MULTIVARIATE STATISTICS

Stats and R workshop 5th – 6th March 2018

Susan Jarvis

Centre for Ecology & Hydrology
NATURAL ENVIRONMENT RESEARCH COUNCIL

NERC SCIENCE OF THE ENVIRONMENT

What is multivariate statistics?

- Univariate statistics – **single response**

- Multivariate statistics – **multiple responses**

What is multivariate statistics?

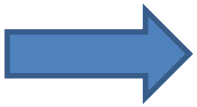
- Univariate statistics – **single response**

- Multivariate statistics – **multiple responses**

What does multivariate statistics do?

- Summarise multivariate data into a smaller number of variables that represent the main sources of variation in the data (ordination)

Site	Sp1	Sp2	Sp3	Sp4	Sp5
1	0	4	3	5	1
2	0	0	0	3	0
3	1	1	2	2	0
4	2	2	0	2	1
5	6	0	4	0	8
6	0	0	6	0	2

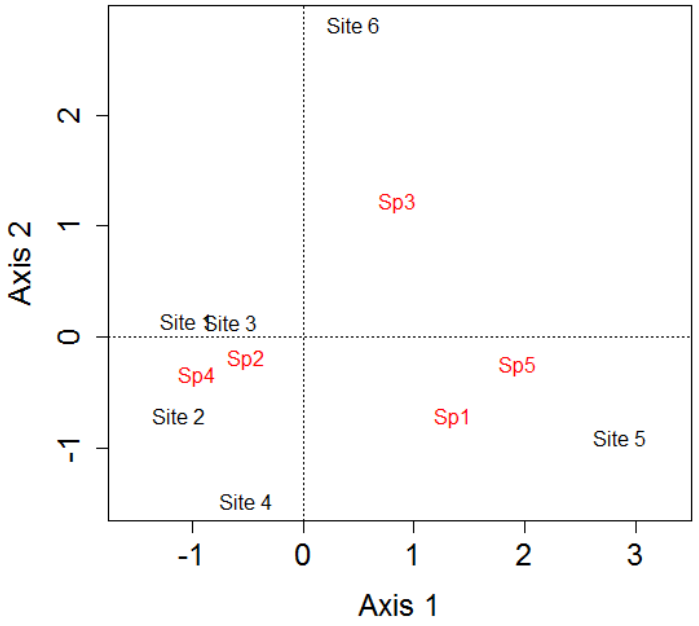


Site	Axis 1	Axis 2
1	-1.04	0.14
2	-1.12	-0.70
3	-0.65	0.13
4	-0.51	-1.48
5	2.87	-0.91
6	0.46	2.82

What does multivariate statistics do?

- Summarise multivariate data into a smaller number of variables that represent the main sources of variation in the data (ordination)

Site	Sp1	Sp2	Sp3	Sp4	Sp5
1	0	4	3	5	1
2	0	0	0	3	0
3	1	1	2	2	0
4	2	2	0	2	1
5	6	0	4	0	8
6	0	0	6	0	2



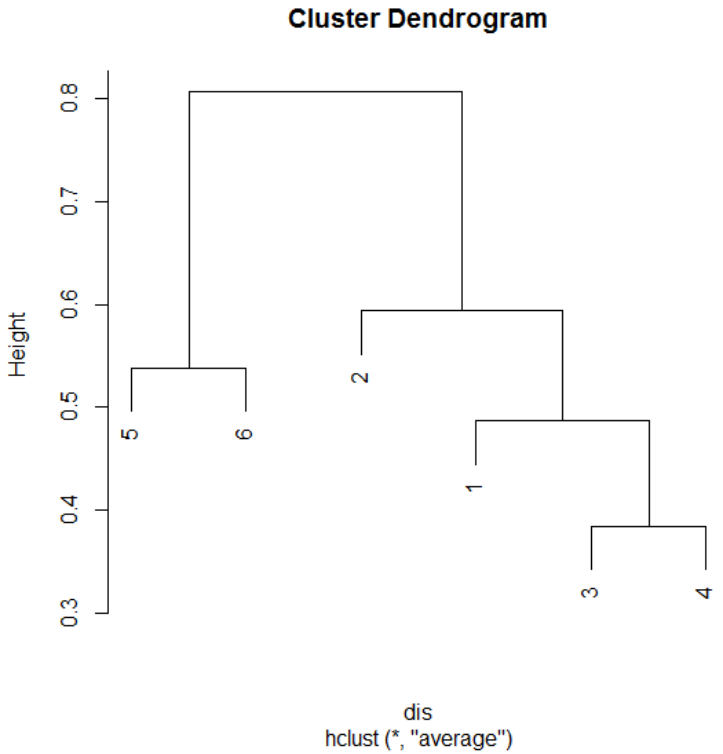
Centre for Ecology & Hydrology
NATURAL ENVIRONMENT RESEARCH COUNCIL

NERC SCIENCE OF THE ENVIRONMENT

What does multivariate statistics do?

- Find groups of samples that share similar properties based on multiple variables (e.g. cluster analysis)

Site	Sp1	Sp2	Sp3	Sp4	Sp5
1	0	4	3	5	1
2	0	0	0	3	0
3	1	1	2	2	0
4	2	2	0	2	1
5	6	0	4	0	8
6	0	0	6	0	2



What does multivariate statistics do?

