

19. Simultaneous Consideration of Several Variables	529
19.1 Scope of Multivariate Methods	529
19.1.1 Essentials of a Multivariate Setup	530
19.1.2 Statistical Limitation on the Number of Variables	530
19.2 Dependent and Independent Sets of Variables	531
19.2.1 Dependents and Independents Both Quantitative	531
19.2.1.1 Canonical Correlation	531
19.2.1.2 Multivariate Multiple Regression	532
19.2.1.3 Path Analysis	535
19.2.2 Quantitative Dependents and Qualitative Independents: Multivariate Analysis of Variance	536
19.2.2.1 Regular MANOVA	537
19.2.2.2 MANOVA for Repeated Measures	539
19.2.3 Classification of Subjects into Known Groups: Discriminant Analysis	539
19.2.3.1 Discriminant Functions	539
19.2.3.2 Classification Rule	540
19.2.3.3 Classification Accuracy	540
19.3 Identification of Structure in the Observations	543
19.3.1 Identification of Clusters of Subjects: Cluster Analysis	543
19.3.1.1 Measures of Similarity	543
19.3.1.2 Hierarchical Agglomerative Algorithm	544
19.3.1.3 Deciding on the Number of Natural Clusters	545
19.3.2 Identification of Unobservable Underlying Factors: Factor Analysis	546
19.3.2.1 Factor Analysis	547
19.3.2.2 Steps for Factor Analysis	548
19.3.2.3 Features of a Successful Factor Analysis	549
19.3.2.4 Factor Scores	550
References	550
Exercises	551
20. Quality Considerations	555
20.1 Statistical Quality Control in Medical Care	555
20.1.1 Statistical Control of Medical Care Errors	556
20.1.1.1 Adverse Patient Outcomes	556
20.1.1.2 Monitoring Fatality	557
20.1.1.3 Limits of Tolerance	557
20.1.2 Quality of Lots	558
20.1.2.1 Lot Quality Method	558
20.1.2.2 LQAS in Health Assessment	558
20.1.3 Quality Control in a Medical Laboratory	559
20.1.3.1 Control Chart	559
20.1.3.2 Cusum Chart	560
20.1.3.3 Other Errors in a Medical Laboratory	561
20.1.3.4 Six Sigma Methodology	561
20.1.3.5 Nonstatistical Issues	561
20.2 Quality of Measurement Instruments	562
20.2.1 Validity of Instruments	562
20.2.1.1 Types of Validity	562
20.2.2 Reliability of Instruments	563
20.2.2.1 Internal Consistency	563
20.2.2.2 Cronbach Alpha	564
20.2.2.3 Test-Retest Reliability	565
20.3 Quality of Statistical Models: Robustness	566
20.3.1 Limitations of Statistical Models	566
20.3.2 Validation of the Models	567

