

Using motifs and to analyse (multiple) bipartite networks

S. Robin

Sorbonne université

Ecole EcoNet, April 2025, CESAB, Montpellier

Outline

Bipartite networks and motifs

A null model

Motif distribution

Network embedding

Goodness-of-fit and network comparison

Distance-based network comparison

Outline

Bipartite networks and motifs

A null model

Motif distribution

Network embedding

Goodness-of-fit and network comparison

Distance-based network comparison

Bipartite network

Two types of actors.

- ▶ Mutualistic: plant-pollinator
- ▶ Antagonistic: host-parasite

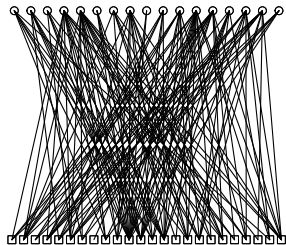
Topological analysis:

understanding the network organisation

Local: node or edge properties (degree, betweenness)

Global: density, connected components, nestedness

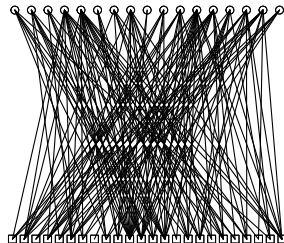
Zackenberg network: [SROB16]



Bipartite network: notations

Species.

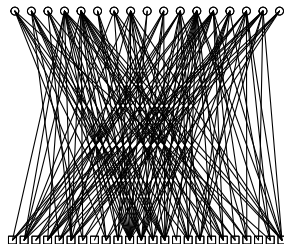
- ▶ $i = 1, \dots, m$ pollinators = rows = bottom nodes
- ▶ $j = 1, \dots, n$ plants = columns = top nodes



Bipartite network: notations

Species.

- ▶ $i = 1, \dots, m$ pollinators = rows = bottom nodes
- ▶ $j = 1, \dots, n$ plants = columns = top nodes



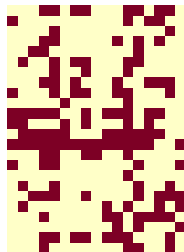
Interactions.

- ▶ $A_{ij} = 1$ if pollinator i interacts with plant j ,
0 otherwise

$$A_{ij} = 1 \Leftrightarrow i \sim j$$

- ▶ adjacency matrix : $m \times n$

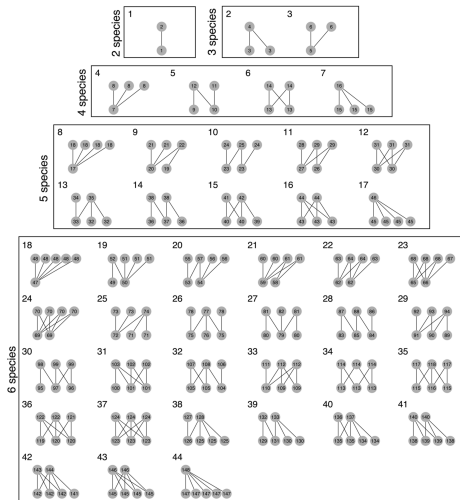
$$A = [A_{ij}]_{1 \leq i \leq m, 1 \leq j \leq n}$$



Bipartite motifs

'Meso-scale' analysis. [SCB⁺19]

- Motifs = 'building-blocks'
- between local (several nodes) and global (sub-graph)



Bipartite motifs

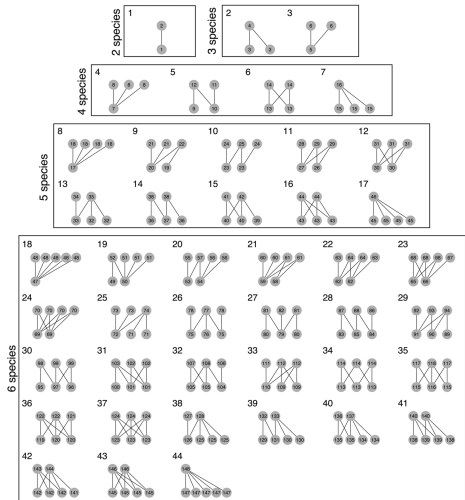
'Meso-scale' analysis. [SCB⁺19]

- Motifs = 'building-blocks'
- between local (several nodes) and global (sub-graph)

Interest.

- Generic description of a network
- Enables network comparison
- Even when the nodes are different

(+ 'species-role': out of the scope here)



Bipartite motifs

'Meso-scale' analysis. [SCB⁺19]

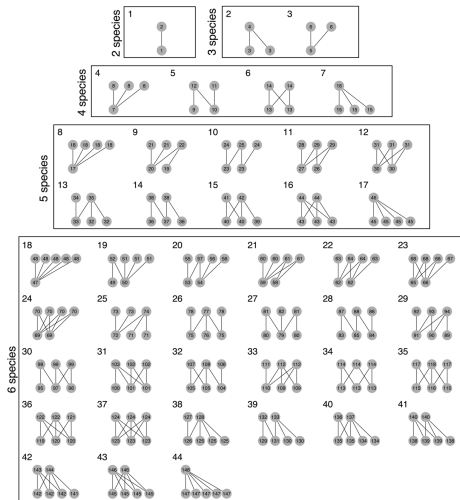
- ▶ Motifs = 'building-blocks'
- ▶ between local (several nodes) and global (sub-graph)

Interest.

- ▶ Generic description of a network
- ▶ Enables network comparison
- ▶ Even when the nodes are different

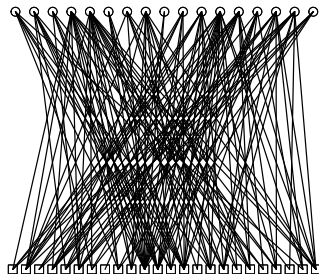
(+ 'species-role': out of the scope here)

Existing tool. `bmotif` package [SSS⁺19]:
counts motif occurrences



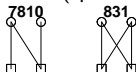
Example

Plant-pollinator network [SROB16]

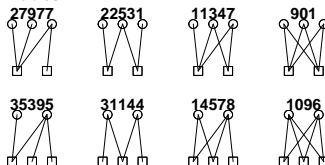


Motif counts.

4 nodes (species)

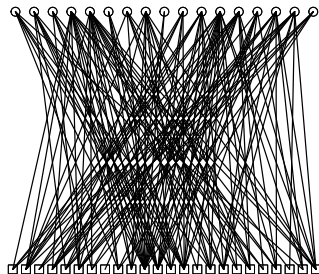


5 nodes



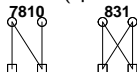
Example

Plant-pollinator network [SROB16]

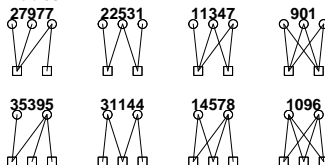


Motif counts.

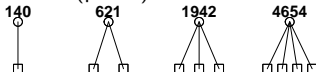
4 nodes (species)



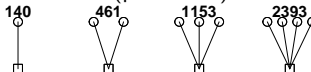
5 nodes



top 'stars' (plants)



bottom 'stars' (pollinators)



Outline

Bipartite networks and motifs

A null model

Motif distribution

Network embedding

Goodness-of-fit and network comparison

Distance-based network comparison

Need for a null model

Motif counts obviously depend on

- ▶ the size of the network: $n \times m$
- ▶ the density of the network
- ▶ the imbalance between bottom-node degrees (specialist vs generalist pollinators)
- ▶ the imbalance between top-node degrees (specialist vs generalist plants)

Need for a null model

Motif counts obviously depend on

- ▶ the size of the network: $n \times m$
- ▶ the density of the network
- ▶ the imbalance between bottom-node degrees (specialist vs generalist pollinators)
- ▶ the imbalance between top-node degrees (specialist vs generalist plants)

Bipartite expected degree distribution (BEDD) model: (in words)

- ▶ Consider m pollinators ($i = 1, \dots, m$):
each plant i has a specific propensity to interact (degree of generalism)
- ▶ Consider n plants ($j = 1, \dots, n$):
each plant j has a specific propensity to interact (idem)
- ▶ The probability for pollinator i and plant j to interact is proportional to the product of their respective propensities.

