Testing and model evaluation

567 Statistical analysis of social networks

Peter Hoff

Statistics, University of Washington

Testing and rejecting models

Descriptive network analysis: Computation of

- graph level statistics: density, degree distribution, centralization
- node level statistics: degrees, centralities
- covariate effects: relative densities and odds ratios

What conclusions can we draw from such statistics?

- Are observed statistics large or small?
  - as compared to other observed networks?
  - as compared to hypothetical networks?
- Are observed statistics consistent with a theoretical model?
  - what model is appropriate for comparison?
  - how are comparisons made?

Testing and rejecting models

Descriptive network analysis: Computation of

- graph level statistics: density, degree distribution, centralization
- node level statistics: degrees, centralities
- covariate effects: relative densities and odds ratios

What conclusions can we draw from such statistics?

- Are observed statistics large or small?
    - as compared to other observed networks?
    - as compared to hypothetical networks?
- Are observed statistics consistent with a theoretical model?
    - what model is appropriate for comparison?
    - how are comparisons made?

Testing and rejecting models

Descriptive network analysis: Computation of

- graph level statistics: density, degree distribution, centralization
- node level statistics: degrees, centralities
- covariate effects: relative densities and odds ratios

What conclusions can we draw from such statistics?

- Are observed statistics large or small?
    - as compared to other observed networks?
    - as compared to hypothetical networks?
- Are observed statistics consistent with a theoretical model?
    - what model is appropriate for comparison?
    - how are comparisons made?

Testing and rejecting models

Descriptive network analysis: Computation of

- graph level statistics: density, degree distribution, centralization
- node level statistics: degrees, centralities
- covariate effects: relative densities and odds ratios

What conclusions can we draw from such statistics?

- Are observed statistics large or small?
  - as compared to other observed networks?
  - as compared to hypothetical networks?
- Are observed statistics consistent with a theoretical model?
  - what model is appropriate for comparison?
  - how are comparisons made?

Testing and rejecting models

Descriptive network analysis: Computation of

- graph level statistics: density, degree distribution, centralization
- node level statistics: degrees, centralities
- covariate effects: relative densities and odds ratios

What conclusions can we draw from such statistics?

- Are observed statistics large or small?
    - as compared to other observed networks?
    - as compared to hypothetical networks?
- Are observed statistics consistent with a theoretical model?
    - what model is appropriate for comparison?
    - how are comparisons made?

Testing and rejecting models

Descriptive network analysis: Computation of

- graph level statistics: density, degree distribution, centralization
- node level statistics: degrees, centralities
- covariate effects: relative densities and odds ratios

What conclusions can we draw from such statistics?

- Are observed statistics large or small?
  - as compared to other observed networks?
  - as compared to hypothetical networks?
- Are observed statistics consistent with a theoretical model?
  - what model is appropriate for comparison?
  - how are comparisons made?

Testing and rejecting models

Descriptive network analysis: Computation of

- graph level statistics: density, degree distribution, centralization
- node level statistics: degrees, centralities
- covariate effects: relative densities and odds ratios

What conclusions can we draw from such statistics?

- Are observed statistics large or small?
    - as compared to other observed networks?
    - as compared to hypothetical networks?
- Are observed statistics consistent with a theoretical model?
    - what model is appropriate for comparison?
    - how are comparisons made?

Testing and rejecting models

Descriptive network analysis: Computation of

- graph level statistics: density, degree distribution, centralization
- node level statistics: degrees, centralities
- covariate effects: relative densities and odds ratios

What conclusions can we draw from such statistics?

- Are observed statistics large or small?
  - as compared to other observed networks?
  - as compared to hypothetical networks?
- Are observed statistics consistent with a theoretical model?
  - what model is appropriate for comparison?
  - how are comparisons made?

Testing and rejecting models

Descriptive network analysis: Computation of

- graph level statistics: density, degree distribution, centralization
- node level statistics: degrees, centralities
- covariate effects: relative densities and odds ratios

What conclusions can we draw from such statistics?

- Are observed statistics large or small?
    - as compared to other observed networks?
    - as compared to hypothetical networks?
- Are observed statistics consistent with a theoretical model?
    - what model is appropriate for comparison?
    - how are comparisons made?

Testing and rejecting models

Descriptive network analysis: Computation of

- graph level statistics: density, degree distribution, centralization
- node level statistics: degrees, centralities
- covariate effects: relative densities and odds ratios

What conclusions can we draw from such statistics?

- Are observed statistics large or small?
  - as compared to other observed networks?
  - as compared to hypothetical networks?
- Are observed statistics consistent with a theoretical model?
  - what model is appropriate for comparison?
  - how are comparisons made?

Testing and rejecting models

Descriptive network analysis: Computation of

- graph level statistics: density, degree distribution, centralization
- node level statistics: degrees, centralities
- covariate effects: relative densities and odds ratios

What conclusions can we draw from such statistics?

- Are observed statistics large or small?
  - as compared to other observed networks?
  - as compared to hypothetical networks?
- Are observed statistics consistent with a theoretical model?
  - what model is appropriate for comparison?
  - how are comparisons made?

Testing and rejecting models

Descriptive network analysis: Computation of

- graph level statistics: density, degree distribution, centralization
- node level statistics: degrees, centralities
- covariate effects: relative densities and odds ratios

What conclusions can we draw from such statistics?

- Are observed statistics large or small?
  - as compared to other observed networks?
  - as compared to hypothetical networks?
- Are observed statistics consistent with a theoretical model?
  - what model is appropriate for comparison?
  - how are comparisons made?

Testing and rejecting models

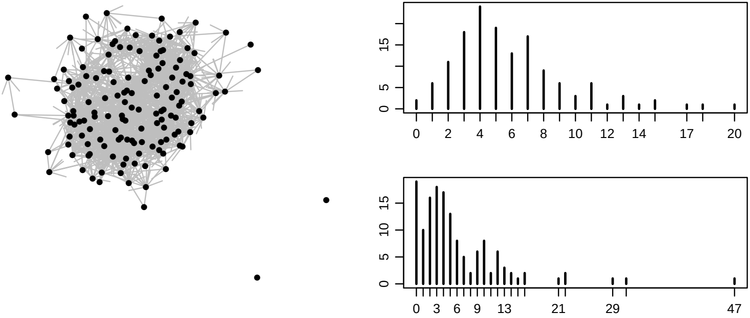
Descriptive network analysis: Computation of

- graph level statistics: density, degree distribution, centralization
- node level statistics: degrees, centralities
- covariate effects: relative densities and odds ratios

What conclusions can we draw from such statistics?

- Are observed statistics large or small?
  - as compared to other observed networks?
  - as compared to hypothetical networks?
- Are observed statistics consistent with a theoretical model?
  - what model is appropriate for comparison?
  - how are comparisons made?

Example: Girls friendships



```
mean(Y,na.rm=TRUE)

## [1] 0.04088967

Cd(Y)

## [1] 0.1003644

Cd(t(Y))

## [1] 0.2918349
```

Covariate effects

We also have data on

- gpa: `hgpa` = indicator of above-average gpa;
- smoking behavior: `hsmoke` = indicator of above-average smoking behavior.

```
p.smoke[2]/p.smoke[1]

## [1] 0.9638469

p.smoke[3]/p.smoke[1]

## [1] 1.008332

( p.smoke[1]*p.smoke[4] ) / ( p.smoke[2]*p.smoke[3] )

## [1] 1.248783

###

p.gpa[2]/p.gpa[1]

## [1] 0.3868937

p.gpa[3]/p.gpa[1]

## [1] 0.5379664

( p.gpa[1]*p.gpa[4] ) / ( p.gpa[2]*p.gpa[3] )

## [1] 5.754995
```

Covariate effects

We also have data on

- gpa: `hgpa` = indicator of above-average gpa;
- smoking behavior: `hsmoke` = indicator of above-average smoking behavior.

```
p.smoke[2]/p.smoke[1]

## [1] 0.9638469

p.smoke[3]/p.smoke[1]

## [1] 1.008332

( p.smoke[1]*p.smoke[4] ) / ( p.smoke[2]*p.smoke[3] )

## [1] 1.248783

###

p.gpa[2]/p.gpa[1]

## [1] 0.3868937

p.gpa[3]/p.gpa[1]

## [1] 0.5379664

( p.gpa[1]*p.gpa[4] ) / ( p.gpa[2]*p.gpa[3] )

## [1] 5.754995
```

Covariate effects

We also have data on

- gpa: `hgpa` = indicator of above-average gpa;
- smoking behavior: `hsmoke` = indicator of above-average smoking behavior.

```
p.smoke[2]/p.smoke[1]

## [1] 0.9638469

p.smoke[3]/p.smoke[1]

## [1] 1.008332

( p.smoke[1]*p.smoke[4] ) / ( p.smoke[2]*p.smoke[3] )

## [1] 1.248783

###

p.gpa[2]/p.gpa[1]

## [1] 0.3868937

p.gpa[3]/p.gpa[1]

## [1] 0.5379664

( p.gpa[1]*p.gpa[4] ) / ( p.gpa[2]*p.gpa[3] )

## [1] 5.754995
```

Covariate effects

We also have data on

- gpa: `hgpa` = indicator of above-average gpa;
- smoking behavior: `hsmoke` = indicator of above-average smoking behavior.

```
p.smoke[2]/p.smoke[1]

## [1] 0.9638469

p.smoke[3]/p.smoke[1]

## [1] 1.008332

( p.smoke[1]*p.smoke[4] ) / ( p.smoke[2]*p.smoke[3] )

## [1] 1.248783

###

p.gpa[2]/p.gpa[1]

## [1] 0.3868937

p.gpa[3]/p.gpa[1]

## [1] 0.5379664

( p.gpa[1]*p.gpa[4] ) / ( p.gpa[2]*p.gpa[3] )

## [1] 5.754995
```

Descriptive results

Summary:

- Density:
  - the overall density of ties is 0.041.
- Centrality:
  - outdegree centrality (.10) is less than indegree centrality (.29)
- Smoking:
  - smokers tend to be less active as senders of ties ($p_{10}/p_{00} = .69$)
  - there is positive homophily for smoking ($\gamma = 1.47$)
- Gpa:
  - students with high and low gpas have similar densities ($p_{10}/p_{00} \approx p_{01}/p_{00} \approx 1$).
  - there is positive homophily for gpa ($\gamma = 1.25$)

What conclusions can we draw?

Can we infer anything about how these people formed ties?

Descriptive results

Summary:

- Density:
  - the overall density of ties is 0.041.
- Centrality:
  - outdegree centrality (.10) is less than indegree centrality (.29)
- Smoking:
  - smokers tend to be less active as senders of ties ($p_{10}/p_{00} = .69$)
  - there is positive homophily for smoking ($\gamma = 1.47$)
- Gpa:
  - students with high and low gpas have similar densities ($p_{10}/p_{00} \approx p_{01}/p_{00} \approx 1$).
  - there is positive homophily for gpa ($\gamma = 1.25$)

What conclusions can we draw?

Can we infer anything about how these people formed ties?

Descriptive results

Summary:

- Density:
  - the overall density of ties is 0.041.
- Centrality:
  - outdegree centrality (.10) is less than indegree centrality (.29)
- Smoking:
  - smokers tend to be less active as senders of ties ($p_{10}/p_{00} = .69$)
  - there is positive homophily for smoking ($\gamma = 1.47$)
- Gpa:
  - students with high and low gpas have similar densities ($p_{10}/p_{00} \approx p_{01}/p_{00} \approx 1$).
  - there is positive homophily for gpa ($\gamma = 1.25$)

What conclusions can we draw?

Can we infer anything about how these people formed ties?

Descriptive results

Summary:

- Density:
  - the overall density of ties is 0.041.
- Centrality:
  - outdegree centrality (.10) is less than indegree centrality (.29)
- Smoking:
  - smokers tend to be less active as senders of ties ($p_{10}/p_{00} = .69$)
  - there is positive homophily for smoking ($\gamma = 1.47$)
- Gpa:
  - students with high and low gpas have similar densities ($p_{10}/p_{00} \approx p_{01}/p_{00} \approx 1$).
  - there is positive homophily for gpa ($\gamma = 1.25$)

What conclusions can we draw?

Can we infer anything about how these people formed ties?

Descriptive results

Summary:

- Density:
  - the overall density of ties is 0.041.
- Centrality:
  - outdegree centrality (.10) is less than indegree centrality (.29)
- Smoking:
  - smokers tend to be less active as senders of ties ($p_{10}/p_{00} = .69$)
  - there is positive homophily for smoking ($\gamma = 1.47$)
- Gpa:
  - students with high and low gpas have similar densities ($p_{10}/p_{00} \approx p_{01}/p_{00} \approx 1$).
  - there is positive homophily for gpa ($\gamma = 1.25$)

What conclusions can we draw?

Can we infer anything about how these people formed ties?

Descriptive results

Summary:

- Density:
  - the overall density of ties is 0.041.
- Centrality:
  - outdegree centrality (.10) is less than indegree centrality (.29)
- Smoking:
  - smokers tend to be less active as senders of ties ($p_{10}/p_{00} = .69$)
  - there is positive homophily for smoking ($\gamma = 1.47$)
- Gpa:
  - students with high and low gpas have similar densities ($p_{10}/p_{00} \approx p_{01}/p_{00} \approx 1$).
  - there is positive homophily for gpa ($\gamma = 1.25$)

What conclusions can we draw?

Can we infer anything about how these people formed ties?

