

HANDBOOK of tables for PROBABILITY and

STATISTICS

SECOND EDITION

EDITOR

WILLIAM H. BEYER, Ph.D.

Professor, Department of Mathematics
University of Akron (Ohio)

Published by
THE CHEMICAL RUBBER CO.

18901 Cranwood Parkway, Cleveland, Ohio 44128

All rights reserved. This book, or any parts thereof, may not be reproduced in any form without written consent from the publisher. Printed in the United States of America.

© 1966, 1968 by THE CHEMICAL RUBBER CO.

Second printing, 1974

Library of Congress Card No. 66-17301

International Standard Book No. 0-87819-692-7

Preface

Statistics is an important component of scientific reasoning, as well as an integral part of academics, business, and technology. As viewed by the late Sir Ronald Fisher, statistics is the key technology of the present day. Practicing statisticians and scientists working in diverse fields need an authoritative reference handbook of statistical tables developed to "aid" in the investigation and solution of many of today's challenging problems. This book has been compiled and arranged to meet the needs of these users of statistics.

This Second Edition of the Handbook of tables for Probability and Statistics brings together in a logically arranged, documented, and readily usable form an extensive collection of relatively standard statistical tables. The general arrangement of the First Edition has been retained. Many of the tables have been expanded and increased in effectiveness. All tables have been corrected of all errors detected. Examples of expanded tables are:

Individual Terms of the Binomial Distribution Cumulative Terms of the Binomial Distribution Confidence Limits for Proportions Tests of Significance in 2 × 2 Contingency Tables Critical Values for Testing Outliers Critical Values of U in the Mann-Whitney Test Distribution of the Total Number-of-Runs Test Number of Combinations

Included in the expository section of the Handbook (Part I) is a completely rewritten section on descriptive statistics.

Additional tables and graphs which enhance the importance of this Second Edition are:

Summary of Significance Tests Summary of Confidence Intervals Table of Signs for Calculating Effects in Factorial Designs up to Six Factors Operating Characteristic (OC) Curves for Tests on the Mean and Standard Deviation(s) of Normal Distributions Cochran's Test for the Homogeneity of Variances Percentage Points of the Maximum F-Ratio Confidence Limits for σ Based on Mean Range Critical Values for Duncan's New Multiple Range Test Critical Values for Rank-Sum Tests for Dispersion Cumulative Sum Control Charts (CSCC) Logarithms of the Binomial Coefficients

Preparation of this enlarged Second Edition has been possible only through the participation of recognized authorities who have taken time from their busy schedules to interpret their thoughts into writing. The Editor has been fortunate indeed to secure the aid of a well coordinated and specially selected advisory board. The names of members of the advisory board are presented in the forefront of this handbook. The Editor is most grateful to them for their continued cooperation and for their invaluable contributions.

The Editor gratefully acknowledges the authors, editors, and publishers who gave permission to reproduce these tables. Reference to the sources of material used in this handbook is indicated in the acknowledgment section. It is quite possible that proper credit has not always been given. Regrets and apologies are offered to the authors of such material.

To the many users of the current edition who sent in suggestions for alterations and additions, the Editorial Staff extends a special thanks. It is hoped that those interested will continue to send in suggestions and comments to assist in the continuous improvement of the contents.

William H. Beyer April, 1968

Advisory Board

Editor and Chairman

WILLIAM H. BEYER, Ph.D.

Professor, Department of Mathematics University of Akron (Ohio)

Rolf E. Bargmann, Ph.D. Department of Statistics University of Georgia Athens, Georgia 30601

Ralph A. Bradley, Ph.D. Professor and Head, Department of Statistics The Florida State University Tallahassee, Florida 32306

Dr. Herbert A. David, Ph.D. Department of Biostatistics University of North Carolina Chapel Hill, North Carolina 27514 Spencer M. Free, Jr., Ph.D.
Section Head, Statistics Section
Smith, Kline and French Laboratories
Research and Development Division
1500 Spring Garden St.
Philadelphia, Pa. 19101

Franklin A. Graybill, Ph.D. Director, Statistical Laboratory Colorado State University Fort Collins, Colorado 80521

Boyd Harshbarger, Ph.D. Head, Department of Statistics Director, Statistical Laboratory Virginia Polytechnic Institute Blacksburg, Virginia 24061

Samuel M. Selby, Ph.D. Head, Department of Mathematics University of Akron 302 E. Buchtel Ave. Akron, Ohio 44304

Contents

Greek Alphabet

Part I-PROBABILITY AND STATISTICS	
Descriptive Statistics	
Weighted Mean (Weighted Arithmetic Mean) Geometric Mean	
Geometric Mean Harmonic Mean	
Harmonic Mean Relation Between Arithmetic Geometric	
Relation Between Arithmetic, Geometric, and Harmonic Mean	
Mode	
Median Empirical Relation Retween Moon Modian and As	
Quartiles	
Decilea	
Deciles. Percentiles	
Percentiles	4
Variance. Range	ì
Quartile Deviation (Somi Internal)	6
Quartile Deviation (Semi-Interquartile Range) Coefficient of Variation	e
Coefficient of Quartile Valeting	6
Coefficient of Quartile Variation. Standardised Variable (Standard Scores)	6
Moments Coefficient of Skewness	6
Coefficient of Skewness Coefficient of Momental Skewness	7
Coefficient of Momental Skewness. Pearson's First Coefficient of Skewness	7
Pearson's First Coefficient of Skewness Pearson's Second Coefficient of Skewness	7
Pearson's Second Coefficient of Skewness. Quartile Coefficient of Skewness.	7
Quartile Coefficient of Skewness Coefficient of Kurtosis	7
Coefficient of Kurtosis. Coefficient of Expess (Kurtosis)	7
	7
Sheppard's Corrections for Grouping Curve Fitting, Regression, and Correlation	7
Curve Fitting, Regression, and Correlation.	8
Curve Fitting. Regression and Correlation	8
and Correlation.	9
Dasic Concept for Algebra of Sets.	10
r rougonity	
	11
	11
Marginal and Conditional Probability Probability Theorems	11
	12
Random Variable. Probability Function (Discrete Case)	12
Probability Function (Discrete Case) Cumulative Distribution Function (Discrete Case)	13
Cumulative Distribution Function (Discrete Case) Probability Density (Continuous Casa)	14
Probability Density (Continuous Case) Cumulative Distribution Function (Continuous Case)	14
Cumulative Distribution Function (Continuous Case) Mathematical Expectation	14
Mathematical Expectation (Continuous Case)	14
A. Expected Value. B. Moments	15
B. Moments.	15
	15

Contents C. Generating Functions. D. Cumulant Generating Function 17 E. Characteristic Functions. 17 Multivariate Distributions. 17 B. Continuous Case. 17 Moments. 18 Marginal and Conditional Distributions 18 Probability Distributions. 19 1. Discrete Uniform Distribution 19 2. Binomial Distribution. 19 3. Geometric Distribution 20 4. Multinomial Distribution 20 5. Poisson Distribution 20 6. Hypergeometric Distribution 20 7. Negative Binomial Distribution. B. Continuous. 1. Uniform Distribution 2. Normal Distribution. 3. Gamma Distribution 4. Exponential Distribution 5. Beta Distribution Sampling Distributions.... 1. Chi-Square Distribution.... Reproductive Property of Chi-Square Distribution 2. Snedecor's F-Distribution 3. Student's t-Distribution Summary of Significance Tests... Testing for the Value of a Specified Parameter Comparison of Two Populations. Summary of Confidence Intervals... Analysis of Variance (ANOVA) Tables.... One-Way Classification Two-Way Classification With One Observation Per Cell Nested Classifications With Unequal Samples Nested Classifications With Equal Samples Fixed Model Two-Factor Factorial Experiment in a One-Way Classification Design Fixed Model Three-Factor Factorial Experiment in a Completely Randomized Design txt Latin Square.... Graeco-Latin Square Youden Square... Balanced Incomplete Block (BIB) General Linear Model.... 2. General Linear Model The Simple Regression Model Multiple Regression Model One-Way Classification Analysis of Variance Two-Way Classification (Two Factors Factorial) 2.5 Analysis of Covariance. 3. Summary of Rules for Matrix Operations Expectation Partitioning of Determinants Inverse of a Partitioned Matrix. Characteristic Roots. Differentiation. Some Additional Definitions and Rules.... 4. Principle of Minimising Quadratic Forms and Gauss-Markov Theorem..... Some Remarks on Multivariate Distributions. The Principle of Least Squares. Minimum Variance Unbiased Estimates

Contents

5. General Linear Hypothesis of Full Deal	
5. General Linear Hypothesis of Full Rank 5.1 Notation 5.2 Simple Linear Regression 5.3 Applying of M	
5.4 Multiple Linear Regression 5.5 Randomized Blocks 5.6 Ougdesig France, One-way Classification	47
5.8 Summary 5.9 Computational Procedure for Testing a Hypothesis 5.10 Regression Significance Test 5.11 Alternate Form of the	. 50
5.10 Regression Significance Test. 5.11 Alternate Form of the Distribution. 6. General Linear Model of L	. 51
5 11 Albarra Significance Test	. 51
6.1 Estimable Function and Estimability 6.2 General Linear Hodel of Less Than Full Rank 6.3 General Linear Hodel and Estimability	. 52
61 Fatimal Model of Less Than Full Rank	. 52
6.3 Constraints and Control Model of Less Than Full Rank	- 53
6.2 General Linear Hypothesis Model of Less Than Full Rank 6.3 Constraints and Conditions Simplified Computations for Multiple Regression	55
Algebraic Procedure. Algebraic Procedure for the Forward Solution of the Abbreviated Docket. Tests of Significance.	. 00
Algebraic Procedure Algebraic Procedure for the Forward Solution of the Abbreviated Doolittle Method Tests of Significance Backward Solution of the Abbreviated Doolittle Method	58
Tests of Significance	· 58
Tests of Significance Backward Solution of the Abbreviated Doolittle Method Plans for Design of the Abbreviated Doolittle Method	60
Backward Solution of the Abbreviated Doolittle Method Plans for Design of Experiments Selected Latin Squares Graeco-Latin Squares	61
Selected I at 1 2 2	62
Greece Latin Squares	64
Selected Latin Squares Graeco-Latin Squares Index to Plans of Factorial Experiments Confounded in Randomized Incomplete Blocks Index to Plans for 2 Factorial Experiments.	64
Confounded in Pactorial Experiments Confounded in Page 1	65
Confounded Designs for Other Factorial Experiments. Index to Plans for 2 ⁿ Factorials in Fractional Replication.	66
Index to Plans for 2" Factorials in Fractional Papeling	67
Index to Plans for 2 ⁿ Factorials in Fractional Replication. Index to Plans, Incomplete Block Designs.	73
Index to Plans, Incomplete Block Designs. Index to Plans, Incomplete Latin Squares. Main Effect and Incomplete Latin Squares.	
	84
	93
Function Builday from Die	106
Finite Differences Function Build-up from Differences Interpolation Newton's Forward Formula Newton's Backward Formula	110
Newton's Famous 1 D	113
Newton's Forward Formula Newton's Backward Formula Gauss' Forward Formula	113
Newton's Backward Formula Gauss' Forward Formula Gauss' Rackward Formula	113
Gauss' Forward Formula Gauss' Backward Formula Stirling's Formula	114
Gauss' Backward Formula Stirling's Formula Steffenson's Formula	114
Stirling's Formula Steffenson's Formula Bessel's Formula	114
Steffenson's Formula Bessel's Formula Everett's Formula	114
Bessel's Formula Everett's Formula Bessel's Formula	114
	114
bessel's Formula (unmodified)	114
Bessel's Formula (unmodified) Everett's Formula (unmodified) Generalized Throwback Symmetric Formula	116
	116
Interpolation Techniques Which Do Not Require the Function to be Tabulated for Equal Interval of the Argument	118
Tabulated for Equal Interval of the Argument. a) Lagrangian Polynomials.	118
A) Lagrangian Dalam	
b) Divided Differences. 1 c) Adjusted Divided Differences. 1 d) Iterative View View View View View View View Vie	19
c) Adjusted Directores. 1 d) Iterative Linear Interpolation. 1 e) Gauss' Trigonometric 1	19
d) Iterative Lives 1	19
e) Gauss' Tricer Interpolation.	20
f) Regional Diametric Interpolation Formula	22
e) Gauss' Trigonometric Interpolation Formula 1 f) Reciprocal Differences 1 Inverse Interpolation	22
	23
	24
Part II—NORMAL DISTRIBUTION	-
II.2 Tolerance Factors for Normal Distributions 12 II.3 Factors for Computing Probable Errors 13 II.4 Probability of O	5
	7
Distribution With Known Standard Deviation	
January Deviation	