BEDD model

Bipartite expected degree distribution (BEDD) model: (formaly)

- ▶ ρ = network density
- ▶ g = top node degree imbalance ($\int g = 1$)
- ▶ h = bottom node degree imbalance ($\int h = 1$)

$$\{U_i\}_{i=1,\dots,m} \text{ iid } \sim \mathcal{U}[0, 1] \qquad \{V_j\}_{j=1,\dots,n} \text{ iid } \sim \mathcal{U}[0, 1]$$

$$\mathbb{P}\{i \sim j \mid U_i, V_j\} = \rho \, g(U_i) \, h(V_j)$$

(Bipartite version of the EDD model [CL02])

BEDD model

Bipartite expected degree distribution (BEDD) model: (formaly)

- ▶ ρ = network density
- ▶ g = top node degree imbalance ($\int g = 1$)
- ▶ h = bottom node degree imbalance ($\int h = 1$)

$$\{U_i\}_{i=1,\dots,m} \text{ iid } \sim \mathcal{U}[0, 1] \qquad \{V_j\}_{j=1,\dots,n} \text{ iid } \sim \mathcal{U}[0, 1]$$

$$\mathbb{P}\{i \sim j \mid U_i, V_j\} = \rho g(U_i) h(V_j)$$

(Bipartite version of the EDD model [CL02])

Model parameters:

$$\theta = (\rho, g, h).$$

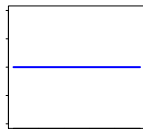
BEDD model

$$\mathbb{P}\{i \sim j \mid U_i, V_j\} = \rho g(U_i) h(V_j)$$

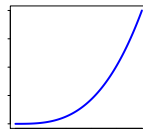
$$\mathbb{E}(D_i \mid U_i) = n \rho g(U_i)$$

$$\mathbb{E}(D_j \mid V_j) = m \rho g(V_j)$$

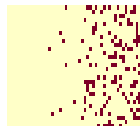
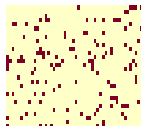
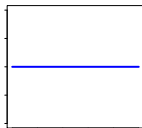
$$h_0(v) =$$



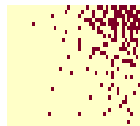
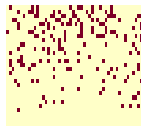
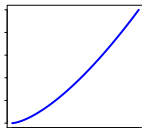
$$h(v) =$$



$$g_0(u) =$$



$$g(u) =$$



Properties of the BEDD model

Assumptions.

- ▶ No preferred or avoided specific connexion
- ▶ **Graph-exchangeable** model: pollinators can be permuted and plants can be permuted

Properties of the BEDD model

Assumptions.

- ▶ No preferred or avoided specific connexion
- ▶ **Graph-exchangeable** model: pollinators can be permuted and plants can be permuted

Properties.

- ▶ Expected degree for pollinator i given $U_i : n \rho g(U_i)$.
- ▶ Expected degree for plant j given $V_j : m \rho h(V_j)$.
- ▶ 'Nested' structure by construction

Properties of the BEDD model

Assumptions.

- ▶ No preferred or avoided specific connexion
- ▶ **Graph-exchangeable** model: pollinators can be permuted and plants can be permuted

Properties.

- ▶ Expected degree for pollinator i given $U_i : n \rho g(U_i)$.
- ▶ Expected degree for plant j given $V_j : m \rho h(V_j)$.
- ▶ 'Nested' structure by construction

Sufficient statistics to fit BEDD:

- ▶ Pollinator degrees + plant degrees
- ▶ or, equivalently, star (single edge, top, bottom) frequencies

Outline

Bipartite networks and motifs

A null model

Motif distribution

Network embedding

Goodness-of-fit and network comparison

Distance-based network comparison

Counting motifs

Number of 'positions'.

- ▶ Choose p nodes among m
- ▶ Choose q nodes among n
- ▶ Try all *automorphisms*

$$c_s := \binom{m}{p} \times \binom{n}{q} \times r_s$$

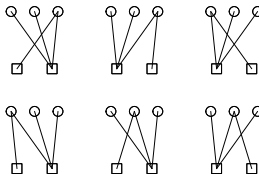
Counting motifs

Number of 'positions'.

- Choose p nodes among m
- Choose q nodes among n
- Try all *automorphisms*

$$c_s := \binom{m}{p} \times \binom{n}{q} \times r_s$$

Automorphisms = non-redundant permutations



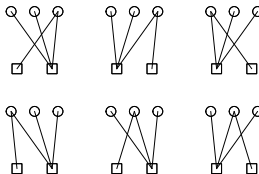
Counting motifs

Number of 'positions'.

- ▶ Choose p nodes among m
- ▶ Choose q nodes among n
- ▶ Try all *automorphisms*

$$c_s := \binom{m}{p} \times \binom{n}{q} \times r_s$$

Automorphisms = non-redundant permutations



Motif count. Try all positions $\alpha = 1, \dots, c_s$, define

$$Y_{s\alpha} = 1 \text{ if match, } \quad 0 \text{ otherwise,}$$

then count the number of matches:

$$N_s = \sum_{\alpha} Y_{s\alpha}$$

→ **Motif frequency:** $F_s := N_s / c_s$

Motif probability

Occurrence probability $\overline{\phi}_s = \mathbb{P}\{Y_{s\alpha} = 1\}$. Under the B-EDD model [OLR22]:

$$\left(\begin{array}{c} \circ \\ \circ \\ \circ \\ \square \\ \square \end{array} \right) = \underline{\hspace{10cm}}$$

Motif probability

Occurrence probability $\bar{\phi}_s = \mathbb{P}\{Y_{s\alpha} = 1\}$. Under the B-EDD model [OLR22]:

$$\left(\begin{array}{ccc} \circ & \circ & \circ \\ & \diagdown & \diagup \\ \square & & \square \end{array} \right) = \frac{\overbrace{\left(\begin{array}{c} \circ \\ | \\ \square \end{array} \right) \left(\begin{array}{c} \circ \\ | \\ \square \end{array} \right) \left(\begin{array}{cc} \circ & / \backslash \\ & \square \quad \square \end{array} \right)}^{\text{top stars}}}{\quad}$$

Motif probability

Occurrence probability $\overline{\phi}_s = \mathbb{P}\{Y_{s\alpha} = 1\}$. Under the B-EDD model [OLR22]:

$$\left(\begin{array}{ccc} \circ & \circ & \circ \\ & \diagdown & \diagup \\ \square & & \square \end{array} \right) = \frac{
 \overbrace{\left(\begin{array}{c} \circ \\ | \\ \square \end{array} \right) \left(\begin{array}{c} \circ \\ | \\ \square \end{array} \right) \left(\begin{array}{ccc} \circ & & \circ \\ & \diagdown & / \\ \square & & \square \end{array} \right)}^{\text{top stars}}
 \overbrace{\left(\begin{array}{c} \circ \\ | \\ \square \end{array} \right) \left(\begin{array}{ccc} \circ & \circ & \circ \\ & \diagdown & / \\ \square & & \square \end{array} \right)}^{\text{bottom stars}}
 }{
 }$$

Motif probability

Occurrence probability $\overline{\phi}_s = \mathbb{P}\{Y_{s\alpha} = 1\}$. Under the B-EDD model [OLR22]:

$$\left(\begin{array}{c} \circ \quad \circ \quad \circ \\ \diagdown \quad \diagup \quad \diagup \\ \square \quad \square \end{array} \right) = \frac{\overbrace{\left(\begin{array}{c} \circ \\ | \\ \square \end{array} \right) \left(\begin{array}{c} \circ \\ | \\ \square \end{array} \right) \left(\begin{array}{c} \circ \\ / \quad \backslash \\ \square \quad \square \end{array} \right)}^{\text{top stars}} \overbrace{\left(\begin{array}{c} \circ \\ | \\ \square \end{array} \right) \left(\begin{array}{c} \circ \quad \circ \quad \circ \\ \backslash \quad / \quad / \\ \square \end{array} \right)}^{\text{bottom stars}}}{\underbrace{\left(\begin{array}{c} \circ \\ | \\ \square \end{array} \right)^4}_{\text{edges}}}$$

