STATISTICS

Eleventh Edition

Robert S. Witte
Emeritus, San Jose State University

John S. Witte
University of California, San Francisco

WILEY

Contents

PREFACE iv

ACKNOWLEDGMENTS vi

1 INTRODUCTION 1

1.1 WHY STUDY STATISTICS? 2
1.2 WHAT IS STATISTICS? 2
1.3 MORE ABOUT INFERENTIAL STATISTICS 3
1.4 THREE TYPES OF DATA 6
1.5 LEVELS OF MEASUREMENT 7
1.6 TYPES OF VARIABLES 11
1.7 HOW TO USE THIS BOOK 15

Summary 16
Important Terms 17
Review Questions 17

PART 1 Descriptive Statistics: Organizing and Summarizing Data 21

2 DESCRIBING DATA WITH TABLES AND GRAPHS 22

TABLES (FREQUENCY DISTRIBUTIONS) 23

2.1 FREQUENCY DISTRIBUTIONS FOR QUANTITATIVE DATA 23
2.2 GUIDELINES 24
2.3 OUTLIERS 27
2.4 RELATIVE FREQUENCY DISTRIBUTIONS 28
2.5 CUMULATIVE FREQUENCY DISTRIBUTIONS 30
2.6 FREQUENCY DISTRIBUTIONS FOR QUALITATIVE (NOMINAL) DATA 31
2.7 INTERPRETING DISTRIBUTIONS CONSTRUCTED BY OTHERS 32

GRAPHS 33

2.8 GRAPHS FOR QUANTITATIVE DATA 33
2.9 TYPICAL SHAPES 37
2.10 A GRAPH FOR QUALITATIVE (NOMINAL) DATA 39
2.11 MISLEADING GRAPHS 40
2.12 DOING IT YOURSELF 41

Summary 42
Important Terms 43
Review Questions 43

3 DESCRIBING DATA WITH AVERAGES 47
3.1 MODE 48
3.2 MEDIAN 49
3.3 MEAN 51
3.4 WHICH AVERAGE? 53
3.5 AVERAGES FOR QUALITATIVE AND RANKED DATA 55
Summary 56
Important Terms 57
Key Equation 57
Review Questions 57

4 DESCRIBING VARIABILITY 60
4.1 INTUITIVE APPROACH 61
4.2 RANGE 62
4.3 VARIANCE 63
4.4 STANDARD DEVIATION 64
4.5 DETAILS: STANDARD DEVIATION 67
4.6 DEGREES OF FREEDOM (*df*) 75
4.7 INTERQUARTILE RANGE (IQR) 76
4.8 MEASURES OF VARIABILITY FOR QUALITATIVE AND RANKED DATA 78
Summary 78
Important Terms 79
Key Equations 79
Review Questions 79

5 NORMAL DISTRIBUTIONS AND STANDARD (*z*) SCORES 82
5.1 THE NORMAL CURVE 83
5.2 *z* SCORES 86
5.3 STANDARD NORMAL CURVE 87
5.4 SOLVING NORMAL CURVE PROBLEMS 89
5.5 FINDING PROPORTIONS 90
5.6 FINDING SCORES 95
5.7 MORE ABOUT *z* SCORES 100
Summary 103
Important Terms 103
Key Equations 103
Review Questions 103

6 DESCRIBING RELATIONSHIPS: CORRELATION 107
6.1 AN INTUITIVE APPROACH 108
6.2 SCATTERPLOTS 109
6.3 A CORRELATION COEFFICIENT FOR QUANTITATIVE DATA: *r* 113
6.4 DETAILS: COMPUTATION FORMULA FOR *r* 117
6.5 OUTLIERS AGAIN 118
6.6 OTHER TYPES OF CORRELATION COEFFICIENTS 119