- Test for differences between groups of samples based on multiple variables (e.g. MANOVA)

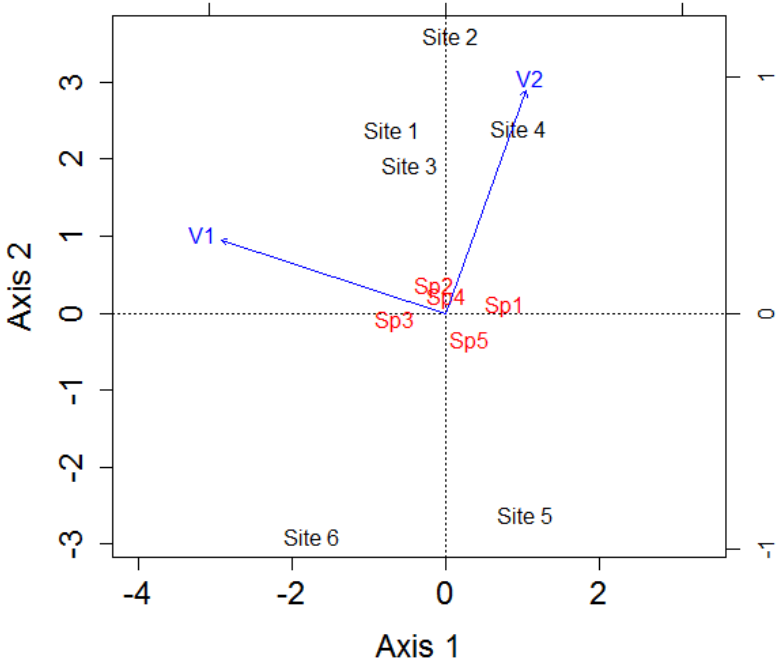
Site	Sp1	Sp2	Sp3	Sp4	Sp5	Group
1	0	4	3	5	1	**A**
2	0	0	0	3	0	**A**
3	1	1	2	2	0	**A**
4	2	2	0	2	1	**A**
5	6	0	4	0	8	**B**
6	0	0	6	0	2	**B**

	Df	SS	MS	F	R2	P
Group	1	0.65	0.65	4.61	0.53	0.06
Residuals	4	0.56	0.14		0.46	
Total	5	1.21				

What does multivariate statistics do?

- Test whether patterns in variation are explained by independent variables (constrained ordination)

Site	Sp1	Sp2	Sp3	Sp4	Sp5	V1	V2
1	0	4	3	5	1	**4.5**	**1.2**
2	0	0	0	3	0	**2.4**	**1.1**
3	1	1	2	2	0	**5.3**	**2.1**
4	2	2	0	2	1	**2.4**	**1.4**
5	6	0	4	0	8	**2.5**	**1.3**
6	0	0	6	0	2	**5.5**	**0.8**



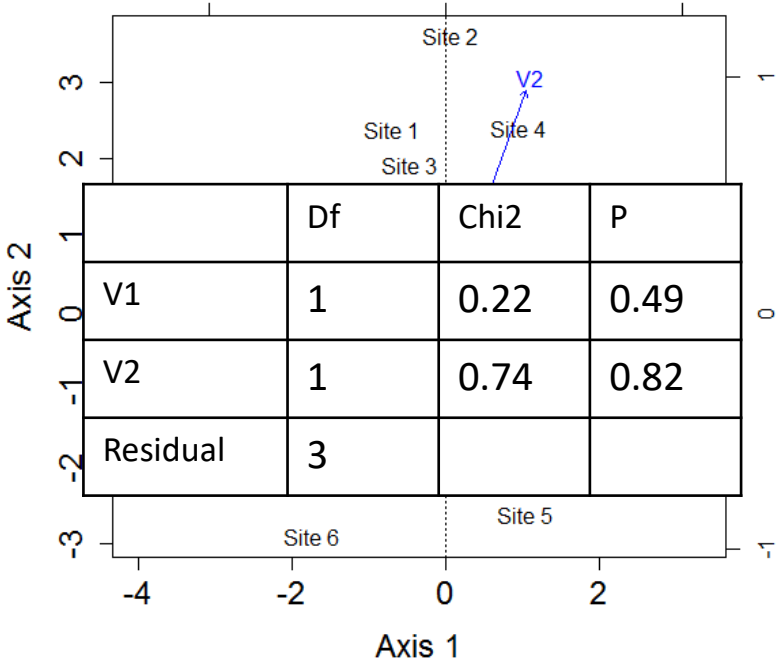
Centre for Ecology & Hydrology
NATURAL ENVIRONMENT RESEARCH COUNCIL

NERC SCIENCE OF THE ENVIRONMENT

What does multivariate statistics do?

- Test whether patterns in variation are explained by independent variables (constrained ordination)

Site	Sp1	Sp2	Sp3	Sp4	Sp5	V1	V2
1	0	4	3	5	1	**4.5**	**1.2**
2	0	0	0	3	0	**2.4**	**1.1**
3	1	1	2	2	0	**5.3**	**2.1**
4	2	2	0	2	1	**2.4**	**1.4**
5	6	0	4	0	8	**2.5**	**1.3**
6	0	0	6	0	2	**5.5**	**0.8**



	Df	Chi2	P
V1	1	0.22	0.49
V2	1	0.74	0.82
Residual	3		

Practical

- Use R package ‘vegan’
- Includes:
  - Ordination
  - Constrained ordination
  - MANOVA with dissimilarities
  - Clustering methods