

Statistics and the Evaluation of Evidence for Forensic Scientists

Second Edition

Colin G.G. Aitken

*School of Mathematics,
The University of Edinburgh, UK*

Franco Taroni

*Ecole des Sciences Criminelles and
Institut de Médecine Légale,
The University of Lausanne, Switzerland*



John Wiley & Sons, Ltd

***Statistics and the
Evaluation of Evidence
for Forensic Scientists***

STATISTICS IN PRACTICE

Founding Editor

Vic Barnett

Nottingham Trent University, UK

Statistics in Practice is an important international series of texts which provide detailed coverage of statistical concepts, methods and worked case studies in specific fields of investigation and study.

With sound motivation and many worked practical examples, the books show in down-to-earth terms how to select and use an appropriate range of statistical techniques in a particular practical field within each title's special topic area.

The books provide statistical support for professionals and research workers across a range of employment fields and research environments. Subject areas covered include medicine and pharmaceuticals; industry, finance and commerce; public services; the earth and environmental sciences, and so on.

The books also provide support to students studying statistical courses applied to the above areas. The demand for graduates to be equipped for the work environment has led to such courses becoming increasingly prevalent at universities and colleges.

It is our aim to present judiciously chosen and well-written workbooks to meet everyday practical needs. Feedback of views from readers will be most valuable to monitor the success of this aim.

A complete list of titles in this series appears at the end of the volume.

Statistics and the Evaluation of Evidence for Forensic Scientists

Second Edition

Colin G.G. Aitken

*School of Mathematics,
The University of Edinburgh, UK*

Franco Taroni

*Ecole des Sciences Criminelles and
Institut de Médecine Légale,
The University of Lausanne, Switzerland*



John Wiley & Sons, Ltd

Copyright © 2004 John Wiley & Sons Ltd, The Atrium, Southern Gate, Chichester,
West Sussex PO19 8SQ, England
Telephone (+44) 1243 779777

Email (for orders and customer service enquiries): cs-books@wiley.co.uk
Visit our Home Page on www.wileyeurope.com or www.wiley.com

All Rights Reserved. No part of this publication may be reproduced, stored in a retrieval system or transmitted in any form or by any means, electronic, mechanical, photocopying, recording, scanning or otherwise, except under the terms of the Copyright, Designs and Patents Act 1988 or under the terms of a licence issued by the Copyright Licensing Agency Ltd, 90 Tottenham Court Road, London W1T 4LP, UK, without the permission in writing of the Publisher. Requests to the Publisher should be addressed to the Permissions Department, John Wiley & Sons Ltd, The Atrium, Southern Gate, Chichester, West Sussex PO19 8SQ, England, or emailed to permreq@wiley.co.uk, or faxed to (+44) 1243 770620.

This publication is designed to provide accurate and authoritative information in regard to the subject matter covered. It is sold on the understanding that the Publisher is not engaged in rendering professional services. If professional advice or other expert assistance is required, the services of a competent professional should be sought.

Other Wiley Editorial Offices

John Wiley & Sons Inc., 111 River Street, Hoboken, NJ 07030, USA

Jossey-Bass, 989 Market Street, San Francisco, CA 94103-1741, USA

Wiley-VCH Verlag GmbH, Boschstr. 12, D-69469 Weinheim, Germany

John Wiley & Sons Australia Ltd, 33 Park Road, Milton, Queensland 4064, Australia

John Wiley & Sons (Asia) Pte Ltd, 2 Clementi Loop #02-01, Jin Xing Distripark, Singapore 129809

John Wiley & Sons Canada Ltd, 22 Worcester Road, Etobicoke, Ontario, Canada M9W 1L1

Wiley also publishes its books in a variety of electronic formats. Some content that appears in print may not be available in electronic books.

Library of Congress Cataloging in Publication Data

Aitken, C.G.G.

Statistics and the evaluation of evidence for forensic scientists / Colin G.G. Aitken,
Franco Taroni. — 2nd ed.

p. cm. — (Statistics in practice)

Includes bibliographical references and index.

ISBN 0-470-84367-5 (alk. paper)

1. Forensic sciences—Statistical methods. 2. Forensic statistics. 3. Evidence, Expert.
I. Taroni, Franco. II. Title. III. Statistics in practice (Chichester, England)

HV8073.A55 2004

363.23'015195—dc22

2004048016

British Library Cataloguing in Publication Data

A catalogue record for this book is available from the British Library

ISBN 0-470-84367-5

Typeset in 10/12pt Photina by Integra Software Services Pvt. Ltd, Pondicherry, India

Printed and bound in Great Britain by TJ International, Padstow, Cornwall

This book is printed on acid-free paper responsibly manufactured from sustainable forestry in which at least two trees are planted for each one used for paper production.

To
Liz, David and Catherine
and
Anna and Wanda

Contents

List of tables	xiii
List of figures	xix
Foreword	xxii
Preface to the first edition	xxv
Preface to the second edition	xxviii
1 Uncertainty in forensic science	1
1.1 Introduction	1
1.2 Statistics and the law	2
1.3 Uncertainty in scientific evidence	5
1.3.1 The frequentist method	6
1.3.2 Stains of body fluids	7
1.3.3 Glass fragments	9
1.4 Terminology	12
1.5 Types of data	15
1.6 Probability	16
1.6.1 Introduction	16
1.6.2 A standard for uncertainty	18
1.6.3 Events	20
1.6.4 Subjective probability	21
1.6.5 Laws of probability	23
1.6.6 Dependent events and background information	25
1.6.7 Law of total probability	29
1.6.8 Updating of probabilities	32
2 Variation	35
2.1 Populations	35
2.2 Samples and estimates	37
2.3 Counts	40
2.3.1 Probabilities	40
2.3.2 Summary measures	41
2.3.3 Binomial distribution	43
2.3.4 Multinomial distribution	44
2.3.5 Hypergeometric distribution	45
2.3.6 Poisson distribution	48
2.3.7 Beta-binomial distribution	51
2.4 Measurements	52
2.4.1 Summary statistics	52
2.4.2 Normal distribution	53

