Weight-of-evidence for Forensic DNA Profiles

David J. Balding

Professor of Statistical Genetics
Imperial College,
London, UK

John Wiley & Sons, Ltd

Weight-of-evidence for Forensic DNA Profiles

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. The subject areas covered include medicine and pharmaceutics; 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. The 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.

Weight-of-evidence for Forensic DNA Profiles

David J. Balding

Professor of Statistical Genetics
Imperial College,
London, UK

John Wiley & Sons, Ltd

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

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

Weight-of-evidence for forensic DNA profiles / David J. Balding.
p. cm. – (Statistics in practice)
Includes bibliographical references and index.
ISBN 0-470-86764-7 (alk. paper)
1. Forensic genetics – Statistical methods. I. Balding, D. J. II. Series.

RA1057.5.A875 2005
614′.1′0727 – dc22

2004061587

British Library Cataloguing in Publication Data

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

ISBN 0-470-86764-7

Produced from LaTeX files supplied by the authors and processed by Laserwords Private Limited, Chennai, India
Printed and bound in Great Britain by Antony Rowe Ltd, Chippenham, Wiltshire
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 Emily Laura

Contents

Sections marked † are at a higher level of specialization and/or are peripheral topics not required to follow the remainder of the book. They may be skipped at first reading.

Preface **xi**

1 Introduction **1**
- 1.1 Weight-of-evidence theory . . . 1
- 1.2 About the book . . . 3
- 1.3 DNA profiling technology . . . 3
- 1.4 What you need to know already . . . 4
- 1.5 Other resources . . . 5

2 Crime on an island **7**
- 2.1 Warm-up examples . . . 7
  - 2.1.1 Disease testing: Positive Predictive Value (PPV) . . . 7
  - 2.1.2 Coloured taxis . . . 9
- 2.2 Rare trait identification evidence . . . 10
  - 2.2.1 The "island" problem . . . 10
  - 2.2.2 A first lesson from the island problem . . . 11
- 2.3 Making the island problem more realistic . . . 13
  - 2.3.1 Uncertainty about p . . . 14
  - 2.3.2 Uncertainty about N . . . 15
  - 2.3.3 Possible typing errors . . . 15
  - 2.3.4 Searches . . . 17
  - 2.3.5 Other evidence . . . 18
  - 2.3.6 Relatives and population subdivision . . . 19
- 2.4 Weight-of-evidence exercises . . . 20

3 Assessing evidence via likelihood ratios **22**
- 3.1 Likelihood ratios . . . 22
- 3.2 The weight-of-evidence formula . . . 24
  - 3.2.1 Application to the island problem . . . 25
  - 3.2.2 The population $\mathcal{P}$ . . . 25

3.3 General application of the formula . . . 27
3.3.1 Several items of evidence . . . 27
3.3.2 Assessing all the evidence . . . 29
3.3.3 The role of the expert witness . . . 30
3.4 Consequences for DNA evidence . . . 31
3.4.1 Many possible culprits . . . 31
3.4.2 Incorporating the non-DNA evidence . . . 31
3.4.3 Relatives . . . 33
3.4.4 Laboratory and handling errors . . . 34
3.4.5 Database searches . . . 35
3.5 Some derivations † . . . 36
3.5.1 Bayes theorem for identification evidence . . . 37
3.5.2 Uncertainty about p and N . . . 38
3.5.3 Grouping the alternative possible culprits . . . 39
3.5.4 Typing errors . . . 40
3.6 Further weight-of-evidence exercises . . . 40

4 Typing technologies 43
4.1 STR typing . . . 44
4.1.1 Anomalies . . . 46
4.1.2 Contamination . . . 49
4.1.3 Low copy number (LCN) profiling . . . 50
4.2 mtDNA typing . . . 50
4.3 Y-chromosome markers . . . 51
4.4 X-chromosome markers † . . . 52
4.5 SNP profiles . . . 53
4.6 Fingerprints † . . . 54