6.7 COMPUTER OUTPUT 120
Summary 123
Important Terms and Symbols 124
Key Equations 124
Review Questions 124

7 REGRESSION 126
7.1 TWO ROUGH PREDICTIONS 127
7.2 A REGRESSION LINE 128
7.3 LEAST SQUARES REGRESSION LINE 130
7.4 STANDARD ERROR OF ESTIMATE, $s_{y|x}$ 133
7.5 ASSUMPTIONS 135
7.6 INTERPRETATION OF r^2 136
7.7 MULTIPLE REGRESSION EQUATIONS 141
7.8 REGRESSION TOWARD THE MEAN 141
Summary 143
Important Terms 144
Key Equations 144
Review Questions 144

PART 2 Inferential Statistics: Generalizing Beyond Data 147

8 POPULATIONS, SAMPLES, AND PROBABILITY 148

POPULATIONS AND SAMPLES 149
8.1 POPULATIONS 149
8.2 SAMPLES 150
8.3 RANDOM SAMPLING 151
8.4 TABLES OF RANDOM NUMBERS 151
8.5 RANDOM ASSIGNMENT OF SUBJECTS 153
8.6 SURVEYS OR EXPERIMENTS? 154

PROBABILITY 155
8.7 DEFINITION 155
8.8 ADDITION RULE 156
8.9 MULTIPLICATION RULE 157
8.10 PROBABILITY AND STATISTICS 161
Summary 162
Important Terms 163
Key Equations 163
Review Questions 163

9 SAMPLING DISTRIBUTION OF THE MEAN 168

9.1 WHAT IS A SAMPLING DISTRIBUTION? 169
9.2 CREATING A SAMPLING DISTRIBUTION FROM SCRATCH 170
9.3 SOME IMPORTANT SYMBOLS 173
9.4 MEAN OF ALL SAMPLE MEANS ($\mu_{\bar{X}}$) 173
9.5 STANDARD ERROR OF THE MEAN ($\sigma_{\bar{X}}$) 174
9.6 SHAPE OF THE SAMPLING DISTRIBUTION 176
9.7 OTHER SAMPLING DISTRIBUTIONS 178

Summary 178
Important Terms 179
Key Equations 179
Review Questions 179

10 INTRODUCTION TO HYPOTHESIS TESTING: THE *z* TEST 182

10.1 TESTING A HYPOTHESIS ABOUT SAT SCORES 183
10.2 *z* TEST FOR A POPULATION MEAN 185
10.3 STEP-BY-STEP PROCEDURE 186
10.4 STATEMENT OF THE RESEARCH PROBLEM 187
10.5 NULL HYPOTHESIS (H_0) 188
10.6 ALTERNATIVE HYPOTHESIS (H_1) 188
10.7 DECISION RULE 189
10.8 CALCULATIONS 190
10.9 DECISION 190
10.10 INTERPRETATION 191

Summary 191
Important Terms 192
Key Equations 192
Review Questions 193

11 MORE ABOUT HYPOTHESIS TESTING 195

11.1 WHY HYPOTHESIS TESTS? 196
11.2 STRONG OR WEAK DECISIONS 197
11.3 ONE-TAILED AND TWO-TAILED TESTS 199
11.4 CHOOSING A LEVEL OF SIGNIFICANCE (α) 202
11.5 TESTING A HYPOTHESIS ABOUT VITAMIN C 203
11.6 FOUR POSSIBLE OUTCOMES 204
11.7 IF H_0 REALLY IS TRUE 206
11.8 IF H_0 REALLY IS FALSE BECAUSE OF A *LARGE* EFFECT 207
11.9 IF H_0 REALLY IS FALSE BECAUSE OF A *SMALL* EFFECT 209
11.10 INFLUENCE OF SAMPLE SIZE 211
11.11 POWER AND SAMPLE SIZE 213

Summary 216
Important Terms 217
Review Questions 218

12 ESTIMATION (CONFIDENCE INTERVALS) 221

12.1 POINT ESTIMATE FOR μ 222
12.2 CONFIDENCE INTERVAL (CI) FOR μ 222
12.3 INTERPRETATION OF A CONFIDENCE INTERVAL 226
12.4 LEVEL OF CONFIDENCE 226
12.5 EFFECT OF SAMPLE SIZE 227
12.6 HYPOTHESIS TESTS OR CONFIDENCE INTERVALS? 228
12.7 CONFIDENCE INTERVAL FOR POPULATION PERCENT 228
Summary 230
Important Terms 230
Key Equation 230
Review Questions 231