2.4.3	Student's <i>t</i> -distribution	60
2.4.4	Beta distribution	62
2.4.5	Dirichlet distribution	63
2.4.6	Multivariate Normal and correlation	64
3	The evaluation of evidence	69
3.1	Odds	69
3.1.1	Complementary events	69
3.1.2	Examples	70
3.1.3	Definition	70
3.2	Bayes' Theorem	72
3.2.1	Statement of the theorem	72
3.2.2	Examples	73
3.3	Errors in interpretation	78
3.3.1	Fallacy of the transposed conditional	79
3.3.2	Source probability error	81
3.3.3	Ultimate issue error	82
3.3.4	Defender's fallacy	82
3.3.5	Probability (another match) error	83
3.3.6	Numerical conversion error	84
3.3.7	False positive fallacy	85
3.3.8	Uniqueness	86
3.3.9	Other difficulties	87
3.3.10	Empirical evidence of errors in interpretation	89
3.4	The odds form of Bayes' theorem	95
3.4.1	Likelihood ratio	95
3.4.2	Logarithm of the likelihood ratio	99
3.5	The value of evidence	101
3.5.1	Evaluation of forensic evidence	101
3.5.2	Summary of competing propositions	105
3.5.3	Qualitative scale for the value of the evidence	107
3.5.4	Misinterpretations	111
3.5.5	Explanation of transposed conditional and defence fallacies	112
3.5.6	The probability of guilt	116
3.6	Summary	118
4	Historical review	119
4.1	Early history	119
4.2	The Dreyfus case	122
4.3	Statistical arguments by early twentieth-century forensic scientists	125
4.4	People v. Collins	126
4.5	Discriminating power	129
4.5.1	Derivation	129
4.5.2	Evaluation of evidence by discriminating power	130
4.5.3	Finite samples	133
4.5.4	Combination of independent systems	135
4.5.5	Correlated attributes	136
4.6	Significance probabilities	141
4.6.1	Calculation of significance probabilities	141
4.6.2	Relationship to likelihood ratio	144
4.6.3	Combination of significance probabilities	147

4.7	Coincidence probabilities	149
	4.7.1 Introduction	149
	4.7.2 Comparison stage	151
	4.7.3 Significance stage	151
4.8	Likelihood ratio	153
5	Bayesian inference	157
5.1	Introduction	157
5.2	Bayesian inference for a Bernoulli probability	160
5.3	Estimation with zero occurrences in a sample	162
5.4	Estimation of products in forensic identification	165
5.5	Bayesian inference for a Normal mean	166
5.6	Interval estimation	170
	5.6.1 Confidence intervals	170
	5.6.2 Highest posterior density intervals	172
	5.6.3 Bootstrap intervals	172
	5.6.4 Likelihood intervals	173
5.7	Odds ratios	175
6	Sampling	179
6.1	Introduction	179
6.2	Choice of sample size	182
	6.2.1 Large consignments	182
	6.2.2 Small consignments	186
6.3	Quantity estimation	190
	6.3.1 Frequentist approach	190
	6.3.2 Bayesian approach	191
6.4	Misleading evidence	197
7	Interpretation	205
7.1	Concepts and court cases	205
	7.1.1 Relevant population	205
	7.1.2 Consideration of odds	206
	7.1.3 Combination of evidence	208
	7.1.4 Specific cases	208
7.2	Pre-assessment and relevant propositions	214
	7.2.1 Levels of proposition	214
	7.2.2 Pre-assessment of the case	217
	7.2.3 Pre-assessment of the evidence	220
7.3	Assessment of value of various evidential types	221
	7.3.1 Earprints	221
	7.3.2 Firearms and toolmarks	223
	7.3.3 Fingerprints	226
	7.3.4 Speaker recognition	228
	7.3.5 Hair	229
	7.3.6 Documents	231
	7.3.7 Envelopes	233
	7.3.8 Handwriting	235
	7.3.9 Paint	239
7.4	Pre-data and post-data questions	239

8	Transfer evidence	245
8.1	The likelihood ratio	245
	8.1.1 Probability of guilt	245
	8.1.2 Justification	246
	8.1.3 Combination of evidence and comparison of more than two propositions	248
8.2	Correspondence probabilities	254
8.3	Direction of transfer	255
	8.3.1 Transfer of evidence from the criminal to the scene	255
	8.3.2 Transfer of evidence from the scene to the criminal	260
	8.3.3 Transfer probabilities	261
	8.3.4 Two-way transfer	270
	8.3.5 Presence of non-matching evidence	271
8.4	Grouping	271
8.5	Relevant populations	274
9	Discrete data	283
9.1	Notation	283
9.2	Single sample	283
	9.2.1 Introduction	283
	9.2.2 General population	286
	9.2.3 Particular population	286
	9.2.4 Examples	286
9.3	Two samples	288
	9.3.1 Two stains, two offenders	288
	9.3.2 DNA profiling	291
9.4	Many samples	292
	9.4.1 Many different profiles	292
	9.4.2 General cases	293
9.5	Relevance of evidence and relevant material	295
	9.5.1 Introduction	295
	9.5.2 Subjective probabilities	296
	9.5.3 Association propositions	296
	9.5.4 Intermediate association propositions	297
	9.5.5 Examples	298
	9.5.6 Two stains, one offender	302
9.6	Summary	304
	9.6.1 Stain known to have been left by offenders	304
	9.6.2 Relevance: stain may not have been left by offenders	305
	9.6.3 Relevance and the crime level	307
9.7	Missing persons	308
	9.7.1 Case 1 (Kuo, 1982)	309
	9.7.2 Case 2 (Ogino and Gregonis, 1981)	309
	9.7.3 Calculation of the likelihood ratio	310
9.8	Paternity: combination of likelihood ratios	312
	9.8.1 Likelihood of paternity	314
	9.8.2 Probability of exclusion in paternity	317

