Contents

Preface xxi
About the authors xxv
Acknowledgments xxvi
SaplingPlus for statistics xxviii
Online resources for students xxix

PART 1 INTRODUCTION TO STATISTICS

1. Statistics and samples 1
   1.1 What is statistics? 1
   1.2 Sampling populations 3
   Example 1.2: Raining cats 3
   1.3 Types of data and variables 11
   1.4 Frequency distributions and probability distributions 13
   1.5 Types of studies 15
   1.6 Summary 17
   Practice problems 18
   Assignment problems 20

INTERLEAF 1 Correlation does not require causation 24

2. Displaying data 27
   2.1 Guidelines for effective graphs 28
   2.2 Showing data for one variable 32
   Example 2.2A: Crouching tiger 33
   Example 2.2B: Effects of Zika virus infection on fetuses 36
   2.3 Showing association between two variables and differences between groups 39
   Example 2.3A: Reproductive effort and avian malaria 40
   Example 2.3B: Blood responses to high elevation 44
   2.4 Showing trends in time and space 46
   2.5 How to make good tables 48

2.6 How to make data files 50
2.7 Summary 52
Practice problems 53
Assignment problems 59

3. Describing data 65
3.1 Arithmetic mean and standard deviation 66
Example 3.1: Gliding snakes 66
3.2 Median and interquartile range 73
Example 3.2: I'd give my right arm for a female 74
3.3 How measures of location and spread compare 77
Example 3.3: Disarming fish 77
3.4 Cumulative frequency distribution 80
3.5 Proportions 82
3.6 Summary 83
3.7 Quick formula summary 84
Practice problems 85
Assignment problems 90

4. Estimating with uncertainty 97
4.1 The sampling distribution of an estimate 98
Example 4.1: The length of human genes 98
4.2 Measuring the uncertainty of an estimate 103
4.3 Confidence intervals 104
4.4 Error bars 107
4.5 Summary 109
4.6 Quick formula summary 109
Practice problems 110
Assignment problems 112

INTERLEAF 2 Pseudoreplication 117

5. Probability 119
5.1 The probability of an event 119
5.2 Venn diagrams 121
5.3 Mutually exclusive events 122
5.4 Probability distributions 122
5.5 Either this or that: adding probabilities 125

5.6 Independence and the multiplication rule 128
Example 5.6A: Smoking and high blood pressure 129
Example 5.6B: Mendel's peas 130
5.7 Probability trees 131
Example 5.7: Sex and birth order 131
5.8 Dependent events 133
Example 5.8: Is this meat taken? 133
5.9 Conditional probability and Bayes' theorem 135
Example 5.9: Detection of Down syndrome 139
5.10 Summary 140
Practice problems 141
Assignment problems 146

6. Hypothesis testing 151
6.1 Making and using statistical hypotheses 152
6.2 Hypothesis testing: an example 155
Example 6.2: The right hand of toad 155
6.3 Errors in hypothesis testing 161
6.4 When the null hypothesis is not rejected 163
Example 6.4: The genetics of mirror-image flowers 163
6.5 One-sided tests 166
6.6 Hypothesis testing versus confidence intervals 168
6.7 Summary 169
Practice problems 170
Assignment problems 172

INTERLEAF 3 Why statistical significance is not the same as biological importance 177

PART 2 PROPORTIONS AND FREQUENCIES

7. Analyzing proportions 179
7.1 The binomial distribution 180
7.2 Testing a proportion: the binomial test 185
Example 7.2: Sex and the X 185
7.3 Estimating proportions 188
Example 7.3: She-turtles 189

7.4 Deriving the binomial distribution 191
7.5 Summary 192
7.6 Quick formula summary 192
Practice problems 194
Assignment problems 198

INTERLEAF 4 Biology and the history of statistics 201

8. Fitting probability models to frequency data 203
8.1 χ^2 Goodness-of-fit test: the proportional model 204
Example 8.1: No weekend getaway 204
8.2 Assumptions of the χ^2 goodness-of-fit test 212
8.3 Goodness-of-fit tests when there are only two categories 213
Example 8.3: Gene content of the human X chromosome 213
8.4 Random in space or time: the Poisson distribution 215
Example 8.4: Mass extinctions 217
8.5 Summary 221
8.6 Quick formula summary 222
Practice problems 223
Assignment problems 225