Motif probability

Occurrence probability $\bar{\phi}_s = \mathbb{P}\{Y_{s\alpha} = 1\}$. Under the B-EDD model [OLR22]:

$$\begin{aligned}
 \left(\begin{array}{c} \circ \quad \circ \quad \circ \\ \diagdown \quad \diagup \quad \diagup \\ \square \quad \square \end{array} \right) &= \frac{\overbrace{\left(\begin{array}{c} \circ \\ | \\ \square \end{array} \right) \left(\begin{array}{c} \circ \\ | \\ \square \end{array} \right) \left(\begin{array}{c} \circ \\ / \quad \backslash \\ \square \quad \square \end{array} \right)}^{\text{top stars}} \overbrace{\left(\begin{array}{c} \circ \\ | \\ \square \end{array} \right) \left(\begin{array}{c} \circ \quad \circ \\ \diagdown \quad \diagup \\ \square \end{array} \right)}^{\text{bottom stars}}}{\underbrace{\left(\begin{array}{c} \circ \\ | \\ \square \end{array} \right)^4}_{\text{edges}}} \\
 \bar{\phi}_s = \mathbb{P}_{BEDD} \left(\begin{array}{c} \circ \quad \circ \quad \circ \\ \diagdown \quad \diagup \quad \diagup \\ \square \quad \square \end{array} \right) &= \frac{(\phi_1^2 \phi_2) (\phi_1 \phi_4)}{(\phi_1)^4} = \frac{\phi_2 \phi_4}{\phi_1}
 \end{aligned}$$

Motif probability

Occurrence probability $\bar{\phi}_s = \mathbb{P}\{Y_{s\alpha} = 1\}$. Under the B-EDD model [OLR22]:

$$\begin{aligned}
 \left(\begin{array}{c} \circ \quad \circ \quad \circ \\ \diagdown \quad \diagup \quad \diagup \\ \square \quad \square \end{array} \right) &= \frac{\overbrace{\left(\begin{array}{c} \circ \\ | \\ \square \end{array} \right) \left(\begin{array}{c} \circ \\ | \\ \square \end{array} \right) \left(\begin{array}{c} \circ \\ / \quad \backslash \\ \square \quad \square \end{array} \right)}^{\text{top stars}} \overbrace{\left(\begin{array}{c} \circ \\ | \\ \square \end{array} \right) \left(\begin{array}{c} \circ \quad \circ \\ \diagdown \quad \diagup \\ \square \end{array} \right)}^{\text{bottom stars}}}{\underbrace{\left(\begin{array}{c} \circ \\ | \\ \square \end{array} \right)^4}_{\text{edges}}} \\
 \bar{\phi}_s = \mathbb{P}_{BEDD} \left(\begin{array}{c} \circ \quad \circ \quad \circ \\ \diagdown \quad \diagup \quad \diagup \\ \square \quad \square \end{array} \right) &= \frac{(\phi_1^2 \phi_2) (\phi_1 \phi_4)}{(\phi_1)^4} = \frac{\phi_2 \phi_4}{\phi_1}
 \end{aligned}$$

Estimated probability \bar{F}_s .

$$\bar{\phi}_s := \frac{\phi_2 \phi_4}{\phi_1} \quad \rightarrow \quad \bar{F}_s := \frac{F_2 F_4}{F_1}$$

where F_1, F_2, F_4 = observed frequencies of edges, top stars and bottom stars.

Moments of the count

- ▶ Number of positions: c_s
- ▶ Mean: $\mathbb{E}_{BEDD}(N_s) = c_s \times \bar{\phi}_s$

Moments of the count

► **Number of positions:** c_s

► **Mean:** $\mathbb{E}_{BEDD}(N_s) = c_s \times \bar{\phi}_s$

► **Variance:** Same game, requires to evaluate $\mathbb{E}_{BEDD}(N_s^2) = \mathbb{E}_{BEDD} \left(\sum_{\alpha} Y_{s\alpha} \right)^2$

→ Need to account for overlap between positions (*super-motifs*: [PDK⁺08] [\[#31\]](#))



→ Compute the respective expected count in the way as for regular motifs

Moments of the count

► **Number of positions:** c_s

► **Mean:** $\mathbb{E}_{BEDD}(N_s) = c_s \times \bar{\phi}_s$

► **Variance:** Same game, requires to evaluate $\mathbb{E}_{BEDD}(N_s^2) = \mathbb{E}_{BEDD} \left(\sum_{\alpha} Y_{s\alpha} \right)^2$
 → Need to account for overlap between positions (*super-motifs*: [PDK⁺08] [\[#31\]](#))



→ Compute the respective expected count in the way as for regular motifs

► **Covariance:** Same game to compute $\mathbb{Cov}(N_s, N_{s'})$

Moments of the count

► **Number of positions:** c_s

► **Mean:** $\mathbb{E}_{BEDD}(N_s) = c_s \times \bar{\phi}_s$

► **Variance:** Same game, requires to evaluate $\mathbb{E}_{BEDD}(N_s^2) = \mathbb{E}_{BEDD} \left(\sum_{\alpha} Y_{s\alpha} \right)^2$
 → Need to account for overlap between positions (*super-motifs*: [PDK⁺08] [#31])



→ Compute the respective expected count in the way as for regular motifs

► **Covariance:** Same game to compute $\mathbb{Cov}(N_s, N_{s'})$

► **Asymptotic normality:** [#31] [#31] $(F_s - \bar{F}_s) / \sqrt{\widehat{\mathbb{V}}(F_s)} \xrightarrow{m, n \rightarrow \infty} \mathcal{N}(0, 1)$

Outline

Bipartite networks and motifs

A null model

Motif distribution

Network embedding

Goodness-of-fit and network comparison

Distance-based network comparison

Network embedding: Multivariate analysis

Analysing multiple networks. Principle

- ▶ 'Embed' each network into a convenient space (e.g. \mathbb{R}^d)
- ▶ Use standard multivariate analysis (clustering, PCA, MDS, ...)

Network embedding: Multivariate analysis

Analysing multiple networks. Principle

- ▶ 'Embed' each network into a convenient space (e.g. \mathbb{R}^d)
- ▶ Use standard multivariate analysis (clustering, PCA, MDS, ...)

Using motifs. K networks

$$(\text{Network})_k \rightarrow (N_1^k, \dots, N_S^k) \in \mathbb{R}^S$$

but need to correct for: network sizes, correlation between motif frequencies, etc...

Network embedding: Multivariate analysis

Analysing multiple networks. Principle

- ▶ 'Embed' each network into a convenient space (e.g. \mathbb{R}^d)
- ▶ Use standard multivariate analysis (clustering, PCA, MDS, ...)

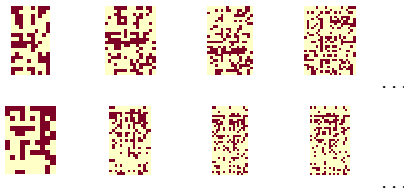
Using motifs. K networks

$$(\text{Network})_k \rightarrow (N_1^k, \dots, N_S^k) \in \mathbb{R}^S$$

but need to correct for: network sizes, correlation between motif frequencies, etc...

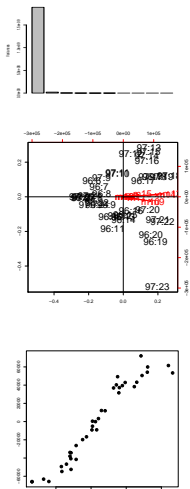
Zackenberg dataset. $K = 46$ networks

- ▶ 2 years
- ▶ 1 network observed every few days



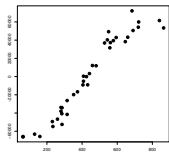
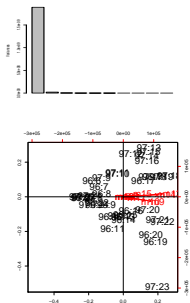
Network embedding: Zackenbergs data [SROB16]

Raw counts

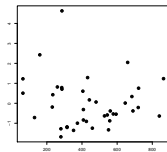
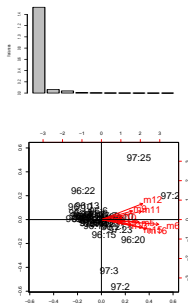


Network embedding: Zackenbergs data [SROB16]

Raw counts

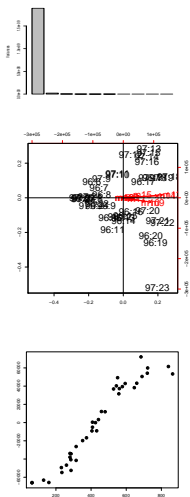


Corrected stat.