Descriptive results

Summary:

- Density:
    - the overall density of ties is 0.041.
- Centrality:
    - outdegree centrality (.10) is less than indegree centrality (.29)
- Smoking:
    - smokers tend to be less active as senders of ties ($p_{10}/p_{00} = .69$)
    - there is positive homophily for smoking ($\gamma = 1.47$)
- Gpa:
    - students with high and low gpas have similar densities ($p_{10}/p_{00} \approx p_{01}/p_{00} \approx 1$).
    - there is positive homophily for gpa ($\gamma = 1.25$)

What conclusions can we draw?

Can we infer anything about how these people formed ties?

Descriptive results

Summary:

- Density:
  - the overall density of ties is 0.041.
- Centrality:
  - outdegree centrality (.10) is less than indegree centrality (.29)
- Smoking:
  - smokers tend to be less active as senders of ties ($p_{10}/p_{00} = .69$)
  - there is positive homophily for smoking ($\gamma = 1.47$)
- Gpa:
  - students with high and low gpas have similar densities ($p_{10}/p_{00} \approx p_{01}/p_{00} \approx 1$).
  - there is positive homophily for gpa ($\gamma = 1.25$)

What conclusions can we draw?

Can we infer anything about how these people formed ties?

Descriptive results

Summary:

- Density:
  - the overall density of ties is 0.041.
- Centrality:
  - outdegree centrality (.10) is less than indegree centrality (.29)
- Smoking:
  - smokers tend to be less active as senders of ties ($p_{10}/p_{00} = .69$)
  - there is positive homophily for smoking ($\gamma = 1.47$)
- Gpa:
  - students with high and low gpas have similar densities ($p_{10}/p_{00} \approx p_{01}/p_{00} \approx 1$).
  - there is positive homophily for gpa ($\gamma = 1.25$)

What conclusions can we draw?

Can we infer anything about how these people formed ties?

Descriptive results

Summary:

- Density:
  - the overall density of ties is 0.041.
- Centrality:
  - outdegree centrality (.10) is less than indegree centrality (.29)
- Smoking:
  - smokers tend to be less active as senders of ties ($p_{10}/p_{00} = .69$)
  - there is positive homophily for smoking ($\gamma = 1.47$)
- Gpa:
  - students with high and low gpas have similar densities ($p_{10}/p_{00} \approx p_{01}/p_{00} \approx 1$).
  - there is positive homophily for gpa ($\gamma = 1.25$)

What conclusions can we draw?

Can we infer anything about how these people formed ties?

Descriptive results

Summary:

- Density:
  - the overall density of ties is 0.041.
- Centrality:
  - outdegree centrality (.10) is less than indegree centrality (.29)
- Smoking:
  - smokers tend to be less active as senders of ties ($p_{10}/p_{00} = .69$)
  - there is positive homophily for smoking ($\gamma = 1.47$)
- Gpa:
  - students with high and low gpas have similar densities ($p_{10}/p_{00} \approx p_{01}/p_{00} \approx 1$).
  - there is positive homophily for gpa ($\gamma = 1.25$)

What conclusions can we draw?

Can we infer anything about how these people formed ties?

Descriptive results

Summary:

- Density:
  - the overall density of ties is 0.041.
- Centrality:
  - outdegree centrality (.10) is less than indegree centrality (.29)
- Smoking:
  - smokers tend to be less active as senders of ties ($p_{10}/p_{00} = .69$)
  - there is positive homophily for smoking ($\gamma = 1.47$)
- Gpa:
  - students with high and low gpas have similar densities ($p_{10}/p_{00} \approx p_{01}/p_{00} \approx 1$).
  - there is positive homophily for gpa ($\gamma = 1.25$)

What conclusions can we draw?

Can we infer anything about how these people formed ties?

Descriptive results

Summary:

- Density:
  - the overall density of ties is 0.041.
- Centrality:
  - outdegree centrality (.10) is less than indegree centrality (.29)
- Smoking:
  - smokers tend to be less active as senders of ties ($p_{10}/p_{00} = .69$)
  - there is positive homophily for smoking ($\gamma = 1.47$)
- Gpa:
  - students with high and low gpas have similar densities ($p_{10}/p_{00} \approx p_{01}/p_{00} \approx 1$).
  - there is positive homophily for gpa ($\gamma = 1.25$)

What conclusions can we draw?

Can we infer anything about how these people formed ties?

Descriptive results

Summary:

- Density:
  - the overall density of ties is 0.041.
- Centrality:
  - outdegree centrality (.10) is less than indegree centrality (.29)
- Smoking:
  - smokers tend to be less active as senders of ties ($p_{10}/p_{00} = .69$)
  - there is positive homophily for smoking ($\gamma = 1.47$)
- Gpa:
  - students with high and low gpas have similar densities ($p_{10}/p_{00} \approx p_{01}/p_{00} \approx 1$).
  - there is positive homophily for gpa ($\gamma = 1.25$)

What conclusions can we draw?

Can we infer anything about how these people formed ties?

Models of tie formation

A **probability model** of tie formation is a probability distribution over sociomatrices.

More specifically, let

$$\mathcal{Y} = \{\mathbf{Y} : y_{i,j} \in \{0,1\}, y_{i,i} = \text{NA}\}$$

be the set of all possible sociomatrics.

A probability model P over $\mathcal{Y}$ assigns a number $P(\mathbf{Y})$ to each $\mathbf{Y} \in \mathcal{Y}$

$$0 \leq P(\mathbf{Y}) \leq 1 \text{ for all } \mathbf{Y} \in \mathcal{Y}$$

$$\sum_{\mathbf{Y} \in \mathcal{Y}} P(\mathbf{Y}) = 1$$

Models of tie formation

A **probability model** of tie formation is a probability distribution over sociomatrices.

More specifically, let

$$\mathcal{Y} = \{\mathbf{Y} : y_{i,j} \in \{0,1\}, y_{i,i} = \text{NA}\}$$

be the set of all possible sociomatrics.

A probability model P over $\mathcal{Y}$ assigns a number $P(\mathbf{Y})$ to each $\mathbf{Y} \in \mathcal{Y}$

$$0 \leq P(\mathbf{Y}) \leq 1 \text{ for all } \mathbf{Y} \in \mathcal{Y}$$

$$\sum_{\mathbf{Y} \in \mathcal{Y}} P(\mathbf{Y}) = 1$$

Models of tie formation

A **probability model** of tie formation is a probability distribution over sociomatrices.

More specifically, let

$$\mathcal{Y} = \{\mathbf{Y} : y_{i,j} \in \{0,1\}, y_{i,i} = \text{NA}\}$$

be the set of all possible sociomatrics.

A probability model P over $\mathcal{Y}$ assigns a number $P(\mathbf{Y})$ to each $\mathbf{Y} \in \mathcal{Y}$

$$0 \leq P(\mathbf{Y}) \leq 1 \text{ for all } \mathbf{Y} \in \mathcal{Y}$$

$$\sum_{\mathbf{Y} \in \mathcal{Y}} P(\mathbf{Y}) = 1$$

Simple random graph

The **simple random graph** model P_θ assumes

- all ties are formed independently of each other;
- each tie exists with some probability θ, common across all ties.

Under P_θ the entries of $\mathbf{Y}$ are **independent and identically distributed**:

$$y_{1,2}, \ldots, y_{n-1,n} \sim \text{ i.i.d. binary}(\theta)$$

Exercise: Compute the probability of each graph under P_θ

$$\begin{pmatrix} NA & 0 & 1 & 1 & 0 & 0 \\ 0 & NA & 0 & 1 & 0 & 0 \\ 0 & 1 & NA & 0 & 0 & 0 \\ 1 & 0 & 0 & NA & 1 & 0 \\ 0 & 1 & 0 & 1 & NA & 0 \\ 0 & 0 & 0 & 0 & 0 & NA \end{pmatrix} \begin{pmatrix} NA & 1 & 1 & 1 & 1 & 0 \\ 1 & NA & 0 & 0 & 0 & 0 \\ 1 & 0 & NA & 0 & 0 & 0 \\ 1 & 0 & 0 & NA & 0 & 0 \\ 1 & 0 & 0 & 0 & NA & 0 \\ 0 & 0 & 0 & 0 & 0 & NA \end{pmatrix}$$

Simple random graph

The **simple random graph** model P_θ assumes

- all ties are formed independently of each other;
- each tie exists with some probability θ, common across all ties.

Under P_θ the entries of $\mathbf{Y}$ are **independent and identically distributed**:

$$y_{1,2}, \ldots, y_{n-1,n} \sim \text{ i.i.d. binary}(\theta)$$

Exercise: Compute the probability of each graph under P_θ

$$\begin{pmatrix} NA & 0 & 1 & 1 & 0 & 0 \\ 0 & NA & 0 & 1 & 0 & 0 \\ 0 & 1 & NA & 0 & 0 & 0 \\ 1 & 0 & 0 & NA & 1 & 0 \\ 0 & 1 & 0 & 1 & NA & 0 \\ 0 & 0 & 0 & 0 & 0 & NA \end{pmatrix} \begin{pmatrix} NA & 1 & 1 & 1 & 1 & 0 \\ 1 & NA & 0 & 0 & 0 & 0 \\ 1 & 0 & NA & 0 & 0 & 0 \\ 1 & 0 & 0 & NA & 0 & 0 \\ 1 & 0 & 0 & 0 & NA & 0 \\ 0 & 0 & 0 & 0 & 0 & NA \end{pmatrix}$$

Simple random graph

The **simple random graph** model P_θ assumes

- all ties are formed independently of each other;
- each tie exists with some probability θ, common across all ties.

Under P_θ the entries of $\mathbf{Y}$ are **independent and identically distributed**:

$$y_{1,2}, \ldots, y_{n-1,n} \sim \text{ i.i.d. binary}(\theta)$$

Exercise: Compute the probability of each graph under P_θ

$$\begin{pmatrix} NA & 0 & 1 & 1 & 0 & 0 \\ 0 & NA & 0 & 1 & 0 & 0 \\ 0 & 1 & NA & 0 & 0 & 0 \\ 1 & 0 & 0 & NA & 1 & 0 \\ 0 & 1 & 0 & 1 & NA & 0 \\ 0 & 0 & 0 & 0 & 0 & NA \end{pmatrix} \begin{pmatrix} NA & 1 & 1 & 1 & 1 & 0 \\ 1 & NA & 0 & 0 & 0 & 0 \\ 1 & 0 & NA & 0 & 0 & 0 \\ 1 & 0 & 0 & NA & 0 & 0 \\ 1 & 0 & 0 & 0 & NA & 0 \\ 0 & 0 & 0 & 0 & 0 & NA \end{pmatrix}$$

Simple random graph

The **simple random graph** model P_θ assumes

- all ties are formed independently of each other;
- each tie exists with some probability θ, common across all ties.

Under P_θ the entries of $\mathbf{Y}$ are **independent and identically distributed**:

$$y_{1,2}, \ldots, y_{n-1,n} \sim \text{ i.i.d. binary}(\theta)$$

Exercise: Compute the probability of each graph under P_θ

$$\begin{pmatrix} NA & 0 & 1 & 1 & 0 & 0 \\ 0 & NA & 0 & 1 & 0 & 0 \\ 0 & 1 & NA & 0 & 0 & 0 \\ 1 & 0 & 0 & NA & 1 & 0 \\ 0 & 1 & 0 & 1 & NA & 0 \\ 0 & 0 & 0 & 0 & 0 & NA \end{pmatrix} \begin{pmatrix} NA & 1 & 1 & 1 & 1 & 0 \\ 1 & NA & 0 & 0 & 0 & 0 \\ 1 & 0 & NA & 0 & 0 & 0 \\ 1 & 0 & 0 & NA & 0 & 0 \\ 1 & 0 & 0 & 0 & NA & 0 \\ 0 & 0 & 0 & 0 & 0 & NA \end{pmatrix}$$

Simple random graph

The **simple random graph** model P_θ assumes

- all ties are formed independently of each other;
- each tie exists with some probability θ, common across all ties.

Under P_θ the entries of $\mathbf{Y}$ are **independent and identically distributed**:

$$y_{1,2}, \dots, y_{n-1,n} \sim \text{ i.i.d. binary}(\theta)$$

Exercise: Compute the probability of each graph under P_θ

$$\begin{pmatrix} NA & 0 & 1 & 1 & 0 & 0 \\ 0 & NA & 0 & 1 & 0 & 0 \\ 0 & 1 & NA & 0 & 0 & 0 \\ 1 & 0 & 0 & NA & 1 & 0 \\ 0 & 1 & 0 & 1 & NA & 0 \\ 0 & 0 & 0 & 0 & 0 & NA \end{pmatrix} \begin{pmatrix} NA & 1 & 1 & 1 & 1 & 0 \\ 1 & NA & 0 & 0 & 0 & 0 \\ 1 & 0 & NA & 0 & 0 & 0 \\ 1 & 0 & 0 & NA & 0 & 0 \\ 1 & 0 & 0 & 0 & NA & 0 \\ 0 & 0 & 0 & 0 & 0 & NA \end{pmatrix}$$

Simple random graph

Under P_θ, the probability of a graph $\mathbf{Y}$ is

$$\begin{aligned} P_\theta(\mathbf{Y}) &= \prod_{i \neq j} \theta^{y_{i,j}} (1-\theta)^{1-y_{i,j}} \\ &= \theta^{\sum y_{i,j}} (1-\theta)^{\sum (1-y_{i,j})} \end{aligned}$$

Would this be a good model for our friendship data?

Are the data **consistent** with this model?

Simple random graph

Under P_θ, the probability of a graph $\mathbf{Y}$ is

$$\begin{aligned} P_\theta(\mathbf{Y}) &= \prod_{i \neq j} \theta^{y_{i,j}} (1-\theta)^{1-y_{i,j}} \\ &= \theta^{\sum y_{i,j}} (1-\theta)^{\sum (1-y_{i,j})} \end{aligned}$$

Would this be a good model for our friendship data?