10	Continuous data	319
10.1	The likelihood ratio	319
10.2	Normal distribution for between-source data	321
	10.2.1 Sources of variation	322
	10.2.2 Derivation of the marginal distribution	322
	10.2.3 Approximate derivation of the likelihood ratio	324
	10.2.4 Lindley's approach	326
	10.2.5 Interpretation of result	327
	10.2.6 Examples	328
10.3	Estimation of a probability density function	330
10.4	Kernel density estimation for between-source data	337
	10.4.1 Application to medullary widths of cat hairs	339
	10.4.2 Refractive index of glass	339
10.5	Probabilities of transfer	342
	10.5.1 Introduction	342
	10.5.2 Single fragment	342
	10.5.3 Two fragments	345
	10.5.4 A practical approach to glass evaluation	349
	10.5.5 Graphical models for the assessment of transfer probabilities	352
10.6	Approach based on <i>t</i> -distribution	353
10.7	Appendix Derivation of <i>V</i> when the between-source measurements are assumed normally distributed	357
11	Multivariate analysis	359
11.1	Introduction	359
11.2	Description of example	360
11.3	Univariate <i>t</i> -tests	362
11.4	Hotelling's T^2	363
11.5	Univariate Normality, two sources of variation	365
11.6	Multivariate Normality, two sources of variation	366
11.7	Caveat lector	371
11.8	Summary	372
11.9	Appendix	373
	11.9.1 Matrix terminology	373
	11.9.2 Determination of a likelihood ratio with an assumption of Normality	377
12	Fibres	381
12.1	Introduction	381
12.2	Likelihood ratios in scenarios involving fibres	381
	12.2.1 Fibres evidence left by an offender	382
	12.2.2 Comments on the fibres scenario	387
	12.2.3 Fibres evidence not left by the offender	388
	12.2.4 Cross-transfer	389
12.3	Pre-assessment in fibres scenarios	392
	12.3.1 The case history	392
	12.3.2 Formulation of the pairs of propositions and events	392
	12.3.3 Assessment of the expected value of the likelihood ratio	394
12.4	The relevant population of fibres	396

13	DNA profiling	399
13.1	Introduction	399
13.2	Hardy–Weinberg equilibrium	401
13.3	DNA likelihood ratio	404
13.4	Uncertainty	405
13.5	Variation in sub-population allele frequencies	405
13.6	Related individuals	409
13.7	More than two propositions	412
13.8	Database searching	414
	13.8.1 Search and selection effect (double counting error)	418
13.9	Island problem	419
13.10	Mixtures	421
13.11	Error rate	424
14	Bayesian networks	429
14.1	Introduction	429
14.2	Bayesian networks	430
	14.2.1 The construction of Bayesian networks	431
14.3	Evidence at the crime level	437
	14.3.1 Preliminaries	437
	14.3.2 Description of probabilities required	437
14.4	Missing evidence	439
	14.4.1 Preliminaries	439
	14.4.2 Determination of a structure for a Bayesian network	440
	14.4.3 Comments	442
14.5	Error rates	442
	14.5.1 Preliminaries	442
	14.5.2 Determination of a structure for a Bayesian network	443
14.6	Transfer evidence	444
	14.6.1 Preliminaries	444
	14.6.2 Determination of a structure for a Bayesian network	444
	14.6.3 Comment on the <i>transfer</i> node	446
14.7	Combination of evidence	447
14.8	Cross-transfer evidence	449
	14.8.1 Description of nodes	451
	14.8.2 Probabilities for nodes	452
14.9	Factors to consider	452
	14.9.1 Parameter choice	453
14.10	Summary	453
	References	455
	Notation	485
	Cases	489
	Author index	491
	Subject index	499

List of Tables

1.1	Genotypic frequencies for locus <i>LDLR</i> amongst Caucasians in Chicago based on a sample of size 200 (from Johnson and Peterson, 1999).	8
1.2	Refractive index measurements.	10
1.3	Genotype probabilities, assuming Hardy–Weinberg equilibrium, for a diallelic system with allele probabilities p and q .	27
1.4	The proportion of people in a population who fall into the four possible categories of genetic markers.	29
2.1	Probabilities for the number of sixes, X , in four rolls of a fair six-sided die	41
2.2	Intermediate calculations for the variance of the number of sixes, x , in four rolls of a fair six-sided die	43
2.3	Probabilities that a random variable X with a Poisson distribution with mean 4 takes values 0, 1, 2, 3, 4 and greater than 4	50
2.4	Values of cumulative distribution function $\Phi(z)$ and its complement $1 - \Phi(z)$ for the standard Normal distribution for given values of z .	57
2.5	Probabilities for absolute values from the standard Normal distribution function	57
2.6	Percentage points $t_{(n-1)}(P)$ for the t -distribution for given values of sample size n , degrees of freedom $(n - 1)$ and P , and the corresponding point $z(P)$ for the standard Normal distribution.	61
3.1	Two-by-two contingency table for frequencies for the tabulation of patients with or without a disease (S or \bar{S}) given a blood test positive or negative (R or \bar{R}).	74
3.2	Two-by-two contingency table for probabilities for the tabulation of patients with or without a disease (S or \bar{S}) given a blood test positive or negative (R or \bar{R}).	75
3.3	Hypothetical results for deaths amongst a population.	76

3.4	Probability θ of at least one match, given a frequency of the trace evidence of γ , in a population of size 1 million.	83
3.5	Evidence occurs with RMP γ . Smallest number ψ of people to be observed before a match with the evidence occurs with a given probability, $Pr(M) = 0.5, 0.9$; $\psi_5 = \log 0.5 / \log(1 - \gamma)$, $\psi_9 = \log 0.1 / \log(1 - \gamma)$, n_5 is the smallest integer greater than ψ_5 , n_9 is the smallest integer greater than ψ_9 .	85
3.6	The probability, θ' , of at least one match with the evidence, which occurs with RMP γ , when $n' = 1/\gamma$ people are tested.	85
3.7	Frequency of Kell and Duffy types by colour in a hypothetical area. (Reproduced by permission of The Forensic Science Society.).	97
3.8	Effect on prior odds in favour of H_p relative to H_d of evidence E with value V of 1 000. Reference to background information I is omitted.	104
3.9	Value of the evidence for each genotype.	107
3.10	Qualitative scale for reporting the value of the support of the evidence for H_p against H_d (Evetts <i>et al.</i> , 2000a).	107
3.11	The values of the logarithm of the posterior odds in favour of an issue determined from the values of the logarithm of the prior odds in favour of guilt ($\log(\text{Prior odds})$) and the logarithm of the likelihood ratio (Aitken and Taroni, 1998). The values in the body of the table are obtained by adding the appropriate row and column values. Logarithms are taken to base 10. The verbal description is taken from Calman and Royston (1997). (Reproduced by permission of The Forensic Science Society.).	110
3.12	Probability of guilt required for proof beyond reasonable doubt (Simon and Mahan, 1971).	116
4.1	Probabilities suggested by the prosecutor for various characteristics of the couple observed in the case of <i>People v. Collins</i> .	127
4.2	Allelic frequencies for <i>TPOX</i> locus for Swiss and NZ Caucasians and the probability Q_{TPOX} of a match.	136
4.3	Allelic frequencies for <i>TH01</i> locus for Swiss and NZ Caucasians and the probability Q_{TH01} of a match.	136
4.4	The calculation of discriminating power (DP) for Normal distributions of p dimensions.	139
4.5	Significance probabilities P for refractive index x of glass for mean $\theta_0 = 1.518458$ and standard deviation $\sigma = 4 \times 10^{-5}$ and decisions assuming a significance level of 5%	143
4.6	Variation in the likelihood ratio V , as given by (4.7), with sample size n , for a standardised distance $z_n = 2$, a result which is significant at the 5% level.	147

