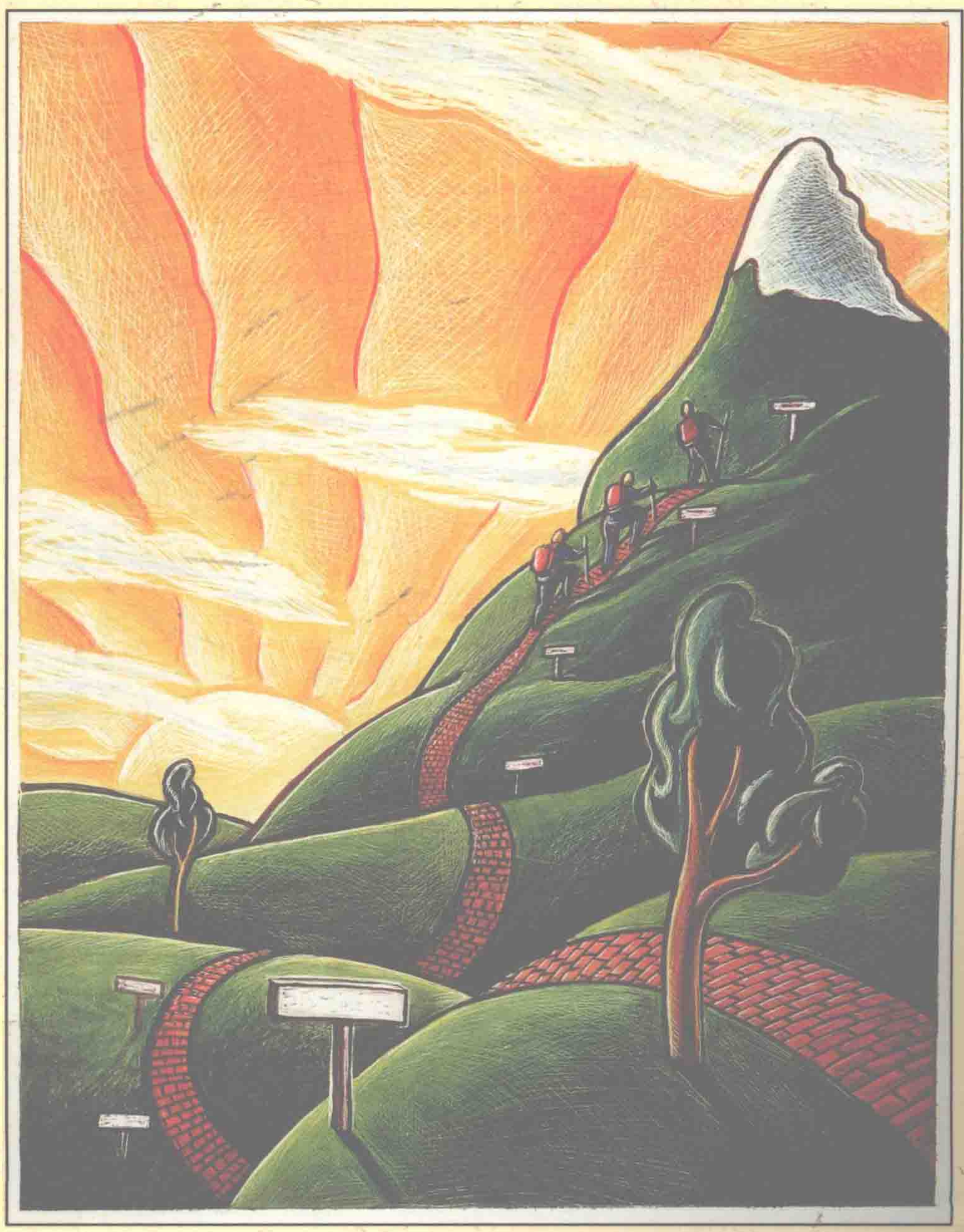
Student Friendly Statistics



Thomas Sanocki

Student Friendly Statistics

Thomas Sanocki

University of South Florida
Tampa, Florida

Upper Saddle River, New Jersey 07458

To Madeline and the innovators who will define their own future

Library of Congress Cataloging-in-Publication Data

Sanocki, Thomas.
Student-friendly statistics / author, Thomas Sanocki.
p. cm.
Includes bibliographical references and index.
ISBN 0-13-026521-7
1. Statistics. I. Title.

