

7. Advanced Topics

Overview

1. Categorical variables
2. Non-linearities
3. Phylogenetic correlations

7.1 Categorical Variables.

- Binary variables = 0 or 1
- Map to continuous variable if ordered factor (e.g., low = 1, medium = 2, high = 3)
- Turn into multiple dummy variables (presence or absence of a treatment level)
- Run separate SEM for each level of factor (assuming sufficient power)
- If categorical variable is exogenous, simply estimate marginal means for each level and represent with multiple paths

Jutila (1997): Grazing in Finnish coastal meadows



Grazing changes the relationship
between elevation, biomass, and
diversity

7.1 Categorical Variables.

```
# Load data
jutila <- read.csv("jutila.csv")

# Create SEM
jutila.sem <- psem(

  lm(mass ~ elev, jutila),

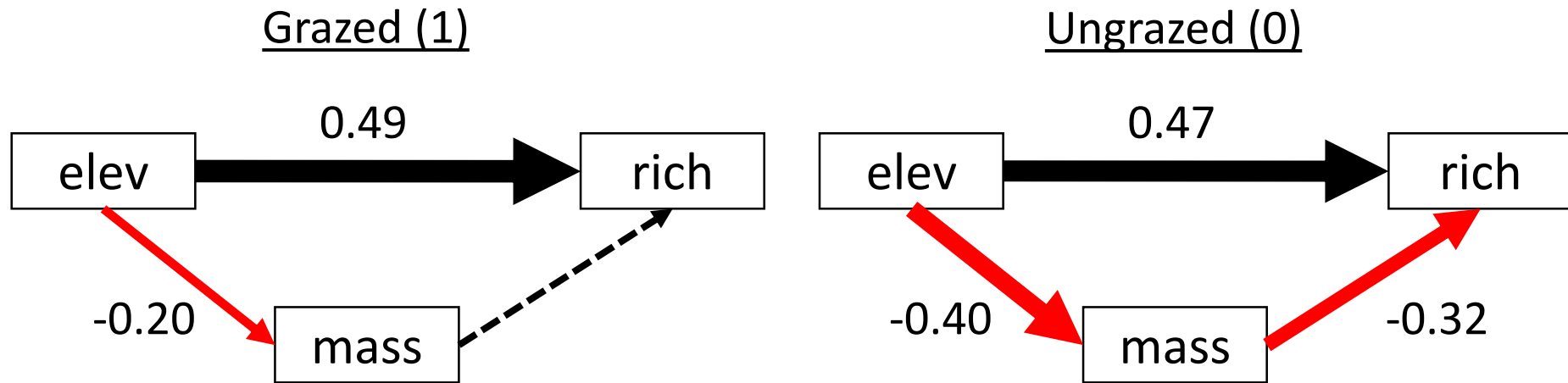
  lm(rich ~ mass + elev, jutila),

  data = jutila

)

# Specifying grouping in summary
summary(jutila.sem, groups = "grazed")
```

