Focus on: Monte Carlo Tests

Modelling Complex Systems

Monte-Carlo tests

Q: Is the test statistic on our network, t^*, expected if network is drawn from the null distribution.

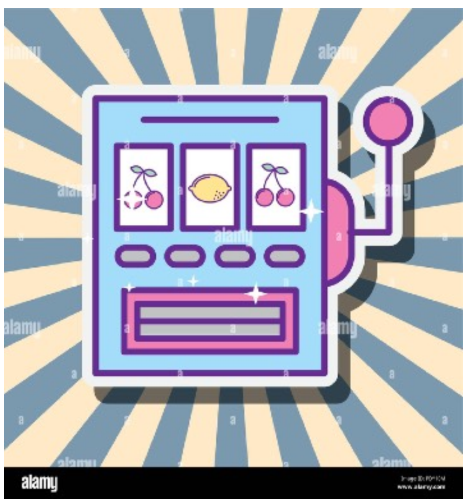
- Only need to be able to sample from the null distribution
- Works a bit like a proof by contradiction - start with a hypothesis (that you think is false), and get to a situation which would be unlikely if the hypothesis is true.

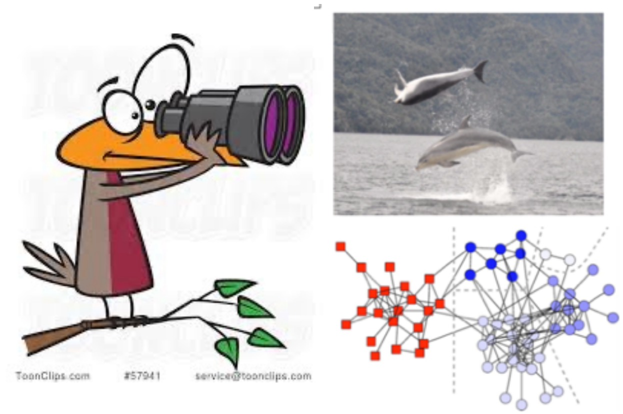
H_0 - hypothesis : - e.g. that observed network an observation from null distribution

Monte-Carlo tests

Want to say - observed network not observation from null distribution

- allow some error





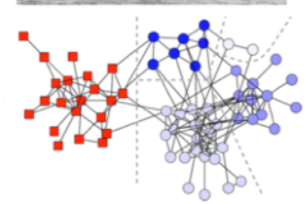
- Sample network from null distribution
- Observed Network

Monte-Carlo tests

H_0 - hypothesis : - e.g. that observed network an observation from null distribution

- Sample network from null distribution
- Observed Network

H_1 - reject hypothesis - e.g. that observed network not observation from null distribution



- Sample network from null distribution
- Observed Network

Monte-Carlo tests

If null hypothesis true all orderings of data are equally likely, the probability that the one you observe is among the top m is $m/(n+1)$

Also called ‘parametric bootstrapping’, ‘conditional uniform graph test’

Method for $\alpha = m/(n+1)$

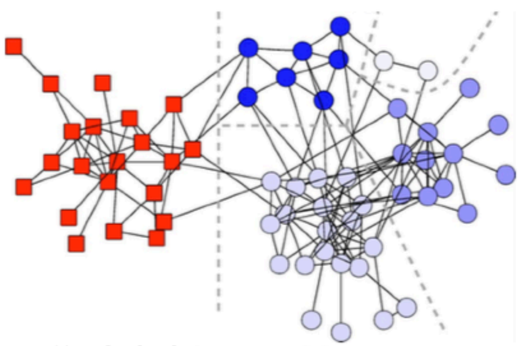
- sample n from the null distribution and calculate test statistic $t_1, \ldots t_n$,

e.g. sample n configuration models and calculate modularity score of each,

- order $t^*, t_1, \ldots, t_n$
- if t^* among top m values reject null hypothesis for distribution of network
- rule of thumb, take m at least 5.

Monte Carlo example

- with Erdos-Renyi random graph as null model.



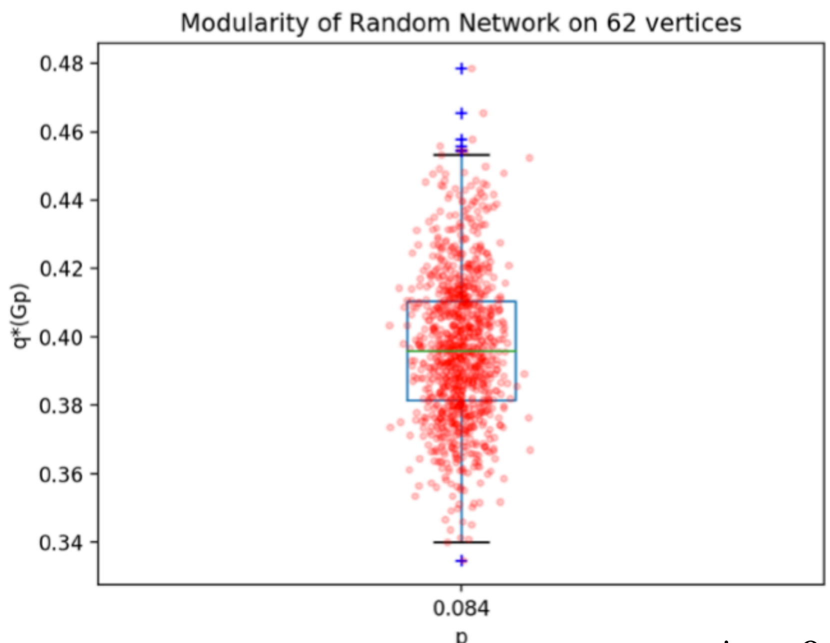
$\# \text{ dolphins} = 62$
$\# \text{ edges} = 159$

