

Statistics and Probability in High School

Carmen Batanero

Universidad de Granada, Spain

and

Manfred Borovcnik

University of Klagenfurt, Austria

Statistics and probability are fascinating fields, tightly interwoven with the context of the problems which have to be modelled. The authors demonstrate how investigations and experiments provide promising teaching strategies to help high-school students acquire statistical and probabilistic literacy.

In the first chapter the authors put into practice the following educational principles, reflecting their views of how these subjects should be taught: a focus on the most relevant ideas and postpone extensions to later stages; illustrating the complementary/dual nature of statistical and probabilistic reasoning; utilising the potential of technology and show its limits; and reflecting on the different levels of formalisation to meet the wide variety of students' previous knowledge, abilities, and learning types.

The remaining chapters deal with exploratory data analysis, modelling information by probabilities, exploring and modelling association, and with sampling and inference. Throughout the book, a modelling view of the concepts guides the presentation.

In each chapter, the development of a cluster of fundamental ideas is centred around a statistical study or a real-world problem that leads to statistical questions requiring data in order to be answered. The concepts developed are designed to lead to meaningful solutions rather than remain abstract entities. For each cluster of ideas, the authors review the relevant research on misconceptions and synthesise the results of research in order to support teaching of statistics and probability in high school.

What makes this book unique is its rich source of worked-through tasks and its focus on the interrelations between teaching and empirical research on understanding statistics and probability.