7.1 Categorical Variables.



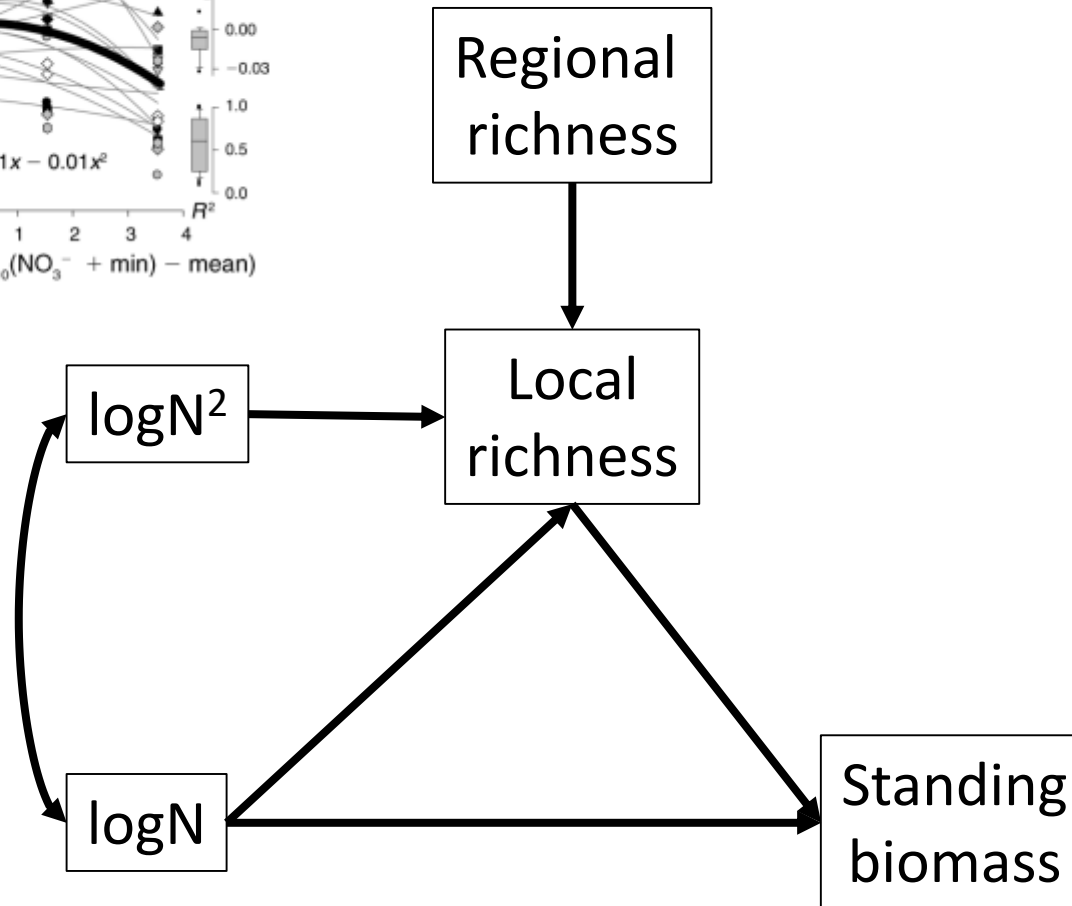
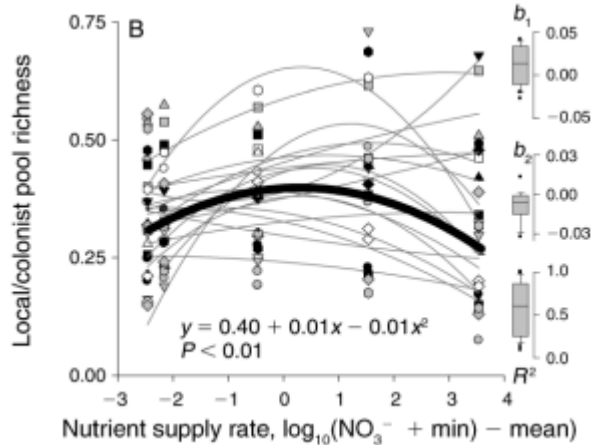
7.2 Non-linearities

- Treat similar to interactions: code as independent linear and squared, cubed, etc. terms
- Currently squirrely if transformed inside model formulae, e.g., `lm(y ~ poly(x, 2))`

Cardinale et al. (2009) - Does diversity drive productivity or vice versa?



7.2 Non-linearities. Cardinale (2009)



7.2 Non-linearities. Cardinale (2009)

Tests of directed separation:

	Independ.	Claim	Estimate	Std.Error	DF	Crit.Value	P.Value
logCh1 ~ SR + ...			0.0088	0.0046	18	1.9176	0.0712

Coefficients:

Response	Predictor	Estimate	Std.Error	DF	Crit.Value	P.Value	Std.Estimate	
SA	logN	-3.2046	1.1307	30	-2.8342	0.0081	-0.5398	**
SA	logN2	-0.5136	0.1783	30	-2.8806	0.0073	-0.5488	**
SA	SR	0.4260	0.0663	18	6.4251	0.0000	0.7651	***
logCh1	SA	0.0117	0.0050	29	2.3528	0.0256	0.2285	*
logCh1	logN	0.0813	0.0781	29	1.0410	0.3065	0.2684	
logCh1	logN2	-0.0022	0.0123	29	-0.1811	0.8576	-0.0467	
~~logN	~~logN2	-0.9685	NA	125	-43.4652	0.0000	-0.9685	***

