Preface v
Acknowledgments vii

PART I

BASIC STATISTICAL CONCEPTS IN GEOGRAPHY
1 Introduction: The Context of Statistical Techniques
   1.1 The Role of Statistics in Geography 4
   1.2 Examples of Statistical Problem Solving in Geography 8
2 Geographic Data: Characteristics and Preparation
   2.1 Selected Dimensions of Geographic Data 21
   2.2 Levels of Measurement 23
   2.3 Measurement Concepts 25
   2.4 Basic Classification Methods 27

PART II

DESCRIPTIVE PROBLEM SOLVING IN GEOGRAPHY
3 Descriptive Statistics and Graphics 39
   3.1 Measures of Central Tendency 39
   3.2 Measures of Dispersion and Variability 44
   3.3 Measures of Shape or Relative Position 52
   3.4 Selected Issues: Spatial Data and Descriptive Statistics 53
4 Descriptive Spatial Statistics 62
   4.1 Spatial Measures of Central Tendency 63
   4.2 Spatial Measures of Dispersion 69

PART III

THE TRANSITION TO INFERENTIAL PROBLEM SOLVING
5 Basics of Probability and Discrete Probability Distributions
   5.1 Basic Probability Processes, Terms, and Concepts 77
   5.2 The Binomial Distribution 81
   5.3 The Geometric Distribution 83
   5.4 The Poisson Distribution 85
6 Continuous Probability Distributions and Probability Mapping
   6.1 The Normal Distribution 93
   6.2 Probability Mapping 96
7 Basic Elements of Sampling 101
   7.1 Sampling Concepts 101
   7.2 Types of Probability Sampling 108
   7.3 Spatial Sampling 111
8 Estimation in Sampling 117
   8.1 Basic Concepts in Estimation 117
   8.2 Confidence Intervals and Estimation 121
   8.3 Geographic Examples of Confidence Interval Estimation 126
   8.4 Sample Size Selection 134
Preface

When contemplating a third edition of this text, we reexamined an interrelated set of questions that we have been asking for years:

• What is the most effective way to introduce statistical methods and techniques to undergraduate and beginning graduate geography students?
• How can we best demonstrate the usefulness and practical benefits of statistics to geography students in a way that holds their interest and generates enthusiasm?
• Can students who believe they are weak in math and statistics achieve a strong fundamental competency in the use of statistical methods?

We believe that these questions can all be answered successfully if the statistical procedures are presented in the context of solving real-world geography problems—problems that already interest geography students. As much as possible, our focus centers on actual geographic problems and issues. Previous editions of this text already pointed strongly in this direction, but we are convinced more than ever that “geographic problem-solving” is the best guiding principle and contextual framework through which to present statistical techniques.

As a result, in this third edition you will see even more emphasis placed on the development of descriptive statements and statistical hypotheses that relate directly to contemporary problems in geography. Chapter 1 illustrates this “geography-first” approach by discussing several contemporary spatial patterns, all presented in the context of the scientific research process. Right from the beginning, students are exposed to critical geographic issues including the distribution of obesity levels across the United States, the global variation of life expectancy levels, the timing of the last spring frost across the southeastern U.S., and the recent changes in population patterns for both states and counties in the U.S. These issues continue to appear as geographic variables to illustrate new statistical procedures as they are introduced throughout the text.

With attention centered on real-world geographic problem-solving using statistics, our text presents more than 50 map examples throughout the following pages. It is worthwhile comparing this text to others that deal with statistics in geography, where maps are virtually nonexistent. We believe that geography students will learn how to use statistics more effectively when presented with real, contemporary problems and maps.

Generally speaking, what must a geography student learn to become skilled in applying statistics to solve real-world problems? While something of an oversimplification, two essential skills should be mastered. First, when presented with a geographic situation or problem to solve, you must be able to select the correct statistical technique (or set of techniques) that allow you to approach that situation or problem in the most effective and productive way. Second, when you have results from a statistical analysis (the computer output, for example), you must fully and properly interpret that output to reach the correct conclusions. Simply stated, you must fully understand how to interpret the results of statistical analysis and use that understanding to recommend appropriate geographic policies and plans.

What began as a fresh coat of paint on what we considered a good book, ended with a complete rewrite over a 3-year period. While the general chapter-by-chapter organization remains relatively unchanged, most of the examples are different. In addition, a number of new statistical methods are presented for the first time. More than a decade has passed since the second edition was published, and many aspects of statistical problem-solving in geography have changed dramatically. The significant advances in Geographic Information Sciences (GIS), for example, have altered the ways geographers conduct research. Consequently, we have placed greater emphasis on a variety of spatial statistics that focus entirely on geographic patterns. In fact, an entire segment of the book (Part Five) is now devoted to inferential spatial statistics, including more detailed discussions of spatial autocorrelation and related concepts such as variograms. Also presented for the first time is a separate chapter dealing specifically with two key multivariate techniques—multivariate regression models and cluster
analysis. Quite simply, it is fair to say that this is a virtual rewrite of the second edition.