Are the data **consistent** with this model?

Simple random graph

Under P_θ, the probability of a graph $\mathbf{Y}$ is

$$\begin{aligned} P_\theta(\mathbf{Y}) &= \prod_{i \neq j} \theta^{y_{i,j}} (1-\theta)^{1-y_{i,j}} \\ &= \theta^{\sum y_{i,j}} (1-\theta)^{\sum (1-y_{i,j})} \end{aligned}$$

Would this be a good model for our friendship data?

Are the data **consistent** with this model?

Evaluating the simple random graph

Specification of a probability model requires specification of θ.

Let's make an ad-hoc selection of $\theta = 0.04$ for now, and ask the question:

Are the data consistent with an SRG model with $\theta = 0.04$?

For now, let's evaluate consistency in terms of a few simple statistics:

- $t_d(\mathbf{Y})$ = tie density;
- $t_{cr}(\mathbf{Y})$ = outdegree centrality;
- $t_{cc}(\mathbf{Y})$ = indegree centrality;

For each **test statistic**, we will ask the question

Is the observed value of our test statistic consistent with the values of the statistic we could have observed, under P_θ?

Evaluating the simple random graph

Specification of a probability model requires specification of θ.

Let's make an ad-hoc selection of $\theta = 0.04$ for now, and ask the question:

> *Are the data consistent with an SRG model with $\theta = 0.04$?*

For now, let's evaluate consistency in terms of a few simple statistics:

- $t_d(\mathbf{Y})$ = tie density;
- $t_{cr}(\mathbf{Y})$ = outdegree centrality;
- $t_{cc}(\mathbf{Y})$ = indegree centrality;

For each **test statistic**, we will ask the question

> *Is the observed value of our test statistic consistent with the values of the statistic we could have observed, under P_θ?*

Evaluating the simple random graph

Specification of a probability model requires specification of θ.

Let's make an ad-hoc selection of $\theta = 0.04$ for now, and ask the question:

> *Are the data consistent with an SRG model with* $\theta = 0.04$*?*

For now, let's evaluate consistency in terms of a few simple statistics:

- $t_d(\mathbf{Y})$ = tie density;
- $t_{cr}(\mathbf{Y})$ = outdegree centrality;
- $t_{cc}(\mathbf{Y})$ = indegree centrality;

For each **test statistic**, we will ask the question

> *Is the observed value of our test statistic consistent with the values of the statistic we could have observed, under* P_θ*?*

Evaluating the simple random graph

Specification of a probability model requires specification of θ.

Let's make an ad-hoc selection of $\theta = 0.04$ for now, and ask the question:

> *Are the data consistent with an SRG model with* $\theta = 0.04$*?*

For now, let's evaluate consistency in terms of a few simple statistics:

- $t_d(\mathbf{Y}) =$ tie density;
- $t_{cr}(\mathbf{Y}) =$ outdegree centrality;
- $t_{cc}(\mathbf{Y}) =$ indegree centrality;

For each **test statistic**, we will ask the question

> *Is the observed value of our test statistic consistent with the values of the statistic we could have observed, under* P_θ*?*

Evaluating the simple random graph

Specification of a probability model requires specification of θ.

Let's make an ad-hoc selection of $\theta = 0.04$ for now, and ask the question:

> *Are the data consistent with an SRG model with $\theta = 0.04$?*

For now, let's evaluate consistency in terms of a few simple statistics:

- $t_d(\mathbf{Y})$ = tie density;
- $t_{cr}(\mathbf{Y})$ = outdegree centrality;
- $t_{cc}(\mathbf{Y})$ = indegree centrality;

For each **test statistic**, we will ask the question

> *Is the observed value of our test statistic consistent with the values of the statistic we could have observed, under P_θ?*

Evaluating the simple random graph

Specification of a probability model requires specification of θ.

Let's make an ad-hoc selection of $\theta = 0.04$ for now, and ask the question:

> *Are the data consistent with an SRG model with $\theta = 0.04$?*

For now, let's evaluate consistency in terms of a few simple statistics:

- $t_d(\mathbf{Y})$ = tie density;
- $t_{cr}(\mathbf{Y})$ = outdegree centrality;
- $t_{cc}(\mathbf{Y})$ = indegree centrality;

For each **test statistic**, we will ask the question

> *Is the observed value of our test statistic consistent with the values of the statistic we could have observed, under P_θ?*

Evaluating the simple random graph

Specification of a probability model requires specification of θ.

Let's make an ad-hoc selection of $\theta = 0.04$ for now, and ask the question:

> *Are the data consistent with an SRG model with* $\theta = 0.04$*?*

For now, let's evaluate consistency in terms of a few simple statistics:

- $t_d(\mathbf{Y})$ = tie density;
- $t_{cr}(\mathbf{Y})$ = outdegree centrality;
- $t_{cc}(\mathbf{Y})$ = indegree centrality;

For each **test statistic**, we will ask the question

> *Is the observed value of our test statistic consistent with the values of the statistic we could have observed, under* P_θ*?*

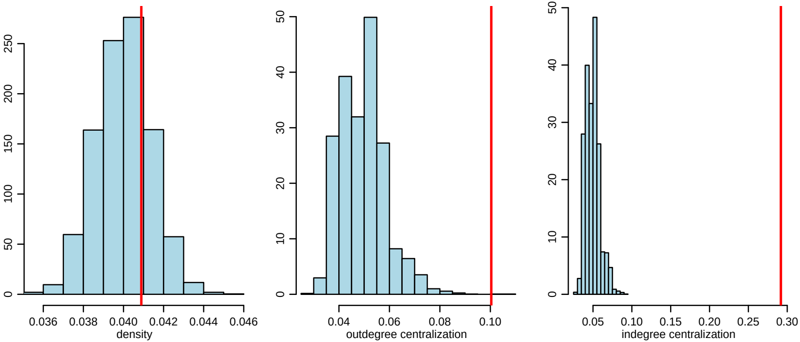
Null distributions

```
t123.stat<-function(Y)
{
  c(mean(Y,na.rm=TRUE), Cd(Y) , Cd(t(Y)) )
}

###

t123.obs<-t123.stat(Y)

t123.obs

## [1] 0.04088967 0.10036442 0.29183493
```

Null distributions

```
theta<-.04

T123.sim<-NULL
for(s in 1:S)
{
  Ys<-matrix(rbinom(nrow(Y)^2,1,theta),nrow(Y),nrow(Y))
  diag(Ys)<-NA
  T123.sim<-rbind(T123.sim, t123.stat(Ys))
}

###

head(T123.sim)

##            [,1]       [,2]       [,3]
## [1,] 0.03996698 0.05165961 0.06584261
## [2,] 0.03957848 0.05914508 0.04496208
## [3,] 0.03962704 0.05909583 0.03782133
## [4,] 0.04059829 0.05101940 0.07229390
## [5,] 0.04030692 0.03713188 0.05131488
## [6,] 0.03753885 0.03993893 0.03993893
```

Null distributions



Hypothesis testing and null distributions

A **pure hypothesis test** is a comparison of the data to a probability model.

Ingredients

- A test statistic t:
  - $t : \mathcal{Y} \rightarrow \mathbb{R}$;
  - t is a known function of the data.
- A null distribution: $\Pr(\cdot|H)$
  - H refers to the hypothesized probability model, i.e. the "null hypothesis"
  - $\Pr(\cdot|H)$ is the probability distribution of t under H.
- A comparison of $t_{\text{obs}} = t(\mathbf{Y})$ to $\Pr(\cdot|H)$.
  - graphical comparison;
  - p-value: $\Pr(t \geq t_{\text{obs}}|H)$.

To make the p-value useful, we usually choose t to be large for values of $\mathbf{Y}$ that are "far away" from H.

Hypothesis testing and null distributions

A **pure hypothesis test** is a comparison of the data to a probability model.

Ingredients

- A test statistic t:
  - $t : \mathcal{Y} \to \mathbb{R}$;
  - t is a known function of the data.
- A null distribution: $\Pr(\cdot|H)$
  - H refers to the hypothesized probability model, i.e. the "null hypothesis."
  - $\Pr(\cdot|H)$ is the probability distribution of t under H.
- A comparison of $t_{\text{obs}} = t(\mathbf{Y})$ to $\Pr(\cdot|H)$.
  - graphical comparison;
  - p-value: $\Pr(t \geq t_{\text{obs}}|H)$.

To make the p-value useful, we usually choose t to be large for values of $\mathbf{Y}$ that are "far away" from H.

Hypothesis testing and null distributions

A **pure hypothesis test** is a comparison of the data to a probability model.

Ingredients

- A test statistic t:
  - $t : \mathcal{Y} \rightarrow \mathbb{R}$;
  - t is a known function of the data.
- A null distribution: $\Pr(\cdot|H)$
  - H refers to the hypothesized probability model, i.e. the "null hypothesis."
  - $\Pr(\cdot|H)$ is the probability distribution of t under H.
- A comparison of $t_{\text{obs}} = t(\mathbf{Y})$ to $\Pr(\cdot|H)$.
  - graphical comparison;
  - p-value: $\Pr(t \geq t_{\text{obs}}|H)$.

To make the p-value useful, we usually choose t to be large for values of $\mathbf{Y}$ that are "far away" from H.

Hypothesis testing and null distributions

A **pure hypothesis test** is a comparison of the data to a probability model.

Ingredients

- A test statistic t:
  - $t : \mathcal{Y} \to \mathbb{R}$;
  - t is a known function of the data.
- A null distribution: $\Pr(\cdot|H)$
  - H refers to the hypothesized probability model, i.e. the "null hypothesis."
  - $\Pr(\cdot|H)$ is the probability distribution of t under H.
- A comparison of $t_{\text{obs}} = t(\mathbf{Y})$ to $\Pr(\cdot|H)$.
  - graphical comparison;
  - p-value: $\Pr(t \geq t_{\text{obs}}|H)$.

To make the p-value useful, we usually choose t to be large for values of $\mathbf{Y}$ that are "far away" from H.

Hypothesis testing and null distributions

A **pure hypothesis test** is a comparison of the data to a probability model.

Ingredients

- A test statistic t:
  - $t : \mathcal{Y} \to \mathbb{R}$;
  - t is a known function of the data.
- A null distribution: $\Pr(\cdot|H)$
  - H refers to the hypothesized probability model, i.e. the "null hypothesis."
  - $\Pr(\cdot|H)$ is the probability distribution of t under H.
- A comparison of $t_{\text{obs}} = t(\mathbf{Y})$ to $\Pr(\cdot|H)$.
  - graphical comparison;
  - p-value: $\Pr(t \geq t_{\text{obs}}|H)$.

To make the p-value useful, we usually choose t to be large for values of $\mathbf{Y}$ that are "far away" from H.

Hypothesis testing and null distributions

A **pure hypothesis test** is a comparison of the data to a probability model.

Ingredients

- A test statistic t:
  - $t : \mathcal{Y} \to \mathbb{R}$;
  - t is a known function of the data.
- A null distribution: $\Pr(\cdot|H)$
  - H refers to the hypothesized probability model, i.e. the "null hypothesis."
  - $\Pr(\cdot|H)$ is the probability distribution of t under H.
- A comparison of $t_{\text{obs}} = t(\mathbf{Y})$ to $\Pr(\cdot|H)$.
  - graphical comparison;
  - p-value: $\Pr(t \geq t_{\text{obs}}|H)$.

To make the p-value useful, we usually choose t to be large for values of $\mathbf{Y}$ that are "far away" from H.

Hypothesis testing and null distributions

A **pure hypothesis test** is a comparison of the data to a probability model.

Ingredients

- A test statistic t:
    - $t : \mathcal{Y} \rightarrow \mathbb{R}$;
    - t is a known function of the data.
- A null distribution: $\Pr(\cdot|H)$
    - H refers to the hypothesized probability model, i.e. the "null hypothesis."
    - $\Pr(\cdot|H)$ is the probability distribution of t under H.
- A comparison of $t_{\text{obs}} = t(\mathbf{Y})$ to $\Pr(\cdot|H)$.
    - graphical comparison;
    - p-value: $\Pr(t \geq t_{\text{obs}}|H)$.

To make the p-value useful, we usually choose t to be large for values of $\mathbf{Y}$ that are "far away" from H.

Hypothesis testing and null distributions

A **pure hypothesis test** is a comparison of the data to a probability model.

Ingredients

- A test statistic t:
  - $t : \mathcal{Y} \to \mathbb{R}$;
  - t is a known function of the data.
- A null distribution: $\Pr(\cdot|H)$
  - H refers to the hypothesized probability model, i.e. the "null hypothesis."
  - $\Pr(\cdot|H)$ is the probability distribution of t under H.
- A comparison of $t_{\text{obs}} = t(\mathbf{Y})$ to $\Pr(\cdot|H)$.
  - graphical comparison;
  - p-value: $\Pr(t \geq t_{\text{obs}}|H)$.

To make the p-value useful, we usually choose t to be large for values of $\mathbf{Y}$ that are "far away" from H.

Hypothesis testing and null distributions

A **pure hypothesis test** is a comparison of the data to a probability model.

Ingredients

- A test statistic t:
  - $t : \mathcal{Y} \to \mathbb{R}$;
  - t is a known function of the data.
- A null distribution: $\Pr(\cdot|H)$
  - H refers to the hypothesized probability model, i.e. the "null hypothesis."
  - $\Pr(\cdot|H)$ is the probability distribution of t under H.
- A comparison of $t_{\text{obs}} = t(\mathbf{Y})$ to $\Pr(\cdot|H)$.
  - graphical comparison;
  - p-value: $\Pr(t \geq t_{\text{obs}}|H)$.