Contents	xiii
 II.6 Bivariate Normal Probabilities II.7 Circular Normal Probabilities II.8 Circular Error Probabilities II.9 Charts of Upper 1%, 2.5%, and 5% Points of the Distribution of the Largest Characteristic Root. II.10 Probit Analysis 	147 151 154
	170
Part III—BINOMIAL, POISSON, HYPERGEOMETRIC, AND NEGATIVE BINOMIAL DISTRIBUTIONS III.1 Individual Terms Binarial Distributions	
 III.1 Individual Terms, Binomial Distribution III.2 Cumulative Terms, Binomial Distribution III.3 Individual Terms, Poisson Distribution III.4 Cumulative Terms, Poisson Distribution III.5 Confidence Limits for Proportions III.6 Confidence Limits for the Expected Value of a Poisson Distribution III.7 Various Functions of p and q = 1 - p III.8 Hypergeometric Distribution III.9 Negative Binomial Distribution III.10 Percentage Points of the Beta Distribution III.11 Tests of Significance in 2 × 2 Contingency Tables 	194 206 212 219 238 ≤ 240 245
Part IV—STUDENT'S t-DISTRIBUTION	. 266
IV.1 Percentage Points, Student's t-Distribution IV.2 Power Function of the t-Test. IV.3 Number of Observations for t-Test of Mean IV.4 Number of Observations for t-Test of Difference Between Two Means. IV.5 Operating Characteristic (OC) Curves for a Test on the Mean of a Normal Distribution With Unknown Standard Deviation.	284 286 288
Part V-CHI-SQUARE DISTRIBUTION	290
 V.1 Percentage Points, Chi-Square Distribution. V.2 Percentage Points, Chi-Square Over Degrees of Freedom Distribution. V.3 Number of Observations for the Comparison of a Population Variance With a Standard Value Using the Chi-Square Test. V.4 Operating Characteristic (OC) Curves for a Test on the Standard Deviation of a Normal Distribution. 	295 299
Part VI-F-DISTRIBUTION	300
VI.2 Percentage Points, F-Distribution VI.2 Power Functions of the Analysis-of-Variance Tests. VI.3 Number of Observations Required for the Comparison of Two Population Variances Using the E.T.	304 311
VI.4 Operating Characteristic (OC) Curves for a Test on the Standard Deviations of Two Normal Distributions. VI.5 Cochran's Test for the Homogeneity of Variances. VI.6 Percentage Points of the Maximum F-Ratio	320 322 325
Part VII—ORDER STATISTICS	3 2 8
VII.1 Expected Values of Order Statistics From a Standard Normal Population VII.2 Variances and Covariances of Order Statistics VII.3 Confidence Intervals for Medians. VII.4 Critical Values for Testing Outliers. VII.5 Percentile Estimates in Large Samples. VII.6 Simple Estimates in Small Samples.	330 332 338 339 346 348
FART VIII—RANGE AND STUDENTIZED RANGE	0120
VIII.1 Probability Integral of the Range. VIII.2 Percentage Points, Distribution of the Range. VIII.3 Percentage Points, Studentized Range. VIII.4 Critical Values for Duncan's New Multiple Range Test. VIII.5 Substitute t-Ratios.	351 360 361 368

Contents	
VIII.6 Substitute F-Ratio VIII.7 Analysis of Variance Based on Range VIII.8 Confidence Intervals for σ Based on Mean Range	···· 382 ··· 385
Part IX—CORRELATION COEFFICIENT	· · · · 387
IX.2 Confidence Limits for the Population Correlation Coefficient when $\rho = 0$. IX.3 The Transformation $z = \tanh^{-1} r$ for the Correlation Coefficient.	389
Part X—NON-PARAMETRIC STATISTICS X.1 Critical Values for the Size T	394
X 2 Critical Value Sign Test.	
X.4 Critical Values of U in the Wilcoxon (Mann-Whitney) Two-Sample Statistic X.5 Critical Values for the Wilcoxon (Mann-Whitney) Two-Sample Statistic X.6 Distribution of the Total Number-of-Runs Test X.7 Critical Values for the Kolmogorov-Smirnov One-Sample Statistic X.8 Critical Values for the Kolmogorov-Smirnov One-Sample Statistic	399 401 405 409
Y 18 Critical Wallis One-Way Analysis of Variance by Ponting	427
X.9 Kruskal-Wallis One-Way Analysis of Variance by Ranks. X.10 Critical Values for a Sum of Ranks Procedure for Relative Spread in Unpaired	430
X.12 Critical Values of Spearman's Rank Correlation Coefficient. X.13 Critical Values of Kendall's Rank Correlation Coefficient. Part XI. Ollah Prince Control of C	. 433
YI.1 Factors for Communication Part XI—QUALITY CONTROL	449
XI.1 Factors for Computing Control Limits. XI.2 Percentage Points of the Distribution of the Mean Deviation. XI.3 Cumulative Sum Control Charts (CSCC). A. CSCC for the Mean. B. CSCC for Sample Ranges. C. CSCC for Sample Variances. D. CSCC for Number of Defectives, np, or Fraction Defective p. E. CSCC for Number of Defects, c. F. Summary of CSCC Limits.	451 455 456 456 456 458 459
Part XII—MISCELL ANDONE	464
XII.1 Number of Permutations.	
XII 5 Penden N	472
	479
XII.7 Random Normal Numbers, $\mu = 2$, $\sigma = 1$. XII.8 Orthogonal Polynomials. XII.9 Percentage Point (1)	484
XII.8 Orthogonal Polynomial. $\mu = 0, \sigma = 2$.	
XII.8 Orthogonal Polynomials. XII.9 Percentage Points of Pearson Curves. Part XIII—MISCELLANFOUS MATHEMATICAL	499
Para VIII	504
Part XIII—MISCELLANEOUS MATHEMATICAL TABLES XIII.1 Miscellaneous Constants	518
XIII.1 Miscellaneous Constants. XIII.2 Numerical Constants. XIII.3 Radians to Degrees Missellaneous MATHEMATICAL TABLES	
XIII.3 Radians to Domesta	527
XIII.2 Numerical Constants XIII.3 Radians to Degrees, Minutes, and Seconds XIII.4 Natural Functions for Angles in Radians XIII.5 Squares Cubes	528
XIII.5 Squares Cubes on I Radians.	529
XIII.6 Exponential E.	530
XIII.7 Six Place I redictions.	532
XIII.8 Natural or Nanorius	549
XIII.8 Natural or Naperian Logarithms XIII.9 Factorials and Their Logarithms XIII.10 Reciprocals of Factorial	557
XIII 10 Positional Then Logarithms.	579
XIII 11 D. A.	587
XIII.12 Sums of Powers of Integers	589
XIII.12 Sums of Powers of Integers XIII.13 Integrals XIII.14 Gamma Function	590
	592
Index	594
	635
	637

XIV

GREEK ALPHABET

Greek	Greek	English	Greek	Greek	English
letter	name	equivalent	letter	name	equivalen
$egin{array}{lll} eta & eta & { m Bet} \\ \Gamma \gamma & { m Gas} \\ \Delta \delta & { m Dei} \\ E \epsilon & { m Eps} \\ Z \zeta & { m Zet} \\ H \eta & { m Eta} \\ \Theta \theta \theta & { m The} \\ I \epsilon & { m Kag} \\ \end{array}$	mma ita silon a ota	a b g d ë s e th i k l m	N ν Ζ ξ Ο ο Π π Ρ ρ Σ σ τ Τ υ Φ φ φ Χ χ Ψ ψ Ω ω	Nu Xi Omicron Pi Rho Sigma Tau Upsilon Phi Chi Psi Omega	n x ö p r s t u ph ch ps ö

DESCRIPTIVE STATISTICS

a) Ungrouped Data

The formulas of this section designated as a) apply to a random sample of size n, denoted by x_i , $i = 1, 2, \ldots, n$.

b) Grouped Data

The formulas of this section designated as b) apply to data grouped into a frequency distribution having class marks x_i , $i = 1, 2, \ldots, k$, and corresponding class frequencies f_i , $i = 1, 2, \ldots, k$. The total number of observations given by

$$n = \sum_{i=1}^{k} f_i$$

In the formulas that follow, c denotes the width of the class interval, x_o denotes one of the class marks taken to be the computing origin, and $u_i = \frac{x_i - x_o}{c}$. Then coded class marks are obtained by replacing the original class marks with the integers \dots , -3, -2, -1, 0, 1, 2, 3, . . . where 0 corresponds to class mark x_0 in the original scale.

Mean (Arithmetic Mean)

a)
$$\bar{x} = \frac{1}{n} \sum_{i=1}^{n} x_i = x_1 + x_2 + \dots + x_n$$

b.1) $\bar{x} = \frac{1}{n} \sum_{i=1}^{k} f_i x_i = \frac{f_1 x_1 + f_2 x_2 + \dots + f_k x_k}{n}$

If data is coded

b.2)
$$\bar{x} = x_o + c \frac{\sum_{i=1}^{k} f_i u_i}{n}$$

Weighted Mean (Weighted Arithmetic Mean)

If with each value x_i is associated a weighting factor $w_i \geq 0$, then $\sum_{i=1}^{n} w_i$ is the total weight, and

a)
$$\bar{x} = \frac{\sum_{i=1}^{n} w_i x_i}{\sum_{i=1}^{n} w_i} = \frac{w_1 x_1 + w_2 x_2 + \cdots + w_n x_n}{w_1 + w_2 + \cdots + w_n}$$

Geometric Mean

a) G.M. =
$$\sqrt[n]{x_1 \cdot x_2 \cdot \cdot \cdot x_n}$$

Probability and Statistics

In logarithmic form

log (G.M.) =
$$\frac{1}{n} \sum_{i=1}^{n} \log x_i = \frac{\log x_1 + \log x_2 + \cdots + \log x_n}{n}$$

b) G.M. =
$$\sqrt[n]{x_1^{f_1} \cdot x_2^{f_2} \cdot \cdot \cdot \cdot x_k^{f_k}}$$

In logarithmic form

$$\log (G.M.) = \frac{1}{n} \sum_{i=1}^{k} f_i \log x_i = \frac{f_1 \log x_1 + f_2 \log x_2 + \cdots + f_k \log x_k}{n}$$

Harmonic Mean

a) H.M. =
$$\frac{n}{\sum_{i=1}^{n} \frac{1}{x_i}} = \frac{n}{\frac{1}{x_1} + \frac{1}{x_2} + \cdots + \frac{1}{x_n}}$$

b) H.M. =
$$\frac{n}{\sum_{i=1}^{k} \frac{f_i}{x_i}} = \frac{n}{\frac{f_1}{x_1} + \frac{f_2}{x_2} + \cdots + \frac{f_k}{x_k}}$$

Relation Between Arithmetic, Geometric, and Harmonic Mean

H.M. \leq G.M. \leq \bar{x} , (Equality sign holds only if all sample values are identical.)

Mode

a) A mode M_o of a sample of size n is a value which occurs with greatest frequency, i.e., it is the most common value. A mode may not exist, and even if it does exist it may

b)
$$M_o = L + C \frac{\Delta_1}{\Delta_1 + \Delta_2}$$

where L is the lower class boundary of the modal class (class containing the mode),

 Δ_1 is the excess of modal frequency over frequency of next lower class, Δ_2 is the excess of modal frequency over frequency of next higher class.

Median

a) If the sample is arranged in ascending order of magnitude, then the median \mathcal{M}_d is given by the $\frac{n+1}{2}$ nd value. When n is odd, the median is the middle value of the set of ordered data; when n is even, the median is usually taken as the mean of the two middle values of the set of ordered data.

b)
$$M_d = L + c \frac{\frac{n}{2} - F_c}{f_m}$$
,

where L is lower class boundary of median class (class containing the median),

 F_c is the sum of the frequencies of all classes lower than the median class,

 f_m is the frequency of the median class.

Empirical Relation Between Mean, Median, and Mode

Mean-Mode = 3 (Mean-Median)

ANALYSIS OF VARIANCE (ANOVA) TABLES

The analysis of variance (ANOVA) table containing the sum of squares, degrees of freedom, mean square, expectations, etc., present the initial analysis in a compact form. This kind of tabular representation is customarily used to set out the results of analysis of variance calculations. Appropriate ANOVA tables for various experimental design models are presented here. In the tables, the use of "dot notation" indicates a summing over all observations in the population, i.e., when summing over a suffix, that suffix is replaced by a dot. Small letters refer to observations, whereas capital letters refer to observation totals.

ANALYSIS OF VARIANCE AND EXPECTED MEAN SQUARES FOR THE ONE-WAY CLASSIFICATION

Model: $y_{ij} = \mu + \alpha_i + \epsilon_{ij}$ $(i = 1, 2, ..., k; j = 1, 2, ..., n_i)$

Source of Variation	Degrees of Freedom	Sum of Squares	Mean Square	Test Statistic
Between groups	k - 1	$S_1 = \sum_i n_i (\hat{g}_i - g_i)^2 = \sum_i \left(\frac{Y_i^2}{n_i} \right) = \frac{Y_i^2}{n_i}$	$s_1^2 = \frac{S_1}{k-1}$	$F = \frac{s_1^2}{s_1^2}$
Within groups	n - k	$S_{*} = \sum_{i} \sum_{j} (g_{ij} - g_{i.})^{2} = \sum_{i} \sum_{j} y_{ij}^{2} = \sum_{i} \left(\frac{Y_{i.}^{2}}{n_{i}} \right)$	$s_e^2 = \frac{S_e}{n_e - k}$	
Total	n 1	$S = \sum_{i} \sum_{j} (y_{ij} - \bar{y}_{})^{2} = \sum_{i} \sum_{j} y_{ij}^{2} - \frac{Y_{}^{2}}{n_{.}}$		

Source of	Degrees of	Mean	Expe	cted Mean Square for
Variation	Freedom	Square	Fixed Model	Random Model
Between groups Within groups Total	k-1 $n-k$ $n-1$	81 82	$\sigma^2 + \frac{\sum_{i} n_{i} \alpha_i^2}{k-1}$	$\sigma^{2} + \frac{1}{k-1} \left(n - \frac{\sum_{i} n_{i}^{2}}{n_{i}} \right) \sigma_{ii}^{2}$ σ^{2}

$$Y_{i,} = \sum_{j} y_{ij}; \quad Y_{..} = \sum_{i'} \sum_{j} y_{ij}; \quad g_{i,} = \frac{1}{n_{i}} \sum_{j} y_{ij} = \frac{1}{n_{i}} Y_{i,};$$

$$n_{.} = \sum_{i} n_{i}; \quad g_{..} = \frac{1}{n_{i}} \sum_{i} \sum_{j} y_{ij} = \frac{Y_{..}}{n_{i}}$$