QA276.12.S27 2000
519.5–dc21 00-061158

VP, Editorial Director: Laura Pearson
Acquisitions Editor: Jayme Heffler
Editorial Assistant: April Dawn Klemm
Managing Editor: Mary Rottino
Production Liaison: Fran Russello
Project Manager: Kelly Ricci, The PRD Group
Prepress and Manufacturing Buyer: Tricia Kenny
Art Director: Jayne Conte
Cover Designer: Karen Sanatar
Marketing Manager: Sharon Cosgrove

This book was set in 9.5/10.5 Baskerville by The PRD Group
and was printed and bound by R. R. Donnelley, Harrisonburg.
The cover was printed by Phoenix Color.

A Division of Pearson Education
Upper Saddle River, New Jersey 07458

Printed in the United States of America

10 9 8 7 6 5 4 3 2 1

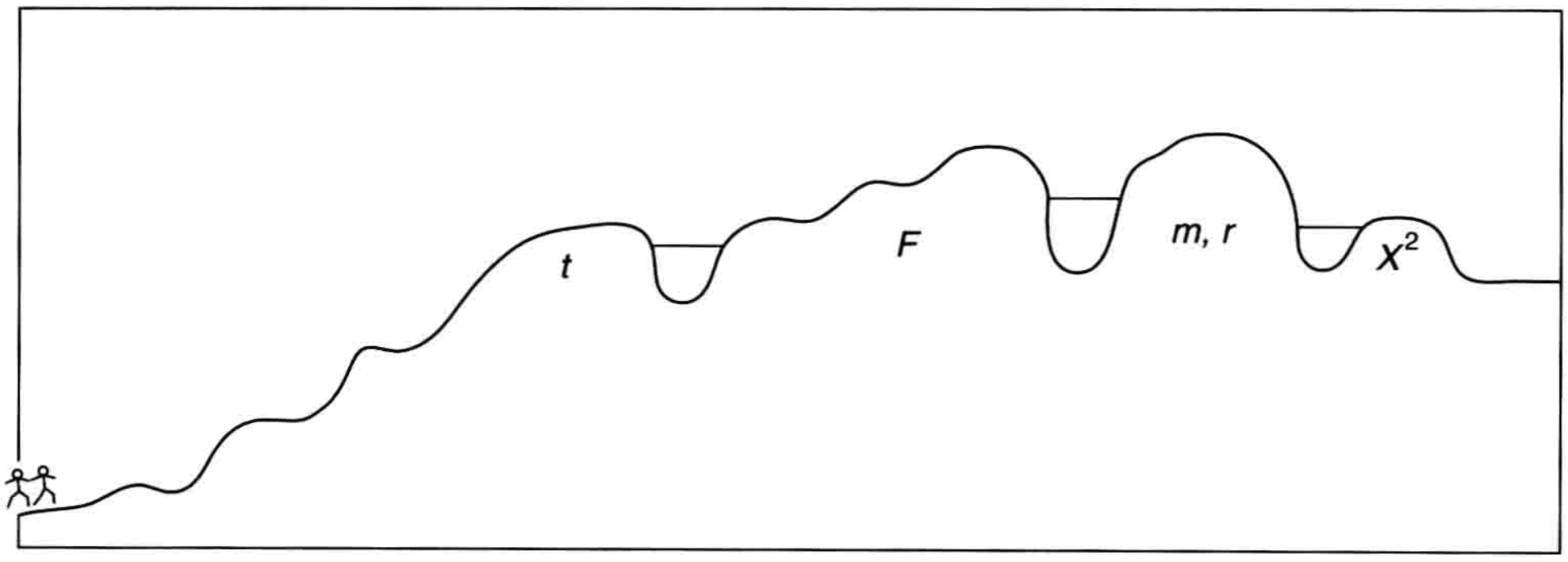
ISBN 0-13-026521-7

Prentice-Hall International (UK) Limited, *London*
Prentice-Hall of Australia Pty. Limited, *Sydney*
Prentice-Hall Canada Inc., *Toronto*
Prentice-Hall Hispanoamericana, S.A., *Mexico*
Prentice-Hall of India Private Limited, *New Delhi*
Prentice-Hall of Japan, Inc., *Tokyo*
Pearson Education Asia Pte. Ltd., *Singapore*
Editora Prentice-Hall do Brasil, Ltda., *Rio de Janeiro*

