Economic Networks

Theory and Empirics

Giorgio Fagiolo
Laboratory of Economics and Management (LEM)
Sant'Anna School of Advanced Studies, Pisa, Italy

http://www.lem.sssup.it/fagiolo/
giorgio.fagiolo@sssup.it

Lecture 6

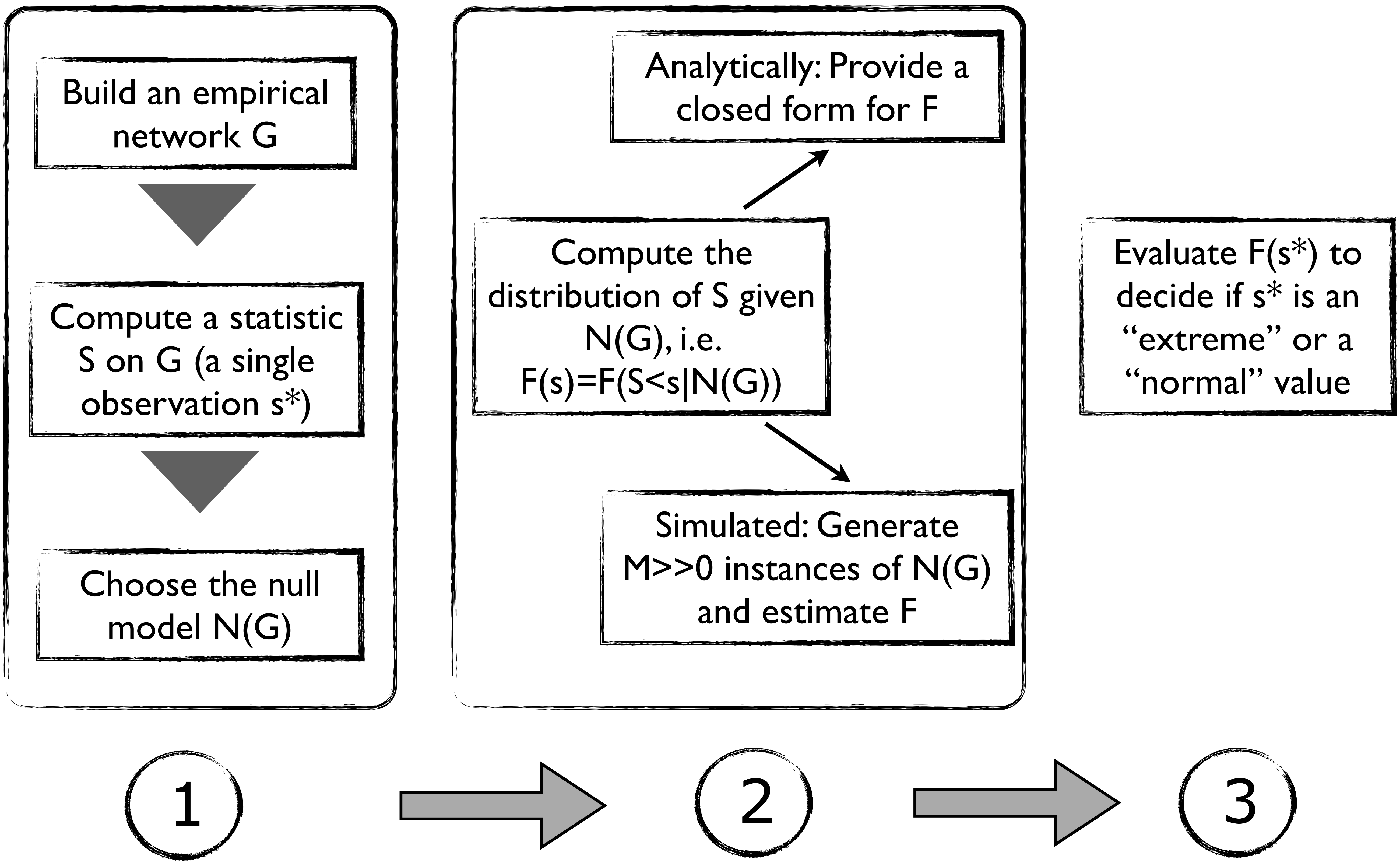
This Lecture

- What is a network? Examples of networks
- Why networks are important for economists?
- Networks and graphs
- Measures and metrics on networks
- Distributions of metrics and measures in large networks
- Models of network formation
- Null statistical network models
- Economic applications

