Universitat Pompeu Fabra
Barcelona

Seminar 4 · Networks, Crowds and Markets

Centrality Measures and Social Networks

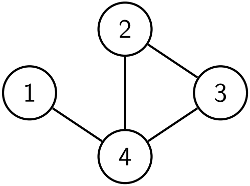
Warm-up · Building a small $G(4, p)$ graph

We consider an Erdős–Rényi model $G(4, p)$ with $p = \frac{2}{3}$. There are $\binom{4}{2} = 6$ possible edges.

1. List all possible pairs of nodes: 12, 13, 14, 23, 24, 34.
2. Roll a fair die once per potential edge (six times). Connect the pair if the result is 1–4 (probability 2/3).
3. Draw the resulting graph.

Example outcome:

pair	12	13	14	23	24	34
die	5	6	2	3	3	1
edge	0	0	1	1	1	1



Compare our experiment results with expected values:

$$\mathbb{E}[L] = p\binom{4}{2} = 4, \qquad \mathbb{E}[\deg(v)] = (N-1)p = \tfrac{2}{3} \times 3 = 2.$$

Midterm review

I went through the solutions of the midterm.

Additional exercises

Exercise · Independence in Erdős–Rényi

Show that in $G(N, p)$:

- Any two different edges are independent.
- Degrees of two nodes are not independent (why?).

Compute:

$$\mathrm{Cov}(\deg(i), \deg(j)) \quad \text{for } i \neq j.$$

Exercise · ER as an Exponential Family

[This was discussed in the lecture] Show that the Erdős–Rényi model

$$P(Y = y) \propto \exp(\theta \cdot s(y)), \quad s(y) = \#\text{edges}(y),\ \theta = \log \frac{p}{1-p}.$$

is an exponential family.

Identify:

- the sufficient statistic,
- the natural parameter,
- the log-partition function.

Exercise · The p_2 Model (undirected)

The p_2 model extends Erdős–Rényi by allowing each node i to have its own random sociality effect α_i:

$$\text{logit} \Pr(Y_{ij} = 1 \mid \alpha_i, \alpha_j) = \theta + \alpha_i + \alpha_j.$$

1. Interpret the roles of θ and α_i.
2. How does the variance σ^2 control network heterogeneity?
3. What does the model reduce to when $\alpha_i = 0$ for all i?
4. Compare this to a mixed-effects logistic regression model.

Exercise · Latent Space Model

In a latent-space model, each node i is embedded at position $z_i \in \mathbb{R}^2$, and edges are independent with

$$\operatorname{logit} P(Y_{ij} = 1) = \alpha - \|z_i - z_j\|.$$

1. Explain how distance $\|z_i - z_j\|$ affects connection probability.
2. What network features can this model capture that $G(N, p)$ or p_2 cannot?
3. Suggest one real-world system where such geometry might be natural.