

research methods for the biosciences

SECOND EDITION

D. Holmes

P. Moody

D. Dine

OXFORD
UNIVERSITY PRESS

Contents

Detailed Contents	vii
Boxes outlining the methods for the statistical tests included in this book	xvii

Section 1 Planning an experiment

1 Where do I begin?	3
2 Planning your experiment	19
3 Questionnaires, focus groups, and interviews	54
4 Research, the law, and you	65

Section 2 Handling your data

5 What to do with raw data	111
6 An introduction to hypothesis testing	147
7 Hypothesis testing: do my data fit an expected ratio?	172
8 Hypothesis testing: associations and relationships	209
9 Hypothesis testing: do my samples come from the same population? Parametric data	255
10 Hypothesis testing: do my samples come from the same population? Non-parametric data	321

Section 3 Reporting your results

11 Reporting your research	361
----------------------------	-----

<i>Appendix a.</i>	<i>How to choose a research project</i>	397
<i>Appendix b.</i>	<i>Planning your experiment</i>	401
<i>Appendix c.</i>	<i>Which statistical test should I choose?</i>	409
<i>Appendix d.</i>	<i>Tables of critical values for statistical tests</i>	416
<i>Appendix e.</i>	<i>Maths and statistics</i>	434
	<i>Glossary</i>	446
	<i>References</i>	452
	<i>Index</i>	455

Detailed Contents

Section 1 Planning an experiment

1	Where do I begin?	3
1.1.	Aims and objectives	3
1.1.1.	The aim	4
1.1.2.	The objective	4
i.	Clumpers or splitters	5
ii.	Only one objective	5
iii.	Personal and experimental objectives	6
1.2.	Data, items, and observations	6
1.3.	Populations	7
1.4.	Sample	8
1.4.1.	Representative samples	8
1.4.2.	How do you obtain a representative sample?	9
i.	Random sampling	9
ii.	Systematic/periodic sampling	9
iii.	Stratified random sampling	9
iv.	Homogeneous stands	10
1.5.	Population parameters and sample statistics	11
1.5.1.	Mathematical notation for populations and samples	11
1.5.2.	Calculations	11
1.6.	Treatments	12
1.7.	Variation and variables	13
1.7.1.	Variation	13
1.7.2.	Variables	13
1.7.3.	Partitioning variation and confounding (non-treatment) variables	14
1.8.	Hypotheses	15
1.8.1.	Null hypotheses	16
1.8.2.	Information-theoretic models	17
2	Planning your experiment	19
2.1.	Evaluating published research	20
2.1.1.	What are the aim and objective(s)?	20
2.1.2.	Strengths of the experimental design	21
2.1.3.	Weaknesses of the experimental design	22
i.	Faults	22
ii.	Limitations	22
iii.	Communication	22

a.	Fault in the design	22
b.	Limitation	23
c.	Communication	23
2.2.	Have a go!	24
2.2.1.	What are my aim and objectives?	25
i.	Background information	25
a.	The context	25
b.	The justification	26
c.	The practicalities	26
d.	Previous designs	27
ii.	Aim and objectives	28
2.2.2.	What is the statistical population?	28
2.2.3.	Which variables am I investigating?	30
2.2.4.	Are there any confounding variables?	31
i.	Identifying confounding variables	31
ii.	Minimizing the effect of non-treatment variation	32
a.	Equalizing the effect	33
b.	Randomization	33
iii.	Mathematically separate out (partition) the variation in your data	34
2.2.5.	Will I need replicates?	34
i.	When should you use replicates?	34
a.	To increase the reliability of your estimate of the population parameters	34
b.	To allow mathematical estimates of non-treatment variation	34
ii.	Features of replicates	35
iii.	How many replicates?	36
2.2.6.	Will I need any controls?	36
2.2.7.	How will I analyse my data?	37
i.	Which type of hypothesis am I testing?	38
ii.	What type of data do I have?	40
iii.	How many variables do I wish to test the effect of?	41
iv.	How many categories do I have? How many observations in each category?	41
v.	Therefore, which statistical test(s) seem appropriate?	42
vi.	Finalizing your aim and objectives	44
vii.	Writing your hypothesis	44
2.2.8.	Do I need to take action to ensure that I comply with UK law?	45
2.2.9.	Are there any causes of possible bias? Have I made any assumptions?	46
i.	Influencing outcomes and causes of bias	46
ii.	Assumptions	46
2.2.10.	Will I repeat the investigation?	47
2.2.11.	Back to the beginning	48
i.	Review your planning	48
ii.	Experimental design and methods	48
2.3.	Managing research	50
2.3.1.	Time management	50
2.3.2.	Space management	51
2.3.3.	Data management	51
i.	Laboratory or field notebooks	51
ii.	Electronic records	52