5.1	Probabilistic upper bounds θ_0 for the proportion θ of occurrences of an event in a population when there have been no occurrences in a sample of size n , for given values of $\epsilon = Pr(\theta > \theta_0)$, with a beta prior for which $\alpha = \beta = 1$.	164
5.2	Numbers of banknotes contaminated.	176
5.3	The value of evidence in comparing two binomial proportions for different values of sample size n and number of ‘successes’ z , where the samples are banknotes and a ‘success’ is a contaminated banknote. The binomial proportion in the numerator is 0.83, estimated from banknotes known to be used for drug dealing, and the binomial proportion in the denominator is 0.57, estimated from banknotes known to be from general circulation.	178
6.1	The probability that the proportion of drugs in a large consignment is greater than 50% for various sample sizes m and prior parameters α and β . (Reprinted with permission from ASTM International.)	184
6.2	The sample size required to be 100% certain that the proportion of units in the consignment which contain drugs is greater than θ , when all the units inspected are found to contain drugs. The prior parameters $\alpha = \beta = 1$. (Reprinted with permission from ASTM International.)	185
6.3	Estimates of quantities q of drugs (in grams), in a consignment of $m + n$ units, according to various possible burdens of proof, expressed as percentages $P = 100 \times Pr(Q > q \mid m, z, n, \bar{x}, s)$ in 26 packages when 6 packages are examined ($m = 6, n = 20$) and $z = 6, 5$, or 4 are found to contain drugs. The mean (\bar{x}) and standard deviation (s) of the quantities found in the packages examined which contain drugs are 0.0425 g and 0.0073 g. The parameters for the beta prior are $\alpha = \beta = 1$. Numbers in brackets are the corresponding frequentist lower bounds using the fpc factor (6.2). (Reprinted with permission from ASTM International.)	194
6.4	Estimates of quantities q of drugs (in grams), in a consignment of $m + n$ units, according to various possible burdens of proof, expressed as percentages $P = 100 \times Pr(Q > q \mid m, z, n, \bar{x}, s)$ in 2600 packages when 6 packages are examined ($m = 6, n = 2594$) and $z = 6, 5$, or 4 are found to contain drugs. The mean (\bar{x}) and standard deviation (s) of the quantities found in the packages examined which contain drugs are 0.0425 g and 0.0073 g. The parameters for the beta prior are $\alpha = \beta = 1$. Numbers in brackets are the corresponding frequentist lower bounds without using the fpc factor (6.8). (Reprinted with permission from ASTM International.).	196

6.5	Probabilities of strong misleading evidence $M(n)$ and weak evidence $W(n)$ for boundary values k of 8 and 32 for strong evidence and sample sizes n of 5, 10 and 20.	203
7.1	Examples of the hierarchy of propositions. (Reproduced by permission of The Forensic Science Society.)	217
7.2	Hypothetical relative frequencies for the maximum CMS count y for bullets fired from the same gun $f(y SG)$ and for bullets fired from different guns $f(y DG)$, and likelihood ratios $V = f(y SG)/f(y DG)$. (Reprinted with permission from ASTM International.)	224
7.3	Poisson probabilities for the maximum CMS count y for bullets fired from the same gun $Pr(Y = y SG)$ (7.1) with mean $\lambda_S = 3.91$ and for bullets fired from different guns $Pr(Y = y DG)$ (7.2) with mean $\lambda_D = 1.325$ and likelihood ratio $V = Pr(Y = y SG)/Pr(Y = y DG)$. (Reprinted with permission from ASTM International.)	225
7.4	Values of a probabilistic γ upper bound for the probability ϕ_0 of a false match for different values of n and γ .	242
7.5	Values of a probabilistic γ lower bound for the likelihood ratio $1/\phi_0$ for different values of n and γ , with values rounded down.	242
8.1	Probability evidence in <i>State v. Klindt</i> (altered for illustrative purposes) from Lenth (1986).	249
8.2	Frequencies of Ruritanians and those of genotype Γ for locus <i>LDLR</i> in a hypothetical population.	256
8.3	Distribution of blood groups of innocently acquired bloodstains on clothing of people of type O , compared with the distribution in the general population.	268
8.4	Critical values $r(\alpha, n)$ for the range of a sample of size n from a standardised Normal distribution, from Owen (1962).	273
8.5	Critical values $\lambda(\alpha, n)$ for samples of size n for the divisive grouping algorithm from Triggs <i>et al.</i> (1997).	274
8.6	Probability that a criminal lived in a particular area, given the crime was committed in Auckland, New Zealand.	280
9.1	Gene frequencies for New Zealand in the <i>ABO</i> system.	287
9.2	Gene frequencies for Chinese in New Zealand in the <i>ABO</i> system.	287
9.3	Phenotypes for two cases of missing persons.	308
9.4	Gene frequencies and phenotypic incidences for Case 1 (Kuo, 1982).	309
9.5	Gene frequencies and phenotypic incidences for Case 2 (Ogino and Gregonis, 1981).	309
9.6	Frequencies for up to three codominant alleles.	310
9.7	Relative frequencies for <i>ABO</i> system.	310