5 Some population genetics for DNA evidence 56
5.1 A brief overview . . . 56
5.1.1 Drift . . . 56
5.1.2 Mutation . . . 59
5.1.3 Migration . . . 60
5.1.4 Selection . . . 60
5.2 θ, or F_{ST} . . . 62
5.3 A statistical model and sampling formula . . . 63
5.3.1 Diallelic loci . . . 63
5.3.2 Multi-allelic loci . . . 68
5.4 Hardy–Weinberg equilibrium . . . 69
5.4.1 Testing for deviations from HWE † . . . 70
5.4.2 Interpretation of test results . . . 74
5.5 Linkage equilibrium . . . 75
5.6 Coancestry † . . . 77
5.7 Likelihood-based estimation of θ † . . . 79
5.8 Population genetics exercises . . . 81

6 Identification **82**
6.1 Choosing the hypotheses . . . 82
6.1.1 Post-data equivalence of hypotheses . . . 84
6.2 Calculating likelihood ratios . . . 85
6.2.1 The match probability . . . 85
6.2.2 One locus . . . 87
6.2.3 Multiple loci: the "product rule" . . . 89
6.2.4 Relatives of s . . . 90
6.2.5 Confidence limits † . . . 92
6.2.6 Other profiled individuals . . . 93
6.3 Application to STR profiles . . . 94
6.3.1 Values for the p_j . . . 95
6.3.2 The value of θ . . . 96
6.3.3 Errors . . . 98
6.4 Application to haploid profiles . . . 99
6.4.1 mtDNA profiles . . . 99
6.4.2 Y-chromosome markers . . . 101
6.5 Mixtures . . . 101
6.5.1 Visual interpretation of mixed profiles . . . 101
6.5.2 Likelihood ratios under qualitative interpretation . . . 103
6.5.3 Quantitative interpretation of mixtures . . . 108
6.6 Identification exercises . . . 109

7 Relatedness **111**
7.1 Paternity . . . 111
7.1.1 Weight of evidence for paternity . . . 111
7.1.2 Prior probabilities . . . 112
7.1.3 Calculating likelihood ratios . . . 113
7.1.4 Multiple loci: the effect of linkage . . . 117
7.1.5 s may be related to c but is not the father . . . 119
7.1.6 Incest . . . 120
7.1.7 Mother unavailable . . . 121
7.1.8 Mutation . . . 122
7.2 Other relatedness between two individuals . . . 126
7.2.1 Only the two individuals profiled . . . 126
7.2.2 Profiled individual close relative of target . . . 127
7.2.3 Profiles of known relatives also available † . . . 128
7.3 Software for relatedness analyses . . . 129
7.4 Inference of ethnicity or phenotype † . . . 131
7.5 Relatedness exercises . . . 133

8 Other approaches to weight of evidence **135**
8.1 Uniqueness . . . 135
8.1.1 Analysis . . . 136
8.1.2 Discussion . . . 138

8.2 Inclusion/exclusion probabilities . . . 138
8.2.1 Random man . . . 138
8.2.2 Inclusion probability of a typing system . . . 139
8.2.3 Case-specific inclusion probability . . . 139
8.3 Hypothesis testing † . . . 141
8.4 Other exercises . . . 143

9 Issues for the courtroom **145**
9.1 Bayesian reasoning in court . . . 145
9.2 Some fallacies . . . 146
9.2.1 The prosecutor's fallacy . . . 146
9.2.2 The defendant's fallacy . . . 147
9.2.3 The uniqueness fallacy . . . 148
9.3 Some UK appeal cases . . . 148
9.3.1 Deen (1993) . . . 148
9.3.2 Dalby (1995) . . . 149
9.3.3 Adams (1996) . . . 149
9.3.4 Doheny/Adams (1996) . . . 151
9.3.5 Watters (2000) . . . 153
9.4 US National Research Council reports . . . 154
9.5 Prosecutor's fallacy exercises . . . 155

10 Solutions to exercises **157**

Bibliography **175**

Index **183**

Preface

Thanks are due to Kathryn Firth for drawing the figure on page 12 and to Karen Ayres and Renuka Sornarajah for providing helpful comments on a draft of the book. Lianne Mayor contributed some of the material for Section 7.2.3. Discussions with many colleagues and friends over more than 10 years have contributed to the ideas in this book: many forensic scientists have helped me towards some understanding of the laboratory techniques, and I thank Peter Donnelly for his stimulating comments on, suggestions about, and criticisms of my statistical ideas during the formative years of my interest in the field.

