

### Overview of Lecture

- Multivariate Analysis of Variance
  - What is MANOVA?
  - Why use MANOVA
  - The Assumptions of MANOVA
  - Example MANOVA
- Discriminant Functions Analysis
  - What is DFA?
  - Why use DFA
  - The Assumptions of DFA
  - Example DFA

### What is MANOVA?

- Multivariate analysis of variance is used to perform an ANOVA style analysis on several dependent variables simultaneously.
- MANOVA answers the question
  - Does the combination of several DVs vary with respect to the IVs?
  - For example, do surgeons and psychiatrists differ in terms of the following personality traits: Abasement, Achievement, Aggression, Dominance, Impulsivity, Nurturance?
- In MANOVA a new DV is created that attempts to maximise the differences between the treatment groups
- The new DV is a linear combination of the DVs

### Advantages of MANOVA

- In comparison to ANOVA, MANOVA has the following advantages
  - The researcher improves their chances of finding what changes as a result of the experimental treatment
  - Since only 'one' DV is tested the researcher is protected against inflating the type 1 error due to multiple comparisons
  - It can show differences that individual ANOVAs do not – it is sometimes more powerful

### Assumptions of MANOVA

- Multivariate Normality
  - The sampling distributions of the DVs and all linear combinations of them are normal.
- Homogeneity of Variance-Covariance Matrices
  - Box's M tests this but it is advised that  $p < 0.001$  is used as criterion
- Linearity
  - It is assumed that linear relationships between all pairs of DVs exist
- Multicollinearity and Singularity
  - Multicollinearity – the relationship between pairs of variables is high ( $r > .80$ )
  - Singularity – a variable is redundant; a variable is a combination of two or more of the other variables.

### Example MANOVA

- A group of children with moderate learning difficulties were assessed on a number of measures
  - IQ, Maths, Reading Accuracy, Reading Comprehension, Communication Skill.
- The children were divided into four groups on the basis of gender (male, female) and season of birth (summer, not summer)
- A MANOVA was performed using gender and season of birth as the IVs and IQ mathematics, reading accuracy, reading comprehension and communication skills as the dependent variables.

Based on Bibby et al (1996)

### Example MANOVA – Descriptive Statistics

3. Gender \* Season of Birth

Dependent Variable	Gender	Season of Birth	Mean	Std. Error	95% Confidence Interval		
					Lower Bound	Upper Bound	
IQ	Female	Not Summer	52.486	2.708	47.059	57.913	
		Summer	57.414	2.141	53.140	61.688	
	Male	Not Summer	60.878	1.626	57.233	64.523	
		Summer	73.779	1.786	70.214	77.344	
	Mathematical Ability	Female	Not Summer	2.218	.665	.891	3.545
			Summer	4.245	.526	3.195	5.294
Male		Not Summer	3.720	.448	2.825	4.615	
		Summer	5.035	.439	4.160	5.911	
Reading Accuracy		Female	Not Summer	7.041	.385	6.273	7.809
			Summer	7.628	.304	7.023	8.233
	Male	Not Summer	7.372	.259	6.854	7.890	
		Summer	7.487	.254	6.991	8.003	
	Reading Comprehension	Female	Not Summer	7.599	.244	7.111	8.087
			Summer	8.441	.193	8.055	8.827
Male		Not Summer	7.971	.165	7.642	8.300	
		Summer	8.754	.161	8.442	9.066	
Communication Skill		Female	Not Summer	6.139	.535	5.071	7.207
			Summer	7.256	.423	6.412	8.101
	Male	Not Summer	6.995	.361	6.275	7.715	
		Summer	8.084	.353	7.380	8.788	

### Example Manova – Testing Assumptions

**Bor's Test of Equality of Covariance Matrices**

Bor's M	28.543
F	.770
df1	30
df2	2990.804
Sig.	.810

Tests the null hypothesis that the observed covariance matrices of the dependent variables are equal across groups.

<sup>a</sup> Design: Intercept+GENDER\*SOB+GENDER\*SOB

**Levene's Test of Equality of Error Variances**

IC	F	df1	df2	Sig.
Mathematical Ability	2.003	3	67	.122
Reading Accuracy	1.250	3	67	.295
Reading Comprehension	1.471	3	67	.230
Communication Skill	1.282	3	67	.284

Tests the null hypothesis that the error variance of the dependent variables is equal across groups.