SensePublishers

DIVS

ISBN 978-94-6300-622-4



9 789463 006224

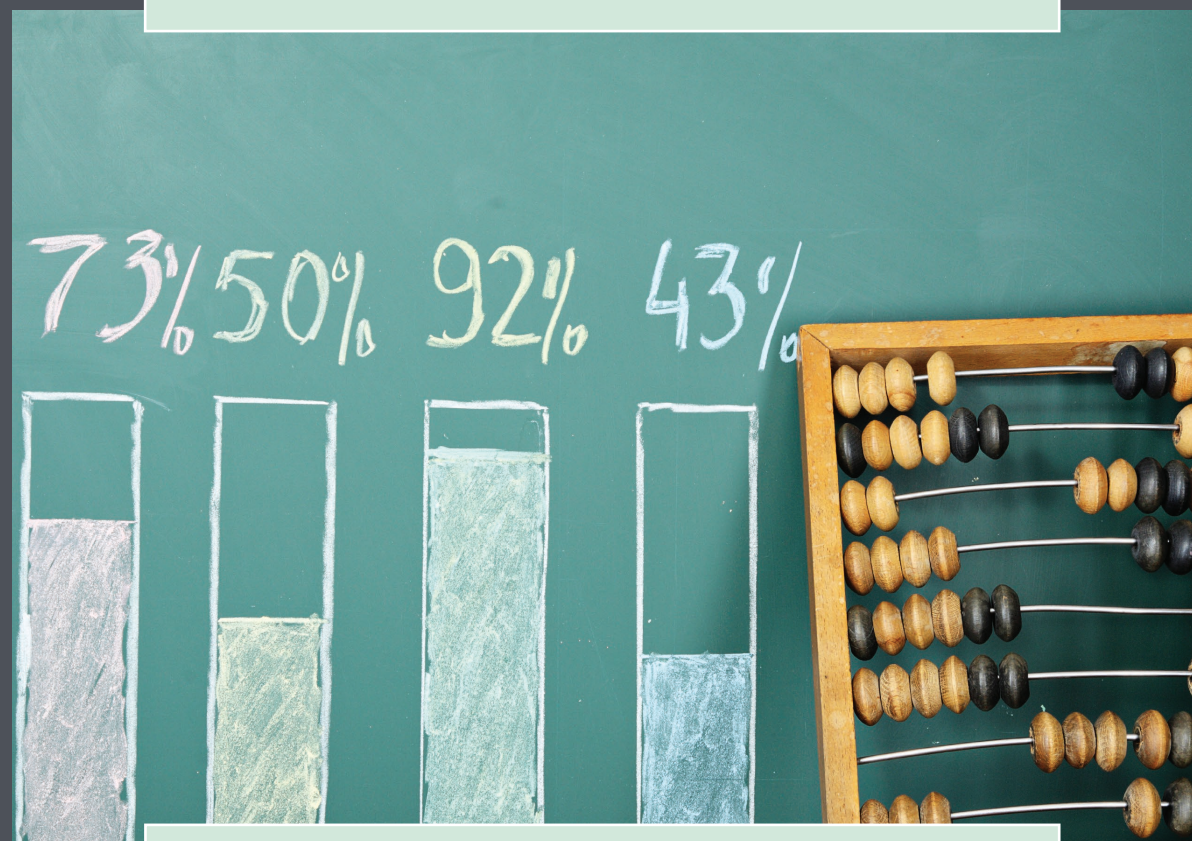


Statistics and Probability in High School

Carmen Batanero and Manfred Borovcnik

Statistics and Probability in High School

Carmen Batanero and
Manfred Borovcnik



SensePublishers

TABLE OF CONTENTS

Preface	ix
1. Educational Principles for Statistics and Probability	1
1.1. Introduction	1
1.2. Fundamental Ideas in Statistics and Probability	2
1.2.1. Exploratory Data Analysis (Chapter 2)	3
1.2.2. Modelling Information by Probabilities (Chapter 3)	4
1.2.3. Exploring and Modelling Association (Chapter 4)	5
1.2.4. Sampling and Inference (Chapter 5)	6
1.3. Complementary Views of Statistics and Mathematics	7
1.4. The Role of Technology	10
1.5. Adapting the Levels of Formalisation to the Diversity of Students	12
1.6. Statistical and Probabilistic Literacy	12
1.6.1. Statistical Literacy	12
1.6.2. Statistical Literacy Components	13
1.6.3. Actions and Resources Directed to Increase Statistical Literacy	14
1.7. Statistical and Probabilistic Thinking	15
1.7.1. Statistical Thinking	16
1.7.2. The Statistical Investigation Cycle	16
1.7.3. Fundamental Types of Statistical Thinking	18
1.7.4. Components of Probabilistic Thinking	20
1.8. Making Sense of Statistics and Probability	21
1.9. Statistical Investigations and Experiments	22
1.10. Final Thoughts	22
2. Exploratory Data Analysis	25
2.1. Introduction	25
2.2. A Teaching Situation to Introduce Elementary Statistical Concepts and Procedures	26
2.2.1. Starting Questions	26
2.2.2. Exploring Qualitative Variables	27
2.2.3. Exploring Numerical Variables	28
2.2.4. Comparing Groups	34
2.3. Additional Activities	38
2.3.1. Exploring Continuous Variables	38
2.3.2. Exploring Bivariate Relationships	45
2.4. Synthesis of Learning Goals	46
2.4.1. Distribution and Different Types of Frequencies	47
2.4.2. Simple Univariate Graphs	48
2.4.3. Simple Summary Statistics	49
2.4.4. Spirit of Exploratory Data Analysis (EDA)	51
2.4.5. Basic Strategies in Data Exploration	52