ANALYSIS OF VARIANCE AND EXPECTED MEAN SQUARES FOR THE TWO-WAY CLASSIFICATION WITH ONE OBSERVATION PER CELL

Model: $y_{ij} = \mu + \alpha_i + \beta_i + \epsilon_{ii}$ (i = 1, 2)

Source of Variation	Degrees of Freedom	Sum of Squares	Mean Square	Test Statistic
Column effects	c - 1	$SSC = \frac{\sum_{i} Y_{i}^{2}}{r} - \frac{Y^{2}}{cr}$	$s_1^2 = \frac{SSC}{c-1}$	83
Row effects Error	r-1 $(c-1)(r-1)$	$SSR = \frac{\sum_{j} Y_{-j}^{2}}{c} - \frac{Y_{-j}^{2}}{cr}$ $SSE = SST - SSC - SSR$	$s_{z}^{2} = \frac{SSR}{r - 1}$ $s_{z}^{2} = \frac{SSE}{r - 1}$	83 82
[otal	cr - 1	$SST = \sum_{i} \sum_{j} y_{ij}^2 - \frac{Y_{}^2}{cr}$	$\frac{s_{i}-c_{i}-c_{i}}{(c-1)(r-1)}$	es til storder i menne det til sy fotolig lindde de en y til til se skeler fotolig

Source of Variation	Degrees of	Mean	Expe	cted Mean Square for	
v acincion	Freedom	Square	Fixed Model	Mixed Model (at)	Randon Model
Column effects	c - 1	82	$\sigma^2 + r \left(\frac{\sum_i \alpha_i^2}{c-1} \right)$	$\sigma^2 + r \left(\frac{\sum_i \alpha_i^2}{c - 1} \right)$	83 d 188
Row effects Error Fotal	r-1 $(c-1)(r-1)$ $cr-1$	s; s;	$\frac{\sigma^2 + c\left(\frac{\sum_j \beta_j^2}{r-1}\right)}{\sigma^2}$	$\sigma^2 + c\sigma_{\beta}^2$ σ^2	o² + co;

ANALYSIS OF VARIANCE AND EXPECTED MEAN SQUARES FOR NESTED CLASSIFICATIONS WITH UNEQUAL SAMPLES

Model: y _{iju}	= \mu + \alpha_i + \delta	$(i=1,2,\ldots,k;j=1,$		
Source of Variation	Degrees of Freedom	1	Mean Square	Expected Mean
Between main groups	k — 1	$S_1 = \sum_{i} \frac{Y_{i}^2}{n_{i}} - \frac{Y_{i}^2}{n_{i}}$	$s_1^2 = \frac{S_1}{k-1}$	~-1
groups (experimental	·	$S_{i} = \sum_{i} \sum_{j} \frac{Y_{ij.}^{s}}{n_{ij}} = \sum_{i} \frac{Y_{i}^{s}}{n_{i.}}$	•	$\sigma^2 + \frac{\sum_{i} \sum_{j} n_{ij} \delta_{ij}^2}{\sum_{i} n_{i} - \dot{x}}$
Within sub- groups (sam- pling error)	$n_{\cdot\cdot} - \sum_{i} n_{i}$	$S_{\nu} = \sum_{i} \sum_{j} \sum_{u} y_{iju}^{z} - \sum_{i} \sum_{j} \frac{Y_{ij}^{z}}{n_{ij}}$	$a_s^2 = \frac{S_3}{n_{} - \sum_i n_i}$	o‡
l'otai	n 1	$S = \sum_{i} \sum_{j} \sum_{u} y_{iju}^{z} - \frac{Y_{}^{z}}{n_{}}$	•	

Source of Variation	Degrees of Freedom	Mean			
,	1	Square	Mixed Model (α)	Mixed Model (8)	Random Model
Between main groups	k - 1	s 13	$\frac{\sum_{\sigma^2 + b\sigma_0^2 + \frac{i}{k-1}} n_{i,\alpha_i^2}}{\sigma^2 + a\sigma_0^2}$	$\sigma^2 + c\delta_a^2$	$\sigma^2 + b\sigma_0^2 + c\sigma_\alpha^2$
Experimental error	$\sum_{i} n_{i} - k$	· 4	$\sigma^2 + \alpha \sigma_i^2$	$\sigma^{4} + \frac{\sum_{i} \sum_{j} n_{ij} \delta_{ij}^{2}}{\sum_{i} n_{i} - k}$	e ² + ae;
	$n_{ii} - \sum_{\ell} n_{\ell}$	e;	ر مع ا	₹ p2	_p a
Total	8., -1				

where

$$\begin{cases} a = \frac{n_{..} - \sum_{i} \frac{n_{ij}^2}{n_{i.}}}{\sum_{i} n_{i} - k} \\ b = \frac{\sum_{i} \frac{\sum_{j} n_{ij}^2}{n_{i.}} - \frac{\sum_{i} \sum_{j} n_{ij}^2}{n_{..}}}{k - 1} \\ c = \frac{n_{..} - \frac{\sum_{i} n_{i.}^2}{n_{..}}}{k - 1} \end{cases}$$

ANALYSIS OF VARIANCE AND EXPECTED MEAN SQUARES FOR NESTED CLASSIFICATIONS WITH EQUAL SAMPLES

	Domes	$+ \delta_{ij} + \epsilon_{iju} \qquad (i = 1, 2, \ldots, k; j = 1,$	1, 2, , n; u	$=1,2,\ldots,r)$
Source of Variation	Degrees of Freedom	Sum of Saus ros	Mean Square	Expected Mean Square for Fixed Model (a, 5)
Between main groups	1	$S_1 = \sum_{i} \frac{Y_{i}^s}{nr} - \frac{Y_{i}^s}{knr}$	$s_1^2 - \frac{S_1}{k-1}$	$\sigma^2 + nr \frac{\sum_{i} \alpha_i^2}{k-1}$
Experimental error		$S_{3} = \frac{\sum_{i} \sum_{j} Y_{ij}^{3}}{r} - \frac{\sum_{i} Y_{i}^{3}}{nr}$	$s_1^2 = \frac{S_1}{k(n-1)}$	$\sigma^{2} + r \frac{\sum_{i} \sum_{j} \delta_{ij}^{2}}{k(n-1)}$
Sampling error	kn(r-1)	$S_r = \sum_{i} \sum_{j} \sum_{u} y_{iju}^2 = \frac{\sum_{i} \sum_{j} Y_{ij}^2}{r}$	$s_s^1 = \frac{S_s}{kn(r-1)}$	3
	1	$S = \sum_{i} \sum_{j} \sum_{w} y_{ijw}^{2} - \frac{Y^{2}}{knr}$		

Source of Variation	Degrees of Freedom	Square	Expected Mean Square for			
Freedom	r rectom		Mixed Model (α)	Mixed Model (8)	Random Model	
Between main groups	k - 1	e <mark>t</mark>	$\sigma^2 + r\sigma_i^2 + nr\left(\frac{\sum \alpha_i^4}{k-1}\right)$	1	$e^2 + re_4^2 + nre_4^2$	
Experimental error Sampling error	k(n-1)	i	$\sigma^2 + r\sigma_i^2$	$\sigma^2 + \frac{r \sum_i \sum_j s_{ij}^2}{k(n-1)}$	o² + rơj	
otal	$\frac{kn(r-1)}{knr-1}$	8,	σ3	g-1	gr ²	

where

$$a = b = 1$$

ANALYSIS OF VARIANCE AND EXPECTED MEAN SQUARES FOR A FIXED MODEL TWO-FACTOR FACTORIAL EXPERIMENT IN A ONE-WAY CLASSIFICATION DESIGN

Model: $y_{iju} = \mu + \alpha_i + \beta_j + (\alpha\beta)_{ij} + \alpha_{iju}$ (i = 1, 2, ..., c; j = 1, 2, ..., r; u = 1, 2, ..., n)

Source of Variation	Degrees of Freedom	Sum of Squares	Mean Square	Expected Mean Square for Fixed
Treatment combination	cr — 1	SSTr	$s_0^2 = \frac{SSTr}{cr - 1}$	Model $[a, \beta, (\alpha\beta)]$ $\sum_{j=1}^{d} \sum_{j=1}^{r} (\mu_{ij} - \mu_{ij})$ $\sigma^{2} + n \frac{i}{\sigma r - 1}$
Factor A	e - 1	884	$s_1^2 = \frac{38A}{c-1}$	$e^2 + rn \frac{\sum_{i=1}^{n} e_i^2}{c-1}$
Factor B	r - 1	SSB	$a_2^2 = \frac{88B}{r-1}$	$\sigma^2 + cn \frac{\sum_{j}^{r} \beta_j^2}{r-1}$
nteraction	(c-1)(r-1)	88AB = 88Tr - 88A - 88B	si =	$\sum_{i=1}^{n}\sum_{j}^{r}\left(\alpha\beta\right)_{i,j}^{2}$
	cr(n-1)	85W = 88T - 88T ₇	$c_{r}^{2} = \frac{SSW}{\sqrt{\frac{1}{2}}}$	$\frac{1}{(c-1)(r-1)}$
otal	crn - 1	58T	<i>α</i> (η − 1)	-

where

$$SST_{r} = \frac{\sum_{i}^{s} \sum_{j}^{r} Y_{i,i}^{s}}{s} - \frac{Y_{i}^{s}}{crn} \qquad SSA = \frac{\sum_{i}^{s} Y_{i,i}^{s}}{rn} - \frac{Y_{i}^{s}}{crn}$$

$$SSB = \frac{\sum_{j}^{r} Y_{i,i}^{s}}{cn} - \frac{Y_{i}^{s}}{crn} \qquad SST = \sum_{i}^{s} \sum_{j}^{r} \sum_{n}^{s} Y_{i,jn}^{s} - \frac{Y_{i}^{s}}{crn}$$

Source of Variation		Expected Mean Square for		
-		Random Model	Mixed Model (a)	
Pactor A	$a_i^2 = \frac{88A}{c-1}$	$\sigma^2 + n\sigma_{a0}^2 + rn\sigma_a^2$	$\sigma^2 + n\sigma_{\alpha\beta_c}^2 + rn \frac{\sum_i \alpha_i^2}{c - 1}$	
Factor #	$a_1^2 = \frac{BBB}{r-1}$	$\sigma^2 + n\sigma_{\alpha\beta}^2 + cn\sigma_{\beta}^2$	$c-1$ $\sigma^2 + cn\sigma_\beta^2$	
Interaction	$s_s^2 = \frac{88AB}{(c-1)(r-1)}$	$\sigma^2 + n\sigma_{\alpha\beta}^2$	r ² + no ²	
Within (error)	$s_n^2 = \frac{88W}{\operatorname{cr}(n-1)}$	<u>د</u>	git.	
Total .	$s_{s}^{2} = \frac{88T}{670 - 1}$			

ANALYSIS OF VARIANCE AND EXPECTED MEAN SQUARES FOR A THREE-FACTOR FACTORIAL EXPERIMENT IN A COMPLETELY RANDOMIZED DESIGN

Model: $y_{ijbu} = \mu + \alpha_i + \beta_j + \gamma_b + (\alpha\beta)_{ij} + (\alpha\gamma)_{ik} + (\beta\gamma)_{jk} + (\alpha\beta\gamma)_{ijk} + \alpha_{ijbu}$ $(i = 1, 2, \ldots, c; j = 1, 2, \ldots, r; k = 1, 2, \ldots, l; u = 1, 2, \ldots, n)$

Source of Variation	Degrees of Freedom	Sum of Squares	Mean Square	l; u = 1, 2,, n) Expected Mean Square for Fixed Model
Factor A	c - 1	88A	£1	$\sigma^2 + rin \frac{\sum_{i} \alpha_i^2}{c - 1}$
Factor B	r-1	SSB	eį.	$\sigma^2 + c \ln \frac{j}{r-1}$
Factor C	l-1°	SSC	e;	$\sigma^2 + cm \frac{\sum_{k} \gamma_k^2}{l-1}$
Interaction A × B	(c-1)(r-1)	SSAB	e;	$\sigma^2 + \ln \frac{\sum_i \sum_j (\alpha \beta)_{ij}^3}{(c-1)(r-1)}$
Interaction A × C	(c-1)(l-1)	88AC	e;	$\sigma^2 + r_B \frac{\sum_i \sum_k (\alpha_i \beta)_{ik}^2}{(c-1)(l-1)}$
Interaction B × C	(r-1)(l-1)	88BC	<i>z</i> ;	$\sigma^2 + cn \frac{\sum\limits_{j} \sum\limits_{k} (\beta \gamma)_{ik}^2}{(r-1)(l-1)}$
nteraction 1 × B × C Vithin (error)	(c-1)(r-1)(l-1) $crl(n-1)$	88ABC	a ² 7	$e^{2} + n \frac{\sum_{i} \sum_{j} \sum_{k} (\alpha \beta \gamma)_{ijk}^{k}}{(c-1)(r-1)(l-1)}$
otal	arin - 1	SSE SST	• ;	(c 1)(1-1)

where

$$SST = \sum_{i} \sum_{j} \sum_{k} \sum_{u} y_{ijku}^{2} - \frac{Y^{2}}{crin}$$

$$SSA = \frac{\sum_{i} Y_{i...}^{2}}{rin} - \frac{Y^{4}}{crin}$$

$$SSB = \frac{\sum_{j} Y_{ij...}^{2}}{cin} - \frac{Y^{2}}{crin}$$

$$SSC = \frac{\sum_{k} Y_{i.k.}^{2}}{crn} - \frac{Y^{2}}{crin}$$

$$SSTr(ABC) = \frac{\sum_{i} \sum_{j} \sum_{k} Y_{ijk...}^{2}}{n} - \frac{Y^{2}}{crin}$$

$$SSTr(AB) = \frac{\sum_{i} \sum_{j} Y_{ij...}^{2}}{in} - \frac{Y^{2}}{crin}$$

$$SSTr(AC) = \frac{\sum_{i} \sum_{k} Y_{i,k}^{2}}{rn} - \frac{Y_{i,k}^{2}}{crin}$$
$$\sum_{i} \sum_{k} Y_{i,k}^{2} - \frac{Y_{i,k}^{2}}{rn}$$
$$SSTr(BC) = \frac{\sum_{i} \sum_{k} Y_{i,k}^{2}}{rn} - \frac{Y_{i,k}^{2}}{rn}$$

