

# Contents

<b>1</b>	<b>Introduction</b>	1
1.1	Random Matrices and Vectors	4
1.2	Multivariate Normal Distributions	7
1.3	Distributions of Quadratic Forms	11
1.3.1	Results for General Covariance Matrices	14
1.4	Generalized Linear Models	16
1.5	Additional Exercises	18
	References	20
<b>2</b>	<b>Estimation</b>	21
2.1	Identifiability and Estimability	22
2.2	Estimation: Least Squares	28
2.3	Estimation: Best Linear Unbiased	33
2.4	Estimation: Maximum Likelihood	34
2.5	Estimation: Minimum Variance Unbiased	35
2.6	Sampling Distributions of Estimates	37
2.7	Generalized Least Squares	38
2.8	Normal Equations	43
2.9	Variance-Bias Tradeoff	44
2.9.1	Estimable Functions	47
2.10	Bayesian Estimation	48
2.10.1	Distribution Theory	52
2.11	Additional Exercises	57
	References	59
<b>3</b>	<b>Testing</b>	61
3.1	More About Models	61
3.2	Testing Models	64
3.2.1	Small Test Statistics	71
3.2.2	A Generalized Test Procedure	72

- 3.3 Testing Linear Parametric Functions . . . . . 74
  - 3.3.1 A Generalized Test Procedure . . . . . 83
  - 3.3.2 Testing an Unusual Class of Hypotheses . . . . . 86
- 3.4 Discussion . . . . . 88
- 3.5 Testing Single Degrees of Freedom in a Given Subspace . . . . . 89
- 3.6 Breaking a Sum of Squares into Independent Components . . . . . 90
  - 3.6.1 General Theory . . . . . 91
  - 3.6.2 Two-Way ANOVA . . . . . 95
- 3.7 Confidence Regions . . . . . 97
- 3.8 Tests for Generalized Least Squares Models . . . . . 98
  - 3.8.1 Conditions for Simpler Procedures . . . . . 101
- 3.9 Additional Exercises . . . . . 103
- References . . . . . 105
- 4 One-Way ANOVA . . . . . 107**
  - 4.1 Analysis of Variance . . . . . 108
  - 4.2 Estimating and Testing Contrasts . . . . . 116
  - 4.3 Additional Exercises . . . . . 120
- 5 Multiple Comparison Techniques . . . . . 123**
  - 5.1 Basic Ideas . . . . . 124
  - 5.2 Scheffé’s Method . . . . . 128
  - 5.3 Least Significant Difference Method . . . . . 133
  - 5.4 Bonferroni Method . . . . . 135
  - 5.5 Tukey’s Method . . . . . 135
  - 5.6 Multiple Range Tests: Newman–Keuls and Duncan . . . . . 137
  - 5.7 Summary . . . . . 139
    - 5.7.1 Fisher Versus Neyman–Pearson . . . . . 142
  - 5.8 Additional Exercises . . . . . 142
  - References . . . . . 143
- 6 Regression Analysis . . . . . 145**
  - 6.1 Simple Linear Regression . . . . . 146
  - 6.2 Multiple Regression . . . . . 148
    - 6.2.1 Partitioned Model . . . . . 151
    - 6.2.2 Nonparametric Regression and Generalized Additive Models . . . . . 153
  - 6.3 General Prediction Theory . . . . . 155
    - 6.3.1 Discussion . . . . . 155
    - 6.3.2 General Prediction . . . . . 156
    - 6.3.3 Best Prediction . . . . . 156
    - 6.3.4 Best Linear Prediction . . . . . 160
    - 6.3.5 Inner Products and Orthogonal Projections in General Spaces . . . . . 164
  - 6.4 Multiple Correlation . . . . . 165

- 6.4.1 Squared Predictive Correlation . . . . . 168
- 6.5 Partial Correlation Coefficients . . . . . 170
- 6.6 Best Linear Unbiased Prediction . . . . . 172
- 6.7 Testing Lack of Fit . . . . . 177
  - 6.7.1 The Traditional Test . . . . . 178
  - 6.7.2 Near Replicate Lack of Fit Tests . . . . . 180
  - 6.7.3 Partitioning Methods . . . . . 182
  - 6.7.4 Nonparametric Methods . . . . . 185
- 6.8 Polynomial Regression and One-Way ANOVA . . . . . 186
- 6.9 Additional Exercises . . . . . 191
- References . . . . . 194
  