To make the p-value useful, we usually choose t to be large for values of $\mathbf{Y}$ that are "far away" from H.

Hypothesis testing and null distributions

A **pure hypothesis test** is a comparison of the data to a probability model.

Ingredients

- A test statistic t:
    - $t : \mathcal{Y} \rightarrow \mathbb{R}$;
    - t is a known function of the data.
- A null distribution: $\Pr(\cdot|H)$
    - H refers to the hypothesized probability model, i.e. the "null hypothesis."
    - $\Pr(\cdot|H)$ is the probability distribution of t under H.
- A comparison of $t_{\text{obs}} = t(\mathbf{Y})$ to $\Pr(\cdot|H)$.
    - graphical comparison;
    - p-value: $\Pr(t \geq t_{\text{obs}}|H)$.

To make the p-value useful, we usually choose t to be large for values of $\mathbf{Y}$ that are "far away" from H.

Hypothesis testing and null distributions

A **pure hypothesis test** is a comparison of the data to a probability model.

Ingredients

- A test statistic t:
  - $t : \mathcal{Y} \rightarrow \mathbb{R}$;
  - t is a known function of the data.
- A null distribution: $\Pr(\cdot|H)$
  - H refers to the hypothesized probability model, i.e. the "null hypothesis."
  - $\Pr(\cdot|H)$ is the probability distribution of t under H.
- A comparison of $t_{\text{obs}} = t(\mathbf{Y})$ to $\Pr(\cdot|H)$.
  - graphical comparison;
  - p-value: $\Pr(t \geq t_{\text{obs}}|H)$.

To make the p-value useful, we usually choose t to be large for values of $\mathbf{Y}$ that are "far away" from H.

Hypothesis testing and null distributions

A **pure hypothesis test** is a comparison of the data to a probability model.

Ingredients

- A test statistic t:
  - $t : \mathcal{Y} \rightarrow \mathbb{R}$;
  - t is a known function of the data.
- A null distribution: $\Pr(\cdot|H)$
  - H refers to the hypothesized probability model, i.e. the "null hypothesis."
  - $\Pr(\cdot|H)$ is the probability distribution of t under H.
- A comparison of $t_{\text{obs}} = t(\mathbf{Y})$ to $\Pr(\cdot|H)$.
  - graphical comparison;
  - p-value: $\Pr(t \geq t_{\text{obs}}|H)$.

To make the p-value useful, we usually choose t to be large for values of $\mathbf{Y}$ that are "far away" from H.

Hypothesis testing and null distributions

A **pure hypothesis test** is a comparison of the data to a probability model.

Ingredients

- A test statistic t:
  - $t : \mathcal{Y} \to \mathbb{R}$;
  - t is a known function of the data.
- A null distribution: $\Pr(\cdot|H)$
  - H refers to the hypothesized probability model, i.e. the "null hypothesis."
  - $\Pr(\cdot|H)$ is the probability distribution of t under H.
- A comparison of $t_{\text{obs}} = t(\mathbf{Y})$ to $\Pr(\cdot|H)$.
  - graphical comparison;
  - p-value: $\Pr(t \geq t_{\text{obs}}|H)$.

To make the p-value useful, we usually choose t to be large for values of $\mathbf{Y}$ that are "far away" from H.

Example: testing with the density statistic

Let's consider testing the SRG model with $\theta = .04$ using the density statistic.

- $H : \{y_{i,j} : i \neq j\} \sim$ i.i.d. binary(0.04)
- $t(\mathbf{Y}) = |\bar{y}_{\cdot\cdot} - 0.04|$

A large value of the test statistic t

- occurs if $\bar{y}$ is very different from 0.04
- suggests something is wrong with H.

Example: testing with the density statistic

Let's consider testing the SRG model with $\theta = .04$ using the density statistic.

- $H : \{y_{i,j} : i \neq j\} \sim$ i.i.d. binary(0.04)
- $t(\mathbf{Y}) = |\bar{y}_{\cdot\cdot} - 0.04|$

A large value of the test statistic t

- occurs if $\bar{y}$ is very different from 0.04
- suggests something is wrong with H.

Example: testing with the density statistic

Let's consider testing the SRG model with $\theta = .04$ using the density statistic.

- $H : \{y_{i,j} : i \neq j\} \sim$ i.i.d. binary(0.04)
- $t(\mathbf{Y}) = |\bar{y}_{\cdot\cdot} - 0.04|$

A large value of the test statistic t

- occurs if $\bar{y}$ is very different from 0.04
- suggests something is wrong with H.

Example: testing with the density statistic

Let's consider testing the SRG model with $\theta = .04$ using the density statistic.

- $H : \{y_{i,j} : i \neq j\} \sim$ i.i.d. binary(0.04)
- $t(\mathbf{Y}) = |\bar{y}_{\cdot\cdot} - 0.04|$

A large value of the test statistic t

- occurs if $\bar{y}$ is very different from 0.04
- suggests something is wrong with H.

Example: testing with the density statistic

Let's consider testing the SRG model with $\theta = .04$ using the density statistic.

- $H : \{y_{i,j} : i \neq j\} \sim$ i.i.d. binary(0.04)
- $t(\mathbf{Y}) = |\bar{y}_{\cdot\cdot} - 0.04|$

A large value of the test statistic t

- occurs if $\bar{y}$ is very different from 0.04
- suggests something is wrong with H.

Example: testing with the density statistic

Let's consider testing the SRG model with $\theta = .04$ using the density statistic.

- $H : \{y_{i,j} : i \neq j\} \sim$ i.i.d. binary(0.04)
- $t(\mathbf{Y}) = |\bar{y}_{\cdot\cdot} - 0.04|$

A large value of the test statistic t

- occurs if $\bar{y}$ is very different from 0.04
- suggests something is wrong with H.

Example: testing with the density statistic

Let's consider testing the SRG model with $\theta = .04$ using the density statistic.

- $H : \{y_{i,j} : i \neq j\} \sim$ i.i.d. binary(0.04)
- $t(\mathbf{Y}) = |\bar{y}_{\cdot\cdot} - 0.04|$

A large value of the test statistic t

- occurs if $\bar{y}$ is very different from 0.04
- suggests something is wrong with H.

Example: testing with the density statistic

Let's consider testing the SRG model with $\theta = .04$ using the density statistic.

- $H : \{y_{i,j} : i \neq j\} \sim$ i.i.d. binary(0.04)
- $t(\mathbf{Y}) = |\bar{y}_{\cdot\cdot} - 0.04|$

A large value of the test statistic t

- occurs if $\bar{y}$ is very different from 0.04
- suggests something is wrong with H.

Example: testing with the density statistic

Let's consider testing the SRG model with $\theta = .04$ using the density statistic.

- $H : \{y_{i,j} : i \neq j\} \sim$ i.i.d. binary(0.04)
- $t(\mathbf{Y}) = |\bar{y}_{\cdot\cdot} - 0.04|$

A large value of the test statistic t

- occurs if $\bar{y}$ is very different from 0.04
- suggests something is wrong with H.

Computing the null distribution

The null distribution of this particular statistic can be computed with

- a normal approximation;
- a Monte Carlo approximation.

For many other statistics, the Monte Carlo approximation will be more accurate.

```
theta<-0.04

t.H<-NULL

for(s in 1:S)
{
  Ysim<-matrix(rbinom(nrow(Y)^2,1,theta),nrow(Y),nrow(Y)) ; diag(Ysim)<-NA
  t.H<-c(t.H, abs( mean(Ysim,na.rm=TRUE) - theta ) )
}
```

Computing the null distribution

The null distribution of this particular statistic can be computed with

- a normal approximation;
- a Monte Carlo approximation.

For many other statistics, the Monte Carlo approximation will be more accurate.

```
theta<-0.04

t.H<-NULL

for(s in 1:S)
{
  Ysim<-matrix(rbinom(nrow(Y)^2,1,theta),nrow(Y),nrow(Y)) ; diag(Ysim)<-NA
  t.H<-c(t.H, abs( mean(Ysim,na.rm=TRUE) - theta ) )
}
```

Computing the null distribution

The null distribution of this particular statistic can be computed with

- a normal approximation;
- a Monte Carlo approximation.

For many other statistics, the Monte Carlo approximation will be more accurate.

```
theta<-0.04

t.H<-NULL

for(s in 1:S)
{
  Ysim<-matrix(rbinom(nrow(Y)^2,1,theta),nrow(Y),nrow(Y)) ; diag(Ysim)<-NA
  t.H<-c(t.H, abs( mean(Ysim,na.rm=TRUE) - theta ) )
}
```

Computing the null distribution

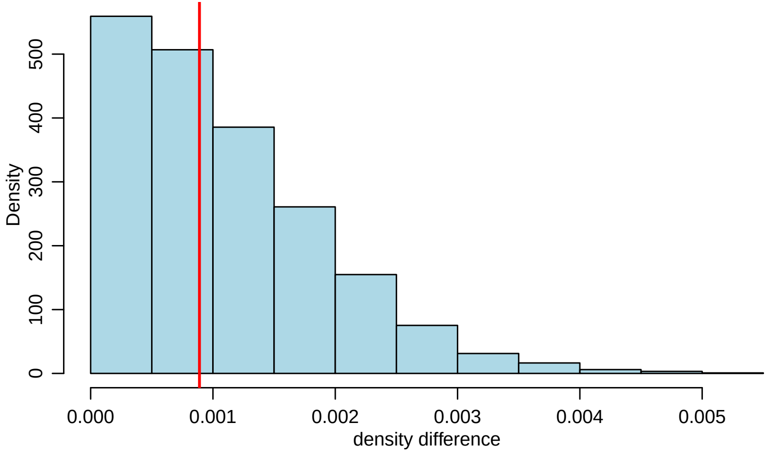
The null distribution of this particular statistic can be computed with

- a normal approximation;
- a Monte Carlo approximation.

For many other statistics, the Monte Carlo approximation will be more accurate.

```
theta<-0.04

t.H<-NULL

for(s in 1:S)
{
  Ysim<-matrix(rbinom(nrow(Y)^2,1,theta),nrow(Y),nrow(Y)) ; diag(Ysim)<-NA
  t.H<-c(t.H, abs( mean(Ysim,na.rm=TRUE) - theta ) )
}
```

Example: testing with the density statistic

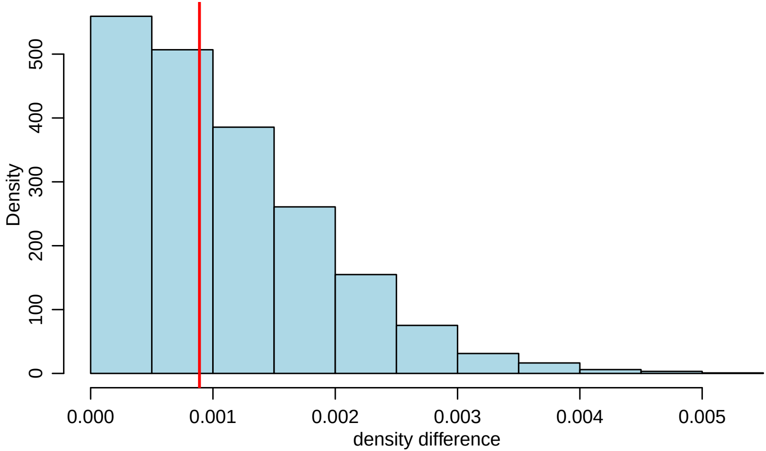


The blue histogram is the (Monte Carlo approximation to) the distribution of t under H.

The red line is the observed value $t_{\text{obs}} = t(\mathbf{Y})$ of the test statistic.

Is there a big discrepancy?
Should we "reject" H?

Example: testing with the density statistic

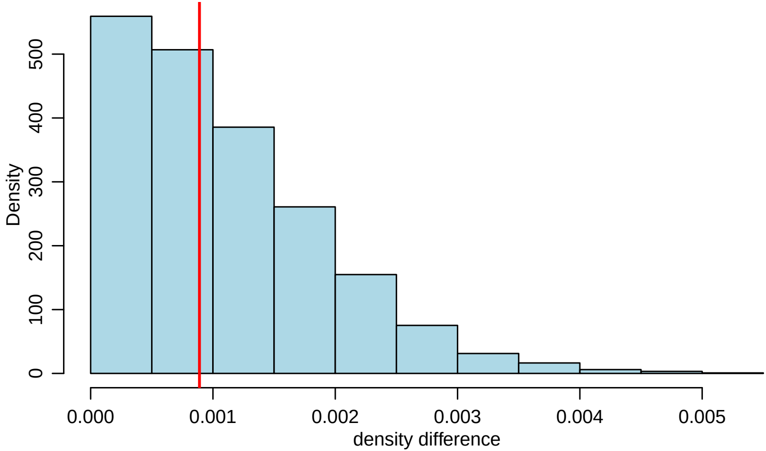


The blue histogram is the (Monte Carlo approximation to) the distribution of t under H.

The red line is the observed value $t_{\text{obs}} = t(\mathbf{Y})$ of the test statistic.

Is there a big discrepancy?
Should we "reject" H?

Example: testing with the density statistic



The blue histogram is the (Monte Carlo approximation to) the distribution of t under H.

The red line is the observed value $t_{\text{obs}} = t(\mathbf{Y})$ of the test statistic.

Is there a big discrepancy?
Should we “reject” H?

Quantifying model-data discrepancy

A popular way of quantifying the discrepancy is with a *p*-value:

$$p = \Pr(t(\tilde{\mathbf{Y}}) \geq t_{\text{obs}} | H)$$

The *p*-value can be approximated via the Monte Carlo method:

```
p.val<-mean( t.H > t.o )
p.val

## [1] 0.51
```

The result says that, if H were true, the probability of observing a value of $t(\mathbf{Y})$ bigger than 9×10^{-4} is about 0.51.