13 *t* TEST FOR ONE SAMPLE 233

13.1 GAS MILEAGE INVESTIGATION 234
13.2 SAMPLING DISTRIBUTION OF t 234
13.3 t TEST 237
13.4 COMMON THEME OF HYPOTHESIS TESTS 238
13.5 REMINDER ABOUT DEGREES OF FREEDOM 238
13.6 DETAILS: ESTIMATING THE STANDARD ERROR ($s_{\overline{X}}$) 238
13.7 DETAILS: CALCULATIONS FOR THE t TEST 239
13.8 CONFIDENCE INTERVALS FOR μ BASED ON t 241
13.9 ASSUMPTIONS 242
Summary 242
Important Terms 243
Key Equations 243
Review Questions 243

14 *t* TEST FOR TWO INDEPENDENT SAMPLES 245

14.1 EPO EXPERIMENT 246
14.2 STATISTICAL HYPOTHESES 247
14.3 SAMPLING DISTRIBUTION OF $\overline{X}_1 - \overline{X}_2$ 248
14.4 t TEST 250
14.5 DETAILS: CALCULATIONS FOR THE t TEST 252
14.6 p-VALUES 255
14.7 STATISTICALLY SIGNIFICANT RESULTS 258
14.8 ESTIMATING EFFECT SIZE: POINT ESTIMATES AND CONFIDENCE INTERVALS 259
14.9 ESTIMATING EFFECT SIZE: COHEN'S d 262
14.10 META-ANALYSIS 264
14.11 IMPORTANCE OF REPLICATION 264
14.12 REPORTS IN THE LITERATURE 265

14.13 ASSUMPTIONS 266
14.14 COMPUTER OUTPUT 267
Summary 268
Important Terms 268
Key Equations 269
Review Questions 269

15 *t* TEST FOR TWO RELATED SAMPLES (REPEATED MEASURES) 273
15.1 EPO EXPERIMENT WITH REPEATED MEASURES 274
15.2 STATISTICAL HYPOTHESES 277
15.3 SAMPLING DISTRIBUTION OF $\bar{D}$ 277
15.4 *t* TEST 278
15.5 DETAILS: CALCULATIONS FOR THE *t* TEST 279
15.6 ESTIMATING EFFECT SIZE 281
15.7 ASSUMPTIONS 283
15.8 OVERVIEW: THREE *t* TESTS FOR POPULATION MEANS 283
15.9 *t* TEST FOR THE POPULATION CORRELATION COEFFICIENT, ρ 285
Summary 287
Important Terms 288
Key Equations 288
Review Questions 288

16 ANALYSIS OF VARIANCE (ONE FACTOR) 292
16.1 TESTING A HYPOTHESIS ABOUT SLEEP DEPRIVATION AND AGGRESSION 293
16.2 TWO SOURCES OF VARIABILITY 294
16.3 *F* TEST 296
16.4 DETAILS: VARIANCE ESTIMATES 299
16.5 DETAILS: MEAN SQUARES (*MS*) AND THE *F* RATIO 304
16.6 TABLE FOR THE *F* DISTRIBUTION 305
16.7 ANOVA SUMMARY TABLES 307
16.8 *F* TEST IS NONDIRECTIONAL 308
16.9 ESTIMATING EFFECT SIZE 308
16.10 MULTIPLE COMPARISONS 311
16.11 OVERVIEW: FLOW CHART FOR ANOVA 315
16.12 REPORTS IN THE LITERATURE 315
16.13 ASSUMPTIONS 316
16.14 COMPUTER OUTPUT 316
Summary 317
Important Terms 318
Key Equations 318
Review Questions 319

17 ANALYSIS OF VARIANCE (REPEATED MEASURES) 322
17.1 SLEEP DEPRIVATION EXPERIMENT WITH REPEATED MEASURES 323
17.2 *F* TEST 324