- 7 Multifactor Analysis of Variance . . . . . 197**
  - 7.1 Balanced Two-Way ANOVA Without Interaction . . . . . 197
    - 7.1.1 Contrasts . . . . . 203
  - 7.2 Balanced Two-Way ANOVA with Interaction . . . . . 204
    - 7.2.1 Interaction Contrasts . . . . . 207
  - 7.3 Polynomial Regression and the Balanced Two-Way ANOVA . . . . . 214
  - 7.4 Two-Way ANOVA with Proportional Numbers . . . . . 217
  - 7.5 Two-Way ANOVA with Unequal Numbers: General Case . . . . . 219
    - 7.5.1 Without Interaction . . . . . 219
    - 7.5.2 Interaction . . . . . 222
    - 7.5.3 Characterizing the Interaction Space . . . . . 228
  - 7.6 Three or More Way Analyses . . . . . 230
    - 7.6.1 Balanced Analyses . . . . . 230
    - 7.6.2 Unbalanced Analyses . . . . . 232
  - 7.7 Additional Exercises . . . . . 238
  - References . . . . . 240
  
- 8 Experimental Design Models . . . . . 241**
  - 8.1 Completely Randomized Designs . . . . . 242
  - 8.2 Randomized Complete Block Designs: Usual Theory . . . . . 242
  - 8.3 Latin Square Designs . . . . . 243
  - 8.4 Factorial Treatment Structures . . . . . 247
  - 8.5 More on Factorial Treatment Structures . . . . . 250
  - 8.6 Additional Exercises . . . . . 253
  - References . . . . . 253
  
- 9 Analysis of Covariance . . . . . 255**
  - 9.1 Estimation of Fixed Effects . . . . . 256
    - 9.1.1 Generalized Least Squares . . . . . 260
  - 9.2 Estimation of Error and Tests of Hypotheses . . . . . 261
  - 9.3 Another Adjusted Model and Missing Data . . . . . 264
  - 9.4 Balanced Incomplete Block Designs . . . . . 267

9.5	Testing a Nonlinear Full Model . . . . .	276
9.6	Additional Exercises . . . . .	277
	References . . . . .	279
<b>10</b>	<b>General Gauss–Markov Models . . . . .</b>	<b>281</b>
10.1	BLUEs with an Arbitrary Covariance Matrix . . . . .	282
10.2	Geometric Aspects of Estimation . . . . .	288
10.3	Hypothesis Testing . . . . .	291
10.4	Least Squares Consistent Estimation . . . . .	296
10.5	Perfect Estimation and More . . . . .	304
	References . . . . .	311
<b>11</b>	<b>Split Plot Models . . . . .</b>	<b>313</b>
11.1	A Cluster Sampling Model . . . . .	314
11.2	Generalized Split Plot Models . . . . .	318
	11.2.1 Estimation and Testing of Estimable Functions . . . . .	322
	11.2.2 Testing Models . . . . .	326
	11.2.3 Unbalanced Subplots . . . . .	328
11.3	The Split Plot Design . . . . .	328
11.4	Identifying the Appropriate Error . . . . .	332
	11.4.1 Subsampling . . . . .	332
	11.4.2 Two-Way ANOVA with Interaction . . . . .	335
11.5	Exercise: An Unusual Split Plot Analysis . . . . .	337
	References . . . . .	338
<b>12</b>	<b>Model Diagnostics . . . . .</b>	<b>341</b>
12.1	Leverage . . . . .	344
	12.1.1 Mahalanobis Distances . . . . .	345
	12.1.2 Diagonal Elements of the Projection Operator . . . . .	347
	12.1.3 Examples . . . . .	348
12.2	Checking Normality . . . . .	354
	12.2.1 Other Applications for Normal Plots . . . . .	360
12.3	Checking Independence . . . . .	362
	12.3.1 Serial Correlation . . . . .	363
12.4	Heteroscedasticity and Lack of Fit . . . . .	369
	12.4.1 Heteroscedasticity . . . . .	369
	12.4.2 Huber–White (Robust) Sandwich Estimator . . . . .	374
	12.4.3 Lack of Fit . . . . .	377
	12.4.4 Residual Plots . . . . .	380
12.5	Updating Formulae and Predicted Residuals . . . . .	380
12.6	Outliers and Influential Observations . . . . .	384
12.7	Transformations . . . . .	388
	References . . . . .	390