TABLE OF CONTENTS

2.5. Students' Reasoning and Potential Difficulties	54
2.5.1. Graphical Competencies and Communication Skills	54
2.5.2. Errors in Producing Graphs	56
2.5.3. Understanding Measures of Central Tendency or Centre	57
2.5.4. Understanding Spread	60
2.5.5. Understanding Order Statistics	61
2.6. Additional Resources	62
2.6.1. Journals and Books	62
2.6.2. Data Sets	63
2.6.3. Internet Resources	63
Appendix: Data	65
3. Modelling Information by Probabilities	67
3.1. Introduction	67
3.2. Teaching Situations to Characterise Probability	69
3.2.1. Frequentist Probability: Investigating Coin Tossing	69
3.2.2. Subjectivist Probability: The Insurance Contract	72
3.2.3. Laplace (A Priori) Probability: Calibrating Weights of Evidence	73
3.3. Teaching Situations Introducing Conditional Probability	74
3.3.1. Conditional Probability and Circumstantial Evidence	75
3.3.2. Conditional Probability and Compound Probability	78
3.4. Additional Teaching Activities	79
3.4.1. Random Variables	79
3.4.2. Additivity of Expected Value and Variance for Repeated Experiments	81
3.4.3. Distribution Models for Standard Situations	83
3.4.4. Central Theorems	88
3.5. Synthesis of Learning Goals	92
3.5.1. Concepts to Model and Investigate Uncertain Situations	92
3.5.2. Different Connotations of Probability	93
3.5.3. Circumstantial Evidence and Bayes' Formula	97
3.5.4. Random Variables and Expectation	99
3.5.5. Standard Models of Distributions	100
3.5.6. Law of Large Numbers and Central Limit Theorem	101
3.6. Students' Reasoning and Potential Difficulties	104
3.6.1. Misconceptions and Heuristics (Strategies) in Probability Situations	104
3.6.2. Independence and Conditional Probability	107
3.6.3. Taking into Account Students' Reasoning to Improve Teaching	108
3.7. Additional Resources and Ideas	109
3.7.1. Investigating Randomness: Generating Coin Tossing from Memory	109
3.7.2. Odds and Bayes' Formula – Revising Weights of Evidence	109
3.7.3. Mediating Tools to Support Teaching	110

4. Exploring and Modelling Association	117
4.1. Introduction	117
4.2. A Teaching Situation to Explore Contingency Tables	119
4.2.1. Exploring Association in 2×2 Contingency Tables	119
4.2.2. Different Distributions in 2×2 Contingency Tables	121
4.2.3. Simple Methods to Evaluate Association in 2×2 Tables	123
4.2.4. Expected Frequencies for Independent Variables	124
4.3. Life Expectancy: A Teaching Situation to Explore Correlation and Regression	126
4.3.1. Exploring and Explaining Correlation	127
4.3.2. Fitting Models to Bivariate Data	134
4.4. Additional Activities	140
4.5. Synthesis of Learning Goals	142
4.5.1. Contingency Tables	143
4.5.2. Correlation and Regression	144
4.6. Some Potential Difficulties	148
4.6.1. Intuitive Strategies in Contingency Tables	148
4.6.2. Linear Regression and Correlation	150
4.6.3. Misconceptions Related to both Association and Correlation	153
4.7. Additional Resources and Ideas	157
4.7.1. Measures of Association in Contingency Tables	157
4.7.2. Introduction to Multivariate Visualisation	158
5. Sampling and Inference	163
5.1. Introduction	163
5.2. A Teaching Situation: The Tea-Tasting Experiment	164
5.2.1. The Lady Tasting Tea	165
5.2.2. Using Experimental Data to Test a Hypothesis	165
5.2.3. Different Approaches to Compute the p -value	169
5.2.4. Sampling Distribution for the Proportion and the Effect of Sample Size	171
5.2.5. Estimating the Population Proportion by a Confidence Interval	172
5.3. Additional Activities	173
5.3.1. Exploring the Central Limit Theorem	173
5.3.2. Inference for Proportions	174
5.3.3. Inference for a Mean	176
5.3.4. Statistical Tests as Decision Rules	177
5.4. Synthesis of Learning Goals	181
5.4.1. Basic Inferential Concepts and Procedures	181
5.4.2. Additional Advanced Learning Goals	184
5.5. Some Potential Difficulties	184
5.5.1. Understanding Random Sampling and Sampling Distributions	185
5.5.2. Understanding Statistical Tests	187
5.5.3. Understanding Confidence Intervals	189

TABLE OF CONTENTS

5.6. Additional Resources and Ideas	191
5.6.1. Developing Informal Ideas of Inference	191
5.6.2. Resampling Methods	191
5.6.3. Updating a Prior Distribution for a Proportion	193
References	197
Author Index	209
Subject Index	213

PREFACE

Research in statistics and probability education has produced a variety of results that would be useful for both secondary and high-school mathematics teachers and the educators of these teachers. Although there are many good textbooks in different countries that describe statistical ideas with a formalisation level adequate for students, usually these textbooks are written in a sequential way so that the different concepts and procedures are introduced in turn, with insufficient connections between them and limited attention to students' underlying intuitions.