In other words,

> If H were true,
>
> values of t as big as t_{obs} are not extremely unlikely.

A *p*-value of 0.51 is not generally seen as

- a strong discrepancy between model and data;
- evidence against H.

Quantifying model-data discrepancy

A popular way of quantifying the discrepancy is with a *p*-value:

$$p = \Pr(t(\tilde{\mathbf{Y}}) \geq t_{\text{obs}} | H)$$

The *p*-value can be approximated via the Monte Carlo method:

```
p.val<-mean( t.H > t.o )
p.val

## [1] 0.51
```

The result says that, if H were true, the probability of observing a value of $t(\mathbf{Y})$ bigger than 9×10^{-4} is about 0.51.

In other words,

> If H were true,
>
> values of t as big as t_{obs} are not extremely unlikely.

A *p*-value of 0.51 is not generally seen as

- a strong discrepancy between model and data;
- evidence against H.

Quantifying model-data discrepancy

A popular way of quantifying the discrepancy is with a p-value:

$$p = \Pr(t(\tilde{\mathbf{Y}}) \geq t_{\text{obs}} | H)$$

The p-value can be approximated via the Monte Carlo method:

```
p.val<-mean( t.H > t.o )
p.val

## [1] 0.51
```

The result says that, if H were true, the probability of observing a value of $t(\mathbf{Y})$ bigger than 9×10^{-4} is about 0.51.

In other words,

> If H were true,
>
> values of t as big as t_{obs} are not extremely unlikely.

A p-value of 0.51 is not generally seen as

- a strong discrepancy between model and data;
- evidence against H.

Quantifying model-data discrepancy

A popular way of quantifying the discrepancy is with a *p*-value:

$$p = \Pr(t(\tilde{\mathbf{Y}}) \geq t_{\text{obs}} | H)$$

The *p*-value can be approximated via the Monte Carlo method:

```
p.val<-mean( t.H > t.o )
p.val

## [1] 0.51
```

The result says that, if H were true, the probability of observing a value of $t(\mathbf{Y})$ bigger than 9×10^{-4} is about 0.51.

In other words,

> If H were true,
>
> values of t as big as t_{obs} are not extremely unlikely.

A *p*-value of 0.51 is not generally seen as

- a strong discrepancy between model and data;
- evidence against H.

Quantifying model-data discrepancy

A popular way of quantifying the discrepancy is with a *p*-value:

$$p = \Pr(t(\tilde{\mathbf{Y}}) \geq t_{\text{obs}} | H)$$

The *p*-value can be approximated via the Monte Carlo method:

```
p.val<-mean( t.H > t.o )
p.val

## [1] 0.51
```

The result says that, if H were true, the probability of observing a value of $t(\mathbf{Y})$ bigger than 9×10^{-4} is about 0.51.

In other words,

> If H were true,
>
> values of t as big as t_{obs} are not extremely unlikely.

A *p*-value of 0.51 is not generally seen as

- a strong discrepancy between model and data;
- evidence against H.

Quantifying model-data discrepancy

A popular way of quantifying the discrepancy is with a *p*-value:

$$p = \Pr(t(\tilde{\mathbf{Y}}) \geq t_{\text{obs}} | H)$$

The *p*-value can be approximated via the Monte Carlo method:

```
p.val<-mean( t.H > t.o )
p.val

## [1] 0.51
```

The result says that, if H were true, the probability of observing a value of $t(\mathbf{Y})$ bigger than 9×10^{-4} is about 0.51.

In other words,

> If H were true,
>
> values of t as big as t_{obs} are not extremely unlikely.

A *p*-value of 0.51 is not generally seen as

- a strong discrepancy between model and data;
- evidence against H.

Quantifying model-data discrepancy

A popular way of quantifying the discrepancy is with a *p*-value:

$$p = \Pr(t(\tilde{\mathbf{Y}}) \geq t_{\text{obs}} | H)$$

The *p*-value can be approximated via the Monte Carlo method:

```
p.val<-mean( t.H > t.o )
p.val

## [1] 0.51
```

The result says that, if H were true, the probability of observing a value of $t(\mathbf{Y})$ bigger than 9×10^{-4} is about 0.51.

In other words,

If H were true,

values of t as big as t_{obs} are not extremely unlikely.

A *p*-value of 0.51 is not generally seen as

- a strong discrepancy between model and data;
- evidence against H.

Quantifying model-data discrepancy

A popular way of quantifying the discrepancy is with a p-value:

$$p = \Pr(t(\tilde{\mathbf{Y}}) \geq t_{\text{obs}} | H)$$

The p-value can be approximated via the Monte Carlo method:

```
p.val<-mean( t.H > t.o )
p.val

## [1] 0.51
```

The result says that, if H were true, the probability of observing a value of $t(\mathbf{Y})$ bigger than 9×10^{-4} is about 0.51.

In other words,

> If H were true,
>
> values of t as big as t_{obs} are not extremely unlikely.

A p-value of 0.51 is not generally seen as

- a strong discrepancy between model and data;
- evidence against H.

Quantifying model-data discrepancy

A popular way of quantifying the discrepancy is with a *p*-value:

$$p = \Pr(t(\tilde{\mathbf{Y}}) \geq t_{\text{obs}} | H)$$

The *p*-value can be approximated via the Monte Carlo method:

```
p.val<-mean( t.H > t.o )
p.val

## [1] 0.51
```

The result says that, if H were true, the probability of observing a value of $t(\mathbf{Y})$ bigger than 9×10^{-4} is about 0.51.

In other words,

If H were true,

values of t as big as t_{obs} are not extremely unlikely.

A *p*-value of 0.51 is not generally seen as

- a strong discrepancy between model and data;
- evidence against H.

Decision making and error rates

	Truth	
Decision	H	not H
accept H	correct	type II error
reject H	type I error	correct

Suppose H is true and you will perform the following procedure:

- Sample $\mathbf{Y}$
- Compute $t_{\text{obs}} = t(\mathbf{Y})$
- Compute $p = \Pr(t(\tilde{\mathbf{Y}}) \geq t_{\text{obs}}|H)$
- Reject H if $p < \alpha$, accept otherwise.

Then your probability of making a type I error is α.

$$\Pr(\text{reject } H|H \text{ is true }) = \Pr(p < \alpha|H) = \alpha.$$

Often people choose $\alpha = 0.05$.
If H is true, then their chance of falsely rejecting H is 0.05.

Decision making and error rates

Decision	Truth H	 not H
accept H	correct	type II error
reject H	type I error	correct

Suppose H is true and you will perform the following procedure:

- Sample $\mathbf{Y}$
- Compute $t_{\text{obs}} = t(\mathbf{Y})$
- Compute $p = \Pr(t(\tilde{\mathbf{Y}}) \geq t_{\text{obs}}|H)$
- Reject H if $p < \alpha$, accept otherwise.

Then your probability of making a type I error is α.

$$\Pr(\text{reject } H|H \text{ is true }) = \Pr(p < \alpha|H) = \alpha.$$

Often people choose $\alpha = 0.05$.
If H is true, then their chance of falsely rejecting H is 0.05.

Decision making and error rates

	Truth	
Decision	H	not H
accept H	correct	type II error
reject H	type I error	correct

Suppose H is true and you will perform the following procedure:

- Sample $\mathbf{Y}$
- Compute $t_{\text{obs}} = t(\mathbf{Y})$
- Compute $p = \Pr(t(\tilde{\mathbf{Y}}) \geq t_{\text{obs}} | H)$
- Reject H if $p < \alpha$, accept otherwise.

Then your probability of making a type I error is α.

$$\Pr(\text{reject } H | H \text{ is true }) = \Pr(p < \alpha | H) = \alpha.$$

Often people choose $\alpha = 0.05$.
If H is true, then their chance of falsely rejecting H is 0.05.

Decision making and error rates

	Truth	
Decision	H	not H
accept H	correct	type II error
reject H	type I error	correct

Suppose H is true and you will perform the following procedure:

- Sample $\mathbf{Y}$
- Compute $t_{\text{obs}} = t(\mathbf{Y})$
- Compute $p = \Pr(t(\tilde{\mathbf{Y}}) \geq t_{\text{obs}} | H)$
- Reject H if $p < \alpha$, accept otherwise.

Then your probability of making a type I error is α.

$$\Pr(\text{reject } H | H \text{ is true }) = \Pr(p < \alpha | H) = \alpha.$$

Often people choose $\alpha = 0.05$.
If H is true, then their chance of falsely rejecting H is 0.05.

Decision making and error rates

	Truth	
Decision	H	not H
accept H	correct	type II error
reject H	type I error	correct

Suppose H is true and you will perform the following procedure:

- Sample $\mathbf{Y}$
- Compute $t_{\text{obs}} = t(\mathbf{Y})$
- Compute $p = \Pr(t(\tilde{\mathbf{Y}}) \geq t_{\text{obs}} | H)$
- Reject H if $p < \alpha$, accept otherwise.

Then your probability of making a type I error is α.

$$\Pr(\text{reject } H | H \text{ is true }) = \Pr(p < \alpha | H) = \alpha.$$

Often people choose $\alpha = 0.05$.
If H is true, then their chance of falsely rejecting H is 0.05.

Decision making and error rates

	Truth	
Decision	H	not H
accept H	correct	type II error
reject H	type I error	correct

Suppose H is true and you will perform the following procedure:

- Sample $\mathbf{Y}$
- Compute $t_{\text{obs}} = t(\mathbf{Y})$
- Compute $p = \Pr(t(\tilde{\mathbf{Y}}) \geq t_{\text{obs}} | H)$
- Reject H if $p < \alpha$, accept otherwise.

Then your probability of making a type I error is α.

$$\Pr(\text{reject } H | H \text{ is true }) = \Pr(p < \alpha | H) = \alpha.$$

Often people choose $\alpha = 0.05$.
If H is true, then their chance of falsely rejecting H is 0.05.

Decision making and error rates

	Truth	
Decision	H	not H
accept H	correct	type II error
reject H	type I error	correct

Suppose H is true and you will perform the following procedure:

- Sample $\mathbf{Y}$
- Compute $t_{\text{obs}} = t(\mathbf{Y})$
- Compute $p = \Pr(t(\tilde{\mathbf{Y}}) \geq t_{\text{obs}} | H)$
- Reject H if $p < \alpha$, accept otherwise.

Then your probability of making a type I error is α.

$$\Pr(\text{reject } H | H \text{ is true }) = \Pr(p < \alpha | H) = \alpha.$$

Often people choose $\alpha = 0.05$.
If H is true, then their chance of falsely rejecting H is 0.05.

Rejecting the SRG

$$H : \{y_{i,j} : i \neq j\} \sim \text{ i.i.d. binomial}(0.04)$$

Most wouldn't reject H based on the density statistic and its p-value.

We might say that the **model is adequate** in terms of $t(\mathbf{Y}) = \bar{y}$...

Is the model adequate in terms of other statistics? Consider

- $t_{odc} = C_{od}(\mathbf{Y})$ (outdegree centralization)
- $t_{idc} = C_{id}(\mathbf{Y})$ (indegree centralization)

Rejecting the SRG

$$H : \{y_{i,j} : i \neq j\} \sim \text{ i.i.d. binomial}(0.04)$$

Most wouldn't reject H based on the density statistic and its p-value.

We might say that the **model is adequate** in terms of $t(\mathbf{Y}) = \bar{y}_{\cdot\cdot}$

Is the model adequate in terms of other statistics? Consider

- $t_{odc} = C_{od}(\mathbf{Y})$ (outdegree centralization)
- $t_{idc} = C_{id}(\mathbf{Y})$ (indegree centralization)

Rejecting the SRG

$$H : \{y_{i,j} : i \neq j\} \sim \text{ i.i.d. binomial}(0.04)$$

Most wouldn't reject H based on the density statistic and its p-value.

We might say that the **model is adequate** in terms of $t(\mathbf{Y}) = \bar{y}_{\cdot\cdot}$.

Is the model adequate in terms of other statistics? Consider

- $t_{odc} = C_{od}(\mathbf{Y})$ (outdegree centralization)
- $t_{idc} = C_{id}(\mathbf{Y})$ (indegree centralization)

Rejecting the SRG

$$H : \{y_{i,j} : i \neq j\} \sim \text{ i.i.d. binomial(0.04)}$$

Most wouldn't reject H based on the density statistic and its p-value.

We might say that the **model is adequate** in terms of $t(\mathbf{Y}) = \bar{y}_{\cdot\cdot}$

Is the model adequate in terms of other statistics? Consider

- $t_{odc} = C_{od}(\mathbf{Y})$ (outdegree centralization)
- $t_{idc} = C_{id}(\mathbf{Y})$ (indegree centralization)

Rejecting the SRG

$$H : \{y_{i,j} : i \neq j\} \sim \text{ i.i.d. binomial}(0.04)$$

Most wouldn't reject H based on the density statistic and its p-value.

We might say that the **model is adequate** in terms of $t(\mathbf{Y}) = \bar{y}_{\cdot\cdot}$

Is the model adequate in terms of other statistics? Consider

- $t_{odc} = C_{od}(\mathbf{Y})$ (outdegree centralization)
- $t_{idc} = C_{id}(\mathbf{Y})$ (indegree centralization)

Rejecting the SRG

$$H : \{y_{i,j} : i \neq j\} \sim \text{ i.i.d. binomial}(0.04)$$

Most wouldn't reject H based on the density statistic and its p-value.