PREFACE TO THE INSTRUCTOR

The purpose of this text is to make the major statistics accessible to all students while revealing the depth of the statistics and their meaning within research design. The statistics are explained completely, with simple terms and an interactive style. The main statistics concepts treated are the logic of statistical inference, the *t*-test, single-factor and multifactor analysis of variance (including interaction), linear regression and correlation, and chi-square. Because the tools of ANOVA, interaction, and correlation/regression are extremely important, I provide a foundation for them in the first parts of the book that is solid yet brief enough to allow major coverage of these advanced statistics in a semester course. Relations to experimental design are covered in order to increase the meaningfulness of statistics.

Any instructor knows that consistent effort is needed to learn statistics and to individuate them from each other. To get this idea across, I have developed the mountain-climbing journey metaphor illustrated. Each statistic is pictured as a mountain that can be unpacked into more basic concepts. The hikers represent student progress along the journey. A pair of hikers indicates the student's position as each chapter begins, and the spacing between hikers corresponds to chapter length. Initial chapters are short and easy for a gradual warm-up, followed by longer chapters as the material becomes more advanced. The journey metaphor also helps students learn to individuate statistics, an issue that is emphasized in the last chapter, as students look back at their path and the differences between statistics.



The exposition of the statistics is motivated by the psychologies of perception, cognition, and learning. The goal is to help students establish a core understanding that will be maintained in memory. All statistics are taught in the context of tables that break the computations into meaningful pieces. Page layouts are designed to group related ideas and figures, and to facilitate student interaction. Definitional formulas are used because they are meaningful; opaque computational formulas are eschewed because they are unnecessary in this age of computers. Major ideas common to most statistics are reinforced throughout the book.

Because context aids learning, the statistics are presented within meaningful research contexts. To simplify exposition, the most prototypical forms of statistics are emphasized. Other important distinctions are covered, but after the main exposition.

Active learning is encouraged throughout. You are invited to have students copy activity pages and hand them in for checking. Encourage students to copy formulas and

tables on their own paper (the tight spacing of this text was dictated by economics). Overheads can be found at the Internet address that follows.

This book is designed to be covered sequentially. However, it is possible to omit any of the more advanced chapters without loss of continuity. (These are Chapters 9 and 10 on multilevel and multifactor ANOVA, respectively, Chapter 11 on regression and correlation, and Chapter 12 on chi-square).

I thank the many students who have found statistics rewarding to learn. And I am happy to acknowledge the helpful comments and advice on this text from Madeline Altabe, Judy Bourgeois-Smith, Jacquelyn Fresenius, Jennifer Kisamore, Evelyn Mott, and Doug Rohrer. I am deeply indebted to Tracy Krueger and Jennifer Perone for writing the Exercises. And I thank Allison Westlake for finding a great cover, and Anne Mattson and the PRD Group for being so helpful in the production of this book. I am sure students will find that I've made some small errors, but (hopefully!) no large ones. I invite reports of errors and will post corrections as well as thanks, at *http://chuma.cas.usf.educ/~sanocki/SFS corrections.html.* Also, look for overheads, larger versions of the statistics worksheets, and step-by-step instructions at *http://chuma.cas.usf.edu/~sanocki/SFS readable.html.*