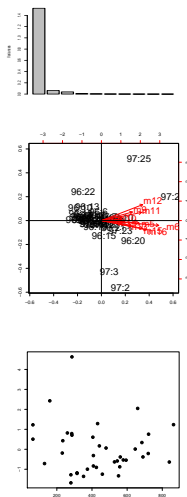


Network embedding: Zackenbergs data [SROB16]

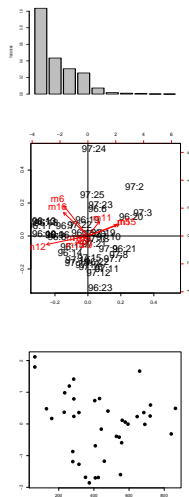
Raw counts



Corrected stat.

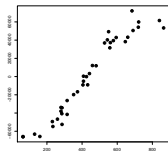
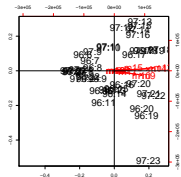
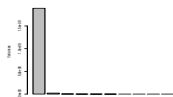


Choleski [#31].

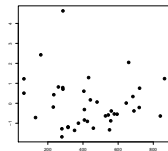
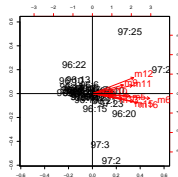
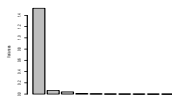


Network embedding: Zackenbergs data [SROB16]

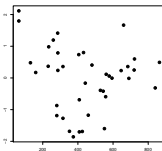
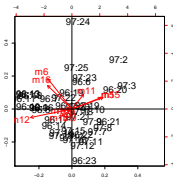
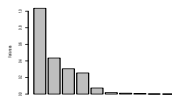
Raw counts



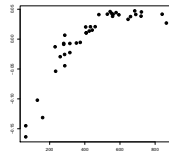
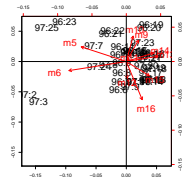
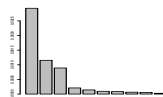
Corrected stat.



Choleski [#31].



Bray-Curtis MDS



Outline

Bipartite networks and motifs

A null model

Motif distribution

Network embedding

Goodness-of-fit and network comparison

Distance-based network comparison

Goodness-of-fit (GOF)

Aim of GOF tests. Test if the observed data arise from a given model.

More cautious: 'if the model fits the data reasonably well'

Goodness-of-fit (GOF)

Aim of GOF tests. Test if the observed data arise from a given model.

More cautious: 'if the model fits the data reasonably well'

Typical approach.

1. Define some statistic (= function of the data) T ,
2. Establish the distribution of T under the model,
3. Compare the observed value of T with its distribution under the model.

Goodness-of-fit (GOF)

Aim of GOF tests. Test if the observed data arise from a given model.

More cautious: 'if the model fits the data reasonably well'

Typical approach.

1. Define some statistic (= function of the data) T ,
2. Establish the distribution of T under the model,
3. Compare the observed value of T with its distribution under the model.

Example.

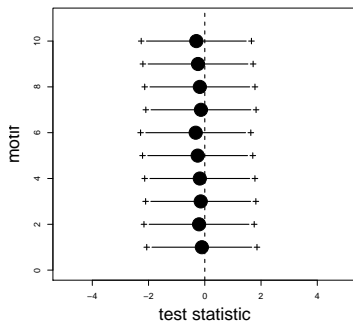
1. Data = observed plant-pollinator network
2. Statistic T = motif count N_5
3. Model = BEDD

Goodness-of-fit (GOF) of the BEDD model

Zackenberg network.

Raw statistic:

$$T_s = \frac{N_s - \hat{\mathbb{E}}N_s}{\sqrt{\hat{\mathbb{V}}N_s}}$$



Goodness-of-fit (GOF) of the BEDD model

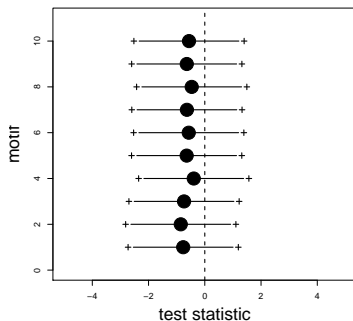
Zackenberg network.

Raw statistic:

$$T_s = \frac{N_s - \hat{\mathbb{E}}N_s}{\sqrt{\hat{\mathbb{V}}N_s}}$$

Corrected stat.: accounts for the estimation error in $\hat{\mathbb{E}}N$ [#31] [#31]

$$T'_s = \frac{N_s - (\hat{\mathbb{E}}N_s - \hat{\mathbb{B}}(\hat{\mathbb{E}}N_s))}{\sqrt{\hat{\mathbb{V}}(N_s - \hat{\mathbb{E}}N_s)}}$$



Testing degree imbalance

Question. Is there some degree imbalance between plants?

Testing degree imbalance

Question. Is there some degree imbalance between plants?

Statistical test.

- Assume $A \sim BEDD(\rho, g, h)$,

$$H_0 = \{h = 1\}$$

- For motif s , evaluate $\hat{\mathbb{E}}_0(N_s)$ and $\hat{\mathbb{V}}_0(N_s)$ and compare

$$W_s = (N_s - \hat{\mathbb{E}}_0(N_s)) / \sqrt{\hat{\mathbb{V}}_0(N_s)}$$

with $\mathcal{N}(0, 1)$

Testing degree imbalance

Question. Is there some degree imbalance between plants?

Statistical test.

- Assume $A \sim BEDD(\rho, g, h)$,

$$H_0 = \{h = 1\}$$

- For motif s , evaluate $\hat{\mathbb{E}}_0(N_s)$ and $\hat{\mathbb{V}}_0(N_s)$ and compare

$$W_s = (N_s - \hat{\mathbb{E}}_0(N_s)) / \sqrt{\hat{\mathbb{V}}_0(N_s)}$$

with $\mathcal{N}(0, 1)$

Example. (only one significant difference)

plant-pollinator					
s	5	6	10	15	16
W_s	$-6.45 \cdot 10^{-2}$	$9.96 \cdot 10^{-1}$	$-6.63 \cdot 10^{-2}$	$7.52 \cdot 10^{-1}$	2.43

seed dispersal					
s	5	6	10	15	16
W_s	$-2.14 \cdot 10^{-1}$	$-2.14 \cdot 10^{-1}$	$-2.93 \cdot 10^{-1}$	$-2.95 \cdot 10^{-1}$	$-3.56 \cdot 10^{-1}$

Comparing network imbalances

Question. Do network A and B share the same imbalance for pollinators?

Comparing network imbalances

Question. Do network A and B share the same imbalance for pollinators?

Statistical test.

- Assume $A \sim BEDD(\rho^A, g^A, h^A)$ and $B \sim BEDD(\rho^B, g^B, h^B)$

$$H_0 = \{g^A = g^B\}$$

- For motif s , evaluate $\hat{\mathbb{E}}_{\hat{\rho}^A, \hat{g}^B, \hat{g}^A}(N_s^A)$ and $\hat{\mathbb{E}}_{\hat{\rho}^B, \hat{g}^A, \hat{g}^B}(N_s^B)$ and compare

$$W_s^{(g)}(A, B) = \frac{(N_s^A - \hat{\mathbb{E}}_0(N_s^A)) - (N_s^B - \hat{\mathbb{E}}_0(N_s^B))}{\sqrt{\hat{\mathbb{V}}_0(N_s^A) + \hat{\mathbb{V}}_0(N_s^B)}}$$

with $\mathcal{N}(0, 1)$

Comparing network imbalances

Question. Do network A and B share the same imbalance for pollinators?

Statistical test.

