

---

# Contents

---

<b>Preface</b>	<b>xv</b>
<b>Authors' Biographies</b>	<b>xix</b>
<b>I Preparation for Analysis</b>	<b>1</b>
<b>1 What is multivariate analysis?</b>	<b>3</b>
1.1 Defining multivariate analysis	3
1.2 Examples of multivariate analyses	3
1.3 Multivariate analyses discussed in this book	7
1.4 Organization and content of the book	12
<b>2 Characterizing data for analysis</b>	<b>15</b>
2.1 Variables: their definition, classification, and use	15
2.2 Defining statistical variables	15
2.3 Stevens's classification of variables	16
2.4 How variables are used in data analysis	19
2.5 Examples of classifying variables	20
2.6 Other characteristics of data	21
2.7 Summary	21
2.8 Problems	22
<b>3 Preparing for data analysis</b>	<b>23</b>
3.1 Processing data so they can be analyzed	23
3.2 Choice of a statistical package	24
3.3 Techniques for data entry	26
3.4 Organizing the data	33
3.5 Example: depression study	40
3.6 Summary	43
3.7 Problems	45

<b>4</b>	<b>Data screening and transformations</b>	<b>47</b>
4.1	Transformations, assessing normality and independence	47
4.2	Common transformations	47
4.3	Selecting appropriate transformations	51
4.4	Assessing independence	61
4.5	Summary	62
4.6	Problems	64
<b>5</b>	<b>Selecting appropriate analyses</b>	<b>67</b>
5.1	Which analyses to perform?	67
5.2	Why selection is often difficult	67
5.3	Appropriate statistical measures	68
5.4	Selecting appropriate multivariate analyses	72
5.5	Summary	74
5.6	Problems	74
<b>II</b>	<b>Applied Regression Analysis</b>	<b>79</b>
<b>6</b>	<b>Simple regression and correlation</b>	<b>81</b>
6.1	Chapter outline	81
6.2	When are regression and correlation used?	81
6.3	Data example	82
6.4	Regression methods: fixed- $X$ case	84
6.5	Regression and correlation: variable- $X$ case	89
6.6	Interpretation: fixed- $X$ case	90
6.7	Interpretation: variable- $X$ case	91
6.8	Other available computer output	95
6.9	Robustness and transformations for regression	103
6.10	Other types of regression	106
6.11	Special applications of regression	108
6.12	Discussion of computer programs	112
6.13	What to watch out for	113
6.14	Summary	115
6.15	Problems	115
<b>7</b>	<b>Multiple regression and correlation</b>	<b>119</b>
7.1	Chapter outline	119
7.2	When are regression and correlation used?	120
7.3	Data example	120
7.4	Regression methods: fixed- $X$ case	123
7.5	Regression and correlation: variable- $X$ case	125

CONTENTS	ix
7.6 Interpretation: fixed- $X$ case	131
7.7 Interpretation: variable- $X$ case	134
7.8 Regression diagnostics and transformations	137
7.9 Other options in computer programs	143
7.10 Discussion of computer programs	148
7.11 What to watch out for	151
7.12 Summary	154
7.13 Problems	154
<b>8 Variable selection in regression</b>	<b>159</b>
8.1 Chapter outline	159
8.2 When are variable selection methods used?	159
8.3 Data example	160
8.4 Criteria for variable selection	163
8.5 A general $F$ test	166
8.6 Stepwise regression	168
8.7 Subset regression	173
8.8 Discussion of computer programs	177
8.9 Discussion of strategies	180
8.10 What to watch out for	183
8.11 Summary	186
8.12 Problems	186
<b>9 Special regression topics</b>	<b>189</b>
9.1 Chapter outline	189
9.2 Missing values in regression analysis	189
9.3 Dummy variables	197
9.4 Constraints on parameters	206
9.5 Regression analysis with multicollinearity	209
9.6 Ridge regression	211
9.7 Summary	215
9.8 Problems	215
<b>III Multivariate Analysis</b>	<b>221</b>
<b>10 Canonical correlation analysis</b>	<b>223</b>
10.1 Chapter outline	223
10.2 When is canonical correlation analysis used?	223
10.3 Data example	224
10.4 Basic concepts of canonical correlation	225
10.5 Other topics in canonical correlation	230