I am grateful to John Buckleton and Colin Aitken for sending me pre-publication manuscripts of Buckleton *et al*. (2004) and Aitken and Taroni (2004), respectively; at the time of writing, the published versions of these works had not appeared.

The statistical figures in this book were produced using `R`, a software package for statistical analysis and graphical display that has been developed by some of the world's leading statisticians. `R` is freely available for multiple platforms, with documentation, at `www.r-project.org`. Other diagrams have been created by the author using `xfig`, interactive drawing freeware that runs under the X Window System on most UNIX-compatible platforms, available at `www.xfig.org`.

1

Introduction

1.1 Weight-of-evidence theory

The introduction of DNA evidence at the end of the 1980s was rightly heralded as a breakthrough for criminal justice, but it had something of a "baptism of fire". In the media and in courts there was substantial controversy over the validity of the technology and the appropriate interpretation of the evidence.

DNA profiling technology has advanced impressively since then, and understanding by lawyers and forensic scientists of the appropriate methods for *interpreting* DNA evidence has also generally improved. Consequently, disputes about the accuracy and reliability of DNA evidence, and about its interpretation, have diminished in number and volume. However, the potential for crucial mistakes and misunderstandings remains. Although DNA evidence is typically very powerful, the circumstances under which it might not lead to satisfactory conclusions about identification or relatedness are not widely appreciated.

The primary goal of this book is to help equip a forensic scientist charged with presenting DNA evidence in court with guiding principles and technical knowledge for

- the preparation of statements that are fair, clear, and helpful to courts, and
- responding to questioning by judges and lawyers.

The prototype application is identification of the (single) culprit whose DNA profile was recovered from a crime scene, but we will also discuss profiles with multiple contributors, as well as paternity and other relatedness testing. The latter arise in both criminal and civil cases, as well as in the identification of human remains. We assume the setting of the US, UK, and Commonwealth legal systems in which decisions on guilt or innocence in criminal cases are made by lay juries, but the general principles should apply to any legal system.

Weight-of-evidence for Forensic DNA Profiles David Balding
 ISBN: 0-470-86764-7

We will introduce and develop a weight-of-evidence theory based on two key tenets:

1. The central question in a criminal trial is whether the defendant is guilty.

2. Evidence is of value inasmuch as it alters the probability that the defendant is guilty.

Although these tenets may seem self-evident, it is surprising how often they are violated. Focussing on the right questions clarifies much of the confusion that has surrounded DNA evidence in the past.

It follows from our tenets that evidential weight can be measured by likelihood ratios and combined to assess the totality of the evidence using the appropriate version of Bayes Theorem. We will discuss how to use this theory in evaluating evidence and give principles for, and examples of, calculating likelihood ratios, including taking into account relevant population genetic factors.

No theory ever describes the real world perfectly, and the analysis of forensic DNA profiles is a complex topic. It follows that the theory developed in this book cannot be applied in a naive, formulaic way to the practical situation faced by lawyers and forensic scientists in court. Nevertheless, a firm grounding in the principles of the theory provides

- grounds for deciding what information a clear-thinking juror needs in order to understand the strength of DNA profile evidence;

- the means to detect and thus avoid serious errors;

- a basis for assessing approximations and simplifications that might be useful in court;

- a framework for deciding how to proceed when the case has unusual features.

Fortunately, we will see that the mathematical aspects of the theory are not too hard. Of course, assessing some of the relevant probabilities – such as the probability that a sample handling error has occurred – can be difficult in practice, reflecting the real-world complexity of the problem. Further complications can arise for example in the case of mixed DNA samples (Section 6.5). However, the same simple rules and principles can give useful guidance in even the most complex settings.

Universal agreement is rare in the academic world, and there exist alternative theories of weight of evidence based on, for example, belief functions or fuzzy sets, rather than probabilities. The theory presented here is the most widely accepted, and its philosophical underpinnings are compelling (Bernardo and Smith 1994; Good 1991). It follows that whatever is actually said in court in connection with DNA evidence should not conflict with this theory.