9.8	Probabilities that the parents will pass the specified stain phenotype, Case 1.	311
9.9	Probabilities that the parents will pass the specified phenotypes, Case 2.	311
9.10	Two pieces of evidence on DNA markers.	312
9.11	Posterior probabilities of paternity for various prior probabilities for evidence for alleged father $E_1 = 11 - 13, E_2 = 18 - 18$.	316
9.12	Hummel's likelihood of paternity.	317
10.1	Likelihood ratio values for varying values of $(y - \bar{x})/\sigma$ and $(y - \mu)/\tau$.	326
10.2	Value of $\tau(2^{1/2}\sigma)^{-1} \exp(-\frac{1}{2}\lambda^2 + \frac{1}{2}\delta^2)$ (10.7) as a function of $\lambda = \bar{x} - \bar{y} / (2^{1/2}\sigma)$ and $\delta = z - \mu / \tau$ for $\tau/\sigma = 100$.	329
10.3	Medullary widths in microns of 220 cat hairs (Peabody <i>et al.</i> , 1983).	332
10.4	Value of the evidence for various values of \bar{x} and \bar{y} , the smoothing parameter λ and the within-cat standard deviation σ ; $m = n = 10$ throughout; $s = 23$ microns.	339
10.5	Refractive index of 2269 fragments of float glass from buildings, Lambert and Evett (1984).	340
10.6	Coincidence probability and value of the evidence (kernel and Lindley approaches) for various values of \bar{x} and \bar{y} and the smoothing parameter λ (for the kernel approach); $m = 10, n = 5$; within-window standard deviation $\sigma = 0.000\ 04$, between-window standard deviation $\tau = 0.004$; overall mean $\mu = 1.5182$.	341
10.7	Distributional parameters for glass problems.	344
10.8	Transfer probabilities for glass problems.	344
10.9	Some values for the likelihood ratio V for the single-fragment case, from Evett (1986).	345
10.10	Possible sources of two fragments.	345
10.11	Refractive indices of glass fragments for Johnston, recovered, and a control set with means, separate and pooled standard deviations (s.d). The number of recovered fragments $n_y = 11$ and the number of control fragments $n_x = 10$. Example presented in Walsh <i>et al.</i> (1996).	355
10.12	Summary statistics for concentration of dye CI 14720 in illicit pills (Goldmann <i>et al.</i> , 2004).	356
12.1	Findings from the analysis of the hair combings.	394
12.2	Events and probabilities relating to findings under H_p and H_d .	395
12.3	Likelihood ratios for the outcomes from Table 12.1 with $t_0 = 0.01, t_s = 0.04, t_l = 0.95; p_0 = 0.78, p_1 = 0.22; s_s = 0.92, s_l = 0.08; m = 0.05$, as proposed by Champod and Jackson (2000).	396
13.1	Hardy-Weinberg proportions for a locus with two alleles, A and a , with frequencies p and q such that $p + q = 1$.	402

13.2	Observed and expected frequencies of <i>HumTH01</i> genotypes based on 95 unrelated Turkish individuals (from Çakir <i>et al.</i> , 2001).	403
13.3	Effects of population structure, as represented by F_{ST} , on the likelihood ratio, the reciprocal of the conditional match probability (13.8) for heterozygotes between alleles with equal frequency p .	408
13.4	Match probability $Pr(G G, H_d, I)$ that a relative has the same genotype as the suspect and the corresponding value for V assuming allelic frequencies of 0.1, from Weir and Hill (1993). (Reproduced by permission of The Forensic Science Society.)	410
13.5	Effects of family relatedness on match probability, $Pr(G_c G_s, H_d, I)$, from Weir (2001a). Note the use of θ for F_{ST} for clarity. ©2001. John Wiley & Sons, Ltd. Reproduced with permission.	411
13.6	General match probability values recommended for use when reporting full SGM-plus profile matches, with an F_{ST} value of 0 in situation 6 and a value of 0.02 for situations 1–5, from Foreman and Evett (2001). (Reproduced by permission of Springer-Verlag.)	411
13.7	Likelihood ratio for three-allele mixed sample, heterozygous victim and homozygous suspect. The victim has alleles a, b , the suspect has allele c . There are three ethnic groups under consideration, labelled 1, 2 and 3. The co-ancestry coefficient F_{ST} for the ethnic group, 1, of the unknown donor is labelled θ_1 for clarity. The allelic frequencies are p_{a1}, p_{b1} and p_{c1} in ethnic group 1. Ethnicities of the unknown donor, the victim and the suspect are considered. From Fung and Hu (2002). (Reproduced by permission of Springer-Verlag.)	423
13.8	Posterior odds that a suspect is the source of a sample that reportedly has a matching DNA profile, as a function of prior odds, random match probability, and false positive probability. Extracted from Thompson <i>et al.</i> (2003). (Reprinted with permission from ASTM international.)	426
14.1	Conditional probabilities for transfer with one group. The features of the control object are X , and the features of the recovered fibres are Y , which may be from 0, 1 or 2 groups. These recovered fibres may be present by chance alone (b_1) or not (b_0), or have been transferred from the victim (t_1) or not (t_0).	445
14.2	Conditional probabilities for transfer with many groups, as illustrated in Section 12.3.3. The transfer may be of no fibres (t_0), a small amount (t_s) or a large amount (t_l). A group may be present (p_1) or not (p_0), and the group may be small (s_s) or large (s_l). The outcome O has five categories: no groups, one not-matching group, one small matching group, one large matching group and two groups.	447

