# Ensembles and statistics

Homework | Agent-based modelling, Konstanz, 2024

Henri Kauhanen

25 June 2024

In the lecture on network simulations, we simulated VL on Watts–Strogatz networks, using either $\beta = 0.1$ or $\beta = 0.5$ as the rewiring probability when the network is constructed. The question was whether one value of β leads to shorter durations of change than the other. The results we saw with 10 simulations per each value of β are illustrated in this plot:
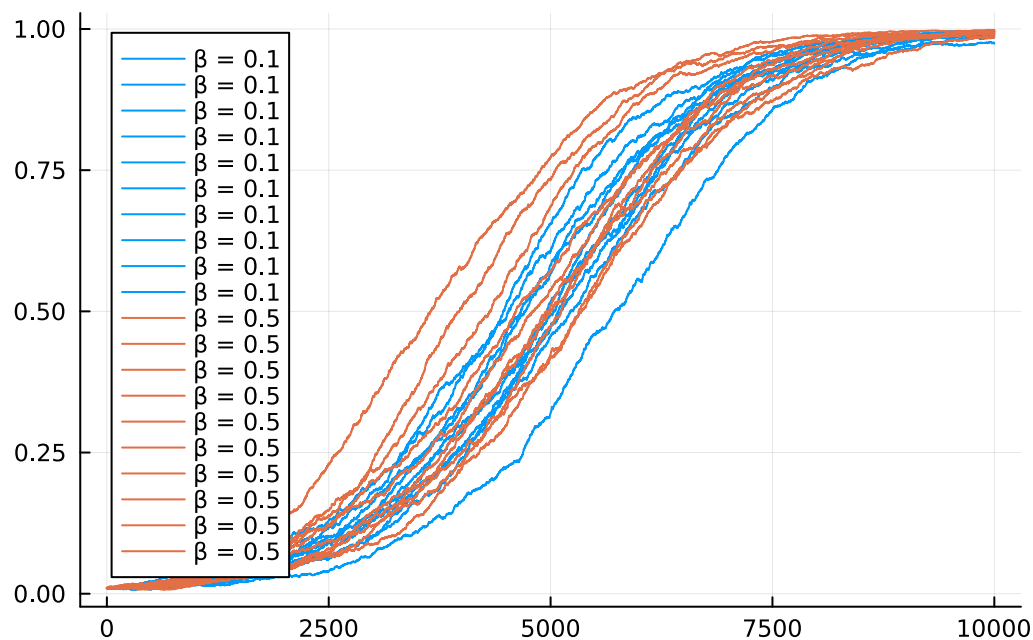
It looks like there may be a small difference in the average duration of change between the β values. However, it is difficult to be convinced about this: firstly, we have only 10 simulations per β, and secondly, all we have done so far is to visually inspect a picture.

In this homework exercise, I want you to think about the following questions:

1. How could you quantify the duration of a change using a single number? In other words, what sort of summary statistic can you use to decide whether one trajectory goes up earlier than another one? Try to go for the *simplest* such summary statistic.
2. Once you have such a number for each trajectory, you have a set of these numbers. What kind of statistical test could you use to decide whether the set of numbers for $\beta = 0.1$ is significantly different from the set of numbers for $\beta = 0.5$? (Hint: you want a test that compares two means from two samples.)
3. Once you have answers to the above questions, you can try and implement the following procedure:

    a. Instead of 10 simulations, use `ensemblerun!` to produce simulated trajectories for 100 repetitions for each β.
    b. Then figure out how to extract your summary statistic from these data.
    c. Finally, carry out the statistical test in order to make a decision.

In this exercise, the most important thing is to think about questions 1 and 2 above. If you struggle with the implementation (question 3), don't worry – I will show you how to do it in the next lecture.

> 💡 Tip
>
> If you do wish to work on the implementation, you will find the following useful:
>
> 1. The documentation for DataFrames.jl will help you figure out how to extract your summary statistic from the simulation data.
> 2. HypothesisTests.jl has functions for a number of statistical tests.
>
> Alternatively, if you do not wish to work in Julia, you can export your simulation data and continue working in R, Python, Excel, or some other software of your choice. The following saves the contents of a dataframe `df` in a file named `outfile.csv` as a comma-separated values file:
>
> ```
> CSV.write("outfile.csv", df)
> ```
>
> If you don't have the CSV.jl package yet, you'll have to install it first:
>
> ```
> using Pkg
> Pkg.add("CSV")
> ```