3	Questionnaires, focus groups, and interviews	54
3.1.	What is a questionnaire, interview, or focus group?	55
3.1.1.	Questionnaires	55
3.1.2.	Focus groups	55
3.1.3.	Interviews	55
3.2.	Open and closed questions	56
3.2.1.	Closed questions	56
3.2.2.	Open questions	57
3.3.	Phrasing questions	57
3.4.	Your participants	59
3.5.	Sample sizes	59
3.5.1.	Closed questions	60
3.5.2.	Open questions	63
3.5.3.	Achieving the required sample size	64
4	Research, the law, and you	65
4.1.	About the law	66
4.1.1.	International law	66
4.1.2.	National law	67
i.	Common law (case law)	67
ii.	Legislation	67
iii.	Local law	68
iv.	Institutional writers	68
v.	Canon law	68
vi.	Custom	69
vii.	Quasi-legislation	69
viii.	Enforcement	69
4.2.	Health and safety	69
4.2.1.	Hazard identification and rating	70
i.	The environment	71
ii.	Chemicals	71
iii.	Biological agents	72
iv.	Rating the hazard	73
4.2.2.	Activity	73
i.	Hazards arising from the activity	74
ii.	Scale of exposure	74
iii.	Storage and disposal	74
4.2.3.	Probability of harm	75
i.	Probability of harm	75
ii.	Variation among individuals	75
iii.	Occupational exposure limits	75
4.2.4.	Minimizing the risk	76
i.	Prevention	76
ii.	Reduction	76
a.	Wear personal protection	76
b.	Reduce quantities	76
c.	Control ventilation	77
d.	Organize space	77
e.	Take a break	77
f.	Inform	77
g.	Company	78
h.	Be informed	79
iii.	Re-evaluate the probability of harm	79
4.2.5.	Risk evaluation	79

4.2.6.	Emergencies	79
i.	Failure of control measures	79
ii.	External emergencies	80
iii.	Information	80
4.2.7.	Action	80
4.3.	Access and sampling	81
4.3.1.	Access	81
i.	Access in England and Wales	81
a.	Public right of way	82
b.	Open country	82
c.	Waterways	82
d.	Foreshore	82
e.	Commons and village and town greens	83
f.	Other areas	83
ii.	Access in Northern Ireland	83
iii.	Access in Scotland	83
4.3.2.	Theft	84
4.3.3.	Plants, animals, and other organisms	85
i.	Wild birds	86
a.	General protection	86
b.	Enhanced protection	87
ii.	Wild animals	87
a.	Badgers	88
b.	Deer	88
c.	Fish	88
d.	Bats	89
iii.	Wild plants	89
a.	General protection	89
b.	Enhanced protection	90
4.3.4.	Protection in special areas	90
i.	Habitats and Natura 2000	91
ii.	Nature reserves	91
a.	Statutory local and national nature reserves	91
b.	Non-statutory nature reserves	92
iii.	National Parks, Areas of Outstanding Natural Beauty, and Sites of Special Scientific Interest	92
iv.	Ministry of Defence land and National Trust land	93
4.3.5.	Movement, import, export, and control	94
4.3.6.	Permits and licences	95
4.4.	Animal welfare	95
4.4.1.	Protection of Animals Act 1911	96
4.4.2.	Animals (Scientific Procedures) Act 1986 and Amendment, 1998	97
4.4.3.	Animal Welfare Act (2006) and Animal Health and Welfare (Scotland) Act 2006	98
4.5.	Working with humans	98
4.5.1.	Harm, risks, and benefits	100
4.5.2.	Consent	101
4.5.3.	Special cases: children and vulnerable people	102
4.5.4.	Equality	103
4.5.5.	Anonymity, confidentiality, information storage, and dissemination	104
4.6.	Discussion topics	105