$q^* = 0.52$

8.4% of possible edges

- Calculate p so that null model has same expected degree: $p = \#\text{edges} / \binom{\#\text{dolphins}}{2}$

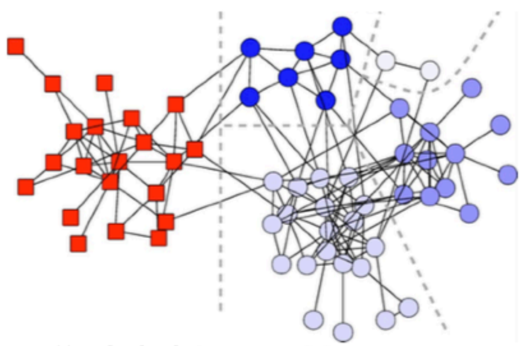
$q^*(\text{dolphins}) > q^*(\text{random network})??$



- The test statistic on our network is $t^* = 0.52$
- Red dots on graph above give test statistic on graphs sampled from the null model. Note t^* is greater than value of test statistic on any generated graph.

Monte Carlo example

- with Erdos-Renyi random graph as null model.

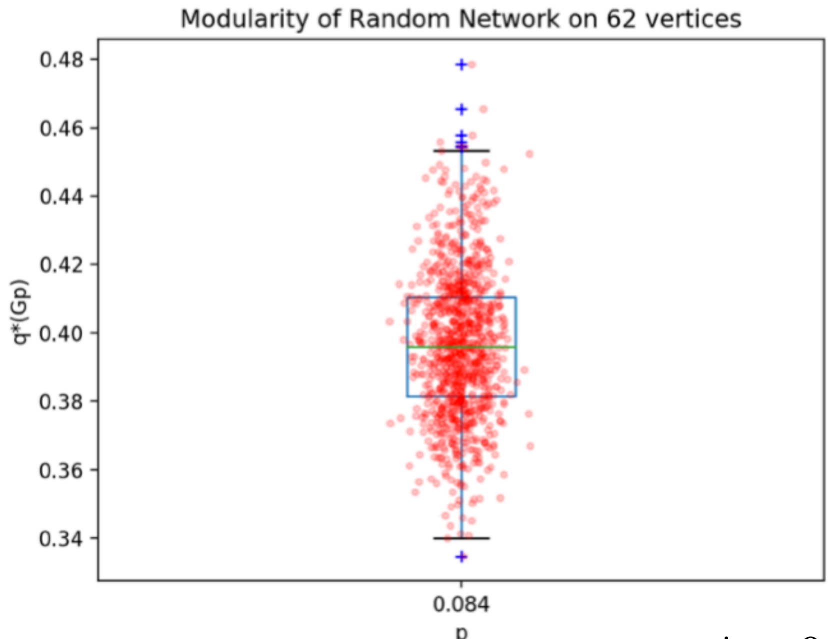


$\#$ dolphins $= 62$
$\#$ edges $= 159$

$q^* = 0.52$

8.4% of possible edges

$q^*(\text{dolphins}) > q^*(\text{random network})??$



- The test statistic on our network is $t^* = 0.52$
- Red dots on graph above give test statistic on graphs sampled from the null model. Note t^* is greater than value of test statistic on any generated graph.

- Calculate p so that null model has same expected degree: $p = \#\text{edges} / \binom{\#\text{dolphins}}{2}$

Monte-Carlo tests

Q: Is the test statistic on our network, t^*, expected if network is drawn from the null distribution.

e.g. modularity of network - compare network to configuration model same degree

- useful when we don’t know the distribution
- need to be able to sample from the null distribution
- discrete data has ties (break randomly and method still valid).

Monte-Carlo tests

Q: Is the test statistic on our network, t^*, expected if network is drawn from the null distribution.

e.g. modularity of network - compare network to configuration model same degree

Method for $\alpha = m/(n+1)$

- sample n from the null distribution and calculate test statistic $t_1, \dots t_n,$

e.g. sample n configuration models and calculate modularity score of each,

- order $t^*, t_1, \dots, t_n$
- if t^* among top m values reject null hypothesis for distribution of network

- rule of thumb, take m at least 5.

