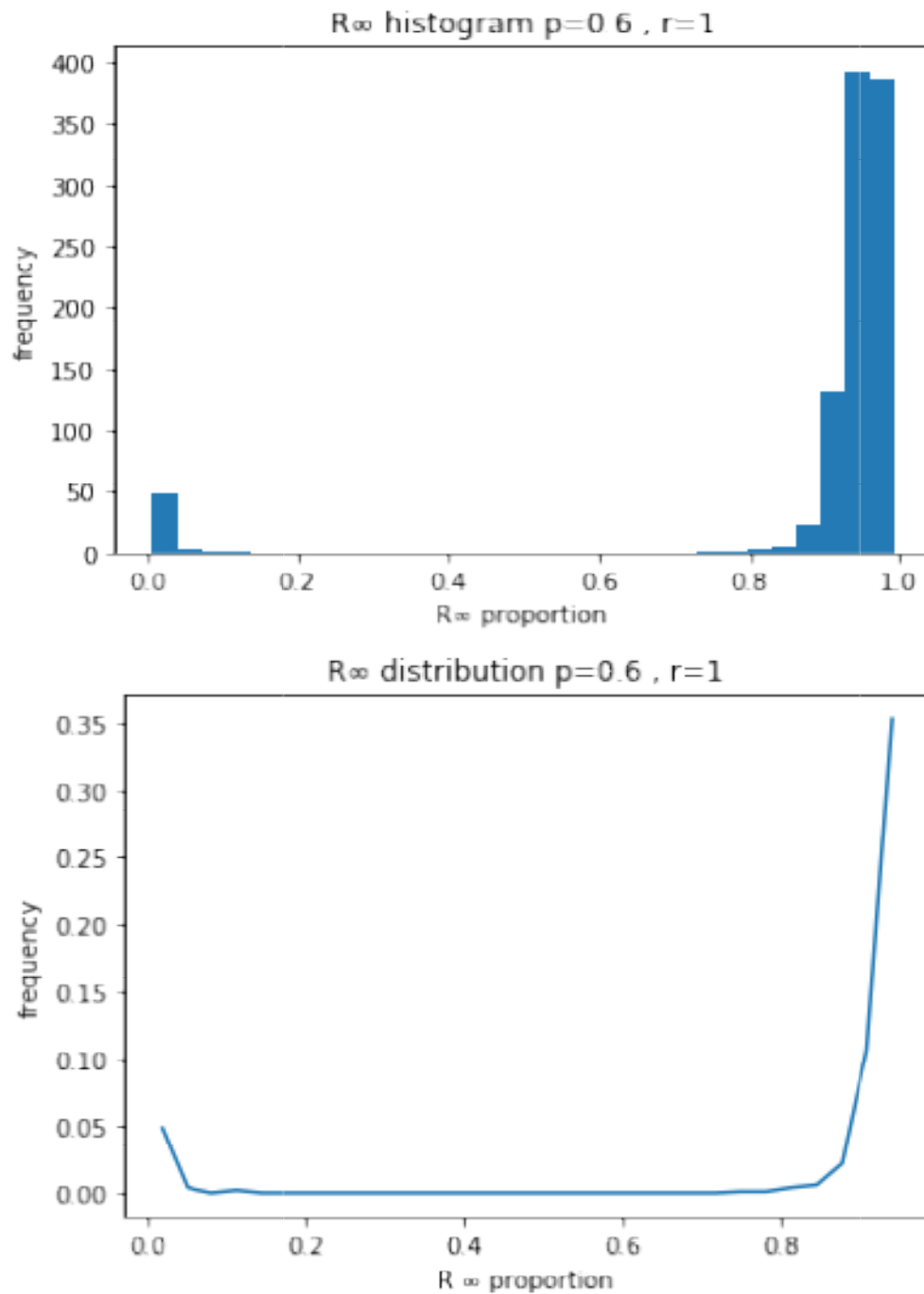
As said before, it's really similar to the condition $p = 0.25$ which it means that the disease spreads out with really small probability and the system collapses in earlier stages of prevalence. It somehow has a strong intuition. We have a disease that with probability smaller than the recovery probability, affects the susceptible population so as expected infected people recover soon and don't give the epidemic disease to others.

- $P=0.5, r=1, I_0 = 1$

When we look at the histogram and also as a function, we might think that this condition is a critical point in these initial parameters because the curve definitely changes before and after this point. For $p < 0.5$ R_∞ stays frequently near zero but vis-à-vis for $p > 0.5$ R_∞ distribution's peak moves to the right and it is more probable that the disease spreads out and becomes epidemic or worldwide. In the curve we have two local maximums. First near zero (which means that the first infected person recovered in the first step) and one near 0.8 (which means the disease will spread out). Generally, I understand that a competition will start between recovering and infection. Sometimes one loses, sometimes another, and the result is just based on the random.