- Assume $A \sim BEDD(\rho^A, g^A, h^A)$ and $B \sim BEDD(\rho^B, g^B, h^B)$

$$H_0 = \{g^A = g^B\}$$

- For motif s , evaluate $\hat{\mathbb{E}}_{\hat{\rho}^A, \hat{g}^B, \hat{g}^A}(N_s^A)$ and $\hat{\mathbb{E}}_{\hat{\rho}^B, \hat{g}^A, \hat{g}^B}(N_s^B)$ and compare

$$W_s^{(g)}(A, B) = \frac{(N_s^A - \hat{\mathbb{E}}_0(N_s^A)) - (N_s^B - \hat{\mathbb{E}}_0(N_s^B))}{\sqrt{\hat{\mathbb{V}}_0(N_s^A) + \hat{\mathbb{V}}_0(N_s^B)}}$$

with $\mathcal{N}(0, 1)$

Example. (no significant difference)

s	5	6	10	15	16
F_s^A	$9.21 \cdot 10^{-5}$	$1.00 \cdot 10^{-5}$	$8.12 \cdot 10^{-6}$	$3.32 \cdot 10^{-7}$	$4.47 \cdot 10^{-8}$
$\hat{\mathbb{E}}_0 F_s^A$	$1.96 \cdot 10^{-4}$	$3.75 \cdot 10^{-5}$	$1.74 \cdot 10^{-5}$	$4.25 \cdot 10^{-6}$	$1.33 \cdot 10^{-6}$
F_s^B	$5.13 \cdot 10^{-4}$	$1.15 \cdot 10^{-4}$	$5.07 \cdot 10^{-5}$	$1.79 \cdot 10^{-5}$	$5.96 \cdot 10^{-6}$
$\hat{\mathbb{E}}_0 F_s^B$	$2.66 \cdot 10^{-4}$	$2.92 \cdot 10^{-5}$	$2.85 \cdot 10^{-5}$	$1.50 \cdot 10^{-6}$	$1.69 \cdot 10^{-7}$
W_s	-1.56	-1.56	-0.97	-1.28	-0.96

Outline

Bipartite networks and motifs

A null model

Motif distribution

Network embedding

Goodness-of-fit and network comparison

Distance-based network comparison

Plant-pollinator networks in space & time

Questions. Does the structure of plant (herbaceous species) pollinator (native wild bees and hoverflies) network differ¹:

- ▶ along the environmental gradient?
- ▶ across sites within the season?

¹Joint work with Natasha de Manincor & François Massol + Sarah Ouadah & Pierre Latouche

Plant-pollinator networks in space & time

Questions. Does the structure of plant (herbaceous species) pollinator (native wild bees and hoverflies) network differ¹:

- ▶ along the environmental gradient?
- ▶ across sites within the season?

Dataset.

- ▶ 6 sites: 3 regions \times 2 sites per region
 - ▶ 2 years
 - ▶ 7 months (April to October) per year
- $n = 82$ networks, $[4, 39]$ plants, $[8, 80]$ insects

¹Joint work with Natasha de Manincor & François Massol + Sarah Ouadah & Pierre Latouche

Plant-pollinator networks in space & time

Questions. Does the structure of plant (herbaceous species) pollinator (native wild bees and hoverflies) network differ¹:

- ▶ along the environmental gradient?
- ▶ across sites within the season?

Dataset.

- ▶ 6 sites: 3 regions \times 2 sites per region
 - ▶ 2 years
 - ▶ 7 months (April to October) per year
- $n = 82$ networks, [4, 39] plants, [8, 80] insects

Approach. Define a motif-based distance between each pair of networks.

¹Joint work with Natasha de Manincor & François Massol + Sarah Ouadah & Pierre Latouche

Distance-based network comparison

Network distance. For a pair of networks (A, B)

- ▶ for a given comparison (insect imbalance):

$$H_0^{(g)} = \{g^A = g^B\},$$

- ▶ for a given motif s : test statistic $W_s^{(g)}$,
- ▶ define the 'distances':

$$D^{(g)}(A, B) = \sqrt{\sum_s \left(W_s^{(g)}\right)^2}$$

Distance-based network comparison

Network distance. For a pair of networks (A, B)

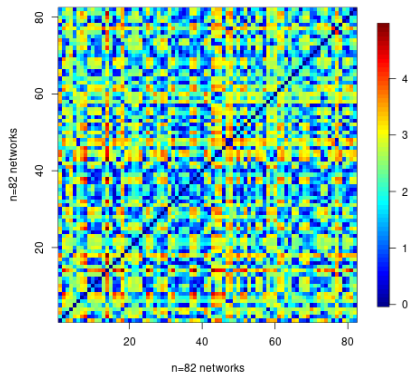
- ▶ for a given comparison (insect imbalance):

$$H_0^{(g)} = \{g^A = g^B\},$$

- ▶ for a given motif s : test statistic $W_s^{(g)}$,
- ▶ define the 'distances':

$$D^{(g)}(A, B) = \sqrt{\sum_s \left(W_s^{(g)}\right)^2}$$

Distance matrix:



Distance-based network comparison

Network distance. For a pair of networks (A, B)

- ▶ for a given comparison (insect imbalance):

$$H_0^{(g)} = \{g^A = g^B\},$$

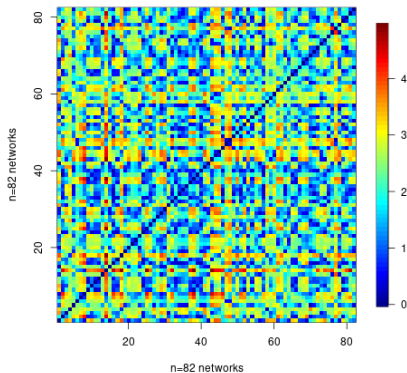
- ▶ for a given motif s : test statistic $W_s^{(g)}$,
- ▶ define the 'distances':

$$D^{(g)}(A, B) = \sqrt{\sum_s \left(W_s^{(g)}\right)^2}$$

The same for

- ▶ plant imbalance: $H_0^{(h)} = \{h^A = h^B\} \rightarrow W_s^{(h)} \rightarrow D^{(h)}(A, B)$
- ▶ both imbalance: $H_0^{(gh)} = \{g^A = g^B, h^A = h^B\} \rightarrow W_s^{(gh)} \rightarrow D^{(gh)}(A, B)$

Distance matrix:



Distance-based analysis of variance ('Adonis')

Data at hand. n networks ($A = 1, \dots, n$)

- ▶ network covariates: e.g. region, year, month, ...
- ▶ distance matrix $D = [D(A, B)]$ for all pairs of networks

²See [MA01, ZS06] for details, #31 for an intuition, R package `vegan` (`adonis2`) for an implementation

Distance-based analysis of variance ('Adonis')

Data at hand. n networks ($A = 1, \dots, n$)

- ▶ network covariates: e.g. region, year, month, ...
- ▶ distance matrix $D = [D(A, B)]$ for all pairs of networks

Analysis of variance for distance matrices [And01]. Briefly speaking²:

- ▶ Think of the regular linear model (regression, analysis of variance)
- ▶ Do as if the distance was computed on pseudo variable rules by a linear model
- ▶ 'Model':

$$D(A, B) = f(\text{region}_A, \text{region}_B, \text{month}_A, \text{month}_B, \\ (\text{region*month})_A, (\text{region*month})_B, \dots)$$

- ▶ Compute a pseudo F statistics for each effect of interest
- ▶ Assess significance using permutation tests.

²See [MA01, ZS06] for details, #31 for an intuition, R package `vegan` (`adonis2`) for an implementation

Result

Insect imbalance $D^{(g)}$.

	Df	Sum Of Sqs	R^2	F	Pr(F)
insectNb	1	69.9	0.2595	42.69	1e-05
plantNb	1	31.17	0.1157	19.04	1e-05
Year	1	2.66	0.0099	1.62	0.22212
Month	6	24.8	0.092	2.52	0.00959
Region	2	8.67	0.0322	2.65	0.04531
Year:Month	6	4.81	0.0179	0.49	0.88756
Year:Region	2	5.51	0.0204	1.68	0.1787
Month:Region	12	32.41	0.1203	1.65	0.06346
Year:Month:Region	12	27.26	0.1012	1.39	0.15884
Residual	38	62.22	0.2309		
Total	81	269.42	1		

Result

Insect imbalance $D^{(g)}$.

	Df	Sum Of Sqs	R^2	F	Pr(F)
insectNb	1	69.9	0.2595	42.69	1e-05
plantNb	1	31.17	0.1157	19.04	1e-05
Year	1	2.66	0.0099	1.62	0.22212
Month	6	24.8	0.092	2.52	0.00959
Region	2	8.67	0.0322	2.65	0.04531
Year:Month	6	4.81	0.0179	0.49	0.88756
Year:Region	2	5.51	0.0204	1.68	0.1787
Month:Region	12	32.41	0.1203	1.65	0.06346
Year:Month:Region	12	27.26	0.1012	1.39	0.15884
Residual	38	62.22	0.2309		
Total	81	269.42	1		

Result

Insect imbalance $D^{(g)}$.