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 1

Goodness-of-fit:

Global model: Fisher's C = 5.285 with P-value = 0.071 and on 2 degrees of freedom

Individual R-squared:

Response	Marginal	Conditional
SA	0.53	0.75
logCh1	0.16	0.48

7.2 Non-linearities. Cardinale (2009)

Center squared term to reduce collinearity:

Coefficients:

Response	Predictor	Estimate	Std.Error	DF	Crit.Value	P.Value	Std.Estimate	
SA	logN.cen	0.4356	0.3279	30	1.3284	0.1941	0.0734	
SA	logN2.cen	-0.5136	0.1783	30	-2.8806	0.0073	-0.1592	**
SA	SR	0.4260	0.0663	18	6.4251	0.0000	0.7651	***
logCh1	SA	0.0117	0.0050	29	2.3528	0.0256	0.2285	*
logCh1	logN.cen	0.0971	0.0223	29	4.3542	0.0002	0.3206	***
logCh1	logN2.cen	-0.0022	0.0123	29	-0.1811	0.8576	-0.0136	

7.2 Non-linearities. Cardinale (2009)

Get orthogonal linear and squared terms using `poly`

https://en.wikipedia.org/wiki/Orthogonal_polynomials

Coefficients:

Response	Predictor	Estimate	Std.Error	DF	Crit.Value	P.Value	Std.Estimate	
SA	Npoly	-14.6986	7.2071	30	-2.0116	0.0533	-0.0952	
SA	N2poly	-15.7023	7.2771	30	-2.1285	0.0416	-0.1017	*
SA	SR	0.4249	0.0664	18	6.3974	0.0000	0.7631	***
logCh1	SA	0.0119	0.0049	29	2.4031	0.0229	0.2331	*
logCh1	Npoly	1.8490	0.5023	29	3.6814	0.0009	0.2348	***
logCh1	N2poly	-1.6239	0.5071	29	-3.2026	0.0033	-0.2062	**

7.3 Phylogenetic Correlations.

- Species are not independent data points = shared evolutionary history
- Phylogenetic least-squares (PGLS) fixes correlation matrix based on phylogenetic tree

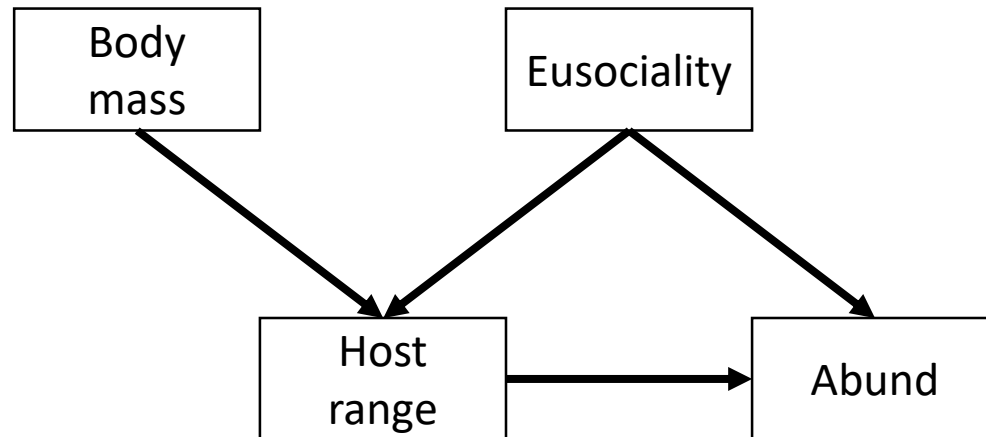
Duffy & Macdonald (2010): Factors driving ecological success of eusocial shrimps in the Caribbean

A close-up photograph of two translucent, eusocial shrimp (likely Hippidae) interacting on a bright yellow, porous sponge. The shrimp are semi-transparent, revealing internal structures. One shrimp is positioned slightly higher and further back than the other, with its head and antennae reaching towards the second shrimp. The background is dark, making the shrimp and the sponge stand out.