Section 2 Handling your data

5	What to do with raw data	111
5.1.	Types of data	112
5.2.	Distributions of data	115
5.2.1.	Normal distribution	115
5.2.2.	Binomial distribution	117
5.2.3.	Poisson distribution	119
5.2.4.	Exponential distribution	120
5.3.	Summary statistics	121
5.4.	Estimates of the central tendency	123
5.4.1.	Mode	123
5.4.2.	Median	124
5.4.3.	Mean	124
	i. Normal distribution	125
	a. Your data are not grouped in categories	125
	b. Your data are grouped into categories	125
	ii. Binomial distribution	127
	iii. Poisson distribution	127
5.4.4.	Skew and kurtosis	128
5.5.	Estimates of variation	131
5.5.1.	Range	131
5.5.2.	Interquartile range and percentiles	132
5.5.3.	Variance	132
	i. Normal distribution	133
	a. Your data are not grouped into classes	133
	b. Your data are grouped into classes	135
	ii. Binomial distribution	136
	iii. Poisson distribution	136
5.6.	Coefficient of variation	137
5.7.	Confidence limits	137
5.8.	Parametric and non-parametric data	139
5.8.1.	Are my data parametric?	139
5.8.2.	Ranking non-parametric data	141
5.9.	Transforming data	141
6	An introduction to hypothesis testing	147
6.1.	What are the hypotheses to be tested?	148
6.1.1.	Types of hypothesis	148
	i. Do the data fit expected values?	148
	ii. Is there an association between two or more variables?	149
	a. The presence or absence of an association	149
	b. The nature of the association	149
	iii. Do samples come from the same or a different population?	149
6.1.2.	How to write hypotheses	151
6.1.3.	Why hypotheses come in pairs	153
	i. Sampling error and chance	153
	ii. Sampling error and experiments	154
	iii. What does the term 'sampling error' mean?	154
	iv. Sampling error and hypotheses	155
	v. Hypothesis testing	155

6.2. Working out the calculated value of the test statistic	156
6.2.1. What is a calculated value or test statistic?	156
i. The distributions	156
ii. Criteria for statistical tests	156
iii. Name of the test statistic	157
6.3. Finding the critical value of the test statistic	157
6.3.1. What is the critical value?	158
6.3.2. p values	158
i. Distributions and p values	159
ii. Hypotheses and p values	161
iii. Which p value?	161
iv. Critical values and p	162
6.3.3. One-tailed and two-tailed tests	163
6.3.4. Degrees of freedom	165
6.3.5. More than two criteria	167
6.3.6. Interpolation	167
6.4. The rule	169
6.5. What does this mean in real terms?	169
7 Hypothesis testing: do my data fit an expected ratio?	172
7.1. Chi-squared goodness-of-fit test	174
7.1.1. Using this test	176
7.1.2. The general calculation	177
7.1.3. How to check whether your data have a normal distribution using the chi-squared goodness-of-fit test	179
i. Arrange your data into a contingency table	179
ii. Do the data meet all the criteria for using a chi-squared goodness-of-fit test?	179
iii. Calculate the expected values	181
7.2. Heterogeneity in a goodness-of-fit test	183
7.2.1. Using this test	185
7.2.2. The calculation	185
7.3. Chi-squared test for association	187
7.3.1. Using this test	188
7.3.2. The calculation	189
7.4. Chi-squared test with one degree of freedom	191
7.4.1. Chi-squared goodness of fit test when there is one degree of freedom	192
7.4.2. Chi-squared test for association when there is only one degree of freedom	195
7.5. G -tests	198
7.5.1. G goodness-of-fit test	198
7.5.2. An $r \times c$ G -test for association	200
7.6. The problem with small numbers	202
7.7. Experimental design and chi-squared and G -tests	204
8 Hypothesis testing: associations and relationships	209
8.1. Correlations	213
8.2. Spearman's rank correlation	215
8.2.1. Using this test	215
8.2.2. The calculation	218
8.3. Pearson's product moment correlation	220
8.3.1. Using this test	220

