
Table of Contents

Preface, xvii

Acknowledgments, xxiii

About the Authors, xxv

How to Contact the Authors and Access the Data Sets, xxvii

CHAPTER 1 ■ Overview: The Role of Statistics in
Transportation Engineering 1

1.1 WHAT IS ENGINEERING? 2

1.2 WHAT IS TRANSPORTATION ENGINEERING? 5

1.3 GOAL OF THE TEXTBOOK 7

1.4 OVERVIEW OF THE TEXTBOOK 7

1.5 WHO IS THE AUDIENCE FOR THIS TEXTBOOK? 8

1.6 RELAX—EVERYTHING IS FINE 8

REFERENCES 9

CHAPTER 2 ■ Graphical Methods for Displaying Data 11

2.1 INTRODUCTION 11

2.2 HISTOGRAM 13

2.3 BOX AND WHISKER PLOT 16

2.4 QUANTILE PLOT 18

2.5 SCATTER PLOT 19

2.6 PARALLEL PLOT 22

2.7 TIME SERIES PLOT 23

2.8	QUALITY CONTROL PLOTS	24
2.9	CONCLUDING REMARKS	26
	HOMEWORK PROBLEMS	26
	REFERENCES	27
CHAPTER 3 ■ Numerical Summary Measures		29
3.1	INTRODUCTION	29
3.2	MEASURES OF CENTRAL TENDENCY	30
3.2.1	Mean	30
3.2.2	Median	32
3.2.3	Trimmed Mean	33
3.2.4	Mode	33
3.2.5	Proportion	34
3.3	MEASURES OF RELATIVE STANDING	35
3.3.1	Percentile	35
3.3.2	Quartile	36
3.4	MEASURES OF VARIABILITY	36
3.4.1	Range	36
3.4.2	Variance	37
3.4.3	Standard Deviation	37
3.4.4	Interquartile Range	38
3.4.5	Coefficient of Variation	38
3.5	MEASURES OF ASSOCIATION	39
3.5.1	Pearson's Sample Correlation Coefficient	40
3.5.2	Spearman's Rank Correlation Coefficient	41
3.6	CONCLUDING REMARKS	43
	HOMEWORK PROBLEMS	44
	REFERENCES	46
CHAPTER 4 ■ Probability and Random Variables		47
4.1	INTRODUCTION	47
4.2	SAMPLE SPACES AND EVENTS	47
4.3	INTERPRETATION OF PROBABILITY	48

4.4	RANDOM VARIABLE	50
4.4.1	Functions of a Random Variable	52
4.5	EXPECTATIONS OF RANDOM VARIABLES	53
4.5.1	Expected Values Are Linear	54
4.6	COVARIANCES AND CORRELATION OF RANDOM VARIABLES	54
4.7	COMPUTING EXPECTED VALUES OF FUNCTIONS OF RANDOM VARIABLES	55
4.8	CONDITIONAL PROBABILITY	57
4.9	BAYES' THEOREM	59
4.10	CONCLUDING REMARKS	60
	HOMEWORK PROBLEMS	61
	REFERENCES	63
CHAPTER 5 ■ Common Probability Distributions		65
5.1	INTRODUCTION	65
5.2	DISCRETE DISTRIBUTIONS	66
5.2.1	Binomial Distribution	66
5.2.2	Poisson Distribution	69
5.2.3	Negative Binomial Distribution	71
5.3	CONTINUOUS DISTRIBUTIONS	73
5.3.1	Normal Distribution	74
5.3.2	<i>t</i> -Distribution	78
5.3.3	Lognormal Distribution	79
5.3.4	Exponential Distribution	80
5.3.5	Gamma Distribution	81
5.3.6	Chi-Square Distribution	82
5.3.7	<i>F</i> -Distribution	84
5.3.8	Beta Distribution	85
5.4	CONCLUDING REMARKS	85
	APPENDIX: TABLE OF THE MOST POPULAR DISTRIBUTIONS IN TRANSPORTATION ENGINEERING	87