In addition to maintaining the successful general organization of the previous editions, we also keep several features that characterized the earlier editions. We continue to stress the importance of written narratives that clearly explain each statistical technique in ways that undergraduate and beginning graduate students in geography can understand. This is accomplished without compromising or oversimplifying the statistical integrity of the material. We also continue to be enthusiastic advocates of the exploratory, investigative approach. Therefore, in many parts of the text you will see the calculation of descriptive summaries and graphics used as intermediate steps in a multi-stage geographic research process.

In some cases, we continue to use inferential statistics as exploratory and descriptive tools to reinforce the basic concept of learning as much as possible about a situation in geography. More specifically, we demonstrate how the magnitudes of multiple test statistic values can be compared to gain geographic insights, even though the statistical test is not being used to make formal probabilistic statements.

Also, in several instances, we “pair together” a slightly stronger parametric inferential test (such as ANOVA) with a slightly weaker non-parametric test (such as Kruskal-Wallis). This “pairing” allows a researcher to gather as much useful information as possible about a geographic problem or situation from appropriate statistical techniques. For example, in some cases it may be uncertain if the data fully meet all of the required assumptions needed to run a stronger parametric test. If both a parametric and corresponding nonparametric test are run, we can examine the differences in test statistic results and better evaluate the comparative advantages and disadvantages each test offers.

We expect you will use a statistical software package to do most of the analysis, rather than making extensive calculations manually. Nevertheless, when presenting many of the statistical techniques, we feel the knowledge and understanding gained through showing the basic steps in the calculation procedure will help you better understand the goals and objectives of the technique.

A CD is available with this text. The CD contains virtually all of the major data sets we use in the various problems and examples. For additional information related to the CD, please contact Waveland Press through their website at www.waveland.com.

Finally, upon review of the final product we realize that too much information is included to cover everything in a single undergraduate course in statistics and geography. We defer to an instructor’s best judgment as to the proper pace to follow when using the text in a course. Nonetheless, we encourage instructors to strive for a positive student experience in the hope that an appreciation for the role of statistics in geography will encourage the student to become a lifetime learner in the discipline. One possible strategy for instructors wishing to reduce the volume of material presented in a single course is to consider using certain chapters in more advanced courses. For instance, the inferential spatial statistics in Part V (chapters 13–15) might be introduced in an undergraduate GIS course. Similarly, the multiple regression and cluster analysis material in chapter 18 could be presented to students in an advanced statistical applications course in geography.
Acknowledgments

Many people have helped with the development of the third edition of this book. As we point out in the preface, this edition is fundamentally a thorough rewrite of the second edition. Of particular significance are the many new examples and considerable strengthening of the maps and graphic support.

In this context, certain individuals deserve special recognition. John Patrick Soderstrom created or revised earlier versions of all the figures and tables that appear in the book. This was a complicated and demanding task, which Patrick performed at the highest standard of excellence. The success of this book will be due in part to the exemplary quality of his graphics and ancillary material.

Also deserving particular mention is Dr. Kevin A. Butler. In the earlier, formative stages of book development and organization, Kevin played a critical role. He created the original ideas for a number of the geographic problems new to this edition, including: the linear directional mean of hurricane and tropical storm tracks across Puerto Rico; the standard deviational ellipse of anti-shipping (pirating) activities off the east coast of Africa; nearest neighbor analysis of public services in Toronto, Canada; and Moran’s I analysis of racial and ethnic groups in Cleveland, Ohio.

A number of faculty members commented critically on aspects of book organization and development of geographic examples. Particular thanks are extended to Dr. Darren Parnell, who graciously provided original data for last spring frost dates at various weather stations in the Southeastern United States.

Several students and former students at Salisbury University also contributed to the creation of geographic examples, data collection, or statistical analysis. Particularly, we would like to recognize the following individuals: Danielle Bruner, Court McGrew, and Denise Tweedale.

The personnel at Waveland Press have been incredibly supportive throughout the entire process. Don Rosso, an Editor and the Production Manager at Waveland, has always been available if we had a question or issue that needed to be addressed, and he has constantly maintained an enthusiastic and positive attitude about our project. Perhaps as important as anything else, he encouraged us to keep improving the manuscript, even if it meant lengthening the amount of time involved. Producing the finest product has always been his stated goal. Dakota West has done a fine job in editing the manuscript and preparing it for distribution. We thank them for their efforts.

Finally, we would be remiss to overlook the support of our respective families. This project took more time, and was substantially more work than any of us originally estimated. Without the support of our families, we could not have devoted the time required to produce a quality textbook.