Modeling Cohesive subgroups

Mark S. Handcock

handcock@ucla.edu

Statistical Analysis of Networks

October 22, 2024

Modeling Cohesive subgroups

- Suppose $1, \ldots, n$ are categorized into g known groups. and the relationship is directed
- Y_{ij} are independent but depend on the category of the actors
- Suppose actors have separate propensities to "mix" with other groups.

$$x_{ij}^{ab} = \begin{cases} 1 & \text{the category of } i \text{ is } a \text{ and the category of } j \text{ is } b \\ 0 & \text{otherwise} \end{cases}$$

Modeling Cohesive subgroups

- Suppose $1, \ldots, n$ are categorized into g known groups. and the relationship is directed
- Y_{ij} are independent but depend on the category of the actors
- Suppose actors have separate propensities to "mix" with other groups.

$$x_{ij}^{ab} = \begin{cases} 1 & \text{the category of } i \text{ is } a \text{ and the category of } j \text{ is } b \\ 0 & \text{otherwise} \end{cases}$$

$$\log \text{odds}(Y_{ij} = 1 | x_{ij}^{ab}) = \eta_{ab} x_{ij}^{ab} \qquad \forall \ i, j = 1, \ldots, n$$

```
ergm(net1 ~ nodemix("color"))
```

Example: Are the Monk's in cohesive groups?

```
Formula:    samplike ~ nodemix("group")
                              Estimate Std.Err p-value
mix.group.Loyal.Loyal            0.1911   0.3100 0.538185
mix.group.Outcasts.Loyal        -3.2958   1.0183 0.001347 **
mix.group.Turks.Loyal           -2.1748   0.4719 < 1e-04 ***
mix.group.Loyal.Outcasts        -2.5649   0.7338 0.000545 ***
mix.group.Outcasts.Outcasts      1.6094   0.7746 0.038589 *
mix.group.Turks.Outcasts        -3.2958   1.0183 0.001347 **
mix.group.Loyal.Turks           -1.4917   0.3689 < 1e-04 ***
mix.group.Outcasts.Turks        -1.0986   0.4364 0.012354 *
mix.group.Turks.Turks            0.9163   0.3416 0.007714 **
---
    Null  Deviance: 424.21  on 306  degrees of freedom
 Residual Deviance: 261.10  on 297  degrees of freedom
          Deviance: 163.10  on   9  degrees of freedom
```

3

Comparing Models for Cohesive subgroups

- The analysis of deviance has been formalized in the anova() function.

```
> anova(fit, fit.mix)
Analysis of Variance Table

Model 1: samplike ~ edges
Model 2: samplike ~ nodemix("group")
        Df Deviance Resid. Df Resid. Dev Pr(>|Chisq|)
NULL                      306     424.21
1        1   57.02        305     367.18  4.307e-14 ***
2        8  106.08        297     261.10  < 2.2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*'
```

Example: Are the Monk's in cohesive groups?

```
> fit.match <- ergm(samplike ~ nodematch("group"))
> summary(fit.match)
=====================
Formula:   samplike ~ edges + nodematch("group")
              Estimate Std. Error MCMC %  p-value
edges         -2.0015     0.2131     NA   <1e-04 ***
nodematch.group  2.6481   0.3026     NA   <1e-04 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05

    Null Deviance: 424.21  on 306  degrees of freedom
 Residual Deviance: 276.86  on 304  degrees of freedom
         Deviance: 147.35  on   2  degrees of freedom

AIC: 280.86     BIC: 288.31
```

5

How does the Homophily model fit?

```
> fit.homo.diff <- ergm(samplike ~
          edges + nodematch("group", diff=TRUE))
> summary(fit.homo.diff)

ergm(formula = samplike ~ edges + nodematch("group", diff=TRUE))

Maximum Likelihood Results:
                        Estimate Std. Error z value Pr(>|z|)
edges                    -2.0015     0.2131  -9.393  <1e-04 ***
nodematch.group.Loyal     2.1925     0.3762   5.828  <1e-04 ***
nodematch.group.Outcasts  3.6109     0.8034   4.495  <1e-04 ***
nodematch.group.Turks     2.9178     0.4026   7.248  <1e-04 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

     Null Deviance: 424.2  on 306  degrees of freedom
 Residual Deviance: 272.2  on 302  degrees of freedom

AIC: 280.2  BIC: 295.1  (Smaller is better.)
```

Homophily model compared to Cohesive subgroups

```
> anova(fit, fit.homo, fit.diff.homo)
Analysis of Variance Table

Model 1: samplike ~ edges
Model 2: samplike ~ edges + nodematch("group")
Model 3: samplike ~ edges + nodematch("group", diff=T)

     Df Deviance Resid. Df Resid. Dev Pr(>|Chisq|)
NULL                        306       424.21
1    1    57.02      305       367.18    4.307e-14 ***
2    1    90.32      304       276.86    < 2.2e-16 ***
3    2     4.64      302       272.22       0.0983 .
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05
```

Homophily model compared to Cohesive subgroups

```
> anova(fit, fit.homo, fit.mix)
Analysis of Variance Table

Model 1: samplike ~ edges
Model 2: samplike ~ edges + nodematch("group")
Model 3: samplike ~ nodemix("group")

     Df Deviance Resid. Df Resid. Dev Pr(>|Chisq|)
NULL                         306     424.21
1   1    57.02       305     367.18    4.307e-14 ***
2   1    90.32       304     276.86    < 2.2e-16 ***
3   7    15.76       297     261.10      0.02741 *
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05
```