HOMEWORK PROBLEMS	87
REFERENCES	88
CHAPTER 6 ■ Sampling Distributions	91
6.1 INTRODUCTION	91
6.2 RANDOM SAMPLING	92
6.3 SAMPLING DISTRIBUTION OF A SAMPLE MEAN	94
6.4 SAMPLING DISTRIBUTION OF A SAMPLE VARIANCE	103
6.5 SAMPLING DISTRIBUTION OF A SAMPLE PROPORTION	103
6.6 CONCLUDING REMARKS	106
HOMEWORK PROBLEMS	107
REFERENCES	109
CHAPTER 7 ■ Inferences: Hypothesis Testing and Interval Estimation	111
7.1 INTRODUCTION	111
7.2 FUNDAMENTALS OF HYPOTHESIS TESTING	112
7.3 INFERENCES ON A SINGLE POPULATION MEAN	114
7.3.1 Hypothesis Tests about a Population Mean	114
7.3.1.1 z -Test	116
7.3.1.2 t -Test	117
7.3.2 Interval Estimates for a Population Mean	119
7.3.2.1 Bias-Adjusted Confidence Intervals	120
7.4 INFERENCES ABOUT TWO POPULATION MEANS	123
7.4.1 Hypothesis Tests about Equality of Two Population Means	123
7.4.1.1 Paired t -Test	123
7.4.1.2 Pooled t -Test	125
7.4.1.3 Unequal Variance t -Test	125

7.4.2	Interval Estimates and Bias Adjustment for Difference of Two Population Means	127
7.5	INFERENCES ABOUT ONE POPULATION VARIANCE	131
7.6	INFERENCES ABOUT TWO POPULATION VARIANCES	133
7.6.1	Confidence Intervals for the Ratio of Two Variances	134
7.6.2	Bias-Corrected Confidence Interval for the Ratio of Two Variances	134
7.6.3	One-Sided Tests	135
7.7	CONCLUDING REMARKS	135
	APPENDIX: WELCH (1938) DEGREES OF FREEDOM FOR THE UNEQUAL VARIANCE t -TEST	136
	HOMEWORK PROBLEMS	136
	REFERENCES	138

CHAPTER 8 ■	Other Inferential Procedures: ANOVA and Distribution-Free Tests	139
8.1	INTRODUCTION	139
8.2	COMPARISONS OF MORE THAN TWO POPULATION MEANS	140
8.3	MULTIPLE COMPARISONS	143
8.4	ONE- AND MULTIWAY ANOVA	144
8.5	ASSUMPTIONS FOR ANOVA	149
8.6	DISTRIBUTION-FREE TESTS	151
8.6.1	The Kolmogorov–Smirnov Goodness-of- Fit Test	152
8.6.2	The Kruskal–Wallis Approach to ANOVA	153
8.7	CONCLUSIONS	154
	HOMEWORK PROBLEMS	154
	REFERENCES	155

CHAPTER 9 ■ Inferences Concerning Categorical Data	157
9.1 INTRODUCTION	157
9.2 TESTS AND CONFIDENCE INTERVALS FOR A SINGLE PROPORTION	157
9.3 TESTS AND CONFIDENCE INTERVALS FOR TWO PROPORTIONS	161
9.4 CHI-SQUARE TESTS CONCERNING MORE THAN TWO POPULATION PROPORTIONS	163
9.4.1 Chi-Square Test for Univariate Categorical Data	163
9.4.2 Tests for Independence of Two Categorical Variables	166
9.4.3 Tests for Homogeneity of Two or More Populations	171
9.5 THE CHI-SQUARE GOODNESS-OF-FIT TEST FOR CHECKING DISTRIBUTIONAL ASSUMPTIONS	174
9.6 CONCLUSIONS	175
HOMEWORK PROBLEMS	176
REFERENCES	177
CHAPTER 10 ■ Linear Regression	179
10.1 INTRODUCTION	179
10.2 SIMPLE LINEAR REGRESSION	180
10.2.1 Correlation Coefficient	181
10.2.2 Fitting a Straight Line	183
10.2.2.1 Estimating Slope and Intercept	185
10.2.2.2 Inferences on Slope and Intercept	186
10.2.3 Prediction	189
10.3 TRANSFORMATIONS	191
10.4 UNDERSTANDING AND CALCULATING R^2	192
10.5 VERIFYING THE MAIN ASSUMPTIONS IN LINEAR REGRESSION	193
10.6 COMPARING TWO REGRESSION LINES AT A POINT AND COMPARING TWO REGRESSION PARAMETERS	194

