

Contents

0	Preface	vii
1	Variability	1
1.1	Sources of Variability	1
1.2	Types of Variables	3
1.3	Survey	5
2	Designing Studies	13
2.1	Comparative Experiments	13
2.2	Scientific Hypotheses	18
2.3	Randomization	19
2.4	Evidence	23
2.5	The Language of Hypothesis Testing	26
3	Visualizing Distributions	31
3.1	Dot Plots	32
3.2	Histograms	36
3.3	Density Plots	39
3.4	Categorical Variables	40
4	Quantiles	49
4.1	Sample Median	49
4.2	Quartiles	51
4.3	Quantile Plots	52
4.4	Box Plots	55
5	Averages	61
5.1	Sample Mean	61

5.2	Sample Standard Deviation	63
5.3	Causality	68
6	Visualizing Relationships	73
6.1	Scatter Plots	73
6.2	Time Plots	75
6.3	Describing Relationships	79
6.4	Two-Way Tables	83
6.5	Three-Way Tables	86
7	Linear Relationships	91
7.1	Pearson Correlation	91
7.2	Least-Squares Lines	94
7.3	Transforming Nonlinear Relationships	100
8	Probability	105
8.1	Populations and Sampling	105
8.2	Probability	107
8.3	Discrete Random Variables	109
8.4	Odds	110
8.5	Surprisal and Information	112
8.6	Continuous Random Variables	113
9	Conditional Probability	119
9.1	Sensitive Questions	119
9.2	Diagnostic Testing	121
9.3	Independence	123
10	Expectation	127
10.1	Expected Value	127
10.2	Variance	129
10.3	Continuous Random Variables	131
10.4	Combining Variables	132
11	Discrete Distributions	137
11.1	Binomial Distribution	137
11.2	Sample Counts and Proportions	146
11.3	Maximum Likelihood Estimation	149
11.4	Hypergeometric Distribution	150
11.5	Poisson Distribution	151

12 The Normal Distribution	159
12.1 Normal Density Curves	159
12.2 Standard Units	161
12.3 Normal Probability Plots	163
12.4 The $1.5 \times$ IQR Rule Revisited	165
12.5 Histograms	166
12.6 Random Numbers	166
13 Sampling Distribution of the Mean	173
13.1 Simulation	173
13.2 Some Algebra	179
14 Confidence Intervals	183
14.1 Standard Error	184
14.2 Student's t Distribution	184
14.3 Confidence Interval for a Mean	191
14.4 Choosing Sample Size	193
14.5 Prediction Intervals	194
14.6 Assumptions for t Methods	196
14.7 Transforming Data	197
15 Decisions	205
15.1 Means	205
15.2 Types of Errors	207
15.3 Improving Power	210
16 Comparing Two Means	219
16.1 Standard Error	219
16.2 Confidence Intervals	220
16.3 Hypothesis Tests	225
16.4 Pooling Standard Deviations	230
16.5 Choosing Sample Sizes	233
17 Inferences for Proportions	241
17.1 Binomial Tests	241
17.2 Normal Approximation	243
17.3 Comparing Two Proportions	250
17.4 Odds Ratios	253
17.5 Relative Risk	256

18 Inferences for Regression	261
18.1 Statistical Models	261
18.2 Inferences for the Slope	263
18.3 Confidence and Prediction Intervals	264
18.4 Checking Assumptions	266
18.5 Inferences for Correlation	268
18.6 A Summary of Parametric Methods	272
19 Analysis of Variance	277
19.1 Analysis of Variance for Regression	277
19.2 F Distribution	283
19.3 ANOVA Tables	289
19.4 Comparing Two Means	290
19.5 Comparing More Than Two Means	293
19.6 Assumptions for ANOVA	297
20 Multiple Comparisons	305
20.1 Least Significant Difference	305
20.2 Bonferroni Correction	310
20.3 Tukey's Honestly Significant Difference	318
21 Multiple Regression	325
21.1 Indicator Variables	325
21.2 Multiple Predictors	326
21.3 Two-Way Analysis of Variance	332
21.4 A Very Simple Model	338
21.5 R-Squared Revisited	340
22 Contingency Tables	345
22.1 Testing Randomness	345
22.2 Parametric Distributions	353
22.3 Relationship to Proportion Test	354
22.4 Two-Way Tables	355
22.5 Simpson's Paradox	358
23 Logistic Regression	363
23.1 Logistic Curves	363
23.2 Odds Ratios	365
23.3 Inference for Logistic Regression	366
23.4 Odds Ratios	369

24 Nonparametric Methods	375
24.1 Effect of Outliers	375
24.2 Sign Test	377
24.3 Signed-Rank Test	379
24.4 Rank-Sum Test	382
24.5 Kruskal-Wallis Test	389
24.6 Spearman's Rank Correlation	392
24.7 Final Words	395
25 Pictures of Point-Oh-Five	399
A Mathematical Miscellany	407
A.1 Logarithms	407
A.2 Arithmetic-Geometric Mean	416
A.3 Choosing Histogram Bins	419
A.4 Desert Island Formulas	420
A.5 Compound Interest and Student's t Distribution	422
A.6 Other Distribution Functions	424
B Historical Data Sets	427
B.1 The Passage Time of Light	427
B.2 The Mean Density of the Earth	428
B.3 Effects of Cross and Self Fertilization	428
B.4 Counting Yeast Cells with a Hemocytometer	429
B.5 The First t Test	431
B.6 Yields of Potatoes	432
B.7 Dimensions of Irises	433
C Notation	437
C.1 Greek Letters	437
C.2 Capital Letters	437
C.3 Other Symbols	438
C.4 Riemann's Zeta	438
References	447
Index	457