	Df	Sum Of Sqs	R^2	F	Pr(F)
insectNb	1	69.9	0.2595	42.69	1e-05
plantNb	1	31.17	0.1157	19.04	1e-05
Year	1	2.66	0.0099	1.62	0.22212
Month	6	24.8	0.092	2.52	0.00959
Region	2	8.67	0.0322	2.65	0.04531
Year:Month	6	4.81	0.0179	0.49	0.88756
Year:Region	2	5.51	0.0204	1.68	0.1787
Month:Region	12	32.41	0.1203	1.65	0.06346
Year:Month:Region	12	27.26	0.1012	1.39	0.15884
Residual	38	62.22	0.2309		
Total	81	269.42	1		

- Because of small network sizes, need to correct for it [\[#31\]](#)

Result

Insect imbalance $D^{(g)}$.

	Df	Sum Of Sqs	R^2	F	Pr(F)
insectNb	1	69.9	0.2595	42.69	1e-05
plantNb	1	31.17	0.1157	19.04	1e-05
Year	1	2.66	0.0099	1.62	0.22212
Month	6	24.8	0.092	2.52	0.00959
Region	2	8.67	0.0322	2.65	0.04531
Year:Month	6	4.81	0.0179	0.49	0.88756
Year:Region	2	5.51	0.0204	1.68	0.1787
Month:Region	12	32.41	0.1203	1.65	0.06346
Year:Month:Region	12	27.26	0.1012	1.39	0.15884
Residual	38	62.22	0.2309		
Total	81	269.42	1		

- ▶ Because of small network sizes, need to correct for it [\[#31\]](#)
- ▶ Significant effect of the region and the month, indicating change of the insect imbalance both in space and time

Result

Insect imbalance $D^{(g)}$.

	Df	Sum Of Sqs	R^2	F	Pr(F)
insectNb	1	69.9	0.2595	42.69	1e-05
plantNb	1	31.17	0.1157	19.04	1e-05
Year	1	2.66	0.0099	1.62	0.22212
Month	6	24.8	0.092	2.52	0.00959
Region	2	8.67	0.0322	2.65	0.04531
Year:Month	6	4.81	0.0179	0.49	0.88756
Year:Region	2	5.51	0.0204	1.68	0.1787
Month:Region	12	32.41	0.1203	1.65	0.06346
Year:Month:Region	12	27.26	0.1012	1.39	0.15884
Residual	38	62.22	0.2309		
Total	81	269.42	1		

- ▶ Because of small network sizes, need to correct for it [\[#31\]](#)
- ▶ Significant effect of the region and the month, indicating change of the insect imbalance both in space and time
- ▶ The pattern is conserved from year to the next (not Year effect)

Result

Insect imbalance $D^{(g)}$.

	Df	Sum Of Sqs	R^2	F	Pr(F)
insectNb	1	69.9	0.2595	42.69	1e-05
plantNb	1	31.17	0.1157	19.04	1e-05
Year	1	2.66	0.0099	1.62	0.22212
Month	6	24.8	0.092	2.52	0.00959
Region	2	8.67	0.0322	2.65	0.04531
Year:Month	6	4.81	0.0179	0.49	0.88756
Year:Region	2	5.51	0.0204	1.68	0.1787
Month:Region	12	32.41	0.1203	1.65	0.06346
Year:Month:Region	12	27.26	0.1012	1.39	0.15884
Residual	38	62.22	0.2309		
Total	81	269.42	1		

- ▶ Because of small network sizes, need to correct for it [#31]
- ▶ Significant effect of the region and the month, indicating change of the insect imbalance both in space and time
- ▶ The pattern is conserved from year to the next (not Year effect)
- ▶ No effect is found for the plant imbalance distance $D^{(h)}$ [#31]

References I



Marti J Anderson.

A new method for non-parametric multivariate analysis of variance.
Austral ecology, 26(1):32–46, 2001.



F. Chung and L. Lu.

Connected components in random graphs with given expected degree sequences.
Annals of Combinatorics, 6(2):125–145, 2002.



C. Gao and J. Lafferty.

Testing for global network structure using small subgraph statistics.
Technical Report 1704.06742, arXiv, 2017.



B. H McArdle and M. J Anderson.

Fitting multivariate models to community data: a comment on distance-based redundancy analysis.
Ecology, 82(1):290–297, 2001.



S. Ouadah, P. Latouche, and S. Robin.

Motif-based tests for bipartite networks.
Electronic Journal of Statistics, 16(1):293 – 330, 2022.



F. Picard, J.-J. Daudin, M. Koskas, S. Schbath, and S. Robin.

Assessing the exceptionality of network motifs.,
J. Comp. Biol., 15(1):1–20, 2008.



B.I. Simmons, A. Cirtwill, N. Baker, L.V. Dicks, D.B. Stouffer, and W.J. Sutherland.

Motifs in bipartite ecological networks: uncovering indirect interactions.
Oikos, 128(2):154–170, 2019.



S. Saavedra, R. P Rohr, J. M Olesen, and J. Bascompte.

Nested species interactions promote feasibility over stability during the assembly of a pollinator community.
Ecology and evolution, 6(4):997–1007, 2016.

References II



B. I Simmons, M. JM Sweering, M. Schillinger, L. V Dicks, W. J Sutherland, and R. Di Clemente.
bmotif: A package for motif analyses of bipartite networks.
Methods in Ecology and Evolution, 10(5):695–701, 2019.



D. Stark.
Compound Poisson approximations of subgraph counts in random graphs.
Random Structures & Algorithms, 18(1):39–60, 2001.



M. A Zapala and N. J Schork.
Multivariate regression analysis of distance matrices for testing associations between gene expression patterns and related variables.
Proceedings of the national academy of sciences, 103(51):19430–19435, 2006.

Super-motifs

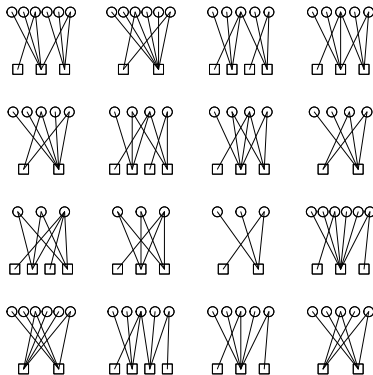
Motif:



Variance:

$$\begin{aligned}
 N_s^2 &= \left(\sum_{\alpha} Y_{s\alpha} \right)^2 \\
 &= \sum_{\alpha, \beta: \alpha \cap \beta = \emptyset} Y_{s\alpha} Y_{s\beta} \\
 &\quad + \underbrace{\sum_{\alpha, \beta: \alpha \cap \beta \neq \emptyset} Y_{s\alpha} Y_{s\beta}}_{\text{occurrence of a super-motif}}
 \end{aligned}$$

Some super-motifs:



... 396 super-motifs

Covariance: same game, for $Y_{s\alpha} Y_{s'\beta}$ with $s \neq s'$ [#3]

Asymptotic distribution of the count

Estimated probability.

$$\bar{\phi}_s := \phi_2 \phi_4 / \phi_1 \quad \rightarrow \quad \bar{F}_s := F_2 F_4 / F_1$$

where F_1, F_2, F_4 = observed frequencies of top stars, bottom stars and edges.

Asymptotic distribution of the count

Estimated probability.

$$\bar{\phi}_s := \phi_2 \phi_4 / \phi_1 \quad \rightarrow \quad \bar{F}_s := F_2 F_4 / F_1$$

where F_1, F_2, F_4 = observed frequencies of top stars, bottom stars and edges.

Asymptotic normality for non-star motifs. Under BEDD (and sparsity conditions):

$$(F_s - \bar{F}_s) / \sqrt{\hat{V}(F_s)} \xrightarrow{m, n \rightarrow \infty} \mathcal{N}(0, 1)$$

Proof:

► decompose

$$F_s - \bar{F}_s = \underbrace{(F_s - \phi_s)}_{\text{random fluctuations}} + \underbrace{(\phi_s - \bar{\phi}_s)}_{\text{null under BEDD}} + \underbrace{(\bar{\phi}_s - \bar{F}_s)}_{\text{estimation error} \rightarrow 0},$$

► construct a counting martingale [GL17] for $F_s - \phi_s$

Asymptotic distribution of the count

Estimated probability.

$$\bar{\phi}_s := \phi_2 \phi_4 / \phi_1 \quad \rightarrow \quad \bar{F}_s := F_2 F_4 / F_1$$

where F_1, F_2, F_4 = observed frequencies of top stars, bottom stars and edges.