10.7	THE REGRESSION DISCONTINUITY DESIGN (RDD)	195
10.8	MULTIPLE LINEAR REGRESSION	195
10.8.1	Confidence Intervals on Regression Parameters	199
10.8.1.1	Extrapolation	199
10.9	VARIABLE SELECTION FOR REGRESSION MODELS	199
10.10	ADDITIONAL COLLINEARITY ISSUES	201
10.11	CONCLUDING REMARKS	202
	HOMEWORK PROBLEMS	203
	REFERENCES	204
CHAPTER 11	■ Regression Models for Count Data	205
11.1	INTRODUCTION	205
11.2	POISSON REGRESSION MODEL	206
11.3	OVERDISPERSION	208
11.4	ASSESSING GOODNESS OF FIT OF POISSON REGRESSION MODELS	209
11.5	NEGATIVE BINOMIAL REGRESSION MODEL	215
11.6	CONCLUDING REMARKS	226
	APPENDIX: MAXIMUM LIKELIHOOD ESTIMATION	226
	HOMEWORK PROBLEMS	228
	REFERENCES	229
CHAPTER 12	■ Experimental Design	231
12.1	INTRODUCTION	231
12.2	COMPARISON OF DIRECT OBSERVATION AND DESIGNED EXPERIMENTS	231
12.3	MOTIVATION FOR EXPERIMENTATION	233
12.3.1	The Good Experiment	233
12.3.2	The Poor Experiment	233
12.4	A THREE-FACTOR, TWO LEVELS PER FACTOR EXPERIMENT	235
12.4.1	A Common Fallacy: Changing One Factor at a Time Experiments	236

12.5	FACTORIAL EXPERIMENTS	237
12.6	FRACTIONAL FACTORIAL EXPERIMENTS	239
12.7	SCREENING DESIGNS	241
12.8	D-OPTIMAL AND I-OPTIMAL DESIGNS	242
12.9	SAMPLE SIZE DETERMINATION	245
12.10	FIELD AND QUASI-EXPERIMENTS	248
12.11	CONCLUDING REMARKS	252
	APPENDIX: CHOICE MODELING OF EXPERIMENTS	253
	HOMEWORK PROBLEMS	254
	REFERENCES	255
CHAPTER 13 ■ Cross-Validation, Jackknife, and Bootstrap Methods for Obtaining Standard Errors		257
13.1	INTRODUCTION	257
13.2	METHODS FOR STANDARD ERROR ESTIMATION WHEN A CLOSED-FORM FORMULA IS NOT AVAILABLE	258
13.3	CROSS-VALIDATION	259
13.4	THE JACKKNIFE METHOD FOR OBTAINING STANDARD ERRORS	260
13.5	BOOTSTRAPPING	262
13.6	CONCLUDING REMARKS	265
	HOMEWORK PROBLEMS	265
	REFERENCES	266
CHAPTER 14 ■ Bayesian Approaches to Transportation Data Analysis		267
14.1	INTRODUCTION	267
14.2	FUNDAMENTALS OF BAYESIAN STATISTICS	268
14.3	BAYESIAN INFERENCE	270
14.3.1	Conjugate Priors	271
14.3.1.1	For Poisson Likelihood	271
14.3.1.2	For Binomial Likelihood	272
14.3.1.3	For Normal Likelihood	273

14.3.2	Point Estimation	277
14.3.3	Uncertainty Estimation	279
14.3.3.1	Posterior Standard Deviation	279
14.3.3.2	Credible Intervals	281
14.4	CONCLUDING REMARKS	284
	HOMEWORK PROBLEMS	284
	REFERENCES	286
CHAPTER 15	■ Microsimulation	287
15.1	INTRODUCTION	287
15.2	OVERVIEW OF TRAFFIC MICROSIMULATION MODELS	287
15.3	ANALYZING MICROSIMULATION OUTPUT	294
15.3.1	Model Calibration Validation and Verification	294
15.3.2	Aggregate Goodness-of-Fit Measurements	295
15.3.3	Statistical Analysis of Microsimulation Output	297
15.3.4	Identifying the Statistically Optimal Number of Simulation Runs	301
15.3.5	Microsimulation Calibration	302
15.3.6	Computer-Based Microsimulation Calibration	304
15.3.7	Absolute vs. Relative Model Accuracy	306
15.4	PERFORMANCE MEASURES	307
15.4.1	Overview of Performance Measures	308
15.4.2	Common Performance Measures Used in Simulation Analysis	310
15.4.3	Travel Time	311
15.4.4	Total Delay	313
15.4.5	Queue Length (Oversaturation)	316
15.4.6	Number of Vehicles Completing Simulation	317
15.4.7	Percentage of Corridor (OD Movements) Congested (PCC) Measure	317
15.4.8	Travel Time Variability	319
15.4.9	Level of Service	320

15.4.10 Travel Rate	321
15.4.11 Performance Indices	321
15.5 CONCLUDING REMARKS	321
HOMEWORK PROBLEMS	322
REFERENCES	322
APPENDIX: SOFT MODELING AND NONPARAMETRIC MODEL BUILDING, 325	
INDEX, 331	