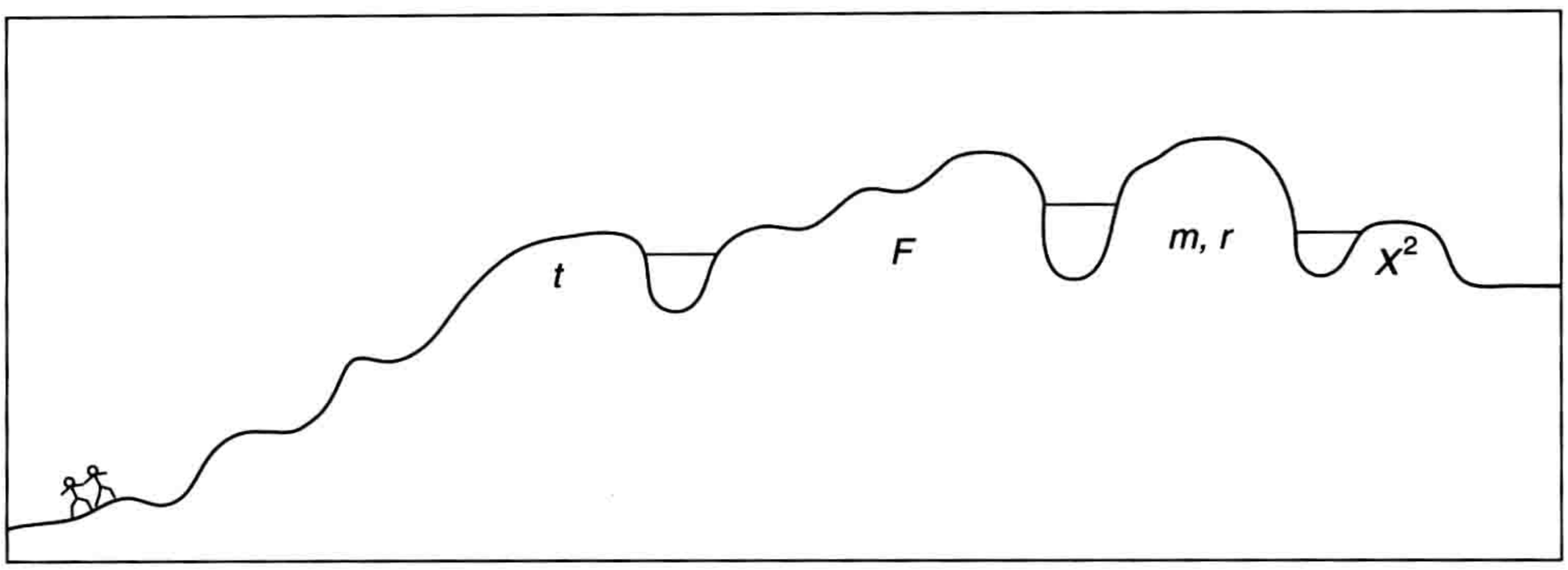
Thomas Sanocki
E-mail: Sanocki@chuma.cas.usf.edu

Trail Head: Start Here!

PREFACE TO THE STUDENT

Learning statistics is a journey that will require steady effort, but that will reward you and sometimes even surprise you pleasantly. This prefactory chapter provides a brief introduction to your journey, including tips for making the journey more pleasant. Statistics does not have to be a scary subject, as you will learn.

To help you maintain perspective along your journey, a "trail map" will be provided at the start of each chapter. It provides important landmarks. For example, the first large goal—the *t*-test—is represented by "Mount *t*." It is built on smaller concepts that we will take one step at a time. The placement of the hikers at the far left indicates that you are just beginning your journey.



WHY DO STATISTICS SEEM DIFFICULT?

It helps to know what you are up against. Contrary to popular belief, behavioral statistics are not difficult because of the mathematics. Good teachers reduce mathematics for students of behavioral science. In this text, the math has been minimized. All concepts are taught in terms of their meaning first, and only simple arithmetic will be required. The real problem with statistics, in my view, is that statistics are very abstract and unlike anything you have met or even thought about previously. They are strange for your brain! This means that they are difficult to learn and very easy to forget.