SSAB = SSTr(AB) - SSA - SSB SSAC = SSTr(AC) - SSA - SSC SSBC = SSTr(BC) - SSB - SSC SSABC = SSTr(ABC) - SSA - SSB - SSC - SSAB - SSAC - SSBC SSE = SST - SSTr(ABC)

	355 - 551 - 5517	<u> </u>	pected Mean Square f	or the
Source of Variation	Mean Square	Random Model	Mixed Model (α)	Mixed Model (\alpha, \beta)
Factor A	$s_1^2 = \frac{SSA}{c-1}$	$\sigma^2 + n\sigma_{\text{eff}}^2 + in\sigma_{\text{eff}}^2 + rn\sigma_{\text{eff}}^2 + rln\sigma_{\text{eff}}^2$	$\sigma^{2} + n\sigma_{\alpha\beta\gamma}^{2} + ln\sigma_{\alpha\beta}^{2} + rn\sigma_{\alpha\gamma}^{2} + rln \frac{\sum_{c} \alpha_{c}^{2}}{c - 1}$	$\sigma^2 + rn\sigma_{\alpha\gamma}^2$ $\sum \alpha_i^2$
Factor B	$s_i^2 = \frac{SSB}{r-1}$	$\sigma^{3} + n\sigma_{\alpha\beta\gamma}^{2} + ln\sigma_{\alpha\beta}^{2} + cn\sigma_{\beta\gamma}^{2} + cln\sigma_{\beta}^{2}$	$\sigma^{2} + cn\sigma_{\beta\gamma}^{2} + cln\sigma_{\beta}^{2}$ $\sigma^{2} + cn\sigma_{\beta\gamma}^{2} + crn\sigma_{\gamma}^{2}$ $\sigma^{2} + cn\sigma_{\beta\gamma}^{2} + crn\sigma_{\gamma}^{2}$ $\sigma^{2} + n\sigma_{\alpha\beta\gamma}^{2} + ln\sigma_{\alpha\beta}^{2}$	$\begin{vmatrix} + rin \frac{7}{c-1} \\ \sigma^2 + cn\sigma^2_{\theta\gamma} \\ \sum \beta^2_1 \end{vmatrix}$
Factor C	$e_i^2 = \frac{88C}{l-1}$	$ e^2 + ne_{np\gamma}^2 + rne_{n\gamma}^2 + cne_{\gamma}^2 + crne_{\gamma}^2$	$\sigma^2 + cn\sigma_{\beta\gamma}^2 + crn\sigma_{\gamma}^2$	$+ cln \frac{7}{r-1}$ $\sigma^{1} + crn\sigma_{\gamma}^{1}$
	$\varepsilon_4^2 = \frac{SSAB}{(c-1)(r-1)}$	$\sigma^2 + n\sigma_{\alpha\beta\dot{\gamma}}^2 + in\sigma_{\alpha\dot{\beta}}^2$	$\sigma^2 + n\sigma_{\alpha\beta\gamma}^3 + in\sigma_{\alpha\beta}^3$	$\frac{\sigma^{3} + n\sigma_{\alpha\beta\gamma}^{2}}{\sum_{i} \sum_{j} (\alpha\beta)_{ij}^{2}}$
4×c	(c-1)(t-1)	THE REPT THE WAY	+ 70 mg + 770	$\sigma^2 + m\sigma^2$
$B \times C$	$\mathbf{r}_{i}^{t} = \frac{88BC}{(r-1)(l-1)}$ $\mathbf{r}_{i}^{t} = \frac{88ABC}{(c-1)(r-1)(l-1)}$	$\sigma^3 + n\sigma_{\alpha\beta\gamma}^2 + cn\sigma_{\beta\gamma}^2$	$\sigma^2 + cn\sigma_{\beta\gamma}^2$	$\sigma^2 + cn\sigma^2_{\beta\gamma}$
XB×C	$c_1^2 = \frac{SSABC}{(c-1)(r-1)(l-1)}$	(4 ³ + 76 ³ 47)	$\sigma^1 + n\sigma^1_{\alpha\beta\gamma}$	$\sigma^2 + n\sigma_{\alpha\beta\gamma}^2$
	$r_{i}^{2} = \frac{SSE}{crl(n-1)}$		دي	g/3
Cotal s	$\frac{88T}{crin - 1}$			

ANALYSIS OF VARIANCE AND EXPECTED MEAN SQUARES FOR A ! X ! LATIN SQUARE

Model: $y_{ij(k)} = \mu + \alpha_i + \beta_j + \gamma_{(k)} + \epsilon_{ij(k)}$ (i = 1, 2, ..., t; j = 1, 2, ..., t; k = 1, 2, ..., t)Source of Degrees of Expected Variation Sum of Squares Freedom Mean Square Mean Square for Fixed Model Columns Rows Treatments |t-1||(t-1)(t-2)| SSE = SST - SSC - SSR - SSTr $|s_t^2| = \frac{.88E}{(t-1)(t-2)}$ Error Total $t^2 - 1$

Source of Variation		Expected Mean Square for			
	Mean Square	Random Model	Mixed Model (γ)	Mixed Model (α, γ)	
Columns Rows	$s_1^2 = \frac{88C}{t-1}$ $88R$	o2 + to2	e ² + lo _{ji}	$\sigma^2 + i \frac{\sum_i \alpha_i^2}{i-1}$	
	$s_1^2 = \frac{88R}{t-1}$	or + toj	$\sigma^2 + i\sigma_{\beta}^2$	$\sigma^2 + i\sigma_{\beta}^2$	
Prestments	$s_1^2 = \frac{88Tr}{t-1}$	$\sigma^2 + t\sigma_\gamma^2$	$\sigma^2 + i \frac{\sum_k \gamma_k^2}{i-1}$	$\sigma^2 + t \frac{\sum_{k} \gamma_k^2}{t-1}$	
Error	$s_t^2 = \frac{SSE}{(t-1)(t-2)}$	92	t - 1	t - 1	

ANALYSIS OF VARIANCE FOR A GRAECO-LATIN SQUARE

Model: $y_{ijuk} = \mu + \alpha_i + \beta_j + \gamma_u + \delta_k + \epsilon_{ijuk}$ $(i, j, u, k = 1, 2, \ldots, n)$

			
Source of Variation	Degrees of Freedom	Sum of Squares	Mean Square
Factor I (Rows)	n-1	$S_{1} = \frac{\sum_{i} Y_{i}^{2}}{n} - \frac{Y_{}^{2}}{n^{2}}$ $S_{2} = \frac{\sum_{j} Y_{.j}^{2}}{n} - \frac{Y_{}^{2}}{n^{2}}$	$s_1^2 = \frac{S_1}{n-1}$
Factor II (Columns)	n – 1	$S_1 = \frac{\sum_{j} Y_{.j}^2}{n} - \frac{Y_{}^2}{n^2}$	$s_1^2 = \frac{S_2}{n-1}$
Factor III (Latin Letters)		Z Y Y.	. 8.
Factor IV (Greek Letters)	n - 1	$S_{4} = \frac{1}{n-1} - \frac{1}{n^{2}}$ $S_{4} = \frac{\sum_{k} Y_{k}^{2}}{n} - \frac{Y_{k}^{2}}{n^{2}}$ $S_{7} = \text{difference}$	$s_4^2 = \frac{S_4}{n-1}$
Residual	(n-1)(n-3)	S_* = difference	$s_e^2 = \frac{S_e}{(n-1)(n-3)}$
Total		$S = \sum_{i} \sum_{j} y_{ijuk}^2 - \frac{Y_{}^2}{n^2}$	

ANALYSIS OF VARIANCE FOR A YOUDEN SQUARE

Model: $y_{iju} = \mu + \alpha_i + \beta_j + \gamma_u + \epsilon_{iju}$ (i = 1, 2, . . . , b; j = 1, 2, . . . , t(=b); u = 1, 2, 21 -411

Source of Variation	Degrees of Freedom	Sum of Squares	Mean Square
Blocks (crude)		$S_1 = \sum_{i} \frac{Y_{i.}^2}{k} - \frac{Y_{}^2}{bk}$	
Treatments (adjusted)	t-1	$S_1 = \frac{t-1}{bk^2(k-1)} \sum_{j} \left(kY_{.j.}^2 - \sum_{i \in j} Y_{i}^2 \right)$	$s_1^2 = \frac{S_2}{t-1}$
Treatments (crude)		$S_1 = \sum_{j} \frac{Y_{,j.}^s}{r} - \frac{Y_{,}^s}{tr}$	
Blocks (adjusted)	b - 1	$S_4 = \frac{b-1}{bk^2(k-1)} \sum_{i} \left(rY_{i,.}^2 - \sum_{j(i)} Y_{.j,.}^2 \right)$	$s_1^2 = \frac{S_4}{b-1}$
Factor II (γ)	k-1	$S_{1} = \frac{\sum_{u} Y_{u}^{2}}{k} - \frac{Y_{}^{2}}{bk}$	$s_1^2 = \frac{S_5}{k-1}$
Residual	bk-t-b-k+2	$S_{\bullet} = S - (S_1 + S_2 + S_4)$ = $S - (S_2 + S_4 + S_5)$	$s_{\epsilon}^{2} = \frac{S_{\epsilon}}{bk - t - b - k}$
Total	bk - 1	$S = \sum_{i} \sum_{j} y_{iju}^{*} - \frac{Y_{}^{*}}{bk}$	

(Note that $S_1 + S_2 = S_1 + S_4$)

ANALYSIS OF VARIANCE FOR BALANCED INCOMPLETE BLOCK (BIB)

Model: $y_{iju} = \mu + \alpha_i + \beta_j + \epsilon_{iju}$ $(i = 1, 2, ..., b; j = 1, 2, ..., t; u = n_{ij})$

Source of Variation	Degrees of Freedom	Sum of Squares	Mean Square
Blocks Treatments (adjusted)	b-1 $t-1$	$S_{1} = \frac{\sum_{i} Y_{i}^{2}}{k} - \frac{Y_{i}^{2}}{bk}$ $S_{2} = \frac{i-1}{bk^{2}(k-1)} \sum_{j} \left[kY_{.j.} - \sum_{i(j)} Y_{i} \right]^{2}$	$s_1^2 = \frac{S_1}{b-1}$ $s_2^3 = \frac{S_3}{t-1}$
Residual	bk-t-b+1	S. = difference	$s_i^2 = \frac{S_i}{bk - i - b + 1}$
Total	bk — 1	$S = \sum_{i} \sum_{j} y_{iju}^2 - \frac{Y^2}{bk}$	

where

t = number of treatment levels

b = number of blocks

k = number of treatment levels per block

r = number of replications of each treatment level

 λ = number of blocks in which any given pair of treatment levels appear together

bk = tr

 $r(k-1) = \lambda(t-1)$

GENERAL LINEAR MODEL

by Dr. Rolf E. Bargmann

1. NOTATION

A matrix will be denoted by bold-face capital letters, e.g., if A has m rows and ncolumns, we may often specify $A = A(m \times n)$

$$\mathbf{A} = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ & \ddots & \ddots & \ddots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}.$$

A' denotes the transpose of A.

(A)_{ij} = a_{ij} denotes the element in the *i*'th row and *j*'th column of A.

I denotes the identity matrix.

 \mathbf{D}_s denotes a diagonal matrix. The subscript indicates the terms in the diagonal.

T, U, i.e., any matrix with a tilde (~) above it will denote a lower triangular matrix.

A column vector will be denoted in general, by a lower-case bold-face letter, e.g.,

$$\mathbf{x}, \mathbf{y}, \mathbf{j}, \mathcal{G}, \xi; \mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$$

Occasionally, capital bold-face letters represent column vectors. Examples are as

T (vector of treatment totals)

B (vector of Block totals).

A lower case letter with a prime denotes a row vector, e.g.,

$$\mathbf{x}' = \begin{bmatrix} x_1 & x_2 & x_3 & \cdots & x_n \end{bmatrix}.$$

2. THE GENERAL LINEAR MODEL

2.1. The Simple Regression Model

$$y_i = \alpha + \beta x_i + e_i$$

where x_i is a fixed concomitant variable whose values are assumed to be known before an experiment is performed and is not subject to chance. Let E denote the expectation operator, var and cov the (population) variances and covariances, respectively.

 $E(e_i) = 0$, var $(e_i) = \sigma^1$, cov $(e_i, e_i) = 0$, i.e., $E(y_i) = \alpha + \beta x_i$.

$$y_1 = \alpha + \beta x_1 + e_1$$

$$y_2 = \alpha + \beta x_2 + e_2$$

$$\vdots$$

$$y_n = \alpha + \beta x_n + e_n$$

we can write, in matrix form

$$\begin{bmatrix} y_1 \\ y_2 \\ y_6 \\ \vdots \\ y_n \end{bmatrix} = \begin{bmatrix} 1 & x_1 \\ 1 & x_2 \\ 1 & x_1 \\ \vdots \\ 1 & x_n \end{bmatrix} \begin{bmatrix} \alpha \\ \beta \end{bmatrix} + \begin{bmatrix} \theta_1 \\ \theta_2 \\ \vdots \\ \theta_n \end{bmatrix}$$

$$y = A \quad \xi + \theta .$$

where A is the design matrix and can be also written in the form

$$\mathbf{A} = \begin{bmatrix} \mathbf{j} & \mathbf{x} \end{bmatrix} (n).$$
(1) (1)

The numbers in parentheses denote the order of the matrix.

j is a column vector containing all ones.

x is a column vector of all concomitant observations.