INTERLEAF 5 Making a plan 231

9. Contingency analysis: Associations between categorical variables 233
9.1 Associating two categorical variables 234
9.2 Estimating association in 2×2 tables: relative risk 234
Example 9.2: Take two aspirin and call me in the morning? 235
9.3 Estimating association in 2×2 tables: the odds ratio 237
Example 9.3: Your litter box and your brain 241
9.4 The χ^2 contingency test 243
Example 9.4: The gnarly worm gets the bird 244
9.5 Fisher's exact test 250
Example 9.5: The feeding habits of vampire bats 250
9.6 Summary 252
9.7 Quick formula summary 253
Practice problems 254
Assignment problems 259

Review Problems 1 267

PART 3 COMPARING NUMERICAL VALUES

10. The normal distribution 273
10.1 Bell-shaped curves and the normal distribution 274
10.2 The formula for the normal distribution 276
10.3 Properties of the normal distribution 277
10.4 The standard normal distribution and statistical tables 278
Example 10.4: One small step for man? 281
10.5 The normal distribution of sample means 283
10.6 Central limit theorem 286
Example 10.6: Young adults and the Spanish flu 286
10.7 Normal approximation to the binomial distribution 289
Example 10.7: The only good bug is a dead bug 289
10.8 Summary 292
10.9 Quick formula summary 293
Practice problems 293
Assignment problems 297

INTERLEAF 6 Controls in medical studies 300

11. Inference for a normal population 303
11.1 The *t*-distribution for sample means 304
11.2 The confidence interval for the mean of a normal distribution 307
Example 11.2: Eye to eye 307
11.3 The one-sample *t*-test 310
Example 11.3: Human body temperature 310
11.4 Assumptions of the one-sample *t*-test 314
11.5 Estimating the standard deviation and variance of a normal population 314
11.6 Summary 317
11.7 Quick formula summary 317
Practice problems 318
Assignment problems 321

12. Comparing two means 327
12.1 Paired sample versus two independent samples 328
12.2 Paired comparison of means 329
Example 12.2: So macho it makes you sick? 330
12.3 Two-sample comparison of means 335
Example 12.3: Spike or be spiked 335

12.4 Using the correct sampling units 342
Example 12.4: So long; thanks to all the fish 342
12.5 The fallacy of indirect comparison 345
Example 12.5: Mommy's baby, Daddy's maybe 345
12.6 Interpreting overlap of confidence intervals 347
12.7 Comparing variances 348
12.8 Summary 349
12.9 Quick formula summary 350
Practice problems 354
Assignment problems 361

INTERLEAF 7 Which test should I use? 368

13. Handling violations of assumptions 371
13.1 Detecting deviations from normality 372
Example 13.1: The benefits of marine reserves 375
13.2 When to ignore violations of assumptions 377
13.3 Data transformations 378
13.4 Nonparametric alternatives to one-sample and paired *t*-tests 384
Example 13.4: Sexual conflict and the origin of new species 385
13.5 Comparing two groups: the Mann–Whitney *U*-test 388
Example 13.5: Sexual cannibalism in sagebrush crickets 389
13.6 Assumptions of nonparametric tests 393
13.7 Type I and Type II error rates of nonparametric methods 394
13.8 Permutation tests 395
13.9 Summary 399
13.10 Quick formula summary 400
Practice problems 402
Assignment problems 410

Review Problems 2 421

14. Designing experiments 427
14.1 Lessons from clinical trials 428
Example 14.1: Reducing HIV transmission 428
14.2 How to reduce bias 430
14.3 How to reduce the influence of sampling error 433
Example 14.3A: Holey waters 438
Example 14.3B: Plastic hormones 439

14.4 Experiments with more than one factor 440
Example 14.4: Lethal combination 441
14.5 What if you can't do experiments? 442
14.6 Choosing a sample size 444
14.7 Summary 448
14.8 Quick formula summary 449
Practice problems 452
Assignment problems 455

