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 to the second edition xi

Preface to the first edition xiii

1 Introduction 1
1.1 Weight-of-evidence theory 1
1.2 About the book 3
1.3 DNA profiling technology 4
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 *People* v. *Collins* (California, 1968) 7
2.1.2 Disease testing: positive predictive value (PPV) 8
2.1.3 Coloured taxis 10
2.2 Rare trait identification evidence 10
2.2.1 The 'island' problem 11
2.2.2 A first lesson from the island problem 12
2.3 Making the island problem more realistic 13
2.3.1 The effect of uncertainty about p 13
2.3.2 Uncertainty about N 14
2.3.3 The effect of possible typing errors 15
2.3.4 The effect of searches 16
2.3.5 The effect of other evidence 17
2.3.6 The effects of relatives and population subdivision 18
2.4 Weight-of-evidence exercises 19

3 Assessing evidence using likelihoods 21
3.1 Likelihoods and their ratios 21
3.2 The weight-of-evidence formula 23
3.2.1 The population $\mathcal{P}$ 23

3.2.2 Grouping the R_X 24
3.2.3 Application to the island problem 25
3.3 General application of the formula 25
3.3.1 Several items of evidence 25
3.3.2 The role of the expert witness 28
3.4 Consequences for DNA evidence 28
3.4.1 Many possible culprits 28
3.4.2 Incorporating the non-DNA evidence 30
3.4.3 Relatives 31
3.4.4 Laboratory and handling errors 31
3.4.5 Database searches 33
3.5 Derivation of the weight-of-evidence formula † 35
3.5.1 Bayes' theorem 35
3.5.2 Uncertainty about p and N 35
3.5.3 Grouping the alternative possible culprits 37
3.5.4 Typing errors 37
3.6 Further weight-of-evidence exercises 38

4 Profiling technologies 41
4.1 STR typing 41
4.1.1 Anomalies 44
4.1.2 Contamination 47
4.1.3 Low-template DNA (LTDNA) profiling 47
4.2 mtDNA typing 48
4.3 Y-chromosome markers 49
4.4 X-chromosome markers † 50
4.5 SNP profiles † 50
4.6 Sequencing † 52
4.7 Methylation † 52
4.8 RNA † 53
4.9 Fingerprints † 53

5 Some population genetics for DNA evidence 55
5.1 A brief overview 55
5.1.1 Drift 55
5.1.2 Mutation 58
5.1.3 Migration 59
5.1.4 Selection 59
5.2 F_{ST} or θ 61
5.2.1 Population genotype probabilities 62
5.3 A statistical model and sampling formula 63
5.3.1 Diallelic loci 63
5.3.2 Multi-allelic loci 66
5.4 Hardy–Weinberg equilibrium 68
5.4.1 Testing for deviations from HWE † 69
5.4.2 Interpretation of test results 73

5.5 Linkage equilibrium 73
5.6 Coancestry † 75
5.6.1 One allele 75
5.6.2 Two alleles 75
5.6.3 Three alleles 76
5.6.4 General proof via recursion 76
5.7 Likelihood-based estimation of F_{ST} † 77
5.8 Population genetics exercises 78

6 Inferences of identity **80**
6.1 Choosing the hypotheses 80
6.1.1 Post-data equivalence of hypotheses 82
6.2 Calculating LRs 83
6.2.1 The match probability 83
6.2.2 Single locus 85
6.2.3 Multiple loci: the 'product rule' 86
6.2.4 Relatives of Q 88
6.2.5 Confidence limits † 90
6.2.6 Other profiled individuals 91
6.3 Application to STR profiles 92
6.3.1 Values for the p_j 92
6.3.2 The value of F_{ST} 94
6.3.3 Choice of population 95
6.3.4 Errors 95
6.4 Application to haploid profiles 96
6.4.1 mtDNA profiles 96
6.4.2 Y-chromosome markers 98
6.5 Mixtures 98
6.5.1 Visual interpretation of mixed profiles 98
6.5.2 Likelihood ratios under qualitative interpretation 99
6.5.3 Quantitative interpretation of mixtures 105
6.6 Identification exercises 106

7 Inferring relatedness **108**
7.1 Paternity 108
7.1.1 Weight of evidence for paternity 108
7.1.2 Prior probabilities 109
7.1.3 Calculating LRs 110
7.1.4 Multiple loci: the effect of linkage 114
7.1.5 Q may be related to c but not the father 115
7.1.6 Incest 116
7.1.7 Mother unavailable 117
7.1.8 Mutation 117
7.2 Other relatedness between two individuals 121
7.2.1 Only the two individuals profiled 121

7.2.2 Profiles of known relatives also available † 122
7.2.3 Software for relatedness analyses 123
7.3 Familial search 125
7.4 Inference of ethnicity † 126
7.5 Inference of phenotype † 128
7.6 Relatedness exercises 128

8 Low-template DNA profiles 130
8.1 Background 130
8.2 Stochastic effects in LTDNA profiles 132
8.2.1 Drop-out 132
8.2.2 Drop-in 133
8.2.3 Peak imbalance 133
8.2.4 Stutter 134
8.3 Computing likelihoods 134
8.3.1 Single contributor allowing for drop-out 134
8.3.2 Profiled contributors not subject to drop-out 135
8.3.3 Modelling drop-in 136
8.3.4 Multi-dose drop-out and degradation 136
8.3.5 Additional contributors subject to drop-out 138
8.3.6 Replicates 138
8.3.7 Using peak heights 139
8.4 Quality of results 141

9 Introduction to `likeLTD` † 143
9.1 Installation and example `R` script 143
9.1.1 Input 144
9.1.2 Allele report 145
9.1.3 Arguments and optimisation 145
9.1.4 Output report 147
9.1.5 Genotype probabilities 148
9.2 Specifics of the package 151
9.2.1 The parameters 151
9.2.2 Key features of `likeLTD` 151
9.2.3 Maximising the penalised likelihood 152
9.2.4 Computing time and memory requirements 153
9.3 Verification 154

10 Other approaches to weight of evidence 157
10.1 Uniqueness 157
10.1.1 Analysis 158
10.1.2 Discussion 159
10.2 Inclusion/exclusion probabilities 160
10.2.1 Identification: single contributor 160
10.2.2 Identification: multiple contributors 161

10.2.3 Paternity 161
10.2.4 Discussion 162
10.3 Hypothesis testing † 162
10.4 Other exercises 163

11 Some issues for the courtroom 165
11.1 The role of the expert witness 165
11.2 Bayesian reasoning in court 166
11.3 Some fallacies 167
11.3.1 The prosecutor's fallacy 167
11.3.2 The defendant's fallacy 168
11.3.3 The uniqueness fallacy 169
11.4 Some UK appeal cases 169
11.4.1 Deen (1993) 169
11.4.2 Adams (1996) 170
11.4.3 Doheny/Adams (1996) 171
11.4.4 Watters (2000) 173
11.4.5 T (2010) 174
11.4.6 Dlugosz (2013) 175
11.5 US National Research Council reports 177
11.6 Prosecutor's fallacy exercises 178

Solutions to exercises 179

References 196

Index 211