There has been debate about the appropriateness in court of using numbers to measure weight of evidence. We only touch on this argument here (Section 6.3.3).

It is currently almost a universal practice to accompany DNA evidence by some sort of numbers to try to measure its weight, and so we focus here on issues such as which numbers are most appropriate in court and how they should be presented.

1.2 About the book

Chapters 2, 3, and 9 are not scientifically technical and, for the most part, are not specific to DNA evidence. I therefore hope that lawyers dealing with scientific evidence, and forensic scientists not principally concerned with DNA evidence, will also find at least these chapters to be useful. Courtroom lawyers ignorant of the weight-of-evidence theory described in Chapters 2 and 3 should be as rare as theatre critics ignorant of Shakespeare, yet, in reality, I suspect that few are able to command its elegance, power, and practical utility.

I first set out the weight-of-evidence theory informally, via a simplified model problem (Chapter 2) and then more formally using likelihood ratios (Chapter 3). In Chapter 4, we briefly survey DNA-based typing technologies, starting with an introduction to autosomal[1] STR typing, emphasizing possibilities for typing error, then moving on to other DNA typing systems, and finishing with a brief digression to discuss fingerprint evidence. Next, we survey some population genetics theory relevant to DNA profile evidence (Chapter 5). These two chapters prepare us for calculating likelihood ratios for DNA evidence, which is covered in Chapters 6 (identification) and 7 (relatedness). In Chapter 8, we discuss some alternative probability-based approaches for assessing evidential strength: none of these methods is recommended but each has its merits, which should be understood and appreciated. In Chapter 9, I draw together ideas from the previous chapters and bring them to bear on the problem of conveying effectively, clearly, and fairly the weight of the DNA profile evidence to the court. To this end, we discuss some basic fallacies and briefly review the opinions of some UK and US legal and scientific authorities.

1.3 DNA profiling technology

For the most part, we will assume that the DNA evidence is summarized for reporting purposes as the lengths of short tandem repeat (STR) alleles at multiple (perhaps 10 or more) autosomal loci. The final result at four of the loci might be reported as

STR locus:	D18	D21	THO1	D8
Genotype:	14, 16	28, 31	9·3, 9·3	10, 13

in which each pair of numbers at a locus indicates the number of repeat units in the individual's two alleles. Although whole repeats are the norm, partial repeats

[1] The nuclear chromosomes excluding X and Y.

sometimes occur (Section 4.1.1); the profile represented here is homozygous for a THO1 allele that includes a partial repeat.

STRs now form the standard DNA typing technology in many countries. Currently in routine use in the United Kingdom and several other countries is an 11-locus system, including the sex-identifying locus Amelogenin, developed by the Forensic Science Service and known as *SGMplus*. CODIS is a 13-locus system developed by the FBI and widely used in the United States. The two systems are similar and indeed have eight loci in common (see Buckleton *et al*. 2004). PowerPlex® is a commercially available 16-locus STR typing system that contains the 13 CODIS loci, two pentanucleotide repeat loci, and Amelogenin.

The process of typing STR profiles is introduced in Section 4.1 but is not covered in great depth in this book. For further details emphasizing the CODIS system, see Butler (2001), and for a UK perspective, see Gill (2002). Rudin and Inman (2002) gives a general introduction both to technical and interpretation issues. Although we emphasize STR profiles, the principles emphasized below apply equally to any DNA profiling system. Interpreting profiles from the haploid parts of the human genome (the Y and mitochondrial chromosomes) raises special difficulties. These systems are introduced in Sections 4.2 and 4.3, and interpretation issues specific to them are discussed briefly in Section 6.4. In Section 4.5, we briefly discuss profiles based on single-nucleotide polymorphism (SNP) markers.