Asymptotic normality for non-star motifs. Under BEDD (and sparsity conditions):

$$(F_s - \bar{F}_s) / \sqrt{\hat{V}(F_s)} \xrightarrow{m, n \rightarrow \infty} \mathcal{N}(0, 1)$$

Proof:

► decompose

$$F_s - \bar{F}_s = \underbrace{(F_s - \phi_s)}_{\text{random fluctuations}} + \underbrace{(\phi_s - \bar{\phi}_s)}_{\text{null under BEDD}} + \underbrace{(\bar{\phi}_s - \bar{F}_s)}_{\text{estimation error} \rightarrow 0},$$

► construct a counting martingale [GL17] for $F_s - \phi_s$

Test statistic. Under BEDD:

$$N_s \approx \mathcal{N}(\hat{\mathbb{E}}(N_s), \hat{V}(N_s)) \quad \Leftrightarrow \quad (N_s - \hat{\mathbb{E}}(N_s)) / \sqrt{\hat{V}(N_s)} \approx \mathcal{N}(0, 1)$$

[#3]

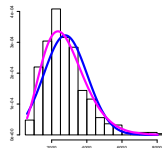
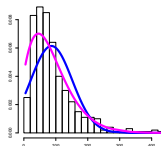
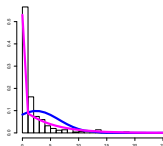
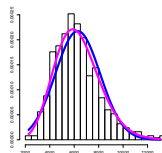
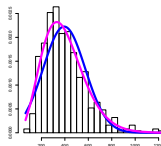
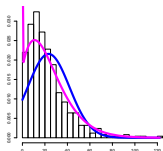
In practice: Asymptotic normality

$$(n = 2m/3)$$

$$m = 50$$

$$m = 100$$

$$m = 200$$



Normal distribution, Poisson-geometric distribution with same mean and variance [Sta01, PDK⁺08] [#3]

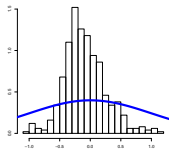
In practice: Test statistic

Need to account for the estimation error of $\hat{\mathbb{E}}N$

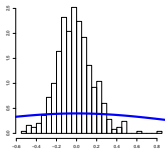
Regular stat.:

$$\frac{N - \hat{\mathbb{E}}N}{\sqrt{\hat{\mathbb{V}}N}}$$

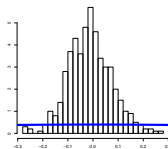
$m = 50$



$m = 100$



$m = 200$

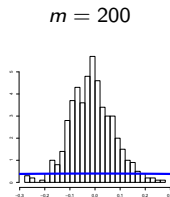
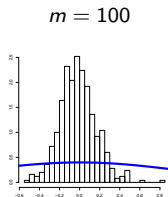
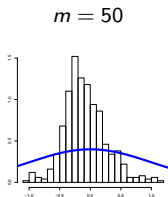


In practice: Test statistic

Need to account for the estimation error of $\hat{\mathbb{E}}N$

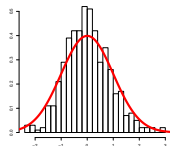
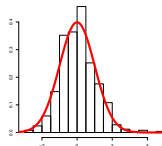
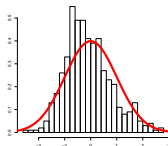
Regular stat.:

$$\frac{N - \hat{\mathbb{E}}N}{\sqrt{\hat{\mathbb{V}}N}}$$



Correction:

$$\frac{N - (\hat{\mathbb{E}}N - \hat{\mathbb{B}}(\hat{\mathbb{E}}N))}{\sqrt{\hat{\mathbb{V}}(N - \hat{\mathbb{E}}N)}}$$



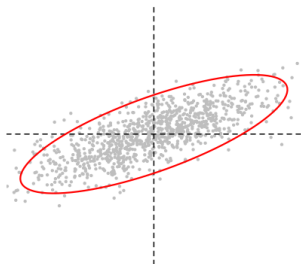
► Need to evaluate $\mathbb{V}(N - \hat{\mathbb{E}}(N))$ and $\mathbb{B}(\hat{\mathbb{E}}N)$: resort to Taylor expansion (Δ -method) [\[#5\]](#)

Choleski transform

Aim: 'Remove' correlation and variance heterogeneity

Covariance matrix of (X_1, X_2) :

$$\Sigma_{X_1, X_2} = \begin{bmatrix} \sigma_1^2 & \sigma_{12} \\ \sigma_{12} & \sigma_2^2 \end{bmatrix}$$



Diagonalization: $\Sigma = P\Lambda P^{-1}$

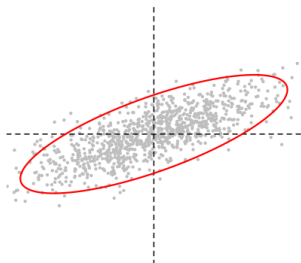
Choleski matrix: $\Sigma^{-1/2} = P\Lambda^{-1/2}P^{-1}$

Choleski transform

Aim: 'Remove' correlation and variance heterogeneity

Covariance matrix of (X_1, X_2) :

$$\Sigma_{X_1, X_2} = \begin{bmatrix} \sigma_1^2 & \sigma_{12} \\ \sigma_{12} & \sigma_2^2 \end{bmatrix}$$

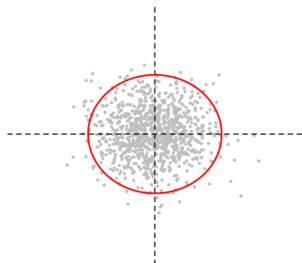


Diagonalization: $\Sigma = P \Lambda P^{-1}$

Choleski matrix: $\Sigma^{-1/2} = P \Lambda^{-1/2} P^{-1}$

Choleski transform:

$$\begin{bmatrix} X'_1 \\ X'_2 \end{bmatrix} = \Sigma^{-1/2} \begin{bmatrix} X_1 \\ X_2 \end{bmatrix}$$



$$\Sigma_{X'_1, X'_2} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

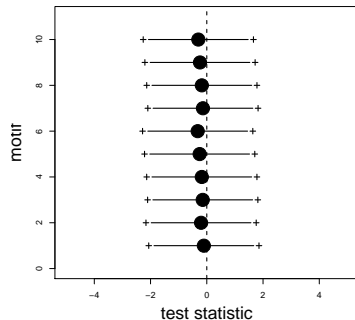
[#4]

Goodness-of-fit (GOF) of the BEDD model

Raw statistic:

$$T_s = \frac{N_s - \hat{\mathbb{E}}N_s}{\sqrt{\hat{\mathbb{V}}N_s}}$$

Zackenberg network.



[#5]

$${}^3\Sigma = P\Lambda P^\top, \Sigma = P\Lambda^{-1/2}P^\top$$

Goodness-of-fit (GOF) of the BEDD model

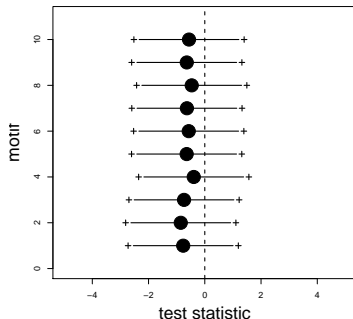
Raw statistic:

$$T_s = \frac{N_s - \hat{\mathbb{E}}N_s}{\sqrt{\hat{\mathbb{V}}N_s}}$$

Corrected stat.: accounts for the estimation error in $\hat{\mathbb{E}}N$

$$T'_s = \frac{N_s - (\hat{\mathbb{E}}N_s - \hat{\mathbb{B}}(\hat{\mathbb{E}}N_s))}{\sqrt{\hat{\mathbb{V}}(N_s - \hat{\mathbb{E}}N_s)}}$$

Zackenberg network.