- 13 Collinearity and Alternative Estimates** . . . . . 393
  - 13.1 Defining Collinearity . . . . . 394
  - 13.2 Tolerance and Variance Inflation Factors . . . . . 398
  - 13.3 Regression in Canonical Form and on Principal Components . . . . . 401
    - 13.3.1 Regression in Canonical Form . . . . . 401
    - 13.3.2 Principal Component Regression . . . . . 404
    - 13.3.3 Generalized Inverse Regression . . . . . 404
  - 13.4 Classical Ridge Regression . . . . . 405
    - 13.4.1 Ridge Applied to Principal Components . . . . . 408
  - 13.5 More on Mean Squared Error . . . . . 410
  - 13.6 Robust Estimation and Alternative Distance Measures . . . . . 410
  - 13.7 Orthogonal Regression . . . . . 413
  - References . . . . . 416
  
- 14 Variable Selection** . . . . . 419
  - 14.1 All Possible Regressions and Best Subset Regression . . . . . 421
    - 14.1.1  $R^2$  . . . . . 421
    - 14.1.2 Adjusted  $R^2$  . . . . . 422
    - 14.1.3 Mallows’s  $C_p$  . . . . . 423
    - 14.1.4 Information Criteria: AIC, BIC . . . . . 425
    - 14.1.5 Cost Complexity Pruning . . . . . 427
  - 14.2 Stepwise Regression . . . . . 428
    - 14.2.1 Traditional Forward Selection . . . . . 428
    - 14.2.2 Backward Elimination . . . . . 431
    - 14.2.3 Other Methods . . . . . 433
  - 14.3 Discussion of Traditional Variable Selection Techniques . . . . . 433
    - 14.3.1  $R^2$  . . . . . 434
    - 14.3.2 Influential Observations . . . . . 435
    - 14.3.3 Exploratory Data Analysis . . . . . 435
    - 14.3.4 Multiplicities . . . . . 436
    - 14.3.5 Predictive Models . . . . . 436
    - 14.3.6 Overfitting . . . . . 437
  - 14.4 Modern Forward Selection: Boosting, Bagging, and Random Forests . . . . . 437
    - 14.4.1 Boosting . . . . . 438
    - 14.4.2 Bagging . . . . . 441
    - 14.4.3 Random Forests . . . . . 445
  - References . . . . . 446

- Appendix A: Vector Spaces** . . . . . 447
- Appendix B: Matrix Results** . . . . . 457
  - B.1 Basic Ideas . . . . . 457
  - B.2 Eigenvalues and Related Results . . . . . 459
  - B.3 Projections . . . . . 463
  - B.4 Miscellaneous Results . . . . . 471
  - B.5 Properties of Kronecker Products and Vec Operators . . . . . 473
  - B.6 Tensors . . . . . 474
  - B.7 Exercises . . . . . 476
- Appendix C: Some Univariate Distributions** . . . . . 481
- Appendix D: Multivariate Distributions** . . . . . 485
  - D.1 Identifiability . . . . . 489
- Appendix E: Inference for One Parameter** . . . . . 491
  - E.1 Testing . . . . . 493
  - E.2 *P* Values . . . . . 495
  - E.3 Confidence Intervals . . . . . 496
  - E.4 Final Comments on Significance Testing . . . . . 497
- Appendix F: Significantly Insignificant Tests** . . . . . 499
  - F.1 Lack of Fit and Small *F* Statistics . . . . . 501
  - F.2 The Effect of Correlation and Heteroscedasticity on *F* Statistics . . . . . 503
- Appendix G: Randomization Theory Models** . . . . . 509
  - G.1 Simple Random Sampling . . . . . 509
  - G.2 Completely Randomized Designs . . . . . 511
  - G.3 Randomized Complete Block Designs . . . . . 513
- References** . . . . . 516
- Author Index** . . . . . 517
- Subject Index** . . . . . 521



<http://www.springer.com/978-3-030-32096-6>

Plane Answers to Complex Questions

The Theory of Linear Models

Christensen, R.

2020, XXII, 529 p. 33 illus., Hardcover

ISBN: 978-3-030-32096-6