

# An Introduction to Linear Discriminant Function Analysis (and a comparison with multiple regression)

## Multiple Regression

### Basic Data:

1 quantitative criterion &  
2+ quantitative, binary or coded predictors

### Does the model work?

F-test of  $H_0: R^2 = 0$

### How well does the model work?

$R^2$

### Which variables contribute to the model?

t-test of  $H_0: b=0$  for each variable

### Relative importance of contributors?

Very "gentle" interpretation of Beta weights  
Relative changes in  $R^2$  when drop different predictors

### Hypothesis-testing

$R^2\Delta$  F-test for nested models  
 $R^2$  difference Z-test for non-nested models

### Using the model

Computation and use of  $y'$

## Linear Discriminant Function

1 qualitative grouping variable &  
2+ quantitative, binary or coded discriminating variables

- 1) Sphericity Test  $X^2$  test - are there between group mean difs?
- 2) F-test of  $H_0: R_c^2 = 0$  bg variance attributable to model
- 3) re-classification accuracy (e.g., t-test of % accuracy)

### How many contributing functions are there?

- 1) Max # is smaller of #grps-1 or #discriminating variables
- 2) "kept #" based on significance tests, relative contribution, interpretability, contribution to %correct reclassification
- 3) distinction between **concentrated** and **diffuse** data structures

$\Lambda$ ,  $R_c^2$  & % correct reclassification

Each "kept" function is "interpreted"

- 1) raw score discriminant/regression weights
- 2) standardized disc/reg weights
- 3) structure weights (correlation of ldf and each var)\*
  - based on a "cutoff," usually of .3 or .4
  - preferred by many, because no colinearity influence

**What does each function "do for us"** - what groups are discriminated by **each** interpreted function?

- 1) Follow-up analyses on the ldf (pairwise comparisons)
- 2) Contribution to re-classification

### What does the set of ldfs "do for us"

Working together, what groups can and cannot be discriminated?

Equally "gentle" interpretation of structure weights  
Same, but compare  $R_c^2$  and % reclassification as well

$R_c^2$  for nested models (same tests)

Test for  $X^2\Delta$  for nested models

McNemar's test for change in % reclassification for either

Computation and use of ldf score(s) & comparisons to cutoffs for Classification decisions