Null Statistical Network Models

- Inference on empirical network properties
  - ✓ Suppose we have observed a network G and we have computed a set of interesting network statistics on G, say $s_1(G),...,s_k(G)$
  - ✓ These may be: the clustering coefficient or the correlation between ND and ANND, or between NS and ANNS
- Problem: how can we say something about whether these observed values are large or small?
  - ✓ We need statistical benchmarks (null models) to assess the distribution of any given statistics given the null model at hand
  - ✓ Many ways to do it: one must choose the most appropriate null model, i.e. decide which properties of the observed graph we want to preserve
  - ✓ The null model generates maximally random graphs satisfying the selected constraints (preserved properties)
- Example: Maximally random graphs given
  - ✓ Binary graphs: Density only, degree distribution, degree sequence, etc.
  - ✓ Weighted graphs: Weight distribution, binary topology, etc.

Design of the Experiment



Binary Networks: Example 1

- Preserving density only
  - ✓ Compute density d on observed graph G with N nodes and L links
  - ✓ Generate a Poisson random graph with density d

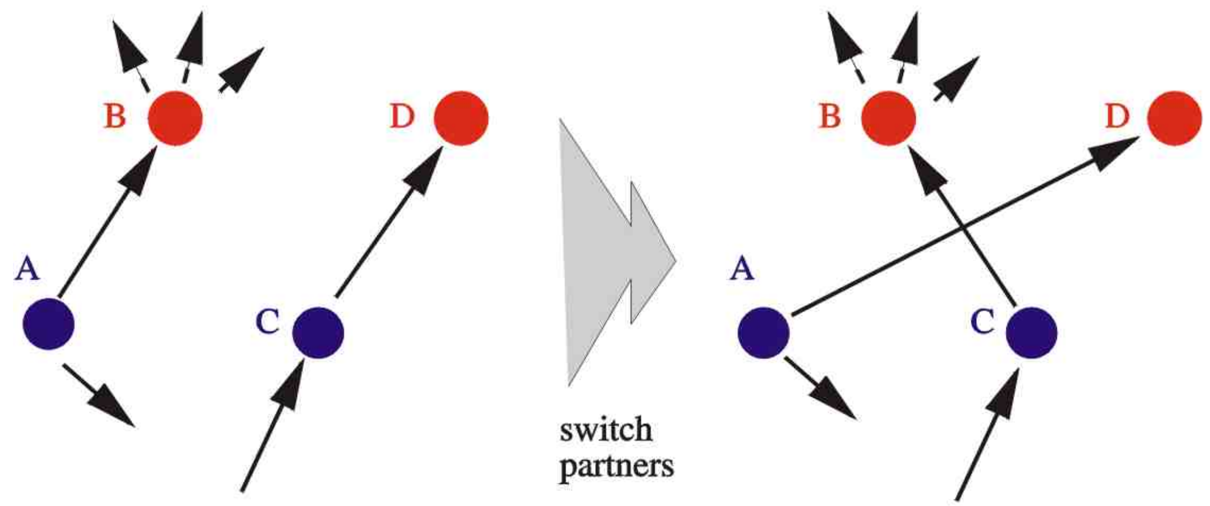
- Two alternatives
  - ✓ Average or exact density: using G(N,p) or G(N,m) models, where p=d and m=L (number of links in G)

- What is not preserved
  - ✓ Actual links, degree sequence and distribution all change

- Extension to digraphs
  - ✓ Generating Poisson random digraphs in such a way to preserve either the total number of links (in- and out- density change) or the exact number of in and out links in the observed graph

Binary Networks: Example 2

- Degree-preserving random rewiring (Maslov & Sneppen, 2002)
  - ✓ Preserve exactly the degree sequence of the observed binary network



A pair of directed edges A->B and C->D is randomly selected. These edges are then rewired in such a way that A becomes connected to D, while C to B, provided that none of these edges already exist in the network, in which case the rewiring step is aborted and a new pair of edges is selected. Note that the above rewiring algorithm conserves both the in- and out-degree of each individual node (and degrees if the graph is undirected)

