

---

# Contents

<b>Preface</b> .....	vii
<b>List of Notations and Terminology</b> .....	xix

---

## Part I Background and Preparations

---

<b>1 Background in Statistics</b> .....	3
1.1 Introduction to R .....	4
1.2 The Binomial, Poisson, and Normal Models .....	7
1.3 Testing Hypothesis .....	9
1.3.1 The Structure of a Statistical Test of Hypotheses .....	9
1.3.2 Testing Genetic Identity by Descent of Affected Siblings	10
1.4 Limit Theorems .....	13
1.5 Testing Equality of Two Binomial Parameters .....	16
1.6 Statistical Power .....	17
1.7 Correlation and Regression .....	24
1.8 Stochastic Processes .....	28
1.9 Likelihood-Based Inference .....	29
1.10 Properties of Expectations and Variances .....	32
1.11 Bibliographical Comments .....	33
Problems .....	33
<b>2 Introduction to Experimental Genetics</b> .....	35
2.1 The Mouse Model .....	36
2.1.1 A Quantitative Trait Locus (QTL) .....	38
2.1.2 Simulation of Phenotypes .....	40
2.2 Segregation of the Trait in Crosses .....	42
2.3 Molecular Genetic Markers .....	48
2.3.1 The SNP Markers .....	49
2.3.2 The SSR Markers .....	49

2.4	Bibliographical Comments .....	50
	Problems .....	50
<b>3</b>	<b>Fundamentals of Genetics</b> .....	<b>53</b>
3.1	Inbreeding .....	53
3.1.1	Segregation of the Phenotype in a Recombinant Inbred Strain .....	55
3.1.2	Dynamic of Inbreeding at a Single Locus .....	58
3.2	The Recombination Fraction .....	62
3.3	Random Mating .....	68
3.4	Inbreeding in Infinite Populations and Identity by Descent ....	70
3.5	Bibliographical Comments .....	71
	Problems .....	71

---

**Part II Experimental Genetics**

---

<b>4</b>	<b>Testing for Linkage with a Single Marker</b> .....	<b>77</b>
4.1	Tests for the Presence of a QTL .....	77
4.1.1	Testing a QTL in the Backcross Design .....	77
4.1.2	The Noncentrality Parameter .....	82
4.1.3	The BC Test as a Test of the Regression Coefficient ...	88
4.2	The Effect of Recombination .....	88
4.2.1	The Recombination Fraction .....	89
4.2.2	The Distribution of a Test Statistic at a Linked Marker	90
4.2.3	*Covariance of (4.4) at Different Marker Loci .....	92
4.2.4	*Sufficiency and the Correlation Structure .....	93
4.3	Intercross and Other Designs .....	94
4.4	Bibliographical Comments .....	96
	Problems .....	96
<b>5</b>	<b>Whole Genome Scans: The Significance Level</b> .....	<b>99</b>
5.1	The Multi-Marker Process .....	100
5.2	Multiple Testing and the Significance Level .....	105
5.3	Mathematical Approximations of the Significance Level .....	109
5.4	Other Methods .....	114
5.5	P-values .....	116
5.6	*The Recombination Fraction in the Haldane Model .....	117
5.7	Bibliographical Comments .....	117
	Problems .....	118
<b>6</b>	<b>Statistical Power and Confidence Regions</b> .....	<b>121</b>
6.1	The Power to Detect a QTL .....	122
6.2	An Analytic Approximation of the Power .....	128
6.3	Designing an Experiment .....	132