Monte-Carlo tests

If null hypothesis true all orderings of data are equally likely, the probability that the one you observe is among the top m is $m/(n+1)$

Also called ‘parametric bootstrapping’, ‘conditional uniform graph test’

Method for $\alpha = m/(n+1)$

- sample n from the null distribution and calculate test statistic $t_1, \ldots t_n$,

e.g. sample n configuration models and calculate modularity score of each,

- order $t^*, t_1, \ldots, t_n$
- if t^* among top m values reject null hypothesis for distribution of network

- rule of thumb, take m at least 5.

Monte-Carlo tests

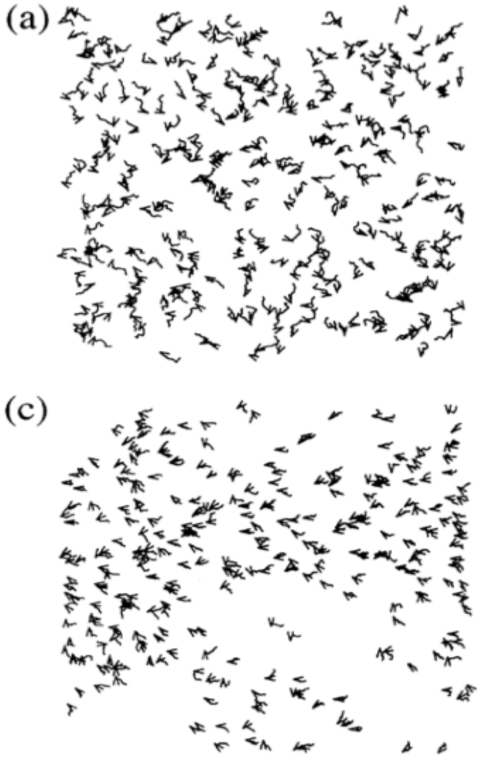
If null hypothesis true all orderings of data are equally likely, the probability that the one you observe is among the top m is $m/(n+1)$

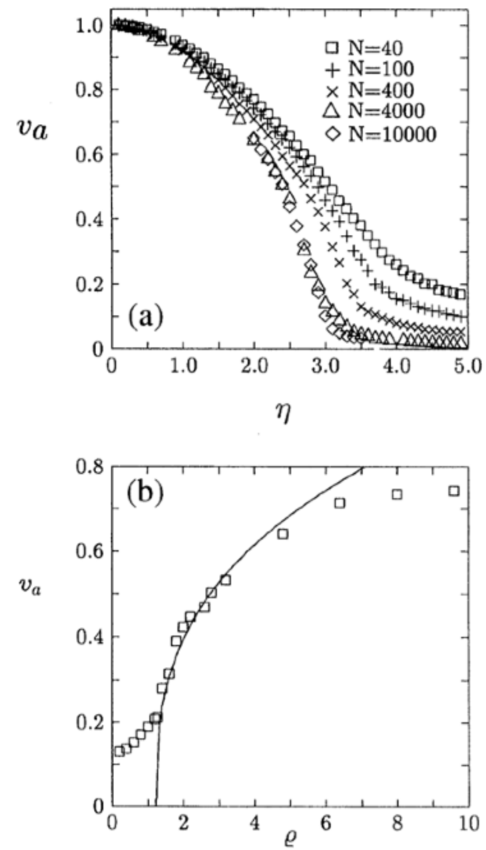
Also called ‘parametric bootstrapping’, ‘conditional uniform graph test’

Watch

- https://www.youtube.com/watch?v=QT2xj9k00q0
- 1:07-1:11 discusses Monte-Carlo test

Vicsek Model

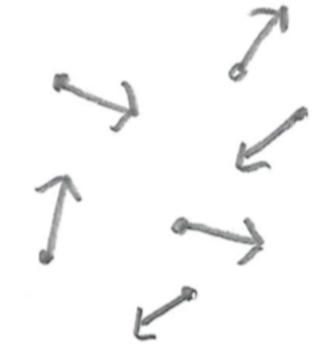




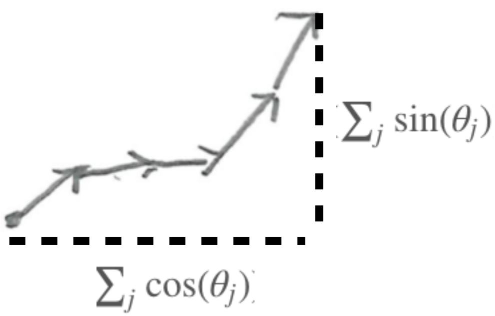
Vicsek et al., PRL 75 (1995)

Measure of Alignment: Polarisation

High polarisation



Low Polarisation



Polarisation of: $\theta_1, \theta_2, \ldots, \theta_N$

$$= \frac{1}{N}\sqrt{\left(\sum_j \sin(\theta_j)\right)^2 + \left(\sum_j \cos(\theta_j)\right)^2}$$