- Problems
  - ✓ To get a single instance we need many such rewirings (at least 4L): this takes time

Weighted Networks

- Random reshuffling preserving density only
  - ✓ Generate Poisson random graphs with exact observed density d
  - ✓ Randomly reshuffle existing (positive) link weights across the randomly generated instance of binary topology using G(N,m)

- Maslov-Sneppen for WUNs and WDNs
  - ✓ Maslov-Sneppen rewiring algorithm works perfectly also for WUNs and WDNs
  - ✓ Just move the link-weight together with the link that is rewired
  - ✓ This does not preserves strength (in/out/tot) sequence (check it)

- Preserving weight distribution and binary topology
  - ✓ Reshuffle weights among existing binary links
  - ✓ This preserves link-weight distributions and binary structure (A)
  - ✓ Therefore also the degree sequence is preserved exactly

Alternative Null Models (1)

- Configuration model
  - ✓ An alternative algorithm to generate random (binary undirected) networks with a given degree sequence $\{k_1,\dots,k_N\}$
  - ✓ Suppose that $\{k_1,\dots,k_N\}$ is graphic, i.e. it is a feasible degree sequence of a graph (e.g. sum of all degrees is even). This is automatically satisfied if the sequence comes from an empirically-observed graph

- Algorithm
  - ✓ Construct a sequence where node i is listed k_i times for all i

$$\underbrace{1,1,\dots,1}_{k_1 \text{ times}} \quad \underbrace{2,2,\dots,2}_{k_2 \text{ times}} \quad \cdots \quad \underbrace{N,N,\dots,N}_{k_N \text{ times}}$$

  - ✓ Randomly pick any two elements from the list and form a link between the nodes corresponding to those entries.
  - ✓ Delete those entries from the list and repeat until we get to the end (note: if the sum of degrees were odd we will remain with a single node)

Alternative Null Models (2)

- Output of the configuration model
  - ✓ A random graph where the degree sequence is preserved
  - ✓ Problems: multiple self-loops and multi-edges are not ruled out
  - ✓ Therefore the configuration model generates multi graphs

- Ways out
  - ✓ Delete multi-edges and all self loops: this destroys degree sequence but if multi edges are not that frequent the resulting degree distribution is close to the observed one
  - ✓ Employ null models preserving degree sequence only on average

- Expected-degree models (and beyond)
  - ✓ Chung-Lu (2002)
  - ✓ Squartini, Garlaschelli (2011) , Squartini, Fagiolo, Garlaschelli (2011a,b)

Alternative Null Models (3)

- Chung-Lu model
  - ✓ Start with the observed degree sequence $\{k_1,...,k_N\}$ and an empty graph
  - ✓ Go through each pair of nodes and form a link with probability

$$\frac{k_i k_j}{\sum_h k_h} = \frac{k_i k_j}{2L}$$

- Notice
  - ✓ Self-loops are still allowed with probability $k_i^2/2L$ but no multiple edges between different nodes
  - ✓ To have well-defined probabilities it must be that $\max\{k_i\}<\text{Sqrt}(2L)$. It can be checked that if the degree sequence is very broad this condition is not satisfied and the ratios above are larger than one. This means that the observed degree sequence cannot be replicated on average.
  - ✓ What is the probability of connecting i and j in the configuration model? Prove that is equal to $k_i k_j/(2L-1)$, i.e. equal to the one in Chung-Lu model for large L

Economic Interpretation of Null Models

- Null models provide a statistical benchmark to compare observed network statistics
  - ✓ Almost no economics in them
  - ✓ Why are they useful in economics?

- Null statistical models in economics
  - ✓ Suppose a given observation is in line with what predicted by a given null statistical model. Then that value of the statistics does not require additional economic explanations. It can be simply the outcome of randomness. If we provide an economic model that reproduces that observation then that model could not be selected against the null random model.
  - ✓ Suppose instead a given observation is found to be an extreme value for the null model at hand. Then the null model must be rejected because it cannot explain that observation. We need to find an explanation elsewhere, probably in the economic realm.

Next Lecture

- What is a network? Examples of networks
- Why networks are important for economists?
- Networks and graphs
- Measures and metrics on networks
- Distributions of metrics and measures in large networks
- Models of network formation
- Null statistical network models
- Economic applications