The simple regression model is frequently written in the form

$$y_i = \mu + \beta(x_i - \bar{x}) + \epsilon_i ,$$

where $\mu = \alpha + \beta \bar{x}$.

This, too, can be written in the general linear model form

where A = [j, (x - zj)].

2.2. Multiple Regression Model

$$y_i = \beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i} + \beta_3 x_{2i} + \cdots + \beta_n x_{ki} + \epsilon_i$$

Assumption: $E(e_i) = 0$,

$$var(y_i) = var(e_i) = \sigma^2,$$

$$cov(y_i, y_i) = 0.$$

If we write

$$y_{1} = \beta_{0} + \beta_{1}x_{11} + \beta_{2}x_{21} + \beta_{3}x_{21} + \cdots + \beta_{5}x_{51} + \epsilon_{1}$$

$$y_{2} = \beta_{0} + \beta_{1}x_{12} + \beta_{2}x_{22} + \beta_{2}x_{22} + \cdots + \beta_{5}x_{52} + \epsilon_{2}$$

$$\vdots$$

$$\vdots$$

$$y_{n} = \beta_{0} + \beta_{n}x_{1n} + \beta_{2}x_{2n} + \beta_{2}x_{2n} + \cdots + \beta_{5}x_{5n} + \epsilon_{n}$$

we can write

$$\begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} = \begin{bmatrix} 1 & x_{11} & x_{21} & x_{21} & \cdots & x_{h1} \\ 1 & x_{12} & x_{22} & x_{22} & \cdots & x_{h2} \\ \vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\ 1 & x_{1n} & x_{2n} & x_{kn} & \cdots & x_{kn} \end{bmatrix} \begin{bmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \\ \beta_2 \\ \vdots \\ \beta_h \end{bmatrix} + \begin{bmatrix} e_1 \\ e_2 \\ e_3 \\ \vdots \\ e_n \end{bmatrix}$$

$$y = A \qquad \qquad \xi + e,$$

where $A = [j \ X](n)$, and X denotes the matrix of all observations on all concomitant variables.

2.3. One-way Classification Analysis of Variance Model

$$y_{ij} = \mu + \tau_i + e_{ii}$$

where μ = general effect

 τ_i = treatment effects

eu = experimental error,

* treatments with effects
$$\tau_1, \tau_2, \ldots, \tau_r; j = 1, 2, 3, \ldots, n_i$$
.

$$y_{11} = \mu + \tau_1 + \epsilon_{11}$$

$$y_{12} = \mu + \tau_1 + \epsilon_{12}$$

$$y_{1n_1} = \mu + \tau_1 + \epsilon_{2n_1}$$

$$y_{21} = \mu + \tau_2 + \epsilon_{2n_1}$$

$$y_{2n_2} = \mu + \tau_2 + \epsilon_{2n_2}$$

$$y_{v_1} = \mu + \tau_v + \epsilon_{v_1}$$

$$y_{v_1} = \mu + \tau_v + \epsilon_{v_1}$$

We can again write this

$$\begin{bmatrix} y_{11} \\ y_{12} \\ \vdots \\ y_{2n_1} \\ y_{2n_1} \\ \vdots \\ y_{rn_s} \end{bmatrix} = \begin{bmatrix} 1 & 1 & 0 & 0 & \cdots & 0 & 0 \\ 1 & 1 & 0 & 0 & \cdots & 0 & 0 \\ 1 & 1 & 0 & 0 & \cdots & 0 & 0 \\ 1 & 0 & 1 & 0 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 1 & 0 & 1 & 0 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 1 & 0 & 0 & 0 & \cdots & 0 & 1 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 1 & 0 & 0 & 0 & \cdots & 0 & 1 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ y_{rn_s} \end{bmatrix} + e$$

$$y = A \qquad \qquad \xi + e.$$

The design matrix A can be written

$$\mathbf{A} = [\mathbf{j}, \mathbf{A}_r](n),$$

$$(1), (v)$$

where

$$A_{r} = \begin{pmatrix} (n_{1}) & j & 0 & 0 & \cdots & 0 \\ 0 & j & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ (n_{r}) & 0 & 0 & \cdots & j \end{pmatrix}$$

and the parameter vector can also be written

$$\xi = \begin{bmatrix} \mu \\ \tau \end{bmatrix}$$

2.4. Two-way Classification (Two Factors Factorial) Model

$$y_{ijk} = \mu + \alpha_i + \beta_j + \delta_{ij} + \epsilon_{ijk} .$$

where # = general effect

 α_i = factor A effects (usually row effects)

 β_i = factor B effects (usually column effects)

 δ_{ii} = interaction effects.

For example:

$$\begin{bmatrix} y_{11} \\ y_{11} \\ y_{12} \\ y_{12} \\ y_{21} \\ y_{21} \\ y_{21} \\ \end{bmatrix} = \begin{pmatrix} (n_{11}) \\ (n_{12}) \\ (n_{12}) \\ (n_{21}) \\ (n_{$$

2.5. Analysis of Covariance

Analysis of covariance is equivalent to analysis of variance with one or more concomitant variables added.

For simplicity, let us take one-way classification and one concomitant variable.

$$y_{ij} = \mu + \tau_i + \beta x_{ij} + e_{ij} .$$

As a vector equation, this model reads

$$y_{11} = \mu + \tau_{1} + \beta x_{11} + \epsilon_{11}$$

$$y_{12} = \mu + \tau_{1} + \beta x_{12} + \epsilon_{12}$$

$$y_{1n_{1}} = \mu + \tau_{1} + \beta x_{1n_{1}} + \epsilon_{1n_{1}}$$

$$y_{21} = \mu + \tau_{2} + \beta x_{21} + \epsilon_{21}$$

$$y_{2n_{1}} = \mu + \tau_{2} + \beta x_{2n_{1}} + \epsilon_{2n_{1}}$$

$$y_{2n_{1}} = \mu + \tau_{2} + \beta x_{2n_{1}} + \epsilon_{2n_{1}}$$

$$y_{2n_{2}} = \mu + \tau_{3} + \beta x_{2n_{1}} + \epsilon_{2n_{1}}$$

$$y_{2n_{2}} = \mu + \tau_{3} + \beta x_{2n_{1}} + \epsilon_{2n_{1}}$$

Let x1, x2, and x2 denote the concomitant observations in each group, then

$$\begin{bmatrix}
y_{11} \\
y_{12} \\
\vdots \\
y_{3n_1} \\
y_{21} \\
\vdots \\
y_{3n_n} \\
y_{31} \\
\vdots \\
y_{3n_n} \\
y_{31} \\
\vdots \\
y_{3n_n} \\
y_{41} \\
\vdots \\
y_{n_n} \\
y_{$$

With the above illustrations, it is clear that we can write a great variety of models in the general linear model form

$$y = A\xi + e$$
.

3. SUMMARY OF RULES FOR MATRIX OPERATIONS

3.1. Let E(y) = y, var $(y) = \Sigma$, a symmetric matrix containing all possible variances and covariances. Then, E(My) = My, var $(My) = M\Sigma M'$, for any conforming matrix M. E(y'M) = y'M, var $(y'M) = M'\Sigma M$.

3.2. Partitioning of Determinants

$$\begin{vmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{C} & \mathbf{D} \end{vmatrix} = |\mathbf{A}| |\mathbf{D} - \mathbf{C} \mathbf{A}^{-1} \mathbf{B}| \text{ if } \mathbf{A}^{-1} \text{ exists }.$$

$$= |\mathbf{D}| |\mathbf{A} - \mathbf{B} \mathbf{D}^{-1} \mathbf{C}| \text{ if } \mathbf{D}^{-1} \text{ exists }.$$

3.3. Inverse of a Partitioned Matrix

$$\begin{bmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{C} & \mathbf{D} \end{bmatrix}^{-1} = \begin{bmatrix} \mathbf{X} & \mathbf{Y} \\ \mathbf{Z} & \mathbf{U} \end{bmatrix}$$
where $\mathbf{X} = [\mathbf{A} - \mathbf{B}\mathbf{D}^{-1}\mathbf{C}]^{-1}$

$$\mathbf{U} = [\mathbf{D} - \mathbf{C}\mathbf{A}^{-1}\mathbf{B}]^{-1}$$

$$\mathbf{Y} = -\mathbf{A}^{-1}\mathbf{B}\mathbf{U}$$

$$\mathbf{Z} = -\mathbf{D}^{-1}\mathbf{C}\mathbf{X}$$

3.3.1. Symmetric Case

$$\begin{bmatrix} A_{11} & A_{12} \\ A'_{12} & A_{22} \end{bmatrix}^{-1} = \begin{bmatrix} A^{11} & A^{12} \\ (A^{12})' & A^{22} \end{bmatrix}$$
where $A^{11} = [A_{11} - A_{12}A_{22}^{-1}A'_{12}]^{-1}$

$$A^{22} = [A_{22} - A'_{12}A_{12}^{-1}A_{12}]^{-1}$$

$$A^{13} = -A^{11}A_{12}A_{22}^{-1}$$
or $A^{13} = -A_{11}^{-1}A_{12}A^{12}$.

Computational steps: Order the sets in such a way that A₂₂ is the smaller matrix.

- a. Obtain An-1
- b. Multiply AnArr
- c. Obtain AssAn Ais
- d. Obtain A₁₁ A₁₂A₂₂-1A'₁₂
- e. Invert the matrix in d.; thus obtain A^{11}
- f. Obtain $A^{11}A_{12}A_{21}^{-1}$ by multiplying matrices from steps c. and b.
- g. Change all signs in f.; thus obtain A12
- h. Obtain (A12)' by transposing s.
- i. Obtain A'12A12, the latter factor from g.
- j. Obtain $I = A'_{19}A^{12}$, i.e., change all signs in off-diagonal elements of the matrix in step i. and complement diagonal elements to 1.
- k. Obtain $A_{11}^{-1}[I A_{12}^{j}A^{13}]$, i.e., premultiply the matrix in j. by A_{21}^{-1} obtained in step a.

3.4. Characteristic Roots

- a. ch (AB) = ch (BA) except, possibly, for zero roots.
- b. Corollary: tr (AB) = tr (BA).
- c. If ch (A) = λ_i , ch (A⁻¹) = $1/\lambda_i$, and ch (I \pm A) = $1 \pm \lambda_i$

3.5. Differentiation

3.5.1. Definitions:

Let f be a scalar function of $x_1, x_2, \ldots x_p$.

Then $\partial f/\partial x$ denotes a column vector whose ith element is $\partial f/\partial x_i$.

Let f be a scalar function of $x_{11}, x_{12} \ldots x_{1q}, x_{21}, x_{22} \ldots x_{2q}, \ldots x_{p1}, x_{p1}, \ldots x_{pq}$. Then $\partial f/\partial X$ denotes a matrix whose (i, j) element is $\partial f/\partial x_{ij}$. Note that, in this definition, x4 denotes the element in the i'th row and j'th column of X. If there are any functional relations between the elements of X (as, for instance, in a symmetric matrix) these relations will be disregarded in the above definition. In other words, x_{ij} denotes the variable in the i'th row and j'th column of X, and x, denotes that in the j'th row and i'th column. If the two happen to be identical, a new symbol will be in order. For example, if $x_{ij} = x_{ji} =$ y_{ij} , say, $\partial f/\partial y_{ij} = \partial f/\partial x_{ij} \cdot \partial x_{ij}/\partial y_{ij} + \partial f/\partial x_{ji} \cdot \partial x_{ji}/\partial y_{ij} = \partial f/\partial x_{ij} + \partial f/\partial x_{ji} = (\partial f/\partial X)_{ij} + \partial f/\partial x_{ij} = (\partial f/\partial X)_{ij} + \partial f/\partial x_{$ (8//8X)4. Here, y4 is the symbol for that distinct variable which occurs in two places in X.

If y_1, y_2, \dots, y_p are functions of x, $\partial y'/\partial x$ denotes the row vector whose i²th element is dui/dx.

Probability and Statistics

If y_{11} , y_{12} , ... y_{16} , y_{21} , ... y_{26} , ... y_{p1} , ... y_{pe} are functions of x, $\partial Y/\partial x$ denotes the matrix whose (i,j) element is $\partial y_{ij}/\partial x$.

If each of the quantities y_1, y_2, \dots, y_q is a function of the variables x_1, x_2, \dots, x_p $\partial y'/\partial x$ denotes a matrix of order (pxq) whose (i,j) element is $\partial y_i/\partial x_i$. Note the interchange of

3.5.2. Rules:

- 1. $\partial(x'x)/\partial x = 2x$
- 2. $\frac{\partial (x'Qx)}{\partial x} = Qx + Q'x$
- 3. $\partial(x'Qx)/\partial x = 2Qx$ if Q is symmetric.
- 4. $\partial(\mathbf{a}'\mathbf{x})/\partial\mathbf{x} = \mathbf{a}$
- 5. $\frac{\partial (\mathbf{a}'\mathbf{Q}\mathbf{x})}{\partial \mathbf{x}} = \mathbf{Q}'\mathbf{a}$
- 6. ∂ tr $(AX)/\partial X = A'$
- 7. $\partial \operatorname{tr} (XA)/\partial X = A'$
- 8. $\partial \log |X|/\partial X = (X')^{-1}$, if X is square and nonsingular.
- 9. "Chain Rule No. 1": $\partial y'/\partial x = \partial z'/\partial x \cdot \partial y'/\partial z$.
- 10. $\partial(\mathbf{x}'\mathbf{A})/\partial\mathbf{x} = \mathbf{A}$
- 11. If e = b A'x, $\partial(e'e)/\partial x = \partial e'/\partial x \cdot \partial(e'e)/\partial e$ (according to rule 9), = -2A'e(by rules 10 and 1).
- 12. "Chain Rule No. 2": If the scalar z is related to a scalar z through variables $y_{ij} (i = 1, 2, \cdots p; j = 1, 2 \cdots q),$

$$\partial z/\partial x = \operatorname{tr} \left[\partial z/\partial Y \cdot \partial Y'/\partial x\right]$$

or
 $\partial z/\partial x = \operatorname{tr} \left[\partial z/\partial Y' \cdot \partial Y/\partial x\right]$

This chain rule is correct regardless of any functional relationships which may exist between the elements of Y.