INTERLEAF 8 Data dredging 459

15. Comparing means of more than two groups 463
15.1 The analysis of variance 464
Example 15.1: The knees who say night 464
15.2 Assumptions and alternatives 474
15.3 Planned comparisons 475
15.4 Unplanned comparisons 477
Example 15.4: Wood wide web 478
15.5 Fixed and random effects 481
15.6 ANOVA with randomly chosen groups 482
Example 15.6: Walking-stick limbs 482
15.7 Summary 486
15.8 Quick formula summary 487
Practice problems 490
Assignment problems 497

INTERLEAF 9 Experimental and statistical mistakes 505

PART 4 REGRESSION AND CORRELATION

16. Correlation between numerical variables 507
16.1 Estimating a linear correlation coefficient 508
Example 16.1: Flipping the bird 511
16.2 Testing the null hypothesis of zero correlation 514
Example 16.2: What big inbreeding coefficients you have 514
16.3 Assumptions 517
16.4 The correlation coefficient depends on the range 518

16.5 Spearman's rank correlation 519
Example 16.5: The miracles of memory 520
16.6 The effects of measurement error on correlation 523
16.7 Summary 524
16.8 Quick formula summary 525
Practice problems 528
Assignment problems 533

INTERLEAF 10 Publication bias 540

17. Regression 545
17.1 Linear regression 546
Example 17.1: The lion's nose 546
17.2 Confidence in predictions 554
17.3 Testing hypotheses about a slope 556
Example 17.3: Prairie Home Campion 557
17.4 Regression toward the mean 560
17.5 Assumptions of regression 562
17.6 Transformations 566
17.7 The effects of measurement error on regression 568
17.8 Regression with nonlinear relationships 569
Example 17.8: The incredible shrinking seal 571
17.9 Logistic regression: fitting a binary response variable 572
17.10 Summary 576
17.11 Quick formula summary 577
Practice problems 581
Assignment problems 588

INTERLEAF 11 Meta-analysis 598

Review Problems 3 601

PART 5 MODERN STATISTICAL METHODS

18. Analyzing multiple factors 609
18.1 ANOVA and linear regression are linear models 610
18.2 Analyzing experiments with blocking 615
Example 18.2: Zooplankton depredation 616

18.3 Analyzing factorial designs 618
Example 18.3: Interaction zone 619
18.4 Adjusting for the effects of a covariate 623
Example 18.4: Mole-rat layabouts 624
18.5 Assumptions of linear models 627
18.6 Summary 629
Practice problems 630
Assignment problems 633

INTERLEAF 12 Using species as data points 639

19. Computer-intensive methods 643
19.1 Hypothesis testing using simulation 644
Example 19.1: How did he know? The non-randomness of haphazard choice 644
19.2 Bootstrap standard errors and confidence intervals 647
Example 19.2: The language center in chimps' brains 648
19.3 Summary 654
Practice problems 655
Assignment problems 659

20. Likelihood 663
20.1 What is likelihood? 664
20.2 Two uses of likelihood in biology 665
20.3 Maximum likelihood estimation 667
Example 20.3: Unruly passengers 667
20.4 Versatility of maximum likelihood estimation 672
Example 20.4: Conservation scoop 673
20.5 Log-likelihood ratio test 676
20.6 Summary 678
20.7 Quick formula summary 679
Practice problems 680
Assignment problems 684

21. Survival analysis 687
21.1 Survival curves 688
Example 21.1: Ebola outbreak 689

21.2 Compare survival curves 694
Example 21.2: Tumor genetics 694
21.3 Summary 701
21.4 Quick formula summary 701
Practice problems 702
Assignment problems 706

Statistical tables 711
Using statistical tables 711
Statistical Table A: The χ^2 distribution 713
Statistical Table B: The standard normal (Z) distribution 716
Statistical Table C: Student's t-distribution 718
Statistical Table D: The F-distribution 721
Statistical Table E: Mann–Whitney U-distribution 728
Statistical Table F: Tukey–Kramer q-distribution 730
Statistical Table G: Critical values for the Spearman's rank correlation 732

Literature cited 733

Answers to practice problems 755

Index 809