

Contents

List of Algorithms and Programs	ix
Notation	xi
1 Introduction	1
1.1 The development of calculating devices in statistics	1
1.2 Programming a computer for statistics	3
1.3 Representation of numbers in a computer	4
1.4 Interactive and batch computing	5
1.5 Exercises	6
2 Some general principles	8
2.1 Computation of a mean	8
2.2 Making a program understandable	10
2.3 Making the output understandable	11
2.4 Program structure	12
2.4.1 Algorithms	13
2.4.2 Properties of a good program	13
2.5 Exercises	14
3 Sorting and ranking	15
3.1 Introduction	15
3.2 The Exchange sort	15
3.3 Shellsort	16
3.4 A comparison of sorting methods	18
3.4.1 Quicksort	18
3.4.2 Monkey Puzzle sort	19
3.4.3 Sorting times	19
3.5 Ranking	20
3.6 Maximum and minimum values	21
3.7 Quantiles	23
3.8 Test data for Algorithms	24
3.9 Exercises	24
4 Inspection and summary of data using tables	27
4.1 Introduction	27
4.2 Input of an array	27
4.3 A six-number summary	27

4.4	Output of the six-number summary	31
4.4.1	The PRINT statement	31
4.4.2	Use of delimiters	31
4.4.3	Use of the TAB function	32
4.4.4	Use of spaces in output	33
4.4.5	Rounding data	33
4.5	Frequency table for discrete variate with repeated values	33
4.6	Frequency table with data grouped in classes	34
4.6.1	Choice of class-intervals	35
4.7	Two-way table from classification data	36
4.8	Test data for Algorithms	38
4.9	Exercises	38
5	Inspection and summary of data using graphical methods	41
5.1	Introduction	41
5.2	Histogram	42
5.3	Stem and leaf diagram	43
5.4	Transformations	44
5.5	Box and whisker plot	44
5.6	Scatter diagram	47
5.7	Scatter diagram using cursor controls	49
5.8	High resolution graphics	50
5.9	Test data for Algorithms	50
5.10	Exercises	50
6	Computation of variance and correlation coefficient	53
6.1	Variance by methods based on standard formulae	53
6.2	An updating procedure for calculating mean and variance	54
6.3*	Proof of recurrence relations for mean and variance	55
6.4	Calculation of sums of squares and products of two variates	56
6.5	Input of bivariate data	57
6.6	Calculation of product-moment correlation coefficient	59
6.7	Calculation of Spearman's rank correlation coefficient	59
6.8	The menu method of assembling Algorithms	59
6.9	Test data for Algorithms	61
6.10	Exercises	61
7	Simulation	64
7.1	Introduction	64
7.2	Simulation of card collecting	64
7.3	Generation of pseudo-random numbers	65
7.3.1	The mid-square method	67
7.4	The congruence method	68
7.5	Example of a linear congruential random number generator	69
7.6	Tests of randomness	70
7.6.1	Frequency test	71

7.6.2	Serial test	71
7.6.3	Serial correlation	71
7.7	Exercises	71
8	Probability functions of random variables	73
8.1	Introduction	73
8.2	Discrete random variables	74
8.2.1	Binomial distribution	74
8.2.2	Poisson distribution	74
8.2.3	Geometric distribution	75
8.3	The exponential distribution	75
8.4	Calculation of integrals	76
8.5	The normal integral	77
8.5.1	Simpson's rule	77
8.5.2	An approximation	78
8.5.3	Inverse normal integral by approximation	78
8.6*	Distributions based on the normal distribution	79
8.6.1	The t distribution	80
8.6.2	The χ^2 distribution	80
8.6.3	The F distribution	82
8.6.4	Approximations	82
8.6.5	Inverse distribution functions based on approximations	84
8.6.6	Inverse F distribution	87
8.7	Test data for Algorithms	87
8.8	Exercises	88
9	Generation of random samples from non-uniform distributions	90
9.1	Introduction	90
9.2	Simulation of the behaviour of a queue	90
9.3	A general method of sampling statistical distributions	92
9.4	Discrete random variables	95
9.4.1	Bernoulli distribution	95
9.4.2	Binomial distribution	96
9.4.3	Poisson distribution	96
9.4.4	Geometric distribution	97
9.5	Normal distribution	98
9.5.1	Use of the central limit theorem	98
9.5.2	The Box–Muller method	99
9.6	Distributions based on the normal distribution	99
9.7	Other continuous distributions	100
9.7.1	Exponential distribution	100
9.7.2	Gamma distribution	100
9.8	Exercises	101
10	Significance tests and confidence intervals	104
10.1	Introduction	104
10.2	Single sample analysis	104

10.2.1	The t -test	105
10.2.2	Confidence interval	106
10.3	Paired samples	106
10.3.1	Wilcoxon signed-rank test	107
10.4	Two unpaired samples	108
10.4.1	The t -test	108
10.4.2	Confidence interval	109
10.4.3	Mann–Whitney U test	109
10.5	Frequency data	111
10.5.1	Chi-squared goodness-of-fit tests	111
10.5.2	Chi-squared contingency table tests	113
10.6	Test data for Algorithms	114
10.7	Exercises	114
11	Regression	117
11.1	Introduction	117
11.2	Fitting a line to pairs of values	117
11.3	Residuals	119
11.4*	Multiple regression	120
11.5	Test data for Algorithms	122
11.6	Exercises	123
12	Analysis of variance	125
12.1	Introduction	125
12.2	Input of data	125
12.3	One-way analysis of variance	126
12.4	Two-way analysis of variance	129
12.5	Test data for Algorithms	131
12.6	Exercises	131
Appendix A	Using the Algorithms in this book	134
A.1	Description of Algorithms	134
A.2	Running an Algorithm	135
A.3	Combining Algorithms	136
A.3.1	Chaining Algorithms	136
A.3.2	Subroutines	137
A.4	The menu approach	137
A.5	Points to note when combining Algorithms	138
Appendix B	Compact Algorithms	140
Appendix C	Student data set	147
C.1	Student data set	147
C.2	Exercises using student data set	149
Bibliography		151
Index		152