Measured body size,
eusociality, range and
abundance of 20
shrimp species

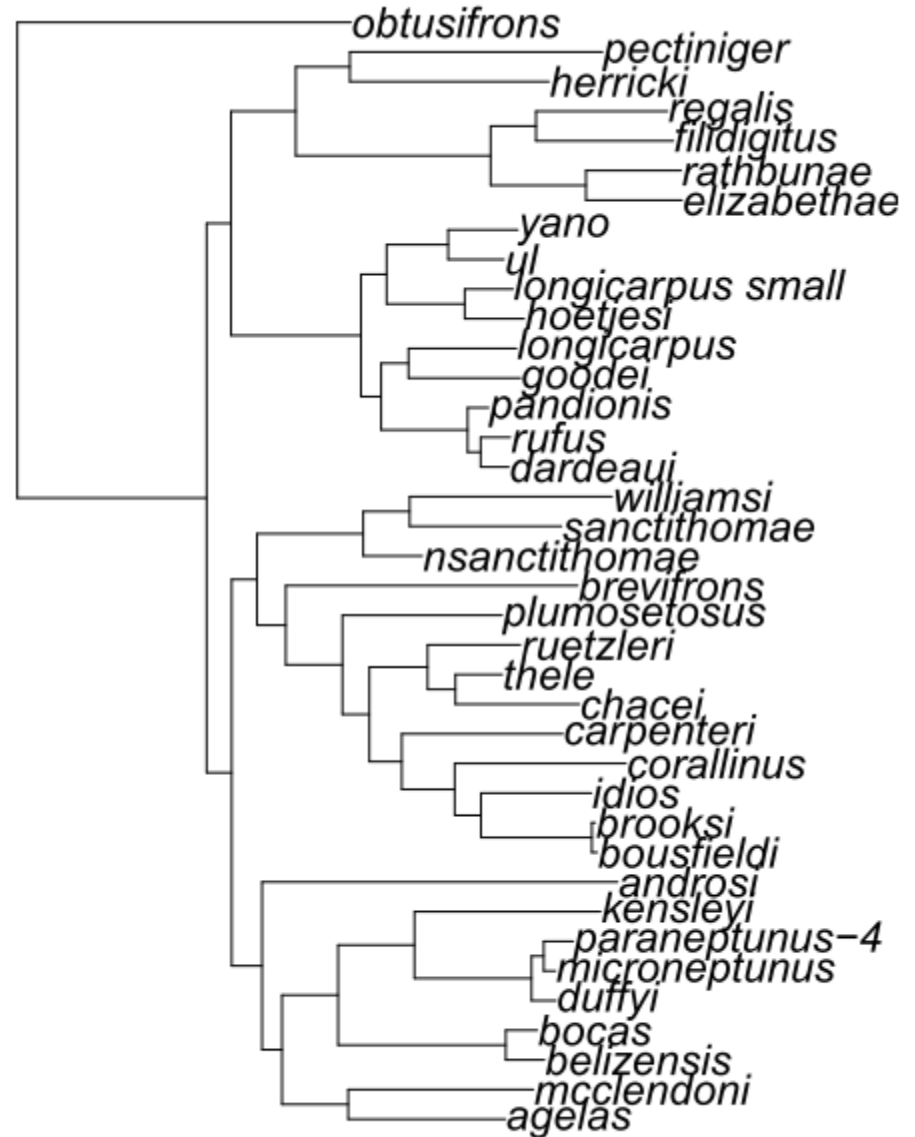
7.3 Phylogenetic Correlations. Duffy example

- Eusociality increases both host range and regional abundance (competitive ability)
- The effect of eusociality on host is mediated by body size (mobility)



7.3 Phylogenetic Correlations. Duffy example

```
duffy <-  
  read.csv("duffy.csv")  
  
tree <- read.tree(  
  "synalpheus_tree.txt")  
  
duffy.merge <-  
  comparative.data(tree,  
    duffy,  
    names.col = "Species")
```



