

Analysis of Variance, Design and Regression

Applied statistical methods

Ronald Christensen

Department of Mathematics and Statistics

University of New Mexico

Albuquerque

New Mexico, USA



CHAPMAN & HALL

London · Weinheim · New York · Tokyo · Melbourne · Madras

Contents

Preface	xiii
1 Introduction	1
1.1 Probability	1
1.2 Random variables and expectations	4
1.2.1 Expected values and variances	6
1.2.2 Chebyshev's inequality	10
1.2.3 Covariances and correlations	10
1.2.4 Rules for expected values and variances	12
1.3 Continuous distributions	14
1.4 The binomial distribution	16
1.5 The multinomial distribution	20
1.6 Exercises	23
2 One sample	27
2.1 Example and introduction	27
2.2 Inference about μ	31
2.2.1 Confidence intervals	34
2.2.2 Hypothesis tests	36
2.3 Prediction intervals	42
2.4 Checking normality	44
2.5 Transformations	51
2.6 Inference about σ^2	55
2.7 Exercises	58
3 A general theory for testing and confidence intervals	61
3.1 Theory for confidence intervals	63
3.2 Theory for hypothesis tests	68
3.3 Validity of tests and confidence intervals	76

3.4	The relationship between confidence intervals and tests	77
3.5	Theory of prediction intervals	78
3.6	Sample size determination and power	81
3.7	Exercises	84
4	Two sample problems	86
4.1	Two correlated samples: paired comparisons	86
4.2	Two independent samples with equal variances	89
4.3	Two independent samples with unequal variances	95
4.4	Testing equality of the variances	100
4.5	Exercises	103
5	One-way analysis of variance	108
5.1	Introduction and examples	108
5.1.1	Theory	114
5.1.2	Balanced ANOVA: introductory example	119
5.1.3	Analytic and enumerative studies	122
5.2	Balanced one-way analysis of variance: theory	123
5.2.1	The analysis of variance table	127
5.3	Unbalanced analysis of variance	130
5.4	Choosing contrasts	132
5.5	Comparing models	137
5.6	The power of the analysis of variance F test	138
5.7	Exercises	140
6	Multiple comparison methods	146
6.1	Fisher's least significant difference method	148
6.2	Bonferroni adjustments	152
6.3	Studentized range methods	155
6.3.1	Tukey's honest significant difference	156
6.3.2	Newman-Keuls multiple range method	157
6.4	Scheffé's method	159
6.5	Other methods	161
6.5.1	Ott's analysis of means method	161
6.5.2	Dunnett's many-one t statistic method	162
6.5.3	Duncan's multiple range method	163
6.6	Summary of multiple comparison procedures	163
6.7	Exercises	164
7	Simple linear and polynomial regression	166
7.1	An example	166
7.2	The simple linear regression model	171
7.3	Estimation of parameters	172
7.4	The analysis of variance table	175
7.5	Inferential procedures	175
7.6	An alternative model	178
7.7	Correlation	178
7.8	Recognizing randomness: simulated data with zero correlation	181
7.9	Checking assumptions: residual analysis	187
7.10	Transformations	197

7.10.1	Box–Cox transformations	199
7.11	Polynomial regression	204
7.12	Polynomial regression and one-way ANOVA	211
7.13	Exercises	222
8	The analysis of count data	227
8.1	One binomial sample	227
8.1.1	The sign test	230
8.2	Two independent binomial samples	231
8.3	One multinomial sample	233
8.4	Two independent multinomial samples	235
8.5	Several independent multinomial samples	239
8.6	Lancaster–Irwin partitioning	242
8.7	Logistic regression	249
8.8	Exercises	253
9	Basic experimental designs	256
9.1	Completely randomized designs	258
9.2	Randomized complete block designs	259
9.3	Latin square designs	271
9.4	Discussion of experimental design	279
9.5	Exercises	279
10	Analysis of covariance	285
10.1	An example	285
10.2	Analysis of covariance in designed experiments	290
10.3	Computations and contrasts	291
10.4	Power transformations and Tukey's one degree of freedom	296
10.5	Exercises	300
11	Factorial treatment structures	303
11.1	Two factors	303
11.2	Two-way analysis of variance with replication	319
11.3	Multifactor structures	323
11.4	Extensions of Latin squares	335
11.5	Exercises	338
12	Split plots, repeated measures, random effects, and subsampling	344
12.1	The analysis of split plot designs	344
12.2	A four-factor split plot analysis	355
12.3	Multivariate analysis of variance	375
12.4	Random effects models	383
12.4.1	Subsampling	383
12.4.2	Random effects	384
12.5	Exercises	386
13	Multiple regression: introduction	391
13.1	Example of inferential procedures	391
13.2	Regression surfaces and prediction	395
13.3	Comparing regression models	397
13.4	Sequential fitting	402

13.5	Reduced models and prediction	404
13.6	Partial correlation coefficients and added variable plots	405
13.7	Collinearity	406
13.8	Exercises	408
14	Regression diagnostics and variable selection	414
14.1	Diagnostics	414
14.2	Best subset model selection methods	421
14.2.1	R^2 statistic	422
14.2.2	Adjusted R^2 statistic	423
14.2.3	Mallows's C_p statistic	424
14.2.4	A combined subset selection table	425
14.3	Stepwise model selection methods	426
14.3.1	Backwards elimination	427
14.3.2	Forward selection	428
14.3.3	Stepwise methods	429
14.4	Model selection and case deletion	429
14.5	Exercises	432
15	Multiple regression: matrix formulation	435
15.1	Random vectors	435
15.2	Matrix formulation of regression models	436
15.3	Least squares estimation of regression parameters	438
15.4	Inferential procedures	442
15.5	Residuals, standardized residuals, and leverage	445
15.6	Principal components regression	446
15.7	Weighted least squares	451
15.8	Exercises	456
16	Unbalanced multifactor analysis of variance	457
16.1	Unbalanced two-way analysis of variance	457
16.1.1	Proportional numbers	457
16.1.2	General case	458
16.2	Balanced incomplete block designs	466
16.3	Unbalanced multifactor analysis of variance	475
16.4	Youden squares	478
16.5	Matrix formulation of analysis of variance	482
16.6	Exercises	486
17	Confounding and fractional replication in 2^n factorial systems	489
17.1	Confounding	492
17.2	Fractional replication	502
17.3	Analysis of unreplicated experiments	507
17.4	More on graphical analysis	515
17.5	Augmenting designs for factors at two levels	518
17.6	Exercises	519
18	Nonlinear regression	521
18.1	Introduction and examples	521
18.2	Estimation	523
18.2.1	The Gauss–Newton algorithm	523

18.2.2 Maximum likelihood estimation	526
18.3 Statistical inference	528
18.4 Linearizable models	537
18.5 Exercises	538
Appendix A: Matrices	540
A.1 Matrix addition and subtraction	541
A.2 Scalar multiplication	541
A.3 Matrix multiplication	542
A.4 Special matrices	543
A.5 Linear dependence and rank	545
A.6 Inverse matrices	546
A.7 A list of useful properties	547
A.8 Eigenvalues and eigenvectors	548
Appendix B: Tables	550
B.1 Tables of the t distribution	551
B.2 Tables of the χ^2 distribution	553
B.3 Tables of the W' statistic	557
B.4 Tables of orthogonal polynomials	558
B.5 Tables of the Studentized range	559
B.6 The Greek alphabet	563
B.7 Tables of the F distribution	564
References	576
Author index	582
Subject index	584