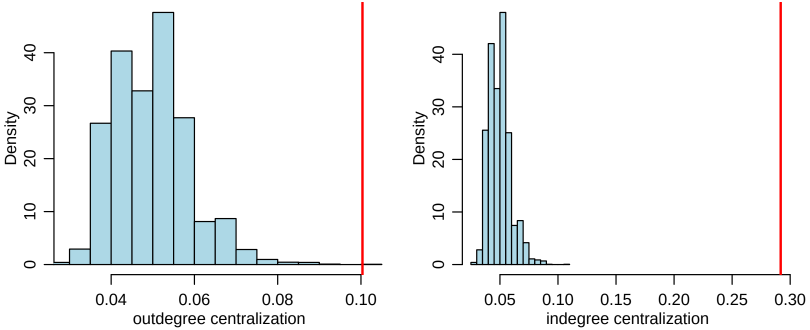
We might say that the **model is adequate** in terms of $t(\mathbf{Y}) = \bar{y}_{\cdot\cdot}$.

Is the model adequate in terms of other statistics? Consider

- $t_{odc} = C_{od}(\mathbf{Y})$ (outdegree centralization)
- $t_{idc} = C_{id}(\mathbf{Y})$ (indegree centralization)

Rejecting the SRG based on centralization

```
t.H<-NULL

for(s in 1:S)
{
  Ysim<-matrix(rbinom(nrow(Y)^2,1,theta),nrow(Y),nrow(Y)) ; diag(Ysim)<-NA
  t.H<-rbind(t.H,  c(Cd(Ysim),Cd(t(Ysim)) ) )
}

t.o<- c(Cd(Y),Cd(t(Y)) )
```



p-values from centralization

```
pval.o<-mean( t.H[,1] >= t.o[1] )
pval.i<-mean( t.H[,2] >= t.o[2] )

pval.o

## [1] 2e-04

pval.i

## [1] 0
```

The plots and *p*-values indicate strong evidence against *H*:

- The binomial(0.04) model predicts much less outdegree centralization than was observed.
- The binomial(0.04) model predicts much much less indegree centralization than was observed.

p-values from centralization

```
pval.o<-mean( t.H[,1] >= t.o[1] )
pval.i<-mean( t.H[,2] >= t.o[2] )

pval.o

## [1] 2e-04

pval.i

## [1] 0
```

The plots and *p*-values indicate strong evidence against *H*:

- The binomial(0.04) model predicts much less outdegree centralization than was observed.
- The binomial(0.04) model predicts much much less indegree centralization than was observed.

p-values from centralization

```
pval.o<-mean( t.H[,1] >= t.o[1] )
pval.i<-mean( t.H[,2] >= t.o[2] )

pval.o

## [1] 2e-04

pval.i

## [1] 0
```

The plots and *p*-values indicate strong evidence against H:

- The binomial(0.04) model predicts much less outdegree centralization than was observed.
- The binomial(0.04) model predicts much much less indegree centralization than was observed.

p-values from centralization

```
pval.o<-mean( t.H[,1] >= t.o[1] )
pval.i<-mean( t.H[,2] >= t.o[2] )

pval.o

## [1] 2e-04

pval.i

## [1] 0
```

The plots and *p*-values indicate strong evidence against *H*:

- The binomial(0.04) model predicts much less outdegree centralization than was observed.
- The binomial(0.04) model predicts much much less indegree centralization than was observed.

Statistical versus probabilistic models

We evaluated evidence against the $SRG(0.040)$ distribution:

$$H : \{y_{i,j} : i \neq j\} \sim \text{ i.i.d. binary}(0.04)\ .$$

Generally, we won't have such a specific hypothesis:

- Rejecting $SRG(0.040)$ doesn't mean we reject $SRG(0.041)$.
- We are more interested in testing *all* SRG distributions.

Statistical versus probabilistic models

We evaluated evidence against the $SRG(0.040)$ distribution:

$$H : \{y_{i,j} : i \neq j\} \sim \text{ i.i.d. binary}(0.04) \ .$$

Generally, we won't have such a specific hypothesis:

- Rejecting $SRG(0.040)$ doesn't mean we reject $SRG(0.041)$.
- We are more interested in testing *all* SRG distributions.

Statistical versus probabilistic models

We evaluated evidence against the $SRG(0.040)$ distribution:

$$H : \{y_{i,j} : i \neq j\} \sim \text{ i.i.d. binary}(0.04) \ .$$

Generally, we won't have such a specific hypothesis:

- Rejecting $SRG(0.040)$ doesn't mean we reject $SRG(0.041)$.
- We are more interested in testing *all* SRG distributions.

Statistical versus probabilistic models

We evaluated evidence against the $SRG(0.040)$ distribution:

$$H : \{y_{i,j} : i \neq j\} \sim \text{ i.i.d. binary}(0.04) \ .$$

Generally, we won't have such a specific hypothesis:

- Rejecting $SRG(0.040)$ doesn't mean we reject $SRG(0.041)$.
- We are more interested in testing *all* SRG distributions.

Statistical model

A statistical model is a collection of probability distributions indexed by an unknown parameter.

$$\mathcal{P} = \{p(\mathbf{Y}|\theta) : \theta \in \Theta\}$$

- θ is the unknown parameter;
- Θ is the parameter space;
- $p(\mathbf{Y}|\theta)$ is the distribution of $\mathbf{Y}$ if θ is correct.

SRG model: The simple random graph **model** is the set of $SRG(\theta)$ distributions:

- $\theta \in \Theta = [0, 1]$
- $p(\mathbf{Y}|\theta)$ is such that $\{y_{i,j} : i \neq j\} \sim$ i.i.d. binary(θ):

$$p(\mathbf{Y}|\theta) = \theta^{\sum y_{i,j}}(1-\theta)^{\sum(1-y_{i,j})}$$

Statistical model

A statistical model is a collection of probability distributions indexed by an unknown parameter.

$$\mathcal{P} = \{p(\mathbf{Y}|\theta) : \theta \in \Theta\}$$

- θ is the unknown parameter;
- Θ is the parameter space;
- $p(\mathbf{Y}|\theta)$ is the distribution of $\mathbf{Y}$ if θ is correct.

SRG model: The simple random graph **model** is the set of $SRG(\theta)$ distributions:

- $\theta \in \Theta = [0, 1]$
- $p(\mathbf{Y}|\theta)$ is such that $\{y_{i,j} : i \neq j\} \sim$ i.i.d. binary(θ):

$$p(\mathbf{Y}|\theta) = \theta^{\sum y_{i,j}}(1-\theta)^{\sum(1-y_{i,j})}$$

Statistical model

A statistical model is a collection of probability distributions indexed by an unknown parameter.

$$\mathcal{P} = \{p(\mathbf{Y}|\theta) : \theta \in \Theta\}$$

- θ is the unknown parameter;
- Θ is the parameter space;
- $p(\mathbf{Y}|\theta)$ is the distribution of $\mathbf{Y}$ if θ is correct.

SRG model: The simple random graph **model** is the set of $SRG(\theta)$ distributions:

- $\theta \in \Theta = [0, 1]$
- $p(\mathbf{Y}|\theta)$ is such that $\{y_{i,j} : i \neq j\} \sim$ i.i.d. binary(θ):

$$p(\mathbf{Y}|\theta) = \theta^{\sum y_{i,j}} (1 - \theta)^{\sum (1 - y_{i,j})}$$

Statistical model

A statistical model is a collection of probability distributions indexed by an unknown parameter.

$$\mathcal{P} = \{p(\mathbf{Y}|\theta) : \theta \in \Theta\}$$

- θ is the unknown parameter;
- Θ is the parameter space;
- $p(\mathbf{Y}|\theta)$ is the distribution of $\mathbf{Y}$ if θ is correct.

SRG model: The simple random graph **model** is the set of $SRG(\theta)$ distributions:

- $\theta \in \Theta = [0, 1]$
- $p(\mathbf{Y}|\theta)$ is such that $\{y_{i,j} : i \neq j\} \sim$ i.i.d. binary(θ):

$$p(\mathbf{Y}|\theta) = \theta^{\sum y_{i,j}}(1-\theta)^{\sum(1-y_{i,j})}$$

Statistical model

A statistical model is a collection of probability distributions indexed by an unknown parameter.

$$\mathcal{P} = \{p(\mathbf{Y}|\theta) : \theta \in \Theta\}$$

- θ is the unknown parameter;
- Θ is the parameter space;
- $p(\mathbf{Y}|\theta)$ is the distribution of $\mathbf{Y}$ if θ is correct.

SRG model: The simple random graph **model** is the set of $SRG(\theta)$ distributions:

- $\theta \in \Theta = [0,1]$
- $p(\mathbf{Y}|\theta)$ is such that $\{y_{i,j} : i \neq j\} \sim$ i.i.d. binary(θ):

$$p(\mathbf{Y}|\theta) = \theta^{\sum y_{i,j}}(1-\theta)^{\sum(1-y_{i,j})}$$

Statistical model

A statistical model is a collection of probability distributions indexed by an unknown parameter.

$$\mathcal{P} = \{p(\mathbf{Y}|\theta) : \theta \in \Theta\}$$

- θ is the unknown parameter;
- Θ is the parameter space;
- $p(\mathbf{Y}|\theta)$ is the distribution of $\mathbf{Y}$ if θ is correct.

SRG model: The simple random graph **model** is the set of $SRG(\theta)$ distributions:

- $\theta \in \Theta = [0,1]$
- $p(\mathbf{Y}|\theta)$ is such that $\{y_{i,j} : i \neq j\} \sim$ i.i.d. binary(θ):

$$p(\mathbf{Y}|\theta) = \theta^{\sum y_{i,j}}(1-\theta)^{\sum(1-y_{i,j})}$$

Statistical model

A statistical model is a collection of probability distributions indexed by an unknown parameter.

$$\mathcal{P} = \{p(\mathbf{Y}|\theta) : \theta \in \Theta\}$$

- θ is the unknown parameter;
- Θ is the parameter space;
- $p(\mathbf{Y}|\theta)$ is the distribution of $\mathbf{Y}$ if θ is correct.

SRG model: The simple random graph **model** is the set of $SRG(\theta)$ distributions:

- $\theta \in \Theta = [0, 1]$
- $p(\mathbf{Y}|\theta)$ is such that $\{y_{i,j} : i \neq j\} \sim$ i.i.d. binary(θ):

$$p(\mathbf{Y}|\theta) = \theta^{\sum y_{i,j}} (1 - \theta)^{\sum (1 - y_{i,j})}$$

Statistical model

A statistical model is a collection of probability distributions indexed by an unknown parameter.

$$\mathcal{P} = \{p(\mathbf{Y}|\theta) : \theta \in \Theta\}$$

- θ is the unknown parameter;
- Θ is the parameter space;
- $p(\mathbf{Y}|\theta)$ is the distribution of $\mathbf{Y}$ if θ is correct.

SRG model: The simple random graph **model** is the set of $SRG(\theta)$ distributions:

- $\theta \in \Theta = [0, 1]$
- $p(\mathbf{Y}|\theta)$ is such that $\{y_{i,j} : i \neq j\} \sim$ i.i.d. binary(θ):

$$p(\mathbf{Y}|\theta) = \theta^{\sum y_{i,j}} (1-\theta)^{\sum (1-y_{i,j})}$$

Statistical model

A statistical model is a collection of probability distributions indexed by an unknown parameter.

$$\mathcal{P} = \{p(\mathbf{Y}|\theta) : \theta \in \Theta\}$$

- θ is the unknown parameter;
- Θ is the parameter space;
- $p(\mathbf{Y}|\theta)$ is the distribution of $\mathbf{Y}$ if θ is correct.

SRG model: The simple random graph **model** is the set of $SRG(\theta)$ distributions:

- $\theta \in \Theta = [0, 1]$
- $p(\mathbf{Y}|\theta)$ is such that $\{y_{i,j} : i \neq j\} \sim$ i.i.d. binary(θ):

$$p(\mathbf{Y}|\theta) = \theta^{\sum y_{i,j}} (1 - \theta)^{\sum (1 - y_{i,j})}$$

Rejecting a statistical model

Is there a procedure that can evaluate *all* SRG distributions at once?

We will discuss two approaches:

- ad-hoc approach: intuitive but not exactly correct in terms of type I error.
- principled: intuitive and correct, but less generalizable.

Rejecting a statistical model

Is there a procedure that can evaluate *all* SRG distributions at once?

We will discuss two approaches:

- ad-hoc approach: intuitive but not exactly correct in terms of type I error.
- principled: intuitive and correct, but less generalizable.

Rejecting a statistical model

Is there a procedure that can evaluate *all* SRG distributions at once?

We will discuss two approaches:

- ad-hoc approach: intuitive but not exactly correct in terms of type I error.
- principled: intuitive and correct, but less generalizable.

Rejecting a statistical model

Is there a procedure that can evaluate *all* SRG distributions at once?

We will discuss two approaches:

- ad-hoc approach: intuitive but not exactly correct in terms of type I error.
- principled: intuitive and correct, but less generalizable.

Rejecting a statistical model

Is there a procedure that can evaluate *all* SRG distributions at once?

We will discuss two approaches:

- ad-hoc approach: intuitive but not exactly correct in terms of type I error.
- principled: intuitive and correct, but less generalizable.

Ad-hoc approach: the best case scenario

Idea:

- We reject the $SRG(\theta)$ distribution if samples from it don't look like $\mathbf{Y}$;
- We would reject all $SRG(\theta)$ distributions if none of them look like $\mathbf{Y}$;
- Instead of comparing $\mathbf{Y}$ to each θ, just compare $\mathbf{Y}$ to the "most similar" θ.

Which value of $\theta \in [0, 1]$ makes $\tilde{\mathbf{Y}} \sim SRG(\theta)$ most similar to $\mathbf{Y}$?