1.4 What you need to know already

Chapters 2, 3, and 9 have essentially no technical prerequisites. To follow Chapter 5, you should know already what an STR profile is, have a rudimentary genetics vocabulary (locus, allele, etc.), and know the basic ideas of Mendelian inheritance. In statistics, you should be familiar at least with the theory of the error in a sample estimate of a population proportion (binomial distribution). The reader with experience in calculating with probabilities will be at an advantage in Chapters 6 and 7, but few technical tools are required from probability theory. In Sections 5.4.1 and 8.3, familiarity with statistical hypothesis testing is assumed, but these sections are labelled with a †, which means that they can be skipped without adverse impact on your understanding of the remainder of the book. The sampling formula (5.16) will at first seem daunting to those without a mathematical background, but the simpler recursive form (5.6) can always be used to build up more complex formulas sequentially. I give examples of its use, which requires only an ability to add and to multiply, and with practice anyone should be able to use it without difficulty.

I do not provide a general introduction to statistics (for an introduction in forensic settings, see Aitken and Taroni 2004) and give only a brief introduction to population genetics (Section 5.1). I strongly believe that many complications and much confusion have arisen unnecessarily in connection with DNA evidence because of a failure to grasp the basic principles of assessing evidential weight. If one focusses on the questions directly relevant to the forensic use of DNA profiles,

the number of ideas and techniques needed from statistics and population genetics will be small.

While the central ideas are not very difficult, inevitably there are special cases with their unique complexities. In addition, new ideas always take some time to absorb. Given some effort, this book should equip you with the basic principles for tackling any problem of interpreting forensic DNA evidence. The details of complex scenarios involving, for example, mixed profiles with missing alleles, will never be straightforward, and no book can replace the need for intelligence, care, and judgement on the part of the forensic scientist. The goal of this book is to complement these with some technical information and bring them to bear on the appropriate questions with guiding principles for assessing weight of evidence.

1.5 Other resources

Part of the reason for writing the book is to synthesize and extend in a coherent manner previous contributions to the forensic science and related literature by myself and co-authors. In particular, Chapter 3 is a development of Balding (2000), Section 7.1 extends the paternity section of Balding and Nichols (1995), and Section 8.1 is based on Balding (1999). Perhaps the most important feature of the book is the introduction of the population genetics sampling formula (Section 5.3) and its systematic application to various identification and relatedness problems. This draws in part on Balding and Donnelly (1995a), Balding and Nichols (1995), and Balding (2003) but some of the development is new here.

There are several other books that deal with the statistical interpretation of DNA and other evidence. Aitken (1995), soon to be superseded by Aitken and Taroni (2004), gives a thorough introduction to the statistical interpretation of scientific evidence in general, including DNA evidence among other evidence types. Robertson and Vignaux (1995) also deals with a range of evidence types and emphasizes interpretation issues from a lawyer's perspective, giving less attention to technical scientific aspects; for example, it does not discuss population genetics. Evett and Weir (1998) is perhaps closest to the present work, but the treatment of population genetics issues by these authors is very different from mine, as is their approach to introducing the relevant statistical issues. At the time of writing, Buckleton *et al.* (2004) is about to appear, offering a more extensive treatment of the interpretation issues raised by STR profile evidence.

As far as I can see, there is no major philosophical difference between myself and these authors: we all embrace the use of likelihood ratios and Bayes Theorem to evaluate evidence. We emphasize different aspects according to our individual perspectives, experience, and target audiences. The present book develops the weight-of-evidence theory in general and from an introductory level, and its approach to population genetics issues is unique, while remaining concisely focussed on DNA profile evidence, without extensive related material.

The December 2003 issue of *International Statistical Review* includes a series of papers dealing with the statistical interpretation of legal evidence, including a

review of the interpretation of DNA evidence from a UK-based, historical perspective (Foreman *et al*. 2003). Other useful references, presented from a somewhat distinct viewpoint, are Kaye and Sensabaugh (2000, 2002). A widely used reference that has much useful background material but also, in my opinion, important flaws, is National Research Council (1996); my criticisms of it are outlined in Section 9.4. Charles Brenner's "Forensic Mathematics" website `dna-view.com` is a rich source of information and discussion, some of which are summarized in an encyclopedia article (Brenner 2003). Weir (2003) offers a more extensive, one-chapter summary of many issues.