7.3 Phylogenetic Correlations. Duffy example

```
# Without phylogenetic correlations

duffy.sem <- psem(

  lm(Host.Range ~ Eusociality.index + Total.Mass.Female,
duffy.merge$data),

  lm(Rubble.Abundance ~ Eusociality.index + Host.Range,
duffy.merge$data)

)
```

7.3 Phylogenetic Correlations. Duffy example

Tests of directed separation:

		Independ.Claim	Estimate	Std.Error	DF	Crit.Value	P.Value
Rubble.Abundance	~	Total.Mass.Female + ...	-0.0925	0.1275	15	-0.7257	0.4792

Coefficients:

	Response	Predictor	Estimate	Std.Error	DF	Crit.Value	P.Value	Std.Estimate
	Host.Range	Eusociality.index	2.2144	0.7723	16	2.8672	0.0112	0.5792 *
	Host.Range	Total.Mass.Female	-0.0101	0.0228	16	-0.4426	0.6640	-0.0894
	Rubble.Abundance	Eusociality.index	8.3001	5.2063	16	1.5942	0.1304	0.3393
	Rubble.Abundance	Host.Range	3.0202	1.3619	16	2.2177	0.0414	0.4720 *

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 1

Goodness-of-fit:

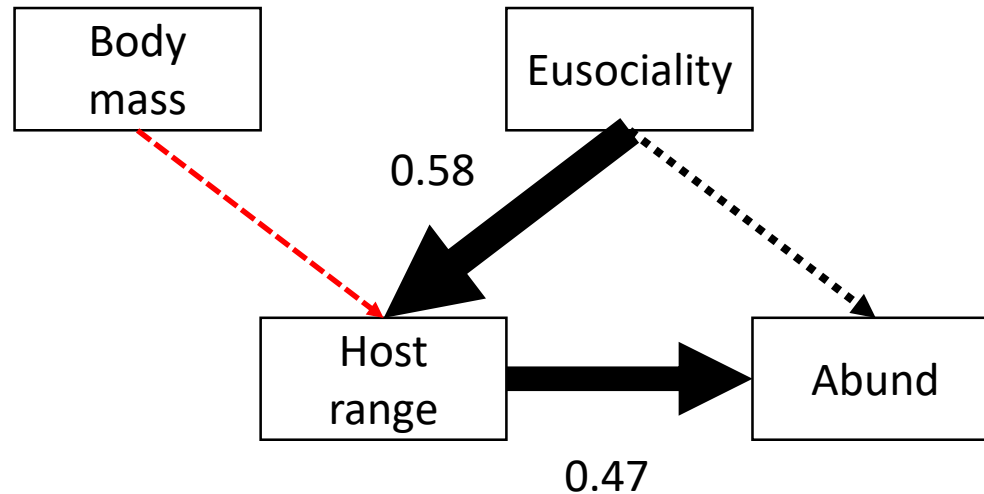
Global model: Fisher's C = 1.471 with P-value = 0.479 and on 2 degrees of freedom

Individual R-squared:

	Response	R.squared
	Host.Range	0.35
	Rubble.Abundance	0.53

7.3 Phylogenetic Correlations. Duffy example

Without phylogenetic covariances



7.3 Phylogenetic Correlations. Duffy example

```
# With phylogenetic correlations

duffy.sem2 <- psem(

  pglis(Host.Range ~ Eusociality.index + Total.Mass.Female,
duffy.merge),

  pglis(Rubble.Abundance ~ Eusociality.index + Host.Range,
duffy.merge),

  data = duffy.merge

)
```

7.3 Phylogenetic Correlations. Duffy example

Tests of directed separation:

		Independ.Claim	Estimate	Std.Error	DF	Crit.Value	P.Value
Rubble.Abundance	~	Total.Mass.Female + ...	0.024	0.0937	NA	0.2565	0.801

Coefficients:

	Response	Predictor	Estimate	Std.Error	DF	Crit.Value	P.Value	Std.Estimate
	Host.Range	Eusociality.index	3.1786	1.0270	19	3.0952	0.0069	0.8314 **
	Host.Range	Total.Mass.Female	-0.0377	0.0142	19	-2.6569	0.0172	-0.3342 *
	Rubble.Abundance	Eusociality.index	2.2544	5.5404	19	0.4069	0.6895	0.0922
	Rubble.Abundance	Host.Range	2.9416	1.1123	19	2.6446	0.0177	0.4597 *

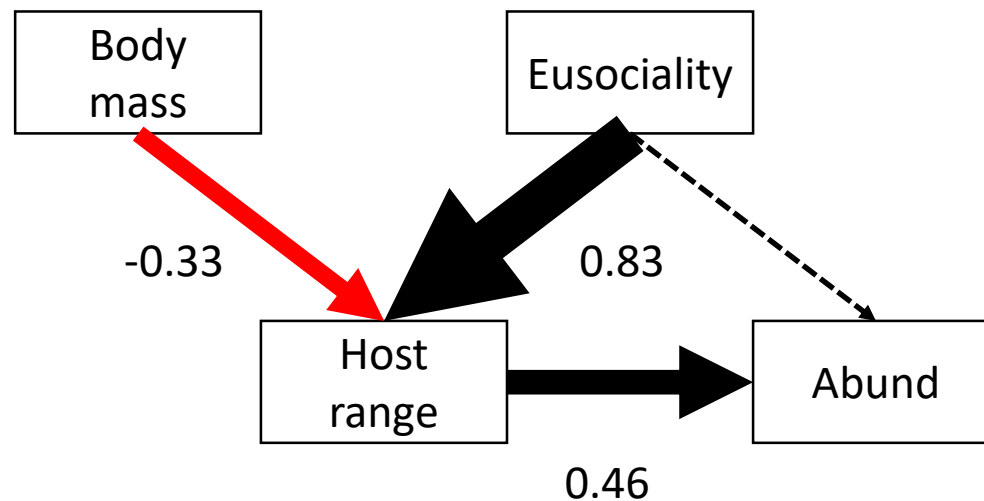
Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 1

Goodness-of-fit:

Global model: Fisher's C = 0.444 with P-value = 0.801 and on 2 degrees of freedom

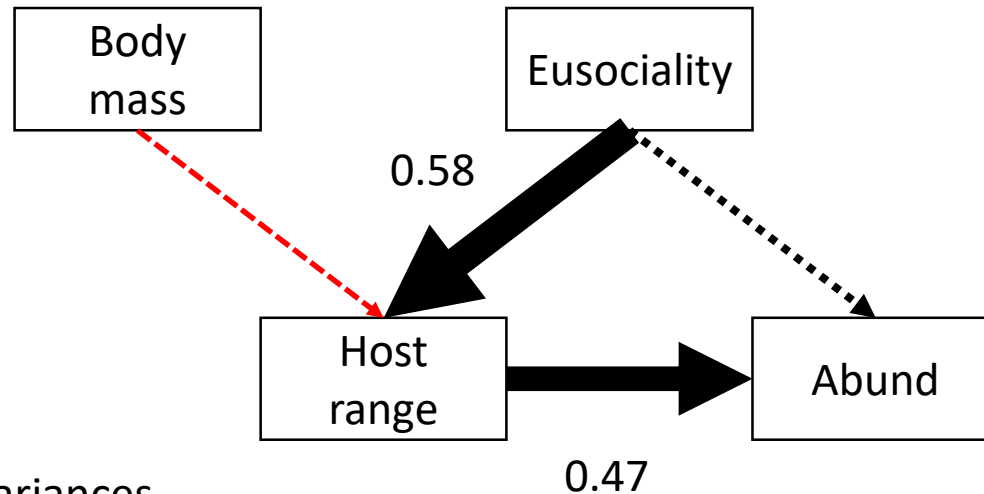
7.3 Phylogenetic Correlations. Duffy example

- Eusociality directly increases host range, indirectly increases regional abundance by increasing host range
- Body mass important



7.3 Phylogenetic Correlations. Duffy example

Without phylogenetic covariances



With phylogenetic covariances