20.3.2.1	Internal Validation	567
20.3.2.2	External Validation	568
20.3.3	Sensitivity Analysis and Uncertainty Analysis	568
20.3.3.1	Sensitivity Analysis	568
20.3.3.2	Uncertainty Analysis	569
20.3.4	Resampling	570
20.3.4.1	Bootstrapping	570
20.3.4.2	Jackknife Resampling	571
20.3.4.3	Optimistic Index	571
20.4	Quality of Data	572
20.4.1	Errors in Measurement	572
20.4.1.1	Lack of Standardization in Definitions	572
20.4.1.2	Lack of Care in Obtaining or Recording Information	572
20.4.1.3	Inability of the Observer to Secure Confidence of the Respondent	573
20.4.1.4	Bias of the Observer	573
20.4.1.5	Variable Competence of the Observers	573
20.4.2	Missing Values	573
20.4.2.1	Approaches for Missing Values	574
20.4.2.2	Handling Nonresponse	575
20.4.2.3	Imputations	576
20.4.2.4	Intention-to-Treat (ITT) Analysis	576
20.4.3	Lack of Standardization in Values	578
20.4.3.1	Standardization Methods Already Described	578
20.4.3.2	Standardization for Calculating Adjusted Rates	578
20.4.3.3	Standardized Mortality Ratio	579
References	580
Exercises	581
21. Statistical Fallacies	583
21.1	Problems with the Sample	583
21.1.1	Biased Sample	583
21.1.1.1	Survivors	584
21.1.1.2	Volunteers	584
21.1.1.3	Clinic Subjects	584
21.1.1.4	Publication Bias	585
21.1.1.5	Inadequate Specification of the Sampling Method	585
21.1.1.6	Abrupt Series	585
21.1.2	Inadequate Size of Sample	585
21.1.2.1	Size of Sample Not Adequate	585
21.1.2.2	Problems with Calculation of Sample Size	586
21.1.3	Incomparable Groups	586
21.1.3.1	Differential in Group Composition	587
21.1.3.2	Differential Compliance	588
21.1.3.3	Variable Periods of Exposure	588
21.1.3.4	Improper Denominator	589
21.1.4	Mixing of Distinct Groups	590
21.1.4.1	Effect on Regression	590
21.1.4.2	Effect on Shape of the Distribution	591
21.1.4.3	Lack of Intragroup Homogeneity	591
21.2	Inadequate Analysis	592
21.2.1	Ignoring Reality	592
21.2.1.1	Looking for Linearity	592
21.2.1.2	Overlooking Assumptions	592
21.2.1.3	Selection of Inappropriate Variables	593
21.2.1.4	Area under the Concentration Curve	593

21.2.1.5	Further Problems with Statistical Analysis.....	594
21.2.1.6	Anomalous Person-Years	594
21.2.1.7	Problems with Intention-to-Treat Analysis and Equivalence	595
21.2.2	Choice of Analysis	595
21.2.2.1	Mean or Proportion?	595
21.2.2.2	Forgetting Baseline Values.....	596
21.2.3	Misuse of Statistical Packages.....	596
21.2.3.1	Overanalysis	597
21.2.3.2	Data Dredging	597
21.2.3.3	Quantitative Analysis of Codes	597
21.2.3.4	Soft Data versus Hard Data	597
21.3	Errors in Presentation of Findings	597
21.3.1	Misuse of Percentages and Means.....	598
21.3.1.1	Misuse of Percentages.....	598
21.3.1.2	Misuse of Means.....	599
21.3.1.3	Unnecessary Decimals	599
21.3.2	Problems in Reporting	600
21.3.2.1	Incomplete Reporting.....	600
21.3.2.2	Overreporting.....	601
21.3.2.3	Selective Reporting	601
21.3.2.4	Self-Reporting versus Objective Measurement	601
21.3.2.5	Misuse of Graphs.....	601
21.4	Misinterpretation	602
21.4.1	Misuse of <i>P</i> -Values	602
21.4.1.1	Magic Threshold of 0.05	602
21.4.1.2	One-Tailed or Two-Tailed <i>P</i> -Values	603
21.4.1.3	Multiple Comparisons	603
21.4.1.4	Dramatic <i>P</i> -Values	603
21.4.1.5	<i>P</i> -Values for Nonrandom Sample.....	603
21.4.1.6	Assessment of "Normal" Condition Involving Several Parameters.....	604
21.4.1.7	Absence of Evidence Is Not Evidence of Absence	604
21.4.2	Correlation versus Cause–Effect Relationship	604
21.4.2.1	Criteria for Cause–Effect	605
21.4.2.2	Other Considerations.....	606
21.4.3	Sundry Issues	606
21.4.3.1	Diagnostic Test is Only an Additional Adjunct.....	606
21.4.3.2	Medical Significance versus Statistical Significance	606
21.4.3.3	Interpretation of Standard Error of <i>p</i>	606
21.4.3.4	Univariate Analysis but Multivariate Conclusions	607
21.4.3.5	Limitation of Relative Risk.....	607
21.4.3.6	Misinterpretation of Improvements	607
21.4.4	Final Comments	608
References	609
Exercises	610
Brief Solutions and Answers to the Selected Exercise	611
Appendix A: Statistical Software	631
Appendix B: Some Statistical Tables	637
Appendix C: Solution Illustrations Using R	643
Index	689