

TABLE OF CONTENTS

1	The story Behind the Statistics	3
2	Comparisons and CARE Solutions	59
3	Understanding Measurements	116
4	Describing Ratios: Percent and Percentage	177
5	Describing Ratios: Rate and Chance Grammar	217
6	Comparing Ratios	234
7	Interpreting Confusing Ratios	270
8	Randomness	321
9	Conclusion	351
10	Social Statistics: Tables, Graphs and Stories	352
11	Guide for Teachers	416
12	Key Word List by Chapter	421
13	List of Figures	423
14	List of Tables	427
15	List of Graphs and Maps	432
16	List of Stories	434
17	List of Charts and Images	435
18	List of Equations and Diagrams	436
19	Glossary	438
20	Index	464
21	Review of Named Ratios	472

Notation used: Bold indicates a term that is being defined. Any word or phrase in bold should be found in the Glossary. Italics indicate an important concept, point or claim.

1 The story Behind the Statistics

TABLE OF CONTENTS

1.1	Learning Outcomes	3
1.2	Introduction	4
1.2.1	Statistics Are Numbers in Context	6
1.2.2	Statistics in Arguments	8
1.2.3	What Kinds of Things Influence Statistics?	10
1.2.4	Summary	11
1.3	Critical Thinking Issues	12
1.3.1	Association	12
1.3.2	Causation	14
1.3.3	ABC Keywords for Association versus Causation	15
1.3.4	Association Is Not Causation	17
1.3.5	Strength of Evidence	20
1.4	Statistical Literacy Issues	21
1.4.1	C = CONFOUNDING	22
1.4.2	A = ASSEMBLY	27
1.4.3	R = RANDOMNESS	33
1.4.4	E = ERRORS/Bias	38
1.5	Putting It All Together: Analyzing Stories	42
1.6	Chapter Summary	45
1.7	Optional/Advanced	46
1.7.1	Arguments	46
1.7.2	Separating Statistical Literacy from Critical Thinking	50
1.7.3	Other Topics	51

1.1 Learning Outcomes

Critical thinking outcomes:

- Distinguish association from causation using an ABC classification.

Statistical literacy outcomes:

- Learn that statistics are numbers in context: they can be influenced.
- Learn that statistics are socially constructed by people with motives and values.
- Learn that "take care" is the best advice when dealing with statistics.
- Learn how all influences on statistics can be grouped into four categories.
- Learn CARE categories: Confounding, Assembly, Randomness, and Error.
- See how each influence can change the size and direction of an association.
- Use "Take CARE" to analyze statistics in the everyday media.

2 Comparisons and CARE Solutions

TABLE OF CONTENTS

2.1	Learning Outcomes	59
2.2	Review	60
2.3	Association	60
2.3.1	Two-Group Comparisons	60
2.3.2	Two-Factor Co-variation	66
2.4	Confounder Solutions	67
2.4.1	Effect Size	68
2.4.2	Control Groups	68
2.4.3	Study Design	69
2.4.4	Selection and Matching	82
2.5	Assembly Solutions	85
2.5.1	Assembly in Definitions	86
2.5.2	Assembly in Comparisons	87
2.6	Randomness Solutions	90
2.6.1	Extremes and Coincidence	90
2.6.2	Big Data and the Law of Very Large Numbers	91
2.6.3	Margin of Error and Statistical Significance	94
2.7	Error (Bias) Solutions	97
2.7.1	Three Kinds of Bias	97
2.7.2	Minimizing Bias	99
2.8	Analyzing Stories	100
2.9	Chapter Summary	104
2.10	Optional/Advanced	105
2.10.1	Associations	105
2.10.2	Comparisons	105
2.10.3	Confounding	106
2.10.4	Randomness	112

2.1 Learning Outcomes

Recognize associations as possible indications of causation

- Analyze associations: two-group comparison and two-factor co-variation
- Create four kinds of arithmetic comparisons. Recognize exceptions

Recognize techniques that deal with confounding

- Distinguish hard science (studies) from soft science (studies)
- Distinguish different kinds of experiments and observational studies
- Recognize how different study designs control different confounders
- Recognize control group and controlled study, control of and control for

Study the techniques that deal with assembly, randomness and bias

- Become sensitized to the variety and importance of assembly
- Test for statistical significance using non-overlapping confidence intervals
- Recognize bias-control: placebo, single and double blind, random sample