3.6. Some Additional Definitions and Rules

j denotes a column vector, each element of which is 1. Hence j'A is a row vector whose elements are the column sums of A, and Aj denotes a column vector whose elements are the row sums of A. j'Aj denotes the sum of all elements in the matrix A.

I denotes the identity matrix. If the order must be stated it will be added in parentheses. Hence I(p) denotes a $(p \times p)$ identity matrix.

If a tilde $(\tilde{\ })$ is placed above a matrix, the matrix is assumed to be triangular. For definiteness, the untransposed matrix

$$\mathbf{T} = \begin{bmatrix} t_{11} & 0 & 0 & \cdots & 0 \\ t_{11} & t_{12} & 0 & \cdots & 0 \\ t_{21} & t_{22} & t_{22} & \cdots & 0 \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ t_{p1} & t_{p2} & t_{p3} & \cdots & t_{pp} \end{bmatrix}$$

is a "lower" triangular matrix, whereas the transposed matrix

$$\mathbf{T}' = \begin{bmatrix} t_{11} & t_{51} & t_{51} & \cdots & t_{p1} \\ 0 & t_{51} & t_{52} & \cdots & t_{p2} \\ 0 & 0 & t_{64} & \cdots & t_{p4} \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & t_{pn} \end{bmatrix}$$

is an "upper" triangular matrix.

45

If T is lower triangular, so is T-1.

If O is a symmetric, positive-definite matrix, we can find, uniquely, a real matrix T. such that O = TT, provided we let the diagonal elements of T be positive. This matrix and its inverse can be readily obtained from the forward Doolittle solution. If, in each cycle, we divide each element of the next-to-last row (the row which is immediately above the one beginning with unity) by the square-root of the "leading" (first) element. we obtain T' on the left and T-1 on the right-hand side.

If O is a $(p \times p)$, symmetric, positive-semidefinite matrix of rank r. the matrix \tilde{T} obtained in the above manner will have zeros to the right of the r'th column. O can then be represented as

$$\mathbf{Q} = \frac{(r)}{(p-r)} \begin{bmatrix} \mathbf{T_1} \\ \mathbf{T_2} \end{bmatrix} \begin{bmatrix} \mathbf{T_1} \\ \mathbf{T_2} \end{bmatrix} (r) \quad \mathbf{T_2} \end{bmatrix},$$

where, of course, only T1 is triangular. This is an important computational device in connection with rule 3.4.a. on characteristic roots. For, if the largest root of AB is desired, where both A and B are symmetric, but A is of low rank, we can obtain the representation

$$\mathbf{A} = \frac{(r)}{(p-r)} \begin{bmatrix} \mathbf{\tilde{T}}_1 \\ \mathbf{T}_2 \end{bmatrix} [\mathbf{\tilde{T}}_1' , \mathbf{T}_2']$$

 $\mathbf{A} = \frac{(r)}{(p-r)} \begin{bmatrix} \mathbf{\tilde{T}}_1 \\ \mathbf{T}_2 \end{bmatrix} [\mathbf{\tilde{T}}_1' \ , \ \mathbf{T}_2']$ by the forward Doolittle solution. Then, by 3.4.a., ch (AB) = ch ($\mathbf{\tilde{T}}_1'$, \mathbf{T}_2']B $\begin{bmatrix} \mathbf{\tilde{T}}_1 \\ \mathbf{T}_2 \end{bmatrix}$), and the matrix in parentheses is of small order and symmetric.

D, denotes a diagonal matrix whose non-zero elements are u1, u2 · · · u2

4. PRINCIPLE OF MINIMIZING QUADRATIC FORMS AND GAUSS MARKOV THEOREM

4.1. Some Remarks on Multivariate Distributions

In univariate situation, suppose we have a random variable x, such that

$$E(x) = \mu$$

$$var(x) = \sigma^2$$

If we want to find a random variable y, such that

$$E(y) = 0 \quad \text{and} \quad \text{var}(y) = 1 ,$$

i.e., we are to find y such that it has mean 0 and variance 1, we perform the "standardization"

$$y = \frac{x - \mu}{\sigma} \cdot$$

We also recall that, if y is normally distributed,

$$y^2 = \frac{(x - \mu)^2}{\sigma^2} = \chi^2 \text{ with } 1 \text{ d.f.}$$

In multivariate situations, we have random variables x such that

$$E(x) = y$$
 and $var(x) = \Sigma$

and we wish to find y such that

$$E(y) = 0$$
 and $var(y) = I$.

To obtain this, we will proceed as follows: Let

$$\Sigma = \Gamma \Gamma'$$
.

where f is a lower triangular matrix, which, given E, can be obtained conveniently as a by-product of the forward Doolittle analysis. Then,

$$\Sigma^{-1} = (\tilde{\Gamma}')^{-1}\tilde{\Gamma}^{-1} .$$

Now, let

$$y = f^{-1}(x - y)$$

 $E(y) = f^{-1}E(x - y) = 0$

and

$$\begin{array}{ll} \operatorname{var} \; (\mathbf{y}) \; = \; \mathbf{f}^{-1} \operatorname{var} \; (\mathbf{x} - \mathbf{y}) (\mathbf{f}^{\mathbf{v}})^{-1} \\ = \; \mathbf{f}^{-1} \operatorname{var} \; (\mathbf{x}) (\mathbf{f}^{\mathbf{v}})^{-1} \\ = \; \mathbf{f}^{-1} \boldsymbol{\Sigma} (\mathbf{f}^{\mathbf{v}})^{-1} \\ = \; \mathbf{f}^{-1} \mathbf{f}^{\mathbf{f}^{\mathbf{v}}} (\mathbf{f}^{\mathbf{v}})^{-1} = \; \mathbf{I} \; . \end{array}$$

Hence v is of the desired standard form. Then

$$y'y = (x' - y')(\Gamma^{-1})'\Gamma^{-1}(x - y)$$
$$= (x' - y')\Sigma^{-1}(x - y).$$

This is called the "Standard Quadratic Form".

Since it is equal to the sum-of-squares of p standard variables, it will be distributed as x^2 with p degrees of freedom, if x has the multivariate normal distribution.

4.2. The Principle of Least Squares

Recall that the General Linear Model is

$$y = A\xi + e.$$

On the assumption, for the time being that A is of full rank, the "least squares" approach tells us to estimate \(\xi \) in such a way that the sum of squares of errors is minimised. Then, e'e is the desired sum of squares and

$$e = y - A\xi$$

$$e' = y' - \xi'A'$$

$$e'e = (y' - \xi'A')(y - A\xi)$$

$$\frac{\partial (e'e)}{\partial \xi} = -2A'(y - A\xi)$$

Setting this equal to zero, we obtain

$$A'(y - A\hat{\xi}) = 0$$
$$(A'A)\hat{\xi} = A'y$$

These are called the Normal Equations for the estimation of ξ .

4.3. Minimum Variance Unbiased Estimates

The minimum variance, unbiased, linear estimate of \$\xi\$ is obtained by the application of a very general form of the Gauss Markov Theorem:

$$y = A\xi + e$$

$$E(e) = 0$$

$$var(y) = var(e) = \sigma^{TV}$$

where V is a matrix (square, symmetric, non-singular) of order $(n \times n)$ with known elements. That is to say that variances of y_i (regardless of i) and covariances between y_i and y are known except for an arbitrary scalar multiplier applied to all of them. Then the

Since

best linear estimate of an arbitrary linear function $\mathbf{l}'\xi$ is equal to $\mathbf{l}'\xi$ where ξ minimizes

$$e'V^{-1}e$$
 .

and

$$E(\mathbf{e}) = \mathbf{0}$$

$$var(e) = var(y) = \sigma^{2}V,$$

the standard quadratic form of e would be

$$(e' - [E(e)'])[var (e)]^{-1}(e - [E(e)]) = (e' - 0')[\sigma^2 V]^{-1}(e - 0)$$

= $\frac{1}{\sigma^2}e'V^{-1}e$.

Minimizing this expression is equivalent to minimizing

$$e'V^{-1}e$$
 .

Hence, the statement, "The best linear estimate of an arbitrary function I' is equal to I's where s is obtained by minimising the quadratic form e'V-1e", as made in the Gauss-Markov Theorem is, in fact, equivalent to the statement . . . where & is obtained by minimising the standard quadratic form due to error. The Normal Equations in this general case are $A'V^{-1}A\xi = A'V^{-1}y$.

5. GENERAL LINEAR HYPOTHESIS OF FULL RANK

In this section, we shall discuss, with illustrations, the problem of testing hypotheses about certain parameters and also derive some necessary distribution in connection with

5.1. Notation

In general, a null hypothesis will be stated as

$$C\xi = k$$
.

where $C = C(n_k \times m)$, $(n_k \le m)$, is called the hypothesis matrix and is of rank n_k ξ is an $(m \times 1)$ column vector of parameters as defined in the general linear model. **k** is a vector of n_k known elements, usually equal to 0.

ns is called degrees of freedom due to hypothesis. Actually it is the number of rows in the hypothesis matrix C. In other words, it is the number of nonredundant statements embodied in the null hypothesis.

n. is called the degrees of freedom due to error and is equal to the number of observations minus the effective number of parameters.

It is important to keep in mind that in stating a composite hypothesis, we should never make:

(1) contradicting statements such as,

$$H_0: \beta_1 = \beta_2$$
 and $\beta_1 = 2\beta_2$ simultaneously

(2) redundant statements such as,

$$H_0: \tau_1 = \tau_2$$
 and $3\tau_1 = 3\tau_2$.

5.2. Simple Linear Regression

Model

$$y_i = \mu + \beta x_i + e_i$$
parameter vector $\xi = \begin{bmatrix} \mu \\ \beta \end{bmatrix}$

EXAMPLE 1

$$H_0$$
: $\mu = 0$
Alt.: $\mu \neq 0$

General linear hypothesis

$$\begin{bmatrix} 1,0 \\ \beta \end{bmatrix} = 0$$

$$C \quad \xi = 0 , \quad n_k = 1$$

EXAMPLE 2

$$H_{\theta}: \beta = 0$$

$$Alt.: \beta \neq 0$$

$$\begin{bmatrix} 0,1 \end{bmatrix} \begin{bmatrix} \mu \\ \beta \end{bmatrix} = 0$$

$$C \quad \xi = 0 , \quad n_{h} = 1$$

EXAMPLE 3

H₀:
$$\mu = 0$$
, $\beta = 0$ simultaneously
Alt.: At least one of the μ and $\beta \neq 0$

$$\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} \mu \\ \beta \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

$$C \qquad \xi = 0$$
, $n_h = 2$

EXAMPLE 4

$$H_{0}: \mu = \beta$$

$$Alt.: \mu \neq \beta$$

$$[1,-1] \begin{bmatrix} \mu \\ \beta \end{bmatrix} = 0 , \quad n_{h} = 1$$

5.3. Analysis of Variance, One-way Classification

$$y_{ij} = \mu + \tau_i + e_{ij}$$
 $(i = 1, 2, 3 \cdot \cdot \cdot v)$
Parameter vector $\xi' = [\mu, \tau_1, \tau_2, \tau_3, \dots \tau_s]$

EXAMPLE 1

$$H_0: \tau_1 = \tau_2 = \tau_3 = \cdots = \tau_s$$

Alt.: $\tau_r \neq \tau_s$ for at least one pair

Keep in mind that we must not make redundant statements. Here we have (v-1)rows in the hypothesis matrix, i.e. $n_{A} = v - 1$

$$(v-1) \begin{bmatrix} 0 & 1 & -1 & 0 & 0 & \cdots & 0 \\ 0 & 1 & 0 & -1 & 0 & \cdots & 0 \\ 0 & 1 & 0 & 0 & -1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & 1 & 0 & 0 & 0 & \cdots & -1 \end{bmatrix} \begin{bmatrix} \mu \\ \tau_1 \\ \tau_2 \\ \tau_3 \\ \vdots \\ \vdots \\ \tau_{v-1} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ \vdots \\ \vdots \\ 0 \end{bmatrix} (v-1)$$

$$C$$

EXAMPLE 2

$$H_{0}: \tau_{1} = \tau_{2} = \tau_{3} = \cdots = \tau_{*} = 0$$

$$Alt.: At least one \tau \neq 0$$

$$\begin{pmatrix} 0 & 1 & 0 & 0 & 0 & \cdots & 0 \\ 0 & 0 & 1 & 0 & 0 & \cdots & 0 \\ 0 & 0 & 0 & 1 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \ddots & \vdots \\ 0 & 0 & 0 & 0 & 0 & \cdots & 1 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ r_{*} & \vdots &$$

Example 3. For simplicity, let us take i = 1, 2, 3, 4

 H_{\bullet} : $-\tau_1 + 2\tau_2 - \tau_1 = 0$ (Quadratic contrast of three effects) Alt.: Quadratic contrast $\neq 0$