How do you get to know a strange, different thing? The best way to get to know anything is to spend time with it, interact with it, and even play with it. Gradually, it can become familiar, even friendly. In the case of statistics, doing the statistics—doing the step-by-step calculation procedures—is an excellent way to get to know them. At the same time, remember that statistics are also meaningful, even deeply meaningful. While doing calculations, think about their meaning and act on their meaning. Write their meanings in your own words in spaces in the book or on your own paper. All of the procedures you will do in this book are designed to make their meaning apparent.

THINGS TO LOOK FOR

Your trail is well marked with signs. Look for them:

> **Important Definitions**
>
> These will be highlighted in boxes.

CONCEPTUALIZATION AHEAD

Especially important idea units are set off by markings like this.

CONCEPT DONE!

Other concepts will be designated by a simple boldface title or italics. These often are important linking concepts. You will also see white spaces; these are great places to summarize your thinking in your own words.

- **Important terms are listed at the end of each chapter.** Learn by actively summarizing their meaning.
- **Exercises are also provided.** No better way to learn!

THE MOST IMPORTANT FEATURE OF THIS BOOK

All of the features of this book pale in comparison to the most important feature—you! You will produce learning by interacting with this book. A substantial body of research on learning and memory indicates that self-relatedness adds greatly to memory. This is vitally important when you are learning something strange like statistics. The more you make the concepts meaningful to yourself, the better you will remember statistics. That means writing in this book, and on your own paper, using your own words to rephrase definitions, and doing calculations while thinking about what they mean.

The content of this book has been formatted to make it easier for you to integrate concepts. As much as possible, related text, figures, and tables have been grouped together on the same pages. This allows you to focus on meaning rather than looking for figures. What is critical is that you are creating a memory image, or an episodic trace (a record of your experience), that will become part of your memory. If the memory is meaningful and well integrated, it will last longer and be easier to recall later. Careful, active study of the concepts will be rewarded with better recall at testing time and in the future when you need to refer back to this book. Note that even advanced statisticians often return to their old textbooks. Their memory needs refreshing too. What a good course and text on statistics should do is make it easy for you to return to your text years later and, after some review, to remember what you had learned.

A PREVIEW OF WHAT IS AHEAD

Your journey will begin with a steady "climb" involving smaller concepts that build on each other. They climax at your first major statistic, the *t*-test (Mount *t*). The *t*-test was quite an intellectual feat when it was created, and your comprehension of it will also be

worthy of praise. Once you understand the *t*-test, you will gain a new perspective on statistics. On the next part of your journey, you will cross a convenient bridge to an important series of statistics called analysis of variance (ANOVA). The most outstanding of these allows you to examine the multidimensional (multicausal) nature of behavior—multifactor ANOVA.

Next will be statistics that allow you to examine relations between variables, including complex relations involving multiple variables (correlation, regression, and multiple regression). The last statistic, chi-square, allows you to examine categories of behavior. Your introduction to statistics will be complete in the last chapter, when you will look back and gain perspective. This chapter compares statistics and begins to handle the question "How do I choose a statistic?"

THE STEP-BY-STEP PROCEDURES

Each statistic taught can be calculated with a step-by-step procedure. You can find these procedures on the following pages:

GET PROVISIONS FIRST

My last advice is to prepare yourself each time you travel on this journey. Your brain will learn much more efficiently if it is relaxed and well nourished. If statistics or test taking makes you anxious, the Appendix provides helpful tips. If this book's type seems small, try an inexpensive pair of reading lenses from a local drugstore—they can make interactions with this book more pleasant. Successful interactions will also generate good feelings, and there are plenty of those ahead.