<sup>a</sup> Design: Intercept+GENDER\*SOB+GENDER\*SOB

- Do not reject the assumption of homogeneity of variance-covariance matrices
- Do not reject the assumption of homogeneity of variance

### Example Manova – Multivariate Tests

**Multivariate Test<sup>a</sup>**

Effect		Value	F	Hypothesis df	Error df	Sig.	Partial Eta Squared
Intercept	Pillai's Trace	.995	2296.239 <sup>b</sup>	5.000	63.000	.000	.995
	Wilks' Lambda	.005	2296.239 <sup>b</sup>	5.000	63.000	.000	.995
	Hotelling's Trace	182.241	2296.239 <sup>b</sup>	5.000	63.000	.000	.995
	Roy's Largest Root	182.241	2296.239 <sup>b</sup>	5.000	63.000	.000	.995
GENDER	Pillai's Trace	.374	7.542 <sup>b</sup>	5.000	63.000	.000	.374
	Wilks' Lambda	.626	7.542 <sup>b</sup>	5.000	63.000	.000	.374
	Hotelling's Trace	.599	7.542 <sup>b</sup>	5.000	63.000	.000	.374
	Roy's Largest Root	.599	7.542 <sup>b</sup>	5.000	63.000	.000	.374
SOB	Pillai's Trace	.388	7.974 <sup>b</sup>	5.000	63.000	.000	.388
	Wilks' Lambda	.612	7.974 <sup>b</sup>	5.000	63.000	.000	.388
	Hotelling's Trace	.633	7.974 <sup>b</sup>	5.000	63.000	.000	.388
	Roy's Largest Root	.633	7.974 <sup>b</sup>	5.000	63.000	.000	.388
GENDER * SOB	Pillai's Trace	.104	1.460 <sup>b</sup>	5.000	63.000	.214	.104
	Wilks' Lambda	.896	1.460 <sup>b</sup>	5.000	63.000	.214	.104
	Hotelling's Trace	.116	1.460 <sup>b</sup>	5.000	63.000	.214	.104
	Roy's Largest Root	.116	1.460 <sup>b</sup>	5.000	63.000	.214	.104

<sup>a</sup> Exact statistic

<sup>b</sup> Design: Intercept+GENDER\*SOB+GENDER\*SOB

- Wilks' Lambda is the statistic of choice for most researchers (and should be reported)

### Example Manova – Univariate Tests

Source	Dependent Variable	Sum of Squares	df	Mean Square	F	Sig.
GENDER	IQ	2441.692	1	2441.692	33.279	.000
	Mathematical Ability	20.893	1	20.893	4.723	.033
	Reading Accuracy	.159	1	.159	.107	.744
	Reading Comprehension	1.922	1	1.922	3.219	.077
SOB	IQ	11.275	1	11.275	3.937	.051
	Mathematical Ability	1267.047	1	1267.047	17.269	.000
	Reading Accuracy	44.414	1	44.414	10.041	.002
	Reading Comprehension	2.017	1	2.017	1.363	.247
GENDER * SOB	IQ	19.350	1	19.350	6.756	.011
	Mathematical Ability	251.550	1	251.550	3.429	.068
	Reading Accuracy	2.009	1	2.009	.454	.503
	Reading Comprehension	.846	1	.846	.572	.452
Error	IQ	9.754E-03	1	9.754E-03	.016	.899
	Mathematical Ability	3.149E-03	1	3.149E-03	.001	.974
	Reading Accuracy	4915.794	67	73.370		
	Reading Comprehension	296.371	67	4.423		
Error	Mathematical Ability	99.134	67	1.480		
	Reading Accuracy	40.018	67	.597		
	Reading Comprehension	191.888	67	2.864		
	Communication Skill					