Exercise contrast
$$\neq 0$$

$$[0 \quad -1 \quad +2 \quad -1 \quad 0]\begin{bmatrix} \mu \\ \tau_1 \\ \tau_2 \\ \tau_3 \\ \tau_4 \end{bmatrix} = 0$$

$$C \qquad \qquad \xi = 0$$

5.4. Multiple Linear Regression

$$y_i = \mu + \beta_1 x_{1i} + \beta_2 x_{2i} + \beta_2 x_{2i} + \cdots + \beta_k x_{ki} + e_i$$

parameter vector $\xi' = [\mu, \beta_1, \beta_2, \beta_3, \dots, \beta_k]$

EXAMPLE 1

$$H_{0}: \beta_{1} = 0$$

$$Alt.: \beta_{1} \neq 0$$

$$[0,1,0,0 \cdots 0] \begin{bmatrix} \mu \\ \beta_{1} \\ \beta_{2} \\ \beta_{3} \\ \vdots \\ \beta_{k} \end{bmatrix} = 0$$

$$C \qquad \xi = 0, \qquad n_{h} = 1.$$

EXAMPLE 2

$$H_{6}: \beta_{1} = \beta_{2} = \beta_{3} = \cdots = \beta_{k} = 0$$
Alt.: At least one $\beta \neq 0$

$$\begin{pmatrix} 0 & 1 & 0 & 0 & \cdots & 0 \\ 0 & 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ 0 & 0 & 0 & 0 & \cdots & 1 \\ & & (k+1) & & & \\ & & & \beta_{k} \end{pmatrix} = 0$$

$$C \qquad \qquad \xi = 0, \qquad n_{h} = k.$$

EXAMPLE 3

H₀:
$$\beta_1 = 0$$
, $\beta_2 = 0$ simultaneously
Alt.: At least one of β_1 and $\beta_2 \neq 0$

$$\begin{bmatrix} 0 & 1 & 0 & 0 & \cdots & 0 \\ 0 & 0 & 0 & 1 & \cdots & 0 \end{bmatrix} \begin{bmatrix} \mu \\ \beta_1 \\ \beta_2 \\ \vdots \\ \beta_k \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

5.5. Randomized Blocks

$$y_{ij} = \mu + \tau_i + \beta_j + \epsilon_{ij}$$

where μ is the general effect

 τ_i are the treatment effects (i = 1, 2, 3) β_i are the block effects (j = 1, 2, 3, 4) ϵ_{ij} is the experimental error

parameter vector $\xi' = [\mu, \tau_1, \tau_2, \tau_3, \beta_1, \beta_2, \beta_3, \beta_4]$

EXAMPLE 1

 H_0 : $\tau_1 = \tau_2 = \tau_3$ (all treatments effects are equal) Alt.: At least one pair $\tau_r \neq \tau_s$

$$\begin{bmatrix} 0 & 1 & -1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & -1 & 0 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} \mu \\ \tau_1 \\ \tau_2 \\ \tau_3 \\ \beta_1 \\ \beta_4 \\ \beta_4 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

EXAMPLE 2

 H_0 : $-\tau_1 + 2\tau_2 - \tau_3 = 0$ (quadratic contrast) Alt.: Quadratic contrast $\neq 0$

$$\begin{bmatrix} 0 & -1 & 2 & -1 & 0 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} \mu \\ \tau_1 \\ \tau_2 \\ \beta_1 \\ \beta_3 \\ \beta_4 \\ \beta_4 \end{bmatrix} = 0$$

$$C \qquad \qquad \xi = 0 \; ; \quad n_{\lambda} = 1 \; .$$

From the above illustrations, it can be seen that we can write a great variety of tests in the form of the General Linear Hypothesis provided that we make no redundant hypothesis statements.

5.6. Quadratic Form due to Hypothesis

So far we have discussed only the model of full rank, i.e., in the normal equations,

$$A'A\xi = A'y$$
,

(A'A) has an inverse. We shall continue to assume this model throughout this chapter. Recall the General Linear Model

$$y = A\xi + e$$

 $E(y) = A\xi$ and $var(y) = \sigma^{2}I$.

We then have the normal equations,

The estimate of E is

$$\xi = (A'A)^{-1}A'y$$

and the variable of the estimate is

*
$$(A'A)^{-1}$$
 var $(A'y)(A'A)^{-1}$
= $(A'A)^{-1}A'$ var $(y)A(A'A)^{-1}$
= $\sigma^2(A'A)^{-1}A'A(A'A)^{-1}$
= $\sigma^2(A'A)^{-1}$

This is the expression for the variance-covariance matrix of the estimates of ξ . Now suppose that we have a null hypothesis

$$H_{\bullet}$$
: $C\xi = 0$.

We have an unbiased estimate of Ct namely Ct, i.e., under the null hypothesis,

$$\begin{split} E(C_{\xi}^{\xi}) &= C_{\xi} = 0 \\ \text{var}(C_{\xi}^{\xi}) &= C \text{ var}(\xi)C' \\ &= \sigma^{2}C(A'A)^{-1}C', \\ [\text{var}(C_{\xi}^{\xi})]^{-1} &= \frac{1}{\sigma^{2}}[C(A'A)^{-1}C']^{-1} \ . \end{split}$$

Thus, under the null hypothesis, the standard quadratic form is

$$\frac{1}{\sigma^2} \xi' C' [C(A'A)^{-1}C']^{-1} C \xi .$$

The expression, $\hat{\xi}'C'[C(A'A)^{-1}C']^{-1}C\hat{\xi}$, is called the sum of squares due to hypothesis, usually denoted by SSH. If y has the multivariate normal distribution, SSH/ σ^2 is distributed as χ^2 with n_k degrees of freedom, since it is a standard quadratic form.

5.7. Sum of Squares due to Error

Recall the general linear model

$$y = A\xi + e$$
.

Let us define $\hat{\mathbf{e}} = \mathbf{y} - \mathbf{A}\hat{\mathbf{\xi}}$, the error of estimation. Then, $\sum_{i=1}^{n} \hat{\mathbf{e}}^{i} = \hat{\mathbf{e}}'\hat{\mathbf{e}}$ is called the sum of squares of errors of estimation. It is customarily denoted by SSE

SSE =
$$\delta'\delta = (y' - \xi'A')(y - A\xi)$$

= $y'y - \xi'A'y - y'A\xi + \xi'A'A\xi$
= $y'y - \xi'A'y - y'A\xi + \xi'A'y$
= $y'y - y'A\xi$.

where y'y is the sum of squares over all observations.

A'y is the column vector on the right hand side of the normal equations.

E is a column vector whose elements are the estimates of E.

In words, SSE is obtained by subtracting from the sum of squares of all observations, the scalar product of the vector of estimates of ξ and the vector on the right-hand side of the normal equations.

It should be noted that SSE can depend only on the model, and is determined once the model is stated; it is entirely independent of any hypothesis which may be stated or tested.

If y is normally distributed, SSE/σ^2 has the χ^2 distribution with n_e degrees of freedom. It is independent of any SSH.

5.8. Summary

We have the general linear model

$$y = A\xi + e$$

 $E(y) = A\xi$.

We assume that the model is of full rank, that is, A'A is non-singular and thus has an inverse. If we further assume that

$$var(y) = \sigma^2 I$$

that is, homoscedasticity plus independence, we will have the normal equations

$$(A'A)\xi = A'y$$

and we can obtain the estimate of E by

$$\dot{\xi} = (A'A)^{-1}A'v .$$

Again, if we further assume that the elements of y are normally distributed, we may test the following hypothesis:

$$H_0$$
: $C\xi = 0$
Alt.: $C\xi = \pi$ $(\neq 0)$

This hypothesis matrix has n_h rows and, if we avoid inconsistency and redundancies in the statement of the hypothesis, n_h will be the "degrees of freedom due to hypothesis".

5.9. Computational Procedure for Testing a Hypothesis

In testing a hypothesis, proceed as follows:

(1) Obtain SSH, the so called "sum of squares due to hypothesis" from the formula

$$88H = \xi'C'[C(A'A)^{-1}C']^{-1}C\xi$$
.

(2) Obtain SSE, the "sum of squares due to error" from

SSE =
$$\sum_{\mathbf{x}|\mathbf{f}} y_{\mathbf{f}^{\mathbf{x}}} - \mathbf{y}' \mathbf{A} \hat{\mathbf{\xi}}$$
.

- (3) Introduce n_e , the "degrees of freedom due to error" which equals n (sample size) minus effective number of parameters in the model.
 - (4) Then, if He is true

$$\frac{\text{SSH}/n_h}{\text{SSE}/n_s} = \mathbf{F}_{(n_h, n_s)} .$$

5.10. Regression Significance Test

Suppose that we have the general linear model

$$E(y) = A\xi - e$$

$$E(y) = A\xi \quad \text{and} \quad var(y) = \sigma^2 I$$

Under this model, we have the normal equations

$$A'A\hat{\xi} = A'y$$

$$SSE = y'y - y'A\hat{\xi}$$
where $\hat{\xi} = (A'A)^{-1}A'y$.

Now, suppose that our hypothesis is of such a nature that we can easily write the reduced

Analogously, after estimating § in the above model (reduced model) we may write

SSE (reduced) =
$$y'y - y'A\xi$$

where ξ is the estimate of ξ in the reduced model. We can then find the sum of squares due

SSE (reduced) and subtracting SSE (the original or general model), i.e.,

5.11. Alternate Form of the Distribution

SSE + SSH has the Beta distribution with parameters
$$(n_{\circ/2}, n_{h/2})$$
.

The beta tests are lower-tail tests, i.e., we reject H_0 if the value of the observed ratio is

rejection region β < constant.

Actually in the tables, percentage points of β are stated as

$$\beta(a,b)$$
,

where a = 2 (second parameter)

b = 2 (first parameter).

Hence, read those tables simply as $\beta(n_h, n_e)$

$$\frac{\text{SSE}}{\text{SSE} \text{ (reduced)}} = \beta \left(\frac{n_s}{2}, \frac{n_b}{2}\right) \qquad \text{(usual notation)}$$

$$= \beta^*(n_b, n_e) \qquad \text{(Tables for Beta percentage points)}$$

$$= I\left(\frac{n_e}{2}, \frac{n_b}{2}\right) \qquad \text{(Tables of the Incomplete Beta Function)}.$$

6. GENERAL LINEAR MODEL OF LESS THAN FULL RANK

So far we have restricted our discussion to models of full rank in the General Linear Model. In practice, many design models are not initially of this form. Models not of full rank are sometimes called singular models. If, in the General Linear Model

$$y = A\xi + e$$

Probability and Statistics

 $A'A\xi = A'y$

the rank of the design matrix A is less than m (r < m) then (A'A) would be singular and has no inverse. We must examine the system to see whether a solution exists. We wish to find functions of the ξ_i 's for which unbiased estimates exist.

6.1. Estimable Function and Estimability

Let us estimate a function $l'\xi$, i.e., find $c'y = l'\xi$ such that the expectation

$$E(c'y) = l'\xi$$
 for all ξ .

and var (c'y) = minimum, i.e., we would like to find a linear function of the y's such that

$$E(\mathbf{c}'\mathbf{y}) = \mathbf{l}'\mathbf{\xi} \ .$$

where I' is a given vector of "weights". The constraints of unbiasedness are

$$E(c'y) = 1'\xi$$

 $c'E(y) = 1'\xi$
 $c'A\xi = 1'\xi$
 $c'A = 1'$
 $c'A - 1' = 0$

Hence, we are minimizing

with normal equations

var (c'y) subject to the constraints c'A = l'

where var $(c'y) = \sigma^t c'c$.

The criterion function ₱ is then

$$\frac{\Phi = \frac{1}{2}e'e - [e'A - 1']\lambda}{\frac{\partial \Phi}{\partial e} = e - A\lambda.$$

Setting the derivative equal to zero, we obtain

$$\mathbf{A}\lambda = \mathbf{c}$$

Premultiplying by A', we have

$$A'A\lambda = A'e$$

which is equal to I under our constraints.

Hence.

$$\mathbf{A'A\lambda} = \mathbf{1}.$$

(6.1.1) and (6.1.2) are called "conjugate normal equations". If A has rank r (< m) we can always select r columns which form a "basis" and take the remaining (m-r)columns as an extension. The latter columns are linear combinations of the former.

$$y = A\xi + e$$

let us order the elements in ξ as well as the columns in A in such a way that

$$\xi' = [\xi'_1 , \xi'_1]$$
 $(r) (m-r)$

and

$$\mathbf{A} = [\mathbf{A}_1 \quad , \quad \mathbf{A}_2](r)$$

$$(r) \quad (m-r)$$

55

and that A_1 is a basis of A. The columns of A_2 must then be linear combinations of those in A_1 . We may express this fact formally by saying there exists

$$Q(r \times \overline{m-r})$$
 such that $A_1 = A_1Q$.

Suppose

$$A_1 = A_1Q$$

$$A'_1A_2 = A'_1A_1Q$$

$$Q = (A'_1A_1)^{-1}A'_1A_2$$

This is one of the ways to determine Q when A_1 and A_2 are given. Usually, however, we would try to find Q by inspection.