3 Understanding Measurements

TABLE OF CONTENTS

3.1	Learning Outcomes	116
3.2	Review	117
3.3	Data and Data Distributions	117
3.3.1	Count (Discrete) Data	118
3.3.2	Measurement (Continuous) Data	119
3.4	Measures of Location	121
3.4.1	Ranks	121
3.4.2	Percentiles	121
3.5	Measures of Center	125
3.5.1	Mean, Median, and Mode	125
3.5.2	Measures of Center versus Skew	127
3.5.3	Describing and Comparing Measures of Centers	129
3.5.4	From Association to Causation	132
3.6	Assembly on Measures of Center	134
3.6.1	Assembly: Defining Groups and Group Averages	134
3.6.2	Assembly: Forming Groups Using a Cut Point	134
3.6.3	Assembly: Choosing a Measure of Center	135
3.7	Confounder Solutions	136
3.7.1	Forming and Comparing Per Ratios	137
3.7.2	Untangling Confounding of Averages Using Selection	139
3.7.3	Standardizing Averages: Mix-Matching	142
3.7.4	Mix-Matching Arithmetically	143
3.7.5	Standardizing Averages Graphically	149
3.8	The Essentials	157
3.9	Chapter Summary	159
3.10	Optional/Advanced	160
3.10.1	Two-Group Comparisons of Measures (Detail)	160
3.10.2	Race-Based Examples	164
3.10.3	Percentage Explained by a Confounder	170
3.10.4	The Cornfield Conditions	172
3.10.5	Other Topics	173

3.1 Learning Outcomes

- Understand ranks and percentiles.
- Distinguish percent, "percentage points," and "percentile points."
- Recognize the order of mean, median, and mode in skewed distributions.
- Compare means, medians, or modes using ordinary English.
- See how assembly can influence the size of a comparison.
- See how ratios can change the size and direction of a comparison.
- *Take into account* the influence of a confounder using standardization.
- Standardize comparisons of averages using arithmetic or graphical techniques.

4 Describing Ratios: Percent and Percentage

TABLE OF CONTENTS

4.1	Learning Outcomes	177
4.2	Review	178
4.3	Part-Whole Ratios	178
4.3.1	Pie Charts of Counts and Ratios: Good and Bad	180
4.3.2	Percent Grammar using Tables of Counts	181
4.4	Percent Grammar Statements	182
4.4.1	Answering Percent Questions using Count Table Data	182
4.4.2	Percent Grammar: Forming Statements	183
4.4.3	Common Activities involving Count Tables	185
4.4.4	100% Ratio Tables: Three Kinds	186
4.4.5	Decoding Problems	191
4.5	Percentage Grammar	193
4.5.1	Reading Percentages in 100% Tables	195
4.5.2	Sports and Portion Grammar	198
4.5.3	Who-Who or Who-What Ambiguity	198
4.6	Count Tables (Percent or Percentage Grammar)	199
4.7	Distribution Grammar: Discrete Data	200
4.8	Half-Tables of Percentages	201
4.8.1	One-way Half-Tables of Percentages	201
4.8.2	Two-Way Half-Tables of Percentages	205
4.8.3	Tables and Graphs with Missing Margins	207
4.9	Chapter Summary	210
4.10	Optional/Advanced	211
4.10.1	Defining Proportion, Fraction, and Percentage	211
4.10.2	Percent and Percentage Grammar: More Details	211
4.10.3	Distribution Grammar with Two Variables	215
4.10.4	Grouping Named Ratio Families by Type of Preposition	216

4.1 Learning Outcomes

Learning outcomes include:

- Identify part and whole in named ratio grammar statements and questions.
- Learn two named-ratio grammars: percent and percentage grammar
- Describe percentages using percent and percentage grammar.
- Translate part-whole statements between percent and percentage grammar.
- Describe percentages in 100% tables and in half-tables.
- Describe percentages in tables with missing margins.
- Recognize a wrong-order error: the confusion of the inverse.

5 Describing Ratios: Rate and Chance Grammar

TABLE OF CONTENTS

5.1	Learning Outcomes	217
5.2	Review	218
5.3	Rate Grammar	219
5.3.1	Introducing Rates	219
5.3.2	Phrase Rate Grammar	222
5.3.3	Clause Rate Grammar	225
5.4	Chance Grammar	227
5.5	Ratio Grammar	230
5.6	Chapter Summary	232
5.7	Optional/Advanced	233

5.1 Learning Outcomes

- Recognize the members of the rate family: prevalence and incidence.
- Recognize that the keyword *rate* is equivocal. It has four forms: frequency, growth, prevalence and incidence.
- Describe ratios using rate grammar in both phrase and a clause form.
- Recognize that the *chance* family of grammars includes odds, risk, likelihood, and probability.
- Describe ratios using chance grammar in both phrase and clause form.
- Recognize that chance grammar is common but often leaves the whole unstated.
- Describe ratios using *ratio* grammar.
- Recognize the confusion of the inverse in rate, chance and ratio grammars.
- Recognize that all ratios are easily influence by the choice of the denominator.
- Graphs: recall that a quintile is one-fifth, quartile is fourth and decile is a tenth.