### Example Manova – Significant Differences

**1. Gender**

Dependent Variable	Gender	Mean	Std. Error	95% Confidence Interval	
				Lower Bound	Upper Bound
IQ	Female	54.940	1.720	51.494	58.386
	Male	67.329	1.277	64.779	69.879
Mathematical Ability	Female	3.232	.424	2.385	4.078
	Male	4.378	.314	3.762	5.003
Reading Accuracy	Female	7.334	.245	6.845	7.824
	Male	4.434	.181	7.072	7.798
Reading Comprehension	Female	6.020	.156	7.709	8.331
	Male	6.368	.115	8.138	8.598
Communication Skill	Female	6.898	.341	6.017	7.379
	Male	7.540	.292	7.036	8.043

**2. Season of Birth**

Dependent Variable	Season of Birth	Mean	Std. Error	95% Confidence Interval	
				Lower Bound	Upper Bound
IQ	Not Summer	56.672	1.633	53.412	59.932
	Summer	65.916	1.296	62.214	69.619
Mathematical Ability	Not Summer	2.969	.401	2.169	3.770
	Summer	4.640	.342	3.957	5.323
Reading Accuracy	Not Summer	7.208	.232	6.743	7.660
	Summer	7.982	.198	7.167	7.968
Reading Comprehension	Not Summer	7.785	.147	7.491	8.079
	Summer	8.603	.128	8.351	8.854
Communication Skill	Not Summer	6.967	.353	5.923	7.211
	Summer	7.670	.278	7.120	8.220

### MANOVA

- The pattern of analysis of a MANOVA is similar to ANOVA
  - If there is a significant multivariate effect then examine the univariate effects (i.e. ANOVA for each DV separately)
  - If there is a significant univariate effect then conduct post hoc tests as necessary

### Discriminant Functions Analysis

- The aim of discriminant functions analysis is to find a set of variables that predict membership of groups.
- It is used when groups are already known and the researcher is trying to find out what the differences are between the groups.
- A DFA is approximately a reversal of a MANOVA
  - The assumptions that underlie a DFA are the same as MANOVA
- Predictors are usually chosen on the basis of theory

### Discriminant Functions Analysis

- The basic principle used in DFA is that groups of subjects can be divided on the basis of functions that are linear combinations of the classifying variables.
- Different functions are calculated that maximise the ability to predict membership of groups
- The maximum number of functions calculated is either
  - The number of levels of the grouping variable less one
  - The number of degrees of freedom of the IV.

### Discriminant Functions Analysis

- The questions that can be answered include
  - Can group membership be predicted reliably from a set of predictors?
  - What are the differences between the predictors that predict group membership?
  - What is the degree of association between the predictors and the groups?
  - What proportion of cases are successfully predicted?

### Example DFA

- Can IQ, mathematical ability, reading accuracy, reading comprehension and communication skills predict who is summer born and who is not?
- See earlier example description.
- Not summer born coded as 0 and summer born coded as 1

### Example Data – Can the groups be separated?

Wilks' Lambda

Test of Function(s)	Wilks' Lambda	Chi-square	df	Sig.
1	.686	25.050	5	.000

Standardized Canonical Discriminant Function Coefficients

	Function
	1
IQ	.235
Mathematical Ability	.388
Reading Accuracy	.141
Reading Comprehension	.640
Communication Skill	.359

- The function successfully separates the groups (see Wilks' Lambda)
- The standardised coefficients show the contribution each variable makes to the function

### Example DFA – The correlations between the predictor variables and the function

Structure Matrix

	Function
	1
Reading Comprehension	.743
IQ	.614
Mathematical Ability	.508
Communication Skill	.442
Reading Accuracy	.175

Pooled within-groups correlations between discriminating variables and standardized canonical discriminant functions. Variables ordered by absolute size of correlation within function.

### Example DFA – How successful is the prediction?

Classification Results<sup>a</sup>

	Original Count	Predicted Group Membership		Total
		Not Summer Born	Summer Born	
Season of Birth				
	Not Summer Born	23	9	32
	Summer Born	7	32	39
%	Not Summer Born	71.9	28.1	100.0
	Summer Born	17.9	82.1	100.0

<sup>a</sup>. 77.5% of original grouped cases correctly classified.

### Reading...

- For ANCOVA, MANOVA and DFA the following book is recommended:

Tabachnick, B., & Fidell, L. (1996) 3rd edition  
Or a later edition