List of Figures

- 2.1 Probability density function for a Normal distribution, with mean 0.7 and variance 0.005. 55
- 2.2 Selected tail area probabilities for a standard Normal random variable Z : (a) $Pr(Z > 1)$, (b) $Pr(Z > 2)$, (c) $Pr(|Z| < 2)$, (d) $Pr(Z > 2.5)$. 58
- 4.1 Discrimination between two groups, identified by Δ and by \circ , with two perfectly correlated variables. 140
- 5.1 Prior density function $f(\theta | \alpha, \beta)$ with $\alpha = 61$, $\beta = 44$, likelihood function $L(\theta | n, x)$ with $n = 372$, $x = 201$, and posterior density function $f(\theta | x + \alpha, n - x + \beta)$ for a Bernoulli parameter. (Reproduced from Weir, 1996a, by permission of Sinauer Associates, Inc.) 162
- 5.2 Standardised likelihood function for the proportion of people of blood group Γ , from a sample of size 30 in which 6 were of group Γ . A likelihood interval, derived from the observed data, of fairly strong support for the values of γ included within it is indicated with the dotted lines: (0.081, 0.372). 174
- 6.1 The prior probability $1 - F(\theta)$ that the proportion of units in a consignment is greater than θ , for various choices of α and β : $\alpha = \beta = 1$ (dot-dashed curve), $\alpha = \beta = 0.5$ (solid), $\alpha = 0.065$, $\beta = 0.935$ (dotted). (Reprinted with permission from ASTM International.) 186
- 6.2 The posterior probability $1 - F(\theta)$ that the proportion of units in a consignment is greater than θ , for various choices of α and β : $\alpha = \beta = 1$ (dot-dashed curve), $\alpha = \beta = 0.5$ (solid), $\alpha = 0.065$, $\beta = 0.935$ (dotted), after observation of four units, all found to be illegal. The corresponding probabilities that at least 50% of the consignment contains illegal units are marked as 0.985 ($\alpha = \beta = 0.5$), 0.970 ($\alpha = \beta = 1$), 0.950 ($\alpha = 0.065$, $\beta = 0.935$). (Reprinted with permission from ASTM International.) 187

6.3	The probability that the total quantity Q of drugs (in grams) in a consignment of 26 packages is greater than q when 6 packages are examined and 6 (solid curve), 5 (dashed), or 4 (dot-dashed) are found to contain drugs. The mean and standard deviation of the quantities found in the packages examined which contain drugs are 0.0425 g and 0.0073 g. The parameters for the beta prior are $\alpha = \beta = 1$. (Reprinted with permission from ASTM International.)	195
6.4	The probability that the total quantity Q of drugs (in grams) in a consignment of 2600 packages is greater than q when 6 packages are examined and 6 (solid curve), 5 (dashed), or 4 (dot-dashed) are found to contain drugs. The mean and standard deviation of the quantities found in the packages examined which contain drugs are 0.0425 g and 0.0073 g. The parameters for the beta prior are $\alpha = \beta = 1$. (Reprinted with permission from ASTM International.)	196
6.5	Bump function for the probability of misleading evidence $Pr_1\left(\frac{f_{2n}}{f_{1n}} > k\right)$ for $k = 8$ and $k = 32$ as a function of c , the distance from the true mean to the alternative, in standard errors. (Royall, 2000; reprinted with permission from the <i>Journal of the American Statistical Association</i> . Copyright 2000 by the American Statistical Association. All rights reserved.)	201
9.1	Variation in the logarithm to base 10 of the likelihood ratio V of the evidence with p , the probability that the stain would have been left by the suspect even though he was innocent of the offence, for various values of r , the probability that the stain would have been left by one of the offenders. The number of offenders, k , equals 4 and the relative frequency of the profile γ is 0.001. Adapted from Evett (1993a), with the inclusion of a curve for $r = 0$. The dotted line at $\log(V) = 0$ indicates where the evidence is equally likely under both propositions.	301
10.1	Refractive index measurements from 2269 fragments of float glass from buildings (from Lambert and Evett, 1984).	323
10.2	Medullary width (in microns) of 220 cat hairs (from Peabody <i>et al.</i> , 1983).	331
10.3	Examples of kernel density estimates showing individual kernels. Smoothing parameter values are (a) $\lambda = 0.5$ and (b) $\lambda = 1$.	333
10.4	Medullary widths, in microns, of cat hairs (Peabody <i>et al.</i> , 1983) and associated kernel density estimate with smoothing parameter equal to 0.09.	335

10.5	Medullary widths, in microns, of cat hairs (Peabody <i>et al.</i> , 1983) and associated kernel density estimate with smoothing parameter equal to 0.50.	335
10.6	Kernel density estimate with smoothing parameter 0.025 of refractive index measurements from 2269 fragments of float glass from buildings (Lambert and Evett, 1984).	336
10.7	Kernel density estimate with smoothing parameter 0.25 of refractive index measurements from 2269 fragments of float glass from buildings (Lambert and Evett, 1984).	337
10.8	The difference in refractive index measurements (higher – lower) for each pair of fragments for individuals who had two fragments of glass on their clothing, from Harrison <i>et al.</i> (1985). (Reproduced from Evett, 1986, by permission of The Forensic Science Society.)	347
10.9	Graphs of $\log_{10} \phi(1)$, $\log_{10} \phi(2)$ and $\log_{10} V = \log_{10}(\phi(1) + \phi(2))$ against $ \bar{y} - \bar{x} / \sigma$ for $\bar{y} = \mu$, for the transfer of two fragments of glass from the scene of the crime to the criminal: (a) $ y_1 - y_2 = \sigma$, (b) $ y_1 - y_2 = 4\sigma$. The value of the evidence is $V = \phi(1) + \phi(2)$. The dotted line is $\log_{10} V$. (Adapted from Evett, 1986.)	348
14.1	Basic connections in Bayesian networks: (a) serial, (b) diverging and (c) converging.	432
14.2	Bayesian network for evidence E and proposition H .	435
14.3	Bayesian network for a serial connection for a reported match RM in a DNA profile, where M denotes a match and H a proposition.	436
14.4	Four-node network for evaluation of evidence at the crime level.	437
14.5	Bayesian network fragments representing the relation between (a) the variables E and H , and (b) the variables M and E ; (c) Bayesian network for missing evidence.	441
14.6	Bayesian network for error rates. The probabilistic dependencies are indicated on the right-hand side of each node.	443
14.7	Five-node network: no outcome.	445
14.8	Five-node network: outcome, converging and serial in Bayesian networks.	446
14.9	Complete network for evaluation of cross-transfer evidence in DNA profiles. (Reprinted from Aitken <i>et al.</i> , 2003, with permission from Elsevier.)	451

Foreword

The statistical evaluation of evidence is part of the scientific method, here applied to forensic circumstances. Karl Pearson was both perceptive and correct when he said that 'the unity of all science consists alone in its method, not in its material'. Science is a way of understanding and influencing the world in which we live. In this view it is not correct to say that physics is a science, whereas history is not: rather that the scientific method has been much used in physics, whereas it is largely absent from history. Scientific method is essentially a tool and, like any tool, is more useful in some fields than in others. If this appreciation of science as a method is correct, one might enquire whether the method could profitably be applied to the law, but before we can answer this it is necessary to understand something of what the scientific method involves. Books have been written on the topic, and here we confine ourselves to the essential ingredients of the method.