Maximum likelihood estimation:

$\mathbf{Y} \sim p(\mathbf{Y}|\theta)$, θ unknown.

$\theta \in \Theta$

The **maximum likelihood estimator(MLE)** of θ is the value $\hat{\theta}$ that maximizes the probability of the observed data:

$$p(\mathbf{Y}|\hat{\theta}) \geq p(\mathbf{Y}|\theta) \quad \text{for all} \quad \theta \in \Theta$$

Ad-hoc approach: the best case scenario

Idea:

- We reject the $SRG(\theta)$ distribution if samples from it don't look like $\mathbf{Y}$;
- We would reject all $SRG(\theta)$ distributions if none of them look like $\mathbf{Y}$;
- Instead of comparing $\mathbf{Y}$ to each θ, just compare $\mathbf{Y}$ to the "most similar" θ.

Which value of $\theta \in [0,1]$ makes $\tilde{\mathbf{Y}} \sim SRG(\theta)$ most similar to $\mathbf{Y}$?

Maximum likelihood estimation:

$\mathbf{Y} \sim p(\mathbf{Y}|\theta)$, θ unknown.

$\theta \in \Theta$

The **maximum likelihood estimator(MLE)** of θ is the value $\hat{\theta}$ that maximizes the probability of the observed data:

$$p(\mathbf{Y}|\hat{\theta}) \geq p(\mathbf{Y}|\theta) \quad \text{for all} \quad \theta \in \Theta$$

Ad-hoc approach: the best case scenario

Idea:

- We reject the $SRG(\theta)$ distribution if samples from it don't look like $\mathbf{Y}$;
- We would reject all $SRG(\theta)$ distributions if none of them look like $\mathbf{Y}$;
- Instead of comparing $\mathbf{Y}$ to each θ, just compare $\mathbf{Y}$ to the "most similar" θ.

Which value of $\theta \in [0, 1]$ makes $\tilde{\mathbf{Y}} \sim SRG(\theta)$ most similar to $\mathbf{Y}$?

Maximum likelihood estimation:

$\mathbf{Y} \sim p(\mathbf{Y}|\theta)$, θ unknown.

$\theta \in \Theta$

The **maximum likelihood estimator(MLE)** of θ is the value $\hat{\theta}$ that maximizes the probability of the observed data:

$$p(\mathbf{Y}|\hat{\theta}) \geq p(\mathbf{Y}|\theta) \quad \text{for all} \quad \theta \in \Theta$$

Ad-hoc approach: the best case scenario

Idea:

- We reject the $SRG(\theta)$ distribution if samples from it don't look like $\mathbf{Y}$;
- We would reject all $SRG(\theta)$ distributions if none of them look like $\mathbf{Y}$;
- Instead of comparing $\mathbf{Y}$ to each θ, just compare $\mathbf{Y}$ to the "most similar" θ.

Which value of $\theta \in [0,1]$ makes $\tilde{\mathbf{Y}} \sim SRG(\theta)$ most similar to $\mathbf{Y}$?

Maximum likelihood estimation:

$\mathbf{Y} \sim p(\mathbf{Y}|\theta)$, θ unknown.

$\theta \in \Theta$

The **maximum likelihood estimator(MLE)** of θ is the value $\hat{\theta}$ that maximizes the probability of the observed data:

$$p(\mathbf{Y}|\hat{\theta}) \geq p(\mathbf{Y}|\theta) \quad \text{for all} \quad \theta \in \Theta$$

Ad-hoc approach: the best case scenario

Idea:

- We reject the $SRG(\theta)$ distribution if samples from it don't look like $\mathbf{Y}$;
- We would reject all $SRG(\theta)$ distributions if none of them look like $\mathbf{Y}$;
- Instead of comparing $\mathbf{Y}$ to each θ, just compare $\mathbf{Y}$ to the "most similar" θ.

Which value of $\theta \in [0,1]$ makes $\tilde{\mathbf{Y}} \sim SRG(\theta)$ most similar to $\mathbf{Y}$?

Maximum likelihood estimation:

$\mathbf{Y} \sim p(\mathbf{Y}|\theta)$, θ unknown.

$\theta \in \Theta$

The **maximum likelihood estimator(MLE)** of θ is the value $\hat{\theta}$ that maximizes the probability of the observed data:

$$p(\mathbf{Y}|\hat{\theta}) \geq p(\mathbf{Y}|\theta) \quad \text{for all} \quad \theta \in \Theta$$

Ad-hoc approach: the best case scenario

Idea:

- We reject the $SRG(\theta)$ distribution if samples from it don't look like $\mathbf{Y}$;
- We would reject all $SRG(\theta)$ distributions if none of them look like $\mathbf{Y}$;
- Instead of comparing $\mathbf{Y}$ to each θ, just compare $\mathbf{Y}$ to the "most similar" θ.

Which value of $\theta \in [0, 1]$ makes $\tilde{\mathbf{Y}} \sim SRG(\theta)$ most similar to $\mathbf{Y}$?

Maximum likelihood estimation:

$\mathbf{Y} \sim p(\mathbf{Y}|\theta)$, θ unknown.

$\theta \in \Theta$

The **maximum likelihood estimator(MLE)** of θ is the value $\hat{\theta}$ that maximizes the probability of the observed data:

$$p(\mathbf{Y}|\hat{\theta}) \geq p(\mathbf{Y}|\theta) \quad \text{for all} \quad \theta \in \Theta$$

Ad-hoc approach: the best case scenario

Idea:

- We reject the $SRG(\theta)$ distribution if samples from it don't look like $\mathbf{Y}$;
- We would reject all $SRG(\theta)$ distributions if none of them look like $\mathbf{Y}$;
- Instead of comparing $\mathbf{Y}$ to each θ, just compare $\mathbf{Y}$ to the "most similar" θ.

Which value of $\theta \in [0,1]$ makes $\tilde{\mathbf{Y}} \sim SRG(\theta)$ most similar to $\mathbf{Y}$?

Maximum likelihood estimation:

$\mathbf{Y} \sim p(\mathbf{Y}|\theta)$, θ unknown.

$\theta \in \Theta$

The **maximum likelihood estimator(MLE)** of θ is the value $\hat{\theta}$ that maximizes the probability of the observed data:

$$p(\mathbf{Y}|\hat{\theta}) \geq p(\mathbf{Y}|\theta) \quad \text{for all} \quad \theta \in \Theta$$

Ad-hoc approach: the best case scenario

Idea:

- We reject the $SRG(\theta)$ distribution if samples from it don't look like $\mathbf{Y}$;
- We would reject all $SRG(\theta)$ distributions if none of them look like $\mathbf{Y}$;
- Instead of comparing $\mathbf{Y}$ to each θ, just compare $\mathbf{Y}$ to the "most similar" θ.

Which value of $\theta \in [0, 1]$ makes $\tilde{\mathbf{Y}} \sim SRG(\theta)$ most similar to $\mathbf{Y}$?

Maximum likelihood estimation:

$\mathbf{Y} \sim p(\mathbf{Y}|\theta)$, θ unknown.

$\theta \in \Theta$

The **maximum likelihood estimator(MLE)** of θ is the value $\hat{\theta}$ that maximizes the probability of the observed data:

$$p(\mathbf{Y}|\hat{\theta}) \geq p(\mathbf{Y}|\theta) \quad \text{for all} \quad \theta \in \Theta$$

Ad-hoc approach: the best case scenario

Idea:

- We reject the $SRG(\theta)$ distribution if samples from it don't look like $\mathbf{Y}$;
- We would reject all $SRG(\theta)$ distributions if none of them look like $\mathbf{Y}$;
- Instead of comparing $\mathbf{Y}$ to each θ, just compare $\mathbf{Y}$ to the "most similar" θ.

Which value of $\theta \in [0, 1]$ makes $\tilde{\mathbf{Y}} \sim SRG(\theta)$ most similar to $\mathbf{Y}$?

Maximum likelihood estimation:

$\mathbf{Y} \sim p(\mathbf{Y}|\theta)$, θ unknown.

$\theta \in \boldsymbol{\Theta}$

The **maximum likelihood estimator(MLE)** of θ is the value $\hat{\theta}$ that maximizes the probability of the observed data:

$$p(\mathbf{Y}|\hat{\theta}) \geq p(\mathbf{Y}|\theta) \quad \text{for all} \quad \theta \in \Theta$$

Ad-hoc approach: the best case scenario

Idea:

- We reject the $SRG(\theta)$ distribution if samples from it don't look like $\mathbf{Y}$;
- We would reject all $SRG(\theta)$ distributions if none of them look like $\mathbf{Y}$;
- Instead of comparing $\mathbf{Y}$ to each θ, just compare $\mathbf{Y}$ to the "most similar" θ.

Which value of $\theta \in [0,1]$ makes $\tilde{\mathbf{Y}} \sim SRG(\theta)$ most similar to $\mathbf{Y}$?

Maximum likelihood estimation:

$\mathbf{Y} \sim p(\mathbf{Y}|\theta)$, θ unknown.

$\theta \in \boldsymbol{\Theta}$

The **maximum likelihood estimator(MLE)** of θ is the value $\hat{\theta}$ that maximizes the probability of the observed data:

$$p(\mathbf{Y}|\hat{\theta}) \geq p(\mathbf{Y}|\theta) \;\; \text{for all} \;\; \theta \in \boldsymbol{\Theta}$$

MLE for SRG

Let's find the MLE for the SRG:

$$\begin{aligned} p(\mathbf{Y}|\theta) &= \theta^{\sum y_{i,j}}(1-\theta)^{\sum(1-y_{i,j})} \\ &= \theta^{m\bar{y}}(1-\theta)^{m(1-\bar{y})} \end{aligned}$$

where

- $m = n(n-1) =$ the number of pairs;
- $\bar{y} = \sum y_{i,j}/m =$ the density.

Recall that $\log x$ is an increasing function of x.

Therefore, the maximizer of $p(\mathbf{Y}|\theta)$ is the maximizer of $\log p(\mathbf{Y}|\theta)$:

$$\begin{aligned} \log p(\mathbf{Y}|\theta) &= m\bar{y}\log\theta + m(1-\bar{y})\log(1-\theta) \\ &= m\left[\bar{y}\log\theta + (1-\bar{y})\log(1-\theta)\right] \end{aligned}$$

MLE for SRG

Let's find the MLE for the SRG:

$$\begin{aligned} p(\mathbf{Y}|\theta) &= \theta^{\sum y_{i,j}}(1-\theta)^{\sum(1-y_{i,j})} \\ &= \theta^{m\bar{y}}(1-\theta)^{m(1-\bar{y})} \end{aligned}$$

where

- $m = n(n-1) =$ the number of pairs;
- $\bar{y} = \sum y_{i,j}/m =$ the density.

Recall that $\log x$ is an increasing function of x.

Therefore, the maximizer of $p(\mathbf{Y}|\theta)$ is the maximizer of $\log p(\mathbf{Y}|\theta)$:

$$\begin{aligned} \log p(\mathbf{Y}|\theta) &= m\bar{y}\log\theta + m(1-\bar{y})\log(1-\theta) \\ &= m\left[\bar{y}\log\theta + (1-\bar{y})\log(1-\theta)\right] \end{aligned}$$

MLE for SRG

Let's find the MLE for the SRG:

$$\begin{aligned} p(\mathbf{Y}|\theta) &= \theta^{\sum y_{i,j}}(1-\theta)^{\sum(1-y_{i,j})} \\ &= \theta^{m\bar{y}}(1-\theta)^{m(1-\bar{y})} \end{aligned}$$

where

- $m = n(n-1) =$ the number of pairs;
- $\bar{y} = \sum y_{i,j}/m =$ the density.

Recall that $\log x$ is an increasing function of x.

Therefore, the maximizer of $p(\mathbf{Y}|\theta)$ is the maximizer of $\log p(\mathbf{Y}|\theta)$:

$$\begin{aligned} \log p(\mathbf{Y}|\theta) &= m\bar{y}\log\theta + m(1-\bar{y})\log(1-\theta) \\ &= m\left[\bar{y}\log\theta + (1-\bar{y})\log(1-\theta)\right] \end{aligned}$$

MLE for SRG

Let's find the MLE for the SRG:

$$\begin{aligned} p(\mathbf{Y}|\theta) &= \theta^{\sum y_{i,j}}(1-\theta)^{\sum(1-y_{i,j})} \\ &= \theta^{m\bar{y}}(1-\theta)^{m(1-\bar{y})} \end{aligned}$$

where

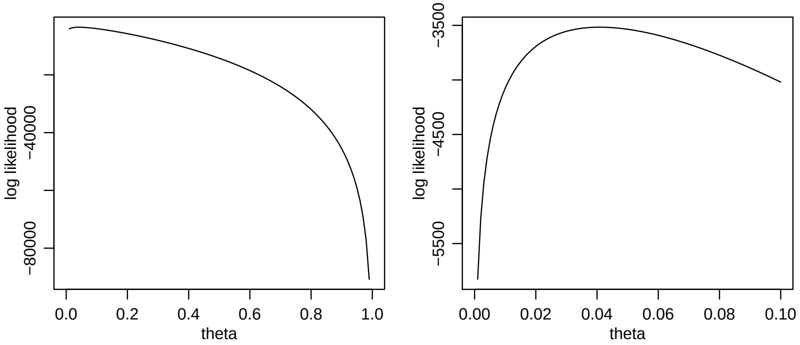
- $m = n(n-1) =$ the number of pairs;
- $\bar{y} = \sum y_{i,j}/m =$ the density.

Recall that $\log x$ is an increasing function of x.