6 Comparing Ratios

TABLE OF CONTENTS

6.1	Learning Outcomes	234
6.2	Review	235
6.3	Introduction to Comparisons of Ratios	235
6.4	Attributable Comparisons	236
6.4.1	Percentage Attributable (Rate Based)	237
6.4.2	Cases (Deaths) Attributable (Rate based)	238
6.5	Comparing Ratios: Common versus Distinct Part	240
6.5.1	Comparing Ratios of 100% Wholes	242
6.6	Comparing Ratios Using Likely	243
6.6.1	Decoding	243
6.6.2	Constructing	246
6.6.3	Likely Comparison Grammar: Examples of All Types	248
6.6.4	Conversions into Likely comparisons	249
6.6.5	Common Pitfalls in Likely Comparisons	250
6.6.6	Examples of Comparison (Toy Table)	253
6.7	Comparing Ratios Using Named Ratios	254
6.7.1	Reading Ratio Comparisons Using Named Ratios	254
6.7.2	Constructing named-ratio comparisons	255
6.7.3	Generating Short Form Comparisons	260
6.7.4	Spotting Errors in Comparisons	261
6.7.5	Comparing Part-Whole Ratios in Graphs:	262
6.8	Comparing Quintiles	263
6.9	Critical Thinking about Comparisons of Ratios	263
6.10	Confounding: A Review	264
6.11	Chapter Summary	265
6.12	Optional/Advanced Topics	266
6.12.1	Miscellaneous: RR, Bayes Comparison	266
6.12.2	Cases (Deaths) Attributable: Multiple Steps	268

6.1 Learning Outcomes

- Calculate the percentage and counts that are attributable to group membership.
- Distinguish common-part and distinct-part comparisons
- Read and write comparisons of part-whole ratios using *likely* grammar.
- Read and write comparisons of part-whole ratios using *named-ratio* grammar.

7 Interpreting Confusing Ratios

TABLE OF CONTENTS

7.1	Learning Outcomes	270
7.2	Review	271
7.3	Diabolical Denominators	271
7.3.1	Complements	274
7.3.2	Scanlan's Paradox	274
7.4	COVID Death Statistics	277
7.5	Errors and Fallacies	278
7.6	Medical Tests: Equivocation	280
7.7	Social Statistics: Prediction and Explanation	286
7.8	Three-Term Ratios	291
7.9	Confounding and Simpson's Paradox	292
7.9.1	Smoking and Lung Cancer	292
7.9.2	Simpson's Paradox for Rates and Percentages	293
7.9.3	Cornfield Conditions	295
7.9.4	Other Examples	296
7.9.5	Controlling for Confounding using Mix-Matching	300
7.10	The Biggest Confounder	304
7.10.1	Confounding in the Humanities and Social Sciences	306
7.11	Chapter Summary	307
7.12	Optional/Advanced	308
7.12.1	Comparing Two Ratios with Related Parts	308
7.12.2	Comparing Two Ratios with Related Wholes	308
7.12.3	Distinct and Common Parts	308
7.12.4	Unclear Wholes	311
7.12.5	Comparing a Change in Ratios with Different Wholes	312
7.12.6	Outstanding Relative Risks	314
7.12.7	Three-Variable Ratios	315
7.12.8	Medical Tests: Sensitivity versus Specificity	315
7.12.9	Cornfield's Second Condition	316
7.12.10	Co-Variation involving Percentages	318
7.12.11	Combining Multiple Groups into Two Groups	319
7.12.12	Number Needed to Treat	319
7.12.13	Standardizing Attributable Statistics	320

Learning Outcomes

- Identify and evaluate inverse ratio claims (the confusion of the inverse)
- Distinguish different kinds of accuracy and error in medical tests.
- Calculate predictive accuracy given test accuracy and disease prevalence.
- Predict the possibility of a reversal using the Cornfield conditions.
- Standardize percentages by mix-matching: algebraic and graphical

8 Randomness

CONTENTS

8.1	Learning Outcomes	321
8.2	Review and Background	322
8.3	Idea of Statistical Significance	322
8.3.1	Good and Bad of Statistical Significance	323
8.4	Influencing Statistical Significance	324
8.4.1	Influenced by Confounding	324
8.4.2	Influenced by Assembly	326
8.4.3	Influenced by Randomness	327
8.4.4	Influenced by Error or Bias	328
8.5	Hypothesis Tests and Criminal Trials	329
8.6	Political Polls	331
8.7	Sampling: Random and Nonrandom	333
8.7.1	Random Sampling	333
8.7.2	Nonrandom Sampling	335
8.7.3	Counting the Uncountable	336
8.8	Statistical Significance Stories	338
8.9	Statistical Significance: Applicability	339
8.9.1	Unexpected if Due to Chance: Evolution	340
8.10	Common Misunderstandings Involving Chance	341
8.11	Chapter Summary	342
8.12	Optional/Advanced	343
8.12.1	Interpreting Statistical Significance in Hypothesis Tests	343
8.12.2	Margin of Error in Subgroups	348

8.1 Learning Outcomes

- Understand the good and bad of statistical significance
- Recognize how statistical significance can be influenced—from significance to insignificance or vice versa—by confounding, assembly, randomness and bias
- Compare hypothesis tests to criminal trials: type 1 and type 2 error
- Distinguish polls from surveys. Polls predict; surveys summarize.
- Estimate unknown size of a population using capture-recapture
- Identify some common misunderstanding involving chance
- Optional: Calculate confirmation error necessary for a statistically-significant result to give a 50% chance that the result is true.
- Optional: Confidence intervals for subgroups given survey margin of error