Two ideas dominate the scientific approach, observation and reasoning. Observation may be passive, as with a study of the motions of the heavenly bodies or the collection of medical records. Often it is active, as when an experiment is performed in a laboratory, or a controlled clinical trial is employed. The next stage is to apply reasoning to the observed data, usually to think of a theory that will account for at least some of the features seen in the data. The theory can then be used to predict further observations which are then made and compared with prediction. It is this see-saw between evidence and theory that characterises the scientific method and which, if successful, leads to a theory that accords with the observational material. Classic examples are Newton's theory to explain the motions of the planets, and Darwin's development of the theory of evolution. It is important to notice that, contrary to what many people think, uncertainty is present throughout any scientific procedure. There will almost always be errors in the measurements, due to variation in the material or limitations of the apparatus. A theory is always uncertain, and that is why it has to be rigorously tested. Only late in the cycle of movement between hard fact and mental activity is a theory admitted as being true. Even then the 'truth' is, in the long run, not absolute, as can be seen in the replacement of

Newton by Einstein. It is important to recognise the key role that uncertainty plays in the scientific method.

If the above analysis is correct, it becomes natural to see connections between the scientific method and legal procedures. All the ingredients are there, although the terminology is different. In a court of law, the data consist of the evidence pertinent to the case, evidence obtained by the police and other bodies, and presented by counsels for the defence and prosecution. There are typically only two theories, that the defendant is guilty or is innocent. As the trial proceeds, the see-saw effect is exhibited as evidence accumulates. Lawyers use an adversarial system, not openly present in scientific practice, but similar to the peer reviews that are employed therein. The most striking similarity between legal and scientific practice lies in the uncertainty that pervades both and the near-certainty that hopefully emerges at the end, the jurors oscillating as the evidence is presented. Indeed, it is striking that the law lightly uses the same term 'probability', as does the scientist, for expressing the uncertainty, for example in the phrase 'the balance of probabilities'.

The case for using the scientific method in a court of law therefore looks promising because the ingredients, in a modified form, are already present. Indeed, the method has had limited use, often by people who did not realise that they were acting in the scientific spirit, but it is only in the second half of the twentieth century that the method has been widely and successfully adopted. The major advances have taken place where the evidence itself is of the form that a scientist would recognise as material for study. Examples are evidence in the form of laboratory measurements on fragments of glass or types of blood, and, more recently and dramatically, DNA data. This book tells the story of this intrusion of science into law and, more importantly, provides the necessary machinery that enables the transition to be effected.

It has been explained how uncertainty plays important roles in both scientific method and courtroom procedure. It is now recognised that the only tool for handling uncertainty is probability, so it is inevitable that probability is to be found on almost every page of this book and must have a role in the courtroom. There are two aspects to probability: firstly, the purely mathematical rules and their manipulation; secondly, the interpretation of probability, that is, the connection between the numbers and the reality. Fortunately for legal applications, the mathematics is mostly rather simple – essentially it is a matter of appreciating the language and the notation, together with the use of the main tool, Bayes' theorem. Lawyers will be familiar with the need for specialist terms, so-called jargon, and will hopefully be appreciative of the need for a little mathematical jargon. Interpretation of probability is a more delicate issue, and difficulties here are experienced both by forensic scientists and lawyers. One miscarriage of justice was influenced by a scientist's flawed use of probability; another by a lack of legal appreciation of Bayes' theorem. A great strength of this book lies in the clear recognition of the interpretative problem and the inclusion of many examples of court cases, for example in Chapter 7. My personal view is

that these problems are reduced by proper use of the mathematical notation and language, and by insisting that every statement of uncertainty is in the form of your probability of something, given clearly-stated assumptions. Thus the probability of the blood match, given that the defendant is innocent. Language that departs from this format can often lead to confusion.

The first edition (1995) of this book gave an admirable account of the subject as it was almost a decade ago. The current edition is much larger, and the enlargement reflects both the success of forensic science, by including recent cases, and also the new methods that have been used. A problem that arises in a courtroom, affecting both lawyers, witnesses and jurors, is that several pieces of evidence have to be put together before a reasoned judgement can be reached; as when motive has to be considered along with material evidence. Probability is designed to effect such combinations but the accumulation of simple rules can produce complicated procedures. Methods of handling sets of evidence have been developed; for example, Bayes nets in Chapter 14 and multivariate methods in Chapter 11. There is a fascinating interplay here between the lawyer and the scientist where they can learn from each other and develop tools that significantly assist in the production of a better judicial system. Another indication of the progress that has been made in a decade is the doubling in the size of the bibliography. There can be no doubt that the appreciation of some evidence in a court of law has been greatly enhanced by the sound use of statistical ideas, and one can be confident that the next decade will see further developments, during which time this book will admirably serve those who have cause to use statistics in forensic science.

D.V. Lindley
January 2004

Preface to the First Edition

In 1977 a paper by Dennis Lindley was published in *Biometrika* with the simple title 'A problem in forensic science'. Using an example based on the refractive indices of glass fragments, Lindley described a method for the evaluation of evidence which combined the two requirements of the forensic scientist, those of comparison and significance, into one statistic with a satisfactorily intuitive interpretation. Not unnaturally the method attracted considerable interest amongst statisticians and forensic scientists interested in seeking good ways of quantifying their evidence. Since then, the methodology and underlying ideas have been developed and extended in theory and application into many areas. These ideas, often with diverse terminology, have been scattered throughout many journals in statistics and forensic science and, with the advent of DNA profiling, in genetics. It is one of the aims of this book to bring these scattered ideas together and, in so doing, to provide a coherent approach to the evaluation of evidence.