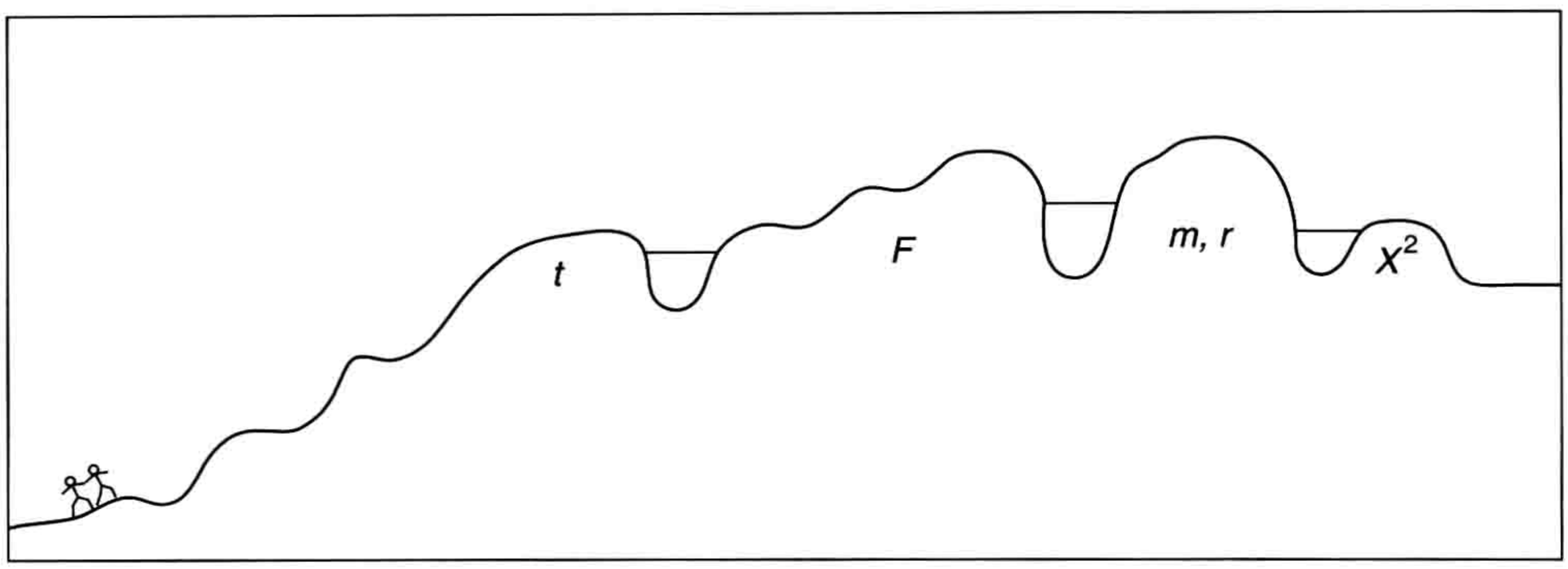
You will need two provisions: a simple calculator and a pencil. Happy trails!

CONTENTS

Chapter 1

The Researcher's Question

Statistics are marvelous tools that make behavioral research possible. A good way to become acquainted with tools is to learn why the tools are necessary. The first goal of this chapter, then, is to illustrate the need for statistics. Also, to make statistics more real, an example is introduced that will be used throughout this book. Get ready, you are starting your journey!



THE RESEARCHER'S QUESTION

Researchers become very interested in their research and often work for months and even years to collect data in an experiment that interests them. After the data have been collected, a big question arises:

Researcher's Question

"Did my experiment work?"

When this question is really understood, it is quite deep. We will be developing an understanding of the question throughout this book.

One way to learn about the meaning of the question is to rephrase it. Let us first rephrase it in terms of the **independent variable** (the variable that the researcher manipulates) and the **dependent variable** (the variable the researcher measures). In the simplest type of experiment, the independent variable is varied between two groups of participants. The dependent variable is measured for each group. We can rephrase the question as follows:

Researcher's Question Rephrased

"Did the independent variable cause a difference in the dependent variable—did it cause one group of participants to behave differently than the other group of participants?"

This question is answered by comparing the results for the two groups. To illustrate, let's consider two examples.

EXAMPLE 1: ISEN, CLARK, AND SCHWARTZ (1976)

The first example—from Isen, Clark, and Schwartz (1976)—was quite new when I first read it as part of an undergraduate research project. It is an interesting instance of research, and I've found that it provides a nice starting point.[1]

Isen, Clark, and Schwartz (1976) examined the relation between good mood and willingness to help. Their hypothesis was that good mood leads to an increased willingness to help other people. The participants in their field study were ordinary people who were in their homes during the day. The independent variable was mood. The experimenter's assistant ("confederate") came to the door, and either gave a sample packet of stationary to the participant (experimental group) or simply demonstrated the stationary (control group). Isen and her colleagues assumed that the free stationary put the experimental group in a good mood, and that the control group was unaffected. (Note that this study was conducted before the age of telemarketing, so interruptions in the home were less bothersome than today.) How did mood influence desire to help?