There are, of course, excellent exceptions such as the books produced by the Schools Council (1980) in Statistical Education Project in the 1980's; yet, even, these textbooks do not include a detailed summary of research related to the teaching of the concepts, which started to get shape only after the first International Conference on Teaching Statistics in Sheffield in 1982.

In the later stages of our careers and, after collaborating and corresponding for many years in different projects, we decided to write a book directed to reinforce the mathematical and didactical knowledge of high-school teachers in statistics and probability. At the same time, we wish to offer examples of potential activities useful to introduce the main statistics and probability concepts and enhance the underlying ideas at this school level.

Consequently, in this book we provide examples of teaching situations, while at the same time we review research on adolescents' stochastic¹ reasoning and literacy, with the aim to provide recommendations and orientations for teaching these topics within high-school mathematics. The expression "high school" relates to different educational levels depending on the country; in this book, we will consider students from ages 14 to 18 (grades 9–12 in the United States of America curriculum). The book is organised in five chapters:

In the first chapter, we present some principles we use to select the content analysed in the book and the approach to teach this content. These principles emerge from:

- a. Our own teaching and research experience;
- b. An analysis of stochastic high-school curricula in several countries (e.g., ACARA, 2010; NCTM, 2000; CCSSI, 2010, MEC, 2007);
- c. The synthesis of available research (as summarised, for example, in Biehler, Ben-Zvi, Bakker, & Makar, 2013; Chernoff and Sriraman, 2014; Garfield & Ben-Zvi, 2008; Jones, 2005; Jones, Langrall, & Money, 2007; Shaughnessy, 1992, 2007; Shaughnessy, Garfield, & Greer, 1996);

¹ In some countries the term *stochastics* is used to highlight the mutual dependence between probabilistic and statistical knowledge and reasoning. Throughout the book we occasionally use *stochastics* for statistics and probability to express our view that these fields are tightly interconnected and should be taught together.

PREFACE

- d. Our own conceptions of statistical and probabilistic literacy, thinking, and reasoning; and
- e. Our extensive experience with strategies that may help support student development in stochastic literacy, thinking, and reasoning.

The first chapter sets out key educational principles. Each of the following chapters (Chapters 2–5) has a focus on a group of related fundamental stochastic ideas, while taking into account that high-school stochastics should be built on basic ideas that students have encountered at primary and middle-school levels. These chapters are organized according to a common structure, including an introduction, with a short analysis of the main stochastic ideas in the particular topic and its place in the curriculum; some initial and more advanced specific examples that may serve to involve learners actively as they progress in their development of the concepts, a summary of what is known about difficulties students encounter with the related concepts, a synthesis of the main learning goals in the chapter, and finally, some additional resources that may help teachers and students. When possible, we make connections between the different chapters and include some historical notes that shed light on ways of thinking about the concepts.

We have tried to give a balanced view on probability and statistics, with a focus on the interrelated nature of the concepts, integrating probabilistic ideas at a level suitable for high school teaching, including the step from descriptive statistics to statistical inference. Where ever we could do it, we have also tried to integrate mathematical concepts and contexts so that the mathematics developed becomes meaningful for the learners. May our exposition contribute to an increase in statistical and probabilistic literacy in our societies.

We hope the book will be both useful for practising teachers, as well as for researchers in statistics education and practitioners in teacher educators (teacher trainers). The different chapters contain original materials, but build upon our extended set of publications, part of which is listed in the references.

We thank our colleagues and students who have commented several drafts of the chapters. Among them we want to name especially two who accompanied us in our research work now for decades: Juan D. Godino and, particularly, Ramesh Kapadia who was also helpful for improving the English. Finally, we would like to express our deepest gratitude to our families and friends for their encouragement and support over the years when we were writing the book.

May, 2016

Carmen Batanero and Manfred Borovcnik