Therefore, the maximizer of $p(\mathbf{Y}|\theta)$ is the maximizer of $\log p(\mathbf{Y}|\theta)$:

$$\begin{aligned} \log p(\mathbf{Y}|\theta) &= m\bar{y}\log\theta + m(1-\bar{y})\log(1-\theta) \\ &= m\left[\bar{y}\log\theta + (1-\bar{y})\log(1-\theta)\right] \end{aligned}$$

MLE for SRG


log likelihood
−40000
−80000
0.0
0.2
0.4
0.6
0.8
1.0
theta
log likelihood
−3500
−4500
−5500
0.00
0.02
0.04
0.06
0.08
0.10
theta

MLE for SRG

Recall, the maximum occurs where the derivative (slope) is zero:

$$\begin{aligned}\tfrac{d}{d\theta}\log p(\mathbf{Y}|\theta) &= m[\frac{\bar{y}}{\theta} - \frac{1-\bar{y}}{1-\theta}] \\ &= m[\frac{\bar{y}}{\hat{\theta}} - \frac{1-\bar{y}}{1-\hat{\theta}}] = 0 \\ \frac{\bar{y}}{\hat{\theta}} &= \frac{1-\bar{y}}{1-\hat{\theta}} \\ \frac{\bar{y}}{1-\bar{y}} &= \frac{\hat{\theta}}{1-\hat{\theta}}\end{aligned}$$

which is satisfied by

$$\hat{\theta} = \bar{y}.$$

Convince yourself that this makes intuitive sense.

Result: By the maximum likelihood criterion, the member of $\mathcal{P} = \{SRG(\theta) : \theta \in [0,1]\}$ that is closest to $\mathbf{Y}$ is the $SRG(\bar{y})$ distribution.

MLE for SRG

Recall, the maximum occurs where the derivative (slope) is zero:

$$\begin{aligned}
\tfrac{d}{d\theta}\log p(\mathbf{Y}|\theta) &= m[\frac{\bar{y}}{\theta} - \frac{1-\bar{y}}{1-\theta}] \\
&= m[\frac{\bar{y}}{\hat{\theta}} - \frac{1-\bar{y}}{1-\hat{\theta}}] = 0 \\
\frac{\bar{y}}{\hat{\theta}} &= \frac{1-\bar{y}}{1-\hat{\theta}} \\
\frac{\bar{y}}{1-\bar{y}} &= \frac{\hat{\theta}}{1-\hat{\theta}}
\end{aligned}$$

which is satisfied by

$$\hat{\theta} = \bar{y}.$$

Convince yourself that this makes intuitive sense.

Result: By the maximum likelihood criterion, the member of $\mathcal{P} = \{SRG(\theta) : \theta \in [0,1]\}$ that is closest to $\mathbf{Y}$ is the $SRG(\bar{y})$ distribution.

Testing with the best case scenario

Intuition:
If we reject $\mathbf{Y} \sim \text{binomial}(\hat{\theta})$,
we should reject $\mathbf{Y} \sim \text{binomial}(\theta)$ for all $\theta \in \Theta$.

Model evaluation procedure: Given a test statistic $t(\mathbf{Y})$,

1. compute $\hat{\theta}$ from $\mathbf{Y}$
2. simulate $\tilde{\mathbf{Y}}_1, \ldots, \tilde{\mathbf{Y}}_S$ from $p(\mathbf{Y}|\hat{\theta})$;
3. compare $t(\mathbf{Y})$ to $t(\tilde{\mathbf{Y}}_1), \ldots, t(\tilde{\mathbf{Y}}_S)$.

Let's do this for our centralization statistics:

- $t_{odc} = C_{od}(\mathbf{Y})$ (outdegree centralization)
- $t_{idc} = C_{id}(\mathbf{Y})$ (indegree centralization)

Testing with the best case scenario

Intuition:
If we reject $\mathbf{Y} \sim \text{binomial}(\hat{\theta})$,
we should reject $\mathbf{Y} \sim \text{binomial}(\theta)$ for all $\theta \in \Theta$.

Model evaluation procedure: Given a test statistic $t(\mathbf{Y})$,

1. compute $\hat{\theta}$ from $\mathbf{Y}$
2. simulate $\tilde{\mathbf{Y}}_1, \ldots, \tilde{\mathbf{Y}}_S$ from $p(\mathbf{Y}|\hat{\theta})$;
3. compare $t(\mathbf{Y})$ to $t(\tilde{\mathbf{Y}}_1), \ldots, t(\tilde{\mathbf{Y}}_S)$.

Let's do this for our centralization statistics:

- $t_{odc} = C_{od}(\mathbf{Y})$ (outdegree centralization)
- $t_{idc} = C_{id}(\mathbf{Y})$ (indegree centralization)

Testing with the best case scenario

Intuition:
If we reject $\mathbf{Y} \sim \text{binomial}(\hat{\theta})$,
we should reject $\mathbf{Y} \sim \text{binomial}(\theta)$ for all $\theta \in \Theta$.

Model evaluation procedure: Given a test statistic $t(\mathbf{Y})$,

1. compute $\hat{\theta}$ from $\mathbf{Y}$
2. simulate $\tilde{\mathbf{Y}}_1, \ldots, \tilde{\mathbf{Y}}_S$ from $p(\mathbf{Y}|\hat{\theta})$;
3. compare $t(\mathbf{Y})$ to $t(\tilde{\mathbf{Y}}_1), \ldots, t(\tilde{\mathbf{Y}}_S)$.

Let's do this for our centralization statistics:

- $t_{odc} = C_{od}(\mathbf{Y})$ (outdegree centralization)
- $t_{idc} = C_{id}(\mathbf{Y})$ (indegree centralization)

Testing with the best case scenario

Intuition:
If we reject $\mathbf{Y} \sim \text{binomial}(\hat{\theta})$,
we should reject $\mathbf{Y} \sim \text{binomial}(\theta)$ for all $\theta \in \Theta$.

Model evaluation procedure: Given a test statistic $t(\mathbf{Y})$,

1. compute $\hat{\theta}$ from $\mathbf{Y}$
2. simulate $\tilde{\mathbf{Y}}_1, \ldots, \tilde{\mathbf{Y}}_S$ from $p(\mathbf{Y}|\hat{\theta})$;
3. compare $t(\mathbf{Y})$ to $t(\tilde{\mathbf{Y}}_1), \ldots, t(\tilde{\mathbf{Y}}_S)$.

Let's do this for our centralization statistics:

- $t_{odc} = C_{od}(\mathbf{Y})$ (outdegree centralization)
- $t_{idc} = C_{id}(\mathbf{Y})$ (indegree centralization)

Testing with the best case scenario

Intuition:
If we reject $\mathbf{Y} \sim \text{binomial}(\hat{\theta})$,
we should reject $\mathbf{Y} \sim \text{binomial}(\theta)$ for all $\theta \in \Theta$.

Model evaluation procedure: Given a test statistic $t(\mathbf{Y})$,

1. compute $\hat{\theta}$ from $\mathbf{Y}$
2. simulate $\tilde{\mathbf{Y}}_1, \ldots, \tilde{\mathbf{Y}}_S$ from $p(\mathbf{Y}|\hat{\theta})$;
3. compare $t(\mathbf{Y})$ to $t(\tilde{\mathbf{Y}}_1), \ldots, t(\tilde{\mathbf{Y}}_S)$.

Let's do this for our centralization statistics:

- $t_{odc} = C_{od}(\mathbf{Y})$ (outdegree centralization)
- $t_{idc} = C_{id}(\mathbf{Y})$ (indegree centralization)

Testing with the best case scenario

Intuition:
If we reject $\mathbf{Y} \sim \text{binomial}(\hat{\theta})$,
we should reject $\mathbf{Y} \sim \text{binomial}(\theta)$ for all $\theta \in \Theta$.

Model evaluation procedure: Given a test statistic $t(\mathbf{Y})$,

1. compute $\hat{\theta}$ from $\mathbf{Y}$
2. simulate $\tilde{\mathbf{Y}}_1, \ldots, \tilde{\mathbf{Y}}_S$ from $p(\mathbf{Y}|\hat{\theta})$;
3. compare $t(\mathbf{Y})$ to $t(\tilde{\mathbf{Y}}_1), \ldots, t(\tilde{\mathbf{Y}}_S)$.

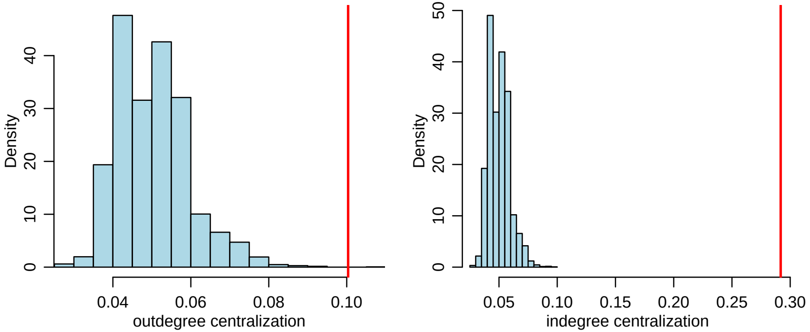
Let's do this for our centralization statistics:

- $t_{odc} = C_{od}(\mathbf{Y})$ (outdegree centralization)
- $t_{idc} = C_{id}(\mathbf{Y})$ (indegree centralization)

Rejecting the best SRG

```
theta<-mean(Y,na.rm=TRUE)
t.H<-NULL

for(s in 1:S)
{
  Ysim<-matrix(rbinom(nrow(Y)^2,1,theta),nrow(Y),nrow(Y)) ; diag(Ysim)<-NA
  t.H<-rbind(t.H,  c(Cd(Ysim),Cd(t(Ysim)) ) )
}

t.o<- c(Cd(Y),Cd(t(Y)) )
```

Rejecting the SRG based on centralization



```
pval.o<-mean( t.H[,1] >= t.o[1] )
pval.i<-mean( t.H[,2] >= t.o[2] )

pval.o

## [1] 2e-04

pval.i

## [1] 0
```

Ad-hockery

The null distributions and *p*-values here are not exactly proper.

Suppose $\mathbf{Y} \sim \text{binary}(\theta)$ for some true but unknown θ.

- θ is only approximated by $\hat{\theta}$.
- the distribution of t under $\text{binary}(\theta)$ is not the same as under $\text{binary}(\hat{\theta})$.
    - Note this latter distribution is "random", as $\hat{\theta}$ is random.

In principle: For some models, we can develop a testing procedure with a correct null distribution.

In practice: It will turn out that our ad-hoc approach often gives the same answers as our exact approach.

Ad-hockery

The null distributions and *p*-values here are not exactly proper.

Suppose $\mathbf{Y} \sim \text{binary}(\theta)$ for some true but unknown θ.

- θ is only approximated by $\hat{\theta}$.
- the distribution of t under binary(θ) is not the same as under binary$(\hat{\theta})$.
  - Note this latter distribution is "random", as $\hat{\theta}$ is random.

In principle: For some models, we can develop a testing procedure with a correct null distribution.

In practice: It will turn out that our ad-hoc approach often gives the same answers as our exact approach.

Ad-hockery

The null distributions and *p*-values here are not exactly proper.

Suppose $\mathbf{Y} \sim \text{binary}(\theta)$ for some true but unknown θ.

- θ is only approximated by $\hat{\theta}$.
- the distribution of t under binary(θ) is not the same as under binary($\hat{\theta}$).
  - Note this latter distribution is "random", as $\hat{\theta}$ is random.

In principle: For some models, we can develop a testing procedure with a correct null distribution.

In practice: It will turn out that our ad-hoc approach often gives the same answers as our exact approach.

Ad-hockery

The null distributions and *p*-values here are not exactly proper.

Suppose $\mathbf{Y} \sim \text{binary}(\theta)$ for some true but unknown θ.

- θ is only approximated by $\hat{\theta}$.
- the distribution of t under binary(θ) is not the same as under binary$(\hat{\theta})$.
    - Note this latter distribution is "random", as $\hat{\theta}$ is random.

In principle: For some models, we can develop a testing procedure with a correct null distribution.

In practice: It will turn out that our ad-hoc approach often gives the same answers as our exact approach.

Ad-hockery

The null distributions and *p*-values here are not exactly proper.

Suppose $\mathbf{Y} \sim \text{binary}(\theta)$ for some true but unknown θ.

- θ is only approximated by $\hat{\theta}$.
- the distribution of t under $\text{binary}(\theta)$ is not the same as under $\text{binary}(\hat{\theta})$.
    - Note this latter distribution is "random", as $\hat{\theta}$ is random.

In principle: For some models, we can develop a testing procedure with a correct null distribution.

In practice: It will turn out that our ad-hoc approach often gives the same answers as our exact approach.

Ad-hockery

The null distributions and *p*-values here are not exactly proper.

Suppose **Y** $\sim$ binary(θ) for some true but unknown θ.

- θ is only approximated by $\hat{\theta}$.
- the distribution of t under binary(θ) is not the same as under binary($\hat{\theta}$).
  - Note this latter distribution is "random", as $\hat{\theta}$ is random.

In principle: For some models, we can develop a testing procedure with a correct null distribution.

In practice: It will turn out that our ad-hoc approach often gives the same answers as our exact approach.

Ad-hockery

The null distributions and *p*-values here are not exactly proper.

Suppose $\mathbf{Y} \sim \text{binary}(\theta)$ for some true but unknown θ.

- θ is only approximated by $\hat{\theta}$.
- the distribution of t under $\text{binary}(\theta)$ is not the same as under $\text{binary}(\hat{\theta})$.
    - Note this latter distribution is "random", as $\hat{\theta}$ is random.

In principle: For some models, we can develop a testing procedure with a correct null distribution.

In practice: It will turn out that our ad-hoc approach often gives the same answers as our exact approach.