[#5]

$${}^3\Sigma = P\Lambda P^\top, \Sigma = P\Lambda^{-1/2}P^\top$$

Goodness-of-fit (GOF) of the BEDD model

Raw statistic:

$$T_s = \frac{N_s - \hat{\mathbb{E}}N_s}{\sqrt{\hat{\mathbb{V}}N_s}}$$

Corrected stat.: accounts for the estimation error in $\hat{\mathbb{E}}N$

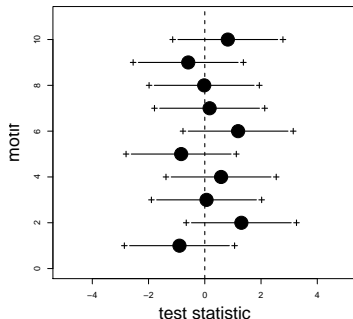
$$T'_s = \frac{N_s - (\hat{\mathbb{E}}N_s - \hat{\mathbb{B}}(\hat{\mathbb{E}}N_s))}{\sqrt{\hat{\mathbb{V}}(N_s - \hat{\mathbb{E}}N_s)}}$$

Cholevski³ transformation: accounts for the correlation between the counts

$$\Sigma_{s,s'} = \text{Cov}(N_s - \hat{\mathbb{E}}N_s, N_{s'} - \hat{\mathbb{E}}N_{s'})$$

$$T'' = \hat{\Sigma}^{-1/2} \left[N_s - (\hat{\mathbb{E}}N_s - \hat{\mathbb{B}}(\hat{\mathbb{E}}N_s)) \right]$$

Zackenberg network.



[#5]

³ $\Sigma = P\Lambda P^T, \Sigma = P\Lambda^{-1/2}P^T$

An intuition for Adonis

Variance. Remind that
$$\sum_{i=1}^n (y_i - \bar{y})^2 = \frac{1}{2n} \sum_{i=1}^n \sum_{i'=1}^n (y_i - y_{i'})^2$$

An intuition for Adonis

Variance. Remind that $\sum_{i=1}^n (y_i - \bar{y})^2 = \frac{1}{2n} \sum_{i=1}^n \sum_{i'=1}^n (y_i - y_{i'})^2$

Analysis of variance. Remind that, for g groups, with r replicates ($n = r g$),

$$\underbrace{\sum_{i=1}^g \sum_{j=1}^r (y_{ij} - \bar{y})^2}_{\text{Total}} = \underbrace{\sum_{i=1}^g r (\bar{y}_i - \bar{\bar{y}})^2}_{\text{Between}} + \underbrace{\sum_{i=1}^g \sum_{j=1}^r (y_{ij} - \bar{y}_i)^2}_{\text{Within}}$$

Anova test statistic $F \propto \text{Between/Within}$

An intuition for Adonis

Variance. Remind that $\sum_{i=1}^n (y_i - \bar{y})^2 = \frac{1}{2n} \sum_{i=1}^n \sum_{i'=1}^n (y_i - y_{i'})^2$

Analysis of variance. Remind that, for g groups, with r replicates ($n = r g$),

$$\underbrace{\sum_{i=1}^g \sum_{j=1}^r (y_{ij} - \bar{y})^2}_{\text{Total}} = \underbrace{\sum_{i=1}^g r (\bar{y}_i - \bar{\bar{y}})^2}_{\text{Between}} + \underbrace{\sum_{i=1}^g \sum_{j=1}^r (y_{ij} - \bar{y}_i)^2}_{\text{Within}}$$

Anova test statistic $F \propto \text{Between/Within}$

Distance. p variables $y_{ij} = (y_{ijk})_{k=1\dots p}$: $D^2(y_{ij}, y_{i'j'}) = \sum_{k=1}^p (y_{ijk} - y_{i'j'k})^2$

An intuition for Adonis

Variance. Remind that $\sum_{i=1}^n (y_i - \bar{y})^2 = \frac{1}{2n} \sum_{i=1}^n \sum_{i'=1}^n (y_i - y_{i'})^2$

Analysis of variance. Remind that, for g groups, with r replicates ($n = r g$),

$$\underbrace{\sum_{i=1}^g \sum_{j=1}^r (y_{ij} - \bar{y})^2}_{\text{Total}} = \underbrace{\sum_{i=1}^g r (\bar{y}_i - \bar{\bar{y}})^2}_{\text{Between}} + \underbrace{\sum_{i=1}^g \sum_{j=1}^r (y_{ij} - \bar{y}_i)^2}_{\text{Within}}$$

Anova test statistic $F \propto \text{Between/Within}$

Distance. p variables $y_{ij} = (y_{ijk})_{k=1\dots p}$: $D^2(y_{ij}, y_{i'j'}) = \sum_{k=1}^p (y_{ijk} - y_{i'j'k})^2$

Adonis. Define Between = Total - Within, where

$$\text{Total} = \frac{1}{2n} \sum_{i=1}^g \sum_{i'=1}^g \sum_{j'=1}^r D^2(y_{ij}, y_{i'j'}), \quad \text{Within} = \frac{1}{2} \sum_{i=1}^g \sum_{j=1}^r \sum_{j'=1}^r D^2(y_{ij}, y_{ij'}),$$

and determine the p -value for $F = \text{Between/Within}$ by permutation. [#6]

Null simulations

100 sets of 83 synthetic networks, with dimensions as the original ones and fixed ρ , g , and h

Without correction.

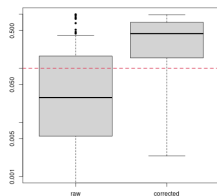
	Df	SumOfSqs	R^2	F	Pr(F)
Year	1	1.5	0.0114	0.94	0.443
Month	6	25.17	0.1914	2.61	0.012
Region	2	0.69	0.0052	0.21	0.845
Year:Month	6	7.46	0.0568	0.77	0.653
Year:Region	2	2.52	0.0191	0.78	0.543
Month:Region	12	12.55	0.0954	0.65	0.831
Year:Month:Region	12	17.39	0.1322	0.9	0.578
Residual	40	64.23	0.4884		
Total	81	131.51	1		

With correction.

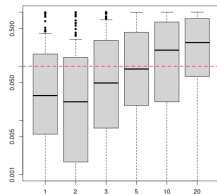
	Df	SumOfSqs	R^2	F	Pr(F)
insectNb	1	10.55	0.0803	7.17	0.003
plantNb	1	7.32	0.0556	4.97	0.015
Year	1	0.77	0.0059	0.53	0.622
Month	6	8.96	0.0681	1.01	0.458
Region	2	4.39	0.0334	1.49	0.241
Year:Month	6	8.54	0.0649	0.97	0.496
Year:Region	2	2.17	0.0165	0.74	0.569
Month:Region	12	13.99	0.1064	0.79	0.699
Year:Month:Region	12	18.9	0.1437	1.07	0.421
Residual	38	55.93	0.4253		
Total	81	131.51	1		

[#6]

Over 100 networks.



Increasing the networks' size.



Results for other motif distances

Plant imbalance $D^{(h)}$.

	Df	Sum Of Sqs	R^2	F	Pr(F)
insectNb	1	9.81	0.1935	19.53	1e-05
plantNb	1	1.42	0.028	2.83	0.06594
Year	1	0.67	0.0133	1.34	0.27518
Month	6	4.14	0.0818	1.38	0.20713
Region	2	0.99	0.0196	0.99	0.43126
Year:Month	6	2.61	0.0514	0.86	0.58258
Year:Region	2	1.87	0.037	1.87	0.1283
Month:Region	12	4.33	0.0854	0.72	0.8062
Year:Month:Region	12	5.75	0.1135	0.95	0.53039
Residual	38	19.09	0.3766		
Total	81	50.69	1		

Both imbalance $D^{(gh)}$.

	Df	Sum Of Sqs	R^2	F	Pr(F)
insectNb	1	56.21	0.265	48.58	1e-05
plantNb	1	16.44	0.0775	14.21	1e-05
Year	1	0.83	0.0039	0.71	0.56559
Month	6	24.41	0.1151	3.52	0.00012
Region	2	8.83	0.0417	3.82	0.00321
Year:Month	6	10.09	0.0476	1.45	0.14299
Year:Region	2	3.15	0.0149	1.36	0.25071
Month:Region	12	29.92	0.1411	2.16	0.0028
Year:Month:Region	12	18.24	0.086	1.31	0.16496
Residual	38	43.96	0.2073		
Total	81	212.09	1		

[#6]