Discriminant Function

- General class of analyses that build a model to predict a single continuous or categorical dependent variable from a set of predictor variables.

- Analysis looks for linear combinations of the data that effectively define or predict the groups.

- There are a variety of methods, but dfa often uses ordination for reducing the dimensionality of the predictor variables.

Discriminant Function

- The result is a predictive model (discriminant function) that can be used to:
  - Emphasize differences among groups
  - Build a predictive model of group membership

- Tests of classification accuracy – use half of the data to build a model, test classification accuracy with the second half

- Function `discrimin.coa` (ade4 package)
  - Provide a matrix of predictor variables and one classification variable
  - Uses CA for data summary then uses classification to form canonical variables
### Discriminant Function

**Common use** – “Can group membership be determined from variables alone?”

- Divide data into “training” and “testing”
- Perform discriminant function with “training” data
  - Function `lda`
    ```r
discrim<-lda(spaeth_train,grouping=spaeth_factors$creek)
```
- Use “testing” data to see if you can predict creek
  ```r
  predict lda(discrim,spaeth_test)
  ```
  - Returns predicted class (creek in this case) for each sample in the testing data.
- Classification accuracy
Discriminant Function – predicted and actual group

Training dataset (37 samples picked randomly):
- big: 4 0 0 0
- bouie: 0 8 0 0
- hayden: 0 0 16 0
- shelton: 0 0 0 9

Testing dataset (other 39 samples):
- big: 3 1 1 0
- bouie: 1 6 4 4
- hayden: 3 2 3 2
- shelton: 0 0 4 5

17 Correct, 22 incorrect
44% accurate, null = 25%
Most accurate for Big (3/5) and Shelton (5/9)
Least accurate for Bouie (6/15) and Hayden (3/10)

Canonical Analyses

- Analysis aimed at identifying the relationship between two multivariate datasets.

- Constrained Ordinations (next week)
  - Redundancy Analysis (RDA)

- Canonical Correspondence Analysis

- Canonical Correlation – measure of association between two sets of variables (function `cancor`).

- Discriminant Function and Canonical Variates (shape analyses)
General Recommendations for Ordination:

A dichotomous key

1. Shape, overlap, size
2. Texture, location
3. Diffuseness, smoothness
4. Petals, color, size
5. Sepals, form, size
6. Bloom, size, color
7. Flower, size, shape
8. Fruit, size, shape, color
9. Seed, size, shape

This page was created and is maintained by Michael Smith.