10.6	Discussion of computer programs	233
10.7	What to watch out for	235
10.8	Summary	236
10.9	Problems	236
<b>11</b>	<b>Discriminant analysis</b>	<b>239</b>
11.1	Chapter outline	239
11.2	When is discriminant analysis used?	240
11.3	Data example	241
11.4	Basic concepts of classification	242
11.5	Theoretical background	249
11.6	Interpretation	251
11.7	Adjusting the dividing point	255
11.8	How good is the discrimination?	258
11.9	Testing variable contributions	260
11.10	Variable selection	262
11.11	Discussion of computer programs	262
11.12	What to watch out for	263
11.13	Summary	266
11.14	Problems	266
<b>12</b>	<b>Logistic regression</b>	<b>269</b>
12.1	Chapter outline	269
12.2	When is logistic regression used?	269
12.3	Data example	270
12.4	Basic concepts of logistic regression	272
12.5	Interpretation: categorical variables	273
12.6	Interpretation: continuous variables	275
12.7	Interpretation: interactions	277
12.8	Refining and evaluating logistic regression	285
12.9	Nominal and ordinal logistic regression	298
12.10	Applications of logistic regression	305
12.11	Poisson regression	309
12.12	Discussion of computer programs	313
12.13	What to watch out for	314
12.14	Summary	317
12.15	Problems	317
<b>13</b>	<b>Regression analysis with survival data</b>	<b>323</b>
13.1	Chapter outline	323
13.2	When is survival analysis used?	324
13.3	Data examples	324

CONTENTS	xi
13.4 Survival functions	325
13.5 Common survival distributions	332
13.6 Comparing survival among groups	334
13.7 The log-linear regression model	335
13.8 The Cox regression model	338
13.9 Comparing regression models	349
13.10 Discussion of computer programs	352
13.11 What to watch out for	354
13.12 Summary	355
13.13 Problems	355
<b>14 Principal components analysis</b>	<b>357</b>
14.1 Chapter outline	357
14.2 When is principal components analysis used?	357
14.3 Data example	358
14.4 Basic concepts	359
14.5 Interpretation	363
14.6 Other uses	371
14.7 Discussion of computer programs	373
14.8 What to watch out for	374
14.9 Summary	376
14.10 Problems	376
<b>15 Factor analysis</b>	<b>379</b>
15.1 Chapter outline	379
15.2 When is factor analysis used?	379
15.3 Data example	380
15.4 Basic concepts	381
15.5 Initial extraction: principal components	383
15.6 Initial extraction: iterated components	386
15.7 Factor rotations	390
15.8 Assigning factor scores	395
15.9 Application of factor analysis	396
15.10 Discussion of computer programs	397
15.11 What to watch out for	400
15.12 Summary	401
15.13 Problems	402
<b>16 Cluster analysis</b>	<b>405</b>
16.1 Chapter outline	405
16.2 When is cluster analysis used?	405
16.3 Data example	407

16.4	Basic concepts: initial analysis	407
16.5	Analytical clustering techniques	414
16.6	Cluster analysis for financial data set	420
16.7	Discussion of computer programs	425
16.8	What to watch out for	428
16.9	Summary	428
16.10	Problems	429
<b>17</b>	<b>Log-linear analysis</b>	<b>431</b>
17.1	Chapter outline	431
17.2	When is log-linear analysis used?	431
17.3	Data example	432
17.4	Notation and sample considerations	434
17.5	Tests and models for two-way tables	436
17.6	Example of a two-way table	440
17.7	Models for multiway tables	442
17.8	Exploratory model building	446
17.9	Assessing specific models	452
17.10	Sample size issues	453
17.11	The logit model	455
17.12	Discussion of computer programs	457
17.13	What to watch out for	458
17.14	Summary	460
17.15	Problems	461
<b>18</b>	<b>Correlated outcomes regression</b>	<b>463</b>
18.1	Chapter outline	463
18.2	When is correlated outcomes regression used?	463
18.3	Data examples	465
18.4	Basic concepts	468
18.5	Regression of clustered data	473
18.6	Regression of longitudinal data	481
18.7	Other analyses of correlated outcomes	485
18.8	Discussion of computer programs	486
18.9	What to watch out for	486
18.10	Summary	488
18.11	Problems	489
<b>Appendix A</b>		<b>491</b>
A.1	Data sets and how to obtain them	491
A.2	Chemical companies financial data	491
A.3	Depression study data	491

CONTENTS	xiii
A.4 Financial performance cluster analysis data	492
A.5 Lung cancer survival data	492
A.6 Lung function data	492
A.7 Parental HIV data	493
A.8 Northridge earthquake data	493
A.9 School data	494
A.10 Mice data	494
<b>References</b>	<b>495</b>
<b>Index</b>	<b>509</b>