17.3 TWO COMPLICATIONS 325
17.4 DETAILS: VARIANCE ESTIMATES 326
17.5 DETAILS: MEAN SQUARE (*MS*) AND THE *F* RATIO 329
17.6 TABLE FOR *F* DISTRIBUTION 331
17.7 ANOVA SUMMARY TABLES 331
17.8 ESTIMATING EFFECT SIZE 333
17.9 MULTIPLE COMPARISONS 333
17.10 REPORTS IN THE LITERATURE 335
17.11 ASSUMPTIONS 336
Summary 336
Important Terms 336
Key Equations 337
Review Questions 337

18 ANALYSIS OF VARIANCE (TWO FACTORS) 339
18.1 A TWO-FACTOR EXPERIMENT: RESPONSIBILITY IN CROWDS 340
18.2 THREE *F* TESTS 342
18.3 INTERACTION 344
18.4 DETAILS: VARIANCE ESTIMATES 347
18.5 DETAILS: MEAN SQUARES (*MS*) AND *F* RATIOS 351
18.6 TABLE FOR THE *F* DISTRIBUTION 353
18.7 ESTIMATING EFFECT SIZE 353
18.8 MULTIPLE COMPARISONS 354
18.9 SIMPLE EFFECTS 355
18.10 OVERVIEW: FLOW CHART FOR TWO-FACTOR ANOVA 358
18.11 REPORTS IN THE LITERATURE 358
18.12 ASSUMPTIONS 360
18.13 OTHER TYPES OF ANOVA 360
Summary 360
Important Terms 361
Key Equations 361
Review Questions 361

19 CHI-SQUARE (χ^2) TEST FOR QUALITATIVE (NOMINAL) DATA 365

ONE-VARIABLE χ^2 TEST 366
19.1 SURVEY OF BLOOD TYPES 366
19.2 STATISTICAL HYPOTHESES 366
19.3 DETAILS: CALCULATING χ^2 367
19.4 TABLE FOR THE χ^2 DISTRIBUTION 369
19.5 χ^2 TEST 370

TWO-VARIABLE χ^2 TEST 372
19.6 LOST LETTER STUDY 372
19.7 STATISTICAL HYPOTHESES 373
19.8 DETAILS: CALCULATING χ^2 373

19.9 TABLE FOR THE χ^2 DISTRIBUTION 376
19.10 χ^2 TEST 376
19.11 ESTIMATING EFFECT SIZE 377
19.12 ODDS RATIOS 378
19.13 REPORTS IN THE LITERATURE 380
19.14 SOME PRECAUTIONS 380
19.15 COMPUTER OUTPUT 381
Summary 382
Important Terms 382
Key Equations 382
Review Questions 382

20 TESTS FOR RANKED (ORDINAL) DATA 386
20.1 USE ONLY WHEN APPROPRIATE 387
20.2 A NOTE ON TERMINOLOGY 387
20.3 MANN–WHITNEY *U* TEST (TWO INDEPENDENT SAMPLES) 387
20.4 WILCOXON *T* TEST (TWO RELATED SAMPLES) 392
20.5 KRUSKAL–WALLIS *H* TEST (THREE OR MORE INDEPENDENT SAMPLES) 396
20.6 GENERAL COMMENT: TIES 400
Summary 400
Important Terms 400
Review Questions 400

21 POSTSCRIPT: WHICH TEST? 403
21.1 DESCRIPTIVE OR INFERENTIAL STATISTICS? 404
21.2 HYPOTHESIS TESTS OR CONFIDENCE INTERVALS? 404
21.3 QUANTITATIVE OR QUALITATIVE DATA? 404
21.4 DISTINGUISHING BETWEEN THE TWO TYPES OF DATA 406
21.5 ONE, TWO, OR MORE GROUPS? 407
21.6 CONCLUDING COMMENTS 408
Review Questions 408

APPENDICES 411
A MATH REVIEW 411
B ANSWERS TO SELECTED QUESTIONS 419
C TABLES 457
D GLOSSARY 471

INDEX 477