NETWORK PARAMETERIZATION

Martina Morris, Ph.D.
Steven M. Goodreau, Ph.D.
Samuel M. Jenness, Ph.D.

Outline

- Model specification and parameterization issues
  - Balance
  - Model specification and degrees of freedom
- Practice with selecting terms and calculating target statistics

Balance

- The idea that the **number** of contacts group A has with group B must equal the **number** that group B has with group A

- Does **not** necessarily mean that the **proportion** of group A's contacts that are with group B equals the **proportion** of group B's contacts that are with group A

- For example, in Seattle the proportion of Black persons' ties that are with White persons is much higher that the proportion of White persons' ties that are with Black persons. Why?
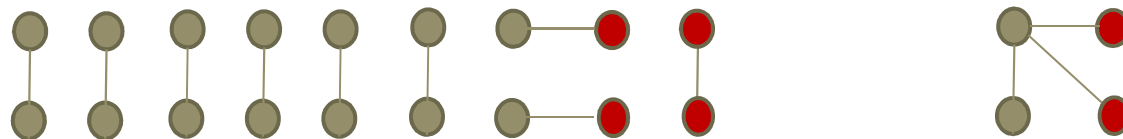
Balance: network models

- E.g. if you are building a purely heterosexual model
  - In the real world, in any population:

    # of relationships/acts that females have with males =

    # of relationships/acts males have with females

  - But this may not be exactly true in **egocentric** data
    - (Random) sampling error
    - Bias (sex ratio of sample does not equal empirical sex ratio, female sex workers are under-sampled)
    - Misreporting (e.g. females may under-report)

  - Nevertheless, one needs to be explicit about balance in the target statistics

Balance in STERGMs

- Occurs initially in the construction of the target stats, and must involve explicitly thinking about data sources. E.g.:

  - Number of ties (pop size * mean degree with other group) must always balance between two contacting groups

  - Imagine a purely heterosexual population (and sample) that both have a 1:1 sex ratio

  - Equal pop sizes implies that mean degree must be equal

  - Males report mean degree of 0.74, females report 0.68

  - You must choose whether to use 0.68, 0.74, 0.71, or something else when calculating target stat

Balance in STERGMs

- Remember: the statnet package `ergm.ego` exists to handle much of this for you.

- But it is also worth understanding what is going on under the hood and what the assumptions are

    - It's always good to understand exactly what you're modeling

    - You may need to communicate this to others (mentors, reviewers, readers)

    - You might want to calculate the statistics yourself

Balance in STERGMs

- Note: once estimation is done, and simulation begins then balance will happen automatically forever, even when we introduce vital dynamics

- This is because the target stats have been converted into parameters based in log-odds

- This is true no matter the nature of complexity of the nodal dynamics

Balance in STERGMs

- Quick quiz:
  - Purely heterosexual population
  - Females have mean sex partner degree of 0.8

  - Males must have mean sex partner degree of:

    A. 0.8

    B. 0.4

    C. 0.89

    D. Not enough information

Balance in STERGMs

- Quick quiz:
  - Purely heterosexual population
  - Females have mean sex partner degree of 0.8
  - There are 200 females and 180 males
 - Males must have mean sex partner degree of:

    A. 0.8

    B. 0.4

    C. 0.89

    D. Not enough information

$$= \frac{200 \times 0.8}{180}$$

Balance

- Balance only applies to numbers of **ties**.

- It can be easy to mistakenly over-apply the concept of balance.

- For instance, imagine a model that considers relational concurrency in heterosexual relationships.

  - Assume equal sex ratio

  - Assume 20% of men report having concurrent partnerships

  - What does that tell us about the % of women having concurrent partnerships?

  - Very little!
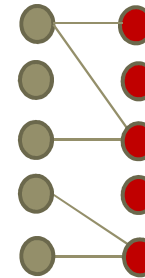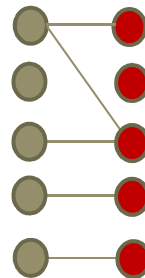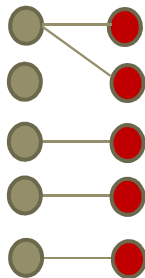
Degrees of freedom

- You can only use as many terms/target stats as you have degrees of freedom
- Can be tricky to identify
- E.g. heterosexual degree distributions
  - You are estimating a model on a network containing 250 females and 250 males
  - You have already included an edges term with target stat 165
  - You have included a constraint that nobody can have more than 3 edges at one time
  - How many more sex–specific degree terms/target stats can you add before your model is fully specified?

Degrees of freedom

Deg	M	F
0	M_0	F_0
1	M_1	F_1
2	M_2	F_2
3	M_3	F_3

Additional constraints:
- 250 males total
- 250 females total
- 165 edges total

- Given that nobody can have degree >3, there are 8 cells that can be filled in.

Degrees of freedom

Deg	M	F
0	M_0	F_0
1	M_1	F_1
2	M_2	F_2
3	M_3	F_3
Total N	250	250
Total Pships	165	165

Additional constraints:
- 250 males total
- 250 females total
- 165 edges total

- Given that nobody can have degree >3, there are 8 cells that can be filled in.

Degrees of freedom

Deg	M	F
0	M_0	F_0
1	M_1	F_1
2	M_2	F_2
3	M_3	F_3
Total N	250	250
Total Pships	165	165

Deg	M	F
0	120	92
1	104	153
2	17	3
3	9	2
Total N	250	250
Total Pships	165	165

- Given that nobody can have degree >3, there are 8 cells that can be filled in.

  - $M_0 + M_1 + M_2 + M_3 = 250$ $M_0 + 104 + 17 + 9 = 250$
  - $F_0 + F_1 + F_2 + F_3 = 250$ $F_0 + 153 + 3 + 2 = 250$
  - $M_1 + 2M_2 + 3M_3 = 165$ $M_1 + 2(17) + 3(9) = 165$
  - $F_1 + 2F_2 + 3F_3 = 165$ $F_1 + 2(3) + 3(2) = 165$

- So users can specify **at most** 2 male degree terms and 2 female degree terms

Practice

- Time for another lab!