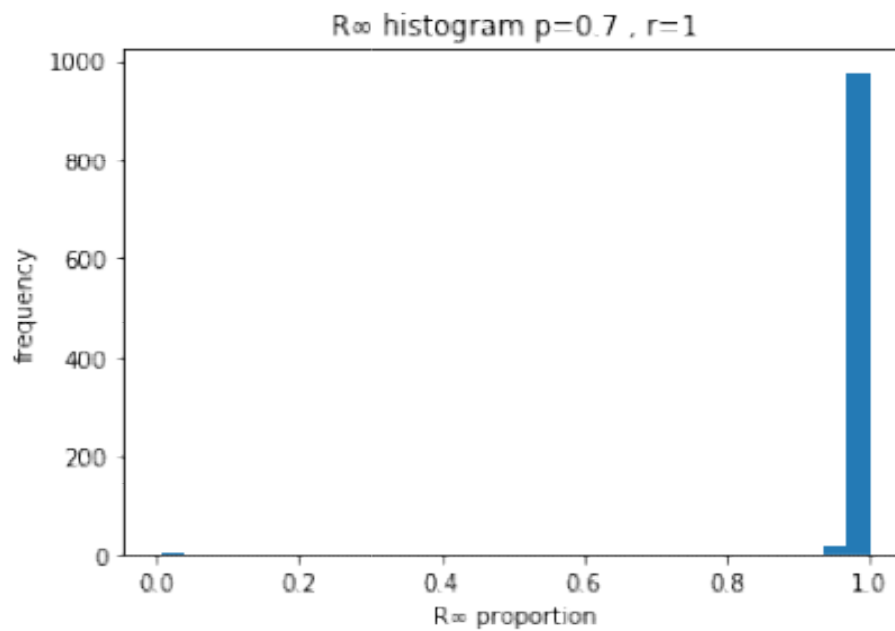


- $P=0.6, r=1, I_0 = 1$



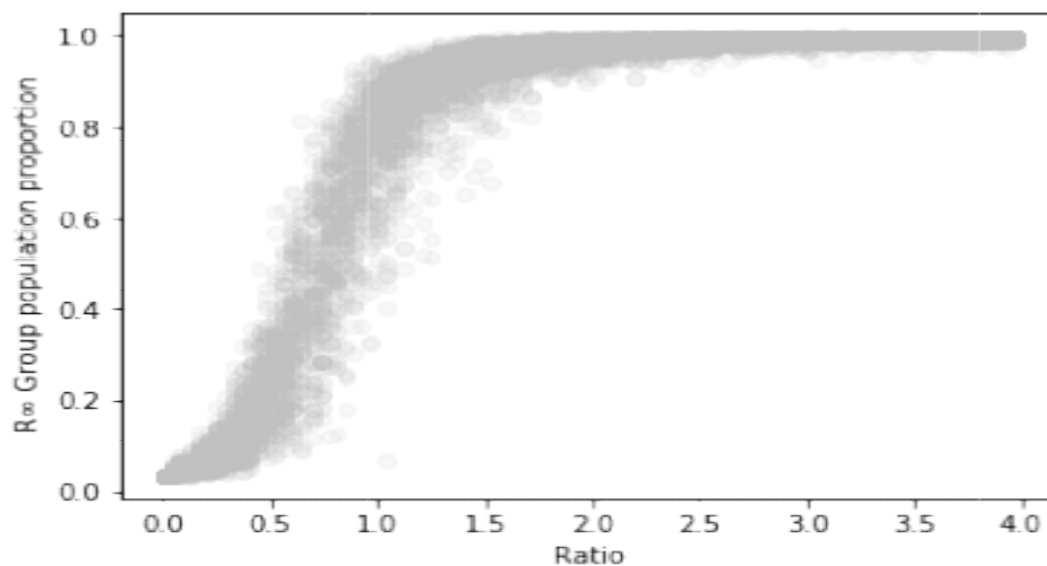
Curves tell all the secrets! Something different happens in this condition. Right and left peaks say to us that with $p = 0.6$ one of teams is much stronger (look at the example $p = 0.5$, infected society is now more powerful). Likewise we have some R_∞ near zero because it is somehow probable that the first infected person recovers and the system collapses.

- $P=0.7, r=1, I_0 = 1$



With $p = 0.7$ all is clear. It is similar to a match between **real Madrid** and Iran's second league (or premier league) team. Real Madrid is a super strong team (infected people) but in some defensive strategies Iranian teams can win (susceptible population) but it's obvious that in special number of matches Real Madrid will win probableier. There is no chance! There is peak near 1 which means that the entire group is recovered but we have a little portion of events near zero.

R_{∞} Distribution with different $\frac{p}{r}$ rate



For this curve, Ratio is changed from near zero to 4 in 100 steps then I added them to an array and collected for the curve. This curve is really good narrator. It tells us that there is a critical point on the ratio =0.5 .Before and after this point something interesting happens. There are dense population of nodes near zero and after 1. There is a good intuition which I've explained above. With smaller ratio, system collapses and with higher ratio disease spreads out and is epidemic. In ratio=0.5 there is a competition based on the random produced and it is near uniformly distributed between zero and 1.

****Other effective parameters**

- **I_0 (Initial infected percent)**

I think that this parameter is effective on the result so I changed it (1, 10 and 100). Then I've gathered the triple and plot all three condition. It definitely has effect and it changes the Behavior of nodes near critical point. When I increases the chance for being epidemic increases too. With small proportion of infected people they recover soon and the system collapses. By increasing I_0 all the people get the sickness. With higher proportion early stage recover doesn't happen.