6.4	Confidence Sets .....	135
6.5	Confidence Bounds for the Genetic Effect of a QTL .....	138
6.6	Bibliographical Remarks .....	140
	Problems .....	141
<b>7</b>	<b>Missing Data and Interval Mapping .....</b>	<b>143</b>
7.1	Missing Marker Genotypes .....	144
7.2	Interval Mapping .....	151
7.2.1	*Approximating the Process of Interval Mapping .....	152
7.2.2	Normal Approximation of the Process for Interval Mapping .....	154
7.2.3	The Statistical Properties of Interval Mapping .....	157
7.3	Bibliographical Comments .....	167
	Problems .....	167
<b>8</b>	<b>Advanced Topics .....</b>	<b>169</b>
8.1	Multivariate Phenotypes, Gene-Environment Interaction, and Longitudinal Studies .....	169
8.1.1	Multivariate Phenotypes .....	169
8.1.2	Gene-Covariate Interaction .....	171
8.1.3	Longitudinal Studies .....	172
8.2	Multiple Gene Models and Gene-Gene Interaction .....	173
8.2.1	Strategies for Detecting Multiple Genes .....	173
8.2.2	Multiple Regression and Model Selection .....	174
8.2.3	Sequential Detection .....	176
8.3	Selective Genotyping .....	177
8.4	Advanced Intercross Lines and Fine Mapping .....	178
8.5	Bibliographical Comments .....	179
	Problems .....	180

---

**Part III Human Genetics**

---

<b>9</b>	<b>Mapping Qualitative Traits in Humans Using Affected Sib Pairs .....</b>	<b>185</b>
9.1	Genetic Models .....	186
9.2	IBD Probabilities at the Candidate Trait Locus .....	189
9.3	A Test for Linkage at a Single Marker Based on a Normal Approximation .....	191
9.4	Genome Scans .....	193
9.5	Parametric Methods .....	197
9.6	Estimating the Number of Alleles Shared IBD .....	200
9.6.1	Simulating Pedigrees .....	201
9.6.2	Computing the Conditional Distribution of IBD .....	206
9.6.3	Statistical Properties of Genome Scans .....	212

9.7	Bibliographical Comments	222
	Problems	222
<b>10</b>	<b>Admixture Mapping</b>	227
10.1	Testing for the Presence of a QTL	229
10.1.1	A Model for an Admixed Population	229
10.1.2	Scanning Statistics and Noncentrality Parameters	231
10.1.3	A Model for the Covariance Structure	233
10.2	Inferring Population Origin from Molecular Markers	235
10.3	Estimating Parameters of the HMM	245
10.4	Discussion	253
10.5	Bibliographical Comments	253
	Problems	253
<b>11</b>	<b>Mapping Complex and Quantitative Traits with Data from Human Pedigrees</b>	257
11.1	Model and Covariances	257
11.2	Statistics to Detect Linkage	259
11.2.1	Regression Statistic	259
11.2.2	Score Statistic	260
11.3	A Two Degree of Freedom Statistic	263
11.4	*Sibships and Larger Pedigrees	264
11.5	*Multiple Gene Models and Gene-Gene Interaction	266
11.6	Using Pedigrees to Map Disease Genes	268
11.7	Discussion	274
11.8	Bibliographical Comments	275
	Problems	276
<b>12</b>	<b>Association Studies</b>	279
12.1	Statistical Models and Statistical Tests	280
12.1.1	A Model for the Trait	281
12.1.2	The Distribution of the Genotypes Among the Cases	282
12.1.3	Statistical Tests for Association when Genotypes at a Trait Locus are Observed	283
12.1.4	The Statistical Properties of the Allele Test	285
12.1.5	The Effect of Linkage Disequilibrium	287
12.1.6	A Demonstration	289
12.2	Population Substructure and Inbreeding	291
12.2.1	A Population Substructure	292
12.2.2	Inbreeding	297
12.3	The Transmission Disequilibrium Test	301
12.4	Discussion	304
12.5	Bibliographical Comments	304
	Problems	304

**13 Inferring Haplotypes from Genotypes and Testing**

**Association** ..... 307

13.1 Determining the Phase Between Two Loci ..... 308

13.2 The Expectation-Maximization Algorithm ..... 312

13.3 Testing Association with Haplotypes ..... 318

13.4 Bibliographical Comments ..... 322

Problems ..... 322

**References** ..... 323

**Subject Index** ..... 329

**Index of R Functions** ..... 333