Now.

$$A = [A_1 , A_1](r)$$

$$(r) (m-r)$$

$$= [A_1, A_1Q] = A_1[I Q](r)$$

$$(r) (m-r)$$

(6.1.2) can be written as

$$\begin{bmatrix} \mathbf{A}_1' \\ \mathbf{Q}' \mathbf{A}_1' \end{bmatrix} [\mathbf{A}_1 , \mathbf{A}_1 \mathbf{Q}] \lambda = \begin{bmatrix} \mathbf{A}_1' \mathbf{A}_1 & \mathbf{A}_1' \mathbf{A}_1 \mathbf{Q} \\ \mathbf{Q}' \mathbf{A}_1' \mathbf{A}_1 & \mathbf{Q}' \mathbf{A}_1' \mathbf{A}_1 \mathbf{Q} \end{bmatrix} \lambda = \begin{bmatrix} \mathbf{I}_1 \\ \mathbf{I}_2 \end{bmatrix} \begin{pmatrix} r \\ (m-r) \end{pmatrix}.$$

Expanding we have

(6.1.3)
$$[A'_1A_1, A'_1A_2] = 1$$
(6.1.4)
$$[Q'A'_1A_1, Q'A'_1A_2] = 1$$

Premultiply (6.1.3) by Q' and obtain

$$[Q'A_1'A_1, Q'A_1'A_1Q]\lambda = Q'I_1.$$

For consistency of the equation system, the condition

$$l_2 = O'l_1$$

must be met.

That is to say, in the function $l'\xi,\,l'$ cannot be chosen arbitrarily but must be of the form

Only an I satisfying this relation can be used in the construction of a function which admits of a linear unbiased (and mathematically consistent) estimate.

(6.1.5) is called the condition of "estimability" of a linear function. Hence, we will call a function I'\(\xi\) estimable if I' can be written as

$$[1'_1, 1'_2]$$
.

where l'_2 is related to l'_1 in the same way as A_2 to A_1 .

We may then define that a parametric function is said to be linearly estimable if there exists a linear combination of the observations whose expected value is equal to the function, i.e., if there exists an unbiased estimate.

Now, if the function I'\xi is estimable, (6.1.3) can be written as

$$(A_1'A_1)[I,O]\lambda = I_1.$$

Notice that l_2 may be disregarded since it is determined by the relation $l_2 = Q'l_1$, hence,

$$(6.1.6) [I,Q]\lambda = (A_1'A_1)^{-1}I_1.$$

The first conjugate normal equations (6.1.1) stated

or,
$$\begin{aligned} \mathbf{A}\lambda &= \mathbf{c} \\ [\mathbf{A}_1, \mathbf{A}_2]\lambda &= \mathbf{c} \\ \mathbf{A}_1[\mathbf{I}, \mathbf{O}]\lambda &= \mathbf{c} \end{aligned}$$

Inserting (6.1.6), we obtain

$$A_1(A_1'A_1)^{-1}I_1 = c$$
.
 $c'y = \widehat{I'}\xi = I'_1(A_1'A_1)^{-1}A'_1y$,

Hence, $\delta' y = I' \xi$

which is of the same form as in the non-singular case, except that A has been replaced by its basis A_1 and in I we consider only the first r elements, i.e., l_1 .

Hence, the normal equations in the Least Squares approach, i.e.,

$$A'A\xi = A'y$$

can be used formally in the reduced statement

$$(A_1'A_1)\xi_1 = A_1'y$$
.

6.2. General Linear Hypothesis Model of Less Than Full Rank

We have the general linear model

$$y = A\xi + e$$

$$= [A_1, A_2] \begin{bmatrix} \xi_1 \\ \xi_2 \end{bmatrix} + e$$

$$= [A_1, A_1Q] \begin{bmatrix} \xi_1 \\ \xi_2 \end{bmatrix} + e$$

$$= A_1\xi_1 + A_1Q\xi_2 + e$$

$$= A_1(\xi_1 + Q\xi_2) + e$$

Hence, we may write the general linear model in the form $y = A_1 \xi^* + e$, where $\xi^* = \xi_1 + Q \xi_2$.

6.2.1. Sum of Squares Due to Error and Its Distribution

Notice that e has not been changed in this model, hence we can set up the normal equations

$$SSE = e'e = y'y - y'A_1\xi^*$$

= y'y - y'A_1\xi^*
= y'y - y'A_1(A_1'A_1)^{-1}A_1'y

Then, as before

$$\frac{\text{SSE}}{\sigma^2} = \chi^2(n-r) \quad \text{where } r \text{ is the rank of A},$$

and replaces m in the non-singular model. The "effective" number of parameters in the singular model is only r, the remaining (m-r) parameters are determined in terms of the first r by the estimability condition.

6.2.2. Sum of Squares Due to Hypothesis and Its Distribution

Suppose that we wish to test

$$H_0: \mathbf{C}\boldsymbol{\xi} = \mathbf{0} ,$$

where $C = [\mathbf{C_1}, \mathbf{C_2}]$. (r) (m-r)

Then $C\xi = 0$ implies that

$$[C_1, C_2] \begin{bmatrix} \xi_1 \\ \xi_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}.$$

Each row on the left-hand side must represent an estimable function, hence, we must have

$$C_2 = C_1 O$$
.

This is called the condition of "testability", i.e., if

$$\mathbf{C} = \begin{bmatrix} \mathbf{c}_2' \\ \mathbf{c}_1' \\ \vdots \\ \vdots \\ \mathbf{c}_{n_k}' \end{bmatrix}$$

where $(c_i'\xi)$ is an estimable function $(i = 1, 2, \ldots, n_h)$. Then the null hypothesis

$$H_0: C_1\xi_1 + C_2\xi_2 = 0$$

can be written as

$$C_1\xi_1+C_1O\xi_2=0$$

or simply $C_1\xi^*=0,$ where $\xi^*=(\xi_1+Q\xi_2).$ Hence, we can formally state that a null hypothesis

$$H_0$$
: $C\xi = 0$

is "testable" if $C\xi$ consists of n_h estimable functions, i.e., if $C_2 = C_1Q$, where $C = [C_1, C_2]$. Consequently,

$$SSH = \hat{\xi}^{*}'C_{1}'[C_{1}(A_{1}'A_{1})^{-1}C_{1}']^{-1}C_{1}\hat{\xi}^{*}$$

where $\hat{\xi}^* = (A_1'A_1)^{-1}A_1y$. As before

$$\frac{\text{SSH}}{\sigma^2} = \chi^2_{(n_k)} .$$

Again, if the null-hypothesis is true, we have the test statistic F

$$\frac{\text{MSH}/n_h}{\text{MSE}/n_e} = \mathbf{F}_{(n_h, n_e)} .$$

6.3. Constraints and Conditions

If the model is singular, of rank r < m, (m - r) constraints on the ξ_i 's (the estimates) may be arbitrarily introduced, for example:

$$\xi_{r+1} = 0, \ldots, \xi_m = 0$$

or

(6.3.2)
$$\sum_{i=1}^{m} \xi_{i} = 0, \sum_{i=1}^{m} n_{i} \xi_{i} = 0.$$

This is called reparametrizing the model. The constraining functions are fairly arbitrary, but they must not be estimable functions, otherwise the resulting model will still be singular.

In effect, this is done by deletion of the last (m-r) rows and columns of A'A and the last (m-r) elements of A'y, for constraints of the type (6.3.1), or by adding a constant to all elements of A'A, for constraints of the type (6.3.2). This has no effect on the value of estimable functions, or test statistics.

An entirely different situation prevails if we place conditions on the parameters of a model, especially on interactions. In the two-way classification model

$$E(y_{ijk}) = \mu + \alpha_i + \beta_i + \delta_{ij}$$

one usually specifies

$$\sum_{i} n_{ij} \delta_{ij} = 0 \quad \text{for all } j$$
's

and

$$\sum_{i} n_{ij} \delta_{ij} = 0 \quad \text{for all } i$$
's ,

where n_{ij} denotes the number of observations in the (i,j) cell. These are sometimes called natural constraints (they are neither natural nor constraints). They simply represent a set of conditions or assumptions on the interactions, minimizing this effect (making SSH for interaction a minimum). After introducing these conditions, one still has a singular model, which can be made nonsingular by introduction of the arbitrary constraints

$$\sum_{i} \alpha_{i} = 0, \sum_{i} \beta_{i} = 0.$$

(Note the carets, for estimates). One could introduce the different assumptions,

All
$$\alpha_i$$
's = 0
All β_i 's = 0

and would have a simple one-way classification model, quite different from the previous one. A classical example is the following. Suppose some organic substance is attacked by sulphuric acid or by sodium hydroxide.

Using, formally, the minimizing interaction conditions, one would obtain effect estimates as means of rows and columns

H₂SO₄ absent: 2 present: 3 NaOH absent: 3 present: 2

and make the ridiculous inference that sodium hydroxide, by itself, has an inhibiting effect. The correct parametric model in this case would be

μ	μ+β	
$\mu + \alpha$	$\mu + \alpha + \beta + \delta$	

i.e., interaction occurs only if both substances are present. This leads to the estimation

$$\begin{aligned}
\mu &= 0 \\
\alpha &= 6 \\
\hat{\beta} &= 4 \\
\delta &= -10
\end{aligned}$$

which is the appropriate neutralization model.

It is usually quite easy to decide whether a constraint or a condition is involved. The (model-changing) conditions are required whenever a hierarchy of effects is present (main effects, interactions, higher-order interactions), while constraints (with no effect on the model) can be introduced within the same kinds of effects (row effect estimates adding to zero, column effect estimates adding to zero). The sum of squares due to a given hypothesis is a good indicator of the situation. If it changes by the introduction of two different sets of combinations, they are conditions, and must be determined in accordance with plausibility of the physical model. If it stays the same, they are usually constraints on the estimates, and thus arbitrary, without effect on the model.

SIMPLIFIED COMPUTATIONS FOR MULTIPLE REGRESSION

by Dr. Clyde Y. Kramer

The following method is especially suited for a research problem which has several dependent variables with one set of independent variables. It allows the worker to decide which dependent variables are explained by regression with the least time and work. Regression coefficients and their variances are not usually wanted unless the regression is significant. This procedure eliminates the need of calculating these quantities when prediction is not good enough to be useful.

The main advantages of this method are:

(a) the numbers of digits to the left of the decimal points of the elements of the sums of squares and sums of products matrix are adjusted to be one or zero which permits the use of a uniform number of decimal places in the calculations;

(b) the multiple correlation coefficient and entries for the analysis of variance table for multiple regression can be found without computing the regression coefficients and the inverse of the sums of squares and sums of products matrix;

(c) the additional reduction due to any regression variable over that obtained for previous ones is obtainable for every regression variable;

(d) the research worker can fit only those regression variables that add a significant additional reduction if he so desires;

(e) time or work is not lost if one wishes to obtain the regression coefficients and their variances; and

(f) numerous checks are employed on the calculations that are required.

Algebraic Procedure

For simplicity, this method will be illustrated by considering four independent variables (x_1, x_2, x_3, x_4) , and one dependent variable, (y). First compute and record the sums of squares and sums of products in the following manner:

where

$$a_{ii} = \sum_{\alpha=1}^{n} x_{i\alpha}^{2} - \frac{\left(\sum_{\alpha=1}^{n} x_{i\alpha}\right)^{2}}{n},$$

$$a_{ij} = \sum_{\alpha=1}^{n} x_{i\alpha}x_{j\alpha} - \frac{\left(\sum_{\alpha=1}^{n} x_{i\alpha}\right)\left(\sum_{\alpha=1}^{n} x_{j\alpha}\right)}{n},$$

$$a_{iy} = \sum_{\alpha=1}^{n} x_{i\alpha}y_{\alpha} - \frac{\left(\sum_{\alpha=1}^{n} x_{i\alpha}\right)\left(\sum_{\alpha=1}^{n} y_{\alpha}\right)}{n},$$

$$a_{yy} = \sum_{\alpha=1}^{n} y_{\alpha}^{2} - \frac{\left(\sum_{\alpha=1}^{n} y_{\alpha}\right)^{2}}{n},$$

i = j = 1, 2, 3, 4, and n is the number of observations.

The first feature of this method is that the sum of squares for the dependent variable, a_{yy} , is recorded as the last entry of the column containing the sums of products of the dependent variable with the independent variables. The addition of a_{yy} to the last column results in a square matrix. This feature will be utilized to adjust the number of digits preceding the decimal points in the elements of the above matrix. The residual sum of squares is also obtained directly by adding the term a_{yy} to the last column of (1).

Then, in order to simplify the calculations, make the diagonal terms $(a_{11}, a_{22}, a_{23}, a_{44}, a_{yy})$ lie between 0.1 and 10 by pre- and post-multiplying (1) by a diagonal matrix of powers of ten which is as follows:

(2)
$$\begin{bmatrix} 10^{\mathfrak{q}_1} & 0 & 0 & 0 & 0 \\ 0 & 10^{\mathfrak{q}_1} & 0 & 0 & 0 \\ 0 & 0 & 10^{\mathfrak{q}_1} & 0 & 0 \\ 0 & 0 & 0 & 10^{\mathfrak{q}_1} & 0 \\ 0 & 0 & 0 & 0 & 10^{\mathfrak{p}} \end{bmatrix}$$

This will also result in the sums of products having at most one digit before the decimal points, thus allowing a uniform number of decimal places in all future calculations. The values of the q's and p are determined as follows:

Consider only a diagonal term and in it the largest even number of places through which the decimal point must be shifted (left or right) to make that term be a number between 0.1 and 10. Then divide this even number by two to get the applicable value of q_1 or p. For example, if $a_{11} = 8,238.93$, q_1 would be -2; if $a_{33} = 2,213,922.00$, q_3 would be -3; and if $a_{yy} = 5,098,35$, p would be -2.

After pre- and post-multiplying (1) by (2), which in effect is accomplished by adding the q's and p according to the term we are adjusting and shifting the decimal point the number of places indicated by the sum, we obtain a matrix of a^* 's. If, as in the above paragraph, $q_1 = -2$ and p = -2, we would shift the decimal point of a_{11} four places to the left and the decimal point of a_{1y} four places to the left, etc.