The evidence to be evaluated is of a particular kind, known as transfer evidence, or sometimes trace evidence. It is evidence which is transferred between the scene of a crime and a criminal. It takes the form of traces – traces of DNA, traces of blood, of glass, of fibres, of cat hairs and so on. It is amenable to statistical analyses because data are available to assist in the assessment of variability. Assessments of other kinds of evidence, for example, eyewitness evidence, is not discussed.

The approach described in this book is based on the determination of a so-called likelihood ratio. This is a ratio of two probabilities, the probability of the evidence under two competing hypotheses. These hypotheses may be that the defendant is guilty and that he is innocent. Other hypotheses may be more suitable in certain circumstances and various of these are mentioned as appropriate throughout the book.

There are broader connections between statistics and matters forensic which could perhaps be covered by the title 'forensic statistics' and which are not covered here, except briefly. These might include the determination of a probability of guilt, both in the dicta 'innocent until proven guilty' and 'guilty

beyond reasonable doubt'. Also, the role of statistical experts as expert witnesses presenting statistical assessments of data or as consultants preparing analyses for counsel is not discussed, nor is the possible involvement of statisticians as independent court assessors. A brief review of books on these other areas in the interface of statistics and the law is given in Chapter 1. There have also been two conferences on forensic statistics (Aitken, 1991, and Kaye, 1993a) with a third to be held in Edinburgh in 1996. These have included forensic science within their programme but have extended beyond this. Papers have also been presented and discussion sessions held at other conferences (e.g., Aitken, 1993, and Fienberg and Finkelstein, 1996).

The role of uncertainty in forensic science is discussed in Chapter 1. The main theme of the book is that the evaluation of evidence is best achieved through consideration of the likelihood ratio. The justification for this and the derivation of the general result is given in Chapter 2. A correct understanding of variation is required in order to derive expressions for the likelihood ratio and variation is the theme for Chapter 3 where statistical models are given for both discrete and continuous data. A review of other ways of evaluating evidence is given in Chapter 4. However, no other appears, to the author at least, to have the same appeal, both mathematically and forensically as the likelihood ratio and the remainder of the book is concerned with applications of the ratio to various forensic science problems. In Chapter 5, transfer evidence is discussed with particular emphasis on the importance of the direction of transfer, whether from the scene of the crime to the criminal or *vice versa*. Chapters 6 and 7 discuss examples for discrete and continuous data, respectively. The final chapter, Chapter 8, is devoted to a review of DNA profiling, though, given the continuing amount of work on the subject, it is of necessity brief and almost certainly not completely up to date at the time of publication.

In keeping with the theme of the Series, *Statistics in Practice*, the book is intended for forensic scientists as well as statisticians. Forensic scientists may find some of the technical details rather too complicated. A complete understanding of these is, to a large extent, unnecessary if all that is required is an ability to implement the results. Technical details in Chapters 7 and 8 have been placed in Appendices to these chapters so as not to interrupt the flow of the text. Statisticians may, in their turn, find some of the theory, for example in Chapter 1, rather elementary and, if this is the case, then they should feel free to skip over this and move on to the more technical parts of the later chapters.

The role of statistics in forensic science is continuing to increase. This is partly because of the debate continuing over DNA profiling which looks as if it will carry on into the foreseeable future. The increase is also because of increasing research by forensic scientists into areas such as transfer and persistence and because of increasing numbers of data sets. Incorporation of subjective probabilities will also increase, particularly through the role of Bayesian belief networks (Aitken and Gammerman, 1989) and knowledge-based systems (Buckleton and Walsh, 1991; Evett, 1993b).

Ian Evett and Dennis Lindley have been at the forefront of research in this area for many years. They have given me invaluable help throughout this time. Both made extremely helpful comments on earlier versions of the book for which I am grateful. I thank Hazel Easey for the assistance she gave with the production of the results in Chapter 8. I am grateful to Ian Evett also for making available the data in Table 7.3. Thanks are due to The University of Edinburgh for granting leave of absence and to my colleagues of the Department of Mathematics and Statistics in particular for shouldering the extra burdens such leave of absence by others entails. I thank also Vic Barnett, the Editor of the Series and the staff of John Wiley and Sons, Ltd for their help throughout the gestation period of this book.

Last, but by no means least, I thank my family for their support and encouragement.

Preface to the Second Edition

In the Preface to the first edition of this book it was commented that the role of statistics in forensic science was continuing to increase and that this was partly because of the debate continuing over DNA profiling which looked as if it would carry on into the foreseeable future. It now appears that the increase is continuing and perhaps at a greater rate than in 1995. The debate over DNA profiling continues unabated. We have left the minutiae of this debate to others, restricting ourselves to an overview of that particular topic. Instead, we elaborate on the many other areas in forensic science in which statistics can play a role.

There has been a tremendous expansion in the work in forensic statistics in the nine years since the first edition of this book was published. This is reflected in the increase in the size of the book. There are about 500 pages now, whereas there were only about 250 in 1995, and the bibliography has increased from 10 pages to 20 pages. The number of chapters has increased from 8 to 14. The title remains the same, yet there is more discussion of interpretation, in addition to new material on evaluation.

The first four chapters are on the same topics as in the first edition, though the order of Chapters 2 and 3 on evaluation and on variation has been reversed. The chapter on variation, the new Chapter 2, has been expanded to include many more probability distributions than mentioned in the first edition. As the subject has expanded so has the need for the use of more distributions. These have to be introduced sooner than before, hence the reversal of order with the chapter on evaluation. Chapter 4 has an additional section on the work of early twentieth-century forensic scientists as it has gradually emerged how far ahead of their time these scientists were in their ideas. Three new chapters have then been introduced before the chapter on transfer evidence. Bayesian inference has an increasing role to play in the evaluation of evidence, yet its use is still controversial and there have been some critical comments in the courts of some of its perceived uses in the legal process. Chapter 5 provides a discussion of Bayesian inference, somewhat separate from the main thrust of the book, in order to emphasise its particular relevance for evidence evaluation and interpretation. Appropriate sampling procedures are becoming ever more important. With