8.3.2. The calculation	221
8.3.3. Coefficient of determination	223
8.4. Regressions	224
8.5. Model I: simple linear regression: only one y for each x	226
8.5.1. Using this test	226
8.5.2. The calculation	228
8.6. Model I: linear regression: more than one y for each value of x, with equal replicates	232
8.6.1. Using this test	234
8.6.2. The calculation	236
8.7. Model II: principal axis regression	241
8.7.1. Using this test	242
8.7.2. The calculation	243
8.8. Model II: ranged principal axis regression	246
8.8.1. Using this test	246
8.8.2. The calculation	247
8.9. About regression lines	250
8.10. Experimental design, correlation, and regression	251
9 Hypothesis testing: do my samples come from the same population? Parametric data	255
9.1. z-test for unmatched data	257
9.1.1. Using this test	258
9.1.2. The calculation	261
9.2. Unequal variance t-test for unmatched data	262
9.2.1. Using this test	263
9.2.2. The calculation	263
9.3. z- and t-tests for matched data	266
9.3.1. Using this test	266
9.3.2. The calculation	268
9.4. Introduction to parametric ANOVA	269
9.5. One-way parametric ANOVA with equal numbers of replicates	272
9.5.1. Using this test	272
9.5.2. The calculation	275
9.6. Tukey's test following a parametric one-way ANOVA with equal replicates	277
9.6.1. Using this test	277
9.6.2. The calculation	277
9.7. One-way parametric ANOVA with unequal replicates	280
9.8. Tukey-Kramer test following a parametric one-way ANOVA with unequal replicates	282
9.8.1. Using this test	283
9.8.2. The calculation	283
9.9. Two-way parametric ANOVA with equal numbers of replicates	285
9.9.1. Using this test	288
9.9.2. The calculation	289
9.10. Tukey's test following a parametric two-way ANOVA with equal replicates	292
9.11. Two-way parametric ANOVA with unequal replicates	293
9.12. Two-way parametric ANOVA with no replicates	296
9.12.1. Using this test	297
9.12.2. The calculation	298

9.13. Two-way nested parametric ANOVA with equal replicates	300
9.13.1. Using this test	302
9.13.2. The calculation	303
9.13.3. Tukey's test for nested parametric ANOVAs	305
9.14. Factorial three-way parametric ANOVA with no replicates	306
9.14.1. Using this test	307
9.14.2. The calculation	308
9.14.3. Tukey's test and a three-way parametric ANOVA	312
9.15. Factorial three-way parametric ANOVA with replicates	312
9.15.1. Using this test	313
9.15.2. The calculation	313
9.16. Experimental design and parametric statistics	314
9.16.1. One variable	314
i. Unmatched	314
ii. Matched	315
iii. Latin square	316
9.16.2. Two variables	317
i. Orthogonal and randomization	317
ii. Confounding variables and repeated measures	318
iii. Replication	318
iv. Fixed and nested	318
9.16.3. Three variables	319
9.16.4. General and specific hypotheses	319
10 Hypothesis testing: do my samples come from the same population? Non-parametric data	321
10.1. Mann–Whitney <i>U</i> test	323
10.1.1. Using this test	323
10.1.2. The calculation	324
10.2. Wilcoxon's matched pairs test	326
10.2.1. Using this test	328
10.2.2. The calculation	328
10.3. One-way non-parametric ANOVA (Kruskal–Wallis test)	330
10.3.1. Using this test	332
10.3.2. The calculation	332
10.4. Multiple comparisons following a non-parametric one-way ANOVA	335
10.4.1. Using this test	336
10.4.2. The calculation	336
10.5. Two-way non-parametric ANOVA	339
10.5.1. Using this test	339
10.5.2. The calculation	340
10.6. Multiple comparisons following a two-way non-parametric ANOVA	347
10.6.1. Using this test	349
10.6.2. The calculation	349
10.7. Scheirer–Ray–Hare test	354
10.7.1. Using this test	354
10.7.2. The calculation	354
10.8. Experimental design and non-parametric tests of hypotheses	356
10.8.1. One variable	356
i. Unmatched data	356
ii. Matched data	356
10.8.2. Two variables	357
10.8.3. Three variables	357

Section 3 Reporting your results

11	Reporting your research	361
11.1.	General format	362
11.2.	Title	363
11.3.	Abstract	364
11.4.	Keywords	365
11.5.	Acknowledgements	365
11.6.	Introduction	366
11.7.	Methods	369
11.7.1.	Site	369
11.7.2.	Methods	370
11.7.3.	Statistical analyses	372
11.8.	Results	373
11.8.1.	Tables	374
11.8.2.	Figures	378
11.8.3.	Reporting statistics	384
i.	When $p = 0.05$	384
ii.	When $p < 0.05$	384
iii.	Statistical software and the exact p value	385
iv.	Regression analysis	385
v.	Tukey's test	386
vi.	Confidence limits	386
11.9.	Discussion	386
11.10.	References	388
11.10.1.	Referencing in the text	389
11.10.2.	Reference list	389
11.11.	Appendix	393
11.12.	Your approach to writing a report	394
Appendix a.	How to choose a research project	397
Appendix b.	Planning your experiment	401
Appendix c.	Which statistical test should I choose?	409
Appendix d.	Tables of critical values for statistical tests	416
Appendix d.	Maths and statistics	434
Glossary		446
References		452
Index		455