We will consider one condition from their experiments. Five minutes after the confederate came to the door, each participant received a phone call from a stranger who asked for help. ("I spent my last coin calling this number and it was the wrong one; could you look up a number and call it for me?") The dependent variable was whether or not the participant helped. The researchers found that participants in a good mood were more likely to help. Table 1.1 shows the results for each of the 11 control and 12 experimental participants in the relevant condition. Examine these results. Now, let me ask you, "Based on these data, did the experimental group behave differently than the control group?" Did the experiment "work"?

If your answer is "Yes!," then we are in agreement. The results are very strong. In the control group, only 1 participant helped, whereas in the experimental group 10 participants helped. (Also, note that Isen *et al.* did a good job of considering and testing alternative explanations of the effects, such as obligations that might be assumed because of the free gift; see Isen *et al.*, 1976.) Isen *et al.*'s effect is very likely to be a real effect, and we don't really need statistics to tell us that. However, this is an unusually effective experiment. In most other cases, we do need statistics because it is not obvious if an experimental effect is real, or if the effect could have occurred because of something called chance variation.

[1] Here is the reference, as well as some more recent references:

Isen, A. M., Clark, M., & Schwartz, M. F. (1976). Duration of the effect of good mood on helping: "Footprints on the sands of time." *Journal of Personality & Social Psychology, 34,* 385–393.

Isen, A. M. (1987). Positive affect, cognitive processes, and social behavior. In L. Berkowitz *et al.* (Eds.), *Advances in experimental social psychology* (Vol. 20, pp. 203–253). San Diego: Academic Press.

Ashby, A. F., Isen, A. M., & Turken, A. U. (1999). A neuropsychological theory of positive affect and its influence on cognition. *Psychological Review, 106,* 529–550.

TABLE 1.1. Each Participant's Response to the Request for Help

CONTROL GROUP		EXPERIMENTAL GROUP	
1.	Did not	1.	Helped
2.	Did not	2.	Helped
3.	Did not	3.	Helped
4.	Did not	4.	Helped
5.	Helped	5.	Did not
6.	Did not	6.	Helped
7.	Did not	7.	Helped
8.	Did not	8.	Helped
9.	Did not	9.	Helped
10.	Did not	10.	Helped
11.	Did not	11.	Did not
		12.	Helped

Source: From Isen *et al.* (1976), Experiment 2.

EXAMPLE 2: A LASTING FRIENDSHIP

Let us now consider a hypothetical example. It is designed to be useful for teaching statistics, and we will use it throughout the book. Therefore, it is a good idea to make friends with it!

Let us assume that the researchers continued their research on mood and helping, but used a more continuous dependent variable (i.e., one with many possible values, rather than the two values of helped/did not). Assume that the experimenters used a confederate who asked the participant to help her do a task. The experimental group was put into a good mood before the request but the control group was not. The dependent variable was how long the participants helped, measured in minutes and seconds.

Some possible results for a total of 12 participants are shown in Table 1.2. The results have been rounded to the nearest minute for simplicity. In this table, I've given each participant a unique number, and I've calculated the mean (average) for each group. (We will cover means in Chapter 3.) Look at the results carefully and tell me, "Did the experimental group behave differently than the control group?"

I hope your answer is something like "Although the experimental group did help more overall, it is difficult to tell for sure if there is a real difference. Some experimental participants helped a lot (e.g., number 10), but so did some control participants

TABLE 1.2. Amount of Time in Minutes That Each of 12 Hypothetical Participants Helped Another Person

CONTROL GROUP		EXPERIMENTAL GROUP	
1.	8	7.	10
2.	9	8.	16
3.	12	9.	12
4.	7	10.	14
5.	13	11.	11
6.	11	12.	15
Mean (average)	10	Mean (average)	13